Université de Montréal

La perception d'attributs visuels de premier et deuxième ordres

par

Rémy Allard

Département de psychologie
Faculté des Arts et des Sciences

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de Philosophiae Doctor (Ph.D.)
en Psychologie
option Sciences cognitives et neuropsychologie

Août, 2008

Université de Montréal

Faculté des études supérieures

Cette thèse intitulée :

La perception d'attributs visuels de premier et deuxième ordres

présentée par :

Rémy Allard

a été évaluée par un jury composé des personnes suivantes :

Denis Cousineau, président-rapporteur

Jocelyn Faubert, directeur de recherche

Frédéric Gosselin, membre du jury

Hugh Wilson, examinateur externe

Jacques Gresset, représentant du doyen de la FES

Résumé

Il n'y a pas de consensus sur la façon dont les stimuli définis par d'autres attributs que la luminance, c'est-à-dire ceux de deuxième ordre, sont traités. Certains auteurs proposent l'existence de mécanismes dédiés au traitement de deuxième ordre alors que d'autres proposent plutôt que les mécanismes traitant la luminance (premier ordre) traiteraient également les attributs de deuxième ordre. Dans le but d'élucider cette problématique, du bruit défini par la luminance et du bruit défini par le contraste ont été conçus pour évaluer le masquage intra- et inter-attribut sur le traitement de luminance (premier ordre) et de contraste (deuxième ordre). Pour la détection de stimuli statiques, l'absence de masquage inter-attribut observée implique que les traitements de stimuli statiques définis par la luminance et le contraste sont, au moins initialement, distincts. L'observation du même ratio signal-bruit nécessaire lors du masquage intra-attribut suggère que les mécanismes extrayant le signal du bruit soient communs aux deux attributs. Cependant, une recension des écrits révélant des doubles dissociations entre le traitement d'un stimulus en présence et en absence de bruit suggère que l'ajout de bruit puisse modifier qualitativement le traitement d'un stimulus. Les mécanismes d'extraction du signal du bruit communs aux traitements de ces deux attributs ne seraient donc pas nécessairement sollicités en absence de bruit. Pour le mouvement à hautes fréquences temporelles, le masquage intra- et inter-attribut avaient le même impact suggérant un traitement commun. De plus, des simulations ont démontré que l'apparente absence d'interaction inter-attribut généralement observée lors du traitement de signaux de luminance et de contraste spatialement superposés peut s'expliquer par la présence de non-linéarités non-uniformes intrinsèques au système visuel. Pour le mouvement à basses fréquences temporelles, le masquage intra- et inter-attribut n'avaient pas le même impact suggérant un traitement distinct. Ces résultats peuvent s'expliquer par l'implication d'un système de mouvement basé sur le suivi d'une caractéristique suite à l'extraction spatiale de la forme. En conclusion, le traitement spatial de premier et deuxième ordres serait distinct, mais il n'est pas nécessaire d'inférer l'existence d'un système de mouvement basé sur l'énergie dédié au deuxième ordre.

Mots-clés : Deuxième ordre, texture, mouvement, attribut visuel, modulation de luminance, modulation de contraste, bruit, masquage, modèle filtre-rectification-filtre, non-linéarité

Abstract

There is no consensus on how stimuli defined by other attributes than luminance, i.e. second-order stimuli, are processed. Some authors suggest the existence of mechanisms dedicated to second-order processing while others propose that the mechanisms processing stimuli defined by luminance (first-order stimuli) also process second-order stimuli. In order to clarify this issue, noise defined by luminance and contrast were used to assess intra- and inter-attribute masking on the processing of luminance and contrast. For the detection of static stimuli, the observed lack of inter-attribute masking implies that the processing of static stimuli defined by luminance and contrast are, at least initially, distinct. The observation of similar signal-to-noise ratio necessary to detect the signal in intra-attribute masking conditions suggests that the mechanisms extracting the signal from noise are common to both attributes. However, a literature review revealed double dissociations between the processing of a stimulus in the presence and absence of noise suggesting that adding noise can qualitatively alter the processing of a stimulus. Therefore, the mechanisms extracting the signal from noise common to the processing of these two attributes would not necessarily operate in noiseless conditions. For motion processing at high temporal frequencies, intra- and inter-attribute masking had the same impact suggesting common processing. In addition, simulations showed that the apparent lack of inter-attribute interaction generally observed when processing spatially superposed luminance and contrast signals can be explained by the presence of non-uniform nonlinearities intrinsic to the visual system. For motion processing at low temporal frequencies, intra- and inter-attribute masking did not have the same impact suggesting separate processing. These results can be explained by the existence of a feature tracking motion system. In conclusion, first- and second-order spatial processing would be distinct, but there is no need to infer the existence of a dedicated second-order energy-based motion system.

Keywords: Second-order, texture, motion, visual attribute, luminance-modulated, contrast-modulated, noise, masking, filter-rectify-filter model, early nonlinearity

Table des matières

Liste des tableaux

Liste des figures

Liste des animations

Liste des abréviations

cd/m^2	candela per square meter
CE	calculation efficiency
cm	centimeter
CM	contrast-modulated
cpd	cycles per degree
DAC	digital to analog converter
Hz	hertz
IEN	internal equivalent noise
LAM	linear amplifier model
LM	luminance-modulated
m	meter
MINS	main internal noise source
ms	milisecond
PTM	perceptual template model
RMS	root-mean-square
s	second
SNR	signal-to-noise ratio
TvC	threshold-versus-contrast

À Judith.

Remerciements

Je tiens sincèrement à remercier tous ceux qui m'ont permis, directement ou indirectement, de mener ce projet à terme. D'abord, mon directeur de recherche, le docteur Jocelyn Faubert, qui m'a initié à la recherche. Jocelyn, je me compte extrêmement chanceux de t'avoir eu comme directeur. Tu m'as laissé une grande liberté dans mon cheminement non-linéaire tout en étant énormément disponible. Je m'ennuierai certainement de nos nombreuses discussions également non-linéaires.

Je tiens également à remercier ma conjointe, Judith, pour sa compréhension, son encouragement et son énergie. En plus de me permettre d'accorder beaucoup de temps à la recherche, elle a su m'écouter et me soutenir durant les périodes les plus exigeantes. Merci Judith pour l'équipe que nous formons.

Un immense merci aux participants. Pour des raisons éthiques, je ne peux pas vous nommer, mais je ne vous oublie pas pour autant. Merci pour le temps et l'énergie que vous m'avez consacrés. Pour ceux que j'ai torturés un peu plus que les autres (vous vous reconnaissez), merci de votre disponibilité lorsque j'étais à la dernière minute et que j'avais besoin de données illico!

Je tiens à saluer tous les membres du Laboratoire de psychophysique et de perception visuelle que j'ai côtoyés quotidiennement au cours de ces sept dernières années. Quiconque fréquente ce laboratoire très disparate réalise que l'ambiance de travail est plaisante. Félicitations Jocelyn pour l'environnement scientifique et social que tu as créé.

Je remercie ma famille pour son encouragement. Merci spécialement à mes parents pour l'éducation et le support que vous m'avez donnés.

Enfin, je remercie le Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG) pour avoir cru en moi et m'avoir octroyé une bourse doctorale.

Chapitre I

Introduction

Les objets qui nous entourent ne réfléchissent pas tous la même quantité de lumière. L'analyse de la variation de luminance en fonction de l'espace et du temps permet au système visuel de percevoir ces objets, de les localiser et d'identifier leurs déplacements. La variation de luminance est donc un attribut essentiel à la perception visuelle. Bien que la variation de luminance permette d'identifier plusieurs objets, celle-ci n'est parfois pas suffisante nécessitant d'autres attributs que la luminance pour percevoir certains objets. Par exemple, un objet rouge présenté sur un fond vert sera clairement perceptible même si le rouge et le vert ont la même luminance. La couleur peut donc nous permettre de distinguer un objet de son arrière-fond. De plus, la texture, c'est-à-dire un attribut de deuxième ordre, peut également permettre de distinguer des objets ayant la même luminance moyenne et la même couleur. La Figure I-1 en présente un exemple. La feuille verte du haut est clairement distinguable de l'arrière-fond composé d'un ciel bleu lumineux. Pour cette feuille, la variation de luminance ainsi que la variation de couleur permettent de ségréger l'objet de son arrière-fond. La feuille plus bas et son arrière-plan ont des luminances et des couleurs moyennes similaires. Seule la texture permet de distinguer cet objet de son arrière-plan. En effet, l'arrière-plan (feuilles au loin) est défini par de plus hautes fréquences spatiales (plus de feuilles pour la même surface) que la feuille en avant-plan.

Figure I-1. Exemple d'attributs de premier et deuxième ordres dans un milieu écologique. La luminance et la couleur permettent de distinguer la feuille du haut de son arrière-plan (ciel bleu clair). Cependant, la luminance et la couleur de la feuille plus bas sont similaires à l'arrière-plan (d'autres feuilles au loin). L'information de deuxième ordre (texture) permet donc de ségréger cet objet de son arrière-plan.

Traitement de premier ordre

Au niveau de l'entrée corticale, l'information visuelle est largement représentée sous forme d'activation de champs récepteurs ayant une forme similaire à des sinus variant en fréquences spatiotemporelles et en orientations. En effet, suite à des études en électrophysiologie, Hubel et Wiesel (1959) ont observé que le stimulus maximisant l'activation de la plupart des neurones du cortex visuel (stimulus préféré) correspondait à des barres variant en luminance et ayant une orientation précise pour un endroit spécifique du champ visuel. Hubel et Weisel ont nommé les cellules répondant maximalement à des barres statiques « simples » et les cellules préférant des barres en translation « complexes ». L'activation de chacun de ces récepteurs est corrélée au contraste de son stimulus préféré minimisant la différence entre le stimulus préféré et l'image projetée sur la rétine. Pour chaque région du champ visuel, le traitement du système visuel décompose donc l'information en différentes fréquences et orientations.

Le stimulus le plus fréquemment utilisé en psychophysique est le sinus (Figure I-2) puisque qu'il correspond au stimulus préféré de plusieurs neurones du cortex visuel. Par exemple, ce stimulus est souvent utilisé pour mesurer la sensibilité au contraste, c'est-à-dire la plus petite variation de luminance détectable. Pour le système visuel, la détection d'un stimulus défini par la luminance peut être effectuée d'une façon computationnellement simple. En effet, l'activation d'un neurone sera plus élevée lors de la présentation d'un sinus correspondant à son stimulus préféré que lorsque le sinus n'est pas présenté.

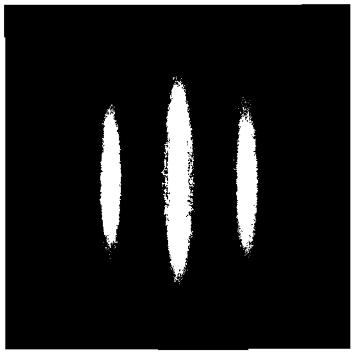


Figure I-2. Gabor : correspond à un sinus modulant la luminance perçue par une ouverture Gaussienne.

Le traitement de signaux définis par la luminance (stimuli de premier ordre) est donc relativement simple pour le système visuel. Mais comment percevons-nous un stimulus défini par un autre attribut que la luminance (stimulus de deuxième ordre (Cavanagh & Mather, 1989; Chubb & Sperling, 1988))? La Figure I-3 présente à gauche un signal défini par la luminance (premier ordre) et à droite, un signal défini par le contraste (le stimulus de deuxième ordre le plus fréquemment utilisé en psychophysique). Notez qu'un stimulus de deuxième ordre est également défini par la luminance puisqu'il est composé d'éléments plus ou moins lumineux. Par contre, la forme du signal (le sinus) n'est pas directement définie par la luminance puisque la luminance moyenne de chaque bande est la même. La forme d'un signal de deuxième ordre est plutôt définie par un attribut dérivé de la luminance, dans notre cas, le contraste.

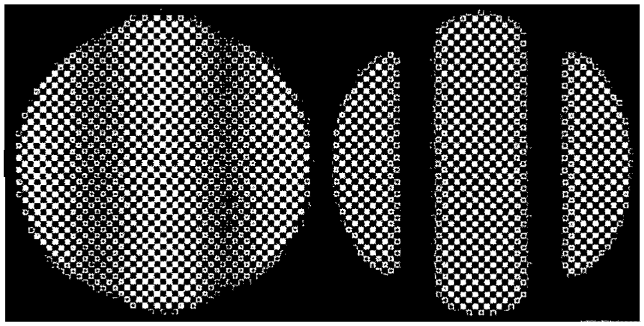


Figure I-3. Signal défini par la luminance (gauche, premier ordre) et le contraste (droite, deuxième ordre).

Traitement de deuxième ordre

Stimuli statiques

Il est évident que dans plusieurs cas, il nous est possible de percevoir la forme d'un signal défini par un attribut de deuxième ordre. Par exemple, la forme du sinus défini par le contraste présenté dans la Figure I-3 est clairement perceptible. La question consiste à savoir comment un signal défini par un attribut de deuxième ordre est perçu. En bref, deux modèles peuvent être proposées. Premièrement, puisqu'un stimulus de deuxième ordre correspond à la variation d'un certain attribut (ex : le contraste), un mécanisme spécifique au traitement de deuxième ordre pourrait être appliqué en deux temps. Un premier mécanisme opérant au niveau local pourrait évaluer un certain attribut (ex : le contraste), suivi d'un mécanisme global évaluant la variation de cet attribut (ex : filtre-rectification-filtre (Wilson, Ferrera & Yo, 1992)). Notez que pour un stimulus de deuxième ordre défini par le contraste, le premier traitement proposé par ce modèle est équivalent à un traitement de premier ordre consistant à évaluer le contraste. Cependant, pour des signaux définis par la luminance et le contraste ayant la même fréquence spatiale, le traitement de premier ordre consisterait à évaluer le contraste à la fréquence du signal alors que la première étape du traitement de deuxième ordre consisterait à évaluer le contraste local, c'est-à-dire à plus hautes fréquences spatiales. Selon un tel modèle, les traitements de premier et deuxième ordres seraient donc distincts et il existerait des mécanismes traitant explicitement les stimuli de deuxième ordre.

Le deuxième modèle serait basé sur l'existence de non-linéarités intrinsèques au système visuel. Ces non-linéarités introduiraient une variation de luminance à un stimulus n'en contenant pas. Pour un signal défini par le contraste, une non-linéarité pourrait, par exemple, diminuer la luminance moyenne d'une région à haut contraste par rapport à une région à bas contraste introduisant ainsi une différence de luminance. Une telle distorsion permettrait au système de luminance de détecter une variation de contraste. Selon un tel modèle, les deux attributs seraient traités par des mécanismes communs.

Il existe donc des modèles proposant que les attributs de premier et deuxième ordres seraient traités par des mécanismes communs alors que d'autres proposent plutôt qu'ils seraient traités par des mécanismes distincts. Savoir si ces attributs sont traités par des mécanismes communs ou distincts aiderait donc à comprendre comment les attributs de deuxième ordre sont traités. Conséquemment, bien que l'objectif ultime de la présente thèse fût de comprendre comment les attributs de deuxième ordre sont traités, le sous-objectif directement étudié était de déterminer si un signal défini par le contraste (le stimulus de deuxième ordre le plus souvent utilisé) est traité par les mêmes mécanismes qu'un signal défini par la luminance (premier ordre). En d'autres termes, existe-il des mécanismes dédiés au traitement de deuxième ordre ou est-ce que les attributs de deuxième ordre sont traités par des mécanismes sensibles aux stimuli de premier ordre?

Mouvement

La même problématique, c'est-à-dire de déterminer si les attributs de premier et deuxième ordres sont traités par des mécanismes distincts ou communs, se pose également pour le mouvement. Cette problématique en lien avec le mouvement est intéressante puisque le cerveau traite l'information visuelle de façon relativement modulaire. En effet, il est généralement reconnu que différentes propriétés (ex : couleur, forme, mouvement) sont traitées en parallèle de façon relativement indépendante. Puisque le mouvement et l'extraction d'une forme définie par un autre attribut que la luminance seraient traités par des aires corticales distinctes, il devient pertinent d'étudier le traitement d'un stimulus nécessitant théoriquement le traitement de deux propriétés : le mouvement et la forme définie par un attribut de deuxième ordre. Plus précisément, la question est de déterminer s'il existe un système de mouvement dédié uniquement au traitement de deuxième ordre (contraste).

Masquage d'un signal par du bruit

La méthodologie adoptée dans la présente thèse pour déterminer si deux stimuli sont traités par des mécanismes communs ou distincts est basée sur le masquage d'un signal par du bruit. L'utilisation de bruit défini par la luminance (ex : Figure I-4 gauche) est largement utilisée en psychophysique. De façon analogue, nous avons conçu un bruit défini par le contraste (Chapitre II, Figure I-4 droite).

Du bruit non corrélé, c'est-à-dire du bruit blanc, contient de l'énergie à toutes les fréquences et peut donc masquer plusieurs stimuli sans être particulièrement sélectif à un stimulus. Dans le but d'affecter le traitement d'un signal en particulier, chacun de ces bruits a été filtré pour ne garder que les fréquences spatiotemporelles près du signal. Le bruit défini par la luminance était donc conçu pour maximiser le masquage sur le traitement du signal défini par la luminance et le bruit défini par le contraste était conçu pour maximiser le masquage sur le traitement du signal défini par le contraste. Dans la présente thèse, le masquage du bruit sur le traitement d'un signal a été évalué dans des conditions intra-attribut (masquage d'un signal de luminance par du bruit de luminance et masquage d'un signal de contraste par du bruit de contraste) et des conditions inter-attribut (masquage d'un signal de luminance par du bruit de contraste et masquage d'un signal de contraste par du bruit de luminance).

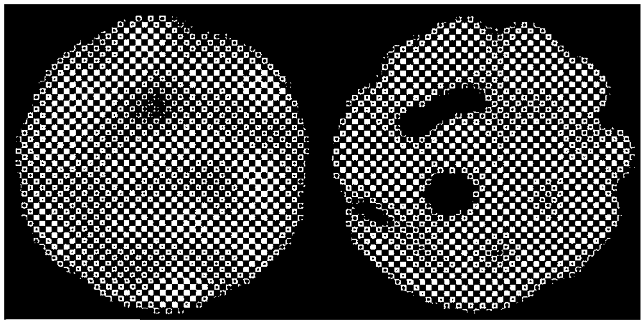


Figure I-4. Bruit défini par la luminance (gauche) et le contraste (droite). Pour le bruit défini par la luminance, certaines régions sont plus sombres que d'autres. Par contre, pour le bruit défini par le contraste, la luminance locale moyenne est constante, seul le contraste local varie d'une région à l'autre.

Masquage intra-attribut

Bien qu'il soit évident que l'ajout de bruit à la même fréquence spatiotemporelle que le signal et défini par le même attribut que le signal masque forcément le signal (à hautes amplitudes), le masquage intra-attribut demeure pertinent puisqu'il permet de déterminer le ratio signal-bruit nécessaire à la perception du signal. De façon générale, lorsque le bruit est défini par le même attribut que le signal, l'amplitude du signal nécessaire pour percevoir le signal est proportionnel à l'amplitude du bruit (Pelli, 1981, 1990; Figure I-5). En d'autres termes, si le contraste du bruit est doublé, le contraste nécessaire pour percevoir le signal devra également être doublé. L'efficacité de calcul du système visuel peut être mesurée par rapport au ratio signal-bruit nécessaire à la perception d'un signal.

Avant d'être perçu, un signal doit être acheminé de l'œil au mécanisme traitant ce signal. Ce cheminement est évidemment sous-optimal et le signal est altéré par le système visuel. Ces altérations représentent le bruit interne déformant le signal pouvant potentiellement limiter la sensibilité à un stimulus. En connaissant le ratio signal-bruit nécessaire pour extraire le signal (obtenu en présence de bruit externe) et en supposant que la tâche consistant à extraire un signal du bruit externe est la même que celle d'extraire un signal dans du bruit interne, il est possible de déduire l'impact relatif du bruit interne. En effet, si le bruit interne est plus important que le bruit externe, alors le bruit externe ne devrait pas avoir d'impact significatif sur la performance du sujet. Par contre, si le bruit externe est plus important que le bruit interne, alors le bruit interne devrait avoir un impact sur la performance du sujet. En évaluant un seuil de contraste en fonction du contraste du bruit externe (Figure I-5), la sensibilité à un signal peut donc être décomposée en deux paramètres : le ratio signal-bruit nécessaire à la perception d'un signal et l'impact du bruit interne.

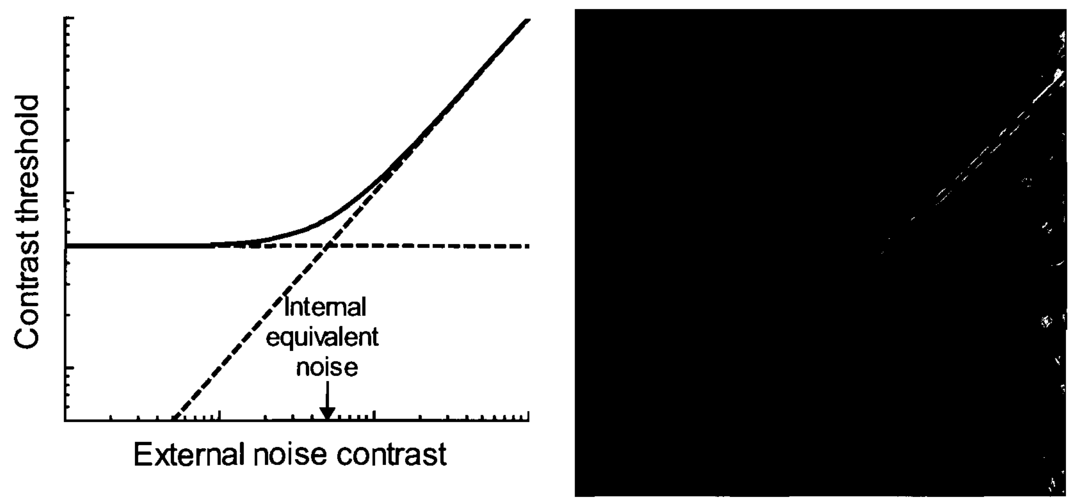


Figure I-5. Seuil de contraste en fonction du contraste du bruit externe. En présence de beaucoup de bruit externe, le bruit interne n'a pas d'impact significatif et la performance ne dépend que du bruit externe (pente de 1 sur le graphe). Par contre, lorsque le bruit externe est plus faible que le bruit interne, alors la performance ne dépend que du bruit interne (pente de 0 sur le graphe). L'impact relatif du bruit interne correspond donc à la jonction de ces deux droites.

Si le même mécanisme extrait le signal du bruit pour différents attributs, le même ratio signal-bruit devrait être nécessaire pour extraire le signal dans les deux cas. Notez que même si un attribut est moins perceptible qu'un autre, le bruit de cet attribut devrait également être moins perceptible. La perceptibilité d'un attribut n'affecte donc pas le ratio signal-bruit nécessaire à la perception du signal. Conséquemment, si, pour différents attributs, le même ratio signal-bruit est nécessaire pour percevoir un signal (lorsque le signal et le bruit sont définis par le même attribut), cela suggèrerait que l'extraction du signal dans le bruit est effectué par les mêmes mécanismes. Par contre, si des mécanismes différents extraits le signal du bruit, alors il ne devrait pas être surprenant d'observer des ratios signal-bruit différents.

Masquage inter-attribut

Si deux stimuli sont traités par des mécanismes distincts, alors il devrait être possible de dissocier leurs traitements. Plus précisément, si une certaine manipulation affecte le traitement du stimulus A sans affecter le traitement du stimulus B, et que vice versa, une autre manipulation affecte le traitement du stimulus B sans affecter le traitement du stimulus A (double dissociation), ceci suggèrerait que les stimuli A et B sont traités, à au moins un niveau, par des mécanismes distincts.

En utilisant des bruits de luminance et de contraste conçus pour sélectivement masquer le traitement d'un attribut, ces bruits peuvent être utilisés pour tenter de dissocier les traitements de signaux définis par la luminance et le contraste. Si le bruit de luminance affecte le traitement de luminance sans affecter le traitement de contraste et que vice versa, le bruit de contraste affecte le traitement de contraste sans affecter le traitement de luminance, alors cela impliquerait que ces deux attributs sont traités, à au moins un niveau, par des mécanismes distincts. Par contre, si chaque bruit affecte le traitement des deux attributs dans des proportions similaires, alors cette indissociabilité suggèrerait plutôt que ces attributs sont traités par des mécanismes communs.

Structure de la thèse

Le corps de la présente thèse se compose d'un chapitre d'introduction (Chapitre I) suivi de deux sections présentant les différents articles et se termine par un chapitre de conclusion (Chapitre VII).

La première section est composée de trois chapitres portant sur la détection de signaux statiques définis par la luminance et le contraste. Le premier chapitre de cette section (Chapitre II) est consacré au masquage intra-attribut, c'est-à-dire au masquage d'un bruit de luminance sur la détection d'un signal de luminance et au masquage d'un bruit de contraste sur la détection d'un signal de contraste. Le but du second chapitre (Chapitre III) était d'évaluer le masquage inter-attribut, c'est-à-dire le masquage d'un bruit de luminance sur la détection d'un signal de contraste et le masquage d'un bruit de contraste sur la détection d'un signal de luminance. Enfin, basé sur une recension des écrits, le Chapitre IV évalue l'hypothèse que l'ajout de bruit modifie qualitativement le traitement d'un stimulus.

La seconde section est composée de deux chapitres portant sur le traitement du mouvement défini par la luminance et le contraste. Le but du Chapitre V était d'évaluer le masquage intra- et inter-attribut en fonction de la fréquence temporelle du signal. Le deuxième chapitre de cette section (Chapitre VI) propose un modèle dans lequel des non-linéarités intrinsèques au système visuel permettraient, pour les hautes fréquences temporelles, au système traitant le mouvement défini par la luminance de traiter également le mouvement défini par le contraste.

Notez que pour les stimuli statiques (Section 1), nous avons utilisé une carte graphique particulière permettant au montage expérimentale d'afficher 1024 niveaux de gris. Cette carte graphique particulière était nécessaire puisque que la plupart des ordinateurs peuvent afficher seulement 256 niveaux de gris, ce qui, dans bien des cas, n'est pas suffisant pour mesurer la plus petite variation de luminance nécessaire pour percevoir un stimulus. Par contre, lors de l'utilisation de stimuli dynamiques (Section 2), 1024 niveaux de gris n'étaient pas suffisants. Nous avons donc développé une méthode permettant à un montage expérimental affichant seulement 256 niveaux de gris d'être équivalent à un montage expérimental affichant un nombre infini de niveaux de gris. Cette méthode consistant à ajouter une faible quantité de bruit au stimulus a fait l'objet d'un article présenté à l'Annexe I.

Contributions des auteurs

Les cinq chapitres représentant le corps de la présente thèse ont été rédigés sous forme d'articles. Pour ces cinq articles ainsi que pour l'article présenté à l'Annexe I, je suis le premier auteur et mon directeur de recherche, le docteur Jocelyn Faubert, est le seul coauteur. Pour chaque article, j'ai développé le protocole expérimental, les programmes informatiques générant les stimuli, effectué la collecte de données, analysé les résultats et rédigé l'article. Toutes ces étapes ont été effectuées sous la supervision du docteur Jocelyn Faubert avec qui j'ai eu des rencontres sur une base régulière. Ces six articles ont été reproduits dans la présente thèse avec l'accord du coauteur (Annexe II) et des éditeurs (Annexe III).

Bibliographie

Cavanagh, P., & Mather, G. (1989). Motion: The long and short of it. *Spat Vis, 4*(2-3), 103-129.

Chubb, C., & Sperling, G. (1988). Drift-balanced random stimuli: A general basis for studying non-fourier motion perception. *J Opt Soc Am A, 5*(11), 1986-2007.

Hubel, D.H., & Wiesel, T.N. (1959). Receptive fields of single neurones in the cat's striate cortex. *Journal of Physiology, 148*, 574-591

Pelli, D. G. (1981). *The effects of visual noise.* Cambridge University, Cambridge.

Pelli, D. G. (1990). The quantum efficiency of vision. In C. Blakemore (Ed.), *Visual coding and efficiency*. Cambridge: Cambridge University Press.

Wilson, H. R., Ferrera, V. P., & Yo, C. (1992). A psychophysically motivated model for two-dimensional motion perception. *Vis Neurosci, 9*(1), 79-97

Section 1

Détection de stimuli statiques définis par la luminance et le contraste

Chapitre II

Same calculation efficiency but different internal noise for luminance- and contrast-modulated stimuli detection

Rémy Allard et Jocelyn Faubert

Laboratoire de psychophysique et perception visuelle

École d'optométrie, Université de Montréal

Le contenu de ce chapitre a précédemment été publié sous forme d'article (©ARVO) :

Allard, R., & Faubert, J. (2006). Same calculation efficiency but different internal noise for luminance- and contrast-modulated stimuli detection. *Journal of Vision, 6* (4), 322-334.

Abstract

There is no consensus on whether luminance- (LM) and contrast-modulated (CM) stimuli are processed by common or separate mechanisms. To investigate this, the sensitivity variations to these stimuli are generally compared as a function of different parameters (ex: sensitivity as a function of the spatial or temporal window sizes) and similar properties have been observed. The present study targets the sensitivity difference between LM and CM stimuli processing. Therefore, instead of studying the variation of sensitivity in different conditions, we propose to decompose the sensitivities in internal equivalent noise (IEN) and calculation efficiency (CE) to evaluate at which processing level the two mechanisms differ. For each stimulus type, the IEN and CE of four observers were evaluated using three different carriers (plaid, checkerboard and binary noise). No significant CE differences were noted in all six conditions (3 carriers x 2 modulation types), but important differences were found between the IEN of the two stimulus types. These data support the hypothesis that the two pathways are initially separate and that the two stimuli may be treated by common mechanisms at a later processing stage. Based on ideal observer analysis, pre-rectification internal noise could explain the difference of IEN between LM and CM stimuli detection when using binary noise as a carrier but not when using a plaid or a checkerboard. We conclude that a sub-optimal rectification process causes higher IEN for CM stimuli detection compared with LM stimuli detection and that the intrinsic noise of the binary carrier had a greater impact on the IEN than the sub-optimal rectification.

Keywords: Contrast, luminance, first-order, second-order, texture, filter-rectify-filter, sensitivity, calculation efficiency, internal equivalent noise

Introduction

Human observers are sensitive to both luminance- (LM) and contrast-modulated (CM) stimuli. In the present study, we define LM stimuli as the addition of an envelope (signal) with a carrier (texture) and CM stimuli as their multiplication. Consequently, for LM stimuli, the local luminance average varies throughout the stimulus according to the envelope while the local contrast remains constant (Figure II-1 left). For CM stimuli, the local luminance average remains constant and the local contrast varies throughout the stimulus according to the envelope (Figure II-1 right). Therefore, since a Fourier transform can directly detect the signal frequency of LM stimuli, this type of stimulus is typically characterized as Fourier, first-order or linear. However, CM stimuli are not considered as Fourier stimuli since the signal frequency is not present in the Fourier domain. Therefore, CM stimuli are characterized to be non-Fourier, second-order or non-linear stimuli (Cavanagh & Mather, 1989; Chubb & Sperling, 1988).

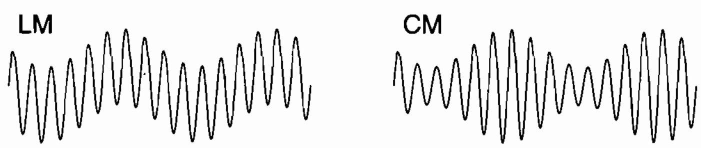


Figure II-1. Luminance profile of luminance- and contrast-modulated stimuli. Luminance profiles presented as a function of space (one-dimension).

Non-linear processing

Wilson, Ferrera and Yo (1992) proposed a 2-stage model for the detection of LM and CM motion. This model may be summarized as follows: both stimuli are initially processed by V1. For LM stimuli, this contrast detection corresponds to the signal or envelope itself and the stimuli do not require anymore processing before MT processes the perceived motion. For CM stimuli, the contrast detection occurs at higher spatial frequencies corresponding to the carrier, therefore the treatment uses another path and the information passes trough V2 for a second-order rectification process before attaining MT.

After a second-order rectification process, a CM stimulus becomes similar to a LM stimulus (Chubb & Sperling, 1988; Solomon & Sperling, 1994; Sperling, Chubb, Solomon, & Lu, 1994). Therefore, the two stimuli could be merged and then treated by the same mechanisms (Baker, 1999). Therefore, the processing of these two stimulus types is initially separated but may be common at a later stage. More recently, many similar models have been developed and are typically referred to as filter-rectify-filter models (Clifford & Vaina, 1999; Nishida & Sato, 1995; Prins & Kingdom, 2003). Figure II-2 shows an example of such a model where an extra process is required for CM stimuli processing. Although this class of model seems to be more popular, other models have also been developed. Some motion models propose that both LM and CM stimuli are treated by common mechanisms (Benton & Johnston, 2001; Johnston & Clifford, 1995a, 1995b; Johnston, McOwan, & Buxton, 1992; Taub, Victor, & Conte, 1997). Consequently, in motion perception, the idea that LM and CM stimuli are processed by common mechanisms is still largely debated. This debate has carried over in spatial vision (which is the object of the present study) where the processing of static LM and CM stimuli has been compared.

Evidence for separate mechanisms

Evidence from spatial vision studies suggests that LM and CM stimuli are, at least initially, processed by separate mechanisms. Nishida, Ledgeway and Edwards (1997) found that after adapting to one type of stimulus (LM or CM) the sensitivity to the same type of stimulus was affected, but not the sensitivity to the other. Schofield and Georgeson (1999) did not find any inter-type sub-threshold summation while intra-type sub-threshold summation was found. The same authors showed strong evidence suggesting that LM and CM stimuli are not merged after a second-order rectification process (Georgeson & Schofield, 2002). They first showed that the recognition of the stimulus type (LM vs CM) was almost as good as the detection of each type (LM vs noise or CM vs noise). They also demonstrated that observers do not confuse the two stimulus types when they are combined since the recognition of the two stimuli combined in-phase or out-phase (LM+CM vs LM-CM) is similar to their detection (LM+CM vs noise or LM-CM vs noise).

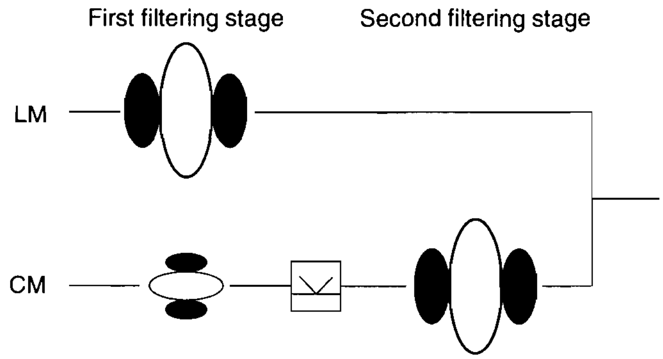


Figure II-2. Filter-rectify-filter model. For contrast-modulated stimuli (CM) the first filtering stage processes the carrier so that, after a rectification process, the second stage can detect the envelope. For luminance-modulated stimuli (LM) no rectification is required.

Evidence for common mechanisms

Although important evidence suggests that LM and CM stimuli are processed by separate mechanisms, there is still evidence suggesting the opposite. First, the processing of LM and CM stimuli by human observers share similar properties. Schofield and Georgeson compared the sensitivity to these stimuli as a function of stimulus size (Schofield & Georgeson, 1999) and presentation time (Schofield & Georgeson, 2000). They found similar spatial and temporal integration for both stimulus types. In their experiments, the sensitivity curves of LM and CM stimuli as a function of a spatial or a temporal window were parallel. These results indicate a similar spatiotemporal integration for both modulation types even though the sensitivity was greater for the LM stimulus. Similar behaviors generally suggest that both stimuli could be processed, at least partially, by common mechanisms. However, the implications of these findings do not necessarily lead to such conclusions. Two separate mechanisms could have similar behaviors.

In another study, Georgeson and Schofield (2002) showed direct evidence of an interaction between the processing of LM and CM stimuli. They found that, after adaptation to a given stimulus type (LM or CM), the perceived contrast of the other stimulus type was reduced almost by the same proportion as the one of the same type.

Purpose of the present study

To investigate if LM and CM stimuli are processed by common or separate mechanisms, the sensitivity variations to these stimuli are generally compared as a function of different parameters. As mentioned above, similar function shapes have been observed in certain conditions, which suggests that the main difference between the two is their sensitivity. Therefore, we propose to decompose the sensitivity to evaluate at which processing level the two mechanisms differ.

Based on the assumptions that the internal noise and calculation are contrast-invariant (see bellow), the sensitivity may be separated into two parameters (Legge, Kersten, & Burgess, 1987; Pelli, 1981, 1990): internal equivalent noise (IEN) and calculation efficiency (CE). The goal of the present study was to elucidate if LM and CM sensitivities differ because of a difference of IEN, CE, or both.

Evaluating the IEN and CE

The internal noise is the signal deterioration introduced by different processing levels that limit the observer's sensitivity (ex: optical noise caused by eye imperfections, photon-noise, neuronal noise...). The calculation is the observer's ability to detect a given signal embedded in noise. Both the internal noise and calculation limit the sensitivity of a noise-free signal. Once the internal noise is added to the stimulus, the observer's task consists in detecting the signal embedded in noise.

Contrast-invariant calculation signifies that the observer's performance only depends on the signal-to-noise ratio (SNR) (Pelli, 1981, 1990). Therefore, equally modifying both the signal contrast (c) and effective noise contrast (N_{eff}) will not affect the observer's performance. The effective noise represents the combination of the internal (observer) and external (stimulus) noise.

Assuming that the observer's calculation is contrast-invariant, the minimum SNR necessary to detect the signal based on a given threshold criterion would be constant:

$$k = \frac{c}{N_{eff}}. \qquad (1)$$

The CE may be defined as the minimum SNR necessary to detect the signal of an ideal observer relative to the observer's SNR:

$$Calculation\ efficiency = \frac{k_{ideal}}{k}, \qquad (2)$$

where k_{ideal} is the k parameter for the ideal observer. An ideal observer is a theoretical observer using all the information available to optimally perform the task. Therefore, k_{ideal} represents the smallest SNR (k) mathematically possible to detect the signal (c) based on a given threshold criterion.

Assuming that the internal noise is also contrast-invariant, the impact of the internal noise will be constant as a function of the signal and external noise contrast (Pelli, 1981, 1990). The root-mean-square (rms) contrast of the effective noise (N_{eff}) will be:

$$N_{eff} = \sqrt{N^2 + N_{eq}^2}, \qquad (3)$$

where N and N_{eq} represent the rms contrasts of the external noise and IEN respectively. The IEN models the impact of the internal noise on the sensitivity. Note that if the external noise contrast is equal to the IEN (N=N_{eq}), the effective noise (N_{eff}) will be √2 times greater. Therefore, assuming that the calculation is contrast-invariant (Equation 1), the IEN will be equal to the external noise contrast that raises the signal contrast threshold (c) by a factor of $\sqrt{2}$.

Based on the two assumptions that the internal noise and calculation are contrast-invariant, the function between the signal contrast (c) threshold and the external noise contrast (N) can be deduce by combining Equations 1 and 3:

$$c = k\sqrt{N^2 + N_{eq}^{\ 2}}\,, \qquad (4)$$

and the relation between the squared signal contrast (c^2) and the external noise variance (N^2) would be linear:

$$c^2 = k^2\left(N^2 + N_{eq}^{\ 2}\right). \quad (5)$$

Such linear relation for the detection grating in gaussian white noise has been found in the past (Pelli, 1981), which supports the hypothesis that the internal noise and calculation are contrast-invariant. For a review on the relation between external noise and detection threshold, which permits to decompose the sensitivity in IEN and CE, refer to Legge, Kersten, and Burgess (1987) and Pelli (1981, 1990).

When the external noise contrast is relatively small compared with the IEN ($N \ll N_{eq}$), varying the external noise does not affect significantly the effective noise (Equation 3) and, therefore, the signal contrast threshold is relatively constant as a function of the external noise. However, in high external noise conditions, varying the external noise has a great impact on the total amount of noise and thereby the signal contrast threshold increases as a function of the external noise.

Given that the processing of two stimuli differs in IEN but not in CE, then in high external noise conditions, the internal noise should not be significant and no important threshold difference should be observed. The left graph of Figure II-3 illustrates this hypothesis. However, if the IENs are equal and the CEs differ, the detection thresholds should be different in all external noise conditions. Indeed, the two curves would be parallel (Figure II-3 right).

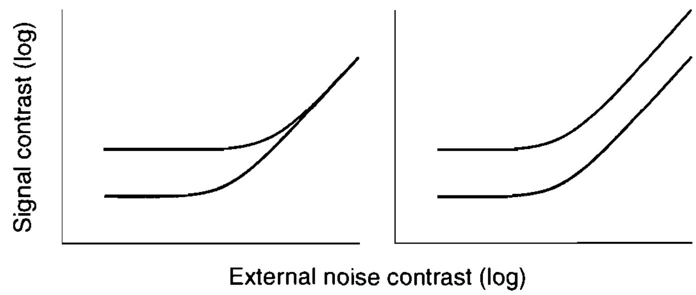


Figure II-3. Detection threshold patterns as a function of external noise contrast. On the left, both patterns have the same calculation efficiencies but the red pattern shows less internal equivalent noises. The graph on the right shows the opposite; both patterns have the same internal equivalent noises but the red pattern has a greater calculation efficiency.

Predictions based on separate mechanisms

If LM and CM stimuli are processed by separate mechanisms, the predictions are straightforward: each stimulus pathway should have its own IEN and CE. Consequently, the probability of having the same IEN or CE for both stimuli would be low. Indeed, the observer's ability to detect a LM signal in LM noise would probably differ from its ability to detect a CM signal in CM noise. In other words, in high external noise conditions, which result in non-significant internal noise, the sensitivity to LM and CM stimuli would probably differ.

Predictions based on common mechanisms

According to the 2-stage (or filter-rectify-filter) model, the difference between LM and CM stimuli processing is that CM stimuli detection requires an extra second-order rectification process. Consequently, deriving its predictions is more complex since the impact that a sub-optimal rectification would have on IEN and CE must be determine.

The CM stimulus may be defined as the multiplication of a modulation ($M(x,y)$) with a texture ($T(x,y)$), where the modulation is defined by lower spatial frequencies relative to the texture, and the texture local rms contrast (T_{rms}) is constant throughout the stimulus. Therefore, the modulation represents the global texture contrast variation. The rms contrast near the position (x,y) is equal to $T_{rms}M(x,y)$. The rectification consists in estimating the local (carrier spatial frequency) rms contrast. This estimation should be applied locally over the entire stimulus (i.e. for all (x,y) positions), which would reconstruct a similar modulation ($T_{rms}M(x,y)$) as the one defining the stimulus ($M(x,y)$).

Consequently, after a rectification, a CM stimulus is converted into an effective stimulus ($T_{rms}M(x,y)$) similar to a LM stimulus without a texture, i.e. the modulation $M(x,y)$. In the LM stimulus, each position represents the luminance intensity. For the rectified CM stimulus, each position of the effective stimulus would represent the local contrast modulation of the CM stimulus. Therefore, after a rectification, a CM stimulus would be analogue to a LM stimulus and both could be treated by common mechanisms.

Since a rectification is likely to be sub-optimal, lets represent the estimation of the local contrast ($T_{rms}M(x,y)$) by a normal distribution centered at $\beta T_{rms}M(x,y)$ and with a standard deviation of N_{rect}. β represents the gain parameter affecting the strength of the rectification output. N_{rect} represents internal noise that could be added during the rectification process.

Suppose a CM stimulus ($M(x,y)T(x,y)$) with a contrast modulation ($M(x,y)$) composed of a signal ($S(x,y)$) with contrast S_{in} embedded in noise ($N(x,y)$) with contrast N_{in}. Using the previously defined rectification, the signal and noise contrast at the output of the rectification would be scaled by a factor of βT_{rms} and noise would also be added (N_{rect}). Consequently, the signal (S_{in}) and noise (N_{in}) contrast of the input of the rectification process would result, after the rectification process, as:

$$S_{out} = \beta T_{rms} S_{in}, \qquad (6)$$

$$N_{out} = \sqrt{(\beta T_{rms} N_{in})^2 + N_{rect}^2}\ , (7)$$

and the SNR would pass from S_{in}/N_{in} to

$$\frac{S_{out}}{N_{out}} = \frac{\beta T_{rms} S_{in}}{\sqrt{(\beta T_{rms} N_{in})^2 + N_{rect}^2}}$$
$$\frac{S_{out}}{N_{out}} = \frac{S_{in}}{\sqrt{N_{in}^2 + (N_{rect}/\beta T_{rms})^2}} \quad . \quad (8)$$

By defining

$$N'_{rect} = N_{rect}/\beta T_{rms}\ , \quad (9)$$

we obtain:

$$\frac{S_{out}}{N_{out}} = \frac{S_{in}}{\sqrt{N_{in}^2 + N_{rect}^{'2}}}\ . (10)$$

Therefore, based on this type of rectification and on the 2-stage model, a sub-optimal rectification would increase the IEN and would not affect the CE. In other words, if the strength of the input noise is relatively low compared with the noise added by the rectification ($N_{in} << N'_{rect}$), then the impact of the input noise would not be significant:

$$N_{out} \approx N'_{rect}\ , \quad (11)$$

and the rectification process would decrease the SNR:

$$\frac{S_{out}}{N_{out}} \approx \frac{S_{in}}{N'_{rect}}$$
$$\frac{S_{out}}{N_{out}} << \frac{S_{in}}{N_{in}} \quad . \ (12)$$

Consequently, in low external noise conditions (small N_{in}) the rectification process would affect the observer's performance. However, if the input CM noise is relatively high ($N_{in} >> N'_{rect}$), then the noise added by the rectification (N'_{rect}) would not be significant:

$$N_{out} \approx N_{in}, \quad (13)$$

and the rectification would not affect the SNR:

$$\frac{S_{out}}{N_{out}} \approx \frac{S_{in}}{N_{in}}. \quad (14)$$

Therefore, if the SNRs of a LM and a CM stimulus are equal, then, in high noise conditions, the rectification should not affect the SNR of the CM stimulus. If both stimuli are treated by common post-rectification mechanisms then the same sensitivity should be observed in high external noise conditions (negligible internal noise) for both stimulus types, which would result in similar CEs.

Methods

Subjects

Four volunteers aged between 26 and 35 years of age participated in the study. Their vision was normal or corrected to normal.

Stimuli

To compare the subjects' performance on the detection of LM and CM stimuli, the luminance average and contrast average over the entire stimulus were the same for both stimulus types. The only difference between the two stimuli was that, for the LM stimulus, the modulation (signal + external noise) was applied to the luminance profile while the contrast remained constant throughout the stimulus and vice versa for the CM stimulus. Mathematically, the luminance of the pixel at position (x,y) for the LM stimuli was defined by the addition of a texture to a luminance modulation:

$$L_{LM}(x,y) = L_0(M(x,y) + T(x,y)), (15)$$

where L_0 is the stimulus luminance average, which was fixed to 59 cd/m^2 for the present study. $M(x,y)$ and $T(x,y)$ represent, respectively, the modulation and the texture of the pixel at position (x,y). The texture was added to the LM stimulus to give both stimuli the same contrast average. Therefore, the two stimulus types were similar with the exception that for the CM stimulus, the modulation was applied to the texture instead of the luminance:

$$L_{CM}(x,y) = L_0(1 + M(x,y)T(x,y)). \qquad (16)$$

Since the modulation and the texture should not affect the stimulus luminance average, the average of $M(x,y)$ and $T(x,y)$ over the entire stimulus (all x and y) must be 1 and 0 respectively.

The present study had the objective of comparing the IEN and CE for LM and CM stimuli sensitivity. To derive the IEN and CE, the signal threshold must be evaluated in different external noise conditions. Indeed, the modulation profile ($M(x,y)$) was composed of a signal ($S(x,y)$), which the subject had to detect, embedded in external noise ($N(x,y)$):

$$M(x,y) = 1 + S(x,y) + N(x,y). \qquad (17)$$

Using this stimulus definition implies that the signal and the external noise are both of the same modulation type. Therefore, the LM and CM IENs were not of the same modulation type. This should remain in mind when comparing the results of the IEN for the LM and CM stimuli detection. Indeed, measuring the IEN actually measures the impact of the internal noise on the task being accomplished, in our case the detection of the signal. However, having the same modulation type for the signal and external noise enables a direct comparison between the LM and CM CEs. Indeed, the CE is the efficiency of detecting the signal embedded in noise. Therefore, the capacity of extracting a LM signal embedded in LM noise can be directly compared with the capacity of extracting a CM signal embedded in CM noise.

Signal

The signal was a Gabor patch (Figure II-4 left), vertically oriented, modulating either the luminance or contrast profile of the stimulus (Equation 18, Figure II-7). Since the CM stimulus requires a texture defined by high spatial frequencies relative to the signal, the spatial frequency (f) of the Gabor patch was set to a low spatial frequency of 1 cycle per degree (cpd). The phase (p) of the sine wave was randomly set at each stimulus presentation. The standard deviation (σ) of the Gabor patch was set to 1 degree of visual angle.

$$S(x,y) = c\sin(fx + p)\exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right), \qquad (18)$$

c represented the contrast of the signal, which corresponds to the Michelson contrast once the signal ($S(x,y)$) is integrated in the modulation ($M(x,y)$, Equation 17). The contrast (c) was the dependant variable.

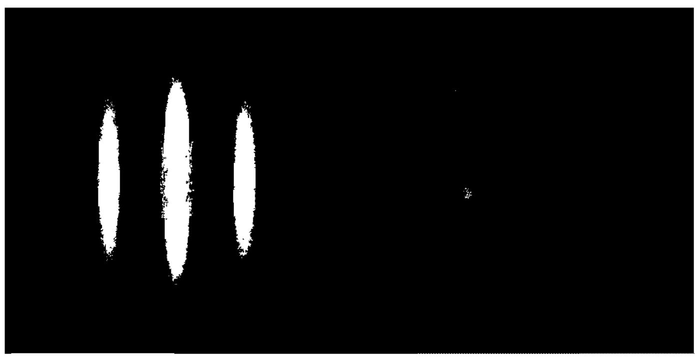


Figure II-4. Gabor patch signal or envelope (left). Gaussian filtered noise (right).

Carriers

Since the present study evaluates the IEN, the carrier used should be chosen to minimize the masking effect on the LM stimulus. Therefore, we chose a carrier for which its spatial frequency did not interfere with the signal spatial frequency (1 cpd). The noise that the carrier will introduce at the signal frequency will be intrinsically present in the IEN evaluated. Therefore, the first carrier used was a plaid composed of two perpendicular oblique sine waves of 7.54 cpd. At each stimulus presentation, the phase of both sine waves varied randomly. The amplitude of the sine waves was set so that the brightest and darkness peeks of the unmodulated texture ($T(x,y)$) was -0.5 and 0.5. Therefore, the difference of luminance between the two peeks was equal to the stimulus luminance average (L_0). In other words, for both stimulus types, the contrast average of the whole stimulus was set equal to the luminance average (L_0) of the whole stimulus (Figure II-5).

A disadvantage of using a plaid as a carrier is that some values are near 0. Since, for the CM stimuli, the carrier is multiplied with the signal and the noise, low values decreases both the signal and the noise, which do not affect the performance of an ideal observer. However, for a human observer, having internal noise makes the detection of low signal strength in low external noise undetectable. To maximize the contrast of the carrier, a checkerboard was also used. Note that a plaid can be seen as a smoothed checkerboard. The element size was 6x6 pixels (0.094 degrees of visual angle).

The third carrier used was the most widely used: binary noise. The element size was also 6x6 pixels. Therefore, the only difference between this carrier and the previous one is that its element positions are randomized.

External Noise

Gaussian distributed white noise was used (Figure II-4 right and Figure II-6). For uncorrelated white noise, the spectral density curve as a function of the spatial frequency is flat (Legge, Kersten, & Burgess, 1987; Pelli, 1981, 1990). For LM stimuli, noise frequencies far from the signal frequency have little impact on the detection of the signal. However, for the CM stimulus, since the noise modulation type is in contrast modulation, the noise should affect the signal (1 cpd) but not the carrier (7.54 cpd). Therefore, a band-pass filter (>0.5 cpd and <2 cpd) was applied to the noise. This filter did not change the noise energy at the signal frequency.

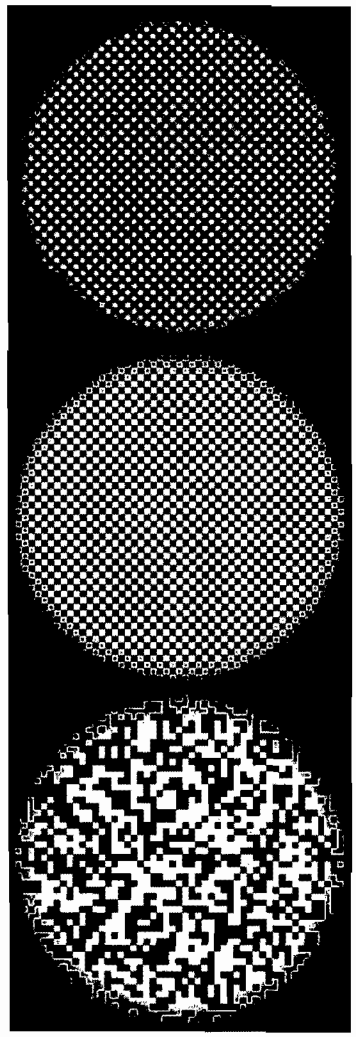


Figure II-5. Carriers. Three different types of carriers: plaid (top), checkerboard (center), and binary noise (bottom).

As shown in the result section, subjects had a high IEN for the CM stimuli detection. To vary significantly the total amount of noise (IEN plus external noise) in order to derive the IEN and CE, a large amount of external noise had to be used. Filtering the external noise also had the benefit of reducing the luminance contrast range used by the noise, which permitted to increase the external noise energy at 1 cpd without truncating luminance values.

The contrast root-mean-square (N) of the noise was set to 0, 0.0125, 0.050 and 2.00 for LM stimuli and 0, 0.050, 2.00 and 4.00 for CM stimuli. More noise was introduced for CM stimuli because of two reasons. First, since the modulation was multiplied with the texture ranging between –0.5 and 0.5, two times more noise could be used without exceeding the monitor luminance range. Second, as mention above, CM had more IEN so greater external noise was required to derive the IEN and CE.

Ideal observer

As stated previously, an ideal observer is a theoretical observer mathematically computing the optimal solution. Consequently, defining an ideal observer generally consists in deriving the smallest SNR (k_{ideal}) sufficient to perform a task for a given threshold criterion. However, the present section will only show that an ideal observer has the same sensitivities to both LM and CM stimuli. Since the CE (k_{ideal}/k) is defined relative to the optimal SNR (k_{ideal}) and that the optimal SNR is the same for both LM and CM stimuli ($k_{idealLM}$=$k_{idealCM}$), deriving the exact optimal SNR is not necessary and is beyond the scope of the present study. The relative difference between CEs of LM and CM stimuli may be compared directly:

$$\frac{CE_{LM}}{CE_{CM}} = \frac{k_{ideal\,LM}/k_{LM}}{k_{ideal\,CM}/k_{CM}} = \frac{k_{CM}}{k_{LM}}. \qquad \textbf{(19)}$$

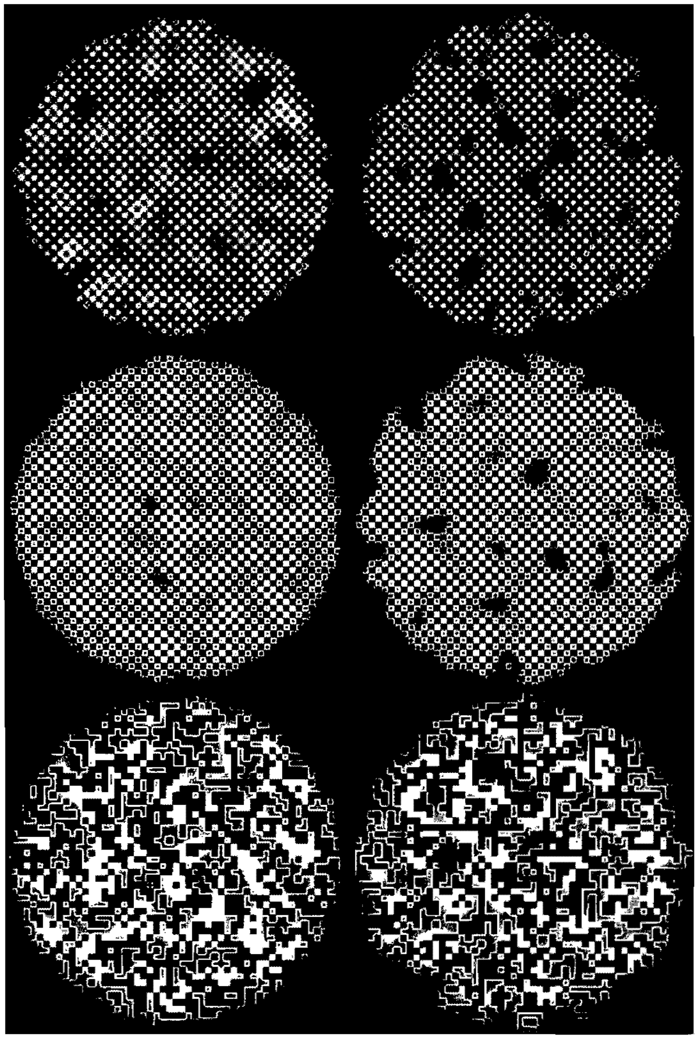


Figure II-6. Luminance- and contrast-modulated noise. Luminance- (left) and contrast-modulated (right) noise with three types of carriers: plaid (top), checkerboard (center), and binary noise (bottom).

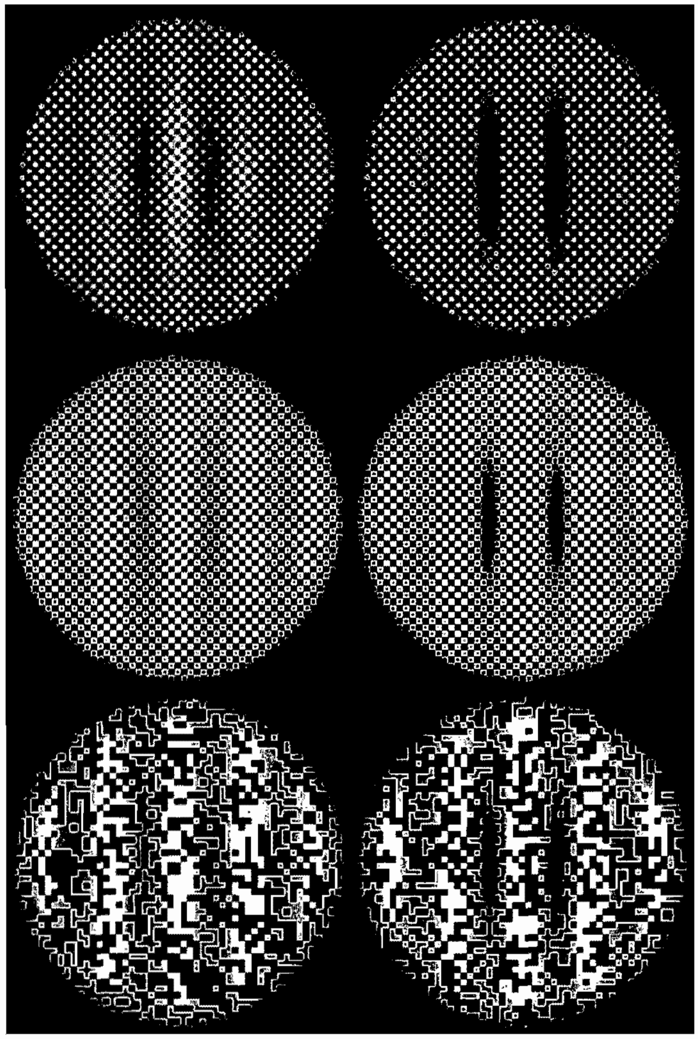


Figure II-7. Luminance- and contrast-modulated signals. Luminance- (left) and contrast-modulated (right) Gabor patch with three types of carriers: plaid (top), checkerboard (center), and binary noise (bottom).

In other words, if the CE for detecting LM stimuli is the same as the CE for detecting CM stimuli ($CE_{LM}=CE_{CM}$), then the SNR required for detecting LM stimuli will be equal to the SNR required for detecting CM stimuli ($k_{LM}=k_{CM}$).

The luminance profile of the LM stimuli may be given by combining Equations 15 and 17:

$$L_{LM}(x,y)=L_0(1+S(x,y)+N(x,y)+T(x,y)), \qquad (20)$$

and for CM stimuli by combining Equations 16 and 17:

$$L_{CM}(x,y)=L_0(1+(1+S(x,y)+N(x,y))T(x,y)). \qquad (21)$$

Since the average luminance (L_0) is constant in all the testing conditions, it may be abstracted from the stimulus equation and the stimuli may be defined by their contrast function $C(x,y)$ (Linfoot, 1964) instead of their luminance function ($L(x,y)$):

$$C(x,y)=L(x,y)/L_0-1. \qquad (22)$$

The contrast functions of LM and CM stimuli are:

$$C_{LM}(x,y)=S(x,y)+N(x,y)+T(x,y), \qquad (23)$$
$$C_{CM}(x,y)=(1+S(x,y)+N(x,y))T(x,y). \qquad (24)$$

The expected contrast profile when there is no signal ($S(x,y)=0$) is $T(x,y)$ and is known to the ideal observer. Therefore, it may be subtracted from the stimulus equation without affecting the ideal observer's performance:

$$C'(x,y)=C(x,y)-T(x,y), \qquad (25)$$
$$C'_{LM}(x,y)=S(x,y)+N(x,y), \qquad (26)$$
$$C'_{CM}(x,y)=(S(x,y)+N(x,y))T(x,y). \qquad (27)$$

Consequently, the performance of an ideal observer to $C'_{LM}(x,y)$ and $C'_{CM}(x,y)$ are identical to the one using $C_{LM}(x,y)$ and $C_{CM}(x,y)$ respectively.

The tasks for an ideal observer may be summarized as detecting a LM signal ($S(x,y)$) in LM noise ($N(x,y)$) or detecting a CM signal ($S(x,y)T(x,y)$) in CM noise ($N(x,y)T(x,y)$).

Since the ideal observer has perfect knowledge of the texture ($T(x,y)$) such an observer may remove it from computation. Therefore, another equivalent CM stimulus may be defined as:

$$\begin{aligned} C''_{CM}(x,y) &= C'_{CM}(x,y)/T(x,y) \\ C''_{CM}(x,y) &= S(x,y)+N(x,y) \end{aligned} \quad . \qquad (28)$$

Consequently, the performance to $C''_{CM}(x,y)$ will be identical to $C'_{CM}(x,y)$, which is the same as using $C_{CM}(x,y)$.

Since $C_{LM}'(x,y) = C_{CM}''(x,y)$ and the ideal observer performance to $L'_{LM}(x,y)$ and $L'_{CM}(x,y)$ are identical to the one using $C'_{LM}(x,y)$ and $C''_{CM}(x,y)$ respectively, the performance of an ideal observer using $L_{LM}(x,y)$ will be equal to the one using $L_{CM}(x,y)$ as long as it has perfect knowledge of the texture ($T(x,y)$). Consequently, for an ideal observer, detecting a LM signal ($S(x,y)$) in LM noise ($N(x,y)$) is equivalent as detecting a CM signal ($S(x,y)T(x,y)$) in CM noise ($S(x,y)N(x,y)$). In other words, an ideal observer will require the same SNR (k_{ideal}) for detecting both stimulus types.

To be abstracted, a texture must be precisely known (i.e. the texture values $T(x,y)$ at all the positions (x,y) must be known). For the plaid carrier, the frequency, amplitude and orientation of the carrier are known except that its phases change randomly at each presentation. However, the phase can easily be deduced precisely since the spatial frequency of the carrier is higher than the rest of the stimulus (signal and noise). Consequently, the local variation only depends on the carrier. For LM stimuli, the signal and noise, which are at lower spatial frequencies, will only change the local mean luminance. For CM stimuli, the signal and the noise will only change the local carrier contrast. Therefore, the phase of the plaid can be detected and the value of the texture

($T(x,y)$) can be precisely computed at each pixel position (x,y) and abstracted from the equation.

For a checkerboard and a binary noise carriers, each position (x,y) may have two possible values: -0.5 and 0.5. Since it is impossible to have negative luminance pixel values, the luminance range of each pixel ($L(x,y)$) must be, for a symmetrical reason, between 0 and $2L_0$. Based on these constraints and on Equations 3 and 4, the modulation ($M(x,y)$) can theoretically range between 0.5 and 1.5 for LM stimuli, and between 0 and 2 for CM stimuli. For both LM and CM stimuli, a texture element ($T(x,y)$) of –0.5 or 0.5 will cause the luminance value at that same position to be bellow or above the luminance average (L_0) respectively. Consequently, an ideal observer can precisely recompute the original texture for all the carriers used in the present study.

Procedure

Hardware

The monitor, which was the only luminance source in the room, was a 19 inches ViewSonic E90FB .25 CRT screen and was calibrated using a Minolta CS100 photometer. A Pentium 4, 3.2 GHz with a 10 bits Matrox Parhelia512 graphic card computed the stimuli. This graphic card was not limited in the number of simultaneously displayed color and therefore could display 1024 different grey levels for a given image. This was necessary since the detection threshold of the LM stimuli in certain conditions was relatively low. The distance between the monitor and the subject was 1.14 meters and each pixel on the screen was 0.016 x 0.016 degrees of visual angle.

Psychophysical methods

The constant stimuli paradigm was used to evaluate the subjects' threshold in different conditions using a two-interval-force-choice procedure. A block was composed of

28 trials: 1 stimulus type (either LM or CM), 4 noise conditions and 7 signal-contrast levels. Five pseudo-random blocks were performed before the subject was free to rest. At that time, the stimulus modulation type was switched and the subject was advised of this change. Therefore, LM and CM stimuli alternated until 20 blocks of each stimulus type were performed.

Data analysis

For each subject, each stimulus type, each carrier and each noise condition, the detection threshold (75% correct) was evaluated by fitting a Weibull function using the bootstrap technique. Afterwards, for each subject, each stimulus type and each carrier, Equation 4 was fitted to deduce the two parameters: k and N_{eq}. The fitting was achieved by minimizing an error function using Excel Solver (Newton method). The error function was the sum over each noise condition of the squared difference in log units between the detection threshold and the predicted threshold by Equation 4.

Results

Figure II-8 shows individual results when a plaid was used as a carrier. All observers had similar patterns. In no- or low-external noise conditions, detection thresholds for CM stimuli were higher than those for LM stimuli. However, the threshold differences were generally not significant in high-noise conditions. Figure II-9 and Figure II-10 show the results for the checkerboard and binary noise carriers. Although the detection threshold differences in no- or low-external noise conditions are smaller, the same pattern was observed. These results strongly suggest that all observers had similar CEs but different IENs for processing LM and CM stimuli. These results are confirmed in Table II-1 where the differences of IEN and CE between LM and CM stimuli processing are listed. The differences of CEs are relatively small (0.02, -0.09 and 0.07 log units using a plaid, checkerboard or binary noise, respectively, as a carrier), compared to the differences of

IENs (1.03, 0.61 and 0.33 log units, respectively). We can also note that the differences in IENs were not the same in all three conditions.

Figure II-11 shows the averaged CE over all observers in all the conditions. As shown, neither the stimulus type nor the carrier type affected the CE. These results are not surprising for LM stimuli since the carrier only plays a masking role. In high noise conditions, the impact of the carrier (or mask) becomes not significant. Therefore, there is no reason to expect a threshold difference in high-noise conditions for LM stimuli. What is more surprising is that, in high-noise conditions, the same thresholds were observed for CM stimuli and this was true for all carriers.

Differences in IEN where noted when using different carriers (Figure II-12). For LM stimulus sensitivities, using a checkerboard as a carrier resulted in slightly more IEN than using the plaid but much less than when using binary noise. These results are not surprising given the carrier's masking role for LM stimuli. For CM stimulus sensitivities, the checkerboard generated the least amount of IEN.

Discussion

Same calculation efficiencies

The results of the current study clearly show that there is no, or very little, difference in CE between LM and CM stimuli detection. This implies that observers are just as efficient for extracting LM signals from LM noise as they are for extracting CM signals from CM noise. Consequently, these results suggest that, after a second-order rectification, both stimulus types are processed by common mechanisms.

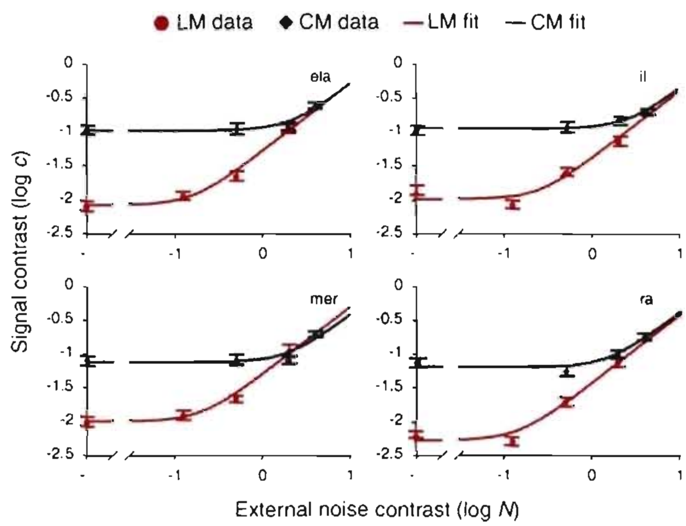


Figure II-8. Results using a plaid as a carrier. The error bars were calculated using bootprob.

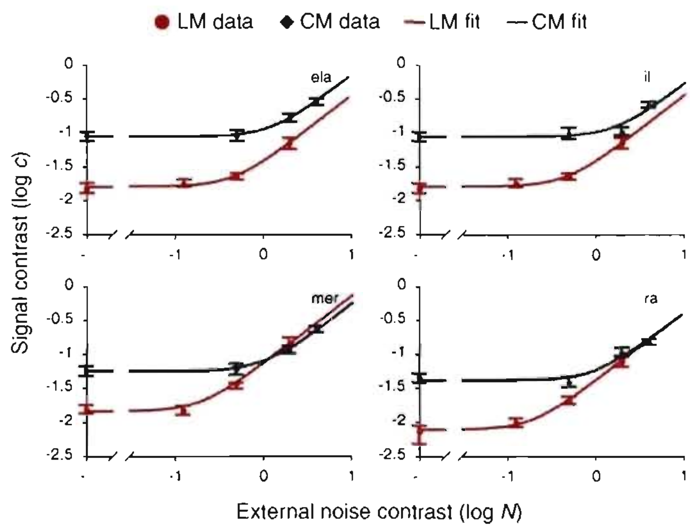


Figure II-9. Results using a checkerboard as a carrier. The error bars were calculated using bootprob.

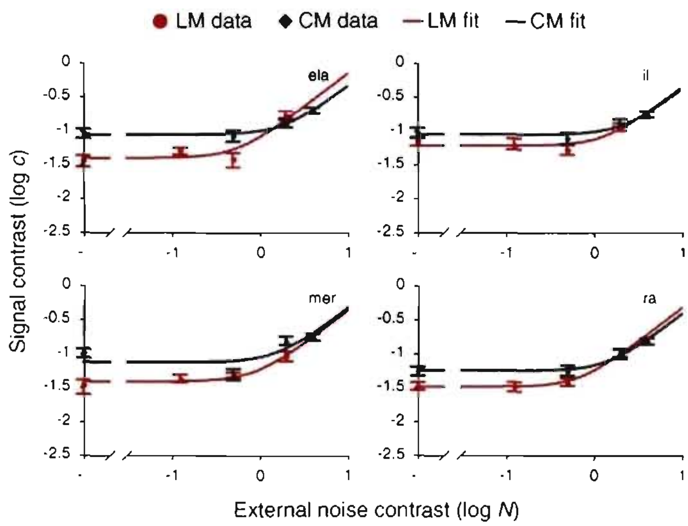


Figure II-10. Results using a binary noise as a carrier. The error bars were calculated using bootprob.

	Plaid		Checkerboard		Binary noise	
	IEN	CE	IEN	CE	IEN	CE
Ela	1,10	0,02	0,44	-0,30	0,53	0,19
Il	0,99	-0,03	0,57	-0,17	0,20	0,04
mer	0,98	0,12	0,71	0,12	0,25	-0,03
Ra	1,04	-0,04	0,72	-0,01	0,33	0,09
Mean	1,03	0,02	0,61	-0,09	0,33	0,07
+/-	0,03	0,04	0,08	0,11	0,08	0,05

Table II-1. IEN and CE differences. Internal equivalent noise (IEN) and calculation efficiency (CE) differences between luminance- (LM) and contrast- (CM) modulated stimuli (CM IEN or CE minus LM IEN or CE in log units). The last row corresponds to the standard error.

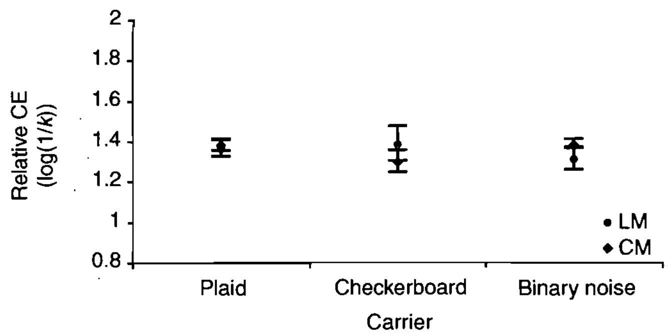


Figure II-11. Mean calculation efficiency. Relative calculation efficiency (CE) for luminance- (LM) and contrast- (CM) modulated stimuli using three different carriers: plaid, checkerboard and binary noise. Error bars show the standard error. Note that for comparative reasons, the same range was used as in Figure II-12.

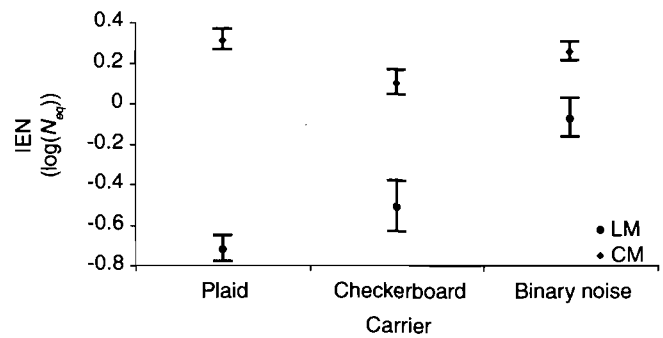


Figure II-12. Mean internal equivalent noise. Internal equivalent noise (IEN) for luminance- (LM) and contrast- (CM) modulated stimuli using three different carriers: plaid, checkerboard and binary noise. Error bars show the standard error.

However, Georgeson and Schofield (2002) found evidence suggesting that LM and CM stimuli are not merged or confused after a rectification process because recognition (LM vs CM) and detection (LM vs noise and CM vs noise) threshold for LM and CM stimuli are similar. We argue that this does not imply that both stimuli are treated by separate post-rectification mechanisms. We suggest that although the same mechanisms are processing two stimuli that were initially treated by separate mechanisms, the different properties of the two stimuli are not necessarily lost.

For example, if we compare the detection and recognition of two LM gratings with opposite phases, there is no reason to expect a difference between the two. Initially, it is not the same neuron population treating the two stimuli. Therefore, the two pathways are initially different. However, it would be unlikely that different mechanisms would be treating the two stimuli at a later stage. Consequently, the properties of the two stimuli are still available after initial processing even though both stimuli are, at a later stage, processed by common mechanisms.

Different internal equivalent noises

The results clearly attribute the difference between LM and CM sensitivities to a difference of IEN. If the same mechanisms treat both stimulus types after a second-order rectification, then the difference in IEN must either come from internal noise prior to the rectification or be caused by a sub-optimal rectification. To address this question an ideal observer with pre-rectification IEN will be considered.

Based on the definition described above for LM and CM stimuli, an ideal observer has the same performance for both stimuli. This is true given that an ideal observer does not have any internal noise. However, when detecting CM stimuli, before the rectification process, some internal noise (ex: optical noise) is likely to be added to the stimulus. In the presence of high external noise (LM or CM), the internal noise will be negligible and the performance will not be significantly affected. Therefore, the CE would be unaffected by pre-rectification noise. In the absence of external noise, however, the CM detection task consists in detecting a CM signal ($S(x,y)T(x,y)$) in LM noise ($N(x,y)$) compared with the LM detection task, which consists in detecting a LM signal ($S(x,y)$) in LM noise ($N(x,y)$). Therefore, the two tasks would only differ by their signal ($S(x,y)$ vs $S(x,y)T(x,y)$).

When the task is to detect a LM signal ($S(x,y)$) in LM noise ($N(x,y)$), the energy of the signal may be defined as (Legge, Kersten, & Burgess, 1987; Pelli, 1981, 1990):

$$E_{LM} = \int_Y \int_X S^2(x,y)dxdy, \qquad (29)$$

where X and Y are the width and height of the image, respectively.

When the task is to detect a CM signal ($S(x,y)$T(x,y)) in LM noise ($N(x,y)$), the energy of the signal may be defined as:

$$E_{CM} = \int_Y \int_X (S(x,y)T(x,y))^2 dxdy.. \qquad (30)$$

If there is no statistical relation between the texture ($T(x,y)$) and the signal ($S(x,y)$), the signal energy of CM stimuli can be approximated by:

$$E_{CM} \approx \frac{\int_Y \int_X S^2(x,y)dxdy \int_Y \int_X T^2(x,y)dxdy}{XY}. \qquad (31)$$

Consequently:

$$E_{CM} \approx E_{LM} \frac{\int_Y \int_X T^2(x,y)dxdy}{XY}.(32)$$

Since the rms contrast of the texture is:

$$T_{rms} = \sqrt{\frac{\int_Y \int_X T^2(x,y)dxdy}{XY}}, \qquad (33)$$

and the energy of the CM stimulus is:

$$E_{CM} \approx E_{LM} T_{rms}^2 . \qquad (34)$$

In other words, for the same signal ($S(x,y)$), the energy of the LM stimulus ($S(x,y)$) will be $1/T^2_{rms}$ times greater ($T^2_{rms}<1$ since $-1<T(x,y)<1$) than the energy of the CM stimulus ($S(x,y)T(x,y)$). Since the energy is proportional to the squared contrast, to have the same energy level as the LM stimulus, the CM contrast must be $1/T_{rms}$ times greater. Therefore, in the same noise condition ($N(x,y)$), the LM sensitivity of the ideal observer's with LM internal noise would be $1/T_{rms}$ times greater than the CM sensitivity.

Consequently, simulating early (pre-rectification) noise causes an ideal observer to have $1/T_{rms}$ times more IEN for CM stimuli detection than for LM stimuli detection. Comparing LM and CM IENs (or sensitivities), if we suppose that the significant internal noise occurs prior to the rectification, a factor of T_{rms} should be considered. Therefore, pre-rectification IEN ($N_{eq\ pre\text{-}rect}$) may be defined as:

$$N_{eq\ pre-rect} = \begin{cases} N_{eq} & for\, LM\, stimuli \\ N_{eq} T_{rms} & for\, CM\, stimuli \end{cases} . \qquad (35)$$

The rms contrast (T_{rms}) of the plaid, checkerboard and binary noise carriers were 0.25, 0.5 and 0.5 respectively. Using binary noise as a carrier, the results (0.33 log) show a difference of IEN (N_{eq}) near the factor predicted by the ideal observer with LM internal noise (2 or 0.30 log). Therefore, by compensating for the texture contrast ($N_{eq\ pre\text{-}rect}$), there was no significant difference between LM and CM IEN (Figure II-13). These results suggest that the significant noise occurred prior to the rectification and was common to both tasks.

One particularity of the binary noise carrier is that it has intrinsic noise. As shown in the method section, this noise does not affect the ideal observer performance. However, it does affect the performance of a human observer since the LM stimuli detection threshold is greater using the binary noise carrier than the checkerboard. Note that the only difference between the checkerboard and binary noise carriers is the randomness of the elements' position. From these results, we conclude that intrinsic noise of the binary noise carrier increases the IEN for LM and CM stimuli detection. Therefore, the IEN measured when using this type of carrier could be caused by the carrier intrinsic noise.

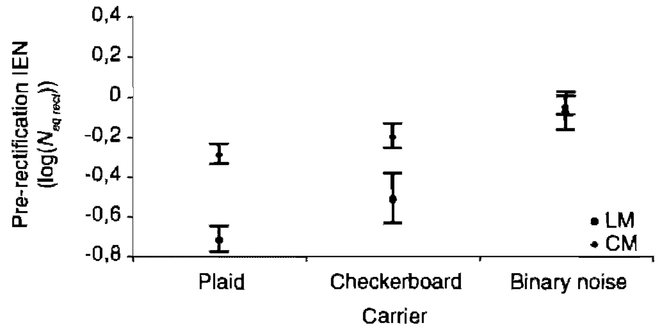


Figure II-13. Mean pre-rectification internal equivalent noise. Pre-rectification internal equivalent noise (IEN) for luminance- (LM) and contrast- (CM) modulated stimuli using three different carriers: plaid, checkerboard and binary noise. Assuming that the main portion of internal noise occurs prior to the rectification, the IEN for CM stimuli was multiplied by the texture rms. Error bars show the standard error.

Using a plaid or a checkerboard as a carrier, the ideal observer with pre-rectification internal noise does not explain the difference between the IEN measured for LM stimuli and the one measured for CM stimuli (Figure II-13). Consequently, the important difference between internal noises does not occur before the rectification. If both stimuli are processed by common mechanisms after the rectification process (same CE) and the pre-rectification noise cannot explain the difference of IEN observed, the difference of IEN must come from a sub-optimal rectification process. This suggests that, for binary noise, the intrinsic noise of the carrier was greater than the IEN caused by the sub-optimal rectification, and therefore, the IEN introduced by the sub-optimal rectification was not significant.

Conclusion

One of the main differences between LM and CM stimuli processing is that the human observer is less sensitive to CM than LM stimuli. We address this question by decomposing the sensitivities in IENs and CEs. The results show no difference of CE and indicate that the IEN is responsible for the sensitivity difference. Based on a rectification model, these results support the hypothesis that the two stimulus types could be treated by common mechanisms after a second-order rectification process.

To investigate the main source of internal noise for CM stimuli detection, an ideal observer with pre-rectification internal noise was built. Based on ideal observer analysis, pre-rectification internal noise could explain the difference of IEN between LM and CM stimuli detection when using binary noise as a carrier but not when using a plaid or a checkerboard. We conclude that a sub-optimal rectification process causes higher IEN for CM stimuli detection compared with LM stimuli detection and that the intrinsic noise of the binary carrier had a greater impact on the IEN than the sub-optimal rectification.

Acknowledgments: This research was supported by NSERC graduate fellowship to RA and NSERC operating grant to JF.

References

Baker, C. L., Jr. (1999). Central neural mechanisms for detecting second-order motion. *Curr Opin Neurobiol, 9*(4), 461-466.

Benton, C. P., & Johnston, A. (2001). A new approach to analysing texture-defined motion. *Proc Biol Sci, 268*(1484), 2435-2443.

Cavanagh, P., & Mather, G. (1989). Motion: The long and short of it. *Spat Vis, 4*(2-3), 103-129.

Chubb, C., & Sperling, G. (1988). Drift-balanced random stimuli: A general basis for studying non-fourier motion perception. *J Opt Soc Am A, 5*(11), 1986-2007.

Clifford, C. W., & Vaina, L. M. (1999). A computational model of selective deficits in first and second-order motion processing. *Vision Res, 39*(1), 113-130.

Georgeson, M. A., & Schofield, A. J. (2002). Shading and texture: Separate information channels with a common adaptation mechanism? *Spat Vis, 16*(1), 59-76.

Johnston, A., & Clifford, C. W. (1995a). Perceived motion of contrast-modulated gratings: Predictions of the multi-channel gradient model and the role of full-wave rectification. *Vision Res, 35*(12), 1771-1783.

Johnston, A., & Clifford, C. W. (1995b). A unified account of three apparent motion illusions. *Vision Res, 35*(8), 1109-1123.

Johnston, A., McOwan, P. W., & Buxton, H. (1992). A computational model of the analysis of some first-order and second-order motion patterns by simple and complex cells. *Proc Biol Sci, 250*(1329), 297-306.

Legge, G. E., Kersten, D., & Burgess, A. E. (1987). Contrast discrimination in noise. *J Opt Soc Am A, 4*(2), 391-404.

Linfoot, E. H. (1964). *Fourier methods in optical image evaluation*. London: Focal Press.

Nishida, S., Ledgeway, T., & Edwards, M. (1997). Dual multiple-scale processing for motion in the human visual system. *Vision Res, 37*(19), 2685-2698.

Nishida, S., & Sato, T. (1995). Motion aftereffect with flickering test patterns reveals higher stages of motion processing. *Vision Res, 35*(4), 477-490.

Pelli, D. G. (1981). *The effects of visual noise.* Cambridge University, Cambridge.

Pelli, D. G. (1990). The quantum efficiency of vision. In C. Blakemore (Ed.), *Visual coding and efficiency.* Cambridge: Cambridge University Press.

Prins, N., & Kingdom, F. A. (2003). Detection and discrimination of texture modulations defined by orientation, spatial frequency, and contrast. *J Opt Soc Am A Opt Image Sci Vis, 20*(3), 401-410.

Schofield, A. J., & Georgeson, M. A. (1999). Sensitivity to modulations of luminance and contrast in visual white noise: Separate mechanisms with similar behaviour. *Vision Res, 39*(16), 2697-2716.

Schofield, A. J., & Georgeson, M. A. (2000). The temporal properties of first- and second-order vision. *Vision Res, 40*(18), 2475-2487.

Solomon, J. A., & Sperling, G. (1994). Full-wave and half-wave rectification in second-order motion perception. *Vision Res, 34*(17), 2239-2257.

Sperling, G., Chubb, C., Solomon, J. A., & Lu, Z. L. (1994). Full-wave and half-wave processes in second-order motion and texture. *Ciba Found Symp, 184*, 287-303; discussion 303-288, 330-288.

Taub, E., Victor, J. D., & Conte, M. M. (1997). Nonlinear preprocessing in short-range motion. *Vision Res, 37*(11), 1459-1477.

Wilson, H. R., Ferrera, V. P., & Yo, C. (1992). A psychophysically motivated model for two-dimensional motion perception. *Vis Neurosci, 9*(1), 79-97.

Chapitre III

Double dissociation between first- and second-order processing

Rémy Allard et Jocelyn Faubert

Laboratoire de psychophysique et perception visuelle

École d'optométrie, Université de Montréal

Le contenu de ce chapitre a précédemment été publié sous forme d'article (©Elsevier) :

Allard, R., & Faubert, J. (2007). Double dissociation between first- and second-order processing. *Vision Research, 47*(9), 1129-1141.

Abstract

To study the difference of sensitivity to luminance- (LM) and contrast-modulated (CM) stimuli, we compared LM and CM detection thresholds in LM- and CM-noise conditions. The results showed a double dissociation (no or little inter-attribute interaction) between the processing of these stimuli, which implies that both stimuli must be processed, at least at some point, by separate mechanisms and that both stimuli are not merged after a rectification process. A second experiment showed that the internal equivalent noise limiting the CM sensitivity was greater than the one limiting the carrier sensitivity, which suggests that the internal noise occurring before the rectification process is not limiting the CM sensitivity. These results support the hypothesis that a suboptimal rectification process partially explains the difference of LM and CM sensitivity.

Keywords: Internal equivalent noise; second-order; rectification process; noise; filter-rectify-filter model

Introduction

Luminance- and contrast-modulated sensitivity

We are less sensitive to contrast-modulated (CM) than to luminance-modulated (LM) stimuli (Figure III-1). Typically, first-order stimuli (ex: LM) are defined by luminance or color, and can be directly detected through Fourier analyses, while second-order stimuli (ex: CM) are defined by other attributes such as texture, orientation or spatial frequency, and cannot be directly detected through Fourier analyses (Baker, 1999; Cavanagh & Mather, 1989; Chubb & Sperling, 1988; Wilson, Ferrera, & Yo, 1992). There is no consensus on the type of nonlinearity enabling the system to process second-order stimuli and, more specifically, some researchers tried to determine whether the same mechanisms are involved in the detection of first- and second-order stimuli. Although this problem is debated in spatial and temporal vision, the present study will focus on the processing of static LM and CM stimuli (spatial vision). As presented by Georgeson and Schofield (2002), the models illustrating the processing of these stimuli may be classified into three groups: Common mechanisms at all processing stages, completely separate mechanisms and initially separate but common late mechanisms.

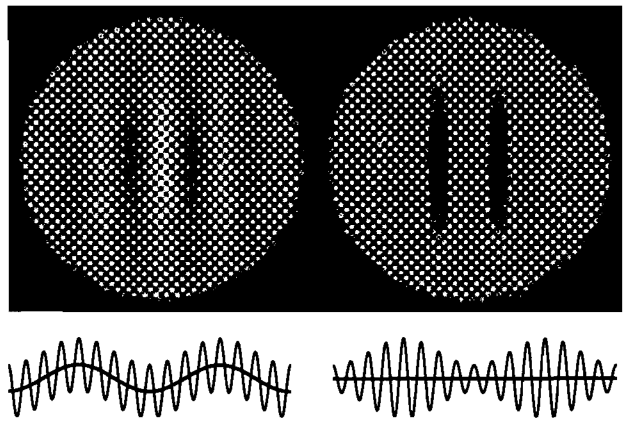


Figure III-1. Luminance- and contrast-modulated stimuli. The luminance- (left) and a contrast-modulated (right) stimuli (top) presented with their luminance profile (bottom, thin line). Thick lines represent the mean luminance variation at the signal spatial frequency.

Common mechanisms

In temporal vision (ex: direction discrimination of dynamic LM or CM stimuli), the common mechanism models suggest that motion detectors are sensitive to both types of stimuli. Although this hypothesis is not the most largely defended, some authors (Benton, 2002; Benton & Johnston, 2001; Benton, Johnston, McOwan, & Victor, 2001; Taub, Victor, & Conte, 1997) have shown that standard motion detection models processing first-order stimuli could also process second-order stimuli in certain conditions. For the detection of static LM or CM stimuli, the common mechanism hypothesis implies an early nonlinearity affecting the luminance profile of the stimulus enabling the LM processing system to also detect CM stimuli (illustrated by a compressive nonlinearity in the top row of Figure III-2 and first suggested in temporal vision by Henning, Hertz and Broadbent (1975)). Although such nonlinearities are known to occur in the visual system (He & Macleod, 1998; Legge & Foley, 1980; MacLeod, Williams, & Makous, 1992), it is generally accepted that it cannot account for CM stimuli sensitivity in all conditions (Derrington & Badcock, 1985; , 1986; Scott-Samuel & Georgeson, 1999; Smith &

Ledgeway, 1997), which suggests the existence of another mechanism specialized in the detection of static CM stimuli.

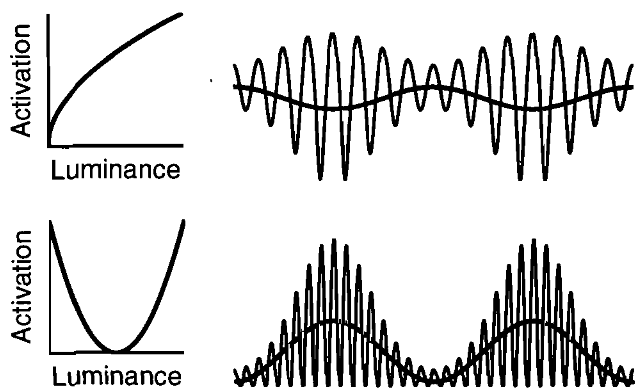


Figure III-2. Nonlinearities enabling the processing of contrast-modulated stimuli. The graphs on the left illustrate two types of nonlinearity that could enable the detection of CM stimuli. The top graph shows a compressive nonlinearity and the bottom one a rectification process. The resulting luminance profiles from passing the luminance profile of a CM stimulus (Figure III-1 on the right) through such functions are shown on the right using the thin lines. As we can observe, both nonlinearities introduce energy near the signal spatial frequency (illustrated by the thick lines showing the mean variation at the signal frequency).

Separate mechanisms

The second hypothesis suggests that separate mechanisms are processing both stimuli. Derrington and Badcock (1985; 1986) evaluated the processing of beat patterns (or CM stimuli) that are composed of two gratings defined at two different high spatial frequencies. The resulting stimulus appears to be one high spatial frequency grating periodically varying in contrast at a low spatial frequency. They found evidence that the processing of beat patterns has different qualitative processing behaviors from the processing of a low spatial frequency grating. For instance, they showed that increasing the temporal frequency reduced the detection threshold to low spatial frequency gratings but not to beat patterns. They also found that adaptation produced a motion aftereffect using low spatial frequency gratings but not using beat patterns. These results have led them to suggest that beat patterns (or CM stimuli) are processed by separated mechanisms evaluating the local contrast increment.

More recently in spatial vision, Georgeson and Schofield (Georgeson & Schofield, 2002; Schofield & Georgeson, 1999) found evidence supporting this hypothesis. They found that, although the processing of both stimuli induces similar responses (e.g. similar spatial (Schofield & Georgeson, 1999) and temporal (Schofield & Georgeson, 2000) integration, similar function shape relative to the spatial frequency (Schofield & Georgeson, 1999)), there is strong evidence suggesting that both stimuli are processed by separate mechanisms. To study this question, they used a facilitation paradigm in which observers were asked to identify a test modulation (LM or CM grating), in the presence of a background modulation (LM or CM grating). In one interval, only the background grating was presented and in the other, the test grating was added to the background grating. The task consisted in identifying the interval containing the test grating. When the modulation depth of the background grating was near detection threshold, the detection of the test grating was facilitated in intra-attribute conditions (test and background gratings of the same modulation type), but not in inter-attribute conditions (gratings of different modulation types). In another experiment (Georgeson & Schofield, 2002), they found that the detection (LM vs noise and CM vs noise) and recognition (LM vs CM) thresholds were similar suggesting that both stimuli are not merged or confused. However, they also found evidence suggesting an interaction between the processing of the two stimuli. Adapting to one type of stimulus affected the perceived modulation depth (difference of luminance or contrast for LM and CM stimuli, respectively) of the other (Georgeson & Schofield, 2002). Based on these results, they concluded that separate mechanisms are processing both stimuli but share a common adaptation mechanism at a late processing stage. Since inter-attribute adaptation effects in high contrast conditions are not very pattern selective (Ross & Speed, 1996; Snowden & Hammett, 1992, 1996), they concluded that both stimuli are processed by separate mechanisms having similar properties with the exception of the common adaptation mechanism.

Initially separate and common late mechanisms

In temporal vision, the model in which both stimuli are initially treated by separate mechanisms but are processed by common motion detection mechanisms at a later stage (usually referred as filter-rectify-filter model, (Wilson, Ferrera, & Yo, 1992)) is the most largely defended (see Baker (1999) for a review). This model illustrated in Figure III-3 suggests that a rectification (or squaring as illustrated in the bottom row of Figure III-2) process locally evaluates the intensity of the carrier (in our case, the local contrast) over the entire stimulus making the second-order information similar to first-order information. This would then enable later mechanisms to process the combination of the first- and second-order information.

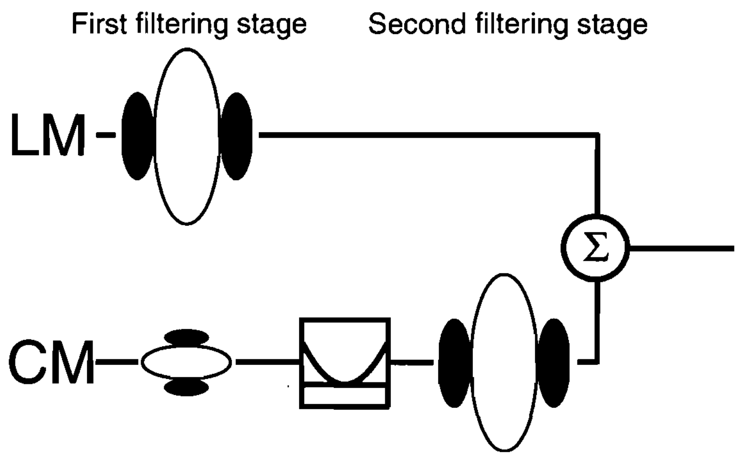


Figure III-3. Filter-rectify-filter model. The filter-rectify-filter model suggests that luminance- and contrast-modulated stimuli are initially processed by separate mechanisms and are combined after a rectification process occurring on the CM pathway. Adapted from (Baker, 1999).

In a recent study on spatial vision (Allard & Faubert, 2006), we found additional evidence showing similar responses between the processing of LM and CM stimuli, which reinforce the hypothesis suggesting the existence of common post-rectification mechanisms other than an adaptation mechanism. Indeed, we decomposed both sensitivities into internal equivalent noise (IEN) and calculation efficiency (CE) (Pelli, 1981). The IEN may be defined as the noise contrast necessary to model the impact of the internal noise on the sensitivity. The CE is inversely proportional to the smallest signal-to-noise ratio (where the noise is composed of internal and external noise) the system needs to detect the signal. To derive the IEN and CE, we need to evaluate the detection threshold as a function of external noise contrast (the TvC function, Figure III-4). In high external noise conditions, the internal noise is not significant and the signal-to-noise ratio (or CE) can be measured since the signal and external noise contrast are known. Our results showed that the detection thresholds of both stimuli did not differ in high noise conditions, i.e. the CEs to these stimuli were similar. In other words, observers were just as efficient at detecting a LM signal embedded in LM noise as detecting a CM signal embedded in CM noise. Based on these similar efficiencies (and other responses such as the ones mentioned above) between the processing of both stimuli, we conclude that both stimuli are probably processed by common mechanisms after a rectification process converting the CM information (signal and noise) to an activation pattern similar to the LM information. In the general discussion below, we will explain more extensively how our conclusion can be compatible with Georgeson and Schofield's results mentioned above leading them to a different conclusion (separate mechanisms).

Since the difference of sensitivity between LM and CM stimuli is due to a difference of IEN and not to a difference of CE, studying the difference of sensitivity can be reduced to studying the difference of IEN, which is the aim of the present study.

Single internal noise source

An internal noise source may be defined as a signal deterioration occurring at any processing level such as photons transduction, signal transmission along the optical nerve, neuronal noise, etc. The IEN corresponds to the external noise quantity necessary to simulate the impact of the internal noise. Consequently, the IEN simulates the impact of the combination of all internal noise sources. Assuming that each noise source may be modeled by a Gaussian distribution centered on 0 with SD of σ_i and that the noise sources are not statistically related, the resulting standard deviation (σ_{total}) of their combination would be:

$$\sigma_{total} = \sqrt{\sum_i \sigma_i^2}$$

Therefore, the resulting SD of the combination of two uncorrelated patterns with SDs of σ_1 and σ_2 is $\sqrt{\sigma_1^2 + \sigma_2^2}$. Hence, if the difference between the two SDs is important, the resulting SD will not largely differ from the greater SD. The greater difference between the greater SD (σ_1 or σ_2) and the resulting SD (σ_{total}) occurs when σ_1=σ_2. In such a case, the resulting SD will be $\sqrt{2}$ times greater.

The typical TvC function (Figure III-4), which may be used to decompose the sensitivity into IEN and CE, is a good example of the impact of the combination of two uncorrelated noise patterns. This function may be separated into three segments. If the external noise is small relative to the IEN, varying the external noise contrast does not significantly alter the effective noise (combination of internal and external noise) and the detection threshold remains relatively constant as a function of the external noise. However, in high noise conditions, the impact of the IEN is not significant and the effective noise contrast mainly depends on the external noise. In such conditions, varying the external noise has a direct impact on the detection threshold, which is then proportional to the noise contrast (slope near 1 in log-log coordinates). The only segment in which the impact of both noise sources is noticeable is when the external noise contrast is near the IEN contrast.

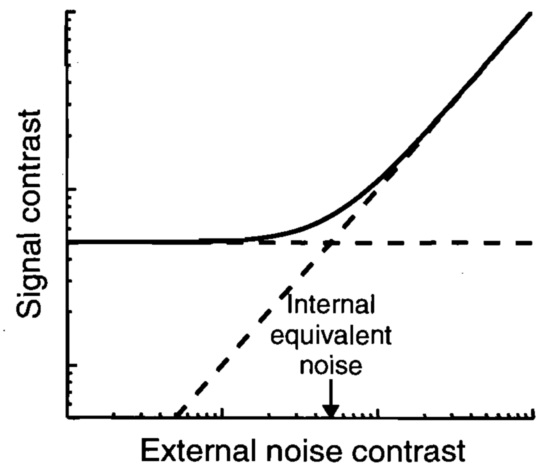


Figure III-4. Threshold versus contrast (TvC) function. The function shows the signal detection threshold as a function of the external noise contrast. Note that the two axes are scaled logarithmically.

Another example demonstrating that the impact of two noise sources generally behave as a winner-take-all rule is the absorbed-photon-noise occurring at the retinal level versus neural-noise occurring after luminance normalization that does not depend on the mean luminance. As presented by Pelli (1990), the IEN is greater in low than in high luminance conditions, he explains these results by two internal noise sources: one limiting the detection in low luminance conditions (absorbed-photon-noise) and another in high luminance conditions (neuronal-noise). The absorbed-photon-noise is not proportional to the stimulus luminance average but the neuronal-noise is since it occurs after luminance normalization. Consequently, if the luminance is sufficiently low, the resulting neuronal-noise is smaller than the absorbed-photon-noise, which does not vary according to the luminance average. As a result, the impact of the absorbed-photon-noise on the relative detection threshold (absolute minimum luminance variation detectable relative to the mean luminance background or Weber fraction) is then inversely proportional to the luminance average. However, in high luminance conditions, the internal noise measured (attributed to neuronal-noise) was proportional to the background luminance so the impact of such noise does not influence the relative detection threshold. Therefore, the visual system has at least two noise sources, one before the luminance normalization (absorbed-photon-noise) limiting the sensitivity in low luminance conditions and another after (neuronal-noise) limiting the sensitivity in high luminance conditions. These two examples show the interaction between two noise sources, where only the greater significantly affects the sensitivity and the smallest has no significant impact if their SDs largely differ.

Different IENs

Assuming the existence of different internal noise sources and that the greater one is significantly greater than the combination of the others, then the IEN measured models the impact of a single noise source. Consequently, when decomposing the sensitivity into IEN and CE, the former probably represents the impact of a single internal noise source. As mentioned above, the difference of LM and CM sensitivity can be attributed to a difference of IEN. Since the IEN probably models the impact of a single internal noise source for each stimulus type, comparing the sensitivity between LM and CM stimuli can be reduced to comparing the impact of their main internal noise sources (MINSs) limiting their sensitivities.

The fact that different IENs were measured for both types of stimuli does not necessarily imply that both MINSs are distinct. Indeed, a low contrast gain (stimulus attenuation, which mathematically corresponds to reducing the contrast of the signal) prior to the MINS increases the impact of the internal noise and, thereby, increases the IEN measured. A low contrast gain after the MINS does not affect the observer's performance since it reduces both the signal and main noise source (external noise), which does not affect the signal-to-noise ratio. Consequently, a low contrast gain does not affect the CE. Stimulus attenuation prior to the MINS increases the impact of this internal noise source and thereby directly affects the IEN measured without affecting the CE. Therefore, the IEN measured represents the combination of the MINS with the contrast gains prior to it. However, these two parameters have the same impact on the TvC function (more specifically on the IEN), which led many authors (Bennett, Sekuler, & Ozin, 1999; Lu & Dosher, 1998) to state that both are mathematically equivalent and therefore cannot be segregated.

Since the IEN depends on the MINS and the contrast gain prior to it, the difference of IEN observed does not necessarily imply that separate mechanisms are processing LM and CM. The difference of IEN could be due to different contrast gains prior to a common noise source. For instance, an early nonlinearity enabling the detection of CM stimuli by the mechanisms processing LM stimuli could explain these results. In such a model, the local variation of luminance (local contrast) introduces an alteration in the local mean luminance. Consequently, the early nonlinearity would introduce a LM grating into the CM grating with a smaller modulation depth than the modulation depth of the CM grating. In other words, both stimuli would be processed by the same mechanisms but would have different contrast gains prior to their MINS. In high external noise conditions, the signal-to-noise ratio would be the same for both types of stimuli since the contrast gain would affect both the signal and the main noise source being the external noise. However, in low external noise conditions, only the signal would be affected by the contrast gain (not the MINS occurring after the nonlinearity) resulting in different signal-to-noise ratios (different detection thresholds). Consequently, we would observe different IENs for both types of stimuli even though they would be processed by the same mechanisms and share a common MINS.

Different types of external noise

Two spatial frequencies are relevant to define CM stimuli, the ones relevant to the carrier and the ones to the signal. The present study evaluates the impact of three different external noise types: LM noise near the signal spatial frequency, CM noise near the signal spatial frequency and LM noise near the carrier spatial frequency (which we will refer to as LM-noise, CM-noise and carrier-noise respectively, Figure III-5). Note that since CM information can only be defined at lower spatial frequencies relative to the carrier, it is not possible to have CM noise near the carrier spatial frequency. Before the nonlinearity enabling the detection of CM stimuli (early nonlinearity or rectification process), the energy of the CM signal is near the carrier spatial frequency and energy near the signal spatial frequency only occurs after the nonlinearity making the CM information (signal and/or external noise) visible. Consequently, the CM detection threshold of an ideal observer is affected by CM- or carrier-noise, but not by LM-noise. LM-noise only affects the mean local luminance without affecting the local contrast defining the CM stimulus.[1] Oppositely, CM- and carrier-noise affects the local contrast without affecting the mean local luminance.

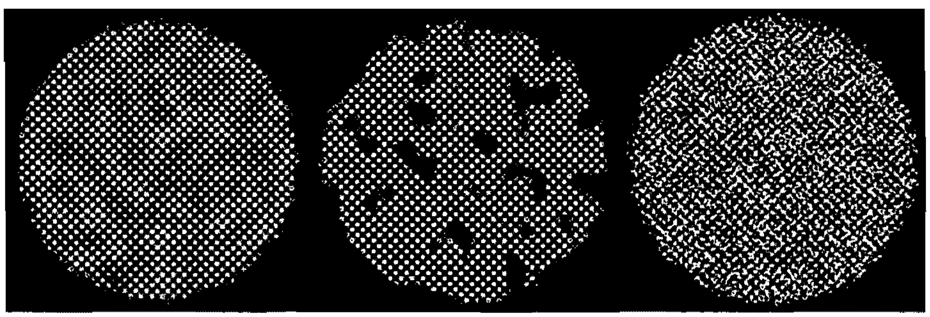


Figure III-5. Noise types. LM- (left), CM- (center) and carrier-noise (right), all three in the presence of a carrier.

Purpose of the present study

The objective of the first experiment was to evaluate inter-attribute interactions between both types of modulations. If an early nonlinearity converts CM information into LM information or if both types of information are merged after a rectification process, we

---

[1] If the local contrast is defined as the local difference of luminance relative to the local mean luminance (as opposed to the local difference of luminance relative to the background mean luminance of the entire stimulus), then altering the mean local luminance would affect the local contrast. As a result, it may be argued that the LM component is "leaking" in the CM component. However, the results obtained in the present study showed that, if such interaction exists, it had no significant impact.

should expect inter-attribute interactions: CM-noise should affect LM signal detection and LM-noise should affect CM signal detection. However, if both attributes are initially processed by separate mechanisms (suggesting that a rectification process evaluates the local contrast of CM stimuli to enable its processing) and they are not merged after the rectification process, then we should observe no or little inter-attribute interaction: noise of one attribute should have little impact on the signal detection of the other.

After showing that both stimuli are initially processed by separate mechanisms (CM detection is due to a rectification process evaluating the carrier contrast and not to an early nonlinearity converting CM information into LM information, see Figure III-2), the second objective was to determine if the MINS limiting CM sensitivity occurs before the rectification or not. Before the rectification process making CM information visible, the processing can be reduced to treating the carrier. The second experiment was aimed at evaluating if the MINS limiting the detection of the carrier also limits the CM detection. To do so, this experiment evaluated the impact of carrier-noise on the carrier and CM detection.

Experiment 1: Inter-attribute interactions

In a previous study (Allard & Faubert, 2006), we showed that observers were just as efficient at detecting LM signal embedded in LM-noise as to detect CM signal embedded in CM-noise. The first objective of the present study was to evaluate if LM and CM stimuli are processed, at least at some point, by separate mechanisms. To do so, the interactions between the processing of LM and CM stimuli were evaluated by measuring LM and CM stimuli detection threshold embedded in LM- and CM-noise (intra- and inter-attribute conditions). The absence of inter-attribute facilitation found by Schofield and Georgeson (1999) using a near threshold signal as background suggests that separate mechanisms are processing these stimuli. We could therefore expect to find no inter-attribute masking effect using noise as a background. Such double dissociation between LM and CM stimuli processing (i.e. LM-noise affecting more LM than CM stimuli detection and CM-noise affecting more CM than LM stimuli detection) would support the hypothesis suggesting the existence of a separate rectification mechanism processing CM stimuli, i.e. separate mechanisms are initially processing both stimuli. It would also imply that both stimuli are not merged to form a single activation pattern after the rectification. On the other hand, if an early nonlinearity in the visual system enables the detection of CM stimuli or if both attributes are merged after a rectification process, then LM- and CM-noise would affect both LM and CM signal detection, since CM information (signal and noise) would be converted into LM information.

Method

Observers

Three subjects aged 26, 27 and 27 years participated to the study. They had normal or corrected to normal vision. One of them (ra) was an author and the others were naive to the purpose of the experiment.

Apparatus

The stimuli were presented using a 19 in ViewSonic E90FB .25 CRT monitor with a mean luminance of 43 cd/m^2 and a refresh rate of 100 Hz, which was powered by a Pentium 4 computer. The 10-bit Matrox Parhelia512 graphic card could produce 1024 gray levels that could all be presented simultaneously. The monitor was the only light source in the room. A Minolta CS100 photometer interfaced with a homemade program calibrated the output intensity of each gun. At the viewing distance of 1.14 m, the width and height of each pixel were 1/64 deg of visual angle.

Stimuli

All the stimuli used in the present experiment are the sum of two terms: a luminance modulation ($M_{LM}(x,y)$) and the multiplication of a contrast modulation ($M_{CM}(x,y)$) with a texture ($T(x,y)$):

$$L(x,y) = L_0[M_{LM}(x,y) + M_{CM}(x,y)T(x,y)]$$

where L_0 represents the luminance average of the stimulus and the background luminance. Both modulations ($M_{LM}(x,y)$ and $M_{CM}(x,y)$) may be defined as

$$M(x,y) = 1 + S(x,y) + N(x,y)$$

where $S(x,y)$ and $N(x,y)$ are the signal and external noise functions respectively.

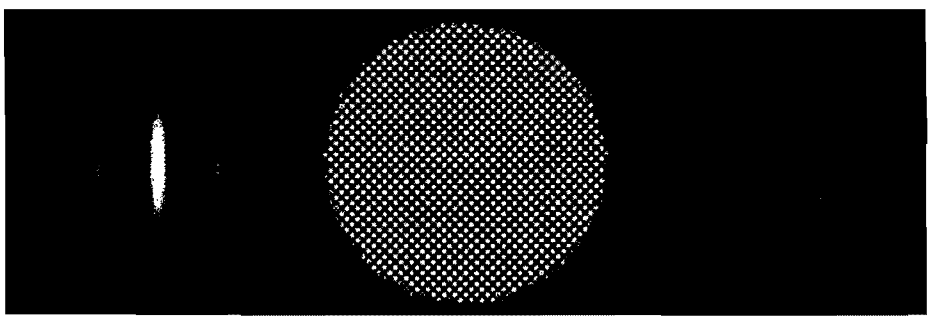


Figure III-6. Signal (Gabor patch, left), carrier (plaid, center) and filtered noise near the signal spatial frequency (right).

Signal function

The signal function ($S(x,y)$) was a Gabor patch (Figure III-6 left) with a center spatial frequency of 1 cpd, a SD of 1 deg, a phase randomized at each stimulus presentation and a Michelson contrast (C_{LM} or C_{CM} depending on the type of modulation) that varied depending on the task (see below).

External noise

The noise function ($N(x,y)$) generated a matrix of 320x320 pixels (5x5 deg), each element being randomly selected from a Gaussian distribution centered on 0. Each noise template was bandpass filtered by applying an ideal circular filter in the Fourier domain to keep all the orientations and only the frequencies within one octave below and above the relevant spatial frequency (1 cpd for the first experiment (Figure III-6 right) and 8 cpd for the second experiment). The SD of the Gaussian distribution before the filtering corresponded to the noise contrast in modulation depth (N_{ExtLM} or N_{ExtCM} depending on the type of modulation), which varied from one task to another.

Carrier

The carrier ($T(x,y)$, Figure III-6 center) was a plaid, i.e. the sum of two sinusoidal gratings. The spatial frequency of the gratings was 8 cpd and their orientations were oblique and perpendicular from one another (±45 deg). Such a carrier has the advantage of being defined within a narrow band spectral frequency. Consequently, the carrier had a limited impact on the sensitivity to LM stimuli defined at a lower spectral frequency. Using noise as a carrier does not have this advantage since it introduces noise at the signal frequency, which may mask the MINS. Another advantage of using a narrow band carrier is that it is easier to introduce noise that will selectively affect the carrier frequencies without affecting the signal frequency as was done in the second experiment. The phases of the two oblique sinusoidal gratings forming the carrier were randomized at each stimulus presentation and the contrast was set so that, in the absence of signal and noise, the luminance peaks were $0.25L_0$ and $0.75L_0$ (i.e. $-0.5 \leq T(x,y) \leq 0.5$).

Procedure

In all the conditions, a 2-interval-forced-choice method was used: one interval contained a carrier modulated by a signal and noise, and the other contained only a carrier modulated by noise. The task was to identify which interval contained the signal (LM or CM Gabor patch). Different noise templates with the same contrast were used in the two intervals. For a given task (detection of a LM or CM signal in LM or CM noise), the signal and noise modulation types were fixed and known to the observer. The stimuli were presented for 500 msec with stimuli intervals of the same duration. The spatial window was circular with a full contrast plateau of 4 deg width and soft edges following a Gaussian distribution with a SD of 0.25 deg. After each trial, a feedback sound indicated to the observer if his response was correct. To evaluate thresholds, a 2-down-1-up procedure was used (Levitt, 1971), that is, after two consecutive correct responses the dependant variable, which varied depending on the task, was decreased (or increased when the dependant variable was a noise contrast) by 10% and increased (or decreased) by the same proportion after each incorrect response resulting in a threshold criterion of 70.7%. For each threshold measured, 100 trials were performed and the threshold was defined as the geometric mean of the last 6 inversions (peaks) of the dependant variable values.

The experiment was conducted in three consecutive steps. The objective of the present experiment (the last step) was to evaluate LM and CM sensitivity in LM- and CM-noise conditions (see Figure III-7 for stimuli examples). Prior to evaluating the impact of different noise types on LM and CM signal detection, we had to determine the noise contrast of the two noise types (second step), which were arbitrarily set to the noise contrast increasing the detection thresholds of their respective stimulus by 0.5 log units. Thus, the first step consisted in measuring the detection thresholds of their respective stimulus, i.e. LM and CM stimulus in noiseless conditions.

Hence, for the first step, the noise contrasts (N_{ExtLM} and N_{ExtCM}) were set to 0. For each modulation signal detection, the signal contrast of the relevant modulation (C_{LM} or C_{CM}) was the dependant variable while the other was set to 0.

The second task consisted in defining the noise contrast for each noise modulation. Therefore, for each modulation signal detection, the dependant variable in the previous task

(C_{LM} or C_{CM}) was fixed to 0.5 log units above the threshold found for each subject and the noise contrast (N_{ExtLM} or N_{ExtCM}) became the dependant variables. Note that the noise contrast increasing the CM detection threshold was so high that the contrast modulation function had to be truncated ($0<M_{CM}(x,y)<2$). Near threshold, this truncation reduced the RMS of the contrast modulation function by less than 1%. We therefore assumed that it had no significant impact on the results.

After fixing the two noise contrasts, the final step consisted in detecting the LM and CM stimuli in these noise contrasts resulting in 4 staircases (2 signal x 2 noise types). For each noise type, the noise contrast was set to the one measured in the previous step while the other was kept to 0. One signal contrast (C_{LM} or C_{CM}) was the dependant variable while the other was fixed to 0.

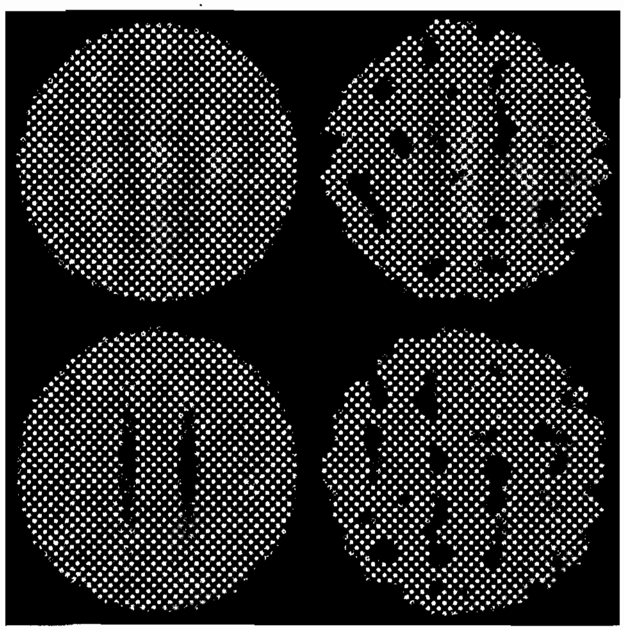


Figure III-7. LM (top row) and CM signals (bottom row) embedded in LM- (left column) and CM-noise (right column).

Results

Table 1 shows the LM and CM detection thresholds in noise free conditions (first step) and the noise contrasts necessary to increase each detection threshold by 0.5 log units respectively (second step) for each subject. In the absence of noise, observers were more sensitive to LM than CM stimuli by a factor of 15, 14 and 11 for subjects il, jmh and ra respectively. In similar proportions (factors of 11, 15 and 10 respectively), greater external noise contrasts were required to affect the CM signal detection because of a greater IEN for CM detection (Allard & Faubert, 2006).

Subjects	Detection threshold LM	Detection threshold CM	Noise threshold LM	Noise threshold CM
il	0.0074 x/÷1.039	0.11 x/÷1.042	0.59 x/÷1.157	6.2 x/÷1.071
jmh	0.0070 x/÷1.057	0.10 x/÷1.035	0.42 x/÷1.042	6.4 x/÷1.032
ra	0.0034 x/÷1.050	0.039 x/÷1.144	0.33 x/÷1.062	3.4 x/÷1.030

Table III-1. Signals and noises contrast thresholds. The first two columns show the detection thresholds for both types of modulations. The last two columns show the noise contrast (prior to the bandpass filtering) that was necessary to increase the respective detection threshold by 0.5 log units. The data are expressed as the geometric mean x/÷ geometric standard error.

Double dissociation between LM and CM stimuli detection

Figure III-8 shows the LM and CM signal detection thresholds measured in LM- and CM-noise. As expected, thresholds in intra-attribute noise conditions (LM and CM signals embedded in LM- and CM-noises respectively) were near 0.5 log units above the ones obtained in noiseless conditions (or close to, learning or measurement errors may explain the small differences). Oppositely, in the inter-attribute conditions, thresholds were similar or slightly above the ones in noiseless conditions. The important results are that, for all three observers, intra-attribute noise had a greater impact than inter-attribute noise for both types of modulation. Consequently, the detection threshold of LM stimuli increased more in LM- than in CM-noise conditions and the detection of CM stimuli was more affected by CM- than LM-noise. These results show a clear double dissociation between LM and CM stimuli processing. It is therefore possible to define a condition that selectively impairs the processing of one attribute while keeping the processing of the other relatively intact.

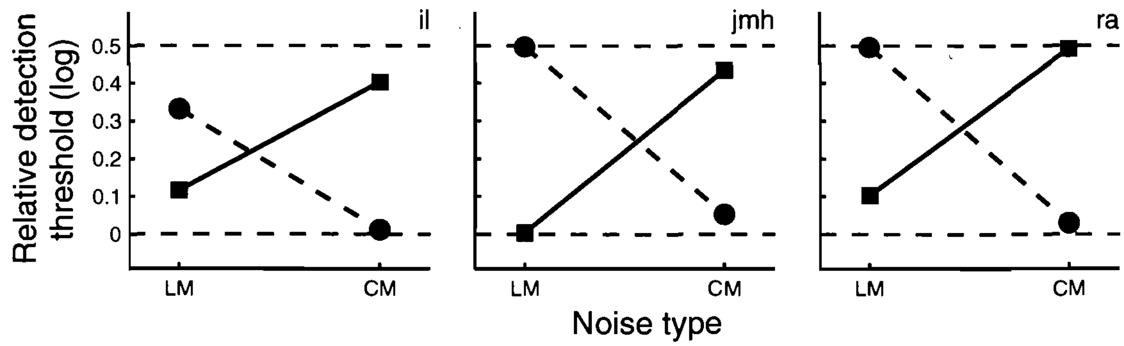


Figure III-8. LM and CM signal detection in LM and CM noise. Y-axis shows the elevation detection threshold relative to the detection threshold in the absence of noise. The x-axis shows the two noise conditions: LM- and CM-noise. The circles show the relative detection thresholds for LM signals and the squares for CM signals.

Discussion

Initially separate mechanisms

This double dissociation between LM and CM stimuli processing implies that both stimuli are, at least at some point, processed by separate mechanisms. Therefore, in agreement with the general consensus in the literature, the detection of static CM stimuli is not due to early nonlinearities (at least not in all conditions) in the visual system making the stimulus detectable through the same mechanisms processing LM stimuli. If this was the case, CM-noise would interfere with LM processing and *vice versa*, and a double dissociation would not be observed. By rejecting the common mechanisms hypothesis, the present data support the existence of a rectification mechanism independent of the mechanisms processing LM stimuli enabling the detection of CM stimuli. In such a model, CM stimuli processing would require an extra processing stage converting CM information into an activation pattern analogous to LM information by evaluating the local contrast over the entire stimulus. Afterwards, both stimuli would be similar and could be processed by common mechanisms (initially separate but common late mechanisms hypothesis) explaining the similarities between the processing of both stimuli or could still be treated by separate post-rectification mechanisms (separate mechanisms hypothesis).

It is worth noting that Schofield and Georgeson (1999) found that a high contrast LM background signal masked the detection of CM signal but not vice-versa. These results differ from ours in which no inter-attribute interaction was observed. However, our results are not necessarily inconsistent with theirs since, as opposed to their methodology, we evaluated the impact of a mask at only one contrast level. Consequently, it is possible that greater noise contrasts would also cause an asymmetrical inter-attribute interaction.

No post-rectification merging

The double dissociation between LM and CM stimuli processing also implies that, after a second-order rectification, both stimuli are not merged to form a single activation pattern. If both stimuli were merged and then processed by common mechanisms, inter-attribute noise would also impair the detection. Consequently, our results reinforce the conclusion emitted by Georgeson and Schofield (2002) that both stimuli are not merged or combined after a second-order rectification process. However, as defended in the general discussion below, we do not agree with their interpretation that the absence of post-rectification merging implies separate post-rectification mechanisms.

Experiment 2: Pre-rectification internal noise

The first experiment suggested that CM detection is due to the existence of a rectification mechanism evaluating the carrier contrast and not to an early nonlinearity converting CM information into LM information. Since the CM detection initially requires the processing of the carrier, the aim of the present experiment was to evaluate whether the MINS limiting the CM sensitivity occurs before the rectification process or later.

As mentioned above, prior to the rectification process, the energy of a CM stimulus is near the carrier spatial frequency. Locally, a CM stimulus affects the local contrast of the carrier and therefore does not affect its spatial frequency. Globally, however, modifying the local contrast of a carrier gives rise to energy slightly off the central spatial frequency of the carrier (side-band components). Since the receptive fields in V1 respond to frequencies one octave above and below their central spatial frequency (Campbell & Robson, 1968), it is generally accepted that the detection of CM stimuli cannot be reduced to the processing of side-band components (Derrington & Badcock, 1985; , 1986). Therfore, we will assume that the carrier central spectral frequency and its side-bands components stimulate the same receptive fields (and thereby the same receptive fields as an unmodulated carrier) and that observers detect CM stimuli by evaluating the local contrast increment rather than by detecting the presence of side-band components.

We can also reasonably assume that an unmodulated carrier and a CM stimulus both stimulating the same receptive fields, are detected using the same receptive fields. Indeed, for efficiency reasons, it would be unlikely to have similar receptive fields using some for the carrier detection and others for the first filtering stage of the CM detection. Based on these assumptions, the processing of an unmodulated carrier and a CM stimulus share the same initial pathways (the first filtering stage of the CM pathway shown in Figure III-3). Thus, the MINS limiting the carrier sensitivity may also be the MINS limiting the CM sensitivity. If this was the case, an external noise greater than the impact of this common MINS should significantly affect the detection thresholds to both stimuli (carrier and CM stimuli). Otherwise, if the external noise is greater than the MINS limiting the sensitivity to the carrier but smaller than the one limiting the CM sensitivity, then the carrier sensitivity would be affected but not the CM sensitivity.

Method

Since the method was very similar to the one used in the previous experiment, the present section only mentions their differences. Two tasks were performed in different noise contrasts: detection of the CM signal and detection of the carrier. Compared to the previous experiment, the noise was filtered near the carrier spatial frequency (carrier-noise, >4cpd and <16cpd, Figure III-9) instead of the signal spatial frequency. No CM-noise was used (N_{CM}=0) and five noise contrasts (SD of the Gaussian distribution before applying the bandpass filter) were used for the LM noise function (which now represents the carrier-noise): N_{LM}=0, 0.0625, 0.125, 0.25 and 0.5 modulation depths. The task consisting in detecting the CM stimuli was identical to the one in the previous experiment with the exception of the noise: the dependant variable was the signal contrast (C_{CM}) and the task consisted in discriminating the interval containing the CM signal from two intervals containing a carrier embedded in noise. For the detection of the carrier, the Michelson contrast of the texture was the dependant variable and the contrasts of both envelopes (C_{LM} and C_{CM}) were fixed to 0. Consequently, one interval contained the carrier embedded in noise and the other contained only noise. Since we were interested in the detection of the carrier near the signal, which was a Gabor patch with a spatial window of 1 deg of standard deviation, the same Gaussian window was used for the carrier detection. The order of the ten staircases (2 tasks x 5 noise levels) was randomized.

To separate the sensitivity into IEN and CE, the typical TvC function fitted to the data was (Legge, Kersten, & Burgess, 1987; Pelli, 1981, 1990):

$$C(N_{ext}) = k\sqrt{N_{eq}^2 + N_{ext}^2}$$

where $C(N_{ext})$ represents the detection threshold in the noise contrast N_{ext}. The two parameters fitted were k and N_{eq}. k is inversely proportional to the CE and N_{eq} represents the IEN. The fit consisted in minimizing the sum of the differences in log units between the evaluated thresholds and the ones estimated by the fit ($C(N_{ext})$).

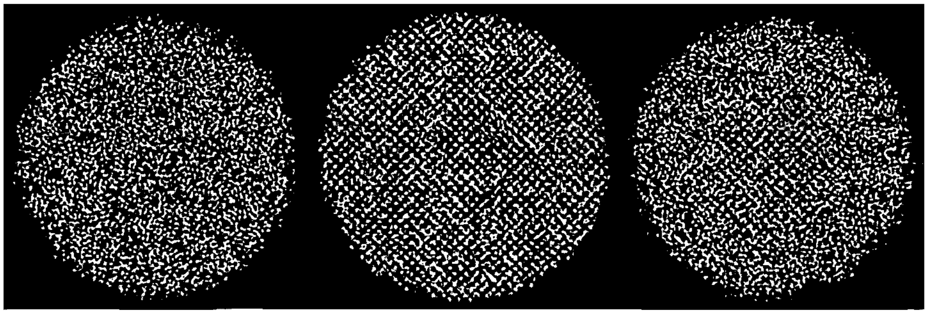


Figure III-9. Carrier-noise without a carrier (left), CM signal embedded in carrier-noise (center) and carrier with a Gaussian envelope with standard deviation of 1 deg in carrier-noise (right). In all three stimuli, the carrier-noise contrast (N_{LM}) is 0.5.

Results

For the carrier detection task, the IENs were 0.085, 0.051 and 0.063 noise contrast for the observers il, jmh and ra respectively. For CM detection task, the IENs were 0.32, 0.56 and 0.15 respectively, although we should consider that the IEN evaluated for observers il and jmh is likely to be inaccurate because of the absence of detection threshold in high noise conditions (considerably above the IEN). However, as it can be observed in Figure III-10, the IEN for CM stimuli detection for these two observers was near the maximum noise contrast used (0.5) or greater since the detection threshold difference between the greater noise contrast condition and the absence of noise is relatively small. The IEN for the detection of CM stimuli was consistently greater than the IEN for the detection of the carrier by a factor of 3.8, 10.1 and 2.4 respectively.

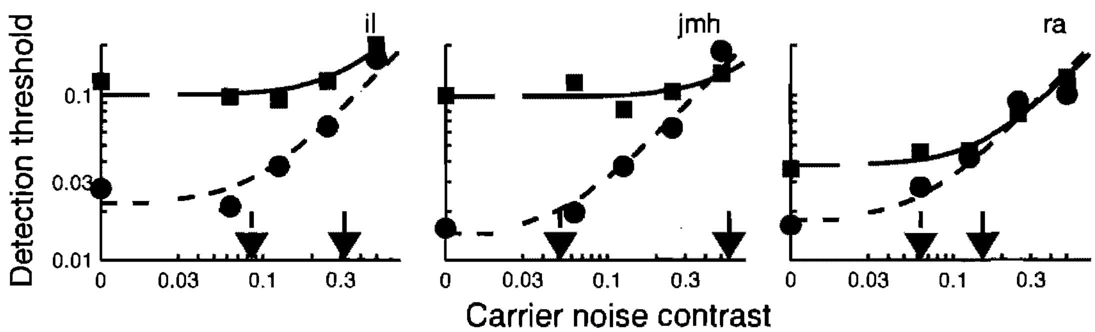


Figure III-10. CM (squares) and carrier (circles) detection thresholds in carrier-noise. Full- and dash-lines show best TvC function fits for CM and carrier detections respectively and arrows corresponds to the IENs.

Discussion

Pre-rectification noise not a limiting factor

Since the IEN for the detection of the carrier was smaller than the IEN measured for the detection of CM stimuli, it is possible to find a given noise condition (carrier-noise with a contrast level between the two IENs) affecting the detection of the carrier without significantly affecting the detection of the CM stimuli. In other words, such a noise level would be greater than the MINS limiting the carrier sensitivity, but smaller than the one limiting the CM sensitivity. We therefore conclude that the MINS limiting the CM sensitivity cannot occur at a processing level common with the carrier detection and must occur after the carrier and CM detection pathways have separated. Since the only processing prior to the rectification is related to the carrier, the present results suggest that the MINS limiting the CM sensitivity does not occur before the rectification but at or after the rectification.

Inter-subject difference

The IENs measured for the carrier sensitivities of the observers were very similar between all three subjects. However, this was not the case for the IENs limiting the CM sensitivity, in which the observer ra (one of the authors) had considerably smaller IEN compared to the two other observers. A possible explanation is that this observer had participated in a greater amount of psychophysical testing using CM stimuli. As shown by Dosher and Lu (2006), learning may reduce the IEN without affecting the CE for CM sensitivity. As a result, the contrast of the carrier-noise necessary to be greater than the MINS would be smaller resulting into a smaller difference between the two IENs for this observer. Note that this observation does not change the fact that, for all three observers, pre-rectification internal noise cannot explain the IEN measured.

General discussion

Common post-rectification mechanisms

Although we agree with Georgeson and Schofield's (2002) conclusion that LM and CM stimuli are not merged after a second-order rectification process applied to a CM stimulus (which thereby rejects, at least for static stimuli, the filter-rectify-filter model illustrated in Figure III-3), we do not agree that this implies that both stimuli are processed by separate post-rectification mechanisms. Post-rectification mechanisms could be able to process both attributes (luminance or contrast) without merging them. Separate processing does not imply separate mechanisms. As illustrated by a modified filter-rectify-filter model in Figure III-11, late mechanisms could process one attribute while ignoring the other. In other words, attentional selection could allow late mechanisms to focus on a single attribute. This modified filter-rectify-filter model can explain the similar responses, such as spatial and temporal integration and CEs, observed during the processing of either attribute and can also explain the lack of interaction between both types of stimuli since ignoring one attribute would limit its impact on the processing of the other.

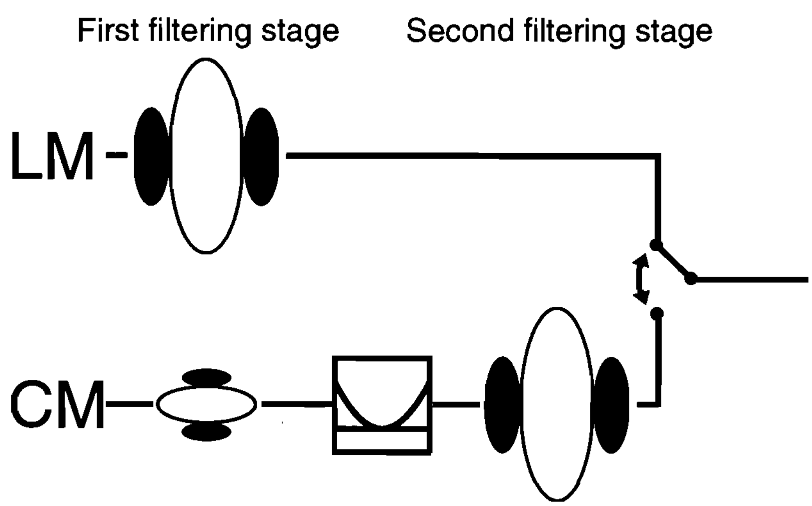


Figure III-11. A modified filter-rectify-filter model in which the late mechanisms can focus on either attribute (LM or CM) and ignore the other compared to the original filter-rectify-filter model (Figure III-3) suggesting that both attributes are combined.

This model could also explain that adapting to one attribute could affect the processing of the other since the adaptation could affect the mechanisms that are common to both pathways. As mentioned in the introduction, some data found by Georgeson and Schofield suggest that common late mechanisms could process both stimuli. They found an important inter-attribute tilt after-effect and an almost complete adaptation transfer effect on the perceived contrast to the cross-attribute stimulus. However, since they found that both stimuli are not merged (because of no sub-threshold summation), they concluded that they must be processed by separate mechanisms with the exception of a common adaptive mechanism. We argue that they share more than adaptive mechanisms and that it is because they share common mechanisms that it is possible to observe cross-type adaptation.

As an analogy demonstrating that common late mechanisms processing two stimuli do not imply merging them, consider a visual search task in which the target is either a red or a green vertical bar within distracters composed of blue vertical bars. We can reasonably assume that the observer's performance will be similar for either target. Now suppose that red horizontal bars are also added as distracters. Based on Treisman and Gelade's (1980) study on visual search, the search of the green target would now require the processing of a single attribute (color) while the search of the red target would require the conjunction of two (color and orientation). As a result, the presence of these two distracters (blue vertical and red horizontal bars) would affect more the observer's ability of searching the red than the green vertical bar. In the presence of green instead of red horizontal bars added as distracters, the opposite results would be obtained. This double dissociation (red horizontal bars affecting the search of the red vertical bar but not the search of the green vertical bar and *vice versa*) would lead to the correct conclusion that green and red bars are processed, at least at some point, by separate mechanisms. This is true because at the retinal level red and green are not absorbed by the same cones. However, it is highly improbable that we have a distinct searching mechanism for each color. Even though the same searching mechanism is used for searching both targets, we will certainly be able to show, using other tasks, that both colors are not merged or confused. Consequently, the visual search mechanism would be common to both colors even though these colors are not merged. Thus, the processing of these stimuli would invoke similar responses without, or with few, interactions. Red and green targets would not be confused, the presence of one would not affect the detection of the other and it would be possible to find two conditions (presence of blue vertical bars combined with the presence of either red or green horizontal bars) that would result in a double dissociation showing that they are processed, at least at some point, by separate mechanisms. Consequently, the fact that a higher-level mechanism is processing two stimuli regardless of their attributes does not imply that these attributes are lost and that the presence of one affects the processing of the other or that we should confuse one with the other.

This analogy shows that the fact that LM and CM stimuli are not merged or confused (lack of inter-attribute interaction) and the presence of a double dissociation does not imply that separate post-rectification mechanisms are processing both stimuli. Oppositely, we argue that processing similarities and inter-attribute adaptation effects suggest that both stimuli are processed by common post-rectification mechanisms able to select either attribute. We find more parsimonious the conclusion that both attributes are processed by common post-rectification mechanisms than the conclusion that they are processed by separate similar mechanisms sharing only an adaptation mechanism.

No impact of pre-rectification internal noise on CM sensitivity

As mentioned in the introduction, the CM detection threshold of an ideal observer would be affected by CM-noise or carrier-noise but not by LM-noise. For human observers, the first experiment showed that LM-noise also had no or little impact on the CM detection threshold. Consequently, pre-rectification internal noise at the signal spatial frequency cannot be a limiting factor, since such noise does not affect the CM sensitivity. The second experiment showed that the MINS limiting the CM sensitivity was greater than the one limiting the detection of the carrier, which implies that the MINS limiting the CM sensitivity must occur once the two pathways have separated. Consequently, pre-rectification internal noise at the carrier spatial frequency cannot be a limiting factor, since it is possible to add external noise greater than this internal noise (which affects the carrier detection) without affecting the CM sensitivity. Since the CM stimuli used in the present study were defined near two spatial frequencies (carrier and signal) and that pre-rectification noise (analogous to LM noise) at either spatial frequency cannot be limiting the CM sensitivity, we conclude that the internal noise occurring before the rectification process does not, in the present conditions, limit the CM sensitivity.

Impact of the first filtering stage

In a previous study, we decomposed the sensitivity to LM and CM stimuli into two IEN and CE, and found similar CEs using both stimuli (Allard & Faubert, 2006). Consequently, although the CE is a factor affecting the CM sensitivity, it does not explain the difference of sensitivity between LM and CM stimuli processing. The present study therefore focused on the difference of IEN. In the introduction, we showed that the IEN may also be separated into two factors: the MINS and a contrast gain prior to the MINS, which affects the signal strength and therefore affects the impact of the MINS.

The last experiment showed that the MINS limiting the CM sensitivity is not analogous to (or cannot be modeled by) adding LM noise to the stimulus either at the signal or carrier spatial frequency. We conclude that the MINS limiting the CM sensitivity must occur either at or after the rectification process. The fact that the MINS occurs after the first filtering stage does not imply that the processing at this filtering stage does not have an impact on the IEN. It rather implies that the internal noise occurring at the first filtering stage does not have a significant impact on the IEN and thereby on the sensitivity. However, contrast gain (signal attenuation or enhancement) prior to the MINS would affect the impact of the MINS and thereby affect the IEN measured. Therefore, the contrast gain occurring at the first filtering stage is a factor determining the IEN and should be considered when comparing LM and CM sensitivity.

For instance, Schofield and Georgeson (1999) have shown that the CM sensitivity is affected by the carrier contrast probably because of a compressive nonlinearity affecting the carrier. A compressive nonlinearity would affect the carrier contrast unevenly depending on the local contrast (defined by the CM signal) and would thereby affect the signal strength of the CM signal. As a result, the CM sensitivity would depend on the carrier contrast and stimulus attenuation prior to the compressive nonlinearity (analogous to lowering the carrier contrast) would influence the signal strength. As stated in the introduction, reducing the signal strength would increase the impact of the MINS without affecting the CE. Consequently, the IEN is not entirely due to second-order processing and, although first-order noise is not a limiting factor, first-order factors such as stimulus attenuation prior to the compressive nonlinearity and the compressive nonlinearity itself also affects the IEN.

Suboptimal second-order processing?

If pre-rectification noise near the carrier spatial frequency would have been the MINS for CM sensitivity, then the difference of IEN (thereby the difference of sensitivity) between LM and CM processing would have been entirely due to first-order limitations since the IEN (the MINS and the contrast gain prior to it) would have occurred at the first filtering stage. We would have been less sensitive to CM than LM stimuli not because they are more complex or require more computation, but simply because CM processing initially requires the processing of the carrier, which introduces noise. Excluding first-order factors is specially important when evaluating clinical populations such as aging (Faubert, 2002; Habak & Faubert, 2000) and autism (Bertone, Mottron, Jelenic, & Faubert, 2003; 2005) in which reduced CM sensitivity has been attributed to second-order processing.

The results of the present study suggest that the MINS limiting the CM sensitivity occurs after the first filtering stage. As we have previously shown (Allard & Faubert, 2006), a suboptimal rectification process evaluating the local contrast would affect the IEN without affecting the CE. In other words, the rectification process could be suboptimal and thereby limit the CM sensitivity by significantly introducing noise (that is, by being the MINS) and/or by attenuating the signal strength. Consequently, the rectification process (half-wave rectification, full-wave rectification or any other type of rectification) evaluating the carrier contrast over the entire stimulus is a potential candidate for the MINS. Further investigations are required to determine the proportion of the IEN due to the signal attenuation at the first filtering stage and the one due to the suboptimal rectification process.

Conclusion

In a previous study, we evaluated the detection of LM and CM stimuli embedded in LM and CM noises, and found that observers had the same sensitivity to both stimuli in high noise conditions. The present study evaluated the detection of LM and CM stimuli embedded in three different noise types: LM-, CM- and carrier-noise. We found a double dissociation between LM and CM stimuli detection in the presence of LM- and CM-noise. LM-noise had a greater impact on LM processing than on CM processing, while CM-noise had a greater impact on CM processing than on LM processing. This double dissociation implies that both stimuli are, at least at some point, processed by separate mechanisms. Combining these results to the ones found in a previous study where similar CEs were observed for LM and CM stimuli detection, we conclude that the processing of CM stimuli requires an extra rectification process but that both stimuli are processed by common post-rectification mechanisms.

Our results also demonstrate that the IEN limiting the sensitivity to the carrier was smaller than the one limiting the sensitivity to CM stimuli. We conclude that pre-rectification noise is small relative to the total amount of internal noise and therefore does not limit the CM sensitivity. We suggest that the internal noise limiting the sensitivity to CM stimuli is caused by a suboptimal rectification.

References

Allard, R., & Faubert, J. (2006). Same calculation efficiency but different internal noise for luminance- and contrast-modulated stimuli detection. *Journal of Vision, 6* (4), 322-334.

Baker, C.L., Jr. (1999). Central neural mechanisms for detecting second-order motion. *Curr Opin Neurobiol, 9* (4), 461-466.

Bennett, P.J., Sekuler, A.B., & Ozin, L. (1999). Effects of aging on calculation efficiency and equivalent noise. *Journal of the Optical Society of America, A, Optics, Image Science, & Vision, 16* (3), 654-668.

Benton, C.P. (2002). Gradient-based analysis of non-Fourier motion. *Vision Research, 42* (26), 2869-2877.

Benton, C.P., & Johnston, A. (2001). A new approach to analysing texture-defined motion. *Proc Biol Sci, 268* (1484), 2435-2443.

Benton, C.P., Johnston, A., McOwan, P.W., & Victor, J.D. (2001). Computational modeling of non-Fourier motion: further evidence for a single luminance-based mechanism. *Journal of the Optical Society of America, A, Optics, Image Science, & Vision, 18* (9), 2204-2208.

Bertone, A., Mottron, L., Jelenic, P., & Faubert, J. (2003). Motion perception in autism: a "complex" issue. *Journal of Cognitive Neuroscience, 15* (2), 218-225.

Bertone, A., Mottron, L., Jelenic, P., & Faubert, J. (2005). Enhanced and diminished visuo-spatial information processing in autism depends on stimulus complexity. *Brain, 128* (Pt 10), 2430-2441.

Campbell, F.W., & Robson, J.G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology, 197* (3), 551-566.

Cavanagh, P., & Mather, G. (1989). Motion: the long and short of it. *Spat Vis, 4* (2-3), 103-129.

Chubb, C., & Sperling, G. (1988). Drift-balanced random stimuli: a general basis for studying non-Fourier motion perception. *J Opt Soc Am A, 5* (11), 1986-2007.

Derrington, A.M., & Badcock, D.R. (1985). Separate detectors for simple and complex grating patterns? *Vision Research, 25* (12), 1869-1878.

Derrington, A.M., & Badcock, D.R. (1986). Detection of spatial beats: non-linearity or contrast increment detection? *Vision Research, 26* (2), 343-348.

Dosher, B.A., & Lu, Z.-L. (2006). Level and mechanisms of perceptual learning: Learning first-order luminance and second-order texture objects. *Vision Research, 46* (12), 1996-2007.

Faubert, J. (2002). Visual perception and aging. *Canadian Journal of Experimental Psychology, 56* (3), 164-176.

Georgeson, M.A., & Schofield, A.J. (2002). Shading and texture: separate information channels with a common adaptation mechanism? *Spat Vis, 16* (1), 59-76.

Habak, C., & Faubert, J. (2000). Larger effect of aging on the perception of higher-order stimuli. *Vision Research, 40* (8), 943-950.

He, S., & Macleod, D.I. (1998). Contrast-modulation flicker: dynamics and spatial resolution of the light adaptation process. *Vision Research, 38* (7), 985-1000.

Henning, G.B., Hertz, B.G., & Broadbent, D.E. (1975). Some experiments bearing on the hypothesis that the visual system analyses spatial patterns in independent bands of spatial frequency. *Vision Research, 15*, 887-897.

Legge, G.E., & Foley, J.M. (1980). Contrast masking in human vision. *Journal of the Optical Society of America, 70* (12), 1458-1471.

Legge, G.E., Kersten, D., & Burgess, A.E. (1987). Contrast discrimination in noise. *J Opt Soc Am A, 4* (2), 391-404.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America, 49* (2), Suppl 2:467+.

Lu, Z.L., & Dosher, B.A. (1998). External noise distinguishes attention mechanisms: *Vision Research, 38* (9), 1183-1198.

MacLeod, D.I., Williams, D.R., & Makous, W. (1992). A visual nonlinearity fed by single cones. *Vision Research, 32* (2), 347-363.

Pelli, D.G. (1981). The effects of visual noise. *Department of Physiology,* Ph.D. (Cambridge: Cambridge University.

Pelli, D.G. (1990). The quantum efficiency of vision. In: C. Blakemore (Ed.) *Visual coding and efficiency* (Cambridge: Cambridge University Press.

Ross, J., & Speed, H.D. (1996). Perceived contrast following adaptation to gratings of different orientations. *Vision Research, 36* (12), 1811-1818.

Schofield, A.J., & Georgeson, M.A. (1999). Sensitivity to modulations of luminance and contrast in visual white noise: separate mechanisms with similar behaviour. *Vision Res, 39* (16), 2697-2716.

Schofield, A.J., & Georgeson, M.A. (2000). The temporal properties of first- and second-order vision. *Vision Res, 40* (18), 2475-2487.

Scott-Samuel, N.E., & Georgeson, M.A. (1999). Does early non-linearity account for second-order motion? *Vision Research, 39* (17), 2853-2865.

Smith, A.T., & Ledgeway, T. (1997). Separate detection of moving luminance and contrast modulations: fact or artifact? *Vision Research, 37* (1), 45-62.

Snowden, R.J., & Hammett, S.T. (1992). Subtractive and divisive adaptation in the human visual system. *Nature, 355* (6357), 248-250.

Snowden, R.J., & Hammett, S.T. (1996). Spatial frequency adaptation: threshold elevation and perceived contrast. *Vision Research, 36* (12), 1797-1809.

Taub, E., Victor, J.D., & Conte, M.M. (1997). Nonlinear preprocessing in short-range motion. *Vision Res, 37* (11), 1459-1477.

Treisman, A.M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12* (1), 97-136.

Wilson, H.R., Ferrera, V.P., & Yo, C. (1992). A psychophysically motivated model for two-dimensional motion perception. *Vis Neurosci, 9* (1), 79-97.

Chapitre IV

Adding noise to a stimulus qualitatively alters its processing

Rémy Allard et Jocelyn Faubert

Laboratoire de psychophysique et perception visuelle

École d'optométrie, Université de Montréal

Le contenu de ce chapitre a été écrit sous forme d'article et sera soumis sous peu.

Abstract

Since the 1980s, many models (e.g. Linear Amplifier Model and the Perceptual Template Model) using external noise have been used extensively to evaluate observers' internal noise. Indeed, such models have characterized various functions (e.g. learning, attention, letter recognition and face recognition) or deficits (e.g. dyslexia, amblyopia and aging). Furthermore, various techniques (e.g. reverse correlation) require adding noise. An important assumption underlying the application of these models and techniques is that adding external noise does not qualitatively alter the processing of the stimulus. In the present review, we argue that taken together, much data in the literature demonstrates a double dissociation (although never presented as such) between the processing in noiseless and noisy conditions. Based on such double dissociations and further evidence, we also argue that the application of popular models measuring the impact of internal noise leads to improbable interpretations. Alternatively, we conclude that adding noise to a stimulus qualitatively alters its processing. We further suggest that the application of these models should be revised. This compromises many interpretations of results when visual noise is used.

Keywords: Signal detection theory; Noise; Internal noise; Linear amplifier model; Perceptual template model

Contrast: A fundamental attribute

Our visual system is effective within environments varying over a large brightness range. Several functions of our visual system enable us to adapt to different intensities. At the input, the pupil partially normalizes the retinal illumination by varying its size. At the retinal level, photoreceptors, which are absorbing photons and producing electrical impulses, are more sensitive once adapted to the mean luminance intensity of the environment explaining why our sensitivity gradually improves over several minutes after walking in a dark room. At a higher retinal processing level, ganglion cells, which indirectly receive their input from several retinal receptors, are sensitive to a luminance difference (i.e. contrast) between their center and surround. As a result, these cells detect luminance variations between their center and surround, and are little affected by absolute luminance variation affecting both the center and surround (Hubel & Wiesel, 1961).

The axons of the ganglion cells form the optic nerve sending the retinal information to the visual cortex via the lateral geniculate nucleus. Consequently, at the optic nerve level, the visual information is mainly represented as center-surround contrast. The processes from the pupil to the ganglion cells enable the visual system to adapt to different environment intensities by converting luminance intensity information (photons entering the eye) into luminance variation (i.e. contrast) information. As a result, contrast is a fundamental attribute for the visual system and its processing has been widely studied and used to characterize various visual functions.

Measuring contrast sensitivity

At the level of the visual cortex, most neurons respond to bars of a particular orientation and frequency (Hubel & Wiesel, 1959). More specifically, the preferred stimulus (i.e. the receptive fields) of these neurons has the shape similar to a Gabor, which is a sine wave grating viewed from a Gaussian window. As a result, sine wave gratings are the most widely used stimuli to measure contrast sensitivity in psychophysics. Contrast sensitivity is defined as the inverse of the detection threshold, i.e. the smallest contrast detectable. Psychophysically, the task used to measure contrast sensitivity typically consists in presenting two stimuli either subsequently (temporal-forced-choice) or simultaneously (spatial-forced-choice) and only one of them contains a signal. The contrast of the other is set to 0 and the task consists in identifying the stimulus containing the signal. The contrast of the signal is manipulated to find the detection threshold corresponding to the smallest contrast detectable based on a given criterion, i.e. a proportion of correct answers.

Since contrast is a fundamental attribute, contrast sensitivity has been widely studied. For instance, contrast sensitivity was measured as a function of various parameters such as spatial frequency, orientation and stimulus size or duration (Campbell & Robson, 1968). Measuring contrast sensitivity as a function of the spatial frequency, which is typically referred to as the "contrast sensitivity function", are probably the most known results. We are typically more sensitive to spatial frequencies near 2 to 4 cycles/degree (cpd), which are therefore defined as medium frequencies. Our sensitivity gradually drops as the spatial frequency gets further from the medium frequencies. Figure IV-1 shows the contrast sensitivity function as a function of aging. Healthy aging affects contrast sensitivity at medium and high spatial frequencies (>2 cpd) but not at low spatial frequencies.

Contrast thresholds for higher cognitive tasks

Since the visual information entering the eye is converted into contrast information at the retinal level, the visibility of an image can be altered by manipulating its contrast. As a result, measuring the stimulus visibility required to perform a given task can be used to characterize visual functions. Thus, contrast thresholds are measured for discrimination, recognition or identification tasks rather than for detection tasks. In other words, instead of asking the observer to detect the stimulus containing a signal (detection task), only one stimulus is presented and the task consists in discriminating, recognizing or identifying a certain property of the stimulus. In either case, the dependant variable is the contrast required to correctly perform the task based on a given criterion level, i.e. proportion of correct answers. Such contrast threshold has been found useful to characterize various visual functions.

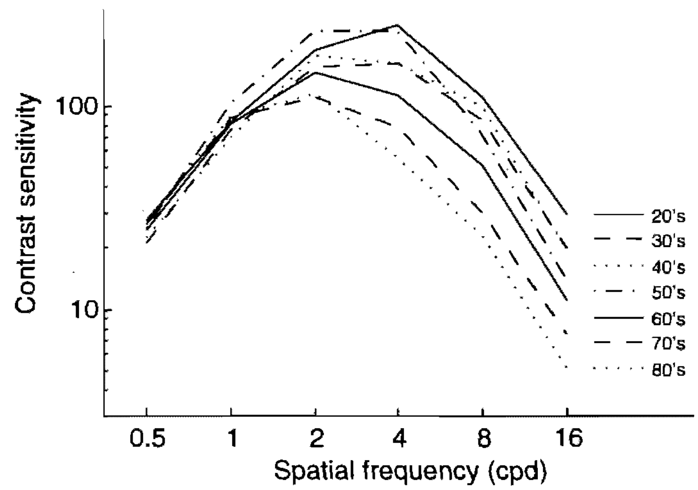


Figure IV-1. Contrast sensitivity as a function of spatial frequency for different age groups. Adapted from Owsley, Sekuler and Siemsen (1983).

For instance, motion processing has been characterized by measuring contrast thresholds. To characterize motion processing, contrast thresholds are often measured using a direction discrimination task. A sine wave grating drifting either to the left or to the right is presented and the task consists in discriminating its direction. The dependant variable is the contrast of the grating. The threshold is typically defined as the smallest contrast required to correctly discriminate the drifting direction. As a result, contrast threshold has often been measured as a function of the temporal frequency, i.e. the drifting speed, to characterize motion processing. Analogously to spatial frequencies, the contrast sensitivity as a function of the temporal frequency has a reverse u-shape typically peaking near 8-10 Hz.

Contrast thresholds are not just useful to characterize low level processing such as direction discrimination; it has also been used to characterize higher cognitive functions such as face identification or letter recognition. For instance, learning was found to have a great impact on contrast thresholds for face identification (Gold, Bennett & Sekuler, 1999). Since learning hardly affects detection thresholds, contrast threshold improvement for face identification as a function of learning is attributed to higher cognitive processing. Consequently, contrast thresholds measuring the visibility required to perform a given task are useful to characterize high cognitive functions, such as face identification or letter recognition, as a function of various parameters such as learning and attention.

Internal noise

As described above, contrast is a fundamental attribute and its processing has therefore been extensively studied. One of the goals pursued has been to identify and measure the factors limiting contrast sensitivity. A common view is that internal noise within the visual system is an important limiting factor. From the light entering the eye to the stimulus percept, the visual information is transformed as it passes from one processing level to another. During each of these transformations the signal is altered. These alterations or signal deteriorations can be seen as internal noise added to the signal limiting contrast sensitivity.

Signal detection theory

The signal detection theory (Green & Swets, 1966) suggests that detection thresholds are limited by internal variation of the inner representation of a given stimulus. Based on this theory, the contrast estimation of a given stimulus (internal representation) can be modeled by a scalar. Due to internal variations (i.e. noise), successive presentations of the same stimulus does not necessarily result in the same inner representation. Typically, it is assumed that the inner representation of a given stimulus at the decision level follows a normal distribution. The mean of this distribution depends on the stimulus intensity (in our case, the contrast) and the standard deviation models the internal noise.

Given two stimuli, there will be an inner distribution for each one (Figure IV-2). For a typical detection task, one stimulus corresponds to a blank field and the other contains a signal with a given contrast. For the blank field, the inner contrast representation depends only on inner variations (or noise). For the stimulus containing the signal, the inner contrast representation depends on the contrast and inner variations. It is generally assumed that both distributions have the same standard deviation. To perform the task, the observer uses a certain decision criterion to discriminate the two stimuli. Given the optimal decision criterion, the observer's performance will depend on two factors: the distance between the two means and the standard deviation of the distributions. The greater the efficiency of the observer to discriminate the two stimuli, the greater the distance between the two means will be. Since the inner contrast representation is not directly observable, the units of the inner representation are unknown and the means and standard deviation of the two distributions are not measurable. However, without knowing the means and standard deviation of the two distributions, it is possible to evaluate their ratio since the observer's performance will directly depend on it. As a result, the distance between the two means in standard deviation units (typically referred as the d') is measurable and corresponds to the signal-to-noise ratio of the inner representation at the decision level.

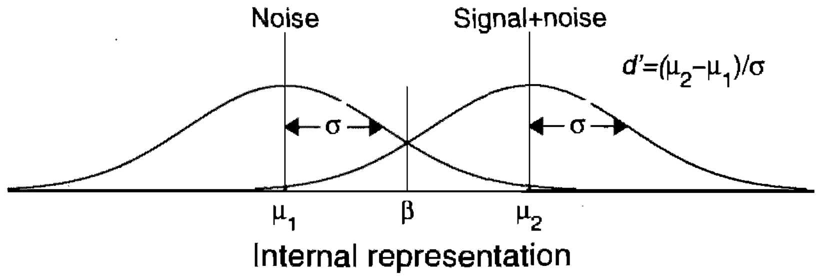


Figure IV-2. Internal representation of the contrast of two stimuli (noise only and signal plus noise) based on the signal detection theory (Green & Swets, 1966). σ **represents the internal variation of the internal representation (i.e. internal noise).** β **represents the optimal decision criterion: if the internal representation of the stimulus is below the decision criterion, the subject responds that the signal was not present, otherwise he respond that it was present.** μ_1 **and** μ_2 **represents the mean of the representation. The performance of the observer (*d'*) depends on the distance between the two means in standard deviation units.**

Measuring internal noise

As suggested by the signal detection theory, it is generally assumed that the performance depends on the signal-to-noise ratio of the internal representation. The visibility of a signal can therefore not only be altered by varying its contrast as described above but also by adding visual noise. Pelli (1981) proposed to measure the internal neural noise by varying the external noise added to the stimulus. If the external noise is lower than the internal noise, then the external noise will have no significant impact and the contrast threshold will not be significantly affected. If the external noise is greater than the internal noise, then internal noise will have no significant impact and contrast threshold will be limited by the external noise and should increase proportionally with the external noise contrast. As a result, contrast threshold as a function of the external noise contrast (TvC function) in log-log units will be initially flat for the portion where the internal noise is greater than the external noise and will rise with a slope of 1 when the external noise is greater than the internal noise as shown in Figure IV-3. The breaking point will correspond to the internal equivalent noise (IEN), that is, the external noise contrast having the same impact as the internal noise.

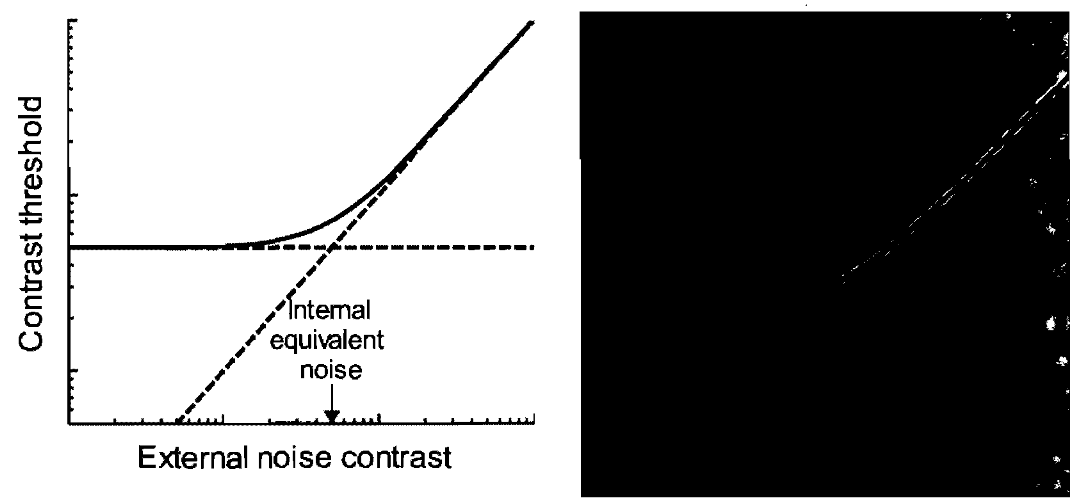


Figure IV-3. Contrast threshold as a function of external noise contrast. The flat asymptote corresponds to the contrast threshold in absence of noise and the other asymptote with a slope of 1 represents the signal-to-noise ratio required in high noise. The image on the right illustrates the visibility of the signal (vertical bars) as a function of the external noise contrast.

Assuming that the visual system adds noise to the stimulus, the contrast sensitivity can be seen as limited by two factors: the noise added to the signal and the signal-to-noise ratio required to detect the signal. In high external noise conditions, the internal noise has no significant impact so the detection threshold only depends on the signal-to-noise ratio required to detect the signal. Pelli (1981) defined the calculation efficiency (CE) as the signal-to-noise ratio of the ideal observer relative to the signal-to-noise ratio of the human observer. As a result, the greatest CE theoretically attainable is 1. By evaluating the detection threshold of a given signal embedded in external noise (e.g. Figure IV-4) as a function of the external noise contrast enables the decomposition of the sensitivity into two factors: the internal noise added to the signal by the visual system and the smallest signal-to-noise ratio the system requires to detect the signal. The former is evaluated by the IEN and the later by the CE. This model is generally referred to as the Linear Amplifier Model (LAM) and is illustrated in Figure IV-5.

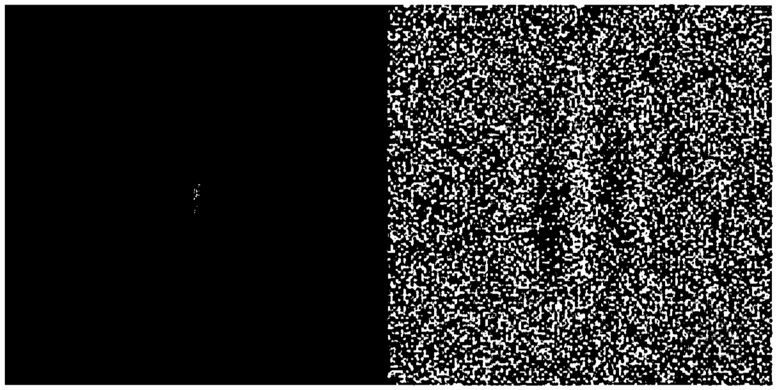


Figure IV-4. A Gabor (i.e. a sine wave grating viewed from a Gaussian window) in noiseless (left) and high noise (right).

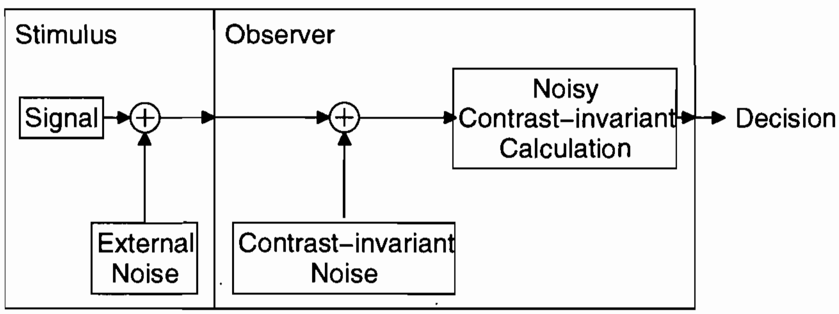


Figure IV-5. Linear amplifier model (adapted from Pelli (1990)). The stimulus is composed of a signal embedded in external noise. The performance of the observer depends on two factors: noise added to the stimulus by the visual system ("contrast-invariant noise") and the signal-to-noise ratio required to detect the signal ("noisy contrast-invariant calculation").

Mathematically, the detection threshold as a function of the external noise contrast is typically defined as (Legge, Kersten & Burgess, 1987, Pelli, 1981, Pelli, 1990):

$$dt(N_{ext}) = k\sqrt{N_{ext}^2 + N_{\text{int}}^2}$$

where N_{int} represents the IEN contrast and k is inversely proportional to the CE. As reviewed by Pelli (1990), this simple model was found to fit adequately many results. Figure IV-3 illustrates this function where in low external noise conditions, the detection thresholds remain relatively unaffected, but in high external noise conditions, the detection thresholds are proportional to the noise contrast (slope of 1 in log-log units). Note that both internal and external noises significantly affect detection thresholds only if the external noise contrast is near the IEN.

As a result, the LAM enables the decomposition of the sensitivity into two independent parameters (IEN and CE) by evaluating the contrast threshold as a function of the external noise contrast. Given a parameter affecting contrast sensitivity, decomposing the sensitivity into two independent parameters is useful to better characterize the sensitivity difference (Pelli & Farell, 1999). Indeed, models evaluating internal noise have been extensively used to characterize various visual functions affecting contrast sensitivity such as learning (Dosher & Lu, 2006, Gold et al., 1999, Gold, Sekuler & Bennett, 2004, Lu & Dosher, 2004b), attention (Dosher & Lu, 2000a, Dosher & Lu, 2000b, Lu & Dosher, 1998), aging (Bennett, Sekuler & Ozin, 1999, Pardhan, 2004, Pardhan, Gilchrist, Elliott & Beh, 1996, Speranza, Moraglia & Schneider, 2001) and dyslexia (Sperling, Lu, Manis & Seidenberg, 2005). It has also been used to characterize the sensitivity to signals defined by other attributes than luminance such as chromaticity (Gegenfurtner & Kiper, 1992) or texture contrast (Allard & Faubert, 2006) or higher cognitive functions such as letter recognition (Oruc, Landy & Pelli, 2006, Parish & Sperling, 1991, Pelli, Farell & Moore, 2003, Pelli, Levi & Chung, 2004) or face identification (Gold et al., 1999, Gold et al., 2004). Most of these results will be extensively described later in the present review.

Patterns of results

Given two differing conditions, three pure patterns of results may emerge when evaluating contrast thresholds as a function of external noise contrast (Lu & Dosher, 1998): contrast thresholds may differ only in low noise (Figure IV-6a), in both low and high noise (Figure IV-6b), or only in high noise (Figure IV-6c). We will refer to these patterns as low-, all- and high-noise patterns respectively.

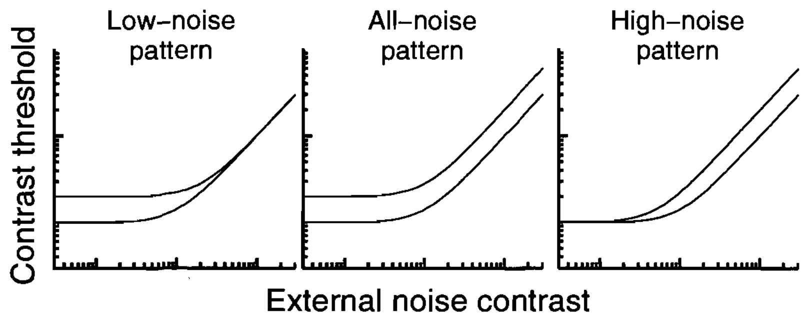


Figure IV-6. Three pure patterns of results for two conditions. Contrast thresholds can differ only in low-noise (left), in both low- and high-noise (middle) or only in high-noise (right).

Low-noise patterns

A low-noise pattern corresponds to a contrast threshold difference in low noise (i.e. difference of sensitivity) but no significant difference in high noise (Figure IV-6a). Based on the LAM, this difference is due to a difference of IEN since the internal noise only has an impact when the external noise is lower than the internal noise. In high external noise conditions, the impact of the internal noise is not significant and no contrast threshold difference is observed.

Such low-noise pattern has often been observed when manipulating the visibility of the stimulus, i.e. the quality of the displayed stimulus. For instance, reducing the mean luminance of the display (Figure IV-7) was found to increase the IEN without affecting the CE (Pelli, 1981) resulting in a low-noise pattern, i.e. contrast thresholds were affected in low- but not in high-noise when manipulating the display luminance. Analogously, the impact of a cataract which affects the retinal illumination was also found to affect contrast thresholds in low- but not in high-noise (Pardhan, Gilchrist & Beh, 1993).

As concluded by Pelli and Farell (1999), "Equivalent noise, being independent of task, invites explanation in terms of known properties of visual neurons: their density, gain, variance, and physiological thresholds" (p. 651-652). In other words, by evaluating detection threshold in high noise (i.e. evaluating the CE) it is possible to abstract many low-level parameters only affecting sensitivity at the level of the IEN.

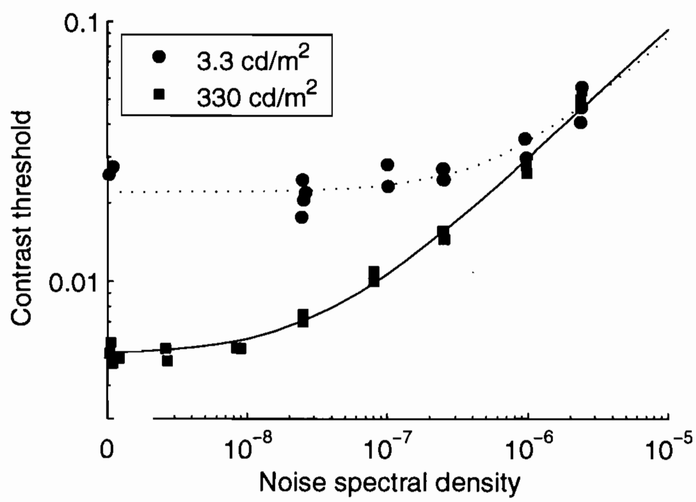


Figure IV-7. Contrast thresholds as a function of external noise contrast for two different background luminances. Adapted from Pelli (1981). Contrast thresholds significantly vary as a function of background luminance in low noise but not in high noise (low-noise pattern).

All-noise patterns

A difference of sensitivity due to different CEs would result in the all-noise pattern (Figure IV-6b). In other words, if, between two conditions, the sensitivity differs because of different signal-to-noise ratio required to detect the signal, then the detection thresholds would differ independently of the external noise contrast, i.e. whether the source of the noise is internal (low noise) or external (high noise). For contrast sensitivity, such pattern was found when modifying the window size (Pelli, 1981). Raghavan (1995) found similar IEN for various tasks such as identifying gratings, letters, or words. Indeed, the contrast thresholds differences between these tasks were found to be relatively constant as a function of the external noise contrast (all-noise patterns). All-noise patterns were also observed for higher perceptual tasks such as learning to identify faces (Gold et al., 1999, Gold et al., 2004). Pelli and Farell (1999) concluded that "Efficiency, being largely independent of viewing conditions, invites explanation in terms of the computation that combines the distributed stimulus and prior information to yield a decision." (page 652).

High-noise patterns

The third possible pattern (high-noise pattern, Figure IV-6c) is rather awkward since there is no difference of sensitivity (same contrast threshold in the absence of noise), yet the IENs and CEs differ. In one condition, the CE is greater and thereby increases contrast sensitivity but the IEN is also greater and decreases contrast sensitivity by the same proportion. Consequently, these two differences cancel one another in low external noise. In high external noise, the internal noise has no significant impact and the contrast threshold only depends on the CE. Thus, the difference of CEs results in different contrast thresholds. For contrast sensitivity, such pattern has been observed for learning (Dosher & Lu, 2006, Lu & Dosher, 2004b) , attention (Dosher & Lu, 2000a, Dosher & Lu, 2000b, Lu & Dosher, 1998), aging (Bennett et al., 1999, Pardhan, 2004, Pardhan et al., 1996, Speranza et al., 2001) and dyslexia (Sperling et al., 2005). We argue below that these results compromise the application of the LAM and the experimental results of most of these studies will therefore be extensively discussed below.

Separate mechanism hypothesis

It is generally assumed that adding external noise to a stimulus does not qualitatively alter its processing. Indeed, it is assumed that internal noise is always added to the stimulus. Consequently, even in the absence of external noise, the task can be seen as equivalent to signal processing in noise. The difference between contrast detection in low and high noise would be the origin of the effective noise source (internal and external respectively) affecting the stimulus visibility without qualitatively affecting the nature of the task. That is, the detection in low and high noise would be processed by the same mechanisms. Indeed, to apply the LAM, Pelli (1990) formulated an assumption proposing a "contrast-invariant calculation" of the "effective stimulus". The effective stimulus is the signal combined with the effective noise (internal + external). Therefore, to apply the LAM, Pelli defines the assumption that the mechanism detecting the signal is the same irrelevant of the noise source (internal or external). In other words, it is assumed that the same signal-to-noise ratio is required to detect whether the noise is internal or external. Consequently, based on this model, adding noise would quantitatively, and not qualitatively, alter the processing of the stimulus.

The present review questions the assumption that the same mechanisms are detecting a signal embedded in either internal or external noise thereby implying that the same signal-to-noise ratio is required to detect the signal embedded either in internal or external noise. Conversely, the present review suggests that adding external noise qualitatively alters the processing of the stimulus. In other words, it is not necessarily the same signal-to-noise ratio required to detect the signal whether the origin of the effective noise source is internal or external. Specifically, we propose that the processing of a signal embedded in high external noise is processed by higher perceptual mechanisms than the processing of the same signal in a noiseless (or low external noise) condition. If different mechanisms are processing signals in internal and external noise, this would force a reinterpretation of many results previously obtained from various experiments using external noise and would also compromise the application of the LAM and its variants.

One of the stronger arguments supporting the hypothesis that two stimuli are processed separately (at least at some point) is the observation of a double dissociation. To support our hypothesis, the present review highlights results reported in the literature that, although never explicitly presented as a double dissociation, actually satisfy its criteria. We will also present results suggesting that at least some stimuli processed in high external noise imply higher perceptual mechanisms than in a noiseless condition.

Double dissociation

To observe a double dissociation, the performance to two tasks (*a* and *b*) needs to be evaluated under two manipulations (*A* and *B*). A classical double dissociation occurs if manipulation *A* affects the performance to task *a* without affecting the performance to task *b*, and manipulation *B* affects the performance to task *b* without affecting the performance to task *a*.

To decompose the sensitivity into IEN and CE, the detection threshold of a given signal (i.e. stimulus) must be evaluated in different external noise contrasts. As mentioned above, it is generally assumed that the nature of the task is unaffected by the external noise contrast. However, the present review rather suggests that varying the external noise contrast qualitatively alters the nature of the task, i.e. qualitatively alters the processing of the stimulus. From a double dissociation perspective, the two tasks correspond to the detection of a given stimulus in low (<<IEN, zero slope on the TvC function presented in Figure IV-3) and high (>>IEN, slope of 1 on the TvC function presented in Figure IV-3) external noise contrast. Our approach consists in showing that it is possible to find two manipulations inducing a double dissociation for the same stimulus: a manipulation affecting contrast thresholds in low but not in high external noise (low-noise pattern, Figure

IV-6a) and another manipulation affecting, for the exact same stimulus, contrast thresholds in high but not in low external noise (high-noise pattern, Figure IV-6c).

Low-noise patterns

It is possible to define a manipulation a priori that would unequivocally result in a low-noise pattern for any task: adding another external noise source to both tasks having a fixed contrast greater than the IEN. The external noise added would affect the contrast threshold in low noise since the added external noise source would be greater than the IEN but would not significantly affect contrast thresholds in high noise since the added external noise source would not be significant when the other external noise source would be greater. Consequently, a low-noise pattern would certainly be observed.

Defining a manipulation a priori resulting into a single dissociation (low-noise pattern) valid for the contrast threshold of any tasks offers a great advantage: it allows inferring a double dissociation from any data resulting in the opposite single dissociation (high-noise pattern). Note however, as mentioned above, that low-noise patterns have been empirically observed in several conditions. For instance, manipulating the background luminance was found to affect thresholds in low but not in high noise contrast (Figure IV-7).

High-noise patterns

Since it is possible to induce a single dissociation (low-noise pattern) for any task consisting in processing a signal embedded in noise, showing a single dissociation in the opposite direction (high-noise pattern) results in a double dissociation between the detection of the stimulus in low and high noise. The objective of the present section is to show that high-noise patterns have empirically been observed as a function of various parameters, for various tasks and in various labs.

Attention

To our knowledge, Dosher and Lu (2000b) were the first to observe a high-noise pattern. They found such pattern when manipulating the observer's spatial attention. They used an orientation discrimination task in which four Gabor patches (spatial frequency of 1.12 cpd) embedded in noise were briefly presented (83 ms) around a centered fixation point. Each Gabor could have one of four orientations: 22.5, 67.5, 112.5 or 157.5 degrees. Observers had to report the orientation of one of the Gabors but they did not know before the stimulus presentation which Gabor was selected to perform the task, i.e. to report orientation. The report cue indicating on which Gabor the orientation should be reported was presented with the stimulus onset. The dependant variable was the contrast of the Gabor necessary to discriminate its orientation. At the stimulus onset, a cue near the fixation point indicated to the observer which Gabor had been selected. To modulate attention, 150 ms before the signal onset an arrow replaced the fixation point indicating to the observer which Gabor was the most likely (5 times out of 8) to be selected. As shown in Figure IV-8, a high-noise pattern was observed between valid and invalid pre-cueing. Contrast thresholds in unattended locations (i.e. when the selected Gabor was not the one previously cued by the arrow) were significantly higher (by a factor of about 1.32) when the task was performed in high noise but not in low noise. Dosher and Lu concluded that attention reduced the impact of the external noise. Their interpretation will be extensively discussed further below.

This high-noise pattern of results was replicated using various, but analogous, methodologies and stimuli. For instance, Lu and Dosher (2000) found similar results using a different target: a "+" sign with one of the four segments missing (" ", " ", " " or " "). The task was to discriminate which symbol was presented. Again, by manipulating spatial attention with a pre-cued arrow, no significant contrast thresholds were observed in the no or low noise, but significant differences were observed in high noise.

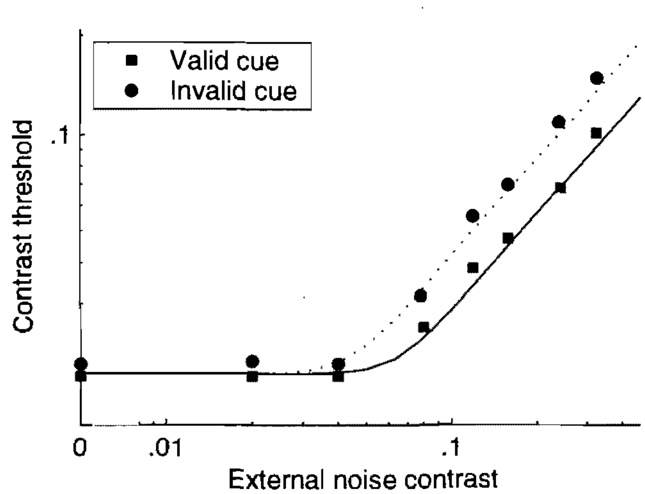


Figure IV-8. Effect of attention on contrast threshold of an orientation discrimination task. Adapted from Lu and Dosher (2004a) who were re-plotting their results from Dosher and Lu (2000b). Contrast thresholds significantly vary as a function of attention in high noise but not in low noise (high-noise pattern).

Learning

Lu and Dosher (2004b) also found a high-noise pattern as a function of learning. Contrast thresholds of a Gabor patch (1.34 cpd) were evaluated in an orientation discrimination task in which the observer had to discriminate between a Gabor tilted to 37 or 53 degrees. In each of the ten learning sessions, contrast thresholds were measured in different noise conditions. Each session was composed of 1440 trials.

Learning gradually improved contrast thresholds in high noise conditions without significantly improving thresholds in low (or no) noise conditions (Figure IV-9). Indeed, in the highest noise condition, the contrast threshold at the last two sessions was reduced by a factor of 1.5 compared to the two first sessions and in no external noise conditions, learning reduced contrast threshold by a factor of 1.04.

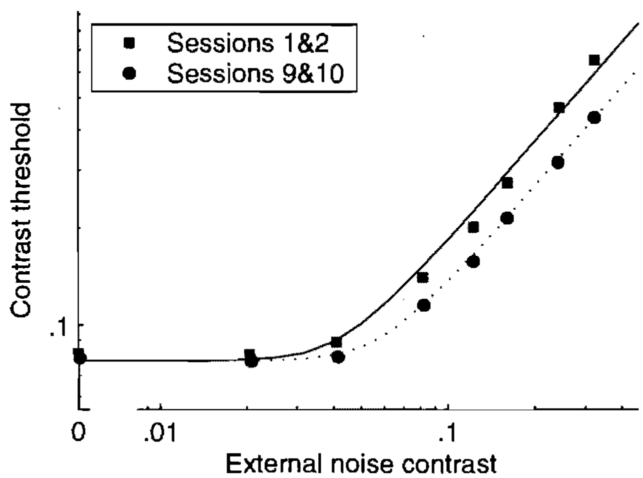


Figure IV-9. Effect of learning on contrast threshold of an orientation discrimination task. Adapted from Lu and Dosher (2004b). Contrast thresholds significantly vary as a function of learnging in high noise but not in low noise (high-noise pattern).

Since learning only significantly improved contrast thresholds in high noise conditions, by applying the LAM Lu and Dosher found that learning increased both the CE and IEN by similar proportions. They therefore conclude that the LAM leads to improbable interpretations, since it would be unlikely that learning increases IEN. They concluded, like when manipulating attention, that learning improves the ability to exclude external noise. Again, their interpretation will be extensively discussed below.

Betts, Sekuler and Bennett (2007) evaluated young and older observers contrast thresholds for an orientation discrimination task. Although there objective was not to evaluate the impact of learning on contrast threshold, they collected their data for each group over two days. A statistical analysis (ANOVA) revealed a simple main effect of the testing day. On average, the IEN and CE increased by similar proportions (a factor of about 1.25) between the two testing days. In other words, practice did not significantly affected contrast thresholds in low noise but improved it in high noise. Following Lu and Dosher, they proposed that practice improved the ability to excluded external noise.

Aging

Bennett, Sekuler and Ozin (1999) evaluated the impact of aging on a contrast detection task in low and high noise for different spatial-frequency uncertainties. The contrast detection to three spatial frequencies (1, 3 and 9 cpd) in a two-spatial-forced-choice procedure, i.e. two stimuli (one containing only noise and the other containing both a signal and noise) were presented simultaneously and the observer had to identify which one contained the signal. In the no frequency-uncertainty conditions, observers knew the frequency of the signal which was constant within a block of trials. In the frequency-uncertainty conditions, observers did not know the spatial frequency of the signal which varied between trials.

No important difference of uncertainty effect was found between the two groups. Because they are beyond the scope of the present review, the small differences within groups due to the frequency-uncertainty will not be discussed. Instead, we will focus on the data they obtained in low and high noise for the two age groups (median of 22 and 68 years old). In low noise, a contrast sensitivity difference (detection threshold in absence of external noise) was observed only at 9 cpd between the two groups. For the 1 and 3 cpd spatial frequencies, the sensitivities were similar between the two groups. These results are not surprising based on the known impact of aging on the contrast sensitivity function: aging affects the contrast sensitivity to high spatial frequencies without affecting contrast sensitivity for low spatial frequencies as presented above (Figure IV-1). In high noise, aging affected the contrast detection thresholds in similar proportions for all three spatial frequencies by factors of 1.45 and 1.66 in the no-uncertainty and the frequency-uncertainty conditions respectively. The results observed for 1 and 3 cpd gratings are evidence of a classical single dissociation (high-noise pattern): aging did not affect contrast thresholds in low-noise and impaired contrast thresholds in high-noise.

Pardhan (2004) compared IEN and CE between a young (17 to 22 years) and an old (60 to 72 years) group. For 1 cpd stimuli, she found that aging had no significant impact on contrast sensitivity (i.e. detection thresholds in noiseless conditions) as it is usually observed for low spatial frequencies (Figure IV-1). By decomposing the sensitivity into IEN and CE, she found a significant decrease of CE with aging but no significant difference of IEN between the two groups. How could that be? If the CE is lower and there is no difference of IEN, then the sensitivity should be reduced. A closer look at her analysis reveals that she compared contrast sensitivities in log units as it is usually (and should be) done, but her fit and statistical analysis on IEN and CE were performed in linear units, which increases the weights of higher thresholds. Indeed, her graph shows a poor fit at the contrast sensitivity for the older group: the fit in noiseless condition estimates a threshold greater than all the thresholds of the older observers. Note that Speranza, Moraglia and Schneider (2001) also found analogous results using similar methodologies and analyses when comparing contrast thresholds of young and old observers for spatial frequencies of 1.1 and 2.2 cpd. For both studies, visually examining their graphs suggests that there is no threshold difference in low noise (as they found by comparing the data in log units) but the older group showed significantly higher thresholds (as they found when comparing CEs). We therefore conclude that their data are consistent with what has been previously observed (see above) with aging for similar spatial frequencies: higher thresholds in high noise without any significant effect in low noise (high-noise pattern).

Dyslexia

Sperling, Lu, Manis and Seidenberg (2005) also found a high-noise pattern when comparing the performance to a contrast detection task between dyslexic (i.e. children with reading disabilities) and non-dyslexic children. The stimuli were vertical sinusoidal gratings viewed through a Gaussian window (i.e. Gabor patterns). Both groups had similar contrast thresholds in noiseless conditions, but the dyslexic group had higher contrast thresholds in high external noise conditions. The authors conclude that dyslexic children are less effective at excluding external noise and that the reduced effectiveness to exclude external noise could contribute to reading disabilities. As mentioned above, such interpretation will be discussed below.

Amblyopia

Amblyopia is a developmental visual disorder characterized by lower visual acuity even though no known ocular anomalies are diagnosed. It is generally suggested that amblyopia is caused by important differences between the two eyes generally due to anisometropia or strabismus. It is now generally admitted that amblyopia causes relatively high processing deficits (i.e. extra-striate) that cannot be explained by low-level deficits such as contrast sensitivity loss (Lerner, Pianka, Azmon, Leiba, Stolovitch, Loewenstein, Harel, Hendler & Malach, 2003, Sharma, Levi & Klein, 2000, Simmers, Ledgeway & Hess, 2005, Simmers, Ledgeway, Hess & McGraw, 2003). Since this deficit is cortical and is known to affect high-order functions, it has been widely studied as a function of various tasks such as first- vs second-order motion processing (Simmers et al., 2003) or multiple-object tracking (Ho, Paul, Asirvatham, Cavanagh, Cline & Giaschi, 2006).

Pelli, Levi and Chung (2004) measured contrast thresholds for letter identification performed either in a noiseless or noisy backgrounds as a function of the letter size. When normalizing for individual acuity (letter size relative to each observer's acuity), they found similar contrast thresholds as a function of acuity-normalized letter size between normal and amblyope observers. This suggests that after compensating for each observer's acuity, letter processing for amblyopes is normal. When performing the exact same task in visual noise, amblyopes had greater (i.e. worst) contrast thresholds. These results cannot be directly interpreted as a high-noise pattern since the task performed by the normal and amblyope observers is not exactly the same since the letter size were scaled relative to the visual acuity which is lower for amblyopes. Consequently, it is theoretically possible that amblyopes have lower CEs for letter identification but also have lower IENs at the tested spatial frequency which depends on the letter size. Nonetheless, the authors also evaluated the IENs and CEs without normalizing for acuity. They found that the CEs were lower for amblyopes and that the difference between the two groups decreased with the letter size. For large letters, amblyopes had similar IEN but for small letters they had less IEN. In other words, for the identification of small letters, amblyopia increased contrast thresholds in high noise more than in low noise. The authors were surprised and had trouble explaining the lower IEN obtained for amplyopes. The authors could think of two possible factors that could explain the lower IEN for amblyopia observers when viewing small letters: the pupil size and modulation transfer function of the optics of the eye. They therefore measured, in the same testing conditions, the pupil size of a few amblyopes which is known to be greater. However, the difference between the pupil size of the amblyopes and normal observers was not significant. It would have had to be large to explain the IEN reduction. They therefore deduced that the modulation transfer function was underestimated. However, they did not directly test this hypothesis.

Even though these results do not correspond to a pure high-noise pattern for small letters (contrast thresholds in low noise was still affected by amblyopia but to a lower extent than in high noise), we conclude that they suggest distinct processing in low and high external noise. Indeed, the fact that amblyopia affects contrast thresholds more, for small letters, in high noise than in low noise suggests that adding noise does not only remove the impact of a factor (internal noise), it also alters the processing.

Discussion

These examples show that, for contrast thresholds, it is possible to empirically find high-noise patterns (single dissociations). In other words, many results from different labs showed that in some conditions, thresholds in high noise are affected but not thresholds in low noise. Combining these results with the single dissociation deduced *a priori* above (low-noise pattern), these examples result in classical double dissociation cases. We therefore conclude that, for such tasks, it is possible to find manipulations affecting contrast thresholds in low but not in high noise and other manipulations affecting contrast thresholds in high but not in low noise, resulting in a classical double dissociation.

Distinct signal-to-noise ratio required in low and high noise

Typically, evidence of a double dissociation is interpreted as a strong argument supporting the hypothesis that the two tasks are processed, at least partially, by separate mechanisms. However, a double dissociation does not necessarily imply this hypothesis. As argued by Dunn and Kirsner (2003), a double dissociation implies that at least two processes are involved but does not necessarily imply that both tasks are processed by separate mechanisms. Indeed, if the accomplishment of either task necessitates two common processes but the effectiveness of each process affects the two tasks by different proportions, it is theoretically possible to observe a double dissociation without inferring that the two tasks are processed by separate mechanisms. The impact of increasing the effectiveness of one process and reducing the effectiveness of the other could completely cancel one another for one task but not for the other resulting into a single dissociation. Modifying the effectiveness of the processes with a different combination could also result in a single dissociation in the opposite direction resulting into a double dissociation.

One of the assumptions of the LAM is that signal-to-noise ratio required to detect the signal is the same whether the noise is internal or external. In other words, a contrast-invariant mechanism would be processing a signal embedded in internal or external noise. Based on this assumption, knowing the signal-to-noise ratio required to detect the signal (CE, i.e. contrast thresholds in high noise) and knowing the signal contrast required to be detected in internal noise (contrast sensitivity, i.e. contrast thresholds in low noise), it is possible to deduce the IEN. The present review questions this assumption and proposes that different mechanisms are extracting the signal from either internal or external noise. This separate mechanism hypothesis can easily explain the double dissociations observed between the processing in low and high noise.

However, since a double dissociation does not necessarily imply the separate mechanism hypothesis, we investigate, in the present section, the plausibility that the same mechanism is extracting the signal from internal and external noise. Using a proof by contradiction, we conclude that it is not the same mechanism extracting the signal from noise in low and high noise. We therefore start by assuming that the same mechanism is extracting the signal from internal and external noise and then show that this assumption leads to improbable interpretations. We therefore conclude that the double dissociations observed suggest that separate mechanisms are extracting the signal from internal and external noise.

A process only effective in high-noise

The double dissociations show that it is possible to affect contrast threshold in low noise without affecting thresholds in high noise (low-noise pattern) and it is also possible to affect contrast threshold in high noise without affecting thresholds in low noise (high-noise pattern). The low-noise pattern can easily be explained by internal noise which would only have an impact in low noise. In other words, as suggested by the LAM, adding noise eliminates a factor influencing contrast thresholds, i.e. internal noise, and thereby enables to decompose contrast sensitivity into two factors: IEN and CE. The LAM therefore assumes that all the factors having an impact in high noise also have the same impact in low noise. Below, we question this assumption.

Indeed, high-noise patterns challenge the application of the LAM. If we suppose that all the factors having an impact on contrast thresholds in high noise also have an impact in low noise, then we must conclude that high-noise patterns are caused by the alteration of at least two factors: one effective in low and high noise (e.g. CE, i.e. the signal-to-noise ratio required to detect the signal) and another canceling the impact of the other in low noise (e.g. IEN). As a result, to explain a high-noise pattern there would be two factors altered in low noise: one decreasing the contrast thresholds and the other increasing it by the same proportion. In other words, a high-noise pattern would be caused by the conjunction of a low- and an all-noise pattern. Consequently, it is theoretically possible to explain the high-noise patterns observed for learning, attention, aging and dyslexia without supposing the existence of a factor having and impact only in high noise. Thereby, it is theoretically possible to explain these double dissociations without inferring the separate mechanism hypothesis. However, it is improbable that learning and attention decrease the effectiveness of a factor (based on the LAM, the IEN), and that aging, dyslexia and amblyopia increase the effectiveness of a factor.

Note that Lu and Dosher (2004b) used similar arguments to criticize the application of the LAM arguing that its application leads to improbable interpretations. Based on the LAM, Lu and Dosher would have had to conclude that, in their conditions, learning and attention increased the CE but also increased, in the same proportions, the IEN. Analogously, we could also naively apply the LAM to the aging and dyslexia results mentioned above. From the LAM's perspective, we would conclude that aging and dyslexia reduced the CE but also reduced, in the same proportion, the IEN. Based on the LAM, we would also conclude that amblyopia reduces the IEN. In agreement with Lu and Dosher, we conclude that, although the LAM could theoretically explain these patterns of results, the interpretations induced are improbable: learning and attention would increase IEN, and aging, dyslexia and amblyopia would decrease IEN.

Instead, we conclude that at least one factor limiting contrast thresholds in high noise has no significant impact in low noise. In other words, high-noise patterns are caused by the alteration of a factor limiting contrast thresholds exclusively in high noise and not by a factor limiting contrast thresholds in low and high noise (all-noise pattern) combined with another factor canceling the effect in low noise (low-noise pattern). This conclusion compromises the application of the LAM which assumes that all the factors limiting contrast thresholds in high noise also limits contrast thresholds in low noise.

External noise exclusion

Concluding that a certain factor only has an impact in high noise does not necessarily imply that different mechanisms are extracting the signal from noise (i.e. different signal-to-noise ratios are required) in low and high external noise. In high noise, the internal noise is not significant and the performance only depends on the signal and external noise contrasts. Since affecting the signal strength would affect contrast thresholds in low noise, a mechanism only affecting thresholds in high noise would have to affect the external noise without affecting the signal strength. Consequently, if we assume that it is the same mechanism extracting the signal from internal and external noise (i.e. the same signal-to-noise ratio is required in low and high noise) and that such mechanism is unaffected (because no contrast threshold difference is observed in low noise) then the difference exclusively observed in high noise (high-noise patterns) must be due to a mechanism affecting external noise without affecting the signal.

Noise is typically defined as random fluctuation. As a result, once combined, two noise sources become indistinguishable form one another. If a given process only affects one of the two noise sources, then it must be effective before the two noise sources are combined. Consequently, a mechanism only affecting external noise must be effective before the main internal noise source is added.

Note that any processing affecting the signal would also affect the noise at the signal spatiotemporal location and frequency by the same proportion since both are merged (even an ideal observer is affected by noise). Indeed, a contrast gain affecting the signal and noise would have no impact on the signal-to-noise ratio and would thereby have no significant impact on contrast thresholds. As a result, the noise exclusion mechanism cannot be effective at the signal (i.e. at its spatiotemporal frequency and spatiotemporal location). However, noise does not only have energy at the signal spatiotemporal frequency or location. A noise exclusion mechanism could reduce the impact of external noise that is not at the spatiotemporal frequency or location without affecting the signal. If the noise that is not at the signal frequency and location affects performance of a human observer (even though it would not affect the performance of an ideal observer), then reducing the impact of this noise would reduce contrast thresholds in high noise. Improving the efficiency of such mechanisms would result in a high-noise pattern. Consequently, the impact of external noise would be reduced in high noise conditions but not in low external noise conditions resulting into a high-noise pattern. In other words, the external noise exclusion mechanism would reduce the impact of irrelevant and potentially disturbing information such as external noise near (but not at) the signal spatiotemporal frequency or location. Based on a priori knowledge of the signal, an external noise exclusion mechanism could have different contrast gains between the signal and the noise near (but not at) the signal.

A model with an external exclusion mechanism has been proposed by Lu and Dosher (1998): the perceptual template model (PTM). As shown in Figure IV-10, the PTM has several parameters and most are irrelevant to the purpose of the present review. We will therefore only consider the two relevant to the purpose of the present review: an external noise exclusion parameter and an additive internal noise. Altering the former results in a high-noise pattern and altering the later results in a low-noise pattern.

According to Lu and Dosher, learning and attention would enable the noise exclusion mechanisms to ignore irrelevant information such as noise at a different spatiotemporal frequency than the signal and only keep the information near the signal. Since this mechanism is effective before the additive internal noise, it only reduces the impact of external noise without altering the impact of additive internal noise. Consequently, this parameter only significantly affects contrast thresholds in high external noise and varying it results into a high-noise pattern (Figure IV-6c). Based on the same reasoning, aging, dyslexia and amblyopia would affect the ability to exclude external noise.

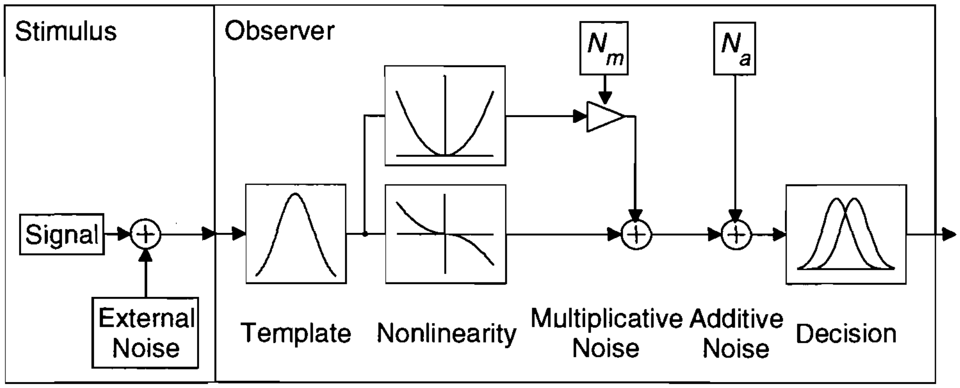


Figure IV-10. The Perceptual Template Model (PTM). Adapted from Lu and Dosher (2000). N_m and N_a represents multiplicative and additive internal noise. The present review focused on two parameters of the observer's performance: the template excluding external noise only has a significant impact in high external noise conditions, and the additive noise, which only has a significant impact in low external noise conditions.

As described above, a priori knowledge about the signal would help the observer to focus only on relevant information by ignoring (decrease the contrast gain of noise near, but not at, the signal) irrelevant information at spatiotemporal frequencies or locations different than the signal. Such filtering performed by the external noise exclusion mechanism could also be artificially applied to the stimulus. Consequently, filtering the external noise to remove irrelevant information would reduce the impact of the external noise exclusion mechanism by reducing the noise to exclude.

In a series of experiments, Lu, Dosher and colleagues systematically investigated this hypothesis (Dosher, Liu, Blair & Lu, 2004, Lu & Dosher, 2004c, Lu, Jeon & Dosher, 2004, Lu, Lesmes & Dosher, 2002). They modulated endogenous attention of observers performing an orientation discrimination task. They evaluated the impact of attention in high noise as a function of various filter applied to the noise. Noise filtering was either spatially (i.e. occurring more or less near the signal in the Euclidian domain), temporally (i.e. occurring more or less at the same time as the signal) or in the spatiotemporal frequency domain (i.e. more or less near the signal in the Fourier domain). In all cases, filtering the noise had no important impact on the contrast thresholds difference between the attended and unattended conditions. In other words, even when filtering the noise as a function of any dimension, attention always improved the performance. If there was an external noise exclusion mechanism, this mechanism should exclude noise along some dimension without affecting the signal strength to explain the high-noise patterns observed. However, along any dimension it was impossible to observe an important external noise exclusion reduction when filtering the external noise.

Lu, Jeon and Dosher (2004) concluded that external noise exclusion mechanisms "uniformly reduce the gain to external noise ... without affecting the gain to the signal stimulus" (p. 1347). For instance, when manipulating the spatial frequency of the external noise, Lu and Dosher (2004c) conclude that the external noise exclusion mechanism "excludes external noise uniformly across all the spatial frequencies without changing the spatial frequency selectivity of the perceptual template" (p. 955). However, the external noise and the signal are merged. We do not see how a mechanism could uniformly affect the noise at all the spatiotemporal locations and frequencies (including the spatiotemporal frequency and location of the signal) without affecting the signal. In other words, they attribute some contrast threshold variations to an external noise exclusion mechanism affecting only the external noise instead of attributing this variation to the mechanism extracting the signal from noise. However, empirical results force them to conclude that the external noise exclusion mechanism is effective at the signal spatial frequency without affecting the signal. We find this conclusion improbable and find more parsimonious the hypothesis that effects observed in high noise is due to the effectiveness of the mechanism extracting the signal from noise. In other words, the different contrast thresholds observed in high noise as a function of various parameters (e.g. attention, learning, aging and dyslexia) would be due to different signal-to-noise ratio required to detect the signal and not to the exclusion of the external noise *at* the signal. If the signal-to-noise ratio is altered in high noise but not in low noise, then we are forced to conclude that it is not the same mechanism extracting the signal from internal and external noise. In other words, high-noise patterns would not be due to a mechanism exclusively affecting external noise without affecting the signal or internal noise, but would rather be due to a different signal-to-noise ratio required to extract the signal from noise which would be exclusively effective in high noise.

Discussion

The goal of the present section was to investigate if the same mechanism is extracting the signal from either internal or external noise. The approach we used was to first assume the common signal-extraction mechanism hypothesis and then show that such hypothesis leads to improbable interpretations.

If it is the same mechanism extracting the signal from internal or external noise, then when a manipulation affects contrast thresholds in high noise and not in low noise, this high-noise pattern can either be due to the combination of a low-noise and all-noise pattern (as suggested by applying the LAM) or due to a factor having an impact on the external noise without affecting the signal or internal noise (as suggested by applying the PTM). By being resistant to a double dissociation (i.e. explaining low- and high-noise patterns without inferring distinct mechanisms), the LAM and the PTM suggest theoretically possible interpretations of high-noise patterns. However, either model leads to improbable interpretations. In certain conditions, the LAM would suggest that learning and attention would increase the IEN, and that aging, dyslexia and amblyopia would decrease the IEN (in many cases, by the same proportion as the CE). The PTM would suggest that it is possible to exclude the external noise at the signal spatiotemporal frequency and location without affecting the signal.

Based on these cumulative evidences, we find it more parsimonious to conclude that both tasks (detection in low and high external noise) are processed by distinct mechanisms, rather than using a model resistant to a double dissociation leading to improbable interpretations. Based on the double dissociation principal, we suggest a simpler hypothesis: separate mechanisms are extracting the signals embedded in internal (i.e. no or low external noise) and high external noise. This new interpretation severely compromises the application of these models used to characterize visual functions and thereby suggests a reinterpretation of various important results.

Contrast thresholds in high noise

The double dissociation between processing in low and high noise suggests that adding noise does not only change the noise source affecting contrast thresholds (external in high noise versus internal in low noise) without affecting the nature of the task but rather suggests that adding noise to the stimulus qualitatively alters the nature of the task. In order to characterize the processing differences relative to the external noise contrast, we examined whether the processing in high noise is more complex and solicits higher perceptual mechanisms than the processing of the same stimulus in low noise.

Independent to low-level alterations

As summarized in the "Low-noise patterns" subsection of the "Patterns of results" section above, by evaluating contrast threshold in high noise it is possible to abstract many low-level parameters affecting contrast threshold in noiseless conditions. In other words, adding external noise provides a task in which distortions added by the visual system (i.e. internal noise) have no significant impact.

Furthermore, any contrast gain affecting the signal and the external noise has no impact on the signal-to-noise ratio and thereby do not affect contrast thresholds. Indeed, the fact that both the signal and external noise utilizes the same pathways implies that any alterations applied to the signal will also be applied to the noise at the signal. As a result, increasing the contrast of the stimulus (i.e. of both the signal and noise) has no impact on the signal-to-noise ratio and thereby does not affect contrast thresholds. This is confirmed by the slope of 1 in log-log units of the TVC function in high-noise (Figure IV-3): increasing the external noise contrast by a given proportion affects the contrast threshold by the same proportion. Alternatively, in noiseless conditions, a contrast gain occurring before the main internal noise source would directly affect the signal-to-noise ratio and thereby affect contrast thresholds.

As a result, contrast thresholds in high noise are unaffected by internal noise and by any contrast gain parameters. Contrast thresholds in high external noise are therefore, as concluded by Pelli and Farell (1999) and discussed above, independent of many low-level alterations affecting the visibility of the stimulus (and thereby contrast thresholds in low noise) such as internal noise or contrast gains. Thus, adding external noise enables to abstract many low-level factors (e.g. ocular factors such as cataracts, display luminance, contrast gains) that influence contrast thresholds in low noise.

High-noise patterns: high-level manipulations

Above, four pure high-noise patterns were observed in certain conditions as a function of learning, attention, aging and dyslexia. In the present section, we investigate, for each of these variables, the processing level at which they have an impact. In other words, we examine whether these conditions typically affect low or high perceptual tasks.

Attention

Above, we cited results obtained by Lu and Dosher about the impact of attention on contrast thresholds. Attention was spatially altered by giving a pointing cue indicating the most likely target location to be the one selected to report orientation. Lu and Dosher (2000, 2004a) described this manipulation to activate the endogenous attention system, i.e. top-down attention. Indeed, the cue indicates the observer where to focus his attention. This cue was not near the stimulus needed to be processed to perform the task. Consequently, to affect the attention level to the pointed target, the pointing sign must first be interpreted. Therefore, as described by Lu and Dosher, such attention corresponds to top-down attention. As a result, endogenous (or top-down) attention reduced contrast thresholds in high but not in low external noise.

Lu and Dosher (2000) also manipulated attention by presenting the cue near the stimulus location which involuntarily increased the spatial attention near the cue. They described such attention as activating more the exogenous attention system (bottom-up). They found that exogenous attention system had an impact in both low and high noise conditions (all-noise pattern). In other words, top-down attention improved thresholds in high noise and bottom-up attention improved thresholds in both low and high noise. Consequently, they attributed threshold reduction in low noise to the exogenous attention system (bottom-up) and threshold reduction in high noise mainly to the endogenous attention system (top-down). This suggests that processing in high noise corresponds to a higher perceptual task than in low noise.

Learning

Although plasticity occurs at all processing levels, it is more prominent for complex tasks than for simple tasks. Consequently, finding an improvement in high but not in low noise (i.e. the high-noise pattern presented above) suggests that the task performed in high noise utilizes higher perceptual mechanisms.

Furthermore, in their study on learning described above, Lu and Dosher (2004b) also evaluated if the learning transferred to a different visual scale. After the 10 days of practice, observers were trained for 6 days to the same task with the exception that the viewing distance was varied by a factor of 2. Recall that learning had no impact on contrast thresholds in low noise but reduced thresholds in high noise. Changing the viewing distance resulted in changing the stimulus' spatial frequency by 1 octave. Over 6 days of training, no substantial improvement was observed in both low and high noise conditions. For the low noise conditions, these results are not surprising since learning did not have a significant impact over the first 10 days. However, the fact that no learning occurred in high noise suggests that the learning completely transferred between the two spatial frequencies. This learning transfer also suggests the implication of relatively high processing mechanisms. The learning transfer observed between the processing of two spatial frequencies suggests the existence of a common mechanism extracting the signal from noise at both scales.

Aging

Aging is known to affect complex tasks (Faubert, 2002) and peripheral or low-level mechanisms. However, the fact that, as described above, aging has no impact on contrast threshold for low spatial frequencies suggests that peripheral or low-level mechanisms are not limiting factors for low spatial frequencies. The observation of high-noise patterns as a function of aging therefore suggests that processing in noise is more complex.

Dyslexia

Sperling, Lu, Manis and Seidenberg (2006) compared poor and good readers when performing a coherent-motion task. In this task, there are a certain percentage of dots moving in a coherent direction (the signal) while the other dots are moving in random directions (the noise). The dependant variable was the percentage of dots moving in a coherent direction necessary to perceive the motion direction. In one conditions, both the signal and noise dots were defined by the same color (hard task), and in the other condition, they were defined by different colors helping the observer to segregate the signal and noise (easy task). They found that the poor readers had the same coherence threshold for the easy task but had high threshold for the hard task. They conclude that the poor readers did not have a motion perception deficit but their ability to exclude external noise (i.e. dots moving in random directions) was reduced. In other words, the lower coherence threshold for the poor readers would not be caused by low level factors limiting the perception of motion, but to the ability of the observer to segregate, at a high processing level, the signal and noise elements. Consequently, the fact that dyslexic children have higher contrast thresholds in high noise and have similar contrast thresholds in low noise as non-dyslexic children is consistent with the hypothesis that contrast thresholds in high noise is more complex.

Discussion

Since these manipulations (learning, attention, aging and dyslexia) are known to affect more high perceptual tasks than low perceptual tasks and that contrast thresholds in noiseless conditions are known to be a relatively low-level task, we conclude that processing in high-noise is more complex and is performed by higher perceptual mechanisms than processing in low (or no) noise. The fact that high-noise patterns were observed for learning, top-down attention, aging and dyslexia, suggests that adding noise increases the complexity of the task.

Attribute-invariant

The present review proposes another strong argument suggesting that contrast detection in noise is processed by relatively high level mechanisms: contrast thresholds are attribute-invariant. Indeed, changing the attribute defining the signal and noise did not affect contrast thresholds as demonstrated by the studies presented below.

Luminance vs chromaticity

Gegenfurtner and Kiper (1992) evaluated luminance and chromatic detection thresholds in luminance and chromatic noise. Based on their data, it is possible to find a noise contrast level for each attribute affecting the detection when the signal and noise are composed of the same attribute without affecting the detection when they are composed of different attributes. Indeed, at a given contrast level, luminance noise affected luminance detection without affecting chromatic detection and chromatic noise affected chromatic detection without affecting luminance detection. This double dissociation confirms, as it is generally admitted, that both attributes are not processed by the same mechanisms.

When both signal and noise were defined by the same attribute, they found the same detection threshold in high noise (low-noise pattern; Figure IV-11). In other words, luminance and chromatic signals necessitated the same signal-to-noise ratio to be detected: observers were just as efficient at detecting a luminance signal in luminance noise as detecting a chromatic signal in chromatic noise. This suggests that the same mechanism is extracting luminance signals embedded in luminance noise and chromatic signals embedded in chromatic noise.

Luminance vs contrast

Analogous results were obtained for first- and second-order processing. Allard and Faubert (2007) evaluated the detection threshold of luminance- (LM) and contrast-modulated (CM) signals in LM and CM noises. They found no (or little) cross-attribute interaction suggesting that both attributes are processed by separate mechanisms: LM noise had no or little impact on CM detection and CM noise had no or little impact on LM detection.

Although the contrast thresholds differed in the absence of noise, they did not significantly differ in high noise when the signal and noise were defined by the same attribute (Allard & Faubert, 2006). In other words, observers had the same detection threshold in high-noise for both attributes (low-noise pattern; Figure IV-12) suggesting that common mechanisms are extracting LM signals in LM noise and CM signals in CM noise.

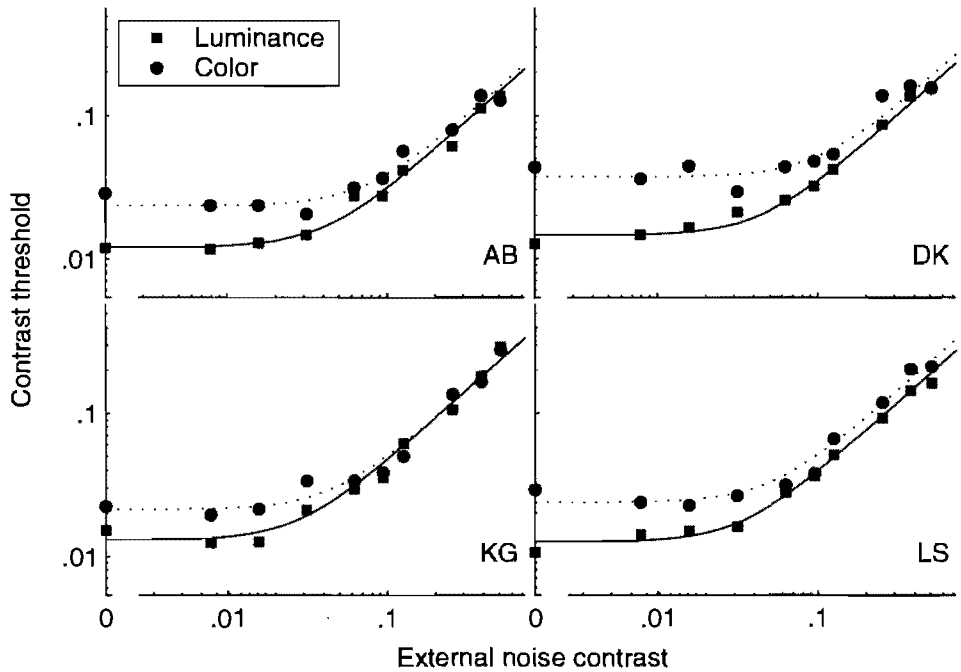


Figure IV-11. Threshold as a function of external noise contrast for luminance and color attributes. Signal and noise are defined by the same attribute. In high external noise conditions, observers are just as efficient at detecting a luminance signal embedded in luminance noise as detecting a chromatic signal embedded in chromatic noise. Adapted from Gegenfurtner and Kiper (1992).

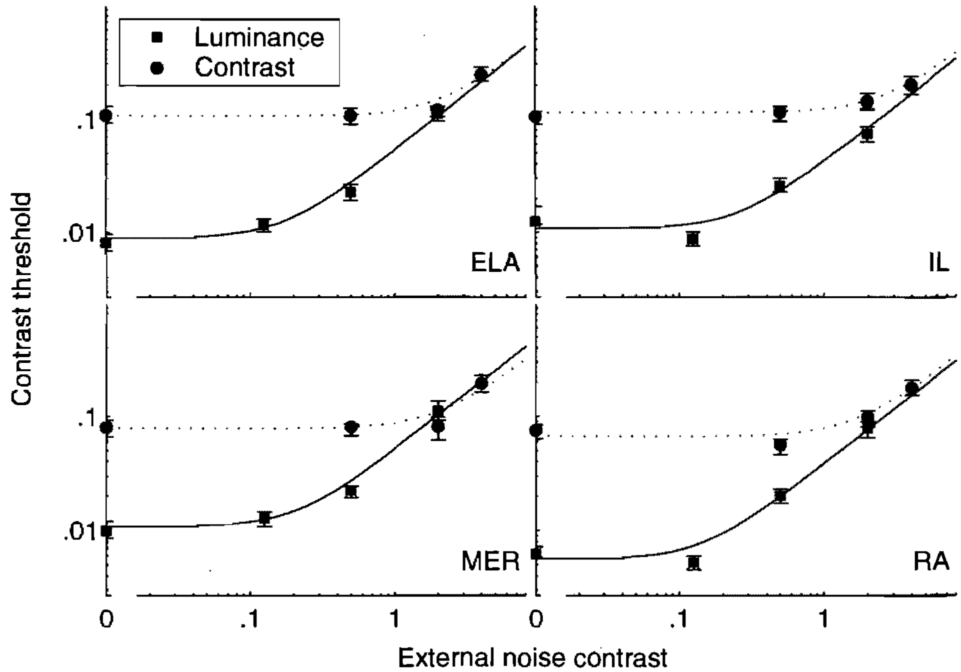


Figure IV-12. Threshold as a function of external noise contrast for luminance and contrast attributes. Signal and noise are defined by the same attribute. In high external noise conditions, observers are just as efficient at detecting a luminance signal embedded in luminance noise as detecting a contrast signal embedded in contrast noise. Adapted from Allard and Faubert (2007).

An attribute-invariant mechanism

The attribute-invariance of contrast detection in high noise (at least for luminance, color and contrast) suggests the implication of common high-level mechanisms extracting the signal from noise. Although important threshold differences in low noise were observed for different attributes, the visibility of the attribute is not a factor in high noise. Indeed, changing the attribute of both the signal and noise to another, changes the visibility to both signal and noise without changing the signal-to-noise ratio. As demonstrated by the slope near 1 in log-log units for contrast detection in high noise (TvC function), increasing (or decreasing) both the signal and noise without altering the signal-to-noise ratio does not alter the visibility of the signal. Consequently, changing the visibility of both the signal and noise by switching attribute does not affect the signal-to-noise ratio in high noise conditions even though the two attributes do not have the same contrast gain. If the same mechanism is extracting the signal from noise for all attributes, then the same detection thresholds will be observed in high noise conditions. Note that similar CEs do not necessarily imply common mechanisms extracting the signal from noise. There could be three separate mechanisms having similar efficiencies extracting the signal from noise for luminance, color and contrast. However, it would be surprising to have three similar mechanisms performing similar tasks with similar efficiencies and yet be distinct.

Discussion

Based on the conclusion that contrast detections in low and high noise are not processed by the same mechanisms, the arguments in the present sections argues that the detection in high noise involves higher perceptual mechanisms than the detection in low noise. Indeed, contrast thresholds in high noise are independent of many low-level parameters, several parameters known to affect higher perceptual task (top-down attention, learning, aging and dyslexia) were found to affect, in certain conditions, contrast thresholds exclusively in high noise and contrast thresholds in high noise were found to be attribute-invariant. These arguments suggest that thresholds in high noise are limited by high level factors.

Furthermore, these arguments can be added to the ones arguing against the hypothesis (as suggested by the PTM) that a mechanism affecting exclusively external noise would affect contrast thresholds in high noise. In the PTM, the mechanism excluding external noise occurs before the internal noise is added. As discussed above, this processing order is necessary to explain a high-noise pattern. Indeed, once the internal and external noises are combined, they become indistinguishable. Consequently, a processing level limiting the impact of the noise after the main internal noise source would affect thresholds in both low and high external noise. Therefore, given that the PTM has a processing level affecting only thresholds in high external noise conditions, this processing level must occur before the internal noise is added. Indeed, the parameters known to affect low-level tasks affect thresholds in low noise and the ones known to affect high-level tasks affect thresholds in high noise. Reversing the order of these mechanisms as suggested by the PTM therefore seems improbable.

For instance, consider the fact that contrast threshold in high noise is attribute-invariant (i.e. same contrast threshold in high noise for different attribute). Again, this suggests the existence of an attribute-invariant mechanism extracting the signal from external noise. According to the PTM, this mechanism would be the external noise exclusion mechanism only influencing thresholds in high noise conditions. Note that no or little cross-type interaction has been observed suggesting that these attributes are processed by separate mechanisms. The difference of sensitivity (threshold in low noise) would be interpreted, according to the PTM, as a difference in additive internal noise occurring after the external noise exclusion mechanism. Consequently, the processing of different attributes would share the same initial attribute-invariant mechanism (external noise exclusion) but would be processed by separate mechanisms at a higher level of processing.

This would be improbable since attribute-invariant mechanisms usually correspond to high level mechanisms.

General discussion

Visual noise is often used in psychophysics for various purposes. It is generally assumed that adding noise to a stimulus does not qualitatively alter its processing. Conversely, the present review concludes that, at least in some conditions, adding external noise qualitatively alter the nature of the task. Consequently, interpretations using external noise could be compromised. Indeed, assuming that a given stimulus is not processed qualitatively differently whether it is presented in noiseless or noisy conditions may have misled some interpretations.

Measuring internal noise

As suggested by the signal detection theory, it is generally assumed that internal noise limits our ability to perform certain tasks. When contrast is the dependant variable of a certain task, the LAM proposes to measure the impact of the internal noise by adding external noise. In high external noise, the internal noise has no significant impact and the signal and noise contrast are both known. By assuming that the same signal-to-noise ratio is required in noiseless conditions (i.e. assuming no qualitative difference between low and high external noise conditions), it is possible to deduce the impact of the internal noise. This assumption was formulated by Pelli (1990) as follows: "the calculation performed is independent of the contrast of the effective stimulus" (p. 6-7). The effective stimulus may be defined as the sum of the signal, external noise and internal noise. In other words, by assuming that the same processing is performed whether the signal is embedded in internal or external noise, the relative impact of the internal noise can be deduced.

Based on double dissociations for processing in low (i.e. internal) and high (i.e. external) noise, we conclude that, at least in some conditions, such an assumption is not valid. Since it is not necessarily the same signal-to-noise ratio required to extract the signal from internal and external noise, it is not possible to deduce the impact of the internal noise based on the signal-to-noise ratio in external noise. This conclusion thereby compromises, at least in some conditions, the application of the models evaluating internal noise such as the LAM or PTM. All things considered, to measure the impact of internal noise using external noise we must assume that the task remains the same when adding external noise. However, the present review argues that, at least for wide variety of conditions, this is not case.

Reverse correlation

The reverse correlation method (Abbey & Eckstein, 2002, Ahumada, 2002, Murray, Bennett & Sekuler, 2002) uses noise to reveal internal representation. This method consists in performing a binary task in high external noise. Suppose that the possible answers are 'a' and 'b'. An observer performs the task for a large number of trials. The difficulty of the task is adjusted (typically by manipulating the stimulus contrast) so that the performance is not perfect but still above chance. As a result, the external noise added to the stimulus affects the observer's performance. In other words, if a certain noise template makes the stimulus 'b' more similar to the stimulus 'a', then such noise templates could lead the observer to an incorrect answer. The reverse correlation consists in analyzing the noise template used as a function of the answer given by the observer. Such analysis is performed by computing the difference between the average of the noise templates used in the trials in which the observer answered 'a' and the average of the noise templates used in the trials in which he answered 'b'. If the observer was biased to answer 'a' when a given region is white, then, on average, this region of the noise templates will be lighter when the observer answered 'a' then 'b'. As a result, the region of the noise templates that bias the observer should appear by applying the reverse correlation. This method therefore reveals the spatial regions on which the observer based his decisions. The reverse correlation reveals some information used by the observer when performing the task in noise. However, there is no guaranty that the observer uses the same regions to perform the task in noiseless conditions. Experimenters using this technique generally assume that the observer uses the same spatial region to perform the task in absence of external noise and thereby generalize their results to the task performed in noiseless conditions. In other words, they assume that adding noise does not qualitatively alter the processing of the stimulus.

For instance, this method has been used to characterize the differences between perceiving upright and inverted faces. We are better at recognizing a face when it is presented upright rather than when it is presented upside-down. This effect has generally been interpreted as evidence that face recognition is holistic and that inverting a face qualitatively alters its processing by making it less holistic (Farah, Tanaka & Drain, 1995a, Farah, Wilson, Drain & Tanaka, 1995b, Murray, Yong & Rhodes, 2000, Tanaka & Farah, 1993, Valentine, 1988). To study the spatial regions used to recognize a face upright or inverted, Sekuler, Gaspar, Gold and Bennett (2004) applied the reverse correlation method for a face recognition task. They found that the same spatial regions were used whether the face was presented upright or inverted. They conclude, as opposed to the general consensus, that inverting faces alters the task quantitatively but not qualitatively. This may be true for their testing conditions, i.e. in high external noise, but may not generalize to noiseless conditions. To generalize this interpretation to noiseless conditions, we must assume that adding noise does not qualitatively alter the nature of the task and, as suggested by the present review, there is no guaranty that this is the case.

Using an EEG protocol, Schneider, DeLong and Busey (2007) recently found evidence that upright and inverted faces are qualitatively processed differently. In noiseless conditions, inverting a face increased the N170 amplitude response attributed to face processing as it has been previously observed. However, when external noise was added to the stimulus, they found that inverting a face decreased the N170 amplitude response.

These two qualitatively different patterns of results support the main conclusion of the present review suggesting that adding noise to a stimulus (it this case, a face) qualitatively affects its processing. Consequently, upright and inverted faces could be qualitatively processed similarly in high external noise (as suggested by Sekuler et al. (2004)), but could be processed qualitatively differently in noiseless conditions as it is generally concluded (Farah et al., 1995a, Farah et al., 1995b, Murray et al., 2000, Tanaka & Farah, 1993, Valentine, 1988).

Contrast detection in noise: A specific perceptual task

As mentioned by Pelli and Farell (1999), adding noise enables to abstract many low-level parameters which they defined as influencing the IEN. We agree that in high noise conditions a given task becomes independent to many low-level factors as discussed above. However, the present review suggests that it does not only eliminate the impact of certain low-level factors it also qualitatively alters the processing, i.e. it also adds other factors limiting contrast thresholds only in high noise. In other words, although the nature of the task is altered by external noise, this new task is also independent of many low-level factors. As a result, contrast thresholds in noise are independent of the background luminance intensity and ocular distortions such as cataracts. More generally, any parameter equivalent to adding noise or to apply a contrast gain to the stimulus has no impact on contrast thresholds in high noise.

For a simple task, such as contrast detection, we have argued above that adding noise qualitatively alters the nature of the task by making it more complex. However, such perceptual task remains cognitively simple. The task simply consists in indicating which interval contained the signal. Consequently, contrast detection in noise is a specific perceptual task independent of most low-level peripheral parameters such as ocular factors, and is also independent of high cognitive factors since the task is cognitively simple. This suggests that contrast detection in noise is a specific perceptual task, i.e. it dependents almost exclusively on perceptual factors. Moreover, this task is perceptually complex enough so that it is significantly affected by various manipulations such as aging, learning, attention, dyslexia and amblyopia. Consequently, contrast detection in noise is a useful task to exclusively assess perceptual functions by evaluating the ability of the visual system to process complex information.

Conclusion

Adding noise to a stimulus has often been used to study visual functions. For visual tasks, adding noise to an image deteriorates its content and generally affects the observer's performance. There are several reasons to add noise. For instance, adding noise enables to null the impact of internal noise (or low-level factors). It has also been used to enable ideal observer analysis which consists in comparing an observer performance with the one of an ideal observer. Furthermore, noise has been used to reveal internal representation by correlating the noise templates with the observer's answers (reverse correlation). All things considered, noise has been widely used to characterize low and high level visual functions.

The present review highlights data present in the literature providing direct evidence of double dissociations between the presence and absence of external noise. This suggests that adding noise to a visual stimulus affects, at least in some conditions, the nature of the task, i.e. different mechanisms are solicited. Therefore, experimenters should be careful when adding noise to there stimuli and they should not presume a priori that adding noise does not qualitatively affect its processing.

Acknowledgements: This research was supported by NSERC graduate fellowship to RA and NSERC operating grant to JF.

References

Abbey, C.K., & Eckstein, M.P. (2002). Classification image analysis: estimation and statistical inference for two-alternative forced-choice experiments. *Journal of Vision, 2* (1), 66-78.

Ahumada, A.J., Jr. (2002). Classification image weights and internal noise level estimation. *Journal of Vision, 2* (1), 121-131.

Allard, R., & Faubert, J. (2006). Same calculation efficiency but different internal noise for luminance- and contrast-modulated stimuli detection. *Journal of Vision, 6* (4), 322-334.

Allard, R., & Faubert, J. (2007). Double dissociation between first- and second-order processing. *Vision Research, 47* (9), 1129-1141.

Bennett, P.J., Sekuler, A.B., & Ozin, L. (1999). Effects of aging on calculation efficiency and equivalent noise. *Journal of the Optical Society of America, A, Optics, Image Science, & Vision, 16* (3), 654-668.

Betts, L.R., Sekuler, A.B., & Bennett, P.J. (2007). The effects of aging on orientation discrimination. *Vision Research, 47* (13), 1769-1780.

Campbell, F.W., & Robson, J.G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology, 197* (3), 551-566.

Dosher, B.A., Liu, S.H., Blair, N., & Lu, Z.L. (2004). The spatial window of the perceptual template and endogenous attention. *Vision Research, 44* (12), 1257-1271.

Dosher, B.A., & Lu, Z.-L. (2006). Level and mechanisms of perceptual learning: Learning first-order luminance and second-order texture objects. *Vision Research, 46* (12), 1996-2007.

Dosher, B.A., & Lu, Z.L. (2000a). Mechanisms of perceptual attention in precuing of location. *Vision Research, 40* (10-12), 1269-1292.

Dosher, B.A., & Lu, Z.L. (2000b). Noise exclusion in spatial attention. *Psychological Science, 11* (2), 139-146.

Dunn, J.C., & Kirsner, K. (2003). What can we infer from double dissociations?[see comment]. *Cortex, 39* (1), 1-7.

Farah, M.J., Tanaka, J.W., & Drain, H.M. (1995a). What causes the face inversion effect? *Journal of Experimental Psychology: Human Perception & Performance, 21* (3), 628-634.

Farah, M.J., Wilson, K.D., Drain, H.M., & Tanaka, J.R. (1995b). The inverted face inversion effect in prosopagnosia: evidence for mandatory, face-specific perceptual mechanisms. *Vision Research, 35* (14), 2089-2093.

Faubert, J. (2002). Visual perception and aging. *Canadian Journal of Experimental Psychology, 56* (3), 164-176.

Gegenfurtner, K.R., & Kiper, D.C. (1992). Contrast detection in luminance and chromatic noise. *Journal of the Optical Society of America A-Optics & Image Science, 9* (11), 1880-1888.

Gold, J., Bennett, P.J., & Sekuler, A.B. (1999). Signal but not noise changes with perceptual learning. *Nature, 402* (6758), 176-178.

Gold, J.M., Sekuler, A.B., & Bennett, P.J. (2004). Characterizing perceptual learning with external noise. *Cognitive Science, 28*, 167-207.

Green, D.M., & Swets, J.A. (1966). Signal detection theory and psychophysics. (New York: John Wiley and sons, inc.

Ho, C.S., Paul, P.S., Asirvatham, A., Cavanagh, P., Cline, R., & Giaschi, D.E. (2006). Abnormal spatial selection and tracking in children with amblyopia. *Vision Research, 46* (19), 3274-3283.

Hubel, D.H., & Wiesel, T.N. (1959). Receptive fields of single neurones in the cat's striate cortex. *Journal of Physiology, 148*, 574-591.

Hubel, D.H., & Wiesel, T.N. (1961). Integrative action in the cat's lateral geniculate body. *Journal of Physiology, 155*, 385-398.

Legge, G.E., & Foley, J.M. (1980). Contrast masking in human vision. *Journal of the Optical Society of America, 70* (12), 1458-1471.

Legge, G.E., Kersten, D., & Burgess, A.E. (1987). Contrast discrimination in noise. *J Opt Soc Am A, 4* (2), 391-404.

Lerner, Y., Pianka, P., Azmon, B., Leiba, H., Stolovitch, C., Loewenstein, A., Harel, M., Hendler, T., & Malach, R. (2003). Area-specific amblyopic effects in human occipitotemporal object representations. *Neuron, 40* (5), 1023-1029.

Lu, Z.L., & Dosher, B.A. (1998). External noise distinguishes attention mechanisms. *Vision Research, 38* (9), 1183-1198.

Lu, Z.L., & Dosher, B.A. (2000). Spatial attention: different mechanisms for central and peripheral temporal precues? *Journal of Experimental Psychology: Human Perception & Performance, 26* (5), 1534-1548.

Lu, Z.L., & Dosher, B.A. (2004a). External noise distinguishes mechanisms of attention. In: L. Itti, & G. Rees (Eds.), *Neurobiology of Attention* (pp. 448-453). San Diego, CA: Academic press.

Lu, Z.L., & Dosher, B.A. (2004b). Perceptual learning retunes the perceptual template in foveal orientation identification. *Journal of Vision, 4* (1), 44-56.

Lu, Z.L., & Dosher, B.A. (2004c). Spatial attention excludes external noise without changing the spatial frequency tuning of the perceptual template. *Journal of Vision, 4* (10), 955-966.

Lu, Z.L., Jeon, S.T., & Dosher, B.A. (2004). Temporal tuning characteristics of the perceptual template and endogenous cuing of spatial attention. *Vision Research, 44* (12), 1333-1350.

Lu, Z.L., Lesmes, L.A., & Dosher, B.A. (2002). Spatial attention excludes external noise at the target location. *Journal of Vision, 2* (4), 312-323.

Murray, J.E., Yong, E., & Rhodes, G. (2000). Revisiting the perception of upside-down faces. *Psychological Science, 11* (6), 492-496.

Murray, R.F., Bennett, P.J., & Sekuler, A.B. (2002). Optimal methods for calculating classification images: weighted sums. *Journal of Vision, 2* (1), 79-104.

Oruc, I., Landy, M.S., & Pelli, D.G. (2006). Noise masking reveals channels for second-order letters. *Vision Research, 46* (8-9), 1493-1506.

Owsley, C., Sekuler, R., & Siemsen, D. (1983). Contrast sensitivity throughout adulthood. *Vision Research, 23* (7), 689-699.

Pardhan, S. (2004). Contrast sensitivity loss with aging: sampling efficiency and equivalent noise at different spatial frequencies. *J Opt Soc Am A Opt Image Sci Vis, 21* (2), 169-175.

Pardhan, S., Gilchrist, J., & Beh, G.K. (1993). Contrast detection in noise: a new method for assessing the visual function in cataract. *Optometry & Vision Science, 70* (11), 914-922.

Pardhan, S., Gilchrist, J., Elliott, D.B., & Beh, G.K. (1996). A comparison of sampling efficiency and internal noise level in young and old subjects. *Vision Research, 36* (11), 1641-1648.

Parish, D.H., & Sperling, G. (1991). Object spatial frequencies, retinal spatial frequencies, noise, and the efficiency of letter discrimination. *Vision Research, 31* (7-8), 1399-1415.

Pelli, D.G. (1981). The effects of visual noise. *Department of Physiology,* Ph.D. (Cambridge: Cambridge University.

Pelli, D.G. (1990). The quantum efficiency of vision. In: C. Blakemore (Ed.) *Visual coding and efficiency* (Cambridge: Cambridge University Press.

Pelli, D.G., & Farell, B. (1999). Why use noise? *Journal of the Optical Society of America, A, Optics, Image Science, & Vision, 16* (3), 647-653.

Pelli, D.G., Farell, B., & Moore, D.C. (2003). The remarkable inefficiency of word recognition.[see comment]. *Nature, 423* (6941), 752-756.

Pelli, D.G., Levi, D.M., & Chung, S.T.L. (2004). Using visual noise to characterize amblyopic letter identification. *Journal of Vision, 4* (10), 904-920.

Raghavan, M. (1995). Sources of visual noise. Doctoral dissertation (Syracuse, NY: Syracuse University.

Schneider, B.L., DeLong, J.E., & Busey, T.A. (2007). Added noise affects the neural correlates of upright and inverted faces differently. *Journal of Vision, 7* (4), 1-24.

Sekuler, A.B., Gaspar, C.M., Gold, J.M., & Bennett, P.J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology, 14* (5), 391-396.

Sharma, V., Levi, D.M., & Klein, S.A. (2000). Undercounting features and missing features: evidence for a high-level deficit in strabismic amblyopia. *Nature Neuroscience, 3* (5), 496-501.

Simmers, A.J., Ledgeway, T., & Hess, R.F. (2005). The influences of visibility and anomalous integration processes on the perception of global spatial form versus motion in human amblyopia. *Vision Research, 45* (4), 449-460.

Simmers, A.J., Ledgeway, T., Hess, R.F., & McGraw, P.V. (2003). Deficits to global motion processing in human amblyopia. *Vision Research, 43* (6), 729-738.

Speranza, F., Moraglia, G., & Schneider, B.A. (2001). Binocular detection of masked patterns in young and old observers. *Psychology & Aging, 16* (2), 281-292.

Sperling, A.J., Lu, Z.-L., Manis, F.R., & Seidenberg, M.S. (2005). Deficits in perceptual noise exclusion in developmental dyslexia. *Nature Neuroscience, 8* (7), 862-863.

Sperling, A.J., Lu, Z.-L., Manis, F.R., & Seidenberg, M.S. (2006). Motion-perception deficits and reading impairment: it's the noise, not the motion. *Psychological Science, 17* (12), 1047-1053.

Tanaka, J.W., & Farah, M.J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology A, 46* (2), 225-245.

Valentine, T. (1988). Upside-down faces: a review of the effect of inversion upon face recognition. *British Journal of Psychology, 79* (Pt 4), 471-491.

Section 2

Traitement du mouvement défini par la luminance et le contraste

Chapitre V

First- and second-order motion mechanisms are distinct at low but common at high temporal frequencies

Rémy Allard et Jocelyn Faubert

Laboratoire de psychophysique et perception visuelle

École d'optométrie, Université de Montréal

Le contenu de ce chapitre a précédemment été publié sous forme d'article (©ARVO) :

Allard, R., & Faubert, J. (2008). First- and second-order motion mechanisms are distinct at low but common at high temporal frequencies. *Journal of Vision, 8*(2), 1-17.

Abstract

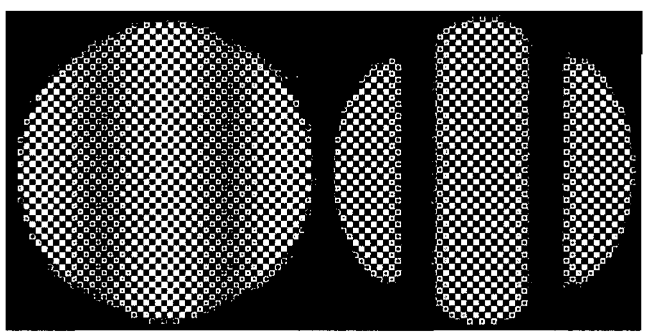
There is no consensus on the type of nonlinearity enabling motion processing of second-order stimuli. Some authors suggest that a nonlinearity specifically applied to second-order stimuli prior to motion processing (e.g. rectification process) recovers the spatial structure of the signal permitting subsequent first-order motion analyses (e.g. filter-rectify-filter model). Others suggest that nonlinearities within motion processing enable first-order sensitive mechanisms to process second-order stimuli (e.g. gradient-based model). In the present study, we evaluated intra- and inter-attribute interactions by measuring the impact of dynamic noise modulators (either luminance (LM) or contrast-modulated (CM)) on the processing of moving LM and CM gratings. When the signal and noise were both of the same type, similar calculation efficiencies but different internal equivalent noises were observed at all temporal frequencies. At high temporal frequencies, each noise type affected both attributes by similar proportions suggesting that both attributes are processed by common mechanisms. Conversely, at low temporal frequencies, each noise type primarily impaired the processing of the attribute of the same type suggesting distinct mechanisms. We therefore conclude that two fundamentally different mechanisms are processing CM stimuli: one lowpass and distinct from the mechanisms processing LM stimuli and the other common to the mechanisms processing LM stimuli.

Keywords: Second-order, contrast-modulated, motion, noise, filter-rectify-filter model, gradient-based model, feature tracking

Introduction

There is no consensus on how second-order stimuli are processed. Typically, first-order stimuli are defined by luminance or color, and second-order stimuli are defined by some other attribute such as contrast, orientation or texture (Baker, 1999; Cavanagh & Mather, 1989; Chubb & Sperling, 1988; Wilson, Ferrera, & Yo, 1992). Second-order stimuli are composed of a carrier and an envelope. The envelope locally defines a certain property of the carrier (e.g. contrast).

Some authors suggest the existence of specialized mechanisms dedicated to second-order stimuli (e.g. filter-rectify-filter model; (Wilson, Ferrera, & Yo, 1992)), while others rather suggest that, at least for some second-order stimuli, nonlinearities within first-order sensitive mechanisms could enable second-order perception (e.g. gradient-based model; (Benton, 2002; Benton & Johnston, 2001; Benton, Johnston, McOwan, & Victor, 2001; Taub, Victor, & Conte, 1997)). Luminance- (LM) and contrast-modulated (CM) patterns (Movie V-1) are the more frequently used profiles to represent first- and second-order stimuli respectively. The present paper investigated whether such motion stimuli are processed by common or separate mechanisms.



Movie V-1. Luminance- (left) and contrast-modulated (right) signals. The temporal frequency is 2 Hz. (Movie available on the CD attached to the thesis.)

Filter-rectify-filter model

The filter-rectify-filter model suggests that LM and CM stimuli are initially processed by separated pathways (see Baker (1999) for a review). Extra processing for CM stimuli (rectification process) would reveal the spatial structure of the envelope, which could then be processed by subsequent mechanisms. Indeed, this model suggests that before perceiving the signal (i.e. envelope), the local property of the carrier should first be evaluated (in this case the contrast) followed by its variation over space (i.e. the signal or envelope). In other words, to perceive a difference of contrast, we first need to evaluate the local contrast of different spatial regions.

Gradient-based model

The fact that we can perceive second-order stimuli containing no spectral energy near the envelope frequency led several authors to suggest the existence of a dedicated mechanism for second-order processing such as the filter-rectify-filter model presented above. However, some authors have proposed models in which both LM and CM would be processed by common mechanisms (Benton, 2002, 2004; Benton & Johnston, 2001; Benton, Johnston, McOwan, & Victor, 2001; Johnston, Benton, & McOwan, 1999; Johnston & Clifford, 1995a; Johnston, McOwan, & Buxton, 1992; Taub, Victor, & Conte, 1997). For instance, a gradient-based algorithm (Benton & Johnston, 2001) computing the temporal derivative relative to the spatial derivative could reveal the motion direction of both LM and CM stimuli. Consequently, although CM stimuli do not have spectral energy near the envelope frequency, nonlinearities (e.g. ratio of temporal vs spatial derivatives) within motion processing could enable LM sensitive mechanisms to also detect CM stimuli. Such an algorithm would not initially recover the spatial property of the stimulus; it would directly process the direction of motion based on the spatial and temporal local variation of luminance. One of the limits of the gradient-based model is that it would not be sensitive to all second-order stimuli (Lu & Sperling, 2001). Nonetheless, the gradient-based approach suggests that, in some conditions, the visual system may not require dedicated motion mechanisms to process second-order stimuli.

Spatial LM and CM processing

In a recent study on spatial vision (Allard & Faubert, 2007), we evaluated inter-attribute interactions between static LM and CM stimuli. We found that LM noise affected LM signal detection but had little or no impact on CM signal detection, and, vice versa, CM noise affected CM signal detection but had little or no impact on LM signal detection. This double dissociation implies that both cues must be processed, at least at some point, by separate mechanisms. These results are in agreement with Schofield and Georgeson's results showing no subthreshold summation between LM and CM stimuli (Schofield & Georgeson, 1999) and similar detection (LM vs noise and CM vs noise) and recognition (LM vs CM) thresholds (Georgeson & Schofield, 2002) suggesting that LM and CM stimuli are processed by two separated pathways.

In another study (Allard & Faubert, 2006), we decomposed the sensitivity to static LM and CM stimuli into internal equivalent noise and calculation efficiency (Legge, Kersten, & Burgess, 1987; Pelli, 1981, 1990). The internal equivalent noise corresponds to the amount of noise added to the stimulus having the same impact as the internal noise. The calculation efficiency is inversely proportional to the smallest signal-to-noise ratio at which the signal may be detected. We found that the difference of sensitivity to LM and CM stimuli was due to a difference of internal equivalent noise and not to a difference of calculation efficiency. In other words, in high noise conditions, observers had similar detection thresholds to both LM and CM stimuli. Indeed, observers were just as efficient at detecting LM signals embedded in LM noise as CM signals embedded in CM noise. This suggests that common mechanisms could be extracting the signal from noise for both LM and CM stimuli. Schofield and Georgeson also found similar responses to static LM and CM stimuli. They observed similar spatial (Schofield & Georgeson, 1999) and temporal

(Schofield & Georgeson, 2000) integration and similar sensitivity function shapes (Schofield & Georgeson, 1999). They also found inter-attribute interactions: adapting to one cue affected the perceived modulation depth of the other (Georgeson & Schofield, 2002). However, since inter-attribute adaptation effects in high contrast conditions are not very pattern selective (Ross & Speed, 1996; Snowden & Hammett, 1992, 1996), they concluded that common adaptation is not strong evidence for common processing.

Schofield and Georgeson (1999) therefore concluded that static LM and CM stimuli are processed by distinct mechanisms with similar properties that share common adaptive mechanisms. Alternatively, we proposed that the detection of static LM and CM stimuli are initially processed by separate pathways, but are processed by common mechanisms at higher levels (Allard & Faubert, 2007). Based on the fact that no inter-attribute interaction was observed near threshold we suggested that common late mechanisms could focus on either attribute without merging them. If late mechanisms were processing both attributes simply by merging them, the noise presented to one pathway would affect the detection of the signal presented to the other. We therefore suggested a gating model in which late mechanisms could select either attribute while ignoring the other.

Purpose of the present study

The main objective of the present study was to apply a similar noise-masking paradigm as the one we have used to study static LM and CM processing in order to investigate whether LM and CM motion stimuli are processed by common or separate mechanisms. We therefore evaluated contrast thresholds of moving LM and CM signals embedded in LM and CM dynamic noise. When the signal and noise were both defined by the same attribute, it was possible to decompose the sensitivity into internal equivalent noise and calculation efficiency. Similar calculation efficiencies, i.e. similar signal-to-noise ratio required to detect the signal, would suggest that, at least at some point, both types of stimuli are processed by common mechanisms. Inter-attribute interactions were evaluated by superimposing a signal and noise defined by different attributes. No or little inter-attribute interaction would imply that both stimuli must be processed, at least at some point, by separate mechanisms. Alternatively, complete inter-attribute interactions (LM and CM noise each affecting LM and CM processing by the same proportions) would suggest that the two cues are processed by common mechanisms.

Experiment 1: Compressive nonlinearity

We are more sensitive to first-order than second-order cues. As a result, small artifacts (introduced either by the display or by the visual system) can enable first-order sensitive mechanisms to process second-order stimuli. Experimenters must therefore assert that first-order artifacts are too small to enable, by themselves, first-order sensitive mechanisms to process second-order stimuli.

Component motion (Scott-Samuel & Georgeson, 1999) can occur when the carrier's spectral energy is not broadband (e.g. periodic or highpass carriers). For such carriers, the spectral energy is concentrated at some frequencies. Adding a contrast modulation to the carrier gives rise to two spectral energy peaks near each energy peak of the un-modulated carrier which are referred to as sidebands. The sideband with the lowest spatial frequency has motion energy in the opposite direction to the CM signal and the other has motion energy in the same direction as the CM signal. If the observer is more sensitive to the sideband with the lowest spatial frequency then he will perceive motion in the opposite direction of the signal. In the present study, the carrier used was defined only by high spatial frequencies so component motion could have been an issue. However, the fact that a direction discrimination task was used ensures that this artifact was not a concern. Indeed, if observers were processing motion due to this artifact, then their response would have been incorrect when motion was perceived and the staircases would not have converged. Since this was not the case in all the conditions, we concluded that the results were unaffected by component motion artifacts.

When using broadband static noise as a carrier, local first-order artifacts could enable CM processing (Smith & Ledgeway, 1997). If the mean luminance of a spatial region of the carrier is not equal to the mean luminance of the entire stimulus, then adding a CM signal causes a local first-order artifact by introducing a LM signal within this region. Local first-order artifacts will be of opposite polarity for spatial regions of the carrier with lower and higher local luminance relative to the mean luminance. As a result, there will be no spectral energy in the Fourier domain since the opposite polarities will on average cancel one another. Nonetheless, there will be local direction biases at various spatial regions that could be used to discriminate the motion direction. Since the carriers used in the present study contained spectral energy only at high spatial frequencies, all local mean luminance were equal to the mean luminance of the display and this type of artifact was therefore not an issue.

The global distortion product artifact is caused by compressive nonlinearities of the visual system (Scott-Samuel & Georgeson, 1999; Smith & Ledgeway, 1997). Indeed, it has been shown that there are early nonlinearities within the visual system prior to LM sensitive mechanisms (He & Macleod, 1998; Legge & Foley, 1980; MacLeod, Williams, & Makous, 1992). These nonlinearities were found to be compressive and generally too weak to explain CM sensitivity (Scott-Samuel & Georgeson, 1999; Smith & Ledgeway, 1997). A compressive nonlinearity would reduce the mean luminance of high-contrast regions. Consequently, although the mean luminance of two regions varying in contrast are the same, early nonlinearities could introduce luminance variations making CM stimuli visible to LM sensitive mechanisms. Therefore, the experimenter could erroneously conclude that a CM stimulus is processed by CM sensitive mechanisms although it is actually processed by LM sensitive mechanisms following an early nonlinearity. The present experiment had two goals. First, measure the early nonlinearity for each tested condition, i.e. for each subject and each temporal frequency. Second, ensure that the processing of CM stimuli is not due to an early compressive nonlinearity and that CM stimuli were processed by CM sensitive mechanisms in all tested conditions.

An early compressive nonlinearity may be cancelled by introducing an expansive nonlinearity of the same magnitude into the stimulus (Scott-Samuel & Georgeson, 1999). The resulting nonlinearity may be defined as the sum of the early nonlinearity introduced by the visual system and the nonlinearity introduce within the stimulus. Our objective was to find the nonlinearity that needs to be introduced within the stimulus to cancel the early nonlinearity caused by the visual system. We supposed that both nonlinearities cancel one another if the same performance is observed whether a LM and a CM signal are combined either in phase (high contrast regions of the CM signal matching with high luminance regions of the LM signal) or in counter-phase (high contrast regions of the CM signal matching with low luminance regions of the LM signal). Indeed, a resulting compressive nonlinearity would lower the contrast of the LM signal in the in-phase condition and would increase the contrast of the LM signal in the counter-phase condition. This would be equivalent to introducing a LM signal in counter-phase with the CM signal. Consequently, the resulting compressive nonlinearity would enhance the LM signal in the counter-phase condition and would reduce it in the in-phase condition. As a result, the performance would be greater in the counter-phase condition. Alternatively, an expansive nonlinearity would cause the opposite pattern resulting in a better performance in the in-phase condition. However, if both nonlinearities (nonlinearity of the stimulus and of the visual system) cancel one another, the same performance level should be observed whether the LM and CM signals are combined in-phase or in counter-phase.

Method

Apparatus

The stimuli were presented on a 19 in ViewSonic E90FB .25 CRT monitor with a mean luminance of 47 cd/m^2 and a refresh rate of 120 Hz powered by a Pentium 4 computer. The 10-bit Matrox Parhelia512 graphic card could produce 1024 gray levels that could all be presented simultaneously. The monitor was the only light source in the room.

A Minolta CS100 photometer interfaced with a homemade program calibrated the output intensity of each gun. At the viewing distance of 1.14 m, the width and height of each pixel were 1/64 deg of visual angle.

DAC precision

Although the setup used could display 1024 levels of grey, in certain conditions the contrast thresholds approached the smallest grey difference (1/1024). Since the desired luminance value for each pixel generally correspond to a continuous value, this value had to be rounded with a precision of 1/1024, i.e. to the nearest DAC value. This procedure can sometimes create sufficiently high artifacts to alter contrast threshold measurement.

Instead of simply rounding to the nearest DAC value, we used a different algorithm consisting in randomly choosing between the two nearest DAC values. The probability distribution between the two values was set so that the expected value was the desired continuous DAC value. That is, the probability of choosing the higher DAC value was equal to the remainder of the continuous desired DAC value. For example, if the desired continuous DAC value was 123.25, then the probability distribution was 0.25 for 124 and 0.75 for 123. This random selection was independently applied to each pixel of each frame.

The advantage of using such a method is that the expected luminance of each pixel of each frame is equal to the desired continuous luminance. Consequently, for a luminance grating, the expected luminance value would vary continuously and it could therefore be possible to present a grating with a difference of luminance smaller than one DAC value (or 1/1024 of the maximal luminance). The spatiotemporal summation of a given region should result into a mean luminance value near the expected luminance. The disadvantage of this method is that it adds noise (random variations) to the presented stimulus. Thus, the noise added to the display may become a limiting factor.

This method of randomly selecting between the two nearest DAC values is mathematically equivalent to rounding to the nearest DAC value after adding dynamic noise in which each element is randomly selected from a uniform distribution varying between -0.5 and 0.5 DAC values. Since the noise sampling varied at 120 Hz and the noise contrast was small (less than 1/1024), the spectral energy of the noise was also small. To ensure that this noise had no significant impact, we measured the noise contrast required to significantly decrease the sensitivity to a LM stimulus. We found that the noise contrast required to affect sensitivity had to be at least 10 DAC values, which is much larger than the noise introduced when randomly selecting between the two nearest DAC values. We therefore concluded that the random variation introduced when randomly selecting the DAC value had no significant impact and enabled the apparatus to display a 1024 grey scale resolution equivalent to a continuous grey scale resolution.

Observers

Two psychophysically experienced observers participated in the study: one of the authors and the other naïve to the purpose of the experiment. They had normal or corrected-to-normal vision.

Stimuli

For CM stimuli, binary noise is typically used as a carrier. However, one of the targets of the present study was to evaluate internal equivalent noise. Intrinsic noise within the stimulus could be a limiting factor affecting the measurement of IEN (Allard & Faubert, 2006). Consequently, we did not want the stimulus to have intrinsic noise so instead of using binary noise we used a static checkerboard as a carrier. Its contrast was set to 50% and each check was composed of 6x6 pixels (5.6x5.6 minutes). To avoid LM cues within a check, the luminance within each check was kept spatially constant (Smith & Ledgeway, 1997). Keeping the luminance within each check spatially constant is equivalent to reducing the stimulus spatial resolution. Since the checks were small relative to the signal

(21 checks per period), such lowering of the spatial resolution was judged to have no significant impact.

All the stimuli used in the present study can be defined as the sum of two terms: a luminance modulation ($M_{LM}(x,y,t)$) and the multiplication of a contrast modulation ($M_{CM}(x,y,t)$) with a static carrier ($T(x,y)$):

$$L(x,y,t) = L_0\left[M_{LM}(x,y,t) + M_{CM}(x,y,t)T(x,y)\right], \quad (1)$$

where L_0 represents the luminance average of the stimulus and the background luminance. In the present experiment, $T(x,y)$ corresponded to a static checkerboard. Its values were -0.5 if $x+y$ was odd and 0.5 otherwise resulting in a Michelson contrast of 0.5. The luminance and contrast modulations were sine wave gratings:

$$M_{LM}(x,y,t) = 1 + \left(C_{LM} + nC_{CM}\right)\sin(sx + ft + p), \quad (2)$$

$$M_{CM}(x,y,t) = 1 + C_{CM}\sin(sx + ft + p), \quad (3)$$

where s, f and p represent, respectively, the spatial frequency (0.5 cpd), the temporal frequency (varying between ±1 and ±16 Hz depending on the testing condition) and the initial phase (randomized at each trial). C_{LM} and C_{CM} represent the contrast of the LM and CM signals which varied according to the condition. n corresponds to the nonlinearity added to the stimulus in order to compensate for early nonlinearities within the visual system. A positive nonlinearity (n>0) corresponds to an expansive nonlinearity (higher luminance for higher contrast regions), while a negative nonlinearity results into a compressive nonlinearity.

A circular spatial window with a diameter of 4 deg and soft edges following a half cosine of 0.5 deg was used. Outside the carrier, the screen remained blanked to the mean luminance (L_0=47 cd/m^2). The presentation time was 500 ms. Between trials, when no stimulus was presented, a non-modulated carrier was shown with a centered fixation point.

Procedure

The task consisted of discriminating the drifting direction (either left or right) of the LM or CM signal by pressing one of two keys. To measure the compressive nonlinearity, the contrast thresholds to LM (manipulating C_{LM} and keeping C_{CM}=0) and CM (manipulating C_{CM} and keeping C_{LM}=0) stimuli were evaluated using a 2-down-1-up procedure (Levitt, 1971). The staircase was interrupted after sixteen inversions and the threshold was estimated by the geometric mean of the contrast (C_{LM} or C_{CM}) at the last 8 inversions. The initial signal contrast (the dependant variable) was set significantly above threshold. The step size before the second inversion was 0.2 log units. Afterwards and until the forth inversions it was set to 0.1 log units. Subsequent to the forth inversion the step size was 0.05 log units.

It was important to properly evaluate the LM and CM contrast thresholds since the settings of the next procedural step depended on them. Large measurement errors could compromise the next procedural step consisting in measuring the compressive nonlinearity of the visual system. To enhance threshold precision, LM and CM thresholds were evaluated three times (three staircases) for each temporal frequency. Each threshold was estimated as the geometric mean of the three staircases.

For each temporal frequency, once the contrast thresholds to LM (C_{LM}) and CM (C_{CM}) stimuli were measured (which we will denote T_{LM} and T_{CM}, respectively), the expansive nonlinearity that needed to be introduced within the stimulus (n) to compensate for the compressive nonlinearity of the visual system was evaluated. To do so, the performance level (proportion of correct answers) was evaluated when superimposing LM and CM signals at threshold either in-phase (C_{LM}=T_{LM} and C_{CM}=T_{CM}) or in counter-phase (C_{LM}=T_{LM} and C_{CM}=-T_{CM}). For each of these two phase conditions, five nonlinearities (n) were added. As mentioned above, the nonlinearity was an additional LM signal in-phase with the CM signal with contrast nC_{CM}. For comparative reasons, we also evaluated the performance level to LM ($C_{LM}=T_{LM}$ and $C_{CM}=0$) and CM ($C_{LM}=0$ and $C_{CM}=T_{CM}$) signals separately. Overall, performance was evaluated for twelve stimuli: the combination of LM and CM signals in-phase using 5 different nonlinearities, the combination of LM and CM signals in counter-phase using the same 5 different nonlinearities, a LM signal alone and a CM signal alone using one nonlinearity. The five nonlinearities for the combined signals and the nonlinearity for the CM signal alone were arbitrarily set based on a pilot study. There values can be seen in the Figure V-1 of the next section. For each performance level evaluated, 50 trials were performed resulting in 600 trials presented in a pseudo-random order.

Fitting the data

Two normalized cumulative Gaussian functions were fitted to the data for each temporal frequency of each subject. For the in-phase condition, the function increased with the nonlinearity, and for the counter-phase condition, the function decreased with the nonlinearity. Both functions were constrained to have the same slope, the same lower bound and an upper bound set to 100% correct response. The lower bound was not fixed because it consisted in the performance for CM stimuli alone.

Results and discussion

Early nonlinearity measured

The estimated stimulus nonlinearities needed to compensate for early nonlinearities of the visual system are presented in Figure V-1. As described above, the same performance observed whether LM and CM signals are combined in phase or in counter-phase (the point at which the two lines cross in Figure V-1) suggests that the early nonlinearity of the visual system was compensated by the nonlinearity added to the stimulus. As previously observed (Scott-Samuel & Georgeson, 1999; Smith & Ledgeway, 1997), at low temporal frequencies, early nonlinearities of the visual system were small

($n \approx 0$). However, at greater temporal frequencies, early nonlinearities were greater and compressive. Indeed, expansive nonlinearities ($n>0$) had to be added to compensate for the early nonlinearities of the visual system.

Modeling the nonlinearity of the visual system

Scott-Samuel and Georgeson (1999) have used the Naka-Rushton equation with the exponent variable set to 1 to model the compressive nonlinearity of the visual system. Using this function, the intensity of the photoreceptor response relative to the maximal response can be modeled by the following function:

$$E(x,y,t) = \frac{L(x,y,t)/L_0}{L(x,y,t)/L_0 + S}, \qquad (4)$$

The compressive nonlinearity of the retinal receptors depends on the parameter S. To fit the parameter S, we have used a similar approach as Scott-Samuel and Georgeson (1999). For each conditions (i.e. each temporal frequency and each subject), we created two stimuli composed of a LM and CM signal either combined in phase ($C_{LM}=T_{LM}$) or in counter-phase ($C_{LM}=-T_{LM}$) using the stimulus nonlinearity (n) at which the same performance was observed whether the signals were combined in phase or in counter-phase. The contrast of the CM signal was also equal to the threshold ($C_{CM}=T_{CM}$). The compressive nonlinearity of the visual system was modeled by applying Equation 4 to each stimulus. The S variable was manipulated until the Fourier transform of the two stimuli gave the same spectral energy at the signal frequency. If the observer has the same performance whether the LM and CM signal are combined in phase or in counter-phase, we concluded that the compressive nonlinearity was compensated by the expansive nonlinearity induced in the stimulus.

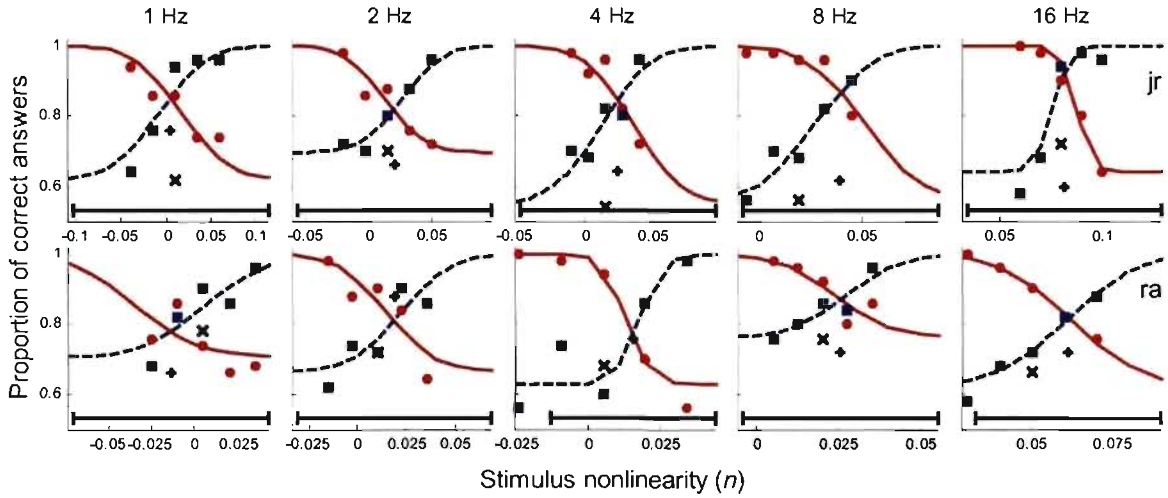


Figure V-1. Nonlinearity results. The blue squares and red dots correspond to the percentage of correct answers when LM and CM stimuli were combined in-phase and in counter-phase, respectively. The black horizontal lines represent the range of nonlinearity added to the stimulus within which CM stimuli must be processed by CM sensitive mechanisms. The + and × signs correspond to the performance level when only LM and CM signals were presented alone, respectively. For LM signals alone, the nonlinearity added is undefined so we arbitrarily set the position of the + signs to the measured nonlinearity compensating for the early compressive nonlinearity. For CM signals, the nonlinearity added was arbitrarily set and is shown by the horizontal position of the × signs.

We did not model the nonlinearity at 1 and 2 Hz since the nonlinearities measured were too low ($n \approx 0$) to have a significant impact on the results of the next experiments. At 4 Hz, the fitted S values were 19 and 20 for observer JR and RA respectively. At 8 Hz they were 7.2 and 11.5 respectively. And at 16 Hz they were 5.2 and 7.1 respectively. Lower values for the S parameter represent higher compressive nonlinearities. The increasing nonlinearity observed when increasing the temporal frequency is consistent with Scott-Samuel and Georgeson's results. However, the nonlinearities observed at 16 Hz were lower than what has been previously observed. Near this temporal frequency, nonlinearities were found to vary between 0.5 and 3.5 (Scott-Samuel & Georgeson, 1999). However, a pilot study showed that leaving the carrier visible between trials reduced the nonlinearity suggesting that the stimulus onset enhances compressive nonlinearities of the visual system. This is consistent with the fact that the nonlinearity decreases with increasing presentation time as suggested by Scott-Samuel and Georgeson (1999).

Once the S parameter was estimated for each testing condition, we compared the difference between using the Naka-Rushton equation (Equation 4) and simply adding a luminance modulation proportional to the contrast modulation using the parameter n (Equation 2). To compare these models, we created stimuli composed of either a unique CM signal at different contrasts (C_{LM}=0 and C_{CM}=0 to 0.9) or a unique LM signal equal to the discrimination threshold ($C_{LM}=T_{LM}$ and C_{CM}=0). The nonlinearity (n) applied to the stimuli was the one estimated by fitting the data as shown in Figure V-1. The modeled nonlinearity was then applied to each one of them using Equation 4 with the estimated S parameter. Afterwards, we evaluated the energy at the envelope spatiotemporal frequency by applying the Fourier transform to each stimulus. In all the conditions, we found that the energy induced by the CM signal was always orders of magnitude lower than the energy induced by the LM signal. In other words, the difference between modeling early nonlinearities using the Naka-Rushton equation and simply adding a luminance modulation proportional to the contrast modulation was too weak to generate a detectable LM signal, and this, at any contrast level. We therefore concluded that, even when CM stimuli were presented well above threshold, they were processed by CM sensitive mechanisms and not LM sensitive mechanisms due to a global distortion artifact.

CM sensitive mechanisms

As mentioned above, the resulting nonlinearity may be defined as the sum of the nonlinearity added to the stimulus and the early nonlinearity added by the visual system. When one is the inverse of the other (one expansive and the other compressive with the same magnitude), the resulting nonlinearity is 0. However, when they differ, a LM signal is added to the effective stimulus (the resulting stimulus after applying the early nonlinearity of the visual system). If the difference is strong enough (contrast of a LM signal induced by the resulting nonlinearity greater than the contrast threshold to LM stimuli), then LM sensitive mechanisms could process such artifact. The black horizontal line in Figure V-1 represents the range within which CM stimuli must be processed by CM sensitive mechanisms. This range was calculated as the stimulus nonlinearity canceling the early nonlinearity of the visual system (where the two lines cross) +/- the LM/CM contrast threshold ratio (T_{LM}/T_{CM}). Consequently, within this range, the performance level observed when presenting a CM signal alone is not due to an early nonlinearity permitting the LM sensitive mechanisms to process CM signals.

The × signs shown in Figure V-1 correspond to the proportion of correct answers to CM stimuli presented alone (C_{LM}=0) with a nonlinearity added to the stimulus. As described above, the nonlinearity added was arbitrarily set based on a pilot study. This value is shown by the × positions on the horizontal axis. As it can be observed, all × signs are within the range in which CM stimuli must be processed by CM sensitive mechanisms (horizontal black line). Furthermore, the proportion of correct answers to CM stimuli presented alone (that is, with a nonlinearity but without a LM signal) are all above chance level (50%). We therefore conclude that at all temporal frequencies tested there are mechanisms sensitive to CM stimuli even if we compensate for early nonlinearities. This does not imply that the mechanisms processing LM and CM stimuli are distinct. A mechanism sensitive to CM stimuli could also be sensitive to LM stimuli. These results rather imply that, in these conditions, CM stimuli were not processed by mechanisms only sensitive to LM stimuli after being distorted by an early nonlinearity introducing a LM signal within a CM stimulus.

Phase independent test

The + signs shown in Figure V-1 correspond to the proportion of correct answers when presenting a LM stimulus alone. For such stimuli, the nonlinearity added to the stimulus (n) has no impact since there was no CM signal (C_{CM}=0). Consequently, there is no defined position on the horizontal axis for LM signals presented alone. We arbitrarily chose to set the position on this axis for LM signals presented alone to the measured compensating nonlinearity (where the two lines cross).

When compensating for the early nonlinearity of the visual system, the proportion of correct answers to the combination of LM and CM signals either in-phase or in counter-phase (performance level where the two fitted lines cross) was generally greater than, or close to, the proportion of correct answers to LM or CM signals alone (+ and × signs). Consequently, in all the temporal frequencies tested, we conclude that there are CM sensitive mechanisms able to discriminate the motion direction. Indeed, if CM stimuli could only be detected due to an early nonlinearity within the visual system, then LM and CM signals should cancel one another either when combined in phase or in counter-phase (Lu & Sperling, 1995, 2001). However, the opposite pattern was observed: combining both either in-phase or in counter-phase generally results in a better performance.

Similar results showing that there are CM sensitive mechanisms even up to 15 Hz have been previously found (Scott-Samuel & Georgeson, 1999). However, it was important to replicate similar experiments to measure the early nonlinearity of the visual system at each temporal frequency of each subject, and to show that with the parameters used and at all temporal frequencies tested, CM stimuli were processed by CM sensitive mechanisms, not by LM sensitive mechanisms following an early nonlinearity. Again, this does not mean that LM and CM stimuli are processed by separate mechanisms. It rather implies that even if we compensate for early nonlinearities, there are mechanisms sensitive to CM stimuli.

Experiment 2: Inter-attribute interaction

We previously found no or little inter-attribute interaction between LM and CM static stimuli processing (Allard & Faubert, 2007). LM noise affected LM signal detection but had little or no impact on CM signal detection and, vice versa, CM noise affected CM signal but had little or no impact on LM signal detection. This double dissociation strongly suggests that LM and CM signals are detected, at least at some point, by separate mechanisms.

In another study (Allard & Faubert, 2006), we also found similar detection thresholds for LM and CM stimuli embedded in high LM and CM noise respectively; that is, similar CEs but different IENs were obtained for detecting static LM and CM stimuli. In other words, observers were just as efficient at detecting LM signals embedded in LM noise as detecting CM signals embedded in CM noise.

The main purpose of the second experiment was to apply a similar noise masking paradigm to LM and CM motion processing. We therefore evaluated the contrast thresholds of LM and CM stimuli embedded in LM and CM noise.

Method

Many aspects of the methodology used in the second experiment were the same as the ones used in the previous experiment. In the current section, only their differences are presented.

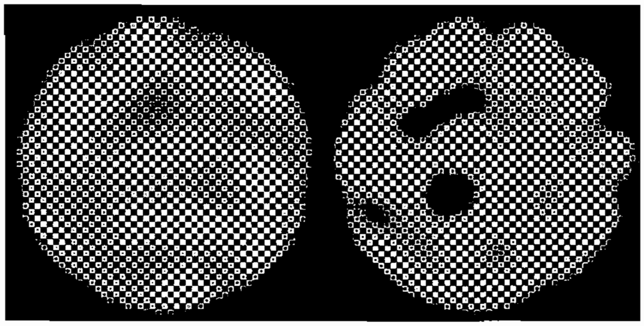
Stimuli

The modulation functions defining the stimuli in the previous experiment (Equation 2 and 3) were altered in order to add LM and CM noise to the stimulus. Consequently, an extra term corresponding to the noise function ($N(x,y,t)$) was added. Similarly to the signal (the two sine wave gratings) the noise could either be LM or CM:

$$M_{LM}(x,y,t) = 1 + (C_{LM} + nC_{CM})\sin(sx + ft + p) + (N_{LM} + nN_{CM})N(x,y,t), \qquad (5)$$

$$M_{CM}(x,y,t) = 1 + C_{CM}\sin(sx + ft + p) + N_{CM}N(x,y,t), \qquad (6)$$

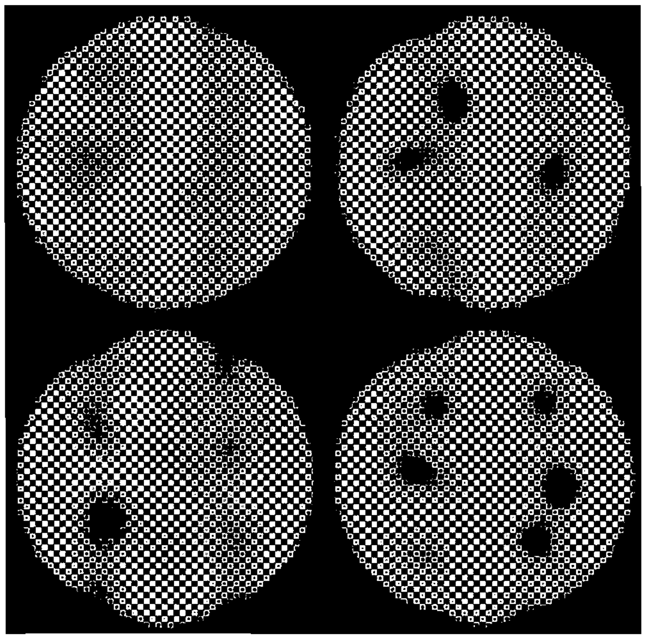
where N_{LM} and N_{CM} correspond to the contrast of the LM and CM noise, respectively. nN_{CM} corresponds to a LM noise added to compensate for the early nonlinearity within the visual system. We supposed that the proportion of CM information being converted into LM information by an early nonlinearity of the visual system is the same for the signal and noise. We therefore simply applied a linear model to compensate the nonlinearity of the visual system in which we supposed that the LM function was proportional (by a factor of n) to the CM function. $N(x,y,t)$ represents the noise function defined as filtered noise following a Gaussian distribution centered on 0 and with a root-mean-square of 1 after being filtered. The noise was filtered in the Fourier domain by an ideal mask keeping only the temporal and spatial frequencies within one octave below and above the frequency of the signal. Movie V-2 illustrates examples of LM and CM noise.



Movie V-2. LM (left) and CM (right) noise. The noise was filtered 1 octave above and below the spatiotemporal frequency of the signal. The spatial frequency of the signal was always 0.5 cpd, i.e. the frequencies within 0.25 and 1 cpd were kept. The temporal frequency varied from one condition to another. In the present example, only the temporal frequencies within 1 and 4 Hz were kept. (Movie available on the CD attached to the thesis.)

Procedure

For each temporal frequency, thresholds for LM and CM signals (C_{LM} and C_{CM}, respectively) were evaluated in five different levels of either LM noise (N_{LM}=0.0088, 0.018, 0.035, 0.071 and 0.14) or CM noise (N_{CM}=0.071, 0.10, 0.14, 0.20 and 0.28). Each threshold was measured using one staircase controlling either C_{LM} or C_{CM} (the other parameter was fixed to 0) as described in the previous experiment. The order of the testing was blocked relative to the temporal frequencies, but the order of the 20 thresholds for one block (4 signal-noise conditions (Movie V-3) and 5 noise levels) was randomized. For each temporal frequency, the nonlinearity (n) induced within the stimulus was the one measured in the previous experiment.



Movie V-3. LM and CM signal embedded in LM and CM noise. In the top row, the signals are LM. In the bottom row, the signals are CM. In the left column, the noise is LM. In the right column, the noise is CM. (Movie available on the CD attached to the thesis.)

Results and discussion

Calculation efficiency

Figure V-2 shows the results for LM and CM thresholds embedded in LM and CM noise, respectively. These results were fitted using the TvC function (see Allard and Faubert (2006) for details) known to give a good fit for contrast thresholds as a function of noise contrast when the signal and noise are of the same type (Legge, Kersten, & Burgess, 1987; Pelli, 1981, 1990; Pelli & Farell, 1999). In certain conditions, even at the highest CM noise level, CM thresholds were hardly affected by the noise (especially for observer JR at 16 Hz). These data result in a good fit for flat portion of the TvC function, but give a poor fit for the rising part of the function. To improve the fit, we introduced an extra constraint: the calculation efficiency to CM stimuli could not be greater than the calculation efficiency to LM stimuli. In other words, the rising parts of the TvC functions of the blue dashed lines fitting CM thresholds in Figure V-2 were constrained to be equal or above the rising parts of the red solid lines fitting LM thresholds. This constraint had no or little impact in almost all the conditions. However, without it at 16 Hz for observer JR the fit resulted in a straight line corresponding to extremely high calculation efficiency and internal equivalent noise. For this condition, the data were only driven by the flat portion of the TvC function, which only gives a lower bound to the internal equivalent noise and calculation efficiency. However, we know of no models suggesting that we could be more sensitive to a given signal embedded in noise when both are CM rather than LM. We therefore think that this extra constraint is justified and enables the fit to set a lower bound to the calculation efficiency to CM stimuli when the highest noise contrast did not significantly affect the threshold.

As expected, LM and CM thresholds largely differed in the absence of noise. However, when the noise was sufficiently high, there was no or little threshold difference between LM and CM signals. In other words, observers were just as efficient at discriminating the direction of a LM signal in LM noise as for a CM signal in CM noise. Consequently, observers had similar calculation efficiencies to both attributes, and this, for a wide temporal frequency range. These results therefore suggest that, for LM and CM stimuli, common mechanisms could be extracting the signal from noise.

Internal equivalent noise

Since the difference of sensitivity to LM and CM stimuli processing was not due to a difference of CE, it was obviously due to a difference of IEN corresponding to the breaking point on the TvC function. As mentioned in the introduction, similar results were obtained for static stimuli (Allard & Faubert, 2006) which led us to suggest that the difference of IEN could be due to a suboptimal rectification process for CM stimuli. However, based on the difference of IEN, one cannot conclude that both attributes are processed by separate mechanisms. If both stimuli are processed by common mechanisms, then the difference of sensitivity would likely be due to different contrast gains. Different contrast gains prior to the main noise source would increase the relative impact of the main noise source and thereby affect the observer's threshold. However, if the main noise source is external (when the external noise is greater than the IEN) the contrast gain would affect both the signal and noise contrasts without affecting the signal-to-noise ratio. Consequently, in high noise conditions, the threshold would be independent of the contrast gain. As a result, one cannot conclude that two stimuli are processed by separate mechanisms simply based on different IENs since the common mechanism hypothesis would also predict this pattern of consequences.

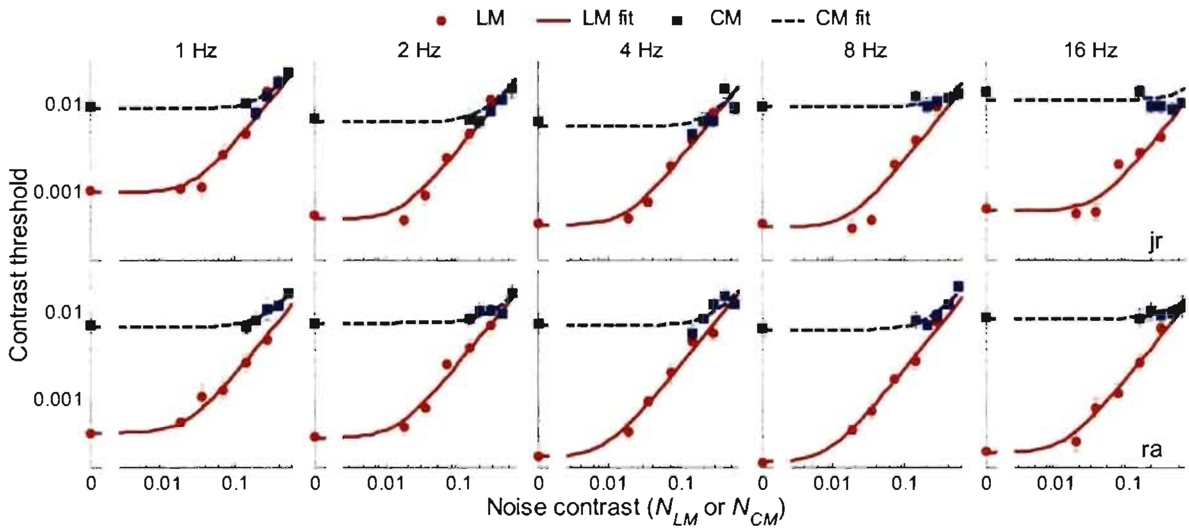


Figure V-2. Motion discrimination in intra-attribute noise. Red dots and solid lines correspond to LM contrast thresholds in LM noise (raw data and fitted TvC function, respectively). Blue squares and dashed lines correspond to CM contrast thresholds in CM noise. Error bars corresponds to the standard deviation from the mean.

CM processing in LM noise

At 16 Hz, LM noise affected both LM and CM thresholds by similar proportions (Figure V-3). In other words, LM noise had the same relative impact on LM and CM thresholds. Although the noise was bandpass at one octave below and above the signal spatiotemporal frequencies, the noise did not selectively impair LM processing without having the same impact on CM processing. In other words, the smallest LM noise contrast significantly affecting CM thresholds matches the one affecting LM thresholds and is highly different from the smallest CM noise contrast significantly affecting CM thresholds shown in Figure V-2. This suggests that, at 16 Hz, LM and CM stimuli are processed by common mechanisms.

At 8 Hz, a similar pattern of results was observed for external noise contrasts affecting LM or CM thresholds by a factor less than 3 or 4 (Figure V-3). This suggests that, at 8 Hz, LM and CM signals are processed by common mechanisms since both attributes were affected by similar proportions. Above this critical value, CM thresholds were less affected by LM noise than LM thresholds. This suggests the existence of separate mechanisms processing LM and CM signals. Taken together, these results suggest that two mechanisms could be processing CM stimuli. The more sensitive one (the one processing CM stimuli at 8 Hz in noiseless conditions) would be common to LM processing. However, the less sensitive one (here by a factor of about 3 or 4) would not be common to LM processing explaining why high LM noise contrasts affected more LM than CM processing at 8 Hz.

At lower temporal frequencies (1, 2 and 4 Hz), CM processing was generally less affected by LM noise than LM processing (Figure V-3). Indeed, the two curves had the tendency to split at the point where thresholds increased by a factor of about 3, 1.5 and 2 for the temporal frequencies 1, 2 and 4 Hz respectively. This suggests that, at least in high LM noise conditions, LM and CM signals are processed by separate mechanisms.

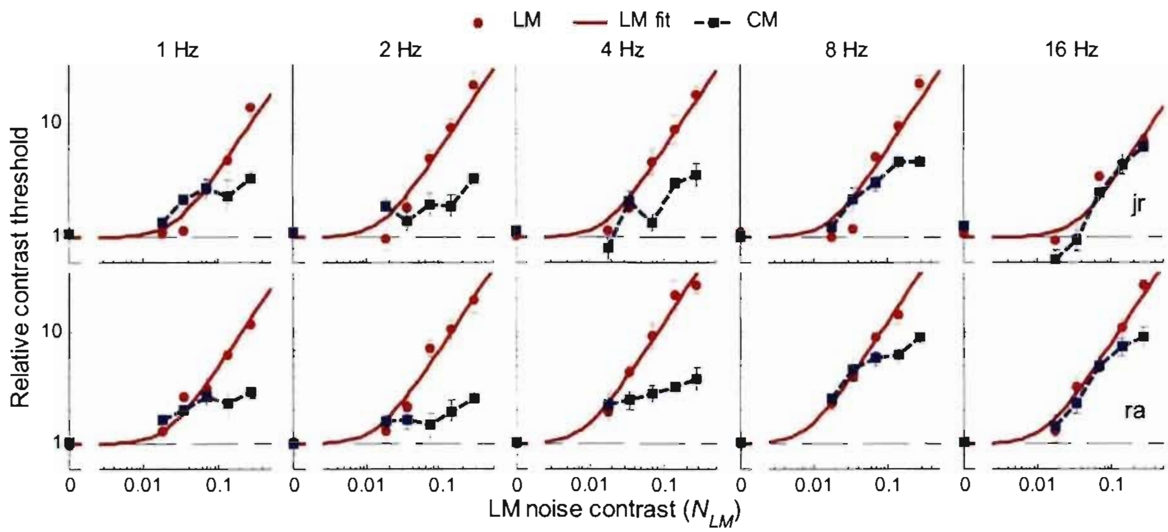


Figure V-3. LM and CM relative contrast thresholds as a function of LM noise contrast. Contrast thresholds are represented relative to their contrast thresholds in absence of noise fitted by the TvC function (where the fitted curves cross the Y-axis in Figure V-2). For LM stimuli, the thresholds and best fitted TvC functions are represented (red dots and solid lines, respectively). For CM stimuli, since the external noise was not of the same type, we could not fit the TvC function and only the evaluated thresholds are represented (blue squares). Error bars corresponds to the standard deviation from the mean.

LM processing in CM noise

Unfortunately, even in high CM noise, the impact of the noise on CM processing was relatively limited (Figure V-4). Consequently, it was not possible to have sufficient noise contrast to severely impair CM processing. However, since at the highest noise contrast CM processing was generally significantly affected, the impact on LM processing could also be evaluated. At temporal frequencies at or above 4 Hz, LM processing seemed to be affected in similar proportions as CM processing (except at 16 Hz for subject JR, who's LM and CM thresholds remained unaffected by the highest CM noise contrast). These results also support the hypothesis that both LM and CM stimuli are processed by common mechanisms at high temporal frequencies.

At lower temporal frequencies (1 and 2 Hz), LM processing was generally less affected than CM processing (Figure V-4). Indeed, at these temporal frequencies, CM noise impaired CM processing more than LM processing suggesting that both attributes are processed by separate mechanisms.

Experiment 3: Carrier and nonlinearity control

Even though component motion could not explain the results obtained in the previous experiment, the choice of a periodic carrier remains an issue. For instance, a checkerboard carrier has constant luminance along the diagonals. Some could therefore argue that CM stimuli have luminance modulations along these diagonals that could be detected by LM sensitive mechanisms. However, the use of relatively small check size (21 checks per signal cycles) should minimize such artifacts. Nonetheless, many experimenters prefer using noise as a carrier rather than a regular structure.

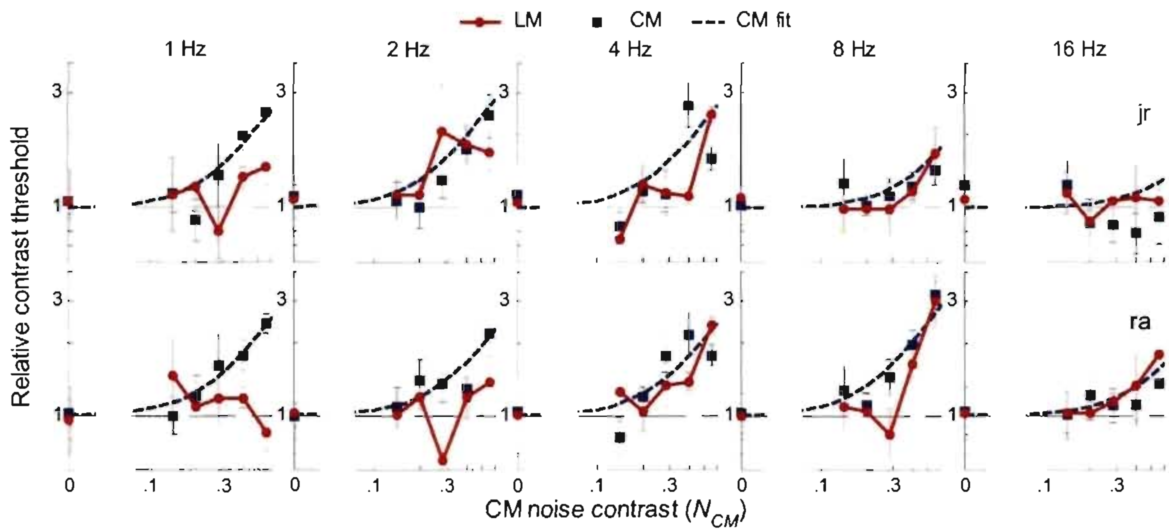


Figure V-4. LM and CM relative contrast thresholds as a function of CM noise contrast. Contrast thresholds are represented relative to their contrast thresholds in absence of noise fitted by the TvC function (where the fitted curves cross the Y-axis inFigure V-2). For CM stimuli, the thresholds and best fitted TvC functions are represented (blue squares and dashed lines, respectively). For LM stimuli, since the external noise was not of the same type, we could not fit the TvC function and only the evaluated thresholds are represented (red circles). Error bars corresponds to the standard deviation from the mean.

Smith and Ledgeway (1997) have shown that using static binary noise carriers with large element size can give rise to local first-order artifacts as described in experiment 1. They suggested the used of either dynamic broadband noise or static highpass noise. An important disadvantage of using dynamic noise for our purpose is that the noise introduced by the carrier would affect contrast thresholds to both LM and CM stimuli which thereby reduces the impact of adding LM or CM noise. We therefore conducted a control experiment using static highpass noise as a carrier.

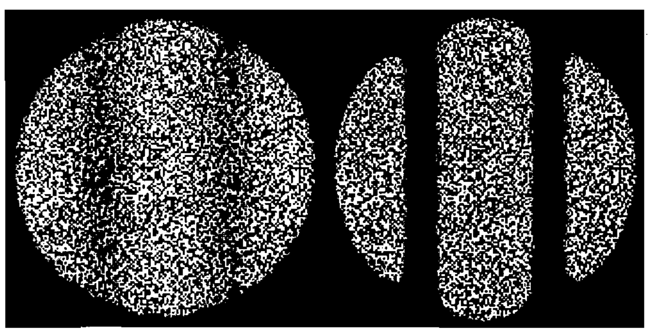
Another artifact that could have influenced our results at high temporal frequencies is if the nonlinearity of the visual system was not well compensated for by the proportional nonlinearity applied to the stimulus resulting into a global distortion product. Indeed, we measured the nonlinearity at threshold and then supposed that the nonlinearity induced by the visual system was proportional to the CM signal contrast. Scott-Samuel and Georgeson (1999) suggested using the Naka-Rushton equation to compensate for the visual system nonlinearities. Even though we have demonstrated using simulations in experiment 1 that the difference between the two models are too small to affect our results, the presence of LM or CM noise could affect the compressive nonlinearity. As a result, it could be argued that our results suggesting common mechanisms at high temporal frequency could be due to a global distortion product that would be different in noise conditions. The present experiment focused on a low (2 Hz) and a high (8 Hz) temporal frequency. We evaluated LM and CM discrimination thresholds in noiseless, LM and CM noise conditions. In order to assert that the results at 8 Hz were not due to some global distortion product, we directly measured the compressive nonlinearity in all three conditions, i.e. no noise, LM noise and CM noise.

Method

Many aspects of the methodology used in the present experiment were the same to the ones used in the previous experiments. In the current section, only their differences are presented.

Stimuli

The stimuli used were similar to the ones in the previous experiment with the exception of the carrier ($T(x,y)$). The carrier was generated by creating a binary noise texture ($T(x,y)$=-0.5 or 0.5) with element size equal to 2x2 pixels, which was then highpass filtered with a cutoff frequency at 4 cpd (Movie V-4). Such filtering had little impact on the RMS contrast of the carrier reducing it by 3%.



Movie V-4. LM (left) and CM (right) stimuli. The gratings are drifting at 2 Hz. The carrier corresponds to binary noise which was filtered to keep only the frequencies above 4 cpd. (Movie available on the CD attached to the thesis.)

Procedure

For each temporal frequency (2 and 8 Hz), LM and CM direction discrimination thresholds were evaluated in noiseless conditions, LM noise (N_{LM}=0.020) and in CM noise (N_{CM}=0.28). Note that the LM noise was not set to its maximal contrast. We chose the noise contrast based on the results obtained at 8 Hz of the previous experiment such that both LM and CM detection thresholds were affected by similar proportions. As discussed in the previous experiment, when the LM noise contrast was too high, CM thresholds were less affected than LM thresholds suggesting that another mechanism was able to process CM stimuli. The same noise contrast was used at both temporal frequencies. Each threshold was evaluated three times using the 2-down-1-up staircase as described in the first experiment.

In experiment 1, we found that the compressive nonlinearity of the visual system was too weak at 2 Hz to introduce global first-order artifacts within CM stimuli that would be detectable by LM sensitive mechanisms. In the current experiment, we therefore did not measure the compressive nonlinearity and assumed that it was null (n=0). The double dissociation observed (see Results and discussion section below) ensures that the nonlinearities of the visual system did not affect the results.

Since we expected complete inter-attribute interactions at 8 Hz, it was important to ensure that the results were not due to compressive nonlinearities introducing LM signals within CM stimuli. Instead of applying the inverse of a given retinal model (e.g. Naka-Rushton equation) to compensate for early nonlinearities, we used the proportional model as in the first two experiments and directly measured the nonlinearity in all three noise conditions (no noise, LM noise and CM noise). We therefore measured the compressive nonlinearity of the visual system for each observer and applied the inverse nonlinearity (parameter n) to the stimulus as done in the previous experiments. Afterwards, the nonlinearity was reevaluated for each noise conditions to show that in all conditions the global distortion product was too low to explain our results.

Results and discussion

Separate mechanisms at low temporal frequencies

At 2 Hz, LM noise had a significantly greater impact on LM processing than on CM processing for both observers and CM noise had a significantly greater impact on CM processing than on LM processing (Figure V-5 left). This double dissociation confirms the results obtained in the previous experiment strongly suggesting that both attributes are processed by separate mechanisms. Indeed, the fact that it was possible to selectively impair either processing implies that both must be processed, at least at some point, by separate mechanisms.

Common mechanisms at high temporal frequencies

At 8 Hz, a completely different pattern of results was observed (Figure V-5right), LM and CM noises each affected LM and CM processing by similar proportions. These complete inter-attribute interactions suggest that both attributes are processed by common mechanisms at high temporal frequencies.

Figure V-6 shows the nonlinearity measured for each of the noise conditions. The stimulus nonlinearities in noiseless conditions found to compensate for early nonlinearities of the visual system (n=0.022 and 0.014 for observers JR and RA respectively) were the ones implemented when measuring contrast thresholds. The nonlinearities in LM and CM noise were measured afterwards to show that the global distortion product could not explain by itself CM thresholds. In order for a global distortion product to cause a sufficiently high artifact detectable by LM sensitive mechanisms, the difference between the nonlinearity applied to the stimulus and the nonlinearity generated by the visual system needs to be greater than the LM/CM threshold ratio. As shown in Figure V-6, this was not the case in all conditions. Indeed, the nonlinearities applied to the stimuli when measuring contrast thresholds in noise (represented by the horizontal position of × signs, i.e. n=0.022 and 0.014 for observers JR and RA respectively) felled within the range CM stimuli must be processed by CM sensitive mechanisms illustrated by the black horizontal lines. We therefore conclude that, in all tested conditions, CM stimuli were processed by CM sensitive mechanisms which, we suggest, are the same as LM sensitive mechanisms.

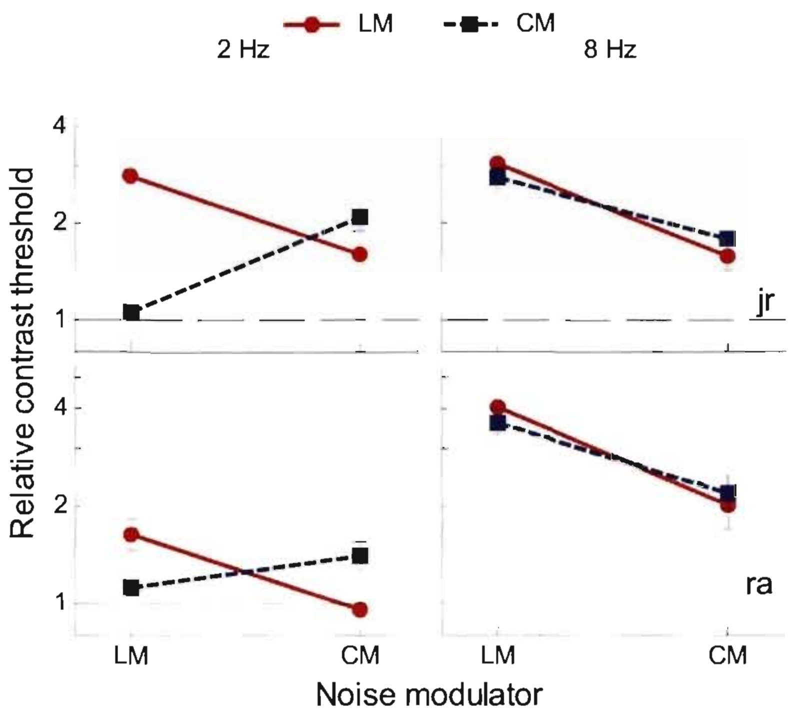


Figure V-5. Impact of LM and CM noise on LM and CM thresholds. The y-axis shows the contrast thresholds relative to contrast threshold in absence of noise. The x-axis represents the type of noise added to the stimulus (LM or CM). Error bars show standard error of the three thresholds measured for each condition.

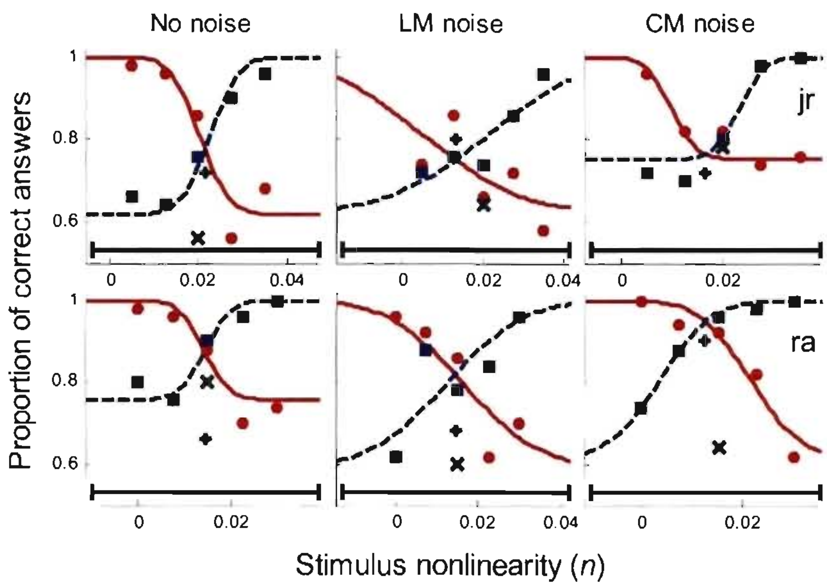


Figure V-6. Nonlinearity results at 8 Hz. The legend is the same as the one described in the caption of Figure V-1.

General discussion

Common mechanisms at high temporal frequencies

The noise-masking paradigm that we developed successfully showed a double dissociation between LM and CM processing for static and low temporal frequencies. Indeed, in these conditions, LM and CM noise were each able to selectively impair same attribute processing with limited impact on cross-attribute processing. However, at high temporal frequencies, LM and CM noise each had similar impact on LM and CM processing. In other words, the noise-masking paradigm failed to impair the processing of either attribute without affecting the other. Consequently, our results did not show a double dissociation between LM and CM processing at high temporal frequencies. The complete interaction between the processing of these attributes suggests that, under these conditions, both attributes are processed by common mechanisms. Note that early nonlinearities prior to motion processing, which could introduce LM signal within CM stimulus, cannot by themselves explain the sensitivity to CM stimuli. Indeed, observers were able to process CM stimuli even when we compensated for such early nonlinearities. Consequently, if LM and CM stimuli are processed by common mechanisms, the nonlinearity enabling CM processing must occur within (and not prior to) motion processing as suggested by the gradient-based model.

There is no reason to think that moving LM gratings are processed by fundamentally different mechanisms (i.e. different processing strategies) depending on the temporal frequency. Indeed, it is generally assumed that there are similar LM sensitive mechanisms tuned to different spatiotemporal frequencies. Consequently, if nonlinearities within the motion processing of LM stimuli at high frequencies enable LM sensitive mechanisms to detect CM stimuli, then similar mechanisms at low temporal frequencies processing LM stimuli but tuned to lower temporal frequencies should also be able to detect CM stimuli. Hence, we do not claim that mechanisms sensitive to LM stimuli at low temporal frequencies cannot process CM stimuli. Instead, we argue that more than one type of mechanisms could be sensitive to CM stimuli and that one of them is common with the one processing LM stimuli. At threshold, CM stimuli would be processed by the most sensitive mechanism depending on the stimulus parameters (temporal frequency, carrier type, carrier contrast, etc).

If CM stimuli are processed by LM sensitive mechanisms and that LM sensitive mechanisms are known to be tuned to a particular spatiotemporal frequency, then one can wonder whether CM stimuli are detected by LM sensitive mechanisms tuned to the envelope or carrier spatial frequency. To our knowledge, the gradient-based model does not clearly specify whether the CM stimuli would be processed by LM sensitive mechanisms tuned to the carrier or envelope spatial frequency. The complete interaction observed at high temporal frequencies suggests that the mechanisms processing LM and CM stimuli would be tuned to the envelope spatial frequency and not the carrier spatial frequency. Indeed, complete interaction was observed at the envelope spatial frequency. Consequently, our results suggest the existence of mechanisms sensitive to both LM and CM stimuli defined by the same signal modulation (i.e. envelope) spatial frequency.

Separate mechanisms at low temporal frequencies

There were no instances in which noise affected the signal processing to a greater extent when the signal and noise were of different attributes compared to when they were of the same attribute. LM noise never significantly affected CM more than LM processing and CM noise never significantly affected LM more than CM processing. However, in certain noise conditions, inter-attribute processing was affected at equivalent levels to intra-attribute processing while in other noise conditions inter-attribute processing was less affected than intra-attribute processing.

In the noise conditions where an asymmetrical impact was observed, we conclude that both attributes must be processed, at least at some point, by separate mechanisms.

Separate mechanisms could either mean similar mechanisms tuned to different spatial frequencies or fundamentally different mechanisms. However, as suggested above for high temporal frequencies, a mechanism sensitive to a LM signal at a given spatial frequency would also be sensitive to CM stimuli based on the envelope spatial frequency. We therefore conclude that both attributes are processed by fundamentally distinct mechanisms when a double dissociation is observed. Indeed, a model suggesting that both attributes are processed by common mechanisms at all processing levels, such as the gradient-based model, could not explain both the double dissociation at low temporal frequencies and the complete interaction at high temporal frequencies.

As seen above, the dissociations between LM and CM processing were more important at low temporal frequencies. This suggests that the dedicated mechanisms processing LM and CM stimuli are more sensitive at low temporal frequencies. Combining these results with the fact that cross-attribute noise has little or no impact on spatial processing (Allard & Faubert, 2006) suggests that the separate mechanisms processing LM and CM stimuli first need to extract spatial information before processing motion. The filter-rectify-filter model suggesting that both attributes are initially processed by separate mechanisms also proposes that the spatial structure of CM stimuli are first extracted (rectification process) before evaluating the motion direction. This initial spatial processing stage could explain why such mechanisms would be more sensitive at low temporal frequencies. Indeed, the sensitivity to most second-order stimuli defined by attributes other than contrast, such as by depth (Lu & Sperling, 1995) or polarity reversals (Bellefeuille & Faubert, 1998; Hutchinson & Ledgeway, 2006), is generally found to be temporally lowpass. After the rectification process (or any process extracting the spatial structure of the envelope), moving CM stimuli could either be processed by an energy-based mechanism dedicated to second-order processing or by a feature tracking mechanism comparing the spatial position of the envelope in time. Our results do not enable us to dissociate these models. Based on other studies, we addressed this question in the following section.

Motion processing of CM stimuli

We are not the firsts to suggest that CM stimuli can be processed by fundamentally different mechanisms depending on the testing parameters (here the temporal frequency). Seiffer and Cavanagh (1999) evaluated whether CM stimuli are processed by an energy-based or a position-based mechanism by measuring motion amplitude thresholds of oscillating gratings. They concluded that CM stimuli could be processed by either mechanism. An energy-based mechanism would be more sensitive to CM stimuli at high temporal frequencies and high carrier contrasts, and a position-based mechanism would be more sensitive at low temporal frequencies and low carrier contrasts. A position-based mechanism would first require spatial processing in order to extract the spatial properties of the signal and then compare its positions at different time intervals. Their results are in agreement with ours given that LM stimuli are always processed by energy-based mechanisms (Zaidi & DeBonet, 2000). At low temporal frequencies, CM stimuli would be processed by a position-based mechanism once the spatial structure has been extracted. At high temporal frequencies, LM and CM stimuli would be processed by common energy-based mechanisms. Note that equivalence between gradient- and energy-based models have been shown (Christopher P. Benton, 2004). Indeed, Benton demonstrated that energy-based models combined with a subsequent contrast normalization process could discriminate the motion direction of CM stimuli.

Ukkonen and Derrington (2000) suggested similar conclusions using the pedestal test (Lu & Sperling, 1995) to evaluate whether the carrier contrast is processed by feature tracking mechanisms or by spatiotemporal filters. Using a low contrast carrier, they found that CM stimuli processing did not pass the pedestal test and was only possible at low temporal frequencies (<4 Hz). They concluded that when using a low contrast carrier, CM stimuli are processed by a feature tracking mechanism. On the other hand, when using a high contrast carrier, CM stimuli processing was unaffected by the pedestal and could be processed at higher temporal frequencies (up to the highest frequency tested of 12 Hz).

They conclude that CM stimuli are processed by spatiotemporal filters when using a high contrast carrier. Again, these results are in agreement with ours. There could be two types of mechanisms processing CM stimuli, one lowpass initially extracting the spatial structure of the envelope and another which, we suggest, could be common to LM processing.

These studies show that CM stimuli could be processed by feature tracking (or position-based) mechanisms and by energy-based mechanisms depending on the stimulus conditions. At high temporal frequencies and high carrier contrasts, the energy-based mechanism would be more sensitive to CM stimuli than the feature tracking mechanism. At low temporal frequencies and low carrier contrasts, the feature tracking mechanism would be more sensitive. Since our results suggest that LM and CM stimuli are processed by common mechanisms at high temporal frequencies ant that LM sensitive mechanisms are known to be energy-based, our results also suggest that CM stimuli are processed by energy-based mechanisms. Since LM stimuli are always processed by energy-based mechanisms, the results obtained by Ukkonnen and Derington, and Seiffer and Cavanagh showing that CM stimuli are processed by a feature tracking (or position-based) mechanism at low temporal frequencies also suggests that LM and CM stimuli are processed by separate mechanism and are consistent with our results. Consequently, this suggests that when a double dissociation was observed, CM stimuli were processed by a feature tracking mechanism. To explain our results, there is no need to define an energy-based mechanism dedicated to second-order stimuli. CM stimuli could either be processed by the same mechanisms processing LM stimuli (which would thereby be energy-based) or by a separate feature tracking mechanism. Nonetheless, before a feature can be tracked, it must be extracted. The filter-rectify-filter model could explain the extraction of the spatial structure of CM stimuli. The second filtering stage of the filter-rectify-filter model could only extract the spatial structure of the envelope and not its spatiotemporal structure (which would result in an energy-based mechanism dedicated to second-order processing) as it is generally suggested. In other words, after the second filtering stage, a feature tracking mechanism could detect spatial changes of position in time. As a result, there would be no dedicated second-order motion mechanisms. In conclusion, therefore, the most heuristic proposition would be that CM stimuli could either be processed by mechanisms able to extract and track the envelope feature of CM stimuli or by energy-based mechanisms common to both LM and CM stimuli.

Conclusion

Cumulative evidence suggests that first- and second-order stimuli are processed by distinct mechanisms. We are partially in agreement with this hypothesis. The dissociations observed at low temporal frequencies (LM noise affecting more LM than CM processing, and CM noise affecting more CM than LM processing) suggest that, at least in some conditions, both pathways are indeed initially distinct. Consequently, we also conclude that a single motion sensitive mechanism cannot explain motion perception in all conditions. However, we do not conclude that first- and second-order stimuli are always processed by distinct mechanisms. We suggest that, at high temporal frequencies, CM stimuli could be processed by the same mechanisms as the ones processing LM stimuli. Some second-order stimuli may not be invisible to first-order sensitive mechanisms and the existence of energy-based mechanisms dedicated to second-order processing is thereby questionable.

Acknowledgments: This research was supported by NSERC graduate fellowship to RA and NSERC operating grant to JF.

References

A llard, R., & Faubert, J. (2006). Same calculation efficiency but different internal noise for luminance- and contrast-modulated stimuli detection. *Journal of Vision, 6*(4), 322-334, http://journalofvision.org/6/4/3/, doi:10.1167/6.4.3.

Allard, R., & Faubert, J. (2007). Double dissociation between first- and second-order processing. *Vision Research, 47*(9), 1129-1141.

Baker, C. L., Jr. (1999). Central neural mechanisms for detecting second-order motion. *Curr Opin Neurobiol, 9*(4), 461-466.

Bellefeuille, A., & Faubert, J. (1998). Independence of contour and biological-motion cues for motion-defined animal shapes. *Perception, 27*(2), 225-235.

Benton, C. P. (2002). Gradient-based analysis of non-Fourier motion. *Vision Research, 42*(26), 2869-2877.

Benton, C. P. (2004). A role for contrast-normalisation in second-order motion perception. *Vision Research, 44*(1), 91-98.

Benton, C. P., & Johnston, A. (2001). A new approach to analysing texture-defined motion. *Proc Biol Sci, 268*(1484), 2435-2443.

Benton, C. P., Johnston, A., McOwan, P. W., & Victor, J. D. (2001). Computational modeling of non-Fourier motion: further evidence for a single luminance-based mechanism. *Journal of the Optical Society of America, A, Optics, Image Science, & Vision, 18*(9), 2204-2208.

Cavanagh, P., & Mather, G. (1989). Motion: the long and short of it. *Spat Vis, 4*(2-3), 103-129.

Chubb, C., & Sperling, G. (1988). Drift-balanced random stimuli: a general basis for studying non-Fourier motion perception. *J Opt Soc Am A, 5*(11), 1986-2007.

Georgeson, M. A., & Schofield, A. J. (2002). Shading and texture: separate information channels with a common adaptation mechanism? *Spat Vis, 16*(1), 59-76.

He, S., & Macleod, D. I. (1998). Contrast-modulation flicker: dynamics and spatial resolution of the light adaptation process. *Vision Research, 38*(7), 985-1000.

Hutchinson, C. V., & Ledgeway, T. (2006). Sensitivity to spatial and temporal modulations of first-order and second-order motion. *Vision Research, 46*(3), 324-335.

Johnston, A., Benton, C. P., & McOwan, P. W. (1999). Induced motion at texture-defined motion boundaries. *Proceedings Biological Sciences/The Royal Society, 266*(1436), 2441-2450.

Johnston, A., & Clifford, C. W. (1995). Perceived motion of contrast-modulated gratings: predictions of the multi-channel gradient model and the role of full-wave rectification. *Vision Res, 35*(12), 1771-1783.

Johnston, A., McOwan, P. W., & Buxton, H. (1992). A computational model of the analysis of some first-order and second-order motion patterns by simple and complex cells. *Proc Biol Sci, 250*(1329), 297-306.

Legge, G. E., & Foley, J. M. (1980). Contrast masking in human vision. *Journal of the Optical Society of America, 70*(12), 1458-1471.

Legge, G. E., Kersten, D., & Burgess, A. E. (1987). Contrast discrimination in noise. *J Opt Soc Am A, 4*(2), 391-404.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America, 49*(2), Suppl 2:467+.

Lu, Z. L., & Sperling, G. (1995). The functional architecture of human visual motion perception. *Vision Research, 35*(19), 2697-2722.

Lu, Z. L., & Sperling, G. (2001). Three-systems theory of human visual motion perception: review and update.[see comment][erratum appears in J Opt Soc Am A Opt Image Sci Vis 2002 Feb;19(2):413]. *Journal of the Optical Society of America, A, Optics, Image Science, & Vision, 18*(9), 2331-2370.

MacLeod, D. I., Williams, D. R., & Makous, W. (1992). A visual nonlinearity fed by single cones. *Vision Research, 32*(2), 347-363.

Pelli, D. G. (1981). *The effects of visual noise.* Cambridge University, Cambridge.

Pelli, D. G. (1990). The quantum efficiency of vision. In C. Blakemore (Ed.), *Visual coding and efficiency*. Cambridge: Cambridge University Press.

Pelli, D. G., & Farell, B. (1999). Why use noise? *Journal of the Optical Society of America, A, Optics, Image Science, & Vision, 16*(3), 647-653.

Ross, J., & Speed, H. D. (1996). Perceived contrast following adaptation to gratings of different orientations. *Vision Research, 36*(12), 1811-1818.

Schofield, A. J., & Georgeson, M. A. (1999). Sensitivity to modulations of luminance and contrast in visual white noise: separate mechanisms with similar behaviour. *Vision Res, 39*(16), 2697-2716.

Schofield, A. J., & Georgeson, M. A. (2000). The temporal properties of first- and second-order vision. *Vision Res, 40*(18), 2475-2487.

Scott-Samuel, N. E., & Georgeson, M. A. (1999). Does early non-linearity account for second-order motion? *Vision Research, 39*(17), 2853-2865.

Seiffert, A. E., & Cavanagh, P. (1999). Position-based motion perception for color and texture stimuli: effects of contrast and speed. *Vision Research, 39*(25), 4172-4185.

Smith, A. T., & Ledgeway, T. (1997). Separate detection of moving luminance and contrast modulations: fact or artifact? *Vision Research, 37*(1), 45-62.

Snowden, R. J., & Hammett, S. T. (1992). Subtractive and divisive adaptation in the human visual system. *Nature, 355*(6357), 248-250.

Snowden, R. J., & Hammett, S. T. (1996). Spatial frequency adaptation: threshold elevation and perceived contrast. *Vision Research, 36*(12), 1797-1809.

Taub, E., Victor, J. D., & Conte, M. M. (1997). Nonlinear preprocessing in short-range motion. *Vision Res, 37*(11), 1459-1477.

Ukkonen, O. I., & Derrington, A. M. (2000). Motion of contrast-modulated gratings is analysed by different mechanisms at low and at high contrasts. *Vision Research, 40*(24), 3359-3371.

Wilson, H. R., Ferrera, V. P., & Yo, C. (1992). A psychophysically motivated model for two-dimensional motion perception. *Vis Neurosci, 9*(1), 79-97.

Zaidi, Q., & DeBonet, J. S. (2000). Motion energy versus position tracking: spatial, temporal, and chromatic parameters. *Vision Research, 40*(26), 3613-3635.

Chapitre VI

Common first- and second-order energy-based motion processing

Rémy Allard et Jocelyn Faubert

Laboratoire de psychophysique et perception visuelle

École d'optométrie, Université de Montréal

Le contenu de ce chapitre a été soumis sous forme d'article à *Journal of Vision*

Abstract

Previous studies found that the most widely studied second-order stimulus (contrast modulation) could be processed by either an energy-based or a feature tracking motion system. The apparent lack of interaction when superposing luminance- (LM) and contrast-modulated (CM) stimuli either in phase or in counter-phase suggests that global distortion products cannot explain by themselves the energy-based processing of CM stimuli. Several authors therefore inferred the existence of a dedicated second-order energy-based motion system (e.g., filter-rectify-filter model) even though LM and CM processing strongly interact in other conditions. By defining early nonlinearities as global distortion products, these authors implicitly assumed that early nonlinearities were uniform across space and time. The present study questions this assumption. Using a noise masking paradigm, we failed at dissociating LM and CM processing at high temporal frequencies suggesting that both stimuli were processed by a common motion system. In a second experiment, simulating early non-uniform nonlinearities showed that the apparent lack of interaction when superposing LM and CM stimuli does not imply separate processing. We conclude that it is not necessary to suppose the existence of a dedicated second-order energy-based motion system. Early non-uniform nonlinearities could enable the luminance energy-based motion system to process CM stimuli.

Keywords: Motion; Second-order; Energy-based; Feature tracking; Contrast-modulated; Early nonlinearities; Filter-rectify-filter model; Gradient-based model

Introduction

Energy-based and feature tracking motion systems

Our perceptual system has at least two motion systems based on fundamentally different computational strategies (for a review, see Lu and Sperling (2001)). One can be described as energy-based directly processing spatiotemporal luminance variation without extracting the feature of the drifting signal. Such a system is often referred to as energy-based, intensity-based, Fourier, linear or first-order motion system, and is typically found to be low-level, monocular and fast (i.e., cutoff frequency near 12 Hz). The other motion system first extracts a feature and then tracks its position change over time. This motion is usually referred to as a feature tracking, high-level or correspondence-based motion system, and is known to be high-level, binocular and slow (i.e., cutoff frequency near 3 Hz). In the present study, we will refer to these two motion systems as energy-based and feature tracking motion systems, respectively.

Most second-order stimuli, which are defined by other attributes than luminance (i.e. first-order attribute) such as contrast (Cavanagh & Mather, 1989; Chubb & Sperling, 1988)), were generally found to be processed by a feature tracking motion system (Lu & Sperling, 1995, 2001; Seiffert & Cavanagh, 1998). However, the most widely studied second-order stimulus consisting in a contrast modulation of given texture (e.g., binary noise) was found to be processed by either an energy-based or a feature tracking motion system depending on the stimulus parameters.

For instance, Smith (1994) dissociated these two motion systems by creating a contrast-modulated (CM) stimulus having energy and a feature moving in opposite directions. For the given testing conditions, he found that CM stimuli were normally processed by an energy-based motion system. However, when disabling the energy-based motion system by introducing a 60 ms delay between images, he found that CM stimuli were processed by a correspondence-based (i.e., feature tracking) motion system. He concluded that, although CM stimuli were normally processed by an energy-based motion system, they could also be processed by a correspondence-based motion system.

Lu and Sperling (1995) developed a pedestal test consisting in adding a stationary grating identical to the drifting grating which theoretically disables the feature tracking motion system without affecting the energy-based motion system. By applying the pedestal test, they found that CM stimuli, like luminance-modulated (LM) stimuli, were processed by an energy-based motion system. By using the pedestal test and modifying the contrast of the carrier, Ukkonen and Derrington (2000) found that the impact of the pedestal test depended on the contrast of the carrier. With a low contrast carrier, CM stimuli were found to be processed by a feature tracking motion system, but with a high contrast carrier, CM stimuli were found to be processed by an energy-based motion system.

Seiffert and Cavanagh (1999) also observed that the motion system processing CM stimuli varied as a function of the stimulus parameter. They measured the smaller displacement of an oscillatory CM grating as a function of the oscillation frequency and the modulation contrast. They found that position displacement was the cue for processing CM gratings at low speeds and low contrasts, but velocity was the cue for processing CM gratings at high speeds and high contrasts. In other words, CM stimuli at high temporal frequencies and high contrasts would be processed by an energy-based motion system and would be processed by a feature tracking motion system otherwise.

When second-order stimuli are processed by a feature tracking motion system, this system is obviously distinct from the one processing first-order stimuli which is normally found to be energy-based (Zaidi & DeBonet, 2000). On the other hand, when second-order stimuli are processed by an energy-based motion system (which has been observed using CM stimuli), one can question whether this energy-based motion system is the same as the one processing first-order (i.e., LM) stimuli. In other words, is there an energy-based motion system dedicated to second-order processing (i.e., distinct from the energy-based motion system processing first-order stimuli)? This question was the target of the present study.

Distinct first- and second-order energy-based motion systems?

Several studies address the question of whether LM and CM stimuli are processed by common or distinct motion systems but few considered that CM stimuli could be processed by two distinct motion systems. Even though some studies found convincing evidence that first- and second-order stimuli were processed separately, this can be explained by the fact that second-order stimuli were processed by a feature tracking motion system. As a result, many studies do not address the question of whether first- and second-order stimuli are processed by common or distinct energy-based motion systems. The present section selectively focused on the studies that considered that CM stimuli can be processed by two fundamentally distinct motion systems.

Filter-rectify-filter model

Lu and Sperling (1995) directly addressed the question of whether first- and second-order stimuli are processed by common or distinct energy-based motion systems. They found evidence that LM and CM stimuli are processed by common energy-based mechanisms and evidence that they are processed by distinct mechanisms. They observed motion cancellation, rather than motion transparency, when superposing an LM and a CM drifting grating in opposite directions suggesting common energy-based motion processing. Furthermore, they observed similar temporal sensitivity function for both stimuli, which was fast (cutoff frequency near 12 Hz). On the other hand, when spatially superposing an LM and a CM grating with the same spatial and temporal frequencies, they found that the performance did not depend on the relative phase separating the two gratings and that this performance was near the one predicted by a simple probability summation model assuming that both attributes are processed separately. They concluded (Lu & Sperling, 1995, 2001) that both stimuli are initially processed by distinct energy-based spatiotemporal filters but are later integrated by summing the motion energy of both pathways (filter-rectify-filter model (Wilson, Ferrera, & Yo, 1992)). Indeed, if the CM grating was rectified and spatially summed with the LM grating prior to common energy-based motion processing, both should either sum or cancel one another depending on the relative phase separating the two gratings. As a result, a performance improvement greater than probability summation or a performance decline should be observed depending on the relative phase. In other words, they concluded that a preprocessing nonlinearity (i.e., rectification) is required to recover the motion energy of the envelope before being processed by an energy-based motion system distinct from the one processing first-order stimuli. The merging of the two pathways after the motion energy analysis could explain the inter-attribute interactions (motion cancellation) and the similar properties (temporal sensitivity function) of the two motion systems.

Gradient-based model

Benton, Johnston and colleagues (Benton, 2002; Benton & Johnston, 2001; Benton, Johnston, & McOwan, 2000; Benton, Johnston, McOwan, & Victor, 2001; Johnston & Clifford, 1995a; Johnston, McOwan, & Buxton, 1992) developed a gradient-based model which computationally demonstrates that CM drifting gratings could be processed by luminance sensitive mechanisms without any preprocessing nonlinearity extracting the spatial information of the envelope. Given that CM stimuli could be processed by a feature tracking motion system or by an energy-based motion system sensitive to LM stimuli, they concluded that it is not necessary to suppose the existence of a distinct energy-based motion system dedicated to second-order stimuli.

The keystone of their model is the calculation of the temporal derivative relative to the spatial derivative. Based on computational simulations, they showed that applying a nonlinearity after such a calculation could extract the drifting direction of a CM stimuli even though the stimulus is drift-balance (same expected motion energy for both direction at all spatiotemporal frequencies, (Chubb & Sperling, 1988)). In other words, given that a CM drifting grating may contain the same expected motion energy in the drifting and the opposite directions at all spatial frequencies, they suggest that first-order processing extracting such motion energy and based on some subsequent nonlinearity could resolve the CM drifting direction.

Since this model proposes the computation of the spatial and temporal derivative of the stimulus, the mechanism processing such stimulus should be tuned to spatial frequencies corresponding to the spectral component of the stimulus which, for CM stimuli, directly depends on the spatial frequency of the carrier and not on the spatial frequency of the envelope which requires a preprocessing nonlinearity to be recovered. Consequently, they suggest that CM stimuli could be initially processed by first-order mechanisms tuned to the energy near the spatial frequency of the carrier which is directly visible through a Fourier (i.e. first-order) analysis. Given a spatially broadband stimulus (e.g., when using broadband noise as a carrier), the mechanisms tuned to such mechanisms could be sensitive to any spatial frequencies. On the other hand, given a band-pass stimulus (e.g., when using band-pass noise as a carrier), the initial motion energy-based mechanisms processing such stimulus should be tuned to the spatial frequency of the stimulus, that is, near the carrier spatial frequency. By proposing mechanisms tuned to the spatial frequency of the carrier, the gradient-based model contrasts with models suggesting that preprocessing nonlinearities prior to the motion energy analysis (e.g., filter-rectify-filter) introduces spectral energy at the envelope spatial frequency before being processed by motion energy-based mechanisms tuned to the envelope spatial frequency.

Early nonlinearities

The filter-rectify-filter model suggests that nonlinearities prior to the energy-based motion processing (i.e. rectification) introduces energy at the spatial frequency of the envelope which can then be processed by energy-based mechanisms similar, though distinct, to the ones processing first-order stimuli. Indeed, almost any preprocessing nonlinearity applied to the stimulus introduces energy at the spatial frequency of the envelope. Nonlinearities prior to standard first-order energy-based processing, typically referred to as early nonlinearities, could enable the processing of second-order stimuli by first-order sensitive mechanisms. In other words, early nonlinearities could introduce artifacts enabling first-order mechanisms to process second-order stimuli. Such artifacts could mislead the experimenter thinking that the stimulus cannot be processed by a first-order motion system because the stimulus does not contain energy at the spatial frequency of the envelope.

Early nonlinearities can either be compressive or expansive. Given a CM stimulus having the same mean local luminance for high and low contrast areas, applying a compressive nonlinearity would result in a relatively greater mean local luminance for low contrast areas whereas an expansive nonlinearity would result in a relatively greater mean local luminance for high contrast areas.

Smith and Ledgeway (1997) considered two potential artifacts enabling LM sensitive mechanisms to process CM stimuli: global distortion products and local first-order motion patches. The former corresponds to nonlinearities (either compressive or expansive) uniformly applied across space and time (i.e. global, Figure VI-1). The later are not uniformly distributed across the stimulus, that is, certain areas are compressive and others are expansive (Figure VI-2).

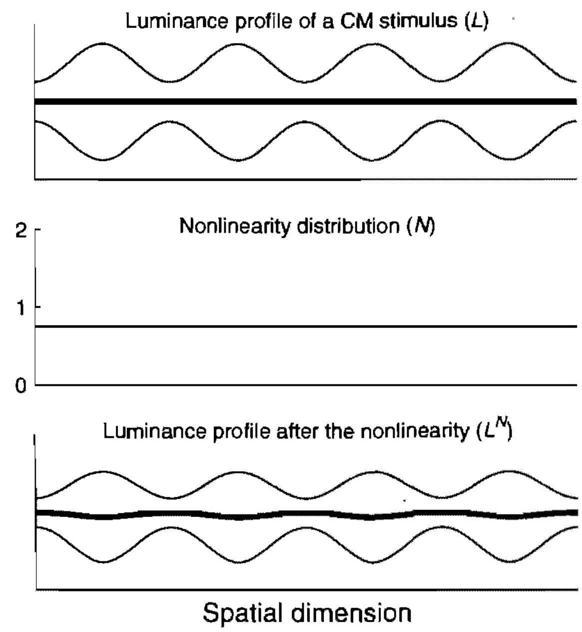


Figure VI-1. Uniform nonlinearities. The top graph shows the lower and upper limits (thin lines) of the luminance profile of a CM stimulus. The dark line shows the local mean luminance. The middle graph shows the distribution of the nonlinearity ($L'=L^N$) which is constant. The bottom graph shows the lower and upper limits (thin lines) of the luminance profile once the nonlinearity applied to the initial stimulus introducing mean luminance variations within the stimulus (dark line).

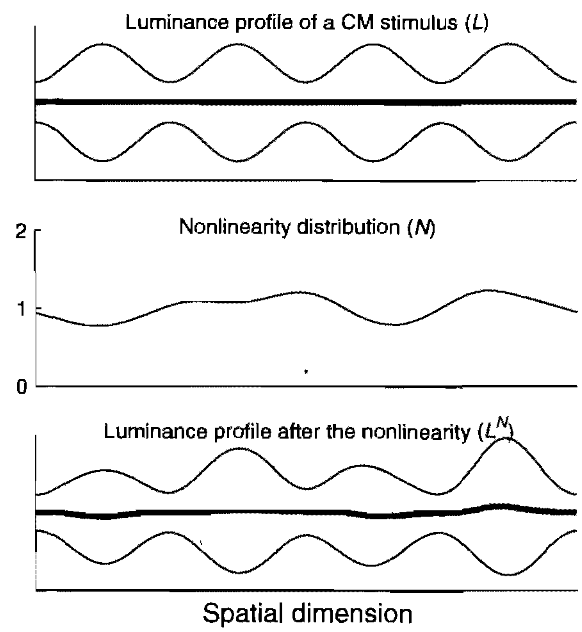


Figure VI-2. Non-uniform nonlinearities. Same as Figure VI-1 with the exception that the nonlinearity applied is not uniform, i.e. varies as a function of space.

The criterion Smith and Ledgeway used to determine whether LM and CM stimuli are processed separately or processed by common mechanisms due to early nonlinearities prior to LM sensitive mechanisms was whether they shared common or distinct properties. More specifically, they measured whether the orientation and direction discrimination thresholds were similar. For LM stimuli, both thresholds were found to be similar but for CM stimuli they were either similar or direction thresholds were about 50% higher depending on the viewing conditions. When orientation and discrimination thresholds were similar for CM stimuli, they concluded that such a stimulus was processed by LM sensitive mechanisms and when they differed they concluded that LM and CM stimuli were processed separately. Since their criterion for distinguishing whether LM and CM stimuli are processed by common or distinct mechanisms was whether they share similar properties or not, they did not directly address the question of whether there is an energy-based motion system dedicated to second-order processing sharing similar properties with the first-order motion system. Alternatively, if LM and CM processing shared similar properties, they concluded that CM stimuli were processed by LM sensitive mechanisms due to early nonlinearities.

The fact that LM and CM were processed by distinct motion systems in certain conditions does not imply the existence of a dedicated energy-based second-order pathway. CM stimuli could be processed by a feature tracking motion system. Furthermore, the fact that orientation discrimination thresholds were found to be lower than direction discrimination thresholds can be viewed as compatible with a feature tracking motion system. Indeed, in such a system, before being tracked, a feature must be extracted. As a result, the process of extracting the feature would be common to both tasks, but the discrimination of the direction would require an extra processing step consisting in tracking the position change over time. Consequently, we could expect the direction sensitivity to be greater than the sensitivity to the feature (e.g., orientation discrimination).

Thus, their study do not directly address the purpose of the present paper which is to determine if LM and CM stimuli are processed by a common or distinct energy-based motion system. Conversely, Smith and Ledgeway (1997) rather suggested that when no distinct properties are observed between the processing of LM and CM stimuli, then they share common pathways. This argument is not compatible with Lu and Sperling's (1995, 2001) conclusion suggesting distinct energy-based pathways sharing similar properties. Nonetheless, the preprocessing nonlinearities potentially enabling LM sensitive mechanisms to process CM stimuli could explain why LM and CM processing share similar properties and should therefore be considered.

Global distortion product

Smith and Ledgeway (1997) described two potential sources for the global distortion product artifacts: the display and the visual system prior to energy-based first-order processing. Obviously, artifacts originating from display nonlinearities must be avoided. Although they can never be completely eliminated, it is relatively easy to determine whether they are great enough to be visible to the first-order motion system (Scott-Samuel & Georgeson, 1999). On the other hand, it is harder to assert that nonlinearities originating from the visual system do not enable the LM sensitive motion system to process CM stimuli. Again, by finding different processing properties for LM and CM stimuli (i.e., greater direction discrimination thresholds than orientation discrimination thresholds for CM stimuli, but similar thresholds for LM stimuli), Smith and Ledgeway (1997) concluded that such artifacts were not an issue.

Scott-Samuel and Georgeson (1999) measured the global distortion product by adding a luminance signal in phase with the CM signal which compensated for the global distortion product of the visual system. They concluded that early global nonlinearities of the visual system can cause CM information to leak within the first-order processing pathway and that such nonlinearities increase with the temporal frequency and with the carrier contrast. More specifically, they found that such nonlinearities were compressive, that is, they reduced the mean luminance of high contrast areas relative to the mean luminance of low contrast areas. When compensating for such nonlinearities by adding an expansive nonlinearity to the stimulus, they found that observers still perceived second-order motion and concluded that first- and second-order motion processing can be processed separately even at the highest temporal frequency tested which was 15 Hz. Note that such a temporal frequency is somewhat faster than the temporal acuity generally observed for the feature tracking motion system which is near 8-10 Hz. Consequently, it is unlikely that CM stimuli were processed by a feature tracking motion system. Furthermore, the processing of CM stimuli could not be null by compensating for the global distortion product. They therefore concluded that CM stimuli were processed by a distinct energy-based motion system.

Local first-order motion patches

Smith and Ledgeway (1997) attributed the source of local artifacts (which they named local first-order motion patches) to local imbalances within the carrier. The most widely used carrier is static binary noise. On average, half of the noise elements are dark and the other half are light. However, for a given local area, the distribution can be greater or lower than 50%. If the distribution is greater than 50%, then increasing the local contrast raises the mean luminance and decreasing it reduces the mean luminance. If the distribution is lower than 50%, then increasing the local contrast reduces the mean luminance and decreasing it raises the mean luminance. Thus, local imbalances introduce artifacts that could enable the first-order motion system to process CM stimuli.

It is not possible to completely eliminate such artifact but Smith and Ledgeway (1997) proposed several ways of decreasing its impact. Reducing the size of the noise elements reduces the impact of this artifact since the mean distribution of light and dark noise elements for a given area will be, on average, closer to 50%. The noise sampling can also be increased by using dynamic instead of static binary noise. Furthermore, by introducing spectral energy at all spatiotemporal frequencies, dynamic noise masks luminance processing (generally increases contrast thresholds to LM stimuli) which also masks artifacts. As a result, artifacts have to be somewhat greater when using dynamic binary noise as a carrier to be an issue. Smith and Ledgeway (1997) concluded that this carrier (dynamic binary noise with small element size) was the safest one to avoid the impact of artifacts but they also proposed another carrier: high-pass static noise. If the low spatial frequencies of the noise are removed, then the mean local luminance will be relatively constant. As opposed to broadband carriers which remain drift-balance (i.e., equal expected motion energy in both directions at all spatiotemporal frequencies) once modulated by a drifting envelope, a disadvantage of using band-pass carrier is that they give rise to sidebands. Indeed, modulating a band-pass carrier with a CM drifting envelope results in a stimulus that does not have the same expected amount of energy in both directions at all spatiotemporal frequencies. Nonetheless, this artifact is usually too weak to be detected but can sometimes be an issue (Scott-Samuel & Georgeson, 1999).

Global and residual early nonlinearities

Smith and Ledgeway did not suggest that the visual system induces local nonlinearities. They defined the nonlinearities caused by the visual system as a global distortion product which can be canceled by applying the inverse nonlinearity to the stimulus. Thereby, they implicitly assumed that the nonlinearities caused by the visual system are uniform across space and time. In the present study, we question this assumption and defined early nonlinearities within the visual system as non-uniform. Such non-uniform nonlinearities can be segregated into two components (Figure VI-3): global nonlinearities, which can be compensated for by applying the inverse nonlinearity, and residual nonlinearities representing the nonlinearities left once compensating for global nonlinearities. As an example, suppose that the visual system introduces an early compressive nonlinearity which varies in strength across space. An expansive nonlinearity uniformly applied across the stimulus (i.e. of constant strength) could be applied to compensate for the mean compressive nonlinearity. Once combining the constant expansive nonlinearity introduced within the stimulus with a varying compressive nonlinearity of the visual system, the residual nonlinearities would be compressive for some areas and expansive for others. In other words, applying a uniform nonlinearity to the stimulus compensates for the global nonlinearities but not for residual nonlinearities.

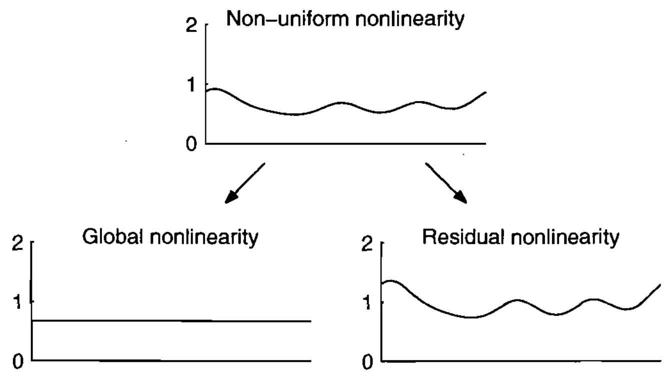


Figure VI-3. Global and residual nonlinearities. Decomposition of non-uniform nonlinearity into a global nonlinearity and residual nonlinearities. The absence of nonlinearity is represented by a value of 1.

At this point, we do not affirm that early nonlinearities are not uniform. We theoretically defined early nonlinearities as composed of global and residual nonlinearities. Empirically, either or both nonlinearities could be null or very weak. Nonetheless, we argue below that the results obtained in the present study suggest that in some conditions, early residual nonlinearities could be strong enough to enable the first-order energy-based motion system to process CM stimuli. Consequently, we argue that it is not necessary to infer the existence of a dedicated second-order energy-based motion system distinct from the luminance sensitive motion system.

Experiment 1: The frequency tuning of CM processing

The objective of the first experiment was to evaluate the frequency tuning of spatiotemporal filters processing CM stimuli by trying to dissociate CM motion processing from luminance processing at the envelope and carrier spatial frequency. If two stimuli are processed by common mechanisms then a manipulation affecting the processing of one stimulus should also affect the processing of the other. On the other hand, a manipulation affecting the processing of one stimulus more than the other suggests that they are processed separately. If CM motion stimuli are processed by luminance sensitive mechanisms, then such luminance sensitive mechanisms can be tuned either to the carrier or envelope spatial frequency. In the present experiment, we tried to dissociate CM motion processing from luminance processing at the carrier and envelope spatial frequency by evaluating the impact of band-pass noise either near the carrier or near the envelope spatial frequencies.

If a preprocessing nonlinearity introducing energy at the spatial frequency of the envelope enables the processing of second-order stimuli (e.g., rectification or early nonlinearities) and that both pathways are ultimately (or entirely) common, then we would expect luminance noise to have the same impact on second-order processing and first-order processing at the spatial frequency of the envelope of the second-order stimulus. On the other hand, if the nonlinearities enabling second-order processing occur within rather than prior to motion processing (e.g., gradient-based model), then we would expect second-order processing to be common to first-order processing at the carrier spatial frequency.

Method

Observers

One of the authors and a naïve observer participated to the study. Both had normal vision and were experienced psychophysical observers.

Apparatus

The stimuli were presented on a 19 in ViewSonic E90FB .25 CRT monitor with a mean luminance of 47 cd/m^2 and a refresh rate of 120 Hz, which was powered by a Pentium 4 computer having a Matrox Parhelia512 graphic card. The Noisy-Bit method (Allard & Faubert, 2008c) implemented with the error of the green color gun inversely correlated with the error of the two other color guns made the 8-bit display perceptually equivalent to an analog display having a continuous luminance resolution. The monitor was the only source of light in the room. A Minolta CS100 photometer interfaced with a homemade program calibrated the output intensity of each gun. At the viewing distance of 114 cm, the width and height of each pixel were 1/64 deg of visual angle.

Stimuli

As described above, the choice of the carrier is important since local motion patches can enable luminance sensitive mechanisms to process CM stimuli. To avoid such an artifact, Smith and Ledgeway (1997) proposed to use dynamic binary noise or high-pass static noise as a carrier. As mentioned above, when using dynamic binary noise, they found different properties between the processing of LM and CM stimuli and found that CM stimuli processing was low-pass suggesting that CM stimuli were processed by a feature tracking motion system. Since the present study focuses on CM processing by an energy-based motion system, we rather used a static carrier only defined by high spatial frequencies. As they noted, the disadvantage of using a carrier that is not broadband is that the resulting CM stimulus is not drift-balance since some spatiotemporal frequencies only contain energy in one direction. To avoid such an artifact, instead of multiplying an envelope with band-pass filtered noise, we applied the filtering operation after the envelope was multiplied with broadband noise. Since multiplying an envelope with broadband noise results in a drift-balance stimulus (i.e., same expected motion energy at all spatiotemporal frequencies, (Chubb & Sperling, 1988)), then band-pass filtering such a stimulus also results in a drift-balance stimulus. Such CM stimulus has the advantages of being drift-balance, minimizes the local motion patches artifact and does not introduce noise at the spatiotemporal frequency of the envelope.

Since the creation of the stimuli required a band-pass filtering operation that cannot be applied independently to each pixel, the stimulus function was defined by a three dimensional matrix (**L**, wherein each element of the matrix (L_{xyt}) represents the luminance at position (x,y) at time t) instead of defining the luminance of each pixel independently as usually formulated ($L(x,y,t)$). The stimulus was defined as a function of three matrices: a signal (**S**), a carrier (**T**) and a noise template (**N**). The carrier (**T**) was static broadband noise: all frames were identical ($T_{xyt}=T_{xy0}$ for all $t>0$) and the value at each pixel (T_{xy0}) was randomly selected from a Gaussian distribution centered on 0. Once filtered, an unmodulated carrier had a RMS contrast of 0.096%. To avoid stimulus onsets the unmodulated carrier was presented between trials. Thus, the same carrier template was used throughout a given staircase, but was randomly regenerated at the beginning of each staircase. **N** represents band-pass dynamic noise with a frequency bandwidth that depended on the testing condition: the temporal frequencies kept were the ones within 1 octave above and below the temporal frequency of the signal (2 or 8 Hz) and the spatial frequencies kept were either 0.5 octave above and below the envelope spatial frequencies (i.e., 0.35 to 0.71 cpd) or at the carrier spatial frequency (i.e., 4 to 8 cpd, see filtering operation (F(**X**)) below). In all cases, the noise templates (**N**) were scaled to have a RMS contrast of 1. **S** represents the drifting sine wave grating where each element S_{xyt} positioned at (x,y) at time t was defined as:

$$S_{xyt} = \sin(f_x x + f_t t + p) \qquad (1)$$

where f_x represents the spatial frequency (0.5 or 5.7 cpd), f_t represents the temporal frequency (±2 or ±8 Hz, the sign representing the drifting direction: left or right) and p represents the initial phase of the sine wave grating which was randomized at each trial.

All stimuli of the first experiment could be defined by the following function:

$$\mathbf{L} = L_0\left[1 + (c_{lm} + dc_{cm})\mathbf{S} + \mathrm{F}((1 + c_{cm}\mathbf{S}) \cdot \mathbf{T}) + n\mathbf{N}\right] \qquad (2)$$

where c_{lm} and c_{cm} represents the LM and CM Michelson contrasts, respectively. The early global nonlinearities of the visual system were compensated for by adding a luminance signal in phase with a contrast proportional to the contrast of the CM signal (d). n corresponds to noise contrast which varied from one condition to another. F(**X**) is the filtering function keeping all temporal frequencies of the three dimensional matrix **X** and only the spatial frequencies between 4 and 8 cpd within the Fourier domain.

Three stimuli (CM motion, luminance motion near the envelope spatial frequency and luminance motion near the carrier spatial frequency, which we referred to as CM0.5; LM0.5 and LM5.7, respectively) were tested under three noise conditions (no noise, noise near the envelope spatial frequency and noise near the carrier spatial frequency) at two temporal frequencies (2 and 8 Hz) resulting in 18 thresholds evaluated. When presenting a luminance signal (either near the carrier or envelope spatial frequency, i.e. LM5.7 or LM0.5 respectively), there was no CM signal (c_{cm}=0) and the dependant variable was the luminance contrast (c_{lm}) and vice versa, when presenting a CM signal (CM0.5), there was no luminance signal (c_{lm}=0) and the dependent variable was CM contrast (c_{cm}). Since we (Allard & Faubert, 2008b) previously found that at 2 Hz early global nonlinearities were too weak to be an issue, we assumed that early global nonlinearities were null without measuring them (d=0). The results described below suggesting separate processing confirm that early nonlinearities were not a factor at 2 Hz. At 8 Hz however, compressive nonlinearities are known to cause the second-order signal to leak within the first-order pathway (Allard & Faubert, 2008b; Scott-Samuel & Georgeson, 1999) but early global nonlinearities can be compensated for by adding a luminance signal. Consequently, at this temporal frequency, a luminance signal at the same phase and same spatial frequency as the envelope ($dc_{cm}\mathbf{S}$) was introduced with a contrast proportional to the contrast of the CM signal (d). The following section briefly describes how this proportion (d) was individually estimated and the second experiment asserts that the measurement of this proportion was accurate enough to insure that CM stimuli were not processed by a luminance sensitive motion system following early global nonlinearities.

The computation of the band-pass filtering operation within the Fourier domain ($\mathrm{F}((1+c_{cm}\mathbf{S})\mathbf{T})$) is time consuming. Since the contrast of the signal (c_{cm}) varies from one trial to another, a disadvantage of applying this filtering operation at each stimulus presentation is that it may introduce delays between trials. However, the time consuming filtering operation can be perform once the carrier (**T**) and signal (**S**) matrices are known independently of the signal contrast (c_{cm}). Indeed, due to the linearity of the filtering operation, Equation (2) can be reformulated as:

$$\mathbf{L} = L_0\left[1 + (c_{lm} + dc_{cm})\mathbf{S} + \mathrm{F}(\mathbf{T}) + c_{cm}\,\mathrm{F}(\mathbf{S}\cdot\mathbf{T}) + n\mathbf{N}\right] \quad (3)$$

Since the carrier template (**T**) does not vary between trials, F(**T**) is constant throughout each staircase. As described above, the signal matrix (**S**) depends on three variables (f_x, f_t and p). For a given staircase, the spatial frequencies of the signal (f_x) is constant, the sign of the temporal frequency (f_t) varies and the initial phase (p) is randomly selected. Since the filtering operation only affects the spatial frequencies without affecting the temporal frequencies, it can be applied independently at each frame. Since only the phase of a given signal varies between frames and between trials (due to the initial phase (p) and the temporal frequency (f_t)), various phases (uniformly distributed between 0 and 2π) of the sine wave signal multiplied with the carrier can be computed (i.e., spatial filtered) before the beginning of the staircase. Thus, by computing the spatial frequency filtering operation (F(**T**) and F(**S·T**)) before performing each staircase avoided unnecessary delays due to the time consuming filtering operation that would otherwise be required at each trial.

Procedure

Contrast threshold measurement. Each contrast threshold (c_{lm} or c_{cm}) was estimated using the geometric mean of three staircases which were performed using a 2down1up procedure (Levitt, 1971) interrupted after 26 inversions. The measure of each staircase was estimated based on the geometric mean of the last 20 inversions during which the contrast step size was fixed to 0.05 log units. The task consisted in discriminating the drifting direction (left or right) of a vertically oriented grating.

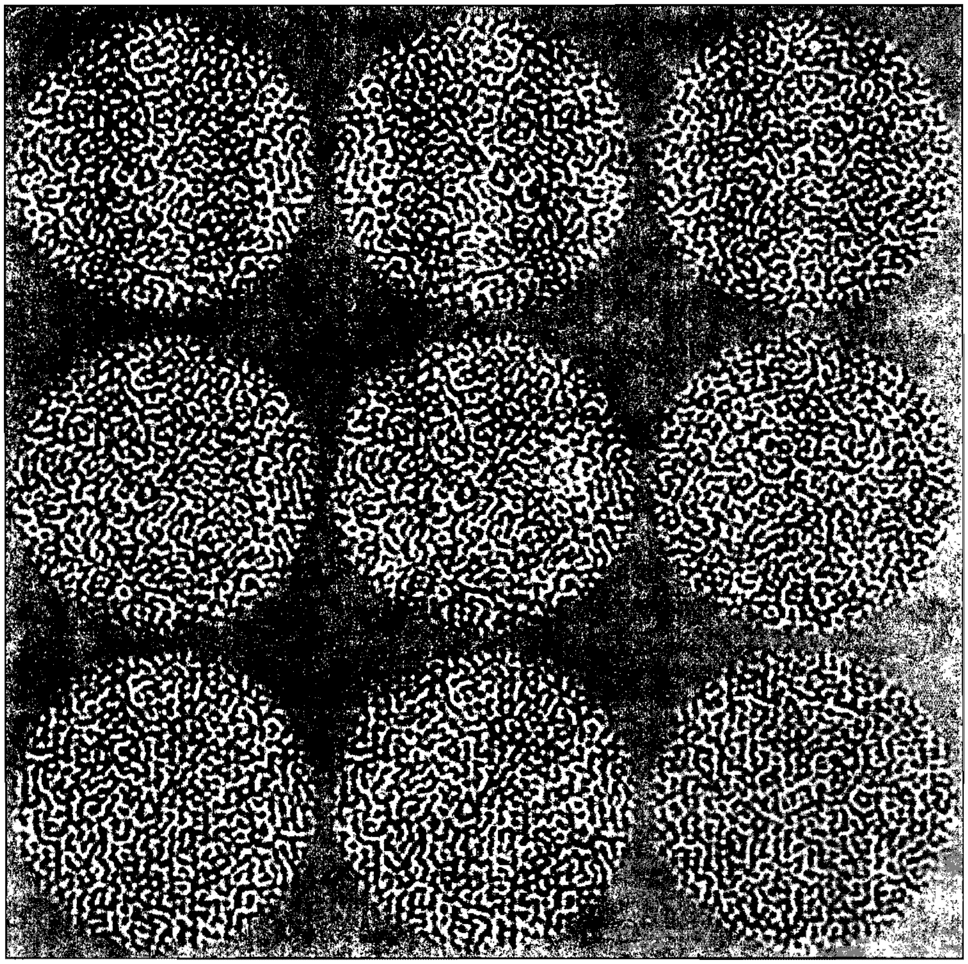
Compensating for early global nonlinearities. As mentioned above, when presenting a CM grating drifting at high temporal frequencies, early compressive nonlinearities of the visual system distort the stimulus. By assuming that such nonlinearities are uniform, they can be cancelled by adding an LM grating in phase with the CM grating (i.e., high luminance matched with high contrast) (Scott-Samuel & Georgeson, 1999). However, without this assumption, only the global early nonlinearities can be cancelled leaving early residual nonlinearities intact. Nonetheless, cancelling global nonlinearities asserts that CM stimuli are not processed by luminance sensitive mechanisms following global nonlinearities. A pilot study was performed to estimate the contrast of the luminance grating relative to the contrast of the CM grating (d) required to cancel the global early nonlinearities of the visual system. When the observer was just as efficient at processing superposed LM and CM gratings either in phase or in counter-phase, we assumed that early global nonlinearities were compensated for. For observer JM and RA, the contrast of the luminance grating added to compensate for early global nonlinearities (d) was 1.5% and 0.71%, respectively. The methodological details of the estimation of the expansive nonlinearity added to the stimulus (d) are described elsewhere (Allard & Faubert, 2008b) and the data are not shown. However, the second experiment asserts that the sensitivity to CM motion was not due to a global nonlinearity prior to luminance sensitive mechanisms.

The impact of the noise. Before evaluating contrast thresholds in noise, the contrast of the noise at the carrier and at the envelope (*n*) was fixed based on another pilot study. For each temporal frequency (2 or 8 Hz) and each observer, the contrasts of the noise near the envelope (i.e. 0.35 to 0.71 cpd) and near the carrier (i.e. 4 to 8 cpd) were adjusted to affect contrast thresholds to LM0.5 and CM0.5 stimuli, respectively, by a factor of approximately 2. We chose a factor of 2 because CM stimuli can be processed by two motion systems (energy-based and feature tracking) and noise can impair one motion system more than another (Allard & Faubert, 2008b). Thus, even though LM and CM stimuli could be processed by common mechanisms at threshold in noiseless conditions, they may be processed separately when noise elevates thresholds. To minimize the probability that distinct motion systems are processing CM stimuli in noiseless and in noisy conditions, we added an amount of noise having a small, though measurable, impact which was defined as a threshold increase of a factor of approximately 2. Noise contrasts used for both observers relative to the testing conditions are presented in Table VI-1.

	Noise near envelope		Noise near carrier	
	2 Hz	8 Hz	2 Hz	8 Hz
JM	0.021	0.0096	0.18	0.12
RA	0.010	0.0062	0.16	0.13

Table VI-1. Noise contrast. Noise RMS contrast (*n*) for each noise condition and each observer.

Once the compressive nonlinearity measured at 8 Hz and the noise contrasts were fixed, contrast thresholds to LM0.5, CM0.5 and LM5.7 drifting gratings in noiseless, noise near the envelope and noise near the carrier conditions were measured in a pseudo-random order at both 2 and 8 Hz resulting into 18 contrast thresholds (3 stimuli X 3 noise conditions X 2 temporal frequencies) each estimated 3 times.



Movie VI-1. Stimuli. Examples of the three stimuli drifting at 2 Hz in the three noise conditions. The signals are a luminance signal at 0.5 cpd (LM0.5, top row), a CM signal at 0.5 cpd (CM0.5, middle row) and a luminance signal at 5.7 cpd (LM5.7, bottom row). The signals are presented either in noiseless conditions (left column), embedded in dynamic noise either spatially filtered near the envelope (0.35 to 0.71 cpd, middle column) or carrier (4 to 8 cpd, right column). In the current examples, the noise was temporally filtered to keep only the temporal frequencies between 1 and 4 Hz since the signal was presented at 2 Hz. At 8 Hz, the temporal frequencies kept were between 4 and 16 Hz. (Movie available on the CD attached to the thesis.)

Results and discussion

Since the target of the present experiment was to measure the impact of the different type of noises on the processing of LM0.5, CM0.5 and LM5.7 stimuli, contrast thresholds to these stimuli embedded in noise were presented relative to the contrast thresholds obtained in noiseless conditions (Figure VI-4).

Distinct luminance and CM processing at 2 Hz

At 2 Hz, none of the three possible pair of stimuli (CM0.5-LM0.5, CM0.5-LM5.7 and LM0.5-LM5.7) showed a similar pattern of results suggesting that they are all processed by distinct mechanisms. It is not surprising that LM5.7 and LM0.5 are processed by distinct mechanisms since they are both luminance motion at markedly different spatial frequencies (0.5 and 5.7 cpd, respectively).

CM motion processing (CM0.5) was more affected by noise at the carrier spatial frequency (4 to 8 cpd) than at the envelope spatial frequency (0.35 to 0.71 cpd) and LM0.5 motion processing was more affected by noise near the envelope spatial frequency than near the carrier spatial frequency. Again, the fact that the processing of these stimuli in noise was unevenly affected suggests that they are processed, at least at some point, separately. Thereby, these results assert that CM0.5 stimuli were not processed by a luminance sensitive motion system due to early nonlinearities which would have predicted similar pattern of results for LM0.5 and CM0.5 stimuli. Furthermore, these results do not imply that CM0.5 stimuli are processed by an energy-based motion system distinct from the one processing LM0.5 stimuli since CM0.5 stimuli could be processed by a feature tracking motion system. Such interpretation is in agreement with previous findings suggesting that CM motion at low temporal frequencies may be processed by a feature tracking motion system (Seiffert & Cavanagh, 1999).

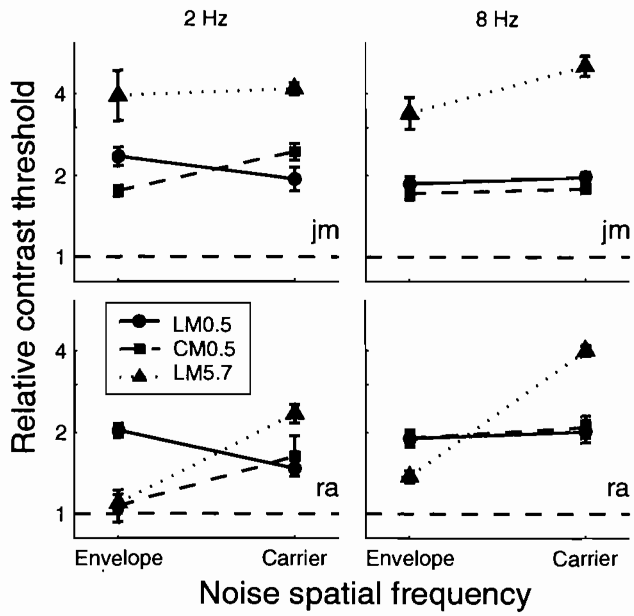


Figure VI-4. Results of experiment 1. Contrast thresholds in different noise conditions relative to contrast threshold in noiseless conditions for both observers (JM and RA) at 2 Hz (left) and 8 Hz (right). Error bars show the standard error.

Common LM and CM processing at 8 Hz

A different pattern of results was observed at 8 Hz. The processing to LM0.5 and CM0.5 were evenly affected by both noises. Conversely, the processing to LM5.7 and CM0.5 were not affected in similar proportions for both noises. The dissociation between CM and luminance processing at the carrier spatial frequency (LM5.7 and CM0.5) and the failure to dissociate CM processing and luminance processing at the envelope spatial frequency (LM0.5 and CM0.5) suggests that the processing pathways of LM0.5 and CM0.5 stimuli are, at least ultimately, common. Indeed, if both stimuli are processed by common mechanisms then we would expect both not to be dissociable: any manipulation affecting the processing of one stimulus would also affect, in a similar proportion, the processing of the other stimulus.

These results support models suggesting that CM stimuli and luminance stimuli at the envelope spatial frequency share, at least at some point, common motion pathways. The two pathways could be entirely common due to early nonlinearities prior to LM sensitive processing or distinct with their energy summed as suggested by the filter-rectify-filter model. In either case, a preprocessing nonlinearity would introduce energy at the envelope spatial frequency causing CM stimuli to be processed by energy-based mechanisms tuned to the spatial frequency of the envelope. On the other hand, by suggesting that CM spatiotemporal filters are tuned to the same spatial frequency as luminance processing at the envelope spatial frequency, these results do not support the gradient-based model suggesting that the energy-based mechanisms processing CM stimuli would be tuned to the spatial frequency of the carrier.

Since early global nonlinearities were compensated for by adding a luminance signal, CM stimuli were not detected by a luminance sensitive mechanism following a global nonlinearity. Although this hypothesis cannot be excluded at the current point (measurement error in the evaluation of the global nonlinearity (*d*) could be responsible for global leaking from the second-order pathway to the first-order pathway), the next experiment asserted that the current results cannot be due to early global nonlinearities. Thus, if nonlinearities prior to luminance sensitive mechanisms enabled the processing of CM stimuli, then such nonlinearities were not uniform, i.e., they cannot be cancelled by adding an LM grating.

As mentioned above, Smith and Ledgeway (1997) proposed that non-uniform nonlinearities (which they referred to as local first-order motion patches) would be caused by using static noise defined by large noise elements. However, this is unlikely the case in the present experiment since low spatial frequencies within the stimulus were removed. Furthermore, Benton and Johnston (1997) found that the impact of this artifact was insignificant, i.e. local imbalance within the carrier distribution increased the expected local motion energy equally in both directions which implies that the stimulus remained microbalanced. We therefore conclude that if the CM stimuli were processed by luminance sensitive motion energy-based mechanisms due to an early nonlinearity, then early nonlinearities of the visual system were not uniform.

All-in-all, two possible models could explain why CM processing (CM0.5) and luminance processing at the envelope spatial frequency (LM0.5) were not dissociable at 8 Hz: residual early nonlinearities of the visual system causing luminance sensitive mechanisms to process CM stimuli or a dedicated energy-based mechanism after a rectification process with converging first- and second-order pathways.

Experiment 2: Simulating residual nonlinearities

The results of the previous experiment suggest that CM motion at 8 Hz was processed by energy-based mechanisms tuned to the envelope spatial frequency following a preprocessing nonlinearity. These mechanisms can either be distinct from first-order energy-based mechanisms with a nonlinearity explicitly applied to second-order stimuli (e.g., filter-rectify-filter) or CM stimuli could be processed by first-order sensitive mechanisms following an early non-uniform nonlinearity as suggested in the present paper.

Based on Lu and Sperling's (2001) review, the main argument suggesting the existence of dedicated energy-based second-order mechanisms is that the performance to superposed LM and CM signals is near the one predicted by a probability summation model assuming distinct processing, and this, independently of the phase separating the two signals. Such apparent lack of interaction was interpreted as strong evidence of distinct processing. Indeed, if an early uniform (i.e. global) nonlinearity enables the processing of CM stimuli by first-order sensitive mechanisms, then the superposition of both stimuli should either cancel or completely sum when presented either in phase or in counter-phase. However, if the nonlinearities are not uniform, then they cannot be completely cancelled by adding a uniform luminance grating. Optimally, the luminance grating would cancel the global nonlinearity leaving residual nonlinearities. Indeed, if the early global nonlinearity of a CM grating is compensated for by adding a luminance grating, then only residual nonlinearities would leak within the luminance pathway. Locally, these residual nonlinearities would be of the same spatiotemporal frequency as the CM grating. Spatially summing these residual nonlinearities (originating from the presented CM stimulus) with a luminance grating of the same spatiotemporal frequency would sum under certain areas and subtract under others resulting in no strong global summation or cancellation. The objective of the present experiment was to empirically demonstrate that the apparent lack of interaction (performance predicted by probability summation) do not necessarily imply distinct processing. If two stimuli subtract one another under half the areas and sum under the other half, then one could expect no important performance summation which may not be greater than the probability summation (i.e., apparently no interaction) even though both stimuli strongly interact.

To simulate residual nonlinearities, we multiplied spatiotemporally low-pass noise (mean of 0) with a luminance grating (Movie VI-2). Negative spatiotemporal blobs inverted the phase of the grating. Thus, half the spatiotemporal areas contained a grating with the same phase as the original grating and the phase is inverted for the other half. Globally, the expected energy of such a stimulus at the spatiotemporal frequency of the grating is null since, on average, the areas cancel one another. Locally, however, each area is defined by a luminance grating. Consequently, there is no reason to suggest that simulated residual nonlinearities and a luminance grating at the same spatiotemporal frequency would be processed by distinct mechanisms. Again, spatially summing these simulated residual nonlinearities with the original grating would not cancel one another. The spatial areas in counter-phase would subtract one another decreasing the visibility of the resulting stimulus, but the spatial areas in phase would sum one another increasing the visibility of the resulting stimulus. The purpose of the present experiment was to demonstrate that two stimuli processed by common mechanisms can result in no apparent interaction. Thus, the apparent lack of interaction between luminance and CM processing may not imply distinct processing. CM stimuli could be processed by luminance sensitive mechanisms following early non-uniform nonlinearities.

Movie VI-2. Simulation of early residual nonlinearities. The resulting signal (right) is the multiplication of a grating (left) with low-pass noise (center). (Movie available on the CD attached to the thesis.)

Method

The same two observers participated to the study and the same apparatus was used.

Stimuli

The stimulus function was similar to the one in the previous experiment with the exception of the last term:

$$\mathbf{L} = L_0\left[1 + \left(c_{lm} + dc_{cm}\right)\mathbf{S} + \mathrm{F}\left(\left(1 + c_{cm}\mathbf{S}\right)\cdot \mathbf{T}\right) + c_{nl}\mathbf{N}\cdot \mathbf{S}\right] (4)$$

That is, no noise was added to the stimulus, but a simulation of residual nonlinearities was added. c_{nl} represents the contrast of the simulated residual nonlinearities and **N** corresponds to spatiotemporally low-pass noise (<0.5 cpd and <8 Hz) centered on zero with a RMS of 1. All other variables were defined as in the previous experiment. Note that we only simulated the impact of early residual nonlinearities at low spatiotemporal frequencies using an ideal filter. Obviously, early residual nonlinearities would certainly also introduce energy at higher spatiotemporal frequencies. Nonetheless, the objective was not to accurately simulate the early nonlinearities of the visual system using a biologically plausible model but rather to show that early residual nonlinearities introducing energy at low spatiotemporal frequencies are sufficient to explain the apparent lack of interaction between LM and CM processing.

Procedure

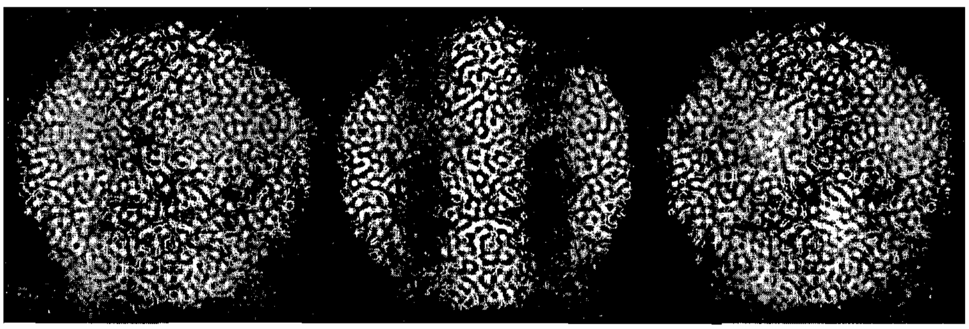
Threshold measurement. Spatially summing two stimuli with equal visibility (same proportion of correct answers) results in a complete summation if the same performance is obtain to either stimulus presented alone with its contrast doubled (assuming that the slopes of the two psychometric functions are identical). For instance, consider the case when the two stimuli are identical (e.g., two luminance gratings). Obviously, there is complete summation between these stimuli since both are processed by the same mechanisms. Spatially superposing both stimuli is equivalent to doubling the contrast. As a result, an observer will certainly have the same performance to either stimulus presented alone with the contrast doubled or with the sum of the two stimuli, since the resulting stimuli are identical.

The first step consisted in measuring the contrast threshold of each grating presented alone (c_{lm}, c_{cm} and c_{nl}). The threshold was defined as half the contrast required to have 87% correct answers (based on a Weibull psychometric function with a typical slope of 3.5, the performance is estimated at 56%). The threshold was evaluated as in the previous experiment with the exception that a 4down1up was used rather than a 2down1up (threshold criterion of 87% instead of 71%) and the contrast measured was then divided by 2. Consequently, complete summation would result into a performance near 87%. We defined the threshold criterion as such to avoid ceiling effects. Based on this definition, all thresholds were significantly lower than perfect performance and above chance level.

Spatial summation. Once the contrast to each grating was fixed, the proportion of correct answers for each grating presented alone at threshold was evaluated (stimuli referred to as LM, CM and NL). Again, based on a typical slope of 3.5, the expected performance at threshold is 56%. The expected performance for complete summation was evaluated by measuring the proportion of correct answers to each grating presented alone at twice the contrast threshold (stimuli referred to as 2LM, 2CM, 2NL). Obviously, the expected performance at twice the threshold was 87%. The interaction between the processing of LM and CM stimuli was evaluated by measuring the proportion of correct answers to the combination of these stimuli either in phase (LM+CM) or in counter-phase (LM-CM). To test the residual nonlinearities model, the performance to the summation of a luminance grating with simulated residual nonlinearities (LM+NL) was also evaluated. For the three interaction cases, if both stimuli completely sum, a performance of 87% would be expected. If both stimuli cancel one another, this should result into a performance near chance level (50%). The proportion of correct answers to each of these 9 stimuli was evaluated over 200 trials presented in a pseudo-random order resulting into 1800 trials which were presented in 4 blocks of 450 trials.

Probability summation. Based on a simple probability summation model supposing that both signals are processed independently (Lu & Sperling, 1995), the probability of not "truly" perceiving a stimulus composed of both signals($1-P^t_{12}$) is equal to the probability of not "truly" perceiving both signals presented separately (($1-P^t_1$) and ($1-P^t_2$)):

$$1-P^t_{12} = (1-P^t_1)(1-P^t_2) \qquad (5)$$



Movie VI-3. Stimuli. Examples of the three stimuli used in the second experiment presented separately: a luminance signal (left), a CM signal (center) and simulated residual nonlinearities (right). (Movie available on the CD attached to the thesis.)

The probability of not having a correct answer ($1-P^c_x$) is equal to the probability of not "truly" perceiving the stimulus ($1-P^t_x$) and not having a correct answer by chance ($1-P_{chance}$, i.e. 50%):

$$1-P^c_x = (1-P^t_x)(1-P_{chance}) \qquad (6)$$

As a result, based on this probability summation model, the probability of having a correct answer when presenting both signals (P^c_{12}) may be defined as a function of the probability of having a correct answer to either signal (P^c_1 and P^c_2) by the following function:

$$P^c_{12} = 1 - \frac{(1-P^c_1)(1-P^c_2)}{(1-P_{chance})} \qquad (7)$$

Supposing a proportion of correct answers of 56% to either stimulus presented alone, the probability summation model would predict a performance of 61% when both signals are presented.

Results

Figure VI-5 summarizes the results. As expected, thresholds for each grating presented alone at twice the contrast thresholds (2LM, 2CM and 2NL) were near 87% and clearly above the performance predicted by the probability summation model as expected due to the sharp slope of the psychometric function. Thresholds for each grating presented alone at threshold (LM, CM and NL) were near 56% as predicted by assuming a slope of 3.5 on the Weibull psychometric function. The most important results were the ones obtain when two different stimuli presented at threshold were superposed (LM+CM, LM-CM and LM+NL). For each of these three stimuli, the proportion of correct answers was near or above the one predicted by the probability summation model.

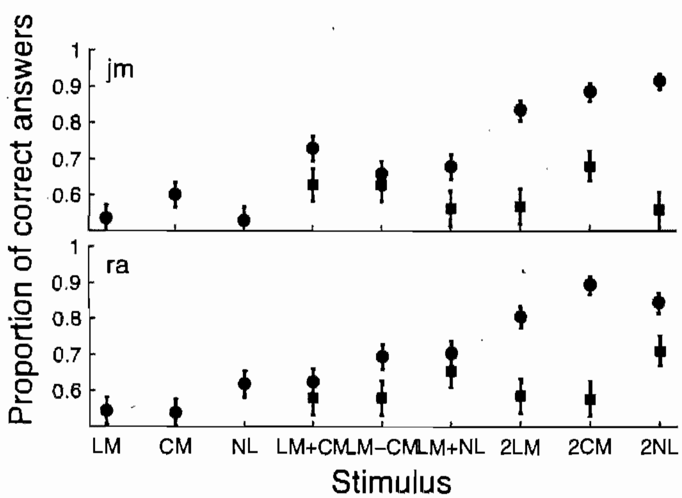


Figure VI-5. Results of experiment 2. Circles represent the proportion of correct answers to the 9 different stimuli. Squares represent the proportion of correct answers predicted by a simple probability summation model (Equation (7)). Error bars shows 68% confidence interval.

Discussion

Global preprocessing nonlinearity

In both experiments, early global nonlinearities were compensated for by adding a luminance grating in phase with the CM grating. Although it is impossible to perfectly compensate for early global nonlinearities, the fact that similar performances were obtain when superposing an LM and a CM stimulus either in phase or in counter-phase asserts that CM stimuli were not processed by luminance sensitive mechanisms following a global preprocessing nonlinearity (rectification or global early nonlinearities). Indeed, if LM and CM stimuli were processed by common energy-based mechanisms due to a global preprocessing nonlinearity either explicitly applied to second-order processing (rectification) or due to early uniform nonlinearities (i.e., global distortion product), then the expected performance should either be complete summation or cancellation depending on whether both stimuli were superposed in phase or in counter-phase. Consequently, these results confirm that the estimated mean global nonlinearity (*d*) was precise enough to assert that the CM stimuli at 8 Hz were not processed by luminance sensitive mechanisms following a global preprocessing nonlinearity.

Common processing without apparent interaction

After concluding that CM stimuli were not processed by luminance sensitive mechanisms due to a global distortion product (i.e., early global nonlinearity), several authors (Lu & Sperling, 1995, 2001; Scott-Samuel & Georgeson, 1999) inferred that LM and CM stimuli were processed by distinct motion systems (e.g., filter-rectify-filter model). Indeed, the performance gain when spatially superposing both stimuli can be explained by a simple probability summation model suggesting that LM and CM stimuli are processed separately. However, by making such deduction, they implicitly assumed that CM stimuli could not be processed by luminance sensitive mechanisms due to early residual nonlinearities. Though CM processing could not be processed by LM sensitive mechanisms following an early global nonlinearity, the present experiment shows that these results could be explained by residual nonlinearities. Indeed, by simulating early residual nonlinearities, we obtained similar results as the ones for CM stimuli. That is, when superposing a NL or CM stimuli with an LM stimuli, the performance obtained was near the one predicted by a probability summation model. Furthermore, the results showed that given two stimuli processed by common motion energy-based mechanisms (LM and NL), no cancellation and no complete summation are necessarily obtained. Consequently, if residual nonlinearities prior to LM sensitive mechanisms enable the processing of CM stimuli, then, based on our simulations, no apparent interactions could be observed even though there is strong interaction (half of the spatiotemporal areas sums and the other half cancels).

General discussion

Failure to dissociate first- and second-order energy-based processing

In a previous study (Allard & Faubert, 2007), we successfully dissociated the detection of static LM and CM stimuli suggesting distinct processing, which is in agreement with the general consensus for static stimuli (Georgeson & Schofield, 2002; Schofield & Georgeson, 1999). By applying a similar noise masking paradigm, we (Allard & Faubert, 2008b) tried to dissociate LM and CM motion processing by evaluating the impact of adding dynamic LM and CM noises on either processing. At low temporal frequencies, a double dissociation was observed: LM noise impaired more LM than CM processing and CM noise impaired more CM than LM processing suggesting that both stimuli are processed by distinct mechanisms. Such dissociation was not observed at high temporal frequencies. That is, either noise had the same impact on both LM and CM processing suggesting that both are processed by common mechanisms. In the current study, another attempt was made at dissociating LM and CM processing by adding luminance noise either at the envelope or carrier spatial frequency. Again, at low temporal frequencies, both processing were dissociable but not at high temporal frequencies.

The ability to dissociate LM and CM processing for static and low temporal frequency stimuli is compatible with the hypothesis that CM stimuli were processed by a feature tracking motion system at low temporal frequencies. Indeed, if LM and CM features are spatially distinguishable (as suggested by the absence of inter-attribute interaction for static stimuli (Allard & Faubert, 2007; Georgeson & Schofield, 2002; Schofield & Georgeson, 1999)), then we could expect a feature tracking motion system that is processing the CM feature to be unaffected by the presence of the LM feature.

At high temporal frequencies, the failure to dissociate LM and CM processing suggests, at least at some point, common processing pathways. Furthermore, Lu and Sperling (1995) found motion cancellation rather than motion transparency when simultaneously presenting an LM and CM grating moving in opposite directions also suggesting that both pathways are, at least ultimately, common. Both pathways could either be entirely common (i.e. early nonlinearities) or initially distinct energy-based mechanism with common late mechanisms summing the output energy of both pathways (e.g., filter-rectify-filter model).

Preprocessing nonlinearity

The development of the gradient-based model suggests that a preprocessing nonlinearity introducing spectral energy at the envelope spatial frequency is not required to process CM motion stimuli. Alternatively, the gradient-based model shows that luminance-based mechanisms tuned to the spectral energy of the stimulus (i.e. the carrier) could enable the processing of CM motion stimuli even though the stimulus is drift-balance. However, the present study found that luminance and CM processing at high temporal frequencies was not dissociable when the spatial frequency of the luminance grating was equal to the CM envelope rather than carrier spatial frequency. This suggests that the mechanisms processing CM stimuli in such conditions are tuned to the spatial frequency of the envelope rather than the carrier. Consequently, CM processing would occur due to a preprocessing nonlinearity introducing spectral energy at the envelope spatial frequency rather than nonlinearities within the motion processing mechanism as suggested by the gradient-based model.

Implicit assumption of the distinct processing hypothesis

Based on the fact that CM motion processing cannot be null by the superposition of luminance grating led several authors (Lu & Sperling, 1995, 2001; Scott-Samuel & Georgeson, 1999) to conclude that early nonlinearities cannot explain by themselves the perception of CM motion and they thereby deduced that LM and CM stimuli are processed by distinct energy-based motion systems. The formulation of this argument only considers early global nonlinearities and implicitly assumes that the preprocessing residual nonlinearities are too weak to cause LM sensitive mechanisms to process CM stimuli. If the preprocessing nonlinearities are not-uniform then some areas could cancel one another and others sum. As shown in the second experiment, such complete interaction could result in no apparent interaction (i.e. a performance similar to the one predicted by a probability summation model).

Early global nonlinearities were measured in various conditions (Allard & Faubert, 2008b; Scott-Samuel & Georgeson, 1999) and were found to vary as a function of the temporal frequency of the stimulus, stimulus duration and carrier contrast. Indeed, at low temporal frequencies, the mean nonlinearities were not found to have a significant impact, but important global nonlinearities were observed at high temporal frequencies. Furthermore, the nonlinearities were also found to vary from one subject to another. In the present study, the nonlinearity measured varied by a factor of about 2 between our two observers. Given that early global nonlinearities vary as a function of many parameters, it is reasonable to assume that nonlinearities could also vary as a function of space or time and that such variation could cause sufficiently great distortions to enable LM sensitive mechanisms to process CM stimuli. On the other hand, concluding that LM and CM stimuli are processed by distinct energy-based motion systems, imply assuming that the variation of the early nonlinearities as a function of space and time are too weak to enable LM sensitive mechanisms to process CM stimuli. Again, since the early nonlinearities were typically defined as global distortion products it was also implicitly assumed that they were uniform, i.e. the residual nonlinearities were too weak to be processed by LM sensitive mechanisms. Such an assumption is necessary to conclude that both stimuli are processed by initially distinct pathways.

Note that almost any variation within early nonlinearities could enable a luminance sensitive motion system to detect a CM stimulus. Here we express such variation either as a function of space or time, but this does not necessarily have to be the case. Variations within a given spatiotemporal area are also possible. For instance, if the two eyes have distinct early global nonlinearities, then it would not be possible to compensate for such nonlinearities simply by adding a nonlinearity to a stimulus viewed binocularly. In any case, there would always be at least one eye which the global nonlinearity would not be properly compensated for. More generally, if two neurons process the same receptive field but do not have the same early nonlinearities (e.g., if nonlinearities occur where the two pathways to these neurons are distinct), then these nonlinearities cannot be compensated for.

Visualizing the non-uniform nonlinearities of the visual system

The present study suggests that the apparent lack of interaction when superposing both LM and CM stimuli is not sufficient to infer the existence of a dedicated energy-based second-order motion system. Alternatively, we suggest that non-uniform early nonlinearities could enable luminance sensitive mechanisms to process CM stimuli. However, we do not directly demonstrate such non-uniform nonlinearities exist.

After viewing a primarily presentation of the current results (Allard & Faubert, 2008a), Harry Orbach (personal communication at the Vision Science Society meeting in 2008) mentioned that when viewing a flickering Ganzfeld at a high contrast and a high temporal frequency, he observed that the field appeared spatially non-uniform even though it was. After this conversation, we created a large flickering stimulus and we observed the same phenomenon. More specifically, the greater the contrast and the higher the temporal frequency, the less uniform the flickering field appeared. This subjective experience directly demonstrates the existence of non-uniform nonlinearities during dynamic presentations and thereby supports the hypothesis that non-uniform nonlinearities could enable the luminance sensitive motion system to process CM stimuli in certain conditions.

Conclusion

Two models can explain the similar properties observed between LM and CM energy-based processing, the failure to dissociate both processing and the partial summation (performance near probability summation) observed when spatially summing both stimuli either in phase or in counter-phase. The filter-rectify-filter model suggests the existence of a rectification process recovering the spectral energy of the envelope with dedicated energy-based second-order mechanisms. The first- and second-order pathways would later sum explaining the similar processing properties and why the processing of both types of stimuli would not be dissociable. As described above, this model implicitly assumes that early residual nonlinearities are too weak to be processed by first-order sensitive mechanisms. On the other hand, the model proposed in the present study suggests that early residual nonlinearities may be large enough in certain conditions to cause CM processing pathway to leak within the LM processing pathway.

To our knowledge, CM stimuli are the only second-order stimuli found to be processed by energy-based mechanisms and this generally occurs at high temporal frequencies with a high contrast carrier. Most other second-order stimuli (e.g. stereo-based) were found to be processed by a feature tracking motion system. Note that as residual nonlinearities (as described above when viewing a flickering field), early global nonlinearities were also found to increase with carrier contrast (Scott-Samuel & Georgeson, 1999) and temporal frequency (Allard & Faubert, 2008b; Scott-Samuel & Georgeson, 1999). Consequently, early nonlinearities are correlated with the probability of CM stimuli to be processed by energy-based mechanisms. Given that we are highly sensitive to LM stimuli at high temporal frequencies, relatively small preprocessing nonlinearities are required to enable the luminance energy-based motion system to process CM stimuli. Thus, the existence of a dedicated second-order energy-based motion system is questionable. We conclude that the model suggesting that residual nonlinearities may introduce sufficient artifact to be processed by luminance sensitive mechanisms is more parsimonious than supposing the existence of a dedicated second-order energy-based motion system and supposing residual nonlinearities are too weak to cause considerable luminance artifacts.

Acknowledgments: This research was supported by an NSERC operating grant to JF and the NSERC-Essilor Research Chair.

References

Allard, R., & Faubert, J. (2007). Double dissociation between first- and second-order processing. *Vision Research, 47*(9), 1129-1141.

Allard, R., & Faubert, J. (2008a). Common first- and second-order motion processing at high temporal frequencies. *Journal of Vision, 8*(6), 18-18.

Allard, R., & Faubert, J. (2008b). First- and second-order motion mechanisms are distinct at low but common at high temporal frequencies. *Journal of Vision, 8*(2), 1-17.

Allard, R., & Faubert, J. (2008c). The noisy-bit method for digital displays: Converting a 256 luminance resolution into a continuous resolution. *Behavior Research Methods, 40*(3), 735-743.

Benton, C. P. (2002). Gradient-based analysis of non-Fourier motion. *Vision Research, 42*(26), 2869-2877.

Benton, C. P., & Johnston, A. (1997). First-order motion from contrast modulated noise? *Vision Research, 37*(22), 3073-3078.

Benton, C. P., & Johnston, A. (2001). A new approach to analysing texture-defined motion. *Proc Biol Sci, 268*(1484), 2435-2443.

Benton, C. P., Johnston, A., & McOwan, P. W. (2000). Computational modelling of interleaved first- and second-order motion sequences and translating 3f+4f beat patterns. *Vision Research, 40*(9), 1135-1142.

Benton, C. P., Johnston, A., McOwan, P. W., & Victor, J. D. (2001). Computational modeling of non-Fourier motion: further evidence for a single luminance-based mechanism. *Journal of the Optical Society of America, A, Optics, Image Science, & Vision, 18*(9), 2204-2208.

Cavanagh, P., & Mather, G. (1989). Motion: the long and short of it. *Spat Vis, 4*(2-3), 103-129.

Chubb, C., & Sperling, G. (1988). Drift-balanced random stimuli: a general basis for studying non-Fourier motion perception. *J Opt Soc Am A, 5*(11), 1986-2007.

Georgeson, M. A., & Schofield, A. J. (2002). Shading and texture: separate information channels with a common adaptation mechanism? *Spat Vis, 16*(1), 59-76.

Johnston, A., & Clifford, C. W. (1995). Perceived motion of contrast-modulated gratings: predictions of the multi-channel gradient model and the role of full-wave rectification. *Vision Res, 35*(12), 1771-1783.

Johnston, A., McOwan, P. W., & Buxton, H. (1992). A computational model of the analysis of some first-order and second-order motion patterns by simple and complex cells. *Proc Biol Sci, 250*(1329), 297-306.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America, 49*(2), Suppl 2:467+.

Lu, Z. L., & Sperling, G. (1995). The functional architecture of human visual motion perception. *Vision Research, 35*(19), 2697-2722.

Lu, Z. L., & Sperling, G. (2001). Three-systems theory of human visual motion perception: review and update.[see comment][erratum appears in J Opt Soc Am A Opt Image Sci Vis 2002 Feb;19(2):413]. *Journal of the Optical Society of America, A, Optics, Image Science, & Vision, 18*(9), 2331-2370.

Schofield, A. J., & Georgeson, M. A. (1999). Sensitivity to modulations of luminance and contrast in visual white noise: separate mechanisms with similar behaviour. *Vision Res, 39*(16), 2697-2716.

Scott-Samuel, N. E., & Georgeson, M. A. (1999). Does early non-linearity account for second-order motion? *Vision Research, 39*(17), 2853-2865.

Seiffert, A. E., & Cavanagh, P. (1998). Position displacement, not velocity, is the cue to motion detection of second-order stimuli. *Vision Research, 38*(22), 3569-3582.

Seiffert, A. E., & Cavanagh, P. (1999). Position-based motion perception for color and texture stimuli: effects of contrast and speed. *Vision Research, 39*(25), 4172-4185.

Smith, A. T. (1994). Correspondence-based and energy-based detection of second-order motion in human vision. *Journal of the Optical Society of America, A, Optics, Image Science, & Vision, 11*(7), 1940-1948.

Smith, A. T., & Ledgeway, T. (1997). Separate detection of moving luminance and contrast modulations: fact or artifact? *Vision Research, 37*(1), 45-62.

Ukkonen, O. I., & Derrington, A. M. (2000). Motion of contrast-modulated gratings is analysed by different mechanisms at low and at high contrasts. *Vision Research, 40*(24), 3359-3371.

Wilson, H. R., Ferrera, V. P., & Yo, C. (1992). A psychophysically motivated model for two-dimensional motion perception. *Vis Neurosci, 9*(1), 79-97.

Zaidi, Q., & DeBonet, J. S. (2000). Motion energy versus position tracking: spatial, temporal, and chromatic parameters. *Vision Research, 40*(26), 3613-3635.

Chapitre VII

Conclusion

Stimuli statiques définis par la luminance et le contraste

Pour un signal statique, le Chapitre III à démontrer que le traitement d'un signal défini par la luminance est plus affecté par le bruit défini par la luminance que par le bruit défini par le contraste et vice versa, que le traitement d'un signal défini par le contraste est plus affecté par le bruit défini par le contraste que par le bruit défini par la luminance. Le fait qu'il soit possible de sélectivement affecter le traitement d'un attribut tout en maintenant le traitement de l'autre relativement intact (c.-à-d., pas ou peu d'interaction inter-attribut) implique que ces attributs sont traités, au moins initialement, par des mécanismes distincts. Ces résultats sont compatibles avec le peu d'interaction inter-attribut généralement observée pour ces attributs dans des conditions statiques (Georgeson & Schofield, 2002; Schofield & Georgeson, 1999).

Cependant, nous avons également observé dans le Chapitre II que l'être humain est aussi efficace à détecter un signal de luminance dans du bruit de luminance qu'à détecter un signal de contraste dans du bruit de contraste. Cette même efficacité de calcul peut s'expliquer par l'existence d'un mécanisme commun extrayant le signal du bruit pour ces deux attributs. De plus, la même efficacité à extraire un signal du bruit a également été observée pour un signal défini par la luminance dans du bruit défini par la luminance comparativement à un signal défini par la couleur dans du bruit défini par la couleur (Gegenfurtner & Kiper, 1992). Ces résultats peuvent s'expliquer par l'existence d'un mécanisme général, c'est-à-dire commun à plusieurs attributs, extrayant le signal du bruit. Il n'est donc pas nécessaire de supposer l'existence de mécanismes distincts extrayant le signal du bruit pour chaque attribut. Étant donné l'absence ou le peu d'interaction inter-attribut, un mécanisme général pourrait sélectionner et traiter un seul attribut en ignorant les autres (Figure VII-1).

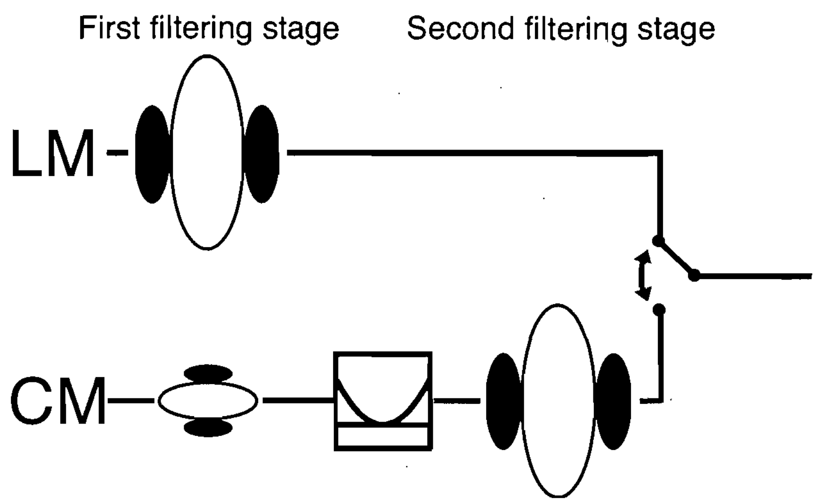


Figure VII-1. Modèle proposant un mécanisme traitant plusieurs attributs de façon sélective.

Il est généralement présupposé que la détectabilité d'un signal dépend du ratio signal-bruit. Le bruit peut provenir du système visuel (bruit interne) ou du stimulus (bruit externe). Lorsque la différence entre ces deux bruits est importante, le bruit le plus faible n'a pas significativement d'impact et la perception du signal dépend seulement du bruit le plus important. En d'autres termes, il est généralement présupposé que l'ajout de bruit externe modifie le traitement seulement quantitativement en altérant le ratio signal-bruit, mais ne modifie pas le traitement qualitativement, c'est-à-dire que le signal est traité de la même façon et par les mêmes mécanismes peu importe la quantité de bruit externe. Selon cette présupposition, le mécanisme général sélectionnant un attribut et extrayant le signal du bruit serait également sollicité lors de la détection d'un signal en absence de bruit tel que suggéré dans le Chapitre III. En d'autres termes, la détection de luminance et de contraste serait initialement traitée de façon distincte, mais impliquerait un traitement général commun (Figure VII-1). Cependant, si l'ajout de bruit modifie qualitativement le traitement d'un stimulus (tel que suggéré dans le Chapitre IV), le modèle suggérant un mécanisme général sélectionnant un attribut et extrayant le signal du bruit de cet attribut pourrait n'être valide qu'en présence de bruit externe. Une recension des écrits a permis de supporter cette dernière alternative en démontrant qu'il est possible de dissocier le traitement d'un stimulus en présence et en absence de bruit externe, c'est-à-dire qu'il existe des conditions affectant seulement le traitement en absence de bruit externe et d'autres conditions affectant seulement le traitement en présence de bruit externe. Cette double dissociation suggère donc que l'ajout de bruit peut modifier qualitativement le traitement d'un stimulus (Chapitre IV). La détection de différents attributs en absence de bruit n'impliquerait donc pas nécessairement des mécanismes communs.

Mouvement défini par la luminance et le contraste

Étant donné que le mouvement défini par le contraste peut être traité par aux moins deux systèmes de mouvement basés sur des stratégies computationnelles différentes (l'un basé sur le suivi d'une caractéristique (*feature tracking*) et l'autre basé sur l'énergie (Lu & Sperling, 1995, 2001; Seiffert & Cavanagh, 1999; Smith, 1994)), la problématique de déterminer si le mouvement défini par la luminance et le contraste sont traités par des mécanismes communs ou distincts peut être posée pour chacune de ces stratégies computationnelles.

Étant donné qu'un stimulus défini par la luminance est normalement traité par un système de mouvement basé sur l'énergie (Zaidi & DeBonet, 2000), lorsqu'un stimulus défini par le contraste est traité par un système de mouvement basé sur le suivi d'une caractéristique, ces deux stimuli ne sont pas traités par des mécanismes communs. De plus, même si un stimulus défini par la luminance était traité par un système de mouvement basé sur le suivi d'une caractéristique, l'absence d'interaction inter-attribut obtenue pour les stimuli statiques (Chapitre III) suggère que les mouvements de stimuli définis par la luminance et le contraste seraient également dissociables. En effet, l'absence d'interaction inter-attribut pour la détection d'un attribut statique suggère qu'une caractéristique spatiale d'un attribut (ex : sa forme) ne soit pas masquée par la présence de bruit d'un autre attribut. Par conséquent, le changement de position en fonction du temps de cette caractéristique devrait être peu affecté par le bruit d'un autre attribut permettant ainsi la perception du mouvement utilisant une stratégie basée sur le suivi d'une caractéristique. Le fait qu'à basses fréquences temporelles le traitement du mouvement défini par la luminance et celui du mouvement défini par le contraste ne soient pas affectés dans la même proportion par différents types de bruit (Chapitre V et Chapitre VI) est donc compatible avec le fait que le mouvement défini par le contraste était traité par un système de mouvement basé sur le suivi d'une caractéristique.

Lorsque les mouvements définis par la luminance et le contraste sont traités par un système de mouvement basé sur l'énergie, déterminer si ces mouvements sont traités par des mécanismes communs ou distincts est plus controversé. Puisqu'un stimulus défini par le contraste ne possède pas d'énergie à la fréquence du signal, le traitement énergétique de celui-ci requière préalablement des nonlinéarités introduisant de l'énergie à cette fréquence. D'une part, il est possible que des nonlinéarités intrinsèques au système visuel précédant le traitement énergétique de luminance permettent également le traitement de stimuli définis par le contraste. Selon cette hypothèse, les mouvements définis par la luminance et le contraste seraient donc traités par des mécanismes communs. D'une autre part, il pourrait exister un mécanisme énergétique dédié au traitement de stimuli définis par le contraste suivant un processus évaluant explicitement le contraste local (rectification) introduisant de l'énergie à la fréquence du signal (Wilson, Ferrera & Yo, 1992). En d'autres termes, des mécanismes de mouvement basés sur l'énergie pourraient traiter du mouvement défini par le contraste suivant l'extraction de la forme du signal. De tels mécanismes seraient distincts d'un mécanisme énergétique traitant la luminance. Dans les Chapitre V et Chapitre VI, le fait que le mouvement défini par la luminance et celui défini par le contraste n'étaient pas dissociables à hautes fréquences temporelles (les deux étaient affectés par la même proportion), suggère l'implication de mécanismes communs. En effet, chacun des bruits définis par la luminance et le contraste avait relativement le même impact sur les traitements de mouvement de luminance et de contraste. Étant donné que l'interaction inter-attribut ne pouvait s'expliquer par un produit de distorsion global (non-linéarité uniforme précédent le traitement de luminance), la conclusion généralement admise est qu'il existe un mécanisme de mouvement basé sur l'énergie dédié au mouvement de deuxième ordre (Lu & Sperling, 1995, 2001; Scott-Samuel & Georgeson, 1999). En effet, en superposant spatialement un signal défini par la luminance avec un signal défini par le contraste, la performance était similaire à celle prédite par un modèle probabiliste supposant que ces deux signaux étaient traités de façon indépendante (Lu & Sperling, 1995). Cependant, en définissant les non-linéarités du système visuel par un produit de distorsion global, ces auteurs ont implicitement présupposé que les non-linéarités du système visuel étaient uniformes en fonction du temps et de l'espace. Le Chapitre VI propose plutôt que des non-linéarités non-uniformes puissent expliquer le fait que ces deux attributs ne soient pas dissociables à hautes fréquences temporelles. Des résultats similaires ont été observés en simulant des non-linéarités non-uniformes, c'est-à-dire que la performance était similaire à celle prédite par le modèle probabiliste lorsqu'un mouvement défini par la luminance était superposé à un mouvement simulant l'impact de non-linéarités non-uniformes appliquées à un mouvement défini par le contraste (également défini par la luminance). Il n'est donc pas nécessaire de supposer l'existence d'un mécanisme de mouvement basé sur l'énergie dédié uniquement au mouvement défini par le contraste. L'existence de non-linéarités non-uniformes intrinsèques au système visuel permettrait d'expliquer l'apparente absence d'interaction inter-attribut lorsque deux signaux sont superposés spatialement et l'interaction inter-attribut lorsque du bruit d'un attribut masque le traitement de l'autre attribut.

Conclusion générale

Due à l'absence d'interaction inter-attribut pour la détection de signaux statiques, nous concluons que les stimuli statiques définis par la luminance et le contraste (premier et deuxième ordres, respectivement) sont, au moins initialement, traités séparément. Il existerait donc des mécanismes distincts pour le traitement spatial de deuxième ordre. Par contre, les mouvements à hautes fréquences temporelles (basés sur l'énergie) définis par la luminance et le contraste n'étaient pas dissociables suggérant que ces deux attributs étaient traités par des mécanismes communs. Des non-linéarités non-uniformes intrinsèques au système visuel pourraient expliquer qu'un mouvement défini par le contraste soit traité par les mécanismes sensibles à la luminance. Il n'est donc pas nécessaire d'inférer l'existence d'un système de mouvement basé sur l'énergie dédié au traitement de deuxième ordre. Le fait que les mouvements définis par la luminance et le contraste soient dissociables à basses fréquences temporelles est compatible avec l'hypothèse que le mouvement défini par le contraste soit traité par un système basé sur le suivi d'une caractéristique (*feature tracking*). En effet, si une caractéristique spatiale n'est pas masquée par le bruit d'un autre attribut (absence d'interaction pour les stimuli statiques), alors il n'est pas surprenant que le suivi de cette caractéristique dans le temps ne soit également pas (ou peu) masqué par le bruit d'un autre attribut. Nous concluons donc qu'il existe des mécanismes dédiés uniquement au traitement spatial de deuxième ordre, mais qu'il n'y a qu'un système de mouvement basé sur l'énergie traitant les stimuli de premier et deuxième ordres.

Bibliographie

Gegenfurtner, K.R., & Kiper, D.C. (1992). Contrast detection in luminance and chromatic noise. *Journal of the Optical Society of America A-Optics & Image Science, 9* (11), 1880-1888.

Georgeson, M.A., & Schofield, A.J. (2002). Shading and texture: separate information channels with a common adaptation mechanism? *Spat Vis, 16* (1), 59-76.

Lu, Z. L., & Sperling, G. (1995). The functional architecture of human visual motion perception. *Vision Research, 35*(19), 2697-2722.

Lu, Z. L., & Sperling, G. (2001). Three-systems theory of human visual motion perception: review and update.[see comment][erratum appears in J Opt Soc Am A Opt Image Sci Vis 2002 Feb;19(2):413]. *Journal of the Optical Society of America, A, Optics, Image Science, & Vision, 18*(9), 2331-2370.

Schofield, A.J., & Georgeson, M.A. (1999). Sensitivity to modulations of luminance and contrast in visual white noise: separate mechanisms with similar behaviour. *Vision Res, 39* (16), 2697-2716.

Scott-Samuel, N. E., & Georgeson, M. A. (1999). Does early non-linearity account for second-order motion? *Vision Research, 39*(17), 2853-2865.

Seiffert, A. E., & Cavanagh, P. (1999). Position-based motion perception for color and texture stimuli: effects of contrast and speed. *Vision Research, 39*(25), 4172-4185.

Smith, A. T. (1994). Correspondence-based and energy-based detection of second-order motion in human vision. *Journal of the Optical Society of America, A, Optics, Image Science, & Vision, 11*(7), 1940-1948.

Wilson, H. R., Ferrera, V. P., & Yo, C. (1992). A psychophysically motivated model for two-dimensional motion perception. *Vis Neurosci, 9*(1), 79-97

Zaidi, Q., & DeBonet, J. S. (2000). Motion energy versus position tracking: spatial, temporal, and chromatic parameters. *Vision Research, 40*(26), 3613-3635.

Annexe I

The noisy-bit method for digital displays: converting a 256 luminance resolution into a continuous resolution

Rémy Allard et Jocelyn Faubert

Laboratoire de psychophysique et perception visuelle

École d'optométrie, Université de Montréal

Le contenu de cet annexe a précédemment été publié sous forme d'article (©Psychonomic Society) :

Allard, R., & Faubert, J. (2008). The noisy-bit method for digital displays: Converting a 256 luminance resolution into a continuous resolution. *Behavior Research Methods, 40*(3), 735-743.

Abstract

Visual psychophysics often manipulates the contrast of the image on a digital display screen. Therefore, the limitation of the number of different luminance intensities displayable for most computers (typically 256) is frequently an issue. To avoid this problem, experimenters generally need to purchase special hardware (graphic cards) and/or develop specific computer programs. Here we describe an easy to implement method consisting in adding noise to the displayed stimulus that we call the "noisy-bit" method. This random dithering method generalized to 256 luminance intensities is equivalent to displaying continuous luminance intensities plus a certain amount of noise. Psychophysical testing using a standard spatiotemporal resolution (60 Hz and 1024x768 pixels) demonstrated that the noise introduced by the noisy-bit method has no significant impact on contrast threshold and is not visible. We conclude that the noisy-bit method combined with the standard 256 luminance levels is perceptually equivalent to an analog display with a continuous luminance intensity resolution when the spatiotemporal resolution is high enough so that the noise becomes negligible (which is easily attainable with typical spatiotemporal resolutions of present-day computers).

Keywords: Grayscale; display; 8-bit; noise; noisy-bit; luminance resolution; dithering

Introduction

Visual psychophysics often manipulates the contrast of the image on a digital display screen. A computer screen can display digital images with a relatively high spatiotemporal resolution. Indeed, most computer displays can produce relatively well defined images (typically at least 1024x768 pixels) at a relatively high temporal resolution (60 to 200 images per seconds). This high spatiotemporal resolution enables computers to display digital images resembling analog images. Indeed, if the spatial and/or temporal resolution is great enough, there will be no significant differences between a digital and analog image. For instance, a luminance grating, which would ideally vary continuously over space and/or time, varies in a discrete manner when presented on a digital display (Figure AnnexeI-1) but appears to vary continuously if the spatial and/or temporal resolution is great enough. Hence, high frequency luminance variations are spatiotemporally summed by the visual system and therefore undetected (Watson, Ahumada & Farrell, 1983).

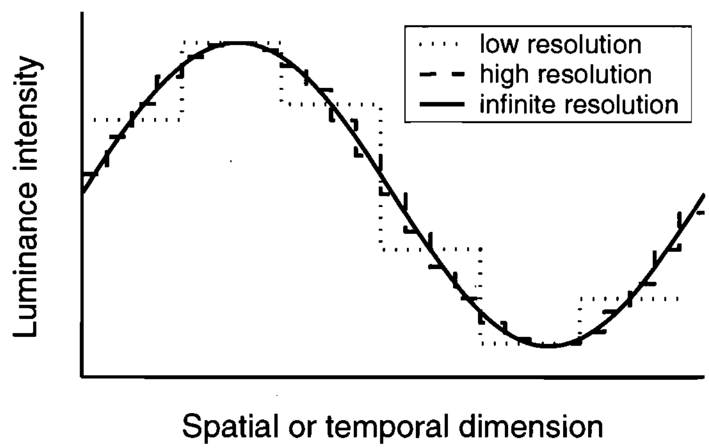


Figure AnnexeI-1. Discrete and continuous resolutions. Luminance variation of two discrete resolutions compared to a continuous resolution as a function of either space or time. If the resolution is high enough, the differences between a discrete and continuous resolution are not detectable.

The grayscale resolution problem

Analogously to the spatiotemporal resolution, the luminance intensity of each pixel is also discrete. The luminance of each pixel of a digital image sent to the display is defined by a digital value typically ranging between 0 and 255, which are called digital-to-analog converter (DAC) values. For sake of simplicity, we will omit that there are three different color guns and we will define each pixel color only by its luminance intensity. In other words, for any given pixel, we will assume that all the three guns are set to the same DAC value. The digital-to-analog converter translates each value into a voltage resulting into a given luminance intensity. Before psychophysical testing, the relation between the DAC value sent to the display (i for an integer) and the luminance intensity produced (d for a discrete value) is typically made linear (i.e. gamma corrected to be proportional to the DAC value):

$$d = \frac{L_{255}}{255} i \qquad (36)$$

where L_{255} represents the luminance intensity emitted when the DAC value is 255. The DAC values are integers ranging between 0 and 255, which limits the number of different displayable luminance intensities. However, the mathematical function defining the luminance intensity of each pixel of the stimulus ($L(x,y,t)$ for the luminance intensity of the pixel spatially positioned at (x,y) at time t) is generally continuous. Knowing the relation between the DAC value and the displayed luminance intensity (equation 36) enables the unit conversion of the luminance intensity of a given pixel (l, for simplicity we will refer to a given pixel which enables us to drop the spatiotemporal position of the pixel (x,y,t) so that $l=L(x,y,t)$) to a continuous DAC value (r for a real value), which is simply the inverse of equation 36:

$$r = \frac{255}{L_{255}} l \qquad (37)$$

However, DAC values are not continuous and must be integers. As a result, the DAC values must be rounded to the nearest integer before being sent to the display:

$$i = \lfloor r + 0.5 \rfloor \qquad (38)$$

where $\lfloor x \rfloor$ represents the floor function (i.e. rounding x to lower integer) and $\lfloor x+0.5 \rfloor$ thereby represents rounding x to the nearest integer. By combining these three equations, the relation between the luminance function defining each pixel (l) and the discrete displayed luminance (d) of this pixel is:

$$d = \frac{L_{255}}{255}\left\lfloor \frac{255}{L_{255}} l + 0.5 \right\rfloor \qquad (39)$$

In other words, there is a limited quantity of displayable luminance intensities (for most computers 256) and the luminance value of each pixel (l) is typically rounded to the nearest DAC value or, which is equivalent, to 1/255 the maximal luminance intensity (L_{255}). This grayscale resolution can often be too low to measure contrast thresholds (Pelli & Zhang, 1991). Indeed, for many spatiotemporal frequencies, the smallest perceptible contrast is less than the smallest luminance intensity difference displayable ($L_{255}/255$). Since contrast thresholds are often measured in psychophysics, the grayscale resolution of most computer displays is frequently an issue.

Current solutions

Hardware

Various methods are used to solve the grayscale resolution problem. The first obvious solution is to purchase a graphic card able to display more than 256 different luminance intensities. This solution does not only require buying special graphic cards, it also necessitates particular programs (generally homemade) interfacing directly with the graphic card. Indeed, under Windows™, very few graphic cards are able to display 1024 grayscales (10 bits) as the majority is limited to 8 bits. To use 10-bit graphic cards, experimenters typically need to develop their own software directly interfacing with the graphic card. On Macintosh computers, today's graphic cards now generally display grayscales with a 10 bit precision and some special graphic cards can display up to 12 or 14 bits. However, they also require specific programming directly interfacing with the lookup table of the graphic card. Moreover, although they can display more than 256 different luminance intensities, only 256 can be displayed simultaneously. Furthermore, 10-bit resolution may still not be sufficient. As a result, the hardware solution could solve the grayscale resolution problem in some conditions but may not be easily applicable for all experimenters.

Bit-stealing

The bit-stealing method (Faubert, 1991; Tyler, 1997) suggests the use of a chromatic jitter to enhance luminance intensity precision. Instead of having the same DAC value for all three color guns, each gun may have slightly different DAC values. Having different DAC values (e.g. 128,129,128) for the three color guns, enables the display of luminance intensities with a greater precision than when all three guns have the same DAC values (e.g. (128,128,128) or (129,129,129)). This alters the chromaticity of the pixel but also displays luminance intensities with a greater resolution. Since we are less sensitive to chromatic variations than luminance intensity variation, such chromatic jitter is generally not detectable. The drawback of this method is that it is relatively complex to implement and still limits the number of grey levels that can be displayed.

Dithering

Some printers, faxes or display devices can only produce 2 colors (typically black and white), and are used to display images typically defined by 256 grayscale intensities. Many techniques, called "halftoning" or "dithering" (Ulichney, 1987), have been developed to artificially display grayscale images using binary output devices. Basically, these techniques consist of using the spatial resolution to give the illusion of presenting grayscale images. For instance, if, within a small spatial region, half of the pixels are black and the others are white, then the spatial integration of the visual system will result in a gray percept.

The simpler dithering algorithm is "random dithering" (Ulichney, 1988). This algorithm proposes to compare the luminance intensity of each pixel of the original image with a cutoff criterion which is randomly selected for each pixel from a uniform distribution varying between the two displayable luminance intensities. If the pixel luminance of the original image is greater than the cutoff criterion, the output pixel is white, otherwise it is black. This method has the advantage of being easy to implement. However, as stated by Ulichney (1988), "the quality of output of this method does not deserve consideration for practical use". Consequently, this method is generally described only for theoretical purposes.

The difference between the original image and the displayed image corresponds to the noise introduced by the dithering. Many algorithms have been developed to minimize the visual impact of this noise. Ulichney (1988) suggested using highpass noise (typically referred as blue noise), which is more "pleasant" than the white noise (random dithering). The "ordered dithering" algorithm (Bayer, 1973) consists in using a given ordered pattern of the cutoff criteria instead of randomly selecting each cutoff criterion. The "error diffusion" algorithm (Floyd & Steinberg, 1976) consists in subtracting the noise introduced at each pixel to adjacent pixels. Mulligan and Ahumada (1992) proposed using knowledge about the contrast sensitivity function of the visual system to minimize the noise at the frequencies we are the more sensitive. All-in-all, many algorithms have been developed with the goal of proposing a dithering technique that would minimize the visual impact of the noise. They all have the advantage of giving a better image quality than random dithering. However, they also all have the disadvantage of being more complicated to implement and generally require more computer resources.

To enhance luminance resolution of typical displays, Mulligan (1990) proposed a simple way of generalizing ordered dithering used for bi-level displays to higher luminance resolution displays. Basically, the algorithm consists in applying ordered dithering independently to each pixel by selecting between the two nearest luminance intensities displayable instead of between the only two luminance intensities available. Note that, even though Mulligan described this generalization for ordered dithering, it could also be applied to any dithering algorithm.

Daly and Feng (2005) implemented such generalization to ordered dithering to enhance apparent luminance intensity resolution, which they named “bit-depth extension”. They developed a sophisticated algorithm with the aim of creating an ordered pattern that would minimize the visibility of the noise. They used the contrast sensitivity function of the visual system to determine at which spatiotemporal frequencies we are less sensitive. They construct their pattern so that the noise introduced by dithering is concentrated to these frequencies. They found that, in some cases, the noise was invisible so that the bit-depth of the original image could be reduced without noticeable impact.

As a result, there is no simple solution to the grayscale resolution problem. Most techniques require special hardware and/or relatively complex programming. The purpose of the present paper is to propose a technique requiring no special hardware and no complex programming that can display stimuli with an infinite number of gray levels. Combining the high spatiotemporal resolution of computer displays with a simple modification of the stimulus function could solve this problem.

The noisy-bit method

As mentioned above, the luminance intensity for each pixel is typically defined by a continuous value (I) that generally has to be rounded with a precision of 1/255 the maximal luminance intensity displayable (L_{255}), i.e. to the nearest DAC value. This procedure can sometimes create sufficiently high artifacts to impair contrast threshold measurement.

Algorithm

Instead of simply rounding to the nearest DAC value (equation 38), we propose a different algorithm consisting in randomly choosing between the two nearest DAC values. The probability distribution between the two values can be set so that the expected value is equal to the continuous DAC value defined by the stimulus function (r). That is, the probability of choosing the higher DAC value is equal to the remainder of the continuous DAC value. For example, if the continuous DAC value is 123.25, then the probability distribution would be 0.25 for 124 and 0.75 for 123 resulting into an expected value of 123.25. Consequently, the noisy-bit method proposes to replace equation 38 by:

$$i = \begin{cases} \lceil r \rceil & \text{if } \mathrm{R}(r) \\ \lfloor r \rfloor & \text{otherwise} \end{cases} \qquad (40)$$

where R(x) returns true with a probability equal to the remainder of x (i.e. x-$\lfloor x \rfloor$) and false otherwise. $\lceil x \rceil$ and $\lfloor x \rfloor$ represent the ceiling and floor functions respectively (i.e. rounding to the upper and lower integer, respectively). Note that this method is equivalent to combining random dithering (Ulichney, 1988) with the generalized application of dithering for 256 instead of 2 luminance intensities (Mulligan, 1990). Indeed, randomly selecting between the two nearest DAC values ($\lfloor r \rfloor$ and $\lceil r \rceil$) with a probability of choosing the highest DAC value ($\lceil r \rceil$) equal to the remainder of the DAC value (r-$\lfloor r \rfloor$) is mathematically equivalent to rounding to one of the two nearest DAC values with a random cutoff criterion selected from a uniform distribution varying between the two nearest DAC values.

The main drawback of the noisy-bit method is that it increases the error between the desired continuous luminance intensity (l) and the displayed luminance intensity (d) (Figure AnnexeI-2). Indeed, this method will choose between the two nearest DAC values so that it will occasionally select a DAC value further than the nearest integer. The error will be equal to 1-x or x depending on whether the continuous DAC value is rounded to the highest or lowest integer respectively, where x represents the remainder of the continuous DAC value (r-$\lfloor r \rfloor$). Assuming that the remainders of the continuous DAC values are uniformly distributed, the root-mean-square (RMS) error can be calculated using the following equation:

$$RMS\ error = \sqrt{\int_0^1 \mathrm{P}(x)(1-x)^2 + (1-\mathrm{P}(x))x^2 dx} \qquad (41)$$

where P(x) corresponds to the probability of selecting the highest integer. When simply rounding to the nearest integer, this probability is equal to 1 if x>0.5 and 0 otherwise, resulting in a RMS error of 0.29 DAC values. Using the noisy-bit method, the probability of rounding to the highest integer is equal to x, resulting in RMS error of 0.41 DAC values. As a result, the RMS error between the continuous DAC value (r) and the DAC value sent to the display (i) will be $\sqrt{2}$ times greater using the noisy-bit method than simply rounding to the nearest integer. However, this random selection will result in an expected value that will be equal to the desired continuous value ($E(i)= r$). Conversely, rounding to the nearest integer will cause the expected displayed value to be equal to the nearest integer which is generally not equal to the desired continuous value ($E(i)=\lfloor r+0.5\rfloor \neq r$).

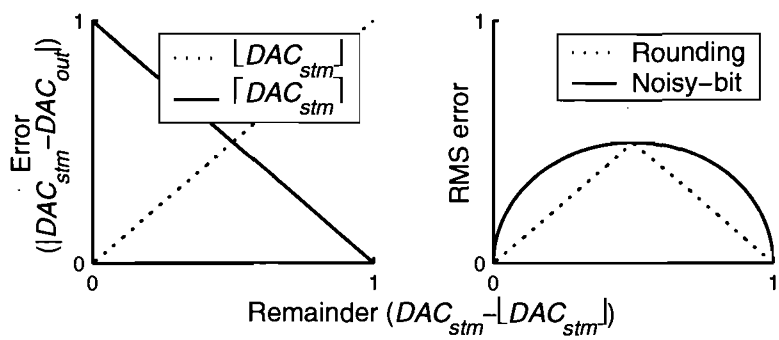


Figure AnnexeI-2. Rounding error. The left graph shows the error (i.e. difference between the continuous DAC value (r) and the displayed DAC value (i)) as a function of the remainder of the continuous DAC value. The two lines show the error for rounding to the lower (dash line) or upper (solid line) DAC values. The right graph shows the root-mean-square (RMS) of this error using two methods of selecting between the lower or upper DAC values. Rounding to the nearest DAC value always results in the lowest error in the left graph. The noisy-bit method rather proposes to randomly select between the two nearest DAC values with a given probability.

Thus, this method is equivalent to displaying the desired luminance intensity (l or r in DAC units) plus a certain amount of noise due to the difference between the desired luminance intensity and the luminance displayed (l-d or r-i). In other words, the noisy-bit method converts an 8-bit (256) grayscale resolution into a continuous grayscale resolution with the drawback of adding noise. The noise energy directly depends on the size and duration of presentation of each element (in our case, pixel) and the impact of the noise decreases when the spatiotemporal resolution increases. If the size of the noise elements is small or its duration is brief, the local luminance variation between each pixel will be summed by the visual system and only the mean value will be perceived. On the other hand, large noise element size and low temporal resolution may result into perceivable noise elements. Using a high spatiotemporal resolution, the impact of the noise should be negligible and the displayed stimulus will be perceptually equivalent to the stimulus defined by continuous luminance values. However, if the spatiotemporal resolution is too poor, the noise introduced could be detected and affect contrast thresholds.

Note that the error between the desired and the displayed DAC values is not constant as a function of the remainder of the desired DAC value. Consider the two extreme cases when the remainder of desired DAC value is either 0 or 0.5 (e.g. desired DAC values of 128 or 128.5). In the first case, all the output DAC values will be 128 and no noise would therefore be added to the display. In the second case, each pixel would be randomly selected between 128 or 129 and the displayed luminance would be noisy. Consequently, these two desired DAC values would generate extremely different displays: no noise or maximum noise (i.e. the luminance error for each pixel would be +/- 0.5 DAC values resulting in a RMS error of 0.5 DAC values). However, in both cases the mean luminance would be near (or equal to) their desired DAC value. Consequently, if the spatiotemporal resolution is high enough so that the noise is not visible, both cases would result in similar percepts: two uniform grays with slightly different intensities.

Implementation

This method of randomly selecting between the two nearest DAC values (equation 40) is mathematically equivalent to rounding to the nearest DAC value after adding a noise value randomly selected from a uniform distribution varying between -0.5 and 0.5 DAC values (N). For instance, if the continuous DAC value (r) is 123.25, then randomly selecting a value between 122.75 and 123.75 and then rounding to the nearest integer results in a probability of selecting the DAC value 123 equal to 0.75 and a probability of selecting 124 equal to 0.25. Consequently, the noisy-bit method can be implemented by replacing equation 40 with:

$$i = \lfloor r + N + 0.5 \rfloor \qquad (42)$$

Matching equation 36, 37 and 42, we obtain

$$d = \frac{L_{255}}{255} \left\lfloor \frac{255}{L_{255}} l + N + 0.5 \right\rfloor \qquad (43)$$

By defining

$$l' = l + \frac{L_{255}}{255} N \quad (44)$$

we obtain the same function as equation 39:

$$d = \frac{L_{255}}{255} \left\lfloor \frac{255}{L_{255}} l' + 0.5 \right\rfloor \qquad (45)$$

Consequently, the noisy-bit method can simply be implemented by adding a small amount of noise to the luminance function (equation 44) rather than by explicitly implementing the random selection between the two nearest DAC values.

As mentioned above, for sake of simplicity we referred to a given pixel which enables us to drop the spatiotemporal position of the pixel (x,y,t). Consequently, equation 44 can be reformulated more generally as:

$$L'(x,y,t) = L(x,y,t) + \frac{L_{255}}{255} N(x,y,t) \quad (46)$$

In other words, the noisy-bit method can be implemented by adding uncorrelated noise ($N(x,y,t)$) with a given contrast ($L_{255}/255$, i.e. 1 DAC value) to the stimulus function ($L(x,y,t)$).

Evaluating the impact of the noise

Contrast thresholds in noise have been widely studied. The threshold-versus-contrast (TvC, Figure AnnexeI-3) function was found to give a reasonably good fit of the contrast threshold as a function of the external noise contrast (n_{ext}) (Legge, Kersten, & Burgess, 1987; Pelli, 1981, 1990; Pelli & Farell, 1999):

$$c(n_{\text{int}}) = k\sqrt{n_{\text{int}}^2 + n_{ext}^2} \quad (47)$$

where n_{int} corresponds to the internal equivalent noise, that is, the contrast of the external noise having the same impact as the internal noise, and k is proportional to the smallest signal-to-noise ratio required to detect the signal.

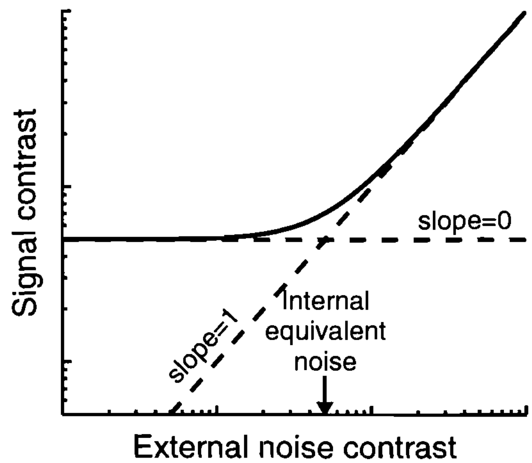


Figure AnnexeI-3. Contrast threshold as a function of external noise contrast (TvC function). The internal equivalent noise corresponds to the breaking point of the curve. When the external noise contrast is significantly lower than the internal equivalent noise, it has no significant impact (slope=0). When the external noise contrast is significantly greater than the internal equivalent noise, the internal noise has no significant impact and the threshold is proportional to the external noise contrast (slope=1). Note that the two axes are scaled logarithmically.

For the purpose of the present study, the important parameter of this function is the internal equivalent noise (n_{int}), which corresponds to the breaking point of the curve. When the external noise is significantly greater than the internal noise ($n_{ext}>>n_{int}$), the internal noise has no significant impact and the contrast threshold is proportional to the external noise contrast (slope of 1 in log-log units). In other words, for this portion of the curve, if you increase the contrast of the external noise by a given factor, the contrast threshold will increase by the same proportion. However, when the external noise is significantly lower than the internal noise ($n_{ext} << n_{int}$), the external noise has no significant impact and the contrast threshold is independent of the external noise contrast (slope of 0).

As mentioned above, the noisy-bit method is equivalent to presenting a stimulus with a continuous grayscale precision ($L(x,y,t)$) plus a certain amount of noise. Using the TvC function, it should be possible to determine if this noise has a significant impact. If the noise introduced by the noisy-bit method is much smaller than the internal noise, then it would be insignificant and would have no impact on the contrast threshold. Again, the noisy-bit method is equivalent to presenting the stimulus with the desired continuous value ($L(x,y,t)$) plus a certain amount of noise. If this noise is not significant, then the noisy-bit method is equivalent to presenting a stimulus with continuous gray levels, i.e. with an infinite number of luminance intensities.

Experiment 1: The impact of the noise

The objective of the present experiment was to evaluate whether a spatiotemporal resolution of a typical digital display (60 Hz and 1024x768 pixels) is great enough to measure contrast thresholds using the noisy-bit method. As mentioned above, the noisy-bit method may be implemented by adding noise to the luminance function defining the stimulus with a uniform distribution ranging between ±0.5 DAC values or, which is equivalent, ranging between $\pm L_{255}/(255\times2)$ luminance intensity. We define the noise contrast, which can be represented in luminance intensity or DAC values, as the range covered by the uniform distribution. Note that equation 37 can be used to pass from luminance intensity units to DAC units. Using the noisy-bit method, the noise contrast added to the stimulus function is 1 DAC value or $L_{255}/255$ luminance intensity.

To assess if the noise introduced within the displayed stimulus by the noisy-bit method affects the contrast threshold, the contrast threshold of a given stimulus was evaluated as a function of the noise contrast. If the noise is a limiting factor, than increasing the noise contrast should affect the contrast threshold by the same proportion (slope of 1 on the TvC function). Alternatively, if the observer's internal noise is greater than the external noise (i.e. the noise introduced by the noisy-bit method), then increasing the external noise will not affect contrast threshold (slope of 0 on the TvC function).

Method

Observers

Two observers participated in the study. One of them was one of the authors and the other was naïve to the purpose of the experiment. Both had normal or corrected-to-normal vision.

Apparatus

The stimuli were presented on a 19 in ViewSonic E90FB .25 CRT monitor powered by a Pentium 4 computer combined with a Matrox Parhelia512 graphic card. All three color guns were constrained to have the same DAC value. As a result, this setup could display 256 different luminance intensities (8-bit luminance depth). The greatest luminance intensity attainable (L_{255}) was 94 cd/m^2. The display was gamma corrected using a Minolta CS100 photometer interfaced with a homemade program to produce a linear relationship between the DAC value and the luminance intensity. The refresh rate was set to 60 Hz, which is typically the lowest refresh rate for most computers. The screen resolution was set to the most standard screen resolution of 1024x768 pixels covering an area of 32x24 cm. At the viewing distance of 114 cm, the width and height of each pixel were 1/64 deg of visual angle. In other words, the spatial resolution of the displayed stimulus was 64 pixels/deg. The monitor was the only light source in the room.

Stimuli

To measure contrast thresholds, sine wave gratings are the most widely used stimuli:

$$L(x,y,t) = L_{128}(1 + c\sin(fx + p)) \qquad (48)$$

where L_{128} corresponds to the mean luminance value (47 cd/m^2, which was the luminance intensity emitted when the DAC values were set to 128). c corresponds to the stimulus Michelson contrast and was the dependent variable. f corresponds to the spatial frequency, which was fixed to 4 cpd (approximately the spatial frequency to which we are the most sensitive). And p represents the phase, which was randomized at each presentation. Notice that the luminance of the grating only depended on the horizontal position (x) and not on the vertical position (y) or the time (t). Consequently, the grating was vertically orientated and static.

To implement the noisy-bit method noise must be added to the stimulus function:

$$L'(x,y,t) = L(x,y,t) + n_{ext}N(x,y,t) \quad (49)$$

where n_{ext} represents the noise contrast. As mentioned above, for the noisy-bit method the contrast of the noise must be fixed to $L_{255}/255$ luminance intensity (equation 46) or, which is equivalent, 1 DAC value. However, in the present experiment we varied the noise contrast so that n_{ext} varied between 1 and 230 DAC values using 7 different noise contrasts. Examples of stimuli are presented in Figure AnnexeI-4.

Figure AnnexeI-4. Sine wave gratings in noise. The contrast of the signal (c) is set to 0.1. From left to right, the noise contrast is 1, 10 and 100 DAC values.

For static stimuli, adding dynamic noise implies passing from a static presentation (an image) to a dynamic presentation composed of several images. A dynamic presentation consumes more computer resources (memory, processing time, etc) than a static presentation which only requires the rendering of a single image. Consequently, passing from a static to a dynamic presentation may not always be convenient and may thereby limit the application of the noisy-bit method. However, the noisy-bit method may also be applied using static noise. That is, the noise template added to the stimulus would not vary over time ($N(x,y)$ instead of $N(x,y,t)$) so that the exact same image would be presented in all frames. For such application, only the spatial summation would permit the integration of the different pixels. If the spatial resolution is high enough, the noise introduced by the noisy-bit method should not affect contrast thresholds. To evaluate if only the spatial resolution could permit the application of the noisy-bit method, we applied the method both spatially (static noise) and spatiotemporally (dynamic noise).

To minimize contrast thresholds, a relatively large spatiotemporal window was used. The presentation time of the stimulus was 500 ms and the spatial window was a disk with a diameter of 2 degrees of visual angle with a soft edge defined by a half cosine of 0.5 degrees.

Procedure

A two alternative-interval-forced-choice task was used, which consisted in identifying the interval in which the sine wave was present by pressing one of two keys. Both intervals contained the same noise contrast (n_{ext}) but were generated by two distinct noise samplings. The delay between the two intervals was 500 ms. Between stimuli presentations, the screen remained blank at the mean luminance level (L_{128}) and a fixation point was presented.

The contrast (c) of the grating in the interval in which the sine wave was presented was manipulated by a 2-down-1-up staircase procedure (Levitt, 1971). In the other interval, the contrast (c) was set to 0. The staircase was interrupted after 10 inversions and the threshold was evaluated as the geometric mean of the last 4 inversions. The step size was fixed to 0.05 log units and the initial contrast (c) was always set well above threshold.

Overall, there were 14 different noise conditions: 7 noise contrasts and the noise was either static or dynamic. These 14 conditions were evaluated 3 times each resulting into 42 staircases performed in a pseudorandom order. For each of these 14 conditions, the resulting threshold was estimated as the geometric mean of the 3 staircases.

Results and discussion

The internal equivalent noise measured was 22 and 16 DAC values in the static noise condition and 71 and 44 DAC values in the dynamic noise condition for observers il and ra, respectively (Figure AnnexeI-5). As a result, the detection thresholds using the noisy-bit method (conditions when the noise contrast was 1 DAC value) were all on the 0-slope portion of the TvC function, i.e. the noise introduced by the noisy-bit method was considerably smaller than the observer's internal noise. Hence, it was possible to significantly increase the external noise contrast without affecting contrast threshold. We therefore conclude that the noise introduced by the noisy-bit method (noise contrast of 1 DAC value) did not affect contrast thresholds either in the static or in the dynamic condition.

As mentioned above, applying the noisy-bit method is equivalent to having a noisy continuous grayscale display. Using this method, the noise corresponds to the luminance variation introduced by randomly selecting between the two nearest DAC values, which corresponded to the conditions when the external noise contrast was 1 DAC value. The present experiment showed that this noise had no significant impact. We therefore conclude that the noisy-bit method enabled a 256 grayscale resolution apparatus to be perceptually equivalent to a continuous (i.e. infinite) grayscale resolution.

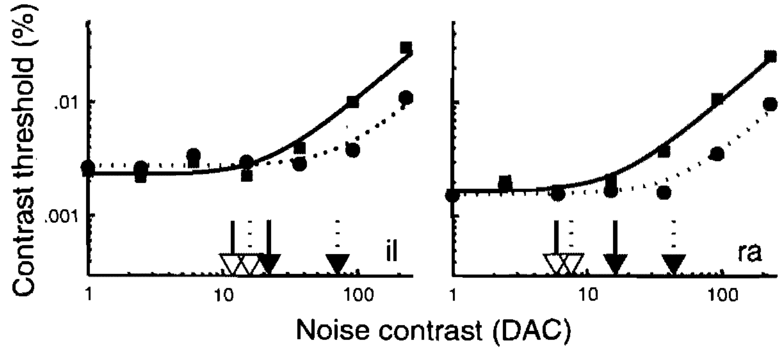


Figure AnnexeI-5. Results. Contrast thresholds as a function of the external noise contrast for the two observers. Squares and circles correspond to thresholds when the noise was static and dynamic respectively. The solid and dashed lines show the best fit of the TvC functions. The filled arrows illustrate the internal equivalent noise corresponding to the breaking point of the TvC functions. Empty arrows illustrate the noise contrast threshold in static (solid) and dynamic (dashed) conditions (results of experiment 2).

Experiment 2: Noise detection

The previous experiment showed that the noise introduced by the noisy-bit method did not significantly affect the contrast threshold of a given task. However, this does not imply that the noise was not detectable. A given noise contrast could be perceived without affecting contrast threshold. This would result into a qualitative difference between a continuous grayscale display and discrete grayscale display combined with the noisy-bit method. The objective of the present experiment was to show that the noise introduced by the noisy-bit method was not perceived even for relatively low spatiotemporal screen resolutions. If the noise is not perceptible, not only would the noisy-bit method enable contrast threshold measurements equivalent to continuous displays, it would also be qualitatively (or perceptively) equivalent. Indeed, the difference between a continuous display and 256 grayscale display would not be measurable nor perceptible.

Method

The same apparatus was used as in the previous experiment and the same two observers participated to the study. The stimuli were composed of noise:

$$L(x,y,t) = L_{128} + n_{ext}N(x,y,t) \qquad (50)$$

The noise detection task consisted in a two-interval-forced-choice procedure. One interval was blanked (n_{ext}=0, that is an even gray) and the other contained noise. A 2-down-1-up staircase procedure as described in the previous experiment was used to measure the noise contrast threshold (n_{ext}). Each threshold was evaluated 3 times in static and dynamic noise conditions resulting in 6 staircases.

Results and discussion

The noise contrast thresholds were 12 and 5.9 DAC values in the static noise condition and 16 and 7.6 DAC values in the dynamic noise condition for observers il and ra, respectively (Figure AnnexeI-5). Below these noise contrasts the observers were unable to differentiate between even gray and noise. Consequently, the noise introduced by the noisy-bit method (1 DAC value) was not perceptible. We therefore conclude that there was no qualitative or perceptible difference between a digital 8-bit grayscale display using the noisy-bit method and an analog display able to display an infinite number of grays. Note that this was true even when using a relatively low spatiotemporal resolution (0 Hz (i.e. static) and 64 pixels/deg) for present-day computers.

General discussion

The noisy-bit method introduces low contrast noise to enhance the luminance intensity precision of digital displays. This method is equivalent to displaying gray level with a continuous precision and adding noise to the displayed image. The two experiments showed that the low contrast noise introduced by the noisy-bit method does not affect contrast threshold and is not perceptible. We therefore conclude that, when the spatiotemporal resolution is high enough (which is easily attainable with typical computers), a discrete 8-bit display combined with the noisy-bit method is perceptually equivalent to an analog display having a continuous grayscale precision.

Evaluating the impact of the noise

Instead of rounding to the nearest DAC value, the noisy-bit method randomly chooses between the two nearest DAC values so that the expected value is equal to the continuous DAC value. As shown above, this method can be implemented in two steps: first add a given amount of noise and then round to the nearest DAC value. The second step is necessarily already implemented (DAC values must be integers) and is the same as the single step when not applying the noisy-bit method. Often, rounding to the nearest integer is already implicitly implemented by the program sending the image to the graphic card. Consequently, the noisy-bit method can be implemented simply by adding a given amount of noise to the stimulus function (equation 46).

Both steps actually affect the luminance profile of the displayed stimulus. Indeed, rounding to the nearest integer could also add noise to the displayed stimulus. However, for high noise contrasts, rounding to the nearest integer has no significant impact. Consequently, if adding noise with a contrast significantly greater than 1 DAC value did not significantly affect contrast threshold, then the luminance variation introduced by the noisy-bit method (that is, adding noise with a contrast of 1 DAC value and rounding to the nearest DAC value) certainly has no significant impact. We therefore conclude that simply adding considerable amount of noise to the luminance function and neglecting the luminance variation introduced by rounding to the nearest integer is an efficient validation to determine whether the noise introduced by the noisy-bit method affects contrast thresholds or not.

Using less than 8-bits

To simulate a display having 7-bit depth, we must add noise to the stimulus function with a contrast of 2 DAC values and then round to the nearest even integer. More generally, to simulate an N-bit display, we must add noise with a contrast of 2^N DAC values and then round to the nearest integer being a multiple of 2^{8-N}. We have measured contrast threshold as a function of the number of bits used to display the stimulus (data not shown) and found that with the current spatiotemporal resolution (64 pixels/deg and 60 Hz) contrast threshold could be measured using only a 5-bit display (noise contrast of 8 DAC values), that is, using only 32 different luminance intensities. Using a higher spatiotemporal resolution (128 pixels/deg and 120 Hz), we found that contrast threshold could efficiently be measured with a 3-bit display (noise contrast of 32 DAC values) that is, using only 8 different luminance intensities.

Reducing the noise of the noisy-bit method

In the present study, we showed that the noisy-bit method can be efficiently implemented using a spatiotemporal resolution that is relatively low for present-day computers (1024x768 pixels at 60 Hz) at a typical viewing distance for psychophysical testing (114 cm). Using these parameters, the noise introduced by the noisy-bit method was found to be too low to affect the contrast threshold of a stimulus at which we are highly sensitive (4 cpd sine wave grating with a relatively large spatiotemporal window). Although the noisy-bit method works well using a relatively low spatiotemporal resolution (1024x768 pixels at a viewing distance of 114 cm at 0 Hz) here we describe different ways to reduce the noise introduced by the noisy-bit method. These methods could be used in the eventuality that the noise introduced by the noisy-bit method becomes a limiting factor for a given condition.

The first obvious way to reduce the noise is to enhance the spatiotemporal resolution. This can be achieved by (1) increasing the temporal resolution of the display (many displays can reach 200 Hz), (2) increasing the spatial resolution of the display (many displays can reach 2048x1536 pixels) and/or (3) increasing the viewing distance.

Although the noisy-bit method was developed for 8-bit displays, it can easily be adapted for displays with more than 8 bits of depth. For instance, the thresholds observed in the current papers were as low as 0.0015 Michelson contrast. A graphic card able to display 1024 gray levels (10 bits) would not be sufficient to properly evaluate such threshold. However, the noisy-bit method could be combined to a 10-bit display to enhance luminance precision. Similarly, the bit-stealing method could also be combined with the noisy-bit method to enhance luminance intensity precision. The noisy-bit method could randomly choose between two DAC value combinations (e.g. (128,128,129) and (128,129,128)) so that the expected luminance intensity would be between the luminance intensities produced by these two combinations.

Alternatively, a simple way of reducing the noise introduced by the noisy-bit method is to apply this method independently to each color gun. Combining the luminance noise of the three guns would reduce the luminance noise without requiring any special hardware or sophisticated programming. Hence, each color gun could have its own noise sampling. When the three color guns are constrained to have the same DAC values, the noise sampling applied to the three guns is perfectly correlated. Without this constraint, the noise sampling is uncorrelated. Consequently, simply applying the noisy-bit method to each gun separately would reduce the luminance noise introduced by the noisy-bit method. Note that since we are less sensitive to chromatic jitter than luminance noise (especially at high spatiotemporal frequencies) if the luminance noise is not detectable, then the chromatic jitter would also not be detectable. Indeed, we found that independently applying the noisy-bit method to the three guns increased the noise contrast threshold (data not shown).

Moreover, instead of simply having an uncorrelated error between the three guns, they could be negatively correlated as suggested by Mulligan (1990) using order dithering. For the noisy-bit method, negatively correlating the error can be implemented by inverting (i.e. subtracting instead of adding) the noise added to the stimulus (equation 11) of one of the three color guns. In other words, the same noise matrix would be used for two guns and the inverted matrix would be used for the other gun. Since the green gun generally produces the highest luminance intensity, we suggest inverting the noise matrix of this gun. As a result, the perfectly correlated luminance errors of the red and blue guns would be partially canceled by the luminance error of the green guns.

If the noise introduced by the noisy-bit method affects the contrast threshold, another modification could be applied to limit its impact. The noise could be filtered to keep only the high spatial and/or temporal frequencies. Indeed, for small details or for high frequency flicker, low contrast stimuli (in our case noise) become undetectable and are therefore spatially and/or temporally summed by the visual system. Note that to add this modification to the noisy-bit method, the contrast of the noise added to the stimulus function would have to be of at least 1 DAC value once filtered.

Conclusion

Although the spatiotemporal resolution of today's computers is relatively high, the luminance intensity resolution is often too low (256 luminance intensities) for many tasks involving contrast manipulation. The noisy-bit method uses the high spatiotemporal resolution of computers to improve the luminance intensity resolution. By randomly selecting between the two nearest DAC values instead of rounding to the nearest DAC value, the noisy-bit method is a powerful tool to bypass the luminance intensity resolution problem. This method can be simply implemented by adding low contrast noise to the luminance function defining the stimulus. By testing the effect of adding higher contrast noise, one can assert that the noise added to the displayed stimulus has no significant impact on a given task. By evaluating the noise contrast detection threshold, one can also assert that the noise is not visible. As a result, the noisy-bit method successfully makes a typical digital display perceptually equivalent to a continuous luminance intensity resolution system.

Acknowledgments: This research was supported by NSERC graduate fellowship to RA and NSERC operating grant to JF.

References

Bayer, B. (1973). An optimum method for two level rendition of continuous tone pictures. *Conference Record, IEEE International Conference on Communications* 1(pp. (26-11)-(26-15)). New York.

Daly, S., & Feng, X.F. (2005). Bit-depth extension: Overcoming LCD-driver limitations by using models of the equivalent input noise of the visual system. *Journal of the Society for Information Display, 13* (1), 51-66.

Faubert, J. (1991). Effect of target size, temporal frequency and luminance on temporal modulation visual fields. In: R.P. Mills, & A. Heijl (Eds.), *Perimetry Update 1990/91* (pp. 381-390). New York.

Floyd, R., & Steinberg, L. (1976). An adaptive algorithm for spatial greyscale. *Proceedings of the Society for Information Display, 17* (2), 75-77.

Legge, G.E., Kersten, D., & Burgess, A.E. (1987). Contrast discrimination in noise. *J Opt Soc Am A, 4* (2), 391-404.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America, 49* (2), Suppl 2:467+.

Mulligan, J.B. (1990). Digital halftoning methods for selectively partitioning error into achromatic and chromatic channels. *Human Vision and Electronic Imaging: Models, Methods, and Applications*, 1249 (pp. 261-270). Santa Clara, CA, USA: SPIE.

Mulligan, J.B., & Ahumada, J.A.J. (1992). Principled halftoning based on human vision models. *Human Vision, Visual Processing, and Digital Display III*, 1666 (pp. 109-121). San Jose, CA, USA: SPIE.

Pelli, D.G. (1981). The effects of visual noise. *Department of Physiology*, Ph.D. (Cambridge: Cambridge University.

Pelli, D.G. (1990). The quantum efficiency of vision. In: C. Blakemore (Ed.) *Visual coding and efficiency* (Cambridge: Cambridge University Press.

Pelli, D.G., & Farell, B. (1999). Why use noise? *Journal of the Optical Society of America, A, Optics, Image Science, & Vision, 16* (3), 647-653.

Pelli, D.G., & Zhang, L. (1991). Accurate control of contrast on microcomputer displays. *Vision Research, 31* (7-8), 1337-1350.

Tyler, C.W. (1997). Colour bit-stealing to enhance the luminance resolution of digital displays on a single pixel basis. *Spatial Vision, 10* (4), 369-377.

Ulichney, R.A. (1987). Digital halftoning. Cambridge, MA: MIT Press.

Ulichney, R.A. (1988). Dithering with blue noise. *Proceedings of the IEEE, 76* (1), 56-79.

Watson, A.B., Ahumada, J.A.J., & Farrell, J.E. (1983). The window of visibility: a psychophysical theory of fidelity in time-sampled visual motion displays. NASA Technical Paper, 2211(National Technical Information Service, Springfield, VA 22161.).