

L'information visuelle efficace pour la reconnaissance des visages

Daniel Fiset

Department of Psychology, University of Victoria

Frédéric Gosselin

Département de psychologie, Université de Montréal

En un coup d'œil, un visage peut nous fournir beaucoup d'informations utiles à nos interactions sociales. Ainsi, en moins de 200 millisecondes, notre système perceptif aura extrait suffisamment d'informations visuelles pour permettre la reconnaissance du sexe, de la race, de l'âge et de l'état émotionnel d'une personne et ce, même si celle-ci nous est inconnue. Pour un visage familier, ces mêmes 200 millisecondes mèneront également à l'activation en mémoire d'une pléiade de renseignements sur cette personne (e.g., son boulot, les circonstances de notre première rencontre, son statut civil) puis peut-être à son identification. À titre d'exemple, tous auront aisément catégorisé l'individu photographié à la figure 1a comme un homme caucasien d'environ 30-45 ans. La majorité des lecteurs l'auront également catégorisé comme une célébrité connue principalement pour ses rôles au cinéma. Les amateurs de films américains l'auront finalement identifié comme Brad Pitt, un acteurs dont la carrière est truffée de films à succès (« Thelma and Louise », « Interview with the Vampire », « Legends of the fall », « Seven », « Twelve Monkeys », « Meet Joe Black », « Troy », etc.) et dont les aventures amoureuses avec de jolies actrices font le tour de la planète « potin ».

Cet exemple illustre l'idée qu'un stimulus visuel peut être classé dans plusieurs catégories différentes (homme, caucasien, acteur, américain, etc.). Quels indices visuels sont nécessaires à la catégorisation d'un visage selon son sexe, son expression faciale et son identité? S'il est impossible de traiter l'ensemble des informations simultanément, dans quel ordre doit-on les extraire afin d'être aussi efficace que possible? L'objectif principal de ce chapitre est de permettre aux lecteurs de mieux comprendre les stratégies perceptives utilisées par l'être humain en reconnaissance des visages.

-----

Figure 1

---

Dans ce chapitre, nous décrirons, d'abord, la théorie dominante dans le domaine de la vision qui propose que la première étape de traitement des visages et, plus généralement, des objets, revient *grosso modo* à effectuer une analyse de Fourier—une décomposition en une somme de grilles sinusoïdales. Ensuite, nous résumerons les différentes études ayant examinés les fréquences spatiales des grilles sinusoïdales nécessaires à la reconnaissance des visages. Puis nous décrirons un ensemble d'études effectuées dans notre laboratoire avec la méthode *Bubbles* (Gosselin & Schyns, 2001) afin de révéler plus précisément les attributs des visages que les participants utilisent « efficacement » pour accomplir la tâche qu'on leur a demandée. Enfin, une méta-analyse des résultats obtenus avec cette méthode suggérera une séquence dans l'extraction de l'information des visages.

### **Psychophysique de la reconnaissance des visages**

À la fin des années 50, Hubel et Wiesel (1959) ont proposé que la fonction des aires visuelles primaires était de détecter et d'extraire les bars et contours des objets environnants. Cette hypothèse quant au fonctionnement des aires visuelles primaires fût dominante en neurosciences pendant quelques décennies. À la fin des années soixante, des travaux importants effectués par Fergus Campbell et ses collaborateurs ont mené à une proposition différente, et aujourd'hui dominante, des premières étapes de traitement visuel effectuées par le cortex strié. En effet, ces auteurs ont proposé que l'atome de la perception visuelle était des grilles sinusoïdales (*sin-wave gratings*, en anglais; voir figure 2 pour des exemples de grilles sinusoïdales avec différentes propriétés). L'idée sous-jacente à cette théorie est qu'une image, quelle qu'elle soit, peut être

décomposée, sans perte d'information, en une somme unique de grilles sinusoïdales, chacune caractérisée uniquement par quatre paramètres : une fréquence spatiale, une orientation, une phase et un contraste (voir figures 2b-e). Par exemple, les « réseaux » (*plaid*, en anglais) sont décomposés en des sommes de deux grilles sinusoïdales possédant une amplitude différente de zéro (celles qui ont une amplitude égale à zéro sont négligeables). En revanche, les images naturelles, comme les images de visage, sont décomposées en des sommes faites de plusieurs milliers de grilles sinusoïdales ayant une amplitude non nulle. L'analyse de Fourier est l'opération mathématique qui permet d'effectuer ces décompositions. Pour la suite de cette section, nous présenterons les études qui ont porté sur le rôle des fréquences spatiales dans la reconnaissance des visages.

-----  
Figure 2  
-----

### ***Les fréquences spatiales***

Dans une grille sinusoïdale, la largeur des barres est relative à la fréquence spatiale : les basses fréquences spatiales sont représentées par de larges barres alors que les hautes fréquences spatiales sont représentées par des barres plus minces. Plus précisément, la notion de *fréquences spatiales* réfère au nombre de cycles (i.e., une barre claire suivit d'une barre foncée) par degré d'angle visuel dans une grille sinusoïdale (voir figure 2a et 2b). Il est habituellement supposé que l'information pertinente à la reconnaissance des visages est invariante à la taille des visages (e.g., Hayes, Morrone & Burr, 1986); l'unité utilisé pour les fréquences spatiales dans ce domaine est donc le nombre de cycles par (la largeur de) visages.

En 1968, Campbell et Robson ont rapporté que la détection ainsi que la discrimination de sommes de quelques grilles sinusoïdales ou de grille rectangulaire (qui se décompose en une somme de toutes les grilles sinusoïdales de fréquences spatiales impaires de contraste décroissant proportionnellement à la fréquence et de phase égale à zéro) pouvaient être expliquées par le contraste de leurs fréquences spatiales. Par la suite, ce même groupe de chercheurs a montré qu'une adaptation à une fréquence spatiale spécifique altérait momentanément la sensibilité au contraste de grilles sinusoïdales ayant des fréquences spatiales similaires à celle de la grille d'adaptation, suggérant que le système visuel primaire réalise une opération qui s'apparente à une analyse de Fourier sur les images naturelles (e.g., Blackmore & Campbell, 1969). Aujourd'hui la plupart des chercheurs s'entendent pour dire que le système visuel primaire effectue une analyse de Fourier locale à l'intérieur de champs récepteurs d'une taille moyenne d'environ 0,5 degré d'angle visuel (DeValois & De Valois, 1988; Westheimer, 2001). Aux quatre paramètres des grilles sinusoïdales mentionnées ci-dessus, il faut donc ajouter la position sur la rétine.

Nous avons décrit, pour l'instant, des expériences portant sur les fréquences spatiales de grilles sinusoïdales extrêmement simples. Qu'en est-il des images naturelles comme des photographies de visages? Il est coutumier en psychologie de parler des basses fréquences spatiales comme reflétant les aspects globaux d'une image (voir figure 1b) alors que les hautes fréquences spatiales donneraient davantage d'informations sur les détails, les parties constituantes (voir figure 1d). En fait, une analyse de Fourier d'un visage sur ses fréquences spatiales constituantes (toutes orientations confondues) montre que l'information présente dans un visage est concentrée dans les basses fréquences alors que l'énergie s'avère de plus en plus basse au fur et à mesure que l'on grimpe dans les hautes fréquences spatiales (la distribution de l'énergie dans les fréquences

spatiales suivant une courbe de type  $1/\text{fréquence spatiale}^2$  ou, plus succinctement,  $1/f^2$ ). Cette distribution en termes d'énergie suggère qu'un système idéal focaliserait son attention sur les basses fréquences spatiales afin de maximiser son efficacité. Il est toutefois important de noter ici que le système visuel humain n'est pas optimal au moins en ce qui a trait à la perception de grilles sinusoïdales. En effet, les études utilisant la perception de grilles sinusoïdales ont montré qu'avec un éclairage ambiant typique, le système visuel humain particulièrement adapté à des fréquences spatiales se situant au autour de 7 cycles par degré d'angle visuel, c'est-à-dire des fréquences spatiales « moyennes ». Il est toutefois rare de rencontrer des grilles sinusoïdales alors que les visages constituent l'une des catégories d'objets les plus fréquemment rencontrés. On peut donc se demander jusqu'où les connaissances acquises en discrimination de grilles sinusoïdales peuvent servir de base pour notre compréhension des mécanismes visuels à la base de la perception d'objets complexes tels que des visages.

Il semble que les visages, en comparaison aux autres types d'objets, soient particulièrement vulnérables aux variations de contenu fréquentiel (Biederman & Kalocsai, 1997). Il est remarquable, par exemple, qu'un dessin d'objet dessiné à partir de traits fins (i.e., hautes fréquences) soit relativement facile à reconnaître alors qu'un visage dessiné de la même manière soit difficilement reconnaissable. Ces observations suggèrent que la reconnaissance des visages serait basée sur une quantité restreinte de fréquences spatiales alors que la reconnaissance d'objets serait plus souple quant à l'information pouvant être utilisée. Plusieurs études ont donc tenté de trouver la gamme de fréquences spatiales utilisées par l'être humain pour reconnaître les visages. Nous décrirons brièvement les différentes techniques utilisées pour étudier le rôle des fréquences spatiales en reconnaissance de visages avant de résumer les résultats des différentes études.

La première méthode proposée afin d'étudier le rôle des fréquences spatiales en reconnaissance de visages consiste à créer un stimulus en superposant une grille à une image et en moyennant les valeurs de luminance des différentes régions (ou pixels) se trouvant dans chacun des carrés de la grille. Cette méthode a mené à quelques études intéressantes dans le domaine de la reconnaissance des visages (Bachmann, 1991; Costen, Parker & Craw, 1994; Harmon & Julesz, 1973; Tieger & Ganz, 1979). Il est toutefois important de noter qu'elle n'est pas sans problème. Par exemple, des hautes fréquences spatiales sont introduites à la jonction de deux carrés de la grille modifiant donc la composition fréquentielle du stimulus et il est possible que ces fréquences ajoutées artificiellement interfèrent avec le traitement normal du stimulus. De plus, il semble que cette méthode induise une difficulté à situer le contour du visage, altérant donc la performance pour des raisons n'ayant rien à voir avec le contenu fréquentiel (Morrone, Burr et Ross, 1983 – mais voir Smilek, Rampel, & Enns, 2006, pour une interprétation plus récente). Aujourd'hui, cette méthode n'est pratiquement plus utilisée.

La technique la plus fréquemment utilisée afin d'étudier le rôle des fréquences spatiales en vision consiste à filtrer l'image originale de sorte que seulement une partie des fréquences spatiales de départ soit conservée. Trois sortes de filtrages sont habituellement effectués : le filtrage passe-bas, le filtrage passe-haut et le filtrage passe-bande. Le filtrage passe-bas laisse passer uniquement les fréquences spatiales se situant en dessous d'une borne choisie (e.g.,  $<6$  cycles par visage). Ce type de filtrage produit une image plus floue que l'original (voir figure 1b). Le filtrage passe-haut laisse passer uniquement les fréquences spatiales se situant au-dessus d'une borne choisie (e.g.,  $>20$  cycles par visage). Ce type de filtrage produit une image où les contours des objets sont définis de façon plus abruptes que dans l'image originale (voir figure 1d). Enfin,

le filtrage passe-bande laisse passer l'information spatiale entre deux bornes (p. ex. 6-20 cycles par visage; voir figure 1c). Ici, l'apparence du stimulus dépend entièrement des seuils choisis : un filtre passe-bande entre 2 et 6 cycles par visage donnera un visage plus flou que l'original, un filtre passe-bande avec des bornes à 6 et 32 cycles par visages changera peu la perception du visage alors qu'un filtre passe-bande entre 32 et 64 cycles par visage soulignera les contours du visage comme un filtre passe-haut.

Une troisième et dernière méthode—le *masquage de la bande critique*—a été utilisée pour la première fois par Pelli et Solomon (1994) pour étudier la reconnaissance des lettres. Cette méthode consiste à mesurer le seuil de sujets humains dans une tâche de discrimination en utilisant du bruit blanc filtré à l'aide d'un filtre passe-haut (ou passe-bas, ceci n'a pas d'importance) à plusieurs bornes de fréquences spatiales. La théorie sous-jacente est relativement simple : si le bruit ajouté au signal contient des fréquences spatiales critiques pour la tâche en cours, ce bruit aura un impact important sur le seuil perceptif; au contraire, si le bruit ne contient peu ou pas d'information fréquentielle importante, le seuil ne sera que peu ou pas touché par le bruit. Les courbes décrites par les seuils obtenus pour les différents bruits passe-hauts servent à produire une estimation de la bande de fréquences spatiales dont les sujets se servent le plus pour faire la tâche.

Les trois méthodes que nous venons de décrire ne constituent en rien une liste exhaustive des techniques pouvant être utilisés pour l'étude du rôle des fréquences spatiales en reconnaissance des visages. Ce sont toutefois les techniques les plus fréquemment utilisées dans le domaine. À partir de ces méthodes, de nombreux chercheurs ont montré qu'une étendue limitée de fréquences

spatiales centrée autour de 10-12 cycles par visage était sous-jacente à la reconnaissance des visages. La section suivante résumera ces études.

### ***Rôle des fréquences spatiales en reconnaissance de visages.***

À notre connaissance, la première étude à avoir explicitement étudié le rôle des fréquences spatiales en reconnaissance de visages nous vient de Tieger et Ganz (1979). Ces auteurs ont superposé des grilles composées de différentes bandes de fréquences spatiales à des visages, et ont mesuré la performance de participants dans une tâche de reconnaissance de visages. Puisque la performance était inférieure lorsque la grille masquait l'information autour de 17.6 cycles par visage, les auteurs ont conclu que cette information était la plus importante pour l'identification des visages. Il est important de noter toutefois que les différentes grilles sinusoïdales utilisées dans cette étude échantillonnaient un nombre limité de fréquences spatiales ( $n = 4$ ), et ne permettaient donc pas une estimation précise de la bande de fréquences spatiales critique pour la reconnaissance des visages.

Fiorentini, Maffei et Sandini (1983) ont étudié l'identification de visages filtrés passe-bas et passe-haut en utilisant un filtre carré leur permettant d'inclure uniquement l'information visuelle en-dessous et au-dessus de la borne mais produisant beaucoup de « halos » (cf. filtre gaussien ou Butterworth). En utilisant un filtre passe-bas, ils ont montré que la performance était inférieure quand l'information fréquentielle conservée se situait sous 5 cycles par visages que sous 8 cycles par visages. Ensuite, les auteurs ont comparé la performance lors d'une condition où les visages étaient filtrés passe-haut et passe-bas à 8 cycles par visages. La performance était similaire dans ces deux conditions et ce même si l'énergie était de beaucoup supérieure avec le filtre passe-bas suggérant ainsi que les fréquences spatiales au-dessus de 8 cycles par visages sont

particulièrement informatives pour l'identification des visages. Enfin, puisque les stimuli étaient trop éloignés pour que l'information fréquentielle se situant au-dessus de 15 cycles par visages soit visible à l'œil humain, leurs résultats nous informent donc que sur la borne inférieure qui se situe à 5 cycles par visage et non sur la borne supérieure qui est d'au moins 15 cycles par visages mais qui pourrait être supérieure. D'autres études sont donc nécessaires afin d'objectiver la borne supérieure permettant la reconnaissance des visages.

Hayes, Morrone et Burr (1986) ont également étudié le rôle des fréquences spatiales en reconnaissance de visage mais cette fois en utilisant des filtres passe-bande carrés de 1.5 octaves (1 octave = 1 doublement de la fréquence spatiale; une bande de fréquences spatiales de 1.5 octave s'étend donc d'une borne  $x$  à une borne  $x * 2^{1.5}$ ) à différentes fréquences spatiales. Dans cette étude, la variable dépendante était la proportion de reconnaissance correcte en fonction de la bande de fréquences spatiales révélées. Selon ces auteurs, la bande optimale pour la reconnaissance des visages se situerait à environ 20 cycles par visage et serait invariante à la taille du stimulus sur la rétine. En effet, plusieurs distances ont été testées dans cette étude et des résultats similaires ont été retrouvés à chaque fois. Cette étude justifie l'utilisation de l'unité « cycles par visage » pour les fréquences spatiales.

Costen, Parker et Craw (1996) ont aussi étudié la performance de sujets en reconnaissance de visages grâce à l'utilisation de filtre passe-bas et passe-haut à plusieurs bornes distinctes. Ils trouvent une baisse rapide des performances lorsque les stimuli sont filtrés passe-bas en dessous de 8 cycles par visages ou filtré passe-haut au-dessus de 16 cycles par visages. Ces résultats suggèrent que l'identification des visages est dépendante d'une bande de fréquences spatiales se situant approximativement entre 8 et 16 cycles par visage avec une moyenne harmonique se

situant à 11.3 cycles par visages. Gold, Bennett et Sekuler (1999) ont également étudié l'information fréquentielle efficace pour la reconnaissance de visages en présentant à 2 sujets des visages filtrés passe-bande à différents seuils situés entre un et soixante-dix cycles par objet. Ils montrent dans cette étude que les visages sont impossibles à reconnaître pour toutes les conditions sauf lorsque le visage est filtré passe-bande à 8 et 17.5 cycles par objets.

Récemment, Näsänen (1999) a obtenu des résultats similaires à ceux de Costen et al. (1996) en utilisant trois méthodes de filtrage distinctes. Dans les deux premières expériences, il mesura le seuil signal/bruit en ajoutant aux stimuli du bruit blanc filtré passe-bande à différentes bornes. Dans la troisième expérience, l'information de phase de bandes spécifiques de fréquence spatiale fut remplacée par des nombres aléatoires. Des seuils furent mesurés pour chacune des bandes remplacées. Dans la dernière expérience, Näsänen a filtré passe-bande le stimulus au moyen d'un filtre Gaussien de 2 octaves pour ensuite mesurer un seuil pour chacune des paires de bornes sélectionnées. À partir de ces différentes méthodes, il propose que la bande critique pour la reconnaissance des visages se situe entre 8 et 13 cycles par visages (voir Tanskanen et al., 2005 pour des résultats similaires mais cette fois en utilisant la M170 comme variable dépendante plutôt que la performance). Puisque des simulations informatiques montrent que de nombreuses fréquences spatiales contiennent de l'information efficace pour la reconnaissance des visages, Näsänen conclut que le *tuning* observé provient de contraintes inhérentes au système visuel humain et non des propriétés de la tâche en soi.

Une bande de fréquences spatiales relativement fine serait donc critique pour l'identification correcte et rapide des visages. La raison demeure toutefois inconnue. Il est par exemple possible que les régions visuelles spécialisées dans la reconnaissance des visages utilisent sélectivement

cette information. Il est également possible que cette bande de fréquences spatiales soit nécessaire afin d'encoder la configuration du visage nécessaire à sa reconnaissance (Boutet, Collin, & Faubert, 2003). Il est enfin possible que cette bande constitue un excellent compromis entre l'utilisation des basses fréquences nécessaires à la reconnaissance rapide du visage et les hautes fréquences idéales pour sa reconnaissance adéquate. Si cette dernière possibilité s'avérait exacte, il serait malheureux que la découverte de l'importance des fréquences moyennes mène à négliger le rôle des basses et des hautes fréquences spatiales en reconnaissance de visages.

En résumé, plusieurs chercheurs ont tenté, au cours de la dernière décennie, d'évaluer quelle bande de fréquences spatiales était optimale pour la reconnaissance. Malgré certaines divergences dans leurs résultats, il semble qu'une bande de fréquences moyennes se situant entre 8 et 20 cycles par visage est optimale pour reconnaître l'identité d'un visage.

### ***Utilisation flexible des fréquences spatiales***

La tâche (i.e., le type de catégorisation) effectuée par l'observateur influence-t-elle la prise d'information à un aussi bas niveau que les fréquences spatiales ou sommes-nous pré câblés de sorte que l'information efficace demeure toujours la même? Oliva et Schyns (1997; Schyns et Oliva, 1999) ont proposé que le système visuel serait très flexible de sorte que l'information fréquentielle efficace change selon le type de catégorisation demandé par la tâche (e.g., jugement d'identité, de sexe et d'émotions, expressif ou non). Afin de tester la flexibilité dans le domaine de la reconnaissance des visages, Schyns et Oliva (1999) ont demandé à des sujets de catégoriser des stimuli *hybrides*.

C'est en utilisant des filtres passe-bas et passe-haut sur deux images différentes et en les combinant qu'on produit des stimuli *hybrides*; la perception des *hybrides* change selon la distance à laquelle se trouve l'observateur. Dans notre exemple (figure 3), nous avons utilisé une photographie de Brad Pitt en basses fréquences spatiales ainsi qu'une photographie de Brandon Fraser en hautes fréquences spatiales. À une distance de lecture (i.e., 30-40 cm), vous devriez identifier le visage du stimulus de gauche de la figure 3 comme étant Brandon Fraser. Par contre, le même stimulus mais en plus petit (figure 3, stimulus de droite) devrait être identifié comme étant Brad Pitt. Si vous voulez vous convaincre que les deux images de la figure 3 sont identiques—sauf pour la taille—éloignez-vous à une distance de un à deux mètres et le stimulus de gauche devrait se 'transformer' en Brad Pitt. Ce type de stimuli peut être utilisé afin de caractériser le rôle des fréquences spatiales ainsi que leurs décours temporels en reconnaissance d'objets et de visages. (En passant, cette démonstration contredit l'idée voulant que les fréquences spatiales utilisées dans les visages soient invariantes à la taille des visages.)

-----  
Figure 3  
-----

Schyns et Oliva ont demandé à un premier groupe de sujets de catégoriser chaque hybride selon son expression (i.e., joie, neutre ou colère) et à un second groupe de décider si le stimulus était expressif ou non (i.e., neutre). Ils avaient déjà montré, dans une étude précédente, que les participants utilisent les hautes fréquences spatiales pour catégoriser l'expression d'un visage, alors qu'ils utilisent les basses fréquences spatiales pour déterminer si un visage est expressif ou non. Ainsi, le premier groupe devait être biaisé vers le traitement des hautes fréquences spatiales

alors que le second groupe devait plutôt favoriser les basses fréquences spatiales. Dans un second temps, chacun des groupes devait juger le genre d'un même ensemble d'hybride présenté pendant 30 ms. Tel que prédit par la théorie de l'usage flexible, les sujets du premier groupe catégorisaient l'hybride selon l'information présente dans les hautes fréquences spatiales alors que les sujets du deuxième groupe les catégorisaient selon l'information présente dans les basses fréquences spatiales. Ces résultats ont permis de démontrer une certaine flexibilité du système visuel et l'influence de la tâche en ce qui a trait à l'information efficace.

### ***L'orientation et son rôle en reconnaissance de visages***

Notre sensibilité à l'orientation de grilles sinusoïdales est bien connue depuis longtemps (Campbell, Kulikowski & Levinson, 1966). Notre système visuel est plus sensible aux grilles verticales et horizontales qu'aux grilles obliques. Notre sensibilité aux fréquences spatiales et à l'orientation pour plusieurs classes de stimuli visuels incluant les scènes naturelles (mais excluant les visages) a récemment été modélisé par Watson et Ahumada (2005).

Le rôle de l'orientation de l'information en reconnaissance des visages n'a été à peu près pas étudié. Il ne faut pas confondre « orientation » avec « inversion des visages » : cette dernière consiste à appliquer une rotation de 180 deg aux images de visages—l'inversion est omniprésente dans la littérature sur les visages—qui n'affecte pas le contenu orientationnel d'une image parce qu'une grille sinusoïdale d'orientation  $\theta$  deg a la même orientation qu'une autre d'orientation  $\theta+180$  deg (une rotation de 180 deg n'affecte pas, non plus, le contenu fréquentiel d'une l'image—i.e., une grille d'une fréquence  $f$  possède toujours une fréquence  $f$  après rotation— mais il affecte son contenu phasique—e.g., une grille d'une fréquence entière devient une grille complémentaire après une rotation de 180 deg).

-----  
Figure 4  
-----

Valérie Goffaux (communication personnelle) a réalisé la seule expérience à notre connaissance sur le rôle de l'orientation en reconnaissance de visages : elle a découvert que les grilles orientées verticalement sont plus efficaces que les grilles orientées horizontalement (voir figure 4). Cette pauvreté expérimentale est surprenante si on considère qu'il est aussi simple de faire varier le contenu orientationnel que le contenu fréquentiel des stimuli visuels.

### **La méthode « Bubbles »**

Les études décrites jusqu'ici sur l'utilisation des fréquences spatiales et de l'orientation sont limitées par rapport à ce qu'elles peuvent nous apprendre sur la reconnaissance d'objets complexes. La position sur la rétine (ou sur l'objet) semble au moins aussi importante; elle peut nous permettre d'inférer les parties ou attributs nécessaires à la reconnaissance. Par exemple, il est évident que certaines parties de lettres, comme la barre horizontale inférieure du E majuscule, sont particulièrement importantes afin de discriminer des lettres visuellement similaires comme le E et le F (Fiset et al., soumis).

Afin de tirer avantage de l'organisation locale (i.e., en champs récepteurs de moins de 0,5 degré d'angle visuel) du système visuel et de mieux comprendre comment un système de ce type reconnaît des objets complexes comme des visages, des objets, des lettres ou des mots, Gosselin et Schyns ont proposé en 2001 une nouvelle méthode—*Bubbles*—pour révéler les régions d'un

stimulus ayant un impact sur la performance en catégorisation. De façon générale, la méthode consiste à présenter aux participants des stimuli partiellement échantillonnés. À chaque essai, les parties révélées sont aléatoirement choisies, de sorte qu'après un nombre relativement élevé d'essais, tout l'espace de recherche est échantillonné. Le concept d'espace de recherche ne se limite pas aux coordonnées cartésiennes  $x$  et  $y$  tel que suggéré dans l'exemple des lettres. En effet, l'idée d'espace de recherche s'étend à tout espace fini tels que, par exemple, les fréquences sonores, les fréquences spatiales, la phase, l'orientation dans l'espace, etc. La méthode *Bubbles* est donc générale et peut être appliquée tant dans les domaines de la vision que de l'audition, par exemple. Chaque élément de cet espace est ensuite corrélé avec la ou les réponses d'intérêt.

L'analyse des données d'une expérience *Bubbles* repose habituellement sur le postulat qu'il existe une corrélation entre la ou les réponses des sujets (la ou les variables d'intérêt—proportion de bonnes réponses, temps de réaction, amplitude d'une onde cérébrale, etc.) et les régions échantillonnées à chaque essai (variables explicatives). Une régression linéaire multiple est donc effectuée sur les masques de bulles et sur la ou les réponses du sujet donnant lieu, dans le cas de l'espace d'échantillonnage illustré à la figure 5, à cinq plans de coefficients de régression—aussi nommés images de classification—indiquant l'efficacité des régions pour la tâche à l'étude. Chauvin et al. (2005) décrivent deux tests statistiques adaptés aux images de classification.

Les avantages de la méthode *Bubbles* sur les méthodes plus classiques discutées plus haut sont :

- (1) La méthode ne requiert peu ou pas d'*a priori* de l'expérimentateur en ce qui a trait à l'information importante—en termes de fréquences spatiales, d'orientation ou de position sur les axes  $x$  et  $y$ —en reconnaissance de visages. Par exemple, pour des bulles dans le domaine de l'espace où les axes  $x$  et  $y$  sont échantillonnés, un nombre suffisant d'essais permet d'évaluer

l'importance relative de chacune des aires locales du visage ainsi que les interactions susceptibles de se produire entre ces différentes régions. Ainsi la méthode permet non seulement d'extraire l'information locale efficace pour la tâche mais également quelle(s) combinaison(s) maximise(nt) la performance des sujets. (2) Puisque l'échantillonnage est aléatoire d'un essai à l'autre, et donc imprédictible, les sujets se voient dans l'impossibilité d'adapter leur stratégie aux spécificités de la tâche.

-----  
Figure 5  
-----

Gosselin et Schyns (2001; voir également Schyns, Bonnar & Gosselin, 2002) ont démontré l'applicabilité de la méthode *Bubbles* en reconnaissance de visages. Pour générer un stimulus, une image est initialement décomposée en six bandes de fréquences spatiales contenant chacune une octave. Par exemple, une image est décomposée en six bandes dont les seuils sont 128-64, 64-32, 32-16, 16-8, 8-4 ainsi que 4-2 cycles par image (voir figure 5; la dernière bande allant de 4-2 n'est pas échantillonnée par des bulles et est présentée comme fond pour l'ensemble des stimuli présentés). Dans ce cas particulier, chacune des échelles spatiales de la plus fine à la plus grossière est échantillonnée aléatoirement par des bulles laissant transparaître un certain nombre de cycle, habituellement trois. Qui plus est, la surface révélée en moyenne à chaque bande de fréquence était maintenue constante. Les bulles sont donc plus petites et plus nombreuses pour les bandes de hautes fréquences et plus grosses et moins nombreuses pour les bandes de basses fréquences spatiales (figure 5). Puisque les bulles sont positionnées aléatoirement sur l'image,

l'ensemble de l'espace de recherche pour chacune des bandes de fréquences spatiales est échantillonné sans biais.

Schyns et al. (2002; voir également Gosselin & Schyns, 2001) ont utilisé cet espace de recherche pour révéler les régions d'un visage utiles pour la reconnaissance de l'identité, du sexe et de l'expression faciale (i.e., heureux vs. neutre). Afin de vérifier quelle bande de fréquences spatiales étaient la plus efficace pour la reconnaissance de visage, Schyns et al. ont calculé la proportion de l'aire du visage efficace pour chacune des cinq bandes de fréquences échantillonnées. En accord avec les résultats rapportés précédemment, la bande allant de 11.25 à 22.5 cycles par visages s'est avérée la plus importante pour la reconnaissance d'identité suivie de près par la bande allant de 5.62 à 11.25 cycles par visages (voir colonne de droite de la figure 6).

-----  
Figure 6  
-----

La figure 6 illustre les régions utiles pour la reconnaissance de l'identité d'un visage, de son genre ainsi que pour une tâche de type expressif ou non. Les résultats pour la tâche d'identité révèlent qu'une grande étendue du visage est efficace pour cette tâche. Une analyse plus fine ainsi que d'autres données obtenues dans les laboratoires de Frédéric Gosselin et de Philippe Schyns montrent l'importance capitale de la région des yeux pour la reconnaissance des visages. Ce résultat est en accord avec l'observation par Caldara et al. (2005) qu'une patiente atteinte d'un déficit sévère en reconnaissance de visage suite à une lésion cérébrale (i.e., patiente prosopagnosique, voir chapitre XX) a de la difficulté à traiter l'information dans la région des

yeux des visages. Les résultats chez le sujet normal et pathologique sont donc consistants avec le rôle important de la région des yeux afin de discriminer les visages entre eux. La région des yeux n'est pas uniquement nécessaire à la reconnaissance de l'identité faciale mais permet également de catégoriser le visage selon son sexe.

Une étude effectuée par Smith, Cottrell, Gosselin et Schyns (2005) a révélé les parties d'un visage qui sont utiles à la reconnaissance des six expressions faciales de base (plus neutre) c'est-à-dire la peur, la joie, le dégoût, la tristesse, la colère et la surprise (Ekman & Friesen, 1975; Izard, 1971; 1994). Les auteurs ont sélectionné 10 identités (dont cinq femmes) dans la *California Facial Expression Database* (Daily, Cottrell & Raily, 2001). Chacun des 14 participants de l'étude a dû reconnaître l'expression faciale de l'un des 70 stimuli de départ présenté aléatoirement (pour un total de 8400 essais par participants) et révélés à travers un masque de bulles échantillonnant les deux dimensions de l'image ainsi que les cinq bandes de fréquences spatiales (comme dans Schyns et al., 2002; voir figure 5).

---

Figure 7

---

Les résultats de Smith et al. (2005; voir figure 7) indiquent que différentes régions du visages sont nécessaires à la reconnaissance de différentes émotions exprimées facialement. En fait, si on parcourait le visage de haut en bas, on reconnaîtrait d'abord la tristesse : les éléments efficaces pour cette émotion sont le plissement du front représenté en hautes fréquences spatiales ainsi que les yeux. Ensuite, nous reconnaîtrions la peur grâce à la région des yeux en hautes fréquences

spatiales. Whalen et al. (2004) ont étudié l'importance des yeux pour l'activation de l'amygdale—une région du cerveau impliquée dans le traitement des émotions, tout particulièrement de la peur (Adolphs et al., 1994) : Whalen et al. ont éliminé toute l'information du visage pour ne conserver que les yeux et ils ont trouvé une corrélation positive entre l'activité hémodynamique de l'amygdale et l'ouverture des yeux. Adolphs et al. (2005) ont montré, en utilisant la méthode *Bubbles*, que les régions d'un visage utilisées par SM, une patiente ayant une lésion complète mais spécifique des amygdales, pour reconnaître—non sans difficulté—l'expression faciale de peur. SM, contrairement aux participants contrôles, n'utilise pas la région des yeux pour reconnaître la peur. Qui plus est, l'enregistrement des mouvements oculaires de SM montre qu'elle ne regarde pas et n'utilisait pas les yeux. Cependant, la performance de SM en reconnaissance de la peur devient normale lorsque les expérimentateurs lui demandent de regarder les yeux. Une lésion de l'amygdale semble donc créer une négligence des attributs nécessaires pour reconnaître la peur sans toutefois créer une incapacité à utiliser ces mêmes attributs.

Selon Smith et al. (2005), les yeux—et les sourcils—sont également impliqués dans la reconnaissance de la colère mais en basses fréquences spatiales. Il s'avère que SM, cette patiente qui ne semble pas porter attention aux yeux, éprouve de la difficulté à reconnaître la tristesse et la colère—cependant, son déficit est moindre pour la reconnaissance de la tristesse et de la colère que pour la reconnaissance de la peur. L'un de nous a récemment évalué la performance de la patiente prosopagnosique PS dans une tâche similaire à celle de Smith et al. (Fiset et al., en préparation). Or PS—comme SM—a de la difficulté à reconnaître la tristesse, la peur et la colère (voir aussi Rossion et al., 2003), ce qui suggère que son déficit n'est pas spécifique à la reconnaissance de l'identité des visages (cf., Caldara et al., 2005) mais touche également de

toutes les tâches perceptuelles nécessitant un traitement du haut du visage et, plus spécifiquement, des yeux. Contrairement à SM, PS ne semble pas bénéficier d'instructions lui demandant explicitement de regarder les yeux; il semble que son problème se situe au niveau de la représentation interne des visages plutôt qu'au niveau attentionnel (Fiset et al., en préparation).

Si on continuait notre exploration du visage vers les attributs inférieurs, on saurait bientôt reconnaître le dégoût, c'est-à-dire qu'on aurait extrait l'information à l'intérieur d'un 'U' comprenant la bouche et se terminant de part et d'autre de la base du nez. Puisque cette expression faciale produit une activation dans deux aires neuro-anatomiques (i.e., putamen et insula; Calder et al., 2000) distinctes de celle activée par la peur (i.e., amygdale), il serait intéressant de tester si une lésion de ces aires telle qu'observée chez les patients souffrant de la maladie de Huntington (Sprengelmeyer, 1996) entraîne une négligence du 'U' naso-bucal des visages (cf. SM) ou entraîne une détérioration de la représentation en mémoire du stimulus distal (cf. PS). Pour terminer notre exploration des résultats de Smith et al. (2005), la région de la bouche nous permettrait de reconnaître deux expressions « positives » : la surprise et la joie.

## **Conclusion**

Dans ce chapitre, nous avons décrit la théorie en vision qui propose que la première étape de traitement des visages et des objets en général, revient à effectuer une analyse de Fourier locale. Ensuite, nous avons passé en revue les différentes études ayant examinés quelles fréquences spatiales sont nécessaires à la reconnaissance des visages. Nous avons enfin résumé un ensemble d'études effectuées dans notre laboratoire avec la méthode *Bubbles* qui permet de révéler plus précisément quels attributs des visages sont efficaces selon la tâche demandée aux participants.

Ainsi, grâce à la méthode *Bubbles*, nous avons précisé quelles régions d'un visage devaient être traitées afin de répondre efficacement à une tâche donnée. Nous n'avons cependant rien écrit à propos de la séquence dans laquelle ces différentes régions des visages sont extraites. Ce qui a été présenté dans ce chapitre ne permet pas de s'attaquer directement à cette question; en faisant quelques suppositions « raisonnables », cependant, il est possible de le faire. Les prochaines lignes décriront une analyse que nous avons effectuée à partir des résultats obtenus pour de nombreuses tâches (i.e. identité, genre, reconnaissance des sept émotions faciales) rapportées dans Smith et al., (2005) et Schyns et al (2002).

Pour simplifier le problème, nous avons sommé les coefficients de régression pour les cinq bandes de fréquences spatiales pour obtenir une seule image de classification par catégorie. Les visages utilisés dans ces deux articles étant de tailles différentes, nous avons aligné les images de classification pour que les attributs faciaux (i.e., les yeux, le nez, la bouche et les sourcils) soient situés aux mêmes endroits. Nous avons ensuite z-scoreé les images de classification tel que décrit dans Chauvin et al. (2005). Chaque pixel d'une image de classification est donc associé à un z-score indiquant l'efficacité de ce pixel pour la tâche effectuée par les sujets. Nous avons ensuite classé en ordre décroissant les z-scores pour chacune des images de classification et compté le nombre minimal de z-scores nécessaires pour atteindre un seuil arbitraire de 3000 (nos résultats ne changent à peu près pas avec d'autres seuils). Nous avons découvert que la catégorie « colère » n'a besoin que de 437 z-scores pour atteindre une somme de 3000; la peur, de 537; la surprise, de 560; le genre, de 599; la joie, de 612; l'identité, de 909; la neutralité, de 923; la tristesse, de 1204; et le dégoût, de 1245. Ceci suggère, par exemple, qu'il est plus facile de placer un visage dans la catégorie « colère » que dans la catégorie « tristesse ».

Nous supposons que le but d'un observateur humain confronté à un visage est de catégoriser ce visage le plus rapidement possible (Rosch, 1978). Un tel observateur humain a donc tout avantage à vérifier l'appartenance du visage à la catégorie « colère ». En balayant les régions du visage efficaces pour la reconnaissance de la colère, l'observateur aura peut-être compilé de l'information « efficaces » pour placer le visage dans d'autres catégories; ces régions n'ont pas besoin d'être revisitées. Il découle que l'ordre de catégorisation prescrit par le nombre de z-scores nécessaires pour atteindre un seuil de 3000 n'est pas la fin de l'histoire. Puisque la région des yeux, explorées pour la catégorisation de la colère, est essentielle pour la reconnaissance de la peur, il n'est pas surprenant que cette catégorie demeure deuxième. Il est intéressant, sur le plan évolutif, que les catégories les plus « négatives » occupent les deux premières positions. Les yeux étant efficace pour la peur et, dans une moindre mesure, pour la colère, les autres catégories aussi définies par ces régions seront avantagées pour la suite. Le sexe, l'identité et la neutralité suivent donc, dans cet ordre. Viennent ensuite la joie puis la surprise, toutes les deux définies par la bouche. Suivent la tristesse puis le dégoût, toutes les deux étant définies par des régions uniques des visages.

Cette analyse nous permet de mieux comprendre l'importance des yeux pour la reconnaissance des visages toutes tâches confondues (e.g., identité, sexe, expressions faciales, dialogue; e.g., Spezio, Huang, Castelli et Adolphs, 2007). Elle permet aussi de mieux comprendre les résultats de Schyns, Petro et Smith (2007) qui ont montré que l'intégration de l'information visuelle dans les visages consistait en un balayage débutant avec les yeux et se terminant avec la bouche en combinant la méthode *Bubbles* et l'électrophysiologie. Notre analyse suggère que l'importance des yeux pour des émotions faciales nécessaires à la survie de l'individu ainsi que leur *efficacité*

pour de nombreuses tâches pourraient avoir influencés les stratégies développées par le cerveau dans le traitement visuel des visages.

## RÉFÉRENCES

Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P. G. & Damasio, A. R. (2005).

A mechanism for impaired fear recognition after amygdala damage. *Nature*, 433, 68-72.

Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1994). Impaired recognition of emotion in

facial expressions following bilateral damage to the human amygdala. *Nature*, 372, 669-672.

Bachmann, T. (1991). Identification of spatially quantised tachistoscopic images of faces: how many pixels does it take to carry identity? Special issue: face recognition. *European Journal of Cognitive Psychology*, 3 (1), 87–103.

Biederman, I., & Kalocsai, P. (1997). Neurocomputational bases of object and face recognition.

*Philosophical Transactions of the Royal Society London: Biological Sciences*, 352, 1203-1219.

Blakemore, C., & Campbell, F. W. (1969). On the existence of neurons in the human visual

system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology (London)*, 203, 237-260.

Boutet, I., Collin, C., & Faubert, J. (2003). Configural face encoding and spatial frequency information. *Perception & Psychophysics*, 65, 1078–1093

Caldara, R., Rossion, B., Mayer, E., Smith, M., Gosselin, F. & Schyns, P. G. (2005). Does prosopagnosia take the eyes out from faces? Evidence for a defect in the use of diagnostic facial information in a single-case. *Journal of Cognitive Neuroscience*, *17*, 1-15

Calder, A. J., Keane, J., Manes, F., Antoun, N. & Young, A. W. (2000). Impaired recognition and experience of disgust following brain injury. *Nature Neuroscience*, *3*, 1077-1078.

Campbell, F. W., Kulikowski, J. J., & Levinson, J. Z. (1966). The effect of orientation on the visual resolution of gratings. *Journal of Physiology*, *187*, 427-436.

Campbell, F.W. & Robson, J.G. (1968). Application of the Fourier analysis to the visibility of gratings. *Journal of Physiology*, **197**, 551-556.

Chauvin, A., Worsley, K. J., Schyns, P. G., Arguin, M., & Gosselin, F. (2005). Accurate statistical tests for smooth classification images. *Journal of Vision*, *5*, 659-667, <http://journalofvision.org/5/9/1/>, doi:10.1167/5.9.1.

Costen, N. P., Parker, D. M., & Craw, I. (1994). Spatial content and spatial quantisation effects in face recognition. *Perception*, *23* (2), 129-146.

Costen, N. P., Parker, D. M., & Craw, I. (1996). Effects of high-pass and low-pass spatial filtering on face identification. *Perception and Psychophysics*, *58* (4), 602-612.

Dailey, M., Cottrell, G.W., & Reilly, J. (2001). California facial expressions, CAFE. Unpublished digital images, University of California, San Diego, Computer Science and Engineering Department.

DeValois, R.L. & DeValois, K.K. (1990). *Spatial Vision*. Oxford University Press: New York.

Ekman, P., & Friesen, W.V. (1975). *Unmasking the face*. Englewood Cliffs, NJ: Prentice Hall.

Fiset, D., Blais, C., Éthier-Majcher, C., Arguin, M., Bub, D. & Gosselin, F. (submitted). Features for uppercase and lowercase letter identification.

Fiset, D., Meyer, E., Blais, C., Scheepers, C., Schyns, P., & Caldara, R. (en préparation). Clarifying the nature of facial identity and facial expression representations with an acquired case of prosopagnosia.

Fiorentini, A., Maffei, L. & Sandini, G. (1983). The role of high spatial frequencies in face perception. *Perception*, **12**, 195-201.

Gibson, B., Lazareva, O. F., Gosselin, F., Schyns, P. G. & Wasserman, E. A. (2007). Non-accidental properties underlie shape recognition in mammalian and non-mammalian vision. *Current Biology*, *17*, 336-340.

Gibson, B. M., Wasserman, E. A., Gosselin, F. & Schyns, P. G. (2005). Applying Bubbles to localize features that control pigeons' visual discrimination behavior. *Journal of Experimental Psychology: Animal Behavior Processes*, 31, 376-382.

Goffaux, V. & Rossion, B. (2006). Faces are "spatial"- holistic face perception is supported by low spatial frequencies. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 1023-1039.

Gold, J., Bennett, P.J. & Sekuler, A.B. (1999). Identification of band-pass filtered letters and faces by human and ideal observers. *Vision Research*, 39, 3537-3560.

Gosselin, F. & Schyns, P. G. (2001). Bubbles: A technique to reveal the use of information in recognition. *Vision Research*, 41, 2261-2271.

Gosselin, F. & Schyns, P. G. (2002). RAP: A new framework for visual categorization. *Trends in Cognitive Science*, 6, 70-77.

Gosselin, F. & Schyns, P. G. (2004a). An image is worth thousands of trials: Rendering the use of visual information from spiking neurons to recognition. *Cognitive Science*, 28, 141-146. [pdf]

Gosselin, F. & Schyns, P. G. (2004b). No troubles with Bubbles: A reply to Murray and Gold. *Vision Research*, 44, 471-477.

Harmon, L. D., & Julesz, B. (1973). Masking in visual recognition: effects of two-dimensional filtered noise. *Science*, 180, 1194–1197

Hayes, T., Morrone, M. C., & Burr, D. C. (1986). Recognition of positive and negative band-pass filtered images. *Perception*, 15, 595–602.

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

Humphreys, K., Gosselin, F., Schyns, P. G., & Johnson, M. H. (2006). Using “Bubbles” with babies: A new technique for investigating the informational basis of infant perception. *Infant Behavior & Development*, 29, 471–475.

Izard, C. (1971). *The face of emotion*. New York: Appleton-Century-Crofts.

Izard, C. (1994). Innate and universal facial expressions—evidence from developmental and cross-cultural research. *Psychological Bulletin*, 115, 288–299.

Morrison, D. J., & Schyns, P. G. (2001). Usage of spatial scales for the categorization of faces, objects, and scenes. *Psychonomic Bulletin & Review*, 8, 454–469.

Morrone, M. C., Burr, D. C., & Ross, J. A. (1983). Added noise restores recognizability of coarse quantized images. *Nature*, 305, 226–228.

Näsänen, R. (1999). Spatial frequency bandwidth used in the recognition of facial images. *Vision Research*, **39**, 3824-33.

Oliva, A. & Schyns, P.G. (1997). Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognitive Psychology*, **34**, 72-107.

Parker, D.M., Lishman, J.R. & Hughes, J. (1992). Temporal integration of spatially filtered visual images. *Perception*, **21**, 147-160.

Parker, D.M., Lishman, J.R. & Hughes, J. (1997). Evidence for the view that temporospatial integration in vision is temporally anisotropic. *Perception*, **26**, 1169-1180.

Parker, D.M. & Costen, N.P. (1999). One extreme or the other or perhaps the golden mean? Issues of spatial resolution in face processing. *Current Psychology*, **18**, 118-127.

Rosch, E. (1978). Principles of categorisation. In E. Rosch & B. B. Lloyd (Eds.), *Semantic factors in cognition* (pp. 137-168). Hillsdale, NJ: Erlbaum.

Rossion, B., Caldara, R., Seghier, M., Schuller, A.-M., Lazeyras, F., Mayer, E. (2003). A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. *Brain*, **126**, 2381-2395.

Solomon, J. A., & Pelli, D. G. (1994). The visual filter mediating letter identification. *Nature*, 369, 395–397.

Schyns, P. G., Bonnar, L. & Gosselin, F. (2002). Show me the features! Understanding recognition from the use of visual information. *Psychological Science*, 13, 402-409.

Schyns, P.G. & Oliva, A. (1994). From blobs to boundary edges: Evidence for time- and spatial-scale- dependent scene recognition. *Psychological Science*, 5, 195-200.

Schyns, P. G. & Oliva, A. (1999). Dr. Angry and Mr. Smile: When categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition*, 69, 243-265.

Schyns, P.G., Jentsch, I., Johnson, M., Schweinberger, S. R., & Gosselin, F. (2003). A principled method for determining the functionality of ERP components. *NeuroReport*, 14, 1665-1669.

Schyns, P.G., Petro, L.S., & Smith, M.L. (2007). Dynamics of visual information integration in the brain for categorizing facial expressions. *Current Biology*, 17, 1580-1585.

Smilek, D., Rempel, m.I., & Enns, J.T. (2006). The illusion of clarity : Image segmentation and edge attribution without filling-in. *Visual Cognition*, 14, 1-36.

Smith, M., Cottrell, G., Gosselin, F. & Schyns, P. G. (2005). Transmitting and decoding facial expressions of emotions. *Psychological Science*, 16, 184-189.

Smith, M. L., Gosselin, F. & Schyns, P. G., (2004). Receptive fields for flexible face categorizations. *Psychological Science*, 15, 753-761.

Spezio, M. L., Huang, P.-Y. S., Castelli, F., Adolphs, R. (2007). Amygdala damage impairs eye contact during conversations with real people. *Journal of Neuroscience*, 27, 3994-3997.

Sprengelmeyer, R., Young, A.W., Calder, A.J., Karnat, A., Lange, H., Hömberg, V., Perrett, D.I. and Rowland, D. (1996). Loss of disgust: perception of faces and emotions in Huntington's disease. *Brain*, 119, 1647-1665.

Tanskanen, T., Näsänen, R., Montez, T., Päällysaho, J. & Hari, R. (2005). Face recognition and cortical responses show similar sensitivity to noise spatial frequency. *Cerebral Cortex*, 15, 526-534.

Tieger, T., & Ganz, L. (1979). Recognition of faces in the presence of two-dimensional sinusoidal masks. *Perception and Psychophysics*, 26 (2), 163–167.

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

Vuilleumier, P., Armony, J.L., Driver, J., & Dolan, R.J. (2003). Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience*, 6, 624–631.

Watson, A. B., & Ahumada, A. J. (2005). A standard model for foveal detection of spatial contrast, *Journal of Vision*, 5 (9), 717-740. <http://journalofvision.org/5/9/6/>.

Westheimer, G. (2001). The Fourier theory of vision. *Perception*, 30, 531–541.

Whalen, P.J. et al. (2004). Human Amygdala Responsivity to Masked Fearful Eye Whites. *Science*, 306, 2061.

Winston, J.S., Vuilleumier, P., & Dolan, R.J. (2003). Effects of lowspatial frequency components of fearful faces on fusiform cortex activity. *Current Biology*, 13, 1824–1829.

### ***Liste des figures***

Figure 1. Photo de l'acteur Brad Pitt : (a) sans filtrage; (b) filtrée « passe-bas » à 12 cycles par visage; (c) filtrée « passe-bande » entre 12 et 36 cycles par visage; et (d) filtrée « passe-haut » à 36 cycles par visage.

Figure 2. (a) Une grille sinusoïdale ayant une fréquence spatiale, une phase, un contraste et une orientation spécifiques.; (b) la même grille avec une fréquence différente; (c) la même grille avec une phase différente; (d) la même grille avec un contraste différent; et (e) la même grille avec une orientation différente.

Figure 3. Hybride comprenant une photo de Brad Pitt filtrée « passe-bas » et une photo de l'acteur Brendan Fraser filtrée « passe-haut ».

Figure 4. Photo de l'acteur Brad Pitt : (a) ne contenant que l'information verticale +/- 5 deg; et (b) ne contenant que l'information horizontale +/- 5 deg.

Figure 5. Photo de l'acteur Brad Pitt échantillonnée aléatoirement avec des « bulles » locales à cinq bandes de fréquences spatiales.

Figure 6. À gauche : information efficace pour trois tâches de discrimination de visages : (a) l'identité, (b) le sexe et (c) l'expression de joie vs. la neutralité. À droite : les graphiques à barre indiquent le nombre de pixels significatifs dans les quatre bandes de fréquences les plus fines pour chacune des tâches.

Figure 7. Information efficace pour la discrimination des six expressions de base (i.e., la joie, la surprise, la peur, la colère, le dégoût et la tristesse) ainsi que la neutralité.

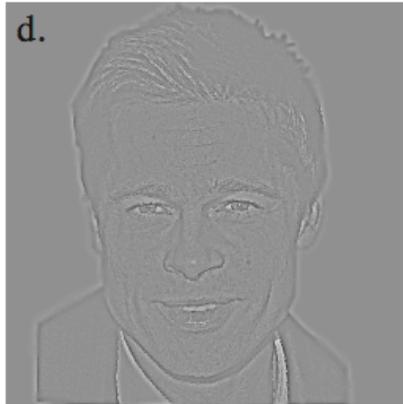
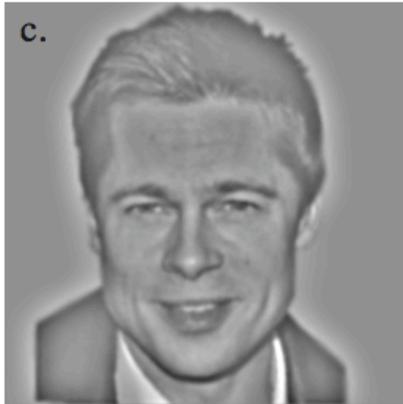
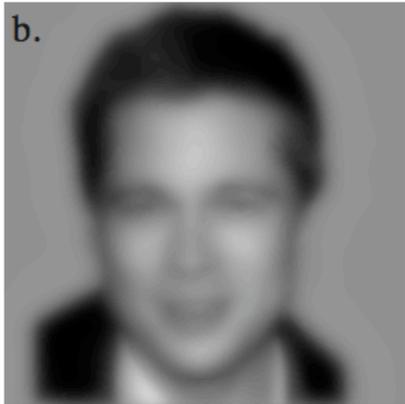
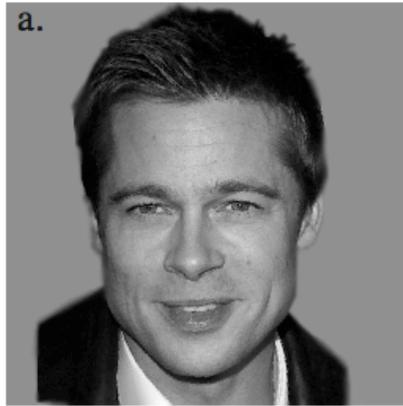


Figure1 - Fiset & Gosselin

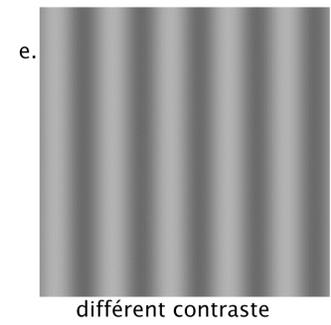
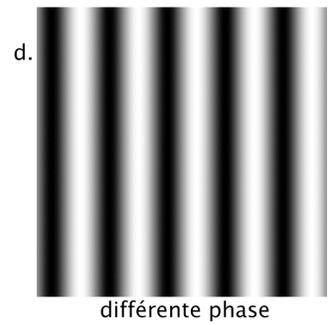
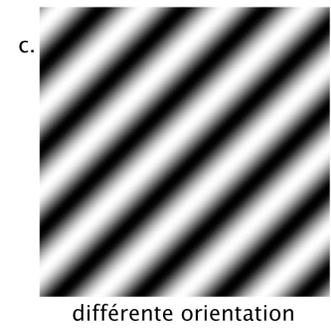
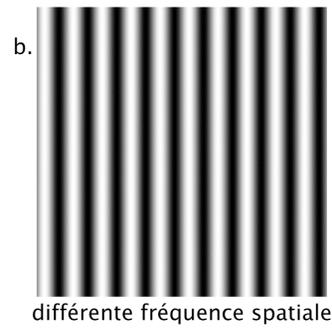
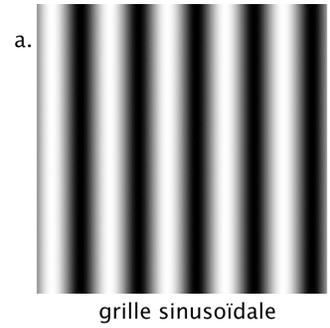


Figure 2 – Fiset & Gosselin



Figure 3 - Fiset & Gosselin

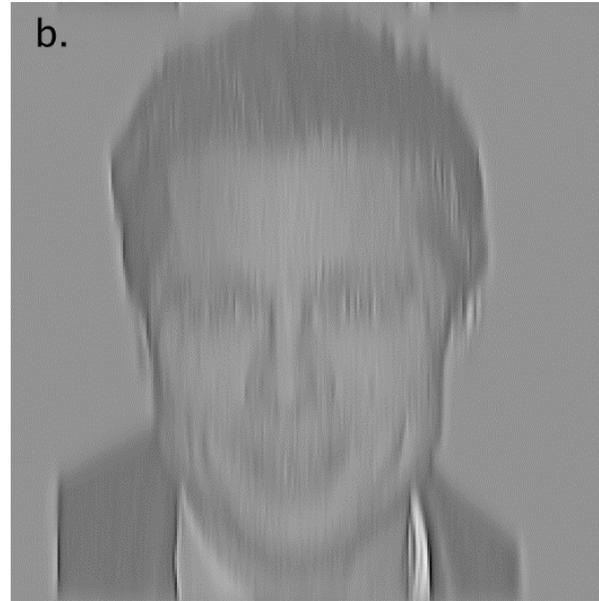
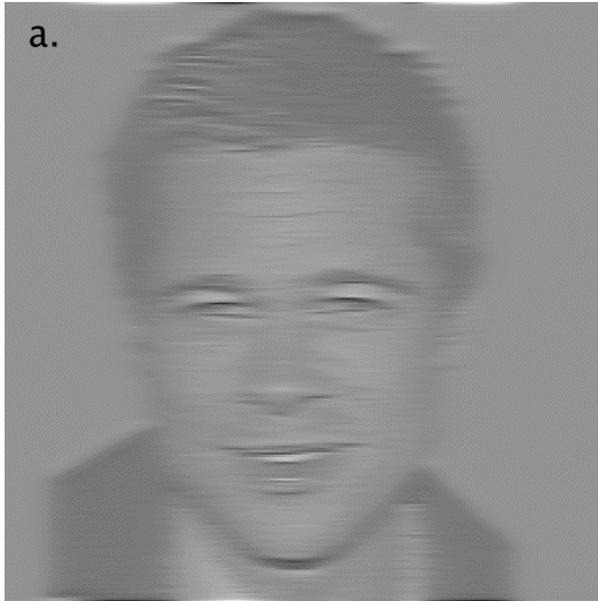


Figure 4 - Fiset & Gosselin

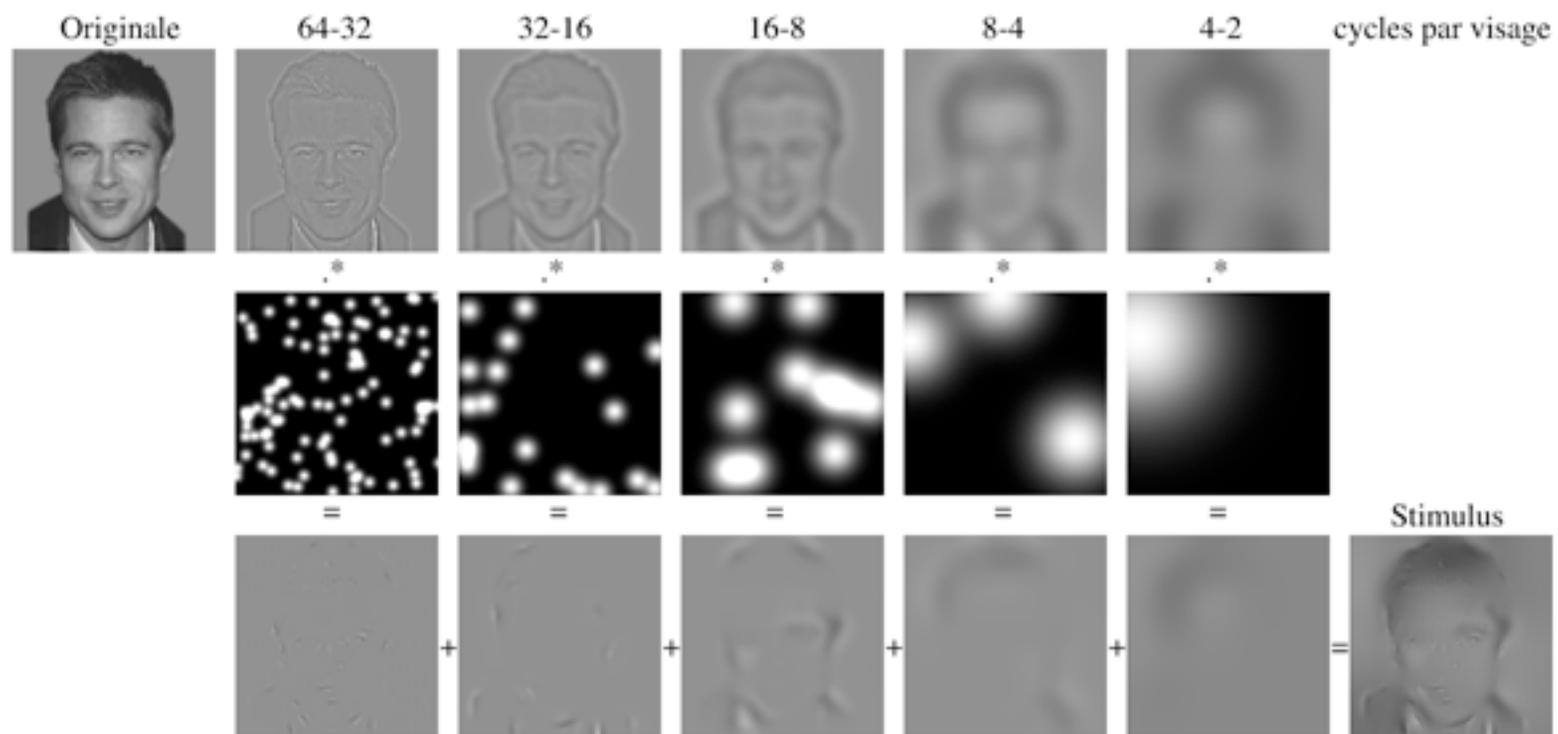


Figure 5 - Fiset & Gosselin

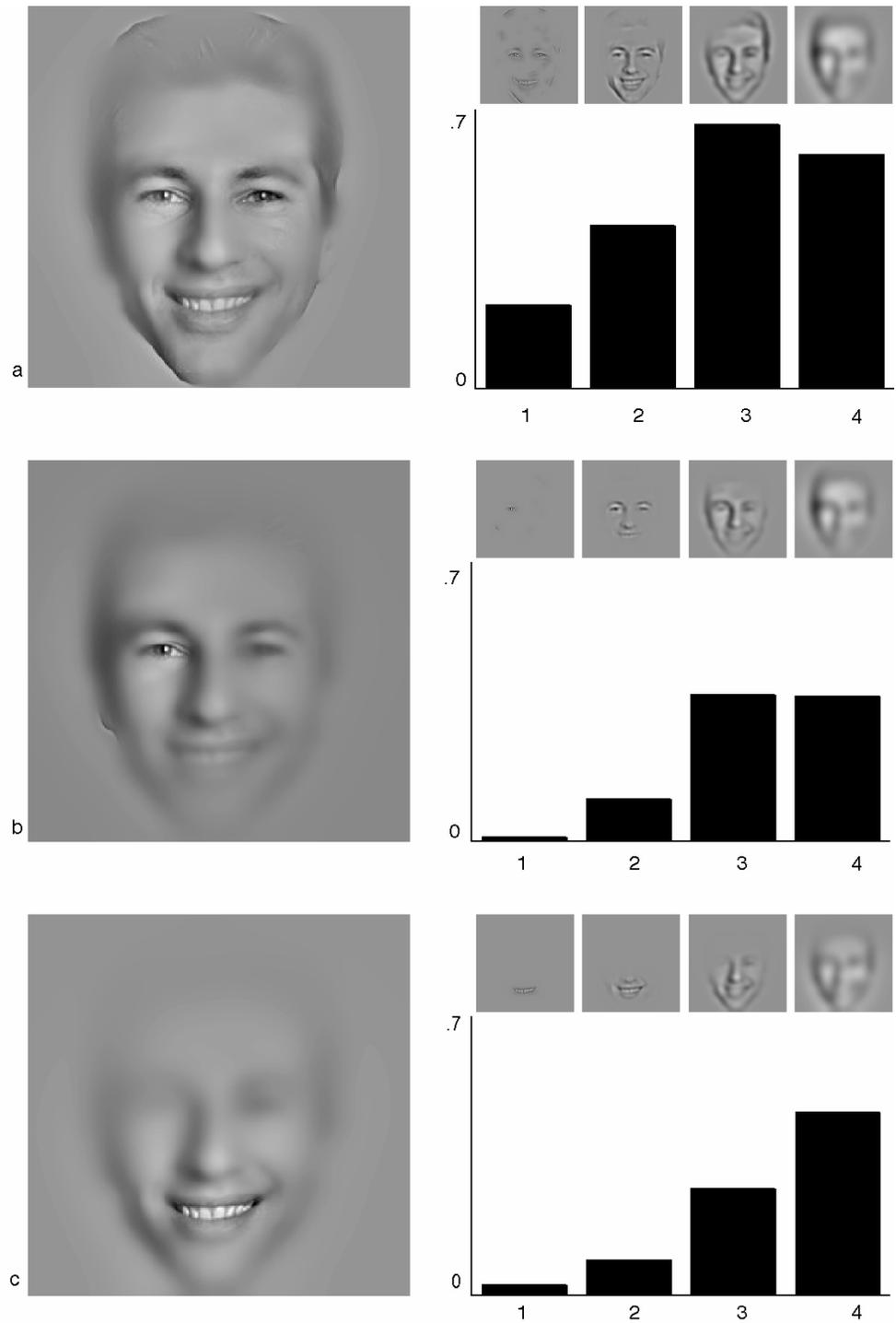


Figure 6 – Fiset & Gosselin

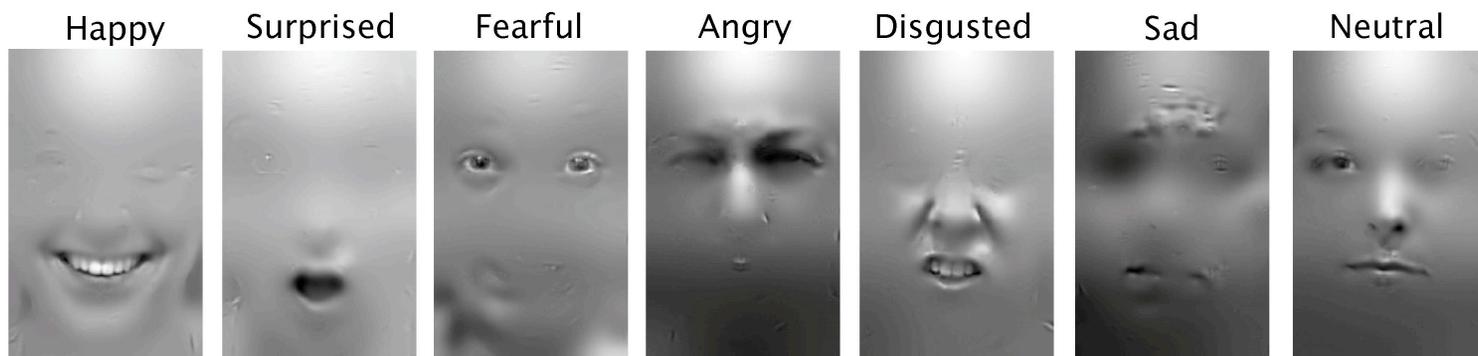


Figure 7 – Fiset & Gosselin