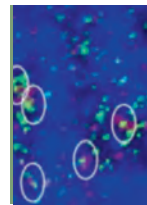


Peut-on comprendre les mécanismes de la perception ?

Jean-Pierre Henry

► La perception est la compréhension des informations que les organes sensoriels envoient au cerveau. Visuelles, elles sont transmises au cortex visuel dans le lobe occipital. L'imagerie par résonance magnétique montre que cette information est traitée dans le cortex temporal. Deux exemples seront discutés dans cette revue : la reconnaissance des mots écrits, et celle des visages. Dans le premier cas, une hypothèse, *la hiérarchie de la combinaison de détecteurs locaux*, explique les propriétés d'invariance de la forme des lettres. Dans le second, le code utilisé dans le cerveau du singe macaque pour reconnaître les visages a été déchiffré : à partir de l'activité électrique de 200 neurones du cortex temporal inférieur, il est en effet possible à un singe de connaître le visage qui lui a été présenté. ◀



Ancien directeur de l'institut de Biologie-Physico-Chimique, Fondation Edmond de Rothschild, 13 rue Pierre et Marie Curie, 75005 Paris
jpierre.henry@orange.fr

est dualiste, opposant l'esprit et la matière. Elle a été développée au XVII^e siècle par René Descartes (1596-1650), mais est toujours vivante chez Henri Bergson (1859-1941). Descartes lui-même ne séparait pas complètement l'esprit et la matière, donnant à une glande nichée dans le cerveau, la glande pinéale (ou épiphyse), la fonction d'effectuer la jonction entre le corps et l'âme, une importance attribuée à l'épiphyse que ne confirme pas la physiologie. Les neurosciences suivent de préférence la première ligne, mais avant-elles de manière significative dans la compréhension de la perception ? Dans un livre récent, *Une Brève Histoire du Cerveau* [1], le neurobiologiste britannique Matthew Cobb (né en 1957) ne voit aucun progrès dans la compréhension de la genèse de la pensée et, pour lui, les neurosciences se sont enfoncées dans une impasse. Dans cette revue, je m'efforcerai de faire le point sur la perception visuelle, dans l'espoir d'atténuer quelque peu le pessimisme de cet auteur.

L'organisation des voies visuelles

Le principe des voies visuelles est simple. Dans la cavité oculaire, une lentille convergente, le cristallin, forme une image sur la rétine. Celle-ci est composée de plusieurs millions de cellules qui, après un certain nombre d'opérations, vont transformer la lumière en signaux électriques qui seront transmis, après un seul relai (le corps genouillé latéral), au cortex visuel primaire situé à l'arrière du cerveau dans les lobes occipitaux. Très grossièrement, le champ visuel d'une cellule ganglionnaire de la rétine (c'est-à-dire la fraction de l'espace observé qui produit un signal électrique dans cette

Pour assurer sa survie, notre organisme doit avoir une connaissance permanente de son environnement. C'est cette information que les organes des sens fournissent au cerveau, qui doit donner un sens à cette masse de données. La perception désigne ce phénomène, d'abord inconscient, puis conscient, c'est-à-dire susceptible d'être « raconté ». Autant le fonctionnement des organes des sens apparaît explicable, relevant de la physique et de la biologie, autant la perception semble mystérieuse. Ce phénomène est fascinant parce qu'il est subjectif, étant la rencontre entre « l'esprit » et le monde objectif. Comme tel, il a intéressé les penseurs à travers les siècles. Pour les uns, le cerveau « sécrète » la pensée, il fabrique une représentation de l'environnement à partir des données sensorielles. D'autres ont gardé et gardent le mystère de l'esprit à l'écart du fonctionnement du cerveau. La première ligne de pensée se retrouve chez Platon (427-347/8 avant J-C), qui oppose le sensible, l'image fugace du concret, à « l'idée », engendrée par « l'âme ». La seconde ligne

Vignette (© Inserm-Delphine Benarroch-Popivker).



cellule) est une cocarde (pour simplifier, nous dirons un disque), et l'organisation générale de l'image est transmise depuis la rétine jusqu'au cortex ; on parle d'organisation rétinotopique.

Qui regarde l'image du cortex visuel ? Comment le cerveau va-t-il traiter cette information ? Comment va-t-il reconnaître un visage, un texte dans ces images ?

L'imagerie donne des premiers éléments de compréhension à ces questions. Lorsque des images sont présentées à un sujet, l'imagerie par résonance magnétique (IRM) nous révèle que le cortex visuel est activé, mais que d'autres aires dont la localisation dépend de la nature de la scène visualisée, s'allument également. La première observation a été réalisée en 1997 par Nancy Kanwisher (*Massachusetts Institute of Technology*, Cambridge, États-Unis) [2] qui observa dans le cortex temporal, une aire possédant une spécificité pour les visages. D'autres aires furent ensuite décrites pour les objets et les paysages [3]. Il existe aussi une aire dite « de la forme des mots » que vous activez en lisant ces lignes [4]. Ces différentes aires, proches les unes des autres, sont situées en avant du cortex occipital, dans le cortex temporal postérieur.

Cette localisation reposant sur une fonction a été une avancée majeure pour les neurosciences. Cependant, il faut aussi en cerner les limites. Dans de nombreux cas, plusieurs aires s'activent, suggérant que la fonction est plutôt associée à un réseau neuronal. Il faut aussi remarquer qu'une corrélation observée n'implique pas la causalité, une remarque souvent faite à propos des résultats d'imagerie. Toutefois, la causalité est apparue dans une expérience [5] réalisée chez un patient auquel une électrode avait été implantée dans l'aire de reconnaissance des visages (pour des raisons thérapeutiques). Lorsque ce patient, dans l'état vigile, recevait une stimulation électrique via cette électrode, sa perception des visages était complètement modifiée (il parlait alors de « métamorphose »), et cela sans qu'aucun autre trouble ne soit détecté.

Si ces expériences suggèrent que l'aire en question participe au traitement de l'image des visages, l'IRM, à elle seule, ne peut néanmoins en aborder le mécanisme ; cette technique d'imagerie est en effet très loin de l'échelle cellulaire. Sans entrer dans des détails techniques, le plus petit volume analysé en IRM, le voxel¹, a des dimensions de l'ordre d'une dizaine de millimètres. Il correspond donc à plus de 5 millions de neurones, soit 30 milliards de synapses et 220 kilomètres d'axones [6] ! Pour avoir une idée des mécanismes impliqués dans le traitement de l'image, il faudra donc descendre à un niveau d'organisation différent, celui du neurone.

Le cortex visuel primaire et le principe de connectivité

Nous avons écrit précédemment que l'organisation de l'information électrique fournie par la rétine était maintenue dans le cortex occipital. C'est ce que pensaient les neurophysiologistes américains David Hubel et Thorsten Wiesel (prix Nobel de physiologie ou médecine 1981) au début de leurs carrières, en se fondant sur des expériences consis-

tant à implanter, chez le chat, des microélectrodes dans son cortex occipital. Mais leur surprise fut grande de découvrir que les neurones implantés ne répondaient pas à une tache circulaire (ou à une cocarde) comme ceux de la rétine. Dans son discours de réception du prix Nobel, David Hubel raconte leur excitation lorsqu'ils ont observé que les neurones avaient des réponses très variables, les plus simples d'entre eux étant activés par des segments d'orientation définie.

Comment passe-t-on d'un point (d'un disque) à un segment orienté ?

Un neurone activé produit un signal électrique de courte durée (environ cinq millisecondes) qui se propage vers l'extrémité de son axone, où il libère un neurotransmetteur. Ce dernier va modifier l'équilibre électrique du neurone cible et, si ce déséquilibre dépasse un seuil, ce neurone sera lui aussi activé. Dans le cortex, les contacts entre neurones (les synapses) sont très nombreux : une cellule est en relation avec 1 000 à 10 000 neurones cibles. L'arrivée d'un signal peut ne pas être suffisante pour activer le neurone cible, si le seuil d'activation n'est pas atteint. Toutefois, si plusieurs neurones contactent la même cellule cible, en même temps, alors les effets s'additionnent et la cellule pourra s'activer. C'est le principe de connectivité.

L'information en provenance de la rétine est relayée dans le corps genouillé, c'est-à-dire qu'il existe des synapses entre le nerf optique, entrant, et les neurones dirigés vers le cortex optique. Les observations de Hubel et Wiesel sont alors explicables [7]. Si trois neurones entrant dans le corps genouillé sont activés, chacun par un point lumineux, si ces trois points sont alignés, et si ces trois neurones contactent le même neurone sortant, alors ce dernier sera sensible au segment formé par les trois points sources excitateurs et activé.

Cet exemple simple montre comment un neurone peut répondre à des signaux complexes, en fonction de son câblage. Pour être complet, il faut ajouter que ce câblage peut se modifier en fonction de l'usage qui en est fait. Cette seconde observation est souvent présentée de la manière suivante : « *des neurones qui s'excitent ensemble se lient entre eux* », un principe énoncé par le neurophysiologiste canadien Donald Hebb en 1949.

Ainsi, le cortex visuel primaire doit-il être vu non comme une simple représentation rétinotopique de ce que voit l'œil, mais comme une base de données des scènes vues : mise en évidence des lignes, de leur orientation, de leur fréquence spatiale, mais aussi de nombreuses autres données concernant les couleurs ou la vision binoculaire. Cette base de données alimente des

¹ Équivalent en trois dimensions du pixel (en deux dimensions).

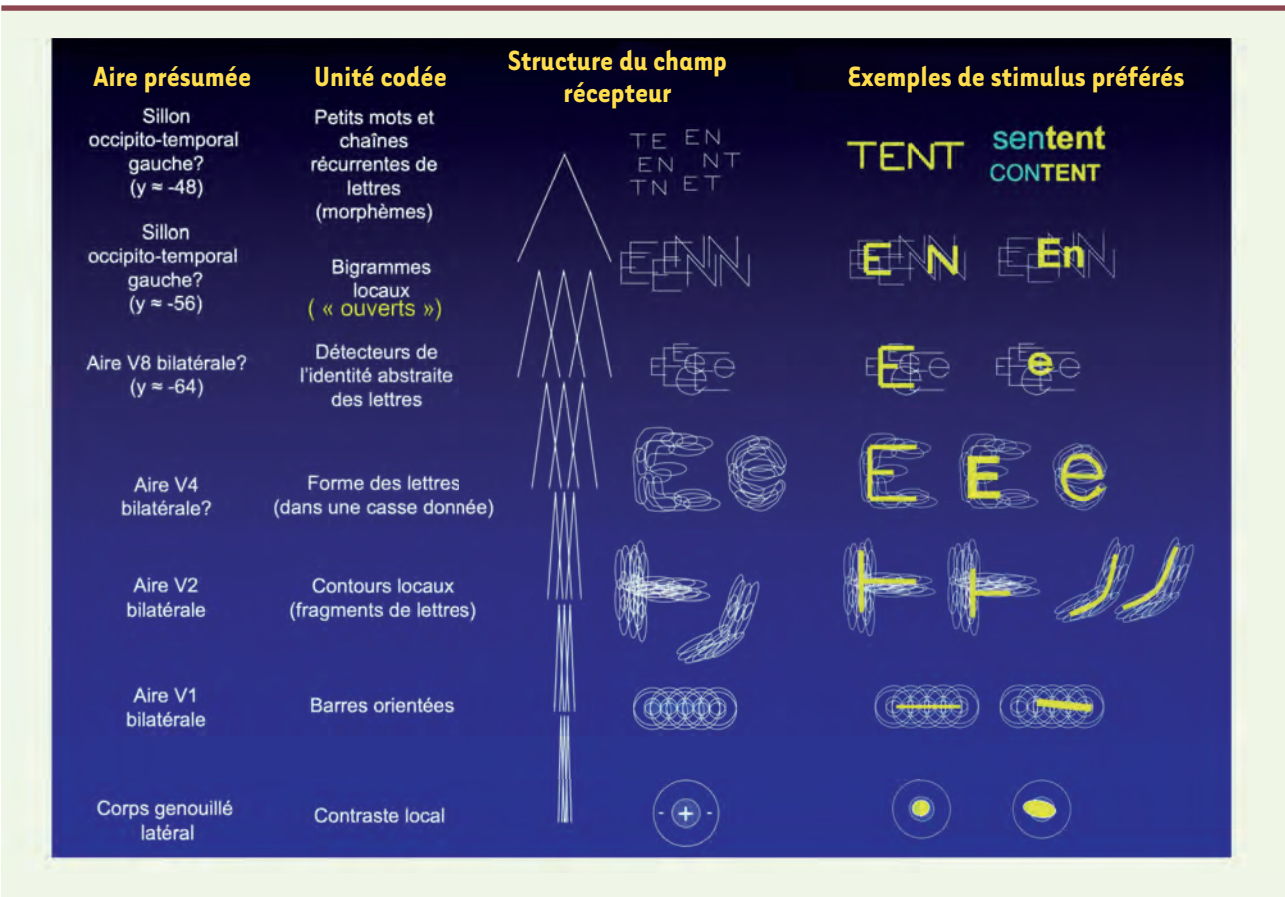


Figure 1. Représentation des lettres dans l'aire de la reconnaissance de la forme des mots (modèle LCD, local combination detectors). Selon l'hypothèse de Dehaene *et al.*, la marche vers l'invariance de la forme est une progression vers des niveaux hiérarchiques de plus en plus complexes, associant d'abord les composants d'une lettre, puis ses différentes représentations. Dans le tableau, la ligne la plus basse (corps genouillé latéral) représente le champ visuel d'un neurone du corps genouillé, la suivante, celui d'une cellule du cortex visuel primaire (V1), puis celui d'aires occipitales plus antérieures (V2, V4, V8) et d'aires proches du cortex temporal (d'après [8]).

cortex secondaires, impliqués dans la compréhension du mouvement (voie dorsale) ou dans la reconnaissance des formes (voie ventrale), ce qui conduira à la perception des images vues.

L'hypothèse des neurones de la lecture

En s'appuyant sur les deux observations précédentes, les neuroscientifiques français Stanislas Dehaene, Laurent Cohen et leur équipe ont proposé une explication des mécanismes de la lecture [4]. Comme pour la perception des visages, une aire située dans le cortex temporal de l'hémisphère gauche est activée pendant la lecture et, comme précédemment, une altération de cette aire (telle qu'un accident vasculaire cérébral), ou sa neutralisation temporaire, par une excitation électrique, entraîne un blocage de la lecture par inhibition de la reconnaissance des mots, ce qui permet de confirmer le rôle de l'aire activée. Mais comment les neurones de cette aire, activée par la lecture, reconnaissent-ils les lettres et les mots [8] ? On peut considérer que les lettres, L, E, V écrites en majuscules, sont des assemblages de segments, de la même manière que les segments sont des assemblages

de points (un segment étant un lien entre deux points). Supposer que des neurones de l'aire de reconnaissance de la forme des mots reçoivent des signaux de neurones de l'aire primaire, codant des segments bien orientés, permet de comprendre comment un neurone peut reconnaître une lettre (via la reconnaissance des différents segments orientés). Le même raisonnement est applicable à l'ensemble des lettres, quelle que soit la forme de l'écriture. La Figure 1 montre ainsi comment les lettres pourraient être définies à partir des données du cortex primaire.

En continuant le même raisonnement, dans une aire supérieure, des neurones sont contactés par les différentes entités représentant les différentes formes de la même lettre : minuscule, majuscule, anglaise, cursive, script, etc. Cette étape est importante, parce que ces nouveaux neurones représentent le premier niveau d'abstraction : ils ne sont plus associés à des dessins, mais à une abstraction, à une lettre.

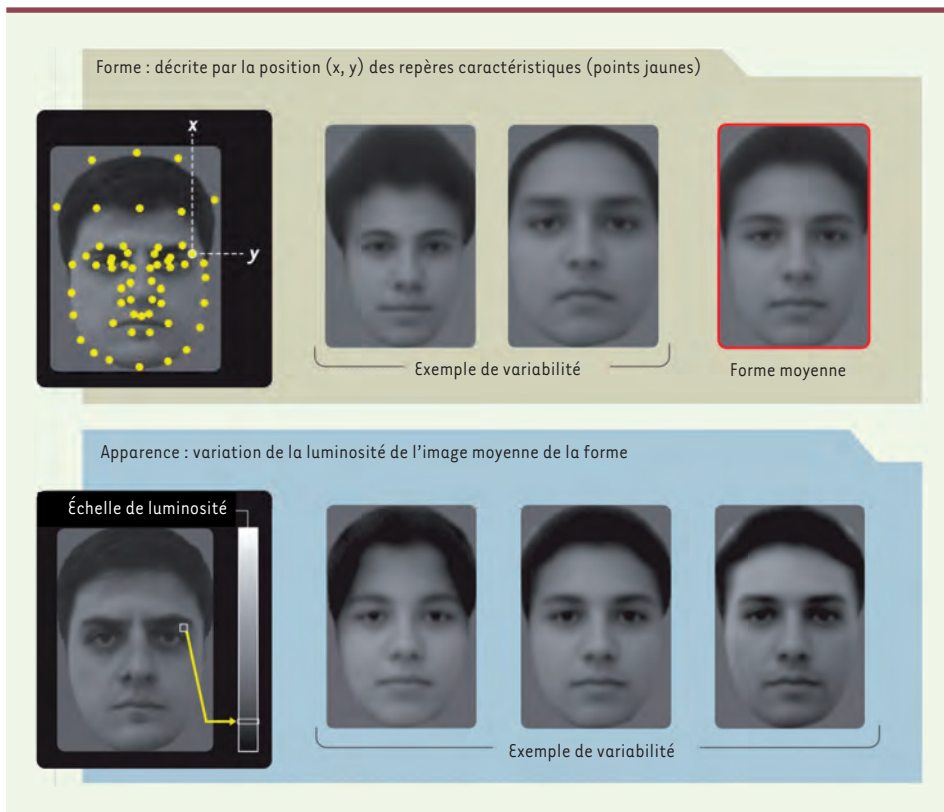


Figure 2. Comment identifier un visage. Les visages ont été numérisés en utilisant deux « descripteurs ». Le premier s'intéresse à la forme. Des points sont placés à des endroits définis et, sur l'image en deux dimensions, les distances entre ces points sont mesurées (selon les axes x et y). Le second considère l'apparence. Pour cela, les différences de contraste, selon une échelle de luminosité, entre différentes parties du visage, sont mesurées (d'après [10]).

L'hypothèse de Dehaene et Cohen propose un mécanisme simple par lequel la vision produit une perception, une « idée » au sens platonicien du mot. Cette hypothèse est difficile à vérifier expérimentalement, mais elle prend toute sa valeur en suggérant un mécanisme d'apprentissage de la lecture. En utilisant la connectivité, certains neurones participeraient à la reconnaissance de syllabes. Ainsi avec quelques règles d'utilisation de ces neurones, Dehaene et Cohen rendent compte des caractéristiques de la lecture : indépendance de la taille des lettres, de leur position, de leur casse (minuscules, majuscules) et du choix des caractères.

Le décodage de la reconnaissance des visages chez le singe macaque

Dépasser la corrélation entre l'imagerie et la perception, c'est l'objectif de Doris Tsao (professeur de biologie au CalTech, *California Institute of Technology*, à Pasadena, Californie, États-Unis). Celle-ci a choisi d'étudier la perception des visages chez le singe macaque, qui identifie parfaitement les visages humains, et sur lequel les approches électrophysiologiques et l'IRM sont possibles.

Dans un premier temps [9], l'imagerie par IRM a révélé la position des aires du cortex temporal inférieur qui sont activées par la vision, par le singe, de visages humains. Ces aires sont formées de six noyaux de neurones (ou groupes de neurones) dans chaque hémisphère. Les neurones de ces noyaux ont ensuite été explorés à l'aide de micro-électrodes. Le résultat de cette expérience a été riche : la grande majorité de ces neurones s'activent lorsque des images de visages, ou des

caricatures de visages, sont présentées au singe, mais ne réagissent pas à d'autres catégories d'objets. Doris Tsao et son équipe étaient ainsi en présence des cellules responsables de la perception, et, immédiatement, ils ont obtenu un résultat important. Les six noyaux activés n'étaient pas équivalents : dans les noyaux les plus postérieurs, on trouvait différents types de neurones, chacun répondant à une orientation différente de la tête (face, profil, etc.) ; le noyau le plus antérieur était en revanche plus homogène, les cellules répondant de manière invariante à l'angle de présentation du visage [10]. Ainsi, les noyaux communiquaient entre eux pour former un réseau travaillant, *in fine*, à la représentation « d'une idée » de visage.

L'étape suivante a été de déterminer le code utilisé par ces neurones. Comment peut-on coder numériquement un visage ? S'appuyant sur les travaux sur la reconnaissance faciale automatique, l'équipe de Doris Tsao a défini deux types de « descripteurs »² (Figure 2). Le premier descripteur considère la forme du visage, qui est imposée par le squelette. Des repères sont posés sur la face d'un sujet. Les distances entre les points fixés au visage permettent de mesurer la largeur du visage, la distance entre les yeux, la place du nez, etc. Le second

² En informatique, un descripteur est une quantité mesurable ou calculable qui permet de décrire en partie un objet.

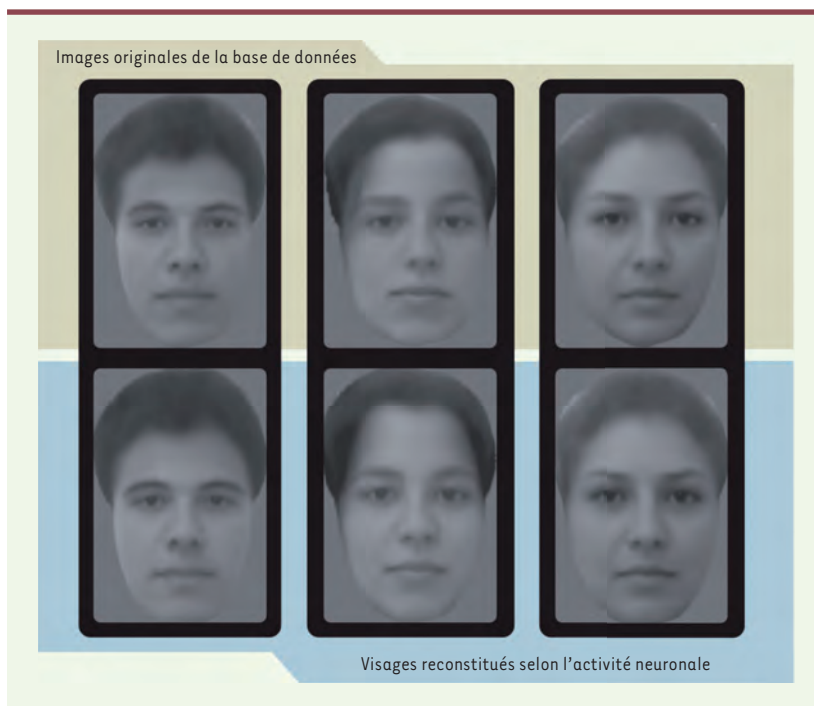


Figure 3. Le déchiffrement du code de reconnaissance des visages. Dans une banque comportant 2 000 visages, trois visages (première ligne) ont été montrés à un singe. Des enregistrements électrophysiologiques ont été effectués et, à partir de ceux de 205 neurones, le code (algorithme) a été appliqué pour reconstruire le visage présenté (visages de la seconde ligne) (d'après [10]).

La majeure partie des neurones répond en fait à plusieurs paramètres, mais les chercheurs ont trouvé la règle simple selon laquelle ces paramètres se composent pour former un axe unique à chaque cellule, la réponse de la cellule variant de manière continue le long de cet axe.

Le raisonnement peut paraître compliqué, mais le résultat est simple et brillant (Figure 3). En analysant la réponse électrique de 200 neurones devant un

visage choisi parmi les 2 000 créés, les chercheurs sont capables de reconstituer avec précision quel visage a été observé. La lecture de l'activité du cerveau permet ainsi d'accéder à la sensation qui l'a créée !

descripteur s'intéresse à l'apparence et à la texture du visage : il prend en compte le teint, la couleur des yeux et des cheveux, en mesurant les contrastes entre les différentes parties du visage. Cette approche est appliquée à un ensemble de visages, obtenus à partir de banques de données, et, à l'aide de techniques mathématiques, les paramètres de description les plus importants, ceux qui différencient le plus les visages, ont été retenus : 25 pour la forme et 25 pour l'apparence. Cela signifie qu'un visage pourra être défini en ne choisissant que 50 valeurs pour ces paramètres ! L'affirmation semble brutale, 50 chiffres pour définir la Joconde... Mais les auteurs pointent que l'ensemble des couleurs, avec les nuances les plus subtiles, peut être défini avec trois couleurs : le jaune, le rouge et le bleu. Si les poids relatifs du jaune, du rouge et du bleu dans une couleur complexe sont portés sur les trois axes x , y et z , alors cette couleur sera représentée par un point dans l'espace à trois dimensions. De la même manière, l'ensemble des visages sera représenté dans un espace à 50 dimensions organisé autour des 50 paramètres retenus. Dans cet espace, les chercheurs ont choisi 2 000 visages, chacun défini par ses valeurs propres des 50 paramètres, et ils ont montré ces visages à un singe dont ils ont enregistré l'activité électrique de neurones des noyaux sensibles. Deux cents cellules ont été enregistrées, chacune ayant vu les 2 000 visages plusieurs fois. Toutes les cellules n'ont pas répondu de la même manière et n'avaient pas la même sensibilité : par exemple, certaines n'ont répondu qu'à la distance entre les yeux, d'autres qu'à la place du nez. Ce point est important parce qu'il indique que les neurones font un travail d'analyse qui est semblable à celui qui a permis aux chercheurs de « numériser » les visages. En reprenant l'analogie des couleurs, certains neurones ne seraient ainsi sensibles qu'au bleu.

Les avancées dans la connaissance de la perception visuelle

L'évolution des hypothèses

Les mécanismes de la perception étaient, jusqu'alors, restés dans le flou. Deux hypothèses étaient avancées. Selon la première, la perception repose sur un grand nombre de cellules, chacune apportant un aspect particulier de la cible. Dans la seconde hypothèse, un nombre limité de cellules sont au contraire impliquées, à la limite, une seule cellule reconnaissant un visage sous toutes ses apparences, une hypothèse connue sous le nom d'*hypothèse de la cellule grand-mère* (*grandmother cell*), une cellule qui reconnaît votre grand-mère. Cette dernière hypothèse trouve son soutien dans des expériences réalisées chez des patients dont le cortex a été implanté par des microélectrodes [11]. Dans ces expériences, des neurones, qui ne répondent qu'à une seule image présentée au patient, par exemple celle de l'actrice Jennifer Aniston, vue sous des angles variés, ont été identifiés. Ces neurones auraient donc individuellement à la fois les propriétés de reconnaissance et d'invariance.

Les résultats de Doris Tsao soutiennent néanmoins la première hypothèse [12]. En effet, si un neurone ne

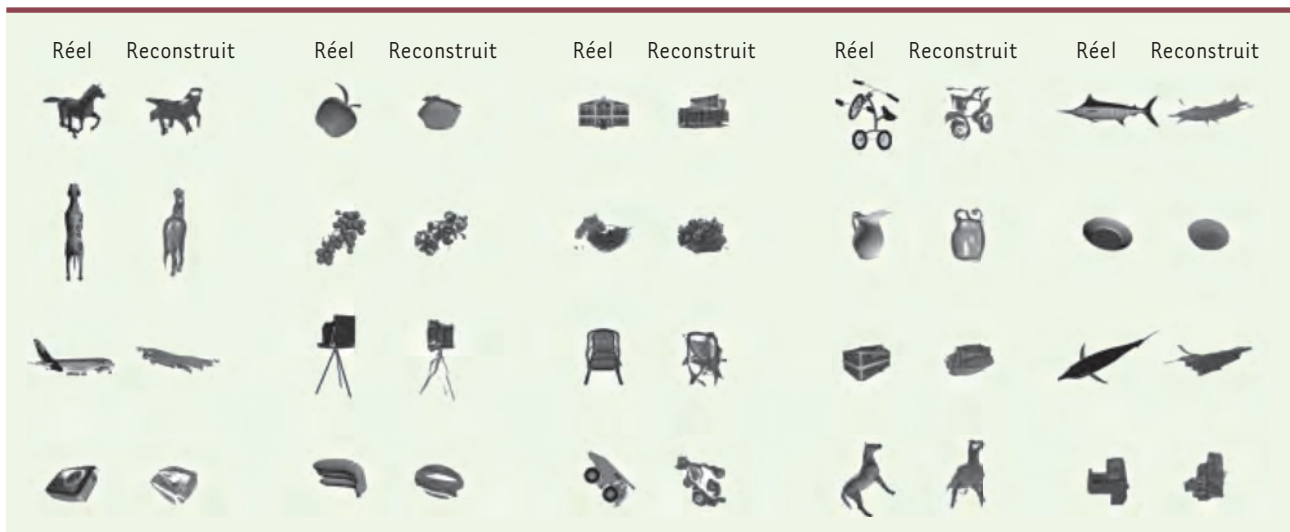


Figure 4. Le code de reconnaissance des objets. Reprenant l'approche développée pour l'identification des visages, plusieurs aires ont été identifiées pour leur participation à la reconnaissance des objets. Des objets ont été présentés à un singe et les signaux de 482 neurones de ces aires ont été traités pour essayer de reconstruire l'image présentée (d'après la figure 6 de [12]).

répond qu'à une seule des 50 coordonnées de l'espace des visages, les 49 autres pourront varier, mais ce neurone sera toujours activé : il répondra donc à un grand nombre de visages différents présentant la même valeur de coordonnée. Cela est toujours vrai si le neurone répond à plusieurs coordonnées, tant que leur nombre est inférieur à 50. Le raisonnement peut être conduit de manière plus abstraite, mais le résultat est clair et il soutient la première hypothèse.

Mais comment alors expliquer le « neurone Jennifer Aniston » ? Ce n'est en fait pas un neurone du cortex temporal inférieur, mais un neurone de l'hippocampe postérieur (lobe temporal médian), siège de la mémoire. On peut donc proposer que la reconnaissance soit faite dans les noyaux du cortex temporal inférieur et que l'information soit ensuite reprise par un réseau comprenant le « neurone Jennifer Aniston », qui correspond en fait à un souvenir.

Ce point est important parce qu'il aborde la question de l'échelon supérieur. Chez l'homme, la reconnaissance des visages joue un rôle important dans les relations sociales, et de nombreux aspects doivent être déchiffrés au-delà de l'identification de la personne : son expression, son genre, son âge, son attractivité, la familiarité avec laquelle une relation est engagée, etc. Ces analyses vont impliquer de manière bidirectionnelle d'autres régions du cerveau, celles participant à la mémoire (hippocampe), aux émotions (amygdale), aux prises de décision (cortex frontal) et à la conscience. Les travaux de Doris Tsao devraient permettre une exploration plus ciblée des relations anatomiques et fonctionnelles entre ces différentes régions.

Reconnaissance faciale par l'intelligence artificielle et perception

La reconnaissance faciale est un domaine de l'informatique en développement rapide. Elle permet par exemple à un ordinateur de reconnaître son propriétaire et de s'allumer. Les difficultés sont grandes : derrière la diversité des images, face profil, tête penchée, visibilité partielle, expressions du visage, le système doit définir un seul visage.

C'est aussi le problème de la perception que le cerveau humain domine depuis la plus tendre enfance. L'approche de décodage proposée par l'équipe de Doris Tsao fait appel à une des premières techniques utilisées, la reconnaissance faciale du modèle d'apparence active (*active appearance model*) qui donne des paramètres de forme et d'apparence aux visages. D'autres techniques sont maintenant utilisées, relevant de l'intelligence artificielle (IA), en particulier l'apprentissage profond (*deep learning*) qui utilise des réseaux de neurones virtuels en couches multiples et d'importantes bases de données pour l'apprentissage. Dans un travail récent [13], des photographies de visages ont été présentées à des singes et différents modèles ont été comparés pour leur capacité à expliquer les résultats électrophysiologiques obtenus après stimulation. Le résultat est surprenant car la méthode de Doris Tsao surpasse tous ces modèles (à une exception près) : elle est en effet meilleure qu'une méthode de *deep learning* (la *visual geometry group [VGG]-Face*) entraînée sur presque un million de visages, et qui comporte des centaines de milliers de paramètres ! Des origines possibles des imperfections de ces modèles ont été proposées, mais ces résultats montrent les performances du cerveau par rapport à l'IA, qui reste une « boîte noire » où la progression du raisonnement est impossible à suivre.

La perception et le lobe temporal inférieur du cortex

Reprenant l'approche conduisant à la caractérisation des noyaux de reconnaissance des visages, combinant imagerie IRM en réponse à des présentations d'images, microstimulations électriques, pour déterminer les

connections en réseaux et analyse électrophysiologique, l'équipe de Doris Tsao a localisé d'autres aires de reconnaissance dans le lobe inférieur, en particulier des aires de reconnaissance des objets [14]. Cette localisation dépend de critères d'aspect (objets pointus ou arrondis), et non de critères sémantiques (chaussures ou avions) qui, pour le macaque, n'auraient aucun sens. Le lobe temporal inférieur apparaît alors comme un centre de reconnaissance visuelle avec des aptitudes différentes (visages, parties du corps, objets pointus ou arrondis) organisées en quatre différents réseaux, chacun avec plusieurs noyaux possédant des capacités différentes de saisir l'invariance de la cible.

La manière dont fonctionnent ces différents réseaux est identique, et en appliquant une approche similaire, il est possible de reconstituer une image, certes très approximative, d'objets d'après l'activité électrique des neurones stimulés lors de la vision d'images réelles (Figure 4). La reconnaissance des visages apparaît alors comme une spécialisation d'un mécanisme plus général, favorisée par son rôle social.

Le travail de Doris Tsao est important parce qu'il décrypte un mode de fonctionnement du cerveau susceptible d'être appliqué à d'autres domaines de la perception visuelle, mais aussi à d'autres modes de la perception, comme la compréhension du langage. Des avancées considérables ont été publiées sur cette question [15, 16]. ♦

SUMMARY

Is understanding the mechanism of perception possible?

Perception is the understanding by the brain of the different sensory information. In humans, information is mostly visual and is conveyed to the occipital cortex. MRI techniques suggest that it is deciphered in different areas of the temporal cortex, depending upon the observed scene. The neural codes used by neurons of these areas will be discussed in two instances: Recognition of written words and of faces. In the first case, an hypothesis using basic properties of neurons and hierarchy of local combination detectors, accounts for the invariance of their visual form. That it is relevant is emphasized by its successful application to reading mechanism. In the second more recent instance, recognition of human faces by macaques is shown to reside in neuron patches in the inferior temporal cortex, forming a system identifying

faces with view invariance. The code allowing facial identification using the electrical activity of 200 neurons was cracked. ♦

LIENS D'INTÉRÊT

L'auteur déclare n'avoir aucun lien d'intérêt concernant les données publiées dans cet article.

RÉFÉRENCES

1. Matthew C. *Une brève histoire du cerveau*. Paris : Dunod, 2021.
2. Kanwisher N, McDermott J, Chun MM. The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception. *J Neurosci* 1997 ; 17 : 4302-11.
3. Haxby JV, Gobbini MI, Furey ML, et al. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 2001 ; 293 : 2425-30.
4. Dehaene S, Cohen L. The unique role of the visual word form area in reading. *Trends Cogn Sci* 2011 ; 15 : 254-62.
5. Parvizi J, Jacques C, Foster BL, et al. Electrical stimulation of human fusiform face-selective regions distorts face perception. *J Neurosci* 2012 ; 32 : 14915-20.
6. Logothetis NK. What we can do and what we cannot do with fMRI. *Nature* 2008 ; 453 : 869-78.
7. Hubel DH, Wiesel TN. Receptive fields of single neurones in the cat's striate complex. *J Physiol* 1959 ; 148 : 574.
8. Dehaene S, Cohen L, Sigman M, Vinckier F. The neural code for written words: a proposal. *Trends Cogn Sci* 2005 ; 9 : 335-41.
9. Tsao DY, Freiwald WA, Tootell RBH, Livingstone MS. A Cortical Region Consisting Entirely of Face-Selective Cells. *Science* 2006 ; 311 : 670-4.
10. Chang L, Tsao DY. The Code for Facial Identity in the Primate Brain. *Cell* 2017 ; 169 : 1013-28.e14.
11. Quiroga RQ, Reddy L, Kreiman G, et al. Invariant visual representation by single neurons in the human brain. *Nature* 2005 ; 435 : 1102-7.
12. Hesse JK, Tsao DY. The macaque face patch system: A turtle's underbelly for the brain. *Nat Rev Neurosci* 2020 ; 21 : 695-716.
13. Chang L, Egger B, Vetter T, Tsao DY. Explaining face representation in the primate brain using different computational models. *Curr Biol* 2021 ; 31 : 2785-95.e4.
14. Bao P, She L, McGill M, Tsao DY. A map of object space in primate inferotemporal cortex. *Nature* 2020 ; 583 : 103-8.
15. Yi HG, Leonard MK, Chang EF. The Encoding of Speech Sounds in the Superior Temporal Gyrus. *Neuron* 2019 ; 102 : 1096-110.
16. Makin JG, Moses DA, Chang EF. Machine translation of cortical activity to text with an encoder-decoder framework. *Nat Neurosci* 2020 ; 23 : 575-82.

TIRÉS À PART

J.P. Henry



Tarifs d'abonnement m/s - 2022

Abonnez-vous
à *médecine/sciences*

> Grâce à m/s, vivez en direct les progrès
des sciences biologiques et médicales

Abonnez-vous sur
www.medecinesciences.org