

CHAPITRE III

La reconnaissance de formes

Chapitre III : La reconnaissance de formes

1 La reconnaissance des formes : un processus complexe et fondamental

La reconnaissance des formes est un mécanisme essentiel du système perceptif humain, nous permettant d'identifier et de comprendre les objets dans notre environnement, souvent sans effort conscient. Pourtant, ce processus, apparemment simple, repose sur des mécanismes cognitifs complexes qui défient souvent les tentatives de reproduction par des systèmes artificiels.

1.1 La reconnaissance des formes dans le quotidien

Le texte introduit la reconnaissance des formes par un exemple simple : lire une étiquette de prix dans un magasin. Cette tâche, souvent banale pour les humains, s'avère être un défi pour les systèmes artificiels, comme les lecteurs optiques de codes-barres, qui peuvent échouer si l'étiquette est mal imprimée ou si le dispositif n'est pas correctement aligné. L'humain, en revanche, est capable de surmonter ces obstacles avec une aisance étonnante, soulignant l'efficacité de notre système de reconnaissance des formes.

1.2 L'efficacité du cerveau humain dans la reconnaissance des formes

Ce qui distingue l'humain des systèmes artificiels, c'est la flexibilité et la robustesse de notre capacité à reconnaître des formes dans une variété de contextes. Nous pouvons identifier des objets indépendamment de leur orientation, de leur position ou des variations dans leur apparence. Par exemple, un livre reste reconnaissable qu'il soit ouvert, fermé, à plat ou rangé dans une bibliothèque. Même si une partie de l'objet est cachée par d'autres objets ou si l'objet est vu sous un angle inhabituel, nous pouvons toujours l'identifier sans grande difficulté.

1.3 Les défis de la reconnaissance artificielle

Les systèmes artificiels, comme les dispositifs de lecture d'étiquettes dans les magasins, illustrent les défis posés par la reconnaissance des formes dans des environnements non contrôlés. Contrairement à l'humain, ces systèmes doivent souvent faire face à des limitations techniques : la qualité de l'impression, l'alignement du lecteur optique, et la précision des capteurs peuvent interférer avec leur capacité à

Chapitre III : La reconnaissance de formes

lire correctement les caractères. Cela met en évidence la complexité du problème de reconnaissance que l'humain résout de manière quasi automatique.

1.4 Les bases cognitives de la reconnaissance des formes

Pour comprendre la reconnaissance des formes chez l'humain, il est nécessaire de considérer ce que représente une « forme » du point de vue cognitif. La reconnaissance des formes commence après l'enregistrement sensoriel des caractéristiques physiques d'un objet (comme ses contours, sa texture, ou sa couleur). Une fois que ces caractéristiques ont été enregistrées, le cerveau peut les analyser et, si l'attention y est portée, les identifier. Cela distingue la reconnaissance des formes d'autres types de perceptions, comme celle du mouvement ou de la distance.

Cependant, la reconnaissance des formes doit également tenir compte du fait que la forme d'un objet peut être modifiée par des facteurs externes tels que le mouvement ou la distance. Par exemple, un objet peut apparaître plus petit ou déformé si on s'en éloigne ou si l'on change d'angle de vue. Malgré ces variations, le cerveau humain parvient à ajuster ses perceptions et à reconnaître l'objet pour ce qu'il est réellement, en s'appuyant sur des mécanismes de correction cognitive.

1.5 L'importance de la reconnaissance des formes dans la perception

La reconnaissance des formes est l'une des capacités les plus fondamentales du cerveau humain. Elle joue un rôle central non seulement dans la perception visuelle, mais aussi dans des activités plus complexes, comme la lecture, la navigation dans un environnement ou la reconnaissance de visages. La flexibilité de cette capacité permet à l'humain d'interagir efficacement avec un monde visuel complexe, où les objets apparaissent sous des formes et des angles variés.

2 La perception des formes dans l'environnement

Dans le domaine de la psychologie cognitive, la perception des formes visuelles est un processus fondamental pour reconnaître et interpréter les objets qui nous entourent. Une forme visuelle est définie comme une région dans l'espace délimitée par des contours. Les contours, eux, représentent des changements marqués dans

Chapitre III : La reconnaissance de formes

l'intensité lumineuse réfléchi par une surface. Ce phénomène est observé dans notre quotidien, notamment lorsque nous lisons du texte imprimé. Les lettres que nous voyons, par exemple, sont perçues comme des figures distinctes grâce aux contrastes entre l'encre noire et le papier blanc qui constituent leurs contours.

Cependant, la perception d'un objet ne se limite pas à ses contours. En effet, l'analyse des variations de **luminance** (la quantité de lumière qu'une surface reflète) joue un rôle essentiel. La luminance dépend de l'éclairage et de la réflectance de la surface, c'est-à-dire sa capacité à renvoyer la lumière. Cela signifie que notre perception d'une couleur ou d'une nuance de gris peut varier en fonction de l'éclairage de la scène. Dans un exemple simple, un gris apparaîtra plus clair ou plus foncé selon la lumière disponible et la capacité de la surface à réfléchir cette lumière. Outre les contours et la luminance, d'autres caractéristiques comme la couleur et la texture peuvent aussi influencer l'analyse d'une image visuelle. Toutefois, dans la plupart des cas, la forme d'un objet suffit souvent à son identification. Cette capacité de reconnaissance se base sur une économie cognitive où l'esprit peut utiliser la forme générale d'un objet pour en déduire son identité, même dans des situations où d'autres informations visuelles (comme la couleur) sont absentes, comme dans le cas des dessins en noir et blanc des bandes dessinées.

2.1 L'appariement à un gabarit

Un modèle classique de la reconnaissance des formes repose sur l'idée de **l'appariement à un gabarit**. Ce modèle stipule que, pour reconnaître une nouvelle forme, nous la comparons à des formes mémorisées en tant que gabarits dans notre mémoire. Par exemple, après avoir appris la forme de la lettre « A », nous pouvons identifier toute nouvelle occurrence de cette lettre en appariant sa forme à ce modèle mémorisé. Ce processus est une forme de **template-matching**, où l'esprit vérifie le degré de correspondance entre une forme donnée et un modèle stocké.

Bien que ce mécanisme fonctionne pour certaines formes et dans des conditions précises, sa généralisation est limitée. Certains systèmes artificiels de reconnaissance de formes, comme ceux utilisés dans la lecture optique des chiffres imprimés sur des

Chapitre III : La reconnaissance de formes

chèques, se basent effectivement sur ce principe. Ces systèmes fonctionnent en comparant les formes standardisées des chiffres à des gabarits pré-enregistrés. Toutefois, pour des tâches plus complexes, comme la reconnaissance de l'écriture manuscrite, ce modèle rencontre des difficultés, car les formes manuscrites varient énormément d'une personne à l'autre, tant par la taille, l'orientation, que la forme des lettres.

2.2 L'analyse des caractéristiques

L'appariement à un gabarit est souvent jugé insuffisant pour expliquer la complexité de la reconnaissance des formes. Une approche alternative est l'**analyse des caractéristiques**. Cette approche propose que les formes sont reconnues non pas par leur contour global, mais par leurs **caractéristiques élémentaires**, c'est-à-dire des traits spécifiques tels que des lignes droites, des courbes, des intersections, ou des orientations particulières. Selon cette théorie, un objet, ou une lettre comme le « T », est identifié par la présence de ses éléments distinctifs : une barre verticale et une barre horizontale qui se croisent.

Ce modèle de caractéristiques, soutenu par des chercheurs comme **Eleanor Gibson** (1969), suggère que la reconnaissance des objets est le résultat d'une analyse des caractéristiques visuelles spécifiques qui sont stockées en mémoire sous forme de listes. Ces caractéristiques sont ensuite comparées avec les stimuli perçus pour identifier les objets. Par exemple, les lettres de l'alphabet peuvent être décrites par des caractéristiques telles que l'orientation des lignes, la présence d'intersections, ou la symétrie, comme illustré dans la figure des lettres de Gibson. Les résultats d'expériences menées par Gibson montrent que le temps de réaction des participants à reconnaître des lettres dépend de la similarité des caractéristiques partagées par les lettres, ce qui renforce l'idée que la reconnaissance des formes se base en grande partie sur l'analyse de caractéristiques visuelles.

Des études en **neurosciences** menées par **David Hubel** et **Torsten Wiesel** (1962) appuient cette théorie en montrant que certaines cellules du cortex visuel réagissent spécifiquement à des caractéristiques visuelles élémentaires comme des barres ou des

Chapitre III : La reconnaissance de formes

bords orientés de manière précise. Ces cellules, appelées **détecteurs de caractéristiques**, montrent que le cerveau est effectivement équipé pour analyser les formes en fonction de leurs caractéristiques visuelles de base.

Cependant, malgré le soutien apporté par ces études, ce modèle présente des limites. Par exemple, il est difficile de décrire précisément une forme naturelle, comme celle d'un cheval, en termes de caractéristiques élémentaires. Les formes naturelles sont souvent trop complexes pour être représentées uniquement par des combinaisons de lignes et de courbes. De plus, la flexibilité et la rapidité avec lesquelles les humains reconnaissent les formes suggèrent que d'autres mécanismes sont également à l'œuvre dans le processus de perception visuelle.

3 Reconnaissance par composantes : Approche théorique et expérimentale

3.1 Introduction

L'étude de la reconnaissance visuelle des objets occupe une place centrale en psychologie cognitive. Une des théories les plus influentes dans ce domaine est celle de **Biederman** (1987), qui propose que les objets peuvent être décrits et reconnus grâce à un nombre limité de formes géométriques élémentaires, appelées *géons*. Cette approche, appelée modèle de **reconnaissance par composantes** (recognition-by-components), postule que la décomposition des objets en un petit ensemble de formes primitives permet leur identification, même dans des conditions d'observation difficiles ou incomplètes. Cette théorie repose sur une analogie avec le langage, où un nombre restreint de phonèmes suffit à générer une grande variété de mots.

3.2 Le Modèle de reconnaissance par composantes

Biederman soutient que tout comme environ 44 phonèmes permettent de coder tous les mots du vocabulaire anglais, il existe environ 36 *géons* qui suffisent pour représenter la majorité des formes rencontrées quotidiennement. Ces *géons* sont des formes géométriques tridimensionnelles simples telles que des cylindres, des cônes et des blocs. Ils jouent un rôle fondamental dans la reconnaissance des objets, car ils sont les "briques" de base combinables pour générer des représentations complexes.

Chapitre III : La reconnaissance de formes

Le processus de reconnaissance, selon cette théorie, commence par l'identification des contours de l'objet observé, qui est ensuite segmenté en *composantes* sur la base de principes d'organisation issus de la psychologie de la Gestalt. Les *géons* identifiés sont ensuite comparés à des représentations stockées en mémoire à long terme, permettant ainsi l'identification de l'objet. Cette approche est similaire au traitement du langage où des unités comme les phonèmes sont reconnues puis combinées pour comprendre des mots et des phrases même dans des environnements bruités ou partiellement masqués.

3.3 Primitives visuelles et reconnaissance d'objets

Les **géons** peuvent être comparés aux couleurs primitives dans le domaine visuel : tout comme un nombre limité de couleurs de base permet de générer l'ensemble des couleurs perceptibles, les *géons* permettent de décrire tous les objets à partir de leurs formes élémentaires. Cette simplification théorique présente plusieurs avantages. D'une part, elle explique la robustesse du processus de reconnaissance visuelle, notamment lorsque les objets sont partiellement cachés ou observés sous différents angles. Par exemple, une tasse peut être définie par deux géons principaux : un cylindre et un arc (l'anse). Peu importe son orientation, l'identification reste possible car les relations entre ces géons demeurent stables.

3.4 Expérimentation sur la perception d'objets dégradés

Pour tester l'importance de la segmentation des objets en *géons*, **Biederman** a mené des expériences visant à comprendre comment la reconnaissance des objets est affectée lorsqu'ils sont dégradés. Dans cette expérience, les objets sont représentés avec certaines parties du contour supprimées. Les résultats de l'expérience montrent que lorsque les éléments critiques pour la segmentation des formes, comme les points de jonction entre les géons, sont préservés, les objets restent reconnaissables. C'est ce que l'on appelle des stimuli *recupérables*. À l'inverse, lorsque ces caractéristiques sont supprimées, les objets deviennent *irrecupérables*, c'est-à-dire difficilement identifiables.

Chapitre III : La reconnaissance de formes

L'expérience consistait à présenter 35 stimuli à des participants, dont la tâche était de les identifier. Parmi ces stimuli, la moitié était récupérable et l'autre moitié non. Les résultats montrent que les objets récupérables sont beaucoup plus facilement identifiables que les objets irrécupérables, même lorsque les participants ont été familiarisés avec les objets dans leur forme intacte avant l'expérience. Ainsi, même après avoir été exposés aux formes intactes, les participants commettaient encore plus de 60 % d'erreurs sur les objets irrécupérables.

3.5 Importance théorique et implications cognitives

Cette expérience souligne le rôle fondamental de la segmentation des objets en composantes pour la reconnaissance visuelle. Lorsque cette segmentation est entravée, l'identification devient difficile, voire impossible. Ces résultats ont des implications importantes pour notre compréhension du fonctionnement cognitif. Ils montrent que notre capacité à identifier les objets repose sur des processus cognitifs qui décomposent visuellement les objets en unités plus petites et plus simples, qui sont ensuite reconstituées à partir de la mémoire à long terme.

De plus, cette théorie a des applications dans divers domaines, comme la reconnaissance automatique des formes en intelligence artificielle, où des systèmes de vision artificielle s'inspirent de ces principes pour segmenter et identifier des objets complexes à partir d'éléments simples.

4 Le rôle du contexte dans la reconnaissance des formes

4.1 Introduction

La reconnaissance des formes est un processus central en psychologie cognitive, car elle permet à un individu d'identifier et de catégoriser les objets qui l'entourent. Cependant, les théories classiques comme l'appariement à un gabarit, l'analyse des caractéristiques, et la segmentation en composantes négligent un aspect fondamental : l'influence du contexte et des attentes sur ce processus. En effet, il est courant que des objets ou des personnes ne soient pas immédiatement reconnus lorsque leur présence

Chapitre III : La reconnaissance de formes

dans un environnement est inattendue, comme l'illustre l'exemple familier de ne pas reconnaître une personne dans un cadre inhabituel.

4.2 L'effet du contexte et des attentes

L'influence du contexte dans la reconnaissance des formes résulte de l'interaction entre ce que l'individu perçoit et ce qu'il s'attend à percevoir. Cela peut être illustré par des situations quotidiennes où la reconnaissance d'objets familiers est retardée lorsque ces derniers apparaissent dans un contexte inattendu. Ce phénomène est expliqué par des processus cognitifs qui ajustent la reconnaissance à partir de schémas préexistants basés sur l'expérience et les attentes de l'individu.

4.3 L'effet d'appréhension du mot

L'effet d'appréhension du mot, rapporté pour la première fois par James McKeen Cattell en 1885, est une manifestation classique de l'influence du contexte dans la reconnaissance. Cet effet décrit le fait que les lettres sont plus facilement reconnues lorsqu'elles forment un mot significatif que lorsqu'elles sont disposées de manière aléatoire. Par exemple, lors d'une présentation très rapide (environ 10 millisecondes), les sujets peuvent identifier les lettres d'un mot entier comme "RAME", alors qu'ils ne sont capables d'identifier que quelques lettres dans une série aléatoire comme "SPBO". Ce phénomène peut être interprété par le rôle facilitateur du contexte : un mot permet au cerveau de compléter les informations manquantes ou peu claires, en s'appuyant sur des connaissances antérieures.

4.4 Explications alternatives de l'effet d'appréhension du mot

Plusieurs explications sont avancées pour interpréter cet effet. D'une part, on peut supposer que les participants devinent les lettres manquantes en fonction des indices contextuels. Par exemple, si un sujet voit "M_T" et a conscience de la possibilité du mot "MAT", il est probable qu'il devine la lettre manquante "A" même s'il ne l'a pas vue directement. Une autre hypothèse est que la mémorisation est facilitée lorsque les lettres sont organisées en mots plutôt qu'en séries aléatoires. Cela signifie que les sujets perçoivent toutes les lettres pendant la présentation, mais l'organisation en mots facilite la rétention d'un plus grand nombre de lettres.

Chapitre III : La reconnaissance de formes

4.5 Expérience de Reicher (1969) : Vérification des hypothèses

Pour tester ces hypothèses, Reicher a mené une expérience en 1969 visant à contrôler les tentatives de deviner et à évaluer l'effet du contexte sur la reconnaissance. Dans cette expérience, les participants devaient identifier des lettres à partir d'une série présentée rapidement. Les résultats ont montré que les performances étaient systématiquement meilleures lorsque les lettres formaient des mots, ce qui confirme l'effet facilitateur du contexte.

4.6 Le modèle d'activation interactive de McClelland et Rumelhart

Pour expliquer en profondeur ce phénomène, James McClelland et David Rumelhart ont proposé le modèle d'activation interactive en 1981, un cadre théorique basé sur l'approche connexionniste. Ce modèle repose sur l'idée que la reconnaissance de formes implique des interactions dynamiques entre différentes unités cognitives, organisées en trois niveaux : les caractéristiques visuelles, les lettres, et les mots.

Le modèle postule que chaque unité (caractéristique visuelle, lettre ou mot) a une certaine valeur d'activation, représentant la probabilité qu'elle soit présente dans le stimulus perçu. L'activation se propage entre les unités à travers des connexions excitatrices et inhibitrices. Par exemple, la perception d'une barre verticale active l'unité correspondant à une caractéristique visuelle simple, qui à son tour active une lettre comme le "A". Si cette lettre fait partie d'un mot significatif comme "AMOUR", l'activation du mot renforce rétroactivement l'activation de la lettre, augmentant ainsi la probabilité que le "A" soit perçu correctement.

Cette rétroactivation est essentielle, car elle permet au système cognitif de compenser des informations incomplètes ou ambiguës. En d'autres termes, si les caractéristiques perçues sont compatibles avec un mot plausible, cela renforce la reconnaissance des lettres constitutives de ce mot.

4.7 Inhibition et interaction

Lorsque les hypothèses sont incohérentes avec les caractéristiques perçues, le système inhibe les unités incompatibles. Par exemple, si le mot "AMOUR" est activé,

Chapitre III : La reconnaissance de formes

les lettres comme "P" ou "W", qui ne peuvent pas former ce mot, seront inhibées. De même, si une caractéristique visuelle spécifique (comme une courbe) est perçue, les lettres incompatibles seront écartées. Ainsi, ce modèle montre que la reconnaissance de formes est un processus interactif, où chaque niveau de perception influence les autres, permettant une identification plus efficace des stimuli.