

Mauro GAIO, Jacques MADELAINE*

Un modèle de l'illusion perceptive : simulation versus expérimentation

Cet article présente une expérimentation où l'ordinateur a le double rôle de sujet et d'outil de simulation. L'ordinateur a été utilisé pour l'élaboration et la validation d'un modèle de fonctionnement de certains principes de structuration de la perception visuelle humaine. Ce modèle est basé sur la théorie de la forme ou *Gestalt*. Le processus de structuration étudié est considéré comme un système dynamique. Ce système dynamique est émulé sur ordinateur grâce à des réseaux récurrents. Cette simulation a permis de concevoir et de valider le modèle par expérimentation sur des classes d'images à illusion perceptive primaire. Les acquis ont été ensuite réinvestis dans la conception d'un module d'identification des structures linéaires prégnantes sur des photographies aériennes.

Mots-clés : perception visuelle, Gestalt, système dynamique, groupements perceptifs; modélisation mathématique.

A model of perceptive illusion: experimentation versus simulation.

This paper presents an experimentation where the computer has the role of subject and simulation tool. The computer has been used for the elaboration and validation of an operating model of a few structural principles of human visual perception. This model is based on the Gestalt theory. The studied structural process is considered as a dynamic system. Wich is emulated by the computer with recurrent networks. Different experiments on a specific class of primary perceptive illusion images allowed the conception and the validation of a model of perception. The principle of this model have been re-used as the basis for the conception of a program for identification of global linear structures in landscape aerial images.

* Mauro Gaio, Jacques Madelaine: GREYC - URA 1526, Université de Caen.
Mail : Mauro.Gaio@info.unicaen.fr, Jacques.Madelaine@info.unicaen.fr
Tél : 31 45 59 66 - Fax : 31 45 58 14

Key Words : visual perception, Gestalt, dynamic system, perceptual grouping, mathematical modelisation.

1. INTRODUCTION

L'expérimentation construite autour de la modélisation de certains mécanismes de structuration de la perception visuelle avait une triple finalité. Elle devait être pour la psychologie expérimentale une aide à la compréhension de l'illusion perceptive. Elle devait également confirmer ou infirmer les présupposés élaborés dans la modélisation de l'interprétation géographique d'images aériennes [Gaio 94]. Enfin, elle devait montrer qu'il était possible de construire un outil informatique capable de simuler certaines propriétés des systèmes dynamiques continus. Dans un tel contexte, l'interaction des trois disciplines concernées, l'informatique, la géographie et la psychologie expérimentale était incontournable.

Pour éclairer les différents aspects de notre démarche nous commencerons par rappeler les principes empruntés à la théorie de la forme qui ont permis de construire puis d'analyser des situations d'observations pertinentes. Nous évoquerons par ailleurs les principes mathématiques qui ont servi pour construire une modélisation opérationnelle des différents processus observés. Les questions suscitées par le choix de l'outil informatique répondant au mieux aux contraintes du modèle seront le fil conducteur pour montrer la place de l'ordinateur dans notre démarche. Place où l'ordinateur joue constamment un double rôle : celui d'outil de simulation et celui de sujet (dans le sens de la psychologie expérimentale).

2. QUELQUES PRINCIPES EMPRUNTES A LA *GESTALT*

Un des concepts fondateurs de la théorie de la forme (*Gestalt*) peut être rapidement résumé par l'idée qu'il faut partir d'une description locale des composants d'une perception, puis établir les interactions potentielles inter-composants.

Ce concept s'appuie sur le constat qu'il n'est pas possible *a priori* de prévoir, à partir de l'analyse isolée des interactions locales, la perception globale que l'on peut avoir d'une image. La *Gestalt* propose donc, par un formalisme global, l'émergence des propriétés collectives des composantes locales de l'image. Pour expliquer la décomposition en objets distincts, les gestaltistes suggèrent que la perception visuelle de l'homme organise les éléments perçus dans des ensembles selon des lois de regroupements. Un autre aspect important qui est mis en évidence est l'explicitation des relations entre l'objet perçu et le fond; on peut ainsi déterminer dans quelles conditions une région est perçue comme un objet ou comme un fond. Autrement dit les gestaltistes offrent un cadre théorique permettant de montrer que les formes

illustrent les propriétés prégnantes de l'image perçue, ce qui implique que les qualités globales d'une sensation transcendent celles de ces composants [Guillaume 79]. Prenons comme exemple la figure 1(a) elle représente quatre formes indépendantes sur un fond visible.

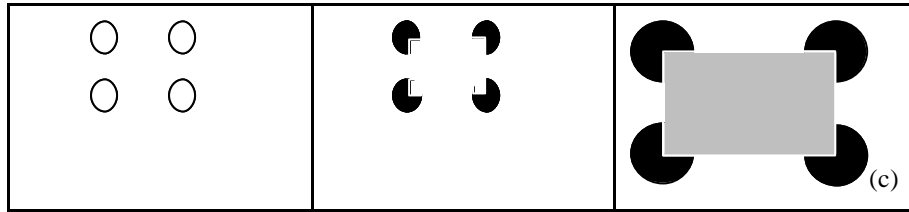


figure 1 : Lois des rapports entre fond et forme

La figure 1(b) représente quatre formes qui par leurs caractères objectifs de structuration provoquent l'émergence d'une nouvelle forme, cette forme est "matérialisée" par un rectangle grisé sur la figure 1(c).

Ces lois, malgré une définition quantitative encore non-formulée, nous ont permis de réaliser une description suffisamment précise des objets. D'autre part, à partir de ces principes théoriques, nous avons pu établir une approximation formelle des effets de champ et des relations nécessaires à notre objectif.

Les principes proposés par les gestaltistes nous permettent également de définir les contraintes qui devront être satisfaites lors du choix de notre terrain d'expérimentation. Après concertation avec les psychologues, nous avons choisi un terrain très spécifique : l'émergence de phénomènes d'illusion perceptive. D'un premier abord celui-ci semble être fort éloigné de nos préoccupations, alors que du point de vue des mécanismes perceptifs stimulés il se retrouve au cœur de notre questionnement.

Ainsi, nous avons choisi de passer par des figures à illusion perceptive primaire pour la mise au point d'un outil d'extraction automatique de structures linéaires prégnantes ; structures qui, présentes dans des images géographiques de paysage, sont apparues comme essentielles dans la démarche d'interprétation de ces images. La manière de s'attaquer à un tel problème n'a pu être élaborée que grâce à la convergence des différentes sensibilités scientifiques impliquées dans ce travail.

2.1. Une terrain d'expérimentation : l'illusion perceptive d'émergence de forme

On appelle figures à illusion perceptive les images qui provoquent une distorsion systématique de l'impression perceptive par rapport à la réalité. Cette classe d'images a été isolée à partir de nombreux travaux en psychologie expérimentale qui portent sur le thème de l'induction de contours et de surfaces. La plupart de ces travaux s'appuient, entre autre, sur le concept de contours illusoires pour élaborer des modélisations de la perception visuelle humaine [Meyer&Petry 86]. Nous disposons ainsi de plusieurs descriptions de l'illusion perceptive. Glenn E. Meyer et Susan Petry dans l'introduction de l'ouvrage qu'ils ont co-dirigé [Meyer&Petry 87] nous en livrent une synthèse qui tient en trois points :

- il existe une impression de surface délimitée qui possède des propriétés telles qu'elle se différencie par rapport aux surfaces voisines ;
- il existe une impression de limite ou de bord autour de cette surface. Le bord peut être net ou diffus ;
- une des propriétés fondamentales est établie par la perception du bord et de la surface induite au-delà des discontinuités des figures inductrices.

D'après ces observations, le message visuel est la résultante de l'action combinée des différentes impressions mais surtout des facteurs, dits *structurants*, qui leurs sont sous-jacents. Le relativement faible nombre de facteurs perceptifs impliqués explique le choix de cette classe d'images comme terrain pour notre expérimentation. Mais ces facteurs peuvent dépendre aussi bien des caractéristiques physiques des objets perçus (facteurs objectifs), que de leur éventuelle signification fonctionnelle (facteurs subjectifs). Nous avons donc réduit encore un peu plus notre terrain expérimental et l'avons limité aux seules figures à illusion perceptive dite *primaire* [Piaget 36]. Ces figures résultent essentiellement des effets de la morphologie et de la répartition spatiale des objets inducteurs : elles permettent ainsi de limiter les effets des facteurs subjectifs. Les travaux réalisés en psychologie expérimentale nous donnent non seulement une définition de l'illusion perceptive, mais mettent également à notre disposition un matériel expérimental approprié qui provoque l'illusion attendue. Ce matériel est constitué par des figures qui, par la pauvreté sémantique de leur message visuel, permettent d'isoler les mécanismes perceptifs responsables de l'induction du bord au delà des discontinuités et donc de *l'illusion*. Les mécanismes que nous voulons appréhender grâce à cette classe d'images sont activés par *les effets de champ*, en d'autres

termes par les relations qui existent entre les différents éléments de champ perceptif.

Les effets de champ sont des phénomènes très fréquents en imagerie aérienne. Les structures prégnantes de ces images sont en effet affectées de nombreuses discontinuités. Ces discontinuités proviennent, soit de la projection en deux dimensions d'objets à trois dimensions, phénomène renforcé par l'angle de la prise de vue (obligatoirement vertical), soit d'un phénomène de mimétisme provoqué par la transformation en niveaux de gris d'un monde en couleur. Par exemple, une route peut disparaître sous un couvert végétal (lors du passage dans un bois), ou alors elle disparaît le temps de son passage entre deux labours (la route ayant acquis le même niveau de gris que les labours qui la côtoient). Ainsi, une part importante dans la démarche interprétative de l'observateur humain passe par l'explicitation de ces discontinuités, et donc par la mise en œuvre des mécanismes d'induction de contours.

Dans notre choix de figures à illusion perceptive, ce n'est donc pas le phénomène de l'illusion qui nous intéresse. Ce sont essentiellement les mécanismes qui sont déclenchés lorsqu'on se trouve en présence des facteurs structurants qui sont sous-jacents à l'illusion.

Nous avons choisi de construire la modélisation mathématique de ces mécanismes par des systèmes dynamiques. Ceci nous permet de représenter certaines des notions dynamiques qui sont contenues dans la théorie de la *Gestalt* et que nous venons de rappeler brièvement.

3. UNE THEORIE MATHÉMATIQUE : LES SYSTEMES DYNAMIQUES

Le modèle mathématique qui va nous servir à représenter de manière opérationnelle les concepts gestaltistes retenus s'appuie sur la théorie des systèmes dynamiques. Nous allons donc d'abord présenter très brièvement les principaux éléments de cette théorie avant de montrer comment l'on peut construire un modèle de la *Gestalt* dans ce cadre.

Un système dynamique est caractérisé par la donnée d'un espace, le plus souvent appelé espace des états, et d'une loi qui décrit l'évolution de chaque état en fonction d'un paramètre t . Quand la loi d'évolution ne dépend pas explicitement du paramètre, on dit que le système dynamique est *autonome* : c'est le seul cas qui nous intéressera ici. Les systèmes dynamiques les plus utilisés dans les sciences physiques sont des systèmes continus : on entend par là que l'espace des états est un espace continu (en toute généralité une variété différentielle), et que le paramètre t , est un réel, qui modélise le plus souvent le temps. La loi

d'évolution est alors entièrement définie par la donnée d'un champ de vecteurs sur l'espace des états, la valeur du champ en chaque point représentant la vitesse à laquelle évolue l'état correspondant.

L'étude d'un système dynamique consiste essentiellement à déterminer les trajectoires dans l'espace des états, c'est-à-dire les différentes courbes que parcourt un point de l'espace des états qui obéit à la loi d'évolution, à partir des différents états initiaux possibles. L'un des aspects les plus importants de l'étude d'un système dynamique concerne son comportement dit asymptotique, c'est-à-dire quand t tend vers l'infini. En effet, si l'on considère que l'on modélise un système qui évolue dans le temps, connaître le comportement asymptotique revient à prédire dans quel(s) état(s) se trouvera, à long terme, un système qui évolue à partir d'un état initial connu. Les comportements asymptotiques les plus simples sont ceux où le système tend vers un état d'équilibre, que l'on appelle un attracteur ponctuel : l'ensemble des conditions initiales conduisant à cet état s'appelle alors *le bassin* de cet attracteur.

Les systèmes dynamiques les plus simples sont caractérisés par l'existence d'un nombre fini d'attracteurs ponctuels isolés dont les bassins forment une partition (à un ensemble de mesure nulle près) de l'espace des états. Dans ce cas, pour presque toutes les conditions initiales, la dynamique du système le conduit à se stabiliser dans l'un de ces états attracteurs. Mais tous les systèmes dynamiques ne sont pas aussi simples. Par exemple, il peut exister des attracteurs cycliques : on entend par là une trajectoire fermée (un cycle) qui est en même temps un attracteur, au sens où les trajectoires issues de points suffisamment proches tendent vers ce cycle. L'évolution d'un système dynamique dont l'état initial est dans le bassin de l'attracteur d'un tel cycle ne le conduit donc pas à se stabiliser dans un état d'équilibre, mais au contraire à osciller indéfiniment en parcourant une trajectoire de plus en plus proche du cycle limite en question. D'autres cas de figure beaucoup plus complexes sont encore possibles : attracteurs constitués de sous-variétés de dimension supérieure à un, attracteurs "étranges", dynamiques chaotiques, etc.¹. Il n'est pas question de nous étendre davantage ici sur ces questions : comme on le verra, nous nous intéresserons essentiellement, pour des raisons d'implémentation informatique, à des approximations discrètes de systèmes dynamiques relativement simples.

¹ Pour une présentation générale des systèmes dynamiques, voir par exemple M.W. Hirsch et S. Smale [Hirsch & Smale 74] ou H. Reinhart [Reinhart 82].

Le seul but de cette présentation très sommaire était de donner une idée du cadre général dans lequel s'inscrit ce genre de modèle, et d'introduire le vocabulaire qui nous sera nécessaire par la suite lorsque sera évoquée la modélisation de la perception visuelle par des systèmes dynamiques.

4. ANALYSE GESTALTISTE, MODELE MATHEMATIQUE ET OUTILS INFORMATIQUES

La modélisation de certains mécanismes perceptifs de structuration que nous proposons consiste à associer à chaque image un système dynamique dans un espace perceptif donné. L'idée principale du modèle est de considérer que, chaque fois que l'on présente une image, se déclenche un processus de perception représenté par l'évolution dynamique d'un "état" perceptif : si ce processus aboutit à un attracteur ponctuel, on dira que cet état stable correspond à la perception de l'image présentée. Bien entendu la loi d'évolution du système dynamique est déterminée dans le modèle par l'application des lois de la *Gestalt*.

4.1. Caractérisation gestaltiste et modélisation mathématique

Prenons un exemple simple. Soit une image constituée simplement de quelques points approximativement alignés verticalement, et supposons que la question qui nous importe soit de savoir si ces points sont perçus comme trois points isolés ou au contraire comme regroupés et constituant alors une ligne pointillée. Si c'est la seule "dimension perceptive" qui nous intéresse, nous pouvons représenter la situation par un espace perceptif uni-dimensionnel, par exemple l'intervalle $[0, 1]$, la valeur sur cet intervalle représentant une sorte de degré de "regroupement perceptif" des points : les valeurs proches de 0 représentant des perceptions de points complètement isolés et les valeurs proches de 1 des perceptions de ligne pointillée. L'idée fondamentale de notre modélisation est de faire correspondre alors aux lois de la *Gestalt* appliquées à ce cas (lois de similarité, de proximité, etc.) la définition d'une dynamique sur cet espace perceptif, qui fasse évoluer l'état perceptif du sujet. Supposons par exemple que les points soient parfaitement alignés, assez proches les uns des autres, régulièrement espacés, et de même taille (figure 2 (a)) : la dynamique correspondante aura alors un seul attracteur situé sur la valeur 1 (perception d'une ligne). Autrement dit, quel que soit l'état perceptif initial du sujet (qui peut dépendre de ce qu'il a vu auparavant, de ce qu'il a anticipé, etc.), la perception qui finira par s'imposer et par se stabiliser est celle d'une ligne pointillée. Au contraire, si les points sont

irrégulièrement espacés, de tailles très dissemblables, et mal alignés (figure 2 (b)), la dynamique aura un seul attracteur situé près de 0 : la perception qui se stabilisera sera celle de points isolés.

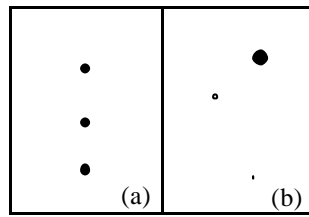


figure 2 : Perception d'une ligne pointillée.

(a) les 3 points sont verticalement alignés, régulièrement espacés et de même taille, (b) les 3 points sont mal alignés, très irrégulièrement espacés et de tailles dissemblables

L'intérêt d'une telle modélisation est double : d'une part, cela permet de rendre compte à l'aide d'un même formalisme, à la fois de cas simples comme la situation perceptive qui vient d'être décrite, et de cas plus complexes : le nombre de dimensions perceptives est bien sûr généralement assez élevé, et de plus on peut avoir des ambiguïtés qui augmentent la complexité de la dynamique sur cet espace multi-dimensionnel.

Ainsi, une ambiguïté pourra être représentée par un système dynamique à deux attracteurs (suivant les conditions initiales, la perception du sujet se stabilise sur l'une ou l'autre des deux significations possibles de la figure présentée), ou encore par une dynamique à attracteur cyclique (quand la perception oscille indéfiniment entre les deux possibilités, comme c'est le cas par exemple pour le cube de Necker, figure 3).

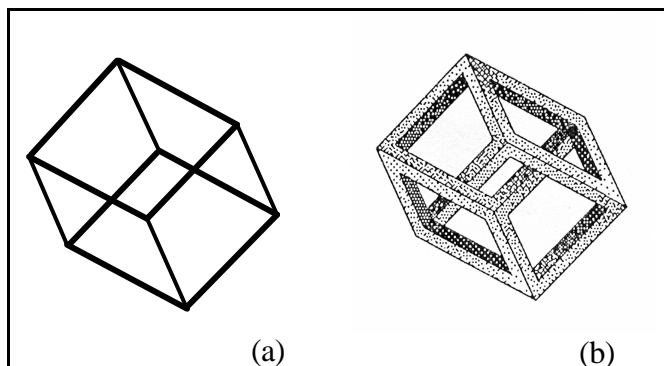


figure 3 : Le cube de Necker (a) et son interprétation par Escher(b)

D'autre part, cette modélisation est bien adaptée à l'analyse gestaltiste de la perception, en particulier au paradigme local/global (relations parties/tout) qui en constitue un aspect fondamental. En effet, s'il est assez facile d'exprimer les lois gestaltistes sous forme de contraintes que l'on peut appliquer localement aux différents éléments qui composent la figure, il est plus difficile de prédire comment ces diverses relations locales vont se combiner pour donner une perception globale. D'autant que cette combinaison doit en principe satisfaire au mieux l'ensemble des contraintes. Avec notre modèle, définir les contraintes locales revient à spécifier la dynamique : la façon dont ces contraintes interagissent globalement est alors donnée par l'évolution du système ainsi spécifié. En somme, le système dynamique peut être vu comme un système de satisfaction de contraintes, les attracteurs de la dynamique correspondant aux solutions optimales du problème de satisfaction de contraintes modélisé.

Il est intéressant de noter que cette représentation du phénomène perceptif par un système dynamique rappelle l'intuition initiale des théoriciens de la *Gestalt*, pour qui les états perceptifs devraient être "isomorphes" à des champs électriques dans le cerveau, ces champs s'établissant à partir de contraintes données par le stimulus sensitif, en fonction des lois perceptives qu'ils avaient découvertes. D'une certaine manière, notre modèle ne fait que reprendre cette idée, en supprimant l'aspect biologique (qui s'est avéré erroné, ce qui a beaucoup nuit à la théorie dans son ensemble).

Le modèle mathématique très prometteur sur le plan théorique aurait peu d'intérêt pour nous, s'il n'y avait pas la possibilité de construire une simulation pour y appliquer ses principes. Autrement dit, si l'on ne pouvait pas disposer d'un outil capable de calculer effectivement la dynamique correspondant à telle ou telle image, pour telle ou telle dimension perceptive. C'est grâce aux réseaux connexionnistes récurrents, que nous allons présenter ci-dessous, que nous avons pu implémenter le modèle théorique, et donc le rendre utilisable.

4.2. Les réseaux récurrents

En informatique, un automate est défini par des entrées, des états internes, des sorties et par deux applications : la fonction de changement d'état qui donne le nouvel état au temps $t+1$ en fonction de l'entrée et de l'état au temps t , et la fonction de sortie qui donne la sortie au temps $t+1$ en fonction de l'entrée au temps t . Il est possible avec ce formalisme de simuler des propriétés collectives comparables qualitativement à celles des systèmes dynamiques continus. Les travaux

de McCulloch et Pitts ont révélé que des équations dynamiques très simples pouvaient engendrer des comportements complexes à condition d'être combinées et incorporées dans un réseau de configuration appropriée [McCulloch & Pitts 1943]².

Un réseau récurrent est constitué d'automates recevant sur leurs entrées les états d'une partie ou de la totalité des autres cellules et parfois, des informations de l'environnement.³ Le réseau de Hopfield est un type particulier de réseau récurrent. Toutes les cellules du réseau sont à la fois cellules d'entrées au moment de l'initialisation et cellules de sorties après sa convergence. De plus, deux cellules quelconques sont liées par deux liens de même pondération. La matrice des liens de pondérations est donc symétrique avec une diagonale nulle. Ces contraintes ont permis à l'auteur de démontrer que le comportement dynamique de ce type de réseau récurrent est gouverné par une fonction de Lyapunov décroissant le long d'une trajectoire, ce qui implique que le réseau converge vers un minimum de cette fonction. Deux ans plus tard, Hopfield [Hopfield 84] a montré qu'un modèle de réseau avec une topologie des liaisons similaire, mais avec des cellules à activité continue dans un intervalle réel borné, avait des propriétés équivalentes au modèle comportant des cellules à réponse binaire. Si le codage dans les liaisons des informations issues des sources de connaissances n'est pas explicite, celui-ci peut s'obtenir implicitement par des techniques d'apprentissage. À titre d'exemple la règle de Hebb [Hebb 49] apporte une des réponses à ce type de problème. Pour notre part la phase d'apprentissage n'est pas nécessaire puisque nous disposons, comme

² Une large présentation de ces techniques est faite dans l'ouvrage collectif dirigé par Rumelhart et McClelland [Rumelhart & McClelland et al. 86]. De façon plus synthétique et sans être exhaustif, Stephen Grossberg [Grossberg 88] propose un large tour d'horizon des modèles de réseaux connexionnistes les plus usités.

³ La dynamique d'un réseau récurrent est entièrement définie par la matrice des pondérations, la fonction de transition et la définition du mode d'itération. Lorsqu'on laisse évoluer ce type de réseau, son comportement peut prendre deux formes. La première forme fait évoluer le système vers une configuration "plus forte" augmentant ainsi sa résistance aux perturbations, on dit alors que les itérations sont convergentes. La seconde forme de comportement mène le système à parcourir à l'infini toutes ou parties des valeurs du domaine. Si son comportement est de la première forme, le système atteint nécessairement un cycle qu'il parcourt ensuite indéfiniment. Nous retrouvons ici les mêmes notions que celles que nous avons évoquées pour les systèmes dynamiques. À savoir, si le cycle ne comporte qu'un seul état, il s'agit d'un état stable. Si un état stable peut être atteint à partir de plusieurs états initiaux : on dit que cet état est un attracteur [Personnaz & Dreyfus 89]. Hopfield a montré en 1982 [Hopfield 82] quelles étaient les conditions nécessaires pour garantir que les itérations soient convergentes.

nous allons le voir dans la section 5 d'une évaluation explicite de la matrice de pondérations des liens.

Dans d'autres travaux sur le groupement perceptif, les réseaux récurrents ont déjà été identifiés comme des paradigmes fort intéressants pour la mise en combinaison d'informations locales et d'informations globales [Rosenfeld & Hummel 76], [Mohan & Nevatia 88 89], [Parent & Zucker 89]. Le modèle de réseau récurrent proposé par Hopfield a lui aussi déjà été utilisé par R. Mohan et R. Nevatia qui l'ont retenu essentiellement pour permettre le passage du niveau local, modélisé par les cellules, au niveau global modélisé par les propriétés dynamiques du réseau. Nous l'avons retenu pour nous permettre d'implémenter le modèle théorique de la perception initialement proposé par les théoriciens de la *Gestalt*. Nous l'avons considéré comme un excellent outil pour nous permettre de combiner les diverses lois perceptives et de les faire opérer de façon graduelle. Nous l'avons également retenu pour ses capacités connues à résister aux modifications ponctuelles de l'environnement auquel il est rattaché, tant que ces variantes locales laissent l'ensemble proche de la référence. Prenons comme illustration le problème des structures linéaires dans une image, si une fraction des segments de droites qui construisent ces structures est occultée il n'en reste pas moins que la plupart des structures restent perceptibles. Cette perception persiste tant que la cohérence globale n'est pas modifiée. On peut considérer ces occultations comme des modifications ponctuelles de l'environnement et la cohérence globale comme la limite maximum d'éloignement de la référence. Autrement dit, il suffit dans notre illustration que lors de l'occultation des segments de droites la nouvelle information permette à la nouvelle configuration initiale du réseau d'être admise par celui-ci comme attracteur.

Si les propriétés d'un réseau récurrent, décrites ci-dessus, apparaissent comme prometteuses, la construction de l'outil de simulation préconisé pour l'expérimentation se base sur un certain nombre d'inconnues qu'il s'agit de découvrir. Pour cela, il a fallu parvenir à répondre aux questions suivantes:

- quels sont les facteurs relationnels les plus discriminants ?
- quels éléments de l'image les cellules vont elles émuler ?
- comment combiner des facteurs qualitatifs à des facteurs quantitatifs ?
- comment lit-on les résultats ?

Comment y sommes nous parvenus, encore une fois en ayant adopté une démarche expérimentale, ce qui suit tente de le montrer.

5. UN OUTIL INFORMATIQUE DE SIMULATION

L'utilisation originale du réseau de Hopfield pratiquée ici tient à fois à l'usage qui est fait du réseau et à la manière dont les pondérations des liaisons sont initialisées (pas de phase d'apprentissage). L'utilisation de figures à illusion perceptive primaire, très simples du point de vue message visuel, nous a permis d'isoler et de modéliser les relations spatiales impliquées dans les phénomènes d'induction de contours. Inductions qui sont, nous le rappelons, le stimulus résultant provoqué par la perception de structures prégnantes. Les liens sont pondérés en fonction de la distance des figures inductrices représentées par les cellules connectées et de leurs relations spatiales deux à deux ; enfin les relations ont été évaluées en s'appuyant, ici encore, sur le formalisme préconisé par les lois de la *Gestalt*.

5.1. Choix d'un matériau d'observation

Les images à illusion perceptive primaire les plus remarquables sont celles élaborées par Gaetano Kanizsa [Kanizsa55], où un certain nombre de "pacmans" produisent de façon très sensible des figures géométriques illusoire. Pacman est le nom le plus en usage pour évoquer les figures inductrices. Dans le cas de la configuration à quatre pacmans retenue pour notre expérimentation, le regroupement illusoire des huit segments de droite disjointes en quatre segments connexes induit la forme d'un rectangle. On remarque, figure 4, que dans aucune des deux configurations les pacmans ne sont connexes, mais que, dans l'une des deux, les pacmans produisent par leurs relations spatiales la perception d'une forme surfacique (figure : 4(b)).

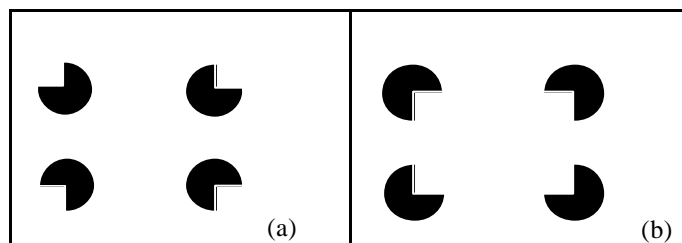


figure 4 : Pacmans et rectangle de Kanizsa

Il faut remarquer que les formes inductrices appartiennent, c'est le cas ici, à une même famille physiologique. Les éléments induits émergent du fond dans la discontinuité spatiale qui sépare les objets inducteurs, discontinuité que nous appellerons désormais : *lacune*. Les relations spatiales dans une figure de Kanizsa sont doubles, l'une est

représentée par la relative proximité d'objets de couleur fortement contrastée par rapport au fond, l'autre est provoquée par l'orientation particulière des quatre objets. Ces influences se combinent pour provoquer l'augmentation de la perception illusoire d'une luminance plus marquée [Dresp et al. 90] dans les lacunes et pour donner une impression de bord autour de cette surface. Les bords formant les limites de la forme induite sont issus d'un prolongement assez conséquent de certains des bords des objets inducteurs.

5.2. Construction de l'outil de simulation

Comme nous l'avons vu précédemment, l'évolution d'un système est conditionnée par les modifications internes qui affectent les composants et par les interactions qui s'établissent entre les relations et l'environnement. Nous avons également vu la bonne capacité des réseaux récurrents à simuler de tels systèmes, c'est donc dans cette optique que nous utilisons une variante du réseau de Hopfield. Son utilisation a été pensée pour simuler les mécanismes de structurations que certaines formes mettent en mouvement, provoquant ainsi la perception d'autres formes (les propriétés de ces dernières sont telles qu'elles n'apparaissent qu'en présence des premières), ceci pour une double finalité de compréhension et d'automatisation. L'information prototype insérée dans cette application n'est autre que la description des relations induisant les phénomènes de l'illusion perceptive. Ainsi, le réseau, n'ayant pas à apprendre les lois qui gèrent les données sur lesquelles il va devoir travailler, voit sa phase d'apprentissage remplacée par un module de définition du nombre de cellules et d'initialisation des pondérations des liens. Ce module travaille sur les éléments descripteurs, obtenus après la phase de segmentation de l'image, et il initialise les pondérations à partir d'une évaluation des relations que chaque élément descripteur d'un objet entretient avec chaque élément descripteur d'un second objet.

Le réseau implémenté utilise des cellules à activité continue et potentiellement entièrement connectées. Les cellules du réseau prennent une signification symbolique en obéissant à certains principes gestaltistes (simplicité, contiguïté et connexité). Par conséquent, si une cellule ne répond pas à un certain principe gestaltiste, celle-ci ne doit plus influencer le réseau; nous avons donc, comme Hopfield, opté pour l'intervalle $[0,1]$ comme domaine d'états pour les vecteurs. D'autre part, dans le conflit des formes possibles, le groupement se fait de façon d'autant plus importante qu'il va dans le sens de la réalisation d'une forme privilégiée (loi de la bonne forme). Dans le souci de pouvoir adhérer à cette définition symbolique de la règle d'activation nous avons

préférée une fonction définie par McClelland et Rumelhart [Rumelhart & McClelland 82] dans le modèle de réseaux IAM (Interactive Activation Model) à celle préconisée par Hopfield. L'activité de la cellule est d'autant plus poussée vers l'une des extrémités de l'intervalle que sa valeur d'entrée est forte. Autrement dit, les valeurs transportées vers chaque cellule par les liens l sont pondérées, puis sommées, pour produire la valeur d'entrée A_i de la cellule c_i . Cette valeur d'entrée est combinée avec la valeur de l'état de la cellule pour produire la nouvelle valeur d'activité selon la règle d'activation ci-après:

$$c_i(u+1) = \begin{cases} c_i(u) + k(1 - c_i(u))A_i(u) & \text{si } A_i(u) > 0 \\ c_i(u) + kc_i(u)A_i(u) & \text{si } A_i(u) \leq 0 \end{cases}$$

avec $A_i(u) = \sum_j l_{ij}c_j(u)$ et $k < \frac{1}{\max_i \left(\sum_j |l_{ij}| \right)}$.

k est une norme qui permet de conserver les valeurs des activités dans l'intervalle $[0,1]$.

Ainsi, les conditions de modifications internes du système se trouvent définies *a priori* à chaque construction du réseau. Le nombre de cellules, leur signification potentielle et leurs différentes relations font l'objet d'une spécification et d'une évaluation *a priori*. La spécification de la signification des cellules et des différentes relations a été établie une fois pour toutes pour les images composées de rectangles de Kanizsa. L'évaluation des relations retenues est pratiquée sur chaque nouvelle image à interpréter. Enfin, par le truchement d'une fonction de mise en correspondance, les liens entre cellules sont influencés par la distance qui sépare les figures inductrices que les cellules représentent. Par conséquent, les liens du réseau récurrent se confondent parfaitement avec la notion de relation définissante des composants d'un système dynamique. L'élaboration de ces données a donc doté le réseau des informations qui lui permettront d'évoluer vers des configurations plus fortes, plus stables.

5.3. Construction de la simulation pour les rectangles de Kanizsa

Le système conçu est entièrement automatisé, il est constitué de deux modules principaux. Le premier extrait une description synthétique (par primitives) des figures inductrices. Il produit également un découpage de l'image pour révéler les lacunes et il retient un point central pour chacune d'entre elles. Enfin, il évalue les relations

spatiales entre les primitives deux à deux, entre les lacunes et les pacmans. Le second module produit une mise en confrontation de l'ensemble de ces relations dans le but de faire émerger les relations structurantes.

Pour identifier, puis extraire une description synthétique et codifiable des pacmans, nous nous sommes appuyés sur les méthodes les plus utilisées en reconnaissance de formes. Pour la modélisation des relations entretenues par les figures inductrices nous nous sommes appuyés sur les lois de la *Gestalt*. Enfin, certains principes gestaltistes ont contribué au choix des différentes significations symboliques possibles que les figures inductrices, ici les "Pacmans", pouvaient acquérir.

5.3.1 Le module de description et d'évaluation des relations.

Description synthétique et codifiable des figures inductrices : les primitives de description des figures inductrices sont succinctes en ce qui concerne les images composées de rectangles de Kanizsa. Il suffit d'évaluer leur position et de déceler l'emplacement de l'ouverture. La valeur moyenne de la couleur est, d'autre part, un descripteur suffisant de leur aspect puisque les pacmans sont construits autour d'une surface homogène. Pour positionner les points servant de représentants des lacunes nous avons construit les images autour d'une grille régulière. Les pacmans se trouvent au centre des cases constituées par la grille. L'emplacement des lacunes est défini soit par l'intersection de quatre cases contenant chacune un pacman, soit par l'intersection des quatre cases ayant un élément commun avec une case contenant un pacman. La figure 5 représente quelques unes des intersections par un petit disque tramé.

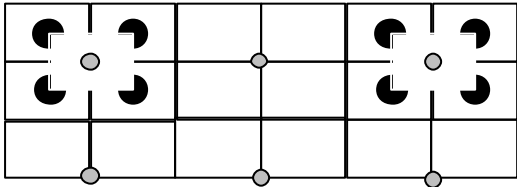


figure 5 : Position des lacunes (disques tramés)

Modélisation des relations : dans le cadre des rectangles de Kanizsa nous avons considéré que les relations définissantes sont dictées par trois propriétés des figures inductrices. Ces propriétés ont été modélisées grâce aux diverses lois de la *Gestalt* [Guillaume 79] :

- a*• la loi de la continuité exprime la capacité du contour des pacmans à induire des bords au-delà de leurs limites physiques ;
- b*• les lois de similarité et de complémentarité expriment la contribution de chaque groupement de pacmans à faire émerger une surface délimitée ;
- c*• la loi de proximité exprime la notion de simple groupement ;
- d*• enfin, la combinaison de ces diverses lois définit le cadre expérimental nécessaire à modéliser l'influence de la localisation des pacmans sur les relations entretenues.

Par conséquent, si les objets sont tous deux des pacmans et que leurs centres se regardent mutuellement, alors les deux objets sont "face à face" ; autrement dit, ils entretiennent une relation prégnante modélisable avec les propriétés *a* et *b*. Sinon les objets entretiennent une relation simplement dictée par *c*. Si l'un des objets est un pacman et l'autre le centre d'une lacune, alors il suffit que l'ouverture du pacman regarde ce centre pour que la relation "face à face" soit vraie. Sinon comme pour l'autre cas les objets entretiennent une simple relation de proximité. Il faut remarquer ici que nous avons privilégié l'orientation des objets inducteurs, ceci conformément aux derniers travaux des néo-gestaltistes qui considèrent que l'orientation est l'un des facteurs d'association les plus prégnants. Enfin, si la relation entretenue entre les deux objets est de classe *c* alors dès que la distance inter-objets augmente la perception de groupement décroît très rapidement. Il n'est pas de même lorsque la relation est de classe *a* ou *b*, l'éloignement des objets fait décroître moins rapidement la perception d'une cohésion entre ces objets. Ce phénomène a été observé et étudié depuis déjà quelques années par les psycho-physiciens [Zucker & Davis 88], [Dresp 90] et il ressort de leurs expériences que la distance inter-objets n'a évidemment pas une influence linéaire sur la perception de groupement. Il ressort plutôt que lorsque la distance est petite (de l'ordre de 10 à 15 mm pour des objets dont le rayon a une taille variant de 4 à 16 mm) celle-ci a une influence négligeable sur la construction d'un groupement perceptif. Il a été également mis en évidence que lorsque la distance augmente, son influence sur la perception augmente rapidement pour infléchir à nouveau et augmenter ensuite très lentement (le point d'inflexion se trouve entre 25 et 30 mm). Ainsi, pour modéliser l'influence de la distance, ces observations nous ont suggéré de choisir des fonctions dont la formule générale est une fraction rationnelle de la forme : $1/\text{distance}^k$.

Concrètement, pour réaliser l'évaluation de ces relations, le module a extrait, au préalable, deux types d'informations : l'une étant la position géographique d'une figure inductrice par rapport aux autres contenues


dans l'image, l'autre constituée par la taille et l'orientation de l'angle provoqué par l'ouverture du disque. Ces deux informations ont été déduites de la description synthétique des contours des pacmans, et notamment des deux segments de droite qui constituent leur ouverture. Le système considère le point de convergence de ces deux segments de droite comme l'emplacement à partir duquel il évalue la distance inter-objets. L'orientation de chacun des deux segments lui fournit, elle, la donnée nécessaire pour que le module puisse évaluer la taille et l'orientation de l'angle d'ouverture.

5.3.2 Le module de mise en confrontation

Signification des cellules : le principe gestaltiste de simplicité a contribué au choix de la signification symbolique potentielle des objets. Les deux unités qui symbolisent ces objets répondent, donc : l'une à un disque incomplet, l'autre à un disque "en occlusion". Un disque en occlusion est un disque entier sur lequel un objet géométrique simple serait surimposé, ici le rectangle⁴. Les principes de la contiguïté et de la connexité ont eux contribué à la définition de la signification des lacunes. Les unités décrivant les lacunes sont au nombre de trois, elles évoquent: l'une le fond, la deuxième un groupement de disques incomplets (pas de forme surimposée), enfin la dernière une forme (surimposée).

La topologie du réseau : le nombre d'objets (objet réel ou lacune) identifiés, la signification potentielle ainsi que les relations entretenues deux à deux par les objets sont les informations nécessaires pour construire le réseau. Chaque signification est représentée par une cellule (c_i) : un ensemble de cellules (C_i) est donc la représentation de l'objet (O_i) dans le réseau. Chaque (c_i) d'un objet (O_i) entretient des relations avec chacune des cellules (c_j) de l'ensemble (C_j), représentation de l'objet (O_j) dans le réseau. Si nous prenons comme exemple le pacman, et que nous nous plaçons au niveau de l'objet celui-ci peut être représenté comme un nœud duquel partent vers tous les autres objets des arcs représentant les relations. La figure 6 représente une vue détaillée des deux premiers pacmans d'une configuration à 4 pacmans (en haut à gauche de la figure) .

⁴ Des configurations avec des angles d'ouverture différents et un nombre de pacmans différents peuvent être aisément envisagées, par exemple, la

configuration à trois pacmans  .

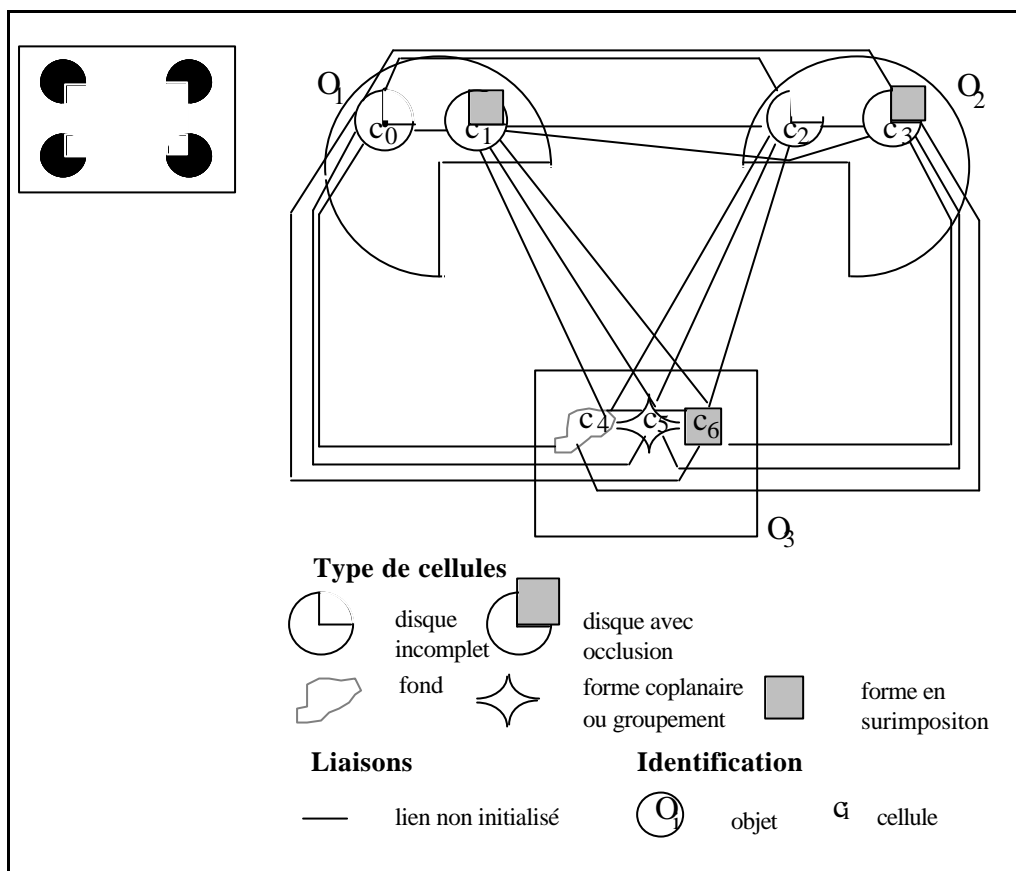


figure 6 : Cartographie du réseau (partie) avec liens non initialisés

Nous pouvons constater que chaque *objet réel* (O_1 , O_2) est représenté par deux cellules (pour O_1 les cellules c_0 et c_1). L'une (c_0) évoque l'appartenance de l'objet à la classe des disques incomplets, l'autre (c_1) à celle des disques avec occlusion. *La lacune* (O_3), elle, est représentée par trois cellules (c_4 , c_5 et c_6) évoquant respectivement l'appartenance de la lacune au fond, au centre d'un groupement d'objets réels, à une forme en surimposition. Les cellules ainsi définies sont potentiellement toutes connectées. Les poids de ces connections (ou liens) vont dépendre des relations d'orientation et de distance des éléments de la figure (pacmans et lacunes).

Initialisation des poids des liens : c'est la forme exacte de l'image vue (orientation des pacmans, distance entre les divers éléments,...) qui permet de calculer les poids. Ce calcul constitue l'étape cruciale du système et a nécessité plusieurs phases de mise au point. Retenons simplement qu'après initialisation, certains liens sont devenus inhibiteurs (pondérations négatives) et d'autres sont devenus excitateurs

(pondérations positives) (pour les spécifications techniques cf. Annexe 3 de [Gaio 94]).

Concrètement, à chaque nouvelle image présentée, et après passage du premier module, il y a création d'un réseau avec deux cellules par figure inductrice et trois cellules par lacune (par exemple : pour une figure à 16 pacmans nous allons avoir : $16*2 + 9*3 = 59$ cellules). Toutes les cellules sont potentiellement connectées, chacune des connections (ou liaisons) est alors initialisée avec un poids fonction de la distance séparant les objets représentés par les cellules extrémités. Selon l'orientation des objets inducteurs la relation est considérée par le système comme prégnante ou simple (par exemple : si le pacman "regarde" l'autre objet alors la relation est prégnante).

5.4. Expérimentation avec le module de simulation des groupements perceptifs

Cette expérience mise au point en collaboration avec des psychophysiciens⁵ a permis de construire un réseau qui émule le comportement de regroupement perceptif humain, face à un échantillon de différentes combinaisons de pacmans pour faire émerger ou non des rectangles. Parallèlement à la mise au point du réseau, les psychophysiciens soumettent à un échantillon représentatif de sujets humains les mêmes figures. La directive est d'identifier les groupements qui "sautent aux yeux" [Zouinar 92]. Nous nous centrerons ici sur la mise au point de l'outil informatique.

Voici, maintenant, les conditions de mise en œuvre et la méthode utilisée. L'échantillon expérimental de base est constitué par quatre figures (figure 7) représentant une combinaison de pacmans.

Dans l'échantillon que nous avons retenu pour une première expérimentation, les figures inductrices (les pacmans) sont de taille constante, leur ouverture est fixée à 90° et ne peuvent être orientées que dans quatre directions. L'image 7(a) représente 4 groupements de 4 pacmans avec, par groupement, une orientation unique de l'ouverture des pacmans. Dans l'image 7(b) nous retrouvons la même disposition en 4 groupements mais cette fois le pacman se trouvant à l'intérieur de chaque groupement a subi une rotation pour que son ouverture puisse se trouver face à face avec les autres pacmans ayant subi une rotation équivalente. Dans l'image 7(c) ce sont les 3 pacmans extérieurs qui ont subi une rotation pour pouvoir se trouver face à face entre pacmans d'un même groupement. Enfin, dans l'image 7(d) tous les pacmans ont subi

⁵ Laboratoire de psychologie expérimentale de l'Université René Descartes, Paris.

une rotation par rapport à la première image, les pacmans internes se trouvant dans la même configuration que dans l'image 7(b) et les extérieurs dans la configuration de l'image 7(c).

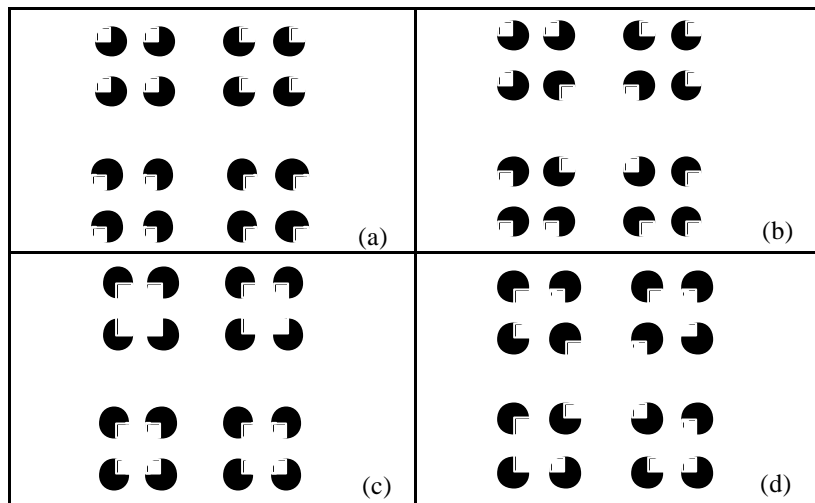


figure 7 : Echantillon expérimental de base

D'autre part, étant donné le manque de précision inhérent à toute discipline qui tente de formaliser le comportement humain, il a été nécessaire de trouver une famille de fonctions pour simuler la perception de la distance sur laquelle il était facile de modifier l'allure de la pente grâce à un paramètre. La mise au point de ces fonctions⁶ a été opérée pour des images où la distance maximum était comprise entre 200 pixels et 2000 pixels. La formule choisie est de la forme :

$$D = 1 / \left(\left(\text{distance}_{i,j} / \text{taillepixel} \right)^k / \left(\text{distance}_{\text{max}} / \text{taillepixel} \right) \right) + 1$$

Les valeurs de k ont été mises au point par une succession de jeux d'essais par le biais de l'outil connexionniste. On a ainsi obtenu : pour les relations perceptives prégnantes k=3, pour les autres relations k=2. Elles permettent au réseau, d'après nos expérimentations, de répondre de manière cohérente sur l'ensemble de l'échantillon et constituent donc des prédictions. L'expérimentation psychologique a

⁶ En première approximation nous considérons la taille du pixel comme invariante du point de vue de sa perception. Pour cela nous nous sommes appuyé sur le postulat que, pour des pixels de taille variable, l'homme adapte la distance à l'image de manière à ce que la taille perçue des pixels soit suffisamment petite. En réalité la relation entre taille du pixel et perception de celui-ci est assez complexe, mais il est vrai que notre approximation est suffisante pour les tailles de pixels utilisées en imagerie aérienne.

permis de mesurer le niveau de précision par comparaison avec les résultats obtenus sur les sujets humains .

Le graphe de ces fonctions a une allure semblable à celles de la figure 8 pour une image de 256*256 pixels.

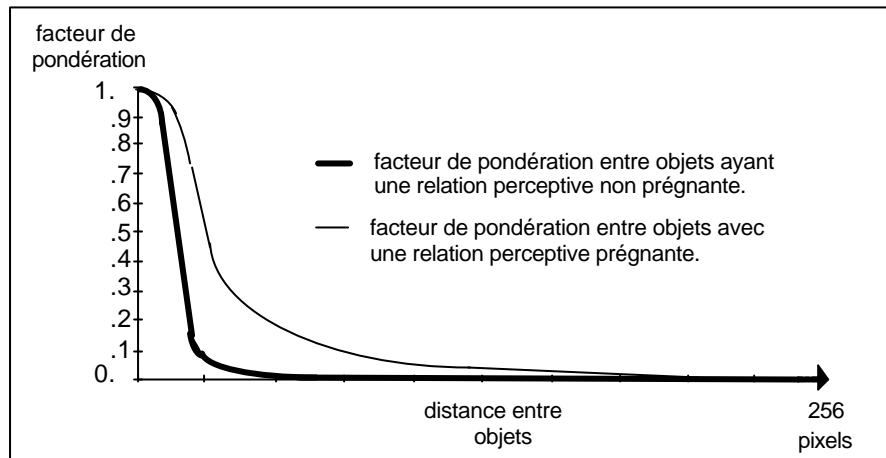


figure 8 : Graphe des fonctions régulant l'influence de la distance

Pour cette mise au point il a fallu réaliser tout un environnement de travail. En particulier, une interface a été conçue pour permettre la visualisation d'un certain nombre d'indicateurs, et faciliter ainsi l'appréhension du comportement du réseau. Une représentation graphique dynamique a été réalisée pour observer les trajectoires dans l'espace des états des différentes cellules du réseau. Cette visualisation est apparue indispensable dès les premiers essais réalisés avec l'outil connexionniste. Les différentes phases, par lesquelles passe le système avant de se stabiliser, se sont avérées aussi riches en enseignements que les états des cellules à système stabilisé. Nous avons ainsi pu observer que plus les combinaisons de pacmans provoquaient des illusions complexes, plus le réseau nécessitait un nombre accru d'itérations pour se stabiliser. Nous avons pu également constater que dans les configurations où l'orientation de certains pacmans provoquait une ambiguïté perceptive, le système avait tendance à osciller entre deux configurations possibles avant de se stabiliser (dans ce cas le nombre d'itérations pour atteindre un état stable était plus important). Cela a confirmé nos prédictions formulées lors de la modélisation mathématique et a permis d'affiner le choix des paramètres.

5.5. Interprétation des résultats et connaissances acquises.

5.5.1. Quelles étaient les regroupements et les illusions qui devaient émerger ?

Comme nous pouvons le constater visuellement les pacmans de l'image 7(a) n'induisent aucune forme, elle constitue en quelque sorte l'image de référence. Nous constatons qu'au contraire dans l'image 7(b) les 4 pacmans réorientés réalisent un rectangle de Kanizsa central. Cette image peut être interprétée comme étant constituée de 12 disques noirs entamés et de 4 disques noirs complets sur lesquels a été posé un rectangle blanc. Nous nous trouvons en présence de 4 rectangles de Kanizsa dans l'image 7(c). Cette image peut être interprétée comme étant constituée de 16 disques noirs complets sur lesquels ont été superposés 4 rectangles blancs. Nous obtenons dans la dernière image 7(d) un Kanizsa central et 4 Kanizsa latéraux. L'interprétation de cette dernière figure pose un problème car celle-ci dépend du nombre de couches d'objets que l'on veut bien supposer. Or l'humain semble rechercher des interprétations simples, c'est-à-dire n'impliquant que 2 couches d'objets, des pacmans ou des disques noirs complets sur lesquels viennent se superposer des rectangles blancs. Il en résulte dans ce cas une ambiguïté sur l'interprétation des 4 disques noirs centraux. Une interprétation en 2 couches donne ainsi la perception d'une image composée de 16 disques noirs. Sur les 12 périphériques se superposent 4 rectangles blancs et sur les 4 centraux se superpose un rectangle de plus grande dimension⁷.

Il a été constaté sur les figures (7(a), 7(b) et 7(c)) testées que dans chaque cas le réseau s'est stabilisé sur une réponse qui correspond à l'interprétation perceptive humaine de la figure. Par contre pour la configuration correspondant à la figure 7(d), le réseau ne converge pas dans une situation limite, mais oscille entre deux positions pour ce qui est des états représentant les quatre pacmans du centre qui sont

⁷ Mais avec une telle interprétation les 4 Kanizsas périphériques provoquent une perception incomplète des rectangles et obligent à émettre 2 hypothèses. L'une conserve le présupposé de seulement 2 couches, mais fait intervenir une construction séquentielle de l'image, l'autre passe à une interprétation à 4 couches. La première suppose ainsi que les 4 disques centraux sont posés après les 12 périphériques et après les 4 rectangles qui s'y superposent. Alors que si on emploie la dernière hypothèse, cela donne une image qui est constituée de 16 disques noirs complets sur lesquels ont été superposés 4 rectangles blancs. Sur les rectangles 4 disques noirs complets ont été placés juste sur les 4 disques centraux. Enfin, sur ces derniers disques un rectangle blanc a été surimposé.

alternativement en occlusion ou disque incomplet suivant que la forme en surimposition principale soit celle du centre ou de la périphérie. On peut voir, sur cet exemple, que le réseau cherche à interpréter les figures par seulement deux couches d'objets superposées. On retrouve ici la situation évoquée dans la section 4.1 à l'aide de la figure de Necker.

À partir des quatre images de la figure 7 nous avons produit par la suite une dizaine d'images dérivées pour pouvoir étudier les différents cas de figure et tester les capacités de notre système de simulation. Les réponses du système sont apparues cohérentes avec la vision des expérimentateurs. La simplicité de mise en œuvre de nouvelles simulations a permis de pousser encore plus loin l'investigation et réfléchir sur des possibilités d'élargissement de l'échantillon expérimental pour permettre d'affiner la modélisation. Une tentative a été réalisée avec des images toujours composées de figures induisant une illusion perceptive mais dans lesquelles un bruit a été ajouté, notamment nous avons fait varier soit la distance entre pacmans, soit l'angle d'ouverture des pacmans, soit les deux paramètres. Les résultats semblent être prometteurs, mais il faudrait un protocole d'expérimentation psychologique complet pour mesurer à nouveau le niveau de précision de notre modèle dans des cas limites.

5.5.2. *Et la géographie?*

Si l'expérimentation sur les figures à illusion primaire peut et doit être continuée, les premiers acquis ont déjà été réinvestis. En effet, nous avons pu partiellement nous affranchir d'une validation plus robuste du modèle grâce à une collaboration avec les géographes. Ainsi, une fois réalisée l'adaptation du système au problème de l'identification des structures linéaires prégnantes dans les images aériennes, les réponses fournies par le système ont pu être mises en perspectives avec les résultats de photo-interprétations "manuelles" réalisées par des géographes. Cette confrontation a permis de tester les propriétés de généralisation du modèle.

Les structures linéaires présentes dans des images aériennes en niveau de gris ont, dans la plupart des cas, une matérialité incomplète et elles obligent ainsi le système visuel à les reconstruire. Si l'on reprend l'exemple du tracé d'une route de campagne "noyé" entre deux labours, la perception que l'on a de ces deux tronçons indépendants est néanmoins une seule route. De manière générale, les structures linéaires dans les images télédéteectées ne présentent aucune forme de continuité physique, ni dans les valeurs radiométriques émises, ni dans

les variations entre les valeurs des pixels les constituant et les valeurs des pixels qui leurs sont contigus. La continuité du message final émis n'est donc qu'une conséquence de la perception visuelle humaine. Divers travaux de recherche, [Mohan et al. 89], [Heraud et al. 90], [Etemadi et al. 93], proposent une description intermédiaire entre le niveau dit de *segmentation* qui opère directement sur la matrice des pixels et le niveau dit d'*interprétation* qui utilise une description plus ou moins sophistiquée des composantes de l'image. Cette description intermédiaire consiste à grouper les éléments issus de la segmentation dans des structures géométriques et/ou relationnelles plus abstraites.

Les segments qui construisent ces structures doivent obéir à un ensemble de propriétés déduites des mêmes relations que nous avons analysées et modélisées grâce aux figures à illusion perceptive primaire.

Le système élaboré pour la détection ces structures linéaires est donc une adaptation des modules mis au point pour l'analyse d'une image de Kanisza. Nous avons conservé les trois propriétés définissant les relations des figures inductrices (cf. 5.3.1) ainsi que les deux fonctions régulant l'influence de la distance (cf. figure 8). L'adaptation a dû être réalisée pour différentes raisons :

- tout d'abord l'image aérienne sémantiquement plus riche a nécessité de spécialiser le module de description par primitives, un module plus élaboré a donc dû être pensé (actuellement seule la description automatique des structures linéaires quasi-rectilignes est opérationnelle).

- D'autre part, le problème de conflit entre deux formes possibles étant plus inhibant dans cette classe d'images, il a fallu modéliser une nouvelle relation basée sur un principe d'opposition à la loi de similarité : le principe de *blocage*.

- Enfin, les segments de droite élémentaires obtenus par la segmentation bas-niveau constituent chacun un objet inducteur et leur signification est intrinsèquement plus simple. Chaque segment de droite se voit finalement codé dans le réseau par une seule cellule. On ne conserve en fin de ce processus que les segments dont la cellule représentante est à l'état haut. Ces segments sont les supports des structures linéaires. Notons que la présence d'un grand nombre de cellules nous a obligés à introduire une étape de simplification du réseau, qui est *a priori* entièrement connecté, en supprimant les liens de poids proche de 0.

Avec le passage à des problèmes d'imageries aériennes, dernière étape de notre démarche, on peut mesurer tout l'intérêt qu'il y a à proposer des modèles, certainement incomplets mais opérationnels.

7. CONCLUSION

Nous avons présenté un travail où l'ordinateur a été utilisé à la fois comme outil d'expérimentation et de modélisation. Il s'agissait de modéliser le mécanisme de perception visuelle. Le principe de base est que la perception est un système dynamique qui "respecte" les principales lois de la *Gestalt*. Chaque système dynamique de perception d'une image a été simulé sur ordinateur par un réseau récurrent.

Le modèle se devait d'être validé sur des images simples, où les mécanismes de base de la perception visuelle étaient faciles à isoler. Cette classe d'images a été constituée à partir d'images à illusion dite primaire, constituées par des rectangles de Kanizsa induits par des disques incomplets. Cette modélisation a permis d'utiliser l'ordinateur pour découvrir les lois de l'influence de la distance entre les pacmans. Cette influence est beaucoup plus fortement décroissante pour des pacmans n'induisant pas de formes surimposées que pour des pacmans positionnés "face à face". L'expérimentation avec le modèle a permis de découvrir une formulation convenable de ces deux fonctions. Ces fonctions demanderaient maintenant à être affinées par une deuxième expérimentation plus poussée. La simulation est déjà mise en œuvre sur machine mais la démarche adoptée nécessite de mettre en perspective les résultats obtenus avec ceux d'une observation sur différents échantillons de sujets humains (par exemple sur un échantillon très générique, puis sur échantillon plus spécifique pris dans une population de géographes).

Néanmoins, les résultats de la première modélisation ont pu être réinvestis avec succès dans l'élaboration du module de reconnaissance automatique de structures linéaires prégnantes dans des images aériennes de paysages ruraux. La situation est apparue à la fois plus simple que dans le cas des figures à illusion perceptive et plus complexe. Plus simple, puisque contrairement à des objets surfaciques (les rectangles de Kanizsa) pour les structures linéaires seules deux lois de la *Gestalt* sont en cause : continuité et proximité. Mais plus complexe car on se trouve avec des réseaux de plusieurs milliers de cellules qui rendent impossible toute recherche exhaustive de solutions.

Bibliographie

- B. Dresp, J. Lorenceau, C. Bonnet (1990) Apparent brightness enhancement in the Kanizsa square with and without contour formation. In *Perception*, 1990, Vol. 19, p. 483-489.
- M. Gaio, J. Madelaine, J. Lorenceau, B. Victorri (1992) L'émergence des formes dans la perception visuelle. In *Emergence dans les modèles de la cognition*. Actes des Journées, ENST.
- M. Gaio (1994) *Une architecture d'aide à la construction de croquis d'interprétation géographique : contribution à l'interprétation des images aériennes*, Thèse de l'université de Caen.
- S. Grossberg (1988) Nonlinear Neural Networks: Principles, Mechanisms and Architectures. In *Neural Networks*, Vol. 1, p. 17-61.
- P. Guillaume (1979) *La psychologie de la forme*. Flammarion, Paris.
- J.J. Hopfield (1982) Neural networks and physical systems with emergent collective computational abilities. In *Proc. Nat. Acad. Science USA*, Vol. 79, p. 2554-2558.
- J.J. Hopfield (1984) Neurons with graded response have collective computational properties like those of two-state neurons, In *Proc. Nat. Acad. Science USA*, Vol. 81, p. 3088-3092.
- D.O. Hebb (1949) *The organization of behavior*, J Wiley & Sons, New York.
- G. Kanizsa (1955) Margini quasi-percettivi in campi con stimolazione omogenea. In *Rivista di psicologia*, n° 49, p. 7-30.
- W.S. McCulloch & W.A. Pitts (1943) A logical calculus of the ideas immanent in nervous activity. In *Bulletin of Mathematical Biophysics*, vol. 5, p. 115-133.
- G.E. Meyer & S. Petry (1986) Adelphi International Conference on Illusory Contours: A report on the conference. In *Perception and Psychophysics*, N° 39, p. 210-222.
- G.E. Meyer & S. Petry (1987) Top-Down and Bottom-Up: The illusory contour as a microcosm of issues in perception. In *The Perception of illusory contours*. Stringer, p. 3-20.
- R. Mohan (1992) Perceptual for Scene Segmentation and Description, In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI 14, n° 5. IEEE Comput. Soc. Press.
- R. Mohan & R. Nevatia (1988) Perceptual grouping for the detection and description of structures in aerial images, In *Proceedings of the DARPA Image Understanding Workshop*, p. 512-516. Morgan-Kaufmann.
- P. Parent & S.W. Zucker (1989) Trace Inference, Curvature Consistency, and Curve Detection, In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 2, n° 8, éd. IEEE Computer. Soc., p. 823-839.
- L. Personnaz & G. Dreyfus (1989) Les réseaux de neurones, situations et perspectives, In *Colloque GRETSI'89*.

- J. Piaget (1936) *La naissance de l'intelligence chez l'enfant*. Neuchâtel, Delachaux et Niestlé.
- A. Rosenfeld, R.A. Hummel & S.W. Zucker (1976) Scene labelling by relaxation operations, In *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 6. IEEE Computer. Soc., p. 420-433.
- D. E. Rumelhart, J.L. McClelland & the PDP Research Group (1986) *Parallel Distributed Processing: explorations in the microstructure of cognition*, Vol. 1, MIT.
- M. Zouinar (1992) *Le groupement perçu peut-il être perturbé par des contours illusoire ?* Mémoire, Paris V.
- S.W. Zucker & S. Davis (1988) Points and endpoints : a Size/Spacing Constraint for Dot Grouping, In *Perception*, n° 17, p. 229-247.