La géométrie des images naturelles et ses algorithmes *

Jacques Froment †, Simon Masnou ‡ et Jean-Michel Morel §

10 novembre 1998

1 Introduction

Pourquoi les stimuli qui arrivent sur nos rétines sont-ils interprétés en terme de droites, d'objets, de formes bien déterminées? cette question paraît bien incongrue : quoi de plus naturel que de reconnaître une droite dans l'image d'une droite, un cube bleu dans l'image d'un cube bleu, une vache paissant dans un pré dans l'image d'une vache paissant dans un pré? C'est que, quand nous croyons voir une droite, les images formées sur nos rétines n'ont en fait "rien à voir" avec celle d'un trait continu et d'épaisseur nulle, s'étendant à l'infini. Ces images ressemblent plutôt à une multitude de petites taches plus ou moins foncées, mais suffisamment proches et alignées pour être perceptuellement regroupées en un tout homogène. Cet exemple est, pour reprendre un terme de G.Kanizsa, une des "énigmes de la perception". Il illustre un problème fondamental résolu par la perception visuelle : l'identification ou la mise en correspondance d'un objet perceptif (le stimulus sur la rétine) avec un objet physique, voire avec un objet abstrait n'ayant aucune réalité physique (comme la droite géométrique). L'étude de la perception visuelle nécessite un examen rigoureux des principes qui permettent la mise en correspondance, encore appelés *principes de la reconstruction visuelle*.

Cette démarche a été initiée par l'école de psychologie de la *Gestalt*, et notamment par le fondateur du "gestaltisme géométrique", M.Wertheimer, dans les années 1920. Mais si, formellement, toute vision est une illusion, il existe quelques situations où la reconstruction visuelle échoue, au sens où la correspondance entre objet perceptif et objet physique est erronée. Il s'agit donc d'*illusions visuelles distorsives*, encore qualifiées d'*optico-géométriques*. Les illusions optico-géométriques sont

*Nous remercions la Revue du Palais de la Découverte qui nous a aimablement autorisés à utiliser en partie le texte et les illustrations de l'article intitulé *Perception visuelle et traitement d'images*, écrit par Jean-Michel Morel et paru dans le volume 25 sous les numéros 243 (décembre 1996) et 244 (janvier 1997).

†U.F.R Maths-Info, Université Paris-V René Descartes, Paris, jf@math-info.univ-paris5.fr

‡Université Paris-IX Dauphine, 75775 Paris Cedex 16, masnou@ceremade.dauphine.fr

§CMLA, Ecole Normale Supérieure de Cachan, Cachan, morel@ceremade.dauphine.fr

les premières à avoir été étudiées, dans le cadre de recherches sur le système visuel. Mais il s'agissait alors, avec le point de vue physiciste qui était adopté notamment par Helmholtz et par Mach, d'apporter des explications physiologiques à l'existence des illusions distorsives et d'en déduire des informations sur l'architecture du cortex visuel. L'illusion distorsive la plus connue est peut-être celle de Hering (figure 1). Elle est construite à l'aide de deux droites parallèles a et b, au centre desquelles un faisceau de droite converge. L'erreur consiste à "voir" les deux droites parallèles se déformer au voisinage des droites convergentes.

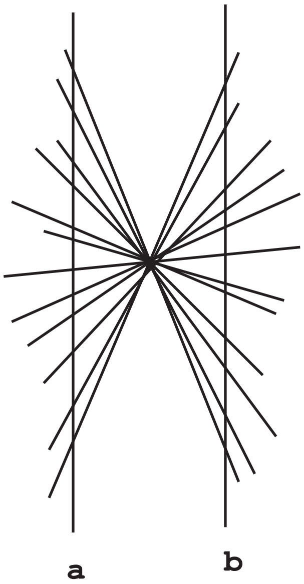

FIG. 1 – *Illusion de Hering : les droites a et b paraissent s'incurver au voisinage des droites convergentes. Notons que le faisceau de droites convergentes introduit un point de fuite dans la scène, et donc une interprétation en terme de perspective.*

Les physicistes et les gestaltistes partagent la même conception géométrique de la vision, et les mêmes expériences d'illusions distorsives. Il diffèrent cependant quant aux objectifs de leurs recherches : les premiers étudient les illusions optico-géométriques comme des révélateurs de mécanismes cachés du système visuel, quand les seconds, une fois admis le caractère non évident de la mise en correspondance, s'attachent à étudier les lois ou les principes qui dirigent la reconstruction visuelle. Pour les gestaltistes, toute figure est alors motif à illusions visuelles et se prête à commentaire. Leur travail essentiel est donc de découvrir, au travers d'expériences visuelles, les différents principes de la reconstruction visuelle. Le recueil le plus complet à ce jour est certainement celui dressé par Metzger dans [13], mais le plus accessible est dû à G.Kanizsa dans [7] et [8]. Ici encore, il est essentiel de réduire les figures, support des expériences visuelles, de façon à bien isoler les différents principes : dans la section 2 de cet article, nous expliquons chaque principe de la reconstruction visuelle en l'illustrant par une figure qui évite le plus possible les interactions avec d'autres principes.

Nous nous attachons, dans la section 3, à montrer comment les principes de la reconstruction visuelle peuvent s'adapter aux figures complexes et notamment aux images naturelles. Deux d'entre eux, le principe "d'invariance par changement de contraste" et le principe de "complétion amodale" sont actuellement utilisés avec succès pour le traitement des images numériques. Cette généralisation est d'importance car elle permet d'espérer la découverte de nouveaux algorithmes de traitement de l'image, au travers de nouvelles représentations atomiques (Cf. [3]). Le saut de complexité entre une image simple (comme celle d'une droite) et une image naturelle (comme celle d'une vache paissant dans un pré) paraît considérable. Si la taille des données (nombre de "pixels" ou points élémentaires constitutifs du stimulus visuel ou de l'image numérique) est effectivement bien supérieure, ce n'est pas le cas de la complexité intrinsèque de ces données : dans chaque cas, l'œil ne reçoit que des nuages de points d'intensité ou de couleur variant presque continûment, sans structure géométrique apparente. Seulement, si la reconstruction visuelle permet l'interprétation d'une figure simple en des termes géométriques (droites, courbes, disques, carrés, cubes, ...), notre habitude est de commenter des figures complexes en nommant les objets, c'est-à-dire par des termes (vache hollandaise, prairie de luzerne, ...) qui échappent à l'analyse mathématique. Comment représenter alors mathématiquement la structure d'une image ? les chercheurs en *Computer Vision* se sont largement inspirés des artistes, qui savent que l'essentiel de l'information est contenu dans le contour des objets (Cf. notamment la technique de la bande dessinée) : les bords, ou lignes de discontinuités de l'intensité lumineuse, sont ainsi souvent considérés comme les "atomes" (objets élémentaires) dans la représentation des images [9]. Or, on peut montrer [3] que les représentations mathématiques actuelles basées sur la détection des bords sont inadaptées, de par la constitution physique des images naturelles. Une solution est apportée par l'école de la *Morphologie Mathématique*, où une image est décomposée en ses ensembles de niveau. Si on suppose par exemple l'image à variation bornée (c'est le cas de toute image numérique), les bords des ensembles de niveau, appelés *lignes de niveau*, sont des courbes de Jordan fermées. Nous montrons comment les principes de la reconstruction visuelle s'appliquent dans le cadre de cette représentation géométrique.

La section 4 présente des exemples d'algorithmes de traitement d'images basés sur les principes gestaltistes. Il est possible de débruiter très simplement certaines images en éliminant des petits ensembles de niveau, sans que cette opération ne détruise d'information importante. La mise en correspondance d'ensembles de niveau permet de repérer un même objet présent sur deux vues, d'un même lieu mais réalisées à des instants differents. Enfin, le principe de complétion amodale peut être simulé en prolongeant les lignes de niveau derrière les jonctions en T. On peut ainsi reconstruire une image plausible qui rend visible les parties cachées de l'image originale. Cet algorithme de désocclusion permet de restaurer des images abimées par la présence de taches.

2 Les principes de la reconstruction visuelle

La reconstruction visuelle conduit à une constitution géométrique des objets visuels, et non à la reconnaissance du contenu de l'image. Nous reprenons ainsi la démarche gestaltiste qui consiste à séparer le processus primaire de constitution géométrique de celui, secondaire, de reconnaissance, ce dernier intégrant nécessairement des processus d'interprétations et de raisonnements. A l'image des travaux de M.Wertheimer et de G.Kanizsa, nous utilisons des expériences visuelles simples pour expliciter les lois qui dirigent la reconstruction. Leur but est de montrer que la reconstruction visuelle obéit à des principes essentiellement géométriques, et qu'elle est donc du ressort de la formalisation mathématique.

Le *principe d'identité chromatique* permet l'unification des régions connexes, où l'intensité lumineuse ne varie pas, ou alors de manière bien régulière. L'unification des régions est un processus de regroupement perceptif qui permet de percevoir les régions comme un tout homogène, une surface qui est souvent interprétée comme la trace d'un objet (notamment lorsque sa frontière est régulière). Ainsi, la figure 2 est composée d'un nuage de points noirs unifiés en un objet visuel. Notons qu'il est impossible de percevoir individuellement les points qui composent cette tache, c'est-à-dire que le tout a occulté les parties. Le principe d'identité chromatique est sûrement le plus fondamental, au sens où la vision repose sur la capacité à distinguer les différences locales de luminosité. Il s'applique de manière systématique, et certainement avant tout autre.

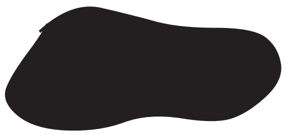

FIG. 2 – *Le principe d'identité chromatique nous permet de "voir" une tache noire là où seul un nuage de points foncés est présent.*

Le *principe de proximité* permet lui aussi un regroupement perceptif; il s'applique lorsque la distance entre les objets unifiés est suffisamment faible, relativement à la taille des objets et aux distances de ces objets avec ceux restants. La figure 3 est reconstruite par application des deux principes précédents : les ellipses noires se constituent visuellement par le principe d'identité chromatique, puis le principe de proximité crée un regroupement en deux objets distincts. Remarquons que le principe de proximité n'a pas pour effet de masquer les parties : les ellipses sont toujours visibles.

Le *principe de similitude* entraîne comme ces prédécesseurs un regroupement perceptif : les objets semblables sont unifiés. Selon la taille et le nombre des éléments similaires, la région formée peut devenir ce que l'on appelle une micro-texture. Dans ce cas, les parties ne sont plus perçues spontanément. On comparera ainsi les figures 4 et 5, qui présentent deux relations tout-parties inverses.

Fig. 3 – *Le principe de proximité implique le regroupement perceptif des ellipses noires en deux objets distincts.*

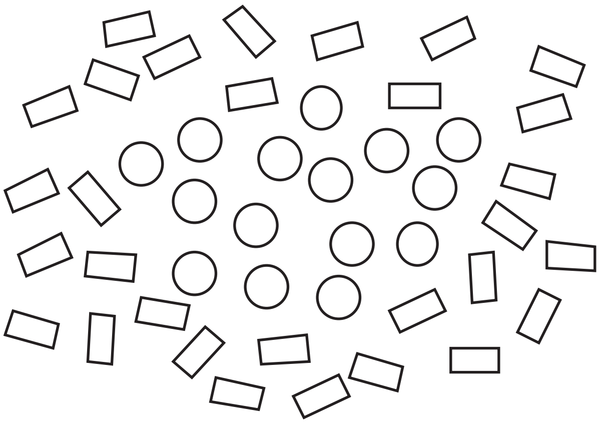

Fig. 4 – *Le principe de similitude permet d'interpréter cette image comme composée de deux régions homogènes : celle centrale des cercles et celle périphérique des rectangles.*

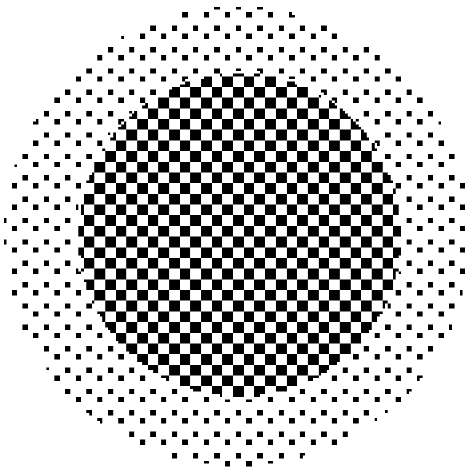

Fig. 5 – *C'est encore le principe de similitude qui sépare cette image en deux régions. Le nombre important d'éléments semblables (les petits carrés) crée un effet de texture de sorte que, contrairement à la figure précédente, le tout occulte les parties.*

Le *principe de continuité de direction* (ou de *bonne continuation*) entraîne une rigidification des objets visuels formés par des courbes régulières, c'est-à-dire un regroupement perceptif dans lequel l'objet possède une qualité de rigidité. Ce principe de bonne continuation nous permet d'appeler courbes des structures géométriques qui n'ont pas de réalité objective. Par exemple, la figure 5 possède une courbe régulière, qui est la frontière entre les deux textures. Il en résulte la formation de deux objets visuels rigides, le disque et l'anneau. Les courbes régulières sont en général créées par des contours occlusifs, des contours d'arête (par exemple le bord de la région noire, figure 2) ou des contours illusoires (comme dans la figure 5).

Le principe de continuité de direction peut se décomposer en trois principes plus élémentaires, que nous allons maintenant aborder : les principes de fermeture, de complétion amodale et de transparence/ombrage.

Le *principe de fermeture* correspond au cas particulier où la courbe est fermée. Un premier exemple a été donné avec la figure 5. La figure 6 illustre le fait qu'il n'est pas nécessaire que l'objet et le fond soient d'intensité différente pour qu'une unification se produise.

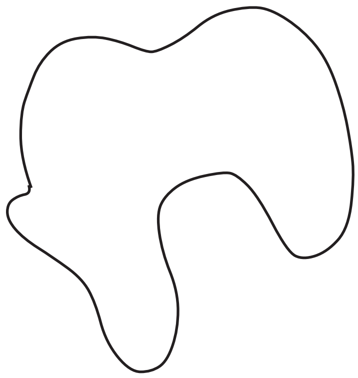

FIG. 6 – *Par fermeture, l'intérieur de la courbe en noir est assimilé à la trace d'un objet. L'extérieur de la courbe est assimilé au fond.*

La *complétion amodale* est certainement le principe de reconstruction visuelle le plus surprenant; nous montrerons dans la section suivante qu'il e st au cœur des illusions visuelles gestaltistes. Ce principe s'applique lorsqu'une courbe régulière, interprétée comme le bord d'un objet par un processus d'identification, s'arrête sur une autre : on observe au voisinage de la jonction une figure en "T", la première courbe correspondant à la barre verticale du "T" et la seconde à la barre horizontale. En conséquence, on nommera ce cas de figure une *jonction en T*. A chaque fois qu'une jonction en T est détectée, notre perception l'interprète comme étant le résultat de l'occlusion d'un objet par un second objet placé plus avant. En général, une jonction en T isolée ne permet pas de déterminer l'ordre dans l'espace des trois régions délimitées par la jonction : l'ordre est choisi de façon à obtenir une interprétation globale cohérente. Afin d'illustrer notre propos, supposons que le bord de l'objet occulté corresponde à la barre verticale du "T". Ce bord est alors complété de manière amodale, c'est-à-dire qu'il est prolongé mentalement derrière la barre horizontale du "T". le prolongement préserve la régularité du contour visible, notamment la continuité de direction et la courbure le long du bord.

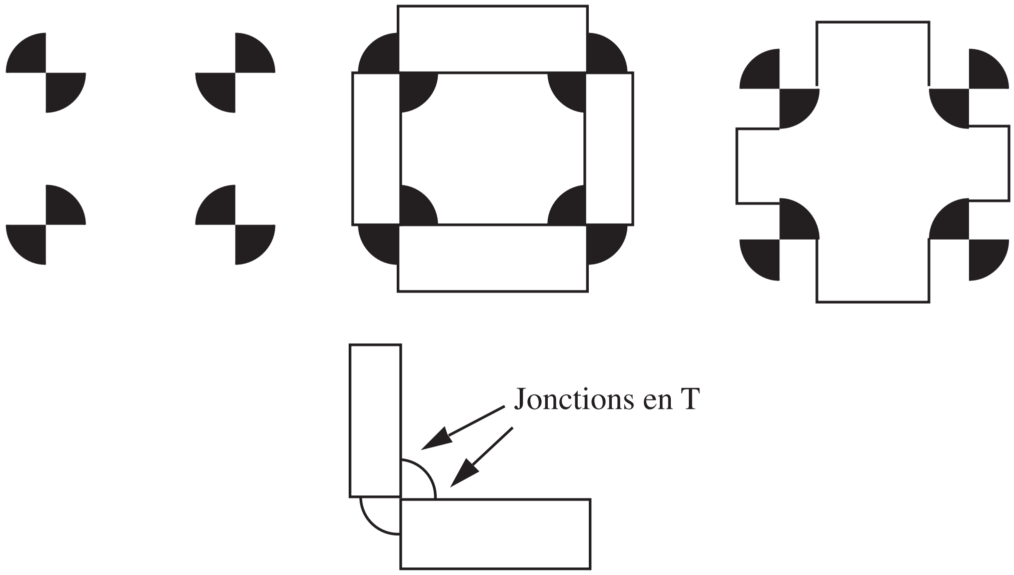

FIG. 7 – *La compétion amodale s'applique lorsqu'une occlusion est détectée, laquelle correspond géométriquement à des jonctions en T.*

La figure 7 illustre à l'aide de quatre dessins le principe de complétion amodale. Le premier dessin représente quatre objets, que l'on pourrait qualifier de papillons noirs. Le second dessin est construit en superposant au premier quatre rectangles blancs. La reconstruction visuelle des objets noirs est alors totalement différente : à la place des papillons noirs, nous voyons des disques noirs partiellement recouverts par les rectangles. Cette différence d'interprétation est due à l'application, pour le second dessin, du principe de complétion amodale. Le quatrième dessin reprend une partie du deuxième dessin, en rendant explicites les courbes de contour. Nous y discernons quatre jonctions en T, et le principe de complétion amodale stipule qu'alors les bords de l'objet occulté doivent être mentalement prolongés en préservant sa courbure. L'objet reconstruit a pour bord un cercle, c'est donc un disque.

Le troisième dessin de la figure 7 présente une alternative à la superposition des quatre rectangles : cette fois-ci, des traits sont ajoutés aux papillons noirs de façon à donner l'impression d'une croix blanche. L'emplacement, à chaque coin, de deux des quatre jonctions en T précedentes est préservé; cependant l'objet mentalement reconstruit est totalement autre : au lieu de voir des disques noirs, nous voyons un grand carré noir. La raison en est que la géométrie de ces jonctions est différente : la continuité de direction implique ici de prolonger le contour par un trait droit.

Un autre type de jonction intervient dans le processus de reconstruction visuelle : dans le cas où deux contours réguliers se croisent, on obtient une figure en "X" que nous appellerons *jonction en X*. Ce phénomène s'observe lorsque deux objets se recouvrent partiellement, l'un étant vu par transparence à travers l'autre. Une telle situation se rencontre également quand les ombres de deux objets "s'additionnent" en partie. Le principe de *transparence/ombrage* stipule ainsi qu'en présence d'une jonction en X, les quatre régions formées par la jonction se regroupent par paires de manière à reconstruire deux objets se recouvrant partiellement.

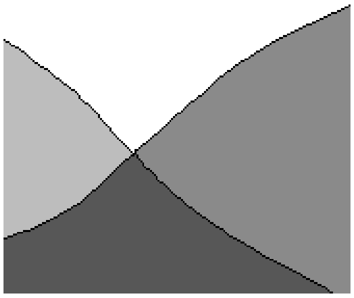

FIG. 8 – *Le phénomène de transparence (ou d'ombrage) apparaît suite à la présence de jonctions en X.*

La figure 8 présente un exemple de jonction en X et d'application du principe de transparence/ombrage : les deux régions les plus claires sont unifiées, ainsi que les deux régions les plus sombres.

Le principe de *constance de largeur* intervient lorsque deux courbes sont parallèles, c'est-à-dire lorsque l'on peut faire glisser entre-elles un disque tangent aux deux courbes, comme sur la figure 9. Les deux courbes sont alors rigidifiées et elle délimitent une région. Remarquons que c'est l'application de ce principe, en collaboration avec le principe d'identité chromatique, qui permet la rigidification des traits comme ceux que l'on trace avec la pointe d'un crayon.

Le principe de *symétrie* entraîne le regroupement d'objets invariants par une symétrie axiale, l'ensemble étant alors rigidifié (figure 10). Il s'agit d'un principe faible, au sens où il s'exprime toujours en collaboration avec d'autres (comme la similitude et la proximité) dont il tend à renforcer l'action.

Le principe de *convexité* est un principe de courbure minimale : il favorise l'interprétation de l'image en des objets convexes. Ainsi, on a tendance à interpréter la figure 11 comme étant composée d'ovales blancs sur fond noir, alors que les principes d'identité chromatique et de fermeture pourraient fournir une articulation fond-forme différente : des formes étoilées noires sur fond blanc.

Le principe de *perspective* entraîne une interprétation de la scène en perspective; il s'applique lorsque l'image est perçue comme la représentation plane d'une scène tridimensionnelle. Cette interprétation est provoquée par la présence de trois segments de droites concourants indiquant soit un point de fuite, soit le coin d'un objet formé par ses contours d'arête (on a alors une jonction en Y). En perspective centrale, un point de fuite est le résultat de la projection d'un ensemble de droites

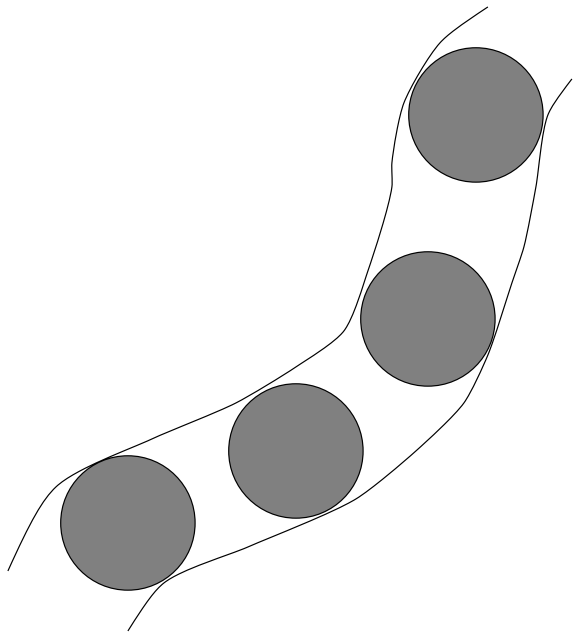

Fig. 9 – *Le principe de constance de largeur s'applique à ces deux courbes car elles sont parallèles.*

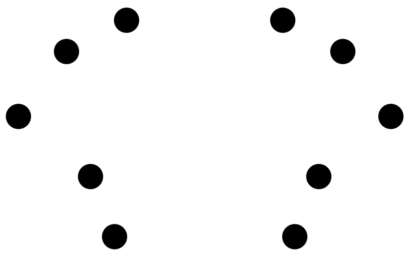

Fig. 10 – *Regroupement perceptif provoqué par une symétrie axiale.*

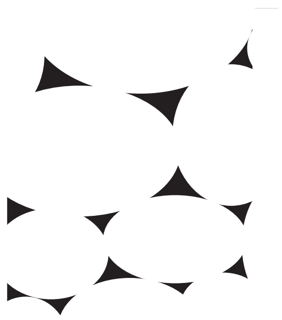

FIG. 11 – *Ovales blancs sur fond noir ou étoiles noires sur fond blanc? le principe de convexité tend à favoriser la première interprétation.*

parallèles sur un plan. Ainsi, la figure 12 possède trois jonctions en Y (a, b et c) et un point de fuite en d.

L'interprétation d'une scène en perspective modifie la compréhension globale de l'image. Tout se passe comme si notre perception mettait en place des procédures de compensation visant à reconstruire la forme et la taille intrinsèques des objets. Ainsi, l'objet de la figure 12 est-il interprété comme une poutre de section carrée constante, mais de grande longueur. Ces mécanismes de compensation peuvent être à l'origine d'illusions visuelles optico-géométriques.

Le principe de perspective tisse ainsi un lien entre les expériences des physicistes et la formalisation gestaltiste de la reconstruction visuelle.

Que le lecteur familier de M.Wertheimer ou de G.Kanizsa ne s'étonne pas que nous ayons passé sous silence certains de leurs principes : ceux qui nous paraissent sortir du domaine de la constitution immédiate des objets visuels sont hors de notre propos (il en est ainsi des principes de "prégnance", "facteur empirique", "régularité", "organisation sans restes", "directionnalité", ...). De même, nous ne nous intéressons pas aux éventuels conflits entre les principes : l'étude de ces conflits par les gestaltistes les a conduit à construire des figures géométriques singulières, régies simultanément par de nombreux principes, avec des résultats peu concluants. Ensuite, elle ne permet pas de conclure de manière certaine à la résolution d'un conflit à l'avantage d'un principe plutôt que d'un autre : cet arbitrage est trop dépendant des conditions particulières à chaque figure.

Après avoir présenté les principes de la reconstruction visuelle, deux questions se posent : peut-on trouver une justification naturelle à l'existence de ces principes géométriques? Peut-on trouver une règle générale qui indique comment les différents

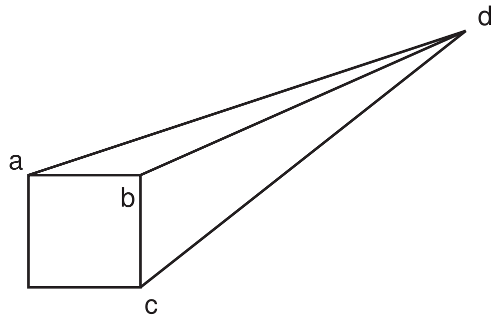

FIG. 12 – *Les jonctions en Y (a,b,c) et le point de fuite d permettent une interprétation en perspective de cette scène.*

principes coopèrent dans la reconstruction ? Nous avons proposé dans [1] une réponse de nature physique à ces problèmes : le principe de *généricité*. Ce principe stipule que la reconstruction visuelle est toujours effectuée en supposant que les objets présents dans le champ visuel ne sont pas perçus dans une situation ou sous un angle de vue particulier ou dégénéré, mais au contraire générique.

3 Le gestaltisme appliqué aux images numériques

Si les principes gestaltistes de la reconstruction visuelle expliquent de manière satisfaisante la mise en correspondance des objets perceptifs avec les objets physiques, ils s'appliquent non seulement dans le cas des figures simples mais aussi dans celui des images naturelles. La théorie gestaltiste ne saurait donc être complète sans une formalisation adaptée aux images naturelles. Ces images, numérisées pour permettre un traitement informatique, se prêtent à l'expérimentation comme les figures tracées sur du papier à l'aide d'un crayon, où l'ordinateur remplacerait ce dernier. L'informatique autorise ainsi une expérimentation dont la finesse était inaccessible aux gestaltistes : ces derniers devaient partir de la représentation atomique d'une image sous forme de traits, alors que la constitution visuelle des traits est déjà une conséquence de la reconstruction visuelle.

Une image numérique que nous désignons par u consiste en une matrice de points ou *pixels*, à chacun de ces points x étant affecté un niveau de gris $u(x)$ (pour une image en noir et blanc) ou un vecteur de trois composantes (pour une image couleur). A l'évidence, la reconstruction visuelle se satisfait pleinement d'une image en noir et blanc, pourvu que la quantification des niveaux de gris soit suffisamment fine. Par la suite, nous nous restreignons aux images en noir et blanc, en sachant que la généralisation à la couleur est possible. La structure mathématique de l'image numérique correspond à peu près à la nature physiologique des stimuli visuels, ce qui signifie qu'une simulation informatique de la reconstruction visuelle devrait permettre une constitution géométrique et informatique des objets présents dans l'image. Sans prétendre donner l'ensemble des clés permettant d'élaborer un algorithme de simulation complet, nous proposons une approche qui suggère de nouveaux algorithmes de traitement de l'image.

Parmi l'ensemble des principes de la reconstruction visuelle présentés à la section 2, seul le principe fondamental *d'identité chromatique* peut facilement se simuler à partir de la structure matricielle des images, et ceci par des techniques de densités locales. Les autres principes font appel à des considérations géométriques portant sur des objets déjà unifiés, ce qui implique une représentation de l'image en des termes géométriques que sont les régions, les courbes, etc. Il faut s'assurer de la complétude d'une telle représentation, c'est-à-dire de la possibilité de reconstruire l'image originale sous sa forme matricielle, et ceci sans dégradation de son sens.

M.Wertheimer, un des fondateurs de la théorie de la forme (*Gestalt*) avait remarqué dès 1923 que notre reconnaissance des objets est indépendante de leur couleur ou luminosité. L'énergie lumineuse reçue et émise par les objets varie en effet considérablement selon l'éclairage ambiant et les objets environnants (de 1 à 10000 entre midi et le crépuscule !). Georges Matheron, de l'Ecole des Mines de Fontainebleau, en a tiré les conséquences pratiques sous la forme d'un principe que doivent vérifier les méthodes d'analyse d'images [12]: le principe d'invariance par changement de contraste. Selon ce principe, que nous appellerons principe de Wertheimer, l'analyse de l'image ne doit pas dépendre de l'échelle arbitraire selon laquelle les niveaux de gris sont mesurés. Nous allons illustrer le principe de Wertheimer par une expérience très commune en traitement d'images : le changement de contraste d'une image numérique. Un changement de contraste consiste à modifier l'échelle des niveaux de gris de l'image. Par exemple, on décide que le “gris foncé” (30) devient plus clair (80), que le “gris clair” devient “gris très clair” (120 devient 180), etc. Un changement de contraste est aisément représenté par un graphique (figure 13). Si par

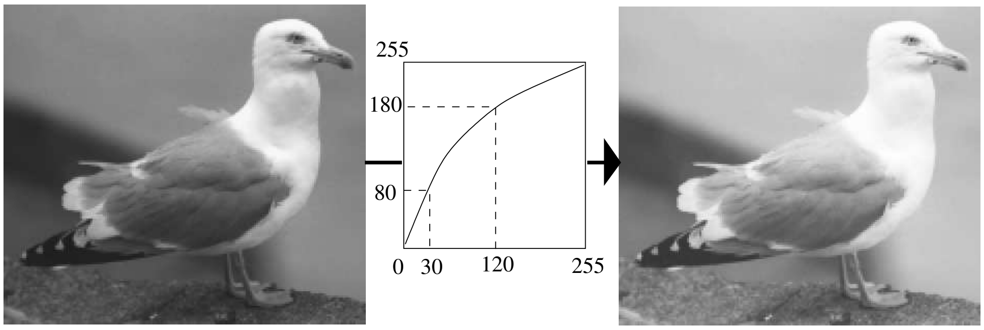

FIG. 13 – *A gauche l'image originale; à droite, l'image obtenue après application du changement de contraste dont nous avons représenté le graphe. Ce changement de contraste accentue le contraste des parties sombres et réduit le contraste entre les parties claires. D'après le principe de Wertheimer, les images de gauche et de droite contiennent la même information sur les formes.*

exemple une image a des parties trop sombres, on fera un changement de contraste qui éclaicit les zones sombres, comme celui que nous venons de décrire. Quand un utilisateur modifie le contraste de son écran d'ordinateur, il applique le principe de Wertheimer : les valeurs effectives des niveaux de gris importent peu, seul compte leur ordre. Le principe de Wertheimer nous oblige à définir pour une image les parties significatives comme des entités qui ne changent pas si on modifie le contraste de l'image. Voyons par un exemple comment définir ces parties significatives.

Nous proposons d'utiliser la *représentation par lignes de niveau*, laquelle est construite à partir des ensembles de niveau. Un ensemble de niveau λ est défini par

$$X_\lambda = \{x/u(x) \geq \lambda\}.$$

La suite des ensembles de niveau (X_λ) permet de reconstruire l'image par la formule

$$u(x) = \arg\sup_\lambda \{x \in X_\lambda\}.$$

La représentation par lignes de niveau contient les bords des composantes connexes des ensembles de niveau, qui sont des courbes de Jordan fermées (orientées afin de différencier l'intérieur de l'ensemble de son extérieur), ainsi que le niveau associé λ. Un simple algorithme de remplissage permet de retrouver à partir de cette information la suite des ensembles de niveau (X_λ) : la représentation par lignes de niveau est donc complète.

On appelle *carte topographique d'une image* une figure qui trace l'emplacement de ses courbes de niveau. Afin de garder une lisibilité à la carte topographique, seuls certains niveaux de gris λ sont représentés, ce qui correspond à une quantification de l'image. Cette quantification, qui préserve le sens de l'image si elle n'est pas trop importante, n'est effectuée que pour la clarté de la présentation : les algorithmes de traitement basés sur la représentation par lignes de niveau ne nécessitent pas de quantification.

La figure 15 donne un exemple simple de carte topographique, l'image numérique associée représentant un dessin géométrique (figure 14). Sur cet exemple, les niveaux de gris λ sont pris de 20 en 20. L'image a été obtenue par numérisation sur 256 niveaux de gris d'une portion de la figure 7 page 7. Dans la figure 17, on a représenté, de gauche à droite et de haut en bas, six ensembles de niveau de l'image (niveau croissant de 40 en 40). La dernière image du tableau est obtenue par reconstruction à partir de ces six ensembles de niveau.

Dans beaucoup de cas, les ensembles de niveau nous donnent effectivement les objets présents dans l'image : c'est le cas pour un document écrit en noir sur blanc. Si les lettres sont noires et donc proches de 0 tandis que le fond est blanc et donc proche de 255, on choisit par exemple l'ensemble de niveau 128; il va séparer les lettres du fond.

La situation se complique dès que l'image à analyser est une photographie du monde "réel". Dans ce cas, les objets photographiés se superposent à un fond lui-même composé d'objets; en réalité la plupart des objets se cachent partiellement les uns les autres. Ainsi dans la figure 18, les ensembles de niveau formés (figures 18-c

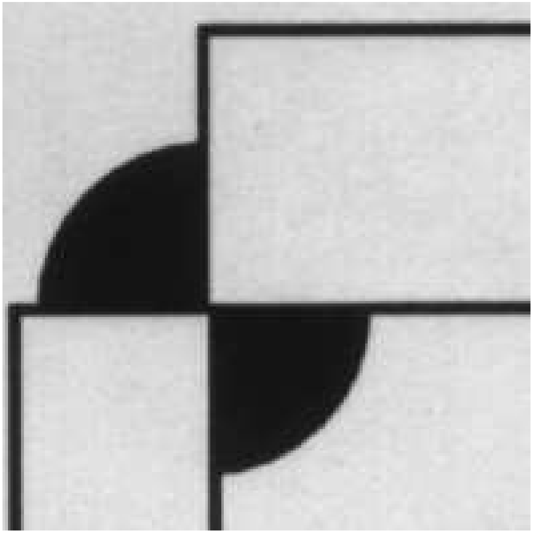

FIG. 14 – *Portion de la figure 7 page 7 numérisée pour donner une image sur* 256 *niveaux de gris.*

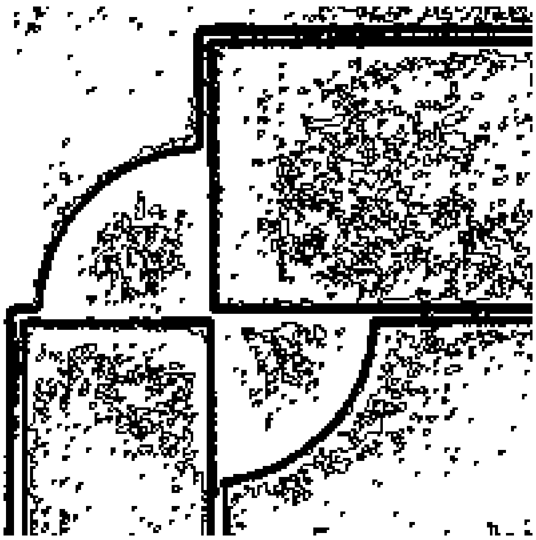

FIG. 15 – *Carte topographique de l'image précédente, les lignes de niveau étant tracées avec un pas de quantification de* 20.

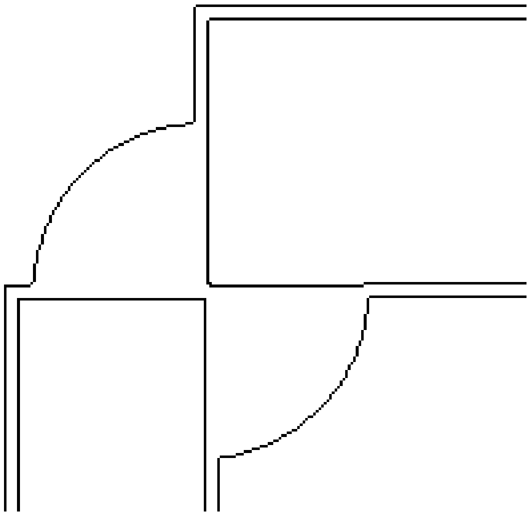

FIG. 16 – *Carte topographique de l'image présentée à la figure 14, seules les lignes de niveau de longueur importante étant conservées (pas de quantification de* 130*).*

FIG. 17 – *De gauche à droite et de haut en bas : l'image originale, les ensembles de niveau 40, 80, 120, 160, 200 et 240 et, enfin, l'image reconstruite à partir de ces ensembles de niveau.*

et 18-d) sont des unions ou des différences ensemblistes des objets que l'on imagine présents dans la scène (un carré et un disque).

En quoi la représentation par lignes de niveau est-elle adaptée à la nature

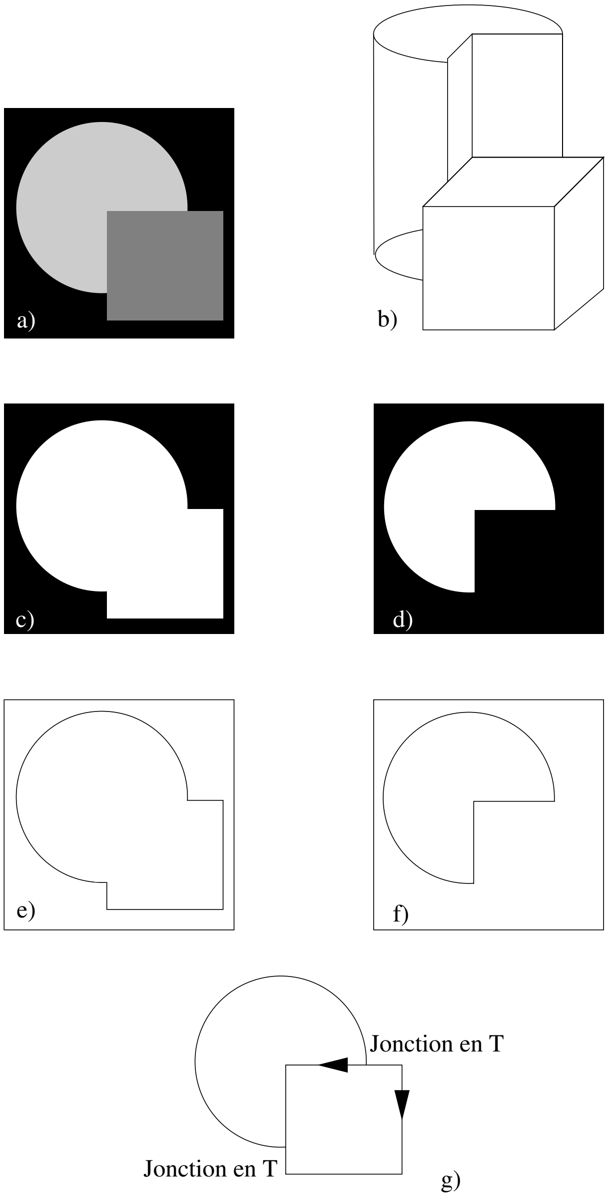

FIG. 18 – *a) Une image représentant un carré "sur" un disque; b) vue en relief; c-d) ensembles de niveau, obtenus en effectuant des coupes horizontales du relief; e-f) lignes de niveau à deux niveaux différents, obtenues comme les bords des ensembles de niveau; g) superposition des lignes de niveau aux deux niveaux différents. Les jonctions en T se forment à l'endroit où les lignes de niveau se séparent. D'après Kanizsa, ces jonctions en T nous indiquent l'endroit où un objet en recouvre un autre, nous permettant d'effectuer une "complétion amodale", c'est-à-dire ici de reconstituer un disque clair derrière le carré sombre.*

géométrique des principes de la reconstruction visuelle? les contours d'arête se reconnaissent aisément sur la carte topographique, de part le contraste localement plus important qui entraîne une densité supérieure des lignes de niveau. D'autre part, on remarque que les lignes de niveau irrégulières sont isolées et qu'elles ne contribuent pas à la constitution des objets : des courbes régulières en densité suffisante doivent donc entraîner la rigidification de l'ensemble délimité par ces bords. Nous retrouvons ici le *principe de continuité de direction*, formulé dans le cadre des images numériques (contours illusoires exclus). Nous allons préciser comment les trois principes élémentaires associés (fermeture, complétion amodale et transparence/ombrage) se retrouvent dans la représentation par lignes de niveau. Nous laissons au lecteur le soin de vérifier qu'une fois acquis la constitution de premiers objets visuels par application des principes d'identité chromatique et de continuité de direction, et leur caractérisation par des lignes de niveau, l'application des autres principes de la reconstruction visuelle résulte de simples exercices de géométrie. Ces exercices sont semblables à ceux que nous avons décrits dans le cadre des figures simples, les courbes formées par les traits étant remplacées par celles des lignes de niveau perceptuellement significatives (c'est-à-dire, régulières et en densité suffisante). Remarquons que les lignes de niveau oscillantes constituent un exemple physique de masquage dans le processus de reconstruction visuelle. La figure 21 présente plusieurs longues lignes de niveau oscillantes qui correspondent à des structures invisibles sur l'image originale (figure 19), quoique bien présentes.

Le *principe de fermeture*, qui a pour fonction d'assimiler à la trace d'un objet l'intérieur d'une courbe fermée, est une conséquence directe de la représentation par lignes de niveau (toute ligne de niveau étant une courbe fermée). Le *principe de complétion amodale* est liée à la détection des jonctions en T, tout comme le *principe de transparence/ombrage* à la détection des jonctions en X. Les cartes topographiques des figures 15- 16 (pour les jonctions en T) et des figures 20, 21 (pour les jonctions en X) illustrent le fait que les lignes de niveau possèdent des jonctions de même nature géométrique que les courbes des figures simples, qu'elles sont clairement identifiables, et qu'elles sont bien la signature de phénomène d'occlusion ou de transparence/ombrage. Dans [3] est décrit un algorithme de détection des jonctions en T et en X sur des images numériques, à partir de la représentation par lignes de niveau.

4 Quelques algorithmes "gestaltistes"

Comme nous pouvons reconstruire une image simplifiée à partir de certains de ses ensembles de niveau, on peut mener quelques expériences surprenantes qui nous prouvent d'abord que l'on peut parfois enlever 80% des ensembles de niveau sans altération flagrante de l'image ! Cette opération de quantification peut être utilisée pour "comprimer" l'image, c'est-à-dire réduire la taille qu'elle occupe dans une mémoire d'ordinateur.

Toutefois, une quantification poussée trop loin conduit à rendre visible des lignes de niveau de l'image qui ne correspondent pas à des bords physiques d'objets. Ce phénomène numérique illustré par la figure 22-2 se produit quand une image trop

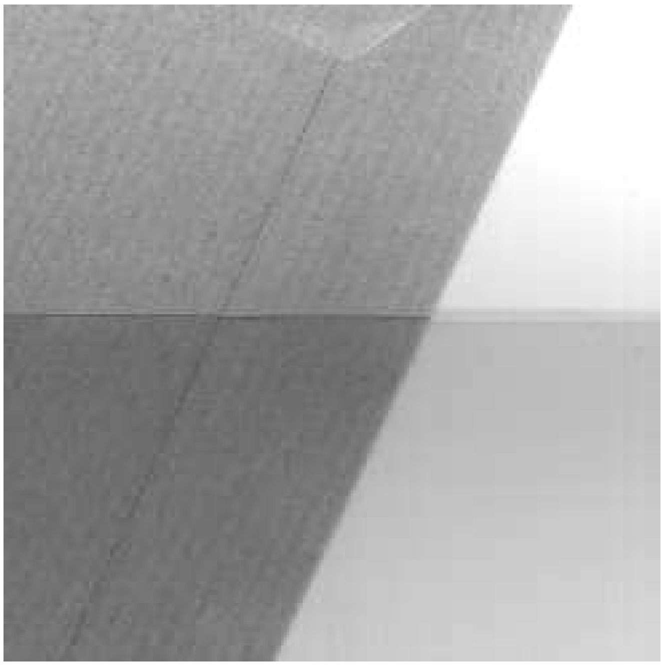

FIG. 19 – *Cette image a été obtenue par numérisation sur 256 niveaux de gris d'une portion de scène représentant une enveloppe partiellement recouverte par une feuille transparente.*

sombre (sous-exposée) ou trop claire (surexposée) a été numérisée sans contrôle. C'est en particulier le cas pour beaucoup d'images prises par les satellites d'observation de la Terre (SPOT). De telles images peuvent être restaurées. Vicent Caselles et Catalina Sbert, de l'Université des Iles Baléares, ainsi que l'un d'entre nous [4][1], ont montré comment on peut reconstruire une image plausible en interpolant entre les lignes de niveau, de manière à restituer les lignes de niveau manquantes. Cet algorithme respecte le principe de Wertheimer, et on peut apprécier l'efficacité du procédé de restauration sur la figure 22-4. Il devient dès lors possible d'écrire un algorithme de compression d'images basé sur la sélection d'un très petit nombre de lignes de niveau significatives [5].

Une autre méthode extrêmement élégante pour réduire la complexité de l'image consiste à éliminer ses "grains", à savoir toutes les lignes de niveau qui entourent

1. Dans cet article, il est montré que la seule méthode d'interpolation isotrope, stable, régulière, vérifiant le principe du maximum et permettant d'interpoler aussi bien des valeurs ponctuelles que des courbes, est donnée par l'équation AMLE (Absolute Minimizing Lipschitz Extension) inventée par G. Arronsson [2]. L'interpolée $u(x)$ vérifie $D^2u(Du, Du) = 0$. Jensen [6] a montré l'unicité de la solution au sens des viscosités de cette équation. La solution de AMLE a la propriété amusante suivante : on ne peut pas pas la réinterpoler sur un sous-domaine par une solution qui aurait une constante de Lipschitz inférieure; cette constante est donc partout la meilleure possible. Un exemple d'interpolation par AMLE est le suivant : si l'on veut interpoler la fonction qui vaut 1 sur le bord du cercle de centre 0 et de rayon 1, et 0 au centre, on obtient un cône.

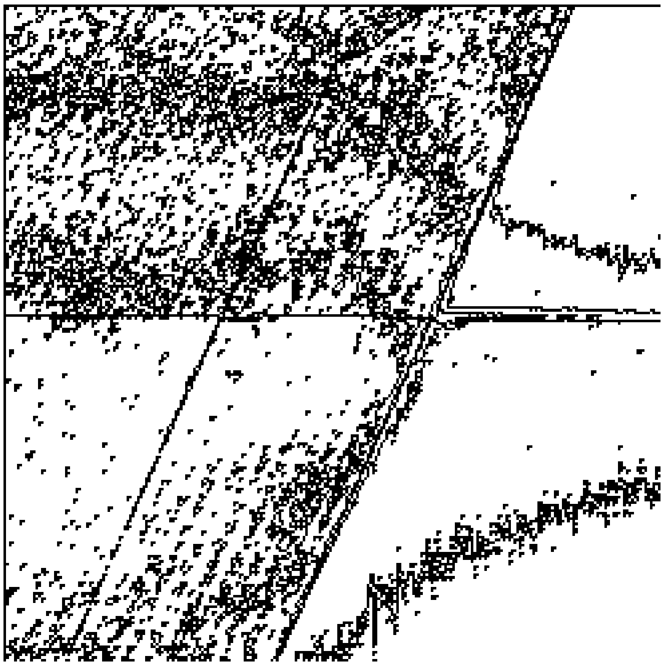

FIG. 20 – *Carte topographique de l'image précédente, les lignes de niveau étant tracées avec un pas de quantification de 20. Deux jonctions en X, caractéristiques des phénomènes de transparence, sont clairement identifiables.*

une aire petite. La figure 23 nous montre l'effet d'une telle opération : l'aire choisie est de 10 pixels, ce qui est déja assez considérable. Toutefois, il est difficile de voir une différence entre l'image brute et l'image "nettoyée" de ses grains inférieurs à 10 pixels, alors que cette même opération appliquée à une image "bruitée" a des résultats spectaculaires. L'algorithme utilisé [10] est une variante d'algorithmes légèrement plus complexes, mais d'effet équivalent, dus à Luc Vincent (de la compagnie Rank Xerox) et Jean Serra (de l'Ecole des Mines de Fontainebleau) d'une part et à Leonid Yaroslavsky (Académie des Sciences de Russie et Université de Tel Aviv) d'autre part.

Ensembles de niveau et comparaison de deux images

La thèse que les formes dans une image sont décrites par les ensembles de niveau conduit à une application pratique importante : la comparaison d'images prises à des dates différentes. La figure 24 montre un bureau photographié à 25 heures d'intervalle. Certains objets ont changé de place (chaise, manteaux, fil électrique). L'éclairage et donc le contraste ont également changé. Ainsi, une comparaison des niveaux de gris en chaque pixel ne permettrait pas de déterminer ce qui est commun aux deux images et ce qui n'a pas bougé. L'algorithme d'intersection d'images développé par C. Ballester et M. Gonzalez de l'Université des Iles Baléares, ainsi

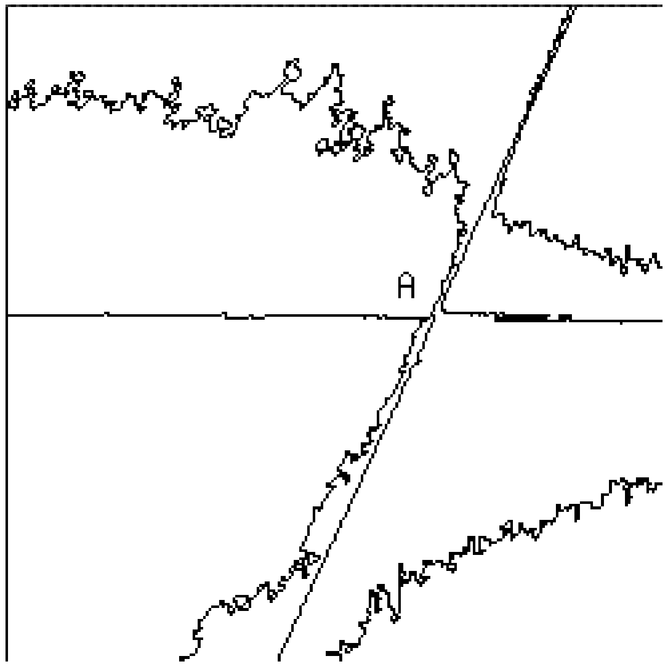

FIG. 21 – *Carte topographique de l'image présentée à la figure 19, seules les lignes de niveau de longueur importante étant conservées (pas de quantification de* 30*). Une seule jonction en X reste visible (au point A) : elle correspond à la plus contrastée des deux.*

que E. Cubero-Castan du CNES, permet une telle comparaison. Il compare deux à deux toutes les composantes connexes des ensembles de points dont l'intensité est comprise entre deux valeurs. Ces ensembles sont appelés sections et vérifient le principe de Wertheimer. Puis, l'algorithme reconstitue une image "commune", qui est une véritable intersection car elle contient toutes les sections communes aux deux images. Les parties noires signalent les endroits où un objet a changé de place. On constate que dans ce bureau de chercheur, tout a bougé d'un jour à l'autre, sauf les parois et les livres !

Pour donner un autre exemple, ce même algorithme s'applique avec succès à des images satellitaires prises à un an d'intervalle, permettant de retrouver les structures géologiques stables.

Revenons maintenant à la figure 18 : les ensembles de niveau nous permettent par différence ensembliste de récupérer un objet de l'image, en l'occurrence le carré. Mais ils ne permettent d'obtenir que les trois quarts du disque, alors que notre impression visuelle est que l'image contient un disque entier, dont un quart est occulté par le carré. En fait, les deux constituants d'une image vérifiant le principe de Wertheimer que sont les ensembles de niveau et les lignes de niveau, ne sont pas suffisants pour décrire les objets tels que nous les "voyons". C'est que la vision humaine s'accompagne d'un phénomème psychophysiologique que nous allons à présent tenter de

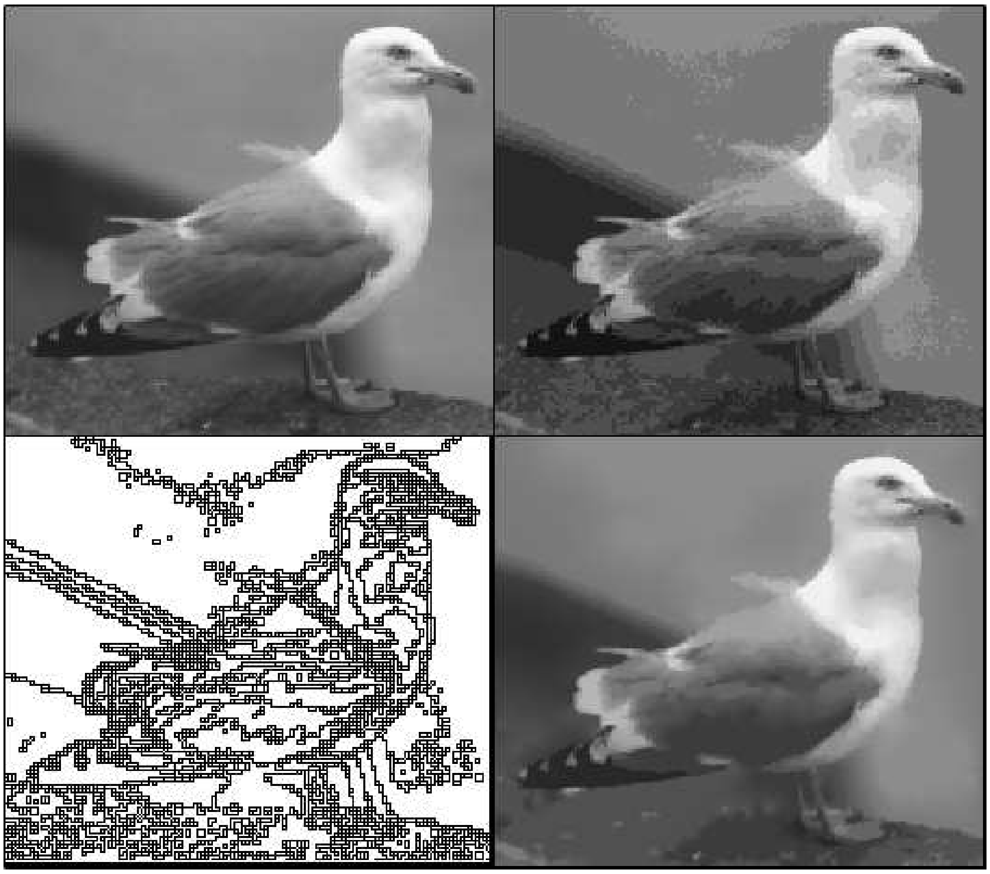

FIG. 22 – *1. Image originale – 2. Image obtenue après quantification (de 20 en 20) – 3. Lignes de niveau de l'image quantifiée – 4. Image restaurée par interpolation entre les lignes de niveau.*

simuler : la complétion amodale.

Un algorithme de désocclusion

Les observations de Kanizsa peuvent faire l'objet d'un algorithme surprenant : l'algorithme de désocclusion. Il se produit souvent qu'une photographie ou une partie d'un film soit abimée, ou tachée. On peut certes atténuer ou gommer la tache, mais qu'y avait-il dessous ? Nous l'ignorons bien sûr, mais le principe de Kanizsa indique comment aller "au delà de l'information donnée" en prolongeant les bords visibles "derrière" l'objet qui les interrompt. Peut-on simuler la complétion amodale de Kanizsa pour éliminer les taches ? La réponse est oui, si les objets cachés par la tache ne le sont que partiellement.

Comme l'avaient déjà remarqué David Mumford, de l'Université Brown, et Marc Nitzberg, de la société Microsoft, nous disposons pour cela d'une information clé : les jonctions en T. En effet, les lignes de niveau de l'image aboutissent au bord de la tache et le principe de complétion amodale nous invite à les prolonger. Toutefois, ce prolongement peut conduire à des conflits insolubles : pour des raisons

FIG. 23 – *En haut à gauche l'image originale ; en haut à droite, l'image obtenue après suppression des grains de taille inférieure à 10 pixels. En bas à gauche l'image originale bruitée et à droite, l'image après suppression des grains de taille inférieure à 10 pixels. Cette expérience montre que le filtre de grains a la propriété remarquable de laisser presque intactes les images de bonne qualité et de restaurer celles qui sont bruitées.*

mathématiques simples, deux lignes de niveaux différents ne peuvent se croiser. En outre, toute ligne de niveau se doit d'être une courbe qui se referme sur elle-même. Un algorithme de "complétion amodale" doit respecter ces deux contraintes.

L'algorithme que nous avons développé [11] consiste à joindre les lignes de même niveau deux à deux de sorte que les courbes de niveau ainsi complétées soient les plus lisses et droites possibles. Notre méthode ne permet pas de reconstruire l'image telle qu'elle était, mais elle reconstruit une image plausible, compatible avec les lois de la perception, ce qui la rend agréable et pour ainsi dire "logique". Pour être plus précis, l'algorithme consiste à trouver parmi tous les recollements possibles deux à deux des lignes de même niveau, celui qui minimise une énergie faisant intervenir la

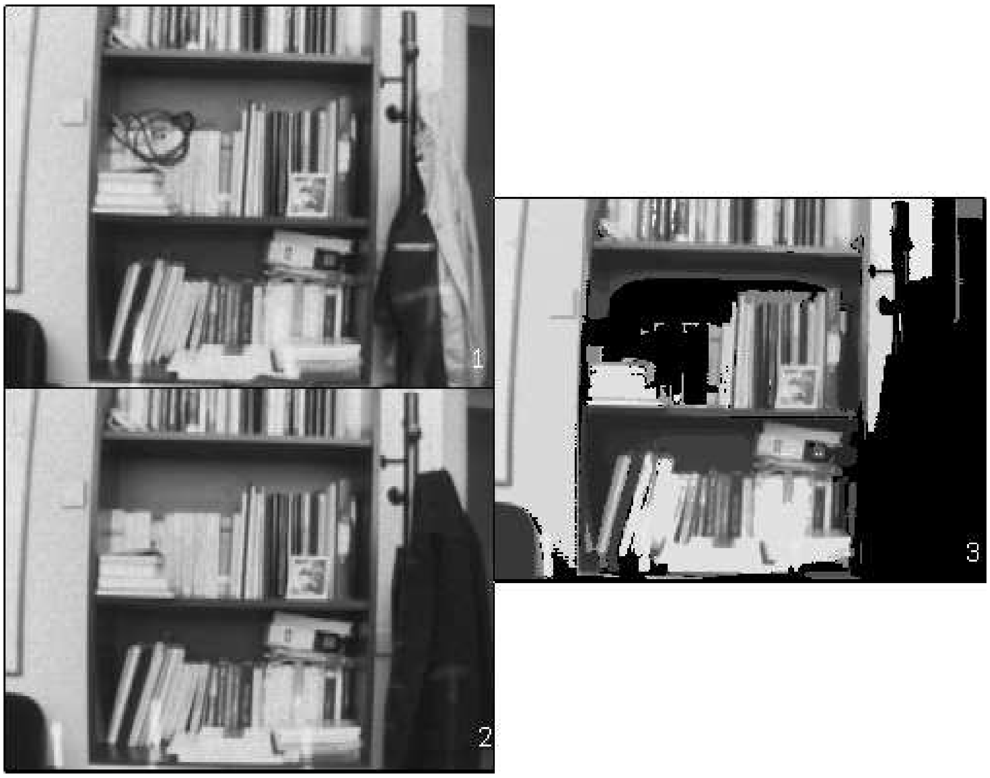

FIG. 24 – *Les deux premières images sont des prises de vue d'un bureau à 25 heures d'intervalle. Le contraste a changé et des objets se sont déplacés. Un algorithme respectant le principe de Wertheimer permet de calculer une véritable intersection des deux images (image 3) où seuls apparaissent les objets qui n'ont pas bougé ou n'ont pas été occultés.*

longueur des courbes de recollement ainsi que la variation totale de la courbure et qui est en fait une relaxée de

$$E(u) = \int_{\Omega} |Du|(1 + |\mathrm{curv}(u)|)dx$$

où Ω désigne la tache à traiter. Les différentes étapes de l'algorithme sont les suivantes :

Etape 1 : détermination de la ligne polygonale décrivant la frontière de l'occlusion.

Etape 2 : énumération dans le sens trigonométrique des jonctions en T appartenant à la frontière; chaque jonction est déterminée par sa position sur la frontière, ses coordonnées, les niveaux de gris auxquels elle est associée et la direction moyenne de la ligne de niveau correspondante.

Etape 3 : triangulation de la frontière.

Etape 4 : détermination par programmation dynamique de l'ensemble des appariements de jonctions en T qui a la plus faible énergie.

Etape 5 : tracé des géodésiques associées à ces appariements.

Etape 6 : propagation géodésique à l'intérieur de l'occlusion des niveaux de gris définis par les lignes reconstruites. Ces deux dernières étapes nécessitent un nombre d'opérations qui dépend linéairement du nombre de pixels de l'occlusion.

L'utilisation de la programmation dynamique permet un gain de temps considérable par rapport à la méthode qui consisterait à énumérer tous les recollements possibles. Son utilisation est rendue possible par le principe de causalité qui régit les connections deux à deux des jonctions en T. Sachant que deux lignes de niveau d'une image ne peuvent pas se croiser, toute décision de recollement d'une jonction a avec une jonction b va contraindre tous les recollements ultérieurs, puisqu'ils ne pourront avoir lieu que d'un côté ou de l'autre de la courbe joignant a et b. Si N est le nombre de jonctions en T, la méthode d'énumération a un coût algorithmique de l'ordre de $N!$ tandis que ce coût devient de l'ordre de N^3 lorsque l'on utilise la programmation géodésique.

On peut juger sur la figure 25 des effets de notre algorithme losqu'on l'applique aux taches blanches présentes dans l'image de gauche. La figure 26 illustre par ailleurs le fait que la désocclusion par interpolation singulière, c'est-à-dire qui permet des discontinuités, est beaucoup plus efficace que la méthode d'interpolation régulière basée sur l'équation AMLE. Notre algorithme peut encore être amélioré mais les premiers résultats obtenus sont très encourageants et prouvent le bien fondé de l'approche gestaltiste.

Il faut savoir que la plupart des dégradations d'images courantes sont des taches, c'est-à-dire que l'image a des parties détruites. Guy Lesthievent, du Centre National d'Etudes Spatiales, a eu l'idée d'utiliser notre algorithme de désocclusion pour éliminer les dégradations d'images du satellite SPOT dues à des pertes d'information lors de la transmission. Ces pertes se manifestent par la présence de taches dans l'image (figure 27) que l'on peut ainsi très efficacement restaurer.

5 Conclusion

L'élaboration d'algorithmes vérifiant les principes gestaltistes ne fait que débuter. La première liste des principes de constitution des formes établie par Wertheimer contenait déjà six principes (convexité, parallélisme, alignement, similitude, symétrie et proximité). Plusieurs de ces principes demanderont une réflexion mathématique intense car ils sont, au contraire de ceux que nous avons appliqués dans cet article, non locaux, c'est-à-dire impliquant des comparaisons entre points très distants d'une même image.

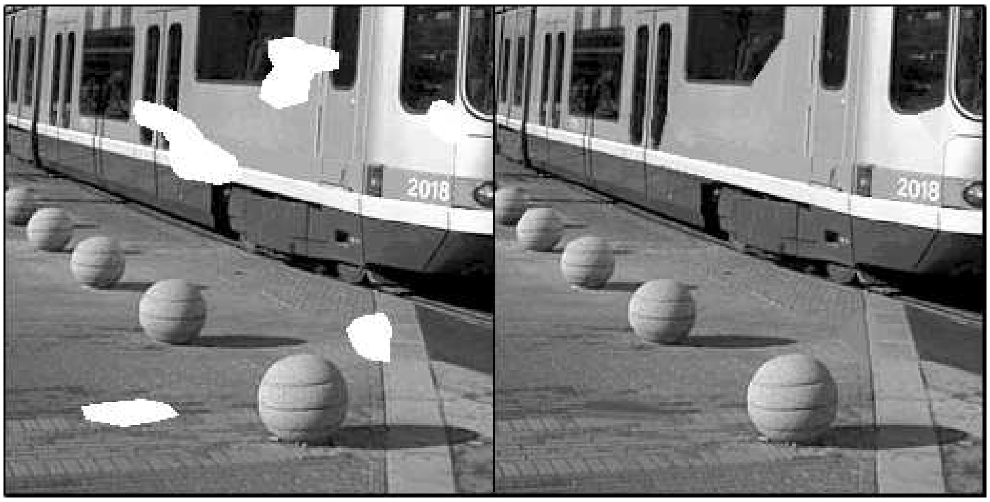

FIG. 25 – *Algorithme de désocclusion ou complétion amodale – à gauche l'image originale et les occlusions en blanc; à droite, l'image obtenue à l'aide d'un algorithme de "complétion amodale". Cet algorithme recolle de la manière la plus plausible les lignes de niveau de l'image "abimée" qui aboutissent à une tache. Ce recollement ne reconstruit pas l'image originale, mais propose une restauration de l'image où les effets des destructions sont presque totalement masqués.*

Références

[1] Jean-Pierre D'Alès, Jacques Froment et Jean-Michel Morel, "Reconstruction visuelle et généricité", preprint CEREMADE, 1997.

[2] G.Arronsson, "Extensions of functions satisfying Lipschitz conditions", *Arkiv für Matematik*, 6, 1967, p. 551-561.

[3] Vicent Caselles, Bartomeu Coll et Jean-Michel Morel, "A Kanizsa programme", Cahiers de Mathématiques de la Décision No 9539, Ceremade, Université Paris-Dauphine, Novembre 1995.

[4] V.Caselles, J.M.Morel et C.Sbert, "An Axiomatic Approach to Image Interpolation", *IEEE Trans. on Image Processing*, vol. 7, no. 3, 1998, p. 376-386.

[5] Jacques Froment, "A Functional Analysis Model for Natural Images Permitting Structured Compression", preprint CMLA, Ecole Normale Supérieure de Cachan, 1998.

[6] R.Jensen, "Uniqueness of Lipschitz extension : minimizing the Sup Norm of the gradient", *Arch. Rat. Mech. Anal.*, 123, 1993, p. 51-74.

[7] Gaetano Kanisza, "Grammatica del Vedere", Il Mulino, Bologna, 1980. Traduction française par Antonin Chambolle, Diderot Editeur des Arts et des Sciences, 1997.

FIG. 26 – *En haut, une image et des occlusions en blanc. En bas à gauche, le résultat obtenu par interpolation régulière à l'aide de l'équation AMLE. En bas à droite, le résultat obtenu par l'algorithme de désocclusion.*

[8] Gaetano Kanisza, "Vedere e pensare", Il Mulino, Bologna, 1991.

[9] David Marr, "Vision", W.H.Freeman and Co., 1982.

[10] Simon Masnou et Jean-Michel Morel, "Image restoration involving connectedness", *DIP'97 Workshop*, SPIE, 3346, 1998.

[11] Simon Masnou et Jean-Michel Morel, "Level lines based disocclusion", *ICIP'98 IEEE Int. Conf. on Image Processing*, 1998.

[12] Georges Matheron, "Random Sets and Integral Geometry", John Wiley and Sons, New York, 1975.

[13] Wolfgang Metzger, "Gesetze des Sehens" (les lois de la vision), Waldemar Kramer 1975.

FIG. 27 – *A gauche une simulation d'une image satellitaire CNES altérée par une perte d'information lors de la transmission; à droite le résultat obtenu après désocclusion de chacune des taches.*

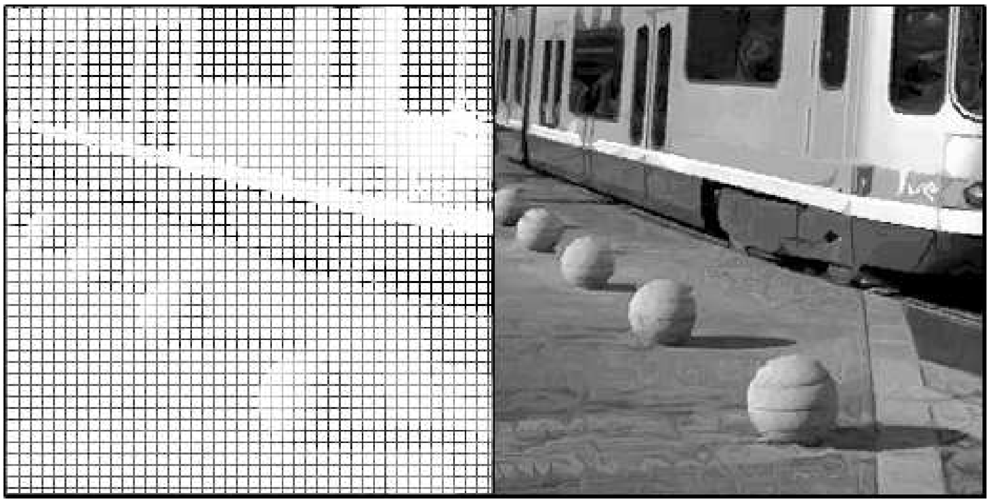

FIG. 28 – *A gauche une image où l'on n'a conservé sur chaque ligne et sur chaque colonne qu'un pixel sur six. Les occlusions sont les carrés blancs de taille 5x5 pixels. A droite, l'image reconstruite par désocclusion*

[14] Jean-Michel Morel, "Perception visuelle et traitement d' images" (I) et (II), Revue du Palais de la Découverte, Novembre 1996 et Février 1997.

[15] Jacques Ninio, "La science des illusions", Odile Jacob, 1998.

[16] Max Wertheimer, "Untersuchungen zur Lehre der Gestalt", Psychologische Forschung, IV, pp. 301-50.