

IMAGE, APPRENTISSAGE, APPROPRIATION VISUELLE DE L'ESPACE

Robert BORODKINE

THEMA UPRESA 6049

Université de Franche-Comté

Résumé

L'image est un élément-clé qui accompagne la démarche géographique dans ses multiples aspects : paysage visible, télédétection ou plus largement appréhension spatiale des phénomènes. De ce point de vue, le traitement automatique de l'image présente un intérêt avéré, mais, malgré des atouts indiscutables, il ne rivalise pas encore en tous points avec la vision humaine dont l'adaptativité reste inégalée. Aussi avons-nous entrepris des travaux plaçant l'apprentissage comme terme central de l'analyse d'image. Ils sont conçus comme une confrontation entre approche cybernétique et approche géographique, cette dernière étant basée sur un rapport à l'espace prenant en compte ses dimensions objectale, fonctionnelle, et subjective. Nous avons notamment approfondi les concepts de représentation et de corrélation. Ces recherches débouchent sur un modèle fondé sur les notions d'apprenabilité et d'appropriation perceptive de l'environnement. Ce modèle éclaire d'un jour nouveau les relations que nous entretenons avec l'espace, objet de perception sensible et de structuration mentale, mais aussi la problématique même de la modélisation du système que constitue l'environnement dans son acception la plus globale.

Abstract

Image is a key-element that accompanies the geographical approach in its multiple aspects : visible landscape, remote sensing or more largely spatial apprehension of phenomena. From this point of view, automatic image processing has an established interest, but, in spite of indisputable advantages, it does not compete yet in all respects with the human vision whose adaptivity remains unequalled. Therefore we undertook works that place learning as central term of image analysis. They have been conceived like a confrontation between cybernetic approach and the geographical one, the latter being based on a report with space that take account of its objectal, functional, and subjective dimension. We in particular looked further into the concepts of representation and correlation. This research leads to a model based on the concepts of learnability and environment visual appropriation. This model clarifies the relations which we deal with space, sensitive object of perception and mental structuring, as well as the problem of the modelisation of the system that constitutes environment in its most total meaning.

Mots-Clés

Analyse d'image, apprentissage, évolution, modélisation, paradigmes naturels et artificiels, perception, représentation, théorie algorithmique de l'information.

Keywords

Algorithmic information theory, image analysis, evolution, learning, modeling, natural and artificial paradigms, perception, representation.

Introduction

Les rapports à l'environnement, la façon même de l'appréhender, sont tributaires de considérations extrêmement diverses, puisque celles-ci s'étendent du pragmatique à l'affectif, de la science au sacré, en passant par l'esthétique. Ces points de vue induisent divers systèmes de représentation et principes d'action. De fait, pour fonctionner avec son environnement, une entité individuelle ou collective conjugue trois éléments : aspirations, connaissance, savoir-faire. Pour l'être vivant en général, il s'agit de survivre grâce à

une certaine « intelligence » de son milieu. Pour le géographe, il est question de fournir à la collectivité des éléments nécessaires à la mise en œuvre des orientations qu'elle se donne, en générant une connaissance de l'environnement et des relations entre actions et impacts sur ce dernier. Dans les deux cas, nous nous attachons spécialement à mettre en évidence des entités à la fois spatiales et temporelles avec lesquelles on est en interaction. Cette analogie pourrait bien ne pas être étrangère au fait que la vision humaine demeure appréciable dans des tâches d'analyse d'image, même dans des contextes qui ne lui sont pas naturellement dévolus. En tout état de cause, les points forts de la vision dans le vivant justifient qu'on veuille s'en inspirer pour réaliser des procédés d'analyse automatique d'images, y compris en géographie où les images considérées sont parfois très différentes de scènes naturelles.

1. Géographie et images

Au cœur des préoccupations géographiques : l'image. On peut s'y intéresser en quelque sorte pour elle-même, car elle est une des composantes du cadre et de la qualité de la vie; on peut aussi s'y intéresser parce qu'elle constitue l'une des sources d'information sur notre environnement.

Dans le premier cas, il va s'agir de considérations esthétiques, psychologiques (le percept ayant un impact direct sur l'humeur, la concentration,...), ou encore pratiques (à rapprocher de la notion d'ergonomie : une image facile à vivre est une image facile à interpréter en tant qu'usager du monde qu'elle reflète; les exemples les plus immédiats sont la route, la cité, et l'habitat). L'appréciation peut se faire à partir d'images réelles, existantes, ou à partir d'images synthétiques qui permettent l'examen depuis des points de vue nombreux ou difficiles à atteindre, ou bien faisant partie des futurs ou des possibles (dans ce cas ce n'est pas seulement l'image qui est virtuelle, mais aussi la scène qu'elle représente).

Dans le second cas, l'image est un signe qui seul ou associé à d'autres permet de constituer une connaissance sur les éléments qui lui ont donné naissance. La modélisation fait appel à des images naturelles de paysages, des photographies aériennes, des images satellitales. La carte, quant à elle, constitue un point de rencontre de ces deux aspects : sa valeur tient à la fois à l'information dont elle est porteuse et à son aptitude à la véhiculer.

Si la vue du dessus est très appréciable pour la connaissance de l'occupation des sols, la vue du dedans – qu'elle soit réelle ou synthétique – est incontournable dans la multitude des problématiques se rapportant à l'évaluation paysagère (composantes de l'attrait touristique d'une région, impacts visuels de l'implantation de lignes électriques, etc.). On se heurte cependant à certaines difficultés dès lors qu'une explicitation suffisante de caractéristiques significatives des éléments à analyser n'est pas immédiate, ce qui est le cas par exemple lorsqu'on s'intéresse aux zones dites sensibles (friches, mitage urbain,...), aux sites favorables à l'accueil d'espèces animales requérant des milieux spécifiques (grand tétras,...), ou à l'une des nombreuses facettes de l'évaluation paysagère.

2. Vision, naturelle et artificielle, et analyse d'images

Le traitement automatique de l'image présente un intérêt avéré – notamment vis-à-vis de maintes problématiques géographiques. Des insuffisances persistent pourtant face à l'adaptativité de la vision humaine, à tel point que cette dernière demeure dans bien des cas un exemple d'efficacité – malgré ses limites – y compris hors de sa vocation première. Les faiblesses de la vision humaine tiennent à ce qu'elle n'a pas une résolution extrême, qu'elle ne travaille que sur une portion du spectre, et qu'elle fait preuve d'une grande variabilité entre individus et pour une même personne.

La vision artificielle n'est efficace que dans des cas bien définis, elle n'est rapide que dans un sous-ensemble restreint de ces cas, elle est très inférieure aux visions humaine et même animale en termes d'adaptabilité; par contre, elle opère sur une large gamme de résolutions et de modalités. Par ailleurs, elle est

reproductible: on peut concevoir des mécanismes qui pour des données identiques fourniront des réponses identiques (ce qui est satisfaisant sur le plan du principe), et elle est automatisable: on peut réaliser des systèmes automatiques destinés à prendre en charge les traitements de façon répétitive (ce qui fait l'intérêt pratique de la chose).

Les travaux que nous avons entrepris sont destinés à contribuer à la réalisation de dispositifs cumulant des avantages des visions naturelle et artificielle. Le rôle de la vision comme de l'analyse d'image est de mettre à disposition une représentation de l'image plus abstraite que cette dernière (faite de couleurs, de lumière, d'ondes, de formes, de textures,...), tout en conservant les informations importantes. Aussi retrouve-t-on cette idée de transformations allant dans le sens d'une abstraction croissante dans la plupart des approches artificielles. Certains procèdent de manière exclusivement ascendante, alors que d'autres prennent également en compte l'influence que peuvent exercer en sens inverse les niveaux symboliques sur l'émergence de ce qui est perçu. Le caractère dynamique des images est lui aussi utilisé, il peut notamment révéler l'unité d'éléments constitutifs d'un même objet – ou d'une même portion d'objet – à travers un principe de destin commun qui présente des analogies avec la notion de cause commune dans les images statiques.

3. Position personnelle et démarche adoptée

Une idée directrice est le principe gestaltiste selon lequel le tout représente autre chose que la somme des parties. Précisons qu'il faut voir la Gestalt Theorie [14] à travers la pensée d'Egon Brunswik [3], qui avait mis en avant la valeur fonctionnelle des phénomènes gestaltistes, le système visuel – façonné par l'évolution – devant permettre – à partir d'informations de nature complexe et incertaine – de survivre au sein d'un environnement avec lequel on est en relation. De Gregory [7] et Marr [11], nous rejetons l'ampleur qu'ils accordent aux aspects reconstitutifs et déductifs de la vision, nous conservons par contre l'idée de transformations allant dans le sens d'une abstraction des entités, au contraire de Gibson [5] dont nous considérons par ailleurs avec le plus grand intérêt les concepts d'invariants et d'extraction d'informations liées à des buts.

Quant à la façon de construire un système de vision artificiel, plutôt que de s'évertuer à vouloir copier le système animal ou à en imaginer un équivalent, il nous semble qu'il serait intéressant de l'obtenir à partir de principes phylogénétiques et ontogénétiques. Les principes que nous tentons de reproduire en priorité ne sont alors plus ceux d'un système de vision, mais ceux qui permettent de le façonner. Soulignons le fait qu'il n'est pas pour autant question d'ignorer les acquis et les perspectives phénoménaux des recherches menées par des cohortes de scientifiques éminents, ne serait ce que parce qu'il paraît difficile d'imaginer des mécanismes de construction de processus visuels de manière déconnectée des problématiques que ces processus visuels sous-tendent. De plus, nous verrons par la suite qu'il n'y a pas de rupture franche entre un aspect "élaboration de systèmes de vision" et un aspect "fonctionnement de ces systèmes" dès lors que l'on pose les questions en termes d'apprenabilité.

Nous émettons l'hypothèse selon laquelle l'adaptativité est le point clé de la construction, du fonctionnement, et de la viabilité des systèmes de vision naturels, et qu'une grande partie de cette adaptativité peut s'exprimer à travers la modulation des interactions et des critères de qualité des indices visuels, puisque ceux-ci conditionnent profondément le déroulement et le résultat des processus d'analyse d'image et de perception. Il s'agit donc d'élaborer un système procédant à leur manière grâce à des méthodes d'apprentissage automatique permettant de définir et préciser ces éléments.

4. Indices visuels

Les indices visuels dits primaires constituent les éléments de base de l'image, il s'agit traditionnellement des régions (zones bidimensionnelles) et de leurs contours. Ce sont des entités fortement cohérentes. Pour une région, la cohérence peut tenir à la similitude entre les luminances ou les couleurs de ses

points; elle peut aussi tenir à des éléments dont l'estimation n'est pas ponctuelle: c'est le cas par exemple des textures. Les contours constituent des limites de régions, des caractéristiques importantes en sont la forme et la discontinuité qu'ils représentent par rapport aux régions associées. Par ailleurs, les indices visuels influent les uns sur les autres quant à la perception que nous en avons. C'est le cas notamment pour la détectabilité, mais aussi moins évidemment pour des éléments comme la localisation, la taille ou encore la couleur.

L'émergence d'éléments étendus ou à des niveaux d'abstraction croissants passe par des regroupements. Outre l'affaiblissement de contraintes d'homogénéité, un facteur intervenant dans ce mécanisme tient au fait que des éléments doivent d'autant plus être groupés que leur réunion produit quelque chose de plus significatif que leur prise en considération isolément: cela peut être lié à la forme qui résulte d'un assemblage et qui n'apparaît que de façon trop partielle dans les morceaux du puzzle pour y être déjà significative; ce peut être aussi lié à un facteur (de la forme ou de la fonction image) qui n'est pas son allure globale mais une caractéristique visible en considérant une autre échelle, une représentation particulière, ou encore un autre niveau d'abstraction.

Le problème est que la définition de mesures d'appréciation de la cohérence des indices visuels n'est pas triviale. Et il est bien évident que pour obtenir des groupements convenables, il faut pouvoir estimer si l'un doit être considéré comme meilleur qu'un autre. Les critères gestaltistes donnent des indications permettant par exemple d'établir un certain nombre de relations d'ordre partiel entre contours (en partie transposables aux régions), mais un ordre total n'en découle pas directement (d'autant plus qu'il faut prendre en compte simultanément formes, couleurs, contrastes, etc.). Concevoir a priori une fonction ordonnant totalement des ensembles d'indices visuels à partir des relations partielles relèverait de l'arbitraire. Pourtant, seul Lowe [10] et ses continuateurs [12] ont une approche systématique de la qualité des indices.

En outre, il existe un problème qui a tendance à passer inaperçu: les divers éléments qu'on assimile couramment à des facteurs de renforcement de la qualité des indices visuels ne peuvent être considérés comme tels indépendamment les uns des autres. De même, des caractéristiques d'indices habituellement sensées inhiber leur perception peuvent au contraire la faciliter. Soient par exemple deux régions texturées qui se touchent dans une image, suivant les types de textures impliquées, il se peut que la continuité de leur limite commune, ou son aspect lisse, ou bien encore son aspect granuleux soit le signe le plus révélateur de la forme du contour en présence duquel on se trouve (ceci étant redevable de la mémorisation de propriétés de l'univers auquel on est confronté). Un intérêt parmi d'autres d'un tel phénomène peut tenir au fait qu'un prédateur est caractérisé par une certaine texture et une certaine morphologie, et que le percevoir est plus aisé – et plus réaliste – avec telle ou telle prépondérance (celui-ci ne se détachant pas de façon criante). Ainsi, sa forme est mieux esquissée, sa détection et sa reconnaissance facilitées. Cette caractéristique peut avoir été apprise en termes d'évolution des espèces, et aussi en termes d'adaptation: prédateur précédemment vu par l'individu, en mouvement par rapport au fond (poursuite visuelle active ou passive par la proie potentielle).

Par ailleurs, les configurations d'indices visuels ne font pas qu'induire une qualité, elles peuvent également donner lieu à des interactions influençant des éléments d'autres natures. Il peut s'agir de phénomènes simples comme certaines modifications de la luminance apparente des zones d'une image, ou de cas nettement moins évidents du point de vue de l'explication physiologique ou même de la définition proprement dite. Par ailleurs, cela peut relever de ce que l'on appelle couramment les illusions, mais aussi de considérations indépendantes de la notion d'authenticité (par exemple: intérêt, priorité). De nombreuses recherches en psychophysiologie visent à enrichir la connaissance des mécanismes régissant la perception et la cognition, et les interactions entre indices visuels sont fréquemment considérées comme des indications sur les mécanismes de la vision. Une telle approche nous paraît tout à fait légitime, mais par ailleurs le fait que certaines de ces interactions constituent en elles-mêmes un élément utile en soulignant ou minimisant certains aspects de l'image nous semble primordial: cela justifie qu'on puisse vouloir les imiter, ce qui en impose un examen suffisamment approfondi que nous ne détaillerons pas ici.

5. Évolution et indices : nature et artifice

Nous sommes parti de l'idée selon laquelle l'apprentissage constituait un élément essentiel dans la problématique de la vision artificielle, or force est de constater qu'aucune des principales théories de ce domaine ne lui accorde la place qui lui revient, la première tentative étant à mettre au crédit de Lev Goldfarb [6]. Pourtant – comme le remarque Goldfarb – des physiologistes illustres insistent depuis longtemps sur son rôle central dans l'ensemble des processus sensoriels. Nous venons de présenter notre façon de considérer la question des indices visuels – entités porteuses d'éléments de la structure de l'univers perceptif – et de leurs interactions, parce qu'elle est au cœur du processus de segmentation, que nous voulons construire de tels processus à l'aide de méthodes d'apprentissage automatique et qu'une façon de procéder consiste naturellement à jouer sur ces éléments.

Le règne animal met en œuvre divers moyens pour réaliser l'adaptation des dispositifs visuels à l'environnement, ou pour être plus précis aux rapports entretenus avec l'environnement (il faut garder à l'esprit les impératifs de base : alimentation, préservation, reproduction, et communication). La nature joue sur l'organe de réception (formes et usages très divers), sur l'étendue du spectre visible, sur la définition et les réglages des processus de coordination et de traitement de l'information. Le représentant d'une espèce hérite de diverses caractéristiques, mais la perception de l'individu doit pouvoir s'accommoder d'environnements n'étant pas identiques pour chacun, et pouvant changer au cours d'une vie. Certains éléments sont donc plus plastiques que ceux cités précédemment, et se modifient au cours du développement et après (divers organes, au premier rang desquels le cerveau).

Du point de vue physiologique, des faits ont été mis en évidence qui indiquent clairement que l'établissement de processus perceptifs est subordonné au recours à certains de leurs homologues. Des éléments théoriques établissent le bien fondé de telles considérations, Aloïmonos a par exemple établi la validité de l'idée intuitive qui veut que le mouvement soit utile, en montrant que la vision dynamique représente un problème bien posé dans de nombreux cas dans lesquels ce n'est pas vrai pour la vision statique [1]. Il faut généraliser le principe d'une participation essentielle de processus perceptifs à la construction d'autres processus perceptifs. Il s'agit là d'une forme de coopération opportuniste, mais l'habituelle acception étriquée de cette expression occulte malheureusement cet aspect, ainsi que la question des indices visuels n'ayant d'existence que les uns par rapport aux autres. Or une telle coopération est un élément incontournable dans les phases de développement et d'adaptation d'un individu : lorsque c'est facile pour une modalité, les autres disposent alors d'une référence, d'un superviseur, et s'il y a matière à apprendre elles le font. Le système se donne ainsi les moyens d'interpréter son environnement dans des cas pour lesquels il ne disposerait pas des types d'informations immédiatement révélateurs. C'est de cette manière qu'il faut adapter les systèmes artificiels, en généralisant à l'ensemble des tâches et sous-tâches visuelles, en ne se contentant pas de principes de coopération déductifs, mais en mettant en avant un principe inductif de supervision opportuniste.

On peut concevoir les systèmes de vision comme un continuum d'apprentissages impliquant divers types de processus à différentes échelles temporelles pour différentes fonctions. L'évolution rend compte d'un environnement au travers de la mémoire de l'espèce, collectivement, génétiquement. L'adaptation fait de même pour l'individu, autorisant une certaine souplesse, affinant les grandes lignes déterminées par l'évolution. Et la répartition entre les mécanismes de l'évolution et ceux de l'adaptation permet d'optimiser la prise en compte de l'environnement. La découverte, enfin, entre en jeu pendant la tâche perceptive, elle correspond à des mécanismes recherchant des configurations présentant des propriétés intéressantes. On a parallèlement une continuité de buts qui se décomposent et interagissent, et qui s'étendent de l'intégration de caractéristiques quasi immuables de l'environnement à la détermination d'un élément de texture qui précédera de quelques millisecondes l'émergence d'une forme, en passant par le repérage d'une proie. Ces multiples buts et sous-butts diffèrent par la durée, le niveau d'abstraction, leur caractère conscient ou non, leurs interactions, etc. Ce positionnement en termes de buts (qu'on peut rapprocher de l'idée de contexte) va de pair avec une notion récemment considérée en intelligence artificielle : l'apprentissage actif.

En vue de produire des répliques artificielles des principes de ces mécanismes, de quel arsenal disposons-nous? Des modèles artificiels existent, mais s'inspirer d'un homologue naturel ne suffit pas à garantir l'adéquation aux problèmes que celui-ci traite avec succès, ce qui est somme toute normal: les équivalents artificiels mettent à profit des propriétés utiles de leurs modèles réels mais n'en sont que des copies très partielles (il n'est qu'à penser aux réseaux de neurones artificiels). Aussi ne peut-on se contenter d'envisager les appariements phénomène naturel - mécanisme artificiel correspondant, et ne doit-on écarter aucune possibilité.

6. Apprentissage et complexité

De notre analyse, il ressort que la vision doit être considérée essentiellement – à travers la notion d'appropriation de l'environnement – comme un ensemble de processus d'apprentissage inductif, continu, actif, multiéchelle en espace et en temps, et basé sur un principe de supervision opportuniste. Encore faut-il pouvoir généraliser et mémoriser les éléments les plus significatifs. Pour favoriser l'émergence de tels éléments, certains se sont référés aux idées gestaltistes en voulant unifier ces dernières autour d'un principe de simplicité descriptive. Celui-ci a été attaqué, sous le prétexte que sa dépendance à un langage de description lui enlèverait toute valeur. Il est temps de récuser cette fable: le langage en question doit tout simplement être vu comme un support de l'apprentissage de l'univers auquel un organisme est confronté, un ordre favorisant des émergences judicieuses d'indices visuels pouvant être représenté au travers du tandem langage de description - principe de simplicité. Pour mesurer la complexité d'un élément, la théorie algorithmique de l'information fournit depuis le milieu des années soixante une notion satisfaisante: la complexité de Kolmogorov [8, 9]. Celle-ci correspond pour une entité à la taille du plus court programme pouvant la produire (formellement: sur une machine de Turing universelle). Il faut remarquer que les travaux qui lui ont donné naissance vont dans le sens de l'idée d'Occam qui établit un lien entre le caractère non arbitraire d'une description ou d'une théorie et sa brièveté, ce qui est une indication pour les manières de généraliser [13]. On dispose aussi d'une mesure de dissemblance entre entités qui est la longueur du plus court programme permettant de transformer l'une en l'autre (comparaison non symétrique, qui fournit également une distance symétrique par addition des mesures des transitions opposées). Il ne s'agit donc plus seulement d'apparier strictement, de transformer de manière sinon fruste du moins contestable, ou de déformer au sens physique, mais de transformer et comparer à un sens informationnel à travers lequel les concepts de corrélation et de similitude sont plus riches qu'à l'accoutumée (dans le domaine de la vision artificielle); une notion calculatoire précise de cause commune peut donc être invoquée pour caractériser des régions homogènes, des contours continus, ou des caractéristiques de la structure d'une image, en somme autant de précurseurs sémantiques. Par rapport à la recherche d'une certaine universalité dans les représentations, il nous faut mentionner Lev Goldfarb qui – bien qu'il n'utilise pas la complexité de Kolmogorov – a récemment envisagé des dispositifs aussi puissants que la machine de Turing dans ce qui est la première théorie (algorithmique) inductive de la vision.

Faisons donc un sort à une seconde fable: les discontinuités ne constituent pas l'information essentielle de l'image, celle-ci est plutôt à rechercher dans les similitudes et les corrélations. Même dans les cas – qu'on dit fréquents – où une image de contours est pertinemment évocatrice, c'est à condition que ceux-ci soient corrects, c'est-à-dire qu'en tant que rupture de modèle ils puissent se référer à une expression appropriée de ce dernier, et qu'en tant que pourvus d'une forme ils soient "réguliers". Ainsi, la question centrale est celle de l'apprenabilité de la fonction image et de la forme des contours. En plus de la représentation, il faut considérer la question de l'approximation, qui réalise implicitement une généralisation. La perception ne comprime pas qu'en éliminant des redondances, elle le fait aussi en supprimant les informations non significatives ou inutiles. Des travaux récents prennent en partie en compte cet aspect des choses (indépendamment de la perception), ainsi une notion baptisée information totale a-t-elle été introduite par M. Gell-Mann [4] elle fait intervenir – en plus de la complexité de Kolmogorov – un terme d'entropie lié à la composante aléatoire de l'entité considérée. Par ailleurs, puisque des équations (notamment différentielles) peuvent rendre compte de phénomènes de manière concise, il nous a semblé intéressant d'introduire la notion de complexité totale qui est une extension de la complexité de Kolmogorov aux descriptions continues, équationnelles, différentielles:

c'est la taille de la plus courte description recourant à la fois aux algorithmes et aux équations. Elle est motivée – indépendamment de la question de la calculabilité – par la recherche de puissance expressive, par la nécessité d'avoir une idée de ce qui peut se concevoir.

7. La segmentation : modélisation, buts, propositions

La segmentation est l'étape fournissant un découpage de l'image. Ce découpage représente une mise en forme de l'information destinée aux étapes des niveaux supérieurs : reconstruction, interprétation et reconnaissance. Outre la mise en forme, il s'agit de fournir des régions bien utilisables par ces étapes.

Approche classique

Modélisation de la segmentation selon Zucker [15]: soit un prédicat P portant sur les régions, une segmentation S est un partitionnement de l'image en n régions tel que

- 1) $S = \{ R_1, R_2, \dots, R_n \}$
- 2) R_i est une région connexe quel que soit i
- 3) Chaque pixel appartient à une région unique
- 4) $P(R_i) = \text{vrai}$ quel que soit R_i
- 5) $P(R_i \cup R_j) = \text{faux}$ quelles que soient les régions connexes R_i et R_j

Approche proposée

Voici à présent quelques propositions concernant l'étape de segmentation, à partir d'une désignation de ses buts plutôt que d'une formalisation contraignante à laquelle nous substituons une nouvelle définition.

But de la segmentation :

Fournir une description des régions approximativement correcte et de faible complexité totale, à l'aide d'un langage dont le choix réalise une prise en compte de l'environnement. La description concerne la restriction de la fonction image aux régions, et la forme de leurs contours. On abandonne la définition classique de la segmentation, celle-ci n'a plus à être une partition. On n'impose plus non plus l'exigence de connexité entre points d'une même zone.

Modélisation de la segmentation :

Soit un prédicat P portant sur les régions, une segmentation S est un recouvrement de l'image en n régions tel que

- 1) $S = \{ R_1, R_2, \dots, R_n \}$
- 2) R_i est une région connexe ou non quel que soit i
- 3) Chaque pixel appartient à une ou plusieurs régions
- 4) $P(R_i) = \text{vrai}$ quel que soit R_i
- 5) $P(R_i \cup R_j) = \text{faux}$ quelles que soient les régions R_i et R_j

Abandon de l'exigence de connexité : on pense aux occultations, aux interruptions, aux dégradations, etc. Une zone peut paraître cohérente bien que ne figurant pas en un seul tenant, l'homogénéité ne doit alors pas y être recherchée de la même manière en tous points. En pareil cas – avec les mécanismes classiques – il y aurait aussi une tendance à malencontreusement prendre en compte certaines limites comme s'il s'agissait de contours dont la forme peut être significative. Par ailleurs, le report du groupement de parties non connexes en aval du processus de segmentation tendrait à le subordonner trop fortement à des considérations relevant de hauts niveaux d'abstraction, alors qu'un tel phénomène relève assez largement de mécanismes pré-attentifs.

Appartenance possible à plusieurs régions : dans les cas d'occultations, il faut pouvoir rendre compte aussi bien des régions occultantes que des régions occultées. Par ailleurs, lorsqu'il y a des effets de transparence, la zone commune à plusieurs régions doit pouvoir être rattachée à chacune d'entre elles dans la mesure où certaines caractéristiques l'en rapprochent suffisamment ; cela est valable dans le cas statique comme dans le cas dynamique. Tout ceci accroît certains problèmes liés à une combinatoire déjà explosive, il faut donc en appeler à des facteurs d'accroissement de l'efficacité.

8. Mise en œuvre, applications, implications

Opter pour un certain nombre de choix destinés à réduire l'impact de biais malencontreux ne va pas sans une certaine contrepartie essentiellement calculatoire. Dans le domaine de l'analyse d'image, c'est peut-être le prix à payer dès lors qu'on envisage de s'attaquer à des questions pertinentes plutôt que de parfaire les réponses apportées à des questions qui ne sont pas les bonnes. Nous ne détaillerons pas ici des difficultés théoriques et techniques trop éloignées des préoccupations habituelles du géographe. Disons simplement que des moyens d'estimer la complexité de Kolmogorov existent et qu'une partie de la communauté scientifique s'attache à les améliorer, et que l'intervention des interactions évoquées dans l'intégration des images semble devoir être un atout permettant de faire face aux autres explosions combinatoires. Ajoutons qu'un élément essentiel, du point de vue du principe comme sur le plan pratique, tient à la mise à disposition de superviseurs, l'idée générale qui s'y rattache est fort simple : la diversité des modalités disponibles – permettant par exemple sur un mode déductif de faire de la fusion d'informations – doit être mise à contribution dans une perspective inductive en tant qu'éventail de candidats à la supervision d'apprentissages portant sur chacune des modalités pouvant en bénéficier.

D'un point de vue expérimental, on est d'une manière générale à la recherche de la fonction qui, à toute image, associe son interprétation. Comme on ne peut pas considérer ce qui se passe pour chaque image possible, on doit se contenter d'examiner le comportement de cette fonction sur une portion de son ensemble de définition. Il faut considérer quelque chose qui soit suffisamment représentatif tout en restant exploitable. Pour cela, nous mettons au point des tests recourant à des modifications d'images réelles, associés à une directivité et un recueil des résultats qui soient adaptés. En l'attente de campagnes suffisamment vastes basées sur cette méthodologie, il serait nécessaire de procéder à une relecture des matériaux collectés au cours d'un certain nombre d'expérimentations menées de par le monde. Par ailleurs, nous ne nous intéressons pas seulement à la prise d'un instantané dans le temps de systèmes de vision, mais nous nous préoccupons aussi de leur évolution. Celle-ci doit être étudiée et enrichir la manière de faire évoluer des entités artificielles caractérisées de la sorte, transformant ainsi des algorithmes de segmentation (basés sur les but, définition et facteur d'efficacité mentionnés plus haut), faisant intervenir des moyens de représentation suffisamment puissants tels que ceux qui ont été proposés dans ce texte.

Intéressons-nous à présent aux applications des méthodes développées en revenant sur quelques cas que nous avons évoqués précédemment. Dans la caractérisation des zones sensibles au moyen d'images satellitales, bon nombre d'aspects significatifs des éléments texturaux et structurels à considérer résistent aux investigations car beaucoup d'aspects ne peuvent être formalisés a priori, les exemples – s'ils ne sont pas rares – présentant une variabilité élevée. On ne voit dès lors guère comment l'essentiel des caractéristiques pertinentes pourrait émaner de méthodes reposant sur des biais trop marqués. A travers les concepts que nous manipulons, nous ne sommes plus tributaires des limitations qu'impliquent certaines approches de la texture, ni du problème de choix allant avec l'abondance des options dans les approches les moins restrictives, et nous pouvons nous accommoder de la coexistence de divers aspects texturaux, structurels, et aléatoires. En ce qui concerne le grand tétras, il se trouve que ce qui lui est nécessaire ne se limite pas à des ressources exprimées en termes de quantités : ce volatile s'en tient à un milieu doté d'une structure particulière – mais dont aucun document ne définit précisément l'architecture – qui est celle de certains types de vieilles futaies entrouvertes. Des tolérances peuvent avoir cours, mais lorsqu'on s'intéresse au coq de bruyère, on ne peut s'affranchir d'une exigence d'irrégularité du milieu, avec les problèmes de caractérisation de régularités que cela

implique. Il est bien évident que de tels cas de figure en appellent à des moyens de représentation et de généralisation puissants. De même, l'évaluation paysagère fait référence à une multitude de problèmes autour de la représentation et de ce qui doit être considéré comme porteur de sens. Soulignons au passage que les nouveaux champs de recherche ouverts par notre approche tant sur le plan conceptuel que méthodologique s'affirment comme un moyen d'enrichir l'approche paysagère de l'équipe bisontine, et confèrent un relief particulier à la notion de schéma systémique du paysage [2]. Avant de conclure ce rapide tour d'horizon des applications et implications de nos travaux, il faut attirer l'attention sur le fait que la dimension perceptive évolutionniste de la vision impose certaines précautions dans de nombreuses situations, dès lors qu'on invoque cette dernière pour faire face à des tâches qui constituent pour elle une extension notable: c'est par exemple le cas en télédétection, en analyse interactive exploratoire des données, et en cartographie. Un domaine extérieur à la géographie est à cet égard symptomatique: en imagerie, l'irruption dans la pratique médicale d'examen sinon nouveaux du moins radicalement différents par leur finesse a parfois entraîné des interventions finalement injustifiées... on avait restreint une partie de l'étape d'appropriation. Pour finir, disons que dans le cas de la géographie l'utilité de l'approfondissement de la question de ce que l'on veut pouvoir concevoir est notamment sensible à travers la notion d'objet géographique, de durabilité, et de relation spatio-temporelle.

9. Conclusion

Nous sommes parti de la volonté d'améliorer des techniques automatiques d'analyse d'images, à la fois sur des plans généraux et du point de vue de leur mise en œuvre dans le domaine de la géographie, à travers l'intégration de principes issus du vivant comme l'adaptation. D'emblée, il était clair que la tâche imposait bien plus qu'une déclinaison de méthodologies existantes relativement aux spécificités d'un domaine, du fait que notre propos incluait l'amélioration de ces méthodologies, mais aussi du fait que l'adaptation à un domaine ne se limite pas toujours à un calque. En outre, nous subodorions que vision et géographie pouvaient entretenir des rapports allant très au-delà de la partie émergée de l'iceberg que constitue l'image. Nous avons dû approfondir quelques points concernant les visions naturelle et artificielle, ainsi que des éléments afférents. En fin de compte, il apparaît que la vision ne se limite pas à l'instant de la vision, elle doit avant tout être considérée – à travers la notion d'appropriation de l'environnement – comme un ensemble de processus d'apprentissage inductif, continu, actif, multiéchelle en espace et en temps, et basé sur un principe de supervision opportuniste. Il faut remarquer qu'en tant que science de l'impermanent, de l'espace, et des systèmes, la géographie est en réalité elle-même essentiellement fondée sur ce paradigme.

En fait, on a là un mécanisme très général de construction de la connaissance n'impliquant pas de rupture dans ses façons d'intégrer l'information, bien que des discontinuités puissent être instaurées dans ses réalisations matérielles (qu'elles soient biologiques ou artificielles). Du point de vue de la vision, il prend le contre-pied du paradigme de Marr à divers égards, mais s'inscrit dans une forme de continuité: en vision artificielle, on a d'abord voulu réaliser des processus à partir de leur explicitation, on a ensuite essayé de reproduire des mécanismes de la perception dans le vivant (sous l'impulsion de Marr), on a encore déplacé le problème d'un cran en cherchant à s'inspirer des principes évolutifs (nous avons généralisé cette idée). Cette dernière approche procédait notamment d'une volonté de s'affranchir de certaines difficultés, et de légitimer les systèmes conçus. Sous l'angle du géographe, notre pensée fait écho à des points de vue qui ont été évoqués de manière informelle et indépendamment les uns des autres au cours du colloque ThéoQuant'99, autour de l'importante part inductive de la construction de la connaissance, autour de la continuité de cette construction, et aussi autour de la durabilité de l'intérêt et de la valeur transconceptuelle de diverses représentations. Au bout du compte, les rapports à l'environnement et la façon même de l'appréhender établissent plus qu'un parallèle entre géographie et perception, que la notion d'apprenabilité réunit dans une même quête du sens.

Références bibliographiques

- [1] ALOÏMONOS, I. WEISS, A. BANDOPADHAY, 1987. Active vision. *International Journal of Computer Vision*, 1(4):333—356, janvier 1987.
- [2] BROSSARD T. et WIEBER J.-C., 1984. Le paysage : trois définitions, un mode d'analyse et de cartographie. *L'espace géographique*, n° 1, pp. 5-12.
- [3] BRUNSWIK E., 1956. *Perception and Representative Design of Psychological Experiments*. Berkeley, University of California Press.
- [4] GELL-MANN M., LLOYD S., 1996. Information Measures, Effective Complexity and Total Information; *Complexity*, 2, 44-52.
- [5] GIBSON J. J., 1979. *The Ecological Approach to Visual Perception*. Houghton Mifflin, Boston.
- [6] GOLDFARB L., DESHPANDE S. S., BHAVSAR V. C., 1996. *Inductive Theory of Vision*. Presented at the Workshop: "What is Inductive Learning?", in conjunction with the AI '96 - Eleventh Biennial Conference on Artificial Intelligence, 21-24 May 1996, Toronto, Ontario, Canada.
- [7] GREGORY R. L., 1974. Perceptions as Hypotheses. In S. C. Brown (Ed.) *Philosophy of Psychology*. Macmillan, London.
- [8] KOLMOGOROV A. N., 1965. Three approaches to the quantitative definition of information. *Problems of Information Transmission* 1:1 — 17.
- [9] LI M., and P.M. VITANYI P.M., 1993. *An introduction to Kolmogorov Complexity and its Application*. Berlin: Springer — Verlag.
- [10] LOWE D. G., 1985. *Perceptual Organization and Visual Recognition*. Kluwer Academic publisher, Hingham, USA.
- [11] MARR D., 1982. *Vision*. W. H. Freeman and Company, San Francisco.
- [12] SARKAR S. et BOYER K. L., 1993. Perceptual organization in computer vision: A review and a proposal for a classificatory structure. *IEEE Transactions on Systems, Man, and Cybernetics*, 23(2):382—399.
- [13] SOLOMONOFF R.J., 1964. A formal theory of inductive inference. *Information and Control* 7:1 — 22,224 —54.
- [14] WERTHEIMER M., 1958. Untersuchungen zur lehre von der gestalt ii, translated as: "principles of perceptual organization". In *Readings in Perception*, pages 115 — 135. Princeton, N.J.
- [15] ZUCKER S., 1976. Region growing: Childhood and adolescence. *Computer Graphics and Image Processing*, 5:382-399.