

Analyse d'images par champs de Markov

Jian-Feng Yao^(*)

Résumé. Nous proposons une introduction aux méthodes mathématiques récentes destinées à l'analyse d'images numériques.

1. Introduction

Jusqu'aux années 1970, les images numériques sont restées une affaire de spécialistes : images radar pour l'aviation, images satellites pour la défense, images à rayon X pour les médecins, etc. À ces utilisations professionnelles se sont ajoutées celles du grand public introduites par les appareils de photo ou les caméras numériques. La capacité toujours croissante de l'informatique conduit à une aisance sans précédent de leur archivage et de leur manipulation.

Cependant lorsque l'on demande à l'ordinateur ou encore à une machine de « regarder » ces images, d'en extraire des « informations utiles », comme par exemple dessiner le contour d'une tumeur suspecte dans une image médicale, l'affaire est loin d'être entendue. Le problème est que l'humain que nous sommes est encore (toujours ?) infiniment plus intelligent qu'une machine pour lire et comprendre une image. On peut alors se poser la question de l'intérêt de demander à une machine de travailler à notre place et moins bien. Beaucoup de réponses à cela : c'est par exemple tout simplement que nous ne sommes pas capables d'être présents partout où au contraire une machine peut très bien se positionner. Il suffit pour cela de songer aux caméras de vidéo-surveillance sur le périphérique de Paris. Il est donc important qu'une machine acquière de l'intelligence que les spécialistes appellent « intelligence artificielle » pour la distinguer de la nôtre.

Le propos de cet exposé est d'illustrer ce que les mathématiques peuvent apporter d'intelligent à un ordinateur dans la compréhension d'une image numérique. Ce domaine d'application des mathématiques s'est vite développé depuis les années 1980. Nous verrons qu'il s'agit d'un domaine riche non seulement pour les applications visées, mais également pour le développement de nouvelles branches modernes des mathématiques.

Dans une première partie de l'exposé, nous présenterons la notion de l'analyse d'images, plus particulièrement la segmentation d'images. Dans une deuxième partie, nous expliciterons une famille de méthodes mathématiques comme solutions à ces problèmes. Ces méthodes feront intervenir les *champs de Markov* comme modèles de base, et nous indiquerons les algorithmes de résolution associés à ces modèles.

(*) Institut de Recherche MATHématique de Rennes Université de Rennes 1.

Les transparents de cet exposé sont disponibles à : <http://www.univ-reriries1.fr/yao>

2. Généralités : images numériques

2.1. Exemples

Nous présentons d'abord quelques exemples d'images numériques.

Exemple 1. Contrôle de qualité

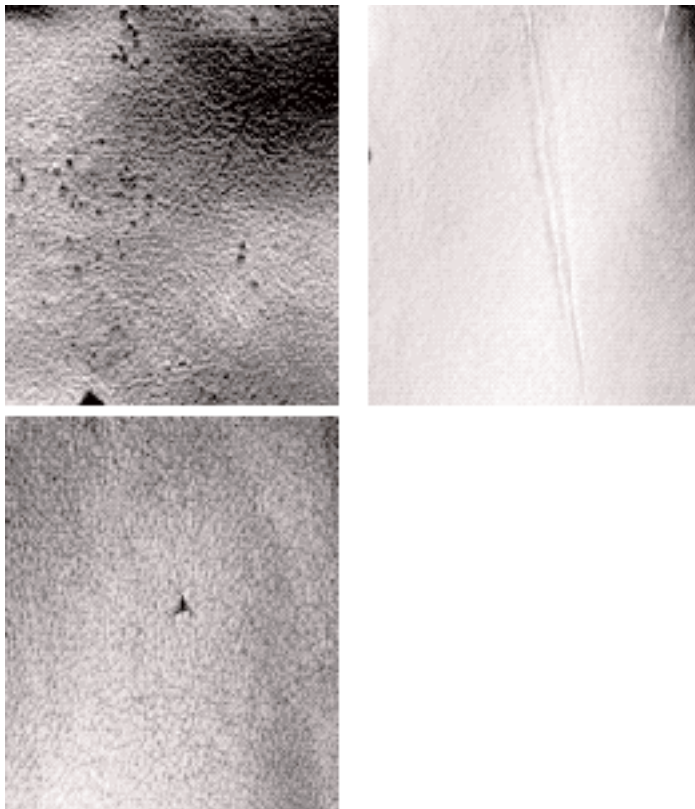


FIG. 1 – Exemple d'images en contrôle de qualité. Source: SUDIMAGE, R. Azencott et J. Yao.

La figure 1 présente des photographies d'échantillons de cuir imprimé, destiné au revêtement de sièges automobiles. Dans le processus de fabrication, l'impression d'un motif de décoration sur le cuir peut provoquer des *défauts* : piqûres, plis, éraflures. L'objectif est de localiser, et mesurer de façon automatisée les défauts afin de donner des indications pour la découpe des zones défectueuses.

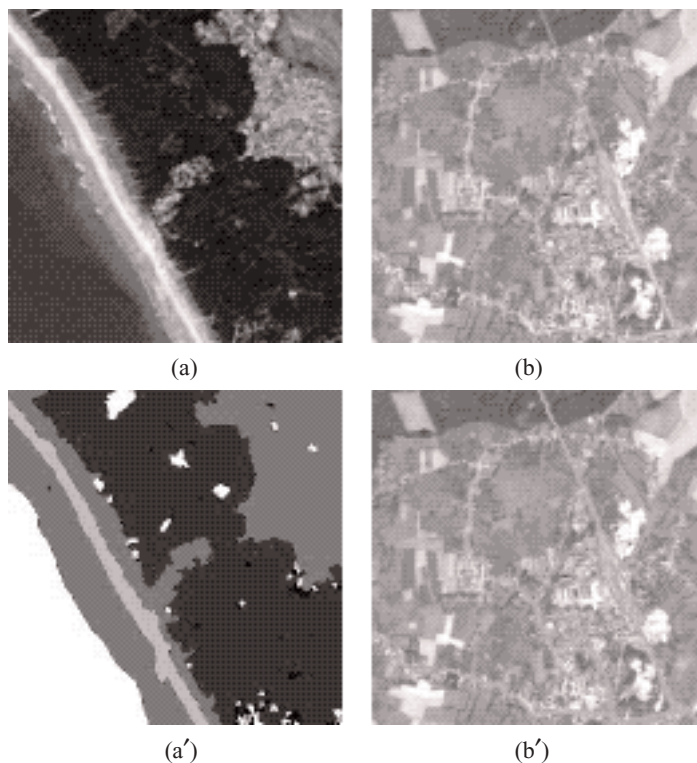
Exemple 2. Images satellite

FIG. 2 – Images satellite. Source : INRIA, Sophia-Antipolis, Équipe de J. ZERUBIA.

Les images (a)-(b) de la figure 2 proviennent du satellite SPOT. L'objectif est de repérer des régions homogènes, des structures linéaires, par exemples les routes sur ces images. Les images (a')-(b') montrent les résultats correspondants.

Exemple 3. Image IRMf

FIG. 3 – Image par résonance magnétique fonctionnelle (IRMf) du cerveau.
Source: INRIA, Sophia-Antipolis, Équipe de O. Faugeras.

Il s'agit d'une image par résonance magnétique fonctionnelle (IRMf). Après injection d'un liquide marqueur, son cheminement et sa concentration dans le cerveau donnent des informations précieuses sur la structure neuro-biologique du cerveau.

Exemple 4. Réflexion sismique

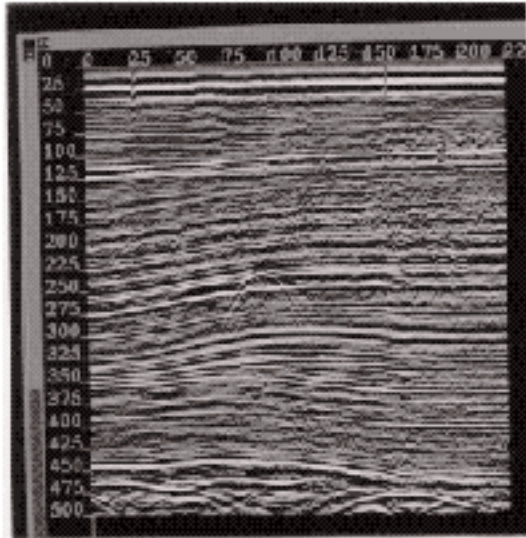


FIG. 4 – Source: SUDIMAGE, R. Azencott et J. Yao.

La figure 4 montre une image de réflexion sismique qui explore les couches géologiques du sous-sol. Le procédé consiste à envoyer une onde forte de vibration en un point du sol ; cette onde traverse le sous-sol, et une partie se réfléchit sur les différentes couches géologiques et finit par remonter à la surface. On installe à cet endroit des *géophones* qui enregistrent les ondes réfléchies. Celles-ci, par leur amplitude et leur instant de retour fournissent des informations sur la topographie du sous-sol. En général, c'est un moyen de prospection préliminaire à un moyen plus lourd comme le forage (on cherche du pétrole !).

2.2. Codage d'une image numérique

L'acquisition ou la numérisation d'une image procède d'abord par introduction d'une grille de points sur l'image. Cette grille est généralement rectangulaire, et nous noterons l'ensemble de ses points par $S = \{1, \dots, M\} \times \{1, \dots, N\}$. Plus la grille est fine, meilleure est la qualité de la numérisation (résolution). En chaque point $s = (i, j) \in S$ de l'image, on enregistre la *luminosité* $Y(s)$ en ce point. En nous restreignant aux images en noir et blanc, $Y(s)$ représente alors le niveau de gris en s : $Y(s)$ est de valeur entière, allant de 0 si le point est d'un noir complet, à 255 si au contraire il est du blanc saturé. Ainsi une *image numérique* est mathématiquement une application

$$S = \{1, \dots, M\} \times \{1, \dots, N\} \rightarrow \{0, \dots, 255\}$$

$$s \mapsto Y(s).$$

Pour une image en couleur, on mesure la luminosité en un point s selon les trois canaux rouge, vert et bleu. Elle est alors représentée par un tableau de vecteurs de dimension 3 : autrement dit, une application de S dans $\{0, \dots, 255\}^3$.

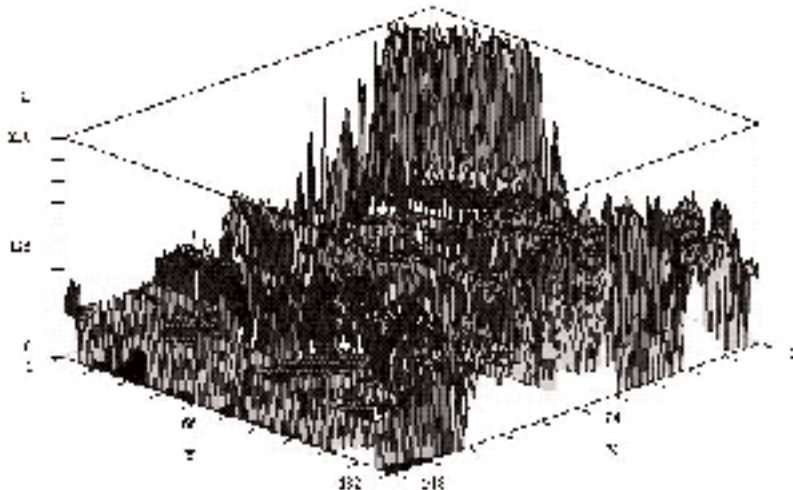


FIG. 5 – De haut en bas : photo d'une locomotive, le cliché de sa tête de 132×148 pixels et les valeurs en niveaux de gris du cliché.

On appelle *pixel*, contraction de l'anglais « picture element », un point s de la grille. Voici une illustration de ce codage. La figure 5 montre une locomotive à vapeur dont on a extrait ensuite le cliché de la tête. Celui-ci correspond à une grille de 132×148 pixels. Les valeurs des niveaux de gris $Y(s)$ sont également sur la même figure. Il n'est pas aisé de reconnaître la tête de la locomotive à partir de ce dernier graphique !

Le tableau suivant donne ces valeurs pour le coin nord-ouest de la tête, d'une taille de 30×15 :

217.	217.	217.	214.	217.	217.	217.	217.	217.	217.	217.	217.	217.	217.	217.	217.	217.	217.	217.	217.
217.	217.	217.	217.	217.	217.	217.	217.	217.	217.	217.	217.	217.	217.	229.	229.	229.	229.	229.	217.
217.	217.	217.	217.	217.	217.	217.	217.	217.	229.	217.	217.	229.	217.	229.	217.	229.	229.	229.	229.
217.	229.	229.	217.	217.	229.	217.	229.	217.	229.	217.	229.	217.	229.	229.	229.	229.	229.	229.	229.
229.	217.	229.	217.	229.	229.	217.	217.	229.	229.	229.	229.	217.	217.	229.	229.	229.	229.	229.	229.
229.	229.	229.	229.	229.	217.	229.	229.	229.	217.	217.	229.	229.	229.	217.	217.	229.	229.	229.	229.
229.	217.	229.	217.	217.	229.	229.	217.	229.	217.	229.	217.	229.	217.	229.	217.	229.	229.	229.	229.
229.	229.	234.	234.	234.	234.	234.	229.	217.	234.	229.	229.	229.	229.	229.	229.	234.	250.	234.	217.
234.	234.	229.	234.	234.	234.	234.	234.	234.	234.	234.	234.	234.	234.	250.	234.	217.	229.	229.	250.
229.	229.	229.	229.	229.	234.	229.	234.	229.	234.	250.	217.	195.	214.	229.					
229.	229.	234.	234.	234.	234.	229.	234.	234.	229.	234.	229.	234.	229.	189.	214.	234.			
229.	234.	229.	234.	229.	229.	229.	229.	234.	234.	234.	250.	250.	150.	153.	173.				
229.	229.	229.	229.	229.	234.	250.	229.	234.	234.	250.	250.	173.	164.	164.					
234.	234.	229.	217.	229.	229.	234.	217.	229.	179.	179.	189.	173.	153.	117.					
229.	229.	229.	250.	229.	229.	234.	229.	250.	150.	127.	102.	102.	102.	102.					
217.	229.	234.	234.	250.	250.	234.	250.	250.	189.	117.	51.	77.	102.	102.					
229.	250.	250.	195.	214.	250.	234.	234.	250.	250.	137.	77.	77.	102.	89.					
229.	234.	250.	173.	166.	250.	250.	250.	250.	250.	127.	89.	89.	77.	77.					
229.	250.	250.	217.	164.	217.	250.	250.	250.	189.	77.	102.	102.	77.	77.					
229.	234.	234.	250.	195.	150.	164.	189.	189.	127.	77.	77.	77.	77.	77.					
217.	214.	217.	214.	166.	137.	137.	204.	117.	77.	77.	77.	51.	77.	51.					
166.	150.	150.	166.	117.	102.	150.	217.	89.	51.	51.	51.	51.	51.	51.					
217.	102.	178.	204.	89.	77.	117.	175.	77.	51.	51.	51.	51.	51.	51.					
234.	77.	102.	89.	153.	77.	51.	0.	0.	51.	51.	51.	51.	51.	51.					
77.	117.	117.	150.	166.	0.	0.	0.	51.	51.	51.	51.	0.	1.	0.					
102.	117.	117.	102.	51.	51.	51.	77.	51.	77.	51.	51.	51.	51.	51.					
102.	117.	153.	173.	0.	51.	51.	51.	51.	51.	51.	51.	0.	51.	51.					
77.	89.	89.	77.	51.	51.	77.	77.	51.	51.	0.	51.	51.	51.	0.					
89.	77.	89.	51.	51.	77.	77.	51.	51.	51.	51.	0.	51.	51.	51.					
117.	175.	195.	51.	77.	77.	51.	51.	0.	51.	51.	51.	51.	51.	51.					

2.3. Analyse d'images

Analyse d'images est un terme générique englobant entre autres :

- *traitement d'images* :
 - dé-bruitage, suppression de flou, rehaussement de contraste.
- *localisation, extraction des éléments intéressants* :
 - événements exceptionnels (contrôle de qualité) ;
 - objet, personne mobile v.s. scène fixe (vidéo surveillance) ;
 - tumeur (imagerie médicale : radio, RMN, scanner, etc.).
- *détection de bords, segmentation* :
 - frontières physiques des différentes régions dans l'image ;
 - frontières plus abstraites, non nécessairement visibles.

Il s'agit là d'une liste non exhaustive. Il importe de la distinguer du terme *Synthèse d'images* qui correspond précisément à la création d'images, statiques ou

animées, d'objets comme des personnages, suivant un scénario précis. Exactement comme au cinéma mais avec des figures virtuelles à la place des acteurs ou d'objets. Pour de plus amples illustrations, nous renvoyons aux derniers films de Hollywood.

Revenons à la photo de la locomotive ci-dessus. Nous, les humains, reconnaissons instantanément la locomotive en voyant la photo ; mais l'ordinateur, ou plus précisément un algorithme de calcul, ne voit que le tableau des chiffres. C'est avec ce tableau qu'on lui demande de reconnaître la locomotive et son environnement ! On peut alors avoir une idée sur les grosses difficultés posées par l'analyse d'images, ou de façon plus générale par ce qu'on appelle la *vision par ordinateur*.

Cela n'empêche pas de fixer des objectifs à l'analyse d'images : imiter la perception humaine le mieux possible ; rendre des services en exécutant de façon automatique des tâches qui sont soit peu intelligentes et répétitives, soit difficilement accessibles aux humains.

Le but de cet exposé est de montrer, dans le cadre des tâches dites de *segmentation/détection*, que l'analyse d'images a pris un bon départ par l'introduction de modèles ou méthodologies mathématiques nouvelles. En particulier deux théories mathématiques ont émergé ou se sont amplifiées dans ce contexte :

1. *Théorie des champs de Markov* : cette théorie a connu un vrai succès en analyse d'images depuis les années 1980. Elle s'est développée à l'aide des disciplines comme les probabilités, la statistique ou encore la mécanique statistique.
2. *Théorie des ondelettes* : celle-ci est plus récente (depuis les années 1990) et basée sur l'analyse de Fourier, l'analyse fonctionnelle ou encore l'analyse numérique.

Nous nous limiterons à la première théorie. Citons deux références très complètes sur le sujet :

1. Bernard CHALMOND, 2000. *Éléments de modélisation pour l'analyse d'images*, Springer-Verlag.
2. Donald GEMAN, 1990. *Random Fields and Inverse Problems in Imaging*, Lectures Notes in Mathematics, N° 1 427, Springer-Verlag.

3. Problème de segmentation d'images

3.1. Formulation du problème

Segmenter une image Y revient à répartir l'ensemble des pixels S en *régions homogènes* :

$$S = S_1 \cup S_2 \cup \dots \cup S_K$$

comme schématisé sur la figure 6.

On introduit la carte des étiquettes $(X_s, s \in S)$ pour représenter une partition :

$$\text{pixel } s \in S_j \Leftrightarrow X_s = j.$$

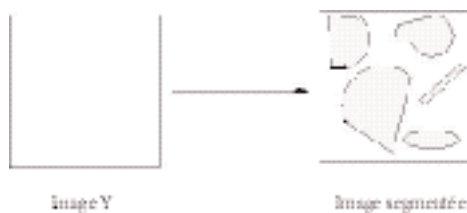


FIG. 6 – Segmentation d'une image.

Citons quelques exemples :

1. détection des défauts de cuir : $K = 2$ étiquettes.
 $X_s = 1 \Leftrightarrow s \in S_1 \Leftrightarrow$ pixel s fait partie d'un défaut.
2. détection d'une tumeur dans une image médicale :
tumeur/non tumeur.
3. vidéo-surveillance :
objets, personnes mobiles/immobiles.
4. détection des bords dans une image : bords/intérieurs.
5. images satellites : $K = 5$ pour {zones urbaines, champs, forêts, routes, mer}.

L'approche par modélisation probabiliste du problème de segmentation consiste à :

- considérer l'image $Y = (Y_s)$, la carte des étiquettes $X = (X_s)$ (à construire) comme des *variables aléatoires* régies par une loi statistique π ;
- proposer une modélisation \equiv définir une telle loi π ;
- X et Y étant liées par la loi π , Y étant donnée, reconstruire ou estimer X à l'aide de π et de Y .

On peut remarquer que si la loi de la formation d'image F :

$$X = (X_s) \mapsto Y = (Y_s) = F(X)$$

était totalement connue, il suffirait d'inverser F ! En réalité, une telle fonction déterministe F est irréaliste, car le mécanisme de la formation d'images est pour le moins complexe d'une part, et entaché du bruit, c'est-à-dire des aléas ou des erreurs de manipulations qui interviennent. L'approche par modèles probabilistes définit les passages $X \mapsto Y$ et $Y \mapsto X$ par des lois statistiques conditionnelles. Et les champs de Markov sont des exemples de telles lois parmi les plus utilisées.

4. Champs de Gibbs, ou encore de Markov pour les étiquettes

Les champs que nous introduisons ici ont été introduits par des physiciens et, pour cette raison, nous altérons légèrement notre langage et appelons un *site* ce que nous avons appelé un pixel jusqu'ici.

On se donne un ensemble de sites $S = [1, m] \times [1, n]$, et une relation de « voisinage » sur S qui est symétrique. On utilise très souvent la relation des 4 plus proches voisins, ou celle des 8 plus proches voisins, comme décrits sur la figure 7. Dans le premier cas par exemple, chaque site a comme voisins les 4 sites N-O, N-E, S-E et S-O. Deux sites s et t voisins seront notés $s \sim t$.

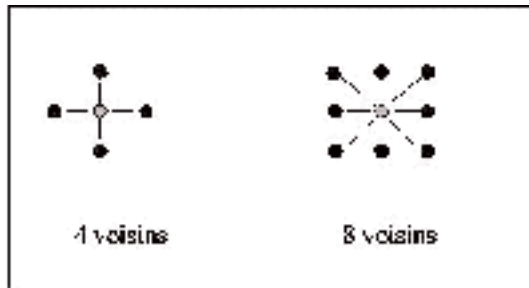


FIG. 7 – Relation de voisinage aux 4 ou 8 plus proches voisins.

Une fois fixée la relation de voisinage, on construit un graphe (non orienté) de voisinage (S, \mathcal{G}) en reliant simplement deux sites s et t s'ils sont voisins. On appelle une clique tout sous-ensemble A de sites qui sont *mutuellement voisins*. Les exemples de cliques sont donnés sur la figure 8.

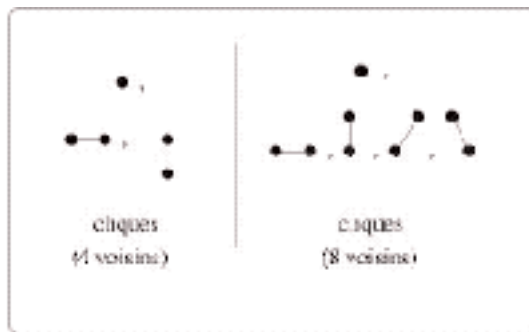


FIG. 8 – Exemples de cliques pour les relations des 4 ou 8 plus proches voisins.

Soit E l'ensemble des étiquettes possibles : $E = \{e_1, \dots, e_K\}$. On appelle l'espace des configurations: $\Omega = E \times \dots \times E = E^S$ et une configuration est notée $x = (x_s)_{s \in S}$ avec $x_s \in E$ la valeur de l'étiquette au site s . Si A est un sous-ensemble de S , la trace de x sur A est $x|_A = (x_s)_{s \in A}$.

Un champ de Gibbs sur (S, \mathcal{G}) est la donnée d'une distribution statistique π sur Ω de la forme :

$$\pi(x) \propto \exp[-H(x)] = \frac{1}{Z} \exp[-H(x)], \quad Z > 0, \quad (1)$$

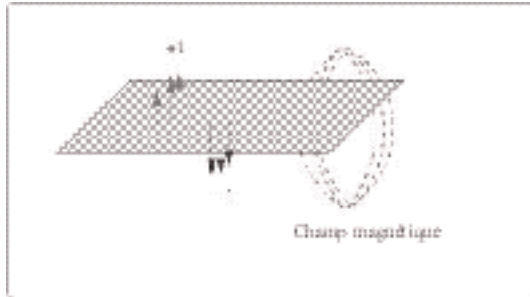
où $H(x)$ est la fonction d'énergie:

$$H(x) = \sum_{\text{cliques } A} J_A(x).$$

Ici J_A est le potentiel de la clique A qui ne dépend que de la trace de x sur A . Autrement dit,

$$J_A(x) = J_A(x') \quad \text{si} \quad x|_A = x'|_A.$$

En résumé, un champ de Gibbs est déterminé par l'ensemble des potentiels $\{J_A\}$ indexés par les cliques A .

Exemple 1 : modèle ferro-magnétique d'Ising

- $x_s = +1$ ou -1 : orientation d'un spin à l'état d'équilibre,
- relation des 4 plus proches voisins,
- fonction d'énergie :

$$H(x) = -h \sum_s x_s - \beta \sum_{s-t} x_s x_t.$$

- $h > 0$: force du champ magnétique extérieur ;
- $\beta > 0$: mesure d'interaction entre les spins :

$$x_s x_t = \begin{cases} +1 & \text{si } (x_s, x_t) = (\uparrow, \uparrow), (\downarrow, \downarrow) \\ -1 & \text{si } (x_s, x_t) = (\uparrow, \downarrow), (\downarrow, \uparrow) \end{cases} \quad \begin{matrix} H(x) \searrow \\ H(x) \nearrow \end{matrix}$$

Les spins voisins ont tendance à orienter de façon identique : on dit que le champ est attractif.

Ce modèle est fréquemment utilisé en analyse d'images. Deux situations peuvent se présenter. Soit les directions horizontales et verticales ne sont pas distinguées, on a alors un modèle isotropique avec

$$H(x) = -\beta \sum_{s-t} x_s x_t.$$

Dans le cas contraire, on a un modèle anisotropique avec

$$H(x) = -\beta_1 \sum_{\substack{s-t \\ \text{horiz}}} x_s x_t - \beta_2 \sum_{\substack{s-t \\ \text{vert}}} x_s x_t.$$

Il est aisé de généraliser ce modèle au voisinage des 8 plus proches voisins.

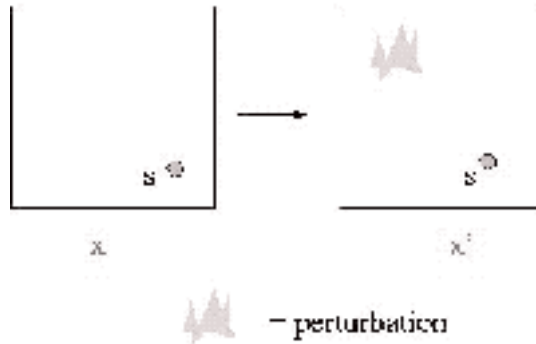
Exemple 2. Modèle de Potts

Il s'agit d'une généralisation du modèle précédent avec $K > 2$ étiquettes possibles au lieu de deux, soit $E = \{e_1, e_2, \dots, e_K\}$ et l'énergie s'écrit :

$$H(x) = -\beta_1 \sum_{\substack{s-t \\ \text{horiz}}} I_{x_s = x_t} - \beta_2 \sum_{\substack{s-t \\ \text{vert}}} I_{x_s = x_t}.$$

4.1. Caractéristiques locales du champ

Une propriété fondamentale d'un champ de Gibbs est que les interactions entre sites sont locales, i.e. limitées à celles entre les voisins. De façon heuristique, cela exprime le fait que si une perturbation se produit, elle a peu d'impact sur les sites éloignés de la perturbation (cf. la figure 9).



$x_s \cong x'_s$ pour s loin de la perturbation
FIG. 9 - Illustration d'une interaction locale.

Nous allons formuler précisément cette propriété. Notons la configuration hors d'un site s , ou encore son extérieur par $x^{(s)} = (x_t, t \neq s)$. Notons la loi conditionnelle au site s , étant donné son extérieur :

$$p_s(e | x^{(s)}) = P(x_s = e | x_t, t \neq s).$$

En partant de le définition de H, il est aisé (voir ci-dessous) de montrer que :

$$p_s(e | x^{(s)}) = \exp \left[- \sum_{A: A \ni s} J_A(x_s = e, x^{(s)}) \right]. \tag{2}$$

On voit que seules les cliques contenant s , c'est-à-dire les voisins de s contribuent à cette loi conditionnelle.

On appelle les *caractéristiques locales du champ* la collection de ces lois conditionnelles :

$$\mathcal{D} = \{p_s(\cdot | x^{(s)}) : s \in S\}.$$

Calcul de la caractéristique locale $p_s(e | x^{(s)})$:

$$\begin{aligned} p_s(e | x^{(s)}) &= P(x_s = e | x_t, t \neq s) \\ &= \text{loi de probabilité sur E} \\ &= \frac{P(x_s = e, x^{(s)})}{P(x^{(s)})} \\ &= \frac{P(x_s = e, x^{(s)})}{\sum_j P(x_s = e_j, x^{(s)})} \\ &= \frac{\exp[-H(x_s = e, x^{(s)})]}{\sum_j \exp[-H(x_s = e_j, x^{(s)})]} \end{aligned}$$

mais:

$$\begin{aligned} H(x_s = e, x^{(s)}) &= \sum_A J_A(x_s = e, x^{(s)}) \\ &\quad \updownarrow \text{ si } x \notin A \\ H(x_s = e_j, x^{(s)}) &= \sum_A J_A(x_s = e_j, x^{(s)}) \end{aligned}$$

D'où, par simplification

$$\begin{aligned} p_s(e | x^{(s)}) &= \frac{\exp\left[-\sum_{A:A \ni s} J_A(x_s = e, x^{(s)})\right]}{\sum_j \exp\left[-\sum_{A:A \ni s} J_A(x_s = e_j, x^{(s)})\right]} \\ &\quad \times \exp\left[-\sum_{A:A \ni s} J_A(e, x^{(s)})\right] \end{aligned}$$

Exemple: modèle d'Ising aux 4 p.p. voisins

$$H(x) = -h \sum_s x_s - \beta \sum_{s-t} x_s x_t.$$

$$\{A: A \ni s\} = \{\{s\}, \{(s,t), t = t_N, t_E, t_S, t_O\}\}.$$

$$\begin{aligned} p_s(e | x^{(s)}) &\times \exp[hx_s + \beta(x_s x_{t_N} + x_s x_{t_E} + x_s x_{t_S} + x_s x_{t_O})] \\ &= \exp[x_s(h + \beta v_s)], \quad v_s = x_{t_N} + x_{t_E} + x_{t_S} + x_{t_O}. \end{aligned}$$

D'où, avec $x_s = e = \pm 1$,

$$p_s(x_s = +1 | x^{(s)}) = \frac{e^{h+\beta v_s}}{e^{-(h+\beta v_s)} + e^{h+\beta v_s}}.$$

$$p_s(x_s = -1 | x^{(s)}) = \frac{e^{-(h+\beta v_s)}}{e^{-(h+\beta v_s)} + e^{h+\beta v_s}}.$$

4.2. Importance des caractéristiques locales

On peut montrer que la famille des caractéristiques locales détermine la loi de Gibbs π (π détermine \mathcal{D} , ce qui est plus évident). Par ailleurs \mathcal{D} est beaucoup plus facile à calculer que π : en effet, jusqu'ici nous n'avons jamais précisé comment faire pour évaluer la constante de normalisation Z qui apparaît dans la définition de la loi π , équation (1), ce qui est en fait difficile.

On peut aussi relier \mathcal{D} à la propriété de Markov. Commençons par rappeler celle-ci en dimension 1. On a une suite de variables aléatoires $W_1, W_2, \dots, W_s, \dots$, que l'on appelle une chaîne de Markov, car elle satisfait la propriété :

$$P(W_s \in \cdot | W_{s-1}, W_{s-2}, \dots, W_1) = P(W_s \in \cdot | W_{s-1})$$

qui dit que la loi conditionnelle de la valeur d'aujourd'hui, étant données toutes les valeurs du passé, ne dépend que de la valeur d'hier. Cette propriété dite de Markov est illustrée sur la figure 10. On peut aussi donner une version symétrique de cette propriété :

$$P(W_s \in \cdot | W_t, t \neq s) = P(W_s \in \cdot | W_{s-1}, W_{s+1}) \quad (3)$$

qui dit que la loi conditionnelle de la valeur en un instant s , étant données toutes les autres valeurs, ne dépend que des valeurs des deux instants voisins $s - 1$ et $s + 1$.

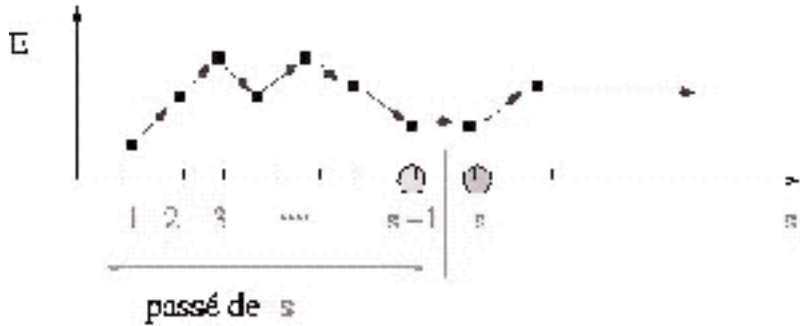


FIG. 10 – Une chaîne de Markov.

Revenons à un champ de Gibbs et on constate que la formule de la caractéristique locale en un site s , équation (2), est une généralisation en dimension 2 de la propriété de Markov (3) comme l'illustre la figure 11. C'est pour cette raison qu'un champ de Gibbs est aussi appelé un champ de Markov.

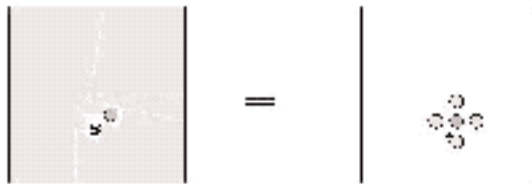


FIG. 11 – Propriété de Markov en dimension 2 : le monde extérieur à un site s n'agit sur lui que par ses plus proches voisins.

5. Modèle probabiliste de segmentation

Une modélisation probabiliste pour le problème de segmentation consiste en les étapes suivantes :

- A. définir un champ de Gibbs, π , pour les étiquettes x ;
- B. construire un modèle, de « formation d'images » :

$$x = (x_s) \mapsto y = (y_s).$$

Cette étape est guidée par la question : *si les étiquettes sont x , que devrait être l'image y ?* \Rightarrow loi statistique de y étant donné x : $q(y|x)$.

Des étapes A et B, on obtient un modèle conjoint pour le couple (x,y) :

$$p(x, y) = \pi(x)q(y|x).$$

- C. Principe statistique de vraisemblance : il s'agit d'une méthode d'estimation générale introduite par le statisticien anglais R. Fisher dans les années 1930 :

Étant donnée l'image observée y , estimer les étiquettes par la configuration \bar{x} qui maximise la distribution conditionnelle des étiquettes, loi a posteriori :

$$\begin{aligned}
 r(x|y) &= \text{probabilité d'un } x \text{ étant donné } y \\
 &= \frac{p(x,y)}{u(y)}, \quad (\text{Formule de Bayes}). \\
 \bar{x} &= \operatorname{argmax}_{x \in \Omega} r(x|y) \\
 &= \operatorname{argmax}_{x \in \Omega} p(x,y)
 \end{aligned}$$

L'étape B de $x \mapsto y$ est en général difficile. Elle nécessite une connaissance approfondie de la loi (au sens physique) de formation de l'image à analyser : compréhension de l'optique, de la physique sous-jacente, etc. C'est dans cette étape notamment que la vision par ordinateur trouve toutes ses limites.

L'étape C est une étape purement mathématique et calculatoire : on sait proposer des solutions satisfaisantes.

Note : la loi *a posteriori* $r(x|y)$ est encore un champ de Gibbs. En effet,

$$\begin{aligned}
 r(x|y) &= \frac{1}{u(y)} p(x,y) \\
 &= \frac{1}{u(y)} \pi(x) q(x|y) \\
 &= \frac{1}{u(y)} \frac{1}{Z} \exp[-H(x)] \prod_{s \in S} f(y_s | x_s) \\
 &= \frac{1}{u(y)Z} \exp \left[-H(x) + \sum_{s \in S} \log f(y_s | x_s) \right] \\
 &\propto \frac{1}{Z'(y)} \exp[-H'(x)]
 \end{aligned}$$

Ce champ possède le même graphe de voisinages que la loi (*a priori*) π des étiquettes.

Détection de défauts de cuir : une illustration

Étape A : étiquettes binaires,

$$x_s = \begin{cases} -1, & \text{pixel } s \text{ appartient à un défaut,} \\ +1, & \text{sinon} \end{cases}$$

Étape B : modèle d'Ising aux 4 ou 8 plus proches voisins,

– Forme produit pour la distribution de la formation d'image :

$$q(y|x) = \prod_{s \in S} f(y_s | x_s).$$

– f est une densité du type « gaussienne », choisie après expérimentation empirique,

- si $x_s = +1$, pixel non défectueux,

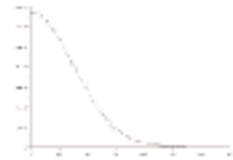
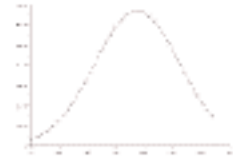
$$f(y_s | +1) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(y_s - m)^2\right\}$$

- si $x_s = -1$, pixel défectueux, piqûres, plis, etc., régions généralement sombres, niveau de gris concentrés sur $[0, c]$, $c < 255$,

$$f(y_s | -1) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}y_s^2\right\}$$

Finalement, la loi de formation d'images est :

$$q(y|x) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^{|S|} \times \exp\left\{-\sum_{s \in S} \frac{x_s}{4\sigma^2} [m^2 - 2my_s + 2\sigma^2 \log 2] - \sum_{s \in S} \frac{1}{4\sigma^2} [y_s^2 + (y_s - m)^2 - 2\sigma^2 \log 2]\right\}.$$



6. Minimisation par relaxation stochastique

Cette section concerne l'étape C de la section précédente où on est amené à résoudre le problème de minimisation suivant:

Étant donné, sur $\Omega = E^S$, une loi de Gibbs

$$\pi(x) = \frac{1}{Z} \exp[-H(x)],$$

trouver les états \hat{x} d'énergie minimale.

L'existence de solution ne fait aucun doute, puisque l'espace des configurations Ω est fini. Malgré cela, ce problème est difficile, car Ω est de grande taille. Par exemple, pour image 256×256 avec des étiquettes binaires

$$|\Omega| = 2^{256 \times 256} \sim 10^{19000}$$

Cette complexité exclut tout algorithme classique. Nous développons une solution élégante basée sur la théorie du recuit simulé.

6.1. Recuit simulé pour un champ de Gibbs

Elle est introduite en analyse d'images par D. GEMAN et S. GEMAN en 1984. Notons

$$H_0 = \min_x H(x),$$

$$\Omega_0 = \{x: H(x) = H_0\}.$$

respectivement l'énergie minimale et l'ensemble des configurations optimales.

L'algorithme de recuit simulé consiste à construire une suite de configurations, de loi d'évolution probabiliste (i.e. non déterministe)

$$\{X(0), X(1), \dots, X(n), \dots\}$$

telle que lorsque $n \rightarrow \infty$, $X(n)$ se concentre sur Ω_0 au sens où

$$P[X(n) \in \Omega_0] \rightarrow 1.$$

Son utilisation pratique consiste à :

- faire tourner l'algorithme jusqu'à n grand, ex. $n = 10\,000$;
- prendre $X(10\,000)$ comme solution minimisante \hat{x} .

6.2. Recuit et recuit simulé

Le *recuit* original est un procédé de refroidissement de matériaux où l'on peut atteindre des états terminaux différents suivant le schéma de refroidissement utilisé. Un refroidissement brutal conduit souvent à des états terminaux méta-stables ; à l'inverse un état stable (d'énergie minimale) ne pourra être atteint que par un refroidissement progressif bien contrôlé.

Les inventeurs du recuit simulé, METROPOLIS, ROSENBLUTH & TELLER⁽¹⁾ se sont inspirés de ce phénomène pour traiter le problème de minimisation ci-dessus. Leur observation de base est la suivante. Soit $T > 0$, un paramètre appelé *température* et considérons la loi de Gibbs tempérée :

$$\pi_T(x) \propto \exp\left[-\frac{H(x)}{T}\right].$$

Soient x, z deux configurations quelconques, de différence d'énergie

$$\delta = H(z) - H(x) > 0.$$

Alors

$$\frac{\pi(x)}{\pi(z)} = \exp[\delta],$$

$$\frac{\pi_T(x)}{\pi_T(z)} = \exp\left[\frac{\delta}{T}\right] \nearrow \infty, \quad T \searrow 0.$$

Autrement dit, quand $T \searrow 0$, la distribution tempérée se concentre de plus en plus sur les états d'énergie faible.

6.3. Algorithme stochastique de recuit simulé:

Voici l'algorithme proposé par les frères GEMAN.

1. fixer une suite de températures $T_n = \frac{c}{\log n} \searrow 0, c > 0$;
2. choisir une configuration initiale $X(0)$ quelconque ;
3. à l'étape n , soit $X(n)$ la configuration actuelle, construire la nouvelle configuration $X(n+1)$ par :
 - parcourir le réseau

(1) *J. Chemical Physics*, 1953.

$$\begin{aligned}
 s = & (1,1) \rightarrow (1,2) \rightarrow \dots \rightarrow (1,N) \\
 & \rightarrow (2,1) \rightarrow (2,2) \rightarrow \dots \rightarrow (2,N) \\
 & \quad \vdots \\
 & \quad \vdots \\
 & \rightarrow (M,1) \rightarrow (M,2) \rightarrow \dots \rightarrow (M,N)
 \end{aligned}$$

- en chaque site s , considérer la caractéristique locale déduite de la loi de Gibbs tempérée π_{T_n} :

$$P_{T_n, s} \{ X^{(s)}(n) \},$$

- en simuler un échantillon ω (*très simple !*)
- *relaxer* la valeur du champ en ce site :

$$X_s(n+1) = \omega.$$

Et l'algorithme fonctionne grâce au

Théorème. *Il existe une constante critique c_0 telle que pour tout $c > c_0$, l'algorithme de recuit simulé converge :*

$$\lim_{n \rightarrow \infty} P[X(n) \in \Omega_0] = 1$$

Il est hors de notre portée d'explicitier davantage cette théorie. Signalons simplement que le problème de segmentation a donné naissance à cet algorithme de recuit simulé stochastique, qui a depuis engendré beaucoup de recherche en mathématiques. Le lecteur intéressé trouvera plus d'explications dans les deux références suivantes :

1. D. Geman et S. Geman 1984. *IEEE trans. Pattern Analysis and Machine Intelligency*, **6**, p. 721-741.
2. R. Azencott, 1988. *Séminaire Bourbaki*, 1987-1988, N° 697.

Une façon de se rendre compte de la popularité de cette méthode est de rentrer le mot clé **simulated annealing** sur [www. google. fr](http://www.google.fr) et de compter (si possible) le nombre de pages www trouvées !

7. Applications

Nous terminons cet exposé par quelques applications.

Application 1. Détection de défauts de cuir

La planche 12 montre en (a) l'image originale d'une section de cuir abîmée. Le modèle de segmentation expliqué précédemment a produit la carte des étiquettes estimées montrée en (b). Un post-traitement géométrique a permis de nettoyer les régions jugées défectueuses et de les entourer par des ellipses. Ce marquage sert enfin à la découpe de régions défectueuses.

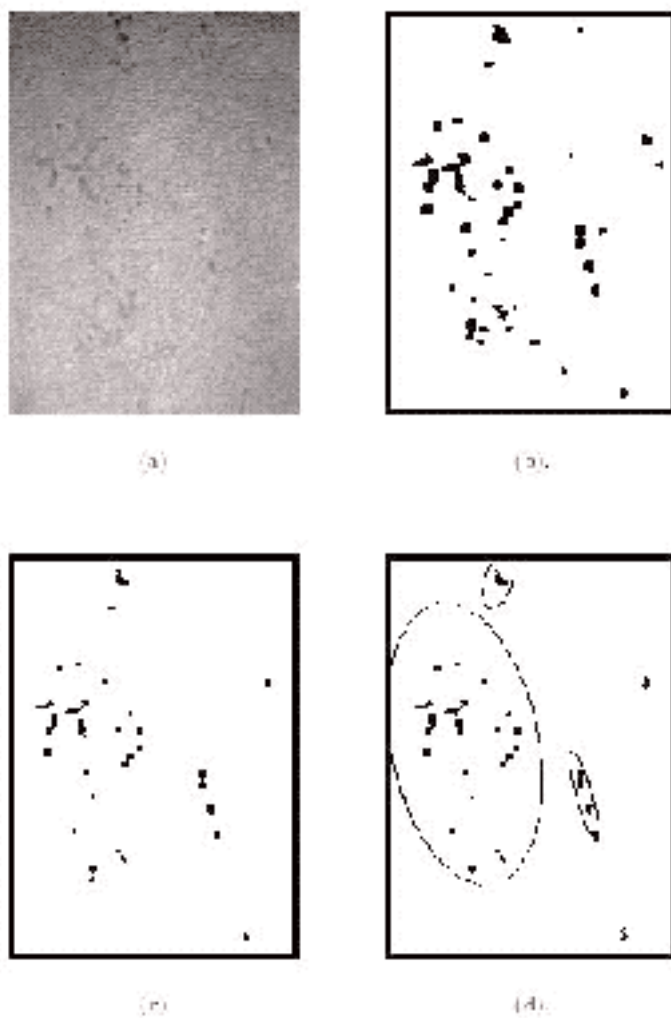
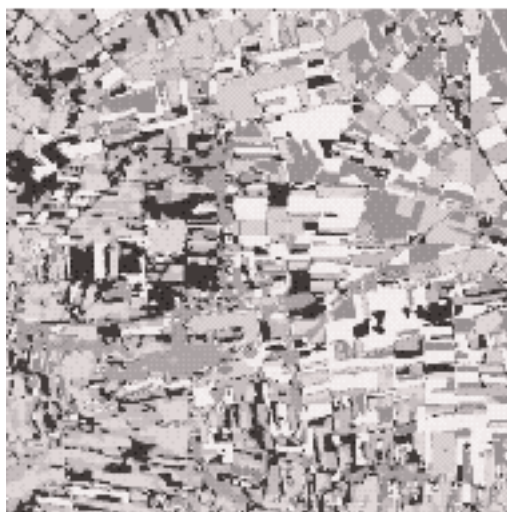


FIG. 12 – Résultat de segmentation et post-traitement des défauts de cuir. Source : SUDIMAGE.

Application 2. Segmentation d'une image Spot.

(a)



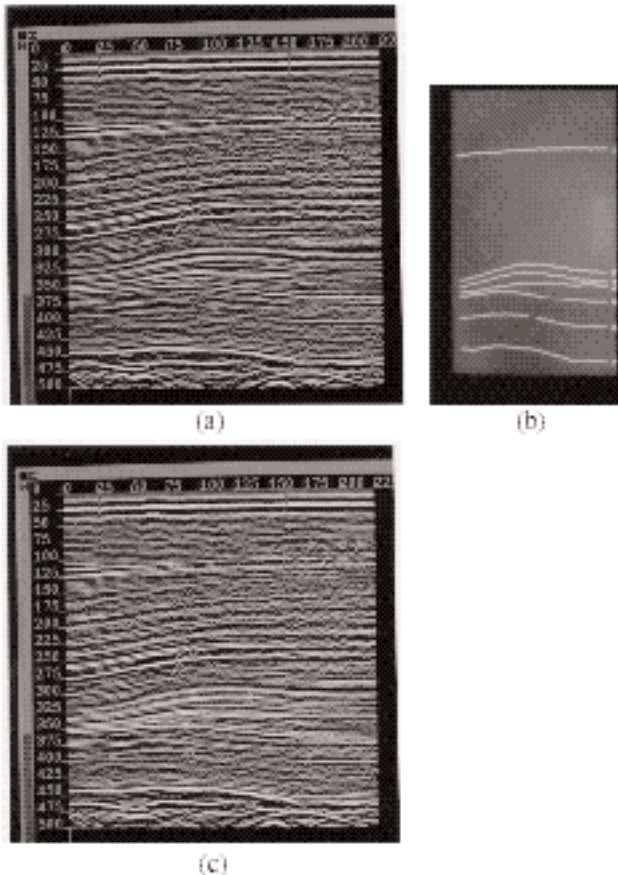
(b)

FIG. 13 – Segmentation d'une image SPOT.

Source : INRIA, Sophia-Antpolis, Équipe de J. Zerubia.

Sur la figure 13, l'image (a) est une image satellite SPOT d'une zone rurale. Sa segmentation en une dizaine de régions de nature différente est donnée en (b) : les étiquettes sont montrées avec des couleurs différentes.

Application 3. Réflexion sismique



Sur l'image originale (a) de cette planche, la détection de 7 séparateurs de couches géologiques est montrée en (b). L'image (c) superpose la détection (b) sur l'image originale (a).

8. Conclusion.

Nous avons décrit dans cet exposé les modèles stochastiques de segmentations en montrant qu'ils conduisent à une bonne formulation mathématique du problème ; que l'on dispose d'algorithmes de résolution efficaces basée sur la théorie du recuit simulé ; qu'ils sont une source de nouveaux développements en mathématiques appliquées.

Nous avons également voulu sensibiliser le enseignants de mathématiques au problème d'analyse d'images, en survolant un petit ε de ce vaste domaine ; en indiquant qu'il demande une coopération de différentes disciplines, allant de la neuro-biologie s'agissant de la compréhension du fonctionnement du cerveau liée à la perception humaine, à l'informatique et aux mathématiques s'agissant de la conception et de la mise en oeuvre des outils théoriques ou pratiques de résolution. La question centrale ici semble être : jusqu'à quel degré la perception humaine d'images peut-elle être comprise par nous les humains ?