

**INTERPRÉTATION COGNITIVE DE LA NON GÉNÉRICITÉ ET DES SIGNIFICATIONS  
INTRINSÈQUES EN TERMES DE PROBABILITÉS**

**Jean Petitot**

École des Hautes Études en Sciences Sociales, Paris

2003

Dans un certain nombre de travaux sur l'analyse d'œuvres d'art picturales ou sculpturales nous avons introduit le concept clé d'une "signification intrinsèque" de certaines compositions fondée sur des critères purement morphologiques et, en particulier sur des critères de "*non généricité*". Nous l'avons appliqué à des phénomènes esthétiques d'un niveau cognitif particulièrement élevé.<sup>1</sup>

On peut considérer une forme globale comme un réseau de parties et de traits morphologiques coactivés et l'analyser au moyen d'un réseau connexionniste de neurones formels. Le problème théorique majeur est alors de comprendre comment le un "principe de clôture" de la composition et les propriétés *non-génériques* de sa structure sélectionnent des oppositions distinctives. Pour cela on peut adapter des modèles de bifurcation perceptive déjà bien connus en ce qui concerne les phénomènes fondamentaux de perception catégorielle.<sup>2</sup> On pondère les relations entre traits locaux et l'on sélectionne les  $n$ -uplés de traits (en général  $n = 2$ ) manifestant un contraste ou une symétrie maximaux. On peut programmer ce processus et extraire ainsi automatiquement les relations pertinentes. En effet, les positions relatives des parties d'un corps articulé sont spontanément processées par la perception. Il est donc facile d'extraire celles qui entrent dans des relations non génériques (alignement, orthogonalité, etc.)

Une fois que l'on a déduit les relations morphologiques pertinentes en faisant usage du principe de clôture, on peut considérer les corrélations (oppositions, etc.) comme une forme de l'expression et les homologuer avec des relations non perceptives et diverses formes du contenu, ce qui les "sémiotise".

Le phénomène morphologique et sémiotique — gestaltiste — d'émergence de significations intrinsèques à partir de critères morphologiques de non généricité se manifeste en fait dès les premiers niveaux de la perception et une modélisation rigoureuse en devient alors possible. Mais le problème demeure difficile car parler de signification intrinsèque à base morphologique signifie que la signification *n'est pas*

---

1 Cf. par exemple Petitot [1988] et [2001].

2 Cf. par exemple Petitot [1984].

déjà donnée. Cela implique qu'il ne s'agit pas d'interpréter conceptuellement une forme donnée mais de dégager de façon *non conceptuelle* ce qui s'y trouve de significatif.

En sciences cognitives, le problème technique est de dépasser les méthodes classiques dites descendantes ou "top down" de "template matching". Pour en donner ici une petite idée, nous allons évoquer un exemple simple relevant des techniques de traitement d'image. Jean-Michel Morel et ses collègues de l'Ecole Normale Supérieure de Cachan ont développé une méthode probabiliste en traitement d'images fondée sur un principe gestaltiste remontant en fait à Helmholtz.<sup>3</sup> L'idée de base est qu'une *grande déviation* par rapport à une situation statistiquement générique engendre une saillance perceptive et une signification *indépendamment de toute connaissance a priori sur la structure de l'image*. Moins une situation est probable et plus elle est intrinsèquement significative. On peut ainsi détecter entre autres les alignements et privilégier ceux qui sont *maximalement* détectables, c'est-à-dire dont les sur- ou sous-configurations sont moins détectables.

L'intérêt de l'approche de Morel est qu'elle arrive à définir des "maximal meaningful events" sans utiliser d'a priori comme on le fait d'habitude dans les approches bayésiennes classiques. Dans un modèle bayésien classique on se donne une image  $I$ , on introduit une classe de modèles satisfaisant des conditions a priori et l'on cherche à abduire le meilleur modèle  $M$  de  $I$  satisfaisant ces contraintes. D'après la règle de Bayes la probabilité conditionnelle  $P(M|I)$  de l'abduction fournissant  $M$  à partir de la donnée  $I$  s'exprime par la formule :

$$P(M|I) = \frac{P(I|M) \times P(M)}{P(I)}$$

La probabilité conditionnelle  $P(I|M)$  de  $I$  étant donné  $M$  correspond au *problème direct* de l'engendrement de l'image  $I$  à partir du modèle  $M$  et est facile à calculer. La probabilité conditionnelle  $P(M|I)$  du modèle  $M$  étant donnée l'image  $I$  correspond quant à elle au *problème inverse* et est en général très difficile à calculer. Or  $M$  correspond à la sa *maximisation*.

Si, par exemple,  $M$  est donné par une distribution de Gibbs :

$$P(M) = \frac{1}{Z} \exp(V(M))$$

où  $Z$  est, comme en physique statistique, une fonction de partition et  $V(M)$  une "énergie" et si  $P(M|I)$  est donnée par une loi gaussienne  $C \exp(-D(M, I))$  où  $D$  est une "distance" entre  $M$  et  $I$ , alors  $M$  est donné par la minimisation de la fonctionnelle  $D + V$ .

---

3 Desolneux, *et al.* [1999].

Les modèles bayésiens correspondent aux techniques top down classiques de template matching. Ils ne sont pas adaptés à notre cas où l'on cherche à définir des *significations non conceptuelles*. Lorsque l'on ne connaît pas d'a priori et que l'on ne cherche pas à trouver un *modèle* interprétant *I* mais seulement à extraire les caractéristiques morphologiques intrinsèquement significatives de *I*, on peut utiliser une *version probabiliste de l'opposition entre générique et non générique* introduite dans ce chapitre.

Jean-Michel Morel a traité l'exemple de l'alignement qui est un exemple typique de regroupement au sens de la Gestalt. Son hypothèse de base est la suivante :

"The main idea is that a meaningful event is an event that, according to probabilistic estimates, should not happen in the image and therefore is significant".

"Significatif" veut donc bien dire ici "non générique". Considérons par exemple dans un cadre de 100×100 pixels un carré noir de 10×10 pixels sur fond blanc. Si la couleur des pixels est tirée au hasard, la probabilité d'une telle configuration est  $90 \times 90 \times (1/2)^{100} \times (1/2)^{900}$ , probabilité de l'ordre de  $10^{-300}$  qui est pratiquement nulle dans tous les sens que l'on peut donner au mot "pratiquement".

Pour la propriété d'alignement, on se donne une grille  $G$  de  $N \times N$  sites et en chaque site  $x = (i, j)$  on considère un petit segment centré sur le point avec une direction  $d_x$ . Cette direction  $d_x$  est une variable aléatoire avec une probabilité uniforme  $p = 1/n$  (i.e. il y a  $n$  directions équiprobables). Soit  $S = (x_1, \dots, x_r)$  un segment de  $r$  points,  $D$  la direction de  $S$  et  $d_i$  les directions des points  $x_i$  de  $S$ . On calcule le nombre maximal  $k(r)$  de points alignés qui doivent être observés pour que l'événement d'alignement soit significatif, i.e. *perceptivement saillant* (ce que l'on appelle un phénomène de "pop out" perceptif).

Plus précisément, soit  $X_i$  la variable aléatoire booléenne de valeur 1 si  $d_i$  est alignée avec  $D$  et de valeur 0 sinon. On a  $P(X_i = 1) = p$  et  $P(X_i = 0) = 1-p$ . La variable aléatoire représentant le nombre des points  $x_i$  de  $S$  d'orientation  $D$  est donc la somme  $S_r$  des  $X_i$ . Les  $X_i$  étant des variables aléatoires indépendantes, on calcule  $P(S_r = k)$  au moyen de la loi binomiale. On définit alors la  $\varepsilon$ -significativité par un seuil  $k(r)$  de significativité :

$$k(r) = \min \left\{ k \text{ entier} \mid P(S_r \geq k) \leq \frac{\varepsilon}{N^4} \right\}$$

---

4 Il y a 90×90 façons de positionner le carré. 1/2 est la probabilité noir/blanc de chaque pixel,  $(1/2)^{100}$  est la probabilité pour que les 100 pixels du petit carré soient noirs et  $(1/2)^{900}$  est la probabilité pour que les 900 pixels restant du grand carré soient blancs.

où  $N^4$  est le nombre de segments orientés (de longueur  $r$  quelconque) dans la grille  $N \times N$ . On obtient par le calcul des estimations précises du seuil  $k(r)$ . Par exemple pour  $N = 512$  et  $p = 1/16$  la longueur minimale d'un segment 1-significatif est de 9. Pour  $N \rightarrow \infty$  et  $r \rightarrow \infty$  "assez vite" on trouve l'estimation :

$$k(r) \simeq pr + \sqrt{2p(1-p) \cdot r \cdot \ln\left(\frac{N^4}{\varepsilon}\right)}$$

On définit ensuite les segments  $S$  maximalelement significatifs qui sont ceux produisant des effets de pop out perceptif.

De telles analyses montrent qu'une approche mathématique rigoureuse des propriétés gestaltistes de la structure morphologique des images est possible. Rien ne s'oppose donc en principe à ce que l'on puisse automatiser la détection des événements intrinsèquement significatifs sur lesquels des significations d'ordre supérieur (conceptuelles et interprétatives) peuvent venir s'accrocher.

#### **BIBLIOGRAPHIE**

- Desolneux, A., Moisan, L., Morel, J.-M., 1999. "Meaningful Alignements", <http://www.cis.ohio-state.edu/~szhu/SCT99.html>
- Petitot, J. 1984. "Paradigme Catastrophique et Perception Catégorielle", *Recherches Sémiotiques / Semiotic Inquiry*, 3, 207-245.
- Petitot, J., 1988. Un mémorialiste du visible (la quête du réel chez Proust), *Protée*, 16, 1-2, 39-52.
- Petitot, J. 2001. "L'espace et sa relation à l'esthétique. L'exemple du Laocoon de Goethe", *Cognition culturelle et cognition spatiale* (A. Quinn éd.), *Visio*, 6/2-3, 263-273.