

## Réseau de Hopfield, un outil pour la mise en correspondance

---

### *Hopfield Neural Network for Stereo Matching*

par **Christophe GARCIA, Saïda BOUAKAZ, Denis VANDORPE**

LIGIA / LISPI — Université Lyon I — Bât. 710  
Boulevard du 11 Novembre 1918  
F-69622 Villeurbanne

#### Résumé

Nous nous intéressons au problème d'appariement de primitives entre deux images. Notre domaine d'application est la mise en correspondance d'un couple d'images stéréoscopiques ou l'identification des parties d'un modèle dans une image observée. Nous proposons dans ce papier une approche utilisant un modèle de réseau de neurones pour résoudre le problème. Nous avons choisi le modèle de Hopfield d'une part parce qu'il est souple et ouvert, d'autre part parce qu'il peut s'implanter aisément sur des calculateurs massivement parallèles.

**Mots clés :** Vision assistée par ordinateur, Vision 3D, Stéréovision, Réseaux de neurones, Mise en correspondance.

#### Abstract

*In this paper, we present an algorithm designed for the stereovision matching problem and 3D objects identification. We use a simulated Hopfield neural network to solve the problem of matching a pair of stereoscopic images. This model is helpful in optimization and it can be implemented on parallel machines easily.*

**Key words :** Computer vision, 3D vision, Stereovision, neural networks, matching.

## 1. Introduction

L'approche que nous présentons s'intègre dans un système de vision « pour la CFAO », système qui doit œuvrer dans le domaine de l'automatisation industrielle et du contrôle qualité [Garcia 93]. Dans de nombreux cas, le contrôle 3D va porter sur des points de la scène dont il faudra mesurer avec précision la position dans un référentiel donné. D'une façon plus générale, la mise en correspondance stéréoscopique nous permettra de rechercher des modèles tridimensionnels dans la scène, ces modèles pouvant provenir directement d'une base de données CFAO. Dans ce cas, le problème est de rechercher rapidement dans la base de données les modèles susceptibles de s'apparier avec les indices perçus. L'étape d'appariement consiste à identifier entre deux images données (deux vues) les informations bidimensionnelles homologues correspondant au même élément physique de la scène observée. Il s'agit clairement d'un problème combinatoire : un point de l'espace se projette dans les deux images en deux endroits différents et comme on ne dispose d'aucune information supplémentaire, il est pratiquement impossible d'établir une correspondance « directe » entre ces deux éléments. Ainsi le choix

des primitives à mettre en correspondance est-il très important. Elles doivent être suffisamment riches pour permettre une mise en correspondance sûre, en nombre suffisant pour permettre d'obtenir une carte 3D suffisamment dense pour être exploitable en modélisation géométrique. Nous donnons ci-après un survol (qui ne prétend pas être exhaustif) de quelques approches proposées dans la littérature et regroupées selon trois critères :

#### • le choix des caractéristiques image :

##### **Appariements de points (*dense stereo matching*)**

Méthodes cherchant à mettre en correspondance les deux images pixel à pixel en minimisant une fonction énergie liée uniquement aux intensités des pixels. Un algorithme pour les images en niveaux de gris peut être trouvé dans [Gennert 88], un algorithme pour les images RVB dans [Jordan 90].

##### **Appariements de points contours (*edge-based stereo matching*)**

Méthodes très utilisées, l'information contenue dans les points contours étant riche pour reconstruire la forme et leur nombre étant réduit. Des algorithmes peuvent être trouvés dans [Mohan et al 89], [Tu et al 90].

### Appariements de régions (*region-based stereo matching*)

Méthodes visant à mettre en correspondance des zones connexes et fermées. Un algorithme intéressant de segmentation utilisant la couleur en format LST peut être trouvé dans [Massen 90].

### Appariements de primitives géométriques (*feature-based stereo matching*)

Méthodes visant à mettre en correspondance des primitives géométriques extraites de chaque image, la plus courante étant le segment de droite, invariant sous la projection perspective. Des algorithmes peuvent être trouvés dans [Ayache 90].

### Appariements de structures (*graph-based stereo matching*)

Méthodes utilisant un groupement de caractéristiques (primitives géométriques liées par des contraintes sémantiques). Des algorithmes peuvent être trouvés dans [Horaud 89].

• **les contraintes utilisées pour réduire la complexité combinatoire :**

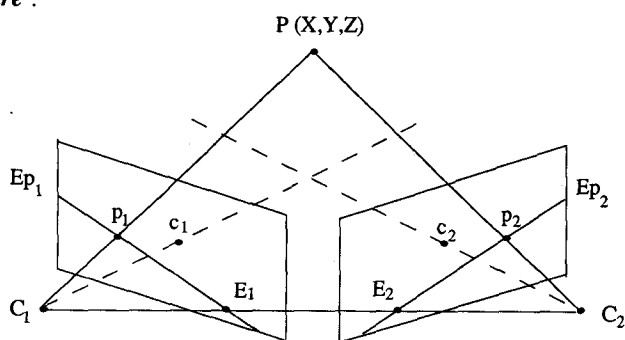


Figure 1. - Système stéréoscopique.

**La contrainte d'unicité :** un point  $p_1$  dans une image a au plus un correspondant  $p_2$  dans l'autre image.

**La contrainte d'ordre :** si un point  $p_{i1}$  se trouve à gauche d'un point  $p_1$  dans une image alors cette relation mutuelle existe dans l'autre image :  $p_{i2}$  est à gauche de  $p_2$ .

**La contrainte de continuité de forme :** une courbe continue de l'espace se projette en des courbes continues dans les deux images.

**La contrainte épipolaire :** un point appartenant à une ligne épipolaire  $Ep_1$  donnée trouve son homologue sur la ligne épipolaire correspondante  $Ep_2$  dans l'autre image. Les lignes épipolaires sont définies comme étant les droites intersection du plan épipolaire  $C_1PC_2$  avec les plans images. Les points  $E_1$  et  $E_2$  sont les points d'intersection de la demi-droite joignant  $C_1$  à  $C_2$  avec chacun des plans image.

• **les stratégies d'appariement :**

Elles dépendent des choix faits dans les deux points précédents. Néanmoins on peut les regrouper en deux classes :

**prédiction / vérification :** voir [Ayache et al 87], [Faugeras 88].

**optimisation d'énergie :** résolution par satisfaction de contraintes (CSP) (voir [FAHLMAN et al 1987], [Mohan 89]).

## 2. Une solution

Pour résoudre le problème de la combinatoire et de l'ambiguïté inhérent à l'appariement, il est nécessaire de faire des hypothèses et d'utiliser les contraintes liées au système. Nous nous retrouvons face à un problème d'optimisation. En effet, nous disposons d'un ensemble d'hypothèses (appariement de caractéristiques) et d'un ensemble de contraintes sur ces hypothèses. Les relations entre ces hypothèses peuvent être coopératives lorsqu'une hypothèse confirme d'autres hypothèses ou en compétition lorsqu'une hypothèse est en conflit avec d'autres. Ce problème est couramment appelé CSP (Constraint Satisfaction Problem) [Mohan 89], [Fahlman et al 87].

### 2.1. MÉTHODE DE RÉOLUTION DU CSP

La méthode consiste à rechercher un sous-ensemble d'hypothèses vérifiées : ces dernières doivent recevoir une coopération maximale des autres hypothèses et entrer en conflit avec un nombre minimum d'autres hypothèses.

Plus précisément, nous pouvons définir :

- un ensemble d'hypothèses :  $H = \{h_i\} \quad |H| = N$
- une fonction VALEUR :  $V(h_i) = V_i, 0 \leq V_i \leq 1$   
 $V_i$  doit être interprétée comme le niveau de confiance accordé à l'hypothèse  $h_i$ . Les hypothèses sont validées ou non selon leur niveau de confiance.
- une fonction SUPPORT :

$$S(h_i, h_j) = T_{ij}V_iV_j, T_{ij} > 0 \quad (1)$$

La valeur de  $S$  donne le niveau d'accord entre les deux hypothèses  $h_i$  et  $h_j$ .

- une fonction CONFLIT :

$$C(h_i, h_j) = T_{ij}V_iV_j, T_{ij} < 0 \quad (2)$$

La valeur de  $C$  donne le niveau de désaccord entre les deux hypothèses  $h_i$  et  $h_j$ .

- une fonction ENTRÉE :

$$I(h_i) = I_iV_i \quad (3)$$

$I_i$  est la somme pondérée des mesures sur les données pour l'hypothèse  $h_i$ . Les mesures qui renforcent l'hypothèse  $h_i$  sont positives et celles qui l'affaiblissent sont négatives.

Résoudre le problème consiste à déterminer les valeurs  $V_i$  qui minimisent une fonction de coût (fonction Énergie) :

$$E = -\sum S - \sum C - \sum I$$

$V_i = 1$  indique que l'hypothèse  $V_i$  est acceptée et  $V_i = 0$  qu'elle est rejetée.

En remplaçant les fonctions  $S$  et  $C$  par les expressions (1), (2) et (3), la fonction énergie  $E$  devient :

$$E = -\sum_i \sum_j T_{ij} V_i V_j - \sum_i V_i I_i \quad (4)$$

Pour déterminer les hypothèses les plus consistantes, les sélections doivent être réalisées simultanément. Aussi, des techniques de résolution séquentielles ne sont pas applicables. Nous avons formulé la fonction coût  $E$  de telle manière que la solution optimale du problème corresponde au minimum de cette fonction. De nombreuses techniques d'optimisation parallèle ont été proposées pour résoudre ce problème, comme le recuit simulé [Kirkpatrick et al. 83] et la machine de Boltzmann [Geman et al. 84].

## 2.2. UN CSN NEURONAL : LE RÉSEAU DE HOPFIELD

Nous avons choisi d'utiliser le réseau de Hopfield ([Hopfield 82]) pour résoudre le problème présenté et cela pour deux raisons principales :

- le modèle de Hopfield correspond directement et naturellement à la formulation CSP
- les neurones formels du modèle de Hopfield sont simples et peuvent être implémentés en Hardware.

Dans le cadre de la résolution CSP, Ce modèle est appelé CSN (Constraint Satisfaction Network).

Dans le problème de mise en correspondance, l'ensemble des appariements potentiels sont les hypothèses et les relations entre ces hypothèses sont les contraintes. Les hypothèses  $V_i$  sont injectées dans les neurones (binaires) du réseau, les contraintes dans les connexions  $T_{ij}$  entre les neurones. Les mesures sur les données (caractéristiques) sont les entrées  $I_i$  (biais) sur chaque neurone (voir figure 2).

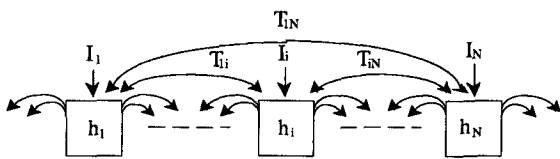


Figure 2. – Réseau CSN présentant  $N$  hypothèses.

Pour assurer la convergence du réseau vers un état stable minimum, les connexions  $T_{ij}$  doivent être symétriques et non réentrantes : Aussi,  $T_{ij} = T_{ji}$  et  $T_{ii} = 0$ .

Le changement d'état  $\Delta V_i$  d'un neurone  $i$  entraîne une variation d'énergie  $\Delta E$  :

$$\Delta E = -(\sum_j T_{ij} V_j + I_i) \Delta V_i$$

Hopfield a montré que la dynamique du système est toujours négative si l'on suit les règles suivantes : [HOPFIELD 82]

$$V_i \longrightarrow 0 \quad \text{lorsque} \quad \sum_j T_{ij} V_j + I_i > 0$$

et

$$V_i \longrightarrow 1 \quad \text{lorsque} \quad \sum_j T_{ij} V_j + I_i < 0$$

Suivant ces règles de transition, il a été démontré que le réseau convergeait vers un état stable, dans le cas d'une mise à jour aléatoire et asynchrone. Dans ce type de mise à jour, chaque neurone s'interroge aléatoirement et se met à jour indépendamment des autres unités. Pour réduire le temps de relaxation, nous avons choisi une approche synchrone qui permet un balayage rapide de tous les neurones. Dans ce cas, la convergence n'est pas démontrée. Dans nos diverses expérimentations, cette convergence a été obtenue, le critère d'arrêt reposant sur la détection de cycles et l'utilisation d'un nombre limité d'itérations.

Zhou et Chellappa [Zhou et al. 88] utilise un modèle de Hopfield tridimensionnel pour appairer point à point les deux images en se basant essentiellement sur la luminance des points. Chaque neurone est relatif aux points à appairer et à la disparité qui leur est associée. Ce modèle est très coûteux d'un point de vue stockage et trop dédié à la reconstruction dense.

Nous avons préféré appairer des points de contour en tenant compte de l'ensemble des contraintes disponibles c'est-à-dire des contraintes photométriques (contraste...) et géométriques (épipolaires, ordre...).

Nous désirons appairer successivement des points extraits sur deux lignes épipolaires correspondantes (voir Figure 1). Nous sélectionnerons  $N$  points dans l'image 1 et  $M$  dans l'image 2.

Dans notre approche, chaque neurone sera de la forme  $V_{ij,p}$  contenant l'hypothèse  $h_{ij}$  : « le  $i^{\text{ème}}$  point de la ligne épipolaire  $p$  de l'image 1 est apparié avec le  $j^{\text{ème}}$  point de la ligne épipolaire  $p$  de l'image 2 ».

Les connexions et les biais prendront la forme  $T_{ij,kl,p}$  et  $I_{ij,p}$ .

L'énergie à minimiser est donc la suivante :

$$E = -\frac{1}{2} \sum_{ij} \sum_{kl} T_{ij,kl,p} V_{ij,p} V_{kl,p} - \sum_{ij} V_{ij,p} I_{ij,p}$$

dans laquelle :

- $T_{ij,kl,p}$  est le poids de la connexion entre les neurones  $(i, j)$  et  $(k, l)$ .
- $V_{ij,p}$  et  $V_{kl,p}$  sont les états (binaires) des neurones  $(i, j)$  et  $(k, l)$ .
- $I_{ij,p}$  est la valeur de l'entrée propre (ou biais) du neurone  $(i, j)$ .

Pour résoudre un problème d'optimisation à l'aide de ce réseau, il nous faut établir une fonction qui obéisse aux règles souhaitées et par identification avec la forme générale donnée précédemment, nous en déduisons les expressions des biais de chaque neurone et des poids de connexion entre neurones.

Dans notre cas, la mise en œuvre s'est faite à partir de la fonction

d'énergie suivante, dans laquelle on notera que chaque terme reprend une contrainte énoncée plus haut :

$$\begin{aligned}
 E = & \left(\frac{\alpha}{2}\right) \sum_i \sum_j \sum_{l(l \neq j)} V_{ij,p} \cdot V_{il,p} \\
 & + \left(\frac{\beta}{2}\right) \sum_i \sum_j \sum_{k(k \neq i)} V_{ij,p} \cdot V_{kj,p} \\
 & + \left(\frac{\chi}{2}\right) \sum_i \sum_k \sum_j \sum_l |\text{ordre}(i, k) - \text{ordre}(j, l)| V_{ij,p} \cdot V_{kl,p} \\
 & + \left(\frac{\delta}{2}\right) \sum_i \sum_k \sum_j \sum_l |\text{disp}(i, j) - \text{disp}(k, l)| V_{ij,p} \cdot V_{kl,p} \\
 & + \varepsilon \sum_i \sum_j C_{ij,p} \cdot V_{ij,p}
 \end{aligned} \quad (5)$$

où

$$\begin{aligned}
 \text{ordre}(m, n) &= 1 \text{ si } m > n \text{ et } \text{disp}(m, n) = m - n \\
 &= 0 \text{ sinon}
 \end{aligned}$$

Dans cette expression,  $\alpha, \beta, \chi, \delta, \varepsilon$  sont des constantes positives qui permettent de pondérer l'influence des différentes contraintes.

- les deux premiers termes correspondent à la contrainte d'unicité, respectivement dans les images 1 et 2 : si les neurones  $(i, j)$  et  $(i, l)$  (resp.  $(i, j)$  et  $(k, j)$ ) sont actifs en même temps, cela signifie qu'un élément de l'image 1 (resp. image 2) peut correspondre à deux éléments de l'image 2 (resp. image 1), ce qui n'est pas possible.
- le troisième terme reprend la contrainte d'ordre : si les couples d'éléments  $(i, j)$  et  $(k, l)$  sont mis en correspondance ( $V_{ij,p}$  et  $V_{kl,p}$  actifs), il faut que ces éléments se retrouvent dans le même ordre dans les deux images.
- le quatrième terme formalise la contrainte de disparité. Si les couples d'éléments  $(i, j)$  et  $(k, l)$  sont tous deux activés, il faut qu'ils aient une disparité voisine.
- le cinquième terme fait intervenir la dissimilarité  $C_{ij,p}$  (définie plus loin), mesure de ressemblance entre les caractéristiques  $i$  et  $j$ .

La fonction énergie du réseau étant posée, on peut déduire les expressions de  $T_{ij,kl,p}$  et  $I_{ij,p}$ , soit :

$$\begin{aligned}
 \bullet T_{ij,kl} = & -\alpha \text{eg}(i, k) \text{diff}(j, l) - \beta \text{eg}(j, l) \text{diff}(i, k) \\
 & - \chi |\text{ordre}(i, k) - \text{ordre}(j, l)| \text{diff}(i, k) \text{diff}(j, l) \\
 & - \delta |\text{disp}(i, j) - \text{disp}(k, l)|
 \end{aligned} \quad (6)$$

où :

$$\begin{aligned}
 \text{eg}(m, n) &= 1 \text{ si } m = n \text{ et } \text{diff}(m, n) = 1 \text{ si } m \neq n \\
 &= 0 \text{ sinon} & \text{diff}(m, n) &= 0 \text{ sinon}
 \end{aligned}$$

$$\bullet I_{ij,p} = \Omega - \varepsilon C_{ij,p}$$

où  $C_{ij,p}$  est la dissimilarité entre les caractéristiques  $i$  et  $j$ .

Les paramètres  $\Omega$  ( $\Omega > 0$ ) et  $\varepsilon$  pondèrent le calcul des biais. Remarquons que le paramètre  $\Omega$  a été ajouté en tant que source d'excitation globale du réseau. Nous avons choisi ici pour le calcul des biais une fonction linéaire des dissimilarités  $C_{ij}$ .

La fonction dissimilarité  $C_{ij}$  permet d'évaluer le niveau de ressemblance entre les caractéristiques des deux images. Seule cette

fonction dépend des caractéristiques images. Elle intervient au niveau des biais des neurones du réseau. Ses valeurs agissent comme des activateurs sur les neurones lorsqu'elles sont proches de 0, et comme des inhibiteurs lorsqu'elles sont fortement positives.

Les valeurs des connexions entre neurones sont indépendantes des caractéristiques images. Elles ne seront donc déterminées qu'une seule fois pour un réseau de taille donnée. Les connexions sont inhibitrices lorsqu'elles sont fortement négatives ( $T_{ij,kl} \ll 0$ ) et activatrices lorsqu'elles sont proches de 0. L'évolution du réseau va permettre de passer par un certain nombre d'états intermédiaires (ou configurations), états qui correspondront à des valeurs décroissantes de l'énergie (dynamique négative) jusqu'à tendre vers un état stable. Cet état, correspondant à une valeur minimale locale de l'énergie, sera alors la solution du problème soumis au réseau. La notion d'état stable ici correspond à la situation où l'énergie du réseau ne varie plus ou que l'on a atteint le nombre maximum d'itérations choisi a priori.

Une nouvelle configuration du réseau se calcule à partir des quantités  $E_{ij,p}$  pour chaque neurone  $(i, j)$ . Elles caractérisent les influences combinées de l'ensemble des neurones  $(k, l)$  du réseau sur chaque neurone  $(i, j)$ .

Pour l'itération  $n$  :

$$E_{ij,p}^{(n)} = \sum_k \sum_l T_{ij,kl,p} \cdot V_{kl,p}^{(n)} \oplus I_{ij,p}$$

Les valeurs  $V_{ij}$  sont mises à jour de manière synchrone à l'aide d'une fonction de transfert  $\phi$  appliquée à  $E_{ij,p}$  :

$$V_{ij,p}^{(n+1)} = \phi(E_{ij,p}^{(n)})$$

Cette fonction est généralement une sigmoïde. Dans le cas binaire, elle correspond à un seuillage. Nous avons préféré l'opérateur « max » pour conserver le caractère adaptatif du réseau. Cet opérateur nous permet d'imposer strictement la contrainte d'unicité, prise en compte uniquement jusque là dans le calcul des connexions.

$$\phi(X_{ij}) = 1 \text{ si}$$

$$X_{ij} = \max\{\{X_{il}/l = 1, \dots, M\} \cup \{X_{kj}/k = 1, \dots, N\}\}$$

$$\phi(X_{ij}) = 0 \text{ sinon}$$

où  $N$  et  $M$  sont respectivement les nombres de lignes et de colonnes de la matrice neuronale.

### 2.3. IMPLÉMENTATION

Nous avons choisi d'extraire des points contour en utilisant le filtre récursif de DERICHE [DERICHE 87]. Les points contour sont sélectionnés dans chaque image des gradients par suppression des non-maxima dans la direction de la normale au contour. Ces points se situent sur des zones de discontinuité des surfaces visibles et seront présents (aux occlusions près) dans les deux images stéréoscopiques.

Chacun de ces points va nous fournir le vecteur de caractéristiques  $p_i$  suivant :

$$p_i[(x_i, y_i), M_i, A_i]$$

où :  $(x_i, y_i)$  est la position du point dans l'image.

$M_i$  est la magnitude du gradient au point  $(x_i, y_i)$ .

$A_i$  est l'angle de la normale au contour au point  $(x_i, y_i)$ .

De façon à utiliser la contrainte épipolaire, nous formons les ensembles suivants à appairer :

$$C_{1,p} = \{p_i \text{ appartient à la ligne épipolaire } p \text{ dans l'image 1}\}$$

et

$$C_{2,p} = \{p_j \text{ appartient à la ligne épipolaire } p \text{ dans l'image 2}\}$$

Nous appliquons la résolution du problème par le CSN pour chaque ligne épipolaire  $p$  dans les deux images.

L'algorithme peut être décrit de la façon suivante :

#### a) Initialisation des neurones

Avant de débiter la relaxation du réseau, nous devons fixer les états initiaux des neurones. Ne disposant pas d'information sur la scène observée, tous les couples de points sont des correspondants potentiels. Par suite, chaque neurone doit être activé. Mais, de façon à se placer dans une configuration plus proche de la configuration finale et ainsi, réduire le nombre d'itérations à effectuer, nous pouvons classiquement fixer une disparité maximum permise  $D_{\max}$ . Aussi,

$$V_{ij,p} = 1 \text{ si } \|\vec{p_i p_j}\| < D_{\max} \text{ pour } p_i \in C_{1,p} \text{ et } p_j \in C_{2,p}$$

$$V_{ij,p} = 0 \text{ sinon}$$

Pour se placer dans une configuration initiale proche de la solution finale et de manière à réduire le coût de la relaxation, il est possible de n'activer, pour un indice  $i$  donné, que le neurone  $V_{ij,p}$  dont la mesure de dissemblance  $C_{ij,p}$  est minimum.

#### b) Calcul des scores de Dissimilarité et des biais

Il nous faut définir une mesure de dissemblance pour chaque couple potentiel. Nous avons choisi la fonction linéaire suivante :

$$C_{ij,p} = a \Delta M_{ij} + b \Delta A_{ij} + c(1 - \rho_{ij})$$

où :

$$\Delta X_{ij} = \frac{|X_i - X_j|}{|X_{\max} - X_{\min}|}$$

avec  $a > 0$ ,  $b > 0$  et  $c > 0$ .

- $\Delta M_{ij} \in [0, 1]$  est la différence normalisée des magnitudes des gradients entre  $p_i$  et  $p_j$ .

- $\Delta A_{ij} \in [0, 1]$  est la différence normalisée des angles des normales aux contours entre  $p_i$  et  $p_j$ .

- $\rho_{ij} \in [-1, 1]$  est le coefficient de la corrélation normalisé entre les voisinages de  $p_i$  et  $p_j$ , qui nous permet de considérer les informations d'intensité lumineuse locales.

Aussi,  $C_{ij,p}$  est une somme pondérée de  $\Delta M_{ij}$ ,  $\Delta A_{ij}$  et  $(1 - \rho_{ij})$ .

Par suite,  $I_{ij,p} = \Omega - \varepsilon(a\Delta M_{ij} + b\Delta A_{ij} + c(1 - \rho_{ij}))$ .

#### c) Calcul des connexions synaptiques

L'expression des connexions synaptiques est donnée en [6].

On remarque que :  $T_{ij,ij,p} = 0$  et  $T_{ij,kl,p} = T_{kl,ij,p}$ .

Pour optimiser l'algorithme, nous ne calculons  $T_{ij,kl,p}$  que si  $V_{kl,p} = 1$ .

Dans le cas où  $V_{kl,p} = 0$ ,  $T_{ij,kl,p}$  n'a pas d'influence sur le réseau.

#### d) Relaxation du réseau

Cette étape est réalisée de manière synchrone dans notre implémentation.

1) Calculer  $E_{ij,p} = \sum_k \sum_l T_{ij,kl,p} V_{kl,p} + I_{ij,p}$   
 $\forall p_i \in C_{1,p} \forall p_j \in C_{2,p}$ .

2) Modifier l'état du neurone  $(i, j, p)$  avec  $\Phi$   
 $V_{ij,p} = \Phi(E_{ij,p}) \forall p_i \in C_{1,p} \forall p_j \in C_{2,p}$

3) Y a-t-il des changements d'états  $V_{ij,p}$ ?  
 Si oui retour à 2).

La configuration des neurones fournit la solution du problème.

Nous présentons un exemple de relaxation du réseau en Figure 3. Les deux images stéréoscopiques correspondent à des couverts posés sur une table. Nous résolvons le problème d'appariement sur la ligne épipolaire n°49 sur laquelle 15 points ont été extraits dans l'image 1 et 13 dans l'image 2.

#### e) Reconstruction 3D

A l'issue du processus de mise en correspondance, nous disposons de paires de points  $(p1, p2)$  appariés provenant des deux plans image. La position du point  $P(X, Y, Z)$  de l'espace physique correspondant à  $(p1, p2)$  est déterminée, d'après la géométrie épipolaire, comme étant à l'intersection des deux droites de supports  $C1p1$  et  $C2p2$  ( $C1$  et  $C2$  centres optiques des caméras). A partir des matrices de calibration des deux caméras ([Faugeras 88], [Tsai 88]), nous obtenons un système linéaire de 4 équations à 3 inconnues  $X, Y, Z$  que nous résolvons par une méthode de moindres carrés surdéterminés.

Le système, malgré l'ensemble des contraintes utilisées, peut hésiter entre plusieurs correspondances pour un même élément lorsque les caractéristiques sont très proches ou fournir de fausses correspondances. On se trouve alors, dans un cas de mise en correspondance ambiguë.

Les causes sont les suivantes :

- les données fournies ne sont pas suffisamment précises pour que le réseau puisse décider sans douter sur le choix qui est fait.
- certaines parties apparaissent dans une image et pas dans l'autre. Les angles de vues des caméras entraînent des zones d'occlusion.

Pour résoudre ce problème, nous pouvons utiliser une valeur de disparité moyenne. On considère les couples d'éléments mis directement en correspondance par le réseau et on calcule la disparité associée : on détermine une disparité moyenne sur l'ensemble de l'image. Celle-ci sera alors utilisée comme une valeur seuil de la disparité pour le reste de l'image.

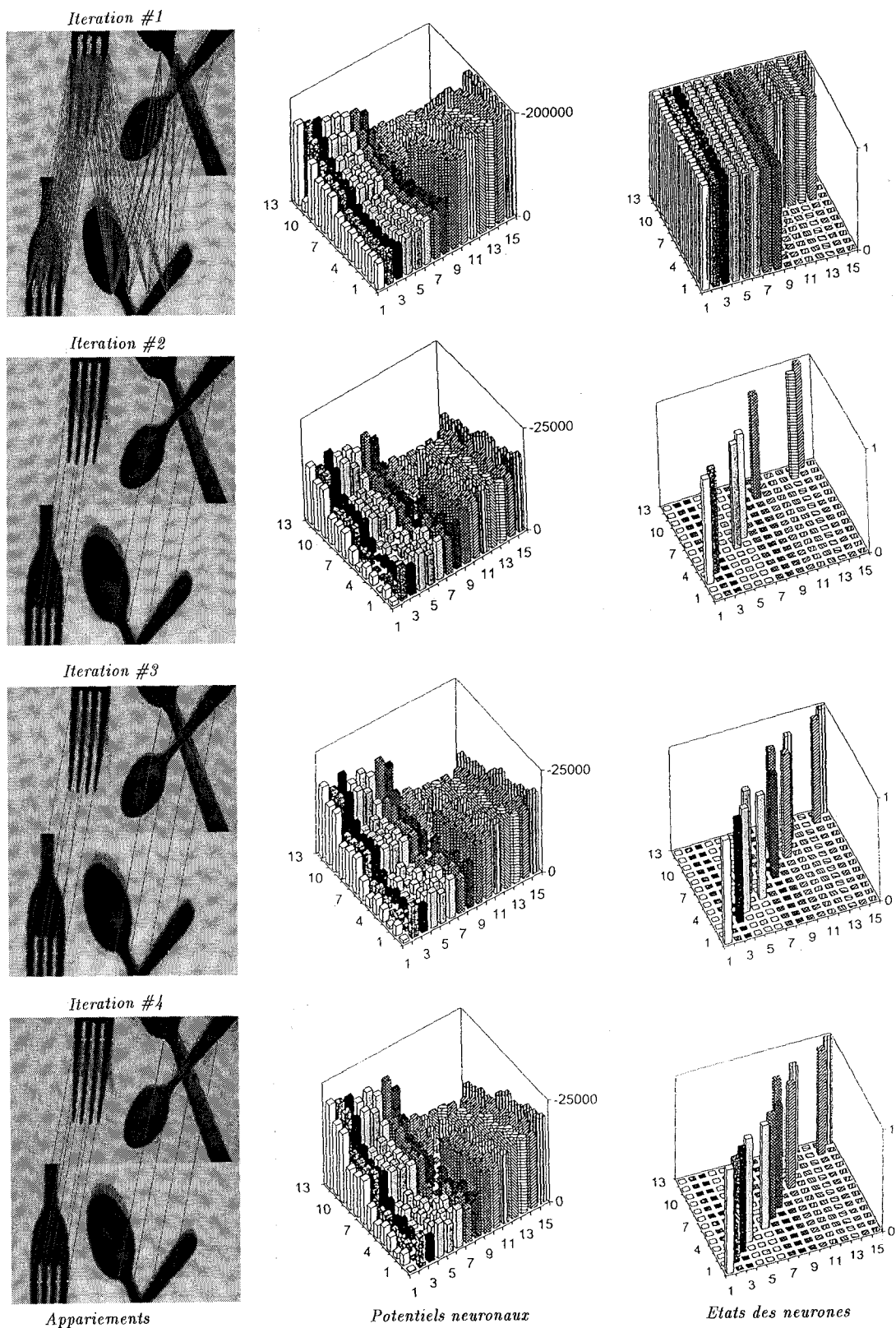


Figure 3. – Relaxation pour la ligne épipolaire # 49 (« couverts »).

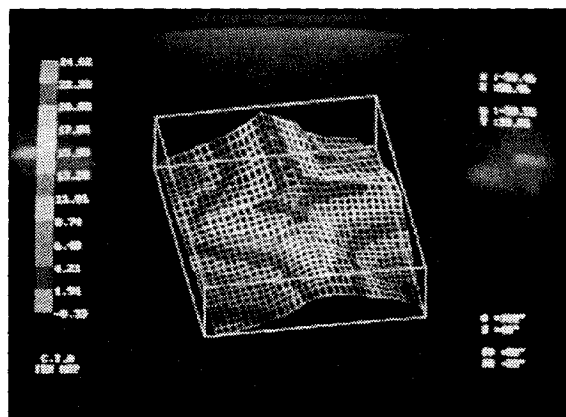


Figure 4. – Système de stéréovision des points 3D reconstruits.

#### 2.4. ERREURS DE RECONSTRUCTION

Des points 3D reconstruits sont erronés : ils proviennent des erreurs de mise en correspondance. La suppression de ces points est possible en utilisant la contrainte de continuité surfacique (en dehors des points situés sur les frontières d'occultation). Un cumul des erreurs apparaît aussi dans le processus de détection des primitives image et dans la résolution des systèmes d'équation.

La reconstruction d'un objet 3D de type polyédrique dont les dimensions sont connues précisément nous a permis de déterminer l'erreur moyenne relative aux points 3D reconstruits. Avec des objectifs de 25 mm, à une distance de 40 cm, l'erreur moyenne mesurée est de 0.19 mm.

Nous présentons une vue de notre système de stéréovision en Figure 4.

#### 2.5. RÉSULTATS EXPÉRIMENTAUX

Nous avons testé notre algorithme sur de nombreuses images de taille  $512 \times 512$ . Les images utilisées proviennent aussi bien de systèmes à géométrie épipolaire parallèle (les images se correspondent ligne à ligne) que de systèmes stéréoscopiques en position générale (cf. Figure 1). Le réseau a été simulé sur une station Silicon Graphic Indigo 2.

Le choix de l'ensemble des constantes utilisées est réalisé actuellement de manière empirique. Les constantes  $\alpha, \beta, \chi, \delta$ , qui pondèrent l'action des contraintes doivent être choisies dans l'intervalle  $[10, 20]$ . Nos nombreux tests ont montré que des valeurs hors de cette plage inhibaient trop le réseau. La valeur de  $\varepsilon$  doit être nettement plus élevée ainsi que celle de  $\Omega$  car toutes deux agissent comme source externe (excitation) sur le réseau. Des valeurs supérieures à 50 donnent de bons résultats.

Les coefficients  $a$ ,  $b$  et  $c$  valent 5. Ils sont identiques car nous ne voulons privilégier aucune caractéristique dans l'espace des éléments image. Avec ces paramètres, le réseau converge rapidement (6 itérations en moyenne).

Nous présentons en Figure 5 des résultats sur trois couples d'images.

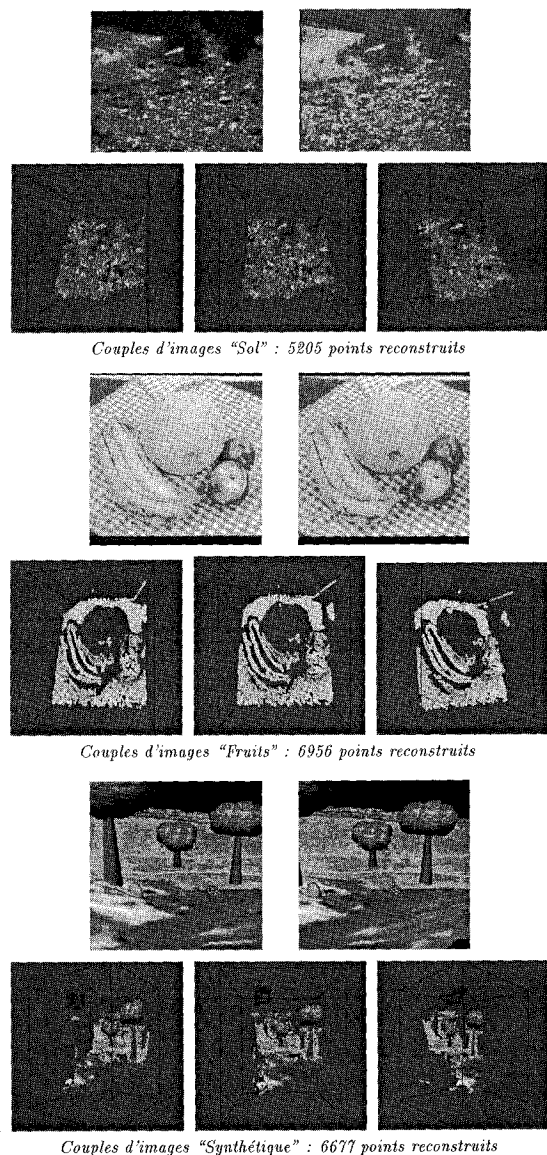


Figure 5. – Quelques résultats.

### 3. Conclusion

Nous avons présenté un système utilisant un réseau neuronal de type Hopfield pour réaliser l'appariement d'informations provenant de deux images stéréoscopiques. L'utilisation d'un tel réseau permet une mise en correspondance globale des caractéristiques : un appariement potentiel est confirmé ou infirmé en fonc-

tion des autres appariements. En outre cet algorithme présente l'avantage de pouvoir s'intégrer directement dans une architecture informatique massivement parallèle. L'avantage de cette approche réside dans le fait que les liens entre neurones sont indépendants des données en entrée (seules les hypothèses sont formulées en fonction des données à traiter). Aussi, cet algorithme est suffisamment générique pour être utilisé dans de nombreux problèmes de vision tels que l'étiquetage, la reconnaissance de forme... Nous travaillons actuellement sur le choix des caractéristiques à appairer (utilisation d'éléments caractéristiques variés associés à des fonctions de ressemblances appropriées), sur le choix de l'ensemble des paramètres du réseau (détermination par apprentissage) et sur une architecture à base de Transputers.

Nous pensons que cette approche neuronale nous permettra de mettre en correspondance des descriptions monoculaires structurées extraites de chaque image pour la reconstruction de modèles 3D s'intégrant dans des systèmes CFAO.

### 4. Remerciements

Une grande partie de ces travaux a été effectuée au Centre des Techniques Avancées de l'usine IBM Montpellier. Nous tenons à remercier en particulier Jean Michel Corrieu pour le financement de ces travaux.

### BIBLIOGRAPHIE

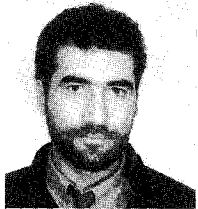
- [AYACHE 90] Ayache N., « Stereovision and sensor fusion », MIT Press, 1990.
- [DERICHE 87] Deriche R., « Using Canny's criteria to derive an optimal edge detector recursively implemented », *International Journal of Computer Vision*, Vol. 2, 1987, p. 167-187.
- [FAHLMAN et al 87] Falhman S.E., Hinton G.E., « Connectionist architecture for artificial intelligence », *IEEE Computer*, January 1982, p. 100-109.
- [FAUGERAS 88] Faugeras O. D., « Quelques pas vers la vision artificielle en trois dimensions », *Technique et Science Informatique*, 1988.
- [GEMAN et al. 84] Geman S., Geman D., « Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images », *IEEE Transaction on Pattern Analysis and Machine Intelligence*, PAMI-6, Novembre 1984, p. 721-741.
- [GARCIA et al 93] Garcia C., Bouakaz S., Vandorpe D., « Un système de stéréovision utilisant le réseau de Hopfield », *Proceedings of the 12th Conference of CAD/CAM, Computer Graphics and Computer aided Technologies*, Paris, Janvier 1993, p. 205-224.
- [GARCIA et al 94] Garcia C., Bouakaz S., Vandorpe D., « Réseau de Hopfield : un outil pour la mise en correspondance », *Proceeding of the 9th Conference RFIA*, Paris, Février 1994, p. 371-378.
- [GENNERT 88] Gennert M.A., « Brightness-based stereo matching », *Proceedings of the SICCV*, Tampa, Florida, p.139-143, 1988.
- [HOPFIELD 82] Hopfield J., « Neural networks and physical systems with emergent collective computational abilities », *Proceedings of National Academy of Sciences*, Vol. 79, 1982, p. 2554-2558.
- [HORAUD 89] Horaud R., Skordas T., « Stereo matching through feature Grouping and maximal clique », *IEEE transactions on PAMI-11*, November 1989.
- [JORDAN 90] Jordan J.R., « Dense stereo correspondence using color », *Proceedings of the SPIE*, vol. 1382, 1990.
- [KIRKPATRICK et al. 83] Kirkpatrick S., Gellat C., Vecchi M., « Optimization by simulated annealing », *SCI*, Vol. 20, 1983, p. 671-680.



- [MASSEN 90] Massen R., « Real time color classification for preprocessing photogrammetry images », *SPIE*, vol. 1395, 1990.
- [MOHAN 89] Mohan R., « Constraint Satisfaction Networks for Vision », *Exploratory Computer Vision Group IBM Thomas J. Watson Research Center Yorktown*, 1989.
- [MOHAN et al 89] Mohan R., Medioni G., Nevatia R., « Stereo Error Detection, Correction, and Evaluation », *IEEE Transactions on PAMI-11*, N°2, February 1989.
- [TSAI 88] Tsai R.Y., « Synopsis of Recent Progress on Camera Calibration », *Research Report 13888, IBM Research Division, Yorktown Heights*, Août 1988.
- [TU et al 90] Tu X., Dubuisson B., « 3-D Information derivation from a pair of binocular images », *Pattern Recognition*, Vol. 23 N°3/4, 1990, p. 223-235.
- [ZHOU et al 88] Zhou Y.T., Chellappa R., « Stereo matching using a neural network », *Proc. Intl. Joint Conf. on Acoustics, Speech, and Signal Processing*, New York, Avril 1988, p. 940-943.

## LES AUTEURS

### Christophe GARCIA



Christophe Garcia a obtenu son diplôme de Doctorat en Informatique devant l'université Claude Bernard Lyon I en Septembre 1994. Durant deux ans, il a travaillé au sein du groupe « Vision Automation » de l'usine IBM de Montpellier plus particulièrement dans le domaine du contrôle qualité et de la robotique mobile. Il est actuellement Attaché Temporaire d'Enseignement et de Recherche à l'université Claude Bernard Lyon I. Ses travaux de recherche sont axés sur la stéréovision, la reconstruction tridimensionnelle et les réseaux de neurones artificiels.

### Saïda BOUAKAZ



Saïda Bouakaz a obtenu son diplôme de Doctorat en Mathématiques Appliquées à Grenoble. Depuis 1990, elle est Maître de Conférence à l'Université Claude Bernard Lyon I. Membre de l'équipe LIGIA, ces travaux de recherche sont axés sur la vision tridimensionnelle et essentiellement la mise en correspondance, la modélisation et la reconstruction 3D.

### Denis VANDORPE



Denis Vandorpe est professeur à l'université Claude Bernard Lyon I. Il dirige l'équipe du LIGIA (Informatique Graphique et Intelligence Artificielle) au sein du Laboratoire d'Informatique des Systèmes de Production Industrielle (LISPI). Il travaille depuis de nombreuses années dans le domaine de la CFAO et plus particulièrement de la modélisation. Il a été durant quatre ans président de l'A.F. Micado.

*Manuscrit reçu le 17 Octobre 1994.*