

Chapitre III

Prototype VLSI-CMOS pour le Processeur Parallèle Optoélectronique Stochastique

1. Introduction

Nous allons présenter maintenant un premier prototype de *processeur parallèle optoélectronique stochastique* réalisé en technologie VLSI-CMOS et dédié au problème du *recuit simulé* sur champ binaire *avec* ou *sans* contrainte de rappel aux données (restauration d'images binaires et modèle d'Ising 2D respectivement). Il s'agit d'une puce CMOS à entrées électroniques *et* optiques, (la sortie est réalisée uniquement par voie électronique)³⁰. Un générateur de nombres aléatoires *optique* rend stochastique la dynamique d'évolution de cette puce, baptisée SPIE600 (pour *Speckle Projection for Ising Emulation*, à $24 \times 25 = 600$ cellules ou processeurs élémentaires).

Deux exemples de démonstration sont possibles avec le prototype SPIE600 à interconnexions *électroniques* ; le premier - démontré lors d'une thèse précédente au laboratoire [Prévos95] - rend compte du bon fonctionnement du circuit pour simuler le comportement d'un réseau de spins bidimensionnel (modèle d'Ising 2D) ; le second est à peine plus élaboré, mais déjà plus intéressant : il s'agit de démontrer la capacité du processeur à traiter le problème du *débruitage d'images binaires à cadence vidéo* (sans prise en compte des bords). Dans ce cas, une séquence d'images binaires détériorées par un bruit de canal est présentée à la rétine. Pour chacune d'elles, la relaxation du champ d'étiquettes est réalisée alors avec *contrainte de rappel aux données* ; les interactions -

³⁰ Le circuit pourrait prétendre au nom de *rétine artificielle* étant donné que l'acquisition et le traitement de l'image sont réalisés sur le même circuit. Cependant, l'image (binaire) à traiter doit être codée au préalable au format *double-voie*, c'est pourquoi on préférera le terme de *processeur optoélectronique* ou bien celui de *matrice de pixels intelligents* (cf. Chap.VII).

électroniques - entre différents PEs demeurent dans ce cas exclusivement positives. Le résultat est une image binaire lissée. Il faut cependant un montage optique additionnel pour projeter l'image codée en mode double-voie sur le circuit SPIE600 (celui-ci sera décrit au Chap.V).

2. Rétine artificielle ou machine de Boltzmann ?

La puce SPIE600 n'a pas été conçue dans le but de traiter les deux cas *d'école* décrits en introduction ; en fait, SPIE600 n'est qu'un premier pas permettant de valider certains éléments constitutifs d'une véritable *rétine artificielle stochastique* dédiée à la restauration d'images à niveaux de gris avec prise en compte des bords. Ce qui suit est destiné à bien placer le prototype SPIE600 dans le cadre de cette recherche sur les rétines artificielles³¹ ; il en ressortira également que SPIE600 peut être considéré à elle seule comme un prototype optoélectronique simplifié d'une machine de traitement du signal beaucoup plus générale connue sous l'appellation de *machine de Boltzmann*.

2.1 Une rétine artificielle pour la restauration d'images...

Le problème de la restauration d'images dégradées a été exposé en détail au Chap.I,§5. Plusieurs auteurs ont proposé des réalisations VLSI permettant de traiter ce problème. Rappelons que la fonction à minimiser était (*cf.* Chap.I, §5.1.3) :

$$U(e, o, l) = \beta_o \sum_{s \in S} (e(s) - o(s))^2 + \beta_s \left\{ \sum_{s \in S} \sum_{r \in V(s)} (e(s) - e(r))^2 \cdot (1 - l_{sr}) + \Psi_B(l) \right\},$$

où o , e et l sont respectivement le champ d'observation (image à niveaux de gris bruitée), le champ d'étiquettes à niveaux de gris, et enfin le champ binaire des bords. Comme on l'a fait remarquer auparavant, la minimisation de la fonction d'énergie par rapport aux variables d'un champ d'étiquettes à niveaux de gris continus (pour un champ de bords donné) est un problème "facile" que l'on peut résoudre par des algorithmes déterministes efficaces. En effet, la fonction d'énergie est quadratique par rapport aux variables du champ ; en outre il est relativement facile de montrer que celle-ci est convexe par rapport aux variables continues [Blake87] : il suffit alors d'implanter un algorithme de type *descente de gradient*, consistant à faire évoluer le champ suivant la direction opposée au

³¹ Nous allons bien sûr nous limiter à l'étude du problème de la restauration d'images avec prise en considération des discontinuités. Il existe cependant un grand nombre de réalisations expérimentales capables de traiter un large éventail de problèmes de vision bas-niveau (traitements spatiaux et/ou spatio-temporels de l'image), allant depuis des véritables rétines artificielles neuromimétiques jusqu'à des simples éléments de capture "améliorés" (*smart sensors*). Ces réalisations font appel à de très diverses technologies micro-électroniques, combinés parfois sur un même prototype (technologies numérique CMOS, réseaux analogiques, composants optoélectroniques, etc.). Il s'agit d'un domaine dans lequel n'y a pas de "solution définitive" ni d'approche "standard", c'est pourquoi l'exposition des différentes idées et prototypes demanderait un chapitre à part entière. Nous renvoyons le lecteur à l'exposé bref et clair dans [Mhani97].

gradient de l'énergie (pour un champ de bords l fixé). L'équation d'évolution du champ est donné alors par :

$$\forall s \in S, \tau \frac{\partial e(s)}{\partial t} = -\frac{\partial U}{\partial e(s)} = F(s,t)$$

(τ est une constante de temps dépendant du mécanisme de relaxation). Dans notre problème, la *force de rappel* s'écrit :

$$F(s,t) = -\frac{\partial U}{\partial e(s)} = -2\beta_o (e(s) - o(s)) - 4\beta_s \sum_{r \in V(s)} (e(s) - e(r))(1 - l_{sr})$$

Si le paramètre β_o est nul - c'est à dire s'il n'y a pas de rappel aux données - alors l'équation d'évolution est une équation de *diffusion* et le système relaxe jusqu'à ce que tout le champ ait le même niveau de gris ; si le terme de rappel n'est pas nul, le système évolue vers une solution qui approchera mieux l'observation, mais qui sera localement *lissée*³² par le deuxième terme de la force.

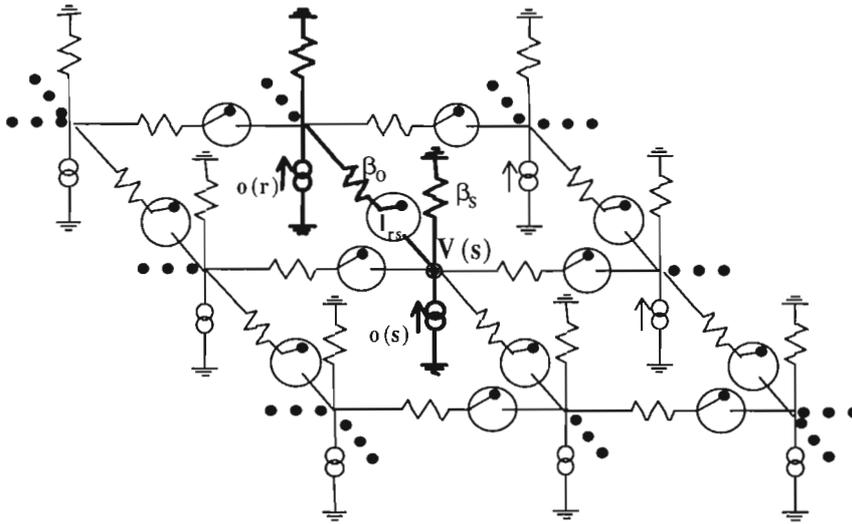


Fig.III.1 : réseau de résistances et d'interrupteurs pour la relaxation conjointe du champ à niveaux de gris et du champ de bords binaire (comparer avec la fig.I.6.)

L'idée de résoudre des problèmes variationnels par le biais du calcul analogique (pas nécessairement électrique) fut proposée très tôt [Poggio85], [Hopfie85], [Marro85b], [Koch89], etc.. Sans rentrer dans les détails (voir [Wyatt92], [Prévos92]) le calcul du minimum de la fonction d'énergie peut être réalisé au moyen d'un *réseau de résistances* bidimensionnel suivant la topologie du voisinage du champ de Markov (voir fig.III.1).

³² L'équation de *diffusion* avec "contrainte" (ou équation *Smoluchowski-Kramers* [Diu][Gidas86]) peut être utilisée en optimisation - comme méthode d'échantillonnage à une température donnée - pour des champs à valeurs continus.

Dans la figure, les résistances "horizontales" doivent être proportionnelles au paramètre β_0 , les résistances "verticales" au paramètre β_S . Il suffit alors d'injecter en chaque nœud s un courant externe proportionnel à la valeur de la contrainte locale $o(s)$. Dans ces conditions, le système relaxe "naturellement" (en une dizaine de nanosecondes! [Luthon95]) vers la solution escomptée, qui est la distribution des potentiels minimisant la puissance dissipée par le réseau (théorème de la dissipation minimale de Maxwell [Millar51]).

Une architecture prenant en compte le processus de bords (cf. Chap.I, §5.1.3) est facilement réalisée par adjonction d'un interrupteur en série sur chaque résistance reliant deux nœuds du réseau (représentés aussi dans la fig.III.1). Pour une configuration de bords *donnée*, la relaxation du champ continu est alors réalisée (quasi) instantanément. Inversement, la relaxation du champ de bords - i.e. la configuration "d'interrupteurs" - pour une configuration du champ continu fixé peut être réalisé en utilisant cette fois-ci la technique de recuit simulé (plus complexe à mettre en œuvre - voir plus loin). Cette procédure d'optimisation "mixte" (en partie stochastique et en partie déterministe) fut proposée par [Marro85b] et est connue dans la littérature sous le nom de *mixed-annealing*. Cependant, elle présente l'inconvénient de rompre l'équilibre thermodynamique du champ couplé (le champ continu est traité de façon déterministe, et donc à température *nulle*, alors que le processus de bords est traité de façon stochastique selon une procédure de recuit simulé, à température non nulle). Dans [Prévos95] il est proposé une architecture optoélectronique permettant de relaxer conjointement le champ continu et le champ de bords à *la même température* grâce à l'utilisation d'un algorithme de mixed-annealing modifié, appelé algorithme de *relaxation quasi-statique* ou *RQS* (il "suffit" pour cela de rajouter des sources de courant aléatoires sur les nœuds, pour maintenir la température du champ continu et éviter qu'il ne "refroidisse" trop vite par rapport au champ de bords). La rétine artificielle proposée pour l'implantation du *mixed-annealing* ou du *RQS* se compose donc essentiellement de deux réseaux "superposés" : un réseau de résistances analogique et un réseau d'interrupteurs numériques. C'est dans le cadre de cette recherche que le circuit SPIE600 fut proposé et réalisé - SPIE600 correspond bien sûr au réseau d'interrupteurs ou bords binaires. Dans le prototype, le potentiel d'interaction entre les bords $\Psi_B(l)$ est élémentaire et se résume à l'expression :

$$\Psi_B(l) = \frac{1}{2} \sum_{s \in S'} \sum_{r' \in V_B(s')} (2.l_{s'} - 1) \cdot (2.l_{r'} - 1),$$

où V_B est un voisinage d'ordre 1 (aux 4 plus proches voisins) invariant par translation sur la grille des bords S' . Pour rendre le prototype plus intéressant (simulation du modèle d'Ising pour le Verre de Spin), SPIE600 a été conçu de sorte à pouvoir prendre en compte des configurations *quelconques* d'interactions *bipolaires* - attractives ou répulsives - entre les bords (la variable binaire codant la présence ou l'absence d'un bord est alors appelée "spin" par analogie avec le modèle du verre de spin - voir §6.1).

2.2 ...une machine de Boltzmann.

Une machine de Boltzmann [Hinton84][Aarts89] est essentiellement un réseau de Hopfield [Hopfield82] - réseau de neurones récurrent - dont l'évolution en chaque site (ou *neurone*) ne se réduit pas à un simple seuillage de la force locale (ou *potentiel neuronal*), mais suit une loi probabiliste de type *sigmoïde* sous le contrôle du paramètre de température T (loi du *Thermostat*, cf. Chap.II, §2.2.2)³³. La dynamique du réseau de neurones est donc *stochastique* ; on a vu au chapitre précédent qu'un tel comportement augmente les chances d'atteindre un minimum absolu de l'énergie (ou *attracteur* du réseau récurrent). En d'autres termes, la machine de Boltzmann (pendant la phase associative du réseau) n'est autre chose que l'implantation matérielle de la procédure de recuit simulé décrite au chapitre précédent pour une fonction d'énergie à variables binaires. Tout comme pour la procédure de recuit simulé, on distingue alors deux sortes de machines de Boltzmann : les machines *séquentielles* et les machines *parallèles* (synchrones ou asynchrones). Les machines de Boltzmann parallèles se prêtent bien à l'implantation dans des machines de type SIMD sur circuit VLSI (appelés parfois *machines-tableaux* ou circuits à *architecture cellulaire*). Dans cette optique, il est clair que SPIE600 est un prototype simple de machine de Boltzmann, pouvant coder uniquement des *poids synaptiques bipolaires* et à interconnexions *locales* [Dupre95b]. Remarquons que - en toute généralité - les poids des interconnexions entre les différents processeurs élémentaires de la machine de Boltzmann (i.e., les *neurones stochastiques*) sont réels et pas nécessairement invariants par "translation" dans le réseau.

L'implantation d'une machine de Boltzmann à interconnexions non invariantes par translation est un problème pratique ayant des applications potentielles très intéressantes. En effet, il est facile de montrer que la machine de Boltzmann permet de traiter le problème du *voyageur du commerce* (voir ci-dessous) ; or, celui-ci est un problème NP-complet, ce qui veut dire que tout problème difficile de la classe NP peut être "traduit" (en un temps polynomial) en une instance du problème du voyageur de commerce³⁴. La

³³ En utilisant la terminologie "neuronal", on dit que la fonction d'activation du neurone est "probabiliste" (le seuil d'activation est une variable aléatoire qui suit la loi sigmoïde à température T).

³⁴ Un problème appartient à la classe P s'il est *soluble* en un temps polynomial par rapport la taille de données (l'ensemble des distances entre les villes dans l'exemple du voyageur de commerce) ; un problème est dit NP si une solution proposée (i.e. un chemin de parcours des villes dans l'exemple) peut être *reconnue ou rejetée* comme étant une solution du problème en un temps polynomial par rapport aux données. Bien sûr, tous les problèmes P sont NP ; par contre on ne sait pas si l'inverse est vraie : à ce jour on ne sait pas si les problèmes "inextricables" existent véritablement (pour ce genre de problème *il ne peut pas exister* d'algorithme permettant de trouver une solution optimale en un temps polynomial par rapport à la taille des données) ou si les problèmes inextricables rencontrés ne le sont qu'en apparence (on ne *connaît* pas d'algorithme efficaces). C'est un problème ouvert et fondamental. Quoi qu'il en soit, dans la pratique on rencontre souvent des problèmes *difficiles* (sont-ils inextricables?) ; il s'avère en général que les seuls algorithmes connus permettant de trouver à coup sûr la bonne solution du problème prennent un temps qui croît *exponentiellement* par rapport à la taille des données (recherche exhaustive dans l'ensemble des parcours possibles du voyageur de commerce ou dans l'ensemble $\Omega_E=2^N$ des configurations possibles du champ de primitives dans le cas du traitement d'images - cf. Chap.II. §1).

machine de Boltzmann permet donc de traiter n'importe quel problème appartenant à la classe NP - à condition de trouver le codage adéquat (c.à.d. les *poids synaptiques* correspondants au problème³⁵).

Exemple de machine de Boltzmann adaptée au problème du voyageur de commerce. Plusieurs fonctions d'énergie ont été proposées pour modéliser le problème du voyageur de commerce ; en voici une due à [Aarts88]:

$$U(e) = \frac{1}{2} \sum_{a,b,k} d_{ab} e_{ak} (e_{b,k+1} + e_{b,k-1}) + \frac{\lambda}{2} \left[\sum_k \left(1 - \sum_a e_{ak} \right)^2 + \sum_a \left(1 - \sum_k e_{ak} \right)^2 \right]$$

où d_{ab} représente la distance entre les "villes" a et b et e_{ak} est une variable binaire qui prend la valeur 1 si la ville a correspond à l'étape k du voyage, (et 0 sinon). Le parcours proposé est ainsi codé sur le "champ d'étiquettes" e . On retrouve dans le premier terme de la somme la distance totale du parcours proposé ; la machine de Boltzmann se doit de minimiser cette distance en tenant compte de deux *contraintes* (termes suivants de la somme) : la première traduit le fait que chaque ville doit apparaître une seule fois dans le parcours ; la deuxième traduit la nécessité (plus abstraite) que chaque étape du voyage se fasse uniquement dans une ville à la fois. Une "bonne" solution du problème peut parfois violer ces contraintes (et devenir alors un parcours "impossible") ; il a été proposé des algorithmes qui respectent ces contraintes de façon *stricte* (c'est la dynamique d'évolution qui change, impliquant par exemple une mise à jour du "champ" par permutation de sites - voir remarque 2 au Chap.II, §2.2.2 et commentaire final dans l'Ann.B, §2).

3. Architectures optoélectroniques pour la machine de Boltzmann.

Plusieurs architectures optoélectroniques ont été proposées pour implanter des réseaux de neurones formels (aussi bien récurrents que propagatifs). Trois éléments doivent être représentés dans l'architecture du prototype, à savoir : le tableau de neurones formels (chaque neurone comporte essentiellement des éléments de transduction optique/électronique, un point mémoire binaire et un comparateur), la topologie d'interconnexion (ou *matrice des connexions*) et les poids correspondant aux interconnexions (ou *matrice des poids synaptiques*). Les prototypes pour lesquels ces éléments sont tous les trois optiques sont plutôt rares ; par contre les réalisations dans lesquelles les interconnexions et les poids synaptiques sont réalisés par des moyens

³⁵ En restauration d'images, la définition du voisinage de Markov est une façon de définir *explicitement* le "codage" (poids synaptiques correspondants au problème) ; on remarquera toutefois que dans la plupart des cas le "codage" direct en termes de poids synaptiques est impossible : c'est le cas pour presque tous les problèmes de *reconnaissance de formes* par exemple, problèmes pour lesquels on ne connaît qu'un nombre plus ou moins grand de *vecteurs d'apprentissage*. Dans ces cas il est encore possible de "trouver" un codage grâce à des méthodes d'*apprentissage* du réseau (supervisé, non supervisé ou automatique). En fait, c'est là le principal intérêt des réseaux des neurones (appelé parfois capacité de *généralisation* du réseau).

optiques et le tableau de neurones par des moyens électroniques sont bien représentées dans la littérature. Il s'opère naturellement une classification des différentes architectures optiques (pour une étude comparative, voir [Duvill96]) :

- architecture optique à *synapses localisées*. Les interconnexions (et les poids associés) sont accessibles de façon indépendante les unes des autres, car elles sont codées dans un modulateur spatial de lumière. L'architecture typique est connue sous le nom de "vecteur×matrice" (on distingue notamment une architecture macro-optique dans laquelle les poids synaptiques sont localisés dans un tableau 2D, et l'architecture plus compacte à microlentilles appelée parfois "matrice×tenseur"). Ces architectures optoélectroniques bénéficient d'une quinzaine d'années d'expérimentation, et s'apparentent essentiellement au dispositif pionnier "vecteur×matrice" de CalTech dans le domaine des réseaux de neurones optoélectroniques [Farhat85].
- architectures optiques à *synapses non-localisées*. Ce type d'architecture est essentiellement représenté par les dispositifs à interconnexions holographiques, pour lesquels on ne peut modifier les poids qu'en recalculant ou en re-enregistrant l'hologramme d'interconnexion. Ces architectures bénéficient de plus d'une trentaine d'années d'expérimentation sur les mémoires associatives holographiques. On distingue notamment les dispositifs utilisant des hologrammes minces (par exemple [Lalann87][Cheval89]) de ceux utilisant des hologrammes en volume (par exemple [Frauel99]).

Dans le cas des réseaux de neurones *récurrents*, cette description peut être complétée par une autre : il est en effet utile de distinguer les systèmes pour lesquels la détection et l'émission des signaux optiques circulant dans la *boucle de rétroaction* optique du réseau récurrent a lieu sur des supports matériels *distincts* (cf. Chap.VII, §2.3.2) reliés par un montage optique plus ou moins compact, de ceux pour lesquels la boucle de rétroaction est *monolithique* (la détection et l'émission des signaux optiques est faite sur le même support optoélectronique).

Dans le cas où la boucle de rétroaction n'est pas monolithique, il est probable qu'une partie du traitement (par exemple la conversion des données au format présentable sur le circuit optique, mais aussi les opérations non linéaires réalisées par le neurone formel) soit réalisée par des moyens électroniques ou informatiques (en l'occurrence sur un ordinateur hôte). Le montage optique pour la boucle de rétroaction n'est pas toujours encombrant : citons par exemple la *neuropuce optoélectronique* à synapses localisées de Mitsubishi, (paire de puces émettrice et réceptrice en technologie GaAs, réunies par *flip-chip bonding*) servant uniquement à coder la matrice des poids synaptiques [Ohta90], et la puce développée par AT&T [Rietma89] où les poids synaptiques peuvent être contrôlés optiquement (il s'agit d'une matrice d'électrodes croisées connectées par du silicium amorphe hydrogéné, sensible à la lumière).

Pour finir, remarquons que l'implantation complète d'une machine de Boltzmann sur circuit *électronique* VLSI ne s'est pas faite attendre avec, par exemple, la *neuropuce* de Mitsubishi [Hirai92] comprenant 336 neurones formels binaires *totale*ment interconnectés par des poids multivalués codés sur 5 bits, et capable d'apprentissage et de généralisation. Des applications concrètes ont aussi vu le jour, par exemple le prototype de rétine artificielle bi-puce (puce pour acquisition et codage et neuropuce parallèle réalisant 100Mops) pour la reconnaissance de caractères imprimés de [Masa99] capable de reconnaître 1000 caractères à la seconde. Une éventuelle réalisation optoélectronique doit donc d'ores et déjà relever de sérieux défis.

4. Une machine de Boltzmann simplifiée pour le traitement d'images : le PPOS.

Néanmoins, l'implantation matérielle d'un réseau de neurones sur circuit OE-VLSI à rétroaction optique monolithique devrait conduire à des réalisations très performantes (*cf.* Chap.VII) ; à ce jour, il n'existe pas -à notre connaissance- de véritable prototype de machine de Boltzmann parallèle optoélectronique à *interconnexions reconfigurables* et *totale*ment interconnectée - à synapses localisées ou non - réalisé sur circuit OEVLSI à entrées et sorties optiques et fonctionnant en rétroaction optique monolithique. Quelques prototypes expérimentaux existent pourtant qui vont dans cette direction. La plupart sont des composants à cristaux liquides sur circuit intégré (prototypes de l'Université de Boulder [Bar95][Mao93]). Dans [Lin93] (équipe de Psaltis à Caltech) est étudiée une matrice optoélectronique de neurones *stochastiques* en substrat GaAs à entrées et sorties optiques (diodes LED et phototransistors), pour laquelle est proposée une architecture d'interconnexion vecteur×matrice ou holographique.

La réalisation de réseaux de neurones formels sur circuit VLSI pourrait bénéficier des nouvelles technologies optoélectroniques très performantes (matrices de VCSEL et MQW en particulier) développées pour la transmission de données entre processeurs ou cartes (*cf.* Chap.VII). Les chapitres V, VI et VII de cette thèse peuvent être vus comme une réflexion sur le sujet, en se rapportant au cas particulier de la machine de Boltzmann appliquée aux problèmes de vision bas niveau. En effet, les problèmes de vision bas niveau qui nous intéressent (débruitage d'images binaires, détection du mouvement) nécessitent un voisinage d'interconnexion *limité* et surtout, *invariant par translation* (*shift-invariant*), c'est pourquoi il est intéressant d'étudier la possibilité d'interconnecter les neurones ou PEs par *voie optique*, l'architecture de choix étant à synapses *non-localisées* (*cf.* Chap.V). Comme on le verra par la suite, on peut également profiter de l'optique pour contrôler l'évolution stochastique du processeur. Nous appellerons ce type particulier de machine de Boltzmann dédié au traitement d'images bas niveau, *processeur parallèle optoélectronique stochastique* (PPOS en abrégé). Au Chap.VII on étudiera des

circuits OE-VLSI à entrées et *sorties* optiques permettant de réaliser une boucle de rétroaction optique *monolithique*.

En résumé, la puce SPIE600 -conçue initialement pour intégrer une véritable rétine artificielle analogique-digitale- constitue donc un premier prototype de PPOS à interconnexions électroniques (le bouclage est monolithique, mais pas optique -les entrées optiques de SPIE600 servent uniquement au contrôle de l'évolution stochastique et à l'introduction de l'information à traiter). Dans le chapitre V nous allons démontrer la possibilité d'étendre le système de voisinage de SPIE600 grâce à des interconnexions optiques (montage de convolution 4-f comportant un SLM et un hologramme de Damman : la rétroaction n'est plus *monolithique*).

5. Présentation du processeur SPIE600

La puce SPIE600 (CMOS 1 μ m), conçue dans le cadre de la thèse de A.Dupret [Dupre95a], est une matrice de processeurs élémentaires (PEs) à mémoire binaire reliés entre eux par des interconnexions électriques aux 4 plus proches voisins (voisinage d'ordre 1). Elle comporte 600 PE, dont 576 fonctionnels rangés sur une matrice carrée 24x24 (la dernière ligne est une ligne de calibration, permettant de tester indépendamment certains éléments de la cellule élémentaire). Chacune des rangées de la matrice comporte un registre à décalage qui permet de lire et charger de proche en proche la configuration de "spins" (i.e. le champ d'étiquettes³⁶) ainsi que les poids des interactions sur la puce (*fig.II.1*). Les différents PEs évoluent de façon quasi-*parallèle*, pour ne pas rompre l'équilibre thermodynamique lors de la relaxation du champ de spins (*cf.* Chap.II, §2.2.5). Deux horloges indépendantes fournissent le signal de mise à jour pour les sites noirs et blancs respectivement (coloriage binaire).

Le circuit électronique ne comporte pas de générateur (pseudo) aléatoire numérique pour la mise à jour stochastique du champ d'étiquettes (en suivant la loi sigmoïde du Thermostat, *cf.* Chap.II, §2.2.2). L'état (binaire) de chaque PE évolue de façon *déterministe* en fonction du voisinage, par simple *seuillage* de la force locale. Celle-ci est calculé électroniquement, grâce à une simple addition analogique des courants électriques. Par contre, le seuil est déterminé *optiquement*, grâce à une paire de photodétecteurs (voir *fig.III.2*).

Tout l'intérêt du circuit SPIE600 réside dans le fait que son comportement peut être rendu facilement *stochastique* : en effet, puisque chaque PE possède un circuit analogique de détection différentielle pour la détermination du seuil, le fonctionnement de chaque processeur élémentaire peut être piloté par un signal aléatoire optique aux

³⁶ L'étiquette *binaire* peut prendre deux valeurs au site s : $e(s) \in \{0,1\}$. La configuration de spins correspondant au champ d'étiquettes est *bipolaire* (± 1) : la valeur du spin au site s est donné alors par $(2.e(s)-1) \in \{-1,1\}$.

propriétés statistiques contrôlées provenant d'un générateur de nombres aléatoires optique basé sur le phénomène de speckle laser [Lalan95a]. SPIE600 est donc un processeur *parallèle* à la fois *numérique* (points mémoire binaires) et *analogique* (détermination du seuil et calcul de la force locale).

5.1 Architecture électronique globale

Comme dit plus haut, SPIE600 est capable de représenter des interactions locales aussi bien attractives que *répulsives* entre PEs, c'est pourquoi il est possible de simuler le comportement du Verre de Spin (cf §6.1). Pour cela, deux points mémoire supplémentaires codant les interactions bipolaires (± 1) ont été ajoutés par cellule (notation : J_r avec $r \in V(s) = \{r_N, r_S, r_E, r_O\}$). La fig.III.2 ci-dessous représente l'architecture électronique de l'ensemble de la matrice ainsi que celle d'un PE individuel. La préparation des interactions avant toute opération ainsi que la récupération du résultat est faite en série sur les colonnes au moyen d'un bus parallèle 24 bits sur les lignes (boucle de registres à décalage).

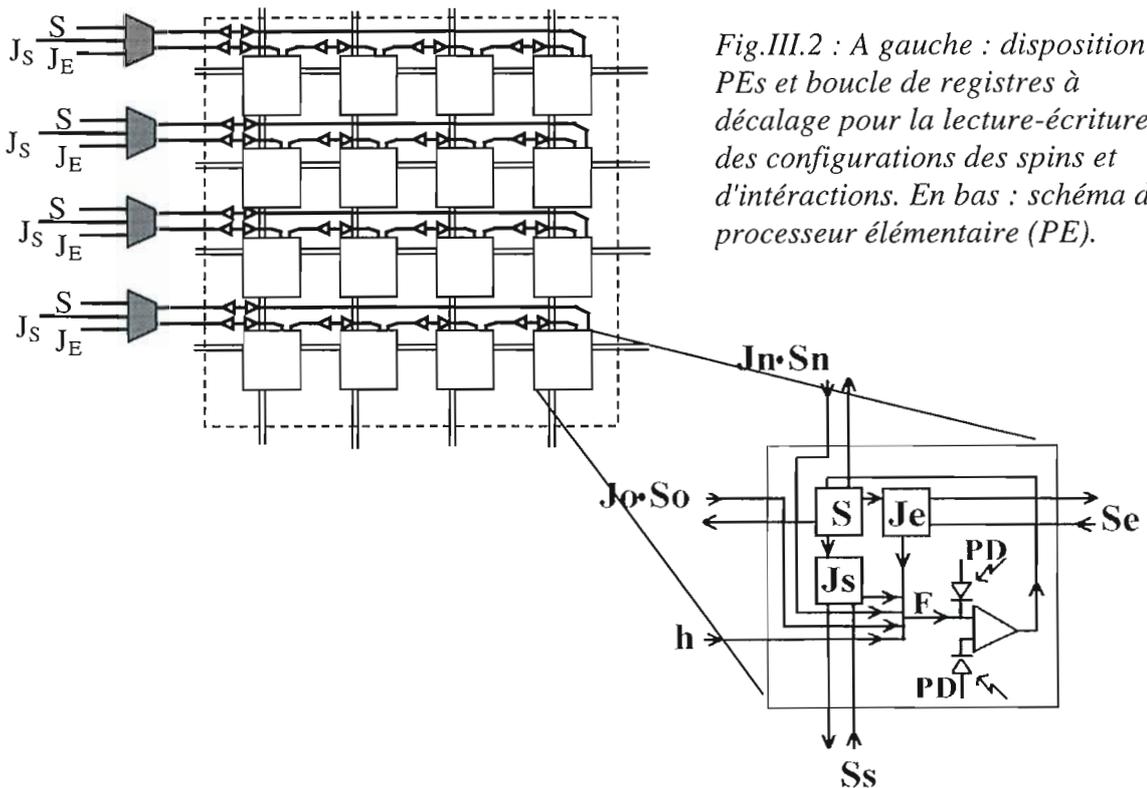


Fig.III.2 : A gauche : disposition des PEs et boucle de registres à décalage pour la lecture-écriture des configurations des spins et d'interactions. En bas : schéma du processeur élémentaire (PE).

5.2 Le processeur élémentaire.

Chaque cellule ou PE comporte 112 transistors et occupe une surface de $215 \times 216 \mu\text{m}^2$. Le calcul de la force locale $F(s)$ est *analogique* (sommateur de courant) et s'effectue par addition de quatre courants bipolaires (et de valeur absolue constante) $\pm I_S$ dont le signe

est déterminé par le produit des "constantes de couplage" J_N, J_S, J_E, J_O (2 points mémoires binaires par PE, voir *fig.III.2*) avec la valeur des spins voisins correspondants (un point mémoire binaire par PE). A cette quantité vient s'ajouter un courant I_h correspondant à un champ extérieur uniforme h . Plus exactement, le *courant* correspondant à la force locale au site s est donné par :

$$I_F(s) = I_S \cdot \sum_{r \in V(s)} J_{rs} (2e(r) - 1) + I_h$$

Ce courant est effectivement mesurable sur un unique site de test, (sortie SPIN_OUT de SPIE600). Le contrôle de la valeur des courants I_h et I_S est fait directement, grâce aux plots de la puce SPIE600 h_pol et f_pol respectivement (de résistance caractéristique de l'ordre du M Ω). Les courants sont de l'ordre du μ A (les *tensions* de commande sont alors de l'ordre de la centaine de mV). Une paire de photodétecteurs (PD dans la figure) - plus précisément des phototransistors - permet la conversion des éclaircissements aléatoires optiques en un photocourant aléatoire différentiel "de seuillage" I_d (voir paragraphe suivant). Enfin, un comparateur (représenté par un triangle dans la figure) permet la mise à jour de l'état du PE par comparaison du courant I_F avec le photocourant différentiel I_d . Pour plus de détails quand aux cellules analogiques et numériques, voir [Dupre95a].

5.3 Le générateur de nombres aléatoires optique (RNG)

5.3.1 Introduction.

La génération de nombres aléatoires statistiquement indépendants et ayant une loi de probabilité donnée est, comme on l'a vu au Chap.II, un élément clé pour l'implantation de la procédure d'échantillonnage du recuit simulé. Par ailleurs, l'utilisation d'un grand nombre de processeurs élémentaires travaillant en parallèle exige la génération d'une quantité considérable de nombres aléatoires indépendants dans le temps *et* dans l'espace. La génération de telles variables aléatoires indépendantes par des moyens purement électroniques est possible [Belhai90], mais soulève le problème de l'encombrement de la surface de la puce et, de ce fait, limite la densité de processeurs élémentaires et/ou la qualité des séries temporelles pseudo-aléatoires -si l'on se limite à l'utilisation des fonctions de l'électronique numériques.

La génération *optique* de nombres aléatoires offre plusieurs avantages : (a) tout d'abord les quantités aléatoires proviennent d'un processus stochastique réel (bruit de photon [Morris86] ou champ de speckle [Marron86] par exemple) ; (b) le signal aléatoire est ensuite véhiculé en espace libre et distribué *en parallèle* sur tous les processeurs de la matrice ; (c) l'interface optoélectronique au sein de chaque PE se limite à des éléments de photodétection et mise en forme du signal (amplification et/ou seuillage) ; enfin, les paramètres statistiques des variables aléatoires dépendent de l'éclaircissement et sont modifiables instantanément sur l'ensemble de la matrice.

Les points (b) et (c) peuvent constituer aussi des désavantages : les défauts du système d'imagerie optique et la dispersion (plus ou moins inévitable) des caractéristiques des photodétecteurs induisent des variations locales des propriétés statistiques des nombres aléatoires sur chaque site (voir Chap.VI, §1.3 et fig.VI.8). Néanmoins, l'approche a été validée dans des travaux précédents dans notre laboratoire dans le cas de la génération massive de courants aléatoires de nature gaussienne [Lalan95a] et sigmoïde [Prémon94], ayant de très bonnes qualités spatiales et temporelles.

5.3.2 Génération de la loi du Thermostat à partir du speckle

L'algorithme de recuit simulé parallèle utilise, pour la mise en équilibre du champ aléatoire à chaque palier de température, la procédure d'échantillonnage de Gibbs. Celle-ci est équivalente dans le cas de variables binaires à la loi du *Thermostat* (cf. Chap.II, §2.2.2) ; la loi d'évolution que chaque PE doit être en mesure de suivre pour la mise à jour de son étiquette est donc la probabilité sigmoïde (Chap.II, §2.2.2, fig.II.1.2) :

$$\Pr[e(s) = 1] = \frac{1}{1 + e^{-\frac{F(s)}{T}}}$$

Cette probabilité de transition dépend du paramètre de contrôle T (appelé température) ainsi que de la différence d'énergie ou force locale $F(s) = -\Delta U(s)$ entre l'état du champ courant E avec localement $e(s)=1$ et l'état courant avec localement $e(s)=0$. Le processeur élémentaire doit générer une sortie (binaire dans le cas qui nous intéresse) selon cette distribution. L'élément de base pour réaliser l'échantillonneur de Gibbs est schématisé ci-dessous (fig.III.3) : il s'agit de ce que l'on pourrait l'appeler un *comparateur stochastique*, le comparateur classique étant un cas particulier quand la température est nulle. Nous allons montrer par la suite - points a), b) et c) - comment on peut s'y prendre pour transformer un comparateur classique en un comparateur stochastique grâce à l'introduction d'un courant aléatoire photoinduit, ayant les bonnes propriétés statistiques.

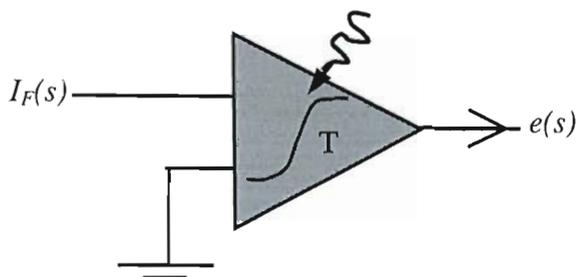


Fig.III.3 : Comparateur "stochastique"(contrôlé optiquement) permettant la mise à jour des sites selon la loi du Thermostat à température T .

a) Statistiques du speckle.

Le *speckle* est un phénomène d'interférences découlant de l'introduction de déphasages aléatoires dans un front d'onde cohérent [Goodman85]. La réflexion d'un faisceau cohérent sur une surface rugueuse ou sa traversée par un milieu présentant des granularités dont la dimension caractéristique est de l'ordre de la longueur d'onde, produisent une figure de speckle caractéristique contenant un grand nombre d'éclairement aléatoires. La distribution de probabilité de l'éclairement I reçue par un photodétecteur placé dans le champ de speckle est en première approximation donné par la loi de distribution gamma [Dainty84] :

$$\text{Pr}_M(I = i) = \frac{1}{\Gamma(M)} \left(\frac{M}{\langle I \rangle} \right)^M i^{M-1} \cdot \exp\left[-\frac{M}{\langle I \rangle} i \right], \quad (i \geq 0).$$

Dans cette expression, le paramètre M désigne le nombre de degrés de liberté associé au champ d'intensité observé et il est donné par le carré du rapport signal sur bruit du champ de speckle : $M = \left(\frac{\langle I \rangle}{\sigma_I} \right)^2$, où $\langle I \rangle$ dénote l'intensité moyenne et σ_I l'écart type de

l'éclairement. En accord avec le théorème de la limite centrale, la densité de probabilité gamma de speckle tend vers une distribution gaussienne lorsque le nombre de degrés de liberté M tend vers l'infini ; le paramètre M peut aussi être interprété comme le nombre moyen de grains de speckle incidents sur le photodétecteur (voire plus loin).

b) Détection différentielle du speckle

L'éclairement aléatoire est converti en photocourant aléatoire grâce à des phototransistors en technologie CMOS sur chaque processeur élémentaire de la rétine (voir *fig.III.4* ci-dessous). A partir d'une détection différentielle sur deux photodétecteurs (suffisamment éloignés pour que les statistiques de speckle soient indépendantes) il est possible d'obtenir une distribution de photocourants presque gaussienne et *centrée* de très bonne qualité.

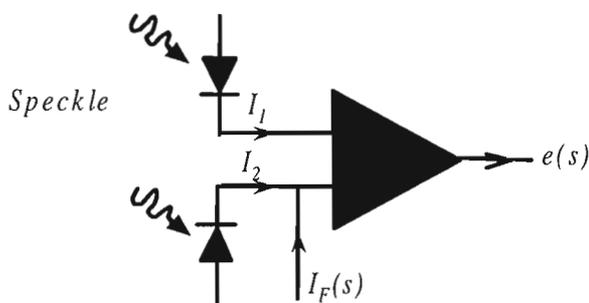


Fig.III.4 : Comparateur classique à deux entrées et détection différentielle de speckle. Le courant photoinduit détermine le seuil du comparateur.

En effet, si l'on suppose que la réponse des photodétecteurs est linéaire et que les éclairements sont indépendants, la densité de probabilité $\text{Pr}(I_d)$ associé à la différence I_d

des photocourants induits I_1 et I_2 est donné par l'autocorrélation de la loi de probabilité gamma de speckle :

$$\Pr(I_d = i) = \Pr(I_1 = i) * \Pr(I_2 = i) = \int_{-\infty}^{+\infty} \Pr(I_1 = x) \cdot \Pr(I_1 = x - i) dx$$

Cette intégrale peut être évaluée analytiquement ; il s'ensuit [Lalan95a] que $\Pr(I_d)$ est quasiment gaussienne même pour des faibles valeurs de M (écart maximal toujours inférieur à 1% pour $M < 5$), ce qui n'était pas le cas pour une détection directe de l'intensité de speckle. De plus, la loi est automatiquement centrée, quel que soit le site considéré ; enfin, l'écart type de la distribution gaussienne ajustée est approximativement donné par :

$$\sigma_{I_d} = \sqrt{2} \sigma_I \approx \sqrt{\frac{2}{M}} \langle I \rangle$$

c) Seuillage de la gaussienne et loi du Thermostat.

La probabilité que la sortie du comparateur de la fig.III.4 soit active est donnée par la probabilité que le courant $I_F(s)$ représentant la force locale $F(s)$ soit inférieur au photocourant différentiel I_d ; autrement dit, la probabilité en question est donnée par la fonction de répartition de la loi différentielle (quasiment gaussienne) :

$$\Pr(e(s) = 1) = \Pr(I_d < I_F(s)) = \int_{-\infty}^{I_F(s)} \Pr(I_d = i) di \approx \text{Erf}_{\sigma_{I_d}} [I_F(s)]$$

Or, on peut montrer que la fonction de répartition d'une gaussienne (fonction noté *Erf* ci-dessus) ne diffère pas plus de 1% de la loi sigmoïde recherchée [Lalann90] :

$$\text{Erf}_{\sigma_{I_d}} [I_F(s)] \approx \frac{1}{1 + \exp\left[-\frac{I_F(s)}{T}\right]}$$

Pour des valeurs élevées de M , la température T de la loi sigmoïde peut être trouvée analytiquement en posant que sa dérivée à l'origine soit égale à la pente d'une loi normale ; on a alors :

$$T = \sqrt{\frac{\pi}{4}} \sigma_{I_d} = \sqrt{\frac{\pi}{4M}} \langle I \rangle$$

Plus généralement, la dépendance de la température avec l'écart type σ_{I_d} de la gaussienne peut être estimée en ajustant la loi de probabilité sigmoïde avec la loi de répartition $\Pr(I_d < i)$; on montre alors que ces quantités sont proportionnelles et que la constante de proportionnalité α dépend du paramètre M et tend vers 0,94 quand M tend vers l'infini [Prémon99] :

$$T = \alpha(M) \cdot \sqrt{\frac{\pi}{4M}} \langle I \rangle, \text{ avec } \lim_{M \rightarrow \infty} \alpha(M) \approx 0,94$$

On retiendra que la température algorithmique est proportionnelle à l'intensité de l'éclairement, et que le facteur de proportionnalité dépend du nombre de degrés de liberté M du speckle.

5.3.3 Montage générateur de speckle : paramètres $\langle I \rangle$ et M .

Le montage expérimental pour générer le champ de speckle est schématisé dans la fig.III.5. Le speckle est produit par l'interférence aléatoire des différents modes de propagation dans une fibre optique multimodale. La diode laser utilisée est une SDL-5412 (Spectra-Physics) de 100mW de puissance maximale, émettant à 847 nm. L'ordinateur hôte contrôle la puissance de la diode par le biais de la tension de commande V_d . Pour assurer l'excitation d'un grand nombre de modes de propagation, un diffuseur est placé devant la face d'entrée de la fibre. La face de sortie de la fibre est imagée ensuite sur la surface du circuit. Le générateur de speckle délivre une puissance optique maximale de 10mW/cm² (assurant une commutation *totale*ment aléatoire des PEs à la

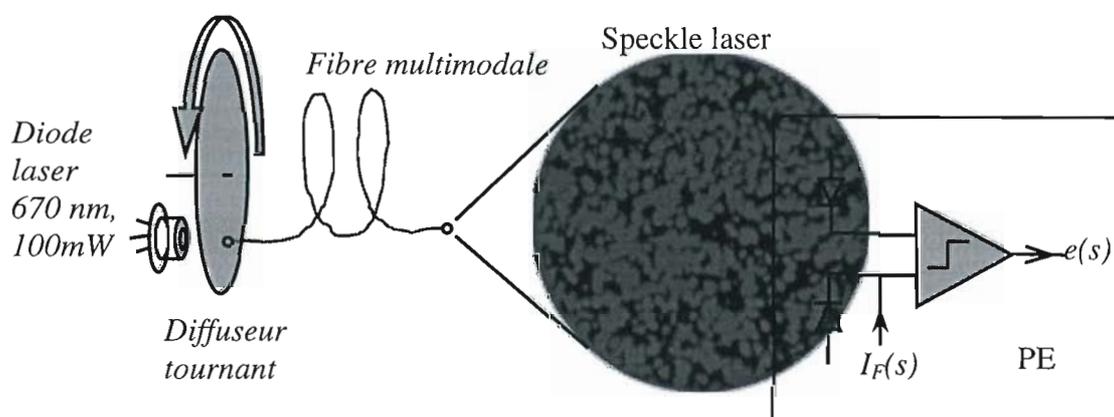


Fig.III.5 : Génération de la loi sigmoïde à partir de la détection différentielle de speckle. Le photocourant aléatoire produit par le speckle rend variable le seuil du comparateur "classique" : celui-ci réalise alors une très bonne approximation du comparateur "stochastique" de la fig.III.3.

fréquence maximale de 250kHz). La complexité du montage optique est comparable à celle d'un lecteur de disques compacts.

Paramètre $\langle I \rangle$: Une uniformité allant jusqu'à $\Delta I / \langle I \rangle = 6\%$ (ΔI représente les variations crête à crête de l'intensité moyenne) à la sortie de la fibre a pu être démontrée sur un tel montage [Lalan95b]. Dans le montage plus complexe qui sera étudié au Chap.V, et pour des raisons d'encombrement, le speckle n'est pas incident à angle droit sur la puce ; ceci induit un gradient spatial d'intensité qui se traduit par un gradient spatial de température algorithmique sur la puce (cf. Chap.IV, §1.3.2).

Paramètre M . Le nombre de degrés de liberté M (défini comme le carré du rapport signal sur bruit) est donné par le nombre de grains de speckle indépendants incidents sur chaque photodétecteur pendant la durée de détection. On a plus précisément $M = 2.M_s.M_t$, où M_s est le nombre maximal de grains de speckle présents

simultanément sur la surface du détecteur, et M_i indique le nombre de champs de speckle indépendants imagés sur le circuit pendant la durée de photodétection - le facteur 2 rend compte des deux polarisations perpendiculaires sortant de la fibre.

Les caractéristiques géométriques du montage optique imposent $M_s \approx 5$ (soit 5 grains de speckle sur la surface - $35 \mu\text{m}^2$ - du photodétecteur) ; l'indépendance spatiale des échantillons de speckle est assurée puisque la distance entre les fenêtres de détection ($128 \mu\text{m}$) reste très supérieure au diamètre du grain de speckle.

Par ailleurs, la décorrélation *temporelle* des figures de speckle est assurée par la vitesse de rotation du diffuseur [Jourda93]. Le dispositif de génération de speckle utilisée peut produire jusqu'à 100.000 champs de speckle indépendants par seconde [Lalan95a]. Cependant, la corrélation temporelle des *courants photoinduits* sur chaque site est conditionnée par la bande passante des phototransistors CMOS : le taux de génération de nombres aléatoires non *corrélés temporellement* est donc borné supérieurement - et cela indépendamment du montage optique (vitesse de rotation du diffuseur). Si l'on modélise le photo-capteur comme un filtre passe-bas de premier ordre de fréquence de coupure $f_c = 1/\tau$, alors le coefficient de corrélation entre deux échantillons aléatoires pris à un intervalle de temps Δt est égal à $\exp(-\Delta t/\tau)$ [Dupret95a]. Si l'on veut - par site - un taux de génération de nombres aléatoires temporellement indépendants (corrélation < 1%) supérieur à 50 kHz ($\Delta t = 20 \mu\text{s}$) nécessaire pour pouvoir réaliser 25 recuits sérieux par seconde (1000 itérations) avec un système de voisinage d'ordre 1 (coloriage binaire), alors la bande passante du circuit de photodétection doit être supérieure à 230 kHz. Les phototransistors associés à la détection différentielle ont une bande passante de l'ordre de 250kHz pour un éclairement de $10 \text{mW}/\text{cm}^2$ (photocourant de l'ordre de $1 \mu\text{A}$). Dans ces conditions, et si l'on fait travailler le circuit à la cadence de 230 kHz, on aura $M_t \approx 1$. Finalement, le nombre de degrés de liberté est de l'ordre de $M \approx 10$, et la détection différentielle seuillée de speckle donnera une bonne approximation de la loi de probabilité sigmoïde de mise à jour (voir Chap.VI, fig.IV.6).

5.4 Conclusion.

Le tableau suivant (tab.III.1) résume les principales caractéristiques de SPIE600. Il sera utile pour établir une comparaison des performances avec d'autres prototypes de processeurs optoélectroniques (voir Chap.VII, §2.4.2).

Caractéristiques globales de SPIE600	
Nombre de PEs fonctionnels	24x24 = 576
Surface totale de la puce	≈5x5 mm ²
Fréquence d'horloge maximale	250 kHz (*)
Puissance électrique consommée	10mW
Caractéristiques d'un PE	
Surface du PE	215x216 μm ²
Transistors par PE	112
Entrées optiques par PE	2
Sorties optiques par PE	0
Détection différentielle (phototransistors MOS cascodés)	
Taille fenêtre optique	35x35 μm ²
Séparation des fenêtres	128 μm (**)
Sensibilité	0,1-0,2 A/W
Gain	≈40
Bande passante	250 kHz (*)

Tab.III.1 : Caractéristiques optoélectroniques du prototype SPIE600.

(*) pour un photocourant induit par une source de speckle de 10mW/cm² sur la surface de la puce (de l'ordre du μA). (**) >50μm pour décorrelation spatiale [Dupret95a].

On remarquera enfin que n'importe quel élément de photodétection différentielle est susceptible d'être employé pour générer des lois sigmoïdes à partir du générateur de speckle. Dans [Prémon99] il est question de photothyristors optiques, composants expérimentaux qui ont donné lieu à des essais préliminaires ; au Chap.VII nous allons envisager la réalisation d'une matrice de PE stochastiques utilisant des diodes p-i-n à puits quantiques multiples qui pourraient être un prolongement au présent travail si l'état de la technologie le permet.

6. Le recuit simulé sur SPIE600.

La validation expérimentale de SPIE600 a fait l'objet d'un travail de thèse au laboratoire [Prévos95]. Nous allons illustrer ici le fonctionnement du circuit SPIE600 tel qu'il peut marcher sans autre montage optique que le générateur de speckle (correspondant à l'expérience du *verre de spin*). Une caractérisation plus approfondie du prototype sera faite au Chap.VI, qui, elle, prendra en compte les particularités d'un montage optique plus élaboré (Chap.V). Nous allons également démontrer en fin de chapitre la possibilité d'utiliser SPIE600 pour réaliser quelques traitements bas niveaux de l'image à *cadence vidéo* (en utilisant la l'interface électronique/informatique décrite dans l'Ann.D et mise en place pour le besoin d'expériences ultérieures décrites au Chap.VI).

6.1 Modèle du Verre de Spin

Le modèle ferromagnétique et anti-ferromagnétique d'Ising a été exposé au Chap.I,§5.1 : les constantes de couplage entre les spins J_{rs} prennent toutes la même valeur (+1 ou -1 respectivement). Le modèle d'Ising pour le *verre de spin* correspond à une configuration de constantes de couplage *aléatoire*, caractérisée par la probabilité $p \in [0,1]$ d'avoir $J_{rs}=+1$. Le modèle ferromagnétique correspond bien sûr au cas particulier $p=1$, et le modèle anti-ferromagnétique au cas $p=0$. Pour $p \in]0,1[$, on ne connaît pas d'algorithme déterministe permettant de trouver la configuration d'énergie minimale qui soit plus efficace que l'énumération exhaustive de toutes les configurations possibles de *spins* : l'utilisation du recuit simulé est alors justifiée. La fonction d'énergie pour le verre de spin est donnée par :

$$U(e) = \beta_s \sum_{s \in S} \sum_{r \in V(s)} (2e(s) - 1)(2e(r) - 1) + h \sum_{s \in S} (2.e(s) - 1)$$

La *fig.III.6* représente le *résultat final* du recuit sur SPIE600 dans les trois cas suivants (de gauche à droite dans la figure) : $p=0,8$ (*verre de spin*), $p=0$ (modèle anti-ferromagnétique) et $p=1$ (modèle ferromagnétique). On observe pour $p=0,8$ la formation de "domaines d'aimantation" en *fin* du recuit. Les cas $p=1$ et $p=0$ (pour lesquels les configurations optimales sont parfaitement connues *a priori*) vont nous permettre d'illustrer quelques particularités du recuit optoélectronique sur SPIE600. En effet, pour $p=1$ un "bon" recuit devrait conduire à une phase finale ferromagnétique *homogène*, (tantôt à +1, tantôt à -1 en fonction de la direction du champ extérieur h "brisant" la symétrie du système). Si le recuit est trop "rapide", alors le système pourra présenter des "domaines d'aimantation", minimisant *localement* l'énergie. En fait, en absence de champ ($h_{pol}=0$ V) la phase finale est presque toujours homogène, et cela même pour un recuit "rapide". Elle possède en outre le signe du *bord* virtuel (voir remarque ci-dessous). Celui-ci agit en brisant la symétrie du système et *force* le signe de la phase finale comme le ferait la présence du champ extérieur.

Remarque : problème des bords. Sur SPIE600, la grille S correspond bien sûr aux 576 sites de la matrice 24×24 , et non pas à un réseau bidimensionnel *infini*. Les bords imposent donc des conditions aux limites particulières. On pourrait rendre ces conditions *périodiques* en interconnectant les PEs des bords opposés de la puce, la topologie torique ainsi créée ne privilégiant aucune position particulière sur la grille. Cependant, le circuit SPIE600 fonctionne en créant un bord "virtuel" de PEs dont l'état est *figé* pendant le recuit (à $+1$ ou à -1). Le résultat du recuit dépend alors fortement du signe de cette "couronne" (on définit d'ailleurs la température critique du système comme la température maximale pour laquelle les conditions aux bords sont "ressenties" - à l'équilibre thermodynamique - au centre d'une grille carrée quelle que soit la taille de celle-ci [Marro85b]). Malheureusement la possibilité de *couper* les interactions électriques avec ce bord virtuel, - minimisant ainsi l'influence des bords - n'a pas été envisagée. Cependant, cette situation correspond naturellement à celle du montage à interconnexions optiques que l'on étudiera au Chap.V.

Pour $p=0$ (tout comme pour $p=1$) il existe deux configurations d'énergie minimale, "diamétralement opposées" dans l'espace des états (i.e. pour passer de l'une à l'autre, il faut inverser *la totalité* des spins). Néanmoins, dans le cas $p=0$, ni le champ extérieur ni les conditions aux bords ne peuvent privilégier aucun de ces deux minimums énergétiques *d'aimantation nulle*. Il est donc vraisemblable que, si le recuit est opéré un peu trop rapidement, la grille S soit partagée spatialement en domaines correspondant chacun à une configuration d'énergie minimale (damier et damier inverse). C'est ce que l'on peut effectivement observer dans la *fig.III.6-(b)*.

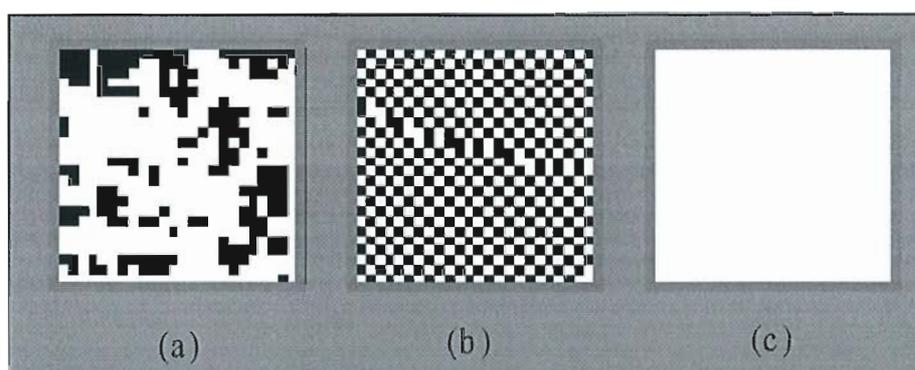


Fig.III.6 : Résultat du recuit pour le modèle d'Ising bidimensionnel (expérience sur SPIE600). (a) Modèle du Verre de Spin avec 80% des couplages à $+1$ (le reste à -1). (b) Modèle anti-ferromagnétique (constantes de couplage négatives). (c) Modèle ferromagnétique (constantes de couplage positives).

L'exemple du verre de spin est étudié plus en détail dans [Prévos95] ; nous aurons cependant l'occasion de reproduire quelques tests validant le prototype en nous appuyant sur l'exemple du *débruitage d'images binaires*.

6.2 Débruitage d'images binaires à cadence vidéo.

Le circuit SPIE600 à *interconnexions électriques* peut fonctionner comme une rétine artificielle pour le débruitage d'images binaires à *cadence vidéo* [Cassin98]. Il suffit pour cela de projeter l'image à traiter sur la surface de la puce (codée en mode double-voie sur chaque photodétecteur différentiel³⁷). Rappelons que la fonction d'énergie à minimiser est dans ce cas :

$$U(e, o) = -\beta_s \sum_{s \in S} \sum_{r \in V(s)} (2e(s) - 1)(2e(r) - 1) - \beta_o \sum_{s \in S} (2.o(s) - 1)(2.e(s) - 1),$$

où $o(s)$, $s \in S$ représente l'image binaire à débruiter (cf. Chap.I, §5.1.1). Celle-ci doit donc être codée en double-voie pour pouvoir représenter la quantité bipolaire $2.o(s)-1$ sur chaque paire de photodétecteurs (cf. fig.V.11).

6.2.2 Test d'échantillonnage avec contrainte aux données optiques.

Comment vérifier que la procédure *d'échantillonnage* opérée sur la puce SPIE600 réalise effectivement la mise en équilibre thermodynamique du champ d'étiquettes, pour une température de speckle et une image binaire de contrainte données? En toute rigueur, il faudrait vérifier que la procédure de mise à jour tend à engendrer, au bout d'un certain temps, des configurations suivant la loi de distribution de Gibbs à la température T (Chap.II, §2.1) :

$$G_T(E = e) = \frac{1}{Z_T} \exp\left(-\frac{U(e)}{T}\right)$$

La vérification exacte de cette distribution à partir d'un large échantillon de configurations s'avère fondamentalement impossible puisque on n'a pas accès à la valeur de la fonction de partition Z_T (il faudrait sommer sur les 2^{576} configurations possibles du champ d'étiquettes pour l'obtenir) ; par ailleurs, il faudrait pouvoir vérifier cette loi sur les éléments de l'espace des configuration Ω_E (et il y en a 2^{576}). On peut tout au plus procéder à la vérification de quelques conditions *nécessaires* - mais pas suffisantes - pour l'obtention de la loi de Gibbs. Tout d'abord, on ne vérifiera pas directement la loi de Gibbs, mais seulement la distribution canonique d'un observable -par exemple l'énergie. On a dans ce cas :

$$\Pr[U = u]_T = C_T \Omega(u) \exp\left[-\frac{u}{T}\right],$$

où $\Omega(u)$ est une fonction indépendante de la température correspondant au degré de dégénérescence de l'énergie u (le spectre est discret), et C_T est une constante qui ne dépend *que* de la température. Ensuite, et puisque $\Omega(u)$ est également une fonction inconnue, nous allons nous en affranchir en travaillant à partir de deux distributions

³⁷ Le dispositif - basé sur des cristaux liquides ferroélectriques - qui permet de réaliser cette opération sera décrit en même temps que le montage optique plus élaboré du démonstrateur final à interconnexions optiques (montage simplifié (E) du Chap.V, §1.1).

correspondant à T_1 et T_2 ; en effet, si l'on suppose que l'expression analytique de la distribution engendrée par le processus d'échantillonnage ne varie pas avec la température³⁸, alors il est possible de s'affranchir de $\Omega(u)$ en calculant le rapport :

$$\frac{\Pr[U = u]_{T_1}}{\Pr[U = u]_{T_2}} = \frac{C_{T_1}}{C_{T_2}} \exp\left[\frac{(T_1 - T_2)}{T_1 T_2} u\right],$$

soit, en prenant le logarithme :

$$\ln \frac{\Pr[U = u]_{T_1}}{\Pr[U = u]_{T_2}} = a.u + b, \text{ avec } a = \frac{(T_1 - T_2)}{T_1 T_2} \text{ et } b = cte(T_1, T_2)$$

Pour deux puissances de speckle données (correspondantes à T_1 et T_2), SPIE600 peut générer rapidement un échantillon large de configurations du champ d'étiquettes ; si l'on calcule pour chacune d'elles l'énergie correspondante (il faut connaître au préalable β_S et β_O) il est alors possible de vérifier la *linéarité* de l'expression ci-dessus et comparer les pentes théorique et expérimentale. La figure *fig.III.7* montre que le logarithme du quotient des distributions canoniques est bien linéaire par rapport à l'énergie ; cependant il existe un écart considérable (de l'ordre de 10% ou plus) entre la pente théorique et expérimentale. Remarquons que le calcul de l'énergie pour une configuration donnée ainsi que la détermination de la température algorithmique dépendent de l'ajustement plus ou moins exact entre les paramètres théoriques et expérimentaux, ajustement qui est fait en *moyennant* sur les différents PEs (Chap.VI. §1.3). La dispersion des caractéristiques entre PEs et le *gradient* de température à la surface de la puce (*cf.* Chap.IV, §1.3.2) peuvent être à l'origine de cet écart. Quoi qu'il en soit, le recuit semble se comporter de façon satisfaisante comme on le verra au paragraphe suivant.

6.2.3 Courbes de recuit.

La relation entre les paramètres théoriques (température T et valeurs β_S et β_O pondérant chaque terme de la fonction d'énergie) et les grandeurs de contrôle expérimentaux correspondants (tension de commande de la diode laser de speckle Vd , tension sur le plot f_{pol} de SPIE600 et enfin courant I_A alimentant la diode laser du système d'imagerie pour la projection de l'image de contrainte – *cf.* Chap.V) est étudiée en détail au Chap.VI, §1.3. Le recuit est contrôlé depuis l'ordinateur hôte ; un exemple de code est donné dans l'Ann.D,§4. La *fig.III.8* permet de comparer le minimum énergétique atteint par recuit simulé avec celui obtenu par minimisation déterministe (ICM) - dans le cas du débruitage d'images binaires. Le recuit est exagérément lent (nombre d'itérations

³⁸ C'est à dire que la dynamique du processus ne change pas ; ceci peut s'avérer faux quand on fait varier le paramètre de température de speckle (non linéarités de la détection différentielle peut être, élargissement du spectre d'émission de la source pour des puissances en dessous du seuil laser, qui change les degrés de liberté M du speckle et donc la *forme* de la loi sigmoïde du thermostat, etc.).

quatre fois plus grand que pour le recuit sérieux), mais les paramètres ont été choisis pour permettre de bien visualiser la courbe d'évolution de l'énergie ; on obtient des résultats finaux comparables pour des recuits comportant seulement une centaine d'itérations. On remarquera que l'énergie finale obtenue par recuit simulé est non seulement inférieure à celle obtenue par minimisation déterministe, mais elle est également (très légèrement) inférieure à l'énergie de l'image originale non bruitée - un modèle de débruitage idéal devrait faire correspondre le minimum énergétique de la fonction d'énergie avec la configuration reproduisant l'image originale.



Image binaire (24x24)
de contrainte aux
données.

Pente théorique :

$$m_{th} = (T_1 - T_2) / (T_1 \cdot T_2) = 0,0099$$

$$m_e / m_{th} = 108 \%$$

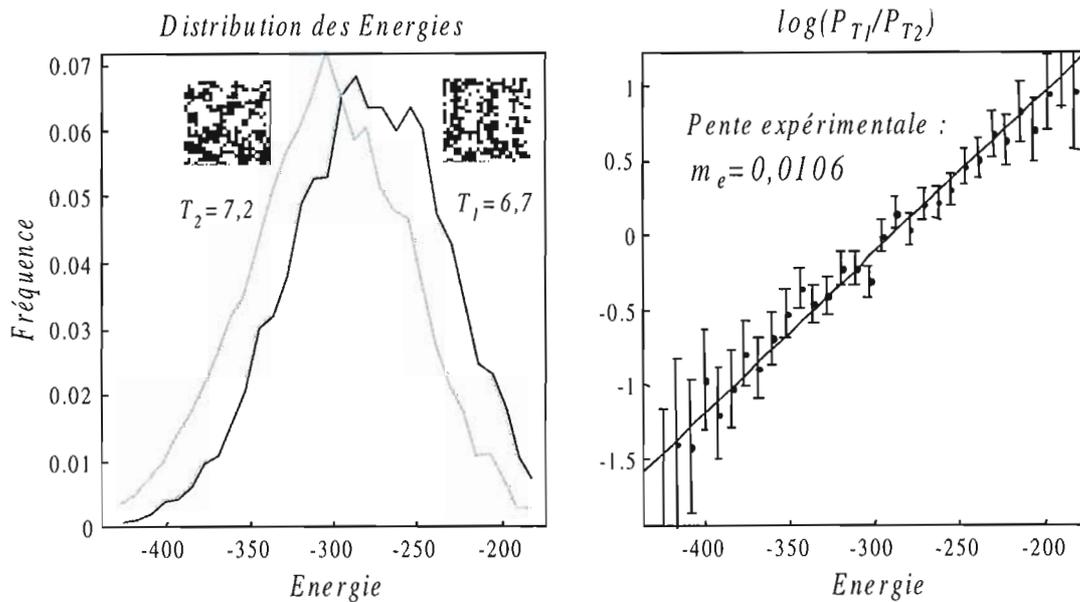


Fig.III.7: test de l'équilibre thermodynamique à partir d'un échantillon de 6000 configuration du champ. Les paramètres expérimentales sont : $I_A=220$ mA, $I_B= 200$ mA, soit $\beta_S=1,02$ et $\beta_C=0,82$.

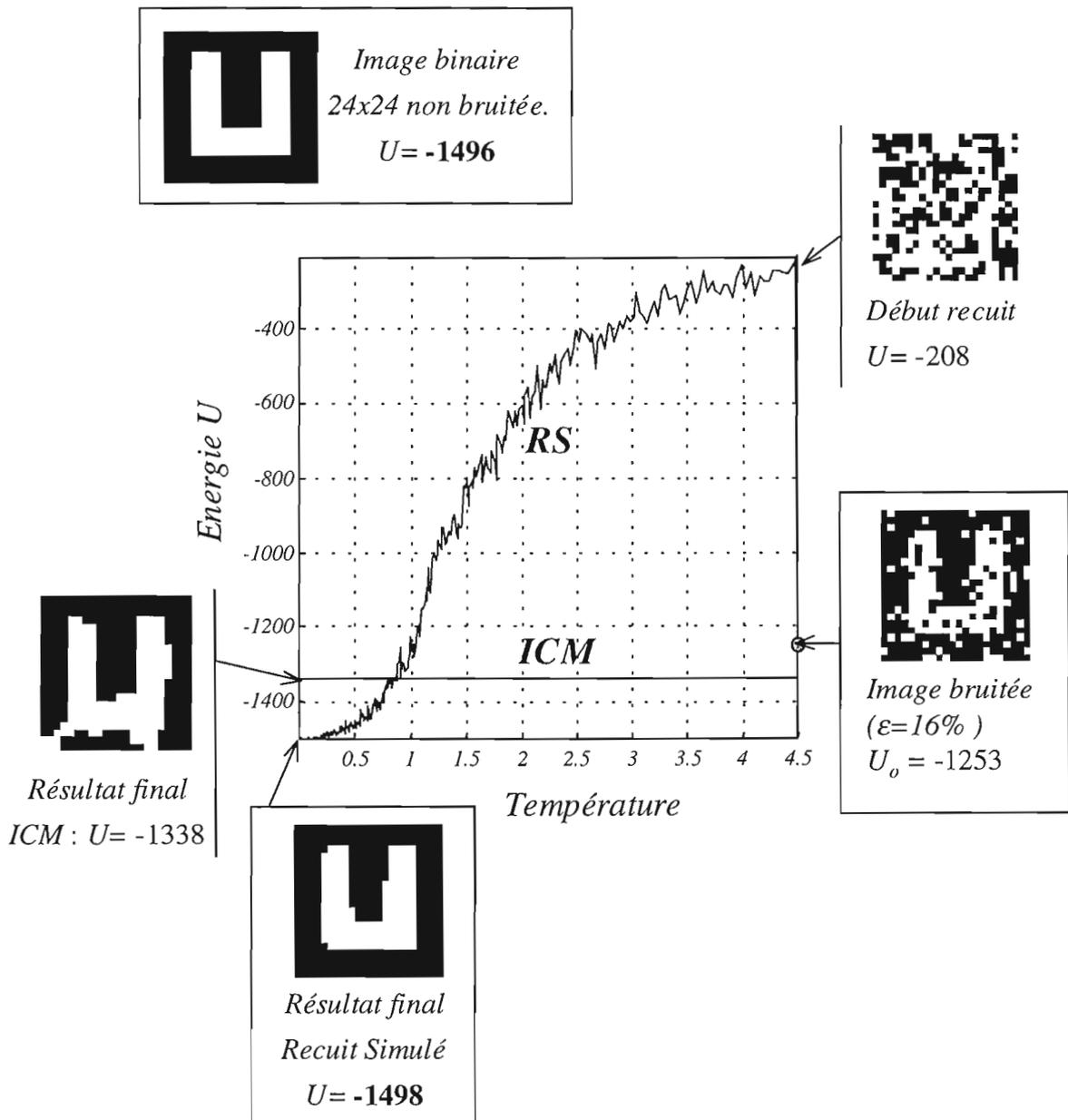


Fig.III.8 : Energie en fonction de la température pendant le recuit simulé (RS) et l'ICM pour la restauration d'une image binaire 24x24 bruitée avec un bruit de canal $\epsilon=16\%$, en utilisant les *interconnexions électriques* aux 4 plus proches voisins du processeur SPIE600. Comparer avec la fig.II.7 du Chap.II.

Paramètres :

Modèle : On a $I_a=220$ mA et $f_{pol}=100$ mV, donc $\beta_C=1,4$ et $\beta_S=1$.

Recuit : $T_i=4,5$; $T_f=0,005$; $r=0,99$ et $nsan=6$ (4062 itérations par recuit).

La procédure ICM correspond à 6 mises à jour du champ avec $T=0$.

6.2.4 Résultats visuels.

La figure *fig.III.9* représente une séquence d'images binaires 24x24 bruitées avec un bruit de canal croissant, ainsi que les résultats correspondant respectivement à la restauration par recuit simulé optoélectronique et par minimisation déterministe (toujours en utilisant le processeur SPIE600). On remarquera que les "conditions aux limites" électriques de la matrice SPIE600 correspondent à un pourtour de PEs virtuels figés dans l'état +1, ce qui perturbe parfois le résultat de la restauration (apparition de pixels parasites aux bords et/ou "fusion" partielle entre l'image et les frontières de l'image).

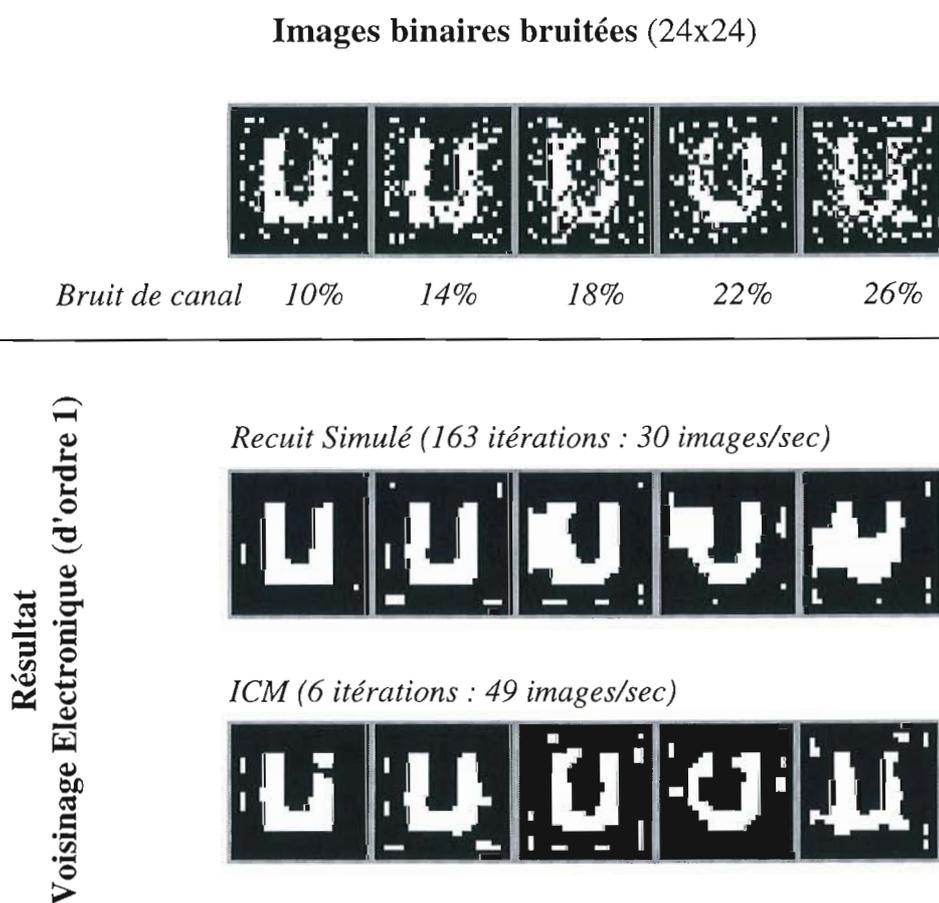


Fig.III.9 : Démonstration pour le débruitage d'images binaires 24x24 en utilisant les interconnexions électriques de SPIE600 (voisinage d'ordre 1).

Paramètres :

Modèle sont : $I_A=220$ mA et $f_{pol}=100$ mV, donc $\beta_C=1,4$ et $\beta_S=1$.

Recuit : $T_i=4,5$; $T_f=0,06$; $r=0,9$ et $nsan=4$, ce qui correspond à 163 itérations par image (au lieu de 1000 pour un recuit sérieux).

7. Conclusion

7.1 Nombre de recuits sérieux à la seconde sur SPIE600.

La fréquence maximale de mise à jour pour un coloriage binaire est de l'ordre de 10 kHz - en utilisant l'interface sous Windows/Matlab décrite dans l'Annexe D (fonction *Boucle_Electronique* pour $nscan=1000$ et $\tau_o=0$). Une dizaine de recuits sérieux par seconde³⁹ est en principe possible avec le prototype SPIE600. A titre de comparaison, nous estimons à 10 secondes le temps nécessaire pour effectuer un seul recuit sérieux en simulation sur un ordinateur Pentium 120MHz, pour un problème de taille 24x24.

7.2 Performances pour le débruitage d'images.

Une centaine d'itérations semblent suffisantes (au lieu de 1000 pour un recuit sérieux) pour obtenir un "bon" minimum énergétique dans le cas du débruitage d'images binaires 24x24. On peut en principe traiter une *centaine d'images binaires* à la seconde. Le dispositif à cristaux liquides ferroélectriques utilisé pour l'affichage des images à traiter (*cf.* Chap.V) peut "suivre" en affichant des images pré-enregistrés (cadence maximale de 1250Hz avec l'interface commerciale Slim II de Displaytech - *cf.* Ann.A). Une démonstration pour le *débruitage d'images à cadence vidéo* a ainsi été possible [Cassin98] - la cadence est réduite car il faut récupérer et sauvegarder le résultat de la restauration.

³⁹ On a pu obtenir des cadences supérieures de mise à jour (jusqu'à 100 kHz) en utilisant une interface informatique plus rudimentaire fonctionnant sous MS-DOS. On arrive alors à une *centaine* de recuits sérieux à la seconde. L'utilisation d'une interface informatique sous Windows 32 bits a eu comme effet de réduire considérablement les performances du démonstrateur, mais s'est avérée indispensable pour le contrôle de nouvelles cartes d'entrée-sortie dans le montage optoélectronique final (*cf.* Chap.V et VI et Ann.D).