

# Chapitre 6

## Groupements de haut niveau et mise en correspondance

---

*Les éléments visuels extraits à partir des groupements précédents peuvent être comparés à un croquis sommaire des formes saillantes de la scène. Nous montrons dans ce chapitre comment établir des relations structurelles plus complexes à partir de ces éléments visuels afin de mettre en correspondance des structures issues de deux scènes. Nous illustrons enfin ce dernier niveau de groupement par une méthode de mise en correspondance de jonctions.*

### 6.1 Mise en correspondance structurelle

La *mise en correspondance*, ou problème d'*appariement*, est l'une des principales tâches que doit accomplir un système visuel. D'une manière générale, nous entendons par "correspondance" l'identification d'attributs soumis à une relation commune. Ces attributs peuvent être extraits à partir d'images ou bien de modèles d'objets à identifier.

Ce problème, parmi les plus difficiles de la vision par ordinateur, se retrouve sous différentes formes en fonction de la nature des attributs ( points, segments, régions ou structures plus élaborées ), de leur origine et des applications envisagées.

- *Reconnaissance de formes*

Etant donné un ensemble de modèles d'objets connus, le problème de la reconnaissance de formes consiste à mettre en correspondance des attributs extraits d'une image avec ceux des modèles afin de détecter la présence d'objets connus dans la scène observée.

- *Stéréo-vision*

La stéréo-vision, ou problème d'appariement stéréoscopique, consiste à identifier les attributs communs à deux scènes afin de déduire le relief à partir de la

disparité entre les deux scènes. Selon les systèmes, la comparaison des images fournies par deux ou trois caméras permet de reconstituer la profondeur et finalement de modéliser la scène observée en trois dimensions.

– *Analyse de séquence d'images.*

Cette dernière variante est une généralisation de la vision stéréoscopique de par l'absence d'information sur le mouvement des objets de la scène et des changements de points de vues de la caméra. La mise en correspondance entre images d'une séquence se heurte aux problèmes posés par la présence d'objets multiples dont les mouvements peuvent être différents. Parmi les applications de ce type de mise en correspondance, on peut citer le suivi automatique d'objets ou encore la déduction de vues intermédiaires entre deux points d'observation clefs.

Pour plus de détails sur les variantes de la mise en correspondance et les méthodes utilisées pour résoudre ce problème, le lecteur pourra se référer aux états de l'art établis par [Zhang, 1993] et [Jones, 1997]. Il ressort de ces deux études une classification des méthodes de mise en correspondance selon trois critères : le choix des attributs à appairer, le type de contraintes utilisées pour comparer ces attributs et enfin, une méthode d'optimisation pour établir les correspondances entre attributs.

Une vue d'ensemble des méthodes existantes soulève trois remarques. Il existe d'une part quelques approches génériques au problème de la mise en correspondance, comme par exemple, l'isomorphisme de sous graphes, directement issues de la théorie des graphes. Ces approches sont cependant de complexité exponentielles et par conséquent difficilement applicables à tout type d'attributs. D'autre part, de nombreux algorithmes existent mais restent adaptés à des applications dans des conditions bien précises. Il en est ainsi, par exemple, de l'utilisation de contraintes épipolaires en vision stéréoscopique.

Enfin, ces études soulignent l'importance des méthodes hiérarchiques, reposant sur la mise en correspondance d'attributs complexes. Ces méthodes sont plus rapides et plus robustes car des structures complexes sont peu nombreuses par comparaison avec des primitives tels que des points d'intérêt ou des segments. La richesse des structures hiérarchiques assure également moins d'ambiguïtés. Enfin, et c'est là le principal avantage de ces méthodes, un appariement entre deux structures complexes peut être propagé facilement à chaque élément de celles-ci.

Pour ces raisons, la mise en correspondance structurelle est un prolongement fréquent des méthodes d'organisation perceptuelle. L'application de règles de groupement complexes permet d'établir des relations hiérarchiques fortes entre éléments visuels. Nous présentons à présent les principales relations structurelles utilisées par ces méthodes. Ces relations, ainsi que des exemples significatifs de mise en correspondance à partir de ces groupements, nous permettront de définir une approche adaptée aux éléments de représentation définis dans le chapitre précédent.

### 6.1.1 Relations structurelles

En pratique, le choix des règles de groupement de haut niveau dépend essentiellement du type de scène observée ou de l'application recherchée. Ces règles ont toutes en commun de produire des structures dont la probabilité d'apparition accidentelle est particulièrement faible. Chacune de ces relations peut être directement appliquée aux hypothèses de segments, d'arcs et de points d'intérêt extraites à partir du réseau de saillance. Parmi les plus utilisées, on peut citer les relations suivantes.

– *Symétrie et parallélisme*

Des groupes de segments localement parallèles [Ylä-Jääski et Ade, 1992] ou constituant des rubans symétriques [Cham et Cipolla, 1995] permettent souvent de prévoir un grand nombre de structures géométriques 2D. Par exemple, [Mohan et Nevatia, 1992] exploitent les configurations entre rubans pour retrouver des axes de symétrie et former des structures convexes telles que des quadrilatères. Ces résultats ont été étendus ensuite à des groupements courbes. Selon les mêmes principes, [Ip et Wong, 1997] utilisent des groupements de courbes parallèles afin de faciliter le suivi de routes sur des séquences vidéo.

– *Convergence et Proximité*

A un niveau local, la relation de convergence et de proximité définit la présence de jonctions entre extrémités de segments ou de courbes. D'une manière plus globale, cette relation est utile pour la détection de points de fuite. Ceux-ci sont détectés en projetant les segments de la scène dans un espace de paramètres exprimés en fonction de l'orientation et de l'équation des droites porteuses de chaque segment. Les segments qui convergent vers un même point correspondent à des points similaires dans cet espace [Straforini *et al.*, 1993] [Tai *et al.*, 1993]

– *Convexité et cycles*

La détection de groupements circulaires entre éléments de représentation est également une structure hiérarchique importante pour la mise en correspondance. A plus forte raison lorsque ces arrangements circulaires peuvent être identifiés à l'aide de modèles tels que des formes cycliques simples (quadrilatères, cercles, ellipses) ou bien à l'aide de propriétés particulières telles que la convexité.

[Jacobs, 1996] propose d'évaluer directement une mesure de convexité et de fermeture en termes de probabilité d'apparition par accident. Il démontre en particulier que le choix de bons critères de groupements permet de réduire la complexité d'une recherche quasi-exhaustive, précisément parce que les structures recherchées ont peu de chances de remplir ces critères par accident.

Pour chaque exemple, des méthodes spécialisées permettent d'extraire efficacement chaque type de groupement. Il en est ainsi de la détection de points de fuite ou de la recherche de cycles et d'ensembles convexes par des méthodes de parcours de graphes.

Il existe également une approche plus générale, qui consiste à prédire des hypothèses à partir de configurations locales entre primitives et de vérifier ensuite ces hypothèses à partir des primitives restantes [Denasi *et al.*, 1992]. Des groupements de haut niveau sont ainsi déduits de manière hiérarchique à partir de groupements plus simples. Par exemple, une succession d'arcs élémentaires co-circulaires constitue une bonne hypothèse pour la présence de cercles ou d'ellipses, qu'il suffit de vérifier ensuite par application d'un modèle.

Cette approche n'utilise pas de connaissance du type de scène observé. La seule connaissance utilisée porte sur les critères d'apparition d'une relation structurelle à partir de groupes plus simples. [Sarkar et Boyer, 1993a] proposent une méthode générique pour modéliser ce type de connaissance et l'exploiter afin de déduire automatiquement l'existence de structures plus complexes.

Leur démarche est divisée en trois parties. A partir de portions de contours de courbure continue, une première partie groupe des fragments compatibles en fonction de règles simples. Pour chaque type de groupement recherché, l'espace des paramètres est discrétisé de manière à ce que chaque segment vote pour les points de cet espace satisfaisant cette relation. Ils constituent ainsi ce qu'ils nomment des "Gestalt graphs" reflétant des propriétés de proximité, continuité, fermeture et région commune. On pourra se reporter à [Sarkar et Boyer, 1992] pour une comparaison plus détaillée de cette méthode avec la transformée de Hough.

La deuxième étape combine les graphes entre eux et utilise des techniques de parcours de graphe et recherche de cliques pour établir des hypothèses. Par exemple, les graphes de "région communes" ET de "proximité" donnent des segments parallèles. Ou encore, la recherche de cycles sur les graphes de jonctions donne des polygones.

Finalement, des hypothèses plus complexes sont déduites dans une troisième partie à l'aide d'un réseau d'inférence Bayésien (*PIN - Perceptual Inference Network*). Ce réseau modélise une connaissance *a priori* de règles de groupement sous la forme d'un graphe de probabilité. La probabilité de détection de rubans, cercles ou rectangles dépend alors des probabilités d'apparition de structures génératrices de ces hypothèses [Sarkar et Boyer, 1994]. Cette démarche constitue l'une des rares tentatives d'approche globale du groupement perceptuel à l'aide d'un seul formalisme.

### 6.1.2 Organisation perceptuelle et mise en correspondance

Les différentes applications du groupement perceptuel à la mise en correspondance structurelle partagent deux principes.

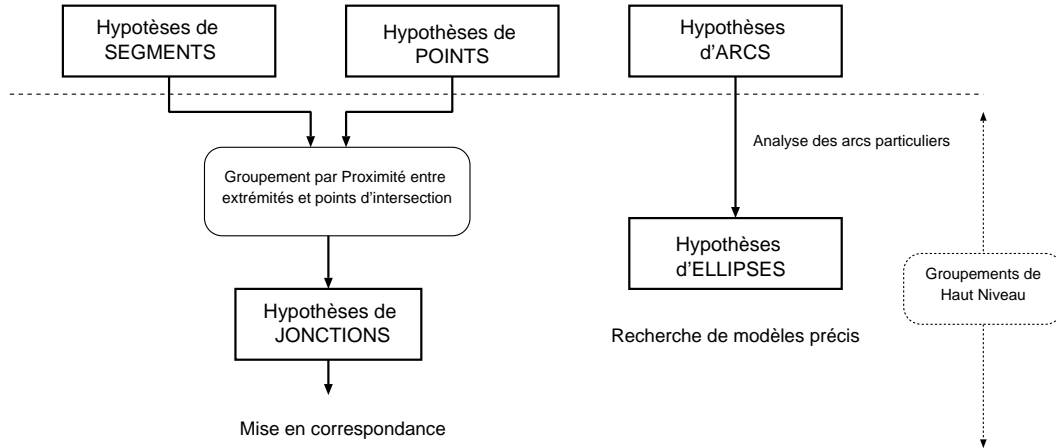
D'une part, la formation de groupement complexes permet un appariement sommaire de structures globales, qui sert ensuite de centre d'attention pour une mise en correspondance plus fine. [Mohan et Nevatia, 1989] démontrent l'intérêt de cette ap-

proche pour l'extraction de structures tridimensionnelles, en particulier pour faciliter la détection de bâtiments en imagerie aérienne. Des paires de segments parallèles sont d'abord extraites à partir de la détection de contours. Des hypothèses intermédiaires de structures en U sont ensuite élaborées à partir de ces paires et complétées sous forme de rectangles. Dans un premier temps, une correspondance sommaire est établie entre les rectangles correspondant aux toits des bâtiments. Les arêtes de chaque rectangle sont ensuite appariées de façon plus précise. Les rectangles donnent ainsi une contrainte forte sur les candidats possibles pour la mise en correspondance précise des contours des toits.

D'autre part, la mise en correspondance est souvent considérée comme un groupement perceptuel particulier. On retrouve en effet dans l'opération d'appariement l'idée d'association entre structures présentant des mouvements similaires. C'est pourquoi ces deux opérations sont souvent présentées avec le même formalisme. [Horaud et Skordas, 1989] ou encore [Sarkar, 1994] représentent les groupements effectués au sein d'une même image sous la forme de graphes relationnels entre primitives. Les appariements possibles sont présentés de la même manière sous la forme de graphes de correspondances. Ce second graphe représente les appariements possibles pour chaque segment de l'image de départ. L'appariement revient ainsi à parcourir ce graphe pour extraire les ensembles de noeuds mutuellement compatibles. Sarkar et Boyer poussent plus loin ce parallèle en étudiant des séquences d'images par superposition de groupements 2D issus d'images consécutives. Cette représentation composite permet d'utiliser des techniques de groupement perceptuel afin de reconstituer les trajectoires des structures de la scène. La mise en correspondance n'est ici rien d'autre qu'un groupement "temporel" entre structures similaires.

Enfin, les approches hiérarchiques élaborées par [Havaldar *et al.*, 1996] ou bien [Venkateswar et Chellappa, 1995] sont autant d'exemples de synthèse de ces deux principes. Dans ce dernier cas, par exemple, les primitives 2D sont organisées selon une hiérarchie de structures de complexité croissante : "segments", "coins", "arêtes" (groupement de segments colinéaires entre deux coins) et enfin "facettes" (succession de contours et de coins consécutifs). La mise en correspondance de structures de haut niveau apporte des contraintes de localisation fortes sur les structures de niveaux inférieurs. Les opérations de groupement et d'appariement sont ici encore présentées selon le même formalisme de graphe.

Nous proposons d'utiliser les groupements définis par les chapitres précédents à la mise en correspondance de jonctions. Dans un premier temps, nous montrons comment détecter et grouper les jonctions à partir des points d'intérêt et des segments. Ces jonctions sont ensuite appariées à l'aide d'une méthode de relaxation stochastique reprenant les deux principes que nous venons d'évoquer.



**Figure 6.1** - Principes des niveaux supérieurs de groupement. Les éléments visuels extraits par les niveaux inférieurs sont soit manipulés directement sous la forme de structures plus complexes (mise en correspondance structurelle), soit utilisés comme centre d'attention pour valider des hypothèses de manière plus précise (prédiction et vérification d'hypothèses).

## 6.2 Extraction et groupement de jonctions

Les jonctions entre primitives géométriques correspondent en général à des sommets ou des occlusions entre arêtes des objets de la scène. Elles peuvent être envisagées de deux manières, soit comme une relation entre segments et/ou arcs, soit comme des entités à part entière, localisées à l'aide de détecteurs spécialisés.

Si les jonctions sont considérées comme des primitives particulières, leur extraction passe d'abord par une localisation de points d'intérêt. En tant qu'intersections entre extrémités de primitives linéaires, les jonctions correspondent en effet à des "coins" (deux branches) ou des "sommets" (trois branches et plus). Les jonctions proprement dites peuvent alors être définies à partir du modèle de coin détecté, ou à l'aide des extrémités de segments ou courbes présentes dans leur voisinage [Lindeberg et Li, 1997].

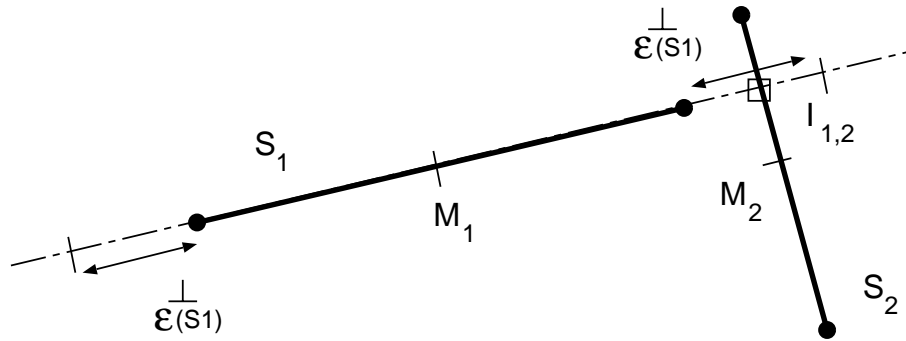
Si au contraire, elles sont considérées comme des relations entre segments ou courbes (relation de connectivité par exemple), les jonctions peuvent être extraites à partir de l'ensemble des intersections possibles entre primitives linéaires, moyennant une certaine zone de recherche autour des extrémités de ces primitives. L'extraction de jonctions cohérentes peut alors se rapporter à un problème d'étiquetage et peut être résolu par un procédé de relaxation sur les probabilités de connexions entre extrémités [Regier, 1991].

Nous conservons, pour l'extraction des jonctions, des principes comparables à ceux appliqués pour le groupement de segments et d'arcs. Dans un premier temps,

des jonctions élémentaires sont détectées à partir des intersections entre hypothèses issues des groupements précédents. Elles sont ensuite groupées entre elles afin d'éliminer les jonctions redondantes et constituer des jonctions multiples. A la différence des groupements précédents, les jonctions sont détectées à partir des primitives géométriques et non plus directement à partir des chaînes sélectionnées par le réseau de saillance.

### 6.2.1 Détection des jonctions élémentaires

Les hypothèses de jonctions élémentaires sont établies à partir de paires de segments non colinéaires répondant à un certain critère de voisinage. A l'image des points dominants détectés lors du groupement d'arcs élémentaires, chaque jonction est associée aux segments qui la génèrent afin de former un triplet  $(I_{1,2}, S_1, S_2)$ .



**Figure 6.2** - Notations utilisées pour une intersection entre deux segments. Une marge d'erreur permet de définir des jonctions "réelles" et "virtuelles". Ici, la jonction entre  $S_1$  et  $S_2$  est virtuelle.

Soient  $S_1$  et  $S_2$  deux segments non colinéaires et  $I_{1,2}$  leur intersection. On note de plus  $M_1$  et  $M_2$  les milieux respectifs de chaque segment.

Malgré l'ajustement des extrémités des segments réalisé en fin de groupements, la détection d'intersections entre segments doit tenir compte d'une marge d'erreur liée à la longueur de chaque segment. Cette distance autour des extrémités permet de rattraper la mauvaise localisation des coins lors de l'extraction des contours.

Notons  $\epsilon^\perp(S_1)$  la marge d'erreur du segment  $S_1$ . En pratique, cet écart est défini par :

$$\epsilon^\perp(S_1) = 0.3 \cdot \|\vec{S}_1\|$$

Cette définition entraîne la détection de trois types de jonctions, résumées par la figure 6.3.

– *Jonctions réelles*

Elles sont caractérisées par les relations suivantes :

$$\left. \begin{array}{l} \text{et} \\ \left\| \overrightarrow{M_1, I_{1,2}} \right\| < \frac{\|\vec{S}_1\|}{2} \\ \left\| \overrightarrow{M_2, I_{1,2}} \right\| < \frac{\|\vec{S}_2\|}{2} \end{array} \right\} \implies (I_{1,2}, S_1, S_2) \text{ Réelle} \quad (6.1)$$

Ce sont les jonctions en L, T ou X selon les positions respectives du point d'intersection sur chaque segment.

– *Jonctions virtuelles*

Ces jonctions en V correspondent au cas inverse où le point d'intersection n'appartient à aucun des segments.

$$\left. \begin{array}{l} \text{et} \\ \left| \left\| \overrightarrow{M_1, I_{1,2}} \right\| - \frac{\|\vec{S}_1\|}{2} \right| < \epsilon^\perp(S_1) \\ \left| \left\| \overrightarrow{M_2, I_{1,2}} \right\| - \frac{\|\vec{S}_2\|}{2} \right| < \epsilon^\perp(S_2) \end{array} \right\} \implies (I_{1,2}, S_1, S_2) \text{ Virtuelle} \quad (6.2)$$

– *Jonctions  $\lambda$*

Cette dernière classe de jonction correspond au cas intermédiaire où le point d'intersection n'appartient qu'à l'un des deux segments.

La recherche des jonctions consiste simplement à comparer les segments deux à deux. Chaque intersection étant symétrique ( $I_{1,2} = I_{2,1}$ ), cette recherche revient à comparer  $\frac{n(n-1)}{2}$  segments, si  $n$  est le nombre de segments.

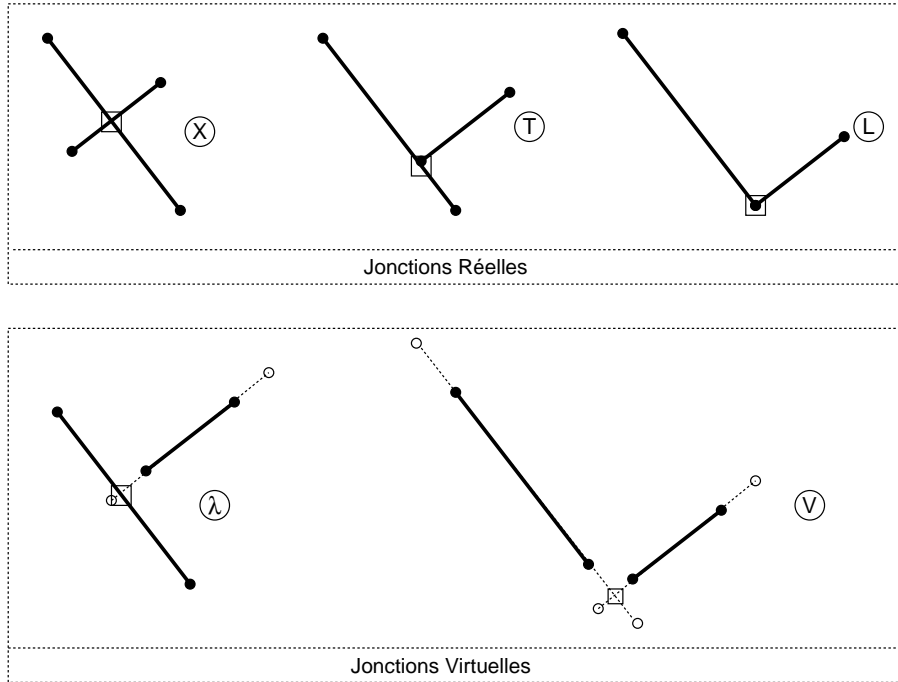
## 6.2.2 Groupement en jonctions complexes

Comme pour la détection de segments et d'arcs, les erreurs de localisation des extrémités de segments entraînent des redondances de jonctions. Les jonctions ainsi extraites sont superposées ou bien très proches les unes des autres, et nécessitent donc une étape de groupement. Le groupement de jonctions consiste donc à produire un ensemble simplifié de  $n$ -uplets  $(I_1^j, S_1^1, \dots, S_j^1)$  constitués d'un point central et d'un ensemble de segments ou "branches".

– Simplification des jonctions superposées

Les règles de groupement entre jonctions concernent à la fois leur position et leurs branches. Deux jonctions  $(I_j^1, S_1^1, \dots, S_j^1)$  et  $(I_k^2, S_1^2, \dots, S_k^2)$  sont fusionnées si leurs centres respectifs sont suffisamment proches. Si, de plus, elles ont





**Figure 6.3** - *Catalogue des différentes classes de jonctions élémentaires entre deux segments.*

des branches en commun, celles-ci sont fusionnées à leur tour.

$$\left\{ \begin{array}{l} \|\overrightarrow{I_j^1}, \overrightarrow{I_k^2}\| < \epsilon^+ \implies I_j^1 \equiv I_k^2 \\ \exists m < j, \exists n < k, \quad |\widehat{S_m^1}, \widehat{S_n^2}| < \epsilon^\theta \implies S_m^1 \equiv S_n^2 \end{array} \right. \quad (6.3)$$

Le centre de la jonction issue du groupement est défini par le milieu du segment  $\overline{I_j^1, I_k^2}$ . Les branches fusionnées sont, quand à elles, remplacées par un segment de longueur :

$$L = \text{Max}(\|\overrightarrow{S_m^1}\|, \|\overrightarrow{S_n^2}\|)$$

et d'orientation :

$$\Theta = \frac{\widehat{S_m^1} + \widehat{S_n^2}}{2}$$

L'algorithme de fusion des branches entre deux jonctions consiste simplement à mettre en commun les branches des deux voisinages et à les trier par ordre croissant d'orientation. Il suffit ensuite de les comparer deux à deux et de fusionner les branches présentant un faible écart angulaire ( inférieur à 5 degrés ).

Enfin, les écarts  $\epsilon^+ = 5$  pixels et  $\epsilon^\theta = 10$  degrés sont définis empiriquement.

– Mesure de saillance à partir des points d'intérêt

Afin de tenir compte des points d'intérêts détectés lors des groupements de segments et d'arcs, une mesure de saillance définie à partir de ces points est associée à chaque jonction. Cette mesure récompense les jonctions situées à proximité de coins ou d'extrema de courbures.

Soit une jonction  $I_j$ . On note  $\{P_1, \dots, P_n\}$  l'ensemble des points d'intérêt situés dans un voisinage de  $\epsilon^+$  pixels du centre de  $I_j$ . La mesure de saillance de la jonction est simplement définie par :

$$\mathcal{S}(I_j) = \sum_{i=1}^n e^{-\frac{d_i}{\sigma^+}}, \quad \text{avec } d_i = \|\overrightarrow{I_j, P_i}\| \quad \text{et } \sigma^+ = \frac{\epsilon^+}{2} \quad (6.4)$$

où  $\sigma^+$  est une constante qui permet d'ajuster l'étendue de l'exponentielle.

Ainsi, la saillance d'une jonction est d'autant plus forte que la localisation du centre est confirmée par des points d'intérêts détectés par ailleurs.

Le résultat de cette opération de groupement est un ensemble simplifié de jonctions, définissant une relation de proximité entre un point d'intérêt et un ensemble de segments. Les branches de ces jonctions sont une liste de segments disjoints. Leur point central est le barycentre des jonctions élémentaires impliquées dans le groupement, auquel est associé une mesure de saillance.

### 6.2.3 Résultats sur les jonctions de segments

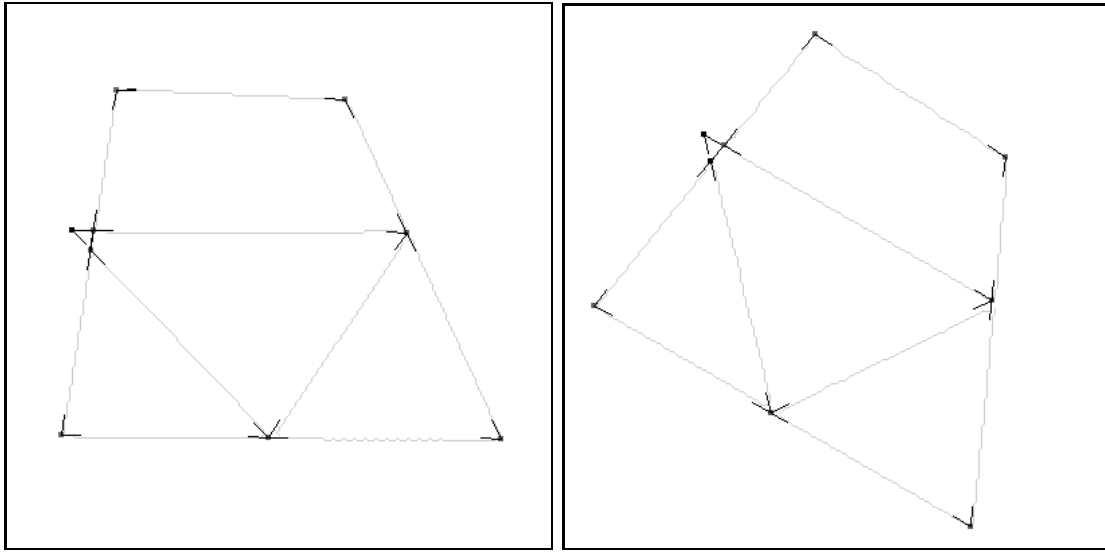
Comme dans le cas des groupements de segments et d'arcs, nous commençons par présenter les résultats du groupement de jonctions dans une scène artificielle de test. Les figures 6.4 et 6.5 illustrent la stabilité en rotation du groupement de jonctions. Pour des raisons de clarté, nous ne représentons que les centres et les directions des branches de chaque jonction. Malgré des irrégularités importantes le long des arêtes des polygones, les jonctions sont correctement groupées. Les jonctions superflues présentes dans les deux images de la figure 6.5 donnent une idée des jonctions résiduelles qui échappent au groupement. Ce type de résidu ne remet pas en cause la détection des autres jonctions, plus stables. En effet, elles seront considérées par l'algorithme de mise en correspondance comme des éléments incohérents, et seront alors rejetées du voisinage des jonctions stables.

Les figures des pages 215 à 216 représentent une scène de bureau. La détection de contours et la sélection des groupements les plus saillants permettent de passer de 440 chaînes de contours à seulement 54 groupements, dont sont extraits les hypothèses de segments (Figure 6.7). Les figures suivantes montrent successivement la détection de 718 jonctions doubles (Figure 6.8) et la simplification de ces hypothèses en 229 jonctions groupées (Figure 6.9). Les jonctions virtuelles sont

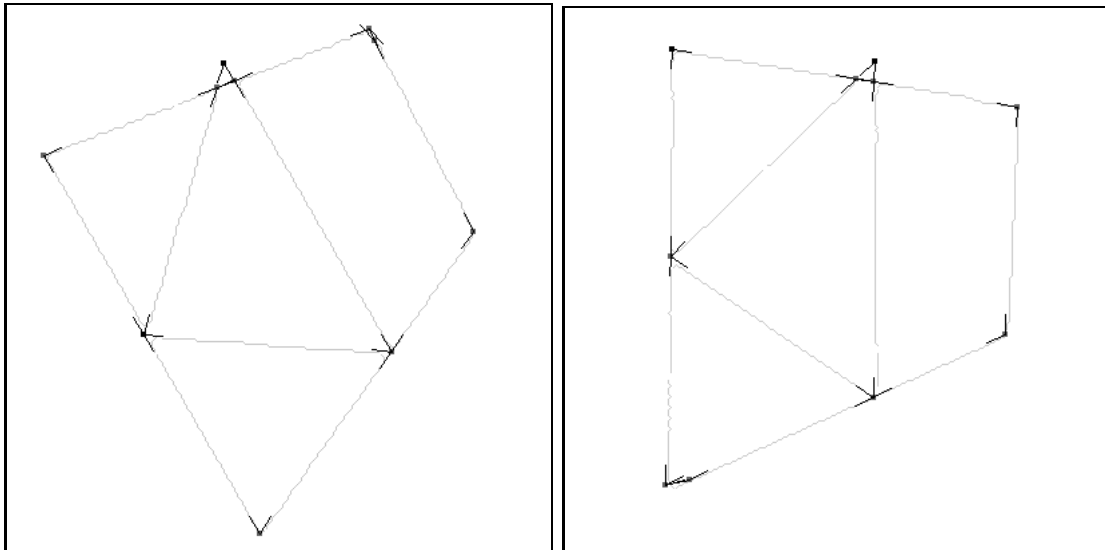
représentées ici en noir. Bien qu'elles n'aient, dans la plupart des cas, aucune signification physique, ces jonctions virtuelles apportent néanmoins des informations sur les positions relatives de segments importants de la scène.

Notre méthode est une approche structurale au problème de la détection de coins. Par comparaison, les méthodes de détection de coins évoquées dans le sous-chapitre 2.3.3 (page 59), abordent le problème à partir des propriétés photométriques de l'intensité lumineuse de l'image d'origine. Celle-ci pourraient être utilisées en complément au groupement de jonctions afin d'obtenir une localisation précise de leur centre. En particulier, le groupement de jonctions pourrait servir d'initialisation au détecteur de coins de [Blaszka et Deriche, 1994b] à l'aide de modèles déformables.

Enfin, les figures des pages 217 et 218 reprennent les scènes utilisées dans le chapitre précédent en présentant, dans chaque cas, la détection et le groupement des jonctions. Malgré une diminution significative du nombre de jonctions superposées, on peut cependant remarquer que les règles de groupement entre jonctions sont insuffisantes pour simplifier toutes les situations. C'est, pour l'instant, la principale limitation de cette méthode. L'un des prolongements immédiats est la définition de règles plus complètes, comme nous l'avons fait avec les groupements de segments. Ces règles pourraient s'inspirer de méthodes de groupement de jonctions par relaxation d'un critère de proximité [Matas et Kittler, 1993] ou encore d'une construction progressive d'un graphe de voisinage entre segments saillants de la scène [Jacot-Descombes et Pun, 1997] .



**Figure 6.4** - *Stabilité du groupement de jonctions en rotation -  $\theta = 0$  et  $\theta = \frac{\pi}{6}$*



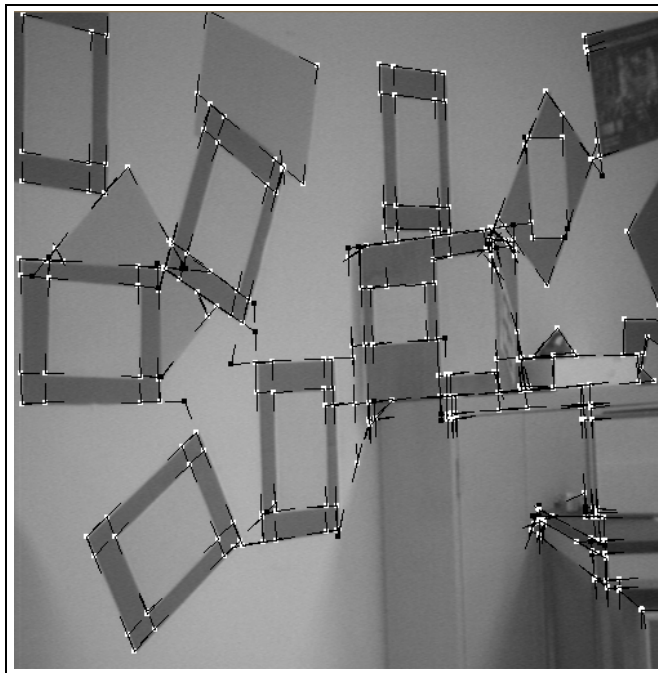
**Figure 6.5** - *Stabilité du groupement de jonctions en rotation -  $\theta = \frac{2\pi}{6}$  et  $\theta = \frac{\pi}{2}$*



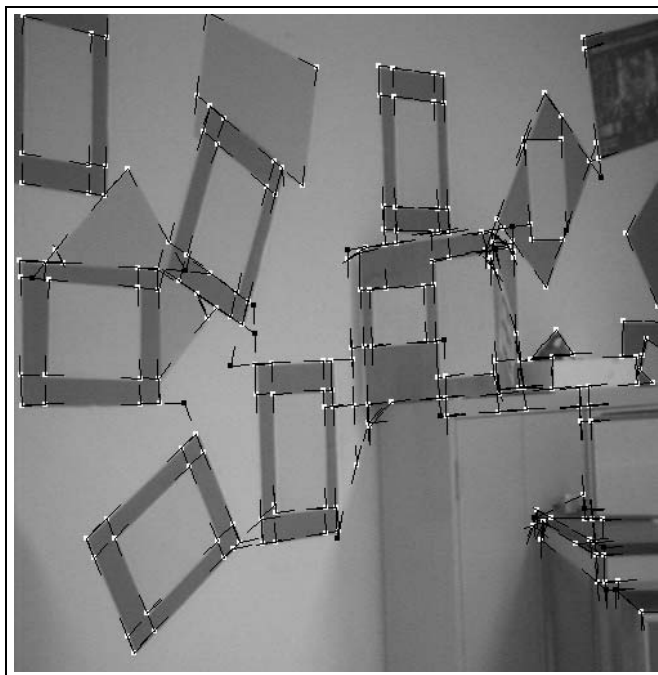
**Figure 6.6** - *Scène de bureau*



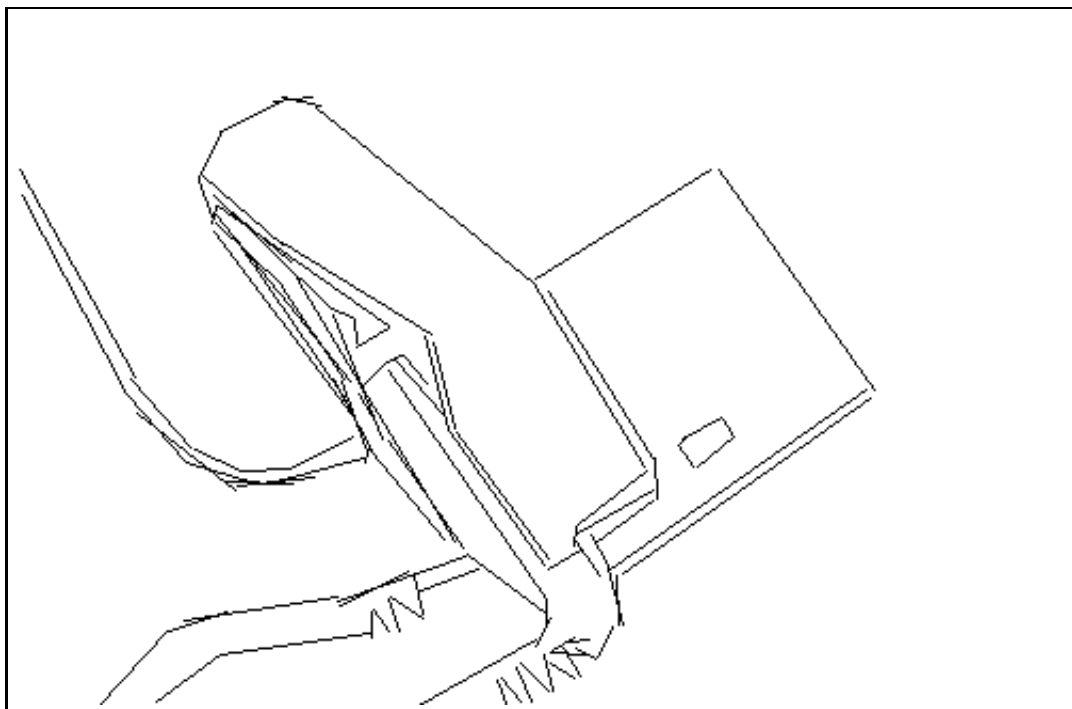
**Figure 6.7** - *Scène de bureau - Détection et groupement de segments - 144 segments extraits à partir de 54 groupements sur 440 chaînes (note : les discontinuités des segments en blanc sont dues à un défaut d'impression).*



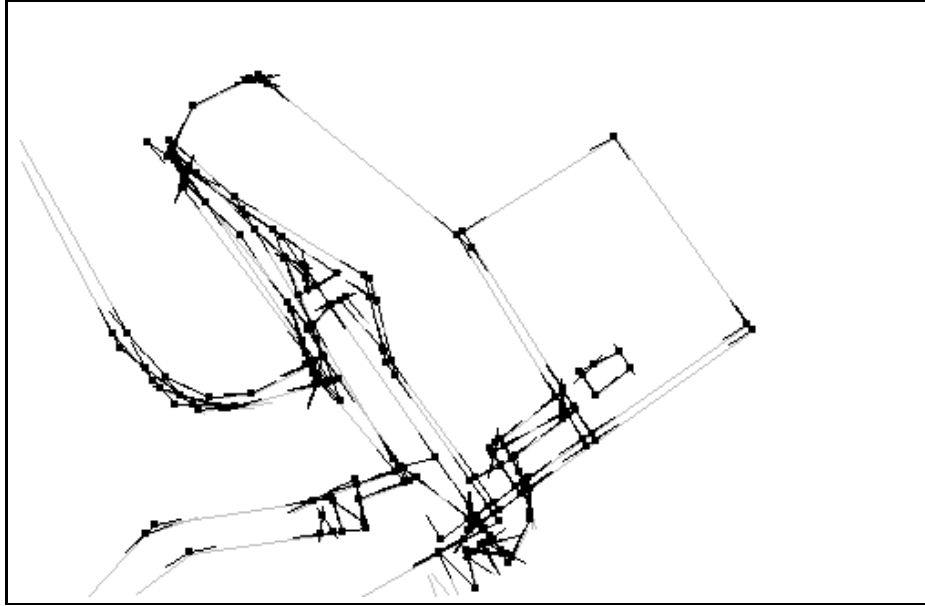
**Figure 6.8** - *Scène de bureau - Détection de 718 jonctions doubles*



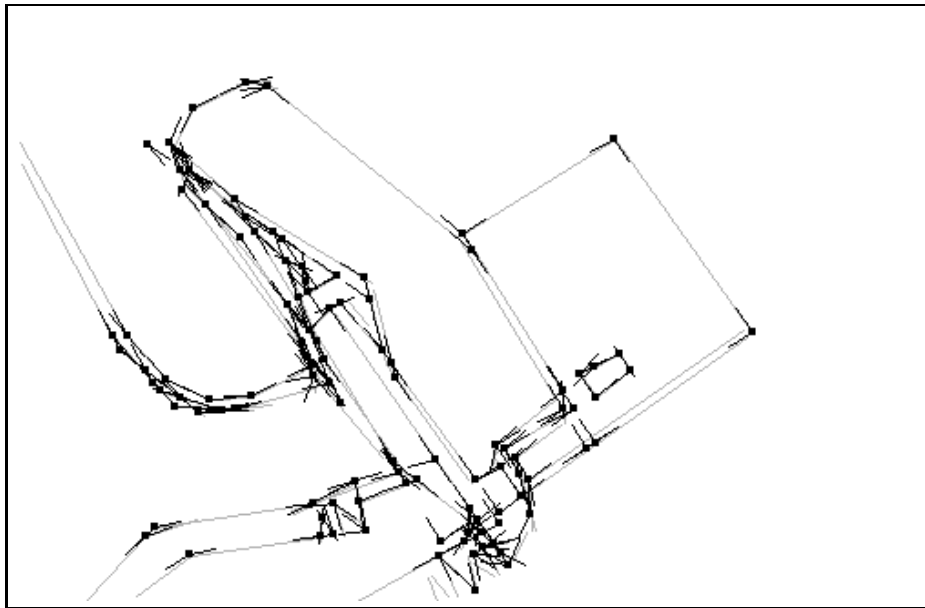
**Figure 6.9** - *Scène de bureau - Groupement de jonctions - restent 229 jonctions groupées*



**Figure 6.10** - *Téléphone - Détection et groupement de segments - 101 segments extraits à partir de 22 groupements sur 550 chaînes*



**Figure 6.11** - *Téléphone - Détection de 482 jonctions doubles*

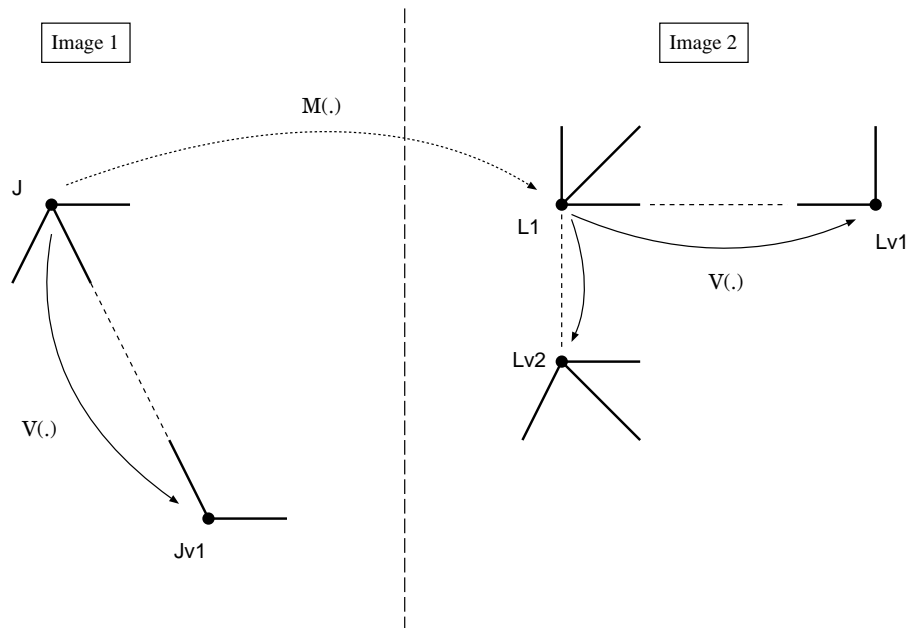


**Figure 6.12** - *Téléphone - Groupement de jonctions - restent 101 jonctions groupées*



## 6.3 Mise en correspondance de jonctions

Nous nous inspirons, pour l'appariement de jonctions, d'un algorithme de coopération entre mise en correspondance et groupement perceptuel proposé par [Chang et Aggarwal, 1997]. Cette méthode tient compte en particulier des principes d'application du groupement perceptuel à la mise en correspondance. Originellement appliquée à la mise en correspondance de segments, elle suit une approche suffisamment générique pour être adaptée simplement à tout type de structure, quel que soit son niveau de hiérarchie ou de complexité. Elle peut donc être utilisée pour guider une mise en correspondance précise à partir d'appariements de structures complexes. Enfin, elle considère la mise en correspondance comme une étape de groupement temporel d'une image vers l'autre, partageant le même formalisme que le groupement perceptuel.



**Figure 6.13** - *Voisinages Temporels et Perceptuels pour une jonction J.*

Cet algorithme part de l'hypothèse d'un mouvement rigide des objets d'une image à l'autre de la scène. Ainsi, les éléments d'un même groupement perceptuel établi dans une image doivent avoir le même mouvement dans l'autre image. Les relations de structures d'un groupement sont donc utilisées comme contraintes sur les appariements possibles.

En retour, la mise en correspondance doit également préserver les relations structurelles. Il est donc possible d'utiliser les hypothèses d'appariements de chaque membre d'un groupement perceptuel afin de rejeter ceux dont le mouvement est incohérent avec celui du groupe.

L'algorithme de mise en correspondance suit une procédure de relaxation stochastique afin d'apparier de manière cohérente chaque élément visuel, tour à tour avec ses correspondants possibles, et ses voisins immédiats.

Nous ne retenons de cet algorithme que ses principes de coopération entre deux types de contraintes pour l'appliquer directement à la mise en correspondance de jonctions. Celles-ci sont en effet moins nombreuses que des segments et surtout, moins ambiguës lorsqu'il s'agit de les comparer.

### 6.3.1 Coopération entre appariement et groupement

Le problème de l'appariement de jonctions consiste donc à établir une correspondance entre un ensemble  $\mathcal{J}_1$  de jonctions détectées dans une image  $I_1$  et un autre ensemble de jonctions  $\mathcal{J}_2$  extraites d'une image  $I_2$ .

La coopération entre mise en correspondance et groupement perceptuel suppose la définition de deux types de voisinages pour chaque jonction  $J \in \mathcal{J}_1$ .

- Un voisinage *temporel*, noté  $\mathcal{M}(J)$ , qui représente l'ensemble des candidats possibles dans  $\mathcal{J}_2$  pour l'appariement avec  $J$ . Chaque élément de ce voisinage vérifie la relation :

$$\begin{array}{ccc} \mathcal{M} : & \mathcal{J}_1 & \longmapsto & \mathcal{J}_2 \\ & J & \longrightarrow & \mathcal{M}(J) = \{L_1, \dots, L_k\}, \end{array} \quad \begin{array}{l} \text{correspondants possibles pour } J \\ (6.5) \end{array}$$

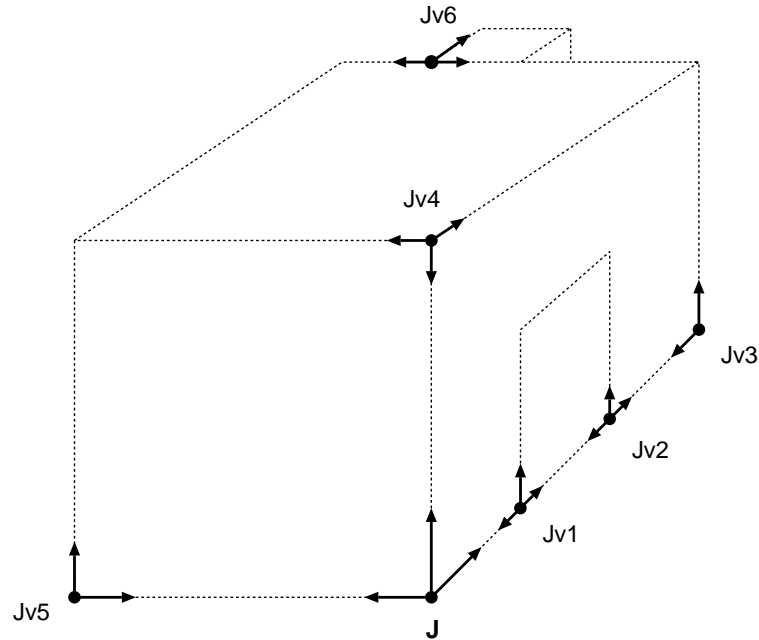
Chaque élément  $L_i$  de  $\mathcal{M}(J)$  définit une probabilité d'appariement avec la jonction  $J$ , notée :  $S_M(J, L_i)$ .

- Un voisinage *perceptuel*, noté  $\mathcal{V}(J)$ , qui représente une relation structurelle entre la jonction  $J$  et un certain nombre de jonctions de  $\mathcal{J}_1$ .

$$\begin{array}{ccc} \mathcal{V} : & \mathcal{J}_1 & \longmapsto & \mathcal{J}_1 \\ & J & \longrightarrow & \mathcal{V}(J) = \{Jv_1, \dots, Jv_n\}, \end{array} \quad \begin{array}{l} \text{voisins de } J \text{ pour la relation } \mathcal{V} \\ (6.6) \end{array}$$

Le voisinage défini par  $\mathcal{V}$  peut être construit, par exemple, à l'aide de l'une des relations définies à la section 6.1.1, ou encore la somme de relations de ce type. Dans le cas présent, le voisinage perceptuel d'une jonction  $J \in \mathcal{J}_1$  est constitué des jonctions de  $\mathcal{J}_1$  dont le centre est aligné avec l'une des branches de  $J$ , comme le montre la figure 6.14. On définit un voisinage similaire pour les jonctions appartenant à  $\mathcal{J}_2$ .

Enfin, chaque élément  $Jv_i$  de  $\mathcal{V}(J)$  définit également un score de groupement avec la jonction  $J$ , noté :  $S_G(J, Jv_i)$ . Contrairement aux probabilités d'appariement, ce score ne représente pas une probabilité mais plutôt un coefficient de compatibilité entre une jonction et ses voisins.



**Figure 6.14** - *Le voisinage perceptuel d'une jonction  $J$  est constitué des jonctions dont le centre se trouve aligné avec l'une des branches de  $J$ .*

**Algorithme 6.1** : Coopération entre Groupement Perceptuel et Mise en Correspondance - algorithme principal

**début**

**pour** *Chaque jonction  $J$  de  $\mathcal{J}_1$  faire*

┌ Initialiser le voisinage perceptuel  $\mathcal{V}(J)$

└ Initialiser le voisinage temporel  $\mathcal{M}(J)$

**pour** *Chaque jonction  $L$  de  $\mathcal{J}_2$  faire*

┌ Initialiser le voisinage perceptuel  $\mathcal{V}(L)$

**répéter**

┌ % Relaxation temporelle - deux passes

Mise à jour des probabilités d'appariement

sur les voisinages temporels des jonctions de  $\mathcal{J}_1$

└ %

┌ % Relaxation perceptuelle

Mise à jour des scores de groupement

sur les voisinages perceptuels des jonctions de  $\mathcal{J}_1$

└ %

**pour** *Chaque jonction  $J$  de  $\mathcal{J}_1$  faire*

┌ Eliminer les voisins de  $\mathcal{V}(J)$  et  $\mathcal{M}(J)$  dont le score est trop faible

**jusqu'à** *stabilité des voisinages temporels des jonctions de  $\mathcal{J}_1$*

**fin**

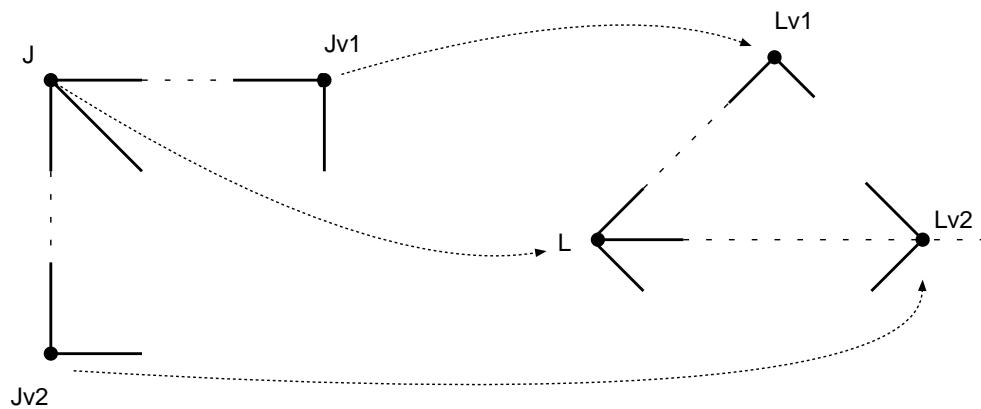
L'algorithme proprement dit se décompose en deux étapes. Dans un premier temps, un processus de relaxation stochastique est appliqué au voisinage temporel de chaque jonction  $J$  afin d'attribuer à chaque hypothèse une probabilité d'appariement. Cette étape compare les voisinages perceptuels de  $J$  et de ses correspondants afin de renforcer itérativement les hypothèses dont le voisinage est cohérent avec celui de  $J$ .

Les probabilités d'appariements sont ensuite utilisées dans la seconde étape de relaxation stochastique afin de renforcer, pour chaque jonction  $J$ , les voisins dont les hypothèses d'appariement sont compatibles avec celles de  $J$ .

Cette seconde relaxation permet de ne conserver que les éléments du voisinage perceptuel de  $J$  dont le mouvement est cohérent avec celui de  $J$ . Les voisins incohérents correspondent en pratique à des groupements accidentels comme ce peut être le cas pour des jonctions redondantes, et sont finalement retirés du voisinage de la jonction. Ainsi, lors de la phase suivante de relaxation "temporelle" seuls les voisins dont le mouvement est cohérent contribueront à renforcer les hypothèses d'appariements, comme le montre la figure 6.15.

Si une jonction d'un objet statique est initialement appariée avec une jonction d'un objet en mouvement à cause de la ressemblance accidentelle de leurs voisinages respectifs, cette hypothèse sera invalidée lorsqu'il s'agira d'estimer la cohérence des mouvements des voisins de cette jonction en fonction de cette hypothèse.

Cette alternance de deux types de relaxation afin de renforcer mutuellement chaque étiquetage permet d'accélérer la convergence et rend la mise en correspondance plus robuste en interdisant à des voisins incohérents de renforcer les hypothèses d'appariement.



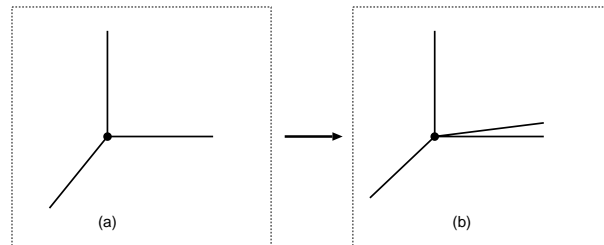
**Figure 6.15** - *Incohérence dans les voisinages perceptuels d'une jonction  $J$  et d'un correspondant possible  $L$ . Dans cette situation, le déplacement de  $Lv_1$  par rapport à  $L$  est cohérent avec celui de  $Jv_1$  et de  $J$ . La jonction  $Lv_2$  se comporte de manière incohérente - elle doit donc être retirée du voisinage de  $L$ .*

En pratique, chaque itération est composée de deux mises à jour des probabilités d'appariement pour une mise à jour des scores de groupement. Les voisinages perceptuels sont en effet jugés plus fiables que les hypothèses d'appariement car ils proviennent de l'environnement direct de chaque jonction. Enfin, toutes les deux itérations, les voisinages temporels et perceptuels sont étudiés afin d'éliminer les hypothèses dont la probabilité est trop faible. Cette dernière étape permet de réduire graduellement la complexité algorithmique des comparaisons entre voisinages.

Nous abordons à présent le détail des algorithmes de mise à jour des scores de groupement et des probabilités d'appariement, en commençant par les différentes mesures utilisées pour comparer les configurations entre jonctions.

### 6.3.2 Mesures de distances entre jonctions

Deux types de mesures sont utiles à la mise en correspondance de jonctions. D'une part, il est nécessaire d'évaluer le déplacement entre une jonction et les éléments de ses voisinages temporels et perceptuels. D'autre part, une mesure de similarité doit permettre, dans chaque cas, de comparer les déplacements de deux couples de jonctions. Enfin, chacune de ces mesures suppose la comparaison préalable d'une jonction avec une autre.



**Figure 6.16** - *Exemple de différences de groupement d'une même jonction dans deux images différentes. La mesure de similarité doit être suffisamment tolérante pour accepter ce genre de distorsion.*

#### – Coefficient de similarité entre deux jonctions - $d_s(J_1, J_2)$

Comparer deux jonctions consiste à déterminer dans quelle mesure leurs branches sont superposables. Le coefficient de similarité doit tenir compte, en particulier, du nombre de branches communes aux deux jonctions.

Comme le montre la figure 6.16, le résultat du groupement d'une même jonction peut être légèrement différent entre deux images. L'apparition, ou la disparition, de branches d'une jonction à l'autre est généralement due à des configurations entre branches à la limite des critères de groupement. La mesure

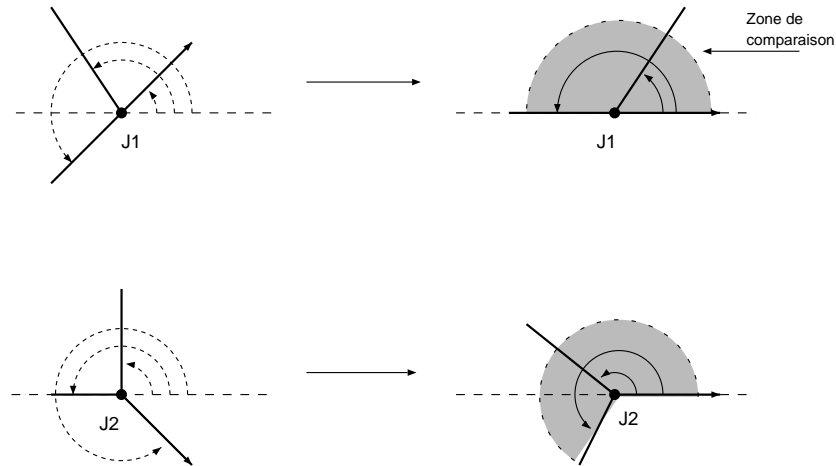
de similarité doit être suffisamment tolérante pour accepter de légères variations entre jonctions. Son principe consiste à évaluer le nombre de branches *consécutives* communes aux deux jonctions.

Chacune des branches est un segment caractérisé par sa longueur  $L$  et son orientation  $\lambda$ , exprimée par rapport à une direction de référence. Comme le montre la figure 6.17, cette direction de référence est, initialement, l'axe horizontal du repère de l'image. Afin de pouvoir prétendre à superposer les deux jonctions, il est nécessaire de les "aligner" de manière à faciliter la comparaison des branches. Une branche de chaque jonction est arbitrairement choisie comme référence par rapport à laquelle sont calculés les angles des branches restantes, noté  $\Theta$ .

On désigne par "zone de comparaison" l'écart angulaire maximum couvert entre la branche de référence et les branches restantes. Afin de faciliter la comparaison, les branches des deux jonctions sont triées par ordre croissant de leur angle  $\Theta$ .

Soit  $\mathcal{B}_1 = \{S_1^1, \dots, S_n^1\}$  et  $\mathcal{B}_2 = \{S_1^2, \dots, S_M^2\}$  les ensembles de branches de  $J_1$  et  $J_2$ , triées par rapport à leur angle. Les branches de référence sont ici  $S_1^1$  et  $S_1^2$ .

L'algorithme de comparaison des deux jonctions consiste, pour chaque branche  $S_i^1$  de  $J_1$ , à rechercher la branche  $S_j^2$  de  $J_2$  qui maximise la similarité entre les angles et longueurs de chaque branche.



**Figure 6.17** - *Alignement de deux jonctions par rapport à une direction de référence commune. Les jonctions sont initialement exprimées par rapport à l'axe horizontal (repère de l'image). La partie en gris signale la zone de comparaison entre les deux jonctions.*

La similarité entre deux branches est déterminée de la manière suivante :

$$\sigma_{i,j} = \frac{\sigma_{\Theta}(S_i^1, S_j^2) + \sigma_L(S_i^1, S_j^2)}{2.0}$$

avec :

$$\sigma_{\Theta}(S_i^1, S_j^2) = \begin{cases} 1 & \text{si } \Theta_i^1 = \Theta_j^2 \\ f_{\alpha}(\Theta_i^1, \Theta_j^2) & \text{sinon} \end{cases} \quad (6.7)$$

et :

$$\sigma_L(S_i^1, S_j^2) = \begin{cases} 1 & \text{si } L_i^1 = L_j^2 \\ f_{\alpha}(L_i^1, L_j^2) & \text{sinon} \end{cases} \quad (6.8)$$

La fonction de similarité  $f(x, y)$  utilisée est la suivante :

$$f_{\alpha}(x, y) = \exp(-\alpha \cdot (1.0 - \frac{\mathbf{Min}(x, y)}{\mathbf{Max}(x, y)})^2)$$

Cette fonction permet de ramener l'écart entre deux valeurs  $x$  et  $y$  à une mesure qui vaut 1 si  $x$  et  $y$  sont similaires et qui tend vers  $e^{-\alpha}$  sinon. Le choix de  $\alpha$  permet de régler la vitesse de l'atténuation. Sa valeur est fixée en pratique à  $\alpha = 2.0$  pour les deux similarités.

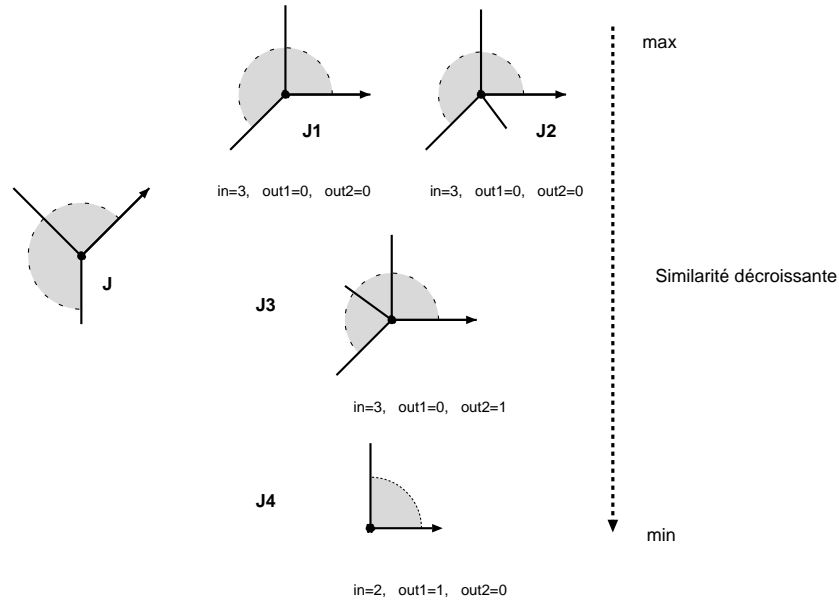
La comparaison se poursuit ensuite entre la branche  $S_{i+1}^1$  et les branches restantes  $\{S_{j+1}^2, \dots, S_M^2\}$  jusqu'à ce qu'il ne reste plus de branche à comparer.

On note alors  $in_{1,2}$  le nombre de branches similaires pour les deux jonctions,  $out_2$  le nombre de branches de  $J_2$  restées à l'intérieur de la zone de comparaison et qui n'ont pas trouvé de correspondant dans  $J_1$ , et enfin,  $out_1$  le nombre de branches de  $J_1$  qui n'ont pas pu être comparées avec celles de  $J_2$ .

La mesure de similarité finale est la somme normalisée des similarités des branches communes aux deux jonctions, pondérée par le nombre de branches délaissées par la comparaison.

$$\sigma(J_1, J_2) = \left( \frac{\sum_{similaires} \sigma_{i,j}}{in_{1,2}} \right) \cdot e^{-out_2} \cdot e^{-out_1}$$

L'idée principale pour cette mesure est d'accepter l'existence de branches superflues entre deux jonctions à comparer tant que ces branches se trouvent hors de la zone de comparaison. De plus, cette mesure établit une relation d'ordre entre jonctions en encourageant les comparaisons entre "petites" et "grandes" jonctions (par rapport au nombre de branches).



**Figure 6.18** - *Similarité entre une jonction  $J$  et quatre candidats  $J_1$  à  $J_4$ . Les directions de référence sont les branches marquées d'une flèche. Dans chaque cas, la zone de comparaison est signalée en gris. Les jonctions sont classées par ordre décroissant de similarité.*

Les exemples de la figure 6.18 montrent l'influence des coefficients  $e^{-out_1}$  et  $e^{-out_2}$ . Les similarités  $\sigma(J, J_1)$  et  $\sigma(J, J_2)$  sont ici identiques. La branche supplémentaire de  $J_2$  ne gêne en rien la superposition de  $J_2$  par  $J$ . Elle n'est donc pas prise en compte dans le calcul de similarité.

Par contre, la jonction  $J_3$  pourrait être superposable avec  $J$  mais sa branche superflue se trouve dans la zone de comparaison et gêne donc la superposition. La mesure de similarité  $\sigma(J, J_3)$  est ainsi plus faible.

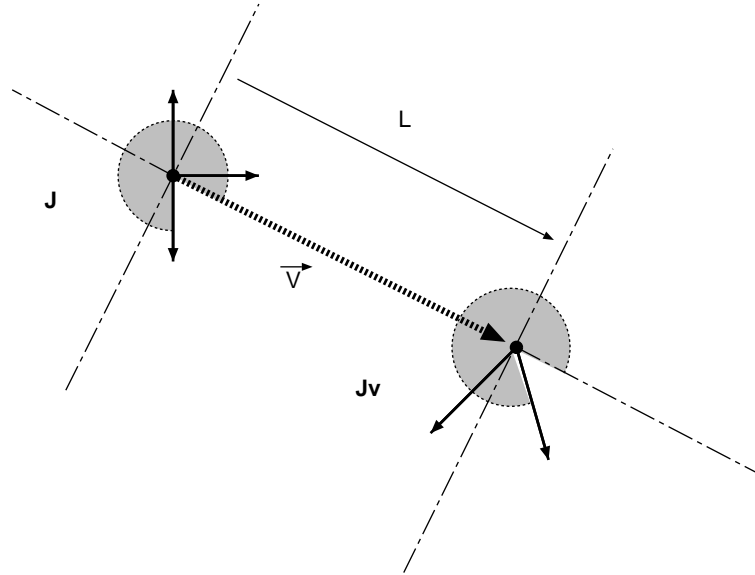
Enfin, la jonction  $J_4$  n'a pas assez de branches pour être correctement comparée à  $J$ . La mesure de similarité  $\sigma(J, J_4)$  est ici inférieure à  $\sigma(J_4, J)$ . Cet exemple illustre la notion d'ordre imposée sur la comparaison de jonctions.

Cette mesure de similarité n'est définie qu'en fonction d'une direction de référence pour chaque jonction. L'algorithme complet de comparaison consiste finalement à prendre successivement comme référence chaque branche de chaque jonction et de conserver la configuration pour laquelle la similarité est maximale.

- Déplacement entre une jonction et un voisin perceptuel -  $d_G(J, Jv)$

Cette mesure permet, lors de la relaxation, de comparer les configurations de couples de jonctions. Une configuration entre deux jonctions est définie selon les mêmes principes que la comparaison de jonctions.





**Figure 6.19** - Configuration entre deux jonctions  $J_1$  et  $J_2$ . Chaque jonction est alignée avec le vecteur  $J_1 \rightarrow J_2$ . Les zones de comparaison sont ici encore signalées en gris.

Soient deux jonctions  $J \in \mathcal{J}_1$  et  $Jv \in \mathcal{J}_1$ . La relation de groupement qui lie ces deux jonctions est notée  $d_G(J, Jv)$ . Cette relation est associée au score de groupement  $S_G(J, Jv)$ .

Comme le montre la figure 6.19, les angles des branches de chaque jonction sont exprimés par rapport au vecteur  $\vec{V} = \overrightarrow{J, Jv}$ . Afin de tenir compte des proportions de la configuration, les longueurs des branches de chaque jonction sont rapportées à la distance  $L = \|\overrightarrow{J, Jv}\|$ .

Exprimés ainsi, les angles et les longueurs des branches sont autant de paramètres propres au couple de jonctions. Les valeurs de ces paramètres sont de plus invariants par transformation du couple de jonctions en rotation, translation et changement d'échelle.

Soient  $(J_1, Jv_1)$  et  $(J_2, Jv_2)$  deux couples de jonctions. La similarité entre ces deux couples est alors simplement définie par le produit des similarités des jonctions prises deux à deux.

$$\sigma_G((J_1, Jv_1), (J_2, Jv_2)) = \sigma(J_1, J_2) \cdot \sigma(Jv_1, Jv_2)$$

Ce produit est comparable à une probabilité conditionnelle. Il représente la compatibilité entre la jonction  $J_1$  et  $J_2$  en supposant que  $Jv_1$  correspond à  $Jv_2$ . Sa valeur est maximale lorsque les jonctions sont effectivement similaires deux à deux.

– Déplacement entre une jonction et un voisin temporel -  $d_M(J, L)$

On désigne par  $d_M(J, L)$  la transformation d'une jonction  $J \in \mathcal{J}_1$  en son correspondant  $L \in \mathcal{J}_2$ . Cette relation est associée à la probabilité d'appariement  $S_M(J, L)$ .

La complexité de la transformation  $d_M(\cdot, \cdot)$  est liée au choix d'hypothèses sur les conditions de mise en correspondance.

Le cas le plus simple correspond à une vision stéréoscopique, avec un faible écart angulaire entre les points de vues et des objets relativement éloignés de l'observateur. La seule transformation envisageable est alors réduite à une simple translation.

Dans un cas plus général où la scène est toujours éloignée de l'observateur mais où les objets peuvent éventuellement bouger librement, la transformation  $d_M$  est la composition d'une translation, d'une rotation et éventuellement d'un changement d'échelle. Il faudrait alors rechercher les paramètres de la plus petite transformation permettant de changer les branches de  $J$  en branches de  $L$ .

Enfin, le cas le plus général est celui d'une transformation projective. Dans ce dernier cas, les angles ne sont plus garantis et la mesure de similarité entre jonctions devrait être redéfinie.

Dans le cadre de notre application, nous nous plaçons dans l'hypothèse d'une paire d'images stéréoscopiques sans mouvement de rotation. Dans ces conditions, résumées par la figure 6.20, la comparaison entre deux appariements possibles est définie de la manière suivante.

Soit  $\vec{V}_{J,L}(\ell_V, \Theta_V)$  la translation qui transforme  $J$  en  $L$ , et  $\vec{U}_{Jv_1, Lv_1}(\ell_U, \Theta_U)$  qui transforme  $Jv_1$  en  $Lv_1$ . La compatibilité entre les deux appariements est donnée par :

$$\sigma_M((J, L), (Jv_1, Lv_1)) = \frac{1}{3}(f_\alpha(\ell_V, \ell_U) + f_\alpha(\Theta_V, \Theta_U) + \sigma_G((J, Jv_1), (L, Lv_1)))$$

Comme dans le cas précédent, cette mesure est comparable à la probabilité conditionnelle que  $J$  soit apparié avec  $L$  en supposant que  $Jv_1$  le soit avec  $Lv_1$ . On retrouve ici le même formalisme entre groupement perceptuel ( mesure de similarité  $\sigma_G$  ) et mise en correspondance ( mesure de similarité  $\sigma_M$  ).

### 6.3.3 Relaxation “temporelle”

Le principe de la relaxation “temporelle” consiste à renforcer la probabilité d’un appariement entre deux jonctions  $J \in \mathcal{J}_1$  et  $L \in \mathcal{J}_2$  si celui-ci est cohérent avec les appariements des voisins de  $J$ . La probabilité d’appariement entre  $J$  et  $L$  est renforcée par les probabilités d’appariement des voisins de  $J$  dont le déplacement est similaire à  $d_M(J, L)$ .

Soient  $Jv_i \in \mathcal{V}(J)$  un élément du voisinage perceptuel de  $J$  et  $L_j \in \mathcal{M}(Jv_i)$  l’un des correspondants de  $Jv_i$ .

La mise à jour des probabilités d’appariement est effectuée de la manière suivante :

$$S_M(J, L) = S_M(J, L) \cdot \left(1.0 + \sum_{(Jv_i, L_j)} \{S_M(Jv_i, L_j) / d_M(J, L) \equiv d_M(Jv_i, L_j)\}\right)$$

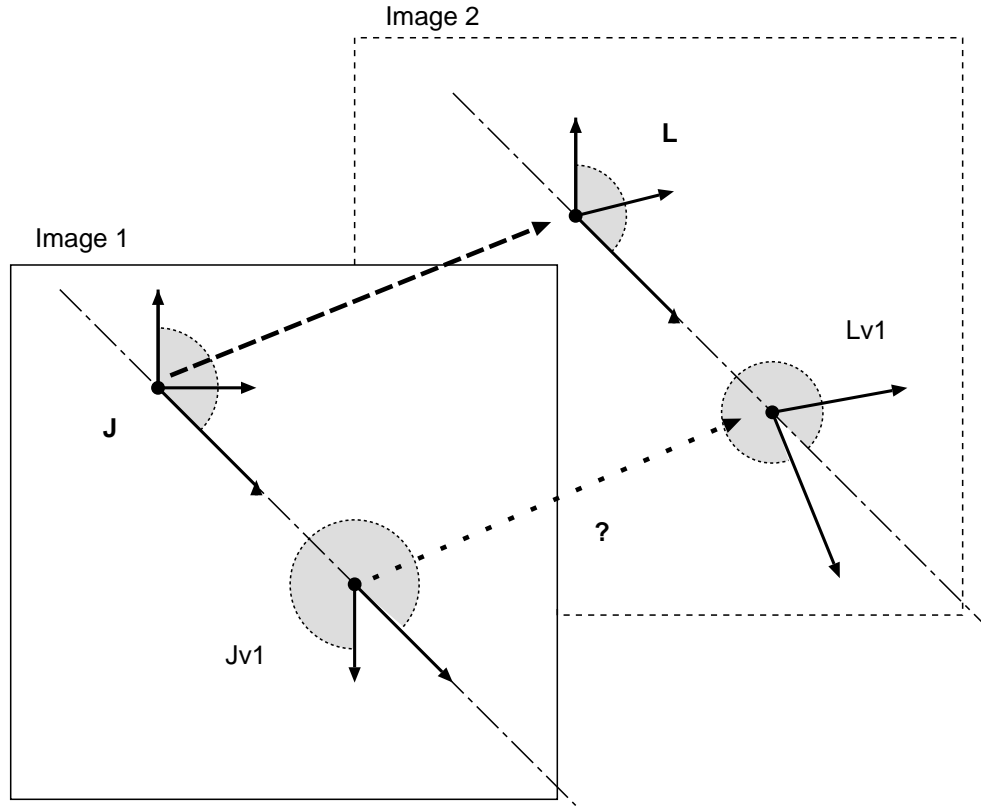
Ainsi, chaque voisin  $Jv_i$  de  $J$  apporte une contribution à la probabilité de l’appariement entre  $J$  et  $L$  si et seulement si il existe un appariement semblable parmi les voisins temporels de  $Jv_i$ .

Le voisinage temporel de chaque jonction  $J \in \mathcal{J}_1$  contient initialement toutes les jonctions de  $\mathcal{J}_2$ . Les appariements initiaux entre jonctions sont initialisés par la mesure de similarité entre jonctions notée  $d_s(J, L)$ .

$$\forall J \in \mathcal{J}_1, \forall L \in \mathcal{J}_2, \longrightarrow L \in \mathcal{V}(J) \text{ et } S_M(J, L) = d_s(J, L)$$

Après avoir reçu les contributions des voisinages des jonctions, les scores d’appariement sont normalisés de manière à ce que leur somme soit égale à 1 pour une jonction donnée.

Le résultat de cet algorithme est, pour chaque jonction  $J \in \mathcal{J}_1$ , un ensemble de correspondants les plus probables et une probabilité d’appariement pour chacun de ces correspondants.



**Figure 6.20** - *Comparaison entre hypothèses d'appariements. En supposant que  $Lv_1$  est apparié avec  $Jv_1$ , dans quelle mesure peut on considérer que  $J$  est associé à  $L$  ?*

**Algorithme 6.2** : Relaxation temporelle

```

début
  pour Chaque jonction  $J \in \mathcal{J}_1$  faire
    pour Chaque hypothèse d'appariement  $L \in \mathcal{M}(J)$  faire
      Contributions  $\leftarrow 0$ 
      pour Chaque voisin  $Jv_i \in \mathcal{V}(J)$  faire
        pour Chaque hypothèse d'appariement  $L_j \in \mathcal{M}(Jv_i)$  faire
          si le déplacement  $d_M(J, L)$  similaire au déplacement  $d_M(Jv_i, L_j)$  alors
            Contributions  $\leftarrow$  Contributions +  $S_M(Jv_i, L_j)$ 
        si Contributions  $\neq 0$  alors
          % Renforcement de la probabilité  $S_M(J, L)$ 
           $S_M(J, L) = S_M(J, L) \cdot (1.0 + \text{Contributions})$ 
fin
  
```

### 6.3.4 Relaxation “perceptuelle”

Cette relaxation sur le voisinage des jonctions est semblable à la phase de relaxation “temporelle”. Pour chaque jonction  $J \in \mathcal{J}_1$ , elle consiste à rechercher parmi les voisins de  $J$  ceux dont les correspondants dans  $\mathcal{J}_2$  sont les plus cohérents avec les correspondants de  $J$ .

Les mécanismes de comparaison des configurations entre jonctions et de mise à jour des scores de groupement sont identiques à ceux de la relaxation “temporelle”. La différence majeure vient du choix des jonctions utiles au renforcement de chaque hypothèse. Pour la relaxation temporelle, il s’agissait des jonctions  $Lv_1 \in \mathcal{V}(L)$ , éléments du voisinage perceptuel des jonctions  $L$ . Dans le cas de la relaxation spatiale, ces jonctions sont  $L_1 \in \mathcal{M}(Jv_1)$ , appariements possibles pour les voisins des jonctions  $J$ .

Le voisinage perceptuel de chaque jonction est défini par une relation de colinéarité avec les branches de chaque jonction. Les groupements sont ici supposés équiprobables. Ils sont donc initialisés par la quantité  $(\frac{1}{n})$ , ‘n’ étant le nombre de voisins du voisinage initial.

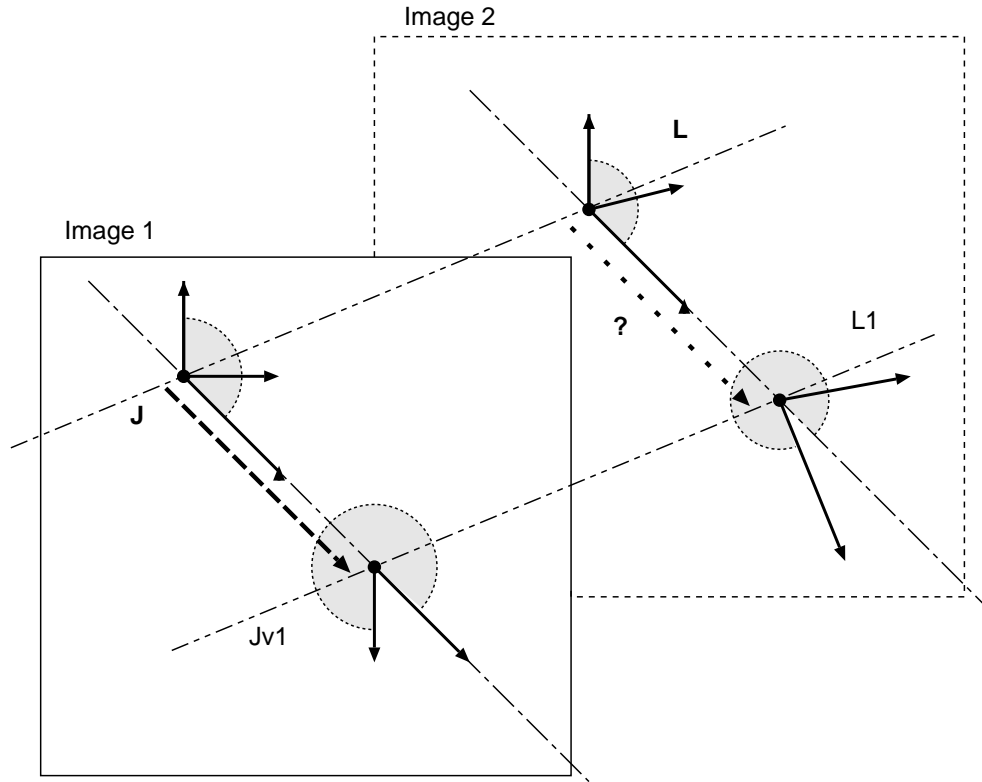
Le renforcement du score de groupement est, quant à lui, différent de celui utilisé pour la relaxation temporelle.

$$S_G(J, Jv_i) = S_G(J, Jv_i) \cdot (0.1 + 2.0 \cdot \sum_{(L, L_j)} \{S_G(L, L_j) \mid d_G(J, Jv_i) \equiv d_G(L, L_j)\})$$

Un coefficient inférieur à 1 permet ici d’atténuer les scores des groupements qui n’ont reçu aucun soutien des jonctions de leur voisinage temporel. Ces scores ne sont pas normalisés de manière à éviter d’éliminer trop vite des voisins dont la contribution serait faible.

La figure 6.14 montre un exemple d’initialisation de ce type de voisinage. Dans cet exemple, le centre de la jonction  $J_6$  est aligné avec une branche de  $J$ . Elle fait donc partie de son voisinage initial. Mais comme il y a peu de chances pour que cet alignement soit présent dans une autre vue de la même scène ( même en conservant un écart faible entre les points de vue ), la relaxation perceptuelle aura pour conséquence d’éliminer  $J_6$  du voisinage de  $J$ .

Le résultat de cette phase de relaxation est donc, pour chaque jonction  $J$ , un ensemble de voisins perceptuels appartenant probablement au même objet que  $J$  ( jonctions cohérentes avec un déplacement rigide ). On retrouve ici le principe Gestaltiste de groupement perceptuel par “comportement commun” (*common fate*). Les scores de groupement traduisent la qualité de chaque association.



**Figure 6.21** - Comparaison entre hypothèses de groupements. En supposant que  $L$  est groupée avec  $L_1$ , dans quelle mesure peut-on considérer que  $J$  est groupée à  $Jv_1$  ?

**Algorithme 6.3** : Relaxation perceptuelle

```

début
  pour Chaque jonction  $J \in \mathcal{J}_1$  faire
    pour Chaque voisin  $Jv_i \in \mathcal{V}(J)$  faire
      Contributions  $\leftarrow 0$ 
      pour Chaque hypothèse d'appariement  $L \in \mathcal{M}(J)$  faire
        pour Chaque voisin  $L_j \in \mathcal{V}(L)$  faire
          si la configuration  $d_G(J, Jv_i)$  similaire à la configuration  $d_G(L, L_j)$ 
            alors
              Contributions  $\leftarrow$  Contributions +  $S_G(L, L_j)$ 
        si Contributions  $\neq 0$  alors
          % Renforcement du score de groupement  $S_G(J, Jv_i)$ 
           $S_G(J, Jv_i) = S_G(J, Jv_i) \cdot (0.1 + 2.0 * Contributions)$ 
fin
  
```

### 6.3.5 Résultats de mise en correspondance

Comme c'était le cas pour les précédentes étapes de groupement, nous avons d'abord appliqué cet algorithme de mise en correspondance de jonctions sur des scènes artificielles puis sur des images réelles. Dans chaque cas, la limite du nombre d'itérations a été fixé à dix cycles de relaxation temporelle et spatiale. Chang et Aggarwal ont montré en effet que la conjugaison des deux relaxations et l'élimination progressive des appariements les plus mauvais renforcent la convergence. En plus de cette condition limite, l'analyse des appariements permet d'arrêter les cycles de relaxation lorsque les voisinages temporels des jonctions sont réduits à un voisin dominant.

Afin d'éliminer une grande partie des ambiguïtés d'appariements, l'algorithme est appliqué successivement de l'image  $I_1$  vers l'image  $I_2$ , puis de  $I_2$  vers  $I_1$ . Seules les jonctions pour lesquelles il existe un appariement réversible entre les deux images sont conservées. Les meilleures hypothèses d'appariements sont enfin extraites selon leur score.

Les figures de la page 234 visualisent les résultats d'appariement pour une scène simple. Les deux rectangles ont ici des déplacements très différents d'une image à l'autre, illustrés par la figure 6.24 des vecteurs de disparité entre les deux ensembles de jonctions. Cet algorithme permet donc de s'affranchir d'une contrainte de déplacement cohérent pour l'ensemble des jonctions grâce au groupement perceptuel des jonctions présentant des déplacements semblables.

Les images des pages 235 et 236 illustrent bien la robustesse de l'appariement malgré des variations importantes dans l'orientation des branches des jonctions. En effet, la reconstitution des arêtes à partir de l'image de contours subit des perturbations importantes dues essentiellement à la discrétisation des contours. C'est le cas, en particulier, pour les arêtes des fenêtres (jonctions 13, 14 et 15).

Cet exemple illustre également l'utilité des jonctions virtuelles. Même si elles sont issues de groupements accidentels, des jonctions apparaissant sur des points de vues différents constituent des points de repères utiles à la mise en correspondance. Les jonctions 9 et 11 sont ainsi appariées bien qu'elles n'aient pas de signification physique particulière.

Les figures des pages 237 et 238 représentent un exemple d'appariement dans des conditions réelles. Malgré l'apparente simplicité de la scène, les différences de contraste et le bruit de chaque image introduisent des différences importantes entre les contours des deux scènes. Comme le montre la figure 6.31, le groupement de segments permet de compenser une majeure partie des discontinuités le long des contours. La principale difficulté de l'appariement dans le cas de scènes réelles vient des jonctions sans correspondant, comme la jonction 3 dans cet exemple. La souplesse de l'algorithme d'appariement attribue tout de même un correspondant à cette jonction. Ce résultat est toutefois intéressant dans la mesure où les jonctions correctement appariées présentent des différences importantes de localisation (représentées par la figure 6.33) et d'orientation de leurs branches.

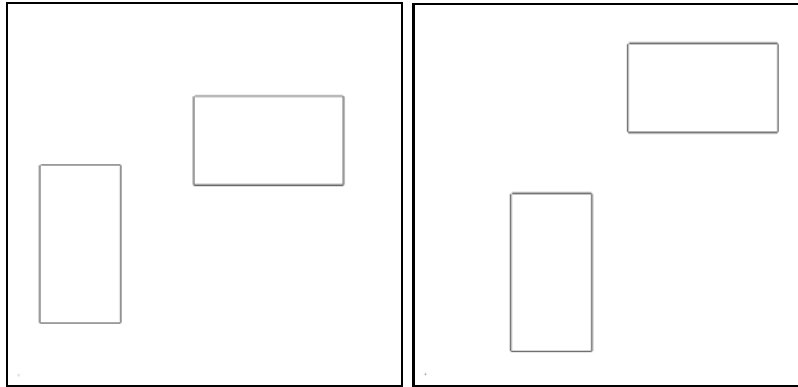


Figure 6.22 - *Appariement simple - rectangles*

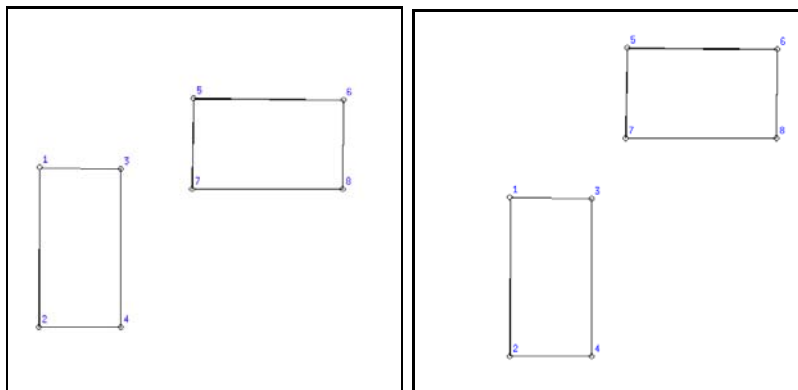


Figure 6.23 - *Appariement simple - rectangles - appariements*

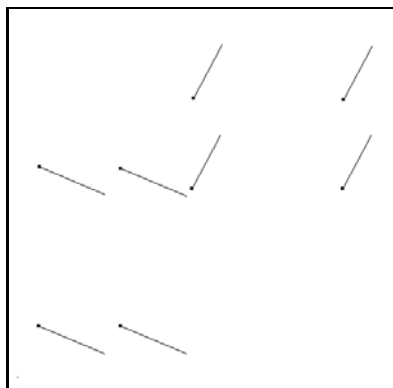


Figure 6.24 - *Appariement simple - rectangles - vecteurs de déplacement*



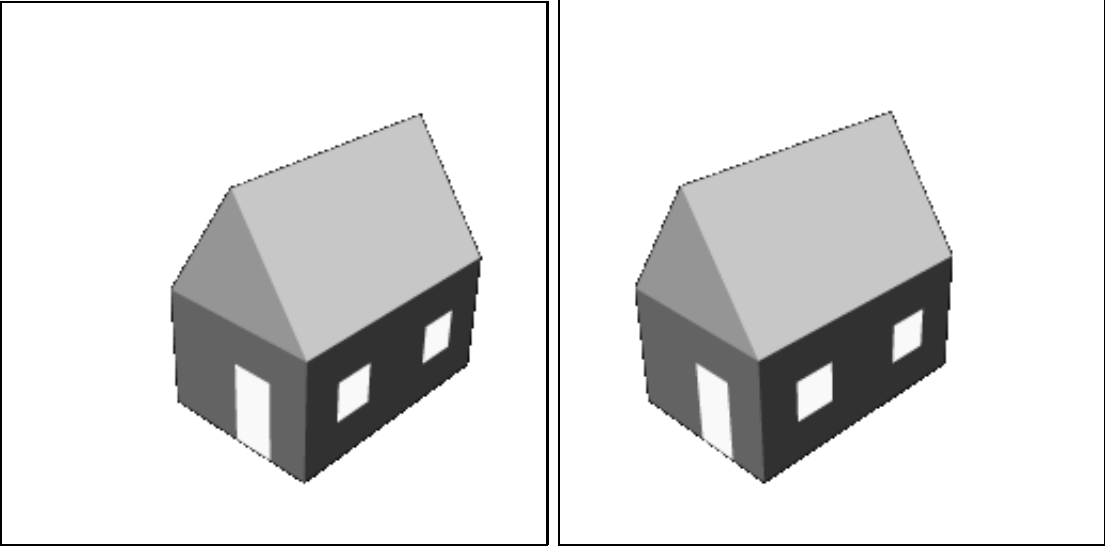


Figure 6.25 - Appariement complexe - maison

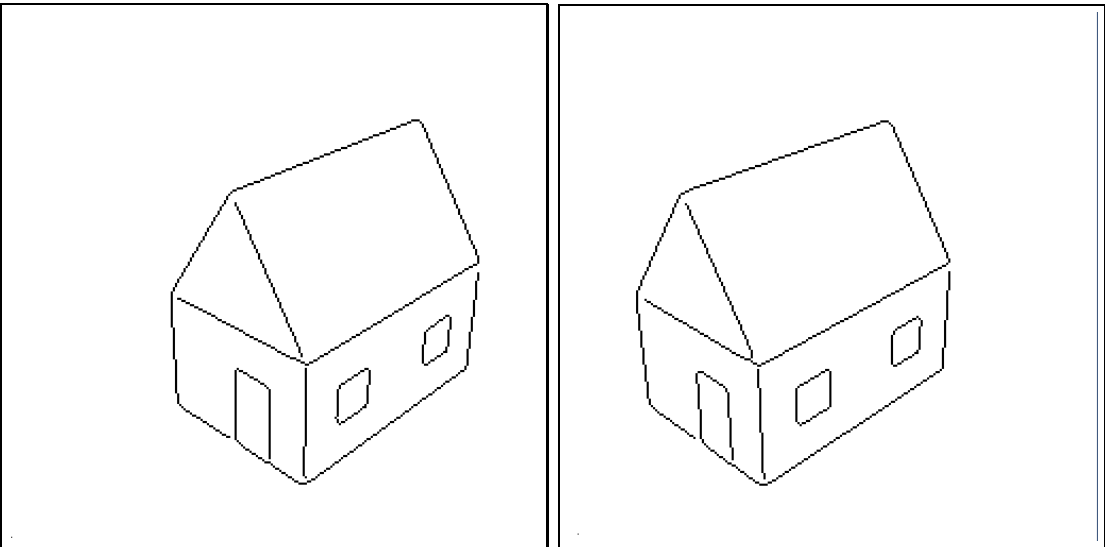
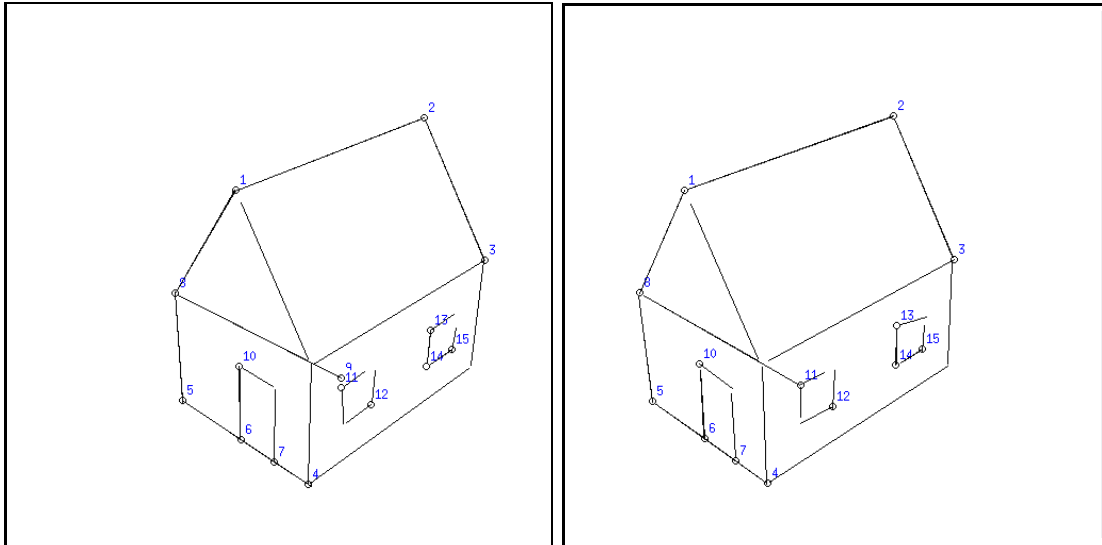
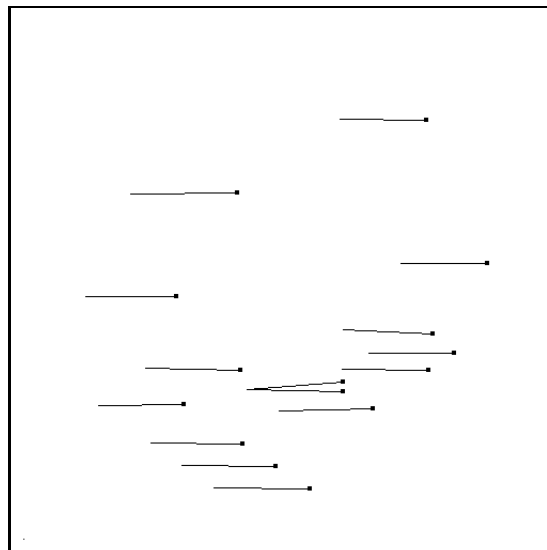


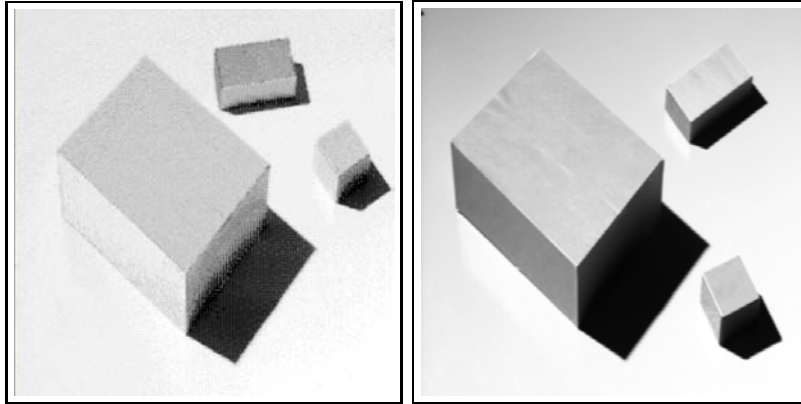
Figure 6.26 - Appariement complexe - maison - détection de contours



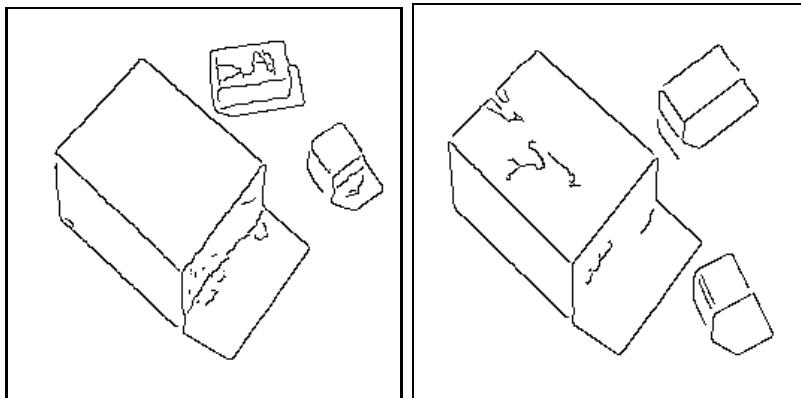
**Figure 6.27** - Appariement complexe - maison. On peut noter l'appariement correct de la jonction 13 dans les deux images, malgré les différences d'orientation des branches. Les jonctions virtuelles 9 et 11 sont superposées dans l'image de droite.



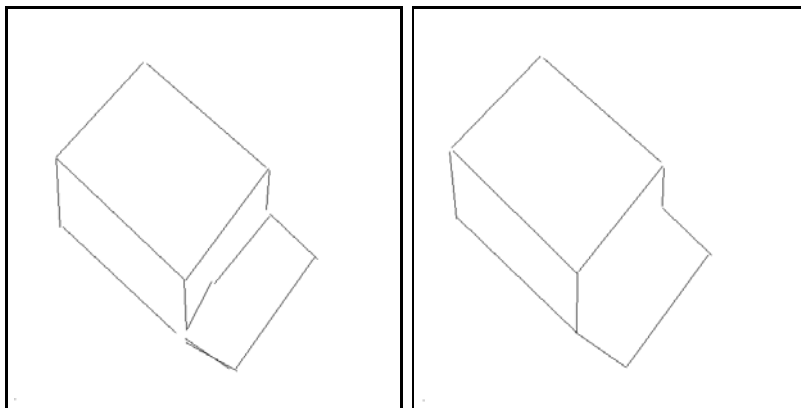
**Figure 6.28** - Appariement complexe - maison - vecteurs de déplacement



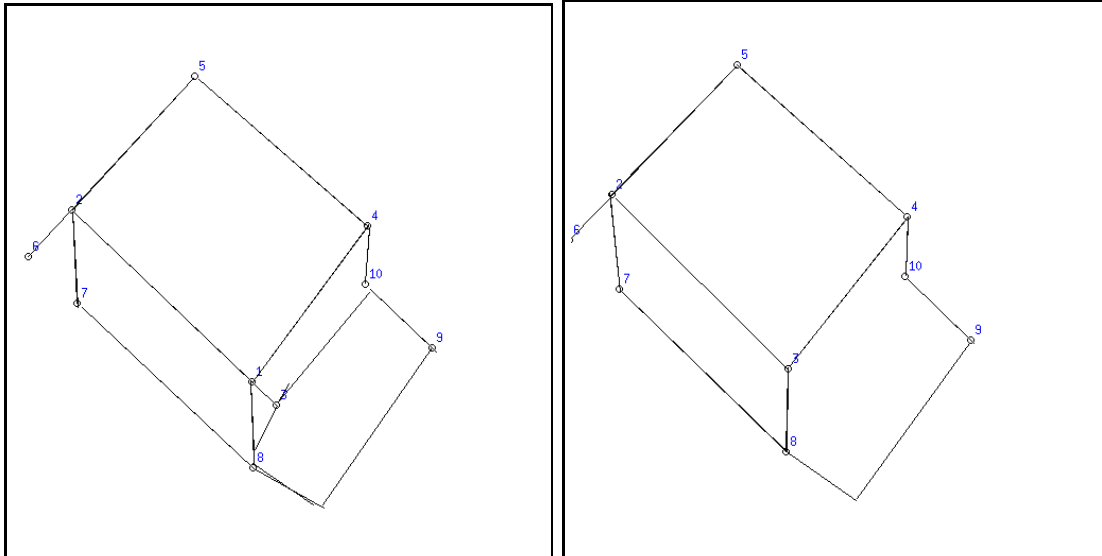
**Figure 6.29** - *Appariement de jonctions - cube*



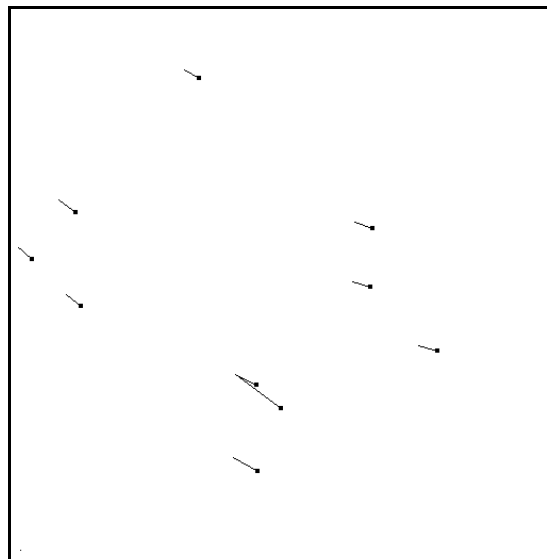
**Figure 6.30** - *Appariement de jonctions - cube - détection de contours*



**Figure 6.31** - *Appariement de jonctions - cube - hypothèses de segments après groupement*



**Figure 6.32** - Appariement de jonctions - cube. On peut noter l'appariement correct de la jonction 10 malgré le passage de 3 à 2 branches d'une scène à l'autre. Les jonctions 1 et 3 sont superposées dans l'image de droite. La jonction 6 est une jonction virtuelle.



**Figure 6.33** - Appariement de jonctions - cube - vecteurs de déplacement

## 6.4 Conclusion

Nous avons montré dans ce chapitre comment utiliser les éléments de représentations extraits à partir des chaînes saillantes d'une image pour constituer des structures plus globales et les manipuler. En particulier, nous avons présenté une méthode de groupement de segments et de points d'intérêt selon un ensemble de jonctions. L'utilité de ce type de groupement de haut niveau est finalement illustré à l'aide d'un algorithme de mise en correspondance de jonctions.

Le principal avantage de cette approche est sa flexibilité et sa robustesse face au bruit grâce à la généralité des méthodes de groupement. Il serait intéressant de la développer selon deux axes privilégiés.

D'une part, la construction d'hypothèses de groupement selon des règles plus complexes. Des propriétés telles que la symétrie ou le parallélisme apportent en effet des contraintes plus fortes pour la mise en correspondance de structures d'intérêt. Leur complexité élevée assure un faible nombre de structures et une recherche d'autant plus rapide.

D'autre part, la mise en correspondance pourrait bénéficier de la hiérarchie de groupements. En effet, l'appariement de jonctions ne peut qu'être le début d'une mise en correspondance hiérarchique. En servant de centre d'intérêt, les jonctions appariées peuvent être utilisées pour mettre en correspondance les jonctions de leur voisinage. Ces appariements seraient ensuite propagés à chaque branche, jusqu'aux pixels de l'image correspondant à ces branches si l'application exige une grande précision. Cette démarche, semblable à celle de [Venkateswar et Chellappa, 1995], consiste en une succession de prédictions et vérifications d'appariements entre structures, les appariements de haut niveau servant de centre d'attention pour niveaux inférieurs.

Enfin, une amélioration plus matérielle est nécessaire à court terme. En tant que structures complexes, les jonctions sont moins nombreuses que de pixels ou des segments, mais leur représentation en termes de mémoire est beaucoup plus volumineuse. Comme pour la manipulation de grands nombres d'hypothèses évoquée en fin du chapitre précédent, il serait utile ici aussi d'adopter des représentations internes plus adaptées à l'utilisation de structures complexes.

