

Résumé

La perception directe est intrinsèquement incomplète pour la reconnaissance: les objets peuvent se présenter dans des positions ambiguës, et les systèmes perceptifs eux-mêmes sont limités dans leur sensibilité. La reconnaissance d'objets nécessite une phase exploratoire.

Les objets peuvent être caractérisés par la structure de variation des apparences par modification dynamique, soit du point de vue sur l'objet, soit des paramètres d'analyse du stimulus visuel. L'ensemble des apparences d'objets peut alors être organisé sous la forme d'un graphe d'aspects.

Les graphes d'aspects classiques sont difficiles à construire, et d'utilisation pratique limitée. Cet article en propose une version probabilisée qui permet de traduire le problème de leur détermination en celui d'une estimation empirique de matrices markoviennes. Des grandeurs statistiques conduisent alors à la définition d'une complexité globale des problèmes de reconnaissance d'objets 3D.

Mots clés

Vision active, graphes d'aspects, chaînes de Markov, grandes déviations des tests d'hypothèses.

Abstract

Direct perception is incomplete: objects may show ambiguous appearances, and sensors have a limited sensitivity. As a consequence, the recognition of complex 3D objects must have an exploratory nature.

Appearance variations when the viewpoint is modified or when the sensor parameters are changed are characteristic of the object. They can be organized in the form of an aspect graph.

Standard geometric aspect graphs are difficult to build. This article presents a generalized probabilistic version of this concept. Tri-dimensional object recognition is transformed here in a problem of Markov chain estimation which yields a statistical characterization of aspect graphs complexity.

Keywords

Active vision, aspect graphs, Markov chains, large deviations of hypotheses testing.

1 Graphes d'aspects géométriques

Les scènes tri-dimensionnelles sont appréhendées par des organes sensoriels bidimensionnels, naturels (rétine) ou artificiels (caméra). Cette réduction de dimensionnalité, génératrice d'ambiguïté, ne rend accessible les objets à reconnaître que selon certaines vues ou certains aspects: deux objets distincts peuvent ainsi présenter des apparences semblables dans des conditions d'observation particulières.

Un deuxième problème associé au point de vue sensoriel concerne la variété infinie des apparences d'un même objet. La multiplicité de configurations liée à la position de l'observateur est encore augmentée par le grand nombre d'objets susceptibles d'être visés, et par les conditions d'observations (illuminations, occlusions, ombres...). Les espaces de vues à manipuler sont ainsi souvent de très grandes dimensions et la difficulté majeure est de les rendre utilisables en pratique.

Ces deux particularités de la vision tri-dimensionnelle — multiplicité et ambiguïté des apparences — doivent être prises en compte explicitement dans les modèles et ont donné lieu à une grande variété d'approches. L'une des approches les plus étudiées a été celle d'une modélisation de la structure des apparences sous la forme d'un *graphe d'aspects*.

1.1 Variation du point de vue

La multiplicité des apparences peut être indexée par un espace de contrôle de dimension finie. L'ensemble des apparences géométriques d'un objet, par exemple, peut être référencé par un espace à deux dimensions — la sphère de vue — dans le cas d'une projection orthographique ou à trois dimensions, pour une projection en perspective.

L'application qui à un point de l'espace de contrôle fait correspondre une apparence est localement continue pour une distance de similarité à définir, sauf en certaines zones frontières très particulières. Celles-ci partitionnent l'espace de contrôle en sous-ensembles pour lesquels les apparences correspondantes sont déclarées équivalentes ou semblables. L'ensemble des apparences équivalentes peut être muni d'une structure de graphe symbolisant les relations entre les différentes classes dans l'espace de contrôle: c'est le *graphe d'aspects*.

L'exemple standard de graphe d'aspects est fourni par la

ceaux. Le contour apparent, correspondant aux points de l'objet pour lesquels la direction de projection sur la rétine de la caméra est tangente à la surface extérieure ou rencontre une arête, se projette en une courbe différentiable, sauf en un nombre fini de points. Pour certaines directions d'observations, le contour d'occlusion subit des variations brusques dans la structure de ses singularités, manifestant ainsi une discontinuité des apparences [20, 19].

Deux contours d'occlusion seront déclarés semblables ou de même type qualitatif s'ils peuvent être transformés l'un dans l'autre par un difféomorphisme. Cette relation de similitude définit des classes d'équivalence structurellement stables qui correspondent à la partition de l'espace de vue par les lieux de catastrophes.

L'origine de la variation des singularités du contour apparent peut être formalisée rigoureusement [28], permettant ainsi une approche systématique de sa détermination pour certaines classes d'objets (polyèdres, solides de révolution, surfaces algébriques).

L'intérêt majeur de cet objet théorique est d'autoriser un rapprochement des modèles vers une perception plus authentique, plus liée à la multiplicité irréductible des apparences tout en conservant un statut objectif qui lui permet d'être formalisé précisément.

Les graphes d'aspects classiques, d'origine géométrique, permettent de réduire l'indétermination du point de vue selon deux directions: par partitionnement de l'espace de vue en classes d'équivalences définies par les propriétés de stabilité structurelle du contour d'occlusion, et par structuration différentielle globale de ces classes d'apparences.

La définition d'une catégorisation intrinsèque des apparences est produite essentiellement par des propriétés locales de la variation du contour d'occlusion. L'ensemble des catégories d'apparences, dans leur globalité, peut également être caractérisé plus spécifiquement: c'est ici qu'intervient la structure de graphe. Elle sera *différentielle*, par opposition à son utilisation plus courante comme structure associative et marquera les distinctions entre aspects.

Les différences entre catégories de contours d'occlusion, à cause de la typologie finie des catastrophes élémentaires, ne sont pas quelconques: un graphe d'aspects est en général de faible degré. Sa structure est donc caractéristique de l'objet et sera porteuse d'information potentiellement discriminante.

En plus d'organiser les catégories, le graphe révèle l'unité de l'objet. L'ensemble des aspects est fini pour des objets délimités par des surfaces lisses par morceaux, et permet ainsi de totaliser, moyennant une relation d'équivalence définie par la stabilité structurelle du contour d'occlusion, la diversité des apparences. La donnée de l'ensemble des aspects et de leur organisation structurée par un graphe détermine qualitativement l'objet observé.

Pour résumer, l'intérêt des graphes d'aspects tient à trois propriétés: il manifeste la possibilité d'une catégorisation intrinsèque des données sensorielles, il décrit la structure différentielle particulière qui la régit, il fournit une unité à la diversité des apparences.

Les propriétés esthétiques des graphes d'aspects se détachent cependant dès que l'on cherche à les exploiter pra-

gérer et à la présence d'autres sources de diversité phénoménale.

1.2 Complexité des graphes d'aspects

La complexité des graphes d'aspects limite leur utilisation pratique pour la reconnaissance pour deux raisons: leur construction est coûteuse et leur comparaison peut exploser combinatoirement.

Les algorithmes de constitution des graphes d'aspects pour des objets quelconques sont coûteux en temps de calcul: les études exhaustives se sont en général consacrées à quelques objets simples [26, 34] révélant les difficultés d'une automatisation.

L'évaluation exacte de la complexité intrinsèque des graphes d'aspects a été calculée par S. Petitjean [25] pour des surfaces algébriques lisses par morceaux. Pour des objets constitués de n quadriques, par exemple, la borne supérieure du nombre d'aspects est donnée par un polynôme dont le monôme dominant est égal à $\frac{64}{9}n^{12}$ dans le cas d'une projection orthographique et à $\frac{512}{27}n^{18}$ dans le cas d'une perspective. Ces bornes prodigieuses expliquent en partie la complexité de détection des aspects.

La complexité de constitution des graphes d'aspects peut être attaquée une fois pour toute pour chaque objet, et stockée dans une mémoire appropriée: telle est du moins une des motivations qui ont poussé à l'étude d'algorithmes génériques de calcul des aspects. Leur utilisation effective pour la reconnaissance sous la forme d'une indexation d'une base de données de graphes touche un autre type de complexité.

On peut suspecter ce problème d'être difficile, au sens d'une complexité algorithmique. La question de décider si deux graphes sont isomorphes n'a jamais été démontrée comme appartenant à la classe des problèmes NP-complets, mais on ne connaît pas non plus d'algorithme en temps polynomial correspondant [13, 18]. La structure de graphe seule ne permettra pas en général de décider quel est l'objet observé.

1.3 Autres sources d'indétermination

La majorité des recherches se sont penchées sur une caractérisation géométrique des apparences avec la notable exception de [8] qui introduit explicitement une notion d'échelle mais comme paramètre d'imprécision de mesure géométrique.

Lorsque les données sensibles sont sous la forme d'images de pixels, la définition de lieux singuliers est problématique. Une solution consiste à généraliser la notion d'aspect en introduisant une échelle d'analyse irréductible qui ne peut être spécifiée à l'avance. Une telle définition des aspects peut être trouvée dans [15].

La changement de la caractérisation des apparences modifie le critère de différenciation des données sensibles. Les aspects pris comme catégories d'apparences ne sont plus définis de manière intrinsèque et dépendent d'un certain nombre de paramètres indéterminés: le nombre d'échelles utilisées, une erreur d'approximation ou le nombre d'aspects prototypiques, par exemple. Les catégories partitionnent certes l'espace des données sensibles par leur

plus d'une différenciation intrinsèque.

L'indétermination de la catégorisation est cependant source d'informations utiles. La variation d'échelle d'analyse, par exemple, donne une indication sur la structure des formes observées: elle est au fondement des théories sur les espaces d'échelles («scale-space») [10, 21, 22, 12]. La détermination d'une échelle optimale d'analyse des signaux sensibles nécessite cependant la définition d'un critère de sélection extrinsèque, dont la validité dépend d'un contexte d'utilisation. Une solution consiste alors à contrôler les sources d'indétermination, à les faire participer directement à la diversité des apparences. La notion d'échelle n'apparaît plus alors comme un paramètre des modes d'analyse mais comme dimension de contrôle de la diversité des apparences. Une catégorisation intrinsèque peut alors être définie par différenciation du contour d'occlusion, selon des critères qui devront être précisés.

1.4 Résumé des points problématiques

Les propriétés intrinsèques du mode de production des apparences fondent une catégorisation dont les éléments sont définis par stabilité structurelle du contour d'occlusion. Un graphe permet de donner une structure caractéristique aux catégories d'apparences et une unité. Dans les graphes d'aspects «classiques», inspirés de la théorie des singularités, la source d'indétermination, unique, est le point de vue perceptif, c'est-à-dire la direction d'observation et la position de la caméra.

Une discussion générale sur l'intérêt des graphes d'aspects a été transcrite en un article déjà «ancien» de CVGIP [9]. Les participants semblaient s'accorder sur trois problèmes caractéristiques:

- *complexité*: la construction exacte des graphes est coûteuse et le nombre d'aspects peut être très grand, même pour des objets simples.
- *imprécision*: les méthodes empiriques de construction des graphes sont peu robustes. Elles demandent une segmentation qui ne peut être réalisée parfaitement, et éludent les questions d'échelles d'analyse.
- *pauvreté*: l'information utile est réduite au contour d'occlusion de l'objet, et néglige d'autres indices visuels, en particulier photométriques, potentiellement caractéristiques.

Ces trois gammes de problèmes restent d'actualité et n'ont pas été résolus dans leur globalité. La section suivante montrera comment en aborder certains en généralisant les graphes à un ensemble de transitions de probabilités.

2 Graphes d'aspects probabilisés

Le fondement mathématique des graphes d'aspects classiques est la théorie des singularités des applications différentiables. Celle-ci impose une approche continue du formalisme qui n'est pas toujours en adéquation avec les contraintes liées à l'utilisation de l'outil informatique: la définition rigoureuse de lieux de singularités, en particulier, est délicate.

La constitution pratique des graphes d'aspects est coûteuse, même pour des objets simples, et sa complexité devient très

Outre ces questions importantes d'ordre pratique, se posent des difficultés conceptuelles: le contour d'occlusion de l'objet peut ne pas porter suffisamment d'information discriminante et toutes les sources de diversité phénoménales ne sont pas contrôlables.

Cette section propose pour aborder ces problèmes de probabiliser la structure du graphe d'aspects sous la forme d'une chaîne de Markov. Elle examinera d'abord l'intérêt de munir les aspects d'une distribution de probabilité, puis justifiera le choix d'une représentation markovienne.

2.1 Probabiliser les aspects

La recherche en vision artificielle s'est principalement intéressée jusqu'à présent à une caractérisation géométrique ou topologique des phénomènes: cette démarche était nécessaire pour formaliser les intuitions et a permis de rendre exploitable dans des applications concrètes des domaines mathématiques relativement abstraits. L'essor récent des invariants géométriques en est la meilleure illustration.

Les théories géométriques sont essentiellement qualitatives: les notions de dimension, en tant que mesure d'extension spatiale, ou d'échelle ne lui sont pas naturelles. Deux types d'études quantitatives cependant ont été menées: elles concernent la robustesse des caractérisations géométriques, et l'évaluation de leur utilité.

Les deux considérations de robustesse et d'utilité se rejoignent dans le concept développé par certains auteurs de vues canoniques ou génériques d'objets. Elles peuvent être caractérisées selon deux axes: stabilité et vraisemblance.

Le principe théorique à la base de la définition d'une vraisemblance de vues est simple et s'appuie sur une propriété géométrique assez intuitive: la distribution de probabilité de la mesure d'un angle entre deux lignes ayant un point commun est maximale lorsque la direction de visée est orthogonale au plan défini par l'angle, si la sphère de vue est échantillonnée uniformément [2, 5]. La distribution de probabilité est très nettement «piquée» à l'endroit de la valeur maximale. Cette propriété de concentration de la distribution a été exploitée dans [33] pour reconnaître des objets polyédriques très simples.

Freeman [11] s'est plutôt penché sur le concept de vue générique dont la stabilité est définie par une formulation bayésienne.

Les deux principes de vraisemblance et de stabilité ont été généralisés par [35] et utilisés par [14] pour déterminer les vues canoniques d'objets lisses simples caractérisés par des mesures de leur contour d'occlusion. A partir de leurs notions formalisées, ils démontrent, dans le cas d'une distribution uniforme des directions d'observation, que les vues stables seront celles qui seront observées le plus vraisemblablement, résultat plutôt intuitif. D'une manière générale, les directions de vues canoniques sont orthogonales au plan défini par les deux axes de plus grande inertie de l'objet.

L'exploitation analytique précise de ces approches, cependant, est difficile dès que la complexité des objets augmente. D'autre part, et plus fondamentalement, l'idée même de la définition d'une vue canonique par un critère de stabilité est contradictoire avec les principes de description des aspects comme région de l'espace de contrôle déli-

aspect est représenté canoniquement par une apparence de stabilité maximale: ce principe de description des objets ne dit rien cependant sur l'organisation générale des aspects, sur les capacités prédictives de leur structure.

Une notion de vue générique ou canonique est utile lorsqu'une seule image d'un objet ou une seule direction d'observation, par exemple, ne sont accessibles. Le cadre d'étude du problème de reconnaissance qui nous intéresse est d'autoriser une variation contrôlable de la direction d'observation: c'est la structure de la catégorisation des apparences qui est la grandeur critique dans ce contexte.

Les variables susceptibles d'être probabilisées sont donc les aspects eux-mêmes, et non les apparences ou les vues. L'origine des variations de leur distribution selon les objets a une origine géométrique, mais vu de l'observateur, la seule réalité accessible est l'observation catégorielle d'une apparence.

Cette réduction de la géométrie à une distribution de probabilité de catégories d'apparences permet de s'affranchir grossièrement des sources de diversité non contrôlables. Mais surtout, elle permet de quantifier l'utilité empirique des vues. Certains aspects seront plus rarement observés que d'autres, en fonction de la géométrie des objets, et leur probabilité conditionnelle définit alors une vraisemblance. Nous verrons dans la section suivante comment les exploiter statistiquement.

2.2 Représentation markovienne des graphes d'aspects

La catégorisation des apparences n'est utilisable que structurée. Dans le cas des aspects définis comme classe d'équivalence d'une relation de stabilité structurelle, leur organisation caractéristique est décrite par un graphe qui indique quelles transitions, quels événements visuels sont susceptibles d'être observés. Le graphe, par la spécificité de sa connectivité, possède ainsi une capacité prédictive élémentaire.

La description probabiliste des aspects a permis de donner une mesure quantitative de l'utilité empirique de leur capacité descriptive. De la même manière, il est possible de munir la structure de leur organisation d'une caractérisation probabiliste: certains événements visuels seront plus utiles que d'autres, plus stables et plus vraisemblables dans leur expérience. Ici encore, nous ne nous intéresserons pas à une analyse géométrique de la stabilité des vues, mais à leur interprétation probabiliste.

Pour accéder véritablement à une formulation probabiliste des graphes d'aspects, il faut se pencher sur leur origine dynamique. Les transitions entre aspects ne peuvent être détectées qu'en changeant la direction d'observation, ou plus généralement en modifiant continûment les paramètres de contrôle. Dit plus trivialement, il est nécessaire d'agir pour qu'il se passe quelque chose. Un aspect ne changera que s'il est poussé à le faire.

Formellement, si s_t désigne l'aspect observé à un certain temps discret t , soumis à une action a_t de l'espace de contrôle, la probabilité d'observer un événement visuel du type $s_t \rightarrow s_{t+1}$ pour un objet o_k sera donné par une probabilité de transition $\mathbf{P}_k[s_{t+1} | s_t, a_t]$. Le graphe d'aspect

contrôlée.

Ce modèle probabilisé de graphes d'aspects peut encore être simplifié en considérant une loi d'action stationnaire, ce qui est toujours possible car laissé au contrôle de l'observateur. Les probabilités de transitions seront alors moyennées en

$$\mathbf{P}_k[s_{t+1} | s_t] = \int_a \mathbf{P}_k[s_{t+1} | s_t, a] d\mu(a)$$

où $\mu(a)$ est la loi d'action stationnaire.

Les probabilités de transition entre aspects $\mathbf{P}_k[s | s']$ seront considérées alors comme des *modèles* markoviens d'objet. Ils ne constituent pas un substitut, ni un résumé des apparences: ils caractérisent leurs relations *a priori* de structuration lorsque l'on modifie les paramètres de contrôle selon une certaine loi fixe sous-jacente et un certain processus de catégorisation.

La structure markovienne des aspects permettra de prendre en considération deux types de caractéristiques: des mesures d'utilité empirique, coinduisant à la définition de vraisemblances, et la mise en évidence d'événements visuels. Les transitions de probabilité nulles et celles de probabilité strictement positives n'ont pas alors la même nature phénoménale. Les premières manifestent une organisation particulière des aspects: elles sont structurantes. Les secondes mesurent quantitativement la fréquence de leur expérience empirique, c'est-à-dire leur potentialité informative. Cette distinction de nature sera exploitée dans la discrimination d'objets, et dans la caractérisation de la complexité d'un problème de reconnaissance d'objet.

2.3 Construction des représentations markoviennes de graphes d'aspects

Dans l'hypothèse de validité d'une représentation markovienne de graphe d'aspects, une méthode naturelle de détermination des paramètres du modèle est une estimation empirique. Le calcul exact du graphe d'aspects se transforme ainsi en un problème d'apprentissage. La complexité à gérer est alors d'une autre nature: problème de dimensionnement, temps d'apprentissage, erreurs d'approximation, erreurs de généralisation. Elle est en général plus directement contrôlable et quantifiable.

L'estimation des matrices markoviennes est conditionnée à une certaine variation des paramètres de contrôle. Sans mouvement, sans dynamique, les transitions ne peuvent être observées. Le principe d'estimation de la chaîne de Markov est alors simplement de produire aléatoirement des actions en présence d'un objet et de collecter les fréquences des transitions d'aspects observées $n(s, s', o_k)$. Les probabilités seront obtenues simplement par normalisation des fréquences selon l'estimateur du maximum de vraisemblance [3]. D'un point de vue informatique, c'est une représentation compacte qui ne contient que des entiers.

Les questions liées à l'estimation des modèles markoviens sont alors de deux types: 1) lorsque l'on dispose d'un temps d'apprentissage limité, comment peut-on exploiter pratiquement l'expérience disponible, et comment juger de la qualité de la reconnaissance produite. 2) lorsque l'on dispose

on décider que l'estimation est de qualité suffisante, et selon quels critères. Ces questions sont abordées dans [16].

2.4 Généralisation des graphes d'aspects

La définition d'un graphe d'aspects demande la spécification de quatre éléments:

- une source contrôlable d'indétermination.
- un mode de caractérisation ou de description des apparences.
- un critère de différenciation de la caractérisation des apparences.
- un processus de catégorisation des données empiriques en aspects fondé sur ce critère.

Dans les graphes d'aspects «classiques», le processus de catégorisation est induit des lieux de discontinuité du critère de différenciation des apparences, caractérisées par la structure différentiable du contour d'occlusion de l'objet. Le graphe décrit l'organisation des variations d'aspects lorsque la source d'indétermination, ici la direction d'observation, est modifiée continûment.

L'espace de contrôle peut être augmenté pour prendre en considération des paramètres caractérisant le mode d'analyse des apparences, par exemple une échelle de filtrage variable. Le processus de catégorisation définissant les aspects nécessitera alors un paramétrage extrinsèque, c'est-à-dire indépendant des données sensorielles, caractéristique de sa complexité ou de sa sensibilité.

Le graphe d'aspects généralisé pourra alors être muni d'une structure probabiliste sous la forme d'une représentation markovienne, mesure de son utilité empirique. Elle permettra de s'affranchir des difficultés techniques de construction du graphe en transposant le problème de caractérisation géométrique en une question d'estimation statistique. La représentation markovienne sera alors toujours relative à une loi d'action stationnaire caractéristique, manifestation de sa nature dynamique sous-jacente.

3 Discrimination de modèles markoviens

Le problème de reconnaissance d'objets se traduit alors en celui de discrimination de modèles markoviens, dont une réalisation de la chaîne est produite par une séquence d'actions générées selon une loi stationnaire.

L'observation de certaines transitions entre aspects sera rare, voire de probabilité nulle, pour certains objets: elles seront ainsi *sélectives*. Quand les informations extraites sont bien adaptées à un problème de reconnaissance spécifique, certains aspects seront sélectifs en éliminant de l'ensemble des objets potentiels ceux qui ne peuvent leur être associés. Dans les situations extrêmes, l'observation de certains aspects permet ainsi l'indexation univoque d'un objet par rejet de tous les autres.

Tant qu'aucune transition sélective n'a été observée, le système continuera son activité exploratoire, et décidera de l'objet observé en fonction d'un seuil selon un test du rapport de vraisemblance.

Ainsi, dans une méthode de reconnaissance active, deux régimes de décision statistique entrent en concurrence: le pre-

modifie par interaction avec l'environnement, le deuxième utilise des données discriminantes *a priori* pour sélectionner dynamiquement les objets potentiellement observables.

3.1 Test d'hypothèses

Le principe de la reconnaissance par discrimination de modèles markoviens est de tester si un objet est plus vraisemblable que tous les autres en examinant chaque couple d'objets potentiellement observables, une fois collectés les formes sensibles générées par l'activité de reconnaissance. Un test d'hypothèse $D_T^\Phi(k, k')$ est une variable aléatoire à valeurs dans $\{0, 1\}$. Elle prend la valeur 1 s'il est estimé en fonction de la séquence d'observations $\Phi_T = (S_0, S_1, \dots, S_t)$ que l'objet o_k est observé préférentiellement par rapport à $o_{k'}$, et 0 dans le cas contraire.

On décidera alors que l'objet o_{k^*} est reconnu si:

$$\forall k \neq k^*, \quad D_T^\Phi(k^*, k) = 1$$

Dans le cas contraire, on prendra la décision de ne pas donner de nom d'objet (rejet d'hypothèse).

On définit les erreurs de première et deuxième espèce $\alpha_T(k, k')$ et $\beta_T(k, k')$ par:

$$\begin{aligned} \alpha_T(k, k') &= \mathbf{P}_k [D_T^\Phi(k, k') = 0] \\ \beta_T(k, k') &= \mathbf{P}_{k'} [D_T^\Phi(k, k') = 1] \end{aligned}$$

qui mesurent l'erreur moyenne que l'on risque de commettre en rejetant l'hypothèse de vraisemblance alors que l'objet o_k est observé et l'erreur moyenne lorsque l'on accepte la vraisemblance de o_k alors que l'on observe l'objet $o_{k'}$.

La comparaison de deux modèles markoviens $p_k(i, j)$ et $p_{k'}(i, j)$, représentations de graphes d'aspects généralisés, sera réalisée par un rapport de vraisemblance. On définit les variables aléatoires $Y_t(k, k')$ comme étant les contributions de chaque nouvelle action au logarithme du rapport de vraisemblance. Elles sont définies par:

$$Y_t(k, k') = r_{kk'}(S_t, S_{t-1}) = \log \left[\frac{p_{k'}(S_t, S_{t-1})}{p_k(S_t, S_{t-1})} \right]$$

Le logarithme normalisé du rapport de vraisemblance vaut alors:

$$L_T(k, k') = \frac{1}{T} \sum_{t=1}^T Y_t(k, k') \quad (1)$$

Si l'on se donne un seuil λ , que l'on peut rendre dépendant de chaque couple d'objet étudié (k, k') , on considérera que l'objet o_k est plus vraisemblablement observé que $o_{k'}$ après avoir collecté T échantillons si:

$$L_T(k, k') < \lambda$$

Les erreurs de première et deuxième espèce sont alors définies par:

$$\begin{aligned} \alpha_T(k, k') &= \mathbf{P}_k [L_T(k, k') \in [\lambda, +\infty[] \\ \beta_T(k, k') &= \mathbf{P}_{k'} [L_T(k, k') \in]-\infty, \lambda] \end{aligned}$$

avons utilisé pour analyser le comportement du modèle de reconnaissance active est le suivant [23, 17, 6]:

Théorème 1 *Le test du rapport de vraisemblance pour les transitions d'une chaîne de Markov de probabilités de transition strictement positives avec seuil constant λ a des erreurs de première et deuxième espèces $\alpha_T(k, k')$ et $\beta_T(k, k')$ vérifiant:*

$$\begin{aligned}\lim_{T \rightarrow \infty} \frac{1}{T} \log \alpha_T(k, k') &= -\Lambda_{kk'}(\lambda) < 0 \\ \lim_{T \rightarrow \infty} \frac{1}{T} \log \beta_T(k, k') &= \lambda - \Lambda_{kk'}(\lambda) < 0\end{aligned}$$

où la fonction de taux $\Lambda_{kk'}(\lambda)$ est définie par

$$\Lambda_{kk'}(\lambda) = \sup_{x \in [0, 1]} (x\lambda - \chi_{kk'}(x)) \quad (2)$$

La fonction $\chi_{kk'}$ est donnée par

$$\chi_{kk'}(x) = \log \rho(\Pi_k(r_{kk'}, x))$$

où $\rho(\cdot)$ est la valeur propre de Perron-Frobenius de la matrice:

$$\Pi_k(r_{kk'}, x) = \left\{ p_k(i, j) \cdot \exp[x \cdot r_{kk'}(i, j)] \right\}_{i, j \in \mathcal{S}}$$

et $r_{kk'}(i, j) = \log[p_{k'}(i, j)/p_k(i, j)]$.

Ce théorème caractérise le fait que le test du rapport de vraisemblance génère des erreurs qui tendent vers 0 avec une vitesse exponentielle. La fonction de taux $\Lambda_{kk'}(\lambda)$ caractérise alors la vitesse logarithmique de convergence.

Un cas intéressant est celui où le seuil de décision λ est nul, rendant les vitesses logarithmiques de convergence égales pour les erreurs de première et deuxième espèces. On peut alors démontrer [6] que le test du rapport de vraisemblance maximise la vitesse de convergence logarithmique du risque de Bayes $R_T(k, k') = \mathbf{P}_k \alpha_T(k, k') + \mathbf{P}_{k'} \beta_T(k, k')$. La fonction de taux pour un seuil nul $\Lambda_{kk'}(0)$ est usuellement appelée borne ou information de Chernoff.

3.2 Cas général

Le résultat de grandes déviations nécessite la stricte positivité des matrices markoviennes. Dans les cas que nous rencontrerons, la structure markovienne générique d'un problème de reconnaissance active sera en pratique constituée de matrices de transitions assez creuses mais irréductibles, dont les états sont en général peu discriminants, mais dont le graphe des transitions est lui caractéristique des objets. Les rapports de vraisemblance ne seront pas définis pour l'observation de certaines transitions, et ne vaudront que conditionnellement à l'ensemble des transitions de probabilité strictement positives.

Ce résultat doit être réaménagé pour prendre en compte explicitement les différences structurelles entre les matrices markoviennes. On définit alors l'ensemble des transitions

$$\mathcal{T}_k = \{(i, j) \in \mathcal{S}^2 / p_k(i, j) > 0\}.$$

En conditionnant par l'ensemble des transitions admissibles pour chaque couple d'objets, $\mathcal{T}_{kk'} = \mathcal{T}_k \cap \mathcal{T}_{k'}$, la décision générale peut être décomposée en deux parties:

$$D_T^\Phi(k, k') = \begin{cases} C_T^\Phi(k, k') & \text{si } \Phi_T \in \mathcal{T}_{kk'}^T \\ S_T^\Phi(k, k') & \text{si } \Phi_T \notin \mathcal{T}_{kk'}^T \end{cases}$$

La première partie, *comparative*, ne considère que les transitions observées strictement de probabilité strictement positives pour deux objets ($\phi_T \in \mathcal{T}_{kk'}^T$), la deuxième, *sélective*, prend une décision de rejet dès qu'une transition observée est de probabilité nulle pour l'un des objets. Les erreurs de reconnaissance peuvent alors être réécrites en:

$$\begin{aligned}\alpha_T(k, k') &= \alpha_T^C(k, k') \mathbf{P}_k[\Phi_T \in \mathcal{T}_{kk'}^T] \\ &\quad + \alpha_T^S(k, k') \mathbf{P}_k[\Phi_T \notin \mathcal{T}_{kk'}^T] \\ \beta_T(k, k') &= \beta_T^C(k, k') \mathbf{P}_{k'}[\Phi_T \in \mathcal{T}_{kk'}^T] \\ &\quad + \beta_T^S(k, k') \mathbf{P}_{k'}[\Phi_T \notin \mathcal{T}_{kk'}^T]\end{aligned}$$

Les décisions sélectives $S_T^\Phi(k, k')$, lorsqu'une transition de probabilité nulle a été observée, permettent d'assurer que la partie correspondante des erreurs de première et deuxième espèces sont nulles: il ne peut y avoir de rejet erroné lorsque les transitions admissibles ont été parfaitement déterminées, ni de fausses attributions. Les erreurs, dans ce cas idéal, ne proviendront alors que de la partie comparative fonction des vraisemblances.

Pour évaluer le comportement global des erreurs avec cette décision composée, on utilise le résultat suivant qui caractérise la vitesse logarithmique de sortie d'un ensemble de transitions d'une chaîne de Markov [31]:

Théorème 2 *Soit $\mathcal{T} \subset \mathcal{S}^2$ un sous-ensemble de transitions d'une chaîne de Markov \mathbf{P}_k , et Φ_T une trajectoire de longueur $T + 1$ de cette chaîne, alors il existe un nombre $\rho_k(\mathcal{T}) < 1$ tel que:*

$$\lim_{T \rightarrow \infty} \frac{1}{T} \log \mathbf{P}_k[\Phi_T \in \mathcal{T}^T] = \log \rho_k(\mathcal{T}) < 0$$

Ce nombre est la valeur propre de Perron-Frobenius de la matrice $p_k(i, j)$ dont on a annulé les coefficients correspondant aux transitions de $\mathcal{S}^2 \setminus \mathcal{T}$.

On en déduit le résultat global sur le comportement asymptotique des erreurs de reconnaissance:

$$\begin{aligned}\alpha_T(k, k') &= \alpha_T^C(k, k') \mathbf{P}_k[\Phi_T \in \mathcal{T}_{kk'}^T] \sim A \cdot e^{-T \cdot v_\alpha(k, k')} \\ \beta_T(k, k') &= \beta_T^C(k, k') \mathbf{P}_{k'}[\Phi_T \in \mathcal{T}_{kk'}^T] \sim B \cdot e^{-T \cdot v_\beta(k, k')}\end{aligned}$$

où les vitesses logarithmiques globales sont données par:

$$\begin{aligned}v_\alpha(k, k') &= \Lambda'_{kk'}(\lambda) - \log \rho_k(\mathcal{T}_{kk'}) > 0 \\ v_\beta(k, k') &= \Lambda'_{kk'}(\lambda) - \lambda - \log \rho_{k'}(\mathcal{T}_{kk'}) > 0\end{aligned}$$

La fonction de taux du test du rapport de vraisemblance $\Lambda'_{kk'}(\lambda)$ sera calculée à partir des expressions (2) par conditionnement de la chaîne de Markov aux éléments de $\mathcal{T}_{kk'}$.

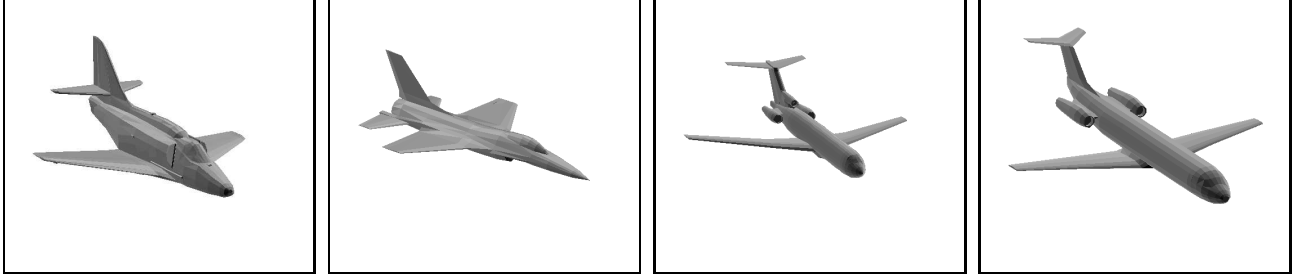


FIG. 1 - Exemple de quatre objets à reconnaître. Les deux objets de gauche sont des avions militaires. Les deux objets de droite sont des avions civils.

Les vitesses logarithmiques globales de convergence des erreurs $v_\alpha(k, k')$ et $v_\beta(k, k')$ donnent une mesure de la complexité du problème de reconnaissance d'objets 3D. La plus petite de ces vitesses permettra alors de caractériser globalement la difficulté du problème, au moins dans sa limite asymptotique.

4 Expérimentations

La séquence d'aspects générés par modification du point de vue selon une loi d'action donnée est une chaîne de Markov d'espace d'états fini, dont la matrice sera caractéristique de l'objet observé. La reconnaissance d'objets 3D est alors assimilable à la discrimination de modèles markoviens. Les résultats théoriques de la section précédente permettront alors de fournir une caractérisation de la complexité du problème de reconnaissance d'objets 3D.

Le problème de reconnaissance sera testé dans le cadre d'une simulation informatique. Ce moyen d'étude permet de s'affranchir de certaines difficultés expérimentales pour concentrer les efforts de modélisation et d'analyse sur une question particulière en réduisant les autres problèmes à des schémas simplifiés. La validité des résultats de simulation est donc toujours restreinte, mais permet néanmoins de révéler certains phénomènes.

Les mesures de complexité des problèmes de reconnaissance seront évaluées sur trois exemples de discrimination choisis parmi 4 objets (FIG. 1): couples d'avions militaires, d'avions civils et couple «mixte». On étudiera plus particulièrement l'influence du nombre d'aspects utilisés sur la complexité effective de la reconnaissance.

4.1 Organisation générale de la simulation

Le problème que nous étudierons dans les simulations est la reconnaissance d'objets tri-dimensionnels par contrôle dynamique des sources de diversité perceptives — direction d'observation — et représentatives — échelle d'analyse du contour d'occlusion.

Quatre modules définissent donc l'activité de reconnaissance proprement dite:

- Un processus d'analyse des apparences par extraction de la structure de points saillants du contour d'occlusion, paramétré par une échelle d'analyse et un nombre maximal d'aspects chargé de déterminer la catégorie d'apparence observée à une certaine échelle.
- Un processus de décision chargé d'analyser la séquence des aspects détectés et de décider de poursuivre la génération d'actions ou de proposer le nom

de l'objet.

- Un actionneur capable de modifier la direction d'observation par mouvement de la caméra.
- Un actionneur susceptible de changer l'échelle d'analyse du contour d'occlusion.

Les deux premiers modules contiennent des éléments qui seront spécifiés par apprentissage: les vecteurs prototypiques d'aspects et les représentations markoviennes de graphes d'aspects. L'environnement d'objets tri-dimensionnels sera assimilé à un générateur de contours d'occlusions d'objets représentés par des modèles CAO. Les caractéristiques utiles étant en première approximation, et dans l'hypothèse d'une capacité idéale de segmentation des formes, insensibles aux variations d'illumination, les seuls sources de diversité sensorielles seront paramétrées par un couple d'angles en coordonnées sphériques, dans le cas d'une projection orthographique.

4.2 Structure des graphes d'aspects

Les matrices stochastiques de représentations markoviennes de graphes d'aspects seront assez creuses. Elle contiennent une diagonale dominante, c'est-à-dire dont les valeurs des coefficients sont nettement supérieures aux autres ($\forall i \neq j, p_k(i, i) \gg p_k(i, j)$). Elle contiennent un assez grand nombre de transitions de probabilité nulle, qui seront potentiellement sélectives. La répartition des valeurs des probabilités strictement positives est en général bimodale: les probabilités correspondant aux éléments de la diagonale de la matrice seront supérieures à 0.2 ou 0.3. La chaîne de Markov aura tendance à rester dans les mêmes états.

Les éléments théoriques présentés dans la section précédente ont porté une attention particulière à la présence de transitions dites sélectives car possédant un pouvoir de rejet infini: elles seront nombreuses en pratique et concerneront tous les aspects.

La répartition des transitions de probabilités nulles est déterminante pour la discrimination entre deux objets: dès qu'une transition sélective a été observée, une décision de rejet peut être prise, sans erreur. Lorsque le nombre d'aspects utilisés augmente, les matrices de transitions ont tendance à se creuser. Elles contiennent un nombre de transitions de probabilités strictement positives qui augmente linéairement ou sous-linéairement en fonction du nombre d'aspects. Cette croissance lente a pour conséquence de faire décroître la densité globale de transitions de probabilités strictement positives.

sélectives est également intéressante à observer et révèle un phénomène intéressant de saturation (FIG. 2). Quand le nombre d'aspects augmente, le degré moyen de chaque matrice de transitions augmente également, mais selon un régime sous-linéaire, indice d'une connectivité décroissante. Le degré moyen du graphe des transitions comparatives, c'est-à-dire le graphe des transitions communes aux deux objets, diminue après avoir atteint un maximum alors que le degré moyen de la partie sélective augmente plus régulièrement, même si une inflexion est observable.

Cette évolution de la connectivité globale des matrices de transition, c'est-à-dire du nombre moyen d'arcs par noeud du graphe des transitions indique que les graphes de transition se séparent pour devenir structurellement discriminants. Le nombre de transitions comparatives entre aspects croît jusqu'à atteindre une valeur limite. Ce noyau limite se répartit ensuite entre les nouveaux aspects créés, sans ajouter de nouvelles transitions sélectives efficaces. Le graphe des transitions comparatives devient ainsi de moins en moins connecté lorsque le nombre d'aspects utilisés augmente.

Dans les analyses menées jusqu'à présent seuls les comportements généraux de la structure des graphes de transition ont été analysés. Des différences entre les couples d'objets sont cependant notables.

Le problème de discrimination d'avions civils contient significativement moins de transitions sélectives pour un nombre sensiblement équivalent de transitions de probabilités non nulles. Le nombre de transitions comparatives est en conséquence plus important. Ce résultat n'est guère étonnant car les objets à discriminer ont de grandes similarités: ils diffèrent essentiellement par les dimensions de leur structure, alors que les avions militaires ont des structures «visiblement» dissemblables.

Le comportement structurellement plus comparatif du problème de discrimination des deux avions civils peut être l'indice d'une plus grande difficulté de discrimination. Pour le vérifier, il devient nécessaire de quantifier le comportement des erreurs de reconnaissance. Les outils de grandes déviations et d'analyse des matrices positives présentés dans la section précédente permettront de réaliser cette évaluation des performances de reconnaissance: ils font l'objet de la section suivante.

4.3 Complexité des graphes d'aspects

Pendant la phase de reconnaissance, deux régimes statistiques entrent en concurrence: un premier qui permet de prendre une décision de rejet dès qu'une transition de probabilité nulle a été observée, un deuxième qui accumule les log-vraisemblances et décide en fonction de sa valeur terminale.

Les capacités de reconnaissance sont bien caractérisées par deux grandeurs: la vitesse logarithmique de sortie de la partie comparative — $\log \rho_k(\mathcal{T}_{k'})$ et l'information de Chernoff de la partie comparative $\Lambda'_{kk'}(0)$. La somme de ces deux nombres donnera $v_\alpha(k, k')$, la vitesse logarithmique globale de convergence des erreurs du test du rapport de vraisemblance.

La complexité du problème de reconnaissance sera dépen-

de la stratégie d'observation des apparences.

Les courbes de FIG. 3 permettent de confirmer la différence de comportement des performances de reconnaissance soupçonnées par les variations structurelles des matrices markoviennes. Les vitesses logarithmiques de sortie de la partie comparative et l'information de Chernoff sont tous deux nettement inférieurs pour le couple d'avions civils. La vitesse logarithmique globale de convergence de l'erreur augmente à peu près linéairement en fonction du nombre d'aspects utilisés. L'écart relatif entre les vitesses logarithmiques du problème de discrimination des avions civils et des avions militaires atteint 40% pour 50 aspect utilisés.

On observe toutefois qu'au delà d'un certain nombre d'aspects (autour de 40), la croissance est essentiellement produite par l'augmentation de la vitesse de sortie de la partie comparative: l'information de Chernoff de la partie comparative sature à une valeur de l'ordre de 0.13 pour les avions militaires pour un nombre d'aspects sensiblement identique à celui provoquant une saturation dans le degré moyen des transitions comparatives (FIG. 2). La structure des matrices semble ici une donnée critique.

5 Autres travaux

Le principe de reconnaissance par accumulation d'évidences a une longue histoire en reconnaissance des formes. Dans le domaine de la vision artificielle, en revanche, la majorité des travaux se sont concentrés sur une caractérisation géométrique globale préalable des objets avant de l'intégrer à une phase de reconnaissance. Un certain nombre d'études, cependant, ont posé la réalisation de «tâches visuelles» comme objectif premier.

Un modèle de reconnaissance, intégrant dans un même système apprentissage, vision 3D et recherche dynamique a été proposé par [30], et repris par [4]. Ils décrivent un schéma complet de constitution empirique d'un graphe d'aspect, et leur utilisation dans la phase de reconnaissance, par un ensemble d'équations différentielles couplées. La génération de nouvelles vues permet d'accumuler les évidences sur le type d'objet observé. L'ambition affichée était de produire un formalisme inspiré de la neurobiologie. Les équations différentielles proposées contiennent cependant un grand nombre de paramètres dont la calibration semble difficile à généraliser, et dont les influences génériques sur les performances d'apprentissage et de reconnaissance sont indéterminées.

[29] décrit un système de vision qui recherche dynamiquement et de manière optimale les informations utiles pour réaliser une tâche donnée, par exemple répondre à la question de savoir si un objet est présent dans une scène. L'évaluation des informations recueillies est produite par un réseau de Bayes [24] qui accumule les informations dans la définition d'une fonction de vraisemblance. Ce type d'approche a été également exploité dans [7] pour la reconnaissance d'objets géométriques simples dans une scène.

[27] décrit une stratégie utilisant deux processus coopératifs, identificateurs et localisateurs, pour détecter la présence d'objets. Les caractéristiques utilisées dans les fonctions de comparaison sont les sorties de filtres gaussiens

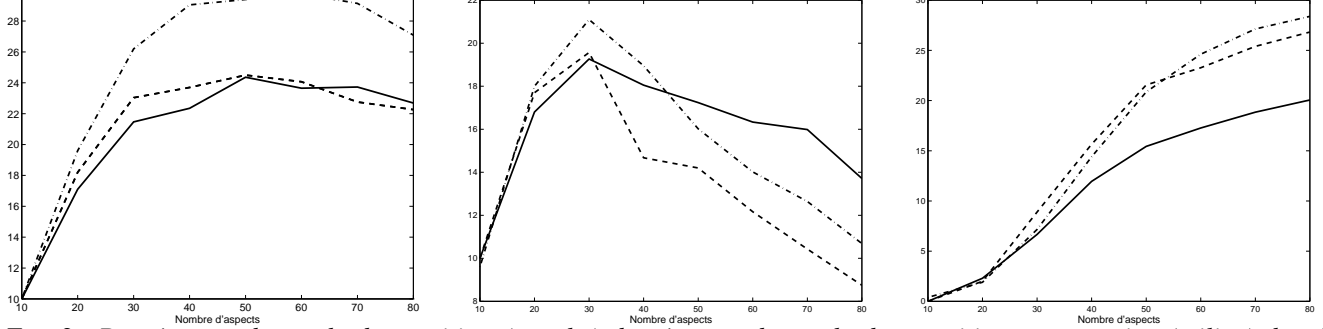


FIG. 2 - Degré moyen du graphe de transitions (gauche), degré moyen du graphe des transitions comparatives (milieu), degré moyen de la partie sélective des transitions (droite) pour les couples d'avions civils (trait continu), militaires (trait pointillé) et mixte (trait semi-pointillé).

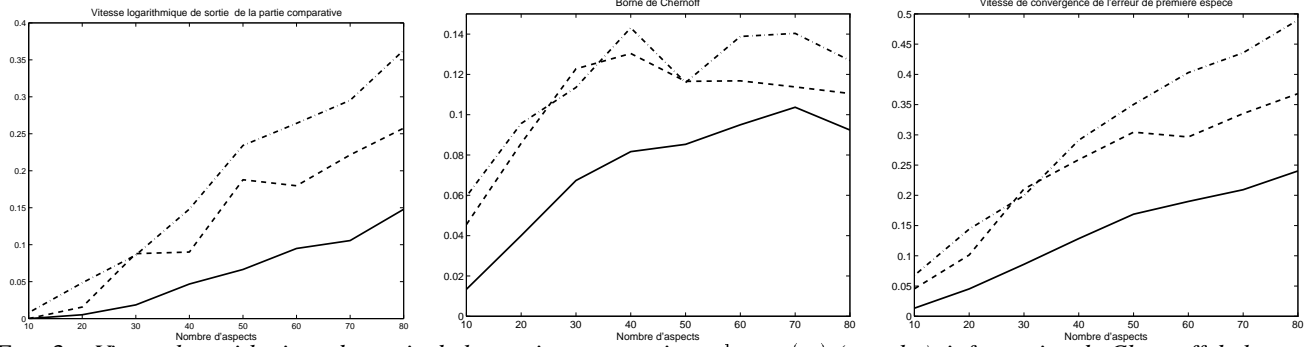


FIG. 3 - Vitesse logarithmique de sortie de la partie comparative – $\log \rho_k(k')$ (gauche), information de Chernoff de la partie comparative (milieu), et vitesse logarithmique de convergence de l'erreur de première espèce (droite) pour les couples d'avions civils (trait continu), militaires (trait pointillé) et mixte (trait semi-pointillé).

différentiels bidimensionnels à plusieurs échelles.

[32] propose un formalisme général pour la sélection spécifique de caractéristiques utiles à la description d'aspects et leur utilisation dynamique dans la vérification d'hypothèses complexes: celles-ci sont composées d'un aspect, d'un objet, d'un mode d'appariement entre données sensorielles et modèle de l'aspect et d'un coefficient caractérisant sa valeur. Lorsqu'une hypothèse est rejetée, le système recherche un nouveau point de vue, sensoriel par mouvement de la caméra, ou représentatif dans la sélection de nouvelles caractéristiques visuelles. L'accumulation d'évidence se produit par restriction incrémentale du nombre d'objets potentiellement observables.

[1] propose une évaluation du caractère informatif des vues par calcul d'une vraisemblance conditionnée par un modèle paramétrique d'objet. Les vraisemblances peuvent être incrémentalement mises à jour par une combinaison de probabilités.

6 Conclusion et discussion

Les grandeurs v_α et v_β mesurent globalement la complexité de discrimination d'objets 3D. Elles sont fonction des modes d'observations des signaux sensoriels et de leurs modes d'acquisition dynamique par l'intermédiaire d'une loi d'action $\mu(a)$: elles caractérisent la difficulté de la reconnaissance *en acte*.

Qualitativement, les vitesses v_α et v_β semblent bien qualifier les complexités des problèmes de reconnaissance: il a été vérifié expérimentalement qu'elle différenciait bien, par exemple, la difficulté de discrimination entre les deux

avions civils et les deux avions militaires.

Les résultats asymptotiques, cependant, requièrent que la chaîne de Markov ait atteint son régime stationnaire. Dans certains cas, cette convergence peut être lente, et limiter la vitesse de décroissance des erreurs lorsque la longueur de trajectoire d'exploration augmente. Les valeurs numériques doivent ainsi être prises avec précaution, particulièrement pour l'évaluation des comportements en temps court.

Les structures probabilistes des représentations markoviennes de graphes d'aspects dépendent de la manière dont sont repérées les transitions entre catégories d'apparences. Elles ne sont valables que pour une loi d'action donnée μ et fixe tout au long de la phase d'apprentissage *et* de reconnaissance. C'est la même stratégie stochastique d'exploration de la sphère de vue qui doit être utilisée dans les deux phases d'estimation des modèles et de leur discrimination empirique. Certains modes de parcours de la sphère de vue peuvent faire converger assez lentement la chaîne de Markov vers son comportement stationnaire.

Remerciements

Ce travail a été conduit dans le cadre d'une thèse sous la direction de Robert Azencott et financé par une bourse BDI du CNRS.

Références

- [1] T. Arbel and F. Ferrie. Informative views and sequential recognition. In *Proc. Fourth European Conference on Computer Vision, Cambridge, England, April 14-18*, pages 1:469–481. 1996.

- angles and distances with applications to 3-D object recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **12**(8):760–774, 1990.
- [3] P. Billingsley. *Statistical Inference for Markov Processes*. The University of Chicago Press, Chicago, London, 1961.
 - [4] G. Bradski and S. Grossberg. Fast learning VIEW-NET architectures for recognizing three-dimensional objects from multiple two-dimensional views. *Neural Networks*, **8**(7/8):1053–1080, 1995.
 - [5] J. Burns, R. Weiss, and E. Riseman. View variation of point-set and line-segment features. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **15**(1):51–68, 1993.
 - [6] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Jones and Bartlett Publishers, Boston, 1993.
 - [7] D. Djian, P. Rives, and P. Probert. Reconnaissance de modèles géométriques simples à l’aide de réseaux bayésiens. In *Actes du 10ème Congrès AFCET-AFIA, RFIA’96, Rennes, France, Janvier*, pages 396–403, 1996.
 - [8] D. W. Eggert, K. W. Bowyer, C. R. Dyer, H. I. Christensen, and D. B. Goldgof. The scale space aspect graph. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **15**(11):1114–1130, Nov. 1993.
 - [9] O. Faugeras, J. Mundy, N. Ahuja, C. Dyer, A. Pentland, R. Jain, and K. Ikeuchi. Why aspect graphs are not (yet) practical for computer vision. *CVGIP: Image Understanding*, **55**(2):212–218, 1992.
 - [10] L. Florack, B. ter Haar Romeny, J. Koenderink, and M. Viergever. Scale and the differential structure of images. *Image and Vision Computing*, **10**:376–388, 1992.
 - [11] W. Freeman. Exploiting the generic viewpoint assumption. *International Journal of Computer Vision*, **20**(3):243, 1996.
 - [12] J. Gårding and T. Lindeberg. Direct computation of shape cues using scale-adapted spatial derivative operators. *International Journal of Computer Vision*, **17**(2):163–191, February 1996.
 - [13] M. Garey and D. Johnson. *Computers and Intractability — A guide to the theory of NP-completeness*. W.H. Freeman, San Francisco, 1979.
 - [14] Y. Gdalyahu and D. Weinshall. Measures for silhouettes resemblance and representative silhouettes of curved objects. In B. Buxton and R. Cipolla, editors, *Proc. Fourth European Conference on Computer Vision, Cambridge, England, April 14-18*, volume 1065 of *Lecture Notes in Computer Science*, pages 363–375. Springer Verlag, Berlin, Heidelberg, New York, 1996.
 - [15] S. Herbin. Recognizing 3D objects by generating random actions. In *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, June 18-20*, pages 35–40, 1996.
 - [16] S. Herbin. *Éléments pour la formalisation d’une reconnaissance active — application à la vision tri-dimensionnelle*. Thèse de mathématiques appliquées, Ecole Normale Supérieure de Cachan, Juillet 1997. In french.
 - [17] D. Kazakos. Asymptotic error probability expressions for multihypothesis testing using multisensor data. *IEEE Transactions on Systems, Man, and Cybernetics*, **21**(5):1101–1114, 1991.
 - phism Problem: Its Structural Complexity*. Birkhauser, Cambridge, MA, 1993.
 - [19] J. Koenderink. *Solid Shape*. MIT Press, Cambridge Ma, 1990.
 - [20] J. Koenderink and A. van Doorn. The singularities of the visual mapping. *Biological Cybernetics*, **24**:51–59, 1976.
 - [21] T. Lindeberg. Scale-space theory: A basic tool for analysing structures at different scales. *Journal of Applied Statistics*, **21**(2):224–270, 1994.
 - [22] T. Lindeberg. *Scale-Space Theory in Computer Vision*. 1994.
 - [23] S. Natarajan. Large deviations, hypotheses testing, and source coding for finite markov chains. *IEEE Transactions on Information Theory*, **31**(3):360–365, 1985.
 - [24] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, San Francisco, second edition, 1988.
 - [25] S. Petitjean. The enumerative geometry of projective algebraic surfaces and the complexity of aspect graphs. *International Journal of Computer Vision*, **19**(3):261–287, 1996.
 - [26] S. Petitjean, J. Ponce, and D. Kriegman. Computing exact aspect graphs of curved objects: algebraic surfaces. *International Journal of Computer Vision*, **9**(3):231–255, 1992.
 - [27] R. Rao and D. Ballard. An active vision architecture based on iconic representations. *Artificial Intelligence*, **78**(1-2):461–505, October 1995.
 - [28] J. Rieger. On the complexity and computation of view graphs of piecewise-smooth algebraic surfaces. Technical Report FBI_HH_M_228/93, Hamburg Universität, 1993.
 - [29] R. Rimey and C. Brown. Control of selective perception using bayes nets and decision-theory. *International Journal of Computer Vision*, **12**(2-3):173–207, April 1994.
 - [30] M. Seibert and A. Waxman. Adaptive 3D-object recognition from multiple views. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **11**:107–124, 1992.
 - [31] E. Seneta. *Non-negative matrices and Markov chains*. Springer Verlag, New York, second edition, 1981.
 - [32] L. Shapiro and M. Costa. Appearance-based 3D object recognition. In M. Hebert, J. Ponce, T. Boulton, A. Gross, and D. Forsyth, editors, *Proc. International Workshop on 3-D Object Representation in Computer Vision, New York City, NY, USA, 5-7 Dec 1994*, volume 994 of *Lecture Notes in Computer Science*, pages 51–64. Springer Verlag, Berlin, 1995.
 - [33] I. Shimshoni and J. Ponce. Probabilistic 3D object recognition. In *Proc. Fifth International Conference on Computer Vision, Cambridge, MA, USA, June 20-22*, pages 488–493, 1995.
 - [34] T. van Effelterre. *Calcul exact du graphe d’aspect de solides de révolution*. Thèse de doctorat, Université de Rennes I, 1995.
 - [35] D. Weinshall, M. Werman, and N. Tishby. Stability and likelihood of views of three dimensional objects. In J.-O. Eklundh, editor, *Proc. Third European Conference on Computer Vision, Stockholm, Sweden, May 2-6*, number 800 in *Lecture Notes in Computer Science*, pages 24–35. Springer Verlag, Berlin, 1994.