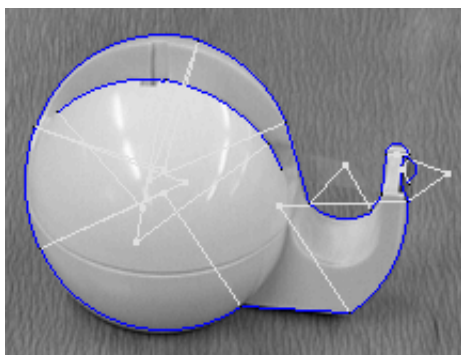
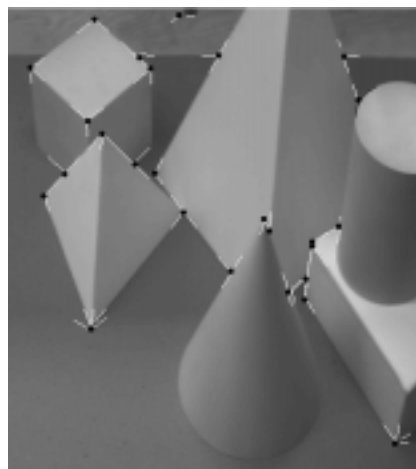
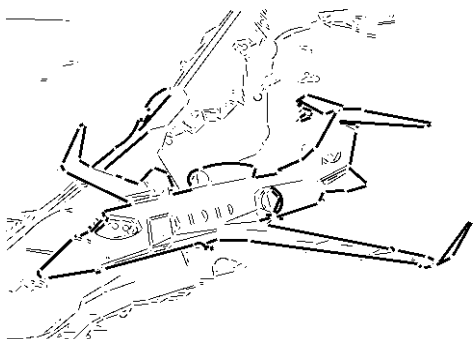


---

## Traitement d'images 2D

## 2D Image Processing





## Segmentation en parties d'un objet 3D à partir d'une seule image 2D complexe

*Une nouvelle approche à la segmentation en parties qui traite des images réelles est présentée. Ces images représentent des objets 3D avec des textures, des ombres et un arrière-plan complexe. L'approche se base sur le contour de l'objet pour guider le regroupement des lignes en utilisant des principes de symétrie et de colinéarité. Ces regroupements de lignes conduisent à l'obtention de formes simples qui peuvent être modélisées par des primitives tridimensionnelles, comme des géons et des cylindres généralisés. Un algorithme basé sur cette approche a été implanté et semble robuste au bruit. De plus, il performe bien dans des conditions génériques.*

---

En reconnaissance d'objets, plusieurs recherches ont été effectuées sur l'extraction des parties d'un objet synthétique. Malheureusement, très peu de tentatives ont été faites pour extraire les parties d'un objet réel manufacturé. Afin que les véhicules automatisés et les robots soient beaucoup plus utiles aux humains, il est essentiel qu'ils soient capables de reconnaître les objets dans leur environnement de travail. L'extraction des parties étant nécessaire à la reconnaissance d'objets, il est donc important d'étudier ce problème. Ainsi, le but de ce projet consiste à extraire les parties d'un objet isolé réel manufacturé, dont l'image a été numérisée à l'aide d'une caméra CCD.

Étant donné que les traitements seront effectués sur une image réelle, plusieurs lignes de l'objet seront manquantes ou incomplètes. De plus, des ombres seront présentes dans l'image, et peut-être des textures sur les surfaces de l'objet. L'approche utilisée devra idéalement être robuste au bruit et être capable de traiter les objets avec de la texture.

Dans le cadre de ce projet, l'approche utilisée pour

## 3D object part segmentation from a single complex 2D image

*A new part segmentation approach is presented which works on real 2D images. These images may contain a complex foreground 3D object with textures and shadows on a cluttered background. The proposed approach relies on the outline of the object to guide the grouping of lines using symmetry and colinearity principles. This grouping of lines leads to simple shapes which could be modeled by 3D primitives such as geons or general cylinders. An algorithm implemented on the basis of this approach appears robust to noise and generic conditions. The results obtained demonstrate the validity of the approach as a means towards 3D generic object recognition from real 2D images.*

---

In object recognition, a lot of research has been done on the extraction of a synthetic object's parts. Unfortunately, very few attempts have been made to extract the parts of a real object. However, to be really useful to us, automatized vehicles and robots need to be able to recognize objects in their working environment. Since part extraction is essential to object recognition, it is important that research be done on this problem. With that in mind, the goal of this project is to extract the parts of a real isolated man-made object in an image obtained by way of digitization with a CCD camera.

Since the processing involved is done on a real image, some lines of the object will be absent or incomplete. Furthermore, shadows will be present in the image and possibly textures on the surfaces of the object. The approach used must ideally be able to take into account noise and textures, and still give the right parts.

For this project, the approach used to resolve this problem is the following:

résoudre ce problème est la suivante :

- 1- Extraction de primitives (arcs et segments) dans l'image d'intensité à l'aide d'un détecteur d'arêtes de type Canny;
- 2- Extraction du contour de l'objet afin d'isoler ses lignes constituantes;
- 3- Détection de caractéristiques entre les primitives formant l'objet;
- 4- Extraction des parties de l'objet à partir des données sur le contour et des caractéristiques trouvées à l'étape 3.

Un logiciel (SE2D [3]) faisant l'extraction de primitives étant disponible au laboratoire, les résultats obtenus par ce logiciel sont utilisés à l'étape 2. La méthode implantée pour extraire le contour de l'objet consiste essentiellement à faire des cycles avec les arcs et les segments de l'objet, mais utilise une caractéristique invariante des objets isolés manufacturés par les humains pour faire le choix des lignes [1][2]. Les résultats obtenus par cet algorithme sont très bons et montrent qu'il est robuste au bruit (Voir figure 1).

Ensuite, la troisième étape consiste à faire l'extraction des lignes parallèles, symétriques et colinéaires. Cette étape est réalisée en analysant les relations géométriques entre les lignes. Les relations géométriques considérées sont par exemple la longueur des lignes, la distance entre celles-ci et le recouvrement. Un facteur de qualité est associé à chaque paire de lignes en fonction de ces relations.

L'étape 4, quant à elle, trouve les deux côtés principaux des parties en se basant sur les caractéristiques extraites à l'étape 3 et leurs facteurs de qualité, et le contour extrait à l'étape 2 [2]. La partie formée des deux côtés est ensuite complétée en faisant des cycles avec les lignes de l'objet pour obtenir son contour complet. Les parties obtenues sont ensuite validées par quelques tests. L'utilisation du contour pour grouper les lignes permet à l'algorithme de cerner avec précision les frontières des parties en contact avec celui-ci.

La figure suivante montre un résultat obtenu par l'implantation de cet algorithme.

- 1- Extraction of the primitives (arcs and segments) present in the image with the help of a Canny type edge detector;
- 2- Extraction of the outline of the object to isolate its constituent lines;
- 3- Detection of the generic features between the primitives composing the object;
- 4- Object parts extraction from the outline data and the features extracted at step 3.

Since there is a software (SE2D [3]) that extracts primitives from a grayscale image already available at our laboratory, the results obtained by this software are used at step 2. The implemented method to extract the outline consists essentially of making cycles with the arcs and the segments of the object, with the help however of an invariant characteristic of real isolated man-made objects [1][2]. The results obtained with this method are very good and show that it is robust to noise (See Figure 1).

The third step consists of making the extraction of parallel, symmetric and colinear lines. This is realized by analyzing the geometric relationships between the lines. The geometric relationships considered are the length of the lines, the distance and the overlap between the lines. From these relationships, a quality factor is calculated and stored for each line pair.

At step 4, the two main sides of the parts are found based on the characteristics extracted at step 3, their quality factors and the outline extracted at step 2 [2]. The part formed by the main sides is completed by making cycles with object lines to obtain its complete contour. The obtained parts are then validated by a few more tests. The use of the outline to pair lines allows the algorithm to circumscribe accurately the frontier of the parts in contact with it.

The following Figure shows results obtained by the implementation of this algorithm.

## References

- [1] G.A. Bilodeau and R. Bergevin, Outline-Based Part Segmentation Using Intermediate-Level Symmetries, in Proceedings of CIPPRS Vision Interface'99, Trois-Rivières, Canada, May 1999.
- [2] G.A. Bilodeau, Segmentation en parties d'objets 3D provenant d'images réelles et complexes, Mémoire de maîtrise, Mars 1999.
- [3] J.L. Arsenaault, R. Bergevin and D. Laurendeau, Extraction of 2D groupings for 3D Object Recognition, SPIE Orlando'94, Visual Information Processing III, Orlando, Florida, April 1994.

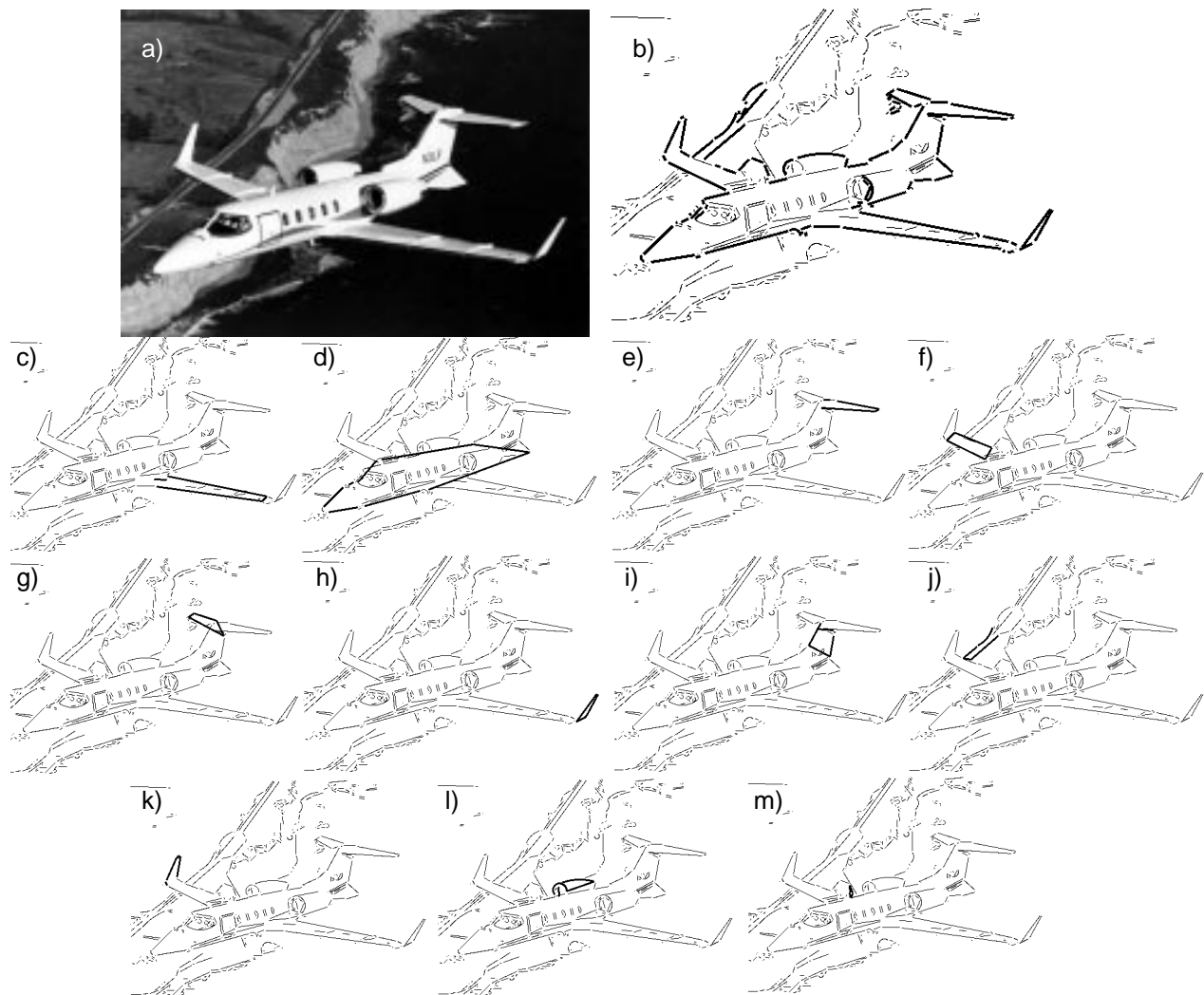


Figure 1. Résultat pour l'image d'un avion. a) L'image originale. b) Le contour obtenu avec notre algorithme. Les images c) à m) montrent les résultats obtenus. *Results on an airplane image. a) The original image. b) The outline obtained from our algorithm. Images c) to m) show the obtained parts.*



Annie Bubel, étudiante au doctorat (Ph.D. student)

Robert Bergevin, directeur de recherche (advisor)

---

## Contribution des jonctions dans l'extraction robuste des contours d'objets 3D dans une image 2D

*Le système de vision dans lequel s'insère ce projet a pour but de détecter et reconnaître l'ensemble des objets 3D présents dans une image 2D d'une scène encombrée. Notre contribution se situe au niveau de la recherche d'une méthode d'extraction robuste des contours significatifs. La méthode proposée se divise en deux étapes principales : (i) un algorithme de validation et de caractérisation de jonctions à partir de la carte des points de contour et, (ii) un processus de formation de la carte complète des contours d'objets 3D de la scène qui exploite les informations fournies par les jonctions extraites.*

L'application que nous visons concerne la réalisation d'un système de description et de reconnaissance des objets 3D présents dans une image 2D. Les scènes considérées impliquent un environnement encombré d'objets manufacturés multi-parties dont on n'a aucune connaissance *a priori*. Ces objets peuvent être vus partiellement ce qui provoque de nombreuses situations d'occlusion. De plus, l'absence de contrôle sur les conditions d'acquisition peut provoquer la présence de discontinuités d'illumination ou de zones de faible contraste qui détériorent la qualité de l'image d'intensité.

Notre contribution à ce projet porte sur l'extraction robuste des contours qui fournissent les données initiales nécessaires à la description géométrique des objets 3D présents dans la scène. Ceci sous-entend la formation de contours qui facilitent la distinction et l'extraction des structures 3D projetées dans l'image. Ainsi, les discontinuités de surface et de profondeur sont recherchées car elles permettent la séparation d'un objet selon ses parties. De même, les occlusions doivent être détectées afin de différencier les contours d'objets distincts.

L'approche classique pour détecter les contours dans une

---

## Junction Contribution in Robust Contour Extraction of 3D Objects in a 2D Image

*This project aims at developing a vision system for the recognition of significant 3D objects in a 2D image of a complex scene. Our contribution to this project is the elaboration of a robust contour extraction scheme. The proposed method is divided in two principal steps: (i) a junction validation and characterization algorithm from an edge point map and, (ii) an extraction process of a complete contour map of 3D objects involved in the scene which exploits information provided by the extracted junctions.*

We aim at realizing a system of description and recognition of 3D objects present in a 2D image. The considered scenes involve a complex scene with man-made multi-part objects for which no *a priori* knowledge is available. These objects may be partially seen leading to many situations of occlusion. Furthermore, the lack of control on acquisition conditions may cause the undesirable presence of illumination discontinuities and low contrast areas which deteriorate the intensity image quality.

Our contribution to this project concerns robust contour extraction, which provides initial data necessary to the geometric description of 3D objects present in the scene. This means the formation of contours to improve the segmentation and the extraction of 3D structures projected in image plane. Thus, surface and depth discontinuities are searched because they allow the separation of an object in its parts. Moreover, occlusion situations have to be detected in order to differentiate edges belonging to different objects.

A classical approach for edge detection in a 2D image aims at applying first a local edge operator on each image pixel, then at validating significant edge pixel data by non maximum suppression and thresholding. Afterwards, a local edge following algorithm is executed

image 2D consiste à appliquer tout d'abord un opérateur local sur chaque pixel de l'image, puis à valider les pixels de contour significatifs par suppression des non maxima et seuillage. Ensuite, un algorithme de suivi local est exécuté pour former des contours complets. Cependant, plusieurs limitations ont été mises en évidence. Les effets les plus courants sont les problèmes de fausse détection ou la formation de contours incomplets. Plus particulièrement, ces opérateurs sont généralement mis en défaut au voisinage des jonctions. De plus, l'utilisation d'informations uniquement locales peut amener la création de contours mal formés en raison d'une mauvaise sélection des points d'arête candidats.

Ces arguments font ressortir qu'une telle approche n'est pas appropriée pour les scènes traitées car celles-ci provoquent la présence fréquente de jonctions dans l'image d'intensité.

Au contraire, plusieurs raisons justifient la nécessité de mettre à contribution les jonctions dans le processus d'extraction des contours.

Tout d'abord, leur présence garantit la position d'un objet à leur voisinage tandis que leur caractérisation donne une bonne indication de la forme des structures sous-jacentes. De plus, les jonctions indiquent de manière naturelle les points de coupure utiles au processus de la segmentation. De même, elles sont des points de délimitation qui permettent de séparer adéquatement les contours selon l'information structurelle des objets 3D. L'ensemble de ces arguments fait ressortir l'apport de la détection et de la caractérisation des jonctions préalablement à l'extraction des contours.

Une nouvelle approche pour l'extraction robuste des contours d'objets 3D est adoptée dans le cadre de ce doctorat. Elle se divise essentiellement en deux étapes : (i) un algorithme de validation et de caractérisation de jonctions à partir de la carte des points de contour et, (ii) un processus d'extraction de la carte complète des contours d'objets 3D de la scène qui exploite les informations fournies par les jonctions extraites.

## Validation et caractérisation de jonctions

Nous avons récemment proposé une méthode non paramétrique de validation et de caractérisation des jonctions [1]. Celle-ci considère en entrée la carte des points de contour et les points d'intérêt fournis simultanément par l'opérateur d'Heitger [2][3]. Le choix de cet opérateur est justifié par son comportement satisfaisant au voisinage des jonctions ou des zones de forte densité de contours de même que par l'estimation précise de l'orientation de chaque point de contour. L'approche

which leads to the formation of complete contours. Several limitations concerning classical edge detectors have been identified. The most common ones are problems of false detection or formation of edges with gaps. In particular, these operators fail nearby junctions, hence producing gaps. Moreover, the use of mainly local information in the edge linking scheme can provide incomplete or false edges due to a wrong selection of candidate points and starting point.

These arguments emphasize that this approach is not appropriate for the involved scenes since they produce many occurrences of junctions in the intensity image.

In contrast, several reasons justify the need to bring junctions to contribute to the contour extraction process.

First, they strongly indicate the presence of an object in their proximity while their characterization gives a good indication of the form of underlying structures. Besides, junctions provide break points useful for the segmentation process. In the same way, they are extremity points which allow to delimitate adequately contours according to structural knowledge of 3D objects. All these arguments clearly point out the contribution of detecting and characterizing junction points before contour extraction.

A new approach for the robust contour extraction of 3D objects is adopted in this thesis. It is mainly divided in two steps: (i) a junction validation and characterization algorithm from an edge point map and, (ii) an extraction process of a complete contour map of 3D objects involved in the scene which exploit information provided by the extracted junctions.

## Junction validation and characterization

We have recently proposed a non parametric method for validating and classifying junctions [1]. This method considers as input the edge point map and keypoints provided by Heitger [2][3]. The choice of this operator is justified by its satisfactory behavior in the proximity of junctions or in regions of high edge density. Besides, it estimates precisely each edge point orientation. This operator provides the additional advantage of defining interest areas where potential junctions may be present. The junction validation approach is divided in three parts:

- An edge point grouping process which considers amplitude and orientation similarity under topology constraints. The output provides structures describing branches associated to a junction.
- A junction validation and position refinement step which considers edge points associated to each branch. This is realized by searching the position



développée comporte trois points principaux :

- Une procédure de regroupement des points de contours selon leur similarité d'amplitude et d'orientation sous des contraintes de topologie. La sortie fournit des structures qui représentent les branches associées à une jonction.
- Une étape de validation et de raffinement de position de la jonction en fonction des points associés à chacune des branches. Ceci est réalisé en cherchant la position qui minimise la distance aux tangentes locales issues des points de contour.
- Une étape d'approximation des branches en segments droits ou courbes qui permet la caractérisation complète des jonctions.

### Extraction des contours

Une nouvelle approche pour la formation des contours est en cours de développement et met à contribution les informations structurelles fournies par les jonctions. Elle cherche à apporter une solution aux limitations énoncées précédemment. Ainsi, les concepts suivants sont retenus:

- Chaque contour est délimité à chacune de ses extrémités par une jonction. Ceci a pour avantage de séparer les contours des différents objets mais aussi des différentes surfaces ou parties de chaque objet.
- Une approche *globale structurelle* est privilégiée où le meilleur contour entre deux jonctions est recherché. Plus précisément, elle est comparable aux méthodes de modèles élastiques (*Snakes*), mais contrairement à celles-ci, l'initialisation se veut entièrement automatique. La notion de globalité apporte plusieurs avantages. Elle permet de contraindre le chemin emprunté par le contour et d'éviter ainsi les problèmes de divergence rencontrés dans les zones de faible contraste, d'ombre, ou de haute densité de contours. D'autre part, elle permet une inférence double de la forme du contour puisque celui-ci est contraint aux deux extrémités par les branches liées à chaque jonction.
- L'algorithme s'adapte au degré de complexité de la forme du contour. Il suit une méthode *hiérarchique descendante*. En effet, le contour est tout d'abord décrit par une primitive simple (droite ou cercle). Si la description n'est pas assez précise, le contour global est divisé selon les intervalles à raffiner et une nouvelle description en primitives est réalisée. Le processus est réitéré jusqu'à ce que la description soit suffisamment précise.

which minimizes the distance to local tangents of edge points.

- An approximation step which describes branches as straight or curved segments and which allows the complete characterization of junctions.

### Contour extraction

A new approach for the formation of 3D object contours is currently under development. It studies the contribution of structural information provided by junctions and searches to offer a solution to the previously expressed limitations. The following concepts are retained:

- Each contour is delimited by a junction at its two extremities. This allows the adequate separation of the contours of different objects but also of the parts of a same object.
- A *global structural* approach is favored since the best contour between two junctions is searched. More precisely, it is comparable to methods involving elastic models such as "Snakes", but contrary to them, the initialization is defined automatically. This notion of globality brings several advantages. First, it constrains the path taken by the contour, which prevents the divergence problems encountered in low contrast, shadowed, or high contour density areas. Furthermore, it defines a double inference of contour shape since the latter is constrained at the two extremities by branches associated to each junction.
- The algorithm is adaptable to the degree of complexity of the contour shape. It follows a *hierarchical descending* rule. Indeed, the contour is initially described by a single primitive (straight line segment or circular arc). If the description is not accurate enough, the global contour is then divided according to the intervals which need to be refined and a new description in terms of primitives is realized. The process is repeated until the description is sufficiently accurate.

### References

- [1] Bubel, A. and R. Bergevin. "Classifying Junctions by Vector Quantization", *VI'98*, Vancouver, June 18-20, p.165-171, 1998.
- [2] Heitger, F. "Feature detection using suppression and enhancement". Technical Report 163, Communication Technology Laboratory, Swiss Federal Institute of Technology ETH, ETH-Zurich, Switzerland, 1995.
- [3] Rosenthaler, L., Heitger, F., Kubler, O. and von der Heydt, R. "Detection of General Edge and Keypoints", *ECCV'92*, p. 78-86, 1992.

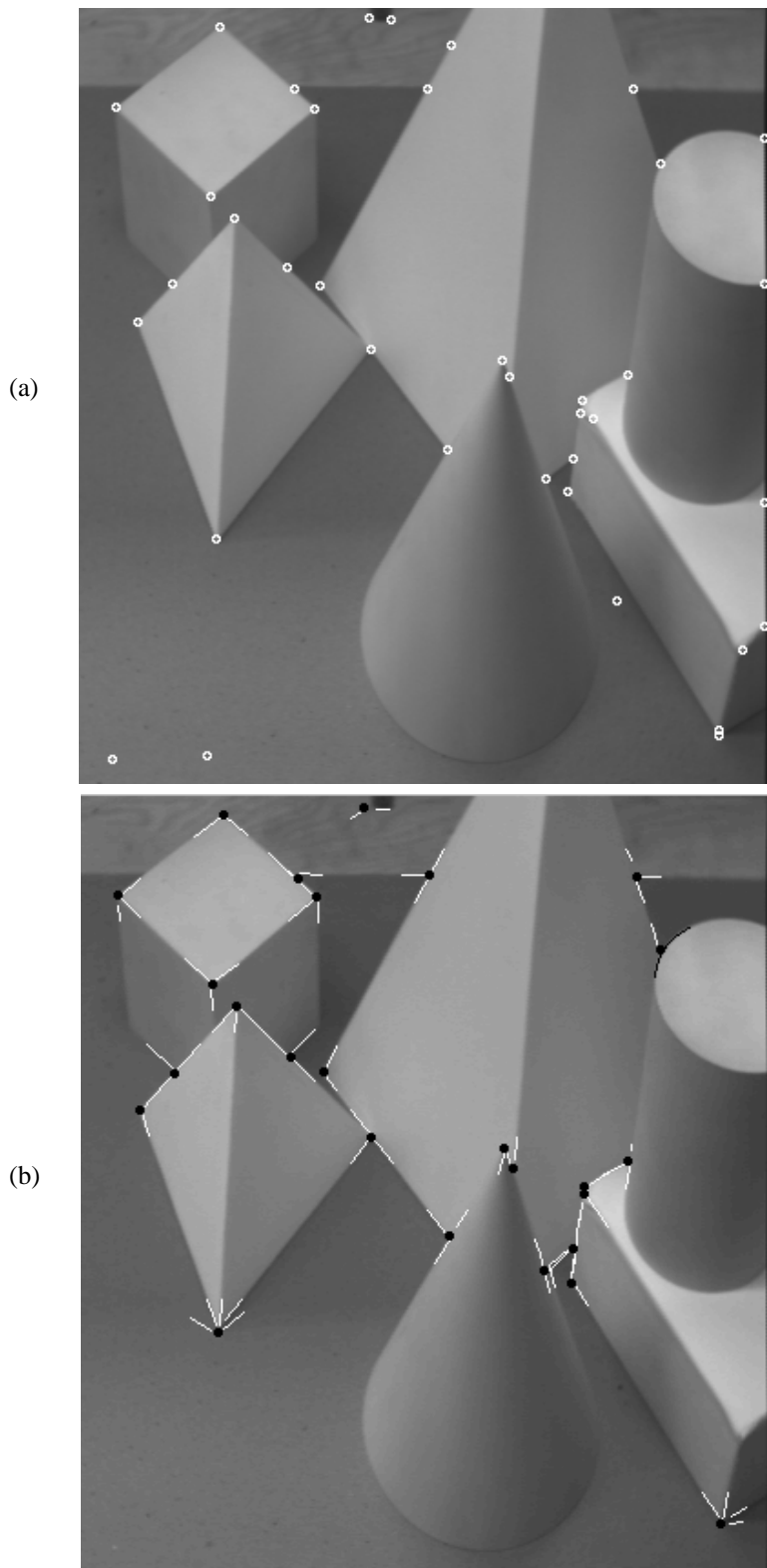


Figure 1. Exemple de validation et de caractérisation des jonctions sur une image réelle. (a) Jonctions potentielles détectées par Rosenthaler. (b) Image résultat des jonctions validées et caractérisées. La nouvelle position des jonctions retenues est affichée et les branches associées aux jonctions sont approximées par des segments droits (en blanc) ou des arcs de cercle (en noir).

Figure 1. Example of validation and characterization of junctions on a real image. (a) Potential junctions detected by Rosenthaler. (b) Result image of validated and characterized junctions. The new position of remaining junctions is displayed. Branches associated to junctions are approximated by straight lines (in white) and circular arcs (in black).

*Christophe Deutsch, étudiant au doctorat (Ph.D. student)*

*André Zaccarin, directeur de recherche (advisor)*

*Denis Laurendeau, codirecteur de recherche (co-advisor)*

---

## **Suivi et localisation d'une cible complexe, connue *a priori*, par vision monoculaire**

*Le suivi et la localisation d'une cible se fait à l'aide d'une caméra 2D. L'objet est connu a priori, mais aucune hypothèse n'est faite sur son mouvement (mouvement saccadé possible). Les six degrés de liberté d'un corps rigide en mouvement accéléré doivent être considérés. La méthode proposée se base sur une modélisation de la cible par graphe d'aspect. Une représentation originale des modèles 2D du graphe d'aspect a été développée. Associée à une analyse en composantes principales, elle permet de retrouver la pose de la cible de manière très efficace. L'algorithme proposé est aussi bien utilisé pour l'initialisation que pour la poursuite. Il permet aussi de bien tenir compte des contraintes très générales qui sont imposées.*

---

Aujourd'hui, la capacité de traitement d'information des ordinateurs permet d'analyser des problèmes complexes en temps réel. La vision par ordinateur a particulièrement profité de ces progrès. En effet, des tâches telles que le téléguidage de robot, la télé-surveillance des autoroutes (pour la vérification automatique du trafic) ou encore la supervision de chaînes de production sont rendues possibles. De nombreux problèmes restent cependant encore du domaine de la recherche. L'un d'eux est le suivi et la localisation d'objet. Il s'agit d'un point très important puisqu'il est indispensable pour toutes sortes d'applications. Le guidage automatique de véhicule, par exemple, nécessite la connaissance exacte de la position des obstacles.

L'estimation de la position exacte d'un objet est un problème générique complexe. Les approches sont très différentes suivant la connaissance *a priori* de l'objet ou non, le type de capteurs utilisés (une ou plusieurs caméras 2D, caméras 3D, etc.), et le type de déplace-

## **Model-based tracking and pose estimation of a complex target by monocular vision**

*Target tracking and pose estimation are done from images obtained with a 2D camera. An object model is given but no hypothesis are made on its motion (a jerky movement is possible). The six degrees of freedom of a rigid object with an accelerated motion have to be considered. The proposed method is based on an aspect graph modeling of the target. An original representation of the aspect graph 2D models has been developed. Associated with a principal component analysis, the target's pose can be computed in a very efficient way. The proposed algorithm is used as well for initialization and tracking. It allows us to deal with the very general constraints of this problem.*

---

Computers are now able to process a large amount of information, thus allowing for real time analysis of complex problems. Computer vision has particularly taken advantage of this progress. In fact, applications such as robot guidance, road surveillance or chain production supervision are now feasible. Many problems, however, have not yet been solved and are active research areas. One of them is the tracking and pose estimation problem. It is a very important research area in computer vision since this task has to be performed in many applications. An automatic car guidance system, for example, has to be able to evaluate the exact location of any obstacle.

Accurate pose estimation is a complex problem. A number of approaches have been proposed and are a function of the prior knowledge of the object, type of sensor used (one or more 2D cameras, 3D cameras, etc.) and assumed type of motion. The purpose of this ongoing research is to find the 3D position of a known, rigid and relative complex object from images obtained with a

ment. Dans la recherche que nous proposons, il s'agit de retrouver la position 3D d'un objet connu, rigide et « relativement complexe », à l'aide d'une caméra 2D fixe. Le mouvement de l'objet peut être saccadé et des occlusions partielles peuvent se produire.

L'application pratique liée à cette étude est le remplacement d'isolateurs sur les lignes électriques. L'opérateur, dans un camion au niveau du sol, commande deux bras manipulateurs rattachés au véhicule par une nacelle, sur laquelle se trouvent des capteurs. Ces capteurs envoient à l'opérateur une image qui lui permet de remplacer les isolateurs. L'image n'est cependant pas stabilisée par rapport au mouvement de la nacelle. L'opérateur peut donc éprouver des difficultés à guider les bras. À l'aide d'une caméra 2D observant la scène, l'algorithme proposé devra permettre de connaître la position de la nacelle et ainsi de compenser le mouvement de celle-ci pour que l'image reçue par l'opérateur soit stable.

Le problème qui se pose se résume donc à trouver la position dans l'espace d'un objet, à partir d'une séquence d'images 2D. Les algorithmes classiques [1,5,6,7], utilisés pour la poursuite de cible, nécessitent toujours une étape de prédiction afin de limiter les recherches dans l'espace des poses possibles. Cette étape de prédiction peut être plus ou moins complexe et va de l'utilisation de la pose de l'objet dans l'image précédente [6] à de pseudo-filtres de Kalman très élaborés [5]. L'estimation initiale de la pose obtenue par cette étape de prédiction sert ensuite de point de départ à des algorithmes de raffinement de pose. Malheureusement, dans notre cas, avec le manque de contraintes imposées au mouvement de la cible, aucune prédiction ne pourrait donner une bonne estimation initiale de la pose. Or les algorithmes itératifs permettant de raffiner la pose convergent d'autant plus rapidement et d'autant mieux que l'estimation initiale est bonne (cf. [2]).

Afin de résoudre ce problème, une approche originale basée sur une modélisation par graphe d'aspect de la cible est proposée. Chaque modèle 2D du graphe d'aspect est représenté d'une part par une image binaire qui contient le contour de la cible (voir figure 1) et d'autre part par une paramétrisation de type « coordonnées polaires ». La paramétrisation permet d'effectuer une analyse en composantes principales et ainsi de diminuer énormément la taille de l'espace de paramétrisation. La topologie naturelle (basée sur la distance euclidienne) de l'espace des paramétrisations représente une bonne méthode de comparaison entre les différentes formes des modèles du graphe d'aspect. Une classification en groupe de modèles de même forme peut donc être effectuée sur le graphe d'aspect. Pratiquement

2D fixed camera. The object motion can be jerky and partial occlusion must be treated.

The practical application of this research deals with the replacement of insulators on electric lines. The operator is on the ground level in a vehicle and controls two articulated arms attached to the vehicle through a traveling crane. The sensors are on the traveling crane and give an image of the scene to the operator so that he can replace the insulators. Unfortunately, the image is not stabilized against the traveling crane movements. Thus the operator may have difficulties guiding the manipulated arms. With a 2D camera observing the scene, the tracking of the traveling crane's position will allow us to stabilize the operator's image.

Thus the problem that has to be solved is to find the 3D position of an object with a 2D image sequence. Almost all the classical tracking algorithms [1,5,6,7] need a prediction step in order to limit the research in the pose space. This prediction step can be more or less complex. It goes from the simple use of the object pose in the former image [6] to elaborate pseudo-Kalman filters [5]. The initial pose estimate, obtained from this prediction step, is then used as the starting point of an iterative pose refinement algorithm. Unfortunately, with the hypothesis of unconstrained motion, it would be impossible in our case to obtain a relevant initial pose estimation. This is a problem since iterative algorithms for pose refinement converge well and rapidly with an accurate initial estimate [2].

To solve this problem, a new approach based on an aspect graph representation of the target has been proposed. Each 2D model of the aspect graph is represented with a binary image containing the target's edges (see figure 1) and with a parametric vector containing "polar coordinates". The parametric representation allows a principal component analysis. This reduces significantly the dimension of the underlying space. The natural topology (based on the euclidean distance) of the parametrization space is a good measure for comparing the different shapes present in the aspect graph. A classification based on this measure can be used to group the aspect graph models in shape classes. This is obtained with a fuzzy K-means algorithm applied in the principal component subspace (see Figure 2). These shape classes are helpful for the target's pose computation.

The parametric representation of the 2D models has an interesting property. An image plane rotation can be represented by a circular permutation of the vector coordinates. Based on this, a discrete Fourier transform is applied to the parametric vectors before the principal component analysis. The effect is a structural invariance

la création de tels groupes se fait à l'aide d'un algorithme de type *fuzzy K-Means* dans l'espace des composantes principales (voir figure 2). Cette classification peut ensuite être utilisée pour simplifier la recherche de la pose de la cible.

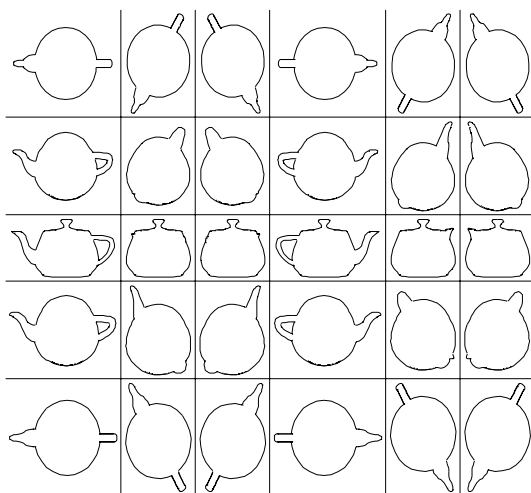


Figure 1. Graphe d'aspect d'une théière avec 30 modèles.  
*Teapot aspect graph with 30 models.*

À cause de la nature « polaire » de la paramétrisation, une rotation dans le plan de l'un des modèles correspond à une permutation circulaire des composantes du vecteur de paramétrisation. Ainsi, en effectuant une transformation de Fourier discrète des vecteurs de paramétrisation avant l'analyse en composantes principales on ajoute une invariance structurelle par rapport aux rotations dans l'espace des composantes principales.

La détermination des paramètres de pose de la cible peut être faite de la manière suivante. Deux des paramètres de translation sont retrouvés à partir de la position du modèle 2D dans l'image. Deux des paramètres de rotations sont obtenus à partir du modèle du graphe d'aspect qui a été retenu. Les deux derniers paramètres (le facteur d'échelle et la rotation dans le plan image) peuvent être obtenus directement à l'aide de la paramétrisation des modèles 2D. En effet, dans l'espace des composantes principales un facteur d'échelle correspond tout simplement à une multiplication par un nombre scalaire. L'angle de rotation peut, quant à lui, être obtenu par la phase de la transformée de Fourier.

À partir de toutes ces remarques, un algorithme a été élaboré permettant de trouver les poses les plus probables dans l'image 2D. Celles-ci sont ensuite vérifiées en utilisant une approche de l'algorithme de H.T. Huttenlo-

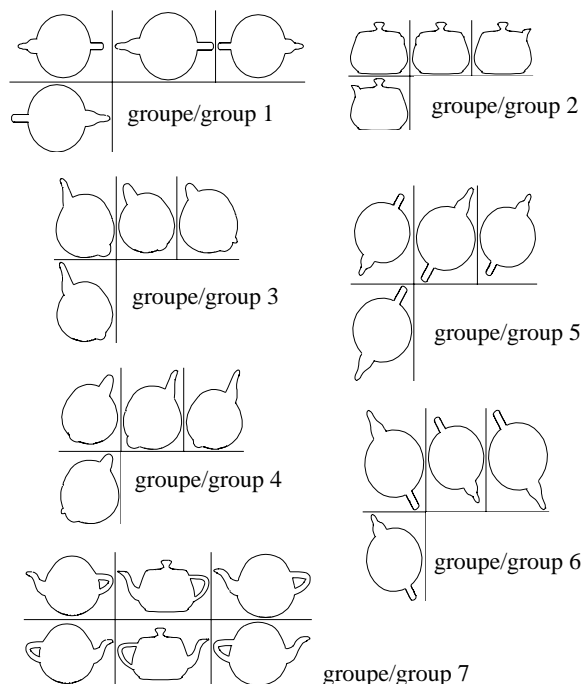


Figure 2. Classification selon la forme pour la théière.  
*Shape groups obtained from the teapot model.*

to image plane rotations in the principal component subspace.

The target's pose parameters can be computed from the information extracted from the 2D image. Two translation parameters are obtained from the position of the 2D model in the image. Two rotation parameters correspond to the position of the selected 2D model in the aspect graph. The last two parameters (scale factor and image plane rotation) can be computed using our parametric representation. In fact, in the principal component subspace, a scale factor corresponds simply to a scalar multiplication and the image plane rotation angle is contained in the Fourier transform phase.

Using all these properties, an algorithm that estimates the most probable target's poses in a 2D image has been developed. These poses are then validated thanks to H.T. Huttenlocher's approach [3]. The pose obtaining the best match is considered as the exact one.

The advantages of the tracking and pose estimation approach we propose is that the motion of the object we are tracking is unconstrained and that the initialization and the tracking processes are solved in the same way. The only difference is that the search is done in the complete pose space for initialization while it is limited to the surrounding of the preceding pose for tracking.

cher [3]. L'hypothèse obtenant la meilleure évaluation est celle qui est retenue.

L'avantage de l'approche que nous proposons réside dans les faibles contraintes imposées au mouvement ou au changement d'aspect de l'objet. De plus, l'algorithme peut être utilisé aussi bien lors de l'initialisation, c'est-à-dire lors du calcul de la pose de la cible dans la première image, que lors de la poursuite proprement dite. La seule différence réside dans le temps de calcul puisque au lieu d'effectuer une recherche complète lors de la poursuite, on se limite à une recherche partielle autour de la pose trouvée pour l'image précédente.

## References

- [1] K.D. Baker, and G.D. Sullivan. *Performance Assessment of Model-Based Tracking*. 1st Workshop on Applications of Computer Vision, Palm Springs, CA, 1992, pp. 28-35.
- [2] C. Deutsch, *Suivi et localisation d'une cible complexe connue par vision monoculaire*, Rapport pré-doctoral (RT-LVSN-97-04), Université Laval, Mai 1997.
- [3] C. Deutsch, A. Zaccarin, D. Laurendeau, *Model-Based Tracking by Monocular Vision*, submitted to imdsp98.
- [4] D.P. Huttenlocher, J. J. Noh, W. J. Rucklidge, *Tracking non-rigid objects in complex scenes*, Tech. Rep. TR92-1320, Computer Science Dep., Cornell University, dec. 92.
- [5] D. Koller, K. Daniilidis, and H.H. Nagel. *Model-Based Object Tracking in Monocular Image Sequences of Road Traffic Scenes*, Int. Journal of Computer Vision, Vol. 10, No. 3, June 1993, pp. 257-281.
- [6] D.G. Lowe, *Robust model-based motion tracking through the integration of search and estimation*, Int. Journal of Computer Vision, 8(2):113-122, 1992.
- [7] G.D. Sullivan, A.D. Worrall, and J.M. Ferryman. *Visual Object Recognition Using Deformable Models of Vehicles*. Proc. Workshop on Context-Based Vision, Cambridge Massachusetts, 19th June 1995, pp. 75-86.

*Jean-François Hébert, étudiant au doctorat (Ph. D. student)*

*Marc Parizeau, directeur de recherche (advisor)*

*Nadia Ghazzali, codirecteur de recherche (co-advisor)*

---

## **Architecture neuronale hybride pour la segmentation d'objets dans une scène 2D : application à l'écriture cursive**

*Pour beaucoup d'applications reliées à la reconnaissance des formes, un intérêt certain existe vis-à-vis la conception d'outils de reconnaissance réalisant un apprentissage incrémental. Principalement, un apprentissage de type incrémental permet de conserver les connaissances déjà acquises par le système lorsque de nouvelles données doivent être apprises. En reconnaissance d'écriture cursive, la taille souvent immense des banques d'entraînement et la possibilité de vouloir tenir compte de l'arrivée de nouveaux scripteurs au fil du temps sont quelques-uns des facteurs qui font qu'un apprentissage incrémental des connaissances est nécessaire.*

---

En reconnaissance des formes, plusieurs raisons justifient la conception de systèmes qui réalisent un apprentissage incrémental des connaissances. Entre autres, un tel type d'apprentissage permet à un système de reconnaissance 1) de réaliser de longues périodes d'apprentissage en plusieurs étapes successives, 2) d'apprendre de nouvelles connaissances, 3) de s'adapter et de réagir suite à une évolution des connaissances, et enfin 4) de renforcer des connaissances déjà acquises.

Ce projet de doctorat vise à implanter un système de reconnaissance capable de localiser la présence d'objets connus (i.e. préalablement modélisés) à l'intérieur de scènes 2D complexes. Ce système devra apprendre d'une manière incrémental, c'est-à-dire qu'il sera possible à tout moment d'ajouter de nouveaux modèles d'objets, ou encore de modifier des modèles existants. Plus particulièrement, nous nous intéressons dans le cadre de ce projet au problème de la localisation de lettres isolées à l'intérieur de mots cursifs.

Dans un tel contexte de reconnaissance d'écriture cursive, notre système réalisera les trois tâches

## **Hybrid neural network architecture for segmentation of objects in 2D scenes: Application to cursive handwriting**

*In the context of pattern recognition applications, it is often of prime interest to design systems that are capable of incremental learning. Mainly, a recognition system based on an incremental learning algorithm is able to learn new data without degrading previous knowledge. For handwriting recognition, where the available training data sets are usually so huge that it becomes difficult to train the network in a single step, and where new writers with very different writing styles may also appear after initial training, an incremental learning is necessary.*

---

There are many reasons for wanting to implement incremental learning in pattern recognition. The main ones are 1) to allow long training periods in many successive steps, 2) to enable learning of new knowledge, 3) to react and consequently adapt to knowledge evolution, 4) to reinforce current knowledge.

This project deals with the design of a recognition system able to locate known objects (i.e. objects that have been previously modeled) into a complex 2D scene. The system should learn incrementally, which means that it will be able to adapt the known object models at any time. More precisely, the localization of isolated characters into cursive word serves as the main application.

In the context of handwriting recognition, the system will realize the following three distinct tasks:

- 1- describe isolated characters in a comprehensive format for the recognition system,
- 2- learn how to model isolated characters,
- 3- learn how to segment the character models in cursive words.

suivantes :

- 1- représenter les caractères isolés dans un format qui soit compréhensible pour le système de reconnaissance,
- 2- apprendre à modéliser des caractères isolés,
- 3- segmenter les modèles de caractères à l'intérieur d'un mot cursif.

## Représentation de l'écriture

Une nouvelle représentation floue est utilisée afin de transformer un segment d'écriture de son espace original (i.e. une séquence de points 2D de longueur variable) vers un nouvel espace vectoriel flou de dimension fixe [1]. Afin d'extraire la représentation floue d'un caractère entier, ce dernier est d'abord segmenté en une séquence d'arcs de cercle, puis divisé en plusieurs régions selon une grille rectangulaire (par exemple, une grille 3x2). Par la suite, chaque arc de chaque région du caractère est « fuzzifié » selon son orientation et sa courbure. Diverses opérations floues sont alors réalisées sur les arcs « fuzzifiés » de chaque région afin de former des vecteurs régionaux flous. La représentation finale, qui permet de décrire le caractère en entier, est obtenue en juxtaposant chacun des vecteurs régionaux. Cette représentation peut par la suite être utilisée afin d'entraîner n'importe quel type de classifieur à reconnaître des caractères isolés tels que des chiffres, des lettres, etc.

## Modélisation de l'écriture

Un réseau de neurones artificiels de type NGCS-P est utilisé pour la modélisation des caractères isolés. Le réseau NGCS-P est une nouvelle architecture neuronale hybride qui combine les propriétés du réseau non-supervisé *Network of Growing Cell Structures* (NGCS) à celles du réseau supervisé *Perceptron* (P) [2]. Au sein du réseau NGCS-P, le rôle du réseau NGCS est de partitionner correctement le réseau P de manière à mieux orienter l'apprentissage de ce dernier et ainsi accroître ses capacités d'apprentissage incrémental. Les caractères issus de la banque internationale UNIPEN [3] sont utilisés pour l'entraînement du réseau NGCS-P. Cette banque contient plus de cinq millions de segments d'écriture de toutes sortes (caractères, mots, phrases, textes, adresses postales, etc.) qui ont été écrits par plus de 2000 scripteurs, ce qui en fait une banque de données complexe et diversifiée. Il en ressort la nécessité de réaliser un apprentissage incrémental des connaissances à l'aide du réseau NGCS-P.

## Handwriting representation

A new fuzzy geometric representation is used to map a character from its original sequence of 2D coordinates into a fixed dimension fuzzy vector space [1]. This fuzzy representation is obtained by first segmenting the original handwritten script into a sequence of circular arcs. Thereafter, the character is decomposed into a rectangular grid of regions (e.g., a 3x2 grid), and each circular arc found in each region of the grid is fuzzified according to its orientation and curvature using some predefined fuzzy sets. Fuzzified arcs are then combined using fuzzy operators to form regional fuzzy vectors, and these vectors are simply concatenated to produce the final representation for the character. The resulting fuzzy representation can thereafter serve as input to any classifier for recognizing isolated characters such as digits, letters, etc.

## Isolated character modeling

The modeling of isolated characters is realized by a NGCS-P artificial neural network. The NGCS-P network is a new neural network which combines an unsupervised "Network of Growing Cell Structures" (NGCS network) with a supervised Perceptron network (P network) [2]. For the NGCS-P network, the goal of the unsupervised NGCS network is to clusterize the supervised P network in order to reduce the risk of forgetting already learned knowledge, and thus improve its incremental learning possibilities. During the NGCS-P supervised training, a very large number of labeled characters are sequentially presented to the network. These characters are extracted from the international UNIPEN database [3] which contains over five million handwritten segments (characters, words, phrases, texts, postal addresses, etc.) written by over 2000 writers. This huge training data set as well as the great variety of handwriting styles found among writers thus makes the cursive script recognition application an appropriate situation for using an incremental learning.

## Character segmentation

In order to segment cursive words, a new strategy has been recently proposed [4] for isolating characters without making any explicit *a priori* segmentation of the handwritten script, and without imposing any lexical or linguistic constraints. The main idea behind the segmentation strategy is to have a window of attention moving around in the cursive word, searching for instances of known characters which had been previously modeled by the NGCS-P network. If one assumes that the current window contains some significant part of a character,



## Segmentation des caractères

Une nouvelle stratégie a été développée afin de pouvoir localiser les lettres d'un mot cursif [4]. L'approche utilisée ne requiert ni segmentation préalable du mot, ni définition de contrainte lexicale ou linguistique. En fait, elle est fondée sur le balayage d'une fenêtre d'attention à l'intérieur d'un mot cursif, ceci dans le but de trouver toutes les instances de caractères préalablement modélisés par le réseau NGCS-P. Si, à un instant quelconque, la fenêtre d'attention contient de l'information significative à propos d'un caractère, elle doit alors subir une translation et une mise à l'échelle afin de converger vers le cadre de ce caractère. Au sein du système, ce processus de localisation est implanté à l'aide d'un réseau *détecteur* ainsi que d'un ensemble de réseaux *locateurs*. Le réseau détecteur est un classifieur, et son rôle est de reconnaître des caractères entiers. Pour sa part, un réseau locateur doit agir localement et se concentrer sur la reconnaissance des parties cruciales de caractères. Le réseau locateur doit alors produire la transformation adéquate des paramètres de la fenêtre d'attention afin de mieux encadrer le caractère recherché.

then the problem is to translate and scale the window of attention in such a way that it converges to the bounding box of that character. This process is implemented using both a *detector* network and a set of *locator* networks. The detector network is responsible for recognizing whole characters of any class, whereas a locator network is assigned the task of recognizing the crucial parts of a given character class and producing the corresponding transformation parameters for the window.

## References

- [1] Hébert J.F., Parizeau M., Ghazzali N., 1998, "A New Fuzzy Geometric Representation for On-Line Isolated Character Recognition", *Int. Conf. on Pattern Recognition (ICPR'98)*, pp. 1121-1123.
- [2] Hébert J.F., Parizeau M., Ghazzali N., 1999, "A New Hybrid ANN Architecture for Active and Incremental Learning: Network of Growing Cell Structures - Perceptron (NGCS-P)", *Int. Joint Conf. on Neural Networks (IJCNN'99)*.
- [3] Guyon I., Schomaker L., Plamondon R., Liberman M., Janet S., 1994, "UNIPEN project of on-line data exchange and recognizer benchmarks", *Int. Conf. on Pattern Recognition (ICPR'94)*, pp. 29-33.
- [4] Hébert J.F., Parizeau M., Ghazzali N., 1999, "Learning to Segment Cursive Words using Isolated Characters", *Vision Interface (VI'99)*.



**Marielle Mokhtari, étudiante au doctorat (Ph.D. student)**

**Robert Bergevin, directeur de recherche (advisor)**

---

**Description d'objets 3D  
dans une image 2D :  
segmentation et approximation  
des contours par segments  
à courbure constante**

*Nous proposons une approche originale pour déterminer la description minimale de courbes planes (ouvertes ou fermées) en ensembles de segments à courbure constante (segments de droites et/ou arcs de cercles). L'algorithme résultant est basé sur la segmentation multi-échelles et l'approximation de courbes, et la classification intra- et inter-échelles de la description obtenue. Son but ultime est de trouver un ensemble de paires adéquates (échelle, ensemble(s) de segments à courbure constante) pour décrire au mieux la forme de la courbe. Ce travail de recherche est partie intégrante d'un projet visant la localisation et la description d'objets 3D et/ou de parties d'objets 3D inclus dans une image 2D, et ce à partir de structures de haut niveau obtenues par groupement perceptuel des segments à courbure constante.*

Dans le contexte d'une application de détection et de description d'objets 3D, la *perception visuelle* d'une scène se matérialise en pratique par la production d'une description de celle-ci selon les caractéristiques de forme géométrique des objets présents. La première phase du module de *perception*, qui en comprend trois, consiste à extraire l'information structurelle de base d'une image 2D à niveaux de gris. Celle-ci représente une scène encombrée, composée d'objets 3D, généralement opaques, en occlusion ou non. Cette première phase est divisée en trois étapes: (i) détection des arêtes selon l'opérateur de Canny; (ii) les contours ouverts et fermés pouvant correspondre aux frontières des objets sont identifiés dans l'image d'arêtes selon une méthode de suivi de contours débutant aux points de jonction et aux extrémités; et, (iii) segmentation/approximation robuste des contours/courbes en segments à courbure constante (segment de droite, arc de cercle). Dans la suite, nous

---

**Description of 3D Objects  
in one 2D Image:  
Segmentation and Approximation  
of Contours using Constant  
Curvature Segments**

*We propose an original approach to find minimal description of planar curves (open or closed) into sets of constant curvature segments (straight line segments and/or circular arcs). The resulting algorithm proceeds according to a multiscale segmentation and approximation of curves, and an intra- and inter-scale classification of this overall process. Its ultimate goal is to find a set of adequate pairs composed of one scale and one or several set(s) of constant curvature segments to best describe the shape of a curve. This research work is part of a more generic project for detecting and describing 3D objects and/or parts of 3D objects in a single 2D image based on high-level structures obtained by perceptual grouping of constant curvature segments.*

Under 3D object detection and description task, the *visual perception* of a scene aims at describing surrounding objects according to their geometric features. The perception module is implemented as a three phase process. The first phase consists of the extraction of structural information from a single grey level 2D image of a scene composed of complex, and possibly occluding man-made objects. This first phase is itself divided in three steps: (i) edge detection with the Canny operator; (ii) identification of open and/or closed contours that may correspond to object boundaries in the edge image, using a contour following algorithm with junctions and endpoints as starting points; and (iii) robust segmentation and approximation of the contours/curves as constant curvature segments (straight line segment, circular arc). In the following, the segmentation and approximation step is described in more details and experimental results on real data are presented.

discuterons de la troisième étape et nous montrerons des résultats de l'application des trois étapes sur des scènes réelles.

### 1. Segmentation et approximation de courbes planes

Dans [1], Bergevin et Mokhtari ont proposé une approche théorique de segmentation multi-échelles et d'approximation de courbes 2D en segments à courbure constante (segments de droites et/ou arcs de cercles). L'algorithme résultant nommé MuscaGrip, pour *MULTIscale Segmentation and Curve Approximation based on the Geometry of Regular Inscribed Polygons*, calcule, à plusieurs échelles, une séquence de deux processus de groupement : (i) un processus de groupement de points conduisant à une approximation polygonale et (ii) un processus de groupement de segments de droites en arcs de cercles lorsque cela est possible. L'algorithme répète le premier processus de groupement à chaque point de la courbe, et le second processus à chaque segment de droite en provenance du premier processus. Ces répétitions conduisent à une description de la courbe composée d'ensembles de segments à courbure constante à différentes échelles, certains ensembles pouvant être identiques et/ou en recouvrement et/ou incluant un débordement. Une approche globale est proposée pour réduire le nombre d'ensembles obtenus à un sous-ensemble non-redondant n'incluant ni recouvrement, ni débordement. La manière d'y parvenir est de classifier les résultats de chaque processus de groupement en utilisant des *étiquettes qualitatives* définies de manière heuristique. Par conséquent, il est possible de trouver, pour chaque courbe, un ensemble de paires (*échelle, ensemble(s) de segments à courbure constante*) qui décrit au mieux la courbe. Pour la suite, une courbe plane est définie par

$$C = \{p_i(x_i, y_i) \mid i \in [1, n] \wedge (x_i \in \mathbb{R} \wedge y_i \in \mathbb{R})\}.$$

### 2. MuscaGrip: approximation polygonale

Le processus d'agrégation de points consécutifs découpe chaque courbe  $C$  en multiples sous-courbes, chacune étant approximée par un segment de droite. Le critère d'uniformité associé à ce groupement de points est équivalent à un critère discret de co-circularité mutuelle parmi les points connectés discrètement sur la courbe. Un *paramètre d'échelle*, agissant comme un critère de déviation maximale, est associé à chaque mesure d'échelle.

Une chaîne de deux points ou moins forment une sous-courbe uniforme. Une chaîne de trois points et plus forment une sous-courbe uniforme si et seulement si la distance perpendiculaire, de chacun des points de la chaîne, par rapport à la droite joignant les extrémités de la chaîne est inférieure au paramètre d'échelle. Cette

### 1. Segmentation/Approximation of Planar Curves

In [1], Bergevin and Mokhtari proposed a theoretical framework for the multiscale segmentation and approximation of 2D curves into constant curvature segments (straight line segments and/or circular arcs). The resulting algorithm, named *MuscaGrip*, which stands for *MULTIscale Segmentation and Curve Approximation based on the Geometry of Regular Inscribed Polygons*, computes, at multiple scales, a sequence of two grouping processes: (i) a *point grouping* process leading to a polygonal approximation and, (ii) a *straight line segment grouping* process leading to a constant curvature approximation which groups adjacent straight line segments into circular arcs whenever feasible. The algorithm repeats the first grouping process using each point on the curve as starting point and the second grouping process using each straight line segment provided by the first process as starting segment. These repetitions lead to a description of the curve composed of sets of constant curvature segments at different scales, some sets being possibly identical and/or overlapping, and/or including overshoot. A global combinatorial method is proposed here to reduce various sets of primitives obtained from a full implementation of the algorithm, to a single non-redundant subset including no-overlapping and no-overshoot. The way to proceed is to classify the results of each grouping process using heuristically-defined *qualitative labels*. Consequently, it is possible to find, for each curve, a set of pairs (*scale, set(s) of constant curvature segments*) which best describe the curve. More formally, a 2D curve is defined by

$$C = \{p_i(x_i, y_i) \mid i \in [1, n] \wedge (x_i \in \mathbb{R} \wedge y_i \in \mathbb{R})\}.$$

### 2. MuscaGrip: Polygonal Approximation

The *point grouping* process splits each curve  $C$  into several subcurves, each of which approximated by a straight line segment. The uniformity criterion associated with point grouping is equivalent to a discrete *mutual co-circularity* criterion among the discretely connected points of the subcurve. A *scale parameter*, acting as a maximum deviation criterion, is associated with a scale measure.

It is assumed that a point chain of one or two points forms a uniform subcurve. A chain of three or more points forms a uniform subcurve if and only if the perpendicular distance of each point of the chain relative to the straight line joining the two endpoints of the chain is less than the scale parameter. The computation of this step is repeated at a number of scales to provide a multi-scale set of polygonal approximations, this using each point of the curve as a starting point.

At the end of the process, at one scale and for each

étape est répétée à plusieurs échelles, en utilisant, à tour de rôle, comme point de départ chaque point de la courbe, ceci afin de fournir un ensemble multi-échelles d'approximations polygonales.

À la fin du processus, à une échelle donnée, et pour un point de départ donné, une approximation polygonale associée à une courbe  $C$  est définie par une liste ordonnée de segments de droites dont les extrémités sont le premier et le dernier points de la sous-courbe de  $C$  (ce processus fait l'objet d'une attention particulière lorsque la courbe est une liste ordonnée de pixels tel que  $C = \{p_i(x_i, y_i) \mid i \in [1, n] \wedge (x_i \in \mathbb{X} \wedge y_i \in \mathbb{Y})\}$ ).

Pour tenir compte de la nature intrinsèque de la courbe (ouverte ou fermée), le processus d'approximation polygonale est effectué de deux façons différentes, [2].

### 3. MuscaGrip: approximation par segments à courbure constante

Pour une approximation polygonale  $PA$ , le processus d'approximation par segments à courbure constante forme le groupement de segments de droites adjacents en arcs de cercles lorsque cela est possible. Le critère d'uniformité associé est basé sur le modèle d'un *polygone régulier* approximant l'arc de cercle dans lequel il est *inscrit*. Le polygone régulier est uniquement défini par  $l$  la longueur de ses côtés (ou  $n$  le nombre de ses côtés) et  $\theta$  l'angle entre deux côtés consécutifs. Soit  $R_{inscribed}$  le rayon de l'arc de cercle dans lequel il est inscrit, et soit  $R_{circumscribed}$  le rayon du cercle qui le circonscrit (correspondant à la longueur de son apothème), la différence entre  $R_{inscribed}$  et  $R_{circumscribed}$  est reliée au paramètre d'échelle.

Soit  $S$  (partition de  $PA$  composée de  $n_{PA}$  segments de droites) une liste ordonnée de  $n_S$  segments de droites (avec  $n_S \leq n_{PA}$ ). Lorsque  $n_S = 1$  ( $S$  n'est composé que d'un seul segment de droite), alors  $S$  est uniforme. Lorsque  $S$  est uniforme, alors l'ensemble  $S'$  composé de  $n_{S'}$  ( $n_{S'} = n_S + 1$ ) segments de droites, et défini par  $S' = \{S \cup \{s'\} \mid s' \in PA\}$  est aussi uniforme si et seulement si:

- 1 •  $n_{S'} \geq 2$ ,
- 2 •  $n_{S'}$  est égal ou inférieur à  $n$ , le nombre de côtés formant le polygone régulier,
- 3 • l'angle  $\theta_{s,s'}$  entre le segment  $s \in S$  (dernier segment de  $S$ ) et  $s'$  (le segment à inclure) est similaire à l'angle  $\theta$  du polygone régulier défini par  $S$  et inscrit dans l'arc de cercle.  $\Delta\theta$ , défini par  $|\theta_{s,s'} - \theta|$ , est le seuil lié à l'angle interne,
- 4 • la longueur  $l_{s'}$  est similaire à la longueur  $l$  des côtés du polygone régulier défini par  $S$  et inscrit dans l'arc de cercle.  $\Delta l$ , défini par  $|l_{s'} - l|$ , est le seuil lié à la longueur des côtés,
- 5 • la différence entre le rayon de l'arc et la longueur de l'apothème associé au polygone défini par  $\theta_{s,s'}$  et  $l_{s'}$ , est égale

starting point, a polygonal approximation associated with a curve  $C$  is defined by an ordered list of straight line segments, whose endpoints are the first and last points of a subcurve of  $C$  (the polygonal approximation process is quite different when the curve is defined by an ordered list of pixels,

$$C = \{p_i(x_i, y_i) \mid i \in [1, n] \wedge (x_i \in \mathbb{X} \wedge y_i \in \mathbb{Y})\}.$$

To take into account the intrinsic nature of the curve, open or closed, the polygonal approximation process is computed in two different ways, [2].

### 3. MuscaGrip: Constant Curvature Approximation

For a polygonal approximation  $PA$ , the constant curvature approximation process aims at grouping adjacent straight line segments into circular arcs whenever feasible. The associated uniformity criterion is based on the model of a *regular polygon* approximating the circular arc into which it is *inscribed*. The regular polygon is uniquely defined by the length  $l$  of its sides (or  $n$ , its number of sides), and  $\theta$ , the angle between two consecutive sides. Let  $R_{inscribed}$  be the radius of the circular arc into which it is inscribed, and  $R_{circumscribed}$  be the radius of the circular arc which it circumscribes (corresponding to the length of its apothem), the difference between  $R_{inscribed}$  and  $R_{circumscribed}$  is related to the scale parameter.

Let  $S$  (a partition of  $PA$  composed of  $n_{PA}$  straight line segments) be an ordered list of  $n_S$  straight line segments (with  $n_S \leq n_{PA}$ ). When  $n_S = 1$  ( $S$  consists of a single straight line segment), then  $S$  is considered as uniform. When  $S$  is uniform, then the set  $S'$  composed of  $n_{S'}$  ( $n_{S'} = n_S + 1$ ) straight line segments, and defined by  $S' = \{S \cup \{s'\} \mid s' \in PA\}$ , is also uniform if and only if:

- 4 •  $n_{S'} \geq 2$ ,
  - 5 •  $n_{S'}$  is equal or less than  $n$ , the number of sides forming the regular polygon,
  - 6 • angle  $\theta_{s,s'}$  between segment  $s \in S$  (the last straight line segment of  $C$ ) and  $s'$  (the straight line segment under test) is similar to angle  $\theta$  of the regular polygon defined by  $S$  and inscribed into the circular arc.  $\Delta\theta$ , defined as  $|\theta_{s,s'} - \theta|$ , is the threshold linked to the interior angle,
  - 7 • length  $l_{s'}$  is similar to the length  $l$  of the sides of the regular polygon defined by  $S$  and inscribed into the circular arc.  $\Delta l$ , defined as  $|l_{s'} - l|$ , is the threshold linked to side length,
  - 8 • the difference between the radius of the arc and the length of the apothem associated to the polygon defined by  $\theta_{s,s'}$  and  $l_{s'}$ , is equal to or less than the scale parameter.
- Once again, this step is repeated using all straight line segments provided by the polygonal approximation as a starting segment. At the end of the process, at each scale, and for each starting segment, a constant curvature approximation, related to a polygonal approximation, is defined by an ordered list of constant curvature

ou inférieure au paramètre d'échelle.

Cette fois encore, l'étape est répétée pour chaque segment de droite formant l'approximation polygonale. À la fin du processus, à une échelle donnée, et pour un segment de droite donné, une approximation en segments à courbure constante, reliée à une approximation polygonale, est définie par une liste ordonnée de segments à courbure constante: segments de droites et/ou arcs de cercles (*produit par le groupement d'une liste ordonnée de segments de droites étant donnés les critères d'uniformité cités ci-dessus*). Encore une fois, pour tenir compte de la nature intrinsèque de la courbe (ouverte ou fermée), le processus d'approximation en segments à courbure constante est effectué de deux façons différentes, [2]. Une étude conduisant à des seuils relatifs à la courbe et à l'échelle de travail est en cours.

#### 4. Classification intra- et inter-échelles

Une charge de calcul importante est requise pour la segmentation multi-échelles et l'approximation d'une courbe  $C$ . Cette approche multi-échelles conduit à un grand nombre de représentations. Parmi celles-ci, seules les plus significatives doivent être conservées. Pour décrire la courbe  $C$ , une classification des résultats obtenus du premier processus de groupement est un bon point de départ pour l'extraction des approximations intéressantes. Pour cela, nous associons une *étiquette qualitative* à chaque approximation polygonale associée à une courbe ouverte ou fermée. Trois étiquettes sont définies, fonction du comportement du processus aux points extrêmes pour une courbe ouverte et à chaque point pour une courbe fermée. De la même manière, une approximation polygonale étiquetée conduit à un ensemble d'approximations en segments à courbure constante, chacune étant à son tour étiquetée de manière qualitative. Trois étiquettes sont définies, fonction des facteurs redondance, chevauchement et dépassement.

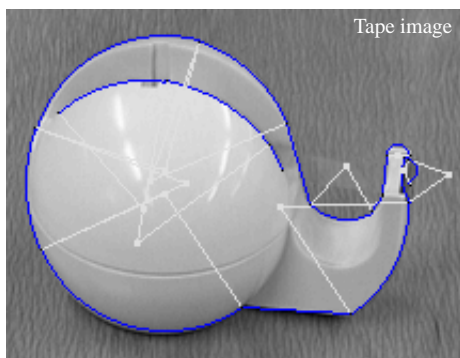


Figure 1. Résultats obtenus sur 2 images à niveaux de gris après application des trois étapes définissant la première phase du module de perception visuelle. Results on 2 grey level 2D images for the first phase of the visual perception module.

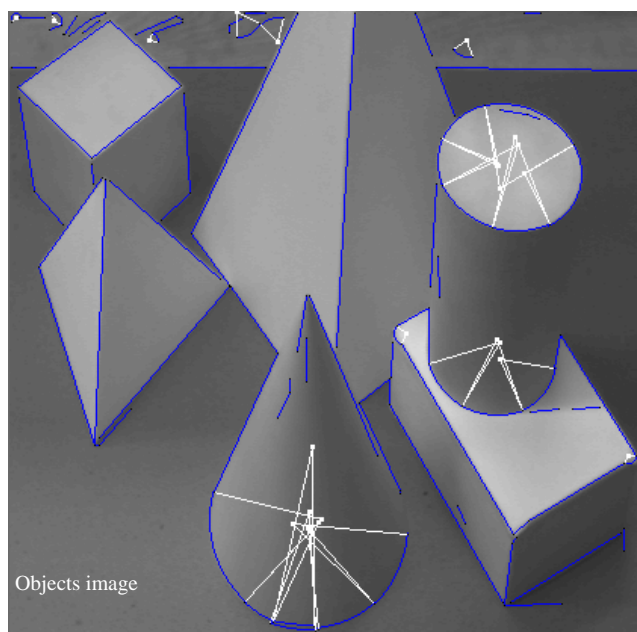
segments: straight line segments and/or circular arcs (*provided by the grouping of an ordered list of straight line segments according to the uniformity criteria listed above*). Again, to take into account the intrinsic nature of the curve (open or closed) the constant curvature approximation process is computed in two different ways, [2]. Future efforts will aim at defining relative thresholds given curve and working scale.

#### 4. Intra- and Inter-Scale Classification

A significant computational load is required for the multiscale segmentation and approximation of a curve  $C$ . This multiscale approach leads to many representations. Among them, only the more salient ones must be considered. To describe the curve  $C$ , a classification of the results obtained from the first grouping process is a good starting point for extracting salient approximations. For that purpose, we associate a *qualitative label* to each polygonal approximation associated with both open or closed curves. Three labels are defined which are function of the behavior of the process at endpoints for an open curve, and at each point for a closed curve. Following the first grouping process, a labeled polygonal approximation leads to a set of constant curvature approximations, each one in turn being qualitatively labeled. Again, three labels are defined which are function of redundancy, overlapping and overshooting.

#### References

- [1] R. Bergevin, and M. Mokhtari, "Multiscale Contour Segmentation and Approximation: an Algorithm Based on the Geometry of Regular Inscribed Polygons", *Computer Vision and Image Understanding*, Vol. 71, No. 1, July, pp. 55-73, 1998.
- [2] M. Mokhtari, and R. Bergevin, "Multiscale Compression of Planar Curves using Constant Curvature Segments", *Proceedings of the 14th International Conference on Pattern Recognition*, Vol. 1, August 16-20, Brisbane, Australia, pp. 744-746, 1998.



*Simon Vaillancourt, étudiant au doctorat (Ph.D. student)*

*Denis Laurendeau, directeur de recherche (advisor)*

*Robert Bergevin, co-directeur de recherche (co-advisor)*

---

## **Groupement perceptuel sur des images à haut degré de complexité**

---

*Le but de ce projet de recherche est de développer une technique pour segmenter des images selon les branchements et l'organisation locale des arêtes de l'image. Ces groupements devront être guidés selon des caractéristiques générales provenant des techniques diverses comme la géométrie intégrale et les probabilités géométriques.*

---

Beaucoup de travaux en vision ont été et sont effectués sur la segmentation et le groupement perceptuel. La plupart de ces travaux réfèrent à des images d'un type spécifique, soit des images d'intérieur dans des environnements plus ou moins encombrés.

Pour utiliser les techniques de groupement perceptuel habituelles, nous avons recours à deux types de primitives, soit des segments de droite et des arcs de cercle. Cette limitation de primitives nous contraint à la segmentation d'images composées d'objets de forme simple.

Nos travaux se situent dans un autre ordre d'idées. Nous visons la segmentation d'images contenant des objets sans aucune limitation due à la forme, à la texture ou toute autre caractéristique photométrique des objets.

Dans un tel cas, la complexité locale de l'image joue un rôle important pour des segmentations de ce type. Peu de travaux ont été faits sur la description de la complexité locale d'une image. C'est pourquoi nous étudierons en détail les caractéristiques déterminantes de la complexité d'une image avant tout. Nous avons fait une recherche sur les techniques de la géométrie fractale, de la géométrie intégrale, de la probabilité géométrique et de la morphologie mathématique sans oublier la topologie. Nous utiliserons ces connaissances pour effectuer des regroupements selon certaines caractéristiques afin de séparer, entre autres, les arêtes provenant des textures et les arêtes provenant des contours.

En bref, il s'agit donc de segmenter des images selon les

## **Perceptual grouping on images of high level complexity**

---

*The purpose of this research project is to achieve a new perceptual grouping technique based on branching and local organization of edges. This grouping will be governed by features from geometric probability and integral geometry.*

---

A large body of work in computer vision focuses on segmentation and perceptual grouping. Most of this work deals with specific type of image which are more or less complex indoor scenes.

Usual grouping techniques are based on two types of primitives, namely straight line segments and circular arcs. This limitation in basic primitives compels us to the segmentation of images containing objects with simple shapes.

Our work deals with another type of image. The objective of this project is the image segmentation without constraints imposed on shape, texture or other photometric features of the objects.

In this case, the local complexity of images has an important role to play for this type of segmentation. A limited amount of work has been done on the description of the local complexity of images. We propose to study in detail the determining features that are embedded in the image complexity. We have conducted a survey on techniques such as fractal geometry, integral geometry, geometric probability, mathematical morphology as well as topology. We intend to use these tools for grouping features in order to separate the edges from textures and the edges from outlines.

In other words, we want to achieve image segmentation based on branching and local organization of edges. This grouping will be governed by features from geometric probability and integral geometry among other tools.

branchements et l'organisation locale des arêtes de l'image. Ces groupements devront être guidés selon des caractéristiques générales provenant des techniques de la géométrie intégrale et des probabilités géométriques.