Détecteur de points caractéristiques dans les images
(Proposition de stage X2005 — 2007)

Frank NIELSEN
(E-mail :`Frank.Nielsen@acm.org`)

Thématique	Traitement d'image, algorithmique
Laboratoire	LIX, École Polytechnique, Paris
Durée	4 à 5 mois (Avril 2008 ∼, anglais ou français)

Positionnement du stage

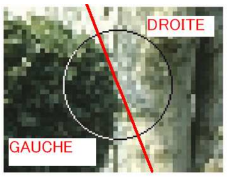
Détecter des points caractéristiques (*features*) dans une image est une procédure fondamentale du traitement d'images aux retombées omniprésentes dans les domaines de la vision par ordinateur (eg., appariemments d'images pour la génération de panoramas, classification de scènes et reconnaissance d'objets, etc.) et des interfaces homme-machine (eg., reconnaissance de mouvements, étiquettage du visuel pour la réalité augmentée, etc.). Le problème fut d'abord abordé simplement dès les années 70 par l'utilisation de filtres (matrices de convolution de Prewitt, Sobel, Harris-Stephens, etc. [1]) malheureusement fort peu robustes et imprécis. Dernièrement, de nouveaux algorithmes efficaces et robustes aux conditions variables d'éclairage (eg. SIFT [2]: *Scale-invariant feature transform*), ont permis de batir des applications fiables et échelonnables (*scalables*) débouchant sur de véritables innovations industrielles (eg., logiciel de reconnaissance des robots domestiques, recherche d'images similaires ou bien encore détection d'objets dans les images, etc.) Bref, la détection de points[1] caractéristiques est un sujet *passionnant* et *d'actualité* comme le prouvent les nombreuses publications dans les conférences annuelles de vision.

Objectifs du stage

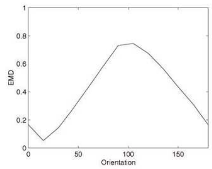
Le stage se propose de tester plusieurs idées *inédites* en étudiant l'opérateur générique compas [3] de Ruzon[2] dans un cadre statistique large.

[1]La notion de "point"/feature doit être prise au sens large. On cherchera notamment à détecter pour chaque feature une matrice de variance-covariance representant sa stabilité , visuellement représentée par une ellipse centrée au point, etc.

[2]Stanford U., travaillant maintenant à Adobe. `http://robotics.stanford.edu/~ruzon/compass/`



Natural Step Edge



$h(\theta)$

L'opérateur compas procéde simplement comme suit : on balaye l'image (de haut en bas, et de gauche à droite) en analysant pour chaque position une fenêtre centrée sur le pixel courant. On cherche alors à mesurer pour toutes les orientations θ de droites passant par le centre et séparant en deux parties la différence $h(\theta)$ entre partie gauche et partie droite. Si le maximum dépasse un seuil, on étiquette le pixel comme "feature" et trouve l'orientation θ de l'arête passant par lui. Dans ce stage, on propose de chercher :

1. à *modéliser* le mieux possible les parties gauches et droites en apprennant une *mixture* statistique jouant le rôle de signature. On étudiera les mixtures dans le cadre général des familles exponentielles [4] en statistique (incluant les distributions usuelles) via l'algorithme de clustering expectation-maximization, que l'on suppose bien meilleures que les traditionelles actuelles mixtures Gaussiennes.

2. à *déterminer* une fonction de *distance* pertinente qui refléte la dissimilarité entre les modèles statistiques des parties gauches/droites. On considèrera des distances sur des distributions/histogrammes, comme la Earth Mover Distance [5] où les distances de Bregman [4].

Notons que plus le voisinage sera grand, plus le feature sera précis à condition que la modélisation statistique soit adéquate. Il est important dêtre robuste aux bruits et de définir aussi un indice de fiabilité pour chaque feature. On regardera comment choisir des tailles de fenêtres variables, et comment affiner le détecteur pour des catégories prédefinies d'objets (bâtiments, voitures, etc.) De plus, si le temps le permet, nous nous attacherons à définir et établir un cadre formel – open benchmark – pour permettre d'évaluer impartiellement les différents algorithmes existants dans la literature (à la manière de ce qui a été fait pour les algorithmes de stéreo-vision ou de segmentations d'images).

Mots clefs: opérateur compas, soft-clustering, images naturelles.
Outils: Java ou C++ (au choix).

Bibliographie

1. F. Nielsen, *Visual Computing: Geometry, Graphics, and Vision*, ISBN 1584504277, Charles River Media, August 2005.

2. D. G. Lowe, Distinctive Image Features from Scale-Invariant Keypoints, International Journal of Computer Vision, 60, 2, pp. 91-110, 2004.

3. M. Ruzon and C. Tomasi, Edge, Junction, and Corner Detection Using Color Distributions, In IEEE Transactions on Pattern Analysis and Machine Intelligence, V. 23, No. 11, pp. 1281-1295, November 2001.

4. A. Banerjee, S. Merugu, I. S. Dhillon, J. Ghosh: Clustering with Bregman Divergences. Journal of Machine Learning Research 6: 1705-1749, 2005.

5. Y. Rubner, C. Tomasi, and L. J. Guibas. A Metric for Distributions with Applications to Image Databases. Proceedings of the IEEE International Conference on Computer Vision, pp. 59-66., Bombay, India, January 1998,