

Modèles Génériques de Représentation et Traitement des Images

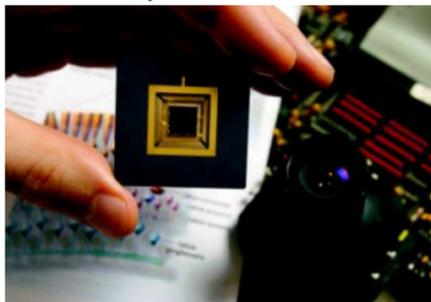
Antoine Manzanera
ENSTA-ParisTech

LRDE - 9 Mai 2012

Contexte et Historique

L'objectif global de notre recherche est la conception de systèmes de vision embarqués efficaces et génériques. Nous nous intéressons depuis longtemps au calcul parallèle, avec un ancrage historique fort pour le calcul cellulaire (rétines programmables).

- ▶ SIMD synchrone / SPMD asynchrone
- ▶ Arithmétique bit-série
- ▶ Morphologie Mathématique
- ▶ Estimation Σ - Δ

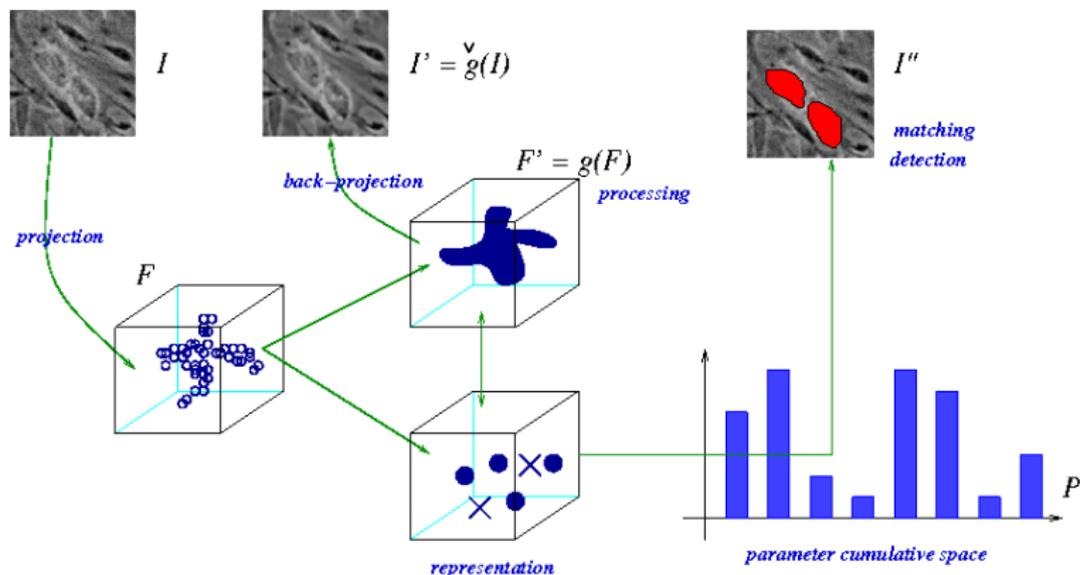


Contexte et Historique

- ▶ Limitations des rétines programmables au bas/moyen niveau
- ▶ Contraintes liées à la conception et fabrication des ASICs
- ▶ Multiplication des modèles et structures de données en traitement d'images
- ▶ Prévalence des systèmes sur étagères (COTS)

Aujourd'hui notre recherche est plus largement orientée vers les algorithmes de vision polyvalents sur systèmes distribués hétérogènes.

Modèle de traitement et de représentation



└ Introduction

└ Vue générale du modèle

Introduction

Contexte et Historique

Vue générale du modèle

Travaux apparentés

Local jets

L'espace LJ

Codages et structures de données

Traitement d'Image

Filtrage NL-means

Flot optique

Détection de mouvement

Représentations visuelles

Singularités et modes de l'espace des caractéristiques

Local jets et transformées de Hough denses

Modèles Implicites de Forme denses

Conclusions

Travaux apparentés

Manifold Image Processing

Les données projetées forment une variété dans l'espace des caractéristiques [Peyré 09]. Un traitement consiste à transformer cette variété, puis à la rétro-projeter dans l'espace image.

Espace d'échelles et dérivées

La décomposition en dérivées multi-échelles est fondée biologiquement [Koenderink 87] et mathématiquement par la théorie des espaces d'échelles [Lindeberg 98].

Bancs de filtres et dictionnaires

Certaines modélisations visuelles pour les textures (p. ex. textons) ou objets (p.ex sacs de mots visuels) combinent bancs de filtres et quantification [Freeman 91], [Rubner 99].

Travaux apparentés

Plus Proche Voisin

Les espaces de caractéristiques sont en général creux et bénéficient de structures telles que les Kd-trees pour réaliser efficacement des opérations de voisinage [Arya and Mount 07].

Espaces d'accumulation

Les techniques d'accumulation dans un espace de paramètres (transformées de Hough) forment un ensemble de méthodes de détection adaptées à la représentation par ensemble de caractéristiques [O'Gorman 76], [Valenti 08].

Introduction

Contexte et Historique

Vue générale du modèle

Travaux apparentés

Local jets

L'espace LJ

Codages et structures de données

Traitement d'Image

Filtrage NL-means

Flot optique

Détection de mouvement

Représentations visuelles

Singularités et modes de l'espace des caractéristiques

Local jets et transformées de Hough denses

Modèles Implicites de Forme denses

Conclusions

Dérivées gaussiennes multi-échelles

Notre espace de caractéristique est l'ensemble des dérivées gaussiennes multi-échelles (r : ordre de dérivation, S : ensemble des échelles).

Projection dans l'espace Local Jet

$$\mathbf{x} \mapsto \hat{\mathbf{x}} = (a_{ij}^\sigma f_{ij}^\sigma(\mathbf{x}))_{i+j \leq r; \sigma \in S}$$

Dérivées gaussiennes multi-échelles

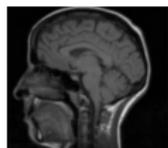
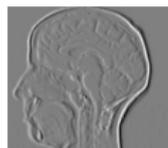
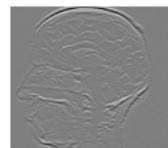
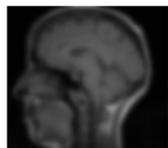
$$f_{ij}^\sigma = f \star \frac{\partial^{i+j} G_\sigma}{\partial x_1^i \partial x_2^j}$$

Local jet normalisé

$$a_{ij}^\sigma = \frac{\sigma^{i+j}}{i+j+1}$$

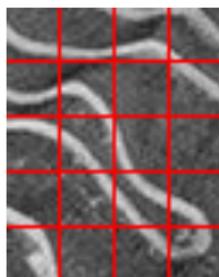
- ▶ G_σ est la fonction gaussienne 2d de variance σ^2 .
- ▶ σ^{i+j} est la normalisation scale space [Lindeberg 98].
- ▶ $i+j+1$ est le nombre de dérivées d'ordre $(i+j)$.

Dérivées gaussiennes multi-échelles

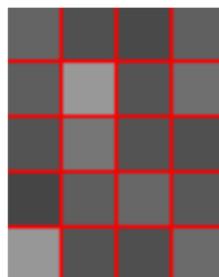
 $f_{00}^{1.0}$  $f_{10}^{1.0}$  $f_{01}^{1.0}$  $f_{20}^{1.0}$  $f_{11}^{1.0}$  $f_{02}^{1.0}$  $f_{00}^{4.0}$  $f_{10}^{4.0}$  $f_{01}^{4.0}$  $f_{20}^{4.0}$  $f_{11}^{4.0}$  $f_{02}^{4.0}$

Le descripteur Local jet

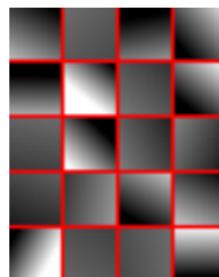
L'utilisation du local jet comme vecteur descripteur dans une mesure de distance peut être interprété en termes d'approximation de la géométrie locale à partir de la formule de Taylor :



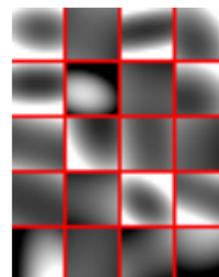
Original



LJ ordre 0



LJ ordre 1



LJ ordre 2

Distances usuelles

Distance mono-échelle

$$d_f^\sigma(\mathbf{x}, \mathbf{y}) = \sum_{i+j \leq r} (f_{ij}^\sigma(\mathbf{x}) - f_{ij}^\sigma(\mathbf{y}))^2$$

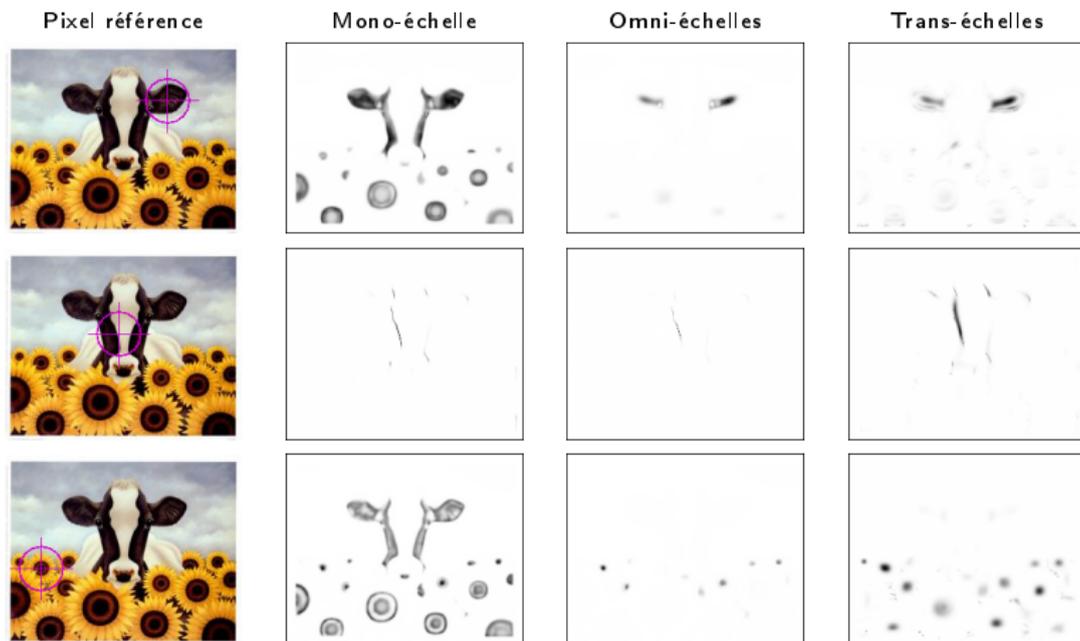
Distance omni-échelles

$$D_f^S(\mathbf{x}, \mathbf{y}) = \sum_{i+j \leq r, \sigma \in S} (f_{ij}^\sigma(\mathbf{x}) - f_{ij}^\sigma(\mathbf{y}))^2$$

Semi-distance trans-échelles

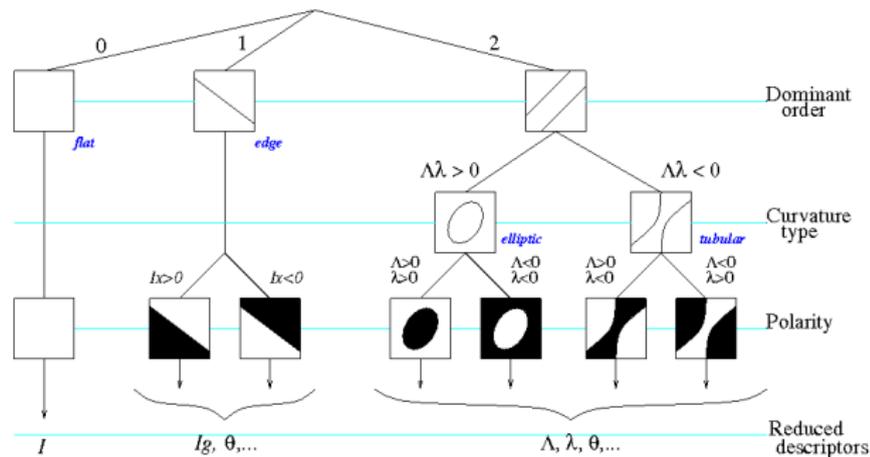
$$d_f^S(\mathbf{x}, \mathbf{y}) = \min_{(\sigma_1, \sigma_2) \in S^2} \sum_{i+j \leq r} (f_{ij}^{\sigma_1}(\mathbf{x}) - f_{ij}^{\sigma_2}(\mathbf{y}))^2$$

Exemples de cartes de similarités



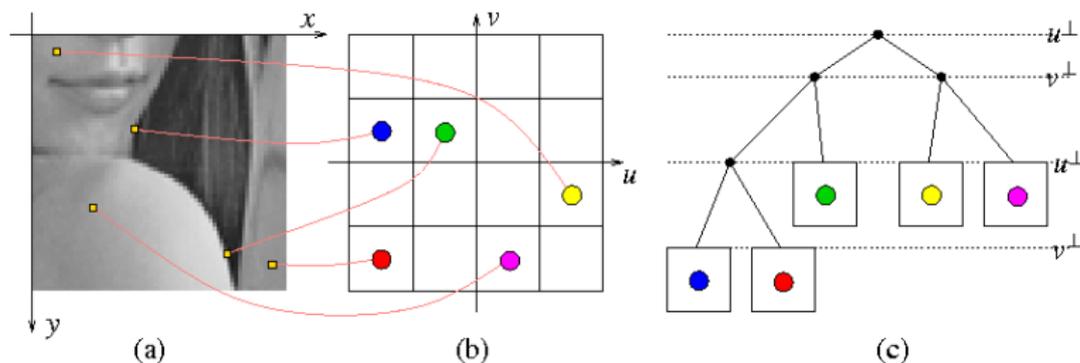
Catégorisation géométrique locale

Un descripteur réduit peut être obtenu en catégorisant le local jet. Voir [Crosier 10].



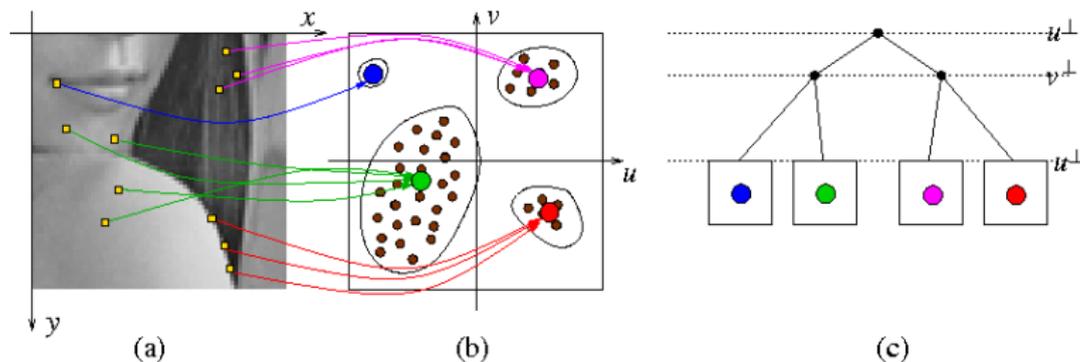
Projection des données et représentation

- ▶ Les données images sont projetées dans l'espace des caractéristiques.
- ▶ Les vecteurs de caractéristiques sont collectés dans un arbre de recherche.
- ▶ Chaque vecteur de caractéristiques est associé à un pixel dans l'espace image.



Quantification de l'espace de caractéristiques

- ▶ La taille de l'espace de caractéristiques peut être réduit par quantification vectorielle.
- ▶ Chaque mot du dictionnaire est associé à un ensemble de pixels dans l'espace image.



Introduction

Contexte et Historique

Vue générale du modèle

Travaux apparentés

Local jets

L'espace LJ

Codages et structures de données

Traitement d'Image

Filtrage NL-means

Flot optique

Détection de mouvement

Représentations visuelles

Singularités et modes de l'espace des caractéristiques

Local jets et transformées de Hough denses

Modèles Implicites de Forme denses

Conclusions

NL-means dans l'espace des caractéristiques

Le filtre NL-means calcule une moyenne pondérée de la valeur de chaque pixel, où les poids sont fonctions de la similarité locale. Ici on utilise simplement une distance dans l'espace LJ...

$$\omega(\mathbf{u}, \mathbf{v}) = e^{-\frac{\|\mathbf{u}-\mathbf{v}\|^2}{h^2}}$$

avec :

$\|\cdot\|$ une norme dans l'espace des caractéristiques.

h un paramètre de décroissance.

et on calcule la moyenne dans un voisinage de \mathbf{x} dans l'espace image (Limited Range),

ou dans un voisinage de $\hat{\mathbf{x}}$ dans l'espace des caractéristiques (Unlimited range).

Limited Range LJ-NL-Means

Les poids (dans l'espace LJ)
sont calculés dans un voisinage
de \mathbf{x} dans l'espace image :

Limited range NL-means

$$f_{LR}^{NL}(\mathbf{x}) = \frac{1}{\zeta(\mathbf{x})} \sum_{\mathbf{y} \in \mathcal{N}(\mathbf{x})} f(\mathbf{y}) \omega(\hat{\mathbf{x}}, \hat{\mathbf{y}})$$

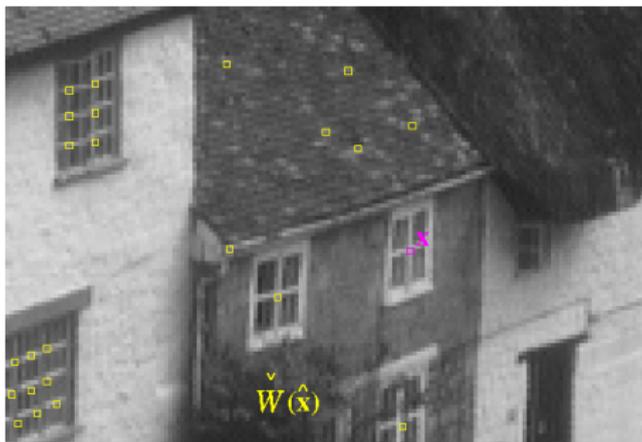


Unlimited Range LJ-NL-Means

Les poids (dans l'espace LJ)
sont calculés dans un voisinage
de $\hat{\mathbf{x}}$ dans l'espace LJ :

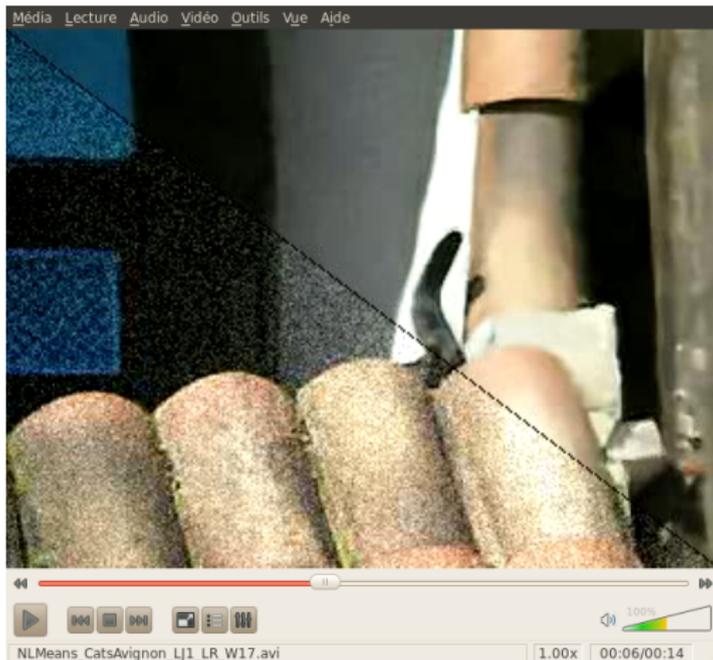
Unlimited range NL-means

$$f_{UR}^{NL}(\mathbf{x}) = \frac{1}{\xi(\mathbf{x})} \sum_{u \in \mathcal{W}(\hat{\mathbf{x}})} f(\check{\mathbf{u}}) \omega(\hat{\mathbf{x}}, \mathbf{u})$$



Débruitage de Vidéo par Local jet NL-Means

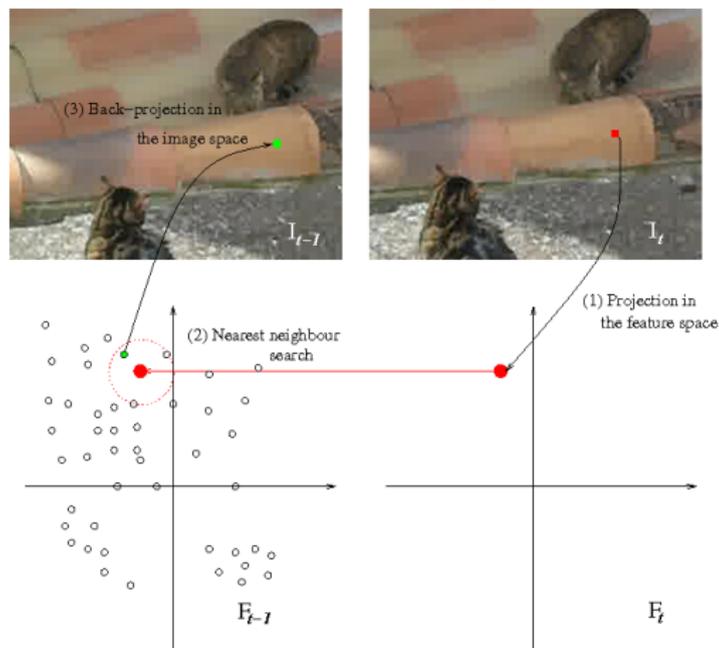
Exemple : Filtrage Local-Jet-NL-Means dans un voisinage espace-temps limité (couleur, ordre 2, une seule échelle).



Mouvement apparent dans l'espace des caractéristiques

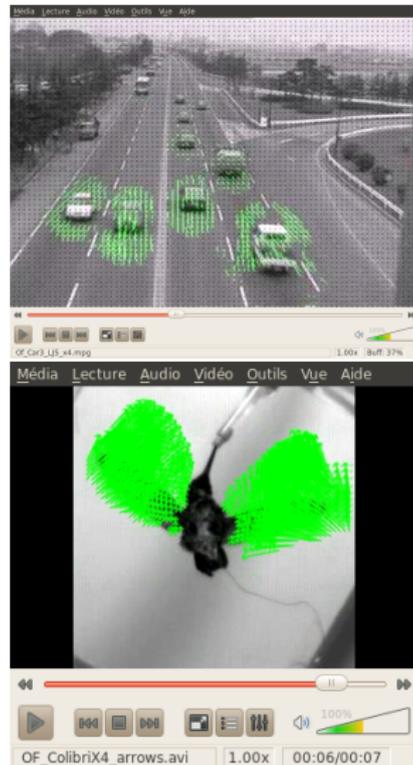
$$u(f_{t-1}, f_t, \mathbf{x}) = \arg \min_{\mathbf{v} \in \mathcal{F}_{f_{t-1}}} d^F(\hat{\mathbf{x}}_{f_t}, \mathbf{v})$$

Dans ce modèle l'estimation du flux optique s'exprime simplement par une recherche du plus proche voisin dans l'espace des caractéristiques :



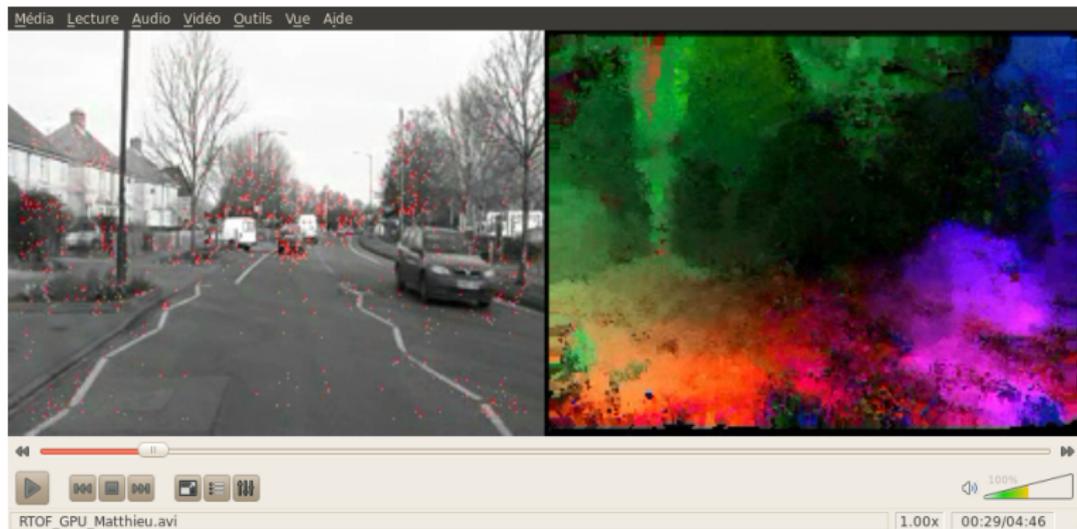
Infinite range optical flow

- + Simplicité conceptuelle
- + Régularisation spatiale implicite
- + Pas de limitation sur la vitesse
- Complexité combinatoire

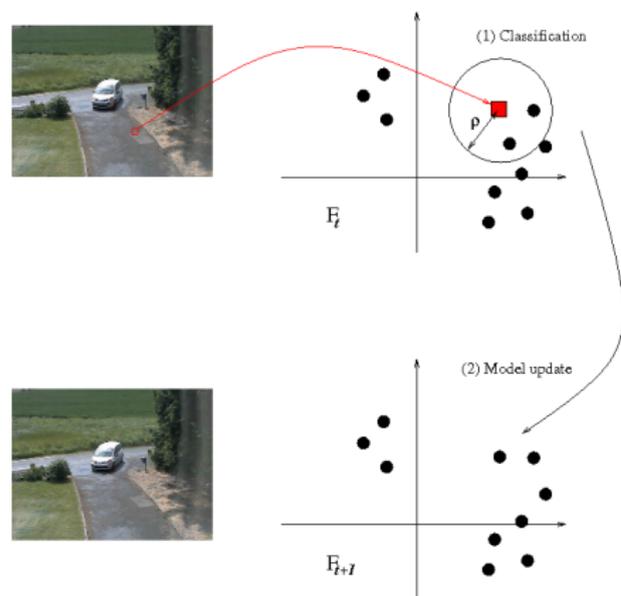


Flux optique temps-réel par plus proche voisin dans l'espace LG sur GPU

Un flux optique dense et sans régularisation spatiale explicite peut être obtenu en implantant une version à portée limitée sur GPU :



Echantillonnage et consensus dans l'espace LJ



- ▶ On modélise le fond statique (background) en échantillonnant les valeurs de chaque pixel dans l'espace du local jet.
- ▶ Les objets mobiles (Foreground) sont classifiés selon un consensus (vote).
- ▶ Le modèle de fond statique est mis à jour en conséquence.

Résultats de la détection d'objets mobiles

- ▶ Espace des caractéristiques : Local jet d'ordre 2, une seule échelle , 3 couleurs.
- ▶ Quantification vectorielle : dictionnaire de 3 000 mots.
- ▶ Nombre d'échantillons : $M = 20$.



Introduction

Contexte et Historique

Vue générale du modèle

Travaux apparentés

Local jets

L'espace LJ

Codages et structures de données

Traitement d'Image

Filtrage NL-means

Flot optique

Détection de mouvement

Représentations visuelles

Singularités et modes de l'espace des caractéristiques

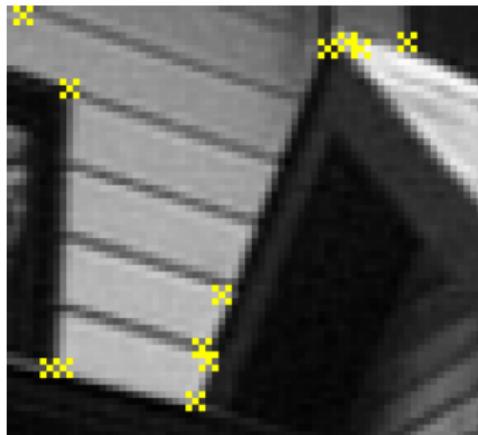
Local jets et transformées de Hough denses

Modèles Implicites de Forme denses

Conclusions

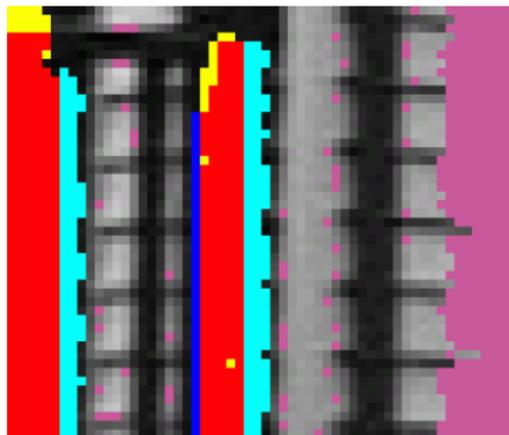
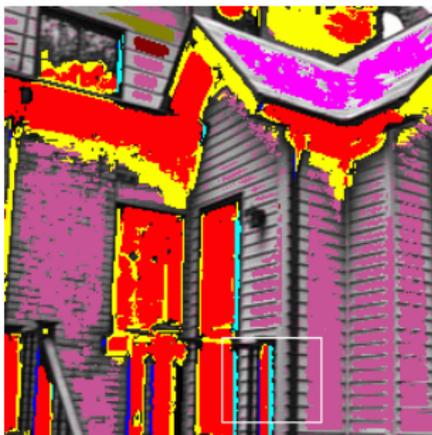
Singularités de l'espace des caractéristiques

La détection des points isolés de l'espace des caractéristiques : fusionner la détection des points d'intérêt et le calcul des descripteurs associés.
Voir aussi [Kervrann 08].



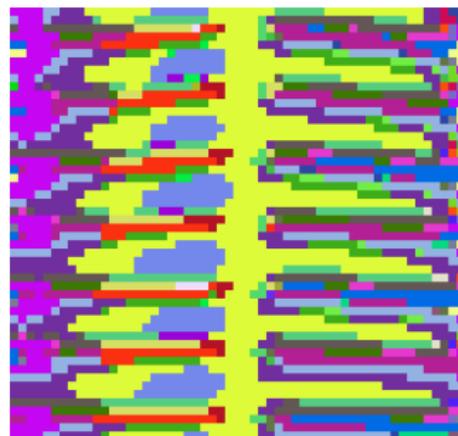
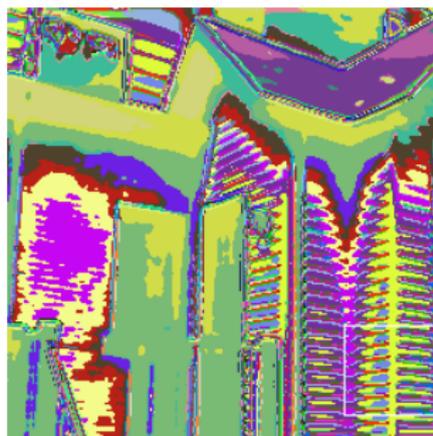
Modes de l'espace des caractéristiques

Les modes (clusters) de l'espace des caractéristiques (voir [Burman 09]), rétro-projetés dans l'espace image, correspondent aux grandes zones homogènes, long contours rectilignes, ou éléments de texture régulière.



Distribution relative au dictionnaire

La distribution des mots du dictionnaire est déjà une représentation possible pour une image ou une catégorie visuelle. Voir par exemple [Rubner 99].



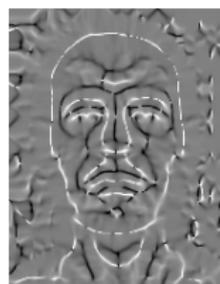
Local jets et transformées de Hough denses

Les dérivées multi-échelles quantifiées peuvent être utilisées avec profit pour indexer un espace de paramètres dans une technique de détection de formes paramétrées fondée sur l'accumulation (vote) dans l'espace des paramètres.



Argument du gradient

($\sigma = 2$)



Courbure de l'isophote

($\sigma = 2$)



Local jet quantifié

(Ordre 2, 5 échelles)

Dérivées partielles et transformées de Hough 1-to-1

Approches traditionnelles

- ▶ éparses : Seules quelques points (contours, points clefs) votent.
- ▶ 1-to-many : Chaque point de l'espace image vote uniformément sur une surface de dimension n dans l'espace des paramètres.
- ▶ many-to-1 : Chaque n -uplet de points de l'espace image vote pour un unique point de l'espace des paramètres.

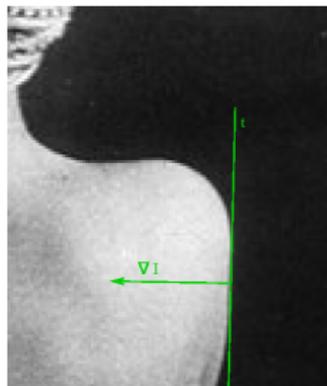
Dérivées partielles et transformées de Hough 1-to-1

TH fondée sur les dérivées partielles

- ▶ denses : Tous les points de l'espace image votent...
- ▶ censitaires : ...mais les votes sont pondérés !
- ▶ 1-to-1 : Chaque point de l'espace image vote pour un point de l'espace des paramètres.

Transformées 1-to-1 : ordre 1

A l'ordre 1, la direction du gradient définit la direction de l'isophote, et donc l'équation de la droite éventuelle. Le poids du vote est le module du gradient.



gradient et droite

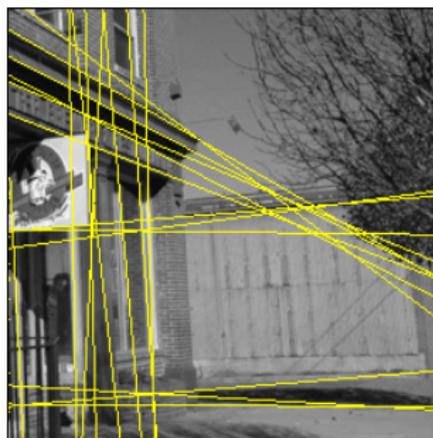


poids du vote

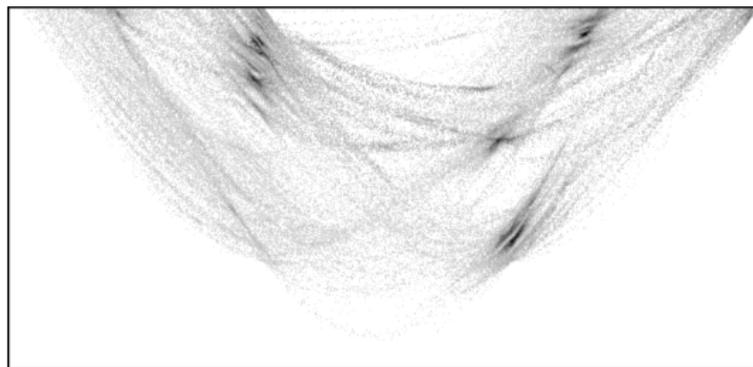


principaux votes

Transformées 1-to-1 : ordre 1



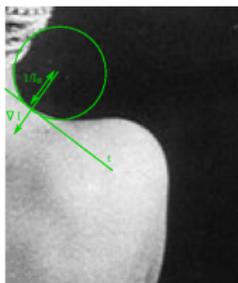
20 meilleures droites



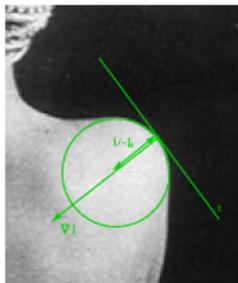
transformée (θ, ρ) 1-to-1

Transformées 1-to-1 : ordre 2

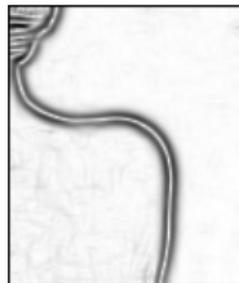
A l'ordre 2, la direction du gradient et la courbure de l'isophote définissent le rayon et le centre du cercle osculateur à la courbe isophote et donc l'équation du cercle éventuel. Le poids du vote est la norme de Frobenius de la matrice hessienne.



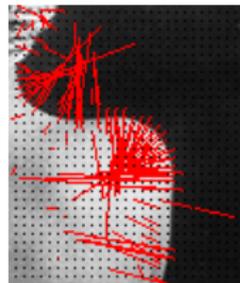
gradient, courbure positive



gradient, courbure négative



poids du vote

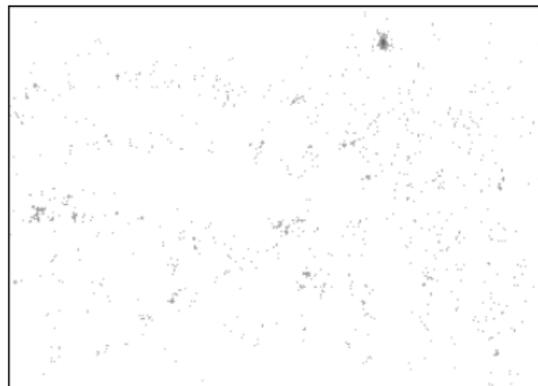


principaux votes

Transformées 1-to-1 : ordre 2



10 meilleures cercles

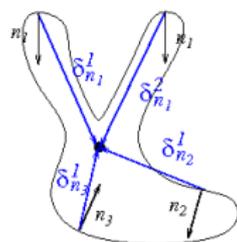


transformée (x, y, ρ) 1-to-1 (coupe $\rho = 19$)

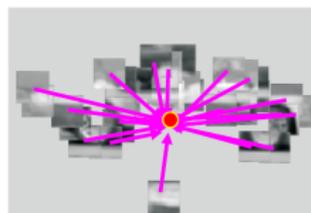
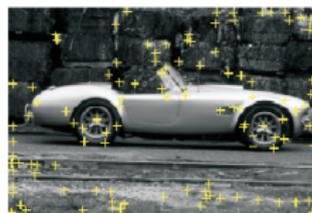
Représentation des objets par R-Tables

Les approches existantes de transformées de Hough généralisées sont *éparses* : elles opèrent sur un ensemble de points caractéristiques (contours, points d'intérêt) [Leibe 04].

$$\text{R-Table} : \{i, \{\vec{\delta}_i^j\}_j\}_i$$

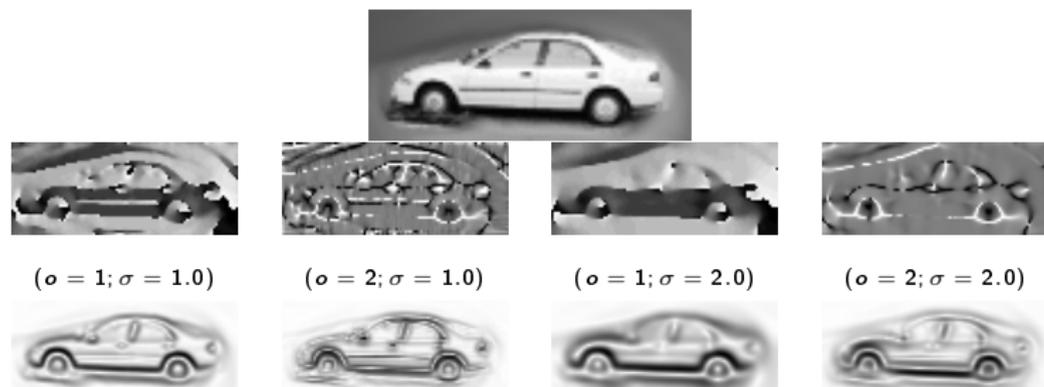


$$\begin{aligned} n_1 &: \{ \delta_{n_1}^1, \delta_{n_1}^2, \dots \}, \\ n_2 &: \{ \delta_{n_2}^1, \dots \}, \\ n_3 &: \{ \delta_{n_3}^1, \dots \}, \\ &\dots \} \end{aligned}$$



R-Tables denses indexées par dérivées

Dans l'approche dense, les index i de la R-Table sont des dérivées multiéchelles quantifiées, disponibles en tout point.



R-Table pondérée : $\{i, \{\delta_i^j, \omega_i^j\}_j\}_i$

Transformée dense généralisée

Détection : pour tout point \mathbf{x} , soit $I(\mathbf{x})$ la mesure quantifiée. Pour toutes les occurrences j de la R-Table associées à $I(\mathbf{x})$, faire :

$$H(\mathbf{x} + \delta_{I(\mathbf{x})}^j) + = \omega_{I(\mathbf{x})}^j$$



Introduction

Contexte et Historique

Vue générale du modèle

Travaux apparentés

Local jets

L'espace LJ

Codages et structures de données

Traitement d'Image

Filtrage NL-means

Flot optique

Détection de mouvement

Représentations visuelles

Singularités et modes de l'espace des caractéristiques

Local jets et transformées de Hough denses

Modèles Implicites de Forme denses

Conclusions

Conclusion

- ▶ **GENERICITE**
 - ▶ Le modèle comprend de nombreux opérateurs de traitement...
 - ▶ ...et recouvre les différents niveaux d'analyse.
- ▶ **INTERET/INNOVATION**
 - ▶ Unification des NL-Means, filtres bilatéraux et autres...
 - ▶ Calcul du flot optique par plus proche voisin.
 - ▶ Point d'intérêt par saillance statistique du descripteur.
 - ▶ Transformée de Hough dense et régulière.
- ▶ **EFFICACITE**
 - ▶ Réduction du descripteur d'apparence (Traitements).
 - ▶ Vote direct à partir de l'index du descripteur (Hough).

Travaux en cours / à faire

- ▶ EVALUATION
 - ▶ Etude comparative.
 - ▶ Etude paramétrique.
- ▶ ACCELERATION
 - ▶ Optimiser la taille du descripteur.
 - ▶ Régulariser et paralléliser la recherche du PPV.
- ▶ DETECTION ET RECONNAISSANCE
 - ▶ Modèles d'objets : généralisation et optimisation des R-Tables denses.
 - ▶ Méthodes cumulatives : estimation de l'ego-mouvement, détection des plans principaux.

Bibliographie (Pour en savoir plus)



[Manzanera 10] A. MANZANERA

Local Jet Based Similarity for NL-Means Filtering
Int. Conf. on Pattern Recognition 2668-2671 (2010)



[Manzanera 11] A. MANZANERA

Local Jet Feature Space Framework for Image Processing and Representation
Int. Conf. on Signal Image Technology and Internet Based Systems
261-268 (2011)



[Manzanera 12] A. MANZANERA

Dense Hough transforms on gray level images using multi-scale derivatives
Working document, ENSTA-ParisTech (2012)

Bibliographie (Concepts et modèles)



[Peyré 09] G. PEYRE

Manifold Models for Signals and Images

Computer Vision and Image Understanding 113(2), 249-260. (2009)



[Lindeberg 98] T. LINDBERG

Feature detection with automatic scale selection

International Journal of Computer Vision 30(2), 77-116. (1998)



[Koenderink 87] J.J. KOENDERINK and A.J. VAN DOORN

Representation of Local Geometry in the Visual System

Biological Cybernetics 55, 367-375. (1987)

Bibliographie (Bancs de filtres et Codebooks)



[Freeman 91] W.T. FREEMAN and E.H. ADELSON

The design and use of Steerable Filters

IEEE Trans. on Pattern Analysis and Machine Intelligence 13(9),
891-906. (1991)



[Crosier 10] M. CROSIER and L.D. GRIFFIN

Using Basic Image Features for Texture Classification

International Journal of Computer Vision 88(3), 447-460. (2010)



[Rubner 99] Y. RUBNER and C. TOMASI

Texture-Based Image Retrieval Without Segmentation

IEEE International Conference on Computer Vision, Kerkyra,
Greece 1018-1024. (1999)

Bibliographie (Outils)



[Mount 97] D.M. MOUNT and S. ARYA

ANN: A Library for Approximate Nearest Neighbor Searching
CGC Workshop on Computational Geometry (1997)

<http://www.cs.umd.edu/~mount/ANN/>



[Burman 09] P. BURMAN and W. POLONIK

Multivariate mode hunting: Data analytic tools with measures of
significance

Journal of Multivariate Analysis 100(6), 1198-1218. (2009)



[Van Vliet 98] L.J. VAN VLIET, I.T. YOUNG and P.W.
VERBEEK

Recursive Gaussian derivative filters

Proc. Int. Conf. on Pattern Recognition vol. 1, 509-514. (1998)

Bibliographie (NL-Means)



[Buades 05] A. BUADES, B. COLL and J.M. MOREL

A non-local algorithm for image denoising

Proc. IEEE Conf. on Computer Vision and Pattern Recognition vol. 2, 60-65. (2005)



[Orchard 08] J. ORCHARD, M. EBRAHIMI and A. WONG

Efficient Non-Local Means Denoising using the SVD

Proc. Int. Conf. on Image Processing 1732-1735. (2008)



[Kervrann 08] C. KERVRANN and J. BOULANGER

Local adaptivity to variable smoothness for exemplar-based image denoising and representation

International Journal of Computer Vision 79(1), 45-69. (2008)

Bibliographie (Détection de mouvement)



[Barnich 09] O. BARNICH and M. VAN DROGENBROECK

ViBe: a powerful random technique to estimate the background in video sequences

International Conference on Acoustics, Speech, and Signal Processing 945-948. (2009)



[Wang 07] H. WANG and D. SUTER

A consensus-based method for tracking: Modelling background scenario and foreground appearance

Pattern Recognition 40(3), 1091-1105. (2007)



[Kim 04] K. KIM, T.H. CHALIDABHONGSE, D. HARDWOOD and L. DAVIS

Background modeling and subtraction by codebook construction
Proc. Int. Conf. on Image Processing 3061-3064. (2004)

Bibliographie (Transformées de Hough)



[O'Gorman 76] F. O'GORMAN AND B. CLOWES

Finding picture edges through collinearity of feature points

IEEE Trans. on Computers C-25 449-456 (1976)



[Leibe 04] B. LEIBE, A. LEONARDIS and B. SCHIELE

Combined object categorization and segmentation with an implicit shape model

ECCV Workshop on Statistical Learning in Computer Vision (2004)



[Valenti 08] R. VALENTI and T. GEVERS

Accurate eye center location and tracking using isophote curvature

Int. Conf. on Computer Vision and Pattern Recognition (2008)