

Bases de données multimédia

Séance 2 : Descripteurs d'images et applications

Valérie Gouet-Brunet
valerie.gouet@cnam.fr

24 Octobre 2007



Plan de la séance

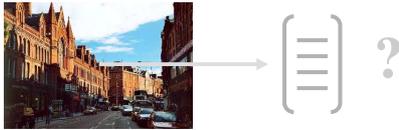
- Rappels, définitions et propriétés
- Descripteurs génériques
 - Description globale de l'image
 - Description locale de l'image
 - Segmentation en régions
 - Points d'intérêt

Bases de données multimédia / Descripteurs d'images et applications

2

● ● ● Définitions

- Indexer une image par son contenu visuel
 - Opération qui consiste à extraire automatiquement d'une image des descripteurs **significatifs** et **compacts** qui seront utilisés et comparés au moment de la recherche ou de la structuration.



- Prise en compte directe du contenu visuel
... mais manifestation fréquente du **gap sémantique**

Bases de données multimédia / Descripteurs d'images et applications 3

● ● ● Définitions

- Descripteur d'image
 - Méthode d'extraction du contenu visuel de l'image (pertinent pour l'application)
 - Exple : histogramme couleur
 - Mesure de similarité associée
 - Exple : intersection d'histogrammes
- Signature d'image
 - Vecteur numérique représentant le contenu visuel de l'image
 - Exple : 1 vecteur de dimension 216 pour l'histogramme couleur
- Espace de description (de représentation) des images
 - 1 image = 1 ou plusieurs points dans un espace multidimensionnel
- Espace de recherche dans la base d'images
 - Structurer l'espace de description pour une recherche efficace (index)

Bases de données multimédia / Descripteurs d'images et applications 4

● ● ● Typologie des descripteurs d'images

- 3 classes de descripteurs
 - Description **globale** de l'image
 - Description *approximative* de toute l'image (recherche d'une *ambiance*)
 - Solutions :
 - Couleur, Texture, Forme
 - 1 point dans un espace de description de grande dimension
 - Description **locale** de l'image
 - Description *précise* des parties de l'image (recherche *précise*)
 - Solutions :
 - Modèles pour la reconnaissance d'objets
 - Génériques (sans modèle) : segmentation en régions, détection de points d'intérêt
 - Plusieurs points dans un espace de description de dimension modérée
 - Descripteurs **spécifiques** (essentiellement biométrie)
 - Empreintes digitales : minuties
 - Visages : EigenFaces, ...

Bases de données multimédia / Descripteurs d'images et applications 5

● ● ● Description globale de l'image



Signature de couleur

↓

Histogramme,...



Signature de texture

↓

Fourier, ondelettes ...



Signature de forme

↓

Moments, EOH, ...

A combiner pour une description plus riche

Bases de données multimédia / Descripteurs d'images et applications 6

● ● ● Signatures de couleur

- Choix d'un espace de représentation des couleurs
 - RVB
 - CMY
 - HSV
 - CIE Lab
 - CIE Luv
 - ...

- Certains espaces colorimétriques sont mieux adaptés à la recherche d'images par contenu
 - Deux couleurs proches en distances le sont aussi pour l'œil.

Bases de données multimédia / Descripteurs d'images et applications 7

● ● ● Les signatures de couleur

Les histogrammes

- Une signature couleur : l'histogramme couleur
 - Échantillonnage des couleurs de l'image
 - Ex : image codée sur 24 bits échantillonnée en 216 couleurs (= 6×6×6)
 - Calcul de l'histogramme pour chaque couleur :

$$\forall c \in C \quad \tilde{h}(c) = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \delta(f(i, j) - c)$$

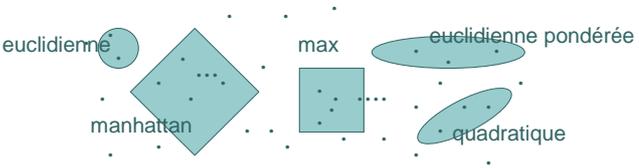
Bases de données multimédia / Descripteurs d'images et applications 8

Les signatures de couleur

Les histogrammes

- Les mesures de similarité en général
 - Distance de Minkowski

$$d(i, j) = \sqrt[q]{(|x_{i_1} - x_{j_1}|^q + |x_{i_2} - x_{j_2}|^q + \dots + |x_{i_d} - x_{j_d}|^q)}$$
 - $q=1$: distance de Manhattan (L_1)
 - $q=2$: distance euclidienne (L_2)
 - $q=\infty$: distance max (L_∞) $d(x, y) = \max_{i=1, d} |x_i - y_i|$



Bases de données multimédia / Descripteurs d'images et applications 9

Les signatures de couleur

Les histogrammes

- Mesures de similarité entre histogrammes couleur
 - Intersection d'histogrammes

$$\cap (H, M) = \frac{\sum_i \min(h_{c_i}, m_{c_i})}{\sum_i m_{c_i}}$$
 - Distances
 - L_1
 - L_2 généralisée $d_{L_2}(H, M) = \sqrt{(H - M)A(H - M)^T}$
 - L_∞ $d_{L_\infty}(H, M) = \max_i |h_{c_i} - m_{c_i}|$

Bases de données multimédia / Descripteurs d'images et applications 10



11

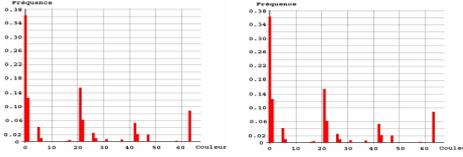
Les signatures de couleur

Les histogrammes

- Avantages des histogrammes
 - Robuste à certaines transformations géométriques de l'image
- Limitations



Contenu visuel différent



Histogrammes similaires

Bases de données multimédia / Descripteurs d'images et applications

12

Les signatures de couleur

Les histogrammes

- Une amélioration : les histogrammes pondérés
 - Principe : tenir compte de la répartition spatiale des couleurs, tout en gardant une signature de même taille
 - Formulation

$$\forall c \in C \quad \tilde{h}(c) = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} w(i, j) \delta(f(i, j) - c)$$
 - Choix du critère de pondération $w(i, j)$
 - Force locale des contours : gradient

$$w(i, j) = (k + \|\nabla(i, j)\|)^\alpha$$
 - $\alpha < 0$ et $k > 0$ (régions uniformes)
 - $\alpha > 0$ et $k > 0$ (contours)
 - Probabilités
 - ...

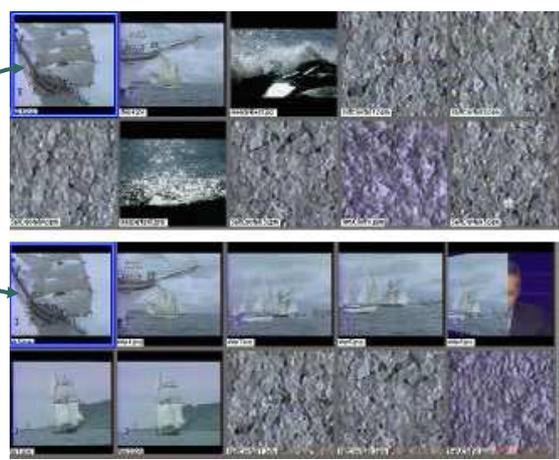
Bases de données multimédia / Descripteurs d'images et applications 13

Les signatures de couleur

Les histogrammes

Requête





Signature couleur classique

Signature intégrant la répartition spatiale de la couleur

<http://www-rocq.inria.fr/imedia>

Bases de données multimédia / Descripteurs d'images et applications 14

Les signatures de couleur

Les histogrammes

- Les histogrammes accumulatifs

0.0 0092.car1.jpg	0.233068 0092.car8.jpg	0.280388 0092.car5.jpg	0.286311 0088.aerialview7.jpg
0.298258 0092.car3.jpg	0.298734 0082.Wood0.jpg	0.299318 0092.car2.jpg	0.30349 0090.argentinabomb7.jpg

Signature histogramme couleur

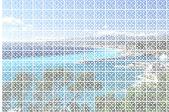
0.0 0092.car1.jpg	0.173526 0092.car8.jpg	0.218645 0092.car5.jpg	0.259065 0088.aerialview7.jpg
0.265694 0092.car3.jpg	0.277073 0092.car6.jpg	0.278409 0088.aerialview7.jpg	0.279707 0092.car4.jpg

Signature histogramme accumulatif

<http://www-rocq.inria.fr/imedia>

Bases de données multimédia / Descripteurs d'images et applications 15

Description globale de l'image



Signature de couleur

↓

Histogramme, ...



Signature de texture

↓

Fourier, ondelettes ...



Signature de forme

↓

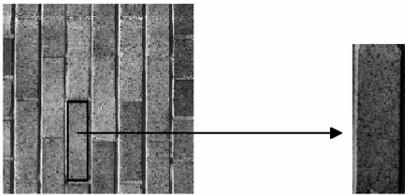
Moments, EOH, ...

A combiner pour une description plus riche

Bases de données multimédia / Descripteurs d'images et applications 16

Les signatures de texture

- Qu'est-ce qu'une texture ?
 - Texture : ensemble de primitives arrangées selon des règles particulières de placement
 - Primitive : ensemble connexe plus ou moins important de pixels de niveaux de gris à peu près semblables (motif de base)
 - Dépend de l'échelle de l'observation



Primitive = brique

Primitive = un ou plusieurs pixels

Bases de données multimédia / Descripteurs d'images et applications 17

Les signatures de texture

- Méthodes de description de la texture
 - Méthodes statistiques
 - Matrice de cooccurrence
 - Longueurs de plage
 - ...
 - Méthodes à base de modèle
 - Décomposition de Wold
 - Modèles Fractals
 - Modèles AR (Autoregressive Models)
 - ...
 - Méthodes fréquentielles (traitement du signal)
 - Fourier
 - Gabor
 - Ondelettes
 - ...

Bases de données multimédia / Descripteurs d'images et applications 18

Les signatures de texture

Méthodes statistiques

- Matrices de cooccurrence

$$MC(c_1, c_2) = \frac{\text{card}(\{(p_1, p_2) / I(p_1) = c_1, I(p_2) = c_2, N(p_1, p_2) = \text{vrai}\})}{\text{card}(\{(p_1, p_2) / N(p_1, p_2) = \text{vrai}\})}$$

- Mesures de Haralick (14 indices en tout)

Contraste $\sum_{i=0}^{n-1} \sum_{j=0}^{n-1} (i-j)^2 MC(i, j)$

Energie $\sum_{i=0}^{n-1} \sum_{j=0}^{n-1} MC(i, j)^2$

Entropie $-\sum_{i=0}^{n-1} \sum_{j=0}^{n-1} MC(i, j) \log(MC(i, j))$

Bases de données multimédia / Descripteurs d'images et applications 19

Les signatures de texture

Méthodes statistiques

- Matrices de cooccurrence



Homogénéité 0.04

Contraste 0.565073650793651

Entropie 0.7786320135903954

Corrélation 0.004844947610446394

Directivité 0.2111111111111111

Uniformité 0.018989888888888889

Bases de données multimédia / D 20

Les signatures de texture

Méthodes fréquentielles

- Distribution spectrale d'énergie
 - Transformée de Fourier
 - Formulation

$$F(u, v) = \iint I(x, y) e^{-i(ux+vy)} dx dy$$
 - Invariance
 - Phase invariante à la luminosité, stable jusqu'à un changement d'échelle de 20%
 - Amplitude invariante à la translation
 - Variantes
 - Transformée de Fourier circulaire (amplitude invariante à la rotation)
 - Transformée de Mellin (amplitude invariante au changement d'échelle)

Bases de données multimédia / Descripteurs d'images et applications 21

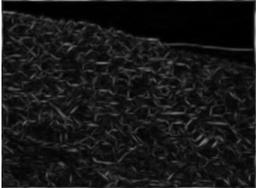
Les signatures de texture

Méthodes fréquentielles

- Filtres de Gabor
 - Filtre passe bande

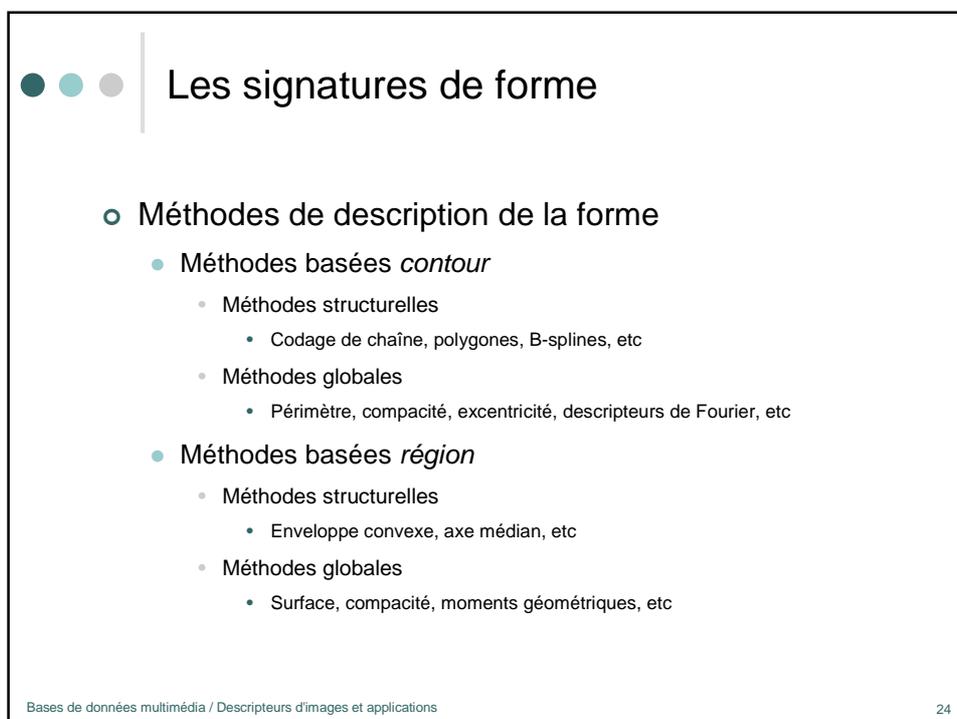
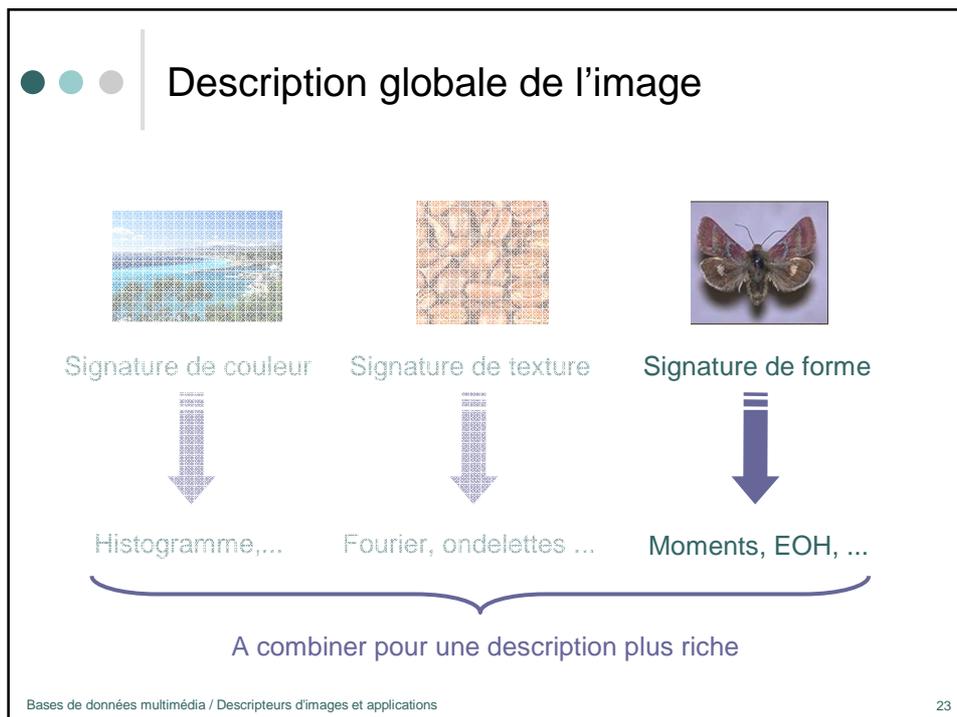
$$G(x, y, \lambda, \theta, \varphi) = \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \cos\left(2\pi\frac{x'}{\lambda} + \varphi\right)$$
 avec $\begin{cases} x' = x \cos \theta + y \sin \theta \\ y' = -x \sin \theta + y \cos \theta \end{cases}$
 - Energie de convolution de l'image avec le filtre

$$S^2(\lambda, \theta) = \sum_{\varphi} \left[\sum_{x,y} G(x, y, \lambda, \theta, \varphi) I(x, y) \right]^2$$



Bases de données multimédia / Descripteurs d'images et applications 22



Les signatures de forme

Descripteurs de Fourier

- Approche globale par descripteurs de Fourier
 1. Linéarisation du contour
 - Courbures
 - Distances au centre (distance centroid method)
 2. Représentation de la série temporelle par transformée de Fourier

Figure 1. Shapes can be converted to time series. A) A bitmap of a human skull. B) The distance from every point on the profile to the center is measured and treated as the Y-axis of a time series of length n (C)

Bases de données multimédia / Descripteurs d'images et applications
25

Les signatures de forme

Codage de Freeman

- Représentation structurale du contour en chaînes de code
 - Codage de Freeman
 - 4, 8 ou 16 orientations

4 directions : 2bits

8 directions : 3bits

16 directions : 4bits

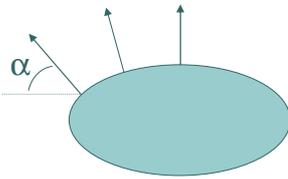
1110100000030303033232303000001010	34x2 = 68bits
221100007777654670000101	24x3 = 72bits
1098815156546788101	16x4 = 64bits

Bases de données multimédia / Descripteurs d'images et applications

Les signatures de forme

EOH

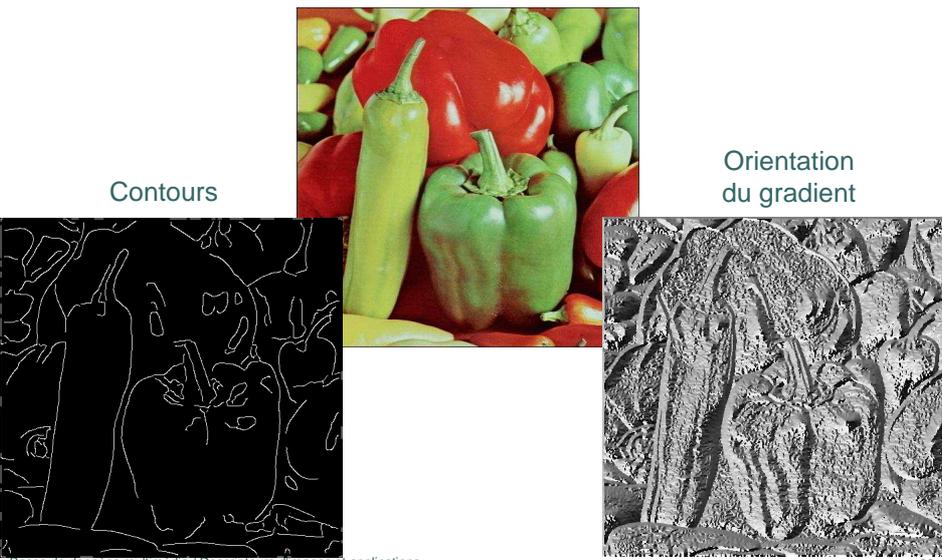
- EOH : Edge Orientation Histogram
 - Principe
 - Extraction des contours de l'image
 - Histogramme de l'angle des gradients sur les contours
 - Invariance aux changements d'illumination affines



Bases de données multimédia / Descripteurs d'images et applications 27

Les signatures de forme

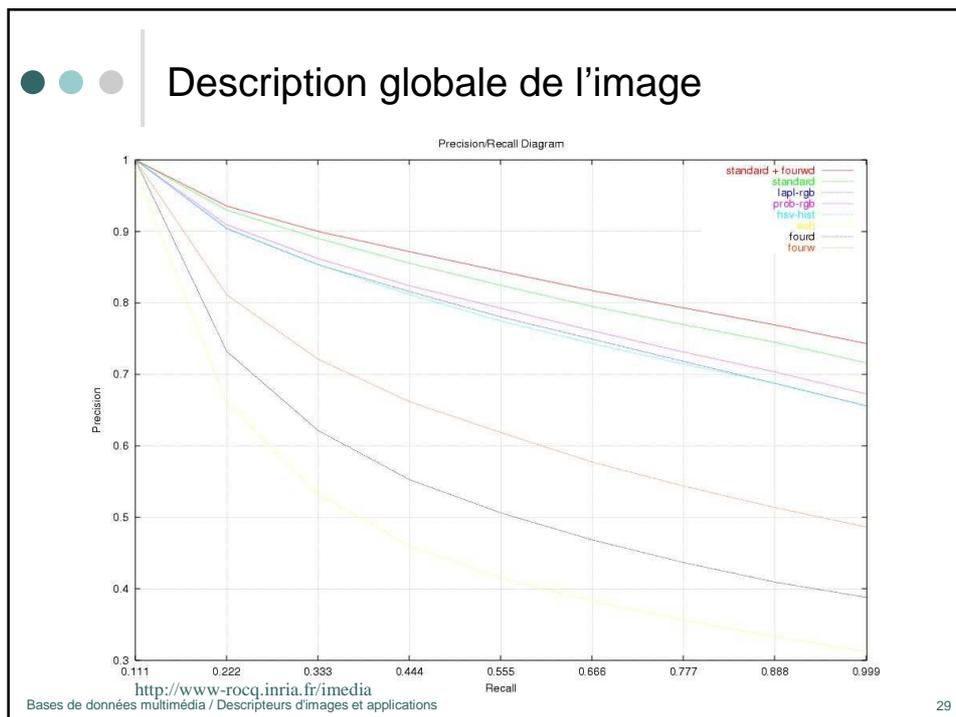
EOH



Contours

Orientation du gradient

Bases de données multimédia / Descripteurs d'images et applications 28



Description locale de l'image

- Objectif
 - Recherche de zones ou d'objets similaires
 - Requêtes dites **partielles**

Je cherche des images contenant des tournesols comme celui-ci

- Solutions
 - Description par le contenu de **zones** de l'image
 - Ajout de relations spatiales
 - Ex : Je cherche des images avec du ciel en haut et de l'herbe en bas

Bases de données multimédia / Descripteurs d'images et applications

30

Description locale de l'image

- Comment choisir les zones à décrire ?
 - Subdivision **systematique**
 - Partitionnement automatique de l'image *indépendante de son contenu*
 - Subdivision **adaptative**
 - Partitionnement automatique de l'image *en fonction de son contenu*

Bases de données multimédia / Descripteurs d'images et applications 31

Description locale de l'image

- Subdivision **systematique**



Requêtes partielles Résultat de la recherche sur la base générique

<http://www-rocq.inria.fr/imedia>

Bases de données multimédia / Descripteurs d'images et applications 32

● ● ● Description locale de l'image

- Subdivision adaptative

Segmentation en régions

Requête à base de régions

Extraction de points d'intérêt

Requête à base de points

$$R = \begin{pmatrix} - \\ - \\ \vdots \\ - \end{pmatrix}$$

$$p = \begin{pmatrix} - \\ - \\ \vdots \\ - \end{pmatrix}$$

Bases de données multimédia / Descripteurs d'images et applications 33

● ● ● Description locale de l'image

- Description de l'image par un ensemble de régions
 - Principe
 - Segmenter **grossièrement** l'image en régions
 - Caractériser **finement** chaque région
 - Objectif
 - Extraction de **zones saillantes** exhibant les **structures dominantes** de l'image
 - Usage
 - Les zones obtenues sont proposées comme requêtes à l'utilisateur

Bases de données multimédia / Descripteurs d'images et applications 34

Description locale de l'image

Segmentation en régions

- Description de l'image par un ensemble de régions
 - Segmentation de l'image en régions
 1. Regroupement des données dans l'espace des couleurs (Luv, Lab)
 - Algorithmes de clustering supervisés/non supervisés
 2. Retro-projection dans l'image
 3. Calcul des ensembles maximaux de pixels connexes
 - Description de chaque région
 - Caractérisation de la distribution des couleurs (couleurs dominantes, couleur moyenne, etc)
 - Surface
 - ...

Bases de données multimédia / Descripteurs d'images et applications 35

Description locale de l'image

Segmentation en régions

Images originales				
Images segmentées				
Couleurs dominantes				

Bases de données multimédia / Descripteurs d'images et applications <http://www.inria.fr/imedia> 36

Description locale de l'image

Segmentation en régions

Zone sélectionnée (cliquée) comme requête

Bases de données multimédia / Descripteurs d'images et applications / <http://www-rocq.inria.fr/imedia>

37

Description locale de l'image

Segmentation en régions

Zone sélectionnée (cliquée) comme requête

Bases de données multimédia / Descripteurs d'images et applications / <http://www-rocq.inria.fr/imedia>

38

Description locale de l'image

Segmentation en régions

- Application à la recherche par image mentale
 - Construction d'un représentant de l'image mentale (la requête) par composition de prototypes de régions (thésaurus de régions)
- Exemple : recherche d'un paysage urbain



Formulation de la requête



Catégorie « régions à dominante verte »

Composition logique correspondante



Bases de données multimédia / Descripteurs d'images et applications
<http://www-rocq.inria.fr/imedia>

39

Description locale de l'image

Segmentation en régions

- Application à la recherche par image mentale (suite)

Résultat de la requête



Exemples d'images rejetées



Bases de données multimédia / Descripteurs d'images et applications
<http://www-rocq.inria.fr/imedia>

40

Description locale de l'image

- Subdivision adaptative

Bases de données multimédia / Descripteurs d'images et applications

41

Description locale de l'image

Les points d'intérêt

- Approche locale par points d'intérêt
 - Extraction automatique de points d'intérêt dans chaque image
 - Un point d'intérêt = site informatif de l'image
 - Défini lorsque les valeurs de l'intensité varient beaucoup dans toutes les directions
 - Description locale de chaque point d'intérêt
 - Exemple : un point peut être décrit par 8 invariants couleur à la translation et à la rotation
 - Mise en place d'une mesure de similarité entre descripteurs de points d'intérêt

Bases de données multimédia / Descripteurs d'images et applications

Description locale de l'image

Les points d'intérêt

Image requête

Points de la requête sélectionnés dynamiquement par l'utilisateur

Recherche des plus proches voisins

Base d'images

Bases de données multimédia / Descripteurs d'images et applications

43

Description locale de l'image

Les points d'intérêt

- Recherche des plus proches voisins :
 - Pour chaque point de la requête, recherche des plus proches voisins (*sphere query*)
 - Seuillage

$$\|p - p_i\| \leq \varepsilon$$
 - Chaque point n'a qu'un seul correspondant par image
 - Problème de la détermination du seuil
 - Algorithme de vote
 - Combinaison des réponses pour les n points de la requête

Bases de données multimédia / Descripteurs d'images et applications

44

Description locale de l'image

Les points d'intérêt

- Algorithme de vote

Image requête

Base d'images

- Distance faible $\leq \epsilon$
- Distance plus élevée $\leq \epsilon$
- Distance $> \epsilon$

Points de la requête sélectionnés dynamiquement par l'utilisateur

Image impliquant le plus grand nombre de points les plus proches de ceux de la requête

Bases de données multimédia / Descripteurs d'images et applications 45

Description locale de l'image

Les points d'intérêt

- Avantages
 - Recherche de parties d'images ou d'objets (élimination de l'arrière-plan)
 - Sélection dynamique de la requête (dite *partielle*)
 - Robuste aux transformations image
 - Robuste aux occultations
- Inconvénients
 - Coûteux en temps de recherche
 - Exple : 500 000 images décrites par 300 points d'intérêt décrits par un vecteur de dimension 17 \Rightarrow recherche dans un espace de dimension 17 contenant 150 millions de points (500 000 \times 300).
 - Nécessité de structurer l'espace de recherche avec un index multidimensionnel

Bases de données multimédia / Descripteurs d'images et applications 46

Description locale de l'image

Les points d'intérêt / extraction

- Extraction de points d'intérêt
 - État de l'art en niveau de gris
 - Moravec (1977)
 - Beaudet (1978)
 - Kitchen et Rosenfeld (1982)
 - Harris et Stephens (1988)
 - Deriche et Faugeras (1990)
 - Heitger (1992)
 - Förstner (1994)
 - Harris précis (1996)
 - Susan (Smith et Brady, 1997)
 - SIFT (1999)
 - État de l'art en couleur
 - Kitchen et Rosenfeld couleur (1998)
 - Harris précis couleur (1998)

Bases de données multimédia / Descripteurs d'images et applications
47

Description locale de l'image

Les points d'intérêt / extraction

- Le détecteur de Kitchen et Rosenfeld
 - Détecteur de coins, ordre 2
 - Principe
 - Un coin est détecté sur un contour aux endroits de plus forte courbure
 - Courbure et gradient élevés
 - Algorithme de détection
 1. Détection des contours
 2. Calcul de la courbure des isophotes
 3. Maximisation de K sur les contours (voisinage local)

$$K = \frac{I_{xx}I_y^2 - 2I_{xy}I_xI_y + I_{yy}I_x^2}{(I_x^2 + I_y^2)^{\frac{3}{2}}}$$

Bases de données multimédia / Descripteurs d'images et applications
48

Description locale de l'image

Les points d'intérêt / extraction

- Exemple de détection



Bases de données multimédia / Descripteurs d'images et applications 49

Description locale de l'image

Les points d'intérêt / caractérisation

- Principe de la caractérisation des points
 - Extraire une information (photométrique, géométrique) décrivant le point d'intérêt, qui soit compacte et invariante d'une image à l'autre
 - Mettre en place une mesure de similarité pour ce descripteur
- Principales transformations envisageables
 - Translation, rotation image
 - Changement d'échelle
 - Changement de point de vue
 - Changement d'illumination



Bases de données multimédia / Descripteurs d'images et applications 50

Description locale de l'image

Les points d'intérêt / caractérisation

- Etat de l'art des méthodes de caractérisation de points
 - Domaine de l'image
 - **En niveau de gris**
 - La corrélation (début des années 80)
 - Le jet local (1987)
 - Les invariants différentiels de Hilbert (1987, 1996)
 - SIFT (1999)
 - ...
 - **En couleur**
 - La corrélation
 - La caractérisation HCP (1998)
 - ...
 - Domaine fréquentiel
 - Transformée de Fourier
 - Transformée de Gabor
 - ...

Bases de données multimédia / Descripteurs d'images et applications 51

Description locale de l'image

Les points d'intérêt / caractérisation

- La corrélation
 - Peut être appliquée sur
 - L'image
 - Ses dérivées

Fenêtres de corrélation

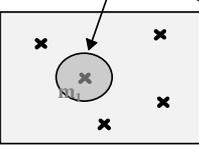


Image 1

Points à mettre en correspondance

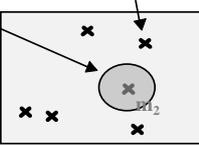


Image 2

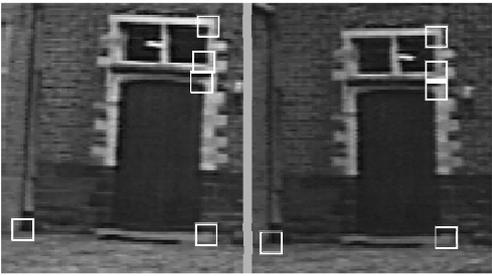
<i>SSD</i>	$scr(F_1, F_2) = \ F_1 - F_2\ ^2$
<i>NSSD</i>	$scr(F_1, F_2) = \ F_1 - F_2\ ^2 / (\ F_1\ \cdot \ F_2\)$ $scr(F_1, F_2) = scr(aF_1, aF_2)$
<i>NCC</i>	$scr(F_1, F_2) = F_1 \cdot F_2 / (\ F_1\ \cdot \ F_2\)$ $scr(F_1, F_2) = scr(aF_1, aF_2)$
<i>ZNSSD</i>	$scr(F_1, F_2) = \ (F_1 - MF_1) - (F_2 - MF_2)\ ^2 / (\ F_1 - MF_1\ \cdot \ F_2 - MF_2\)$ $scr(F_1, F_2) = scr(aF_1 + b_1, aF_2 + b_2)$
<i>ZNCC</i>	$scr(F_1, F_2) = (F_1 - MF_1) \cdot (F_2 - MF_2) / (\ F_1 - MF_1\ \cdot \ F_2 - MF_2\)$ $scr(F_1, F_2) = scr(aF_1 + b_1, aF_2 + b_2)$

Bases de données multimédia / Descripteurs d'images et applications 52

Description locale de l'image

Les points d'intérêt / caractérisation

- La corrélation (suite)



0.9639	-0.3994	-0.1627	-0.3868	0.1914
-0.0533	0.7503	-0.4677	0.5115	0.7193
-0.1826	-0.3905	0.7730	0.1475	-0.7457
-0.2724	0.4878	0.1640	0.7862	0.2077
0.0835	0.5044	-0.4541	0.2802	0.9876

Fenêtres et scores de corrélation obtenus pour les 5 coins en correspondance

Bases de données multimédia / Descripteurs d'images et applications 53

Description locale de l'image

Les points d'intérêt / caractérisation

- Le jet local
 - Approximation par une série de Taylor à l'ordre n

$$f(x_0+x, y_0+y) = f(x_0, y_0) + x \frac{\partial}{\partial x} f(x_0, y_0) + y \frac{\partial}{\partial y} f(x_0, y_0) + \dots + \sum_{p=1}^n x^p y^{n-p} \frac{\partial^n}{\partial x^p \partial y^{n-p}} f(x_0, y_0) + O(x^n, y^n)$$

- Le vecteur jet local

$$J_n(x, y, \sigma) = \{L_{i_1 \dots i_k}(x, y, \sigma) / k = 0, \dots, n\}$$

- Invariance à la rotation (1991) :

$$\begin{cases} L'(\theta) = L_x \cos \theta + L_y \sin \theta \\ L''(\theta) = L_{xx} \cos^2 \theta + 2L_{xy} \sin \theta \cos \theta + L_{yy} \sin^2 \theta \end{cases}$$

Bases de données multimédia / Descripteurs d'images et applications 54

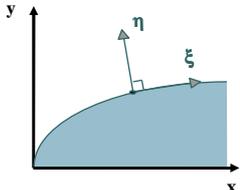
●
●
●

Description locale de l'image

Les points d'intérêt / caractérisation

- Les invariants différentiels de Hilbert
 - Invariance au groupe des déplacements SO(2)
 - Exemple à l'ordre 2 :

$$I \quad I_\eta \quad I_{\eta\eta} \quad I_{\xi\eta} \quad I_{\xi\xi}$$



Coordonnées de Gauge

$$\begin{cases} \eta = \frac{\nabla I}{|\nabla I|} \\ \xi \perp \eta \end{cases}$$

Bases de données multimédia / Descripteurs d'images et applications
55

●
●
●

Description locale de l'image

Les points d'intérêt / caractérisation

- Les invariants différentiels de Hilbert (suite)
 - Une combinaison de ces invariants :
 - Meilleure stabilité numérique

I	I
$I_{\eta\eta} + I_{\xi\xi}$	$I_{xx} + I_{yy}$
I_η	$ \nabla I $
$\frac{I_{\xi\xi}}{I_\eta}$	$\frac{-I_x^2 I_{yy} + 2I_x I_y I_{xy} - I_y^2 I_{xx}}{(I_x^2 + I_y^2)^{\frac{3}{2}}}$
$\frac{I_{\xi\eta}}{I_\eta}$	$\frac{I_x I_y (I_{yy} - I_{xx}) + I_{xy} (I_x^2 - I_y^2)}{(I_x^2 + I_y^2)^{\frac{3}{2}}}$

Bases de données multimédia / Descripteurs d'images et applications
56

Description locale de l'image

Les points d'intérêt / caractérisation

- Les invariants différentiels de Hilbert (suite)
 - Invariants obtenus à l'ordre 3 :

$$\left\{ \begin{array}{l} I \\ I_x^2 + I_y^2 \\ I_{xx}I_x^2 + 2I_{xy}I_xI_y + I_{yy}I_y^2 \\ I_{xx} + I_{yy} \\ I_{xx}^2 + 2I_{xy}^2 + I_{yy}^2 \\ I_{xxx}I_y^3 + 3I_{xxy}I_x^2I_y - 3I_{xxy}I_xI_y^2 - I_{yyy}I_x^3 \\ I_{xxx}I_xI_y^2 + I_{xxy}(I_y^3 - 2I_x^2I_y) + I_{xyy}(I_x^3 - 2I_xI_y^2) + I_{yyy}I_x^2I_y \\ I_{xxy}(2I_xI_y^2 - I_x^3) + I_{xyy}(I_y^3 - 2I_x^2I_y) - I_{yyy}I_xI_y^2 + I_{xxx}I_x^2I_y \\ I_{xxx}I_x^3 + 3I_{xxy}I_x^2I_y + 3I_{xyy}I_xI_y^2 + I_{yyy}I_y^3 \end{array} \right.$$

Bases de données multimédia / Descripteurs d'images et applications 57

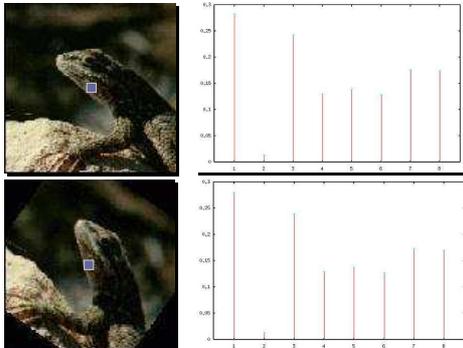
Description locale de l'image

Les points d'intérêt / caractérisation

- Les invariants différentiels de Hilbert (suite)
 - Invariants obtenus à l'ordre 1 généralisés à la couleur

$$\vec{v}_{col}(x, y) = \begin{pmatrix} R \\ \|\nabla R\|^2 \\ G \\ \|\nabla G\|^2 \\ B \\ \|\nabla B\|^2 \\ \nabla R \cdot \nabla G \\ \nabla R \cdot \nabla B \end{pmatrix}$$

Invariants inter-plans



Invariance à la rotation

Bases de données multimédia / Descripteurs d'images et applications 58

●
●
●

Description locale de l'image

Les points d'intérêt / caractérisation

- Robustesse face aux transformations de l'image
 - Adaptation aux changements d'échelle
 - Adaptation aux changements d'illumination
 - Robustesse aux changements de point de vue

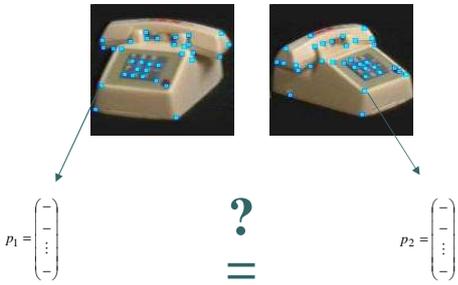
Bases de données multimédia / Descripteurs d'images et applications
59

●
●
●

Description locale de l'image

Les points d'intérêt / caractérisation

○ Comment comparer les vecteurs caractéristiques ?



$$p_1 = \begin{pmatrix} - \\ - \\ \vdots \\ - \end{pmatrix}$$
?

$$p_2 = \begin{pmatrix} - \\ - \\ \vdots \\ - \end{pmatrix}$$

$$=$$

➤ Détermination d'une métrique dans l'espace de représentation des points d'intérêt

Bases de données multimédia / Descripteurs d'images et applications
60

●
●
●

Description locale de l'image

Les points d'intérêt / caractérisation

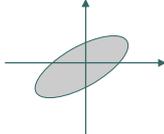
○ Définition d'une mesure de similarité

- Distance Euclidienne :
 - Ne tient pas compte :
 - De l'échelle de chaque composante
 - De la corrélation des données
 - Du bruit contenu dans les données

- Distance pondérée :
 - Ne tient pas compte des corrélations

- Distance de Mahalanobis

$$dist(p, q) = \sqrt{\sum_i (p_i - q_i)^2}$$



$$dist(p, q) = \sqrt{\sum_i w_i (p_i - q_i)^2}$$

Bases de données multimédia / Descripteurs d'images et applications
61

●
●
●

Description locale de l'image

Les points d'intérêt / caractérisation

○ Définition d'une mesure de similarité (suite)

- La distance de Mahalanobis
 - Définition

$$\delta(\vec{v}_1, \vec{v}_2)^2 = (\vec{v}_1 - \vec{v}_2)^T \Lambda^{-1} (\vec{v}_1 - \vec{v}_2)$$
 - Matrice de covariance des composantes

$$\Lambda = \{ \Lambda_{lm}(X_l, X_m) \}_{p \times p}$$

$$\Lambda_{lm}(X_l, X_m) = \frac{1}{n} \sum_{i=1}^n (x_l^i - \bar{x}_l) \cdot (x_m^i - \bar{x}_m)$$

	<i>Obs₁</i>	<i>Obs₂</i>	...	<i>Obs_n</i>
X_1	$\begin{bmatrix} x_1^1 \\ \vdots \\ x_1^n \end{bmatrix}$	$\begin{bmatrix} x_2^1 \\ \vdots \\ x_2^n \end{bmatrix}$...	$\begin{bmatrix} x_1^n \\ x_2^n \\ \vdots \\ x_p^n \end{bmatrix}$
X_2	$\begin{bmatrix} x_1^1 \\ \vdots \\ x_1^n \end{bmatrix}$	$\begin{bmatrix} x_2^1 \\ \vdots \\ x_2^n \end{bmatrix}$...	$\begin{bmatrix} x_1^n \\ x_2^n \\ \vdots \\ x_p^n \end{bmatrix}$
\vdots	$\begin{bmatrix} x_1^1 \\ \vdots \\ x_1^n \end{bmatrix}$	$\begin{bmatrix} x_2^1 \\ \vdots \\ x_2^n \end{bmatrix}$...	$\begin{bmatrix} x_1^n \\ x_2^n \\ \vdots \\ x_p^n \end{bmatrix}$
X_p	$\begin{bmatrix} x_1^1 \\ \vdots \\ x_1^n \end{bmatrix}$	$\begin{bmatrix} x_2^1 \\ \vdots \\ x_2^n \end{bmatrix}$...	$\begin{bmatrix} x_1^n \\ x_2^n \\ \vdots \\ x_p^n \end{bmatrix}$

- Seuillage
 - δ^2 suit la loi du Khi 2

Bases de données multimédia / Descripteurs d'images et applications
62

Description locale de l'image

Les points d'intérêt / caractérisation

- Définition d'une mesure de similarité (suite)
 - Difficultés
 - Complexité du calcul
 - Solution : changement de base

$$\delta(\vec{v}_1, \vec{v}_2)^2 = \left\| \sqrt{D}P\vec{v}_2 - \sqrt{D}P\vec{v}_1 \right\|^2 \quad \text{avec} \quad \Lambda^{-1} = P^T D P$$

- Estimation de Λ :
 - A partir de quels échantillons ?
 - Sur l'ensemble des points ? Λ non discriminante.
 - Sur plusieurs séries de points ? Comment sélectionner les échantillons ?
 - Solutions :
 - Suivi manuel des points
 - Classification automatique de l'espace de représentation

Bases de données multimédia / Descripteurs d'images et applications 63

Description locale de l'image

Les points d'intérêt / caractérisation

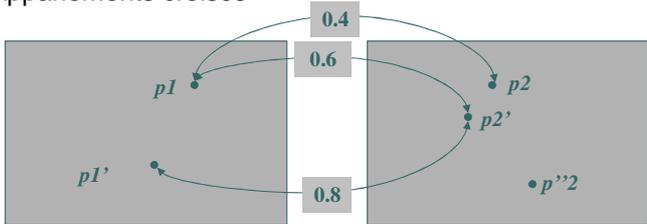
- Procédure d'appariement
 - Si aucune information sur les positions des caméras :
 - Si faible déplacement d'une caméra à l'autre :

Bases de données multimédia / Descripteurs d'images et applications 64

Description locale de l'image

Les points d'intérêt / caractérisation

- Procédure d'appariement (suite)
 - Seuillage
 - Algorithme « Winner takes all »
 - Appariements croisés



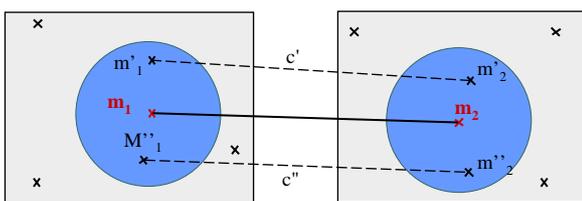
- Dans la pratique, on garde plusieurs appariements par point, puis on élimine les appariements ambigus en utilisant des contraintes géométriques

Bases de données multimédia / Descripteurs d'images et applications 65

Description locale de l'image

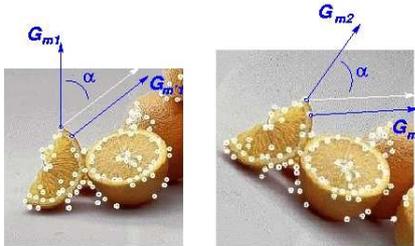
Les points d'intérêt

Contraintes de voisinage



Contraintes de rigidité locale :

Angle α entre 2 gradients couleur



Bases de données multimédia / Descripteurs d'images et applications 66

Description locale de l'image

Les points d'intérêt

Requête sur une partie d'image

Première réponse en utilisant une description globale



Première réponse en utilisant la description locale basée sur les points



Bases de données multimédia / Descripteurs d'images et applications 67

Description locale de l'image

Les points d'intérêt

Requête



IKONIA v0.3 (Main Window)

File Options Help

 0 0 flore_a_000213_90.jpg	 35 561295 flore_a_000214_25.jpg	 39 603794 flore.ids_178a_117.jpg	 45 370834 flore_a_000302_04.jpg
 45 445229 flore_a_000468_84.jpg	 46 139389 flore_a_000135_17.jpg	 46 661137 flore_a_000426_40.jpg	 46 956783 flore_a_000427_54.jpg
 47 038002 516072.jpg	 48 149837 516057.jpg	 48 158958 flore.00031553.jpg	 48 490685 flore_a_000488_83.jpg
 48 725494 516081.jpg	 48 641308 516018.jpg	 48 87936 flore.ids_180a_132.jpg	 48 997254 516063.jpg

Navigation Panel: Prev Next Back Shuffle

Query Panel: Visual Feedback Object Quit Keywords

[Database: Demo_500ppi] [Filename: 516072.jpg] [Distance: 47.038002]

<http://www-rocq.inria.fr/imedia>

Bases de données multimédia / Descripteurs d'images

Description locale de l'image

Les points d'intérêt

Requête





Bases de données multimédia / Descripteurs d'images <http://www-rocq.inria.fr/imedia>

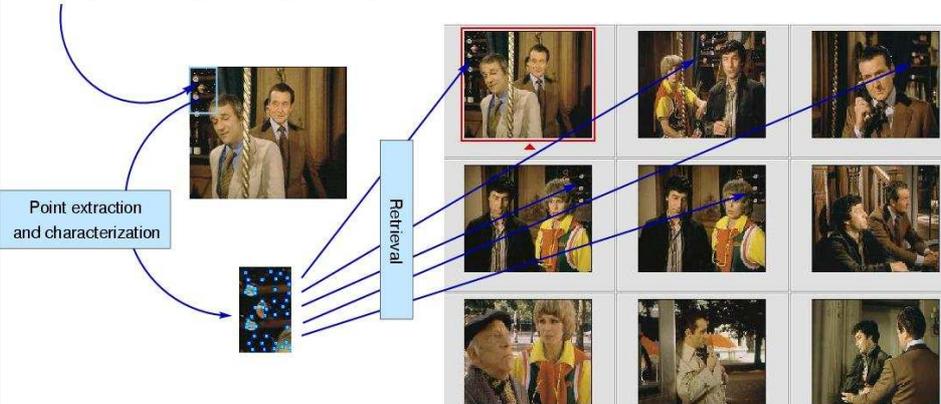
Description locale de l'image

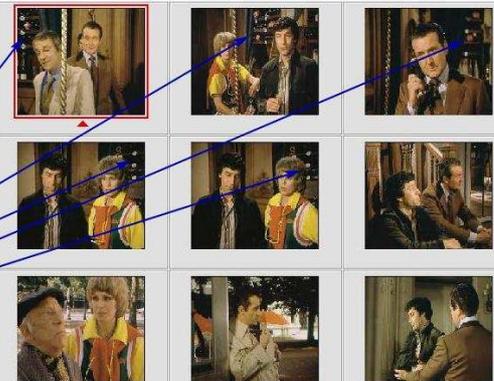
Les points d'intérêt

- Support d'enquête (Police Judiciaire française)

The query : *"I am looking for the images involving the room where there is this wine storeroom".*

Point extraction and characterization





<http://www-rocq.inria.fr/imedia>

Description locale de l'image

Les points

Recherche de motifs particuliers dans une collection d'œuvres d'art

Navigation Panel: Prev, Next, Back, Shuffle

Query Panel: Retrieve, Feedback, Region, Keywords

Bases de données multimédia / Descripteurs d'images et applications
<http://www-rocq.inria.fr/imedia>

Description locale de l'image

Les points

Recherche d'objets particuliers dans une collection d'œuvres d'art (bases d'objets volés)

Navigation Panel: Prev, Next, Back, Shuffle

Query Panel: Retrieve, Feedback, Region, Keywords

Bases de données multimédia / Descripteurs d'images et applications
<http://www-rocq.inria.fr/imedia>

Description locale de l'image

Les points d'intérêt

- Détection et identification de logos (TF1)
 - Aide à l'authentification des images (droits)
 - Jusqu'à présent déterminée manuellement par les documentalistes
 - Ici estimée automatiquement et proposée au documentaliste sous forme d'annotation textuelle (hybridation texte-image)

Bases de données multimédia / Descripteurs d'images et applications 73

Description locale de l'image

Les points d'intérêt

- Détection et identification de logos (TF1)

	 Saab (85%)  L'Equipe (91%)  TF1 (84%)	  
	-	-
	 Saab (85%)  Saab (76%)  Saab (68%)  Saab (54%)	

<http://www-rocq.inria.fr/imedia>
 Bases de données multimédia / Descripteurs d'images et applications 74

Description locale de l'image

Les points d'intérêt dans les vidéos

- Les points d'intérêt dans les vidéos
 - Approches spatio-temporelles pures
 - Points d'intérêt spatio-temporels
 - Pour quoi faire ? Reconnaissance d'actions, de comportements
 - Approches basées sur l'estimation de la trajectoire des points d'intérêt
 - Modélisation de la variabilité des points d'intérêt
 - Analyse du comportement des points dans la vidéo
 - Pour quoi faire ? Reconnaissance d'objets, détection de copies

Bases de données multimédia / Descripteurs d'images et applications 75

Description locale de l'image

Les points d'intérêt dans les vidéos

- Points d'intérêt spatio-temporels
 - Généralisation des points de Harris et Stephens (2004)
 1. Calcul des dérivées premières dans l'espace image et dans le temps (lissages gaussiens de paramètres σ et τ)
 2. Calcul de M avec un lissage gaussien de paramètres $\sigma\sigma$ et $\tau\tau$
$$M = S_{\sigma\sigma, \tau\tau} * \begin{bmatrix} I_x^2 & I_x I_y & I_x I_t \\ I_x I_y & I_y^2 & I_y I_t \\ I_x I_t & I_y I_t & I_t^2 \end{bmatrix}$$
 3. Calcul des maxima locaux de l'opérateur suivant :

$$\text{Det}(M) - k \cdot \text{Trace}^3(M)$$
 - Adaptation indépendante au mouvement de la caméra

Bases de données multimédia / Descripteurs d'images et applications 76

Description locale de l'image

Les points d'intérêt dans les vidéos

- Points d'intérêt spatio-temporels (suite)
 - Généralisation des points de Harris et Stephens (2004)
 - Application à la reconnaissance d'actions



<http://www.nada.kth.se/~laptev/>



<http://www.nada.kth.se/~laptev/>

Bases de données multimédia / Descripteurs d'images et applications 77

Description locale de l'image

Les points d'intérêt dans les vidéos

- Approches basées sur l'estimation de la trajectoire des points d'intérêt
 - Motivations
 - Les vidéos montrent des objets avec des apparences visuelles *variables* d'une trame à l'autre mais facilement *reliables* par des approches de suivi
 - Les points suivis sur plusieurs trames sont plus robustes
 - L'information de mouvement est très discriminante

Bases de données multimédia / Descripteurs d'images et applications 78

Description locale de l'image

Les points d'intérêt dans les vidéos

- Approches basées sur l'estimation de la trajectoire des points d'intérêt (suite)
 - Principe de l'indexation pour la reconnaissance d'objets
 1. *Suivi* de points d'intérêt spatiaux classiques
 - Appariement entre trames consécutives
 - Utilisation de modèles de trajectoires possible



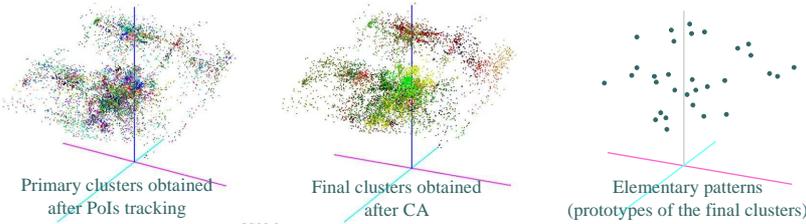
2. *Elimination* (ou traitement particulier) des points peu robustes

Bases de données multimédia / Descripteurs d'images et applications 79

Description locale de l'image

Les points d'intérêt dans les vidéos

- Approches basées sur l'estimation de la trajectoire des points d'intérêt (suite)
 - Principe de l'indexation pour la reconnaissance d'objets
 3. *Structuration* de l'espace de description des points
 - Clusters modélisant la variabilité de la description le long de la séquence
 4. *Compression* de l'information
 - Remplacer les points d'intérêt par les prototypes des clusters



Bases de données multimédia / Descripteurs d'images et applications 80

Description locale de l'image

Les points d'intérêt dans les vidéos

- Approches basées sur l'estimation de la trajectoire des points d'intérêt (suite)
 - Application à la vidéosurveillance / annotation automatique de vidéos (2005)



Duck -> Right view



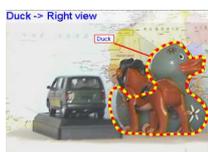
Duck -> Right view



Duck -> Left view



Duck -> Right view



Bases de données multimédia / Descripteurs d'images et applications 81

Description locale de l'image

Les points d'intérêt dans les vidéos

- Approches basées sur l'estimation de la trajectoire des points d'intérêt (suite)
 - Application à la détection de copies (INA, 2005)





Vidéos similaires mais qui ne sont pas des copies

Vidéos aux contenus différents mais qui sont des copies

<http://www-rocq.inria.fr/imedia/> / <http://www.ina.fr>

Bases de données multimédia / Descripteurs d'images et applications 82

Description locale de l'image

Les points d'intérêt dans les vidéos

- Approches basées sur l'estimation de la trajectoire des points d'intérêt (suite)
 - Application à la détection de copies (INA, 2005)
 - Caractérisation des trajectoires de points
 - Attribution d'un *label de comportement*

Points en mouvement

Points immobiles

Point rare

<http://www-rocq.inria.fr/imedia> / <http://www.ina.fr>

Bases de données multimédia / Descripteurs d'images et applications 83

Description locale de l'image

Les points d'intérêt dans les vidéos

- Approches basées sur l'estimation de la trajectoire des points d'intérêt (suite)
 - Application à la détection de copies (INA, 2005)

Requêtes

Copies

Fausse alarmes (images similaires)

Réponses

<http://www-rocq.inria.fr/imedia> / <http://www.ina.fr>

Bases de données multimédia / Descripteurs d'images et applications 84

Description locale de l'image

Les points d'intérêt dans les vidéos

- Approches basées sur l'estimation de la trajectoire des points d'intérêt (suite)
 - Application à la détection de copies (INA, 2005)



Classification de vidéos aux contenus thématiques similaires

<http://www-rocq.inria.fr/imedia/> / <http://www.ina.fr>