

Bases de données multimédia V – Description globale

ENSIMAG

2014-2015

Matthijs Douze & Karteek Alahari

Inria



Description globale

- **Une description globale** est une représentation de l'image dans son ensemble, sous la forme d'un vecteur de taille fixe
- Elle se contente le plus souvent d'exploiter un indice visuel unique
- Caractéristiques
 - ▶ un type de vecteur de description par objet
 - ▶ mesure de (dis-)similarité définie sur l'espace de ces descripteurs

Inria



Exemple de description globale : histogramme de couleur

- Chaque pixel est décrit par un vecteur de couleur
 - ▶ par exemple un vecteur $RGB \in \mathbb{R}^3$, mais le plus souvent on utilise un autre espace de couleur plus approprié
- L'ensemble des vecteurs de couleurs forme une distribution
 - ▶ description de la distribution avec un histogramme
 - ▶ Nécessite discrétisation de l'espace + normalisation
- Comparaison de deux histogrammes par une mesure de dissimilarité, par exemple la « distance » du Chi-2 ou earth mover's distance (au tableau)



Color indexing, Swain & Ballard, IJCV 1991

The earth mover's distance, multi-dimensional scaling, and color-based image retrieval Y Rubner, LJ Guibas, C Tomasi - Proceedings of DARPA Image, 1997

Inria



Distance -> représentation

- Multi-dimentional scaling

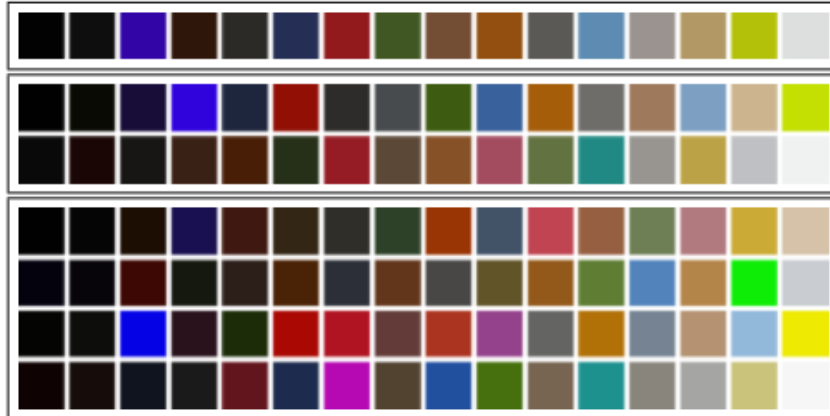


Inria



Palette adaptative

- Apprentissage de la palette par k-means
 - ▶ échantillonnage qui évite d'écraser les couleurs rares
- normalisation d'histogramme



Inria

« bag-of-Colors for improved image search », C. Wengert, M. Douze, H. Jégou, ACM MM 2011

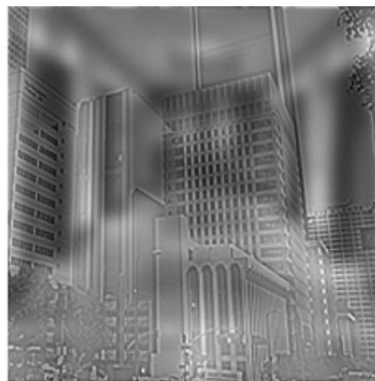


« apparence globale » de l'image : descripteur GIST

- Similaire à SIFT sur une pyramide d'images, avec image = patch.
- Peut prendre en compte la couleur si appliqué sur les canaux R, G, B séparément.



(a)



(b)

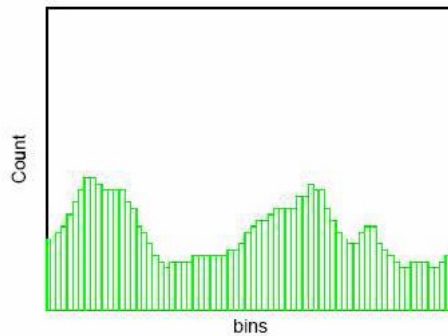
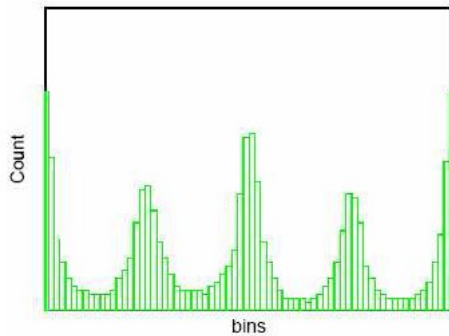
- [Modeling the shape of the scene: a holistic representation of the spatial envelope, Aude Oliva, Antonio Torralba, IJCV 2001], [Gist of the scene, A. Oliva, Neurobiology of attention, 2005]

Inria



Exemple de description globale : contours

- Exemple de descripteur de texture
- Descripteur = histogramme d'orientation des contours



Inria



Aggrégation de descripteurs locaux -- BoW

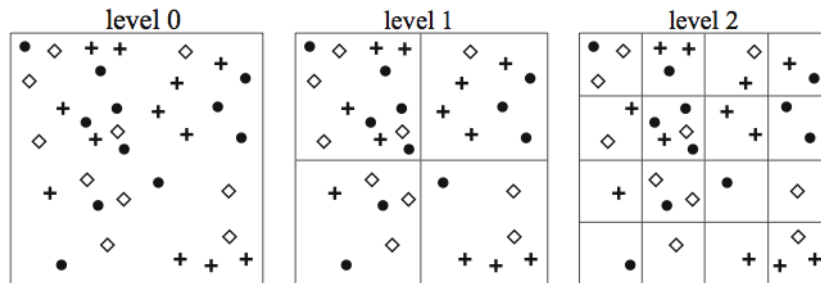
- Définir un descripteur global à partir de l'ensemble des descripteurs locaux d'une image
 - ▶ compact...
 - ▶ facile à manipuler, comparer: un vecteur par image
- La base: "bag of features"
 - ▶ bag = on perd l'ordre, la géométrie
- Quantification: "bag of words"
 - ▶ 1 descripteur local -> un entier (par kmeans quantification vectorielle)
 - ▶ histogramme des entiers = descripteur global

Inria



“pyramide” spatiale: comment garder la géométrie

- Calculer des agrégats de descripteurs sur des sous-images
 - ▶ géométrie rigide
 - ▶ taille des vecteurs multipliée



Beyond Bags of Features: Spatial Pyramid Matching for Recognizing Natural Scene Categories, S. Lazebnik, C. Schmid, J. Ponce, CVPR06

Inria



Autres méthodes local → global

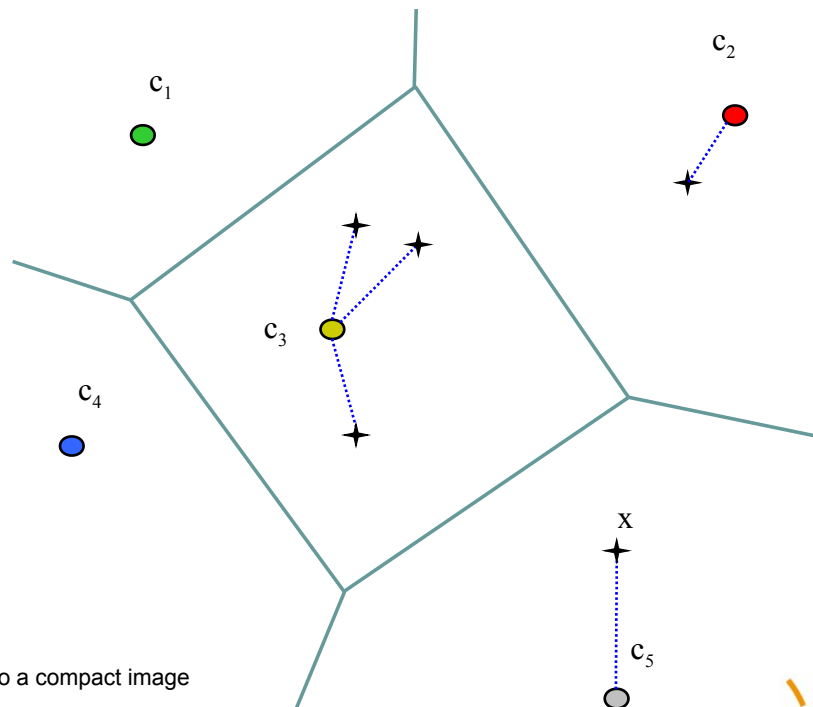
- Bag of features
 - ▶ entrée : ensemble de descripteurs locaux
 - ▶ sortie : vecteur descripteur global de l'image
 - ▶ géométrie supprimée, sauf spatial pyramid
- changement de méthode local → global

Inria



VLAD : Vector of Locally Aggregated Descriptors

- D-dimensional descriptor space (SIFT: $D=128$)
- k centroids : c_1, \dots, c_k



Aggregating local descriptor into a compact image

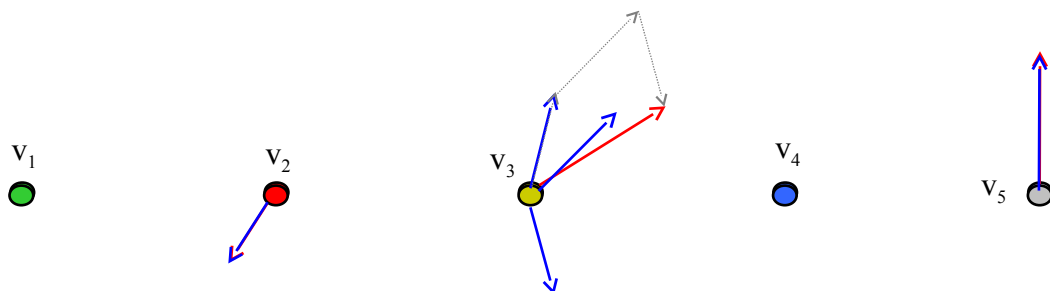
representaton, H. Jegou & al. CVPR 2010

Inria

Grenoble INP
ensimag

VLAD : Vector of Locally Aggregated Descriptors

- D-dimensional descriptor space (SIFT: $D=128$)
- k centroids : c_1, \dots, c_k



- Output: $v_1 \dots v_k =$ descriptor of size $k \cdot D$
- L2-normalized
- Typical $k = 16$ or 64 : descriptor in 2048 or 8192 D
- Similarity measure = L2 distance.

Inria

Grenoble INP
ensimag

Exercice

- calculer les descripteurs globaux BoW et VLAD

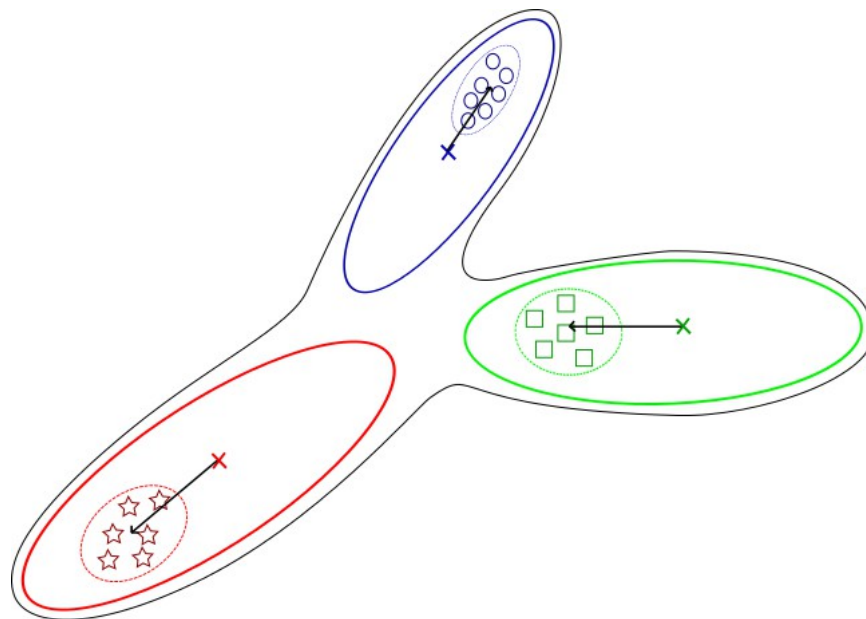
$$c_1=(1,1), c_2=(5,2), c_3=(1,4)$$

$$x_1=(3,0), x_2=(2,2), x_3=(3,4), x_4=(0,4), x_5=(0,1)$$

Inria

Grenoble INP
ensimag

Fisher vector



"Fisher kernels on visual vocabularies for image Categorization". Perronnin & Dance, CVPR 2007

Inria

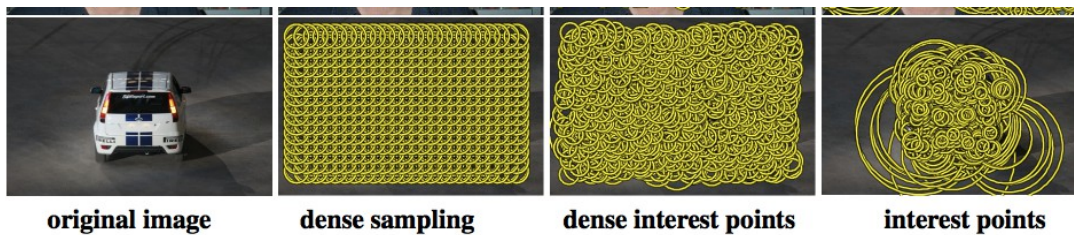
Grenoble INP
ensimag

Descripteurs denses

- Aggrégation de descripteurs
 - ▶ taille du descripteur global indépendant du nb de descripteurs locaux
 - ▶ performance croissante avec le nb de descripteurs

"Sampling strategies for bag-of-features image classification", E. Nowak, F. Jurie, B. Triggs, ECCV 2006

- détecteur aléatoire marche bien
- dense = on se passe de détecteur
- important: récupérer une information statistique sur l'image
 - ▶ ~ texture



"Dense interest points" T. Tuytelaars, CVPR 2010

Inria



Résistance du Fisher Vector à la dilution

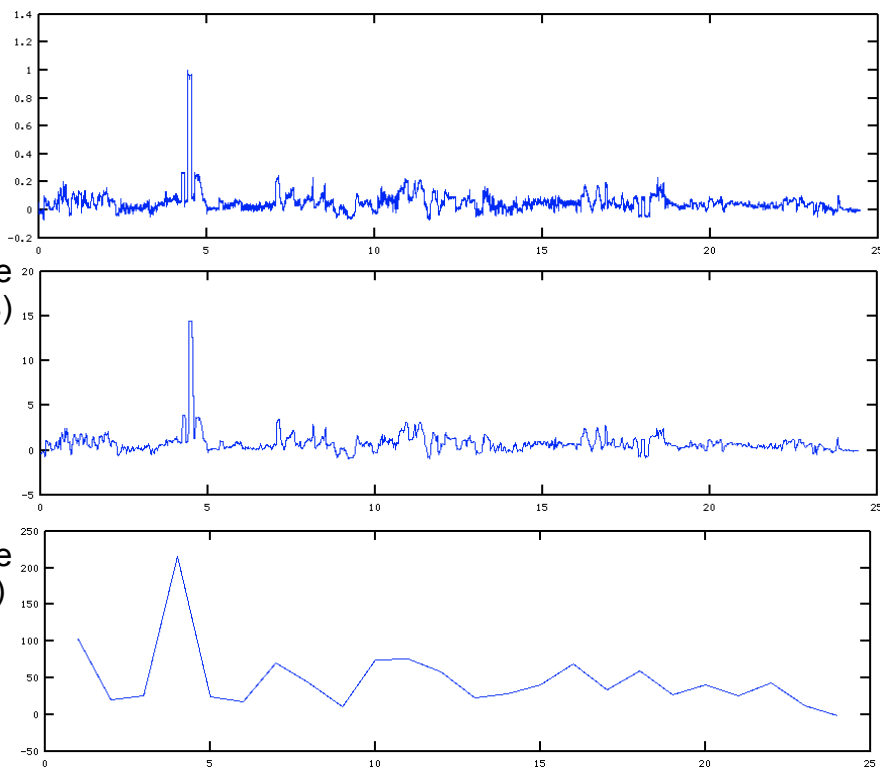
- Test sur une vidéo = suite d'images
 - ▶ +/- de cohérence temporelle
- comparaison d'une image aux autres
 - ▶ plot de corrélation sur un axe des temps
 - ▶ "pic" pour la frame en question
- puis "dilution" de la frame de la vidéo
 - ▶ en faisant la moyenne des descripteurs de Fisher sur le voisinage de l'image



Inria

Résistance du descripteur de Fisher à la dilution

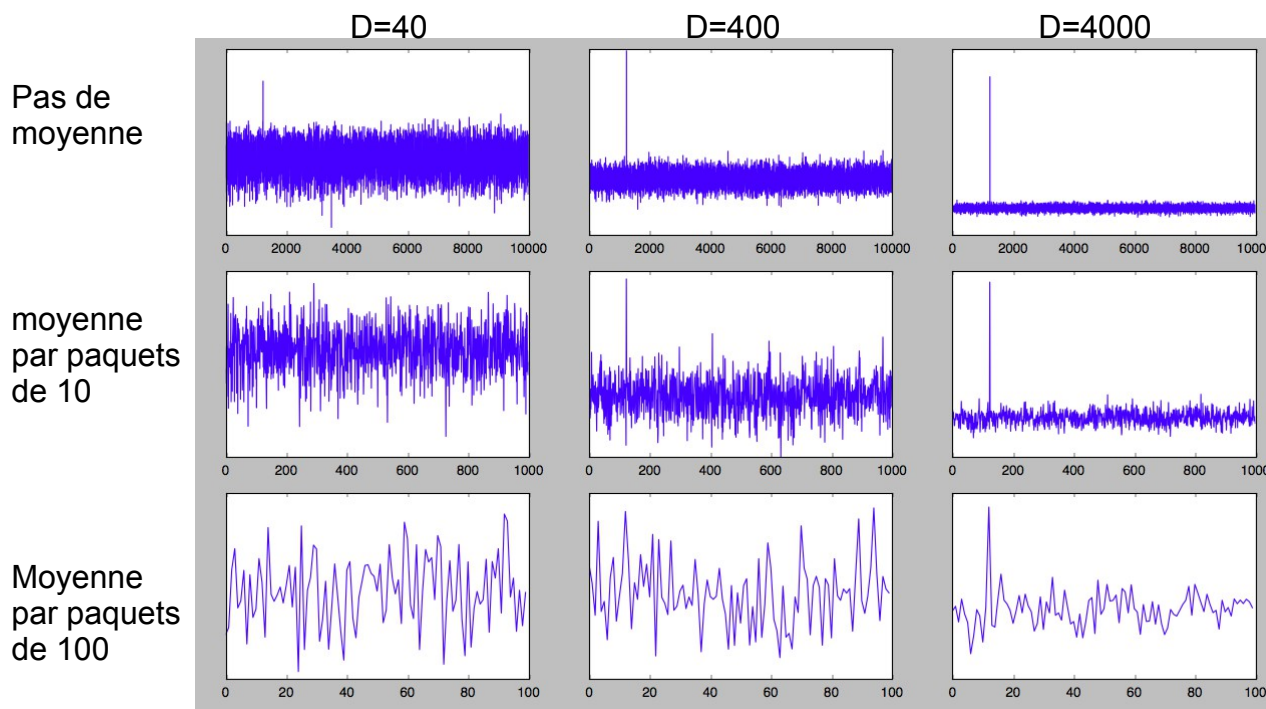
- pas de moyenne (1! fps)
- moyenne sur chaque seconde (15 images)
- moyenne sur chaque minute (900 images)



Inria

Grenoble INP
ensimag

Idem avec des descripteurs aléatoires (10000 vecteurs uniformes)

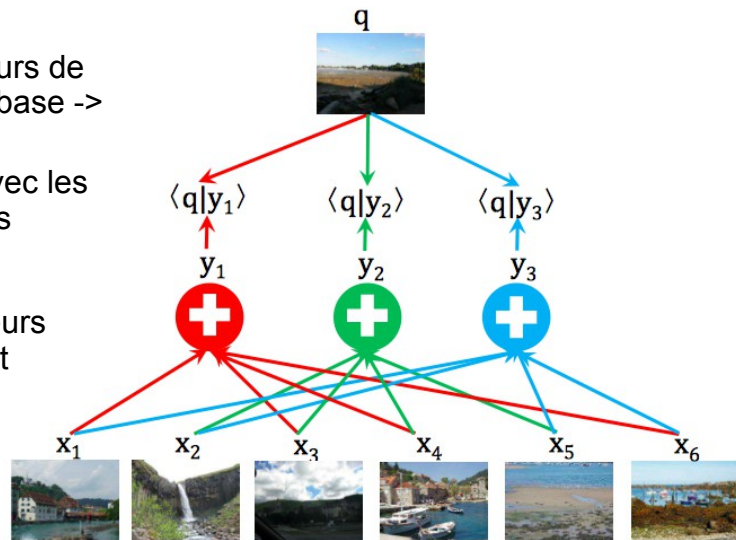


Inria

Grenoble INP
ensimag

Utilisation de la propriété de dilution pour l'indexation

- Group testing
 - ▶ Combiner les descripteurs de plusieurs images de la base -> plus compact
 - ▶ Comparer la requête avec les descripteurs de groupes
 - ▶ Raffiner...
- Marche avec des descripteurs de dimension suffisamment grande



"A Group Testing Framework for Similarity Search in High-dimensional Spaces", Miaoqing Shi, Teddy Furon, Hervé Jégou, ACM multimedia 2014

Inria

Grenoble INP
ensimag

Conclusion

- Descripteurs globaux
 - ▶ souvent version agrégée de descripteurs locaux
 - ▶ statistique sur les textures (surtout avec des descripteurs denses)
 - ▶ multi-échelle = la disposition globale des images apparaît aussi
- taille = aussi grand qu'on veut (1000 – 1M dimensions)
 - ▶ BoW épars
 - ▶ Fisher / VLAD dense

Inria

Grenoble INP
ensimag