Structuration de séquences d'images par diffusion géométrique sur graphe

Youssef Chahir,
chahir@info.unicaen.fr
GREYC - UMR CNRS 6072, Université de Caen
Campus II, BP 5186 , 14032 Caen Cedex

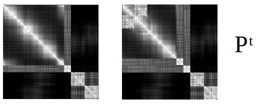
- **Objectif:** Analyse et structuration de séquences d'images
- **Méthode:** Diffusion géométrique par marches aléatoires sur graphe.
- **Résultats:** ⇒ Meilleurs résultats que d'autres approches
  - ⇒ Capturer la similarité entre les séquences
  - ⇒ Validation sur la catégorisation d'images/vidéos

Exploration du graphe par marches aléatoires locales sur graphe

Idées principales :

- Calcul des matrices de similarité / transition

Matrice de transition et ses puissances permettent d'explorer le graphe (données) à différentes échelles

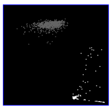


P^t

Matrice de similarité / transition

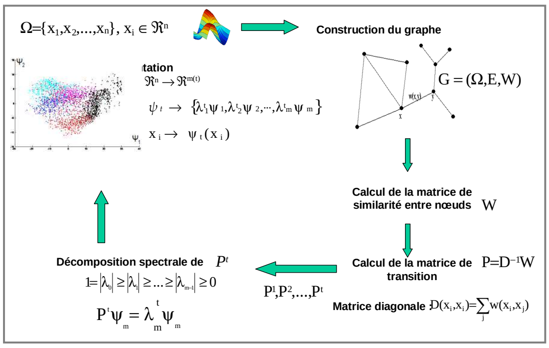
- Décomposition spectrale ⇒ espace réduit
- Clustering dans l'espace réduit

K-means clustering (K=2)



Analyse spectrale

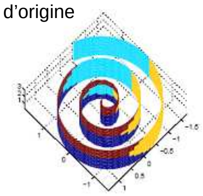
- Exploration de graphe par marches aléatoires
- Exploitation des propriétés spectrales



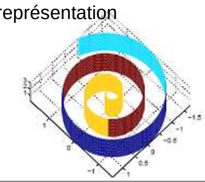
- Cadre unifié pour l'analyse de données de grande dimension
- Noyau gaussien: $w(x_i,x_j)=\exp\left(\frac{d(x_i,x_j)^2}{2\varepsilon}\right)$ ε : facteur d'échelle

Analyse spectrale capture l'organisation perceptuelle

K-Means dans l'espace d'origine



K-Means dans l'espace de représentation



Structuration de séquences vidéos

Approche

Entrée : $X=\{x_1,x_2,\ldots,x_N\}\subset R^d, t, \varepsilon, m$
Sortie : $Y=\{y_1,y_2,\ldots,y_N\}\subset R^m$

- Construction de la matrice de similarité

$$w_\varepsilon(x_i,x_j)=\exp\left(-\|x_i-x_j\|^2/\varepsilon\right)$$

- Normalisation en utilisant l'opérateur de Laplace-Beltrami

$$\tilde{w}_\varepsilon(x_i,x_j)=\frac{w_\varepsilon(x_i,x_j)}{q_\varepsilon(x_i)q_\varepsilon(x_j)}$$

- Matrice de transition

$$p(x_i,x_j)=\frac{\tilde{w}_\varepsilon(x_i,x_j)}{\sqrt{\tilde{q}_\varepsilon(x_i)\tilde{q}_\varepsilon(x_j)}}$$

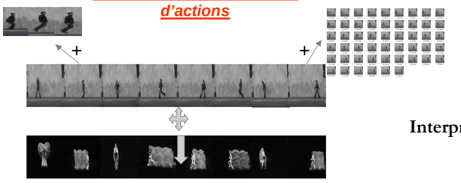
avec $q_\varepsilon(x_i)=\sum_{x_k\in X} w_\varepsilon(x_i,x_k)$

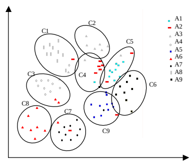
- Diagonalisation de la matrice P
- Espace de diffusion

$$x \to y=(\lambda_1^{t/2}\varphi_1(x),\cdots,\lambda_{m(}^{t/2}\varphi_{m(}(x))^T$$

- Séparation Objet (volume) / Fond
- Caractéristiques 2D/3D d'une action:
  - Moment géométrique 3D
  - Représentation de la silhouette
  - Énergie du Mouvement(MEI)
  - Historique du mouvement (MHI)
- Mesure d'affinité appropriée

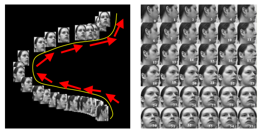
Catégorisation d'un corpus d'actions



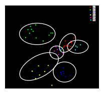


Interprétation visuelle

Réorganisation d'une base



Segmentation en Plans / Scènes



- Choix de mesures de similarité appropriées (ex.): $d_S(x,y)=\|hcum(x_i)-hcum(x_j)\|^2$
- Exploitation des informations des coordonnées : $\{\lambda_1^t\varphi_1,\lambda_2^t\varphi_2,\ldots,\lambda_i^t\varphi_i\ldots\}$
- Image représentative: valeur la plus élevée du vecteur φ_1

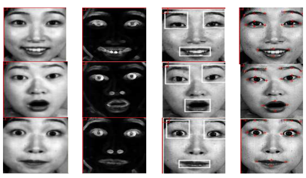
Catégorisation d'objets visuels

Démarche générale

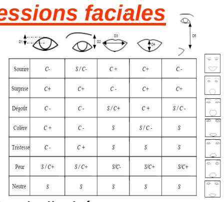
- Transformation de l'image en un graphe de Grande dimension
- Caractérisation des objets (textures) par exploration du graphe par marches aléatoires
- Classification(K-Means, Fuzzy Clustering, ..etc)
- Intérêt :
  - ⇒Approche générale s'appliquant à toute population d'objets discrets : image, nuage de points, ...
  - ⇒ Différentes mesures de similarité possibles

Composantes faciales

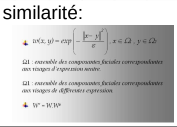
- Exploration du graphe de grande dimension
- Exploitation d'autres critères visuels: luminance/chrominance

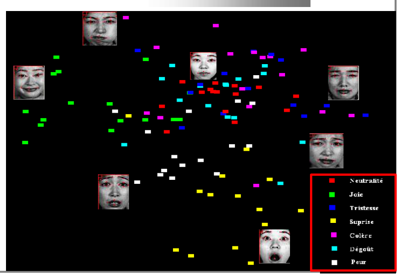


Expressions faciales



- Matrice de similarité:





Conclusion

- Utilité de la décomposition spectrale de la matrice de transition P pour la structuration et la visualisation des actions.
- Noyau gaussien
  - Stratégie pour fixer le paramètre ε: $\varepsilon=\frac{1}{N}\sum_{i=1}^{N}\min\left\{\|x_i-x_j\|^2 / \|x_i-x_j\|^2>0, j=1,2,\ldots N\right\}$
- Contexte semi- supervisé
- Réduction et fusion de données multimédias hétérogènes
  - Exploitation de signaux, audio, …
- Perspectives:
  - ⇒Fusion de données