

Représentation objective et pertinente des expressions faciales émotionnelles

Application à la reconnaissance et la synthèse

Nicolas STOIBER.

Doctorant 3^o année Supélec / Orange Labs

- Focus: Expressions faciales émotionnelles

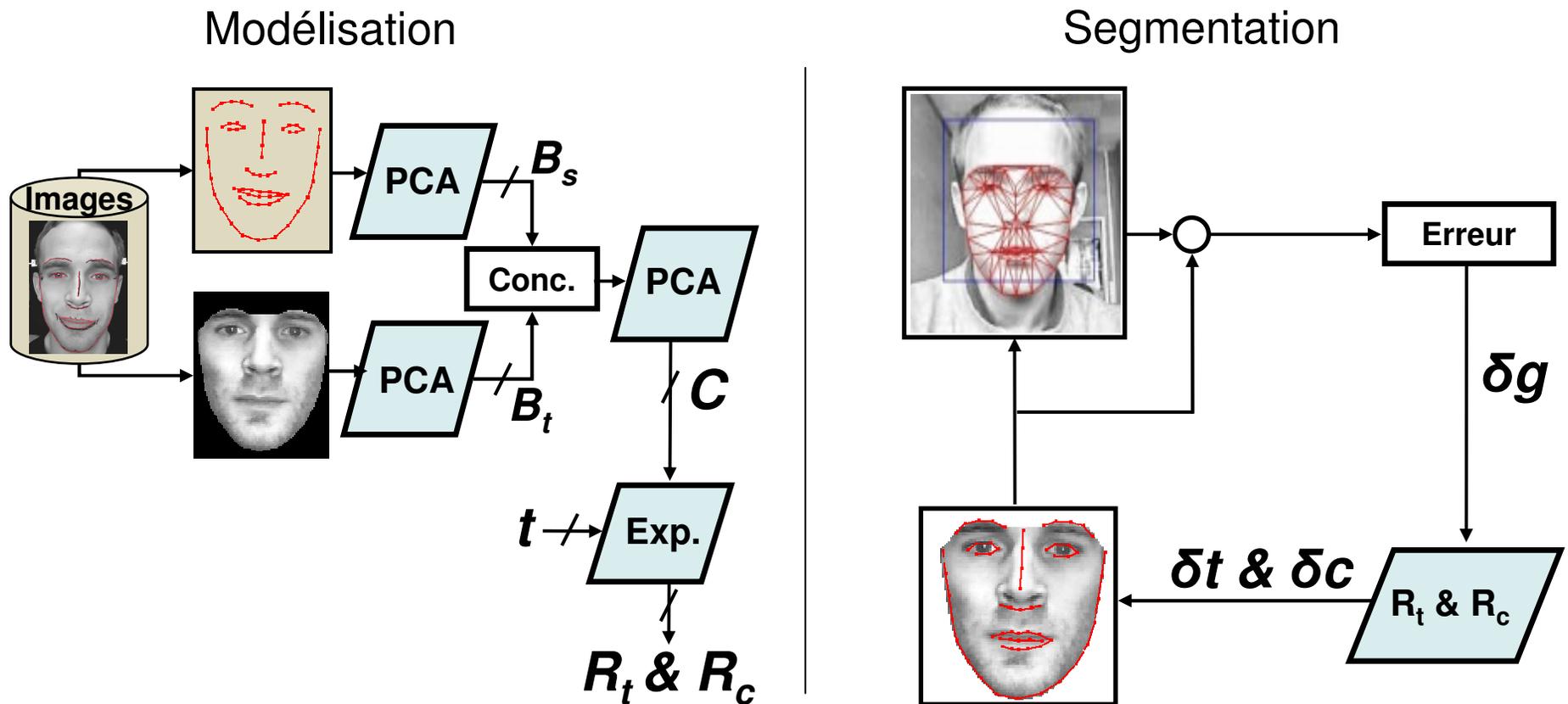


- Communication non-verbale.
- Meilleure intelligibilité, enrichissement de l'information transmise.
- Meilleure interactivité, échange naturel.

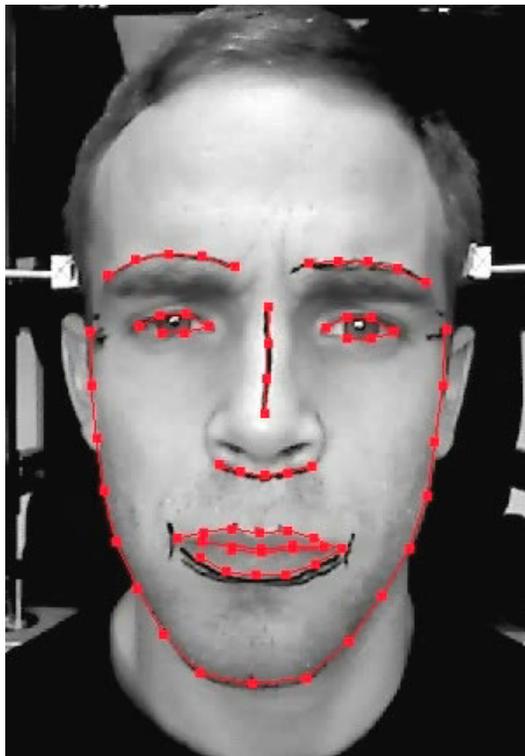
- Objectif:

- Acquérir, comprendre, manipuler l'information expressive des mouvements faciaux.
- Analyse/Synthèse

Modèles Actifs d'Apparence

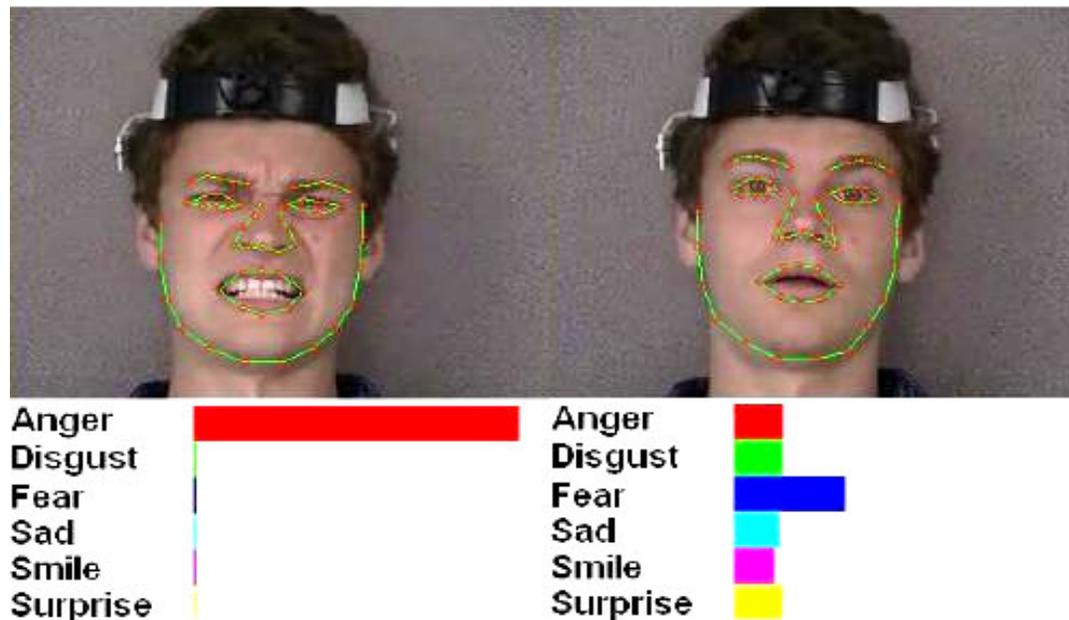


- AAM appliqués aux déformations faciales
 - Base d'apprentissage mono-personne.
 - Pas de mouvements rigides de la tête.



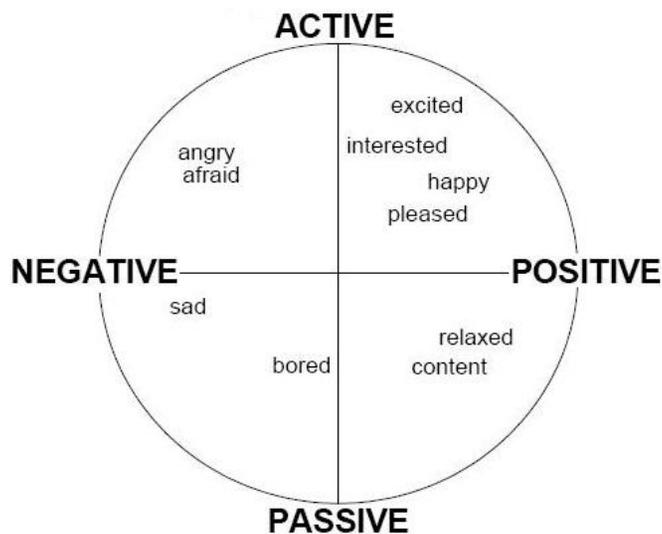
Interpétation ?
Manipulation ?

- Représentations catégorielles
 - Classes **discrètes**
 - Classes émotionnelles (Ekman)

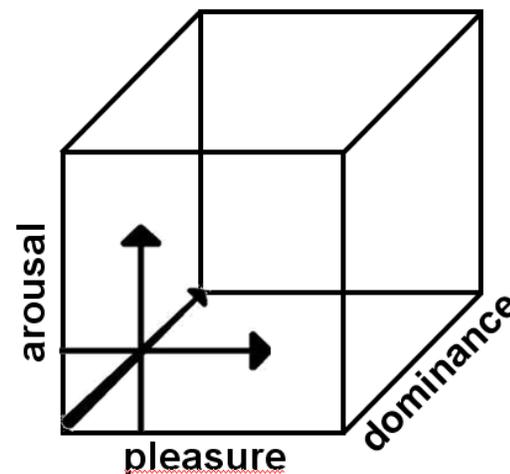


[Hu 2004, Shan 2005, Suzuki 2007]

- Représentations dimensionnelles
 - Espaces **continus**.
 - Considérations psychologiques.



[Cowie 2000]



[Mehrabian 1996]



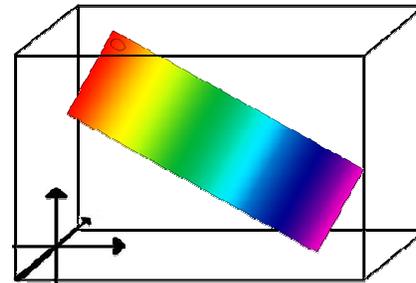
[Plutchik 2001]

- Représentations basées données

- Déformations faciales **non-linéaires** [Chang 03, ElGammal 05, Lee 06]

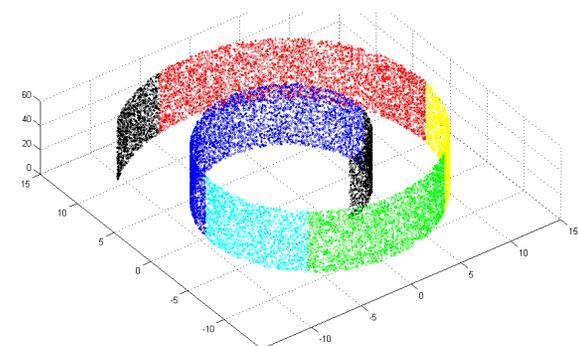
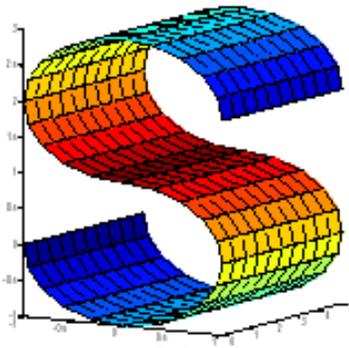
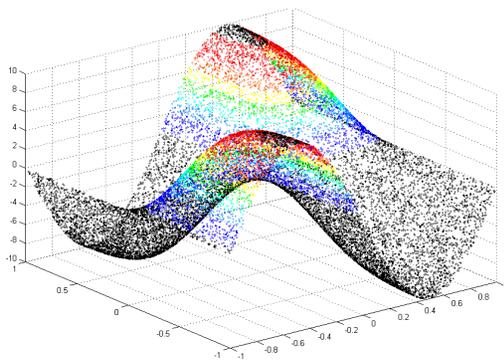
- **Réduction linéaire**

- PCA, ICA, ...



- **Réduction non-linéaire**

- Variétés géométriques



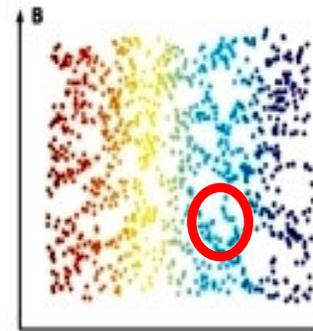
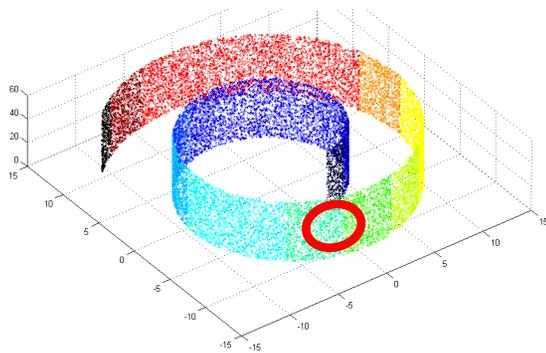


Expressions Faciales: Représentation

- Représentation basées données.
 - Réduction **non-linéaire**, **non-supervisée**.
 - Représentation **continue**, **objective**.

- Représentation basées données.
 - Réduction **non-linéaire**, **non-supervisée**.
 - Représentation **continue**, **objective**.

Locally Linear Embedding (LLE)

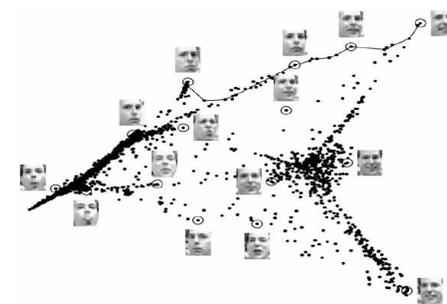


Préservation linéaire
des voisinages

$$\text{Voisinages: } \mathcal{E}(\mathbf{W}) = \sum_i \left\| \vec{X}_i - \sum_j W_{ij} \vec{X}_j \right\|^2$$

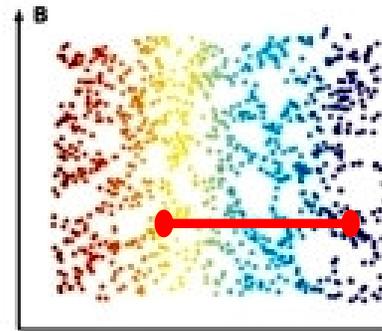
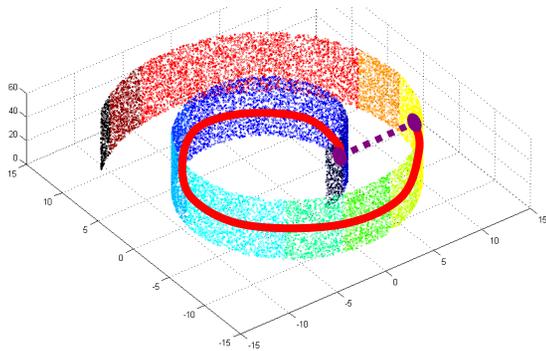
$$\text{Réduction: } \Phi(\mathbf{Y}) = \sum_i \left\| \vec{Y}_i - \sum_j W_{ij} \vec{Y}_j \right\|^2$$

Exemple:



- Représentation basées données.
 - Réduction **non-linéaire**, **non-supervisée**.
 - Représentation **continue**, **objective**.

Isometric Map (ISOMAP)



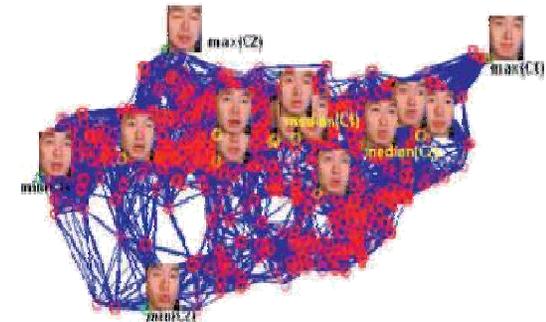
Préservation des
distances
géodésiques

Distances: $(\mathbf{D}_G)_{i,j} = \text{geod}(\vec{X}_i, \vec{X}_j)$

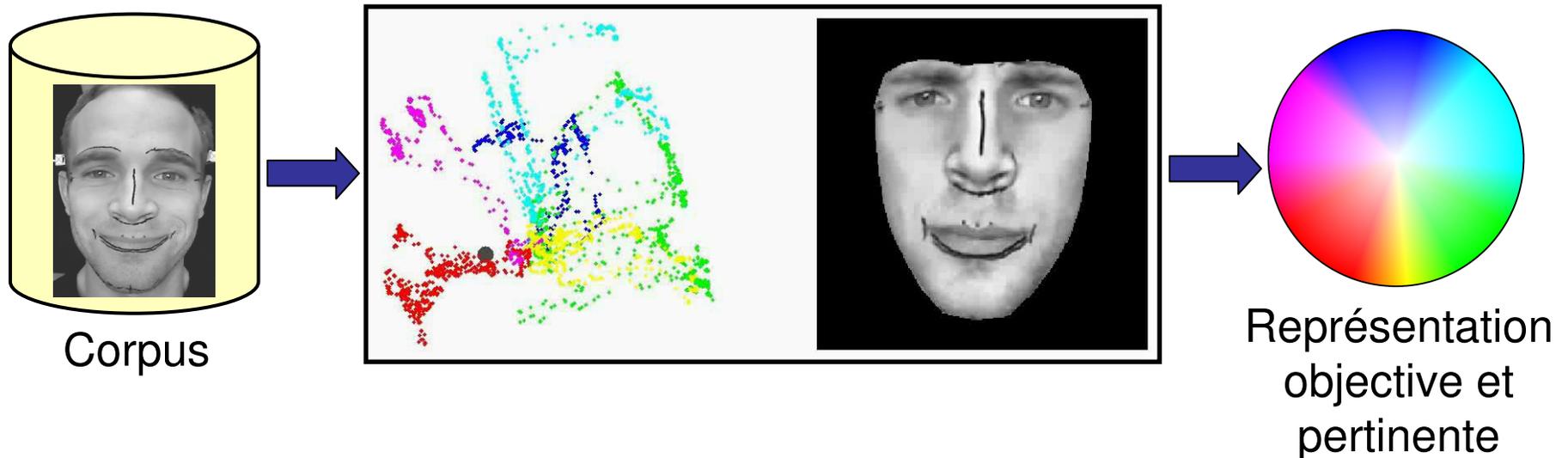
$(\mathbf{D}_Y)_{i,j} = \|\vec{Y}_i - \vec{Y}_j\|^2$

Réduction: $\Phi(\mathbf{Y}) = \|\tau(\mathbf{D}_G) - \tau(\mathbf{D}_Y)\|^2$

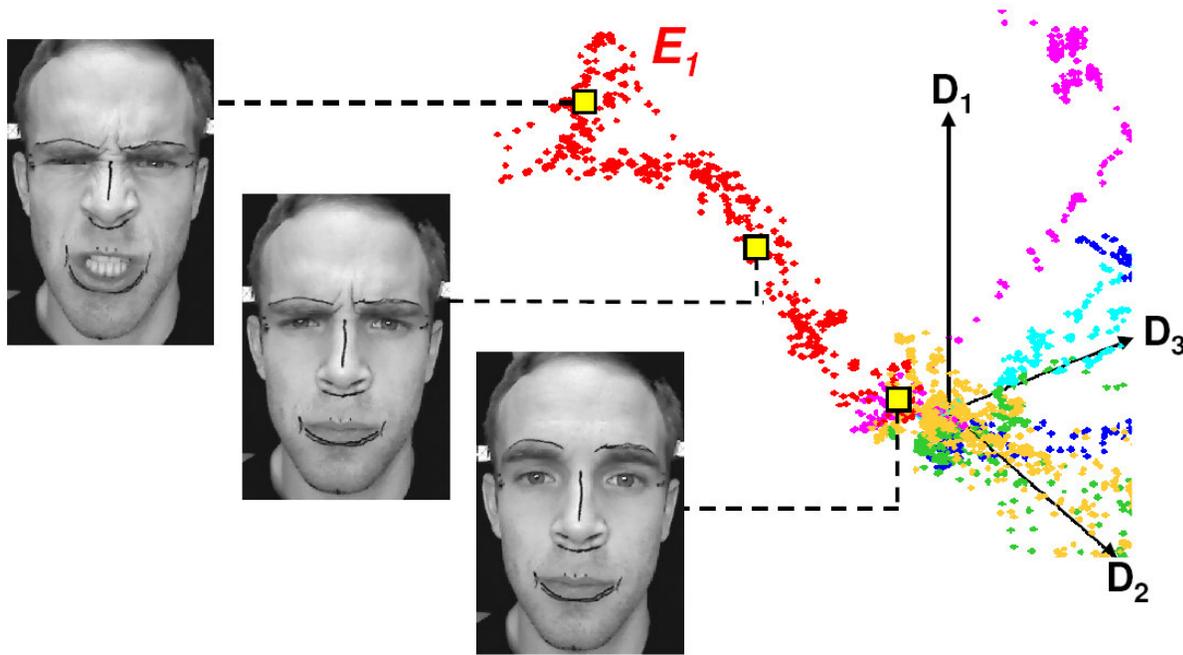
Exemple:



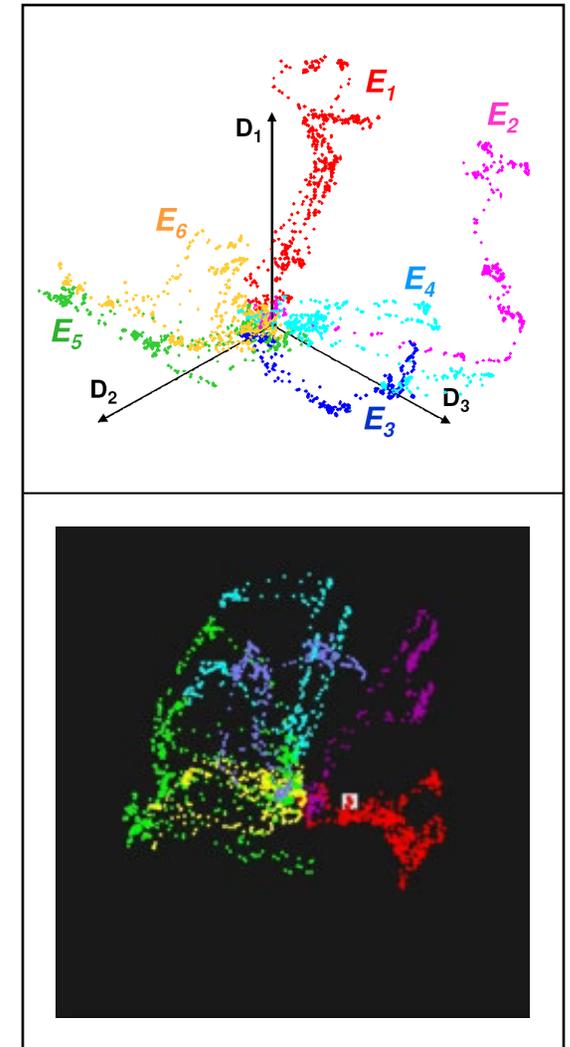
- Notre approche
 - Représentation **basée données**.
 - Réduction **non-linéaire, non-supervisée**.
 - **Adaptée à la topologie** des données faciales expressives.



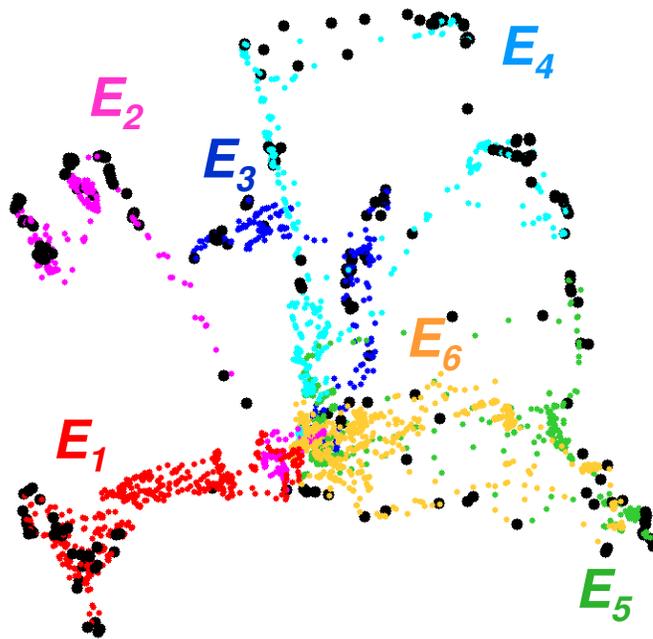
- Nature des données



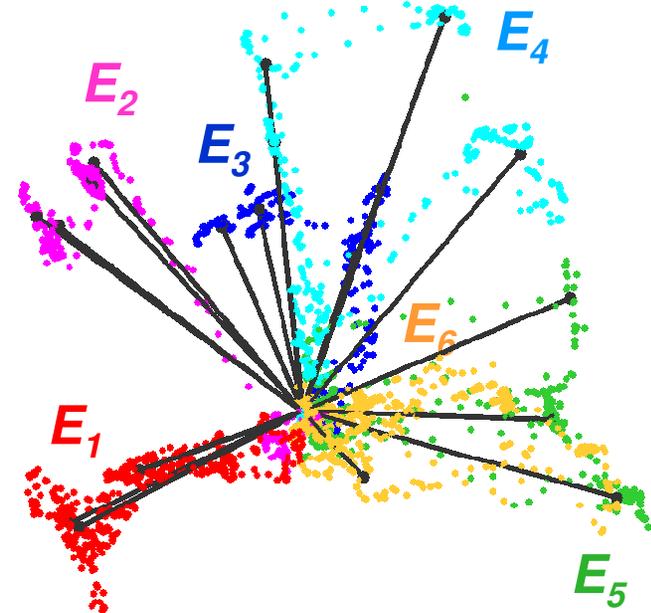
Présence de **directions dominantes**
dans les données



- Détection des directions dominantes
 - Utilisation de l'enveloppe convexe des données

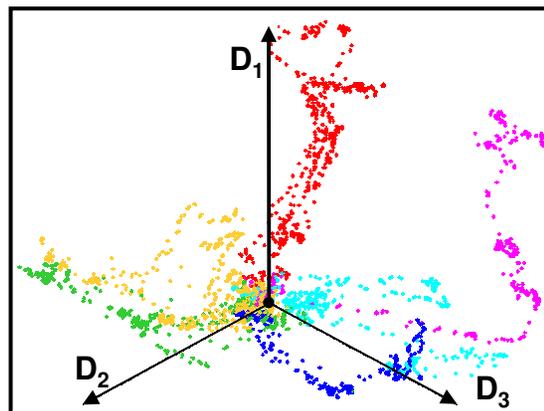


Détection de l'enveloppe
(algorithme Qhull)

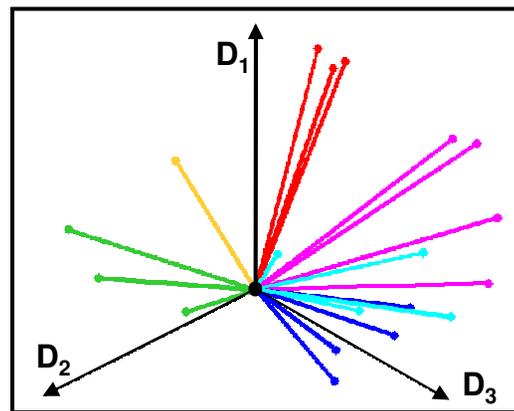


Elagage des candidats
(algorithme Mean Shift)

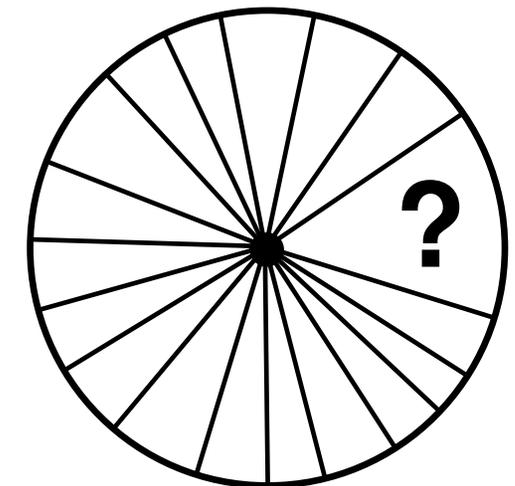
- Réduction de la variété non-linéaire
 - Utilisation des directions dominantes



Données de déformations faciales



Directions dominantes



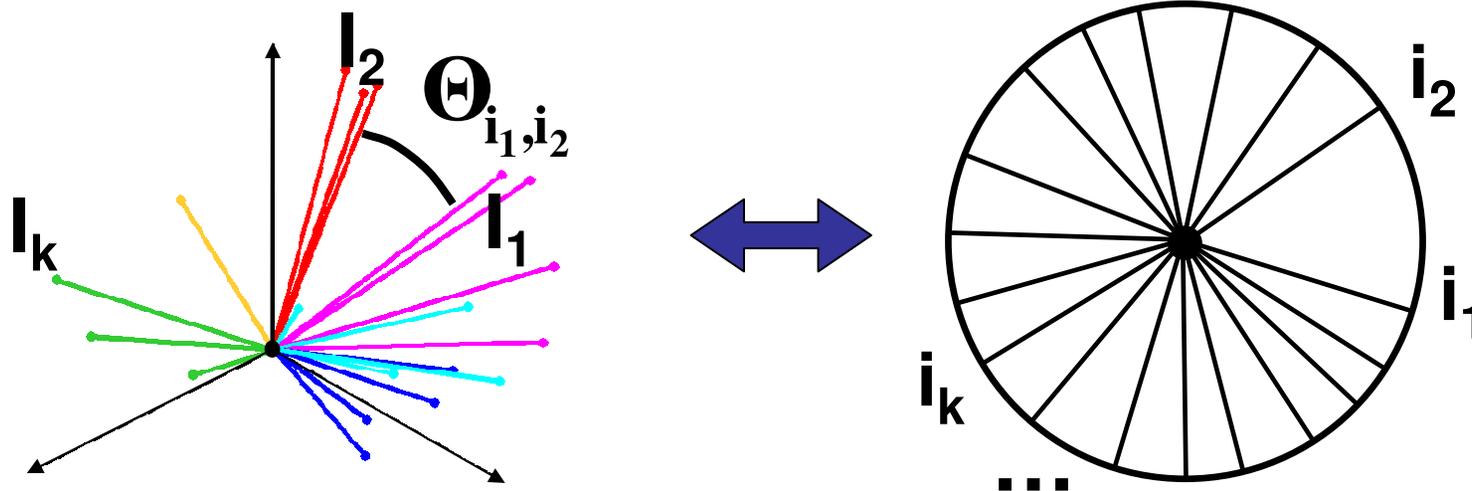
Représentation réduite

- Problème d'optimisation

Critère:

Minimiser la distorsion entre les distributions des directions \mathbf{l}_k et \mathbf{i}_k

Mesure de la distorsion: Angles

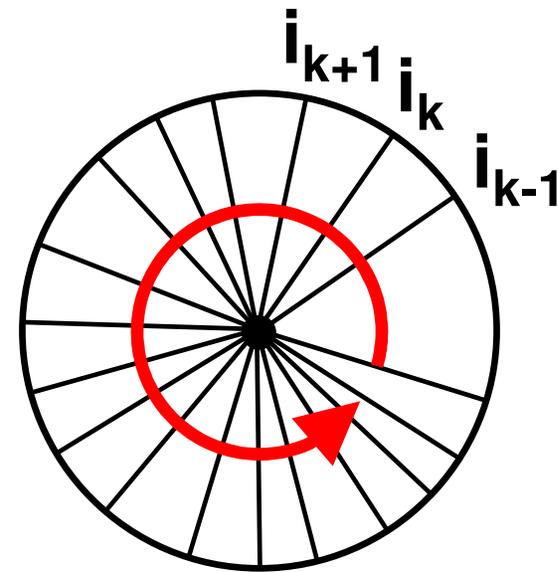


- Problème d'optimisation

Reduction 2D:

Les directions réduites forment
une **suite**.

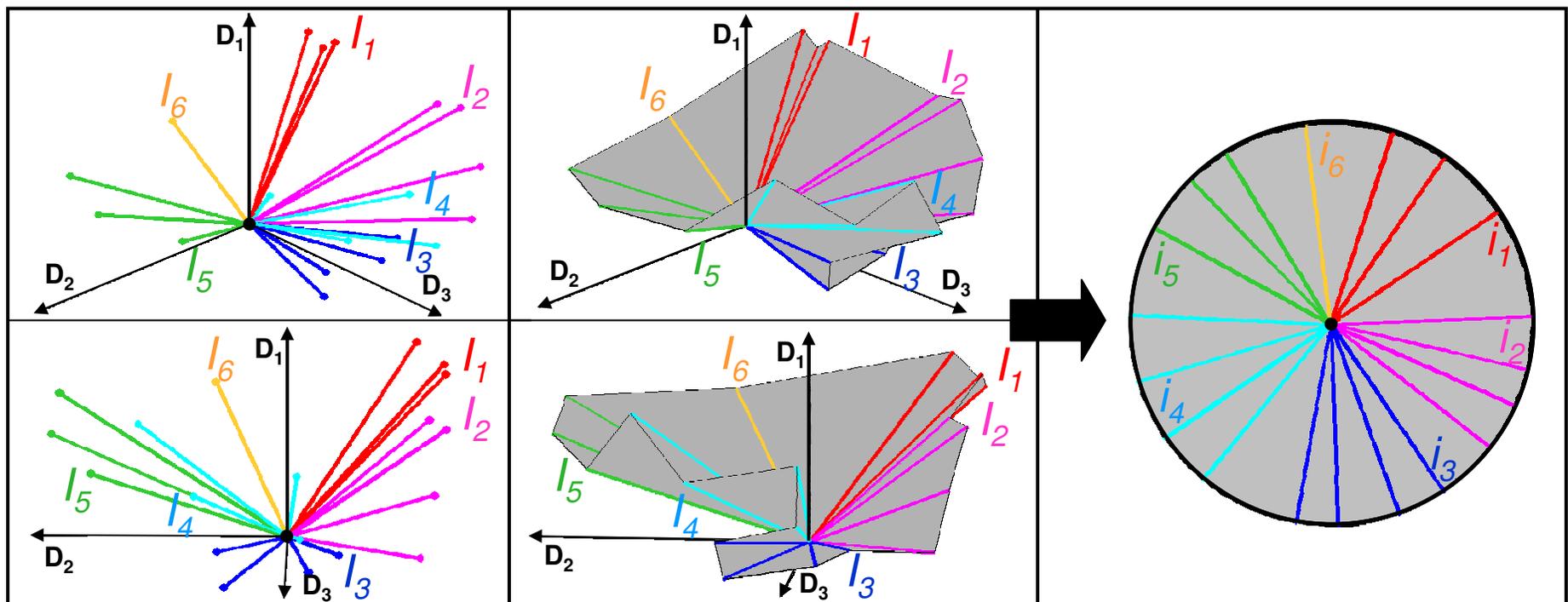
2 voisins uniquement.



**Problème du voyageur
de commerce!**

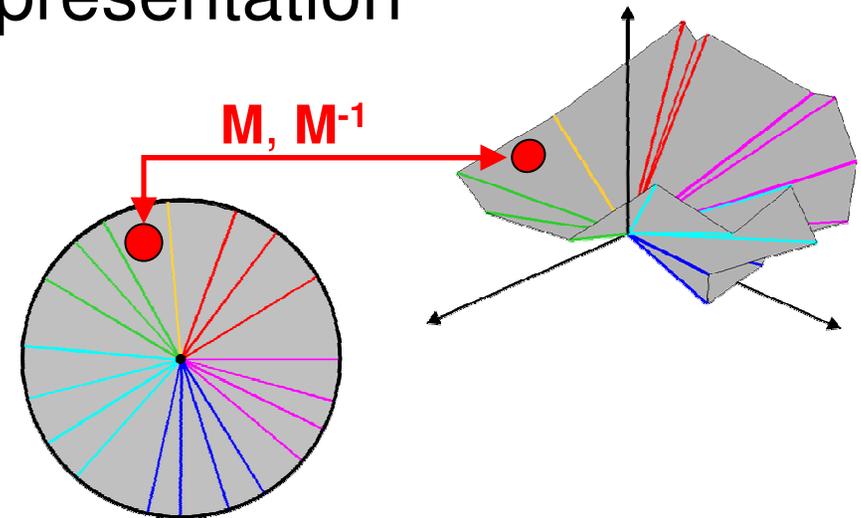
$$\arg \min_{(i_1, \dots, i_{n_d})} \left(\sum_{k=1}^{n_d} \Theta_{i_k, i_{k+1}} \right)$$

- Résultat de l'optimisation
 - Recuit simulé
 - Algorithme génétique
 - ...

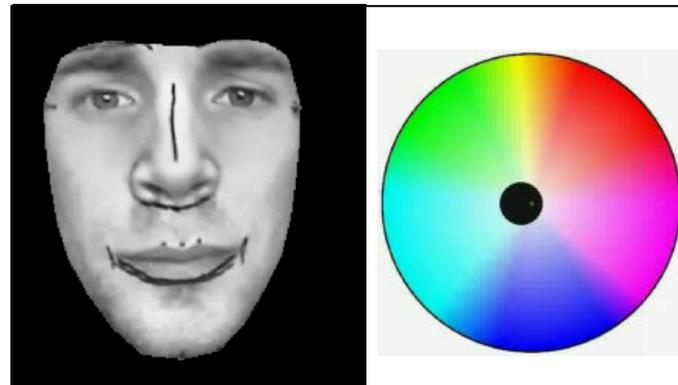
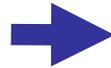


- Lien analytique données/représentation
Regression

- Linéaire par morceau
- Thin-Plate Splines
- Gaussian Process Regression

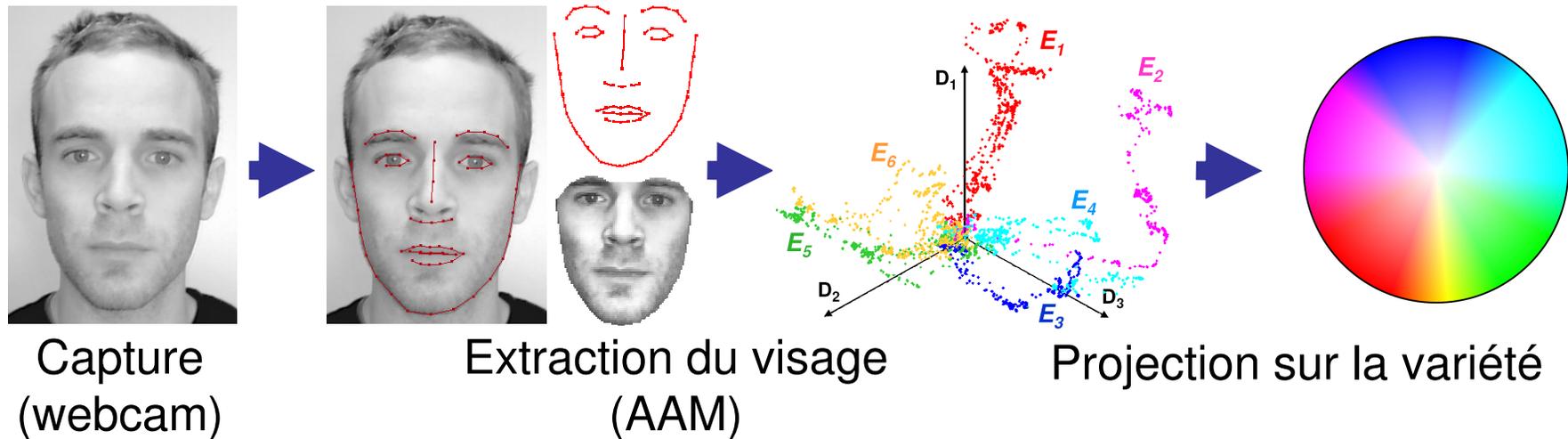


Parcours de
la variété

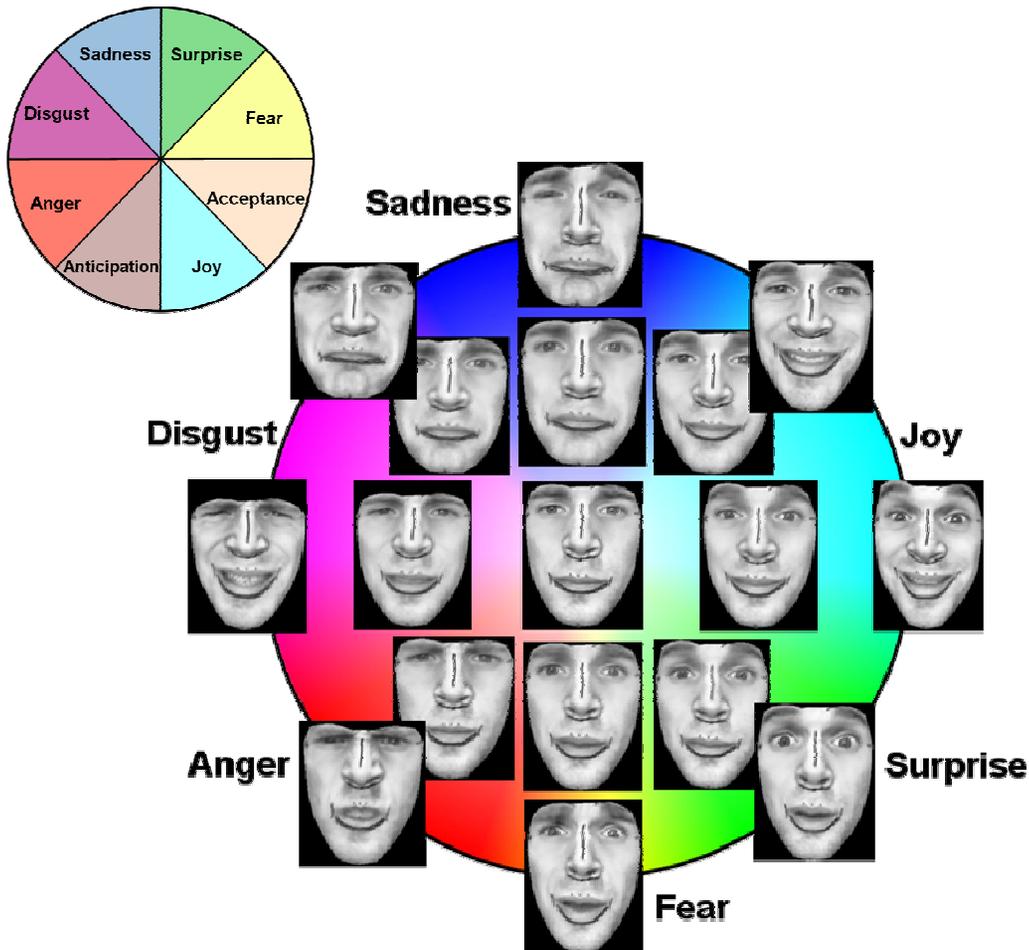


Applications

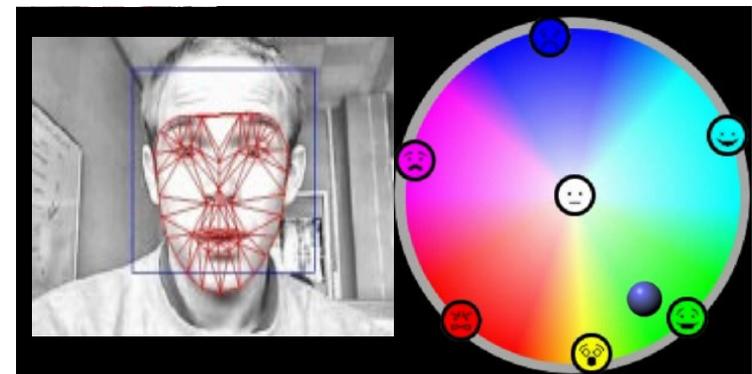
- Reconnaissance d'expressions faciales



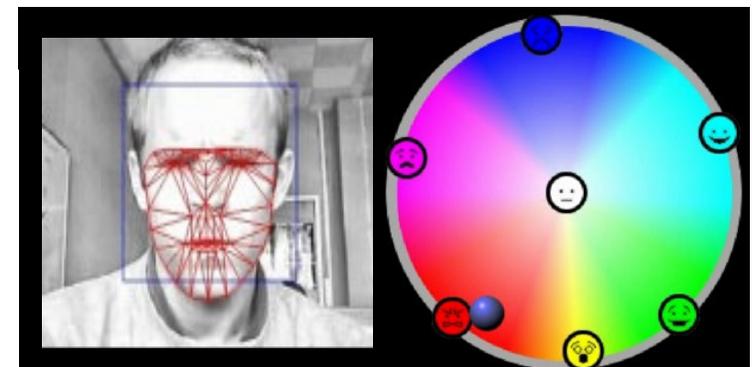
- Reconnaissance d'expressions faciales



Reconnaissance temps réel

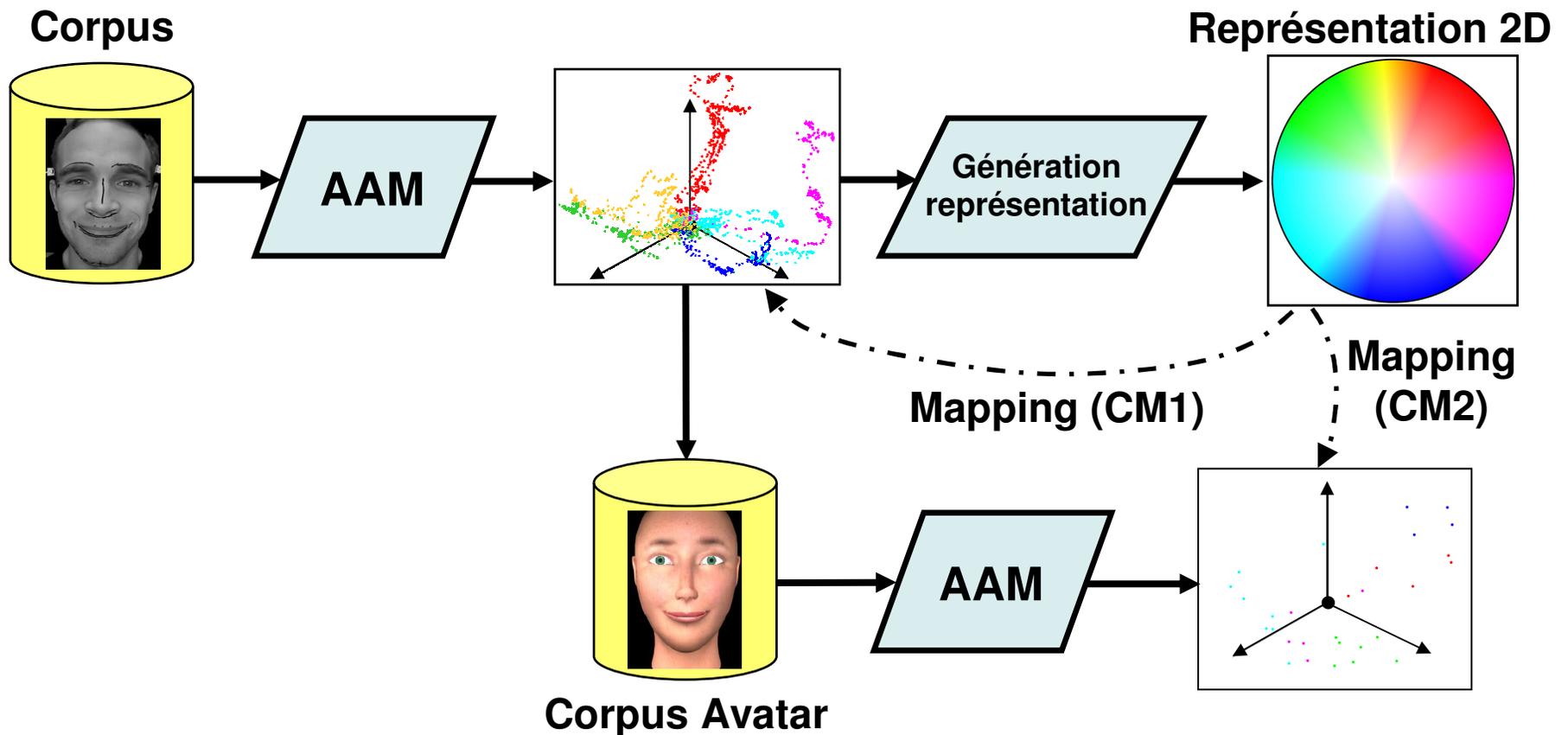


surprise

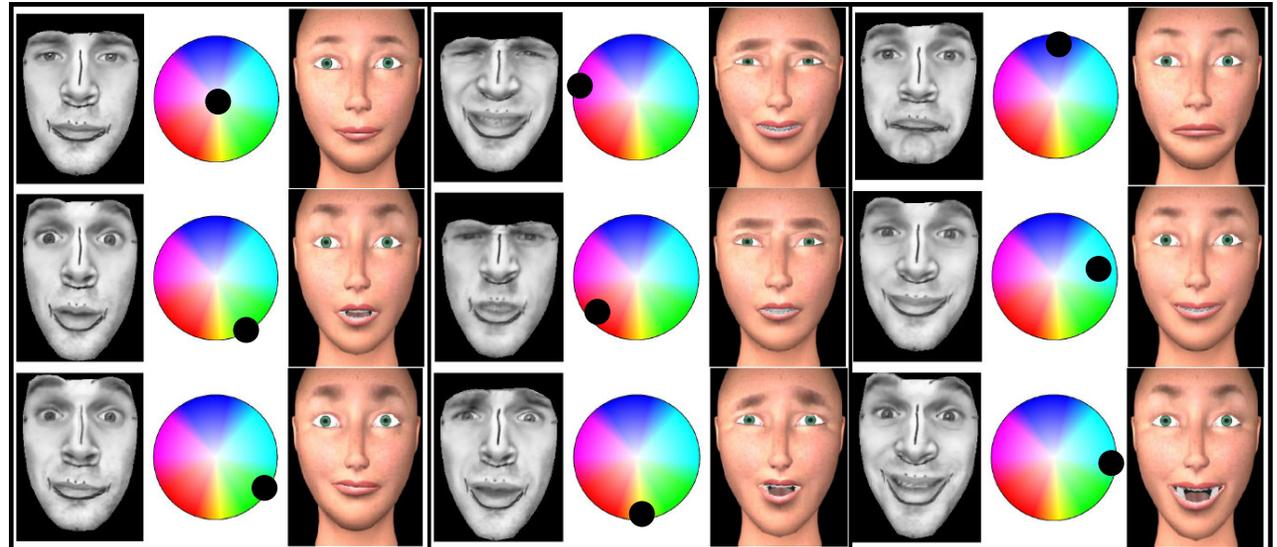


colère

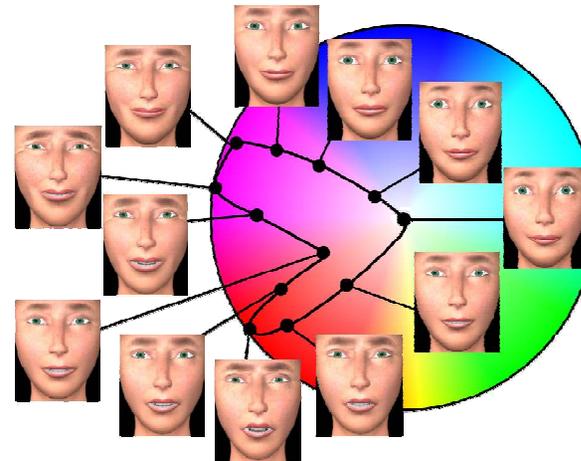
- Synthèse d'expressions faciales



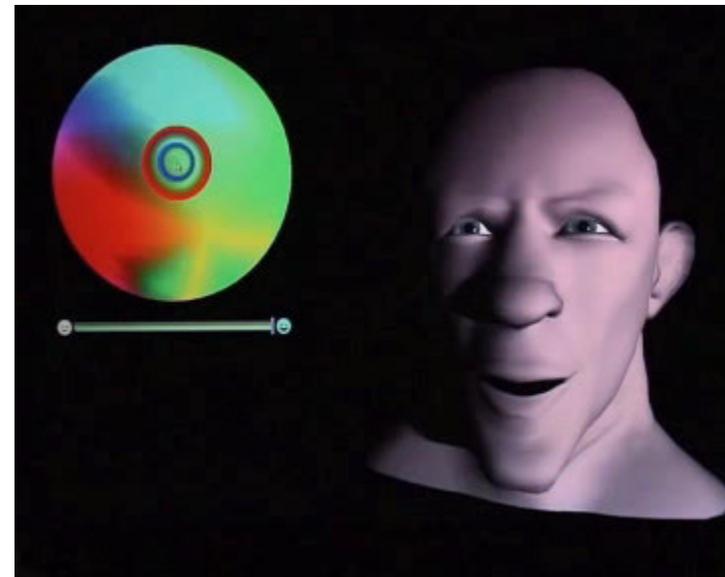
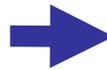
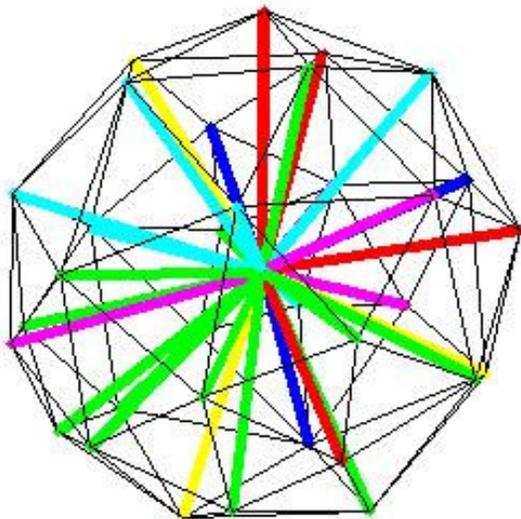
Synthèse
d'expressions fixes



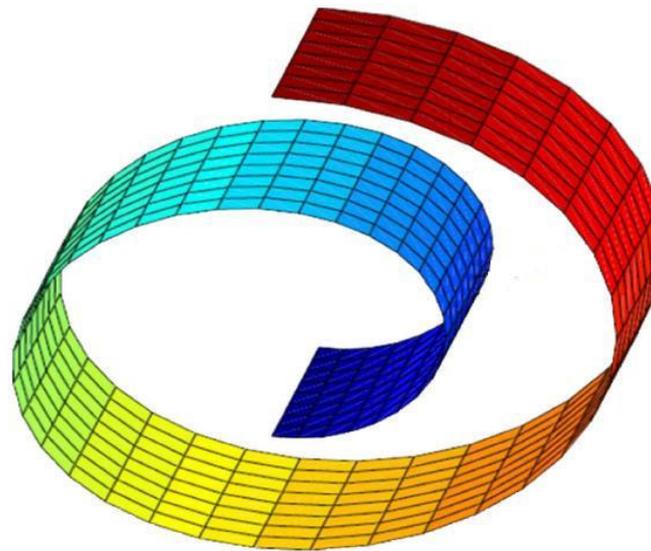
Synthèse de
séquences d'animation



- Représentations de plus grande dimensionnalité
 - Même principe qu'en 2D
 - L'optimisation devient un **TSP d'ordre supérieur**
 - *3D : Volume de surface minimale*
 - Compromis **richesse/simplicité**



- *Reconnaissance*: **généricité** des visages
 - Pour l'instant: mono-personne.
 - Problème: différences morphologiques.
- **Aspect dynamique**
 - Signature dynamique dans les expressions.
 - Peut être exploitée en *reconnaissance/synthèse*.
- **Autres types de mouvements faciaux**
 - Visemes



Merci de votre attention