Moyennes et statistiques de formes

Guillaume Charpiat

28 juin 2005

Équipe Odyssée

ÉCOLE NORMALE SUPÉRIEURE

ÉCOLE NATIONALE DES PONTS ET CHAUSSÉES

INRIA

Pourquoi des moyennes et des statistiques de formes ?

$\hookrightarrow$ pour trouver plus facilement un objet dans une image : segmentation a priori

$\hookrightarrow$ pour classifier des ensembles de formes.

Segmentation avec a priori

- $\hookrightarrow$ donnée : une image A
- $\hookrightarrow$ but : trouver le contour d'un objet dans cette image
- $\hookrightarrow$ moyen : faire évoluer un contour C de manière à minimiser une certaine énergie $E(C)$
- $\hookrightarrow$ énergie $E(C)$: basée sur des descripteurs de l'image (gradient de l'intensité, cohérence de la structure à l'intérieur du contour)
- $\hookrightarrow$ on voudrait en plus : avoir des indications sur la forme à rechercher
- $\hookrightarrow$ moyen : ajouter à l'énergie un terme sur la forme
- $\hookrightarrow$ nouveau but : exprimer la probabilité qu'une courbe quelconque C appartienne à l'ensemble de courbes données en apprentissage.

Exemple de résultat

On dispose de n exemples d'images d'un même objet, ou d'une même catégorie d'objets, déjà segmentées.

Nouvelle image à segmenter :

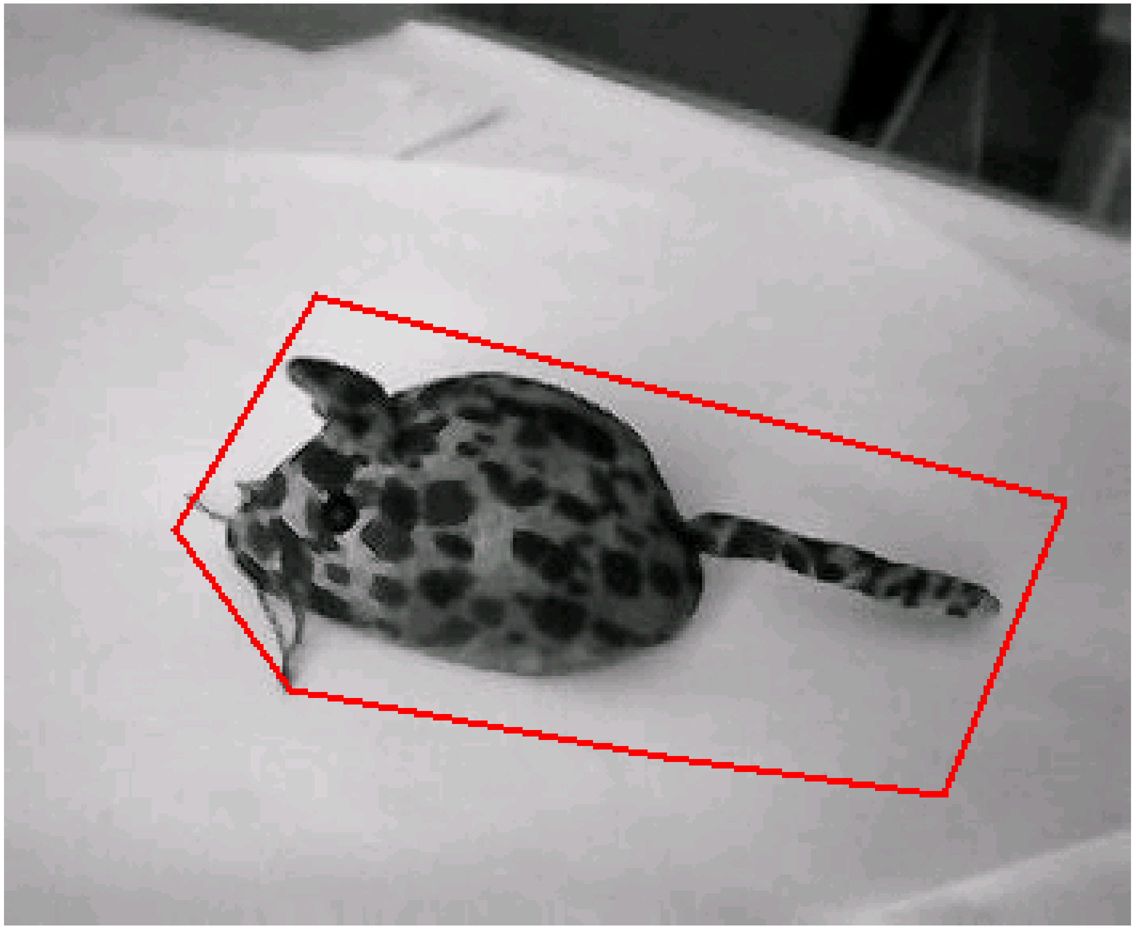


Image segmentée sans a priori avec un algorithme ne considérant que des histogrammes de région :

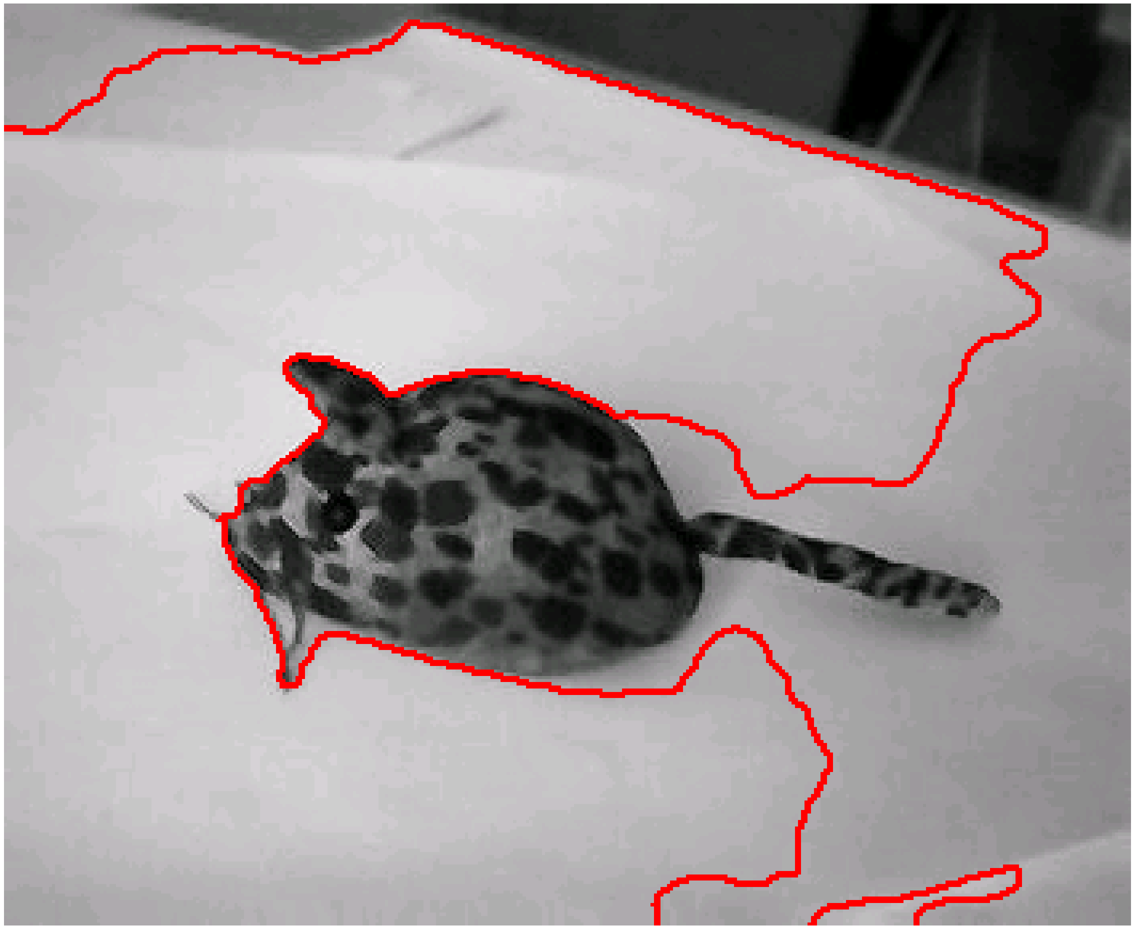
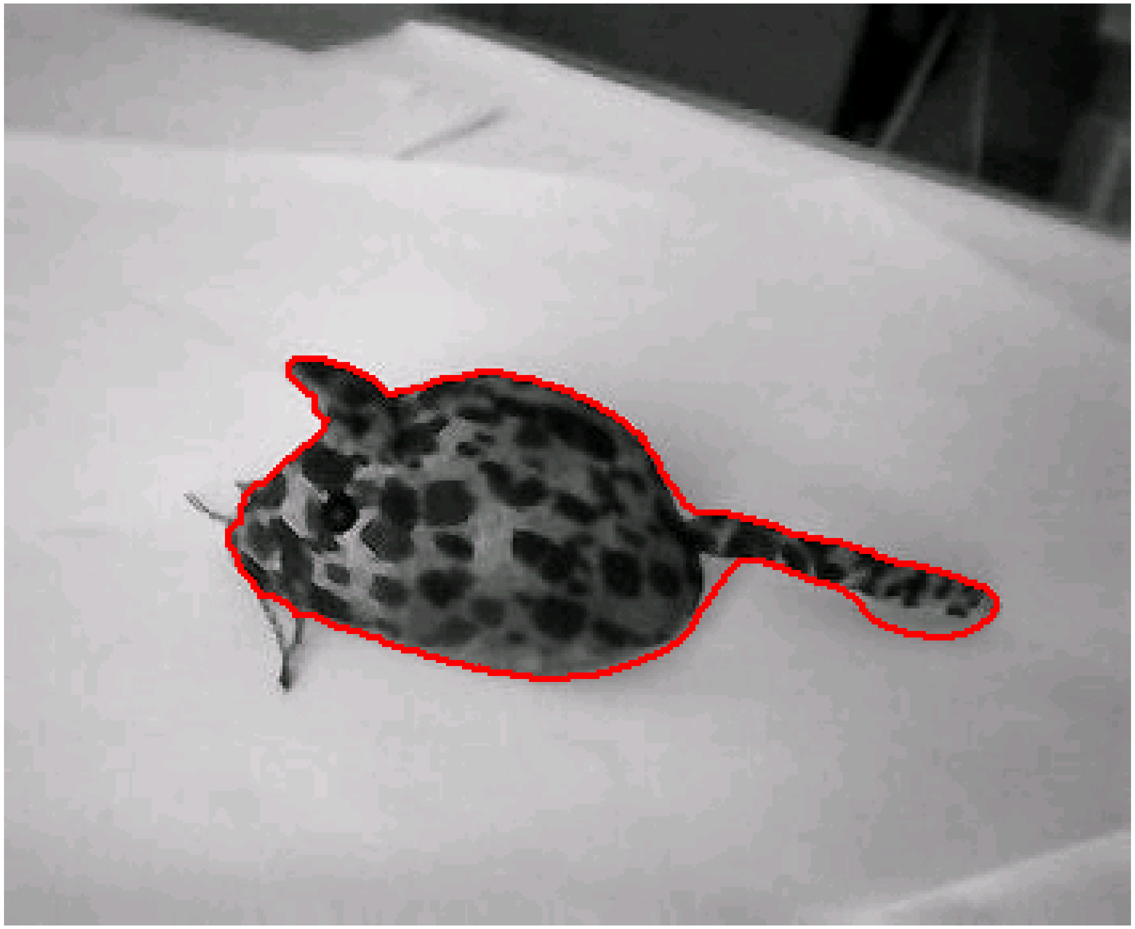


Image segmentée avec le même algorithme, mais avec un a priori sur la forme :



Exemple bruité

Image à segmenter :

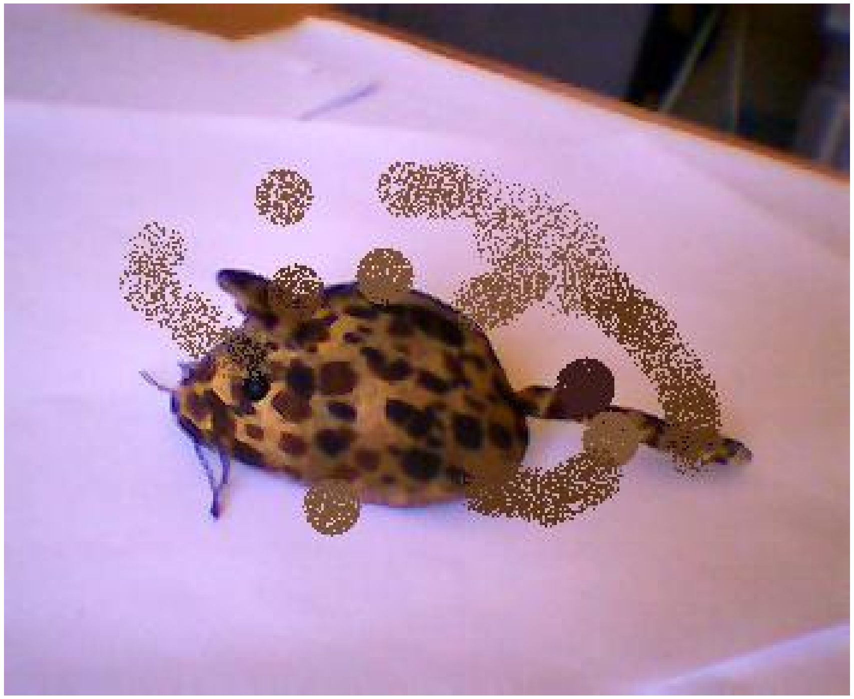
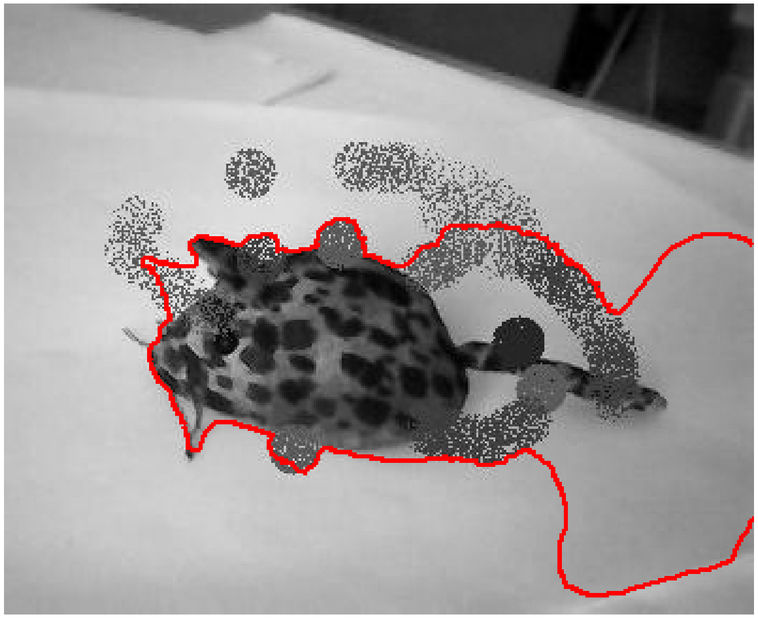


Image segmentée sans a priori :



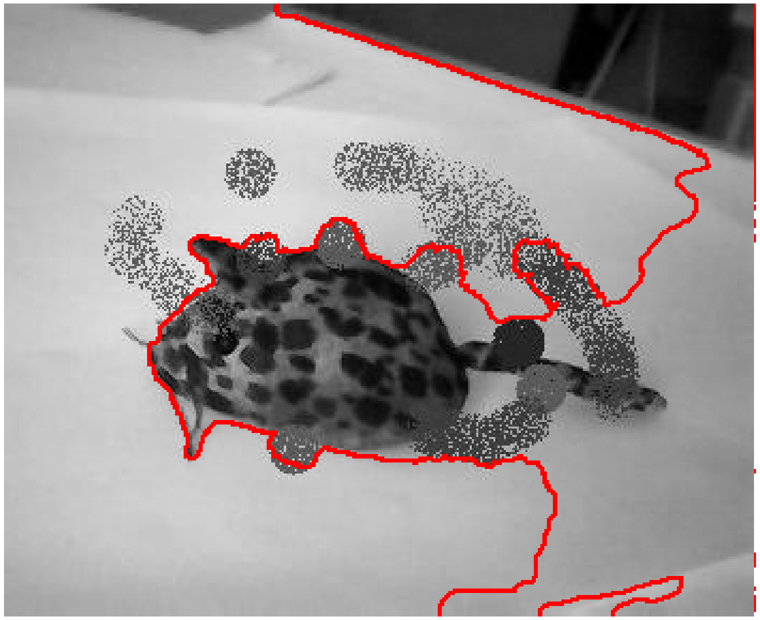
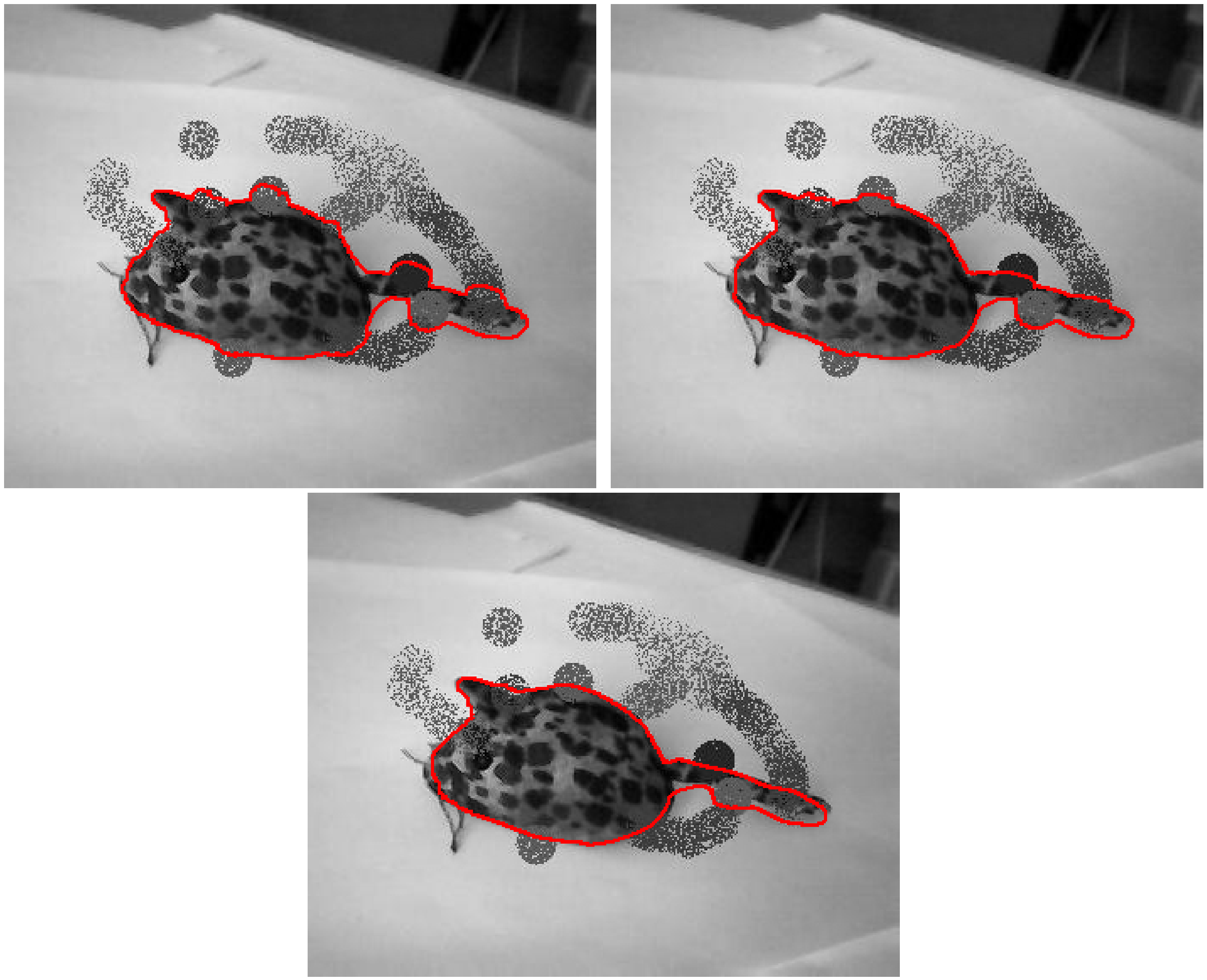


Image segmentée avec a priori (avec une importance croissante du terme sur la forme) :



Retour à la théorie

Quel terme choisir pour exprimer une contrainte sur la forme ?

$\hookrightarrow$ données $\mathcal{D} = \{C_i\}$: ensemble de courbes C_i déjà segmentées dans d'autres images

$\hookrightarrow$ variable à ajuster : la forme courante C, qui évolue

$\hookrightarrow$ critère : probabilité d'appartenance de C à $\mathcal{D}$,
degré de ressemblance de C aux échantillons C_i.

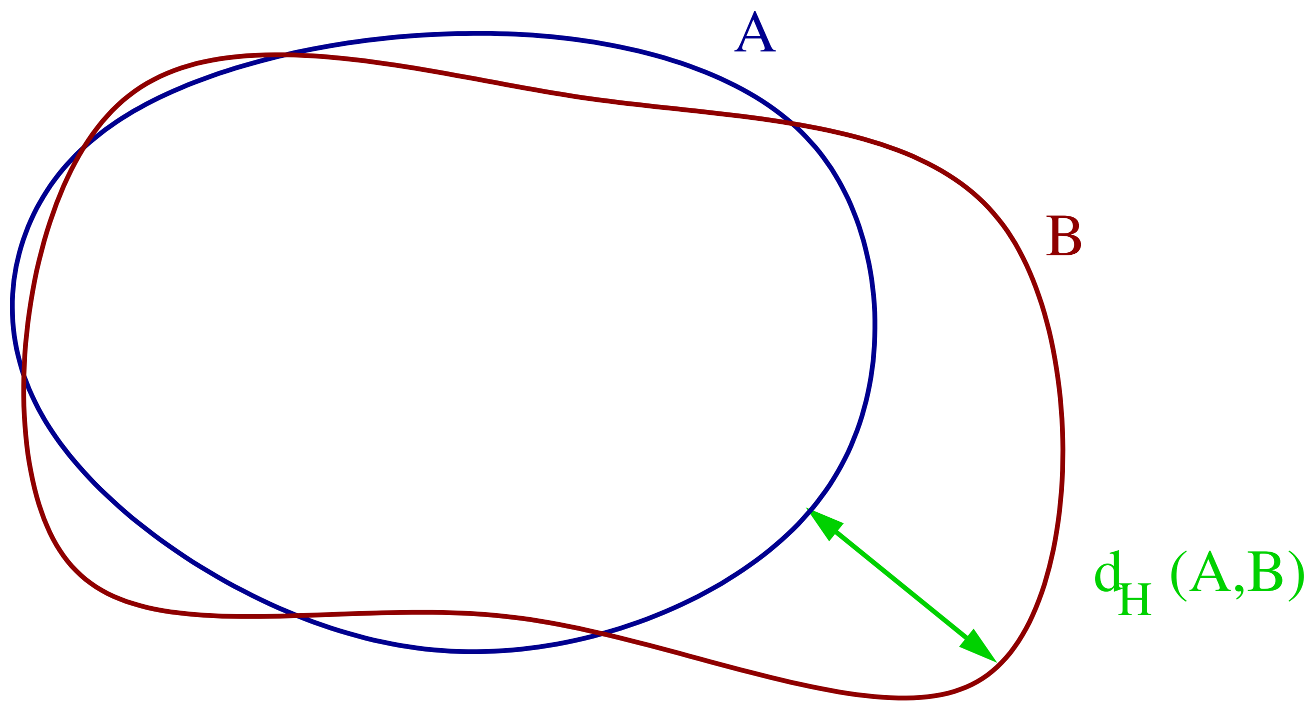
$\hookrightarrow$ statistiques sur les courbes C_i, forme moyenne, formes caractéristiques ?

Moyenne de courbes planes

Nous disposons d'un ensemble de n courbes, dont on voudrait calculer la moyenne.

$\hookrightarrow$ Qu'est-ce que la moyenne de plusieurs courbes ?
C'est la courbe qui ressemble le plus à toutes les autres à la fois.

$\hookrightarrow$ Comment exprimer la ressemblance entre deux courbes ?
Par un critère qui à deux courbes quelconques associe leur distance mutuelle, par exemple la distance de Hausdorff.

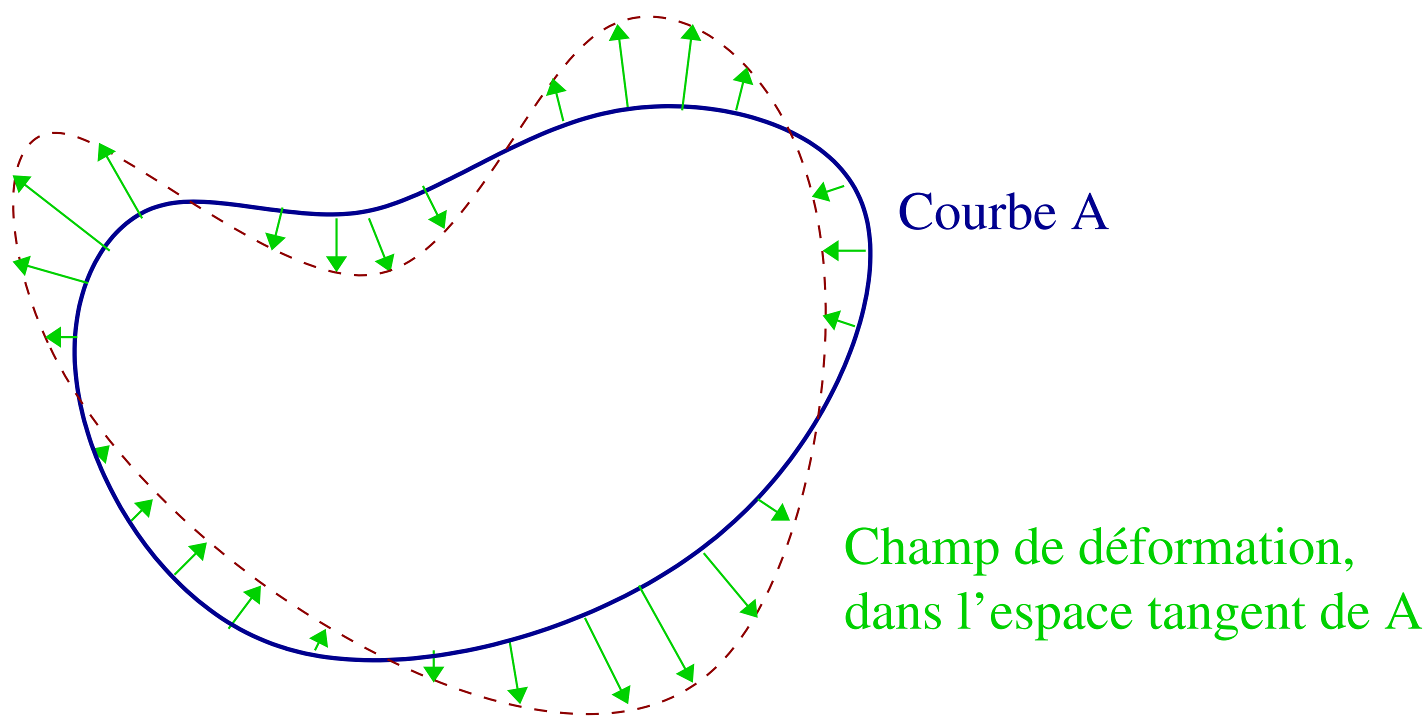


$$d_H(A,B) = \sup_{x\in A} \inf_{y\in B} d(x,y) + \sup_{x\in B} \inf_{y\in A} d(x,y)$$

Mettre 2 courbes en correspondance

Si l'on minimise $d_H(A, B)$ (par rapport à la courbe A), A va se déformer progressivement jusqu'à devenir B.

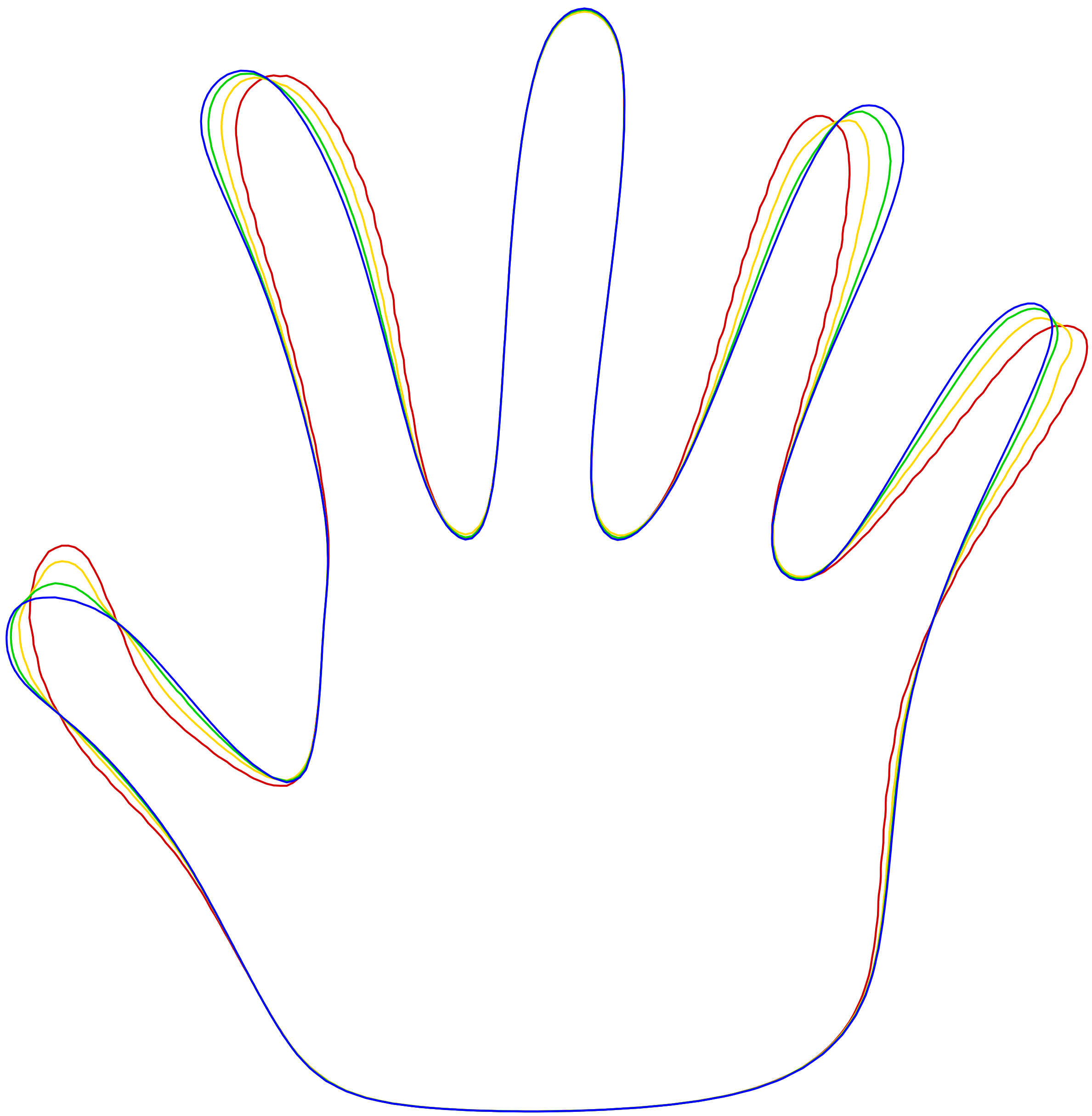
$\hookrightarrow$ descente de gradient : on applique à la courbe A à chaque pas de temps de l'évolution le champ $-\partial_A \, d_H(A, B)$



Problèmes !

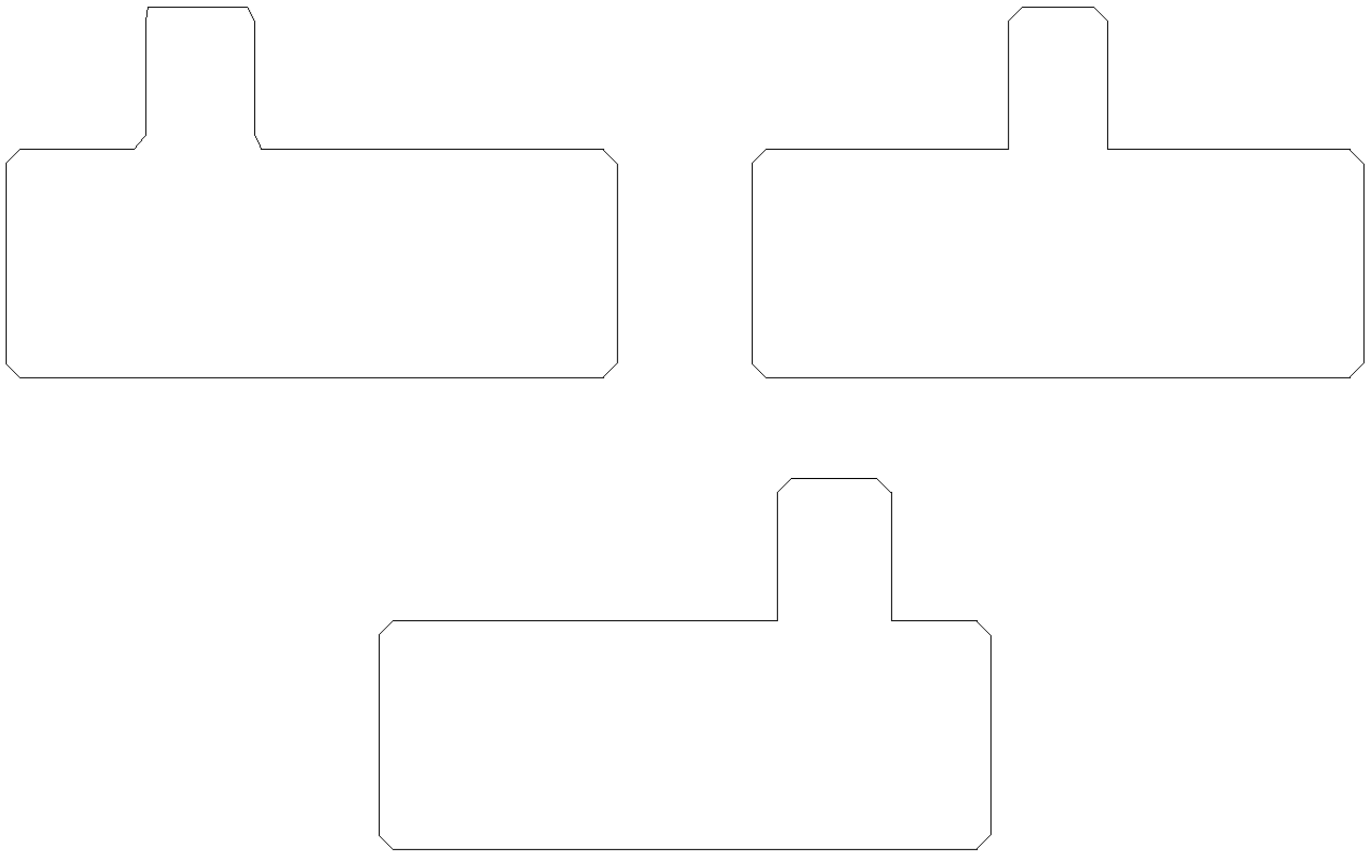
$\hookrightarrow$ la distance de Hausdorff n'est vraiment pas dérivable.
On utilise une approximation lisse de la distance de Hausdorff à la place.

$\hookrightarrow$ l'approximation lisse n'est vraiment plus une distance.
La précision de l'approximation est réglable : l'écart avec la vraie distance est inférieure à un ε *arbitrairement petit en fonction des paramètres de l'approximation.*



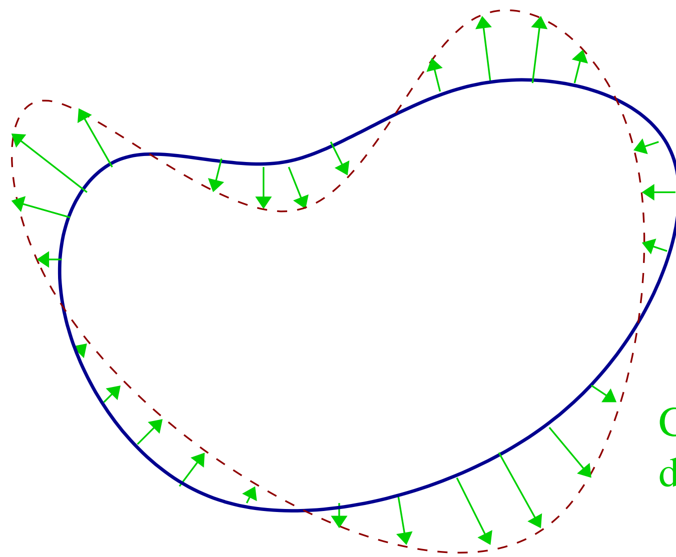
Exemple d’évolution

Comportement qualitatif



Interlude

Descente de gradient, espace tangent, produit scalaire



Minimiser $E(C)$, fonctionnelle d'une courbe.

$\hookrightarrow$ dérivée : forme linéaire

$$DE(C)(\delta C) = \lim_{\varepsilon \to 0} \frac{E(C + \varepsilon\, \delta C) - E(C)}{\varepsilon}$$

Produit scalaire canonique :

$$\langle \delta_1 \,|\delta_2 \rangle_{L^2} = \int_C \delta_1(x)\delta_2(x)dx$$

$\hookrightarrow$ produit scalaire H^1 :

$$\langle \delta_1 \,|\delta_2 \rangle_{H^1} = \int_C \delta_1(x)\delta_2(x)dx + \int_C \partial_x\delta_1(x)\ \partial_x\delta_2(x)dx$$

Gradient : défini pour un produit scalaire donné

$\hookrightarrow$ $DE(C)(\delta C) = \langle \nabla E(C) \,|\delta C \rangle$

Le choix du produit scalaire détermine le type de chemin suivi.

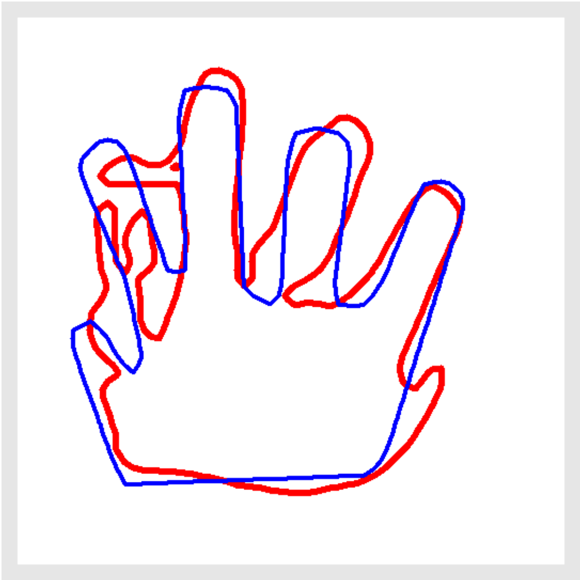
Rigidifier le mouvement

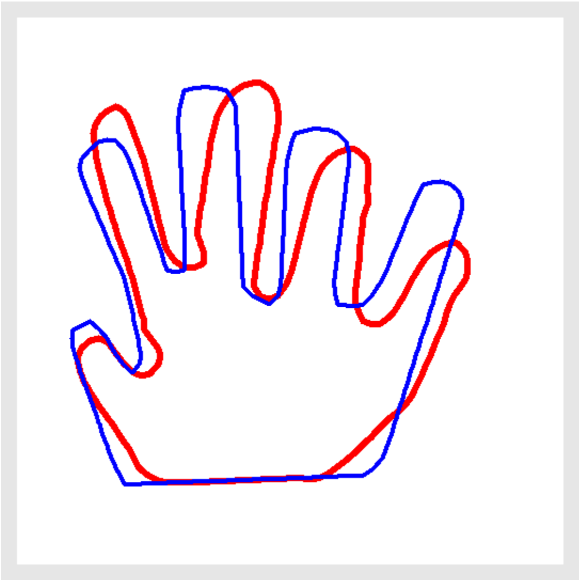
Grâce à un changement de produit scalaire, on peut favoriser les similitudes (translations, rotations, homothéties).

$\hookrightarrow$ changer le type d'évolution (le chemin suivi)

$\hookrightarrow$ tout en minimisant la même énergie

$\hookrightarrow$ permet d'éviter certains minima locaux



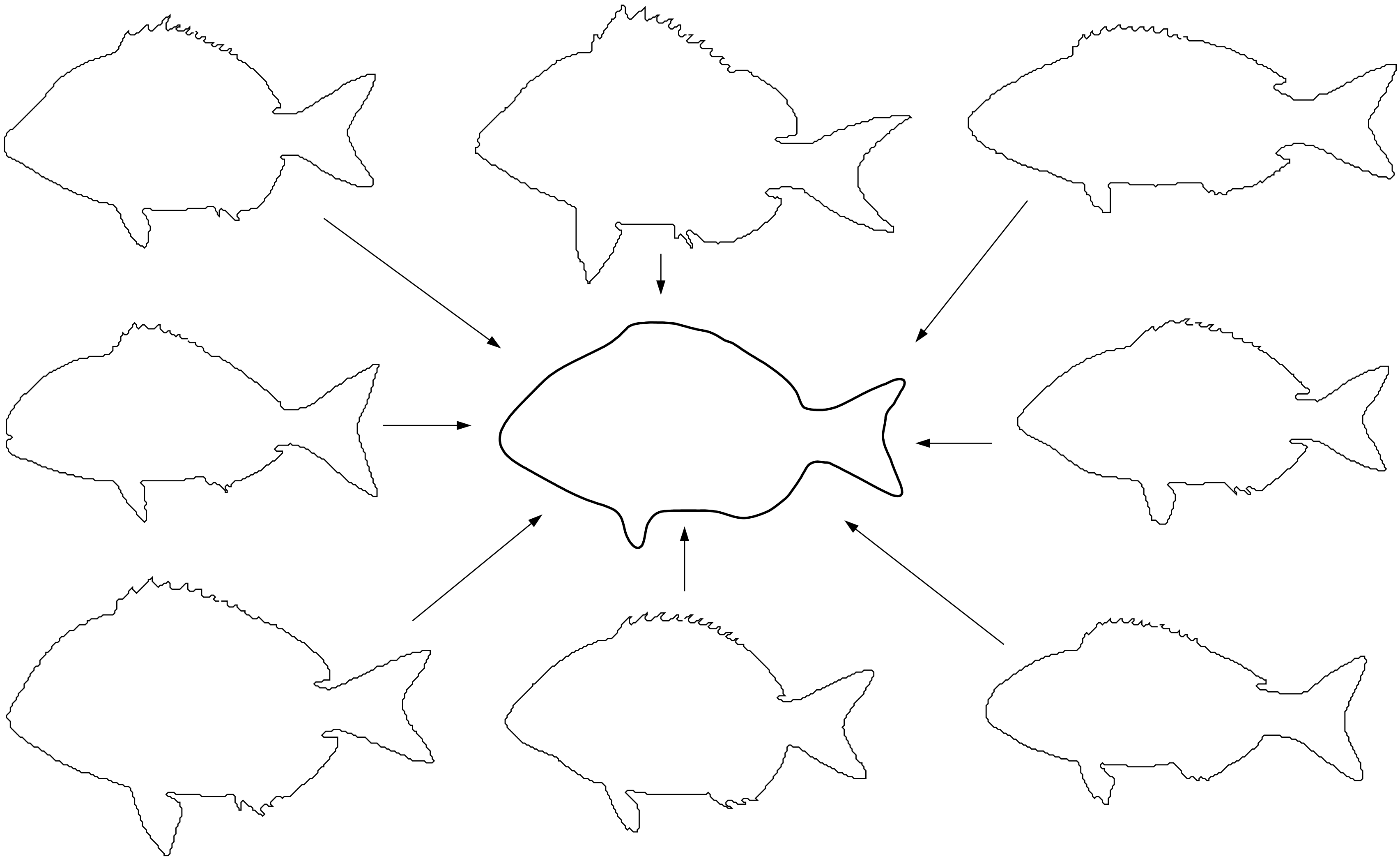


Moyenne M de n courbes A_i

La moyenne M est la courbe qui minimise $\sum_i d(M, A_i)^2$.

$\hookrightarrow$ *barycentre de* n *points*

Exemple de moyenne



Statistiques

Nous disposons :

$\hookrightarrow$ des n courbes A_i

$\hookrightarrow$ de leur moyenne M

$\hookrightarrow$ de n champs de déformations $c_i = -\partial_M E^2(M, A_i)$ à appliquer à M pour que celle-ci se rapproche de la courbe A_i.

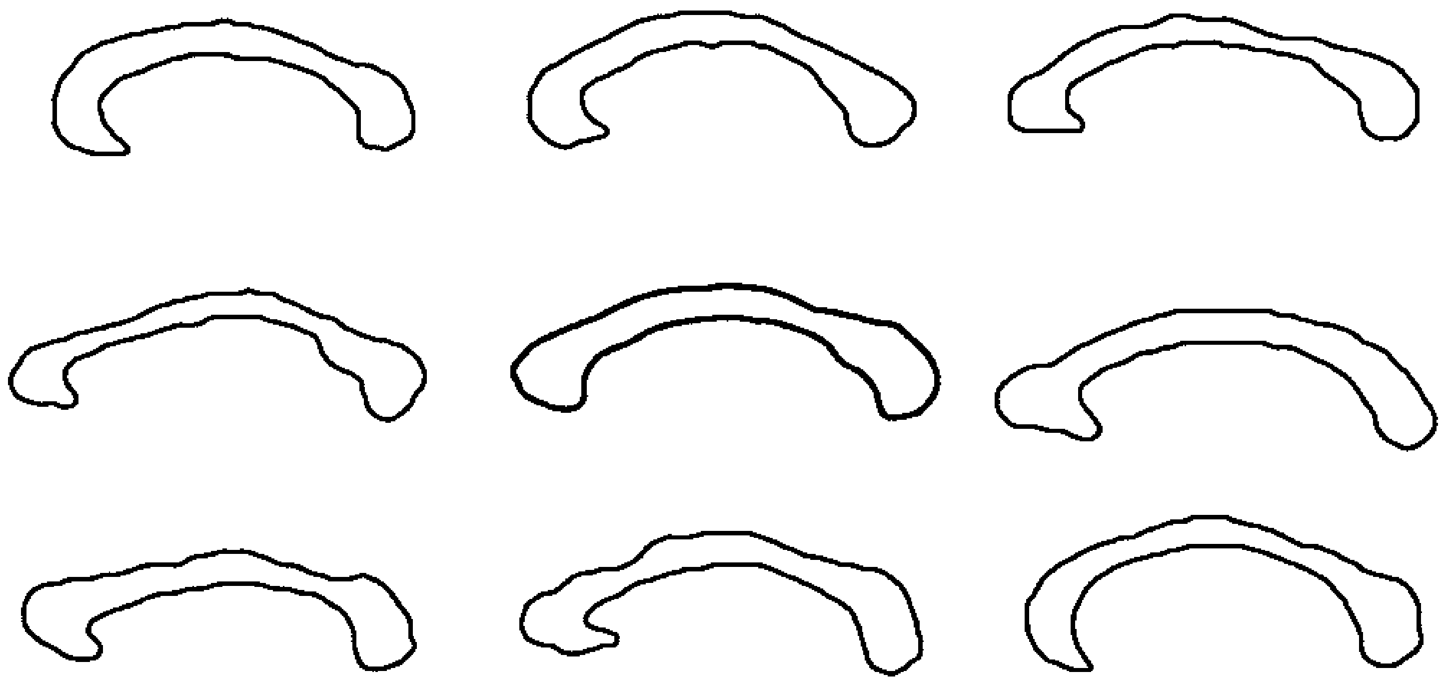
On calcule les statistiques de ces champs de déformation c_i :

$\hookrightarrow$ matrice P définie par $P_{i,j} = \left\langle c_i \middle| c_j \right\rangle_2 = \int_M c_i(x) c_j(x) dx$

$\hookrightarrow$ on diagonalise M

$\hookrightarrow$ on en extrait les vecteurs propres qui sont les modes de déformation caractéristiques

Exemple : les corpi callosi



modes : 1 2 3 4

Probabilité d'une courbe

Données :

$\hookrightarrow$ C : courbe que l'on fait évoluer pour segmenter l'image

$\hookrightarrow$ $\mathcal{D} = \{C_i\}$: ensemble de courbes donné en apprentissage

On cherche :

$\hookrightarrow$ $E(C)$? : mesure de la dissimilarité entre C et $\mathcal{D}$

Méthode simpliste

On dispose d'une distance d entre les courbes :

$\hookrightarrow$ $E(C) = \sum_i d^2(C, C_i)$

$\hookrightarrow$ le minimum étant atteint en M, moyenne des courbes C_i,
autant utiliser $E(C) = d(C, M)$

$\hookrightarrow$ ne tient pas compte des variations de formes dans l'ensemble $\mathcal{D}$

Méthode Parzen

On dispose d'une distance d entre les courbes :

$\hookrightarrow$ $E(C) = \sum_i \exp^{-\frac{d^2(C,C_i)}{2\sigma^2}}$

$\hookrightarrow$ cas d'un échantillon clairsemé pour l'échelle σ :
attraction de C vers un puits de potentiel autour de la courbe la plus proche

$\hookrightarrow$ cas d'un échantillon plus dense :
attraction de C vers une moyenne locale des formes les plus proches

Méthode par modes propres (ACP sur le gradient)

On dispose des statistiques sur $\mathcal{D}$, du champ $c = -\nabla d^2(M, C)$ et on suppose une répartition gaussienne autour des modes propres m_k :

$\hookrightarrow$ $E(C) = \sum_k \frac{1}{\sigma_k^2} \langle c \,|m_k\rangle^2 + \frac{1}{\sigma_{\text{bruit}}^2} \|\text{Reste}(c)\|_2^2$

$\hookrightarrow$ diminution des degrés de liberté, ellipsoïde.

Méthode Parzen sur le gradient

On dispose du champ $c = -\nabla d^2(M, C)$, des champs $c_i = -\nabla d^2(M, C_i)$, et d'une distance $\| \cdot \|_2$ dans cet espace des champs de déformation :

$\hookrightarrow$ $E(C) = \sum_i \exp^{-\frac{\|c - c_i\|_2^2}{2\sigma^2}}$

$\hookrightarrow$ $\nabla_C E(C)$ nécessite le calcul de la dérivée seconde $D_C \, D_M \, d^2(M, C)$

Statistiques d'images

Problème similaire, approche similaire.

Moyenne d'images

On dispose de n images I_i de même taille.

$\hookrightarrow$ définition d'un critère de ressemblance $E(A, B)$ entre deux images A et B
on choisit la corrélation croisée locale avec une approche multi-échelle (cf Gérardo Hermosillo)

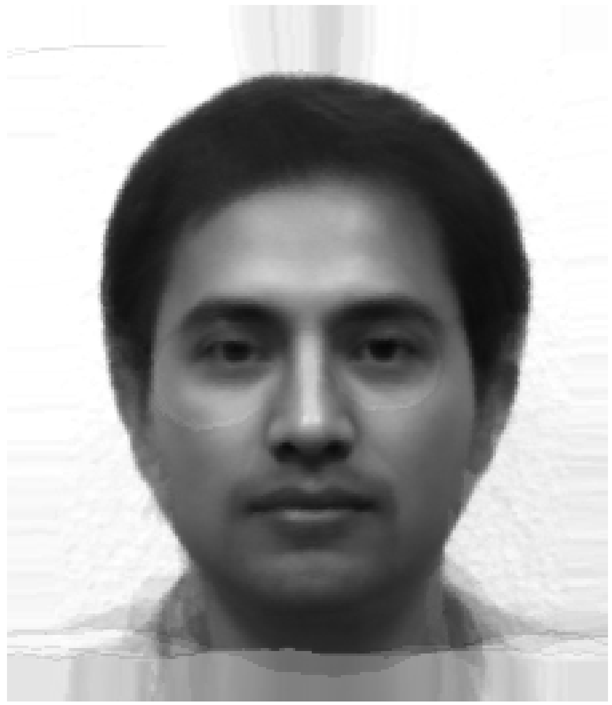
- $\hookrightarrow$ définition de la moyenne M de n images I_i comme étant l'image minimisant $\sum_i E(M, I_i)$
  problèmes de minima locaux très importants
- $\hookrightarrow$ on introduit n difféomorphismes h_i tels que chaque image déformée $A_i \circ h_i$ est censée ressembler à M
- $\hookrightarrow$ on minimise en les h_i : $\sum_{i,j} E(I_i \circ h_i, I_j \circ h_j) \; + \; \sum_i R(h_i)$
- $\hookrightarrow$ la moyenne est $M = \frac{1}{n} \sum_i I_i \circ h_i$.

Exemple : moyenne de photos de visages

Quelques unes des images I_i

Les images déformées $I_i \circ h_i$

La moyenne



Les modes de déformations

Pour passer de la moyenne M des images à une image particulière I_i, il y a deux chose à faire :

$\hookrightarrow$ déformer l'image M (en lui appliquant le difféomorphisme h_i^{-1})

$\hookrightarrow$ changer l'intensité de l'image M (en lui ajoutant $I_i \circ h_i - M$)

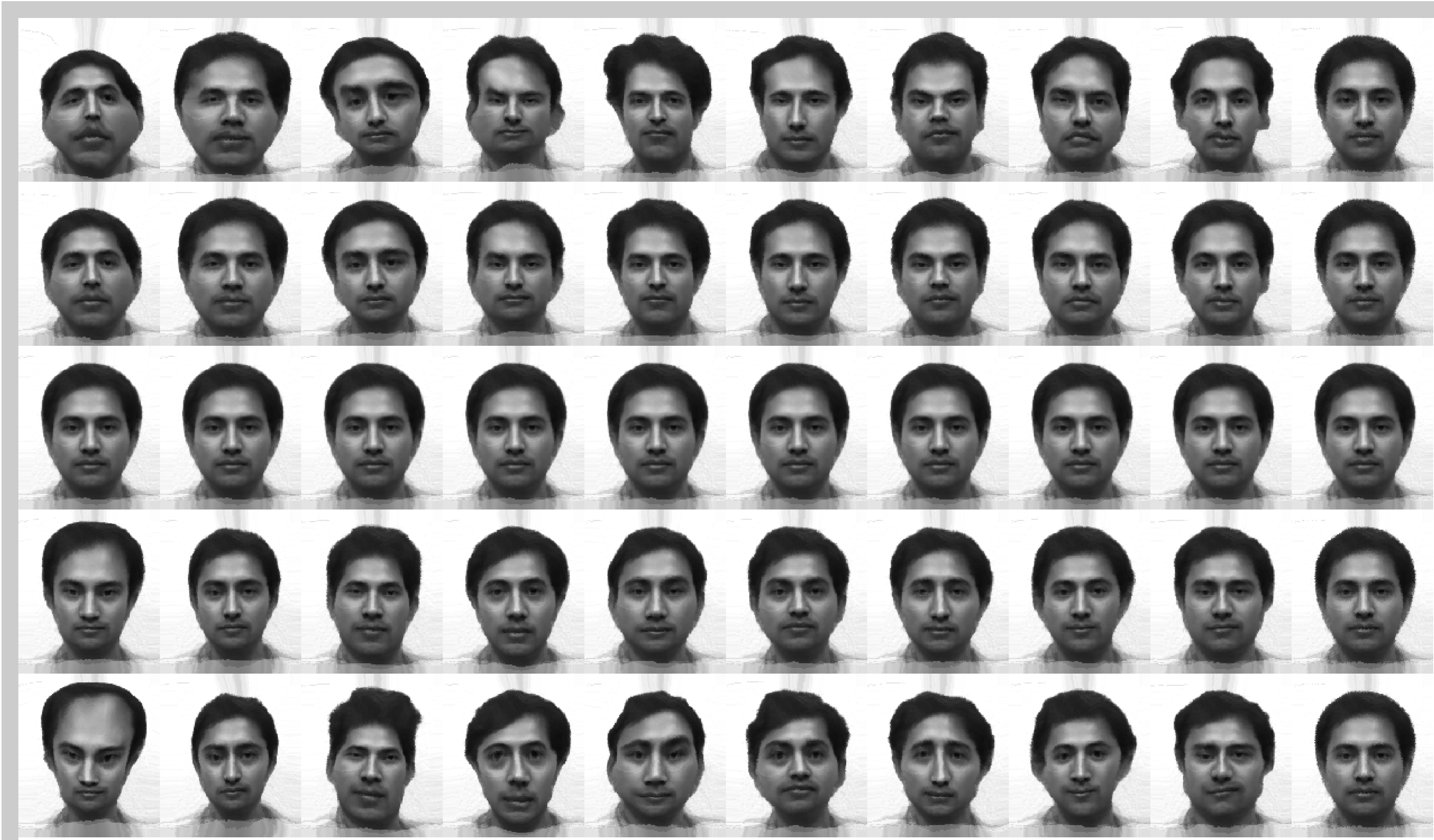
Il y a donc deux types de statistiques à faire à partir de la moyenne M :

$\hookrightarrow$ 1- statistiques sur les déformations h_i^{-1}

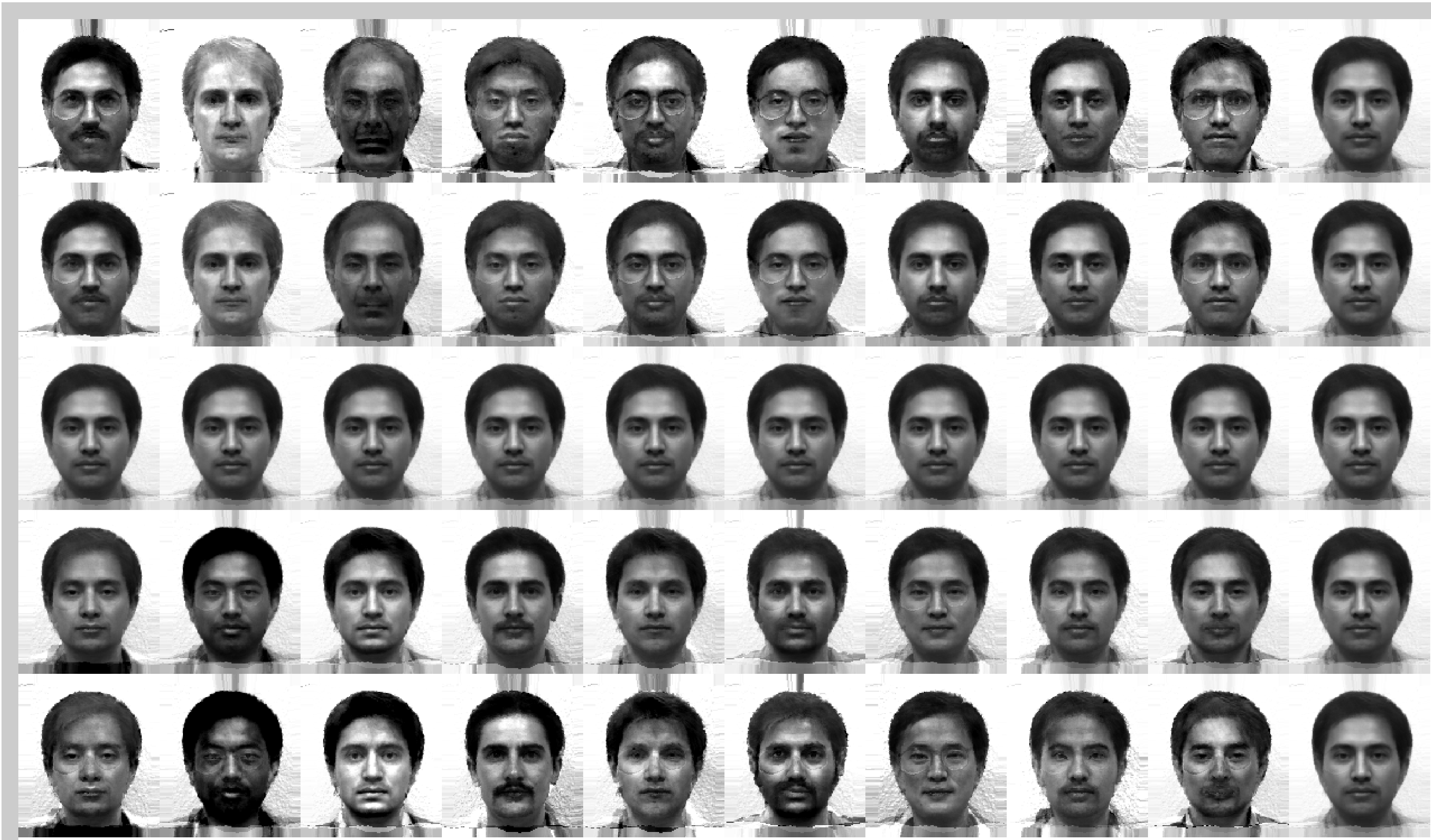
$\hookrightarrow$ 2- statistiques sur les variations d'intensité $I_i \circ h_i - M$

$\hookrightarrow$ 3- statistiques couplées déformations/intensités.

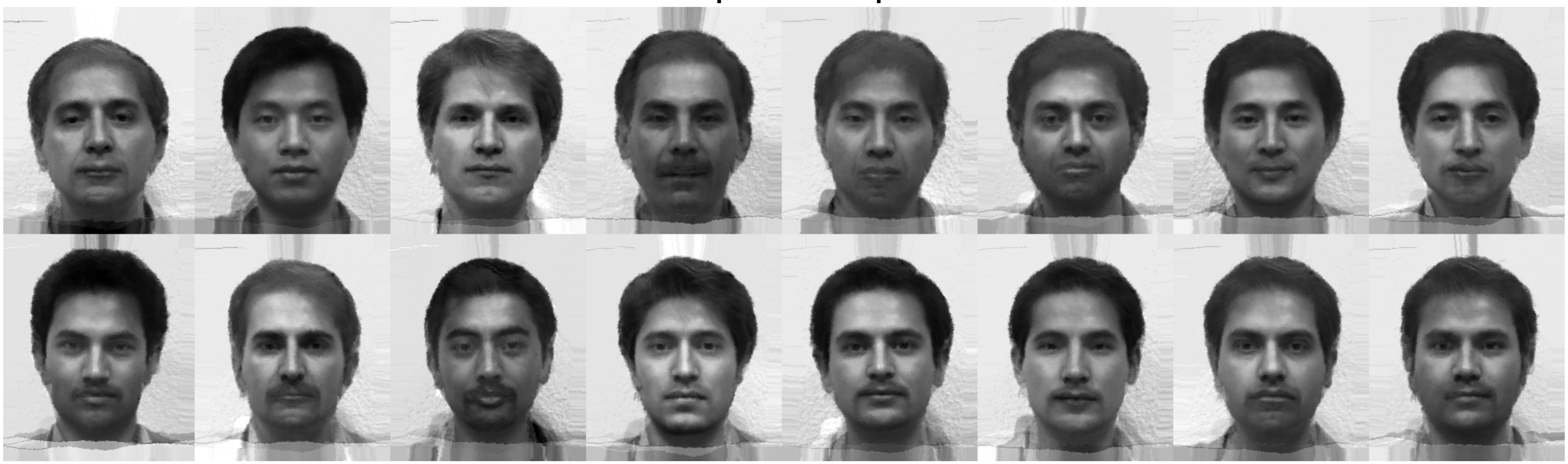
Statistiques sur les déformations



Statistiques sur les variations d’intensité



Statistiques couplées



modes 1 2 3 4 5 6 7 8

Utilisation des statistiques : reconnaissance d'expressions

On dispose d'une base de données de visages, de n personnes avec chacun m attitudes différentes (étiquetées).

On nous donne deux nouvelles photos d'une même personne, l'une dans son état "normal", l'autre avec une expression particulière à déterminer.

Comment faire ?

- $\hookrightarrow$ calculer la moyenne des photos "normales" de la base (ce qui met en correspondance la moyenne avec chacune des photos "normales")
- $\hookrightarrow$ pour chaque personne de la base, mettre en correspondance sa photo "normale" avec chacune des photos expressives
- $\hookrightarrow$ ramener ces déformations/variations d'intensité à la moyenne (afin de pouvoir les comparer)
- $\hookrightarrow$ faire des statistiques sur ces déformations/variations d'intensité pour chacune des expressions sur la base de données

↪ calculer la déformation/variation d'intensité qui envoie la nouvelle photo “normale” sur celle avec une expression

↪ rapatrier ces variations sur la moyenne (via une mise en correspondance de la moyenne et de la nouvelle image “normale”)

↪ comparer aux statistiques et conclure.

Conclusion

Diverses méthodes pour un a priori sur la forme.