

Applications de grandes déformations aléatoires pour l'estimation de la moyenne d'images bruitées

J. Bigot, S. Gadat, J.M. Loubès

Institut de Mathématiques de Toulouse

7 Avril 2008

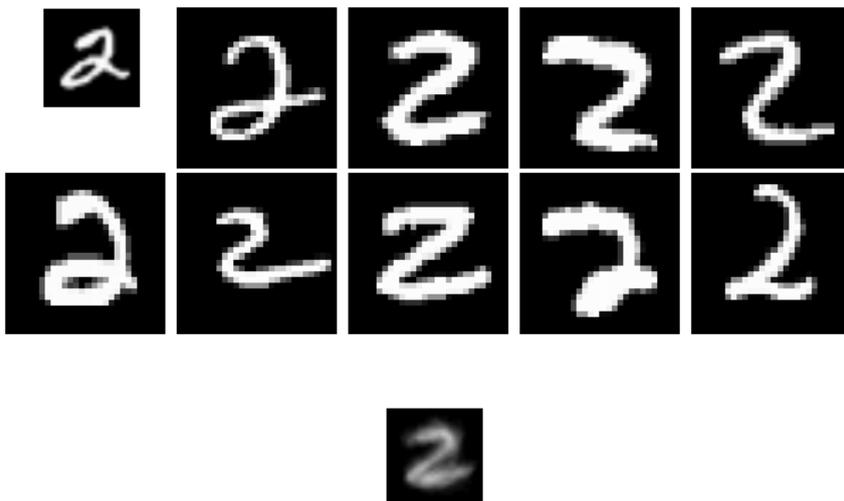
Formulation du problème

- Données : $I_i : \Omega \rightarrow \mathbb{R}, i = 1, \dots, n$ avec $\Omega \subset \mathbb{R}^2$, comment calculer une bonne moyenne de l'image ?
- **Estimateur naïf** : Prendre la moyenne arithmétique

$$\hat{I}_n = \frac{1}{n} \sum_{i=1}^n I_i$$

- **Problème** : Même si les images sont similaires, \hat{I}_n est en général un estimateur approprié de la forme moyenne typique des images.

Formulation du problème



Moyenne naïve de 10 images (base mnist, 28×28 pixels)

Image warping

Soient $I_i : \Omega \rightarrow \mathbb{R}, i = 1, \dots, n$ un ensemble d'images.

- **Objectif de la déformation d'image** : trouver n transformations $\Phi_i : \Omega \rightarrow \Omega, i = 1, \dots, n$ telles que
 - $I_1(\Phi_1(x)) \approx I_2(\Phi_2(x)) \approx \dots \approx I_n(\Phi_n(x))$ pour tout $x \in \Omega$
 - Φ_i inversible et lisse.
- Finalement, la déformation d'image est une tâche préliminaire souvent nécessaire avant de faire une moyenne ou d'autres analyses statistiques

Applications : Reconnaissance de formes, estimation de moyenne, clustering, calcul d'Atlas, ACP, etc...

- 1 Modèle statistique de déformation aléatoires d'images
- 2 M-estimation et déformation pour la moyennisation d'images
- 3 Simulations
- 4 Quelques perspectives

Plusieurs modèles de déformations

Transformation rigides

- Translation : $\Phi(x) = x - b$
- Rotation + homothétie : $\Phi(x) = \frac{1}{a}A_\theta x$ where $a \in \mathbb{R}^+$ and

$$A_\theta = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}$$

- Affine (Translation + rotation + homothétie) : $\Phi(x) = \frac{1}{a}A_\theta(x - b)$

Plusieurs modèles de déformations

Transformations Non-rigides

- Petites déformations : $\Phi(x) = x + h(x)$ avec $h : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ est une fonction "petite" non contrainte . **Problème** : Φ n'est plus bijective

Travaux de Faugeras, Amit, Trouvé...

- Grandes déformations : $\Phi^1(x) = x + \int_0^1 v_t(\Phi^t(x))dt$ i.e.

$$\frac{\partial \Phi^t}{\partial t} = v_t \circ \Phi^t \text{ avec } \Phi^0 = x,$$

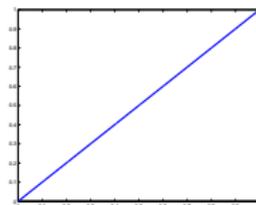
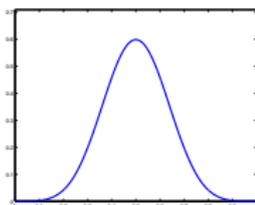
avec $v_t : \Omega \rightarrow \mathbb{R}^2, t \in [0, 1]$ est un champ de vecteur inhomogène.
Sous certaines conditions, $(v_t)_{t \in [0,1]}$ (principalement v_t et ses dérivées s'annulent aux frontières de Ω) alors :

Φ^1 est un difféomorphisme de $\Omega \rightarrow \Omega$.

Travaux de Grenander, Trouvé, Younes, Miller,...

Formulation du problème

Exemple de grande déformation en 1D

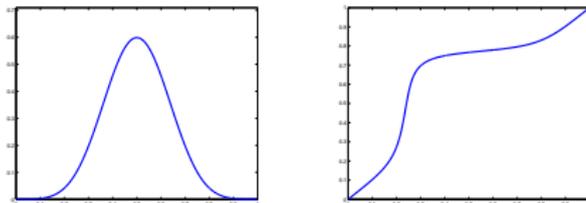


Choisir $v_t(x) = e(x)$ pour tout $t \in [0, 1]$ avec e fonction $e : [0, 1] \rightarrow \mathbb{R}$

$$\Phi^1(x) = x + \int_0^1 e(\Phi^t(x)) dt$$

Formulation du problème

Exemple de grande déformation en 1D

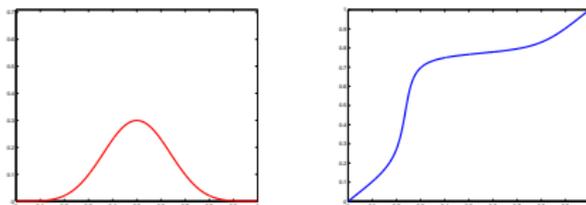


Choisir $v_t(x) = e(x)$ pour tout $t \in [0, 1]$ avec e fonction $e : [0, 1] \rightarrow \mathbb{R}$

$$\Phi^1(x) = x + \int_0^1 e(\Phi^t(x)) dt$$

Formulation du problème

Exemple de grande déformation en 1D

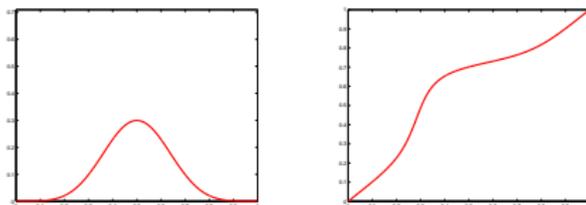


Choisir $v_t(x) = e(x)$ pour tout $t \in [0, 1]$ avec e fonction $e : [0, 1] \rightarrow \mathbb{R}$

$$\Phi^1(x) = x + \int_0^1 e(\Phi^t(x)) dt$$

Formulation du problème

Exemple de grande déformation en 1D



Choisir $v_t(x) = e(x)$ pour tout $t \in [0, 1]$ avec e fonction $e : [0, 1] \rightarrow \mathbb{R}$

$$\Phi^1(x) = x + \int_0^1 e(\Phi^t(x)) dt$$

Un modèle de grande déformation aléatoire

- Soit $\Omega = [0, 1]^2$ et $A > 0$ on tire aléatoirement $a_k^{(j)} \sim_{i.i.d.} \mathcal{U}[-A, A]$, $k = 1, \dots, K$, $j = 1, 2$. On définit pour $x \in [0, 1]^2$

$$v_a(x) = \left(\sum_{k=1}^K a_k^{(1)} e_k(x), \sum_{k=1}^K a_k^{(2)} e_k(x) \right),$$

avec $e_k : [0, 1]^2 \rightarrow \mathbb{R}$ ensemble de fonctions linéairement indépendantes .

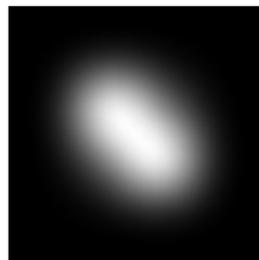
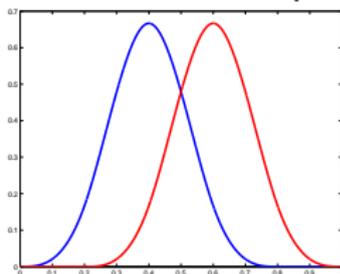
- Alors $\Phi_a(x) = \Phi_{v_a}(x)$ est la solution au temps $t = 1$ de l'équation :

$$\Phi_{v_a}^1(x) = x + \int_0^1 v_a(\Phi_{v_a}^t(x)) dt.$$

Remarque : le champ de vecteur est ici homogène en temps

Un modèle de grande déformation aléatoire

Simulations : Pour les $e_k : [0, 1]^2 \rightarrow \mathbb{R}$, $k = 1, \dots, K$, prendre des produits tensoriels de B-splines 1D qui s'annulent aux frontières of $[0, 1]$. Ex : Nombre de B-splines 1D = 2, $K = 2 \times 2 = 4$



Deux B-splines 1D / Un exemple de $\sum_{k=1}^K a_k^{(1)} e_k(x)$ avec $A = 0.5$.

Un modèle de grande déformation aléatoire

Un exemple de déformation aléatoire obtenue



Lena - 256 × 256 pixels

Un modèle de grande déformation aléatoire

Un exemple de déformation aléatoire obtenue



Déformation avec $A = 0.1$ (amplitude des $a_k^{(j)}$'s)

Un modèle de grande déformation aléatoire

Un exemple de déformation aléatoire obtenue



Déformation avec $A = 0.5$ (amplitude des $a_k^{(j)}$'s)

Un modèle de grande déformation aléatoire

On suppose que les observations proviennent du modèle de déformation.

On observe n images sur une grille de $N \times N$ pixels. Sur chaque pixel p on mesure une réalisation $I_i, i = 1, \dots, n$ donnée par le modèle

$$I_{a_i}(p) = I^* \circ \Phi_{a_i}^1(p) + \varepsilon_i(p), \quad i = 1, \dots, n,$$

avec

- $I^* : [0, 1]^2 \rightarrow \mathbb{R}$ est l'image inconnue à estimer,
- a_i sont des vecteurs i.i.d (in $[-A; A]^{2K}$),
- $\varepsilon_i(p) \in \mathbb{R}$ sont des bruits i.i.d. centrés de variance finie.

Problème : l'estimation de I^* peut être vue comme un problème inverse dont l'opérateur Φ^1 est inconnu et doit être estimé.

- 1 Modèle statistique de déformation aléatoires d'images
- 2 M-estimation et déformation pour la moyennisation d'images
- 3 Simulations
- 4 Quelques perspectives

M-estimateurs en statistique

- **Exemple** : soient $X_1, \dots, X_n \sim_{i.i.d.} P$ avec $\alpha^* = \mathbb{E}_P(X_1)$, on définit

$$\hat{F}_n(\alpha) = \sum_{i=1}^n (X_i - \alpha)^2 \text{ et } F(\alpha) = \int (x - \alpha)^2 dP(x) \text{ pour } \alpha \in \mathbb{R}, \text{ alors}$$

$$\hat{\alpha}_n = \frac{1}{n} \sum_{i=1}^n X_i = \arg \min_{\alpha \in \mathbb{R}} \hat{F}_n(\alpha) \text{ et } \alpha^* = \arg \min_{\alpha \in \mathbb{R}} F(\alpha)$$

- Si l'on veut estimer α^* tel que

$$\alpha^* = \arg \min_{\alpha \in \mathcal{A}} F(\alpha)$$

avec $F : \mathcal{A} \rightarrow \mathbb{R}$ et \mathcal{A} est un ensemble (compact) de paramètres.

Avec n observations, on construit une fonction de contraste $\hat{F}_n : \mathcal{A} \rightarrow \mathbb{R}$, et un M-estimateur de α^* donné par

$$\hat{\alpha}_n = \arg \min_{\alpha \in \mathcal{A}} \hat{F}_n(\alpha)$$

see e.g. van der Vaart (1998)

Fonction de contraste pour l'estimation de la moyenne

- Soit $\mathcal{Z} = \{Z : [0, 1]^2 \rightarrow \mathbb{R}\}$ l'ensemble des images, on définit

$$\mathcal{V} = \{v : \mathbb{R}^2 \rightarrow \mathbb{R}^2; v(x) = \left(\sum_{k=1}^K a_k^{(1)} e_k(x), \sum_{k=1}^K a_k^{(2)} e_k(x) \right), \text{ pour } a_k^{(j)} \in [-A, A]\}$$

- Soit $f(a, \varepsilon, Z) = \min_{v \in \mathcal{V}} \sum_{p=1}^{N^2} (I_a(p) - Z \circ \Phi_v^1(p))^2$, avec

$$I_a(p) = I^* \circ \Phi_a^1(p) + \varepsilon(p) \text{ avec } a \in [-A, A]^{2K} \text{ et } \varepsilon \in \mathbb{R}^{N^2}$$

- Soit $F(Z) = \int_{[-A, A]^{2K} \times \mathbb{R}^{N^2}} f(a, \varepsilon, Z) dP(a, \varepsilon)$ et son pendant empirique

$$\hat{F}_n(Z) = \frac{1}{n} \sum_{i=1}^n f(a_i, \varepsilon_i, Z)$$

Fonction de contraste pour l'estimation de la moyenne

On définit l'ensemble des minimiseurs (l'unicité n'est pas garantie...)

$$\hat{Q}_n = \arg \min_{Z \in \mathcal{Z}} \hat{F}_n(Z) \text{ et } Q_0 = \arg \min_{Z \in \mathcal{Z}} F(Z)$$

Soit $Z_{n_k} \in \hat{Q}_{n_k}$ une sous-suite convergente de minimiseurs et $\hat{Z} = \lim_k Z_{n_k} \in \hat{Q}$ est l'ensemble des limites possibles

Theorem

Si les e_k sont bornés, que \mathcal{Z} est compact et I^ est uniformément Lipschitz sur $[0, 1]^2$, on a*

$$\hat{Q} \neq \emptyset \text{ and } \hat{Q} \subset Q_0 \text{ presque sûrement .}$$

Fonction de contraste pour l'estimation de la moyenne

- La compacité de \mathcal{Z} sert juste à garantir l'existence des minimiseurs
- La condition Lipschitz sur I^* permet un contrôle uniforme de $\sup_{Z \in \mathcal{Z}} |\hat{F}_n(Z) - F(Z)| \rightarrow 0$ p.s. lorsque $n \rightarrow +\infty$
- La consistance a lieu pour les éléments de l'argmin de \hat{F}_n !
- Conclusion : chercher à minimiser \hat{F}_n

Calcul du minimiseur de la fonction de contraste

Procédure itérative : Initialiser avec $Z^{(1)} = \hat{I}_n$ (moyenne naïve). Pour $m = 2, \dots, M$ répéter les étapes suivantes :

- pour $i = 1, \dots, n$ recuit simulé ou une descente de gradient pour calculer la déformation optimale $\Phi_{\hat{\alpha}_i^m}$ qui correspond au champ de vecteur

$$v_{\hat{\alpha}_i^m} = \arg \min_{v \in \mathcal{V}} \sum_{p=1}^{N^2} \left(I_i(p) - Z^{(m-1)} \circ \Phi_v^1(p) \right)^2$$

- calculer $Z^{(m)} = \arg \min_{Z \in \mathcal{Z}} \sum_{i=1}^n \sum_{p=1}^{N^2} \left(I_i(p) - Z^{(m-1)} \circ \Phi_{\hat{\alpha}_i^m}(p) \right)^2$ donné par

$$Z^{(m)}(p) = \frac{\sum_{i=1}^n w_i(p) I_i \circ \Phi_{\hat{\alpha}_i^m}^{-1}(p)}{\sum_{i=1}^n w_i(p)} \text{ avec } w_i(p) = |\det \text{Jac}(\Phi_{\hat{\alpha}_i^m}^{-1})(p)|.$$

Quelques commentaires

- On construit ainsi une suite Z^m telle que $\hat{F}_n(Z^m)$ décroît, mais pas nécessairement vers un minimum de $\hat{F}_n \dots$
- Utilisation de méthode Bayésienne type EM/SA-EM pour approcher ce minimiseur (cf cet am !) pour palier aux problèmes de convergence de l'algorithme type 'min-min'.
- L'approche Bayésienne est légèrement différente puisque la fonction de contraste est de la forme

$$\mathbb{E}_v \|I_a - Z \circ \Phi_v^1\|^2 + \lambda \|v\|$$

Quelques commentaires

- On peut espérer la normalité asymptotique de \hat{Q}_n vers Q_0 en $n^{-1/2}$ par des méthodes classiques de Z-estimation.
- Lorsque le bruit est nul ($\epsilon = 0$), la vraie image I^* est dans Q_0 , c'est faux en approche Bayésienne (convolution avec la loi des déformations).
- Vitesse d'approximation à horizon fini ?
- Que se passe-t-il si la loi des observations est inconnue ?

Modèle plus général

- (I_i) , i.i.d. de loi P inconnue, modèle d'image $(Z_\theta)_{\theta \in \Theta}$ continu ou discret :

$$Z_\theta = \sum_{\lambda \in \Lambda} \theta_\lambda \psi_\lambda$$

- Fonction de contraste :

$$f(I, Z_\theta) = \min_{v \in \mathcal{V}_A} [\|I - Z_\theta \circ \Phi_v^1\|^2 + \lambda_1 \text{pen}_1(v)] + \lambda_2 \text{pen}_2(\theta)$$

- $\lambda_1 \text{pen}(v)$ contraint les déformations, $\lambda_2 \text{pen}(\theta)$ contraint les images Z .

Modèle plus général

- On a consistance du M -estimateur sous certaines hypothèses.
- On suppose que $F(Z_\theta)$ admet un unique minimum en Z_{θ^*} et que Λ est fini, Θ compact et \mathcal{V} décrit par un ensemble de paramètres compact.

Theorem

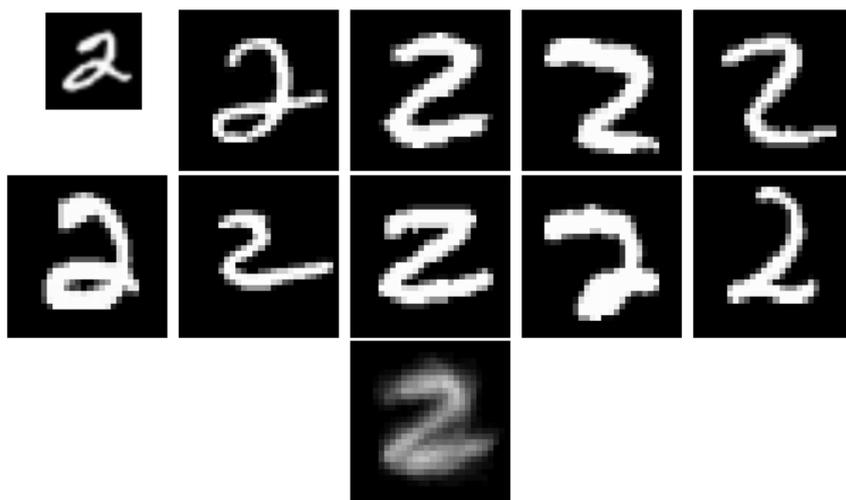
Si p_{en_1} et p_{en_2} sont continues, que P a un moment d'ordre 2 et $\|Z_\theta\| \sim \|\theta\|$, alors le M -estimateur est consistant :

$$\lim_{n \rightarrow \infty} \|Z_{\theta_n} - Z_{\theta^*}\|_\infty = 0 \quad a.s.$$

- Idée de la preuve : utiliser les classes de Glivenko-Cantelli pour $\{f(\cdot, Z_\theta), \theta \in \Theta\}$

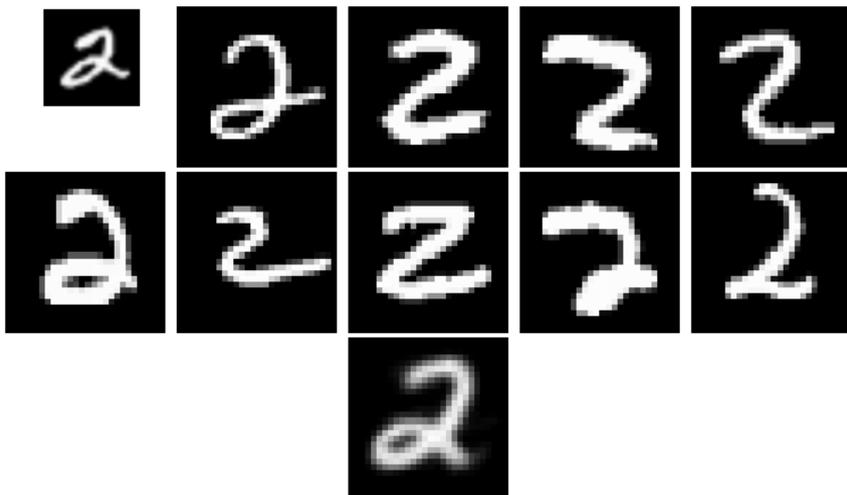
- 1 Modèle statistique de déformation aléatoires d'images
- 2 M-estimation et déformation pour la moyennisation d'images
- 3 Simulations**
- 4 Quelques perspectives

Estimation moyenne de chiffres



Moyenne naïve de 20 images (mnist, 28×28 pixels)

Estimation moyenne de chiffres



*Moyenne par M-estimation (schéma min-min) de 20 images (mnist,
28 × 28 pixels)*

Reconnaissance de chiffres manuscrits de 0 à 9

Ensemble d'apprentissage : 20 images par classe (i.e. 200 images)

Ensemble de test : 100 images par classe (i.e. 1000 images)

- **Classification naïve** : calcul de la moyenne naïve $\hat{I}^k, k = 0, \dots, 9$ pour chaque classe de l'ensemble d'apprentissage. Pour une nouvelle image I , choisir

$$\hat{k} = \arg \min_{0 \leq k \leq 9} \sum_{p=1}^{N^2} (I(p) - \hat{I}^k(p))^2$$

- **Classification avec déformation** : calcul de la moyenne $\hat{Z}^k, k = 0, \dots, 9$ pour chaque classe par M-estimation type min-min avec déformation. Pour une nouvelle image I , la déformer sur la moyenne \hat{Z}^k de chaque classe, et choisir

$$\hat{k} = \arg \min_{0 \leq k < 9} \sum_{p=1}^{N^2} (I \circ \Phi_{\hat{a}_k}(p) - \hat{Z}^k(p))^2$$

Reconnaissance de chiffres manuscrits de 0 à 9

Taux d'erreurs sur l'ensemble test :

Classification Naïve	Classification avec déformation	SVM
30.2 %	15.2 %	21.3 %

Pour le chiffre 2, le taux d'erreur de classification avec déformation est très grand (65 %) :



En fait, il y a au moins deux sous-classes de 2 !

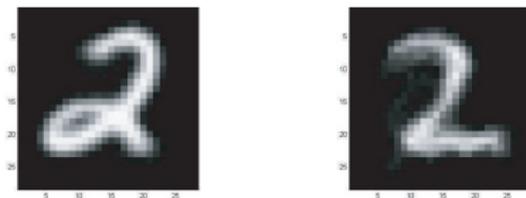
Reconnaissance de chiffres manuscrits de 0 à 9

K-means clustering pour 20 images de la classe 2 de l'ensemble d'apprentissage

- Soit Y_1, \dots, Y_n vecteurs d'un espace \mathcal{D} . L'algorithme du K-means à deux classes trouve une partition $C_1 \cup C_2$ de $\{1, \dots, n\}$ et $G_1, G_2 \in \mathcal{D}$ d'inertie intra-classe minimale pour une fonction distance d donnée.
- Pour chaque image, calculer G_1 et G_2 les moyennes de chaque sous-classe C_1 and C_2 et définir la distance L^2 après déformation optimale de Y_i sur G_k $d^2(Y_i, G_k)$.

Reconnaissance de chiffres manuscrits de 0 à 9

Exemple de K-means clustering pour 20 images par classe de chiffres 2 de l'ensemble d'apprentissage



Clusters G_1 and G_2

Avec ces deux sous-classes, le taux d'erreur de classification des 2 descend à 15 %, et le taux global à 12.5 %

Reconnaissance de chiffres manuscrits de 0 à 9

K-means clustering pour 20 images de la classe des chiffres 3 de l'ensemble d'apprentissage - Première sous-classe C_1



Reconnaissance de chiffres manuscrits de 0 à 9

K-means clustering pour 20 images de la classe des chiffres 3 de l'ensemble d'apprentissage - Seconde sous-classe C_2



Choix du nombre de classes par validation croisée.
Classification des chiffres après clustering sur chacune des 10 classes : 8.6 % (C'est loin d'être le meilleur taux existant).

- 1 Modèle statistique de déformation aléatoires d'images
- 2 M-estimation et déformation pour la moyennisation d'images
- 3 Simulations
- 4 Quelques perspectives

Quelques perspectives

- Unicité du minimum de $F(Z) = \int_{[-A,A]^k \times \mathbb{R}^{N^2}} f(a, \varepsilon, Z) dP(a, \varepsilon)$ avec $f(a, \varepsilon, Z) = \min_{v \in \mathcal{V}} \sum_{p=1}^{N^2} (I_a(p) - Z \circ \Phi_v^1(p))^2$, et On souhaiterait que $I^* = \arg \min_{Z \in \mathcal{Z}} F(Z) ??$
- Critère pénalisé pour l'estimation de I^*

$$\hat{Z}_n = \arg \min_{Z \in \mathcal{Z}} \frac{1}{n} \sum_{i=1}^n f(a_i, \varepsilon_i, Z) + \lambda pen_n(Z)$$

et contrôle de $\mathbb{E} \|\hat{Z}_n - I^*\|^2$ non asymptotique, que mettre dans pen_n ?

- Une première version en 1-D, pour des translations aléatoires de courbes bruitées.
- Donner un sens à une ACP avec déformation.

Merci de votre attention !