

Compressed Sensing

Reconnaissance faciale par représentation *sparse*

Benjamin Phan & Jean-Michel Roufousse

ENSAE ParisTech

Mars 2018

- 1 Introduction
- 2 Théorie
- 3 Implémentation et Résultats
 - Méthodologie
 - Visages avec Variation de Lumière
 - Visages avec lunettes de soleil ou écharpes
 - Visage avec du bruit
 - Rejet d'image
- 4 Conclusion

Introduction

- Idée datant de 2009
- Novateur dans :
 - *Feature Extraction*
 - Reconnaissance d'image bruitées, occultées
 - Computation rapide
 - Index SCI pour invalider certaines images

Théorie



(a) Base d'entraînement



(b) Image à tester

Figure: À partir d'une base d'image, reconnaître une image test

Théorie

- k : le nombre de classes dans la base d'entraînement (de visages)
- Les images ont une taille $m = w \times h$, où chaque pixel est empilé pour faire un vecteur $v \in \mathbb{R}^m$
- Pour une personne i donnée, on a n_i photos, alors la matrice

$$A_i = [v_{i,1}, v_{i,2}, \dots, v_{i,n_i}] \in \mathbb{R}^{m \times n_i}$$

contient chaque images en colonnes

- On concatène toutes ces classes en $A = [A_1, \dots, A_k]$

Théorie

L'idée de l'article est de décrire une image test $y \in \mathbb{R}^m$ comme combinaison linéaire des images de la base d'entraînement : $y = AX$

- La première idée qui consiste à résoudre avec la norme ℓ_2 ne permet pas de retrouver la bonne personne
- Nouvelle idée : trouver la solutions la plus *sparse* :

$$\hat{x}_0 = \arg \min ||x||_0 \text{ s.c. } Ax = y$$

où la 'norme' ℓ_0 compte le nombre de coefficients non nuls d'un vecteur

- Cette optimisation est NP-Hard, alors la norme ℓ_1 est utilisée si le nombre de de coefficients non nuls correspond à une petite fraction de la dimension m

1 Introduction

2 Théorie

3 Implémentation et Résultats

• Méthodologie

- Visages avec Variation de Lumière
- Visages avec lunettes de soleil ou écharpes
- Visage avec du bruit
- Rejet d'image

4 Conclusion

Sparse Representation Based Classification

- Input : Matrice $A = [A_1, \dots, A_k] \in \mathbb{R}^{m \times n}$ avec k classes et une image test $y \in \mathbb{R}^m$
- Les colonnes de A doivent être normalisées
- On résout l'optimisation suivant :

$$\hat{x}_1 = \arg \min ||x||_1 \text{ s.c. } Ax = y$$

- Les résidus sont calculés pour chaque classe $i = 1, \dots, k$

$$r_i(y) = ||y - A\delta_i(\hat{x}_1)||_2$$

avec δ_i une fonction qui choisit toutes les colonnes de la i -ème classe

- La bonne classe est : $\arg \min_i r_i(y)$

SRC avec du bruit

- À cause du bruit, il faut modéliser :

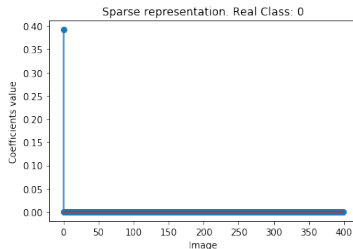
$$y = Ax_0 + z$$

avec $z \in \mathbb{R}^m$ mais avec $\|z\|_2 < \epsilon$

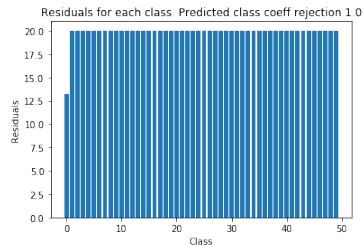
- La minimisation devient

$$\hat{x}_1 = \arg \min \|x\|_1 \text{ s.c. } \|Ax - y\|_2 \leq \epsilon$$

Exemple



(a) Vecteur *sparse* des coefficients de \hat{x}_1



(b) Résidus par classe

Figure: Résultat pour une image test donnée

Exemple

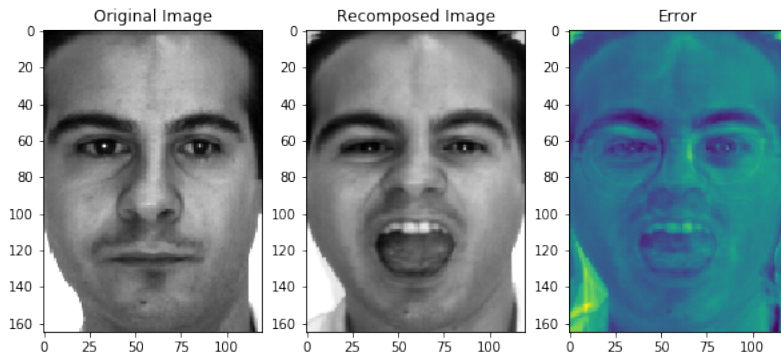


Figure: Sujet bien identifié, recomposé avec plusieurs images de la base d'entraînement

1 Introduction

2 Théorie

3 Implémentation et Résultats

- Méthodologie
- Visages avec Variation de Lumière
- Visages avec lunettes de soleil ou écharpes
- Visage avec du bruit
- Rejet d'image

4 Conclusion

Variation de Lumière

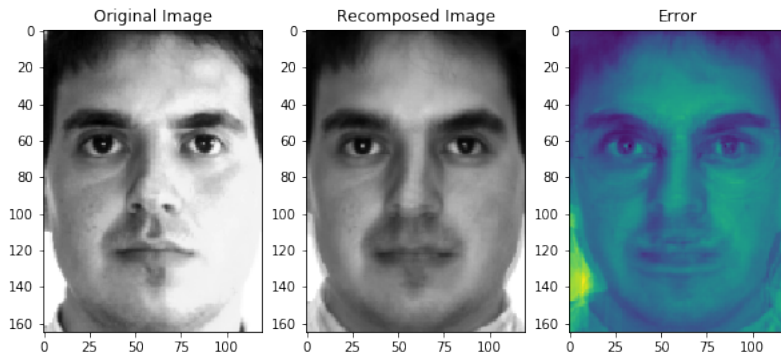


Figure: Sujet bien identifié, recomposé avec plusieurs images de la base d'entraînement ; 65,3 % de la base de test identifié

- 1 Introduction
- 2 Théorie
- 3 Implémentation et Résultats
 - Méthodologie
 - Visages avec Variation de Lumière
 - Visages avec lunettes de soleil ou écharpes
 - Visage avec du bruit
 - Rejet d'image
- 4 Conclusion

Lunette de Soleil

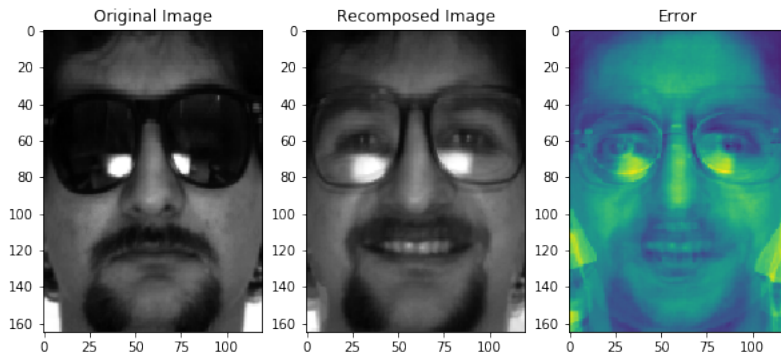


Figure: Sujet bien identifié, recomposé avec plusieurs images de la base d'entraînement ; 50 % de la base de test identifié

Lunette de Soleil

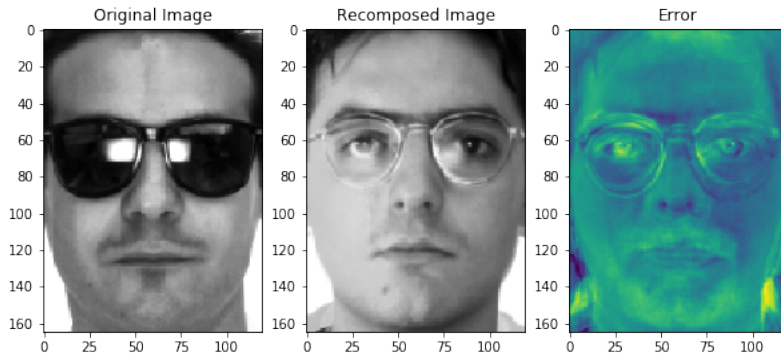


Figure: Sujet mal identifié

Écharpes

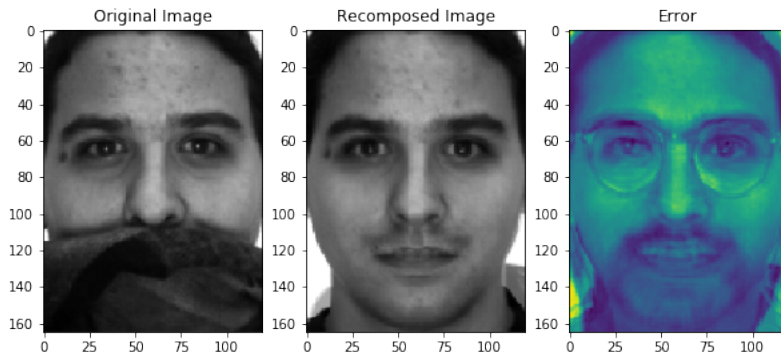


Figure: Sujet bien identifié, recomposé avec plusieurs images de la base d'entraînement ; 40 % de la base de test identifié

Écharpes

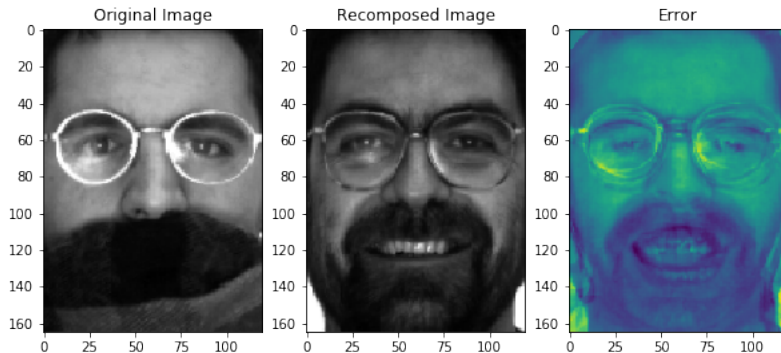


Figure: Sujet mal identifié (la barbe doit être prise pour une écharpe)

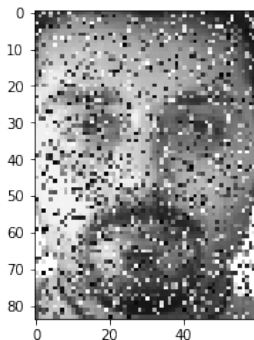
1 Introduction

2 Théorie

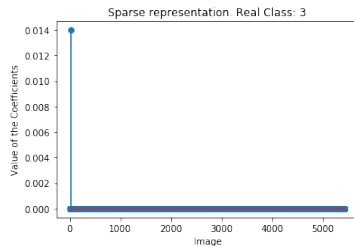
3 Implémentation et Résultats

- Méthodologie
- Visages avec Variation de Lumière
- Visages avec lunettes de soleil ou écharpes
- Visage avec du bruit
- Rejet d'image

4 Conclusion



(a) Image ainsi pixelisée



(b) Vecteurs des coefficients

Figure: Résultat pour une image Pixelisée - Reconnaissable par l'algorithme jusqu'à 65 % pour une photo sans occultation

1 Introduction

2 Théorie

3 Implémentation et Résultats

- Méthodologie
- Visages avec Variation de Lumière
- Visages avec lunettes de soleil ou écharpes
- Visage avec du bruit
- Rejet d'image

4 Conclusion

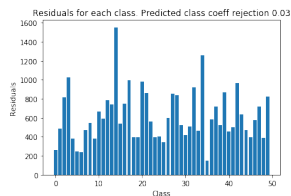
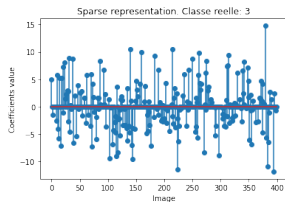


Figure: Résultat pour une image choisie au hasard sur Internet - SCI de 0,04

Conclusion

- 90,5% de reconnaissance dans des conditions idoines (sans occultation, bonne luminosité)
- Tester sur d'autres bases de données (Extended Yale Database)
- Les algorithmes de l'article sont plus performants
- Plus de comparaisons de méthodes (Nearest Neighbours, Nearest Subspace, SVM) et des réseaux de neurones
- Parfois les réseaux de neurones consomment beaucoup d'énergie, utiliser cette méthode permet d'être fiable tout en étant rapide et plus économe en énergie

References I



A.M. Martinez and R. Benavente. The AR Face Database. CVC Technical Report #24, June 1998.



J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma. Robust face recognition via sparse representation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 31(2):210 – 227, 2009



Ganesh, A., Wagner, A., Zhou, Z., Yang, A., Ma, Y., & Wright, J. (2012). Face recognition by sparse representation. In Y. Eldar & G. Kutyniok (Eds.), Compressed Sensing: Theory and Applications (pp. 515-539). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511794308.013

References II



MicrosoftResearch. "Robust Face Recognition via Sparse Representation." YouTube, YouTube, 6 Sept. 2016, www.youtube.com/watch?v=ZoCwNIXU-Hk.