

Dématriçage par représentations parcimonieuses

Olivier Losson, Ludovic Macaire

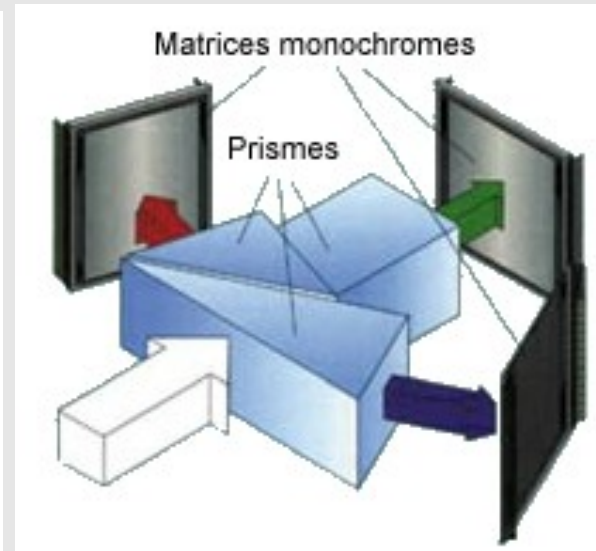
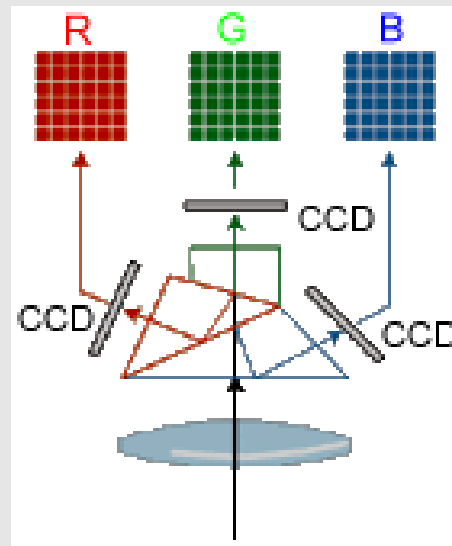
Université Lille, Laboratoire CRISAL

`olivier.losson, ludovic.macaire@univ-lille1.fr`

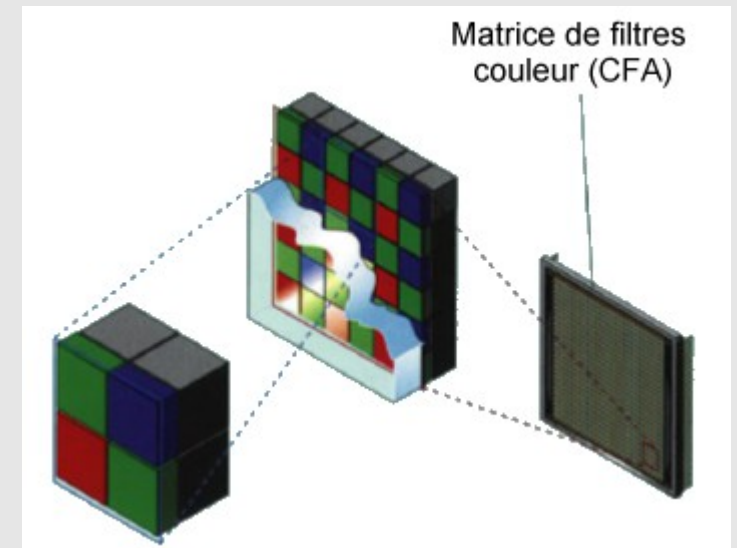
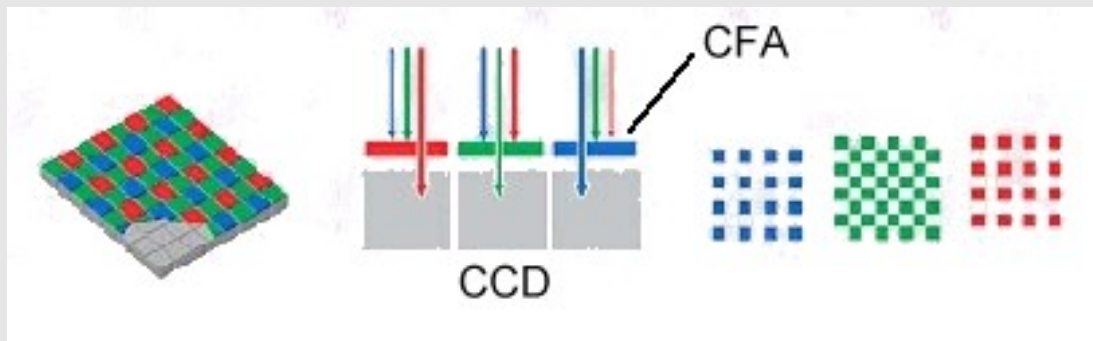
Cadre de l'étude

- **Caméras couleur**

- **Tri-capteurs**

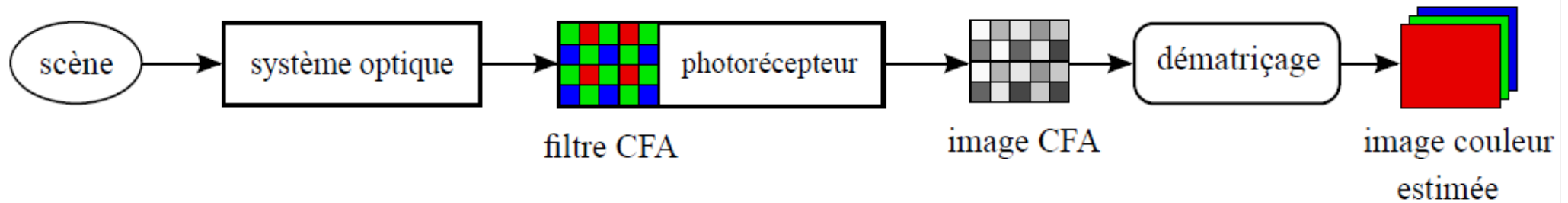


- **Mono-capteur**



Problème

• Dématriçage



CFA de Bayer

$[R_{i,j} \ G_{i,j} \ B_{i,j}]$

$[\alpha_{i,j} \ \beta_{i,j} \ \gamma_{i,j}] = [0 \ 1 \ 0]$

$\alpha_{i,j}R_{i,j} + \beta_{i,j}G_{i,j} + \gamma_{i,j}B_{i,j} = G_{i,j}$

$[\hat{R}_{i,j} \ G_{i,j} \ \hat{B}_{i,j}]$

Dématriçage par reconstruction

- **Formalisation**

- **Image CFA** : $y = \alpha \odot R + \beta \odot G + \gamma \odot B$
- **But du dématriçage** : estimer R , G et B à partir de y

- **Approches**

- Heuristiques
- Guidées par les contours
- Par sélection de fréquences
- **Reconstruction** sous hypothèse *a priori*

L'image cible x admet une représentation parcimonieuse $x = \Psi\xi$

- **Bibliographie**

- **M. Rossi**, “*Color Image Reconstruction via Sparse Signal Representation*”, Ms. dissertation, U. Padova (Italie), 2013
- A. A. Moghadam, M. Aghagolzadeh, M. Kumar, H. Radha, “*Compressive framework for demosaicing of natural images*”, IEEE Trans. on Image Processing, vol. 22, no. 6, 2013

Plan

- **1 – Formalisation et cadre**
 - ➔ Formalisation du problème
 - ➔ Reconstruction par régularisation
- **2 – Dématriçage par représentation parcimonieuse**
 - ➔ Principe général
 - ➔ Construction du dictionnaire
 - ➔ Division en patches et meilleure estimation
- **3 – Résultats**
 - ➔ Protocole expérimental
 - ➔ Résultats de dématriçage
 - ➔ Limites et améliorations

Formalisation du problème

• Image CFA

→ Forme matricielle

$$y = \alpha \odot R + \beta \odot G + \gamma \odot B \quad \text{avec} \quad y, \alpha, R, \dots, B \in \mathbb{R}^{s_1 \times s_2}$$

→ Forme vectorielle

$$y, \alpha, R, \dots, B \in \mathbb{R}^N, N \hat{=} s_1 s_2 \quad (\text{taille de l'image})$$

$$y = [\bar{\alpha} \quad \bar{\beta} \quad \bar{\gamma}] \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad \text{avec} \quad \bar{\alpha} \hat{=} \text{diag}(\alpha), \dots, \bar{\gamma} \hat{=} \text{diag}(\gamma) \in \mathbb{R}^{N \times N}$$

$$= \Phi x \quad \text{avec} \quad \Phi \hat{=} [\bar{\alpha} \quad \bar{\beta} \quad \bar{\gamma}] \in \mathbb{R}^{N \times 3N} \quad \text{et} \quad x \hat{=} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \in \mathbb{R}^{3N}$$

Φ est la **matrice CFA** et x l'**image cible** à reconstruire

Reconstruction par régularisation

• Régularisation

- N équations, $3N$ inconnues → système sous-déterminé
- Résolution par régularisation : x est solution du problème d'optimisation

$$(P_J) : \quad \min_x J(x) \quad \text{tel que} \quad y = \Phi x$$

- Fonction de coût classique $J(x) = \|x\|_2^2$ inadaptée aux images naturelles.

• Approche parcimonieuse

- Forte corrélations inter-pixels et inter-canaux → signal compressible
- Il existe un repère (dictionnaire) $\Psi \in \mathbb{R}^{3N \times M}$, $M \geq 3N$, tel que $x = \Psi \xi$ où
 - soit $\xi \in \mathbb{R}^M$ est k -parcimonieux ($k \ll 3N$ éléments non nuls),
 - soit ξ est compressible ($x \approx \Psi \xi_k$, où $\xi_k \in \mathbb{R}^M$ est k -parcimonieux et contient les $k \ll 3N$ plus grands éléments de ξ).
- Dématriçage : estimation de ξ solution de

$$(P_0) : \quad \min_{\xi} \|\xi\|_0 \quad \text{tel que} \quad y = \Phi \Psi \xi$$

Principe général

Approches existantes

→ Mairal et al. 08, 09 :

- Dictionnaire appris à l'exécution (adaptatif dans 09) → lent
- Seulement CFA de Bayer

→ Moghadam et al. 13

- Dictionnaire appris hors ligne
- Adapté à tout CFA

Étapes

→ 1. Diviser l'image CFA en patchs carrés $y^{(i)}$ avec chevauchement.

→ 2. Reconstruire chaque patch grâce au dictionnaire $\Psi^{(i)}$

- 2.1 Résoudre $(P_1) : \min_{\xi^{(i)}} \|\xi^{(i)}\|_1$ tel que $y^{(i)} = \Phi^{(i)} \Psi^{(i)} \xi^{(i)}$

par poursuite de base pour obtenir $\hat{\xi}^{(i)}$

- 2.2 Estimer le patch cible $\hat{x}^{(i)} = \Psi^{(i)} \hat{\xi}^{(i)}$

→ 3. Choisir la meilleure estimation en chaque pixel puis post-traiter l'image

Construction du dictionnaire (1)

Principe

- **Un dictionnaire par patch (i)** (indice désormais omis) de taille $b \times b$
- **Prise en compte des corrélations**
 - spatiale : décorrélation par DCT 2D
 - spectrale : décorrélation utilisant un repère couleur adapté
- **2 versions du dictionnaire : « séparable » (composantes séparées) ou non**

Dictionnaire « séparable »

→ Décorrélation spatiale

- Soit $D \in \mathbb{R}^{b \times b}$ la matrice DCT 1D et $\tilde{R}, \tilde{G}, \tilde{B}$ les matrices des composantes transformées par DCT 2D.
- Soit $\varphi \hat{=} D^T \otimes D^T$ alors, en notation vectorisée,

$$R = \varphi \tilde{R} \quad , \quad G = \varphi \tilde{G} \quad , \quad B = \varphi \tilde{B} \quad \in \mathbb{R}^n, n = b^2$$

- D'où

$$x = \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} \varphi & 0 & 0 \\ 0 & \varphi & 0 \\ 0 & 0 & \varphi \end{bmatrix} \begin{bmatrix} \tilde{R} \\ \tilde{G} \\ \tilde{B} \end{bmatrix} = (\mathbf{I}_3 \otimes \varphi) \begin{bmatrix} \tilde{R} \\ \tilde{G} \\ \tilde{B} \end{bmatrix}$$

Construction du dictionnaire (2)

• Dictionnaire « séparable » (suite)

→ Réécriture $x = [I_3 \otimes \varphi_{:,1} \cdots I_3 \otimes \varphi_{:,n}]$

$$\begin{bmatrix} \tilde{R}_1 \\ \tilde{G}_1 \\ \tilde{B}_1 \\ \vdots \\ \tilde{R}_n \\ \tilde{G}_n \\ \tilde{B}_n \end{bmatrix}$$

→ Décorrélation spectrale

- Corrélation couleur dans chaque vecteur $[\tilde{R}_k \ \tilde{G}_k \ \tilde{B}_k]^T, k = 1..n$

- On exprime chacun dans le repère $\theta \hat{=} [\theta_{ETF} \ \theta_{YUV}]$

$$\theta_{ETF} = \frac{1}{\sqrt{1+\lambda^2}} \begin{bmatrix} 0 & 0 & 1 & 1 & \lambda & -\lambda \\ 1 & 1 & \lambda & -\lambda & 0 & 0 \\ \lambda & -\lambda & 0 & 0 & 1 & 1 \end{bmatrix}, \theta_{YUV} = \begin{bmatrix} 1 & 0 & 1.14 \\ 1 & -0.39 & -0.58 \\ 1 & 2.03 & 0 \end{bmatrix}$$

par :

$$\begin{bmatrix} \tilde{R}_k \\ \tilde{G}_k \\ \tilde{B}_k \end{bmatrix} = \theta \underbrace{\xi_k}_{\in \mathbb{R}^9} \text{ d'où } x = [I_3 \otimes \varphi_{:,1} \cdots I_3 \otimes \varphi_{:,n}] \begin{bmatrix} \theta & \dots & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \theta \end{bmatrix} \underbrace{\xi}_{\in \mathbb{R}^{9n}}$$

$$= \underbrace{[\theta \otimes \varphi_{:,1} \cdots \theta \otimes \varphi_{:,n}]}_{\Psi \in \mathbb{R}^{3n \times 9n}} \xi$$

Construction du dictionnaire (3)

• Dictionnaire « séparable » (fin)

→ Après permutation des colonnes de Ψ (*atomes*), $x = (\theta \otimes \varphi)\eta$

- φ assure la décorrélation spatiale
- θ assure la décorrélation spectrale

→ **Problème**

- utilise le même θ pour toutes les fréquences spatiales $[\tilde{R}_k \ \tilde{G}_k \ \tilde{B}_k]^T$
- → ne fournit pas une représentation très parcimonieuse pour toutes les fréquences

• Dictionnaire « non séparable »

→ Utilise différents « repères couleur » θ_k

$$\Psi \hat{=} [\theta_1 \otimes \varphi_{:,1} \ \theta_2 \otimes \varphi_{:,2} \ \cdots \ \theta_n \otimes \varphi_{:,n}]$$

→ Appris hors ligne à partir d'une collection de patches

→ Alternative : 1 repère par sous-ensemble de fréquences

Division de l'image CFA en patches

• 2 stratégies

→ **Taille fixe $b \times b$ (ex. $b=8$ ou 16) et chevauchement de l pixels, $l \in [0 \dots b - 1]$**

→ Adaptative

■ Constats

- Petits patches adaptés aux zones de singularités
- Grands patches adaptés aux zones texturées



original

patches 8×8 patches 16×16

- La base DCT 2D est peu adaptée pour représenter des contours diagonaux

■ Proposition

- Division en patches 16×16 avec chevauchement de l pixels
- Détection des patches contenant des détails diagonaux par filtrage passe-haut
- Les patches sans détails diagonaux sont divisés en patches 8×8 se chevauchant

Choix de la meilleure estimation

- **Principe**

- Chevauchement des patches → plusieurs estimations en chaque pixel
- Estimation d'un pixel : moyenne des estimations pondérée par leur qualité

- **Qualité d'une estimation**

- La qualité de l'estimation i est la compressibilité de sa représentation $\hat{\xi}^{(i)}$ estimée par le taux de coefficients d'amplitude inférieure à un seuil

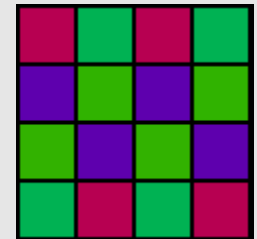
Protocole expérimental

Méthodes comparées

- DL [Zhang et Wu, IEEE IP 05]
- LPA [Paliy et al., Int. J. Imaging Sys. Technol. 07]
- LSSC [Mairal et al., ICCV 09]
- *CD* [Moghadam et al., IEEE IP 13]

Protocole

- Base Kodak
- Plusieurs CFA
 - Bayer CD-B
 - « Second Generation » CD-2G (« optimal », panchromatique)
- Taille des patchs fixe CD-8, CD-16, ou adaptative CD-Ad
- Mesures : PSNR ou visuelle



Influence du CFA



original

CD-B-8

CD-2G-8

Comparaison avec les autres méthodes

- PSNR moyen (dB)**

DL	LPA	LSSC	CD-B-8	CD-B-16	CD-2G-8	CD-2G-16	CD-2G-Ad
40,03	40,47	41,40	38,89	39,03	39,54	39,75	40,94

CD utilise 4 bandes de fréquences spatiales : *DC*, *LF*, *MF* et *HF* avec

30, 10 et 4 atomes pour θ_{DC} , θ_{LF} et θ_{MF} , et $\theta_{HF} = \left[\frac{1}{\sqrt{3}} \quad \frac{1}{\sqrt{3}} \quad \frac{1}{\sqrt{3}} \right]^T$

→ mauvaise représentation de la chrominance en HF

→ erreur d'estimation importante ... mais peu visible

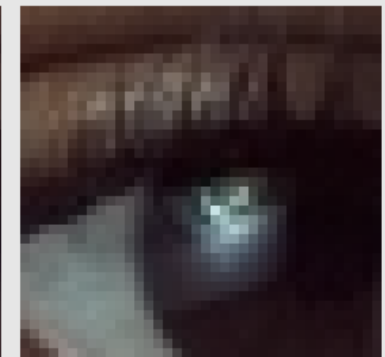
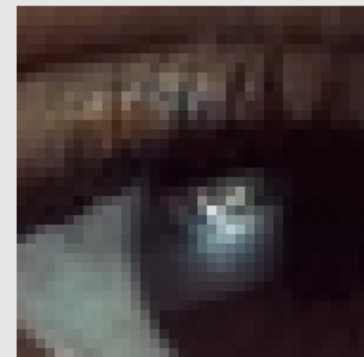
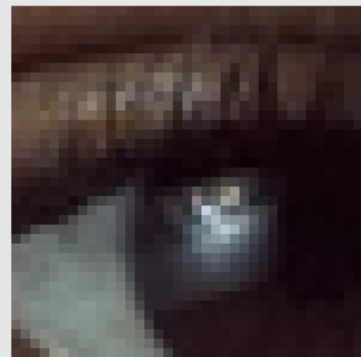
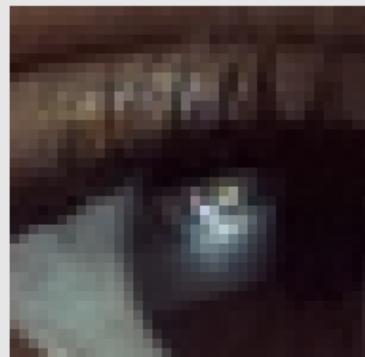
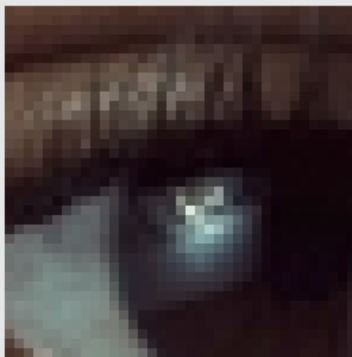
- Visuellement**

39,82 dB

40,09 dB

41,74 dB

39,77 dB



original

DL

LPA

LSSC

CD-2G-Ad

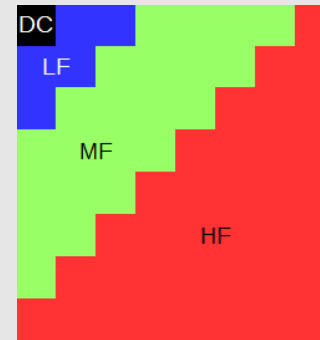
Implémentation de Rossi

• Précisions d'implémentation

→ Uniquement CD-2G-8 et -16

→ Bandes de fréquences

- Patch 8×8 : DC, LF (2), MF (4), HF
- Patch 16×16 : DC, LF (5), MF (9), HF



→ Apprentissage des repères couleur θ_{DC} , θ_{LF} et θ_{MF}

- à partir d'un ensemble de patches aléatoires
- exclusion de l'image à estimer

→ Ajout de la base canonique de \mathbb{R}^3 à θ_{DC}

- Apparition du motif CFA dans les zones saturées
- Zones uniformes surtout caractérisées par θ_{DC}



→ Chevauchement des patches

- 8×8 : $l=6$ pixels, 16×16 : $l=12$ pixels
- → 8 estimations de chaque pixel dans les 2 cas

Résultats de Rossi

original



DL



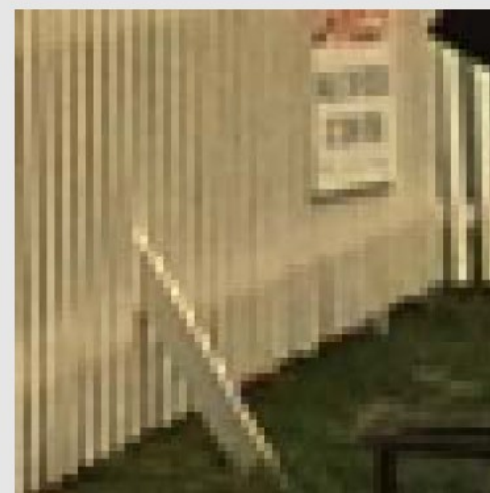
LPA



LSSC



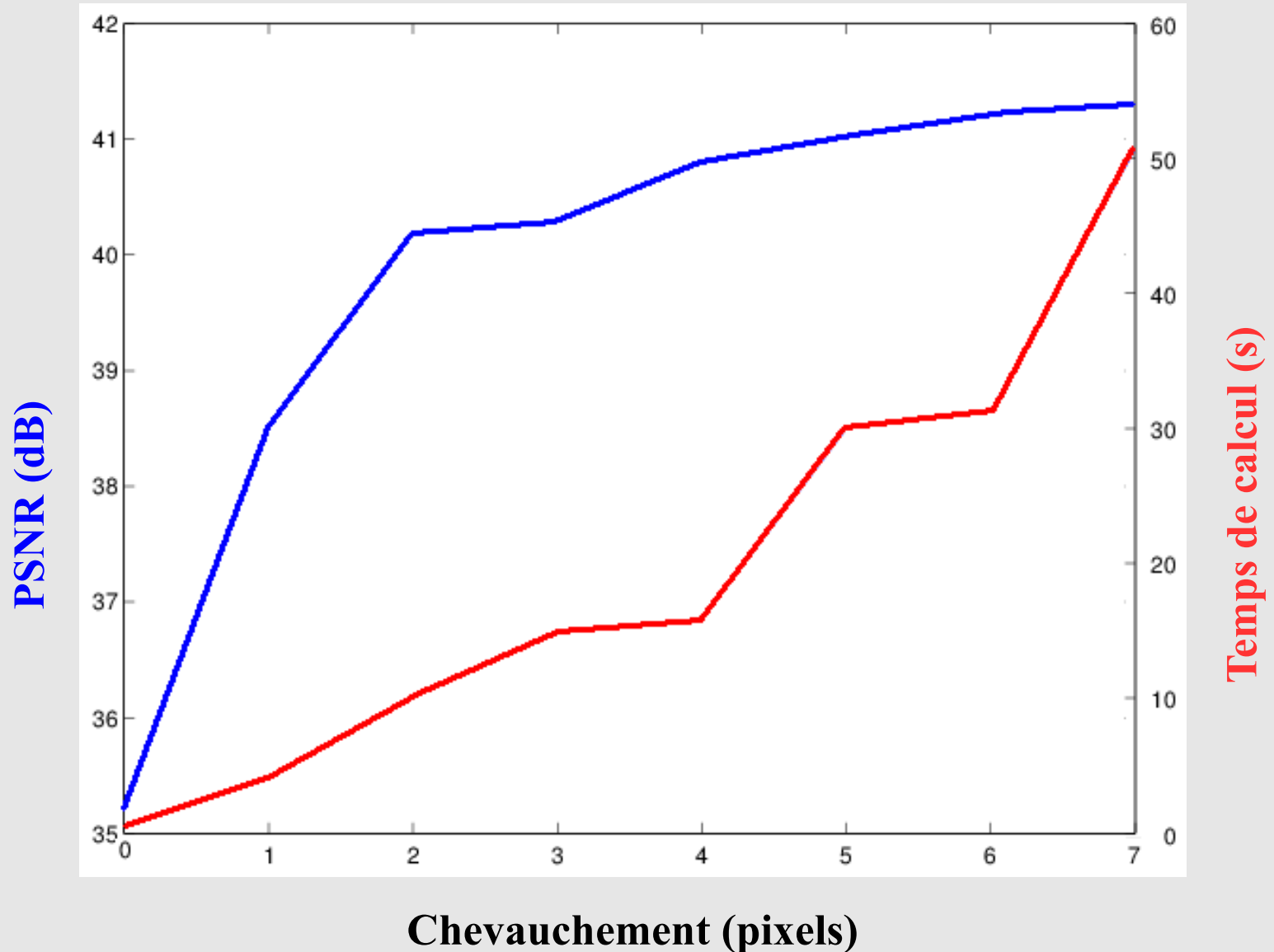
CD-2G-Ad



CD-SG-8 Rossi

Influence du chevauchement des patches

- Image 19
- Patches 8×8



Conclusions et perspectives

• Conclusions

- **La qualité d'estimation dépend principalement du dictionnaire**
- **Apprentissage en ligne efficace mais coûteux**
- **Moghadam et al. 2013**
 - apprentissage hors ligne
 - prise en compte des corrélations spatiale et spectrale
 - résultats (PSNR et visuels) très intéressants

• Améliorations

- **Dictionnaire : compléter, apprendre des repères couleur universels**
- **Stratégie de division adaptative des patches**
- **DCT 2D → transformée en ondelettes**

• Perspectives

- **Tests d'autres CFA**
- **Comparaison avec d'autres méthodes**
- **Extension au dématriçage multispectral**