NL-Means, reprojections et patchs adaptatifs

Groupe de travail (MAP5)
Modélisation numérique et Images

Joseph Salmon, LPMA, Paris Diderot
Collaborateurs : Yann Strozecki (Reprojections)
Charles Deledalle et Vincent Duval (Patchs adaptatifs)

Yann


Charles


Vincent


Plan

Images et bruit

Modèle du bruit additif

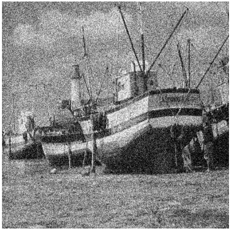


Image observée

Images et bruit

Modèle du bruit additif

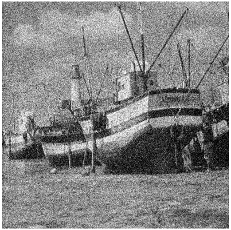


Image observée

$=$

Image idéale

Images et bruit

Modèle du bruit additif

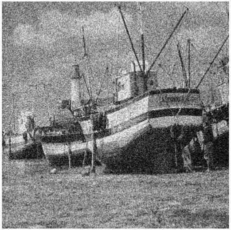
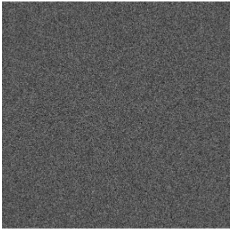


Image observée

$=$

Image idéale

$+$



Bruit

Notation : $\boxed{Y = f + \varepsilon}$

Images, bruit et estimation

f : Image idéale (non disponible) $N \times N$

- Pour un pixel $i = (i_1, i_2) \in \Omega = [\![1, N]\!]^2, f(i) \in \mathbb{R}$

Images, bruit et estimation

f : Image idéale (non disponible) $N \times N$

- Pour un pixel $i = (i_1, i_2) \in \Omega = [\![1, N]\!]^2, f(i) \in \mathbb{R}$

Observation bruitée : modèle de régression

- $Y(i) = f(i) + \varepsilon(i)$
- $\varepsilon(i)$ gaussiennes centrées, i.i.d. de variance σ^2 connue
- modèle le plus simple à étudier

Images, bruit et estimation

f : Image idéale (non disponible) $N \times N$

- Pour un pixel $i = (i_1, i_2) \in \Omega = [\![1, N]\!]^2, f(i) \in \mathbb{R}$

Observation bruitée : modèle de régression

- $Y(i) = f(i) + \varepsilon(i)$
- $\varepsilon(i)$ gaussiennes centrées, i.i.d. de variance $\underline{\sigma^2 \text{ connue}}$
- modèle le plus simple à étudier

Estimation

- Estimer $f(i)$ par $\widehat{f}(i)$ créee à partir de l'observée Y

Images, bruit et estimation

f : Image idéale (non disponible) $N \times N$

- Pour un pixel $i = (i_1, i_2) \in \Omega = [\![1, N]\!]^2, f(i) \in \mathbb{R}$

Observation bruitée : modèle de régression

- $Y(i) = f(i) + \varepsilon(i)$
- $\varepsilon(i)$ gaussiennes centrées, i.i.d. de variance $\underline{\sigma^2 \text{ connue}}$
- modèle le plus simple à étudier

Estimation

- Estimer $f(i)$ par $\widehat{f}(i)$ créee à partir de l'observée Y

Filtrage par moyennes

Cadre général

- Estimer $f(i)$ comme moyenne des pixels bruités

$$\boxed{\widehat{f}(i) = \sum_{k \in \Omega} \lambda_{i,k} Y(k)}$$

- Les poids $\lambda_{i,k}$ peuvent dépendre de la position et/ou des valeurs observées de Y

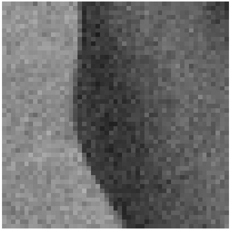
Filtrage par moyennes

Cadre général

- Estimer $f(i)$ comme moyenne des pixels bruités

$$\boxed{\widehat{f}(i) = \sum_{k \in \Omega} \lambda_{i,k} Y(k)}$$

- Les poids $\lambda_{i,k}$ peuvent dépendre de la position et/ou des valeurs observées de Y



Pixel cible i : centre du voisinage

Divers filtres classiques : $\widehat{f}(i) = \sum_k \lambda_{i,k} Y(k)$

Noyau classique Nadaraya [64], Watson [64]

$$\lambda_{i,k} = \frac{K_h(i,k)}{\sum_{k' \in \Omega} K_h(i,k')} \text{ (pas de dépendance en } Y\text{)}$$

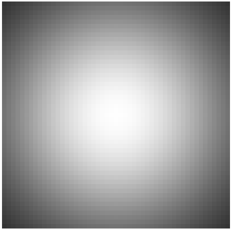
- K : noyau et h : fenêtre ou paramètre de lissage
- Exemple gaussien : $K_h(i,j) = e^{-((i_1-k_1)^2+(i_2-k_2)^2)/2h^2}$

Divers filtres classiques : $\widehat{f}(i) = \sum_k \lambda_{i,k} Y(k)$

Noyau classique Nadaraya [64], Watson [64]

$$\lambda_{i,k} = \frac{K_h(i,k)}{\sum_{k' \in \Omega} K_h(i,k')} \quad \text{(pas de dépendance en } Y\text{)}$$

- K : noyau et h : fenêtre ou paramètre de lissage
- Exemple gaussien : $K_h(i,j) = e^{-((i_1-k_1)^2+(i_2-k_2)^2)/2h^2}$



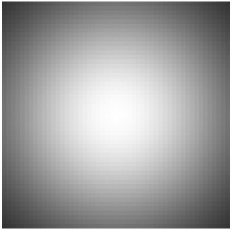
$\lambda_{i,k}$: Spatial

Divers filtres classiques : $\widehat{f}(i) = \sum_k \lambda_{i,k} Y(k)$

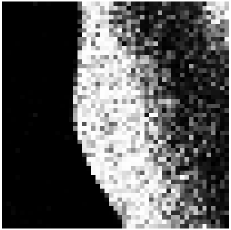
Filtre de Yaroslavsky Yaroslavsky [85], Lee [83]

$$\lambda_{i,k} = \frac{L_g(Y(i), Y(k))}{\sum_{k' \in \Omega} L_g(Y(i), Y(k'))} \quad \text{(dépendance en } Y\text{)}$$

- Utilise la proximité photométrique uniquement
- L : noyau et g : fenêtre ou paramètre de lissage



$\lambda_{i,k}$: Spatial



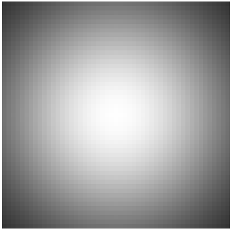
$\lambda_{i,k}$: Yaroslavsky

Divers filtres classiques : $\widehat{f}(i) = \sum_k \lambda_{i,k} Y(k)$

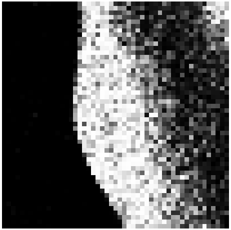
Filtre Bilatéral Tomasi et Manduchi [98]

$$\lambda_{i,k} = \frac{K_h(i,k) L_g(Y(i), Y(k))}{\sum_{k' \in \Omega} K_h(i,k') L_g(Y(i), Y(k'))}$$

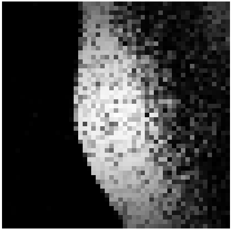
- Utilise la proximité spatiale et photométrique
- K, L : noyaux ; h, g : fenêtres ou paramètres de lissage



$\lambda_{i,k}$: Spatial



$\lambda_{i,k}$: Yaroslavsky



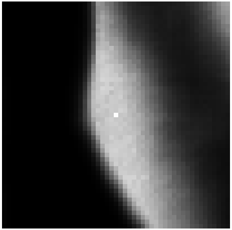
$\lambda_{i,k}$: Bilatéral

Divers filtres classiques : $\widehat{f}(i) = \sum_k \lambda_{i,k} Y(k)$

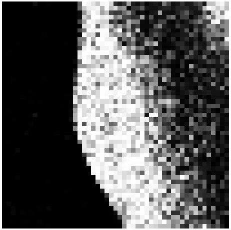
Filtre Bilatéral Tomasi et Manduchi [98]

$$\lambda_{i,k} = \frac{K_h(i,k) L_g(Y(i), Y(k))}{\sum_{k' \in \Omega} K_h(i,k') L_g(Y(i), Y(k'))}$$

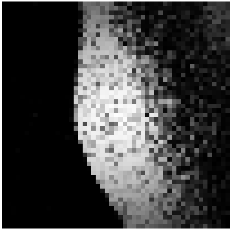
- Utilise la proximité spatiale et photométrique
- K, L : noyaux ; h, g : fenêtres ou paramètres de lissage



$\lambda_{i,k}$: NL-Means



$\lambda_{i,k}$: Yaroslavsky



$\lambda_{i,k}$: Bilatéral

Méthodes à patchs et redondance

Poids tenant compte de la similarité entre patchs

Introduction des patchs : Synthèse de texture Efros et Leung [99] , Débruitage Buades et al. [05], Awate et Whitaker [06], Kervrann et Boulanger [06]

- Patch à débruiter
- Patchs similaires : poids importants
- Patchs peu similaires : poids faibles
- Patchs très différents : poids quasi nuls

Méthodes à patchs et redondance

Poids tenant compte de la similarité entre patchs

Introduction des patchs : Synthèse de texture Efros et Leung [99] , Débruitage Buades et al. [05], Awate et Whitaker [06], Kervrann et Boulanger [06]

- Patch à débruiter
- Patchs similaires : poids importants
- Patchs peu similaires : poids faibles
- Patchs très différents : poids quasi nuls

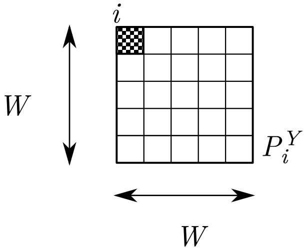
Définition des patchs

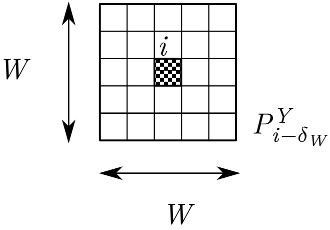
Patchs

- Patch P_i^Y : Sous-image de Y, dont i est le coin haut-gauche
- W : largeur des patchs, en général fixée une fois pour toute

$$P_i^Y(j) = Y(i_1 + j_1, i_2 + j_2) \text{ avec } 0 \leq j_1, j_2 \leq W - 1$$

- Dans le cas W impair, i est le centre du patch $P_{i-\delta_W}^Y$ avec $\delta_W = (\frac{W-1}{2}, \frac{W-1}{2})$: $Y(i) = P_{i-\delta_W}^Y(\delta_W)$





Définition des Non-Local Means (NL-Means)

Méthodes à noyau et à patchs

- Estimation des patchs par moyenne :

$$\widehat{P_i} = \sum_{k \in \Omega} \lambda_{i,k} P_k^Y$$

- Reprojection centrale (W impair)

$$\widehat{f}(i) = \widehat{P}_{i-\delta_W}(\delta_W)$$

NL-Means Buades et al. [05]

- Choisir des poids tenant compte de la similarité des patchs :

$$\lambda_{i,k} = \frac{K_h(P_i^Y, P_k^Y)}{\sum_{k' \in \Omega} K_h(P_i^Y, P_{k'}^Y)}$$

- Noyau gaussien : $K_h(P, Q) = \exp(-\|P - Q\|_2^2 / 2h^2)$
- Noyau plat : $K_h(P, Q) = \mathbb{1}_{\|P-Q\|_2 < h/2}$

Définition des Non-Local Means (NL-Means)

Méthodes à noyau et à patchs

- Estimation des patchs par moyenne :

$$\widehat{P_i} = \sum_{k \in \Omega} \lambda_{i,k} P_k^Y$$

- Reprojection centrale (W impair)

$$\widehat{f}(i) = \widehat{P}_{i-\delta_W}(\delta_W)$$

NL-Means Buades et al. [05]

- Choisir des poids tenant compte de la similarité des patchs :

$$\lambda_{i,k} = \frac{K_h(P_i^Y, P_k^Y)}{\sum_{k' \in \Omega} K_h(P_i^Y, P_{k'}^Y)}$$

- Noyau gaussien : $K_h(P, Q) = \exp(-\|P - Q\|_2^2 / 2h^2)$
- Noyau plat : $K_h(P, Q) = \mathbb{1}_{\|P-Q\|_2 < h/2}$

Autour des NL-Means

Variations

- ACP et patchs Azzabou et al. [07], Tasdizen [09]
- Itérer le processus Brox et al. [08]
- NL-Means + variation totale (TV) Louchet et Moisan [10]
- BM3D Dabov et al. [07,09]
- Apprentissage de dictionnaire Mairal et al. [08,09]

Autour des NL-Means

Variations

- ACP et patchs Azzabou et al. [07], Tasdizen [09]
- Itérer le processus Brox et al. [08]
- NL-Means + variation totale (TV) Louchet et Moisan [10]
- BM3D Dabov et al. [07,09]
- Apprentissage de dictionnaire Mairal et al. [08,09]

Interprétation de la méthode

- Méthode à Noyau Buades [06] , sur des graphes Peyre [08] ou sur des variétés Tschumperle et Brun [08]
- Minimisation d'énergies Gilboa et Osher [07], Brox et al. [08]
- Point de vue bayésien Kervrann et al. [07] ou PAC-bayésien S. et Le Pennec [09]

Autour des NL-Means

Variations

- ACP et patchs Azzabou et al. [07], Tasdizen [09]
- Itérer le processus Brox et al. [08]
- NL-Means + variation totale (TV) Louchet et Moisan [10]
- BM3D Dabov et al. [07,09]
- Apprentissage de dictionnaire Mairal et al. [08,09]

Interprétation de la méthode

- Méthode à Noyau Buades [06] , sur des graphes Peyre [08] ou sur des variétés Tschumperle et Brun [08]
- Minimisation d'énergies Gilboa et Osher [07], Brox et al. [08]
- Point de vue bayésien Kervrann et al. [07] ou PAC-bayésien S. et Le Pennec [09]

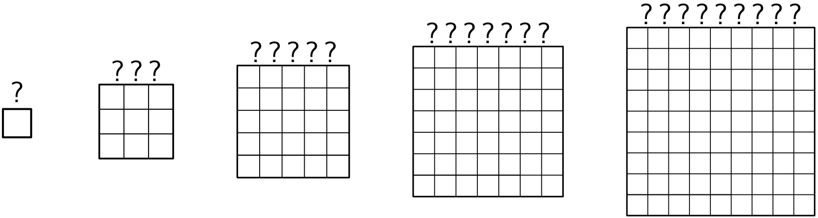
Les trois paramètres des NL-Means

$$\boxed{\widehat{f}(i) = \sum_{k \in \Omega_R(i)} \frac{e^{-d^2(P_i^Y, P_k^Y)/2h^2}}{\sum_{k' \in \Omega_R(i)} e^{-d^2(P_i^Y, P_{k'}^Y)/2h^2}} Y(k)}$$

Les trois paramètres

- la taille des patchs : W (entre 5 et 10)
- la taille de la zone de recherche : R (entre 9 et 21)
- le paramètre de régularisation : h (proportionnel à σ)

La tailles des patchs : W



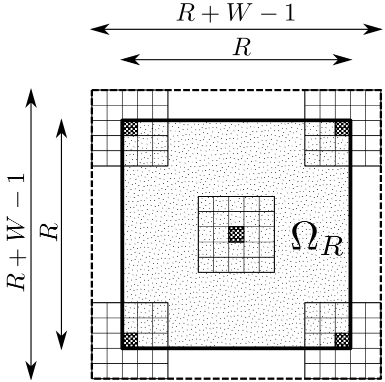
Peu d'approches étudiées

- Apprentissage de patchs : mélanger localement deux tailles de patch Mairal et al. [08] avec des SVM
- BM3D avec des formes adaptées en chaque pixel Dabov et al. [09] grâce à la règle de Lespki [90]

Zone de recherche $\Omega \to \Omega_R(i)$

$$\boxed{\widehat{f}(i) = \sum_{k \in \Omega_R(i)} \frac{e^{-d^2(P_i^Y, P_k^Y)/2h^2}}{\sum_{k' \in \Omega_R(i)} e^{-d^2(P_i^Y, P_{k'}^Y)/2h^2}} Y(k)}$$

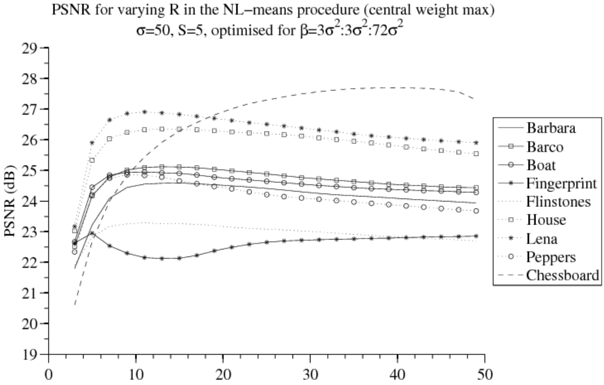
Zone de recherche



- Recherche de patchs dans l'image entière trop long, $O(N^4)$: 1500s (image entière)
- Restriction à une petite zone Ω_R de taille $R \times R$: 10s ($R = 21$)

Zone de recherche et R (II)

Non Local à « portée limitée » Tasdizen [09],S.[10]



Rem : PSNR élevé $\Longleftrightarrow$ erreur (quadratique) faible

- Optimisation locale de la taille de la zone de recherche Kervrann et Boulanger [06]

Quand le paramètre de lissage augmente...

Le choix du paramètre de lissage : h

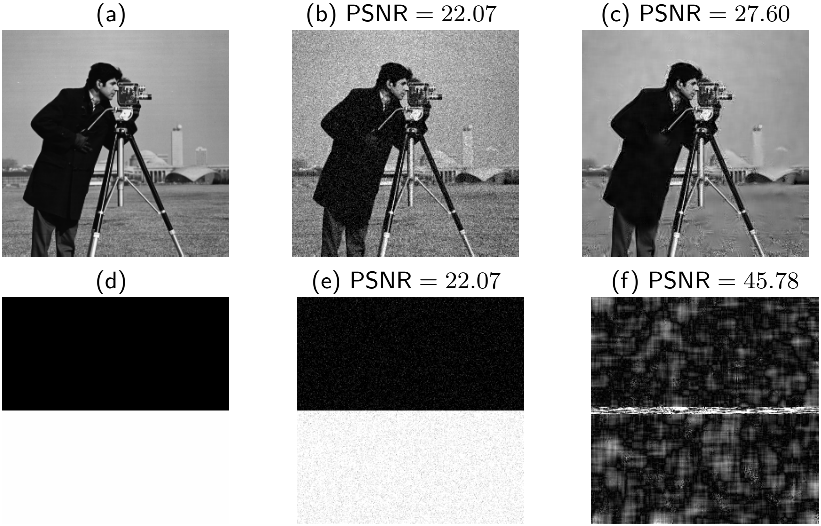
Beaucoup de travaux...

- Choix global par utilisation de SURE (Stein Unbiased Risk Estimate) Van de Ville et Kocher [09]
- Choix local par utilisation de SURE Duval et al. [10]

Point de vue des tests (notamment pour le noyau plat)

- $\|P_i^Y - P_j^Y\|^2/(2\sigma^2) \sim \chi^2(W^2)$ si $P_i^f = P_j^f$ et si les patchs ne se recouvrent pas
- Considérer les quantiles du χ^2 pour choisir h Kervrann et Boulanger [06]

Reprojection centrale en pratique



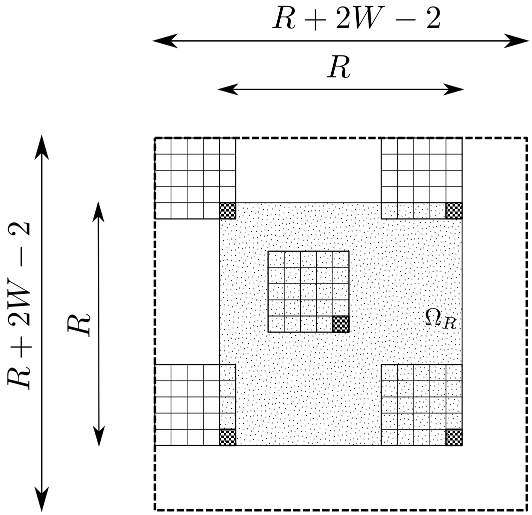
(a,d) Images originales, (b,e) Images bruitées ($\sigma = 20$), (c) Image débruitée, (f) différence avec l'originale, $R = 9$ (haut), $R = 21$ (bas), $W = 9, h^2 = 2\sigma^2 q_{0.99}^{W^2}$.

Artefacts le long des arrêtes

Artefacts le long des arrêtes

Zone de recherche et patchs décentrés

Largeur des patchs W=5, zone de recherche R=15



Reprojection par moyenne uniforme

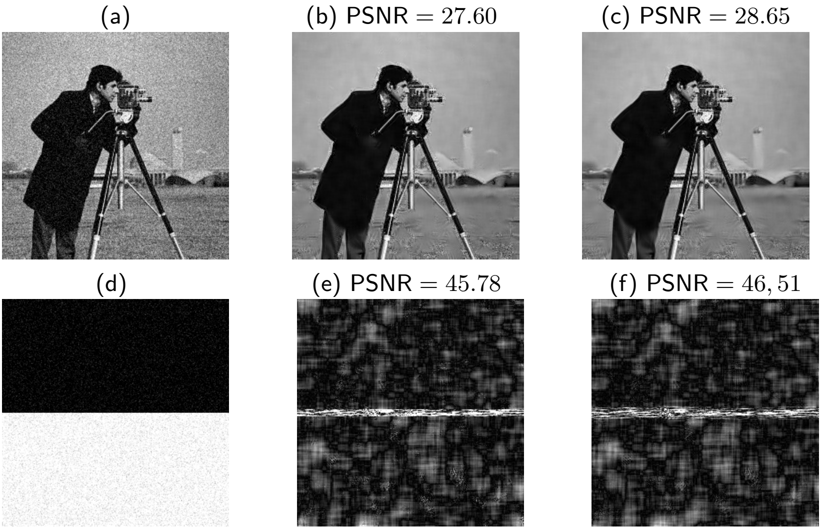
Chaque pixel appartient à W^2 patchs (configurations) : mélanger–agréger–mixer tous ces estimateurs

Cas uniforme Buades et al. [05], Kervrann et al. [07]

$$\widehat{f}^{\mathrm{Unif}}(i) = \frac{1}{W^2} \sum_{\delta \in [\![0, W-1]\!]^2} \widehat{P_{i-\delta}}(\delta)$$

- W peut être pair
- Gains numériques importants

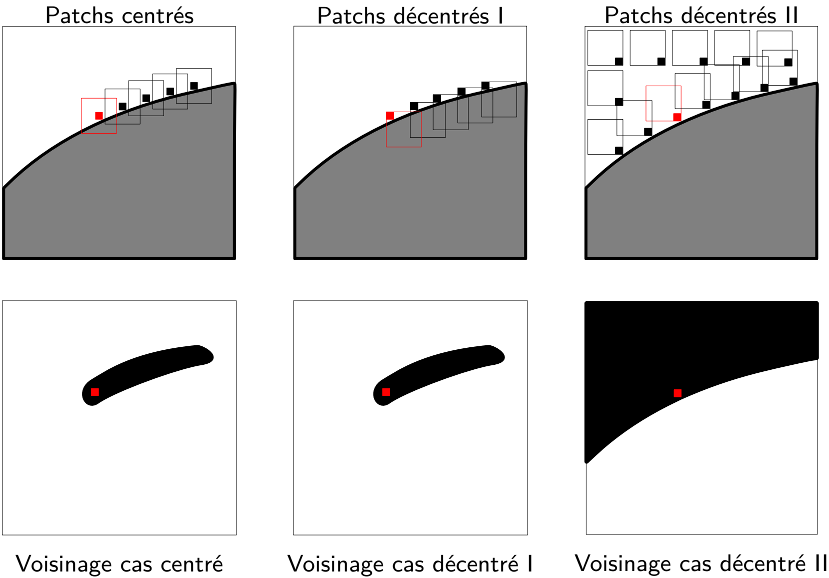
Reprojection moyenne en pratique



(a-d) Images bruitées ($\sigma = 20$), (b-e) Reprojection centrale, (c-f) Reprojection Moyenne, $R = 9$ (haut), $R = 21$ (bas), $W = 9, h^2 = 2\sigma^2 q_{0.99}^{W^2}$.

Artefacts le long des arrêtes (le retour)

Le décentrage peut payer



Reprojection par minimisation de la variance

- Problème : variance trop grande le long des bords
- Idée naïve : supposer le biais négligeable et sélectionner l'estimateur de variance estimée minimale

Variance Minimale

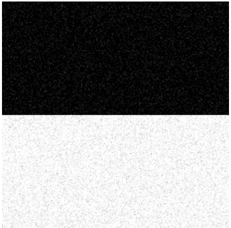
$$\widehat{f}_{\mathrm{Min}}(i) = \widehat{P}_{i-\hat{\delta}}(\hat{\delta}))$$

$$\text{avec} \qquad \hat{\delta} = \underset{\delta \in [\![0, W-1]\!]^2}{\arg\min} \operatorname{Var}\left(\widehat{P}_{i-\delta}(\delta)\right)$$

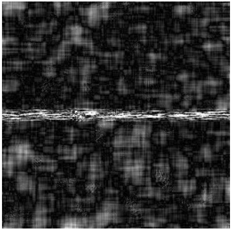
- Noyau plat : choisir la configuration fournissant le plus de voisins

Reprojection par minimisation de la variance

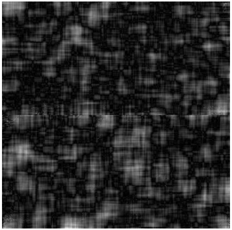
(a)



(b) PSNR $= 46,51$



(c) PSNR $= 48.10$



(d)

(e) PSNR $= 28.65$

(f) PSNR $= 27.09$

(a-d) Images bruitées ($\sigma = 20$), (b-e) Reprojection moyenne, (c-f) Reprojection Minimisant la variance, $R = 9$ (haut), $R = 21$ (bas), $W = 9, h^2 = 2\sigma^2 q_{0.99}^{W^2}$.

Sélection trop brusque : bords crénelés

Agréger les estimateurs selon leur variance
S. et Strozecki [10]

- Supposer le biais négligeable
- Choisir la **combinaison convexe** de variance minimale

Moyenne pondérée par la variance (Wav)

$$\widehat{f}_{\mathrm{Wav}}(i) = \sum_{\delta \in [\![0, W-1]\!]^2} \alpha^\star_\delta \widehat{P_{i-\delta}}(\delta)$$

où $\quad (\alpha^\star_\delta)_{\delta \in [\![0, W-1]\!]^2} = \underset{\alpha \in \mathbb{R}^{W^2}}{\arg\min} \operatorname{Var}\left(\sum_{\delta \in [\![0, W-1]\!]^2} \alpha_\delta \widehat{P_{i-\delta}}(\delta) \right)$

sous la contrainte $\quad \sum_{\delta \in [\![0, W-1]\!]^2} \alpha_\delta = 1$

Rem : contrainte obligatoire sinon solution dégénérée

Reprojection pondérée selon la variance

Solution sous l'hypothèse de décorrélation des estimateurs :

$$\boxed{\alpha_\delta^\star \propto \left[\operatorname{Var}\left(\widehat{P_{i-\delta}}(\delta)\right)\right]^{-1}}$$

- Idée apparue sous le nom de *Stacked Generalization*, *Stacked Regression* ou *Generalized Ensemble Methods* Wolpert [92], Breiman [96],Perrone [93] pour combiner des estimateurs non biaisés
- Cas des méthodes à noyau Goldenshluger et Nemirovski [97], Katkovnik et al. [04], Foi [05] limiter les erreurs de sélection de la méthode de Lepski

- Noyau plat : mélanger les configurations en pondérant par le nombre de candidats trouvés

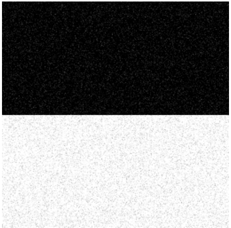
Gains numériques et visuels

(a)

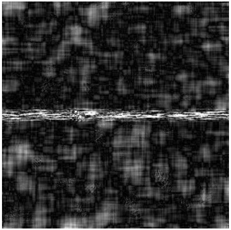
(b) PSNR $= 28.65$

(c) PSNR $= 29.08$

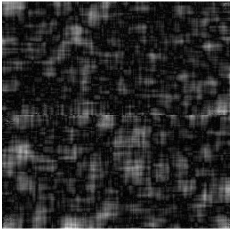
(d)



(e) PSNR $= 46,51$



(f) PSNR $= 47,77$



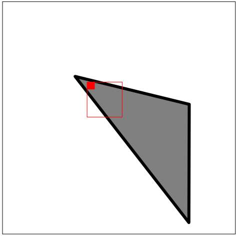
(a,d) Images bruitées ($\sigma = 20$), (b,e) Reprojection Moyenne, (c,f) Reprojection Wav, $R = 9$ (haut), $R = 21$ (bas), $W = 9, h^2 = 2\sigma^2 q_{0.99}^{W^2}$

Avant

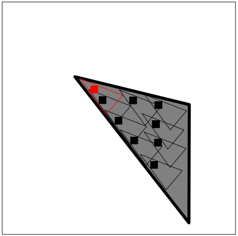
Après : Wav S. et Strozecki [10]

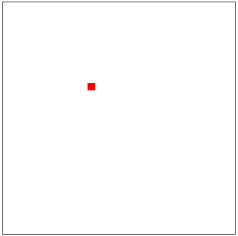
Intérêt des formes

Carré (bien) décentré

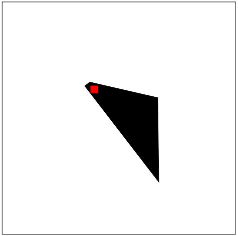


Forme adaptée





Voisinage, carré décentré

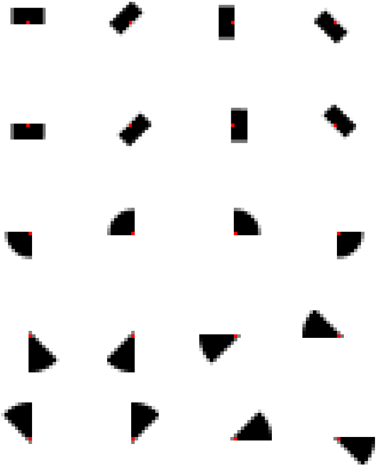


Voisinage, forme adaptée

Des patch carrés aux formes générales

Reprojection : peut etre vu comme l'utilisation de carrés décentrés.

Généralisation : utiliser des formes/tailles plus variées



Forme S et mesure de ressemblance

NL-Means revisitée : $$\boxed{\widehat{f}_{\boldsymbol{S}}(i) = \sum_{k \in \Omega_R(i)} \frac{e^{-d_{\boldsymbol{S}}^2(i,k)^2/2h^2}}{\displaystyle\sum_{k' \in \Omega_R(i)} e^{-d_{\boldsymbol{S}}^2(i,k')^2/2h^2}} Y(k)}$$

$$d_{\boldsymbol{S}}^2(i,k) = \sum_{\tau \in \Omega} \boldsymbol{S}(\tau)(Y(i+\tau) - Y(k+\tau))^2 \,,$$

Exemples classiques

Carré : $$\boldsymbol{S}(\tau) = \begin{cases} 1, & \text{si } \|\tau\|_\infty \leq \frac{W-1}{2} \\ 0, & \text{sinon,} \end{cases}$$

Gaussienne : $$\boldsymbol{S}(\tau) = \begin{cases} \exp(-(\tau_1^2 + \tau_2^2)/2a^2), & \text{si } \|\tau\|_\infty \leq \frac{W-1}{2} \\ 0, & \text{sinon.} \end{cases}$$

FFT

Calcul de la distance par “formes”

$$d^2_{\boldsymbol{S}}(i, i+\delta) = \sum_{\tau \in \Omega} \boldsymbol{S}(\tau) \left(Y(i+\tau) - Y(i+\delta+\tau) \right)^2$$

$$d^2_{\boldsymbol{S}}(i, i+\delta) = (\check{\boldsymbol{S}} \star \boldsymbol{\Delta}_\delta)(i)\,,$$

où $\check{\boldsymbol{S}}(\tau) = \boldsymbol{S}(-\tau)$, $\boldsymbol{\Delta}_\delta(i) = \left(Y(i) - Y(i+\delta) \right)^2$, $\star$: convolution discrète.

$$\check{\boldsymbol{S}} \star \boldsymbol{\Delta}_\delta = \mathcal{F}^{-1}(\mathcal{F}(\check{\boldsymbol{S}})\mathcal{F}(\boldsymbol{\Delta}_\delta)) = \mathcal{F}^{-1}(\overline{\mathcal{F}(\boldsymbol{S})}\mathcal{F}(\boldsymbol{\Delta}_\delta))\,,$$

$\mathcal{F}$: FFT-2D et $\mathcal{F}^{-1}$ la transformation inverse
Complexité de l’algorithme $O(|W| \cdot |\Omega| \cdot \log(|\Omega|))$.

Algorithme

Initialiser $\mathbf{A}$ (numérateur) et $\mathbf{B}$ (dénominateur) à zéro
pour tout les vecteurs de translation δ avec $\|\delta\|_\infty \leq R$ **faire**
Calculer l'image des distances au carré $\boldsymbol{\Delta}_\delta$

$$\boldsymbol{\Delta}_\delta(i) := (\,Y(i) - \,Y(i+\delta))^2 \text{ pour tous les pixels} i$$

Calculer les FFT-2D $\mathcal{F}(\boldsymbol{\Delta}_\delta)$ et $\mathcal{F}(\check{\boldsymbol{S}})$
Calculer $d^2_{\boldsymbol{S}}(\cdot, \cdot + \delta) \leftarrow \left(\mathcal{F}^{-1}\left(\overline{\mathcal{F}(\boldsymbol{S})}\mathcal{F}(\boldsymbol{\Delta}_\delta)\right)\right)(\cdot)$
pour tout pixels i in Ω **faire**

Calculer les poids $\qquad \lambda_{i,i+\delta} = \exp\left(-d^2_{\boldsymbol{S}}(i, i+\delta)/2h^2\right)$

Mettre à jour $\qquad \mathbf{A}(i) \leftarrow \mathbf{A}(i) + \lambda_{i,i+\delta}\, Y(i+\delta)$

$\qquad \mathbf{B}(i) \leftarrow \mathbf{B}(i) + \lambda_{i,i+\delta}$

fin pour
fin pour
Estimateur (normalisé) $\quad \hat{f}(i) = \frac{\mathbf{A}(i)}{\mathbf{B}(i)}$ pour tous les pixels i

Limites d'une seule forme

Mélange - Choix des formes

On dispose initialament de K formes $S_1, \cdots, S_K$
Utiliser des estimateurs (sans biais) du risque $\hat{r}_{\boldsymbol{S}_1}, \cdots, \hat{r}_{\boldsymbol{S}_K}$ pour
Agréger/sélectionner $\hat{f}_{S_1}, \cdots, \hat{f}_{S_K}$

Procédures prossibles

- Moyenne uniforme $\hat{f}^{\mathrm{Unif}}(i) = \frac{1}{K}\sum_{k=1}^{K}\hat{f}_{\boldsymbol{S}}(i)$
- Minimiseur du risque empirique : $\hat{f}^{*}(i) = \hat{f}_{\boldsymbol{S}^*}(i)$, avec $r_{\boldsymbol{S}^*} = \arg\min_{k=1,\cdots,K}\hat{r}_{\boldsymbol{S}_k}$
- Agrégation à poids exponentiels : pour un certain β :

$$\hat{f}^{exp}(i) = \frac{\sum_{k=1}^{K}\exp(-\hat{r}_{\boldsymbol{S}_k}/\beta)\hat{f}_{\boldsymbol{S}_k}(i)}{\sum_{k'=1}^{K}\exp(-\hat{r}_{\boldsymbol{S}_{k'}}/\beta)}$$

Rem sur β (souvent appelé température)

Si $\beta \to 0$, alors $\hat{f}^{exp}(i) \to \hat{f}^{*}(i)$

Si $\beta \to \infty$, alors $\hat{f}^{exp}(i) \to \hat{f}^{\mathrm{Unif}}(i)$

Estimation sans biais du risque (SURE)

Lemme de Stein (cas gaussien) Stein [81]

Si $\hat{f}$ est C^1 alors on a la formule suivante

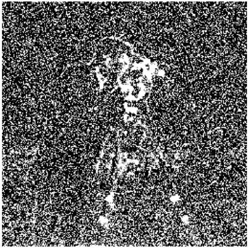
$$\mathbb{E}(\hat{f}_{\boldsymbol{S}}(i) - f(i))^2 = \mathbb{E}\Big(\underbrace{(\hat{f}_{\boldsymbol{S}}(i) - Y(i))^2 + 2\sigma^2 \frac{\partial \hat{f}_{\boldsymbol{S}}(i)}{\partial \boldsymbol{\varepsilon}(i)} - \sigma^2}_{\hat{r}_{\boldsymbol{S}}} \Big).$$

$\hat{r}_{\boldsymbol{S}}$: estimateur sans biais du risque local

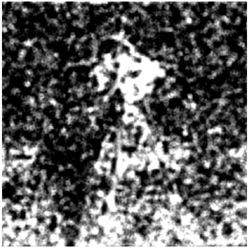
Rem : Pour l'estimateur NL-Means formule fermée pour $\frac{\partial \hat{f}_{\boldsymbol{S}}(i)}{\partial \boldsymbol{\varepsilon}(i)}$ fournie par Van de Ville et Kocher [09],Duval et al. 2010

« Carte » des risques

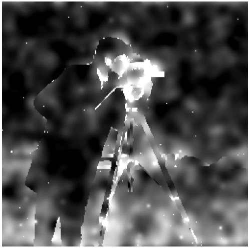
(a) Original



(b) Risque bruité

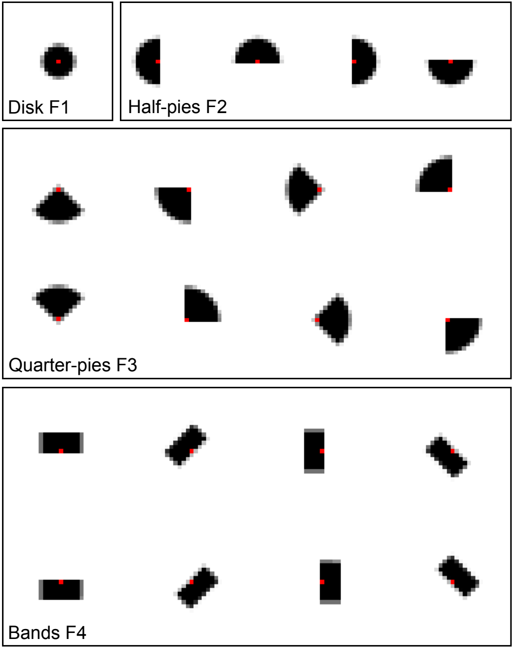


(c) Risque régularisé par convolution



(d) Risque régularisé par diffusion anisotrope

Famille utilisée



Résultats : ($\sigma = 20$)

(e) NL Means

(f) Reprojection Uniforme

(g) NL-Shape

(h) BM3D

Résultats : PSNR/SSIM

	Cameraman	City	Windmill
Noisy ($\sigma = 5$)	34.17/0.845	34.17/0.901	34.17/0.827
NLM [BCM05]	36.92/0.951	35.87/0.965	38.10/0.972
UWA NLM [BCM05]	37.19/0.953	36.19/0.967	38.50/0.974
SURE NLM [DAG10]	37.46/0.956	36.76/0.975	39.14/0.978
BM3D [DFKE07]	**38.17**/**0.962**	**37.48**/**0.978**	**39.91**/**0.983**
NLM-SAP	37.80/0.957	37.26/0.975	39.60/0.979
Noisy ($\sigma = 10$)	28.15/0.633	28.15/0.757	28.15/0.607
NL-Means [BCM05]	32.46/0.905	31.11/0.932	33.62/0.945
UWA NLM [BCM05]	32.80/0.908	31.50/0.934	34.06/0.946
SURE NLM [DAG10]	33.11/0.918	32.11/0.948	34.78/0.954
BM3D [DFKE07]	**34.06**/**0.931**	**33.15**/**0.956**	**35.84**/**0.966**
NLM-SAP	33.44/0.914	32.84/0.950	35.28/0.955
Noisy ($\sigma = 20$)	22.13/0.400	22.13/0.567	22.13/0.385
NLM [BCM05]	28.72/0.820	27.11/0.870	30.04/0.897
UWA NLM [BCM05]	28.89/0.822	27.34/0.872	30.17/0.899
SURE NLM [DAG10]	29.49/0.845	27.85/0.889	30.96/0.906
BM3D [DFKE07]	**30.35**/**0.871**	**29.07**/**0.912**	**32.07**/**0.936**
NLM-SAP	29.50/0.833	28.21/0.887	31.11/0.907

Conclusion

Contributions autour des NL-Means

- Prise en compte de meilleurs forme que les patchs centrés
- Atténuation de certains artefacts
- Amélioration en terme de PSNR/SSIM

Ouvertures possibles et travaux en cours

- Résultats théoriques non asymptotiques : Inégalités d'Oracles
- Liens avec les travaux sur l'apprentissage de patchs
- Traitement d'autres modèles de bruits : modèle de Poisson notamment

http ://people.math.jussieu.fr/ $\sim$ salmon/

Conclusion

Contributions autour des NL-Means

- Prise en compte de meilleurs forme que les patchs centrés
- Atténuation de certains artefacts
- Amélioration en terme de PSNR/SSIM

Ouvertures possibles et travaux en cours

- Résultats théoriques non asymptotiques : Inégalités d'Oracles
- Liens avec les travaux sur l'apprentissage de patchs
- Traitement d'autres modèles de bruits : modèle de Poisson notamment

http ://people.math.jussieu.fr/ $\sim$ salmon/

Références I

- N. Azzabou, N. Paragios, and F. Guichard.
  Image denoising based on adapted dictionary computation.
  In *ICIP*, pages 109–112, 2007.
- S. P. Awate and R. T. Whitaker.
  Unsupervised, information-theoretic, adaptive image filtering for image restoration.
  IEEE Trans. Pattern Anal. Mach. Intell., 28(3) :364–376, 2006.
- A. Buades, B. Coll, and J-M. Morel.
  A review of image denoising algorithms, with a new one.
  Multiscale Model. Simul., 4(2) :490–530, 2005.
- T. Brox, O. Kleinschmidt, and D. Cremers.
  Efficient nonlocal means for denoising of textural patterns.
  IEEE Trans. Image Process., 17(7) :1083–1092, 2008.
- L. Breiman.
  Stacked regressions.
  Mach. Learn., 24(1) :49–64, 1996.

Références II

- A. Buades.
  Image and movie denoising by non local means.
  PhD thesis, Universitat de les Illes Balears, 2006.
- V. Duval, J-F. Aujol, and Y. Gousseau.
  On the parameter choice for the non-local means.
  Technical Report hal-00468856, HAL, Mars 2010.
- K. Dabov, A. Foi, V. Katkovnik, and K. O. Egiazarian.
  Image denoising by sparse 3-D transform-domain collaborative filtering.
  IEEE Trans. Image Process., 16(8) :2080–2095, 2007.
- K. Dabov, A. Foi, V. Katkovnik, and K. O. Egiazarian.
  BM3D image denoising with shape-adaptive principal component analysis.
  In *Proc. Workshop on Signal Processing with Adaptive Sparse Structured Representations (SPARS'09)*, 2009.
- A. A. Efros and T.K. Leung.
  Texture synthesis by non-parametric sampling.
  In *ICCV*, pages 1033–1038, 1999.

Références III

- A. Foi.

  Anisotropic nonparametric image processing : theory, algorithms and applications.

  PhD thesis, Politecnico di Milano, 2005.

- A. Goldenshluger and A. S. Nemirovski.

  On spatially adaptive estimation of nonparametric regression.

  Math. Methods Statist., 6(2) :135–170, 1997.

- G. Gilboa and S. Osher.

  Nonlocal linear image regularization and supervised segmentation.

  Multiscale Model. Simul., 6(2) :595–630, 2007.

- Ch. Kervrann and J. Boulanger.

  Optimal spatial adaptation for patch-based image denoising.

  IEEE Trans. Image Process., 15(10) :2866–2878, 2006.

Références IV

- Ch. Kervrann, J. Boulanger, and P. Coupé.
  Bayesian non-local means filter, image redundancy and adaptive dictionaries for noise removal.
  In *SSVM*, volume 4485, pages 520–532, 2007.
- V. Katkovnik, A. Foi, K. O. Egiazarian, and J. T. Astola.
  Directional varying scale approximations for anisotropic signal processing.
  In *EUSIPCO*, pages 101–104, 2004.
- J-S. Lee.
  Digital image smoothing and the sigma filter.
  Computer Vision, Graphics, and Image Processing, 24(2) :255–269, 1983.
- O. V. Lepski.
  On a problem of adaptive estimation in gaussian white noise.
  Theory of Probability and its Applications, 35(3) :454–466, 1990.
- C. Louchet and L. Moisan.
  Total variation as a local filter.
  To appear, 2010.

Références V

- J. Mairal, G. Sapiro, and M. Elad.
  Learning multiscale sparse representations for image and video restoration.
  Multiscale Model. Simul., 7(1) :214–241, 2008.

- E. A. Nadaraya.
  On estimating regression.
  Theory of Probability and its Applications, 9(1) :141–142, 1964.

- M. P. Perrone.
  Improving regression estimation : Averaging methods for variance reduction with extensions to general convex measure optimization.
  PhD thesis, Brown University, Providence, RI, USA, 1993.

- G. Peyré.
  Image processing with nonlocal spectral bases.
  Multiscale Model. Simul., 7(2) :703–730, 2008.

- J. Salmon.
  On two parameters for denoising with Non-Local Means.
  IEEE Signal Process. Lett., 17 :269–272, 2010.

Références VI

- J. Salmon and E. Le Pennec.

  An aggregator point of view on NL-Means.

  In *Proceedings of the SPIE Conference on Mathematical Imaging : Wavelet XIII*, volume 7446, page 74461E. SPIE, 2009.

- J. Salmon and Y. Strozecki.

  From patches to pixels in non-local methods : Weighted-Average reprojection.

  In *ICIP*, 2010.

- C. M. Stein.

  Estimation of the mean of a multivariate normal distribution.

  Ann. Statist., 9(6) :1135–1151, 1981.

- T. Tasdizen.

  Principal neighborhood dictionaries for nonlocal means image denoising.

  IEEE Trans. Image Process., 18(12) :2649–2660, 2009.

Références VII

- D. Tschumperlé and L. Brun.
  Image denoising and registration by pde's on the space of patches.
  In *LNLA*, pages 32–40, Lausanne, 2008.
- C. Tomasi and R. Manduchi.
  Bilateral filtering for gray and color images.
  In *ICCV*, pages 839–846, 1998.
- D. Van De Ville and M. Kocher.
  SURE-based Non-Local Means.
  IEEE Signal Process. Lett., 16 :973–976, 2009.
- G. S. Watson.
  Smooth regression analysis.
  Sankhya : The Indian Journal of Statistics, Series A, 26(4) :359–372, 1964.
- D. H. Wolpert.
  Stacked generalization.
  Neural Networks, 5(2) :241–259, 1992.

Références VIII

- L. P. Yaroslavsky.

  Digital picture processing, volume 9 of *Springer Series in Information Sciences*.

  Springer-Verlag, Berlin, 1985.