

Chapitre 4.3 - Introduction aux images Super-résolution

9 avril 2014

Plan de la présentation

1 Idée générale de la Super-résolution

- Introduction
- Comment augmenter la résolution ?

2 Super-resolution à plusieurs images

- Définition du problème
- Approche par interpolation non uniforme
- Approche par transformée de Fourier
- Reconstruction SR régularisée

3 Super-resolution à une seule image

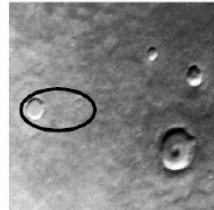
- Approche générale
- Ensemble d'entraînement
- SR par réseaux de Markov

La densité de pixel pour une image est basse, ce qui offre moins de détail.

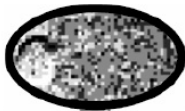
La densité de pixel pour une image est haute, ce qui offre plus de détail.

L'action d'obtenir une image HR à partir d'une ou plusieurs images LR.

- Imagerie médicale
- Image satellite
- Video surveillance (zoom sur plaque d'immatriculation).
- Télévision HD (conversion de contenu de vidéo NTSC en haute résolution).
- ...



Super-résolution





A bright yellow Lotus Elise is shown from a front-three-quarter view, driving on a paved road. The car has a black front bumper and a license plate that reads 'ORV 756V'. The background consists of a grassy area and large trees.

- 1 Diminuer la taille des pixels → Acquisition
 - + Augmente la résolution spatiale.
 - Introduction de bruit → Quantité de lumière réduite.
- 2 Augmenter la taille des capteurs CCD → Acquisition
 - + Augmente la résolution spatiale.
 - Capteurs coûtent très cher.
- 3 Super-résolution (SR) → Traitement
 - + Ne coûte rien.
 - + On obtient à la fois des images LR et HR.

1 Approche à plusieurs images (MFSR) :

- Extrapolation de l'information HR des différentes images LR.

② Approche à une image (SFMR) :

- Interpole l'information HR, donc PAS d'ajout d'information HR réelle à l'image résultante.

- Introduction
- Comment augmenter la résolution ?

- Définition du problème
- Approche par interpolation non uniforme
- Approche par transformée de Fourier
- Reconstruction SR régularisée

- Approche générale
- Ensemble d'entraînement
- SR par réseaux de Markov

Comment peut-on obtenir une image HR à l'aide d'une séquence de LR ?

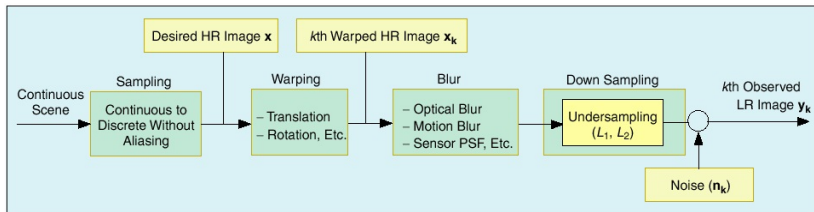
- On doit d'abord avoir fait l'acquisition de plusieurs LR représentant la même scène ;
 - Ces LR doivent donner un différent point de vue de la scène, et ce **avec un décalage sous-pixel**.
- Si les images ont un décalage unitaire (décalage de un, deux... pixels), alors **les images contiennent la même information**. SR ne sera pas possible.
- Si les images ont un décalage sous-pixel différent, alors SR est possible.



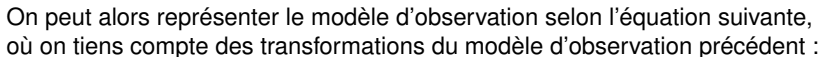


Modèle d'observation

Pour bien comprendre la SR, il faut bien modéliser le système d'observation afin de trouver un lien entre les images LR et l'image désirée HR.



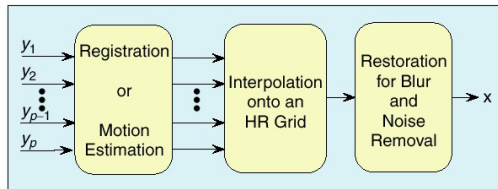
Modèle d'observation



- M_k est une matrice de warping (translation-rotation) ;
- B_k représente une matrice de flou ;
- D est une matrice de sous-échantillonnage ;
- n_k représente une matrice de bruit.

L'**approche par interpolation non uniforme** est l'approche la plus intuitive pour la reconstruction SR. Elle est composée de trois étapes :

- 1 Recalage (ou alignement) ;
- 2 Interpolation sur une grille HR ;
- 3 Réduction de bruit et de flou.



Si un faible mouvement non uniforme perturbe la scène, on doit faire une **interpolation non uniforme** sur les différentes images pour obtenir l'espacement uniforme requis :



Dépendamment de notre connaissance *a priori* du **bruit** et de la **dispersion du flou**, on utilisera les algorithmes/filtres suivants :

- Cas sans bruit, flou connu
Filtre inverse
- Cas bruit connu, flou connu
Filtre Weiner
- Cas bruit et flou inconnu
Filtre par Expection-Maximisation

Le flou créé par le système est modélisé comme étant l'effet d'un **filtre moyennneur**, qui cause une **perte de haute fréquence seulement** :



La modélisation du flou sera donc de faible intensité dans la cas d'une scène statique. Dans le cas d'une scène en mouvement, nous parlerons alors de **flou de mouvement** qui exige un traitement plus avancé.

Nous utiliserons donc un **filtre Wiener**, qui suppose une connaissance *a priori* du flou et du bruit. Le filtre Wiener est basé sur le filtre inverse.

$$\begin{aligned}\hat{x}(i,j) &= x(i,j) * h(i,j) && \text{Domaine spatial} \\ \hat{X}(u,v) &= X(u,v) \cdot H(u,v) && \text{Domaine fréquentiel}\end{aligned}$$

◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡ ≡ ≡ ↺ 🔍 ↻

$$\hat{x}(i,j) = x(i,j) * h(i,j) + n(i,j)$$

Le résultat du filtre inverse $G(u, v) = \frac{1}{H(u, v)}$ sera :



$$E(i, j) = (x(i, j) - \hat{x}(i, j))^2$$

Dans le domaine de fourrier, le filtre $G(u, v)$ minimisant l'équation de l'erreur ci-dessus est donnée par :

$$\begin{aligned} G(u, v) &= \frac{H^*(u, v) \cdot (X(u, v)X^*(u, v))}{(X(u, v)X^*(u, v)) \cdot (H(u, v)H^*(u, v)) + N(u, v)N^*(u, v)} \\ &= \frac{H^*(u, v)}{|H(u, v)|^2 + \frac{|N(u, v)|^2}{|X(u, v)|^2}} \\ &= \frac{1}{H(u, v)} \frac{|H(u, v)|^2}{|H(u, v)|^2 + \frac{|N(u, v)|^2}{|X(u, v)|^2}} \end{aligned}$$

En pratique, on ne connaît ni l'image sans corruption, ni le bruit. Donc, on approxime le rapport avec un facteur K (par exemple, $K = 0.04$) :

$$G(u, v) = \frac{1}{H(u, v)} \frac{|H(u, v)|^2}{|H(u, v)|^2 + K} \quad (2)$$



- Introduction
- Comment augmenter la résolution ?

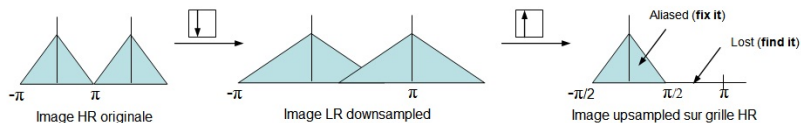
- Définition du problème
- Approche par interpolation non uniforme
- **Approche par transformée de Fourier**
- Reconstruction SR régularisée

- Approche générale
- Ensemble d'entraînement
- SR par réseaux de Markov

Approche par transformée de Fourier

Tout comme l'approche non-uniforme, l'**approche par TF** utilise les même trois étapes : recalage, projection sur la grille HR et réduction. Cependant, la projection sur la grille utilise le domaine fréquentiel.

Étant donné que les LR sont des images sous-échantillonnées, on voudrait retrouver les valeurs de haute fréquence perdues. On a donc de l'aliasing dans nos images sous-échantillonnées LR.



Approche par transformée de Fourier

Projection sur grille HR - Algorithme de Papoulis-Gerchberg

La méthode de projection par Papoulis-Gerchberg assume deux choses :

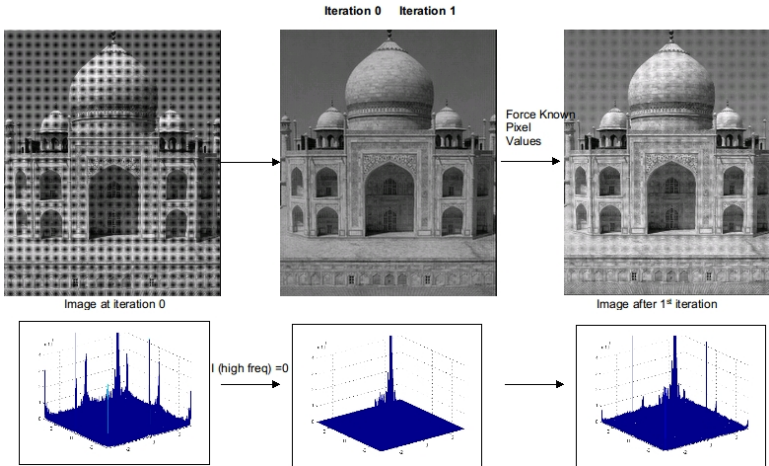
- Quelques valeurs de la grille haute résolution sont connus ;
- Aucune image LR upsampled dans la grille HR n'a de contenu haute fréquence.

◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡ 🔍 ↺

1. *Journal of Management Studies*, 1990, 27, 1.

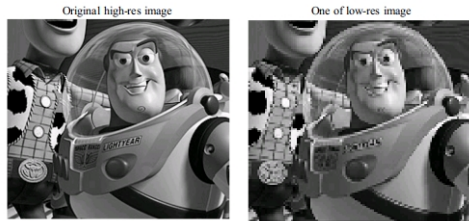
Projection sur grille HR - Algorithme de Papoulis-Gerschberg

100



1. *Journal of the American Medical Association*, 1997; 278: 1039-1044.

Projection sur grille HR - Algorithme de Papoulis-Gerschberg



Bicubic Interpolation

Superresolution Result

Approche par transformée de Fourier

Projection sur grille HR - Algorithme de Papoulis-Gerschberg

Cas où les LR ont un décalage d'un pixel :



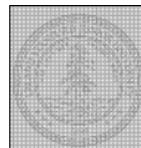
Original high-res image



One of low-res image



Bicubic Interpolation

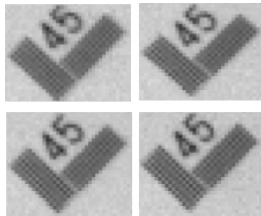


Superresolution result

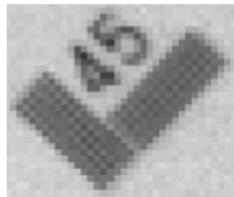
Approche par transformée de Fourier

Projection sur grille HR - Algorithme de Papoulis-Gerchberg

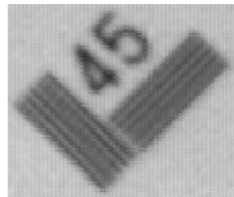
Cas où les images un décalage sous-pixel “répandu uniformément”



Original Low-res images



Bicubic Interpolation (1st image)



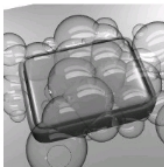
Superresolution Result

Approche par transformée de Fourier

Approche par transformée de Fourier

Projection sur grille HR - Algorithme de Papoulis-Gerchberg

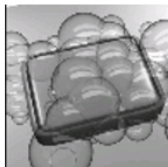
Cas où les images un décalage sous-pixel “répandu uniformément”



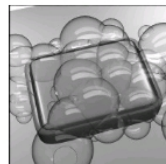
Original high-res image



One of low-res image



Bicubic Interpolation



Superresolution result

Plan de la présentation

- 1 Idée générale de la Super-résolution
 - Introduction
 - Comment augmenter la résolution ?
- 2 Super-resolution à plusieurs images
 - Définition du problème
 - Approche par interpolation non uniforme
 - Approche par transformée de Fourier
 - Reconstruction SR régularisée
- 3 Super-resolution à une seule image
 - Approche générale
 - Ensemble d'entraînement
 - SR par réseaux de Markov

Reconstruction SR régularisée

Avec une bonne quantité d'images LR, on peut résoudre l'équation du système d'observation :

$$\begin{aligned} y_k &= D \cdot B_k \cdot M_k \quad x + n_k \quad \forall 1 \leq k \leq p \\ &= W_k \quad x + n_k \quad \forall 1 \leq k \leq p \end{aligned}$$

En réalité, il est difficile d'avoir un nombre suffisant de LR, car le système est mal posé. On utilise alors une **regression linéaire** pour stabiliser l'inversion de problème. On tentera alors d'approcher la solution de façon itération par des techniques **statistiques**.

Reconstruction SR régularisée

Maximisation a posteriori (MAP)

On cherche à savoir quelle est la meilleure image HR X connaissant les p images Y_k . D'un point de vue statistique, on cherche donc X qui maximisera la probabilité conditionnelle $P(x|y_1 \dots y_n)$:

$$E(X) = \arg \max P(X|Y_1 \dots Y_n)$$

Reconstruction SR régularisée

Maximisation a posteriori (MAP)

Par la loi de Bayes :

$$\begin{aligned}
 E(X) &= \arg \max && P(X|Y_1 \dots Y_n) \\
 &\simeq \arg \max && \log P(X|Y_1 \dots Y_n) \\
 &= \arg \max && \log(P(Y_1 \dots Y_n|X)) + \log(P(X)) \\
 &= \arg \max && \log(e^{-\frac{1}{2\sigma}(\sum_{k=1}^p \|Y_k - W_k \cdot X\|^2)}) + \log(\frac{1}{Z} e^{-\lambda A(X)}) \\
 &= \arg \min && \sum_{k=1}^p \|Y_k - W_k \cdot X\|^2 + \lambda A(X)
 \end{aligned} \tag{3}$$

Où $P(y_1 \dots y_n|x)$ la distribution (influence) de chacune des image y_k sur X et $P(X)$ est un *a priori* de l'image résultante. $A(X)$ est une fonction dépendant du contenu de l'image résultante (*a posteriori*)

→ L'avantage d'un modèle Bayésien réside dans représentation d'un *a priori* de préservation de contour.

Reconstruction SR régularisée

Maximum de vraisemblance

Supposons $A(X)$ uniforme (ou nul). Le **maximum de vraisemblance** (le terme de gauche) sera alors le seul auteur de la réponse optimale pour X

$$E(X) = \arg \min \sum_{k=1}^p \|Y_k - W_k \cdot X\|^2 \quad (4)$$

En dérivant le terme, on obtient :

$$\frac{\delta E(X)}{X} = \sum_{k=1}^p (W_k^T W_k)^{-1} W_k^T X \quad (5)$$

Où $(W_k^T W_k)^{-1}$ doit être inversible, ce qui est rarement le cas.

Reconstruction SR régularisée

Maximum de vraisemblance - IBP

Pour pallier au problème, on peut utiliser une méthode appelée *Iterative Back projection*, qui consiste à produire une image basse résolution simulée \hat{Y}_k à partir de X^t courant :

$$X^{t+1} = X^t + \alpha \left(\sum_{k=1}^p \hat{Y}_k - Y_k \right) \quad (6)$$

Cette méthode est très semblable à Papoulis (domaine fréquentiel).

Supposons maintenant que $A(X)$ soit une méthode dépendante du gradient de l'image.

Soit W_k la matrice qui combine le sous-échantillonnage, le mouvement et le bruit de la k image Y_k . On intègre alors un terme de régularisation basé sur les hautes fréquences, soit le gradient de l'image X .

$$E(x) = \sum_{k=1}^p ||Y_k - W_k \cdot X||^2 + \lambda ||\nabla x||^2 \quad (7)$$

λ est un terme de régularisation. Plus λ est élevé, plus le résultat sera lissé (utile pour peu de LR ou beaucoup de bruit).

Par Euler-Lagrange, on peut trouver X qui minimise la fonction de coût (Eq.7) de façon itérative :

$$\frac{\delta E(X)}{X} = \sum_{k=1}^p W_k^T \cdot (W_k \cdot X - Y_k) - \Delta X \quad (8)$$

Par la descente de gradient, on trouve alors la forme itérative suivante :

$$X^{t+1} = X^t + \alpha \left(\sum_{k=1}^p W_k^T \cdot (W_k \cdot X^t - Y_k) - \lambda \Delta X^t \right) \quad (9)$$

44 / 66

Reconstruction SR régularisée

Approche variationnelle



L^2 superresolution



L¹ superresolution



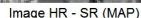
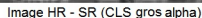
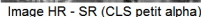
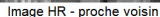
input

L² superresolution

L^1 superresolution

Reconstruction SR régularisée

Approche stochastique (MAP)



Semblable à Papoulis

Qui utilise les différences de focus plutôt que les mouvements subpixels.

Qui intègre l'information temporelle dans le système d'observation.

<https://www.youtube.com/watch?v=9G69WcHgRfo>

- Introduction
- Comment augmenter la résolution ?

- Définition du problème
- Approche par interpolation non uniforme
- Approche par transformée de Fourier
- Reconstruction SR régularisée

3

- Approche générale
- Ensemble d'entraînement
- SR par réseaux de Markov

- Introduction
- Comment augmenter la résolution ?

- Définition du problème
- Approche par interpolation non uniforme
- Approche par transformée de Fourier
- Reconstruction SR régularisée

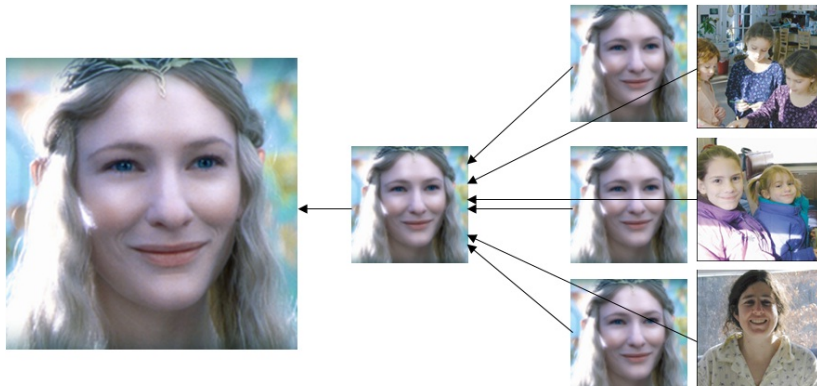
- Approche générale
- Ensemble d'entraînement
- SR par réseaux de Markov

- Basé sur le flou : Filtre de Weiner, filtre median, ...
- Basé sur l'interpolation : Plus proche voisin, bilinéaire, bicubique, ...
- Basé sur l'affutage : Sharpening filter (amplification de détail sans amplifier le bruit), ...

La **SR à un seul frame** estime les détails HR manquants **qui ne sont pas présents dans l'image**, ce qui n'est pas possible avec un filtre basé sur l'affutage.

Ensemble d'entraînement

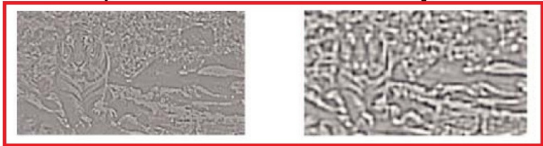
Bien que cela puisse paraître contre-intuitif, la SR à un frame n'utilise pas seulement le frame LR. Il utilise un ensemble d'entraînement pour apprendre les détails HR à basse résolution.



Il utilise ensuite ces relations acquises en phase d'apprentissage pour prédire les détails HR dans d'autres images (une seule image à la fois).

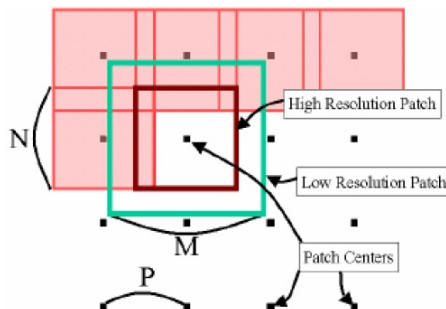
Nous verrons l'algorithme utilisant les **réseaux de Markov**.

- 1 On prend une collection d'images HR ayant un petit lien sémantique avec notre image LR (par exemples, des tigres et des chat pour notre image LR de tigre)
- 2 Pour toutes les images HR de la collection, on dégrade la qualité de l'image (sous-échantillonnage de facteur 4 et flou global).
- 3 Redimensionner en HR en interpolant (cubic spline), ce qui causera la perte des données HR.
- 4 Filtre passe-haut et normalisation de l'intensité pour l'image interpolée *et* l'image HR originale.



Ensemble d'entraînement

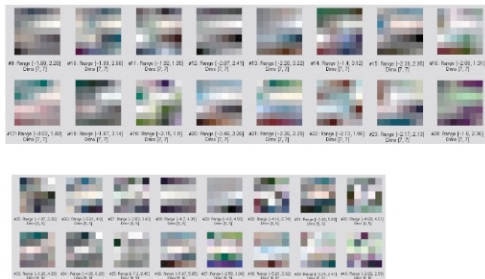
On décrit alors à la fois l'image LR (x) et HR (y) obtenue précédemment selon le système de patch. Soit P la distance entre chaque acquisition de patch, $N \times N$ la dimension de la patch LR et $M \times M$ la dimension de la patch HR :



Ensemble d'entraînement



w: 1024 [x: 0', 2.00]
[0x2 [0', 1]



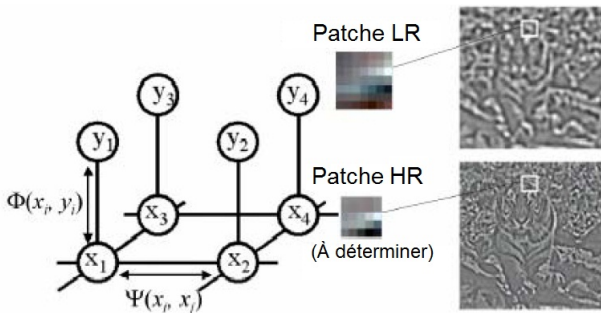
- Introduction
- Comment augmenter la résolution ?

- Définition du problème
- Approche par interpolation non uniforme
- Approche par transformée de Fourier
- Reconstruction SR régularisée

- Approche générale
- Ensemble d'entraînement
- **SR par réseaux de Markov**

Réseaux de Markov

Dans un **réseau de Markov**, les cercles représentent les noeuds du réseau, et les lignes sont les **dépendances statistiques** entre les noeuds. Les noeuds y représentent les patches LR de l'image d'entrée, et x représente les patches HR recherchées.

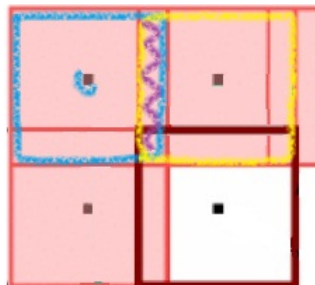


ψ représente le lien entre les patches désirées avoisinantes de l'image et ϕ représente la relation entre la patches d'entrée et la patche HR désirée.

$$P(x|y) = \frac{1}{Z} \cdot \prod_{(i,j)} \psi_{ij}(x_i, x_j) \cdot \prod_i \phi_i(x_i, y_j) \quad (12)$$

On établit une relation de distance entre les patches voisins. Comme **on s'arrange pour avoir un overlay de 1 ou 2 pixels entre les patches HR**, on définira la distance $d_{ij}(x_i, x_j)$ comme étant une relation de différence d'intensité de chaque pixel **dans la zone commune** :

Où σ est un facteur de bruit.



ϕ utilise le même principe que ψ , c'est-à-dire une relation de distance euclidienne entre les pixels de la zone commune :

Une variante des réseaux de Markov, l'**algorithme à une passe**, permet d'obtenir plus rapidement la SR sans avoir trop de perte de la qualité.

Cette variante n'utilise, dans le calcul de la probabilité (Eq.12), que les patches x_i précédemment calculés. C'est donc une sorte de *balayage* de l'Image d'entrée patche LR par patche LR.





Best-match patches
from training data





any illegal bundling, or "can
anelvacatedarulingbythefed-
system,andsentitdowntoanew
finedastandardforweighing
raproduct-bundlingdecision.
softsayssthatthenewfeature:
andpersonalidentification:
soft's'view, butusersandthe
adedwithconsumerinnovation
rePCindustryislookingforw.



