

Motivation. Exemple sur la recherche biomédicale

- 2009 : David Donoho met en évidence une crise de crédibilité dans la recherche scientific
- 2011: le directeur de la division oncologie d'Amgen a essayé de reproduire 53 des plus importants articles sur le cancer. Il n'a pas réussi à en reproduire 47
- Bayer HealthCare Germany (re)confirme: seulement 25% de la recherche sur le cancer est reproductible

Mise en œuvre de la Recherche Reproductible

Sciences non exactes (biologie, médecine, ...): difficile (mais souhaitable).
 Difficile d'avoir exactement les mêmes conditions le long des expériences.

Sciences computationnelles : aucune excuse!

La contribution du Centre Borelli : le journal IPOL

- Lancé en 2009 à l'initiative de Nicolas Limare, Jean-Michel Morel et d'un groupe de rédacteurs fondateurs
- Un journal initialement sur le traitement des images (Image Processing On Line). Accès diamant.
- Très vite, d'autres types de données ont été ajoutés : vidéo, audio, ...
- Aujourd'hui, il s'agit d'un journal général sur les algorithmes reproductibles
 - IPOL → Information Processing On Line

Publications IPOL

- Révision par les pairs. Tant l'article (PDF) que le code source
 - O Reproductibilité: les réviseurs vérifient soigneusement que le code source correspond au pseudo-code.
- Chaque publication :
 - O Un article décrivant la méthode en détail, y compris les pseudo-codes
 - O Le code source, sous une licence de logiciel libre
 - O Une **démo en ligne** qui permet aux utilisateurs de **tester** la méthode ave<u>c leurs **propres données** </u>
 - O Une archive d'expériences
- Il n'est pas nécessaire que ce soit un travail original. Nous sommes intéressés par les détails mathématiques, la reproductibilité et la compréhension
- ISSN, DOI, indexé par SCOPUS. Pas encore d'Impact Factor
- Jetons un coup d'œil! → http://www.ipol.im/pub/art/2017/201/

Article





ipol.im/pub/art/2017/201/



IPOL Journal · Image Processing On Line

HOME · ABOUT · ARTICLES · PREPRINTS · WORKSHOPS · NEWS · SEARCH

Multi-Scale DCT Denoising

Nicola Pierazzo, Jean-Michel Morel, Gabriele Facciolo

article demo archive

published + 2017-10-29

reference * Nicola Pierazzo, Jean-Michel Morel, and Garriele Facciolo, Multi-Scale DCT Denoising, Image Processing On Line, 7 (2017). pp. 288-308. https://doi.org/10.5201/ipol.2017.201

BibTeX info

Communicated by Julie Delon Demo edited by Gabriele Facciolo



This IPOL article is related to a companion publication in the SIAM Journal on Imaging Sciences;

G. Facciolo, N. Pierazzo, J-M. Morel, "Conservative Scale Recomposition for Multiscale Denoising (The Devil is in the High Frequency Detail)" SIAM Journal on Imaging Sciences 10(3):1603-1626, 2017, http://dx.doi.org/10.1137/17M1111826

Abstract

DCT denoising is a classic low complexity method built in the JPEG compression norm. Once made translation invariant, this algorithm was still proven to be competitive at the beginning of this century. Since then, it has been outperformed by patch based methods, which are far more complex. This paper proposes a two-step multi-scale version of the algorithm that boosts its performance and reduces its artifacts. The multiscale strategy decomposes the image in a dyadic DCT pyramid, which keeps noise white at all scales. The single scale denoising is then applied to all scales, thus giving multiple denoised versions of the low frequency coefficients of the denoised image. A 'multi-scale fusion' of these multiple estimates avoids the ringing artifacts resulting from the pyramid recomposition. The final algorithm attains a good PNSR and much improved visual image quality. It is shown to have a deficit of only 1dB with respect to state of the art algorithms, but its complexity is two orders of magnitude lower.

Download

- full text manuscript; PDF low-res, (1.7M) PDF (15.2M) [7]
- source code: Q ZIP SWHID info

Preview

Loading takes a few seconds. Images and graphics are degraded here for faster rendering. See the downloadable PDF documents for original high-quality versions.

We justify in detail our choice of aggregation weights.

For the hard thresholding pass of the algorithm the aggregation weights are set, as in [3], by counting the number N_P of nonzero DCT coefficients (excluding the zero frequency) in the patch after thresholding. These aggregation weights are then given by



where the one is added to prevent the dividing by zero (but it is an arbitrary choice). Indeed, the number of non-zero coefficients will be small for the flat patches, compared to patches containing

290

LOW RESOLUTION PDF: Images may show compression artifacts. A full resolution PDF is available at www.ipol.im.

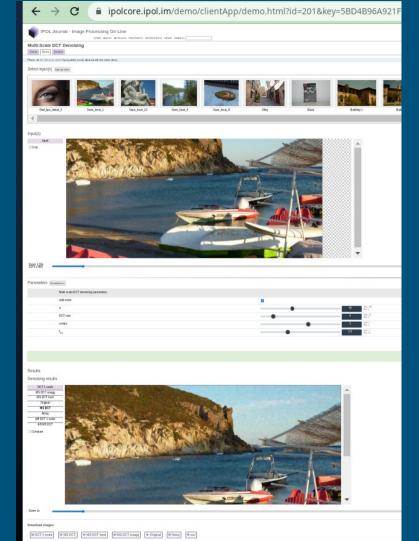
MULTI-SCALE DCT DENOISING

Algorithm 2: DCT Denoising - Hard thresholding

input : noisy image Y, noise level σ , and patch size s

1 Function DCTDENOISINGHARD(Y, σ, s)

Demo





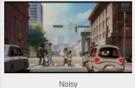
Archive











Pourquoi ne faisons-nous pas tous des recherches reproductibles ?

- Certains chercheurs ne veulent pas rendre public le code de travail
 - Ne correspond à aucune version des pseudo-codes
 - Faible qualité des logiciels
 - Un logiciel de qualité prend plus de temps à produire : tests, documentation, mesures objectives de la qualité,
- Les résultats ne sont pas généralisables
- ..

- Pas vraiment considéré pour l'avancement de la carrière
 - Métrique classique : "Nombre de publications classiques à facteur d'impact élevé"
 - Le logiciel est considéré comme un citoyen de 2ème classe

Avantages du RR, d'après notre expérience dans IPOL

- Si une méthode en vaut la peine, l'impact est important
 - O Les utilisateurs du monde entier peuvent tester l'algorithme avec leurs propres données.
 - O Augmentation du nombre de citations : les autres chercheurs peuvent désormais se comparer à vous

 Accélération scientifique : les autres chercheurs peuvent réutiliser le texte, le code source et les données

Utile pour le Centre Borelli pour montrer un paysage de notre activité scientifique

Une façon de travailler

Conclusions

- La recherche reproductible est nécessaire pour l'avancement de la science
- L'article, le logiciel et les données doivent tous faire partie de la même publication
- Elle a un impact plus important, mais elle demande aussi un plus grand effort.
- Les chercheurs doivent être récompensés s'ils choisissent de l'effectuer
- De nombreuses initiatives dans ce sens :
 - Européen : European Open Science Cloud (EOSC)
 - National : Le "Plan pour la science ouverte"
 - O Centre Borelli: IPOL, OVD-SaaS, track MVA sur la reproductibilité

Probablement le seul moyen de réaliser une science de confiance!

Merci de votre attention