

La chasse au bruit dans les images

Antoni Buades ^{*} [†] Bartomeu Coll ^{*} Jean Michel Morel [†]

Résumé. *Le bruit sonore est de plus en plus considéré comme une nuisance intolérable, mais qu'en est-il du bruit visuel? Toutes les images et les films digitaux comportent un bruit, qui gêne la vision et impose des contraintes technologiques de prix, de taille et d'énergie aux caméras. Des progrès récents dans la compréhension de la structure des images permettent d'éliminer le bruit des images et des films sans les abîmer.*

En quelques années, les images digitales ont envahi notre vie quotidienne. Les caméras numériques permettent en effet de saisir et de manipuler bien plus facilement les images et les films. Leur qualité est maintenant équivalente et souvent supérieure à celles des images obtenues par les procédés photochimiques. Les images digitales offrent de plus l'immense avantage de pouvoir être transmises, améliorées et stockées sur des supports informatiques.

Beaucoup des altérations des images sur papier ou pellicule ne seront bientôt plus d'actualité. Les déchirures, craquelures et altérations de la couleur du support papier se corrigent très bien en scannant les images pour en faire des images digitales et en les restaurant par des algorithmes numériques.

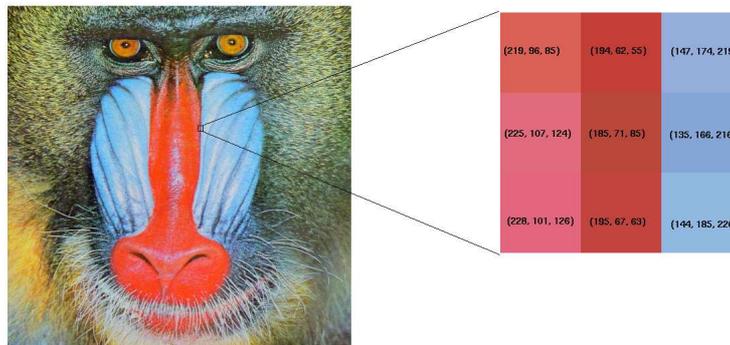


Figure 1: Comme le montre cette image, une image digitale est un ensemble de carrés, appelés pixels. Dans chaque carré, un triplet de trois valeurs indique la couleur par ses composantes rouge, vert et bleu. Chaque valeur peut aller de 0 à 255. On voit neuf pixels de cette image de babouin, à un endroit où rouge et vert contrastent fortement. Dans le rouge, c'est la première composante rouge qui est forte (plus de 210) et les autres sont moyennes. Dans les pixels bleus, c'est la troisième composante, bleue, qui domine.

Toutefois, les images numériques continuent de présenter certains défauts des photographies. Elles ont en particulier un bruit de fond qui limite leur lisibilité. La technologie digitale cherche à réduire au maximum ce bruit. Pour comprendre la nature du bruit visuel, il faut rappeler qu'une image

^{*}Universitat de les Illes Balears, Ctra. Valldemossa Km. 7.5, 07122 Palma de Mallorca, Spain

[†]Centre de Mathématiques et Leurs Applications. ENS Cachan 61, Av du Président Wilson 94235 Cachan, France

numérique se présente comme un rectangle divisé en petits carrés, les pixels. Chaque pixel est le résultat d'une mesure, généralement faite par une matrice de capteurs appelée CCD (pour *Charge Couple Device*). Un pixel correspond à un petit carré de la matrice CCD dans lequel le nombre de photons arrivant est compté. Chaque pixel contient trois nombres allant d'habitude de 0 à 255 et indiquant la quantité de rouge, vert et bleu contenues dans le pixel. Une combinaison adéquate de ces trois nombres permet en principe de reproduire n'importe quelle couleur sur écran d'ordinateur. La figure 1 montre un tout petit morceau de neuf pixels extrait d'une image digitale. Dans les pixels rouges, la première composante est élevée, autour de 210, et dans les pixels bleus c'est la troisième composante, bleue, qui domine.

Dans toute image digitale, les trois valeurs de couleur observées présentent une incertitude due au bruit. Cette incertitude est due aux aléas du comptage des photons arrivant dans chaque capteur. Les valeurs de couleur mesurées sont perturbées car les capteurs reçoivent des photons parasites et subissent des fluctuations électrostatiques lors de leurs charges et décharges. Quand un capteur reçoit beaucoup de photons venant d'une scène bien éclairée, les parasites sont négligeables par rapport au flux de vrais photons. Mais, même dans une photo d'exposition suffisante, les pixels sombres reçoivent très peu de photons et sont donc "bruités".

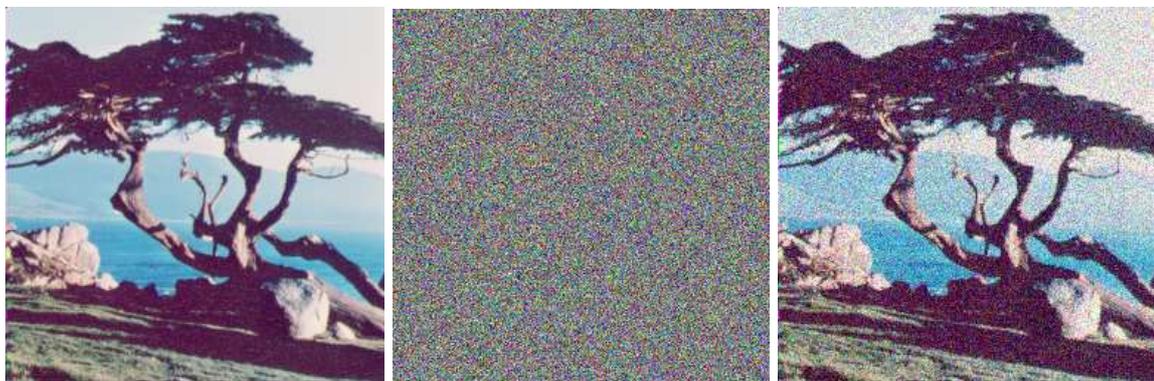


Figure 2: A l'image de gauche, on a ajouté l'image du milieu, une image de bruit dit "blanc", pour obtenir l'image bruitée de droite. Toutes les images ont du bruit mais il est en général moins accentué que dans cet exemple. Le bruit est dû aux fluctuations de la lumière et à l'activité photoélectronique des capteurs. Notre propre perception génère aussi un bruit, que l'on peut voir si on ouvre les yeux dans l'obscurité, ou simplement en fermant les yeux.

Si le temps d'exposition d'une photo est trop court, c'est toute l'image qui est bruitée. En effet, une photo papier sous-exposée serait simplement sombre. Mais une caméra digitale peut la corriger en augmentant le contraste numérique. L'ennui, c'est que cette opération augmente dans la même proportion le bruit. L'utilisateur voit alors une image claire, mais bruitée. En fait, d'autres opérations de restauration des images digitales amplifient le bruit et demandent donc à être couplées avec un débruitage. C'est notamment le cas pour les opérations enlevant le flou d'une image.

Même si l'image et son bruit sont produits en même temps, on peut fabriquer du bruit pur pour voir quel aspect il a et le reconnaître dans les images. Dans l'image de la figure 2, nous avons simulé un tel bruit. Nous l'avons ensuite ajouté à une photographie numérique, afin de simuler une image bruitée.

La figure ?? illustre aussi une des grandes énigmes de la perception humaine. Bien que l'image de droite de la figure 2 soit vraiment très bruitée, il n'y a pas vraiment de perte de visibilité, tout au plus une certaine gêne visuelle. Aussi les chercheurs en traitement d'images ont-ils pendant longtemps cru qu'il serait facile d'éliminer le bruit des images. Malheureusement, il est très difficile de séparer

le bruit des petits détails de l'image. Aussi, les méthodes de débruitage commerciales sont-elles très prudentes : elles enlèvent peu de bruit, afin d'être sûres de n'enlever aucun détail. Le défi est d'être capables de débruiter beaucoup plus, mais en enlevant le bruit et rien que le bruit.

La réduction du bruit dans les images digitales est liée à plusieurs enjeux technologiques. En effet, la taille des capteurs CCD est, elle-même, dictée en partie par l'exigence de réduction de bruit. Le nombre de photons qu'un capteur reçoit dans une prise de vue est proportionnel à sa surface. Un capteur trop petit recevrait donc peu de photons venant de la scène photographiée et serait bruité. Si on disposait d'un procédé de débruitage plus efficace que les actuels, on pourrait construire des caméras encore plus petites, en proportion de la diminution de taille du CCD, ou des caméras avec encore plus de pixels. Mais peut-être l'amélioration la plus importante serait-elle de pouvoir prendre des photos de bonne qualité avec un éclairage médiocre et sans flash. Ces enjeux technologiques sont présents dans pratiquement tout dispositif créant des images : en microscopie, en imagerie médicale ou satellitaire et bien sûr pour les caméras grand public.

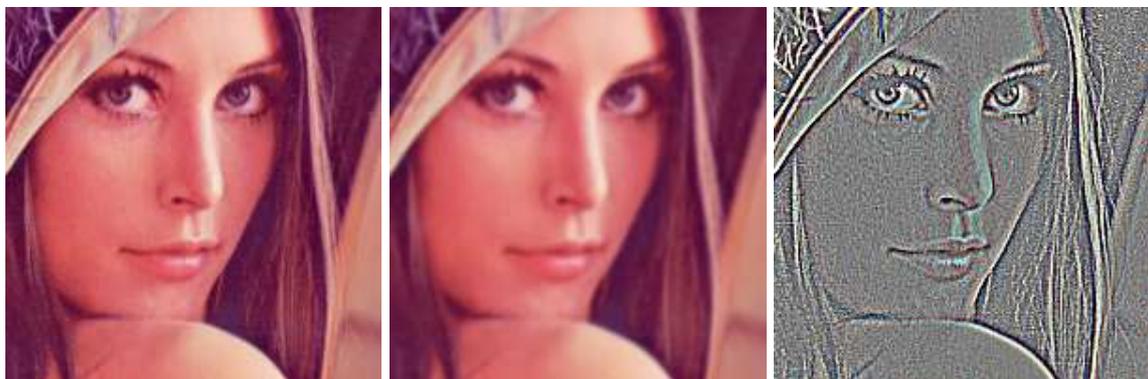


Figure 3: La première image, "Lena", est une image digitale fétiche des chercheurs en traitement d'images. Elle a été scannée en 1973 dans un exemplaire de Playboy et elle est toujours utilisée pour vérifier la validité des algorithmes de traitement ou de compression d'images. On la trouve sur le site de l'University of Southern California, (<http://sipi.usc.edu/database/>). La seconde image est obtenue en remplaçant chaque pixel par une moyenne des pixels voisins pour réduire le bruit. Cette méthode serait excellente si l'image n'en devenait pas floue. La troisième image montre la différence entre les deux images précédentes : elle contient visiblement le bruit, mais aussi des tas de détails de Lena que l'on reconnaît parfaitement. Cette perte d'information n'est pas acceptable.

Selon les experts, un progrès notable semble s'être accompli à la suite de la publication d'une nouvelle méthode de débruitage d'image, la "méthode des moyennes non locales". Grâce à cette méthode on réussit à enlever le bruit sans altérer l'image. Cette méthode peut être expliquée en quelques phrases.

Le principe de la plupart des méthodes de débruitage est simple : Il consiste à remplacer la couleur d'un pixel par la valeur moyenne des pixels voisins. Cette méthode est bien connue des statisticiens, qui savent que pour connaître le résultat probable d'un vote, il faut interroger suffisamment de personnes et faire la moyenne des résultats. Mettons que le résultat d'un vote soit proche de 40-60. Si on sonde une dizaine de personnes, le hasard peut très bien faire que sept de ces dix personnes soient de la première opinion. On conclura alors faussement que le résultat du vote est probablement 70-30. Si on interroge quarante personnes, on se rapprochera d'une évaluation correcte. Un calcul simple issu de la théorie des probabilités nous assure que si on interroge quatre fois plus de personnes, la marge d'incertitude du sondage est divisée par deux : celle-ci varie comme la racine carrée du nombre de sondés.



Figure 4: Dans cette expérience, on voit l'image Lena de la figure 3 débruitée par un filtre classique: le filtre sigma (*sigma-filter*) inventé en 1983. Pour voir si un filtre de débruitage marche bien, on fait la différence arithmétique entre l'image avant et après débruitage. Cette image devrait ressembler à un bruit. Dans le cas du filtre sigma, cette différence contient encore beaucoup de détails : on y reconnaît toujours Lena.

De même, si on calcule la couleur moyenne de quatre pixels voisins d'un pixel donné, cela divise par deux le bruit.

Malheureusement, ce raisonnement n'est valable que si les pixels voisins d'un pixel donné de l'image ont la même couleur. Ce n'est pas le cas partout dans une image. Il y a dans toute image des zones de couleur assez homogènes où la moyenne locale fonctionne. Pour voir quel problème peut surgir, regardons par exemple le pixel rouge placé au milieu de la figure 1. Ce pixel a cinq voisins rouges et trois bleus. Si on remplace la couleur de ce pixel par la moyenne des couleurs de ses voisins, il va bleuir. Le même procédé fera aussi rougir les pixels bleus de cette figure. Donc la frontière du rouge et du bleu deviendra moins nette. Par le procédé consistant à faire la moyenne des pixels voisins, une image perd donc une partie de son bruit mais devient perceptuellement floue, comme le montre la figure 2. Dans cette figure, nous avons appliqué le procédé des moyennes à une image test classique en traitement d'images, "Lena", scannée en 1973. L'effet de flou est dû à l'opération numérique de "moyenne locale".

Revenons à la figure 1. Il est parfaitement clair que pour débruiter le pixel rouge du centre, il faudrait ne faire une moyenne de la couleur de ce pixel qu'avec les pixels rouges voisins, en excluant les bleus. C'est ce que fait un algorithme classique, le filtre sigma (*sigma-filter*). Il est difficile de dire quand ce filtre a été inventé. Dans la littérature anglosaxonne, il est généralement attribué à J.S. Lee, en 1983. Selon Leonid Yaroslavsky, du Technion Institute en Israël, ce filtre était déjà connu des savants soviétiques. Le filtre sigma, s'il est bien réglé, remplacera la valeur du pixel rouge par la moyenne de celui-ci et des pixels rouges situés juste au dessus et au dessous du pixel rouge central de la figure 1. La figure 4 montre que ce filtre marche très bien.

Toutefois, les experts en photographie lui reprochent la création de certaines irrégularités, notamment des aplats et de faux détails, qui sont visibles dans la figure 4. On peut tester très facilement si un filtre de débruitage d'image abîme ou non l'image : il suffit de faire la différence arithmétique entre l'image avant et après filtrage. Elle doit montrer le bruit qui a été enlevé. Dans l'image de droite de la même figure, on voit que ce "bruit" contient en fait plein de détails de l'image. Le débruitage a donc dégradé l'image, ce qui est inacceptable.

La solution que nous avons proposée consiste à chercher dans une vaste portion de l'image tous les pixels qui ressemblent vraiment au pixel que l'on veut débruiter, et à en faire la moyenne.

Ceci est un changement important de conception. Tous les algorithmes antérieurs combinent les

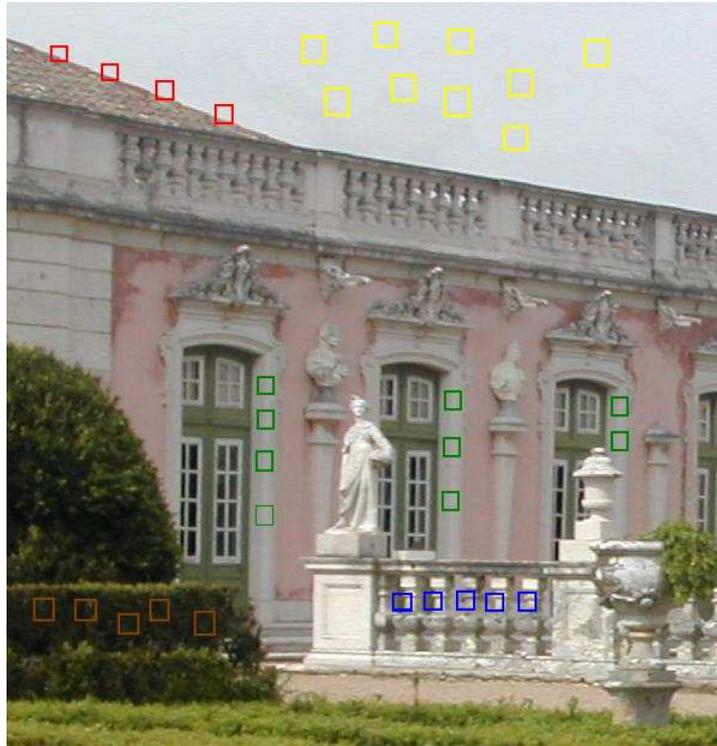


Figure 5: Il est rare qu'un détail dans une image ne soit pas répété plusieurs fois. Les groupes de carré de couleurs indiquent quelques groupes de petites zones de l'image qui sont presque indistinguables entre elles, en rouge sur la bordure du toit, en jaune dans le ciel bleu, très homogène, en vert le long des bordures des fenêtres et en orange dans une haie. Cette autosimilarité des images n'a commencé à être prise en compte que très récemment pour les restaurer. Ici, elle est utilisée pour enlever le bruit des images, grosso modo en faisant la moyenne des carrés qui se ressemblent. Ce moyennage préserve l'intégrité de l'image mais réduit ses petites fluctuations, pour l'essentiel dues au bruit d'enregistrement.

pixels *voisins* dans l'image : Or, on ne cherche plus des pixels proches au sens spatial, mais juste des pixels ressemblants. D'où le nom de "moyennes non locales" qui a été donné à cet algorithme.

La figure 5 permet de bien comprendre le procédé. Il est rare qu'un détail dans une image ne soit pas répété plusieurs fois. Les carrés de couleurs indiquent les petites zones de l'image qui sont presque indistinguables entre elles, en rouge sur la bordure du toit, en jaune dans le ciel bleu, très homogène, en vert le long des bordures des fenêtres, en orange dans une haie et en bleu sur une balustrade. Cette remarque très simple que les images sont "autosimilaires" n'a commencé à être prise en compte qu'assez récemment pour restaurer les images. En 1999, deux chercheurs de l'Université de Californie à Berkeley, Alexei Efros et Thomas Leung, l'ont utilisée pour synthétiser des textures et boucher des trous dans les images. Ici, elle est utilisée pour enlever le bruit des images, *grosso modo* en faisant la moyenne des carrés qui se ressemblent. Ce moyennage préserve l'intégrité de l'image mais réduit ses petites fluctuations, pour l'essentiel dues au bruit d'enregistrement.

L'analogie avec un sondage permet de bien comprendre le procédé: si on veut prévoir le résultat d'un vote national dans une ville, on doit interroger un échantillon de chaque catégorie d'habitants de la ville en respectant les proportions de chaque catégorie socio-économique. Si on interroge des



Figure 6: Dans cette expérience, on voit l'image Lena de la figure 3 débruitée par l'algorithme des moyennes non locales. L'image de droite montre ce qui a été enlevé à l'image originale par l'algorithme. Cette différence ressemble beaucoup plus à un bruit que les précédentes.

habitants d'une autre ville, on arrivera à un résultat comparable en prenant les mêmes catégories de personnes. Il n'est pas nécessairement important de faire le sondage dans la même ville ; ce qui compte, c'est d'interroger ce que les statisticiens appellent un échantillon représentatif. De même, ici, on cherche pour chaque pixel à trouver un groupe de pixels qui lui ressemblent sous tout rapport.

L'algorithme des moyennes non locales s'applique particulièrement bien aux films. Les chercheurs en restauration de films ont depuis longtemps proposer de débruiter chaque pixel d'une image d'un film en suivant sa trajectoire dans les images suivantes et précédentes. Cela revient à faire une moyenne avec "compensation du mouvement". Malheureusement, cette idée, pour excellente qu'elle soit, marche mal en pratique car personne à ce jour n'a réussi à concevoir des algorithmes suivant correctement les trajectoires dans un film. Ceci peut paraître surprenant à un profane : il nous semble que nous suivons très bien du regard les mouvements dans un film. L'ennui, c'est que notre compréhension est plutôt globale: nous suivons un personnage ou un objet en mouvement. Mais nous serions bien en peine de dire où va chaque pixel d'une image à l'autre.

Appliqué à un film, l'algorithme des moyennes non locales n'a pas besoin de calculer les trajectoires. Il cherche simplement les pixels ressemblants où qu'ils se trouvent dans le film. La figure 9 compare les résultats de deux algorithmes des plus sophistiqués compensant le mouvement et celui des moyennes non locales. Seul ce dernier algorithme extrait du film un bruit qui ressemble à un bruit.

References

- [1] Alexei A. Efros et Thomas K. Leung, *Texture synthesis by non-parametric sampling*, Int. Conf. on Computer Vision, Corfu, Greece, September 1999.
- [2] Toni Buades, Bartomeu Coll et Jean-Michel Morel *A review of image denoising algorithms, with a new one*, à paraître dans Multiscale Mathematical Modelling, 2004, prépublication N° 2004-15 du CMLA <http://www.cmla.ens-cachan.fr/Cmla/>, N° 2004-15. L'algorithme et ses applications font l'objet d'une demande de brevet déposée le 5 Mai 2004.
- [3] Dirk Borghys et Marc Acheroy, *Noise reduction in image sequences with sparse temporal sampling*, Proceedings of the LASTED International Conference on Signal and Image Processing-SIP95, Las Vegas (NV), pp. 452-455, 20-23 Novembre 1995.

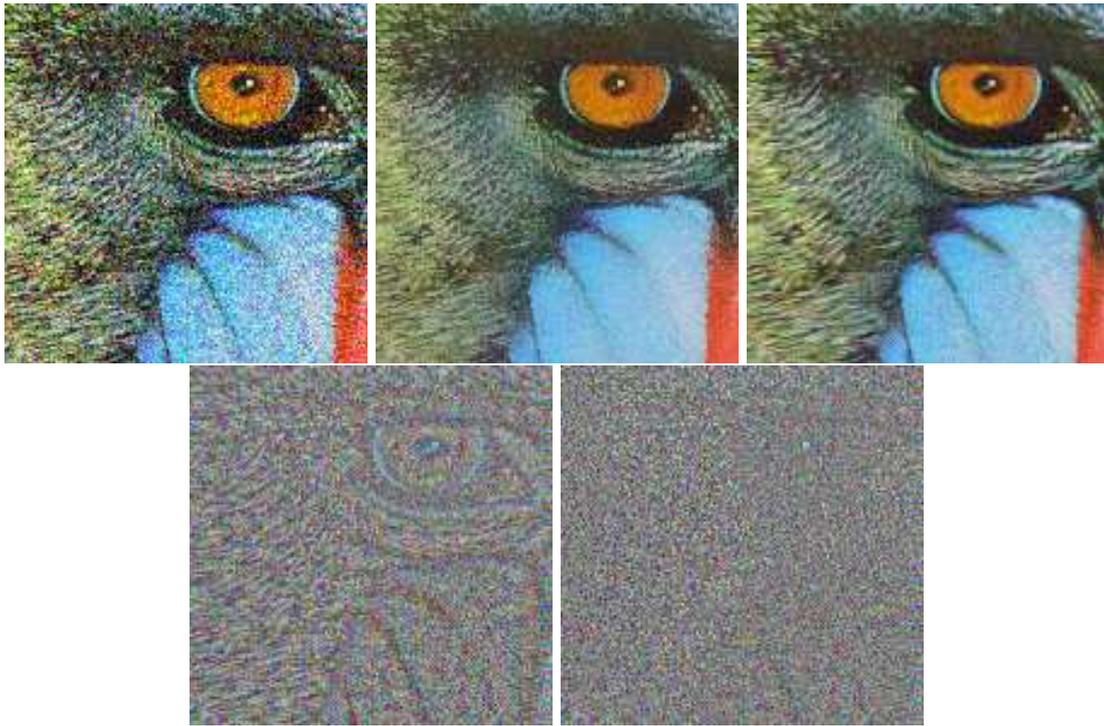


Figure 7: L'image de gauche a été bruitée par le procédé de la figure 2. La seconde image montre le résultat d'un débruitage par le filtre sigma et la troisième par moyennes non locales. La difficulté dans une image comme celle-ci est que la fourrure crée une texture oscillante qui ressemble beaucoup à un bruit et il faut réussir à distinguer les deux. Les deux images du bas montrent la différence entre image originale et image filtrée pour les deux méthodes. Le filtre sigma a inclus dans le bruit une bonne partie des détails, et en particulier la fourrure.



Figure 8: La première image est un exemplaire scanné d'un morceau de vieille photographie trouvée sur le Web. La seconde montre le résultat de son débruitage par l'application des moyennes non-locales. La troisième image est le "bruit" supposé et enlevé par un logiciel commercial répandu et la quatrième celui enlevé par l'algorithme des moyennes non locales. Une méthode est d'autant plus efficace que l'on voit moins de détails de l'image dans le bruit.

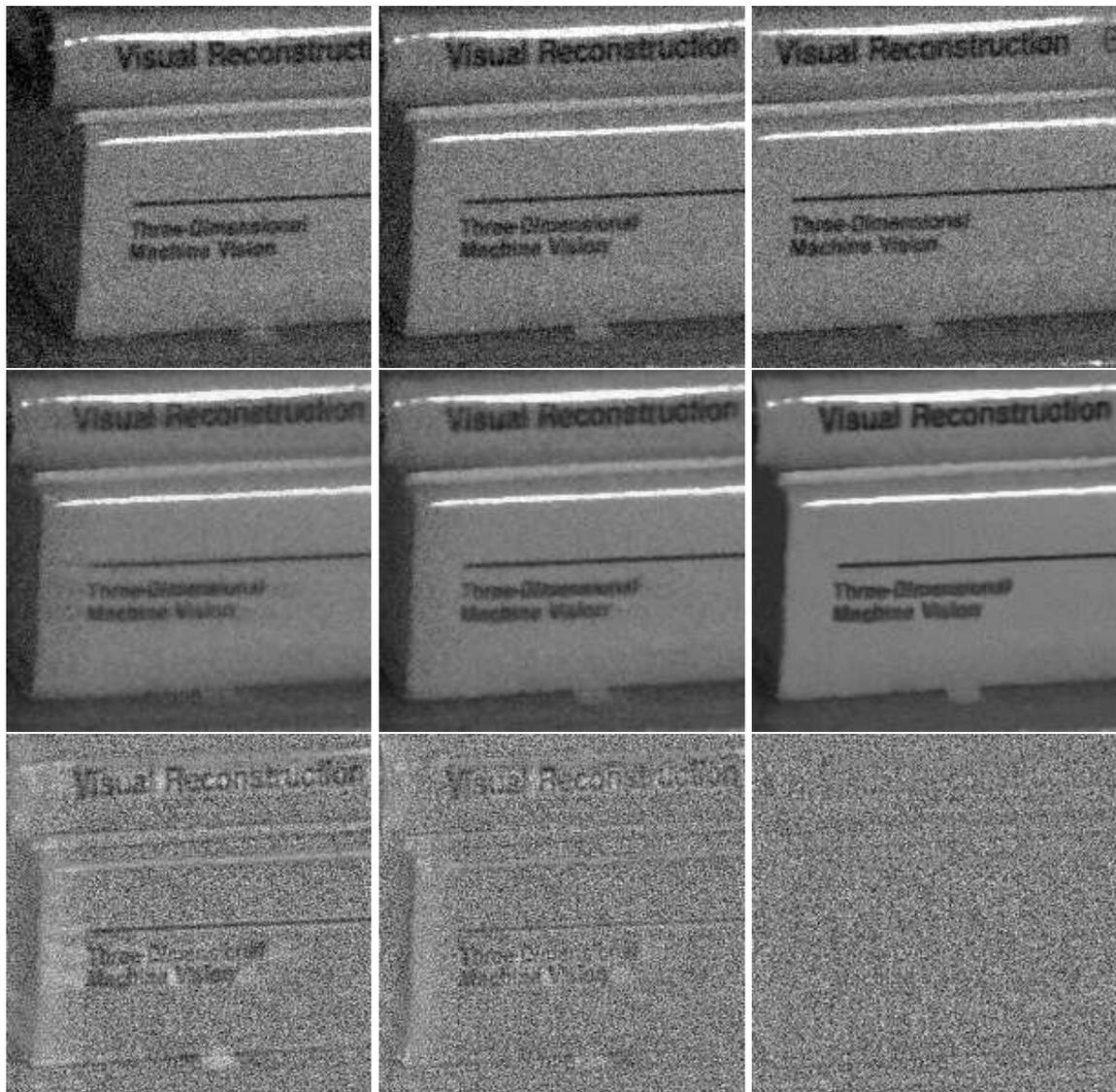


Figure 9: Cette planche compare l'application de la méthode des moyennes non locales avec deux filtres classiques de film ou de vidéo. En première ligne, on voit trois images successives d'un film test, avec un mouvement important de droite vers la gauche. En deuxième ligne, le résultat du débruitage du film par trois filtres. On montre le résultat du débruitage sur la deuxième image du film. Le résultat est obtenu avec AWA, un filtre "compensant le mouvement". Le second filtre dû à Borghys et Acheroy compense aussi le mouvement et calcule les moyennes non locales le long des trajectoires du film. Le troisième est la méthode des moyennes non locales, qui ne calcule pas de trajectoire et utilise donc toutes les ressemblances dans chaque image et entre les images. La troisième ligne montre le "bruit" trouvé par les trois méthodes. Les deux premières méthodes ont provoqué une perte d'information, qui est presque nulle avec la troisième.