

# Amélioration d'images sous-marines

JEAN CAILLÉ, FLORIAN DENIS

Télécom ParisTech - SI241

jean.caille@polytechnique.edu - florian.denis@polytechnique.edu

## Abstract

*Les images sous marines sont par définitions acquises dans un environnement où les conditions de visibilité sont particulièrement mauvaises. Ainsi, les images obtenues par des appareils classiques sont fortement dégradées. L'hétérogénéité du milieu, la diffusion de la lumière par les particules en suspension et la non-transparence de l'eau contribuent chacun à une prise de vue s'éloignant de la réalité. L'article que nous avons étudié montre qu'il est possible de supprimer ces dégradations dans certains cas, en utilisant des hypothèses minimales quand à la scène observée. Cette méthode est de plus applicable à un grand nombre d'exemple car elle ne nécessite qu'une image en entrée.*

*Nous avons étudié, implémenté et mesuré les performances de la méthode proposée par les auteurs de l'article. Dans ce rapport, nous nous attacherons à décrire d'une part la technique employée, et nous décrirons les outils utilisés pour l'implémentation. Nous analyserons ensuite nos résultats et les comparerons avec ceux obtenus par les auteurs. Finalement, nous tenterons de suggérer des améliorations à cette méthode.*

## I. INTRODUCTION ET MOTIVATION

Aujourd'hui, le grand public commence à s'équiper d'appareils photos résistants à l'eau (GoPro, ...). Les images acquises dans ces conditions difficiles sont souvent dégradées. Toutefois, ce même matériel est suffisamment sensible pour permettre la restauration des images, permettant alors d'améliorer grandement la qualité des images rendues.

Nous avons choisi d'effectuer ce projet car les résultats promis par l'article et la description qui en était faite nous semblaient intéressants. Plusieurs concepts variés, très utilisés en images sont introduits (Balance des blancs et Gray-World, Fusion par pyramide, ...). De plus, les techniques décrites dans le papier sont applicables à de nombreux problèmes de restauration d'images, en particulier les images dégradées par un brouillard ou par la pollution aérienne. Finalement, ce projet nous permettait de tester notre implémentation dans des cas réelles (typiquement : des images sous-marines trouvées sur internet). En effet, contrairement à certains projets où les données d'entrées sont nombreuses, ou dans un format particulier, la technique décrite dans l'article peut s'appliquer

aux photos individuelles.

L'article propose un post-traitement basé sur la fusion d'images afin d'améliorer la qualité de la prise de vue. Les applications d'un tel traitement sont nombreuses, tant pour l'affichage et la visualisation de ces images que pour des problématiques de visions plus intéressantes (tels l'extraction de points clés, la détection de contours et la reconnaissance de forme).

## II. MÉTHODE

La méthode décrite par l'article consiste en la fusion de deux images générées (appelées images d'entrées) à partir de l'image initiale. Les auteurs supposent que les dégradations de la prise de vue sous-marine sont du d'une part la balance des couleurs ainsi que les contrastes de la scène. Ainsi les deux images qui seront fusionnées sont chacune créées afin de résoudre un de ces problèmes.

## I. Balance des blancs

La première image du processus de fusion est créée afin de corriger les problèmes de balance de blancs qui existent lors des prises de vue sous marines. Comme nous pouvons le voir sur la figure, les informations de couleurs semblent à première vue être perdues. [[FIGURE]]

Nous pouvons toutefois en post-traitement corriger la balance des blancs. L'article suggère plusieurs méthodes, mais aucune n'est décrite en particulier. Nous avons donc choisi d'implémenter la méthode dite de *Gray-World*. L'idée sous-jacente de cette technique est de trouver l'*illuminant*  $i$ , c'est à dire la couleur de la lumière éclairant la scène. Une fois  $i$  trouvé, nous pouvons "soustraire" cette couleur à l'ensemble des pixels de la scène pour retrouver les couleurs réelles (*i.e.* en lumière blanche). L'hypothèse *Gray-World* suppose que la couleur moyenne d'une scène est un niveau de gris  $g$ . Ainsi si on mesure  $m$  la moyenne réelle des pixels de l'image :

$$m = \frac{1}{N} \sum_{x \in \text{image}} I(x)$$

On peut calculer  $g$  comme

$$g = \frac{m_r + m_g + m_b}{3}$$

L'article suggère une correction de la luminosité de scène pour compenser la perte de luminosité due au milieu

$$g' = 0.5 + \lambda g$$

où  $\lambda$  est un paramètre dont la valeur recommandée par les auteurs est 0.2. On en déduit alors la couleur de l'illuminant

$$i = \frac{m}{g}$$

Finalement, on corrige la couleur de la scène en modifiant la valeur du pixel

$$I(x) = \frac{I(x)}{i}$$

Nous avons ainsi corrigé l'image de telle sorte que la somme des couleurs soit effectivement une nuance de gris. L'image obtenu sera la première entrée du processus de fusion.

## II. Amélioration du Contraste

La seconde entrée du processus de fusion est quand à elle créée pour corriger les problèmes de contrastes dus à la diffusion de la lumière dans le milieu aquatique. Elle est obtenue à partir de l'image où les couleurs ont été corrigées. Elle est obtenue en deux étapes : réduction du bruit puis amélioration du contraste.

La réduction du bruit est effectuée à l'aide d'un simple filtre bilatéral. Ce filtre non linéaire remplace l'intensité de chaque pixel par une moyenne pondérée des pixels avoisinants, où les poids dépendent à la fois de la distance au pixel central, mais également de la différence d'intensité entre les pixels considérés. Ce bruit a la propriété de conserver les contours tout en supprimant correctement le bruit. Il est important de noter que si l'article conseille un débruitage cohérent dans le temps (pouvant donc s'appliquer à la vidéo), cela n'est pas nécessaire pour les applications à des images simples. [[FIGURE]]

Il existe de nombreuses méthodes permettant de réhausser le contraste de l'image, en particulier la méthode de *Local Adaptive Histogram Equalization* qui égalise localement des histogrammes, c'est-à-dire une transformation locale du contraste basée sur le voisinage du pixel. Contrairement à une méthode globale, on ne court pas le risque de saturer l'image tout en gardant une plage dynamique maximum. Toutefois, il est important de noter que cette technique est très sensible au bruit. Si le filtre bilatéral appliqué n'est pas assez puissant, il peut être nécessaire d'employer une technique plus robuste pour augmenter le contraste. Une solution consiste à employer une méthode similaire nommée *Contrast Limited Adaptive Histogram Equalization (CLAHE)*. Dans cette méthode, le contraste est limité, et les pics de l'histogrammes sont redistribués uniformément dans l'histogramme.

## III. Poids de fusions

Une fois les deux entrées générées, il est nécessaire de les fusionner de manière intelligente. Une simple moyenne des images (fusion naïve) n'est pas adaptée, en effet, sur certaines zones (contours, zones visuellement prééminentes, ...)

il est nécessaire de faire ressortir le contraste (et donc l'entrée numéro 2) au détriment des couleurs, alors que dans les zones constantes (aplats de couleur), il vaut mieux mettre en avant la correction de couleur. La fusion va donc s'effectuer en fonction d'un ou plusieurs poids, qui dénoteront l'importance que l'on souhaite appliquer à chaque image d'entrée dans l'image résultante, et ce pour tous les pixels considérés.

### III.1 Poids Laplacien

Il se calcul en appliquant un filtre laplacien sur chaque entrées du filtre. Il est censé mettre en valeur les contours et zones texturées, mais est peu adapté aux problèmes de restaurations d'images sous-marines, car il ne fait pas la différence entre zones de luminosité constante et zones de luminosité légèrement croissante (ou décroissante)

### III.2 Poids de Contraste Local

Ce poids a été créé pour mettre en avant les zones contrastées. Il se calcul en chaque pixel par la différence de la valeur de l'image I avec la valeur de I filtré par un passe-bas :

$$W_{LC}(x, y) = \|I(x, y) - I_f(x, y)\|$$

L'article recommande l'utilisation pour le filtre passe-bas d'un noyau binomial séparable de taille 5 x 5. Ce filtre est en effet une bonne approximation d'une Gaussienne, tout en restant algorithmiquement peu coûteux à calculer.

### III.3 Poids de Saillance

Le poids de Saillance cherche à mettre en valeur les objets visuellement proéminants qui peuvent exister dans une scène sous-marines. Une technique pour trouver la saillance d'un point et de comparer sa valeur avec une moyenne des valeurs des pixels l'entourant. Ce comportement est similaire au concept biologique de *center-surround contrast*, ou contraste centre-contour. Toutefois, l'article précise que ce poids favorise les zones lumineuses, et pour protéger les tons moyens, les auteurs suggèrent l'utilisation du poids d'exposition.

### III.4 Poids d'exposition

Cette carte de poids cherche à estimer si le pixel considéré est bien exposé, i.e. si sa luminosité est proche de la valeur moyenne 0.5. On utilise ensuite une gaussienne pour obtenir la valeur du poids :

$$W_{exp}(x, y) = \exp - \frac{(I(x, y) - 0.5)^2}{2\sigma^2}$$

### III.5 Normalisation des poids

Plusieurs poids sont suggérés, mais seulement deux (un par entrée) sont utilisées dans le processus de fusion. Pour normaliser la contribution de chaque image, on modifie les poids de tel sorte que la somme des deux cartes de poids soit constante et égales à 1. Cette modification est purement locale, et ne dépend que de la valeur des poids au pixel considéré.

## IV. Fusion par pyramide Laplacienne

L'article souligne qu'une fusion naïve des deux images, même en utilisant les poids donne des résultats visuellement peu plaisant, avec par exemple l'apparition de halos ou la distortion de couleurs. Les auteurs suggèrent donc l'utilisation d'une méthode classique de fusion d'image : la fusion laplacienne, ou fusion par pyramide laplacienne. Cette fusion s'effectue dans un espace échelle, et l'image finale est obtenue en combinant les deux images d'entrées et les deux poids dans cet espace échelle. En particulier, on crée des pyramides laplaciennes pour les images d'entrées, tandis que les poids sont utilisés pour créer des pyramides gaussiennes. finalement, les contributions de chaque niveau de la pyramide sont sommées pour obtenir l'image finale.

## III. IMPLÉMENTATION

Nous avons choisi pour l'implémentation de la méthode décrite dans l'article le langage C++, et utilisé la librairie de vision par ordinateur open-source **OpenCV**. Cette librairie fournit de nombreux outils accélérant la mise en place des différentes méthodes décrites par l'article. Le code finale fait en lui même un peu plus de 400 lignes, et est disponible en ligne sur [github](https://github.com).

## I. Choix effectués

Comme nous l'avons mentionné, l'article propose souvent plusieurs méthodes pour arriver au même résultat. Par exemple, la partie concernant la balance des blancs mentionne en particulier l'algorithme *Gray-World*, l'algorithme *Shades-Of-Grey* ainsi que les approches de Finlayson et Trezzi et l'hypothèse de *Grey-Edge*. Nous avons dans ces cas là fait le choix de n'implémenter qu'une des méthodes décrites pour pouvoir arriver le plus vite possible au résultat final, et comparer nos résultats à ceux montrés par les auteurs. Si nous avions eu un peu plus de temps, nous aurions probablement essayé de comparer les différentes méthodes proposées dans l'article.

Nous avons toutefois pris le soin d'implémenter tous les poids décrits par les auteurs, afin de pouvoir comparer leur influence sur le résultat final.

## IV. RÉSULTATS

### I. Aspect visuel et comparaison avec l'article

Nous avons retrouvé et testé notre implémentation sur les différentes images utilisées dans l'article. Nous avons également trouvé sur le site des auteurs leurs résultats en Haute Définition. Notre implémentation produit des résultats correctes, mais il existe quelques différences avec ceux obtenus par les auteurs de l'article.

Tout d'abord, les couleurs dans notre implémentation sont de moins bonne qualité. On voit souvent apparaître des tâches rouges, et la dynamique de couleur est moins importante. Cela est peut-être dû à un mauvais choix du paramètre  $\lambda$  (que les auteurs conseillent de fixer à 0.2), ou à un mauvais choix de l'implémentation de l'algorithme de balance des blancs (comme précisé précédemment, plusieurs options étaient mentionnées dans l'article, et aucune n'était clairement mise en avant).

Nous observons ensuite un contraste moins important dans nos images, mais également un bruit réduit. Ceci est peut-être dû à l'utilisation d'un simple filtre bilatéral pour le débruitage initial, qui limite alors l'augmentation du contraste. Un choix judicieux parmi les différentes

combinaisons de poids proposés par les auteurs pourrait également potentiellement améliorer les résultats de ce point de vue. En conclusion, nous obtenons des résultats comparables, mais de qualité inférieure. Ceci est attendu, car nous n'avons pas tenté d'optimiser pour chaque image les paramètres ( $\lambda$ , débruitage, poids, ...) pour obtenir le meilleur résultat, mais avons essayé de fixer les paramètres donnant un résultat correct sur chaque image du jeu de test.

### II. Vitesse d'exécution

La vitesse d'exécution est correcte. En moyenne, pour traiter une image de taille 1037x768, l'algorithme nécessite 1.9 secondes. Nous n'avons ainsi pas cherché à optimiser particulièrement la vitesse d'exécution de notre code. Plusieurs améliorations sont possibles à ce niveau. Nous aurions pu utiliser la programmation sur GPU, pour par exemple accélérer la fusion laplacienne ou la mise en place des *CLAHE*. Nous aurions pu également accélérer l'étape de balance des blancs en calculant la valeur de l'illuminant sur une plus petite portion de l'image.

### III. Mesure de la qualité du résultat

Si visuellement parlant, les résultats sont satisfaisants, une mesure de la qualité de nos résultats indépendante est nécessaire. Pour cela, l'article évoque que leur technique permet également d'améliorer les résultats d'autres méthodes, tels que la détection et le matching de features. Les auteurs remarquent que peut de points remarquables sont détectés sur les images initiales, mais que de nombreux points de ce types étaient obtenus sur les images corrigées. Nous avons donc voulu mesurer la qualité de nos résultats en comptant le nombre de SURF (*Speeded Up Robust Features*) trouvés sur les images.

# de SURF :	Init.	Amélioré	
DSCN6522	33	136	312 %
DSCN6523	25	66	164 %
DSCN6532	48	93	94 %
DSCN6533	22	86	291 %
DSCN6538_1	25	60	140 %
DSCN6539	70	97	39 %
DSCN6540	27	74	174 %
PICT0400	19	65	242 %
PICT0403	20	47	135 %
PICT0422	260	425	64 %
<b>MOYENNE</b>	<b>54,9</b>	<b>114,9</b>	<b>165,4%</b>

**Table 1:** *Nombre de SURF trouvés avant et après amélioration de l'image par notre implémentation.*

Nous voyons donc qu'en moyenne, 165% fois plus de features sont détectés après l'amélioration de l'image. Toutefois, ces résul-

tats sont faibles par rapport aux résultats de l'article, ou des chiffres de l'ordre de 800% sont atteints. On peut supposer que la mauvaise amélioration du contraste dans notre implémentation mène à une moins bonne détection des SURF, qui se basent notamment sur les contours, et la matrices Hessiennes, fortement sensible au contraste de l'image.

## V. OUVERTURES ET AMÉLIORATIONS POSSIBLES

### REFERENCES

- [Figueredo and Wolf, 2009] Figueredo, A. J. and Wolf, P. S. A. (2009). Assortative pairing and life history strategy - a cross-cultural study. *Human Nature*, 20:317–330.