

# Une solution d'acquisition d'images à multiples fonctions de réponse

Nicolas HAUTIERE<sup>1</sup>, Jérémie BOSSU<sup>1</sup>, Cyril BRAND<sup>2</sup>

<sup>1</sup>Université Paris-Est, LEPSiS, INRETS-LCPC  
58 boulevard Lefèbvre, 75015 Paris, France

<sup>2</sup>Université de Nice - Sophia Antipolis  
Polytech'Nice-Sophia

Nicolas.Hautiere@lcpc.fr, Jeremie.Bossu@lcpc.fr  
Brand@polytech.unice.fr

**Résumé** – Les capteurs CCD ont une réponse linéaire et produisent des images de bonne qualité, ce qui les rend encore incontournables pour certaines applications de vision. En revanche, leur dynamique est limitée par rapport aux capteurs CMOS dont la réponse est logarithmique. Pour contourner cette limitation, nous proposons une solution logicielle permettant d'exploiter une caméra CCD dépourvue d'objectif auto-iris en toutes conditions d'éclairage naturel. La méthode proposée est fondée sur la détermination du temps d'exposition optimal et sa régulation temporelle par un correcteur proportionnel. Le temps d'exposition optimal déterminé et régulé est utilisé pour construire des images à la dynamique souhaitée en fusionnant des images acquises à différents temps d'exposition. Nous disposons ainsi d'une solution d'acquisition d'images à multiples fonctions de réponse permettant d'alimenter différents types d'applications de vision.

**Abstract** – The quality of the images delivered by CCD sensors as well as their linearity with respect to the received light quantity still make them essential for some vision applications. However, their dynamic range is limited with respect to CMOS logarithmic sensors. To avoid this limitation, we propose a software solution allowing to use a CCD camera in outdoor conditions without any auto-iris lens. In this aim, a method to determine the optimum exposure time is proposed as well as a proportional correction rule to control it over time. Once this optimal exposure is found, we use it to build images with the desired dynamic range by fusing images grabbed at different exposure times. We have thus at our disposal a solution to grab images with multiple response functions allowing to feed different vision applications.

## 1 Introduction

Les capteurs d'imagerie de type CCD possèdent pour la plupart une dynamique limitée. Les constructeurs de capteurs et de caméras ont recours à différentes approches pour répondre à ce problème. D'un côté, ils équipent leurs caméras d'objectifs auto-iris qui permettent de réguler automatiquement la quantité de lumière parvenant sur les capteurs photosensibles. Malheureusement, ces dispositifs sont incompatibles avec un calibrage radiométrique correct du capteur [7]. Ils posent également des problèmes aux applications de vidéo-surveillance pour le calcul des images de fond [1]. D'un autre côté, ces fabricants développent des capteurs CMOS qui offrent une solution attractive bon marché. En effet, ces capteurs possèdent une réponse logarithmique, ce qui leur confère une dynamique plus élevée. Malheureusement, la qualité d'image reste encore inférieure à celle des capteurs CCD du fait de la présence d'un Bruit Spatial Fixe (BSF). De nombreux procédés sont publiés pour réduire ce problème comme dans [5], mais il limite néanmoins l'utilisation de ce type de capteur dans des applications de vision. Enfin, ces fabricants développent des capteurs à dynamique étendue (WDR) destinés à des marchés spécifiques, comme le marché de l'automobile en raison de leur coût élevé [4]. En outre, la non-linéarité des pré-traitements effectués en interne rend plus

difficile le calibrage radiométrique du capteur. En effet, la fonction de réponse du capteur peut varier dynamiquement en fonction des éléments présents dans la scène.

Dans le domaine de la photographie, la solution généralement adoptée est celle d'une optimisation du temps d'exposition, consistant à faire coïncider la moyenne de l'intensité de l'image à un gris moyen dans l'image résultante [8]. Les variantes de ce genre de procédé font l'objet de très nombreux brevets industriels.

Dans ce papier, nous proposons de construire une solution logicielle permettant simultanément de remplacer avantageusement un auto-iris par un réglage optimal de la caméra régulé dans le temps. Une fois la caméra réglée, nous proposons d'en augmenter la dynamique en fusionnant différentes images aux temps d'exposition différents. Ainsi, à partir d'une simple caméra 8 bits, nous proposons une solution d'acquisition d'image à multiples fonctions de réponse. Ce genre d'architecture permet de supporter plusieurs applications nécessitant différents réglages de la caméra. Ainsi, par rapport à l'approche classique où l'application s'appuie sur l'image par défaut que lui fournit le capteur, nous proposons une approche inverse où le réglage du capteur (temps d'exposition, gain, dynamique, cadence) est au service de l'application. Nous démontrons notre architecture à travers différentes acquisitions d'images alimentant des

applications de détection de brouillard ou de pluie par caméra fixe, applications nécessitant un calibrage radiométrique des caméras [3].

## 2 Réglage optimal d'une caméra

Nous concevons le contrôle automatique d'une caméra comme un procédé en deux étapes. Une première étape consiste à choisir le temps d'exposition optimal. Une deuxième étape consiste à réguler celui-ci dans le temps. Selon l'application, jouer sur le gain du capteur peut s'avérer plus intéressant, notamment à cause du flou de bougé. Dans notre cas, la caméra est fixe. Le flou de bougé n'est donc pas vraiment problématique. Nous préférons alors jouer sur le temps d'exposition de façon à minimiser le bruit d'acquisition.

### 2.1 Choix du temps d'exposition

Classiquement, les algorithmes de réglage du temps d'exposition sont basés sur la luminance moyenne de la scène. Les caméras CCD peuvent généralement être considérées comme linéaires. La luminance en entrée  $l$  est donc proportionnelle à la valeur d'intensité en sortie  $i$  :

$$l = \alpha i \quad (1)$$

Le réglage d'une caméra consiste à modifier la valeur de  $\alpha$  de façon à s'adapter aux niveaux de luminance présents dans la scène filmée. Cela revient à déplacer la dynamique du capteur  $d_{cam}$  le long de l'histogramme des niveaux de luminance (cf. figure 1).

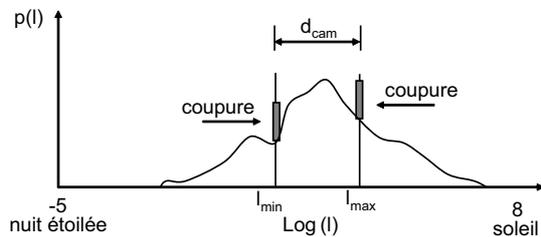


FIG. 1 – Histogramme logarithmique  $p(l)$  des niveaux de luminance d'une scène.

La dynamique de la caméra  $d_{cam}$  ne dépend pas de la valeur de  $\alpha$ . Nous définissons  $l_{min}$  et  $l_{max}$  comme les valeurs où la caméra atteint les intensités  $i_{min}$  et  $i_{max}$  pour une valeur de  $\alpha$  donnée. De façon à choisir une valeur optimale en terme de maximisation de l'information enregistrée, nous devons minimiser la sous-exposition et la sur-exposition. Cela revient à maximiser l'aire de l'histogramme des niveaux de luminance entre les bornes de la dynamique de la caméra [11] :

$$\alpha^* = \operatorname{argmax}_{\alpha} \int_{l_{min}}^{l_{max}} p(l) dl \quad (2)$$

Ainsi à l'initialisation, l'algorithme proposé consiste simplement à balayer la gamme des temps d'exposition possibles et

à retenir celui qui permet d'obtenir l'histogramme possédant une aire non saturée maximale. Ainsi, pour notre caméra qui permet de faire varier le temps d'exposition de 100 à 16384 microsecondes à une cadence de quinze images par seconde, la phase d'initialisation dure environ une seconde pour un pas d'échantillonnage d'une milliseconde.

### 2.2 Contrôle du temps d'exposition

Une fois le temps d'exposition optimal  $\alpha^*$  trouvé, il faut le réguler dans le temps de façon à s'adapter aux changements de conditions d'éclairage. Pour cela, nous nous appuyons sur la moyenne de l'image optimale désignée par  $I_m^*$  et la moyenne de l'image courante désignée par  $I_m(t)$ .

Ainsi, si la moyenne de l'image courante  $I_m(t)$  est proche de celle de l'image au pas de temps précédent  $I_m(t-1)$ , nous pouvons considérer que le réglage de la caméra est toujours correct. Aucune correction du temps d'exposition n'est donc nécessaire. Dans le cas contraire, nous appliquons une correction proportionnelle du temps d'exposition. Il reste à s'assurer la continuité de la solution.

Pour cela, nous proposons d'utiliser un correcteur proportionnel  $f$  possédant les propriétés suivantes :

$$\begin{aligned} f(0) &= 0 \\ f'(x) &= 0 \\ \lim_{x \rightarrow +\infty} f(x) &= ax + b \end{aligned} \quad (3)$$

$x$  représente ici l'écart absolu entre  $I_m(t)$  et  $I_m(t-1)$ . Une solution possible est donnée par :

$$y = f(x) = \sqrt{[d(ax - b)]^2 + 2abd}x + (1 - d)ax + bd \quad (4)$$

$a$  et  $b$  sont les coefficients de l'asymptote d'équation  $y = ax + b$ .  $d$  permet de régler l'éloignement aux deux asymptotes.  $d$  doit toujours être supérieur à 0,5 et plus il est proche de 0,5 plus la fonction colle à l'asymptote. Ainsi, lorsque l'écart entre la consigne et la mesure est faible (inférieur à  $-b$ ), la correction est nulle. Lorsque l'écart augmente (supérieur à  $-b$ ), le correcteur est assimilable à un correcteur proportionnel. La formule proposée répond bien aux exigences formulées précédemment. Des courbes représentatives de cette formule sont tracées sur la figure 2 pour différentes valeurs de  $d$ .

Ainsi l'algorithme de contrôle du temps d'exposition proposé s'écrit de la manière suivante :

$$\begin{cases} \alpha_0 = \alpha^* \\ |\alpha_{t+1} - \alpha_t| = f[I_m(t-1) - I_m(t)]. \end{cases} \quad (5)$$

Cette solution a pour avantages d'offrir une régulation plus stable que celle obtenue par l'emploi d'un auto-iris et de prendre en compte le calibrage radiométrique correspondant du capteur.

Pour éviter une éventuelle divergence de l'algorithme, un seuil sur l'écart absolu entre  $I_m(t)$  et  $I_m^*$  peut être utilisé afin de réinitialiser la chaîne de traitement complète (choix et contrôle du temps d'exposition).

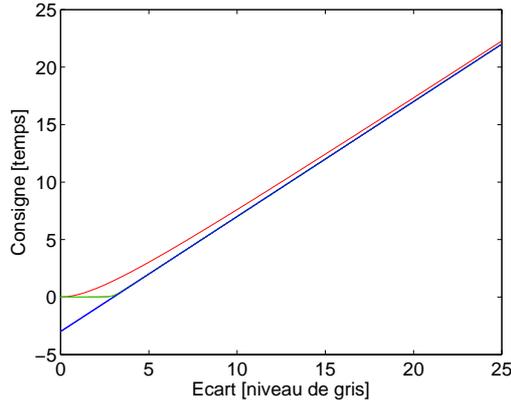


FIG. 2 – Courbes représentatives du correcteur proposé pour  $a = 1$ ,  $b = -3$  et  $d = 2,501$  (courbe rouge),  $d = 0,501$  (courbe verte). L'asymptote  $y = ax + b$  est tracée en bleu.

### 3 Imagerie à grande dynamique

La méthode précédente permet d'obtenir des images à faible dynamique (LDR) qui contiennent un maximum d'information au regard de la dynamique limitée du capteur utilisé. Néanmoins, selon l'application, la dynamique peut s'avérer insuffisante. Dans cette section, nous proposons d'augmenter la quantité d'information capturée par des techniques d'imagerie à grande dynamique (HDR).

#### 3.1 Principe

Pour construire une image HDR, l'approche classique consiste à fusionner plusieurs images acquises à des temps d'exposition  $\alpha_i$  différents [10] :

$$x_j = \frac{\sum_i w(x_i) \alpha_i g^{-1}(x_i)}{\sum_i w(x_i) \alpha_i^2} \quad (6)$$

Dans cette formule,  $g$  désigne la fonction de réponse de la caméra [7]. Dans le cas d'une caméra CCD, la fonction  $g$  est généralement la fonction identité, ce qui simplifie fortement les calculs. L'équation (6) devient en effet :

$$x_j = \frac{\sum_i w(x_i) \alpha_i x_i}{\sum_i w(x_i) \alpha_i^2} \quad (7)$$

Dans le cas d'une caméra linéaire, le calcul de la valeur d'un pixel d'une image HDR se résume donc simplement au calcul d'un barycentre, ce qui est très peu coûteux en temps de calcul.  $w(x_i)$  désigne la pondération associée à un pixel d'intensité  $x_i$ . Elle permet de ne pas prendre en compte les pixels saturés ou d'intensité nulle dans le calcul. Nous avons choisi d'utiliser la fonction suivante issue de [6] :

$$w(x) = 1 - \left(2 \frac{x}{255} - 1\right)^{12} \quad (8)$$

Cette fonction est tracée sur la figure 3.

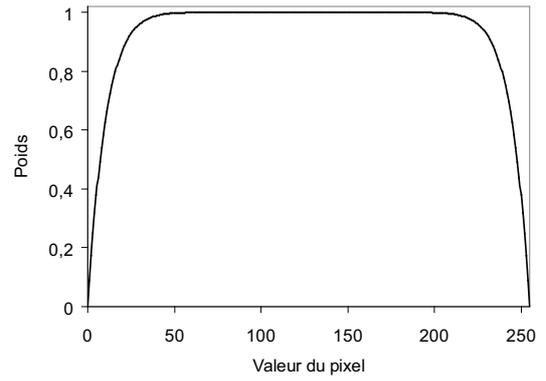


FIG. 3 – Poids des pixels dans le calcul de l'image HDR.

#### 3.2 Mise en œuvre

Pour mettre en œuvre la formule précédente, il faut choisir combien d'images on souhaite combiner et déterminer les temps d'exposition adéquats. Nous proposons différentes approches permettant de minimiser le nombre d'images à acquérir et donc de maximiser la cadence d'acquisition. La solution la plus simple consiste à combiner la séquence de temps d'exposition suivante :

$$(\alpha_i)_{i=-n\dots n} = (\alpha^* \times 2^i)_{i=-n\dots n} \quad (9)$$

où  $2n+1$  désigne le nombre d'images que l'on désire combiner.

Une autre approche consiste à effectuer une dichotomie de l'histogramme des niveaux de luminance présents dans la scène (cf. figure 1). La dynamique du nouveau capteur obtenu s'en déduit. Inversement, si l'on souhaite une dynamique ou une fonction de réponse précise, [2] donne des jeux de temps d'exposition à combiner.

### 4 Résultats



FIG. 4 – Exemples d'acquisitions d'images LDR (a) de nuit et (b) de jour.

Nous avons conçu une architecture logicielle qui nous permet de contrôler le temps d'exposition de notre caméra Dalsa Genie M1400 dépourvue de connecteur auto-iris. En effet, grâce à la liaison GigE entre la caméra et le PC d'acquisition, nous pilotons directement en boucle fermée le temps d'exposition

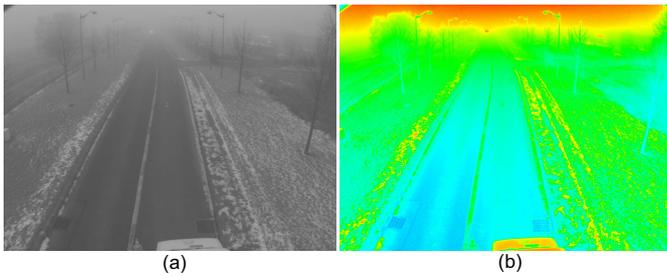


FIG. 5 – Exemple d’acquisition d’image HDR (a) compression  $\gamma = 2$  (b) fausses couleurs.

dans un module de traitement d’images. La caméra possède une cadence d’acquisition maximale de 30 images par seconde. Ainsi, si l’on construit des images HDR à partir de 3 images LDR, nous disposons d’un capteur fournissant 10 images LDR optimales et 10 images HDR par seconde.

La figure 4 présente différentes acquisitions d’images, respectivement une acquisition LDR de nuit par temps de brouillard et une acquisition LDR de jour par temps clair. Ces deux images ont été acquises lors de la même matinée à deux heures d’intervalle. Pendant cette période, le système n’a pas été réinitialisé. La qualité des acquisitions montre que le système s’est correctement adapté à la fois à la présence de l’éclairage public, à son extinction et à la levée du jour, sans parler des véhicules circulant sur la chaussée. La figure 5 présente une acquisition HDR à la levée du jour par temps de brouillard. La première figure présente l’acquisition réalisée avec une transformation  $\gamma = 2$  pour pouvoir être affichée. La deuxième image présente l’acquisition réalisée au moyen de fausses couleurs.

Les résultats obtenus sont satisfaisants tant qu’il n’y a pas trop d’objets mobiles dans la scène pendant la capture des images aux temps d’exposition variés. En effet, le mouvement dans la scène pendant la capture des images fait apparaître dans l’image HDR finale des éléments dits fantômes. Des solutions sont proposées pour filtrer ces fantômes, comme dans [6] ou plus récemment dans [9]. Les méthodes récentes assimilent le problème à une détection de mouvement. Il semble donc envisageable de bâtir une méthode de vidéo-surveillance à grande dynamique sur cette simple idée. En effet, il y a une certaine analogie entre la correction des éléments fantômes en imagerie HDR et le calcul des images de fond en vidéo-surveillance [12].

## 5 Conclusion et Perspectives

Dans ce papier, nous proposons une solution logicielle permettant simultanément de remplacer avantageusement un auto-iris et de disposer d’un capteur HDR à partir d’une simple caméra CCD 8 bits pilotable dépourvue d’auto-iris. Premièrement, nous proposons une méthode de détermination du temps d’exposition optimal ainsi qu’un correcteur proportionnel permettant de le réguler dans le temps. Nous utilisons ce temps

d’exposition optimal régulé pour construire des images HDR en fusionnant des images acquises à des temps d’exposition différents. Dans un proche avenir, nous souhaitons appliquer un algorithme de calcul d’images de fond à nos images HDR pour voir s’il est possible de construire une méthode de vidéo-surveillance HDR. La solution algorithmique proposée est également à même de piloter les caméras à double capteurs CCD qui commencent à apparaître sur le marché.

## Références

- [1] R. Cucchiara, R. Melli, and A. Prati. Auto-iris compensation for traffic surveillance systems. In *IEEE Conference on Intelligent Transportation Systems, Vienna, Austria, 2005*.
- [2] M. Grossberg and S. Nayar. High dynamic range from multiple images : Which exposures to combine ? In *ICCV Workshop on Color and Photometric Methods in Computer Vision, 2003*.
- [3] N. Hautière, E. Bigorgne, J. Bossu, and D. Aubert. Meteorological conditions processing for vision-based traffic monitoring. In *ECCV International Workshop on Visual Surveillance, Marseille, France, 2008*.
- [4] D. Hertel, A. Betts, R. Hicks, and Machiel ten Brinke. An adaptive multiple-reset CMOS wide dynamic range imager for automotive vision applications. In *IEEE Intelligent Vehicles Symposium, Eindhoven, The Netherlands, 2008*.
- [5] S. Kavadias, B. Dierickx, D. Scheffer, A. Alaerts, D. Uwaerts, and J. Bogaerts. A logarithmic response CMOS image sensor with on-chip calibration. *IEEE Journal of Solid-State Circuits*, 35(8) :1146–1152, August 2000.
- [6] E. Khan, A. Akyuz, and E. Reinhard. Ghost removal in high dynamic range images. In *IEEE International Conference on Image Processing, Atlanta, USA, 2006*.
- [7] T. Mitsunaga and S. Nayar. Radiometric self calibration. In *IEEE Conference on Computer Vision and Pattern Recognition, 1999*.
- [8] N. Nourani-Vatani and J. Roberts. Automatic camera exposure control. In *Australasian Conference on Robotics and Automation, Brisbane, Australia, 2007*.
- [9] M. Pedone and J. Heikkilä. Constrain propagation for ghost removal in high dynamic range images. In *International Conference on Computer Vision Theory and Applications, 2008*.
- [10] M. Robertson, S. Borman, and R. Stevenson. Dynamic range improvement through multiple exposures. In *IEEE International Conference on Image Processing, Kobe, Japan, 1999*.
- [11] S. Schulz, M. Grimm, and R. Grigat. Using brightness histogram to perform optimum auto exposure. *WSEAS Transactions on Systems and Control*, 2(2) :93–100, February 2007.
- [12] C. Stauffer and W. Grimson. Learning patterns of activity using real-time tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8) :747–757, 2000.