



De la caméra numérique à la photographie computationnelle

Henri Maître

2013D006

septembre 2013

Département Traitement du Signal et des Images
Groupe TII : Traitement et Interprétation des Images

De la caméra numérique à la photographie computationnelle

Henri Maître

Institut Mines-Télécom – Télécom-ParisTech – CNRS LTCI
46 rue Barrault – 75 013 Paris – France

9 septembre 2013

Résumé¹ : Les caméras numériques ont maintenant plus de 30 ans d'existence, elles ont atteint une popularité exceptionnelle à laquelle les appareils photographiques sur films ne sont jamais parvenus. Les appareils numériques se distinguent des appareils analogiques par un nombre important de particularités, certaines constituant de réels progrès face à la photo analogique, d'autres demeurant encore aujourd'hui des éléments de frein à leur généralisation dans tous les domaines techniques. Cependant ces particularités de l'imagerie numérique voient naître depuis quelques années des techniques nouvelles d'accès au monde qui nous entoure, techniques qui reposent fondamentalement sur les capacités offertes par ces capteurs. C'est le domaine de la photographie computationnelle. Avant de faire un tour d'horizon des perspectives les plus prometteuses de la photographie computationnelle, on fera un inventaire des spécificités des caméras numériques afin de dégager les caractéristiques qui permettent des ouvertures techniques et scientifiques déterminantes.

Mots-clefs : Appareil photographique, caméra, traitement d'images, photographie computationnelle

1 Définitions

Nous désignerons sous le nom d'**appareils photographiques** (ou de **caméras**) **analogiques** les systèmes classiques d'acquisition sur film, systèmes dont le prototype emblématique est le reflex 24×36 mm, mais qui couvrent également une grande variété de boîtiers mono-objectif bon marché ou, à l'inverse, de boîtiers de grand format (4×4 cm, 6×6 cm, etc.) ou de chambres professionnelles.

Nous désignons sous le nom d'**appareil numérique** tous les systèmes autonomes d'acquisition d'images par un capteur solide (CCD ou CMOS) et donc la très large famille allant des téléphones portables aux chambres professionnelles, en passant par les compacts, les bridges, les hybrides et les reflex.

Nous désignons par **traitement numérique** des images² les méthodes d'amélioration et de transformation qui partent d'une image numérique pour fournir une nouvelle image de même nature mais où certaines composantes sont mises en évidence. C'est le cas de toutes les techniques d'augmentation de contraste ou de résolution, de sélection chromatique, de filtrage fréquentiel, de détection de contours ou de régions, de reconnaissance des formes, etc.

Par opposition, ce que l'on appelle aujourd'hui l'**imagerie computationnelle** ou **photographie computationnelle** [16, 41] a pour objectif de fournir, à partir d'un appareil numérique et de traitements afférents, des images contenant des informations de nature différente de celles délivrées par le

1. Ce document est la version en français d'un exposé invité à la conférence *SPIE Eighth International Symposium on Multispectral Image Processing and Pattern Recognition - 26 -27 October 2013 Wuhan, China*.

2. Nous délaisserons très vite le terme « numérique » car notre propos ne concernera que ces traitements faits par ordinateur.

capteur lui-même, ou des informations dans un domaine différent de celui des mesures nominales du capteur. A ce titre, la tomographie de l'imagerie médicale, qui procède par inversion de la transformée de Radon, relève clairement de la photographie computationnelle. C'est aussi le cas de l'imagerie de radar à synthèse d'ouverture ainsi que de multiples techniques utilisées depuis des décennies en astronomie ou en géophysique par exemple. L'imagerie computationnelle a néanmoins pour objectif, à la différence de ces exemples juste cités, de rapprocher ces techniques de l'utilisateur en banalisant, autant que possible, le système d'acquisition, jusqu'à, pourquoi pas ?, les introduire sous forme de progiciel au sein des appareils commerciaux.

Comme on le voit, la limite entre traitement des images et imagerie computationnelle est très imprécise. Il n'y a pas de bon traitement d'image qui ne s'appuie sur la connaissance fine du capteur et il y aura dans toute technique d'imagerie computationnelle des briques logicielles qui seront directement issues du traitement numérique des images, mais l'on y fera également appel à des techniques développées en synthèse d'images ou en reconnaissance des formes.

2 Les appareils photographiques numériques *vs.* analogiques

Les appareils photo analogiques ont atteint, au cours des 150 ans qu'a duré leur suprématie sur la représentation de l'univers qui nous entoure, un perfectionnement remarquable et une qualité unanimement appréciée au regard de critères très variés, qu'ils soient scientifiques (résolution, sensibilité, fidélité, précision), techniques (robustesse, ergonomie, interopérabilité), voire artistiques ou sociétaux.

Il n'est donc pas étonnant que l'architecture des appareils numériques en découle aujourd'hui très fidèlement : en photographie, il n'y a pas la rupture de continuité entre numérique et analogique que l'on retrouve par exemple dans le domaine très voisin de la téléphonie.

Si l'on examine cependant attentivement ce qui est aujourd'hui disponible sur le marché, on constate que, sous des apparences très semblables, toutes les fonctions de l'appareil photo sont profondément modifiées. Ces évolutions laissent très largement la place à une évolution profonde des pratiques du photographe, même si cette évolution n'a pas encore véritablement eu lieu. Notons aussi que, si de nombreuses évolutions sont apparues, elles ne sont pas toutes compatibles entre elles et que, bien souvent même, elles s'excluent. Le monstrueux hybride qui tirerait profit de tous les avantages que nous allons souligner ci-dessous n'est probablement pas viable et ne se trouve bien sûr aujourd'hui dans aucun catalogue.

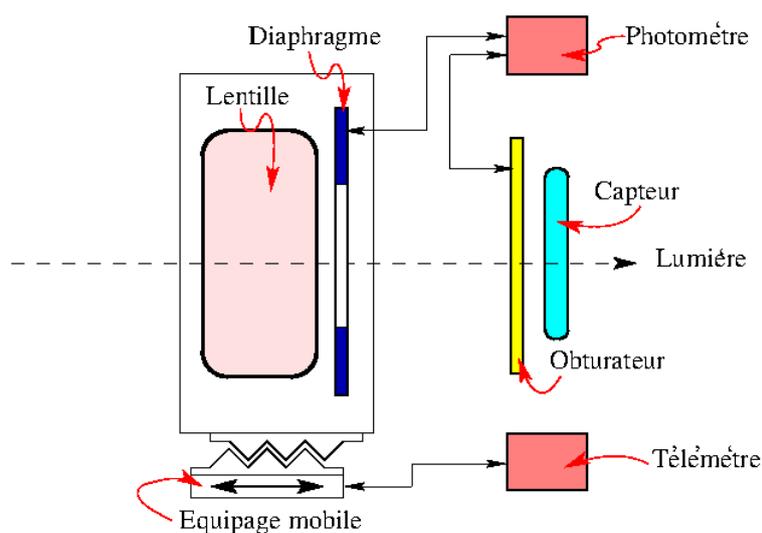


Figure 1 – Schéma de principe d'un appareil photographique, qu'il soit analogique (capture par film) ou numérique (capteur solide).

Très schématiquement, tout appareil photo est constitué d'éléments génériques qui sont toujours présents (voir la figure 1) :

- une lentille ou un objectif en charge de faire l'image réduite de la scène ;
- un capteur (film ou matrice de semiconducteur) ;
- un obturateur qui règle l'exposition ;
- un télémètre qui mesure la distance de mise au point ;
- un équipage mobile qui assure la conjugaison optique entre l'objet observé et l'image faite sur le capteur conformément aux instructions fournies par le télémètre ;
- un photomètre qui mesure l'éclairement reçu de la scène ;
- un diaphragme qui, associé à l'obturateur, contrôle la quantité de lumière reçue par le capteur pendant l'exposition.

Parmi les accessoires qui sont souvent associés à l'appareil et que nous utiliserons par la suite (voir la figure 2), notons :

- les filtres optiques (infra-rouge et ultraviolet) et anti-aliasing, placés devant le capteur ;
- le flash qui est une source de lumière additionnelle et contrôlée ;
- les traitements associés, qu'ils soient photographiques (tirage en laboratoire) ou numériques (les logiciels disponibles dans les appareils ou, éventuellement dans des logiciels déportés (type DxO ou PhotoShop)).

Nous examinons brièvement maintenant les évolutions majeures de chacune de ces fonctions.

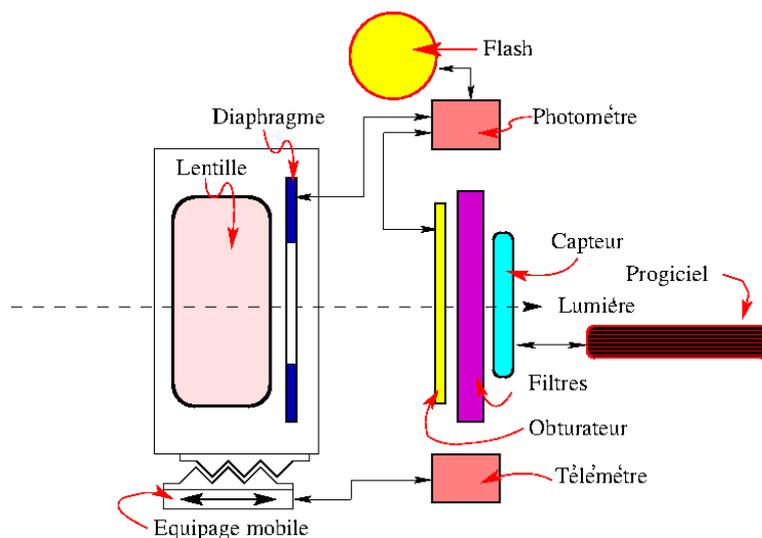


Figure 2 – Les compléments les plus souvent associés à l'appareil : des filtres optiques, un éclairage d'appoint, des logiciels de traitement d'image généralement portés par un microprocesseur embarqué.

2.1 Objectif et lentille

C'est en apparence l'un des domaines où les évolutions sont les moins notables. L'industrie de la photographie, réduite à un petit nombre d'acteurs, s'est efforcée de valoriser un existant très riche et très complet développé pour les appareils argentiques et d'assurer une raisonnable compatibilité avec les boîtiers numériques.

On a ainsi pérennisé une gamme d'objectifs (focales fixes ou zooms), étudiés pour les 24×36 , quand bien même ils étaient mal adaptés aux petits capteurs qui occupaient pendant longtemps le marché du numérique. L'émergence des capteurs numériques « plein-format » (c'est-à-dire également 24×36) permet aujourd'hui de pleinement justifier cette démarche. Notons quand même que pour les nombreux capteurs de petites dimensions, des optiques adaptées au champ couvert permettent généralement de gagner en qualité optique (piqué, ouverture, aberrations) à coût équivalent.

Nous reviendrons plus loin sur les conséquences de petits capteurs sur la qualité du signal acquis, mais toutes choses égales par ailleurs, on peut indiquer leur influence sur les optiques. Un plus petit champ couvert par le capteur pour une image identique, impose une focale plus courte ce qui a pour conséquence une optique plus légère et moins encombrante (aspect pratique) mais aussi, pour des raisons techniques, des grandes ouvertures plus faciles à concevoir (donc des optiques plus lumineuses) et une profondeur de champ plus grande (mise au point plus aisée et plus tolérante).

Poussant ces avantages à leur extrême, les systèmes photographiques les plus simples (dont beaucoup de téléphones) utilisent des lentilles fixes, sans mise au point, tirant profit de la très faible valeur de l'hyperfocale³.

Un tel choix (qui existait déjà dans des boîtiers argentiques très populaires comme le Brownie Kodak) simplifie bien sûr le montage de la figure 1.

D'autres évolutions ont tiré profit de la spécificité des montages optiques numériques non-réflex en proposant l'usage de lentilles additionnelles placées entre l'objectif photo et le boîtier. Alors que les lentilles additionnelles placées en amont de l'objectif (les «*bonnettes*» qui allongent les focales) font partie de l'attirail classique du photographe, ces lentilles ont le rôle inverse de raccourcir la focale (pour cela elles sont généralement incompatibles avec les montages réflex qui ne permettent pas de tirage assez court). Tirant très largement profit du champ trop large des optiques conventionnelles pour les capteurs de taille réduite, ces lentilles proposent souvent commercialement d'augmenter l'ouverture numérique de l'objectif utilisé (ce qui n'est bien sûr pas possible), tirant profit de la définition du nombre d'ouverture : $N = f/D$. Le diamètre D du diaphragme étant constant, passer de f à $f/2$ permet d'utiliser un objectif limité à une ouverture de 2,8 comme un objectif ouvert à 1,4 et donc de diviser par 2 le temps d'exposition nécessaire à une photo (ce qui est normal puisque la même énergie est regroupée sur moins de capteurs, chacun en recevant donc deux fois plus)⁴.

L'évolution majeure des systèmes optiques d'appareils photos est néanmoins à chercher dans les lentilles déformables qui font l'objet d'intenses recherches. L'idée de base pour assurer la mise au point n'est plus de déplacer des trains de lentilles les uns par rapport aux autres, mais de changer dynamiquement la courbure de la lentille⁵. Les méthodes mises en œuvre aujourd'hui utilisent des forces mécaniques ou électrostatiques agissant sur un matériau liquide, visqueux ou déformable. On espère ainsi gagner en complexité (pas de pièce mobile) et en rapidité dans la mise au point ou dans le choix du zoom.

2.2 Le capteur

C'est bien sûr la pièce maîtresse de la caméra numérique et celle qui a vu les plus grands progrès ces dernières années. Rien n'indique que les progrès soient achevés et l'on peut s'attendre à des évolutions aussi importantes dans un proche avenir que celles que nous avons vécues. Les progrès les plus notables concernent :

- la dimension des capteurs (taille du champ récepteur) ;
- la résolution, c'est-à-dire le pas entre deux sites voisins ;
- le nombre total de sites, c'est-à-dire le rapport des deux nombres ci-dessus ;
- la dynamique du signal ;
- le rapport signal sur bruit et donc plus ou moins la sensibilité ISO de l'appareil ;
- la cadence de prise de vue.

3. L'hyperfocale f_h est la distance focale minimale qui assure une mise au point jusqu'à l'infini. Un appareil réglé à l'hyperfocale est net de $f_h/2$ à l'infini. L'hyperfocale d'une lentille de focale f ouverte à N est donnée par : $f_h \approx \frac{2f^2}{\varepsilon N}$ où ε mesure l'erreur tolérée sur la mise au point.

4. Ces compléments d'optique sont souvent vendus sous des noms commerciaux de «*Turbo*» ou «*Booster*» <http://www.dpreview.com/articles/2667195592/first-impressions-metabones-speed-booster>.

5. Rappelons qu'une lentille sphérique dont les faces ont des rayons de courbures R et R' et dont l'indice est n a une focale f qui vérifie :

$$\frac{1}{f} = (n - 1) \left(\frac{1}{R} + \frac{1}{R'} \right) \quad (1)$$

La taille et la résolution des capteurs Brièvement, les capteurs sont de tailles très variées : de quelques millimètres dans les téléphones portables à plusieurs centimètres pour les plus grands, ils couvrent maintenant pratiquement toute la gamme des films photographiques jusqu'au moyen format. Cependant les capteurs de taille supérieure au 24×36 sont encore très honorés et réservés à des usages professionnels (voir la figure 3). La taille des sites photo sensibles s'est considérablement réduite, conduisant à de très fortes densités et une très forte intégration. Les capteurs de moins de 5 Mpixels ne se trouvent pratiquement plus que dans les téléphones, les appareils photo de qualité affichant tous au moins 15 Mpixels et parfois nettement plus.

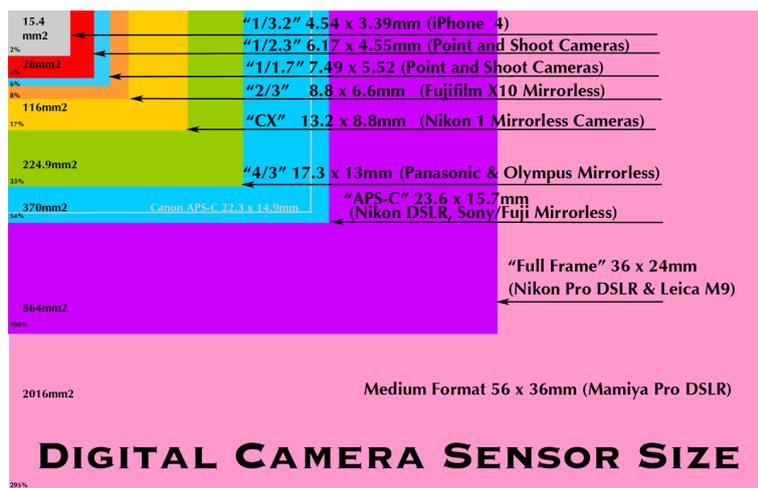


Figure 3 – La variété des formats des capteurs d'image disponibles en 2013 s'étale de 15 mm^2 à $2\,000 \text{ mm}^2$ (d'après Matt Sassamatt - 2011).

Le rapport signal sur bruit La dynamique du signal a beaucoup progressé également, avec des matériaux plus performants, des convertisseurs de meilleure qualité et d'excellents traitements de signal⁶. Des sensibilités ISO de 12 800 donnent aujourd'hui des images très correctes et certains appareils proposent des acquisitions au-delà de 100 000 ISO (rappelons que les films ne s'aventuraient guère au-delà de 1 600 ISO). Parallèlement, les 8 bits de quantification traditionnellement consacrés au pixel apparaissent étroits et nombre d'appareils proposent des images sur 10, 12 ou 14 bits (en mode RAW).

Le photodétecteur L'usage de capteurs CMOS en place des traditionnels CCD a permis de modifier les architectures d'acquisition en implantant les traitements en bout de ligne des photosites et donc en parallélisant la collecte et la conversion des signaux, ce qui permet d'accélérer notablement la cadence de prise de vue. On atteint ainsi aisément les rythmes de la Haute Définition en mode vidéo, ou des rafales d'une dizaine d'images en mode image. Notons ici que les réserves que soulevaient le CMOS quant à la qualité du signal d'image sont aujourd'hui très fortement contrebalancées par les avantages qu'il offre face au CCD en matière de circuiterie interne. Contre toute attente, c'est l'avenir des CCD qui est aujourd'hui incertain.

L'acquisition de la couleur Ce champ est commercialement stabilisé mais scientifiquement toujours ouvert. Le choix d'un maillage spatial de la mesure chromatique a été adopté pratiquement universellement et la matrice de Bayer (qui attribue 2 sites au Vert pour 1 au Bleu et 1 au Rouge) s'impose très majoritairement. Pourtant les solutions en pleine couleur (capteur Foveon de Sigma ou détecteur "True Color" de Nikon) présentent bien des atouts en accroissant notablement la densité

6. Ces améliorations sont indispensables compte tenu de la miniaturisation de plus en plus forte des sites, qui conduit à la détection d'un très faible nombre de photons.

d'échantillons trichromatiques. Moins radicales, des variantes de la matrice de Bayer (voir les figures 4 et 5) ont aussi montré des qualités remarquables, soit en proposant des distributions différentes des sites (solution en bande ou matricage EXR de Fuji), soit en proposant des sélections chromatiques différentes (capteur Kodak incluant des sites « panchromatiques », capteur DSCF628 de Sony utilisant quatre primaires).

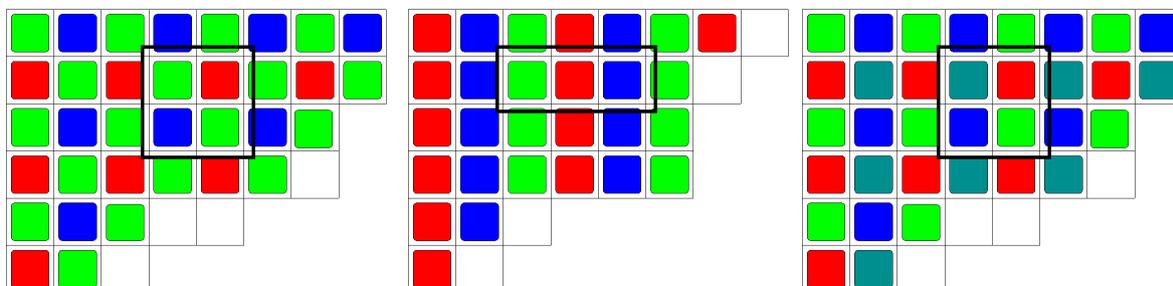


Figure 4 – De gauche à droite : matrice de Bayer, matrice bande, matrice de Sony à quatre primaires (un site Vert sur deux est changé en Vert-Jaune - Emerald en anglais.) Les zones marquées en noir indiquent la cellule élémentaire de périodicité du capteur.

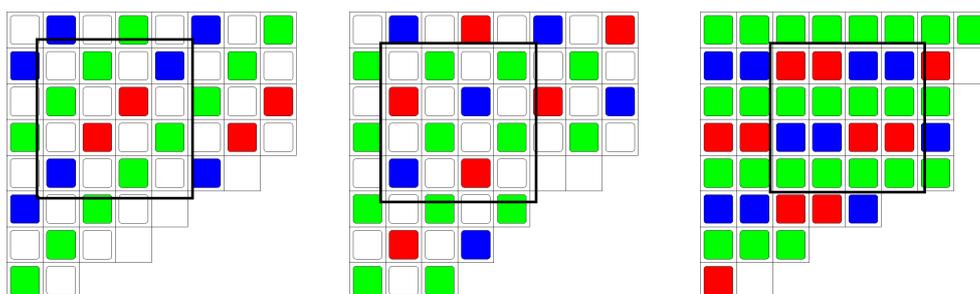


Figure 5 – De gauche à droite : deux matrices Kodak avec sites panchromatiques (en blanc), matrice ESR de Fuji, reconfigurable selon trois modes.

2.3 L'obturateur

Traditionnellement mécanique, il est devenu très souvent aujourd'hui électronique, ouvrant (n'est-ce pas le rôle de l'obturateur comme son nom ne l'indique pas?) des perspectives nouvelles et encore peu exploitées. Une commande électronique permet de traiter séparément les pixels pour peu que l'on trouve des lois de commande intéressantes. Elle permet également de proposer des temps d'exposition plus courts et des synchronisations plus faciles avec le flash. Enfin la combinaison d'ouvertures mécaniques et numériques permet encore d'autres effets particuliers. Il y a là un vaste espace de recherche encore peu exploré.

2.4 Le diaphragme

C'est pour le coup l'élément qui a le moins évolué et qui demeure le plus semblable à celui de la photo conventionnelle. Nous verrons cependant plus loin qu'il offre des ressources qui pourraient être développées plus encore [31].

2.5 Le télémètre

Nous appelons ainsi toutes les techniques permettant de déterminer la distance à l'objet qu'elles soient par superposition d'image, par contraste, par phase ou par mesure de vol. L'exploitation du signal numérique (sur le capteur lui-même ou plus sûrement sur des capteurs annexes associés) a permis de proposer des outils extrêmement puissants de mesure de la distance, en adoptant des stratégies complexes d'optimisation : en un site précis, autour d'un site ou dans l'ensemble de l'image. L'utilisation conjointe ou alternée de modes différents de mesure (actifs ou passifs) permet de résoudre des problèmes difficiles de mise au point en très faible lumière ou sur des surfaces lisses ou périodiques. De nombreux appareils proposent plus de 100 points de mesure dans l'image.

2.6 Le photomètre

Utilisant des solutions assez semblables à celles du télémètre (des capteurs numériques répartis dans le champ de l'image), la photométrie conduit cependant à des solutions plus globales où une pondération harmonieuse des diverses zones du champ est recherchée. Sa tâche se complique cependant car il associe souvent à ses objectifs de mesure énergétique la recherche plus fine de la balance chromatique en déterminant, à la volée, la nature de la source dominante de lumière.

2.7 Les filtres

En plus des filtres chromatiques (que nous avons associés au capteur), il est indispensable d'utiliser des filtres « chaud » et « froid » qui écartent les rayonnements ultra-violet et infra-rouges (surtout) du capteur car le CCD comme le CMOS présentent une très forte sensibilité hors du domaine visible (voir la figure 6).

Moins indispensables sont les filtres anti-aliasing qui élargissent la réponse impulsionnelle du système imageur pour se rapprocher des conditions d'échantillonnage de Shannon.

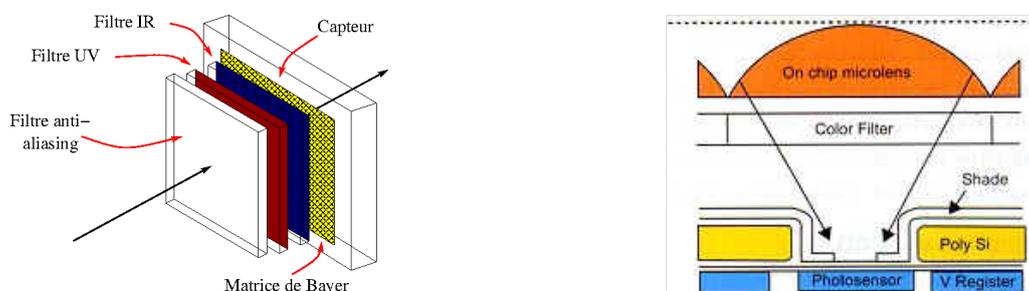


Figure 6 – Les filtres optiques placés devant le capteur : filtre « chaud » (IR), filtre « froid » (UV), filtre anti-aliasing (à gauche). La matrice de sélection chromatique est généralement constituée de filtres passe-bandes en longueur d'onde, associés à des micro-lentilles (à droite).

2.8 Logiciel et progiciel

L'importance des fonctions de calcul dans un appareil de photo numérique est immense. Le processeur est en charge de la gestion de l'interface avec l'utilisateur (menus, courbes, affichages), de la mémoire ainsi que les divers capteurs. Il gère les fonctions d'acquisition d'image, ainsi que celles de compression (Ziv-Lampel, JPEG, MPEG4, etc.). Il prend la plupart des décisions concernant l'ouverture, le temps de pose, la balance des blancs, la mise au point, la stabilisation de l'image à partir des capteurs qu'il commande et sans recours à l'utilisateur le plus souvent ! Il assure un premier niveau de traitement d'images : correction d'aberration, de vignettage, de colorimétrie. Il propose des fonctions de poursuite de cibles reconnues, de reconnaissance de visages, de contexte (intérieur/extérieur,

paysage/portraits, etc.) et l'on voit régulièrement apparaître de nouvelles fonctionnalités⁷. La puissance embarquée aujourd'hui dans un appareil numérique haut de gamme réclame un processeur très performant, souvent à plusieurs cœurs, parfois dédoublé. L'introduction de fonctionnalités nouvelles (GPS, gyromètres, liaison Wifi ou BlueTooth) rendra cette puissance de plus en plus nécessaire.

Notons, et cela fera notre transition avec la section concernant la photographie computationnelle, que beaucoup d'appareils modernes sont l'objet de mises à jour, en ligne, de leur logiciel ; ces mises à jour, assurées par le fabricant, visent toutes les fonctions de l'appareil mais aussi l'optique (commande de mise au point, de diaphragme, de zoom, éventuellement stabilisation). Un appareil photo peut ainsi offrir des compressions JPEG différentes, ou des stratégies de mise au point plus raffinées au cours de sa vie. Ils offrent donc la capacité de recevoir des logiciels complémentaires ou substitutifs et de nombreux professionnels ou amateurs se sont engouffrés dans la brèche. On trouve ainsi en ligne des fournisseurs d'applications spécifiques, qui proposent des algorithmes de mosaïquage ou de mise au point ou des effets spéciaux (solarisation, inversion de contraste, etc.). Par ailleurs, en vue de favoriser l'accès libre aux divers organes de l'appareil photo, des plateformes sont proposées par certains fournisseurs qui sont en fait de vrais appareils mais où toutes les fonctions sont rendues très accessibles et où les composants optiques sont aisément remplaçables. C'est en quelque sorte une version *open source* de l'appareil photo⁸.

3 La photographie computationnelle

Accroissement de la résolution, du nombre de pixels, de la sensibilité, accroissement du nombre de niveaux de quantification du signal, souplesse dans la mesure de la distance, de l'énergie incidente, souplesse dans l'analyse du plan image, contrôle centralisé des fonctions d'ouverture, de diaphragme, de zoom, stabilisation du plan image, intégration de prétraitements, interfaçage convivial avec l'utilisateur et avec le monde numérique, tel est le bilan aujourd'hui de la photo numérique.

Ce n'est pas peu et c'est en tous cas suffisant pour prendre une très grosse part du marché actuel de la photo (mais pas toute!), pour s'introduire dans des produits nouveaux (tablette, ordinateur, téléphone), pour bouleverser l'instrumentation scientifique ainsi que les produits industriels. C'est incontestablement une très grande évolution technologique conduisant à des évolutions économiques et sociétales considérables. Pourquoi alors l'émergence d'une *Computational Photography* qui semble arriver bien tard et qui ne fera rien de plus au succès de la photographie numérique?

hormis les vertus commerciales de l'innovation technologique, la réponse est peut-être à chercher dans les perspectives nouvelles offertes par ces fonctionnalités en dehors du marché immédiat de l'image, dans le domaine scientifique tout d'abord, dans le domaine instrumental ensuite.

L'émergence de la photographie computationnelle provient du constat que ces fonctionnalités nouvelles – admirablement exploitées par le monde de la photographie – peuvent faire plus que « de belles photos » et qu'il est possible de les utiliser, en détournant un peu les composants et les accessoires, pour accéder à des informations nouvelles qui elles-mêmes apporteront dans l'avenir des progrès importants : dans le domaine de l'image peut-être mais aussi dans d'autres domaines encore très mal identifiés. La Photographie Computationnelle, c'est donc un pari scientifique et il est loin d'être gagné. Nous allons en examiner ici quelques facettes révélant certains des objectifs visés et esquissant les pistes envisagées.

3.1 « Full light field »

Une très grande attention a été apportée ces dernières années à deux problèmes très liés, mais cependant bien différents :

7. en 2013 : reconnaissance simultanément de plusieurs visages, reconnaissance de sourires, reconnaissance de chiens et de chats, etc.

8. voir par exemple *The Franken camera* <http://graphics.stanford.edu/papers/fcam/>.

- l’obtention d’une mise au point dans tout le champ de l’image, quelle que soit la position des objets (ce que l’on appelle l’imagerie plénoptique),
- la détermination simultanée de l’image d’un pixel et de sa distance à la caméra (image 3D).

Il s’agit ainsi de mesurer par l’image tout le volume observé, à la façon que le fait l’holographie. Cet objectif ne manque pas d’applications potentielles et à ce titre il a conduit déjà à des produits industriels (encore balbutiants) Raytrix [39] ou Lytro [18], et à des prototypes proches de l’industrialisation (caméras ADOBE, MERL ou CAFIDIS).

Parmi les solutions proposées pour ces deux problèmes, la mesure d’une pile d’images, chacune à profondeur de champ limitée, couvrant tout le champ des mises au point des divers objets de la scène, est apparu comme un objectif intermédiaire intéressant. Ce point a en particulier été exploré dans [24, 38]. Les piles d’images peuvent être obtenues de diverses façons :

- en suivant la proposition pionnière de G. Lippman en 1908, par des matrices de micro-lentilles [33] : c’est ce que proposent les auteurs de [29],
- par déplacement du capteur le long de l’axe dans l’espace image [32], ou symétriquement par déplacement de l’optique de la caméra [36],
- par utilisation d’ouvertures codées sur la lentille qui permettent de sélectionner des rayons d’une orientation particulière par rapport à l’axe optique [31, 32, 40],
- par détermination du flou de mise au point, sujet classique [37], abondamment revisité ces dernières années [14, 27],
- par l’exploitation des aberrations des systèmes optiques [43] ou par codage chromatique [5, 22, 45],
- par la combinaison de plusieurs de ces approches.

Les problématiques associées à ces piles d’images sont de divers types. Comme nous l’avons indiqué, le calcul d’une image parfaitement focalisée en tout point du champ ainsi que la détermination de la distance des objets à la caméra sont les demandes les plus fortes. D’autres applications concernent la reconstruction d’un plan image particulier, le suivi spécifique d’objets en approche, la reconstruction d’images sur des surfaces non planes, etc.

3.2 Images à haute dynamique : « *High Dynamic Range = HDR* »

Les scènes observées pour de nombreuses applications présentent une dynamique très forte entre les plages les plus claires et les plages les plus sombres. Le système visuel humain est remarquablement adapté à suivre les excursions de luminosité importantes, encore faut-il que ces informations lui soient transmises. La quantification sur 8 bits par canal que proposent la plupart des appareils, ainsi que les standards de transmission sont insuffisants pour reproduire ces fortes dynamiques. Deux grandes familles de travaux sont donc attachés à ces signaux HDR : comment les acquérir, et comment les visualiser en maintenant la dynamique (problématique dénommée *tone mapping*).

De nombreux travaux, comparés dans [2], s’attachent spécifiquement à étendre ou à reconstruire la réponse des capteurs par diverses stratégies : utilisation de voies séparées pour les fortes et les faibles luminosités, multiplexage temporel utilisant des séquences d’obturation contrôlées, dynamiques aléatoires suivies de reconstructions, expositions codées pixel par pixel, etc. [4, 21, 26].

Notons que les architectures nouvelles des matrices de capteurs (Kodak ou Fuji, figure 5) ont été conçues tout autant pour améliorer la résolution que pour faciliter l’obtention d’images à hautes densités ou d’images dans des conditions d’illumination difficiles.

3.3 Résolution étendue

Quoique les capteurs soient de plus en plus performants, une importante famille de travaux concerne l’amélioration de leur résolution. Ils abordent le problème avec des angles d’attaque différents : certains s’intéressent à l’opération de démosaïquage, d’autres à la restauration au-delà de la fréquence de coupure à partir d’une connaissance fine (sub-pixellaire) de la réponse impulsionnelle et du bruit,

d'autres enfin au filtrage optique de l'aliasing.

Démosaïquage Les principes et les limitations de l'acquisition par une matrice de photo-détecteurs s'appliquent à toutes les caméras numériques, et interdisent l'emploi des outils classiques issus de l'approche de Nyquist-Shannon. En l'absence d'un filtrage optique fort, tout capteur numérique conduit à du repliement de spectre (figure 7 à gauche), et si l'on choisit un matricage chromatique de type Bayer, on accroît encore considérablement ce repliement (en particulier pour les canaux Rouge et Bleu) (figure 7 à droite).

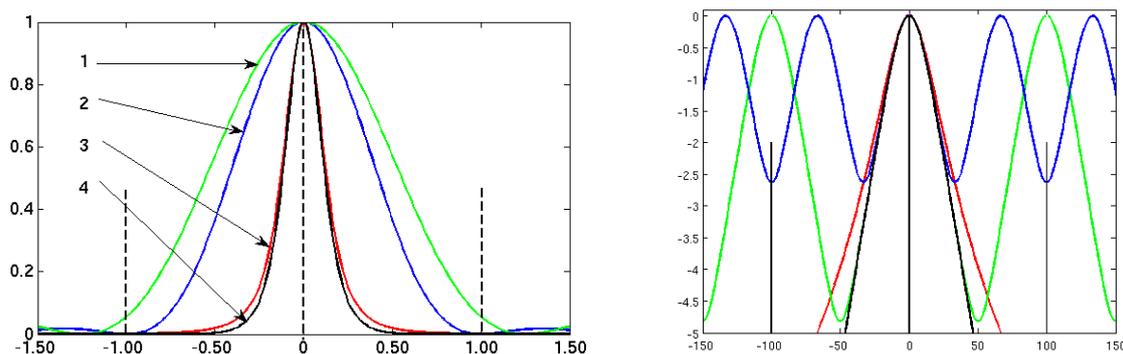


Figure 7 – Spectre de densité de puissance de l'image. A gauche : avant périodisation par la matrice de photosites : le spectre du signal de l'image (courbe 1), maximale à la fréquence nulle, le filtre imposé par le diaphragme (courbe 2) et celui imposé par l'intégration sur le capteur (courbe 3). La période du capteur définit les fréquences utiles dans la bande $u \in [-1, 1]$ par rapport à la fréquence d'échantillonnage par les photosites. Le signal résultant des deux filtrages, avant périodisation par la matrice des capteurs, est représenté par la courbe 4. Nous avons choisi un spectre lorentzien correspondant à une auto-corrélation exponentielle dans l'image, un diaphragme dont la réponse impulsionnelle s'annule à $1,22$ le pas d'échantillonnage et une cellule sensible qui occupe les $3/4$ de la cellule de répétition de la mosaïque.

A droite : après matricage (et en échelle logarithmique) : En noir et rouge, les deux signaux (en noir et rouge) de la courbe de gauche. En vert, l'image telle qu'elle se présenterait après intégration sur un photosite (sans zone de garde) d'une matrice de Bayer (canal R ou B) et échantillonnage au pas des photosites : il y a un repliement de spectre de $1/2$. En bleu : spectre après un matricage chromatique en bandes : le repliement est de $2/3$.

Une attention très forte a été portée au problème de démosaïquage. Tout d'abord les fabricants d'appareils proposent des algorithmes généralement d'excellente qualité (mais brevetés), tirant profit des différences d'échantillonnage et des corrélations des canaux, mais aussi des motifs présents dans l'image et du bruit estimé (connu en fonction des conditions d'exposition et des paramètres choisis). Au-delà de ces solutions qui fonctionnent naturellement en temps réel et de façon native dans les appareils, des travaux récents proposent d'étendre encore ces résultats par des démarches computationnellement plus lourdes mais particulièrement puissantes qui examinent très précisément la continuité figurale des divers canaux et tirent profit de la régularité des surfaces et des textures par la mise en œuvre par exemple de filtrages non-locaux (voir figure 8) [7, 19].

Superrésolution Quoique le terme soit un peu tombé en désuétude, l'objectif de la superrésolution reste, comme nous venons de le voir, en permanence en ligne de mire des travaux sur la résolution ultime des capteurs. Par le fait que les capteurs ne satisfont pas aux conditions d'échantillonnage de Shannon, ils laissent dans le signal numérique des informations caractéristiques des hautes fréquences. Lorsque cette information est pernicieuse, elle est appelée aliasing, lorsqu'on réussit à l'exploiter on fait de la superrésolution. La superrésolution apparaît vraiment accessible lorsque l'on dispose d'une diversité de mesures : c'est le cas par exemple avec les trois canaux chromatiques et les algorithmes aujourd'hui disponibles font effectivement pour chaque canal une vraie superrésolution. Lorsqu'on

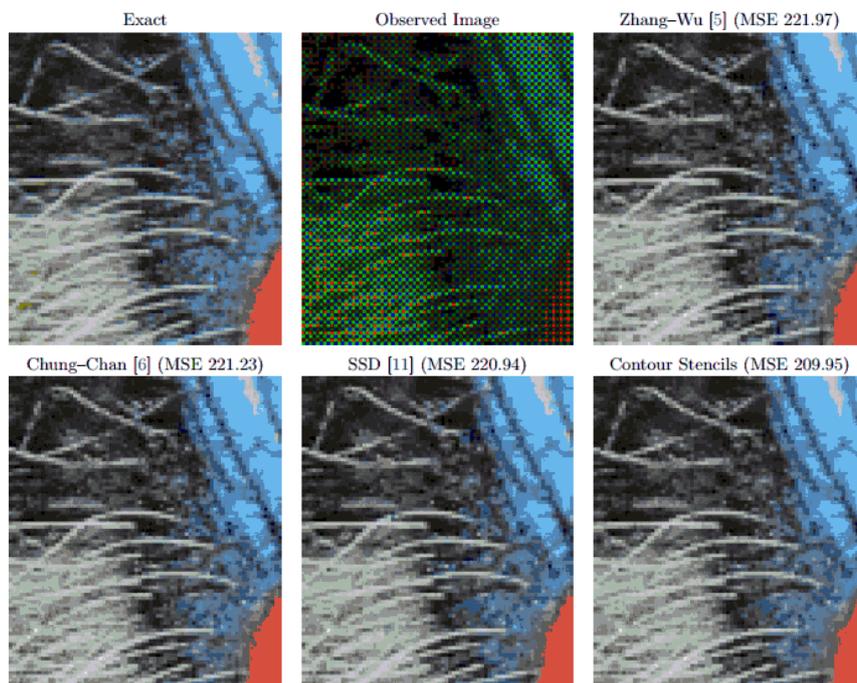


Figure 8 – *Démosaïquage* - de gauche à droite : en haut image originale, image brute après acquisition par une mosaïque de Bayer, image reconstruite par l'approche de Zhang et Wu [46], en bas : par l'approche de Chung et Chan [12], de Buadès [7], de Getreuer [19]. [d'après la bibliothèque IPOL [19]]

dispose d'acquisitions multiples, mettant en avant des différences dans le signal, soit par mouvement de l'objet, soit par déplacement du capteur, soit par balayage en mise au point, on se remet dans des conditions favorables qui sont aujourd'hui abondamment exploitées. À côté de nouvelles approches théoriques s'appuyant sur les représentations parcimonieuses [8], des propositions intéressantes testées sur des schémas existants, utilisent des distributions aléatoires des capteurs ou des formes aléatoires des cellules des photosites pour assurer l'orthogonalité des mesures et préserver des capacités de reconstruction avec super-résolution [42].

Réponse impulsionnelle et bruit D'importants travaux sont consacrés aujourd'hui à la caractérisation fine du bruit dans les capteurs et conduisent à des algorithmes de filtrage, souvent propriété des fabricants d'appareils, d'excellente qualité. L'amélioration de la sensibilité des capteurs à « hauts ISO » que l'on constate aujourd'hui en est directement le produit. La caractérisation du bruit dans l'image [3, 34], ainsi que l'identification de la réponse impulsionnelle avec une précision sub-pixel [14] sont des étapes importantes pour cela, qui permettent ensuite la mise en œuvre d'algorithmes puissants de reconstruction. L'utilisation de mires adaptées à ces étapes de mesure de la réponse impulsionnelle semble aujourd'hui une voie simple vers l'optimalité de ces mesures [14, 25] (voir illustration figure 9). Il faudra probablement, à terme, se passer de ces assistances et déterminer la réponse impulsionnelle directement sur l'image à partir des seuls motifs présents [13].

Une autre piste consiste à adapter les filtres anti-aliasing (voir paragraphe 2.7) à la matrice de détecteurs. Ces filtres sont aujourd'hui montés en série sur les appareils ; ils exploitent les propriétés biréfringentes de matériaux comme le niobate de lithium pour dégrader la réponse impulsionnelle. Deux lames biréfringentes croisées permettent ainsi de remplacer la figure de diffraction par 4 taches, judicieusement décalées au regard du pas du capteur. De nombreuses propositions visent à les modifier ou à les supprimer pour permettre un contrôle meilleur de la réponse impulsionnelle (voir la figure 10).

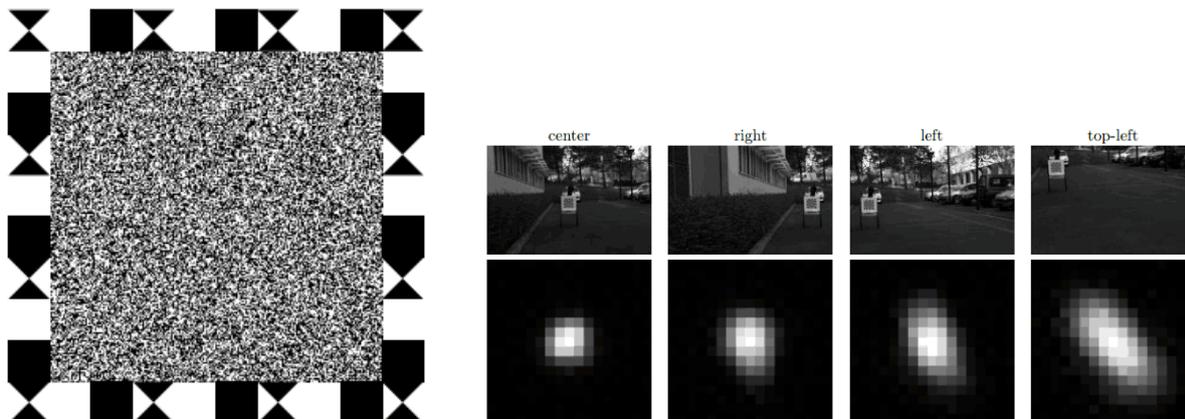


Figure 9 – Mire de résolution (à gauche) et exemples d'estimation de la réponse impulsionnelle à droite, d'après [14].

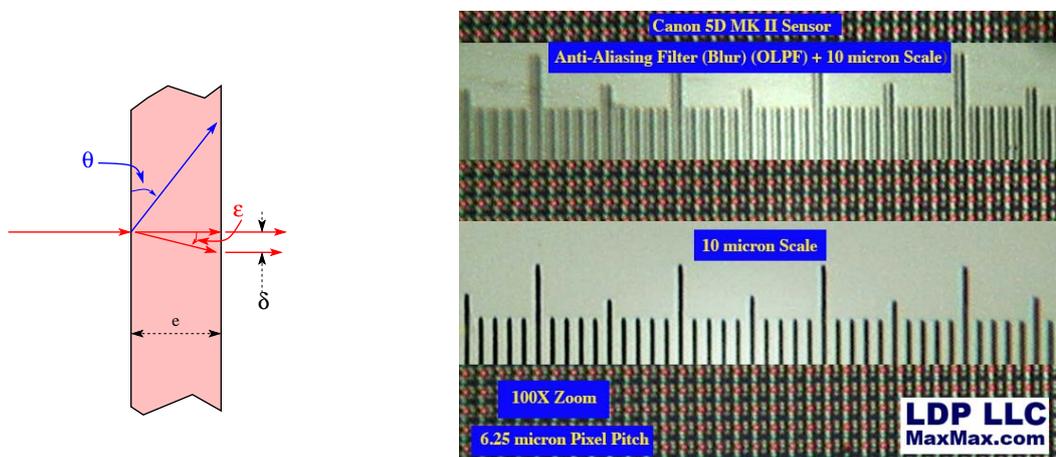


Figure 10 – Filtre anti-aliasing : à gauche, traversée du cristal de niobate de lithium par le faisceau, donnant naissance à deux faisceaux décalés. A droite : effet du filtre anti-aliasing. La partie haute est en présence d'une épaisseur de filtre anti-aliasing : l'image est alors dédoublée à une distance très voisine du pas des photodétecteurs. En bas l'image est sans filtre anti-aliasing et présente donc une image beaucoup plus nette (©MaxMax : <http://www.maxmax.com/>).

3.4 La couleur

C'est probablement l'un des points les plus délicats de la formation de l'image et ce n'est malheureusement pas celui qui retient le plus l'attention des chercheurs actuels. Alors que les produits disponibles sur le marché donnent des résultats souvent très bons, et couvrent en particulier fort bien l'essentiel des demandes du grand-public, des points faibles demeurent dans l'acquisition des images, comme la reproduction des couleurs (rendu des couleurs rares : bleus profonds, violets ou pourpres), la détermination automatique de la nature des sources ou la balance des blancs. Ces faiblesses rendent très difficiles certaines applications comme la reproduction de haute qualité, la caractérisation colorimétrique des matériaux, le changement d'illuminant, ainsi que certaines animations graphiques. Des travaux permettent de déterminer la nature des illuminants [17, 20, 35]; ceux entrepris sur les capteurs (avec des canaux panchromatiques (Kodak) ou des primaires plus nombreux (capteur Sony, voir figures 4 et 5) ouvrent des pistes malheureusement très peu suivies. Le recours aux techniques hyperspectrales [28, 10, 11] qui font le pendant dans le domaine des longueurs d'ondes des approches plénoptiques pour la focalisation sont encore très ignorées de la communauté de photographie com-

putationnelle.

3.5 Gestion du mouvement (*flutter shutter*)

La gestion du mouvement de la scène est un problème classique de la photographie qui se résout traditionnellement en raccourcissant le temps de pose de façon à « figer » l'objet dans une position donnée. Cela se fait au détriment du bilan énergétique, ce qui limite en pratique son emploi, sauf à recourir à des sources de lumière intenses (les flashes) qui fixent alors par leur durée le déplacement de l'objet. D'intéressants effets (surtout artistiques) ont été créés en photographie classique en choisissant des obturateurs particuliers (obturateurs à rideau par exemple), ou en combinant exposition longue (par l'obturateur) et éclairage bref (par le flash), ainsi qu'en choisissant la position particulière du flash (ou sa répétition) à l'intérieur de l'ouverture. Ces travaux pourraient être repris aujourd'hui au regard des progrès des techniques électroniques de contrôle de l'obturateur qui permettent de programmer des « trajectoires » particulières de l'obturation.

L'introduction des techniques de stabilisation mécanique (du capteur ou de l'optique) est une autre solution qui permet aujourd'hui de repousser notablement les problèmes de stabilité de la prise de vue et d'aborder corrélativement les mouvements de la cible dans cette stabilisation.

Couplées avec les algorithmes de poursuite qui sont également embarqués dans les appareils (par exemple pour le *block matching* des vidéos), ces techniques ouvrent la voie à des méthodes dynamiques d'acquisition qui pourraient être moins sensibles aux déplacements des cibles.

C'est un peu sur cette piste que se sont engagées des études sur les « obturateurs battants » (*flutter shutters*) [1, 23] ou encore ceux sur la photographie invariante au mouvement de [30] qui ouvrent très largement les possibilités de commander l'ordre et la durée d'ouverture des obturateurs de façon à introduire un flou de mouvement contrôlé dont l'inversion est « régulière ». Les premières travaux, qui proposent des séquences d'enregistrement identiques pour tous les pixels, montrent des résultats très encourageants mais cependant paradoxaux [44]. Ils proposent des ouvertures contrôlées sur des intervalles bien choisis, de façon à permettre l'inversion aisée du noyau de flou. On pourrait imaginer que la loi d'ouverture dépende également du pixel, les architectures CMOS actuellement employées le permettraient.

Mis au point pour s'affranchir du bougé, les obturateurs battants ont été exploités ensuite pour de nombreuses autres applications que la simple détection de mouvement : pour entrelacer des images vidéo acquises par des sources différentes, en stéréo vision, pour compresser des images.

3.6 Parcimonie *Compressed Sensing*

Derrière beaucoup de ces travaux, se cache une préoccupation majeure de la communauté de l'image, la détermination de représentations parcimonieuses adaptées au contenu de la scène. Ces représentations sont activement attendues pour accéder à des facteurs de compression supérieurs à ceux (pourtant déjà excellents) auxquels nous ont habitués les standards (JPEG2000, MPEG4), ou tout au moins pour conduire à des paradigmes de transmission où la tâche de compression est très allégée (au détriment probable de la tâche de décompression qui risque d'être allourdie). On en espère également des capacités de modélisation puissantes autorisant des descriptions symboliques et des représentations efficaces pour l'archivage, l'indexation et la requête dans les bases d'images.

Toutes les approches citées plus haut se caractérisent par un recours intensif à des représentations multiples et fortement redondantes qui sont loin d'être efficaces selon ces critères. L'espoir d'aboutir à un petit nombre de signaux suffisants pour reconstruire la grande diversité de la scène est sous-jacent à beaucoup de ces approches proposées qui, remettant l'image dans un espace de dimension supérieure (l'image plénoptique ou le *light field*), permettront de dégager les éléments les plus fondamentaux d'une représentation intrinsèque (c'est ce que recouvrent les termes d'acquisition comprimée ou *compressed sensing* [9, 15, 6]).

4 Conclusion

Nous n'avons fait qu'un rapide survol des développements récents visant à repousser les limites actuelles des capteurs. Un très large éventail d'applications tire actuellement ces travaux, le domaine de la vision par ordinateur étant concerné au premier chef, mais aussi celui de la télévision et du multimédia ainsi que celui de la synthèse d'images. Comme on l'a vu, les méthodes proposées prolongent très naturellement celles que l'on range usuellement dans les méthodes de traitement d'images et, à cette aune, il n'y a pas de différence de principe entre traitement d'image et imagerie computationnelle, tout au plus constate-t-on une remise en cause de l'outil d'acquisition (l'appareil photo) et un recours systématique au traitement numérique pour reconstruire le résultat. Sous cette lumière, la photographie computationnelle n'a peut-être pas vocation à devenir un domaine scientifique à part entière. C'est cependant actuellement un secteur très dynamique du traitement d'images, qui regroupe des acteurs particulièrement innovants qui déplacent très sensiblement les frontières de leur champ d'investigation.

Remerciements : Ce travail a bénéficié de nombreuses discussions avec Maria Cecilia AGUERREBERE, Andrès ALMANSA, Julie DELON, Yann GOUSSEAU et Saïd LADJAL.

Références

- [1] A. Agrawal and R. Raskar. Resolving objects at higher resolution from a single motion-blurred image. In *IEEE CVPR'07*, pages 1–8, 2007.
- [2] C. Aguerrebere, J. Delon, Y. Gousseau, and P. Musé. Best algorithms for HDR image generation. A study of performance bounds. HAL-00733853-v2, 2012.
- [3] C. Aguerrebere, J. Delon, Y. Gousseau, and P. Musé. Study of the digital camera acquisition process and statistical modeling of the sensor raw data. HAL-00733538-v3, 2012.
- [4] C. Aguerrebere, J. Delon, Y. Gousseau, and P. Musé. Simultaneous HDR image reconstruction and denoising for dynamic scene analysis. In *IEEE Int. Conference on Computational Photography, ICCP*, 2013.
- [5] Y. Bando, B.Y. Chen, and T. Nishita. Extracting depth and matte using a color filtered aperture. In *ACM SIGGRAPH Asia 2008*, page 134, 2008.
- [6] C.S. Baraniuk. Compressive sensing. *IEEE Signal Processing Magazine*, pages 118–121, July 2007.
- [7] A. Buades, B. Coll, J.M. Morel, and Sbert C. Self-similarity driven color demosaicking. *IEEE trans on PAMI*, 18 (6) :1192–1202, 2009.
- [8] E.J. Candès and C. Fernandez-Granda. Towards a mathematical theory of super-resolution. *Pure & Applied Mathematics*, doi 10.1002/cpa21455, 2013.
- [9] E.J. Candès and M.B. Wakin. An introduction to compressive sampling. *IEEE Signal Processing Magazine*, 25 (2) :21–30, 2008.
- [10] C.I. Chang. *Hyperspectral imaging : techniques for spectral detection and classification*. Springer, 2003.
- [11] C.I. Chang. *Hyperspectral data exploitation : theory & Applications*. Wiley, 2007.
- [12] K.H. Chung and Y.H. Chan. Color demosaicking using variance of color differences. *IEEE trans on Image Processing*, 15 (10) :2944–2955, 2006.
- [13] M. Delbracio, A. Almansa, P. P. Musé, and J.M. Morel. Subpixel point spread function estimation from two photographs at different distances. *SIAM Journal on Imaging Science*, 5 (4) :1234–1260, november 2012.
- [14] M. Delbracio-Bentancor. *Two problems of digital image formation*. PhD thesis, Ecole Normale Supérieure de Cachan, France, 2013.

- [15] D.L. Donoho. Compressed sensing. *IEEE trans on Information Theory*, 52 (4) :1289–1306, 2006.
- [16] F. Durand and R. Szeliski. Guest editors’ introduction : Computational photography. *IEEE Computer Graphics and Applications*, 27 (2), 2007.
- [17] M. Ebner. *Color Constancy*. Wiley, 2007.
- [18] T. Georgiev, A. Lumsdaine, and S. Goma. Lytro camera technology : theory, algorithms, performance analysis. In *Multimedia content and mobile devices, SPIE 8667*, 2013.
- [19] P. Getreuer. Image demosaicking with contour stencils. IPOL Library : www.ipol.im/pub/art/2012/g-dwcs/, March 2012.
- [20] A. Gijsenij, T. Gevers, and J. van de Weijer. Computational color constancy : Survey and experiments. *IEEE trans on Image Processing*, 20 (9) :2475 – 2489, 2011.
- [21] M. Granados, M. Adjin, C. Theobalt, H. Seidel, and H. Lensch. Optimal HDR reconstruction with linear digital cameras. In *Computer Vision and Pattern Recognition CVPR*, 2010.
- [22] F. Guichard, H.-P. Nguyen, R. Tessières, M. Pyanet, I. Tarchouna, and F. Cao. Extended depth-of-field using sharpness transport across color channels. In B. G. Rodricks and S. E. Susstrunk (Eds.), editors, *IST / SPIE Electronic Imaging (p. 72500N-72500N-12)*, 2009.
- [23] J Holloway, A. Sankaranarayanan, A. Verraraghavan, and S. Tambe. Flutter shutter video camera for compressive sensing of videos. In *IEEE int. Conference on Computational Photography, ICCP*, 2012.
- [24] A. Isaksen, L. McMillan, and S.J. Gortler. Dynamically reparametrized light fields. In *SIGGRAPH 2000*, pages 297–306, 2000.
- [25] N. Joshi, R. Szeliski, and D. Kriegman. PSF estimation using sharp edge prediction. In *IEEE Conf. on Comp. Vision & Pattern Recognition*, 2008.
- [26] J. Kronander, S. Gustavson, G. Bonnet, and J. Unger. Unified HDR reconstruction from raw CFA data. In *IEEE int. Conference on Computational Photography, ICCP*, 2013.
- [27] S. Ladjal. *Flou et quantification dans les images numériques*. PhD thesis, Télécom ParisTech (France), 2005.
- [28] D. Landgrebe. Hyperspectral image data analysis. *IEEE Signal Processing Magazine*, 19 (1) :17–28, 2002.
- [29] A. Levin, S.W. Hasinoff, Durand F. Green, P., and W.T. Freeman. 4D frequency analysis of computational cameras for depth of field extension. In *ACM SIGGRAPH*, 2009.
- [30] A. Levin, P. Sand, T.S. Cho, F. Durand, and W.T. Freeman. Motion invariant photography. *ACM trans on Graphics*, 27 :71, 2008.
- [31] C.-K. Liang, T.-H. Lin, B.-Y. Wong, C. Liu, and H.H. Chen. Programmable aperture photography : Multiplexed light field acquisition. *ACM Transactions on Graphics, proc. of Siggraph*, 27 (3) :55, 2008.
- [32] X. Lin, J.-L. Suo, Wetzstein G., Dai Q., and R. Raskar. Coded focal stack photography. In *IEEE int. Conference on Computational Photography, ICCP*, 2013.
- [33] G. Lippman. La photographie intégrale. *Comptes Rendus de l’Académie des Sciences*, 146 :446–551, 1908.
- [34] C. Liu, W.T Freeman, Szeliski R., and S.B. Kang. Noise estimation with single image. In *IEEE CVPR’06*, pages 901–908, 2006.
- [35] B. Mazin, J. Delon, and Y. Gousseau. Illuminant estimation from projections on the Planckian locus. In *CPCV workshop, ECCV 2012*, 2012.
- [36] D. Miao, O. Cossairt, and S. K. Nayar. Focal sweep Videography with deformable optics. In *IEEE int. Conference on Computational Photography, ICCP*, 2013.
- [37] S. K. Nayar and Y. Nakagawa. Shape from focus. *IEEE trans on PAMI*, 16 (8) :824–831, 1994.
- [38] R. Ng. Fourier slice photography. *ACM Transactions on Graphics, Proc. of ACM SIGGRAPH*, 24 (3) :735–744, 2005.

- [39] C. Perwass and L. Wietzke. Single lens 3Dcamera with extended depth of field. In *Human Vision and Electronic Imaging, SPIE 8291*, 2012.
- [40] T. Portz, L. Zhang, and H. Jiang. Random coded sampling for high speed HDR video. In *IEEE int. Conference on Computational Photography, ICCP*, 2013.
- [41] R. Raskar and J. Tumblin. *Computational Photography : Mastering New Techniques for Lenses, Lighting, and Sensors*. A. K. Peters, Ltd. Natick, MA, USA, 2009.
- [42] T. Sasao, S. Hiura, and K. Sato. Super-resolution with random shaped pixels and sparse regularization. In *IEEE int. Conference on Computational Photography, ICCP 2013*, 2013.
- [43] H. Tang and K. Kutulakos. What does an aberrated photo tell us about the lens and the scene? In *IEEE Int.Conf. on Computational Photography, ICCP*, 2013.
- [44] Y Tendero, J.M. Morel, and B. Rougé. The flutter shutter paradox. *SIAM Journal Imaging Science*, 6 (2) :813–845, 2013.
- [45] P. Trouvé. *Conception conjointe optique/traitement pour un imageur compact à capacité 3D*. PhD thesis, Ecole Centrale de Nantes (France), 2012.
- [46] L. Zhang and X. Wu. Color demosaicking via directional linear m.s.e. estimation. *IEEE trans on Image Processing*, 14 (12) :2167–2178, 2005.

