



HAL
open science

Procédé de datation des images et de caractérisation de la fréquence d'acquisition des caméras vidéo

Michel Grimaldi, Philippe Arlotto

► **To cite this version:**

Michel Grimaldi, Philippe Arlotto. Procédé de datation des images et de caractérisation de la fréquence d'acquisition des caméras vidéo. CNRIUT 2019, Jun 2019, TOULON, France. hal-02999449

HAL Id: hal-02999449

<https://hal.science/hal-02999449>

Submitted on 10 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Procédé de datation des images et de caractérisation de la fréquence d'acquisition des caméras vidéo

Michel GRIMALDI¹, Philippe ARLOTTO², Olivier CARON³

¹I.U.T de Toulon - département GEII - Laboratoire MIO

²I.U.T de Toulon - département GEII - plateforme SMIO T

³ Université de Toulon - Laboratoire LAMHESS

INFORMATIQUE - ROBOTIQUE - ELECTRONIQUE (61^{ème} section CNU)

RESUME - Depuis les années 80 et l'avènement des caméras numériques, une multitude de techniques de capture ou de quantification du mouvement par traitement d'images vidéo ont émergées. Toutes ces méthodes, très différentes les unes des autres, ont pourtant toute un point commun : la dimension temporelle est directement déduite de la fréquence d'acquisition des images vidéo qui, par hypothèse, est toujours supposée connue et stable, tout au long de la séquence d'images analysée. Or, il s'avère que cette condition, fondamentale pour une bonne quantification du mouvement, n'est pas toujours scrupuleusement respectée par les systèmes d'acquisition. Dans cet article, nous proposons un dispositif permettant de montrer que pour certaines caméras numériques, les instants d'acquisition des images ne sont pas aussi réguliers qu'on aurait pu l'espérer. Nous verrons ensuite comment dater précisément chaque image de la séquence pour compenser ce défaut lors de la détermination du modèle cinématique du mouvement observé.

Mots-clés—*Traitement d'images, capture de mouvement.*

1. INTRODUCTION

La capture de mouvement par traitement d'images vidéo et les analyses qui en découlent sont à l'origine de nombreux progrès technologiques, scientifiques et même artistiques. La capture du mouvement humain [1][2], par exemple, a très nettement fait progresser le réalisme des animations de personnages virtuels dans le cinéma de fiction. Dans le domaine du sport de haut niveau, en permettant des analyses biomécaniques fines du mouvement des sportifs, elle a contribué à l'amélioration des performances et de la récupération. Plus généralement, en physique expérimentale, ou en mécanique des fluides, elle a permis de vérifier des modèles théoriques de mouvement ou d'étudier certains phénomènes de turbulence complexes. Quelles que soient les techniques utilisées, les paramètres cinématiques du mouvement sont déduits du déplacement de points, ou de zones, caractéristiques entre deux ou plusieurs images successives [3][4][5]. Les paramètres cinématiques de vitesse ou d'accélération sont ensuite évalués par dérivation de ces déplacements par rapport au temps. La composante temporelle est elle, déduite de la fréquence d'échantillonnage du flux vidéo, supposée parfaitement constante, par hypothèse [9][10][11].

Un peu par hasard, en travaillant sur des images de biomécanique, nous nous sommes rendus compte que la majorité des caméras, même de bonne qualité, ne respectaient par toujours cette condition fondamentale. Tous les modèles qui découlent alors du traitement de la séquence d'images

peuvent, dans ce cas là, considérablement différer de la réalité, et mener à des conclusions erronées.

Dans un premier temps, nous allons décomposer le processus d'acquisition des caméras numériques pour tenter d'identifier les sources d'erreur pouvant conduire à une instabilité de la fréquence d'échantillonnage. Nous présenterons ensuite un dispositif optoélectronique permettant de mettre évidence ce défaut, que l'on pourra utiliser, par exemple, avant toute expérience pour spécifier les caméras. N'ayant souvent aucune action possible sur cette dérive de la fréquence d'échantillonnage, nous présenterons une méthode de datation des images de la séquence, afin de tenir compte des décalages temporels réels dans les modèles cinématiques.

2. STRUCTURE D'UNE CHAÎNE D'ACQUISITION D'IMAGE

D'une manière générale, une caméra numérique délivre une image n à partir d'une phase d'intégration des photons par le capteur, Ψ_n , suivi d'une macro phase de mise en forme, transfert et stockage sur la machine hôte.



Fig.1 chronogramme d'acquisition typique d'une caméra numérique

Quelle que soit la durée du transfert, le contenu de l'image est figé à la fin de la phase d'intégration. Dans le cas de caméras asynchrones, cette durée $a_{n+1} - b_n$ conditionne la prise de l'image suivante. Elle dépend du type de bus utilisé (USB, FireWire, IP, ou autre), mais également de la capacité de la machine hôte à récupérer les données. Pour les applications d'analyse de mouvement, on évite en général les modes de fonctionnement dans lesquels la durée d'intégration est variable (contrôle automatique de la luminosité). Pour les caméras synchrones, pour lesquelles le déclenchement de l'intégration est extérieur, nous avons remarqué que celles-ci peuvent avoir différentes stratégies, en fonction de leur *firmware*, en cas d'allongement de la phase de transfert/stockage, pouvant conduire à des images « doublons » ou à des images manquantes.

3. DESCRIPTION DU DISPOSITIF PROPOSÉ

Pour spécifier la caméra, nous effectuons l'acquisition vidéo de l'image d'une carte électronique constituée d'une matrice de diodes électro-luminescentes, *led*, divisée en deux blocs. Le premier, constitué de k *led*, est contrôlé par un registre à décalage délivrant un motif, formé par l'éclairement d'une *led*, qui se déplace à chaque période d'horloge à la fréquence f_0 parfaitement connue. Le registre étant rebouclé, le motif est périodique (*k-cycle*). Le second, constitué de N *led*

représente le numéro n du cycle précédent, en binaire. En fonction du nombre de bits, ce numéro est également cyclique (g -cycle).

nota : Pour éviter des aléas dans la lecture des N led nous utilisons du binaire réfléchi (code Gray).

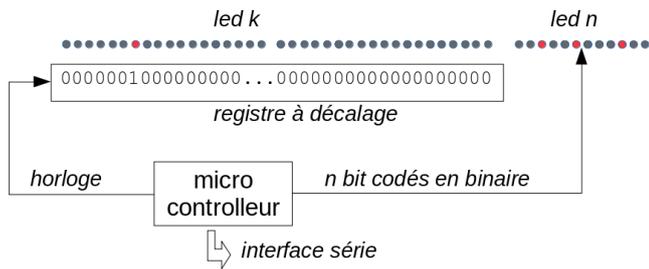


Fig.2 commande de la matrice de led

T_0 , la période de l'horloge principale

$T_1 = kT_0$, la période du k -cycle

$T_2 = 2^N T_1$, la période du g -cycle

L'instant exact t_i d'un front d'horloge i est donné par :

$$t_i = iT_0 + nkT_0 \quad [1]$$

avec i le numéro du front, n le numéro du k -cycle au sein du g -cycle, $0 < n < 2^{N-1}$

notons que t_i est défini à une période T_2 près, avec

$$T_2 = \frac{k2^N}{f_0}$$

La figure 3, ci-dessous, montre un exemple de chronogramme sur deux images successives avec $k=8$. Les rectangles représentent l'allumage des 8 led, la zone grisée les led intégrées par la caméra.

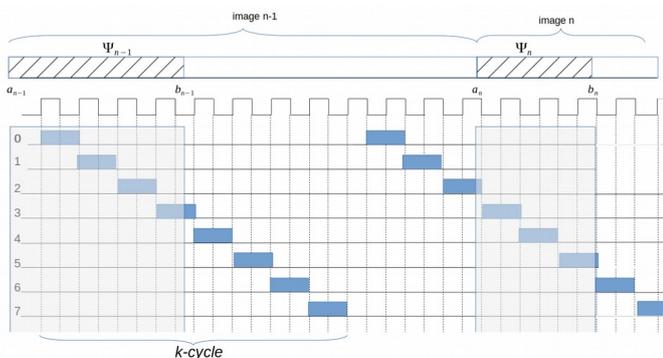


Fig.3 exemple d'intégration des led avec $k=8$

On remarquera que pour que la méthode s'applique il faut impérativement que la durée d'intégration soit inférieure à la durée d'un k -cycle, $\Psi < T_1$ et que la durée maximale entre deux images successives soit inférieure à la durée d'un g -cycle, $a_n - a_{n-1} < T_2, \forall n$. Dans ces conditions, la résolution temporelle est de l'ordre de T_0 .

4. TRAITEMENT D'IMAGE

La tâche principale du programme de traitement des images de la séquence vidéo consiste à localiser chaque led et à en déterminer sa luminosité. Nous utilisons des images monochromes codées sur 256 niveaux de gris, avec la bibliothèque OpenCV [7].

Sur la première image, l'utilisateur est invité à désigner un référentiel, par simples clics sur les quatre coins de la carte électronique. Les dimensions réelles de cette dernière étant

parfaitement connus, la transformation perspective représentant l'inverse de la déformation du système optique, est évaluée. Après correction, les positions approximatives des led sont encadrées par une zone carrée (*region of interest, ROI*) et le niveau maximum dans cette zone leur est associé.

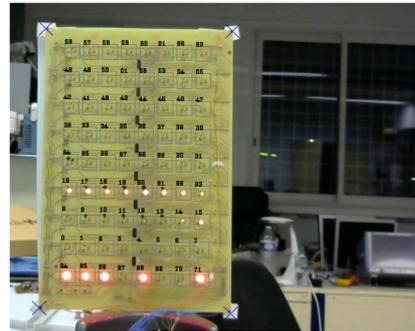


Fig.4 exemple d'image d'une scène incluant le dispositif. On peut remarquer les repères du référentiel et les zones ROI de chaque led.

5. EXPÉRIMENTATION ET RÉSULTATS

Pour vérifier la technique, nous avons réalisé une carte électronique comportant 72 led, ($k=64$ et $N=8$). L'horloge est fournie par un microcontrôleur Atmel (Arduino), permettant de régler la fréquence f_0 et d'afficher le numéro du k -cycle en binaire réfléchi sur 8 bits. Le code a été optimisé pour garantir une bonne stabilité. Une étude indépendante [12] évalue l'écart type des variations fréquentielles de ce type de carte à 0.7Hz sur 16MHz sur une période de 18 heures, ce qui est largement suffisant pour notre application.

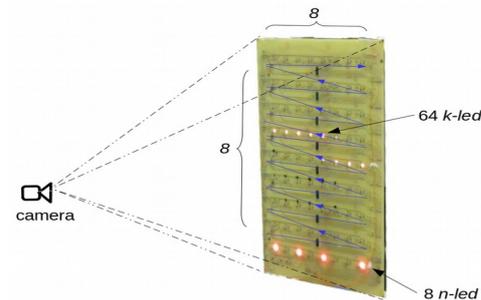


Fig.5 dispositif expérimental $k=64$ et $N=8$. Dimension : 20 cm x 30 cm. L'ordre d'allumage est matérialisé par les lignes bleu, de gauche à droite en montant

Pour valider le dispositif, nous utilisons une caméra, Basler ACE – AC1920-155UM [8], synchronisée par une horloge externe, donnée pour 250fps @ 1200x740 en USB-3.0., avec les paramètres suivants :

($T_0=218\mu s$) la période de l'horloge principale, $f_0 \approx 4587$ Hz, $T_1=13952\mu s$ la durée du k -cycle et $T_2 \approx 3,5s$ la durée du g -cycle.

La figure 6, ci dessous, représente le chronogramme des intensités des 64 k -led, horizontalement de gauche à droite, en fonction des images, verticalement de haut en bas.

On remarquera les débuts d'intégration (repères 1, 3, 5, .etc.), les fins d'intégration (repères 2, 4, 6, .etc.) et les transferts de données vers la machine hôte, dans les périodes intermédiaires (repères 2 à 3, 4 à 5, 6 à 7, .etc.), pour environ 9 périodes T_0 d'horloge. Il est à noter qu'à la 5^{ème} ligne, l'intégration commençant en fin de cycle (13), elle se poursuit en début (14) conformément au balayage modulo 64 du k -cycle. La durée d'intégration est d'environ 10 périodes de $218\mu s$, soit

$\approx 2,18\text{ms}$. La durée d'une image est comprise entre 18 et 19 périodes, ce qui correspond bien à 250 images par seconde (18,35 périodes). Les seules variations observables sont dues à l'erreur sur la résolution temporelle : T_0

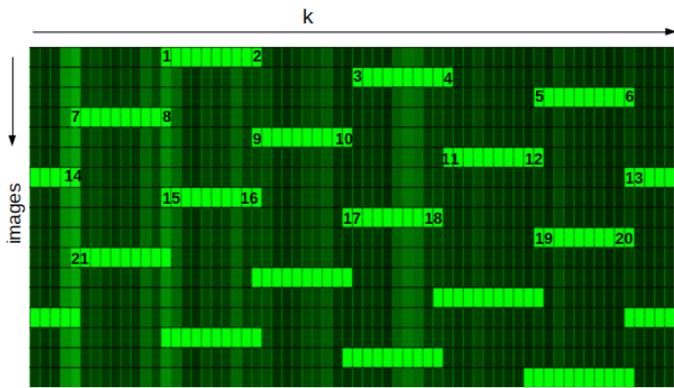


Fig.6 Chronogramme de la Basler ACE : chaque bande horizontale représente l'intensité des 64 k_led ($K=64$) observées sur une même image.

Une deuxième expérience a été menée avec une Webcam Logitech C270 de 3 MPix, 30fps @ 1280x720 avec les paramètres suivants :

$T_0=2\text{ms}$ la période de l'horloge principale, $f_0=500\text{Hz}$, $T_1=128\text{ms}$ la durée du k_cycle et $T_2\approx 32,7\text{s}$ la durée du g_cycle .

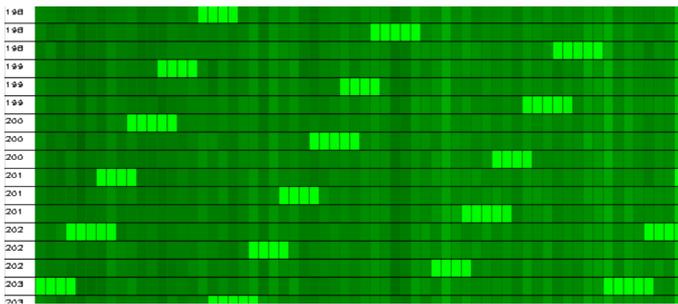


Fig.7 Chronogramme de la Logitech C270: chaque bande horizontale représente l'intensité des 64 k_led ($K=64$) observées sur une même image. Le numéro du k_cycle est reporté en ordonnée.

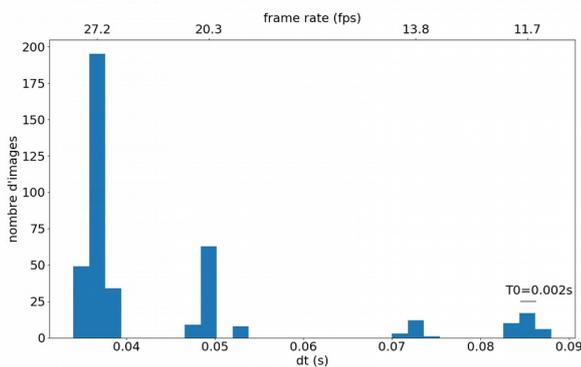


Fig.8 Histogramme des durées d'acquisition/transfert de la Logitech C270. On remarque un caractère multimodal avec des variations de part et d'autre correspondant exactement à la résolution temporelle T_0 .

Nous remarquons que les durées des images ont une variabilité beaucoup plus importante, entre 11 et 27 fps, mais surtout que certaines configurations ne correspondent à aucune image réelle, figure 7 avant dernière ligne, avec des diodes non contiguës allumées. Ceci nous conduit à penser que cette caméra ne respecte par le modèle de la figure 1. Le *firmware*

doit très certainement effectuer les intégrations et les transferts de façon asynchrone.

6. CONCLUSIONS

La technique que nous avons présentée, ainsi que le dispositif associé, permettront de quantifier un phénomène qui a, très certainement, été observé par les personnes qui étudient le mouvement à partir d'images numériques. Cette irrégularité dans la chronologie d'acquisition des images est pourtant très peu évoquée dans la littérature, ou dans de domaines très éloignés [6], alors qu'elle peut conduire à des erreurs importantes dans les modèles cinématiques qui en découlent.

Nous montrons que certaines caméras à bas coût varient énormément et peuvent même induire des images d'artefacts. Elles sont donc à proscrire pour l'analyse du mouvement.

La datation précise et systématique de chaque image telle qu'elle a été présentée peut servir, dans un premier temps, comme moyen de caractérisation de caméras avant leur mise en œuvre dans un système d'analyse de mouvement. N'ayant aucune action sur ces irrégularités, le seul intérêt est alors d'apprécier la véracité des modèles obtenus. En revanche, si l'application le permet, l'inclusion du dispositif directement dans la scène à traiter, pendant l'acquisition, peut apporter un avantage considérable en procédant à une datation, *a posteriori*, qui pourra être prise en compte dans les modèles cinématiques.

La suite logique de ce travail, est de réaliser une étude comparative sur un grand nombre de caméras, dans des conditions d'utilisation aussi larges que possible.

La description du matériel et les logiciels associés à ce dispositif seront prochainement mis à la disposition de la communauté sous licence *Creative Commons*.

7. RÉFÉRENCES

- [1] J.K. Aggarwal and Q. Cai, « Human Motion Analysis: A Review, Computer Vision and Image Understanding », Volume 73, Issue 3, 1 March 1999, Pages 428-440
- [2] Ormonet H., Sidenbladh M.J., Black T. and Hastie D., « Learning and Tracking Human Motion Using Functional Analysis », Proc.IEEE Workshop on Human Modeling, Analysis and Synthesis, June 2000.
- [3] Philip H.S. Torr and Andrew Zisserman: « Feature Based Methods for Structure and Motion Estimation », ICCV Workshop on Vision Algorithms, pages 278-294, 1999
- [4] B. Horn and B. Schunck. « Determining optical flow », Artificial Intelligence, 16(1-3):185-203, August 1981.
- [5] Jianbo Shi and Carlo Tomasi, « Good Features to Track », Conference Paper, 1994 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'94), pp. 593 - 600, January, 1994
- [6] Trang Nguyen, Nam Tuan Le, and Yeong Min Jang, « Asynchronous Scheme for Optical Camera Communication-Based Infrastructure-to-Vehicle Communication », International Journal of Distributed Sensor Networks, Volume 2015, Article ID 908139
- [7] Open Source Computer Vision Library, <https://opencv.org/>
- [8] Basler, site du constructeur, <https://www.baslerweb.com>
- [9] Grigg, J., Haakonssen, E., Rathbone, E., Orr, R., & Keogh, J. W. L. (2018). « The validity and intra-tester reliability of markerless motion capture to analyse kinematics of the BMX Supercross gate start ». Sports Biomechanics, 17(3), 383-401.
- [10] Sayers, M. G. L. (2011). « Kinematic analysis of line-out throwing in elite international rugby union ». Journal of Sports Science & Medicine, 10(3), 553-558.
- [11] Trewartha, G., Casanova, R., & Wilson, C. (2008). « A kinematic analysis of rugby lineout throwing ». Journal of Sports Sciences, 26(8), 845-854.
- [12] Joris van Rantwijk, « Arduino clock frequency accuracy », <http://jorisvr.nl/article/arduino-frequency>.