

Détection statistique avec une caméra biomimétique

Lionel FILLATRE¹

¹Laboratoire I3S - UMR 7271 - Université Nice Sophia Antipolis - CNRS
CS 40121 - 06903 Sophia Antipolis CEDEX, France

lionel.fillatre@i3s.unice.fr

Résumé – Les caméras biomimétiques utilisent le codage temporel pour coder les intensités des pixels. Un codage temporel génère un flux aléatoire d'événements asynchrones. Un événement est défini comme un couple composé d'un instant d'émission de l'événement et du signe de la variation du signal d'entrée depuis le dernier événement généré. Le but de cet article est de détecter un objet connu à partir de l'observation du flux d'événements. Cet article propose un modèle statistique du flux aléatoire d'événements basé sur le modèle physique de la caméra biomimétique. Il propose également un détecteur optimal qui détecte si un objet est présent dans le champ visuel de la caméra. Les performances statistiques du détecteur sont évaluées sur des données artificielles.

Abstract – Asynchronous event-based cameras use time encoding to code the pixel intensity values. A time encoding of an input pattern generates a random stream of asynchronous events. An event is defined as a pair composed of a timestamp and the variation sign of the input signal since the last emitted event. The goal of this paper is the detection of a known input object from the observation of the event stream. This paper proposes a statistical model of the random event stream based on the physical model of the event-based camera. It also proposes an optimal test which detects the input object. Finally, the statistical performance of the test is numerically evaluated on simulated data.

1 Introduction

Les caméras biomimétiques, aussi appelées caméras événementielles, sont en train d'émerger dans le domaine du traitement des images et de la vidéo [1]. Contrairement aux technologies d'acquisition et de traitement d'images à base de trames, les capteurs biomimétiques fournissent une nouvelle approche pour encoder la luminance et ses variations temporelles. Les caméras biomimétiques encodent les localisations spatiales (position du pixel) et les instants de changement (horodatage) du flux lumineux en entrée de la caméra. Les caméras biomimétiques ont une résolution temporelle très élevée dépassant la vitesse de la plupart des caméras classiques à base de trames. Elles permettent d'économiser de l'énergie. Ce nouveau paradigme d'acquisition de données requiert de nouvelles méthodes en traitement du signal et des images [2, 3].

Les caméras biomimétiques encodent temporellement le logarithme $x_k(t)$ des intensités des pixels pour $t > 0$ et tout $k = 1, \dots, K$ où K est le nombre de pixels [4, 5]. La scène en entrée de la caméra est donc représentée comme une séquence aléatoire, appelé le flux d'événements, $(t_{k,j}, y_{k,j})$ où $(t_{k,j})$, $j \in \mathbb{N}$, est une suite d'instants strictement croissants et $y_{k,j} \in \{-1, 1\}$ désigne la polarité de l'événement. De nombreux modèles [6] existent pour coder un signal dans le domaine temporel. Cet article considère l'échantillonnage par franchissement de seuil qui reste l'une des techniques les plus utilisées dans les capteurs biomimétiques [3].

Les travaux actuels [1, 7] s'intéressent à la modélisation électronique et au traitement (filtrage, reconstruction, etc.) des flux d'événements asynchrones produits par une caméra bio-

mimétique. La détection statistique d'objets n'est pas abordée dans ces travaux. Par conséquent, cet article propose deux contributions majeures. Tout d'abord, le flux d'événements est modélisé comme une séquence d'événements aléatoires. Le caractère aléatoire du flux d'événements est dû au bruit naturellement présent à l'intérieur du capteur biomimétique. Ensuite, nous proposons un détecteur statistique optimal basé sur le rapport de vraisemblance afin de détecter un objet en entrée du système. En absence d'objet dans le champ de vision de la caméra, la probabilité de déclarer la présence de l'objet d'intérêt doit être bornée supérieurement.

La section 2 décrit les caméras biomimétiques et propose un modèle statistique du flux d'événements. La section 3 présente le détecteur optimal qui exploite le modèle statistique de la caméra biomimétique. La section 4 propose une étude numérique du détecteur. Enfin, la section 5 conclut cet article.

2 Les caméras biomimétiques

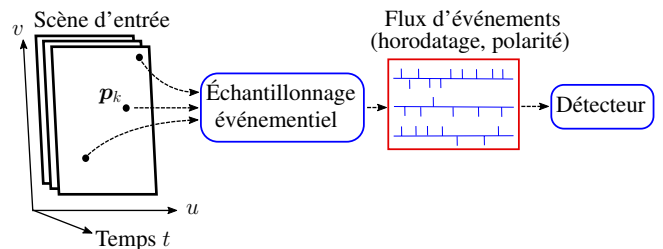


FIGURE 1 – Détection à partir d'un flux d'événements.

Le but de cette section est de modéliser le flux d'événements produit par l'échantillonnage à franchissement de seuil d'une scène variant dans le temps. Ce flux sera utilisé pour détecter un objet connu situé dans le champ de vision de la caméra. Le processus de détection est décrit dans la figure 1. Un événement est défini comme une paire contenant un instant de changement et la polarité de ce changement pour un pixel donné. La polarité d'un changement est $+1$ si la variation de la luminance en entrée du pixel considéré est positive, sinon la polarité vaut -1 . L'instant de changement correspond au moment où la variation mesurée dépasse un certain niveau. Il est important de noter que chaque événement est aléatoire puisque chaque pixel est dégradé par un bruit aléatoire. Cet article ne considère pas d'autres sources d'aléas et évite ainsi l'utilisation de modèles doublement stochastiques.

2.1 Modèle physique de la caméra

Un capteur biomimétique est composé de capteurs qui sont sensibles à la luminance de la scène en entrée de la caméra. Il est admis que la luminance de la scène est caractérisée par une fonction spatio-temporelle $(u, v, t) \mapsto s(u, v, t)$ où (u, v) appartient à $\mathcal{D} \subset \mathbb{R}^2$ et $t \in [0, T]$ où T est la durée de l'acquisition en microsecondes. L'ensemble des pixels $\mathbf{p}_k = (u_k, v_k)$ correspond à certains points d'échantillonnage de \mathcal{D} pour $k = 1, \dots, K$. Chaque pixel est associé à un photorécepteur qui mesure la luminance.

Soit $I_k(t) > 0$ le photo-courant mesuré par le photo-récepteur du pixel \mathbf{p}_k . En temps continu, le pixel est sensible à la dérivée $\dot{x}_k(t)$ du logarithme $x_k(t) = \ln I_k(t)$ du photo-courant :

$$\dot{x}_k(t) = \frac{d \ln I_k(t)}{dt} = \frac{1}{I_k(t)} \frac{dI_k(t)}{dt}. \quad (1)$$

En pratique, le capteur biomimétique fonctionne en temps discret avec une résolution temporelle typique de $1 \mu s$. Soit $\{t_{k,1}, \dots, t_{k,n_k}\}$ les instants de $[0, T]$ où le pixel \mathbf{p}_k émet des événements. Un événement produit par le pixel \mathbf{p}_k est un couple $e_{k,j} = (t_{k,j}, y_{k,j})$ où $t_{k,j}$ est le temps du j -ème événement et $y_{k,j}$ est sa polarité, soit -1 (OFF event) ou $+1$ (ON event). En temps discret $t \geq t_{k,j}$, la dérivée du logarithme de la réponse est approximativement donnée par :

$$\begin{aligned} \delta_k(t, t_{k,j}) &= x_k(t) - x_k(t_{k,j}) + \xi_k(t) \\ &= \dot{x}_k(t, t_{k,j}) + \xi_k(t), \end{aligned} \quad (2)$$

où $\xi_k(t)$ est un bruit. La différence $\dot{x}_k(t, t_{k,j}) = x_k(t) - x_k(t_{k,j})$ est appelée le contraste temporel après le j -ème événement.

Le contraste temporel aléatoire $\delta_k(t, t_{k,j})$ est alors comparé à deux seuils Δ^+ et Δ^- tels que $\Delta^- < \Delta^+$ (typiquement nous avons $\Delta^+ = \Delta = -\Delta^-$). Ces seuils sont supposés être les mêmes pour tous les pixels. Si la différence dépasse le seuil supérieur Δ^+ , le capteur émet un événement ON; si elle est au dessous du seuil inférieure Δ^- , le capteur émet un événement OFF; enfin, si la différence est comprise entre les deux seuils, aucun événement n'est produit. Après l'émission de l'événement $j + 1$, la valeur d'entrée $x_k(t_{k,j+1})$ est sauvegardée dans

le capteur biomimétique pour calculer $\delta_k(t, t_{k,j+1})$ pour le prochain événement. En toute rigueur, les seuils Δ^- et Δ^+ sont aléatoires puisqu'ils dépendent de l'électronique du système. Pour simplifier le modèle, le caractère aléatoire de Δ^- et Δ^+ est inclut dans $\xi_k(t)$.

Les expériences physiques menées dans [1, 8, 9, 10] montrent que $\xi_k(t)$ résulte de plusieurs sources de bruit : bruit photonique, non-uniformité des réponses des pixels, etc. Ces expériences évaluent la sensibilité du capteur au contraste du signal d'entrée en mesurant la probabilité de réponse du capteur en fonction de l'augmentation du contraste. En absence de bruit (dans un monde idéal), la réponse du capteur aurait la forme d'une fonction en échelon : l'émission d'un événement est certaine ou, au contraire, impossible pour un niveau de contraste donné. En pratique, cette courbe a l'allure d'une courbe en "S" (voir Fig. 14 dans [8]) qui se rapproche de la fonction de répartition d'une variable Gaussienne. Ainsi, le bruit $\xi_k(t)$ peut être modélisé par un bruit blanc Gaussien tel que $\{\xi_k(t), 1 \leq k \leq K, 0 \leq t \leq T\}$ est une famille de variables aléatoires indépendantes et identiquement distribuées $\xi_k(t) \sim \mathcal{N}(0, \sigma^2)$ où $\sigma > 0$ est l'écart-type du bruit. Il est admis que la caméra est bien calibrée, donc σ est connu.

2.2 Modèle statistique d'un flux d'événements

Notons $p_{k,j+1}(t, -1)$, resp. $p_{k,j+1}(t, +1)$, la probabilité que le $j + 1$ -ème événement soit un événement OFF, resp. ON, émis à l'instant t par le pixel \mathbf{p}_k . La probabilité de non-émission du $j + 1$ -ème événement à l'instant t au pixel \mathbf{p}_k est notée $p_{k,j+1}(t, 0)$. La distribution du bruit est Gaussienne donc ces probabilités sont données par :

$$p_{k,j+1}(t, 0) = \Pr(\Delta^- \leq \delta_k(t, t_{k,j}) \leq \Delta^+) = \Phi\left(\frac{\Delta^+ - \dot{x}_k(t, t_{k,j})}{\sigma}\right) - \Phi\left(\frac{\Delta^- - \dot{x}_k(t, t_{k,j})}{\sigma}\right), \quad (3)$$

$$\begin{aligned} p_{k,j+1}(t, -1) &= \Pr(\delta_k(t, t_{k,j}) \leq \Delta^-) \\ &= \Phi\left(\frac{\Delta^- - \dot{x}_k(t, t_{k,j})}{\sigma}\right), \end{aligned} \quad (4)$$

$$\begin{aligned} p_{k,j+1}(t, +1) &= \Pr(\delta_k(t, t_{k,j}) \geq \Delta^+) \\ &= 1 - \Phi\left(\frac{\Delta^+ - \dot{x}_k(t, t_{k,j})}{\sigma}\right), \end{aligned} \quad (5)$$

où $\Phi(\cdot)$ est la fonction de répartition de la distribution normale centrée-réduite. On peut remarquer que la probabilité du $j + 1$ -ème événement est seulement conditionnée par le j -ème événement et non pas par les événements plus anciens.

Le flux d'événements du pixel \mathbf{p}_k est un vecteur, noté S_k , composé de n_k couples $(t_{k,j}, y_{k,j})$:

$$S_k = \{(t_{k,j}, y_{k,j}), 1 \leq j \leq n_k\}.$$

En temps discret $t \in \mathbb{N}$, la probabilité du flux S_k est donnée par

$$\Pr(S_k) = B_{k,T} \prod_{q=1}^{q=n_k} A_{k,q} \quad (6)$$

où $A_{k,q}$ est la probabilité de la séquence d'événements commençant à l'instant $t_{k,q-1}+1$ et se terminant à l'instant $t_{k,q}$:

$$A_{k,q} = p_{k,q}(t_{k,q}, y_{k,q}) \prod_{i=t_{k,q-1}+1}^{t_{k,q}-1} p_{k,q}(i, 0), \quad (7)$$

et $B_{k,T}$ est la probabilité de ne pas émettre d'événements après l'instant t_{k,n_k} :

$$B_{k,T} = \prod_{\ell=t_{k,n_k}+1}^T p_{k,n_k+1}(\ell, 0). \quad (8)$$

Dans les équations précédentes, on adopte la convention $t_{k,0} = -1$ pour tous les pixels k . Le flux d'événements pour tous les pixels est noté

$$S_1^K = \{S_k, k = 1, \dots, K\},$$

et \mathcal{S} désigne l'ensemble de tous les flux d'événements sur $[0, T]$. Il est à noter que la probabilité d'un flux d'événements dépend du signal X en entrée de la caméra, défini par

$$X = \{x_k(t), k = 1, \dots, K, 0 \leq t \leq T\}, \quad (9)$$

au travers des contrastes $\hat{x}_k(t, t_{k,j})$.

3 Détecteur de Neyman-Pearson

Détecter un objet consiste à choisir entre les deux hypothèses $\mathcal{H}_0: \{X = X_0\}$ et $\mathcal{H}_1: \{X = X_1\}$ où $X_0 = 0$ et X_1 est un signal connu. Soit $\Pr_\ell(S_1^K)$ la probabilité du flux S_1^K lorsque le signal d'entrée est X_ℓ avec $\ell \in \{0, 1\}$. Un détecteur est une fonction $\phi: \mathcal{S} \mapsto \{0, 1\}$ telle que X_ℓ est choisi quand $\phi(S_1^K) = \ell$. Le détecteur optimal est le test du rapport de vraisemblance [11], aussi appelé test de Neyman-Pearson (NP). Il consiste à comparer la fonction $\Lambda(S_1^K)$ à un seuil h :

$$\phi(S_1^K) = \begin{cases} 0 & \text{si } \Lambda(S_1^K) \leq h, \\ 1 & \text{sinon,} \end{cases} \quad (10)$$

où la fonction de décision $\Lambda(S_1^K)$ est donnée par :

$$\Lambda(S_1^K) = \sum_{k=1}^K \ln \Pr_1(S_k) - \sum_{k=1}^K \ln \Pr_0(S_k),$$

$\Pr_\ell(S_k)$ est décrit dans (6) lorsque $X = X_\ell$. Le seuil h est choisi de façon à ce que

$$\Pr_0(\Lambda(S_1^K) > h) = \alpha$$

où $\alpha \in]0, 1[$ est la probabilité de fausse alarme souhaitée. En toute rigueur, le test devrait être randomisé [11] puisque les distributions sont discrètes. En pratique, la probabilité $\Pr_0(\Lambda(S_1^K) = h)$ est très petite. De ce fait, la randomisation du test ne change pas de façon significative les performances du test. La probabilité de détection β du test, aussi appelée la puissance du test, est définie par

$$\beta = \Pr_1(\Lambda(S_1^K) > h).$$

Cette probabilité dépend clairement des seuils Δ^+ et Δ^- de la caméra biomimétique. Ces seuils peuvent être choisis par l'utilisateur pour contrôler le nombre moyen d'événements générés par la caméra.

4 Résultats numériques

Cette section propose une étude numérique des performances du détecteur NP sur des données simulées. La scène d'entrée est soit vide, soit composée d'un disque en rotation autour du centre de l'image. Tous les paramètres du disque (rayon, intensité et vitesse de rotation) sont connus. Dans tous les cas, le fond de la scène est d'intensité constante supposée nulle. Chaque image est composée de 64×64 pixels.

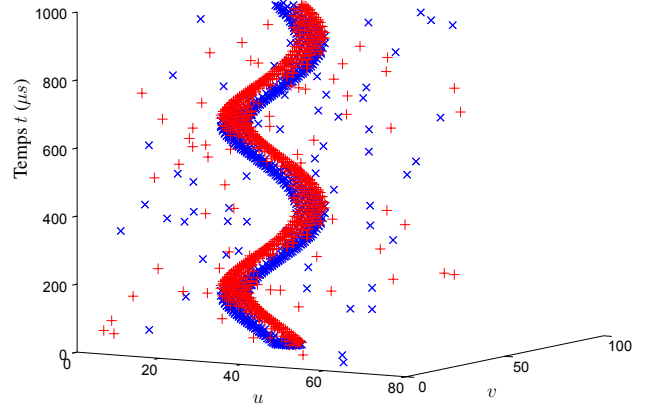


FIGURE 2 – Flux d'événements produit par un disque en rotation. Les croix bleues, resp. les plus rouges, correspondent aux événements ON, resp. OFF.

Un flux typique d'événements est visible sur la figure 2 avec $\Delta^+ = -\Delta^- = 0.5 = \Delta$ et $\sigma = 0.1$. Cette simulation dure $T = 1000 \mu s$. Le disque effectue 2 tours complets pendant la simulation. La rotation est très rapide afin de réduire la durée de la simulation. La figure montre clairement le disque qui est en rotation et des événements purement aléatoires dus au bruit des capteurs.

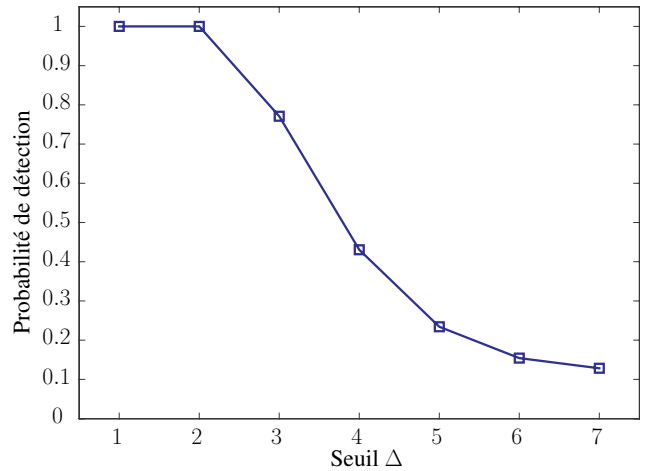


FIGURE 3 – Probabilité de détection en fonction du seuil Δ .

La puissance β du test NP est évaluée à partir d'une simulation Monte-Carlo. Sous l'hypothèse \mathcal{H}_0 , nous générons

1000 flux aléatoires de $10 \mu s$ suivant le modèle décrit dans la sous-section 2.1 pour chacune des valeurs possibles de $\Delta \in \{1, 2, 3, 4, 5, 6, 7\}$. L'écart-type vaut $\sigma = \Delta/4$. La fonction de décision Λ est évaluée pour chaque flux. Ceci nous permet alors d'estimer le seuil $h(\Delta)$, pour chaque valeur de Δ , qui permet de satisfaire la probabilité de fausse alarme $\alpha = 0.02$. Ensuite nous générons 1000 flux aléatoires sous l'hypothèse \mathcal{H}_1 pour les mêmes valeurs de Δ . La fonction de décision Λ est de nouveau évaluée pour chaque flux, ce qui nous permet d'estimer la puissance du test $\beta(\Delta)$ pour chaque valeur de Δ . La figure 3 montre l'évolution de $\beta(\Delta)$ en fonction de Δ . Elle met en évidence que $\beta(\Delta)$ est une fonction décroissante. Ceci est tout à fait logique puisqu'une grande valeur de Δ limite considérablement la génération des événements aléatoires. Il y a donc peu d'événements générés et donc peu d'informations exploitables par le détecteur. La puissance du détecteur devient plus faible.

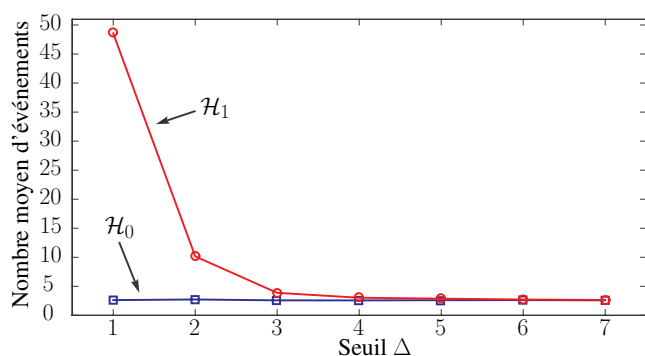


FIGURE 4 – Nombre moyen d'événements générés selon l'hypothèse \mathcal{H}_0 ou \mathcal{H}_1 en fonction du seuil Δ .

Pour chaque valeur de Δ , on calcule également le nombre moyen d'événements générés selon chaque hypothèse. Le résultat est affiché sur la figure 4. Comme souligné précédemment, on constate que, sous l'hypothèse \mathcal{H}_1 , le nombre moyen d'événements générés décroît lorsque le seuil Δ . Sous l'hypothèse \mathcal{H}_0 , les seuls événements émis sont dus au bruit des capteurs puisque le contraste du signal d'entrée est toujours nul (il n'y a pas d'objet dans la scène). De ce fait, le nombre d'événements émis ne change pas significativement en fonction de Δ . En conclusion, pour fixer le seuil Δ d'émission des événements, il s'agit de trouver un compromis entre la puissance du test et l'énergie économisée dans l'émission des événements.

5 Conclusion

Ce papier propose un modèle statistique du flux aléatoire d'événements produit par une caméra biomimétique et un test statistique optimal pour détecter un objet connu. Le test obtenu possède une structure plus complexe que les tests classiques qui utilisent des caméras conventionnelles. Néanmoins, cette complexité est compensée par le fait que les caméras biomimétiques ont une consommation énergétique très réduite par rapport aux caméras conventionnelles.

6 Remerciements

Ces travaux ont été financés en partie par la région Provence Alpes Côte d'Azur dans le cadre du projet COBRA.

Références

- [1] S.C. Liu, T. Delbruck, G. Indiveri, A. Whatley, and R. Douglas, *Event-Based Neuromorphic Systems*, John Wiley and Sons Ltd., UK, 2014.
- [2] Ryad Benosman, Sio-Hoi Ieng, Charles Clercq, Chiara Bartolozzi, and Mandyam Srinivasan, "Asynchronous frameless event-based optical flow," *Neural Networks*, vol. 27, pp. 32–37, 2012.
- [3] Sio-Hoi Ieng, C. Posch, and R. Benosman, "Asynchronous neuromorphic event-driven image filtering," *Proceedings of the IEEE*, vol. 102, no. 10, pp. 1485–1499, Oct 2014.
- [4] A.A. Lazar and L.T. Toth, "Perfect recovery and sensitivity analysis of time encoded bandlimited signals," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 51, no. 10, pp. 2060–2073, oct. 2004.
- [5] P. Lichtsteiner, C. Posch, and T. Delbruck, "A 128×128 120 dB 15 μs Latency Asynchronous Temporal Contrast Vision Sensor," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 2, pp. 566–576, 2008.
- [6] A. A. Lazar, E. K. Simonyi, and L. T. Toth, "Time encoding of bandlimited signals, an overview," in *Proceedings of the Conference on Telecommunication Systems, Modeling and Analysis*, Nov. 2005.
- [7] G. Orchard, C. Meyer, R. Etienne-Cummings, C. Posch, N. Thakor, and R. Benosman, "Hfirst : A temporal approach to object recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PP, no. 99, pp. 1–1, 2015.
- [8] C. Posch, D. Matolin, and R. Wohlgenannt, "A QVGA 143 dB Dynamic Range Frame-Free PWM Image Sensor With Lossless Pixel-Level Video Compression and Time-Domain CDS," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 1, pp. 259–275, 2011.
- [9] Christoph Posch and Daniel Matolin, "Sensitivity and uniformity of a $0.18 \mu m$ cmos temporal contrast pixel array," in *International Symposium on Circuits and Systems*, 2011.
- [10] Teresa Serrano-Gotarredona and Bernabé Linares-Barranco, "A 128×128 1.5% Contrast Sensitivity 0.9% FPN 3 μs Latency 4 mW Asynchronous Frame-Free Dynamic Vision Sensor Using Transimpedance Preamplifiers," *J. Solid-State Circuits*, vol. 48, no. 3, pp. 827–838, 2013.
- [11] A. A. Borovkov, *Mathematical Statistics*, Gordon and Breach Sciences Publishers, Amsterdam, 1998.