

Event-based coding of still images using a bio-inspired frame

Effrosyni Doutsi, Lionel Fillatre, Marc Antonini, Julien Gaulmin

▶ **To cite this version:**

Effrosyni Doutsi, Lionel Fillatre, Marc Antonini, Julien Gaulmin. Event-based coding of still images using a bio-inspired frame. International Conference on Event-based Control, Communication, and Signal Processing (EBCCSP), 2015, Krakow, Poland. hal-01303086

HAL Id: hal-01303086
https://hal.science/hal-01303086

Submitted on 19 Jan 2021

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.

EVENT-BASED CODING OF STILL IMAGES USING A BIO-INSPIRED FRAME

Effrosyni Doutsi[1,2]*, Lionel Fillatre*[1]*, Marc Antonini*[1] *and Julien Gaulmin*[2]

[1]Univ. Nice Sophia Antipolis, CNRS, I3S, UMR 7271, 06900 Sophia Antipolis, France
[2]4G-TECHNOLOGY, 460 avenue de la Quiera 06370 Mouans Sartoux - France.

ABSTRACT

The growing interest of event-based systems has influenced image processing field and is expected to be a groundbreaking architecture for image and video compression. The human visual system is an efficient event-based biological system which is a source of inspiration to design an event-based codec. This paper introduces a bio-inspired encoding architecture for still images. A bio-inspired filter is first applied to the input image and then the retinal-filtered image is fed to an Integrate And Fire (IAF) sampler. The spike train produced by the IAF sampler is decoded in order to reconstruct the still image at the receiver. We prove that the retinal filter has a frame structure, so the reconstruction is almost perfect provided that the IAF sampling is sufficiently fine.

***Index Terms*—** Event-based systems, bio-inspired processing, spatiotemporal filter, Integrate And Fire sampling.

1. INTRODUCTION

During the last few years many standards have been released in image and video compression like JPEG, JPEG 2000 [1], AVC [2] and HEVC [3], which are currently used in most of the technological devices. However, the needs in real time processing and compression change as well as the amount and the resolution of data. As a result, it is necessary to start using other kind of techniques to provide alternative coding results. For instance, compression algorithms using an event-based sampling could be a relevant alternative to minimize the power consumption of the compression chain.

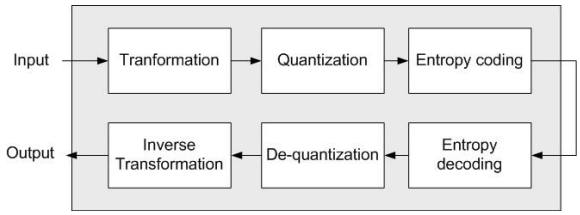


Fig. 1: Conventional coding principle.

The principle in signal compression is given by Fig. 1 where the input can be considered as 1D signal, 2D signal (still-image) or 3D signal (video). The transformation of the signal is the first and the most important part which provides more information about the time and the frequency of the signal. There are several transformations utilized in signal processing like wavelets, Fourier transform, Discrete Cosine Transform (DCT), etc. The quantization step defines the loss of information by eliminating some coefficients. The entropy coding that follows the quantization, is a lossless compression function that translates the intensity of pixels into codewords whose lengths vary inversely to the frequency of pixels occurrence. The goal in compression is always to transmit or store data in order to be able to reconstruct them later. Consequently, there is always the synthesis pathway to deliver an approximation of the input signal. The synthesis pathway consists of entropy decoding, de-quantization and inverse transformation to recover the input signal.

Instead of this conventional architecture, we propose an other way to process and code the input signal using an event-based sampling. In this paper, an event consists of a change in the input signal, typically an augmentation or a reduction of the input signal value. A biological event-based code is pro-

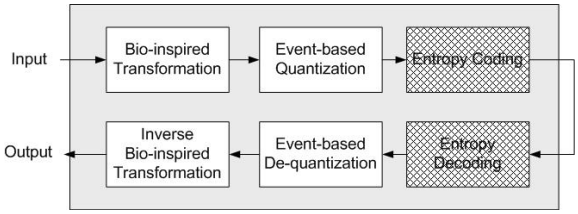


Fig. 2: Bio-inspired coding principle.

duced by the visual system which is part of the central nervous system. This code results from the way the luminance of light is captured, transformed and compressed by the inner part of the eye, the retina. The retina seems to follow the basic principles of compression by sending a spike (electrical impulse) to the cortex when the activity level reaches a threshold [4]. The very active regions of the input scene (those with the most changes) will send a lot of spikes to the brain. Each spike is a spatiotemporally correlated event. This spatiotemporal correlation is due to the layered structure of the retina which consists of different kinds of cells. The amount of cells decreases while they are nearer to the optic nerve [5, 6]. The

Thanks to 4G-TECHNOLOGY and ANRT for funding.

visual cortex is sensitive to the spatial position and also to the time when the input intensity is high enough to emit a spike. A sequence of spikes, which is called a spike train, seems to be analogous to the binary code of the computers.

The goal of this paper is to study the retinal-filter from the signal processing point of view and to produce an event-based code exploiting the Integrate And Fire (IAF) sampling (Fig. 2) [7]. In this paper, we are not dealing with an entropy coding. An approximation of the retinal filtering has been studied and applied on natural images in [8]. This is a separable spatiotemporal filter structured as a Difference of Gaussians (DoG) pyramid based on [9] and [10], each layer of which is delayed due to an exponential temporal function. We improve this filter by introducing a dependency between time and space.

In section 2, we introduce our bio-inspired filter, both in continuous and discrete time and space, which is applied on natural images. The IAF sampling which produces the spike trains is studied in section 3. The spike trains are decoded in section 4. The reconstruction of the input image is studied in section 5. The numerical reconstruction results are discussed in section 6. In the last section, we conclude the paper.

2. BIO-INSPIRED TRANSFORMATION

The aim of this section is to model a bio-inspired non-Separable sPAtioteMporal filter (non-SPAM). This filter is inspired by the mechanism of photo-receptors and horizontal cells which lie inside the retina. These cells act as edge detectors and at the same time as movement detectors due to the way they connect to each other. These are the features that the non-SPAM filter tries to mimic by having a spatial behavior which varies with respect to time. The non-separability of space and time enables the filter to detect temporal variations of luminance even in a uniform spatial region, which is not the case for a separable spatiotemporal filter.

2.1. Bio-inspired Filter

The input image $f(x,t), x \in \mathbb{R}^2, t \in \mathbb{R}^+$ is spatiotemporally convolved with the non-SPAM filter $K(x,t)$. This spatiotemporal convolution results in the function $A(x,t)$ which is called the "activation degree":

$$A(x,t) = K(x,t) \overset{x,t}{*} f(x,t) \tag{1}$$

where $\overset{x,t}{*}$ is the convolution with respect to space and time. The non-SPAM filter in continuous time and space is defined as follows:

$$K(x,t) \quad = C(x,t) - S(x,t), \tag{2}$$

where $C(x,t)$ and $S(x,t)$ are the center and the surround spatiotemporal filters given by equations (3) and (4) respectively:

$$C(x,t) \quad = w_c G_{\sigma_C}(x) W(t), \tag{3}$$

$$S(x,t) \quad = w_s G_{\sigma_S}(x) \left(W \overset{t}{*} E_{\tau_S} \right)(t), \tag{4}$$

w_c and w_s are constant parameters, G_{σ_C} and G_{σ_S} are spatial Gaussian filters standing for the center and surround areas respectively, and E_{τ_S} is an exponential temporal filter. The center temporal filter $W(t)$ is given by:

$$W(t) = E_{\tau_G,n} \overset{t}{*} (\delta_0 - w_c E_{\tau_C})(t), \tag{5}$$

where the gamma temporal filter $E_{\tau_G,n}(t)$ is defined by

$$E_{\tau,n}(t) = \frac{t^n \exp(-t/\tau)}{\tau^{n+1}}, \tag{6}$$

where $n \in \mathbb{N}$ and τ is a constant parameter ($E_{\tau,n}(t) = 0$ for $t < 0$), E_{τ_C} is an exponential temporal filter, $\delta_0(t)$ is the dirac function and $\overset{t}{*}$ stands for the temporal convolution. In case that $n = 0$, the gamma filter turns to an exponential temporal filter. The convolution of the temporal filter $W(t)$ with the exponential filter E_{τ_S} is related to the delay in the appearance of the surround temporal filter with respect to the center one.

The input signal is a still-image which exists for a long time, i.e., $f(x,t) = f(x)\mathbb{1}_{[0,\infty]}(t)$ where $f(x)$ is the still-image and $\mathbb{1}_{[0,\infty]}$ is the indicator function such that $\mathbb{1}_{[0,\infty]}(t) = 1$ if $0 \leq t \leq \infty$, otherwise 0. The following proposition shows that the spatiotemporal convolution (1) turns into a spatial convolution.

Proposition 2.1. *For a still-image $f(x,t) = f(x)\mathbb{1}_{[0,\infty]}(t)$, (1) can be rewritten as:*

$$A(x,t) \quad = \phi(x,t) \overset{x}{*} f(x), \tag{7}$$

where $\phi(x,t)$ is a spatial DoG filter weighted by two temporal filters $R_c(t)$ and $R_s(t)$:

$$\phi(x,t) = w_c R_c(t) G_{\sigma_C}(x) - w_s R_s(t) G_{\sigma_S}(x), \tag{8}$$

$$R_c(t) = \int_{u=0}^{t} W(u)du, \tag{9}$$

$$R_s(t) = \int_{u=0}^{t} (W \overset{t}{*} E_{\tau_S})(u)du. \tag{10}$$

Proof. The proof is based on the straightforward calculation of the spatiotemporal convolution (1). □

The above proposition is crucial for the reason that it enables the simplification and representation of the non-SPAM filter like a time-varying DoG given in (8). The DoG filters have been extensively studied in the past [10, 11, 12]. Proposition 2.1 shows that the retinal-inspired filter can be modeled by a spatial DoG filter which is multiplied by the temporal filters $R_c(t)$ and $R_s(t)$, shown in Fig. 3, which act like weights and modify its spatial spectrum with respect to time. Let $\hat{\phi}(\xi,t)$ be the Fourier transform of the non-SPAM filter $\phi(x,t)$ with respect to x. Fig. 4 shows 5 different time samples of $\hat{\phi}(\xi,t)$ at time $t \in \{t_1,t_2,t_3,t_4,t_5\}$ to illustrate the frequency variations of the non-SPAM filter spectrum.

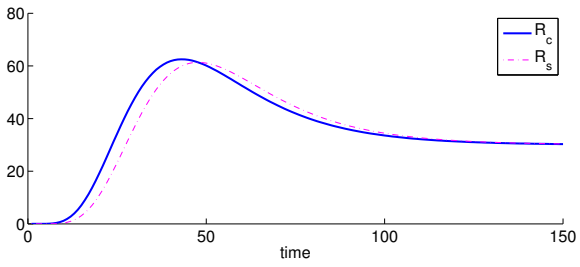


Fig. 3: Temporal filters $R_c(t)$ and $R_s(t)$.

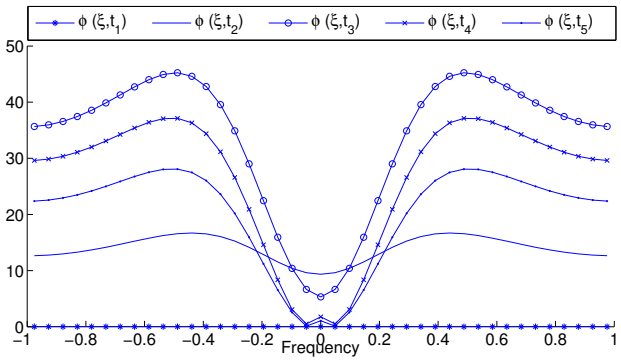


Fig. 4: Non-SPAM spectrum. The non-SPAM is a 2D spatially symmetric filter. This is a transversal cut of its spectrum for 5 different time samples of $\hat{\phi}(\xi, t)$.

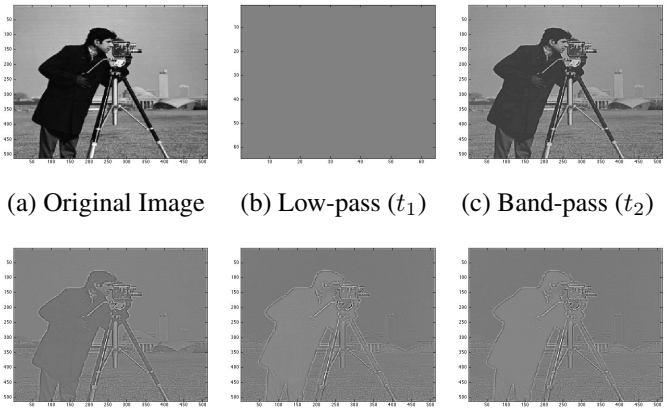
Proposition 2.2. *The function $\phi(x,t)$ is continuous and infinitely differentiable such that $\phi(x,0) = 0$ and*

$$\lim_{t \to +\infty} \phi(x,t) = \phi(x)$$

where $\phi(x)$ is a DoG filter.

Proof. The proof is given in Appendix A. □

In practice, $\phi(x,t)$ almost converges within a short time delay Δt. Hence, we assume that $\phi(x, \Delta t) \approx \phi(x)$ for all x. The time interval $[0, \Delta t]$ represents for how long the non-SPAM filter spatially evolves with respect to time. Let us consider that the time interval is decomposed into several time bins and that the non-SPAM filter $\phi(x,t)$ is (almost) constant over this time bin. It follows that the non-SPAM filter decomposes the signal in a specific way depending on the value of $\phi(x,t)$ over this time bin. Hence, the non-SPAM filter is capable to decompose the input image into different time subbands. After Δt, all the time subbands have been used and all the necessary information is obtained. The non-SPAM decomposition of the input image, for 5 different time bins, is illustrated in Fig. 5 representing the spatial evolution with respect to time which starts with the extraction of the low frequencies and ends up with the high frequencies.



(a) Original Image (b) Low-pass (t_1) (c) Band-pass (t_2)

(d) Band-pass (t_3) (e) Band-pass (t_4) (f) Band-pass (t_5)

Fig. 5: Image decomposition obtained by the non-SPAM filter for 5 different time bins.

2.2. Space and Time Discretization

For numerical purpose, we need to discretize the non-SPAM filter. Let $x_1, \ldots, x_n \in \mathbb{R}^2$ and $t_1, \ldots, t_m \in \mathbb{R}^+$ be some sets of spatial and temporal sampling points. As a consequence, the continuous spatial convolution is approximated by the discrete convolution for all x_k and t_j:

$$\begin{aligned} A(x_k, t_j) &= \phi(x_k, t_j) \circledast f(x_k) \\ &= \sum_{i=1}^{n} \phi(x_k - x_i, t_j) f(x_i), \end{aligned} \tag{11}$$

where $f = (f(x_1), \ldots, f(x_n))$ is the discretized image.

3. EVENT-BASED CODING

The activation degree $A(x_k, t)$ of x_k is fed to an IAF sampler which is responsible for the construction of the neural code. The IAF sampler produces some spikes which correspond to events in our coding chain. This model is inspired by the ganglion cells [4, 6, 13]. The ganglions are able to generate spikes when the signal they receive exits a given threshold θ. These cells are connected to each other and/or self-connected such as they construct a neural network for sending feedback messages concerning the stimulus. However, in this paper we are using the "perfect IAF integrator" without considering any feedback or leakage mechanism [14]. In addition, we assume that there is one IAF sampler for each pixel.

Let $(u_{k,i})_{i=1,\ldots,n_k}$ be the set of n_k spike times at the output of the IAF sampler of pixel x_k. In continuous time, the spike times satisfy:

$$\frac{1}{\kappa} \int_{u_{k,i}}^{u_{k,i+1}} A(x_k, t) dt = \theta \tag{12}$$

with the convention that $u_{k,0} = 0$ for all k. The positive value κ is the constant capacitance. The model (12) describes that, for a given pixel x_k, the activation degree $A(x_k, t)$ is integrated until the integral reaches the threshold θ when a spike is generated at time $u_{k,i+1}$. After firing, the integration is reset to zero and the process goes on until the next spike is emitted, etc. The set of n_k spike times $(u_{k,i})_{i=1,\ldots,n_k}$ is called a spike train. The produced code is asynchronous in the sense that the time delays between the spikes (or events) are not constant. Furthermore, each IAF sampler fires almost independently of the other IAF samplers. Without any loss of generality, the refractory period between two spikes is neglected in this study.

Let us consider the time interval $[t_j, t_{j+1}]$ which corresponds to the j-th time subband. Let us assume that $\Delta t_j = t_{j+1} - t_j$ is small enough to ensure that the activation degree $A(x_k, t) \approx A(x_k, t_j) = a_{k,j}$ is almost constant over this time interval. Then, it is clear from (12) that

$$u_{k,i+1} - u_{k,i} \approx \frac{\kappa\,\theta}{a_{k,j}}. \tag{13}$$

Hence, the spike train is periodic with the period $T_{k,j} = \frac{\kappa\,\theta}{a_{k,j}}$, which is equivalent to say the firing rate $r_{k,j} = 1/T_{k,j}$ is also constant. Fig. 6 shows the behavior of the IAF sampler for a

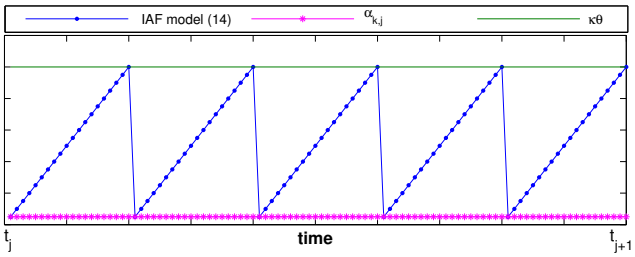


Fig. 6: Firing rate based on the activation degree.

constant input signal $A(x_k, t) \approx a_{k,j}$ during the time interval $[t_j, t_{j+1}]$. The higher the activation degree is, the sooner the IAF model will reach the threshold increasing the number of spikes. This is why the IAF model stands as a quantizer since the number of spikes produced during a given time interval codes the magnitude of $A(x_k, t_j)$.

In discrete time, the time interval $[t_j, t_{j+1}]$ is oversampled into a finer set of time samples $t_{j,1}$, $t_{j,2}$, …, $t_{j,N}$ where N is a positive integer. Each spike time $u_{k,i}$ corresponds to a discrete times $t_{j,\ell}$. Given a spike at time $u_{k,i}$, the next spike time is the first time such that the sum

$$\sum_{t_{j,\ell}=u_{k,i}}^{u_{k,i+1}} A(x_k, t_{j,\ell}) \geq \kappa\,\theta \tag{14}$$

exceeds the threshold $\kappa\,\theta$. This mechanism is clearly a discrete approximation of (12). The overlapping of spike generation between two consecutive time bins, say $[t_j, t_{j+1}]$ and $[t_{j+1}, t_{j+2}]$, are assumed to be negligible.

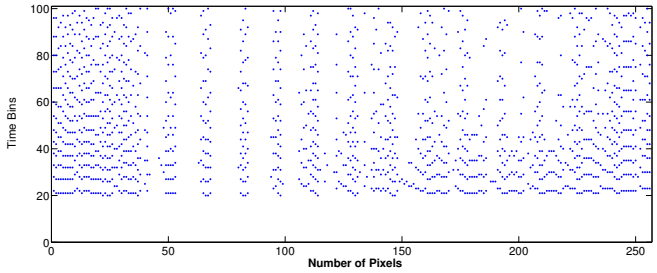


Fig. 7: Raster plot of the bio-inspired event-based code. Each dot represents a spike. For each pixel, the total number of dots forms its spike train.

Fig. 7 shows a raster plot in discrete time. It is obtained by considering 100 different time bins to compute the spike train of each pixel. As it is expected for the very first milliseconds, there are no spikes since the intensity of the activation degree is very low. Then, the firing rate becomes high because the non-SPAM subband which is coded has been filter by a low-pass filter: the main information in the input image is coded. While time increases the number of spikes decreases because they correspond to the details of the input image and most of the data have already been coded.

4. EVENT-BASED DE-QUANTIZATION

The decoding pathway based on the event-based code is performed by estimating the magnitude of the activation degree $\hat{A}(x_k, t_j)$ based on the number of spikes produced during the time interval $[t_j, t_{j+1}]$ for pixel x_k. Let us assume that $\Delta t = \Delta t_j = t_{j+1} - t_j$ is constant for all j and that $A(x_k, t_j)$ is almost constant over $[t_j, t_{j+1}]$. Let $n_{k,j}$ be the number of spikes counted during the time interval $[t_j, t_{j+1}]$ for pixel x_k. From (13), it follows that

$$n_{k,j} \approx \frac{A(x_k, t_j)\,\Delta t}{\kappa\,\theta}$$

Then a good estimation $\tilde{A}(x_k, t_j)$ over the time interval is given by

$$\tilde{A}(x_k, t_j) = n_{k,j}\,\frac{\kappa\,\theta}{\Delta t} \tag{15}$$

since κ, θ and Δt are known. Of course, the estimation $\tilde{A}(x_k, t_j)$ differs from the original activation degree $A(x_k, t_j)$ because of the quantization due to the spike generation mechanism. It is clear that $\tilde{A}(x_k, t_j) \approx A(x_k, t_j)$ when the number of spikes is sufficiently high to ensure that the ratio (15) converges to the value $A(x_k, t_j)$. This means that the threshold θ should be small enough with respect to the time interval length Δt and to the value $A(x_k, t_j)$. In the rest of the paper, these two conditions are assumed to be satisfied.

In practice, following the decoding approach proposed in [8], the relation between $n_{k,j}$ and $\tilde{A}(x_k, t_j)$ is stored in a Look-Up-Table (LUT) constructed off-line. Hence, when we receive $n_{k,j}$ spikes, we look for the corresponding value $\tilde{A}(x_k, t_j)$ in the LUT.

5. INVERSE BIO-INSPIRED TRANSFORMATION

Once the activation degrees $A(x_k, t_j)$ are estimated from the spike trains, the reconstruction of the input still-image requires to inverse the non-SPAM filter. Due to the big redundancy of the transformed coefficients, we exploit the frame theory in order to prove that the filter is invertible. Let us denote $\hat{\phi}_{t_j}(\xi)$ the discrete Fourier transform of the vector $(\phi(x_1, t_j), \ldots, \phi(x_n, t_j))$ with respect to space $x_1, \ldots, x_n$, $\|f\|$ the Euclidean norm of the discrete input signal f, $\varphi_{k,j}$ the row vector of $\mathbb{R}^n$ defined by

$$\varphi_{k,j} = \Big(\phi(x_k - x_1, t_j), \ldots, \phi(x_k - x_n, t_j)\Big),$$

and Φ is the family of all vectors $\varphi_{k,j}$.

Proposition 5.1. *The family of vectors Φ is a frame, i.e., there exist two scalars $0 < \alpha \leq \beta < \infty$ such that:*

$$\alpha\|f\|^2 \quad \leq \sum_{j=1}^{m}\sum_{k=1}^{n} |A(x_k, t_j)|^2 \quad \leq \beta\|f\|^2, \tag{16}$$

where

$$\alpha = \min_{\xi} \left\{ \frac{1}{n} \sum_{j=1}^{m} \left| \hat{\phi}_{t_j}(\xi) \right|^2 \right\} > 0,$$

$$\beta = \sum_{j=1}^{m}\sum_{k=1}^{n}\sum_{i=1}^{n} \phi^2(x_k - x_i, t_j).$$

Proof. The proof is given in Appendix B. □

Proposition 5.1 proves that the non-SPAM filter is a frame, hence the filter is invertible and we can reconstruct the input image. The optimal reconstruction results are given when all the coefficients $A(x_k, t_j)$ are available at final discrete time t_m.

In practice, the reconstruction is processed as follows. Let us define $\tilde{A} = [\tilde{A}_{t_1}, \ldots, \tilde{A}_{t_m}]$ as a vector of size nm and $\Phi = [\phi_1, \ldots, \phi_m]$ a matrix of size $nm \times n$, where

$$\tilde{A}_{t_j} = \begin{bmatrix} \tilde{A}(x_1, t_j) \\ \vdots \\ \tilde{A}(x_n, t_j) \end{bmatrix} \quad \text{and} \quad \phi_j = \begin{bmatrix} \varphi_{1,j} \\ \vdots \\ \varphi_{n,j} \end{bmatrix}. \tag{17}$$

At time t_m, we propose to compute $\hat{f}_{t_m}$ which is the estimation of f given by:

$$\tilde{f}_{t_m} = (\Phi^\top \Phi)^{-1} \Phi^\top \tilde{A}, \tag{18}$$

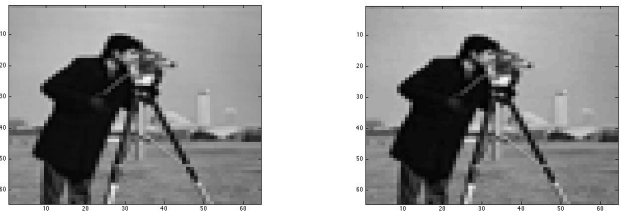
where M^{-1} denotes the inverse of a matrix M and $M^\top$ denotes its transpose. The dual frame, which is necessary to have a perfect decoding at time t_m [12, 15], is $(\Phi^\top \Phi)^{-1}\Phi^\top$. Instead of computing the above matrix operator which can be time consuming and resource demanding, we can note that (18) is a solution of the following least squares problem:

$$\tilde{f}_{t_m} = \arg\min_{f \in \mathbb{R}^n} \left(\sum_{j=1}^{m} \|\phi_j \circledast f - \tilde{A}_{t_j}\|^2 \right). \tag{19}$$

Such a problem can be easily solved by using a gradient descent algorithm.

6. NUMERICAL RESULTS

The numerical results are given in Fig 8. We define the Mean Square Error as $\text{MSE}(f, \tilde{f}_{t_m}) = ||f - \tilde{f}_{t_m}||^2/n$, which measures the distortion between the original image f and the reconstructed image $\tilde{f}_{t_m}$. For these experiments we used an image of size $n = 64 \times 64$ on which we applied a non-SPAM filter of $m = 5$ subbands. In these experiments, the parameters of the IAF sampler are chosen such that the spikes lead to a lossless quantization of the input signal (threshold $\theta = 10^4$ and the constant capacitance $\kappa = 1$). Hence, we focus on the results of the inverse bio-inspired transformation. As it is expected the reconstructed image is very close to the original one and the distortion is very small due to the proof in section 5. The $\text{MSE}(f, \tilde{f}_{t_m}) = 0.3963$ confirms the quality of the reconstruction.



(a) Original Image (64x64) (b) Reconstructed Image

Fig. 8: Frame-based reconstruction of the input image.

7. CONCLUSION

In this paper, we discussed a novel event-based architecture for compressing a signal. We introduced a bio-inspired filter which is invertible and able to extract all the necessary information of a still-image. This filter is the input of an integrate and fire sampler which produces a spike train for each pixel. It is shown that the decoder allows a perfect reconstruction of the input image.

In our future work we aim to improve this model utilizing a Leaky Integrate and Fire (LIF) sampler as a quantizer in our model. The LIF is necessary to be adapted and studied in our architecture because of the big redundancy of the information that the non-SPAM filter produces and it needs to be eliminated.

A. APPENDIX A

This appendix proves Proposition 2.2. It is shown that both $R_c(t)$ and $R_s(t)$ can be defined in a closed form as polynomial functions which are attenuated by exponential ones. The calculation of $R_c(t)$ and $R_s(t)$ are based on the following lemma.

Lemma A.1. *Let ω a real value and $t \geq 0$. Using an integration by parts, we obtain the following equality:*

$$\begin{aligned}
&\int_0^t u^n \exp(-\omega u)\, du \\
&= \left(\sum_{k=0}^{n} -\frac{(n!)}{(n-k)!\omega^{k+1}} t^{n-k} \exp(-\omega t)\right) + \frac{n!}{\omega^{n+1}} \\
&= P_n(t)\exp(-\omega t) + \Omega,
\end{aligned}$$

where $P_n(t)$ is a polynomial function in t of order n depending of ω and Ω is a constant value.

A.1. Closed-form of $R_c(t)$

A straightforward calculation based on Lemma A.1 shows that:

$$\begin{aligned}
R_c(t) &= \int_{u=0}^{t} W(u)du \\
&= \int_{u=0}^{t} E_{\tau_G,n} \overset{t}{*} (\delta_0 - w_c E_{\tau_C})(u)du \\
&= \int_{u=0}^{t} \frac{u^n \exp\left(\frac{-u}{\tau_G}\right)}{\tau_G^{n+1}} du - w_c \int_{u=0}^{t} \frac{\exp\left(\frac{-t}{\tau_c}\right)}{\tau_c} du \\
&= \sum_{k=0}^{n} -\frac{(n!)}{(n-k)!a^{k+1}\tau_G^{n+1}} t^{n-k}\exp(-at) + n! \\
&\quad -\sum_{k=0}^{n}\sum_{l=0}^{m} \frac{n! w_c}{(m-l)!b^{k+1}a^{l+1}\tau_G^{n+1}\tau_C} t^{m-l}\exp(-at) \\
&\quad +\sum_{k=0}^{n} \frac{n! w_c}{b^{k+1}a^{m+1}\tau_G^{n+1}\tau_C} - \frac{n! w_c}{b^{n+1}\tau_G^{n+1}} \\
&\quad +\frac{n! w_c}{b^{n+1}\tau_G^{n+1}} \exp\left(\frac{-t}{\tau_C}\right),
\end{aligned}$$

where $a = \frac{1}{\tau_G}, b = \frac{\tau_A - \tau_G}{\tau_G \tau_A}$ and $m = n - k$. Hence,

$$R_c(t) = P_n(t)\exp\left(\frac{-t}{\tau_G}\right) + \alpha_c \exp\left(\frac{-t}{\tau_C}\right) + \gamma_c$$

where $P_n(t)$ is a polynomial function in t of order n and α_c and γ_c are two reals.

A.2. Closed-form of $R_s(t)$

A straightforward calculation based on Lemma A.1 shows that:

$$\begin{aligned}
R_s(t) &= \int_{u=0}^{t} (W \overset{t}{*} E_{\tau_S})(u)du \\
&= \int_{u=0}^{t} E_{\tau_G} * (\delta_0 - w_A E_{\tau_A}) * E_{\tau_S}(u)du \\
&= \int_{u=0}^{t} E_{\tau_G} * E_{\tau_S}(u)du - w_A \int_{u=0}^{t} E_{\tau_G} * E_{\tau_A} * E_{\tau_S}(u)du \\
&= \frac{1}{\tau_G^{n+1}\tau_S}\left(\sum_{k=0}^{n}\sum_{l=0}^{m} \frac{n!}{g^{k+1}a^{l+1}(m-l)!} t^{m-l}\exp(-at)\right. \\
&\quad -\sum_{k=0}^{n} \frac{n!}{g^{k+1}a^{m+1}} + \frac{n!\tau_S}{g^{k+1}} - \frac{n!\tau_S}{g^{k+1}}\exp\left(\frac{-t}{\tau_S}\right)\Big) \\
&\quad -\frac{w_c}{\tau_G^{n+1}\tau_C\tau_S}\left(\sum_{k=0}^{n}\sum_{l=0}^{m}\sum_{r=0}^{p} -\frac{n!}{(p-r)!b^{k+1}g^{l+1}a^{r+1}} t^{p-r}\exp(-at)\right. \\
&\quad +\sum_{k=0}^{n}\sum_{l=0}^{m} \frac{n!}{b^{k+1}g^{l+1}a^{p+1}} - \sum_{k=0}^{n} \frac{n!\tau_S}{b^{k+1}g^{m+1}}\left(1-\exp\left(\frac{-t}{\tau_S}\right)\right) \\
&\quad \left.\left. +\frac{n!\tau_S}{b^{n+1}\phi}\left(1-\exp\left(\frac{-t}{\tau_S}\right)\right) - \frac{n!\tau_C}{b^{n+1}\phi}\left(1-\exp\left(\frac{-t}{\tau_C}\right)\right)\right)\right),
\end{aligned}$$

with

$$g = \frac{\tau_S - \tau_G}{\tau_G\tau_S},\ \phi = \frac{\tau_S - \tau_C}{\tau_C\tau_S},\ \xi = \frac{1}{\tau_S} + \phi = \frac{1}{\tau_C},$$

and $p = m - l$, $m = n - k$. It follows that

$$R_s(t) = Q_n(t)\exp\left(\frac{-t}{\tau_G}\right) + \alpha_s \exp\left(\frac{-t}{\tau_S}\right) + \beta_s \exp\left(\frac{-t}{\tau_C}\right) + \gamma_s$$

where $Q_n(t)$ is a polynomial function in t of order n and α_s, β_s and γ_s are some reals. We can verify that $\gamma_s = \gamma_c$. Hence, $R_c(t)$ and $R_s(t)$ converge to the same value as $t \to +\infty$.

B. APPENDIX B

This appendix establishes that the non-SPAM filter is a frame. First, we study the upper bound and next the lower bound.

B.1. Upper bound

We are using the Cauchy-Schwarz inequality to calculate the upper bound of the non-SPAM frame:

$$\sum_{j=1}^{m}\sum_{k=1}^{n}\Big|A(x_k,t_j)\Big|^2=\sum_{j=1}^{m}\sum_{k=1}^{n}\Big|\phi(x_k,t_j)\circledast f(x_k)\Big|^2$$

$$=\sum_{j=1}^{m}\sum_{k=1}^{n}\Big|\sum_{i=1}^{n}\phi(x_k-x_i,t_j)f(x_i)\Big|^2$$

$$=\sum_{j=1}^{m}\sum_{k=1}^{n}\Big|\sum_{i=1}^{n}\varphi_{k,j}(x_i)f(x_i)\Big|^2$$

$$\leq\sum_{j=1}^{m}\sum_{k=1}^{n}\Big(\Big|\sum_{i=1}^{n}\varphi_{k,j}^2(x_i)\Big|\Big|\sum_{i=1}^{n}f^2(x_i)\Big|\Big)$$

$$=\Big(\sum_{j=1}^{m}\sum_{k=1}^{n}\Big|\sum_{i=1}^{n}\varphi_{k,j}^2(x_i)\Big|\Big)\Big|\sum_{i=1}^{n}f^2(x_i)\Big|$$

$$=\Big(\Big|\sum_{j=1}^{m}\sum_{k=1}^{n}\sum_{i=1}^{n}\phi^2(x_k-x_i,t_j)\Big|\Big)\Big|\sum_{i=1}^{n}f^2(x_i)\Big|$$

$$=\beta\Big\|f\Big\|^2.$$

B.2. Lower bound

First of all, we use the Parserval Theorem to transform the activation degree coefficients in Fourier domain:

$$\sum_{j=1}^{m}\sum_{k=1}^{n}\Big|A(x_k,t_j)\Big|^2=\sum_{j=1}^{m}\sum_{k=1}^{n}\Big|A_{t_j}(x_k)\Big|^2$$

$$=\sum_{j=1}^{m}\sum_{\xi=1}^{n}\frac{1}{n}\Big|\hat{A}_{t_j}(\xi)\Big|^2=\frac{1}{n}\sum_{j=1}^{m}\sum_{\xi=1}^{n}\Big|\hat{\phi}_{t_j}(\xi)\hat{f}(\xi)\Big|^2$$

$$=\frac{1}{n}\sum_{\xi=1}^{n}\sum_{j=1}^{m}\Big|\hat{\phi}_{t_j}(\xi)\Big|^2\Big|\hat{f}(\xi)\Big|^2.$$

Let us prove by contradiction that $\sum_{j=1}^{m}\Big|\hat{\phi}_{t_j}(\xi)\Big|^2\neq 0$ for all t_j. By definition, the equality

$$\sum_{j=1}^{m}\Big|\hat{\phi}_{t_j}(\xi)\Big|^2=0$$

is equivalent to

$$\sum_{j=1}^{m}\Big|w_cR_c(t_j)\hat{G}_c(\xi)-w_sR_s(t_j)\hat{G}_s(\xi)\Big|^2=0,$$

which means that

$$w_cR_c(t_j)\hat{G}_c(\xi)=w_sR_s(t_j)\hat{G}_s(\xi),\ \forall t_j. \tag{20}$$

From Appendix A, one can show that, for $t=0$, $R_c(0)>0$. In addition, the Fourier transform of a Gaussian is again a Gaussian which means that $\forall\xi,\ \hat{G}_c(\xi)\geq 0$ and $\hat{G}_s(\xi)\geq 0$. As a result, we can rewrite (20) as following:

$$\frac{w_c\hat{G}_c(\xi)}{w_s\hat{G}_s(\xi)}=\frac{R_s(t_j)}{R_c(t_j)},\ \forall t_j. \tag{21}$$

Let t_{eq} be the time when $R_c(t_{eq})=R_s(t_{eq})$. We have the following results:

- For $0\leq t_j\leq t_{eq}$, one can show that $R_c(t_j)>R_s(t_j)$. As a result, $\frac{R_s(t_j)}{R_c(t_j)}<1$.
- For $t_{eq}\leq t_j$, one can show that $R_c(t_j)<R_s(t_j)$. As a result, $\frac{R_s(t_j)}{R_c(t_j)}>1$.

It follows from these results that $\frac{R_s(t_j)}{R_c(t_j)}$ is not constant with respect to t_j. Hence, the ratio on the left in (21) should vary with respect to t_j. This contradicts the fact that this ratio, which is just a function of ξ, does not depend on t_j. It can be concluded that

$$\sum_{j=1}^{m}\Big|\hat{\phi}_{t_j}(\xi)\Big|^2>0$$

for all ξ. Hence, it follows that

$$\alpha=\min_{\xi}\Big\{\frac{1}{n}\sum_{j=1}^{m}\Big|\hat{\phi}_{t_j}(\xi)\Big|^2\Big\}>0.$$

The value α is the lower bound given in Proposition 5.1.

C. REFERENCES

[1] C. Christopoulos, A. Skodras, and T. Ebrahimi, "The JPEG2000 still iamge coding system: An overview," *Transactions on Consumer Electronics*, vol. 46, no. 4, pp. 1103–1127, November 2000.

[2] G. Bjontegaard T. Wiegand, G. J. Sullivan and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Transaction on circuits and systems for video technology*, vol. 13, no. 7, pp. 560–576, July 2003.

[3] A. Mulayoff D. Grois, D. Marpe and O. Hadar, "Performance comparison of H.265/MPEG-HEVC, VP9, and H.264/MPEG-AVC encoders," *30th Picture Coding Symposium 2013 (PCS 2013)*, December 2013.

[4] A. Wohrer and P. Kornprobst, "Virtual retina: A biological retina model and simulator, with constrast gain control.," *Journal of Computational Neuroscience*, vol. 26, no. 2, pp. 219–249, 2009.

[5] T. Gollisch and M. Meister, "Eye smarter than scientists believed: Neural computations in circuits of the retina," *Neuron*, vol. 65, no. 2, pp. 150–164, 2010.

[6] R. Masland, "The fundamental plan of the retina.," *Natural Neuroscience*, vol. 4, no. 9, pp. 877–886, 2001.

[7] Hans G. Feichtinger, José Príncipe, José Luis Romero, Alexander Singh Alvarado, and Gino Angelo Velasco, "Approximate reconstruction of bandlimited functions for the integrate and fire sampler," *Advances in Computational Mathematics*, vol. 36, no. 1, pp. 67–78, 2012.

[8] K. Masmoudi, M. Antonini, and P. Kornprobst, "Streaming an image through the eye: The retina seen as a dithered scalable image coder," *Signal processing Image Communication*, vol. 28, no. 8, pp. 856–869, 2013.

[9] P. Burt and E. Andelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Commun*, vol. 31, no. 4, pp. 532–540, 1983.

[10] David J. Field, "What is the goal of sensory coding?," *Neural Computation*, vol. 6, pp. 559–601, 1994.

[11] G. DeAngelis D. Cai and R. Freeman, "Spatiotemporal receptive field organization in the Lateral Geniculate Nucleus of cats and kittens," *The American Physiological Society*, vol. 22, no. 3077, pp. 1045–1061, 1997.

[12] K. Masmoudi, M. Antonini, and P. Kornprobst, "Frames for exact inversion of the rank order coder," *IEEE Transactin on Neural Networks*, vol. 23, no. 2, pp. 353–359, February 2012.

[13] S. J. Thorpe and J. Gautrais, "Rank Order Coding: A new coding scheme for rapid processing in neural network," *Computational Neuroscience: Trends in Research*, , no. 113-118, 1998.

[14] A. N. Burkitt, "A review of the integrate-and-fire neuron model: I. homogeneous synaptic input," *Biological Cybernetics*, vol. 95, no. 2, pp. 97–112, 2006.

[15] J. Kovacevic and A. Chebina, "An introduction to frames," *Signal Processing*, vol. 2, no. 1, pp. 1–94, 2008.