Signal reconstruction from a bio-inspired event-based code

Effrosyn DOUTSI[1,2], Lionel FILLATRE[1], Marc ANTONINI[1], 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.

doutsi@i3s.unice.fr, lionel.fillatre@i3s.unice.fr,
am@i3s.unice.fr, julien.gaulmin@4g-technology.eu.

Résumé – Cet article présente une nouvelle architecture bio-inspirée pour encoder, puis décoder, un signal 1D. Un filtrage spatio-temporel est appliqué au signal d'entrée afin de le décomposer en plusieurs composantes. Ensuite, chaque composante est échantillonnée par un neurone du type "intègre-et-tire". Cet échantillonneur produit ainsi un train d'impulsions pour chaque composante, ce qui équivaut à quantifier le signal d'entrée. Nous prouvons alors que le signal d'entrée peut être reconstruit à partir des trains d'impulsions. Cette reconstruction s'appuie principalement sur le fait que le filtre spatio-temporel proposé a une structure de frame. Des simulations numériques montrent l'efficacité de cette nouvelle architecture bio-inspirée.

Abstract – This paper introduces a novel and complete architecture of a bio-inspired codec which is applied to a 1D-signal. The result of this filtering is a spatiotemporal decomposition of each component of the input signal which is fed to a Non Leaky Integrate and Fire (NLIF) sampler. The NLIF sampler is a quantizer of the input signal. Using a threshold we are able to turn the quantized signal into a sequence of spikes, which is called spike train. The number of spikes is related to the amplitude of the signal. We prove that we can estimate the quantized signal based on the spike trains. In addition, based on frame theory we prove that the retinal filter has a frame structure, so it is invertible and enables the reconstruction of the input signal. Numerical results show the performance of our architecture.

1 Introduction

Compression has been one of the most challenging research fields for the last decades. An efficient compression algorithm manages to convert the input signal into another and more compact form of data for transmition or storage needs. At the same time, this kind of algorithms guarantee high reconstruction quality. The more complicated the signal is (i.e video) the more the constraints and the difficulties one has to encounter to find the trade-off between the compression ratio of the code and the quality of the reconstruction.

There have been many qualitative standards for image and video compression which have been released during the last decades [1, 2]. However, the most recent the standard is, the more its complexity increases without a dramatic improvement in the compression ratio. At the same time, it is required a significant number of years in order to release a new standard. As a result, there is a general feeling that a novel and groundbreaking architecture should be released, which is going to minimize the energy of the system while at the same time its compression ratio will be high.

There are noticeable similarities between the compression principle of image and video codecs and the way the mammalian visual system captures, transforms and compresses the luminence of light at the inner part of the eye, the retina. The output of the retina is a very compact and informative code of spikes (electrical impulse) which is produced when the activity level reaches a threshold [3]. This code is sent to the visual cortex. The very active regions of the input scene will send a lot of spikes to the brain. Each spike is a spatiotemporal correlated event.

The goal of this paper is to study the retinal compression principle from the signal processing point of view. Our contributions are the construction of the non-Separable sPAtioteMporal (non-SPAM) transformation, which is an improvement of previous models like [4, 5] and the use of the Non Leaky-Integrate and Fire (NLIF) sampler [6]. We apply this novel compression architecture to a 1D-signal and we prove that we are able to reconstruct based on the code of spikes.

In section 2 we introduce our bio-inspired filter, both in continuous and discrete time and space, which is applied on an 1D signal. The NLIF sampling which produces the spike trains is studied in section 3. The reconstruction of the input image is introduced in section 4 and the numerical reconstruction results are discussed in section 5. In the last section, we conclude the paper with a discussion about the future work.

2 Bio-inspired Transformation

The aim of this section is to introduce a wavelet-like retinal-inspired decomposition of 1D-signal. For this reason, we have studied a non-Separable sPAtioteMporal (non-SPAM) filter which has been modeled in [7] and mimics the mechanism of photoreceptors and horizontal cells which lie inside the retina. The non-separability of space and time means that the spatial behavior of the filter varies with respect to time. As a result, for a given input signal which varies with respect to time, when the non-SPAM filter is applied, it is possible to succeed in extracting it the beginning basic information of the signal and while time increases these data are enriched with more details. The non-SPAM filter $K(x,t)$, where $x \in \mathbb{R}$ and $t \in \mathbb{R}^+$, has

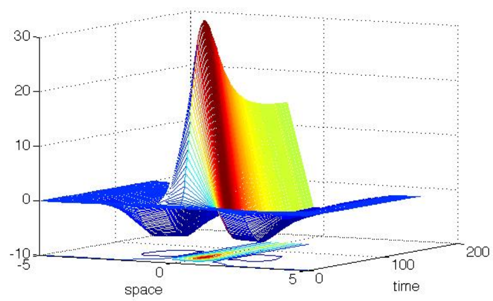

FIGURE 1 – The non-SPAM filter $K(x,t)$.

been introduced in [7] and studied in [8]. This filter is a combination of Gaussian spatial filters and exponential temporal filters which ends up to a structure of a time depending Difference of Gaussian (DoG) function (Figure 1). This function $K(x,t)$ is convolved with a 1D-spatial signal which exists for a long time (Figure 2), hence $f(x,t) = f(x)\mathbb{1}_{[0,\Delta T]}(t)$ where $f(x)$ is the 1D-signal and $\mathbb{1}$ is the indicator function such that $\mathbb{1}_{[0,\infty]}(t) = 1$, if $0 \leq t \leq \infty$, otherwise 0 :

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

where $\overset{x,t}{*}$ is the convolution with respect to space and time. The function $A(x,t)$ is called *activation degree* and consists of the transformed coefficients (Figure 2). In our previous work [8]

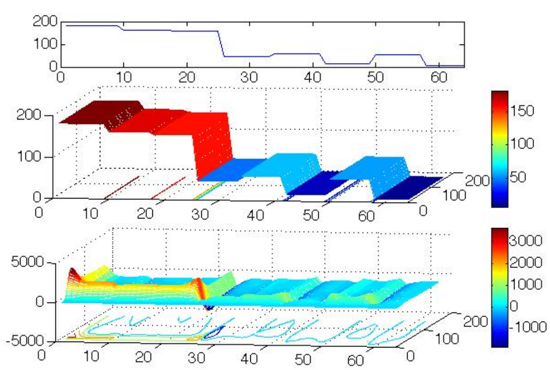

FIGURE 2 – Top : 1D spatial signal $f(x)$. Middle : 1D spatial filter $f(x)\mathbb{1}_{[0,\Delta T]}(t)$ flashed for a given time ΔT. Bottom : Spatiotemporally filtered signal $A(x_k,t_j)$.

we have also proven that the spatiotemporal convolution (1) turns into a spatial convolution :

$$A(x,t) = \phi(x,t) \overset{x,t}{*} f(x), \tag{2}$$

$$\phi(x,t) = w_c R_c(t) G_{\sigma_C}(x) - w_s R_s(t) G_{\sigma_S}(x), \tag{3}$$

where $R_c(t)$ and $R_s(t)$ are temporal filters and G_{σ_C} and G_{σ_S} are Gaussian spatial filters. We have mathematically proven in [9] that the temporal filters are polynomial functions attenuated by exponential ones. 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 (1) is approximated by the discrete spatial convolution :

$$\begin{aligned} A(x_k,t_j) &= \phi(x_k,t_j) \circledast f(x_k) \\ &= \sum_{i=1}^{n} \phi(x_k - x_i, t_j) f(x_i), \end{aligned} \tag{4}$$

for all k and j, where $f = (f(x_1), \ldots, f(x_n))$ and $\phi(x_k,t_j)$ are the discretized signal and non-SPAM filter respectively.

3 Event-based Quantization

The bio-inspired code is generated in this section based on the NLIF model which is inspired by the ganglion cells [3, 10, 11]. The NLIF model, which is applied to every component of the input signal x_k, comes up to the following outputs : the sparse code of spikes and the quantized intensity of the spatiotemporal coefficients.

The discretized activation degree $A(x_k,t_j)$ is fed to the NLIF sampler. The NLIF integrates the spatiotemporally transformed coefficients. When the integration value reaches a given threshold θ it emits a spike (Figure 4). Let $\delta_{k,l} : 1 \leq l \leq n_k$ the set of time instances when the component x_k spikes :

$$\sum_{t=\delta_{k,l}}^{\delta_{k,l+1}} A(x_k,t) \geq \theta, \;\; \text{for} \;\; t \in \{t_1, \ldots, t_m\}, \tag{5}$$

The raster plot according to the NLIF sampler is given by (6) and is illustrated in Figure 3 :

$$R_{k,j} = \begin{cases} 1, & \text{if} \;\; t_j \in \{\delta_{k,1}, \ldots, \delta_{k,n_k}\} \quad \forall k, \forall j \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

As expected, for small time the firing rate is high since the non-SPAM filter has a low-pass structure. However, while time increases the non-SPAM filter turns into a high-pass. As a result, there are finer data which are going to be coded. After each firing, the integration value is reset to zero and the process goes on until the next spike is emitted.

The NLIF model is considered as a deterministic quantizer which gets as an input a signal and always transforms it into the same sparse code of events for a given threshold. The time when each event happens is related to the integration value. Hence, a spike which is emitted very soon means that the integration value reached the threshold θ very fast and as a result, there were only a few spatiotemporal coefficients $A(x_k,t_j)$ counted. We propose that the NLIF model is able to quantize the input signal with respect to the emitted spikes and turn the input signal into a piecewise constant function according to the following assumption :

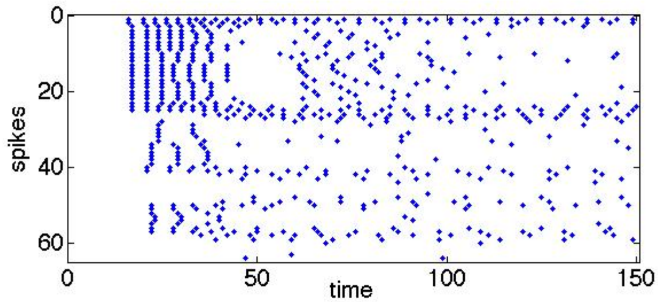

FIGURE 3 – NLIF-based quantization of the input signal.

Assumption 3.1. *It is assumed that* $\forall t_j \in [\delta_{k,l}, \delta_{k,l+1}],\ \forall k, \forall l,$ *there is* θ *which is chosen such that :*

$$A(x_k, t_j) \simeq A_{k,j} = A_{k,j-1} + \frac{A(x_k, t_l) - A(x_k, t_{l-1})}{2},$$

where $A_{k,j}$ *is a constant,* $A(x_k, t_l)$ *is the intensity of the spatiotemporal frame coefficient* x_k *at* $\delta_{k,l}$*, when the* l^{th} *spike is emitted and* $A_{k,0} = 0$.

Assumption 3.1 means that, between two sequential spikes $\delta_{k,l}, \delta_{k,l+1}$, we define the value of the piecewise constant function as the sum of the arithmetic mean of the activation degree $A(x_k, t_j)$ at time t_l, t_{l+1} and the previous constant value of the piecewise function. The above assumption is illustrated in Figure 4 and its goal is to represent the loss of information due to the NLIF. In addition, this assumption enables to evaluate the quality of the reconstructed signal, which is a piecewise constant approximation of the original 1D-signal.

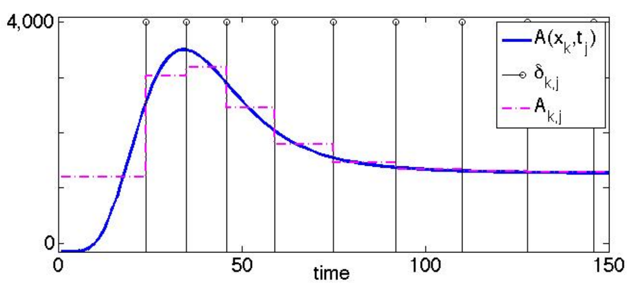

FIGURE 4 – The spatiotemporal evolution $A(x_k, t_j)$ is plotted together with its piecewise constant approximation $A_{k,j}$ and its raster plot.

4 Reconstruction

The reconstruction architecture consists of two steps : first of all it is necessary to decode the signal based on the raster plot and secondly, the non-SPAM filter should be inverted to reconstruct the input signal by its quantized activation degree.

4.1 NLIF-based De-quantization (De-NLIF)

The amplitude of the quantized activation degree can be estimated by the spike code if we know the threshold, the delay $\delta_{k,l+1} - \delta_{k,l}$ which is required for each spike to be emitted [12]. Hence, the input signal is quantized before it is fed to the NLIF according to the following equation :

$$(\delta_{k,l+1} - \delta_{k,l}) * A_{k,j} \simeq \theta. \quad (7)$$

Hence an estimation of $A_{k,j}$ is given by :

$$\tilde{A}_{k,j} \simeq \frac{\theta}{(\delta_{k,j+1} - \delta_{k,j})}. \quad (8)$$

Figure 5 illustrates the estimation of the quantized spatiotemporal coefficients of one component of the input signal according to the reconstruction which has been just introduced in (8).

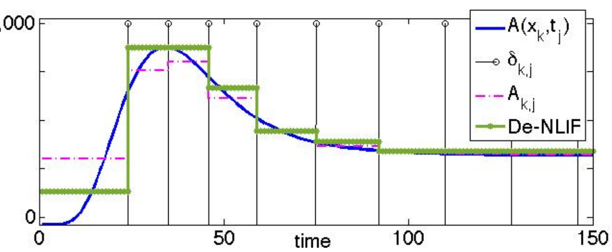

FIGURE 5 – NLIF-based de-quantization.

4.2 Inverting the non-SPAM

Once the activation degree of each component x_k is estimated from the spike trains, the reconstruction of the input signal 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 [13, 14]. We have proven in [9] that the non-SPAM filter is a frame and such that we are able to reconstruct. 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 :

$$\tilde{A}_{t_j} = \begin{bmatrix} \tilde{A}_{1,j} \\ \vdots \\ \tilde{A}_{k,j} \end{bmatrix} \quad \text{and} \quad \phi_j = \begin{bmatrix} \varphi_{1,j} \\ \vdots \\ \varphi_{n,j} \end{bmatrix}, \quad (9)$$

where $\varphi_{k,j} = \Big(\phi(x_k - x_1, t_j), \ldots, \phi(x_k - x_n, t_j)\Big)$. At time t_m, we propose to compute $\tilde{f}_{t_m}$ which is the estimation of f given by :

$$\tilde{f}_{t_m} = (\Phi^\top \Phi)^{-1} \Phi^\top \tilde{A}, \quad (10)$$

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, is $(\Phi^\top \Phi)^{-1}\Phi^\top$. We can invert $\Phi^\top \Phi$ because the non-SPAM filter is a frame.

5 Numerical Results

The numerical results are given in Figure 6. We define the Mean Square Error as $\text{MSE}(f, \tilde{f}_{t_m}) = ||f - \tilde{f}_{t_m}||^2$ which measures the distortion between the original signal f and the reconstructed signal $\tilde{f}_{t_m}$. For these experiments we used a 1D-signal (1x64) on which we applied a non-SPAM filter of m=150 subbands. As it is expected there is a loss of information due to the quantization of the activation degree before the NLIF sampler and the NLIF-based de-quantization which increases the distortion.

We need to highlight first of all that the reconstruction of the spatiotemporal coefficients without the use of the NLIF model is optimal according to the frame proof which is confirmed by the numerical results. In addition, the value of threshold θ is related to the maximum value of the non-SPAM frame ($\max(A(x_k, t_j)) = 3658$). According to Figure 6, while the threshold increases the quality of the reconstruction results is getting closer to the original signal.

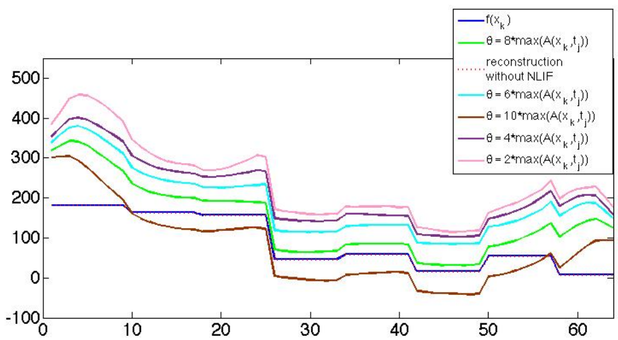

FIGURE 6 – Frame-based reconstruction of the input signal.

6 Conclusion

In this paper we have introduced a complete bio-inspired codec of 1D-signals. The non-SPAM filter and the NLIF sampler are the basic components of this system which produces a code of spikes. Based on this code we reconstruct an lossy approximation of the input signal.

In the future, we aim to study more complicated bio-inspired quantizers like the Leaky Integrate and Fire (LIF) model or noisy-LIF which model better the biological functions and is expected to give us more robust and reliable results.

Références

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

[2] D. Grois, D. Marpe, A.Mulayoff, 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.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] A. Wohrer, P. Kornprobst, and M. Antonini, "Retinal filtering and image reconstruction.," Tech. Rep., Inria Research ReportRR- 6960, 2009.

[8] E. Doutsi, L. Fillatre, M. Antonini, and J. Gaulmin, "Retina-inspired filtering for dynamic image coding," *IEEE International Conference in Image Processing (ICIP)*, 2015.

[9] E. Doutsi, L. Fillatre, M. Antonini, and J. Gaulmin, "Event-based coding of images using a bio-inspired frame," *International Conference on Event-Based Control, Communication and Signal Processing (EBCCSP)*, 2015.

[10] 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.

[11] 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.

[12] R. Van Rullen and S. J. Thorpe, "Rate coding versus temporal order coding : What the retinal ganglion cells tell the visual cortex," *Neural Computation*, vol. 13, pp. 1255–1283, 2001.

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

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