A Retina-inspired encoder: An innovative step on image coding using Leaky Integrate-and-Fire neurons

Melpomeni Dimopoulou, Effrosyni Doutsi, Marc Antonini

HAL Id: hal-01901140

https://hal.science/hal-01901140

A RETINA-INSPIRED ENCODER: AN INNOVATIVE STEP ON IMAGE CODING USING LEAKY INTEGRATE-AND-FIRE NEURONS

Melpomeni Dimopoulou, Effrosyni Doutsi, Marc Antonini

Université Côte d'Azur, I3S, CNRS, UMR7271, 06900 Sophia Antipolis, France
dimopoulou@i3s.unice.fr
doutsi@i3s.unice.fr
am@i3s.unice.fr

ABSTRACT

This paper aims to build an image coding system based on a model of the mammalian retina. The retina is the light-sensitive layer of tissue located on the inner coat of the eye and it is responsible for vision. Inspired by the way the retina handles and compresses visual information and based on previous studies we aim to build and analytically study a retinal-inspired image quantizer, based on the Leaky Integrate-and-Fire (LIF) model, a neural model approximating the behavior of the ganglion cells of the Ganglionic retinal layer that is responsible for visual data compression. In order to have a more concrete view of the encoder's behavior, in our experiments, we make use of the spatiotemporal decomposition layers provided by extensive studies on a previous retinal layer, the Outer Plexiform Layer (OPL). The decomposition layers produced by the OPL, are being encoded using our LIF image encoder and then, they are reconstructed to observe the encoder's efficiency.

Index Terms— Retina, Ganglion cells, Leaky Integrate-and-Fire (LIF) model, neural coding, image coding, weighted difference of Gaussians.

1. INTRODUCTION

As technology advances, the need for finding new ways for the efficient transmission and storage of information augments dramatically. Living in the age of the social networks, the media to be stored and transmitted grows rapidly. However, despite the fact that during the past few decades compression standards kept evolving, the compression ratio does not evolve accordingly to the needs. Consequently, the urge for finding new means of compression remains to be of a great importance. With this paper, we aim to propose a different, bio-inspired, dynamic approach for the encoding of images.

Our work is being inspired by the mammalian retina. The retina can be divided into three basic layers. The Outer Plexiform Layer (OPL) which acts as a spatiotemporal filter, the Inner Plexiform Layer (IPL) that performs a non-linear rectification, and the Ganglionic Layer which is responsible for the encoding of the visual data. The Ganglionic layer consists of the ganglion cells, a type of neuron that compresses visual information according to the Leaky Integrate-and-Fire (LIF) neural model which encodes intensity values into spikes. Under the main belief that nature performs in an optimal way, and based on previous works on the OPL filtering in [1], we built a quantization system making use of the LIF properties to compress images already filtered by the OPL. Unlike the already existing static encoding algorithms, this quantization scheme encodes images in a dynamic way and then using an inverse function the encoded information provides an estimation of the initial image.

2. THE LIF

2.1. Background

As described in [2], the LIF is a neural model which is described by the relation:

$$\tau_m \frac{du}{dt} = -u(t) + RI(t). \tag{1}$$

where $u(t)$ represents the membrane potential at time t, τ_m is the membrane time constant and R is the membrane resistance. In the LIF model, spikes are generated at a firing time $t^{(f)}$ defined by the following threshold criterion:

$$t^{(f)} : u(t^{(f)}) = \theta. \tag{2}$$

Immediately after $t^{(f)}$ the potential is set to a new value $u_r < \theta$,

$$\lim_{t \to t^{(f)}; t > t^{(f)}} u(t) = u_r. \tag{3}$$

While $t < t^{(f)}$ the dynamics is given by eq. (1) until the next threshold crossing occurs. The LIF neuron may also incorporate a refractory period. In this case, if u reaches the threshold at time $t = t^{(f)}$, the dynamics is interrupted during an absolute refractory time Δ^{abs} and the integration restarts at time $t^{(f)} + \Delta^{abs}$ with a new initial condition.

Let's consider the simple case of a constant input current stimulus $I(t) = I_0$. For the sake of simplicity we will assume
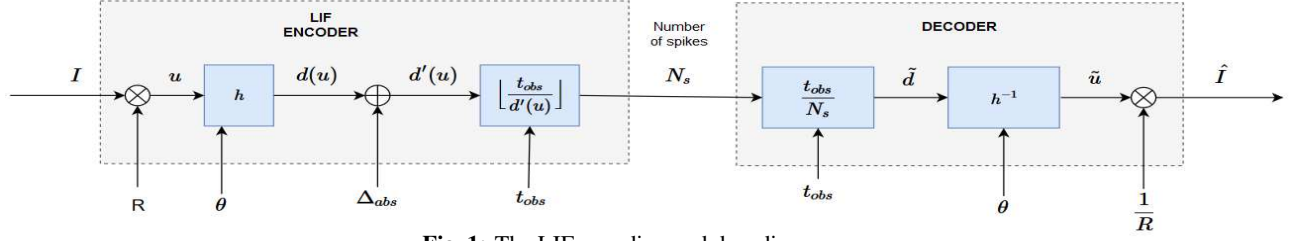
Fig. 1: The LIF encoding and decoding process

a reset potential $u_r = 0$. Assuming that the kth spike has occurred at time $t = t^k$ when the trajectory of the membrane potential is given by integrating eq. (1) with the initial condition $u(t) = u_r = 0$. The solution is given by the relation:

$$u_k(t) = RI_0 \left[1 - exp\left(-\frac{t - t^k}{\tau_m}\right) \right]. \qquad (4)$$

After each spike, the potential is reset to the value $u_r = 0$ and the integration process starts again. The condition $u(t) = \theta$ is satisfied for $t = t^{k+1}$, where t^{k+1} denotes the time when the next spike occurs. Then, eq. (2) can be written as following:

$$u(t^{k+1}) = \theta = RI_0 \left[1 - \exp\left(-\frac{t^{k+1} - t^k}{\tau_m}\right) \right]. \qquad (5)$$

We assume $d(RI_0) = t^{k+1} - t^k$, the inter-spike delay of an integrate-and-fire neuron with no refractory period, which depends on the input current I_0. Consequently, solving (5) for the delay $d(RI_0)$ and simplifying the notation setting $u = u(I_0) = RI_0$ yields:

$$d(u) = \begin{cases} \infty, & u < \theta \\ \\ h(u;\theta) = \tau_m \ln(\frac{u}{u - \theta}), & u \geq \theta, \end{cases} \qquad (6)$$

The firing rate of the LIF neuron, is then given by the relation $\nu = 1/d(u)$.

At this point, it is important to denote that for the case of a neuron with an absolute refractory period, the occurrence of the next spike will be delayed by the duration of the refractory period Δ^{abs}. So, in this case, the inter-spike delay $d'(u)$ is given by:

$$d'(u) = d(u) + \Delta^{abs} = t^{k+1} - t^k + \Delta^{abs}, \qquad (7)$$

where $t = t^{k+1} - t^k + \Delta^{abs}$ is the time instance when the next integration will start after the emission of the $(k + 1)^{th}$ spike.

2.2. The LIF Quantizer

The LIF quantizer, which has analytically been studied and explained in [3], uses the LIF properties to encode input intensities into numbers of spikes within an observation time t_obs. In the encoder, according to Ohm's law, we compute the action potential of the LIF neuron from the relation $u = RI$. Then, using eq. 6, we compute the integration delay $d(u)$ and by adding the refractory period Δ_{abs} we get the interspike delay $d'(u)$. Finally, the floor value of the division of the parameter t_{obs} by the interspike delay $d'(u)$ gives the number of spikes.

In the decoder, following the inverse procedure, we can get an approximation of the delay by dividing the parameter t_{obs} by the number of spikes produced by the encoder. Using this decoded delay and thanks to the inverse function h^{-1} we get an approximation of the action potential $\hat{u}$ and dividing by R we finally get the quantized values of the input current $\hat{I}$ The process of the quantization using the LIF is shown in Figure 1.

There were some previous attempts to use similar bio-inspired coding schemes [4, 5]. However, both these works tried to approximate the LIF model by conventional compression tools. The originality of this work lies in the implementation of a real LIF quantizer with a refractory period.

3. THE OUTER PLEXIFORM LAYER (OPL)

3.1. Background

The Outer Plexiform Layer (OPL) is the first layer of the retina which consists of a dense network of synapses between the photoreceptor, horizontal and bipolar cells.

The input of the OPL cells is the visual stimulus $f(x, t), x \in \mathbb{R}^2, t \in \mathbb{R}^+$ which is spatiotemporally transformed into an electrical signal. This signal is dynamically encoded by the neurons into a code of spikes and is then propagated to the visual cortex. Our goal is to build a bio-inspired encoding/decoding system for images. Figure 2 shows the proposed architecture.

The dynamic OPL transformation has been approximated by the the retina-inspired filter [1, 6]. The retina-inspired filter $\phi(x, t)$ is a novel Weighted Difference of Gaussian (WDoG) [6] which models the center-surround structure of the receptive field of the bipolar cells:

$$\phi(x, t) = a(t)G_{\sigma_C}(x) - b(t)G_{\sigma_S}(x), \qquad (8)$$

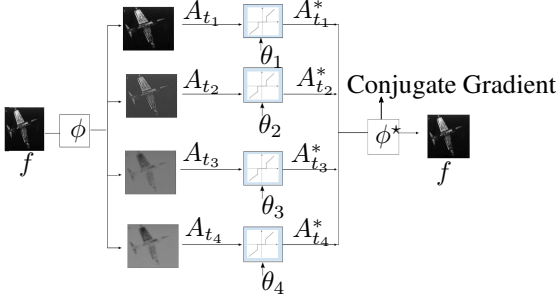where $a(t)$ and $b(t)$ are two time-varying weights which tune

Fig. 2: Retina-inspired image codec. The figure describes the decomposition of the input image f using the retina-inspired filter ϕ into several layers A_{t_j} each one of which is quantized by the LIF quantizer Q. Based on the quantized layers $A_{t_j}^*$, one is able to reconstruct $\tilde{f}$ and compute the distortion comparing to f.

Fig. 3: Subband generation rate. Case 1, corresponds to a subband generation rate with dense middle bandpass frames. Case 2, corresponds to a subband generation rate with sparse middle bandpass frames.

the shape of the DoG, σ_c and σ_s are the standard deviations of the center and the surround Gaussians respectively with $\sigma_c < \sigma_s$.

The retina-inspired filtering, which is a frame, is applied to temporally constant input signals $f(x, t) = f(x)\mathbf{1}_{[0 \leq t \leq T]}(t)$ resulting in:

$$A(x, t) = \phi(x, t) \overset{x}{*} f(x), \tag{9}$$

where $\overset{x}{*}$ is a spatial convolution. Let $t_1, \ldots, t_m$ some temporal samples. For each time instant $t_j, j = 1, \ldots, m$ there is a different decomposition layer $A_{t_j} = A(x, t_j)$. This redundancy is sufficient to perfectly reconstruct the input signal $\tilde{f}$. It is proven in [1] that the retina-inspired filter is a frame hence, the filter is invertible meaning that it is possible to reconstruct the input image.

4. EXPERIMENTS

4.1. Results on one subband

Lets assume $x_1, \ldots, x_n$ some spatial samples such that $A_{t_j} = \big(A(x_1, t_j), \ldots, A(x_n, t_j)\big), j = 1, \ldots, m$ a discrete decomposition layer. The LIF quantizer is applied to every single spatiotemporal sample $A(x_k, t_j)$ where $k = 1, \ldots, n$. For the experiment, we have chosen grayscale images of the size $n = 512 \times 512$ pixels taken from USC-SIPI database [7]. In our tests, we are going to use a LIF neuron with a random refractory period which follows a half-Gaussian distribution.

In [3], it is proven that the presence of a refractory period introduces overload noise to the LIF quantizer. This yields the existence of an optimal θ value which minimizes the Mean Squared Error (MSE). In this work we select this optimal value for each of the OPL subbands.

4.2. Subband Generation using the OPL

The purpose of this paper is to experiment on the application of the LIF quantizer on each of the subbands produced by
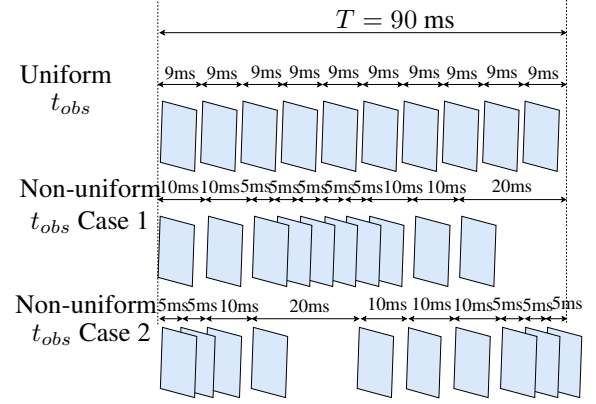
the retina-inspired filter and evaluate the quality and the efficiency of the extended system depicted in Figure 2. In [6] it has been shown that the amount of information on the subbands produced by the OPL decomposition varies while time evolves. More specifically, according to the bio-plausible filtering parameters given in [6], in the very first subbands the range of the intensity values is very small while in the last subbands (i.e. $t \geq 120ms$) there is no big change in the subbands' content. Consequently, in order to reduce the redundancy of the latest subbands, we are going to generate 10 subbands in the range $0 \leq t \leq 90ms$. As a first step, we tested the generation of 10 subbands uniformly distributed in the total filtering range, observing each of the produced layers for a $t_{obs} = 9ms$ as described in Figure 3. Moving on, we experimented on the non-uniform case, trying the two different non-uniform schemes shown in Figure 3. The first one, corresponds to an attempt to keep most of the middle and most informative subbands in the bandpass range $25ms \leq t \leq 50ms$. In this case, although we keep most of the middle informative subbands, we observe each layer for a shorter observation time t_{obs}. Then we also tried the second non-uniform scheme, depicted in Figure 3, which corresponds to a subband generation with sparser layers in the bandpass range of observation times. In this second case, while we keep less of the informative subbands, they are better encoded, as we observe them for a longer observation time. At this point, we should mention the fact that this is only a first experimental attempt to apply the LIF to the layers produced by the OPL filter, in order to evaluate and better understand the properties of our proposed encoder. As a result, our subband selection for the non-uniform sampling cases has been experimentally achieved, without using some specific function.

For our experiments, we have used two different images. After the subband generation we apply the LIF quantization on each of the generated subbands, we reconstruct the en-

Fig. 4: Comparison of visual results and quality metrics of PSNR, SSIM and Entropy for the original image (first on the left), the uniform subband generation(second from left to right), the non-Uniform subband generation with a sparser middle (second from left to right), and the non-uniform scheme with a denser middle sumbbands (first on the right)

coded layers and evaluate the quality of the reconstructed image compared to the original one as described in Figure 2. In Figure 4 we present the visual results of our experiments showing also the values of the Entropy, the Peak Signal to Noise Ratio (PSNR), and the Structural Similarity Index (SSIM)[8].

We observe that for the first image, the non-Uniform subband generation with the denser subbands in the bandpass area provides a better value of PSNR and SSIM compared to the uniform case, while the entropy is being slightly reduced. On the contrary, the nonuniform generation with sparser subbands in the middle observation times behaves poorly in comparison to the uniform generation. For the second image though, depicted in the lower part of Figure 4, we observe that both non-uniform cases of subband generation provides better results of PSNR and SSIM than the uniform case, with the denser middle subband generation behaving slightly better than the sparser middle subbands case.

Consequently, we can assume that the selection of the good θ value according to the observation time as well as the good rate of subband generation in the OPL filtering, can provide very promising results and significantly improve the rate-distortion trade off. In addition to this, we conclude that the good rate of subband generation varies according to the

image characteristics (statistics, content).

4.3. Conclusions

In this work we have implemented an extended retina-inspired compression system. This is an innovative approach which uses a dynamic way of quantization adapted to the needs of the encoding process unlike the existing encoding algorithms. Our study, reveals the fact that this bio-inspired dynamic encoding process can provide very promising results. The good choice of layers produced by the OPL filter, plays an important role to the quality of the image reconstruction and gives a strong motive to further study the behavior of the model according to the total observation time. Since this is a very first attempt to apply this extended encoding system on images, we should underline the significance of improving these results by further experimenting and studying the system's behavior. Furthermore, since in our experiments we used an experimental way of non-uniform subband generation, the use of a particular function that will be able to minimize the rate distortion trade-off according to the image characteristics, is a very important future step that should be studied and implemented.

5. REFERENCES

[1] E. Doutsi, L. Fillatre, M. Antonini, and J. Gaulmin, "Retina-inspired filtering for dynamic image coding," *IEEE International Conference in Image Processing (ICIP)*, pp. 3505–3509, 2015.

[2] W. Gerstner and W. Kistler, *Spiking neuron models: Single Neurons, Populations, Plasitcity*, Cambridge University Press, 2002.

[3] M. Dimopoulou and M. Antonini, "Signal Quantization using a Leaky Integrate-and-Fire neuron," in *GRETSI*, 2017.

[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] E. Doutsi, L. Fillatre, M. Antonini, and J. Gaulmin, "Bio-inspired Sparse Representation of Images," in *Gretsi*, 2017, number 1, pp. 2–5.

[6] E. Doutsi, L. Fillatre, M. Antonini, and J. Gaulmin, "Retina-inspired Filtering," *hal-01350686*, 2016.

[7] A. Weber, "The USC-SIPI Image Database," 1977.

[8] A Horé and D Ziou, "Image quality metrics: PSNR vs. SSIM," *20th International Conference on Pattern Recognition, ICPR*, pp. 2366–2369, 2010.