Image coding using Leaky Integrate-and-Fire neurons

Melpomeni Dimopoulou, Effrosyni Doutsi, Marc Antonini

Image coding using Leaky Integrate-and-Fire neurons

Melpomeni Dimopoulou, Effrosyni Doutsi, Marc Antonini
Université Côte d'Azur, CNRS, I3S, France

mel.dimopoulou@gmail.com
doutsi@i3s.unice.fr
am@i3s.unice.fr

Résumé

This paper aims to build an image coding system based on the 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 the 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 according to which function the ganglion cells of the Ganglionic retinal layer that is responsible for the 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 previous 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.

Mots clefs

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 high 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 visual system and more specifically, by the way the retina works for the perception and compression of natural images. The retina can be divided into three basic layers. The Outer Plexiform Layer (OPL) which acts as a spatiotemporal filter on images, the Inner Plexiform Layer (IPL) that performs a non-linear rectification, and the Ganglionic Layer which is responsible for the encoding of the data. The Ganglionic layer consists of the ganglion cells, a type of neuron which compresses visual information, in response to the brightness of light. The Ganglion cells function 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 be applied on the compression of images already filtered by the OPL. This quantization scheme, unlike the already existing static encoding algorithms, encodes images in a dynamic way and then using an inverse function the encoded information provides an estimation of the initial image.

In section 2, we are going to provide the theoretical background for the LIF, explaining the physical and biological function of the LIF neural model. We also describe the LIF encoding and decoding process performed by the LIF quantizer. In section 3, we discuss about the OPL and the way it acts as a spatiotemporal filter on input images producing decomposition layers. Furthermore we analyze the procedure for the reconstruction. Finally, in section 4, we present our experiments on the extended encoding system for the case of uniform selection of OPL subbands and a non-uniform one which emphasizes on the most informative layers.

2 The LIF

2.1 Background

As described in [2], the LIF is a neural model which is described by the circuit shown in Figure 1. The input current $I(t), t \in \mathbb{R}^+$ is being divided in the current I_R, which passes through the resistor and the current I_C which charges the capacitor. Given the Ohm's law for I_R and the definition of capacity as $C = q/u$ (where q is the charge and u the voltage) the total current can be written as :

$$I(t) = I_R + I_C = \frac{u(t)}{R} + C\frac{du}{dt}. \quad (1)$$

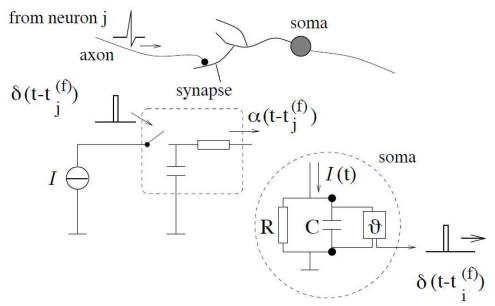

Figure 1 – *The LIF neuron circuit which consists of a resistance R in parallel with a capacitor C (Figure taken from [2]).*

By multiplying eq. (1) by R and by introducing a time constant $\tau_m = RC$ the equation becomes :

$$\tau_m \frac{du}{dt} = -u(t) + RI(t). \tag{2}$$

In the integrate-and-fire model, the form of an action potential is not described explicitly. Spikes are generated at a firing time $t^{(f)}$. This firing time is defined by the following threshold criterion :

$$t^{(f)} : u(t^{(f)}) = \theta. \tag{3}$$

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

$$\lim_{t \to t^{(f)}; t > t^{(f)}} u(t) = u_r. \tag{4}$$

While $t < t^{(f)}$ the dynamics is given by eq. (2) 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 a reset potential $u_r = 0$. Assuming that the k^{th} spike has occurred at time $t = t^k$ when the trajectory of the membrane potential is given by integrating eq. (2) with the initial condition $u(t) = u_r = 0$. The solution is given by the relation :

$$u_k(t) = RI_0 \left[1 - exp\left(-\frac{t - t^k}{\tau_m}\right)\right]. \tag{5}$$

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. (3) can be written as following :

$$u(t^{k+1}) = \theta = RI_0 \left[1 - \exp\left(-\frac{t^{k+1} - t^k}{\tau_m}\right)\right]. \tag{6}$$

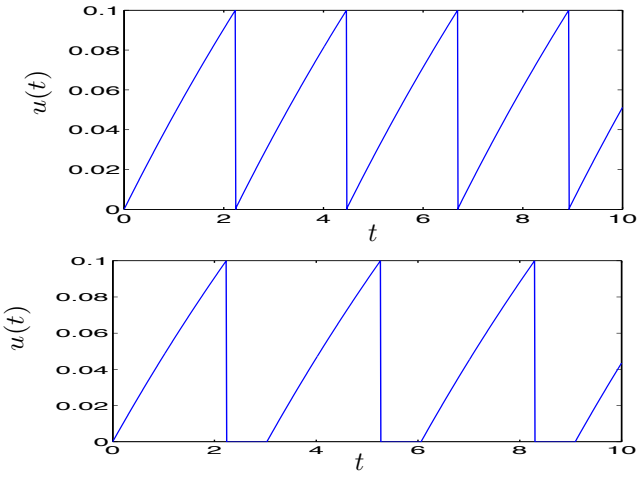

Figure 2 – *The LIF firing process - Top : model with no refractory period. Bottom : model with a constant refractory period $\Delta^{abs} = 2$ ms. In the model without a refractory period the next integration starts exactly after the spike emission while in the presence of a constant refractory period the new integration begins delayed by a time period equal to Δ^{abs} after the latest spike emission time.*

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. For example, the higher the amplitude of the current is, the smaller the delay.
Consequently, solving (6) for the delay $d(RI_0)$ and simplifying the notation setting $u = u(I_0) = RI_0$ yields :

$$d(u) = \begin{cases} \infty, & u < \theta \\ \\ h(u;\theta) = \tau_m \ln(\dfrac{u}{u - \theta}), & u \geq \theta, \end{cases} \tag{7}$$

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 :

$$d'(u) = d(u) + \Delta^{abs} = t^{k+1} - t^k + \Delta^{abs}, \tag{8}$$

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. The firing process of a LIF neuron is described by Figure 2.

2.2 The LIF Quantizer

The LIF quantizer, which has analytically been studied and explained in [3], uses the LIF properties described in the previous section to encode input intensities into numbers of spikes. More specifically, the LIF Quantizer works according to the following procedure. In the encoder, according to Ohm's law, we compute the action potential of the

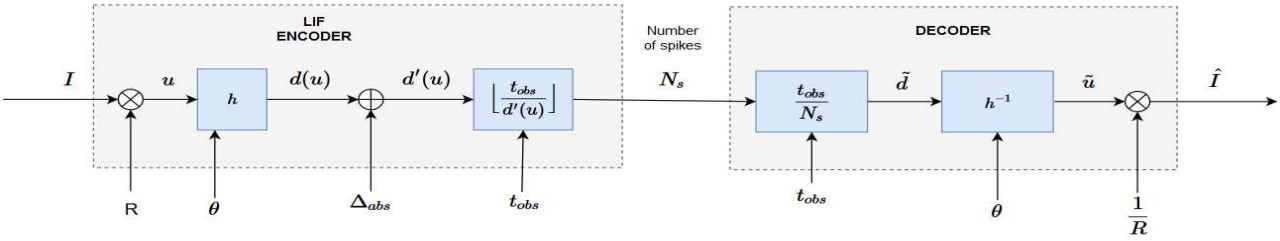

Figure 3 – *The LIF encoding and decoding process*

LIF neuron from the relation $u = RI$. Then using eq. 7 we compute the integration delay $d(u)$. By adding the refractory period we get the interspike delay $d'(u)$.
In the decoder, the floor value of the division of the parameter t_{obs} by the interspike delay $d'(u)$ gives the number of spikes. Following the inverse procedure, dividing the parameter t_{obs} by the encoded number of spikes and using the estimated delay for the computation of the membrane potential estimation, we finally get the quantized values of the input current $\hat{I}$ by dividing by the resistance R. The process of the quantization using the LIF is shown in Figure 3.
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 with respect to the other ones is related to the implementation of a real LIF quantizer with refractory period.

2.3 Refractory period

As described in section 2.1, after the production of a spike in the LIF model there can exist a refractory period. For our Quantizer this refractory period works as an additive noise. In this paper we are going to use a half-Gaussian refractory period. This means that after each spike we generate a random positive refractory period of a specified variance. More specifically the refractory period of our model behaves according to the following distribution :

$$P(x) = A \exp\left(-B(x^2)\right), \qquad (9)$$

where $A = \dfrac{\sqrt{2}}{\sigma\sqrt{\pi}}$ and $B = \dfrac{1}{2\sigma^2}$ for a given standard deviation σ.

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 (see Fig. 4) [6]. 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 the equivalent electrical signal. This electrical signal is dynamically encoded by the neurons into a code of spikes which is propagated to the analysis center; the visual cortex. The neural code is not used to reconstruct

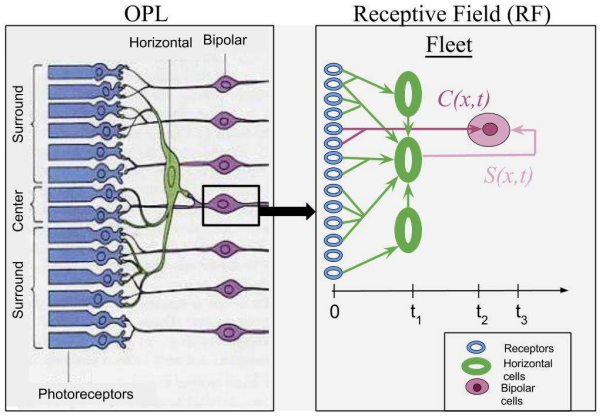

Figure 4 – *The OPL structure (Figure taken from [7]).*

the input signal but to learn and take decisions. However, our goal is to build a bio-inspired coding/decoding system for images. Figure 5 shows the proposed coding/decoding architecture we aim to build.

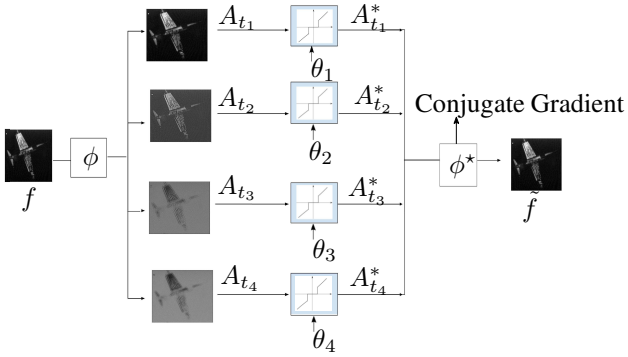

Figure 5 – *Retina-inspired image codec. The figure describes the decomposition of the input image f using the retina-inspired filter ϕ into several layers A_{t_j} each one of which is sparsified by the LIF quantizer Q. Based on the quantized layers $A^*_{t_j}$, one is able to reconstruct $\tilde{f}$ and compute the distortion comparing to f.*

The dynamic OPL transformation which is caused to the visual stimulus while it is captured and propagated through the OPL cells has been approximated by the the retina-

inspired filter [1, 7]. The retina-inspired filter $\phi(x,t)$ is a novel Weighted Difference of Gaussian (WDoG) [7] which models the center-surround structure of the receptive field of the bipolar cells :

$$\phi(x,t) = a(t)G_{\sigma_C}(x) - b(t)G_{\sigma_S}(x), \tag{10}$$

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

$$A(x,t) = \phi(x,t) \overset{x}{*} f(x), \tag{11}$$

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 $\tilde{f}$.

3.2 Reconstruction

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. In practice, one needs to solve the linear system $A = \Phi f$ and reconstruct $\tilde{f}$. At time $t = T$, the exact estimation of $\tilde{f} = f$ according to :

$$\tilde{f} = (\Phi^\top \Phi)^{-1} \Phi^\top A, \tag{12}$$

where Φ^{-1} denotes the inverse of a matrix Φ and $\Phi^\top$ denotes its transpose. The retina-inspired filter Φ is a frame, as a result, we can define as $\Phi^T\Phi$ its frame operator. However, in practice, due to the large size of matrix Φ and in order to avoid a time consuming and resource demanding reconstruction processing, we used the conjugate gradient descent which is one of the most efficient iterative methods [8]. We are interested in reducing this redundancy and discard all the coefficients of low energy keeping only the most informative ones for the reconstruction.

4 Experiments

4.1 Results on one subband

In our experiments we first tested the LIF quantizer on a single subband, in order to understand and evaluate the quantizer's behavior. Let $x_1, \ldots, x_n$ 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 [9]. As described in section 2.3, in our tests, we are going to use a random refractory period which follows a half-Gaussian distribution.

Figure 6, shows the evolution of the Mean Squared Error (MSE) between the original image and the decoded one using the LIF, in function of the value of θ for different values of observation time t_{obs}. It is clear that, the refractory period introduces overload noise on the input image which yields the existence of an optimal threshold value θ that minimizes the MSE. This optimal θ value is different according to the value of the observation time t_{obs}.

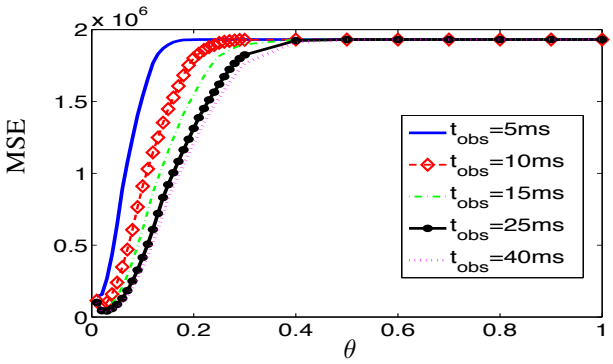

Figure 6 – *The MSE curve in function of the threshold parameter θ for different observation times.*

In order for our quantizer to be adaptive to the needs of the quantization process, we are going to select for each realisation the appropriate value of theta for the LIF.

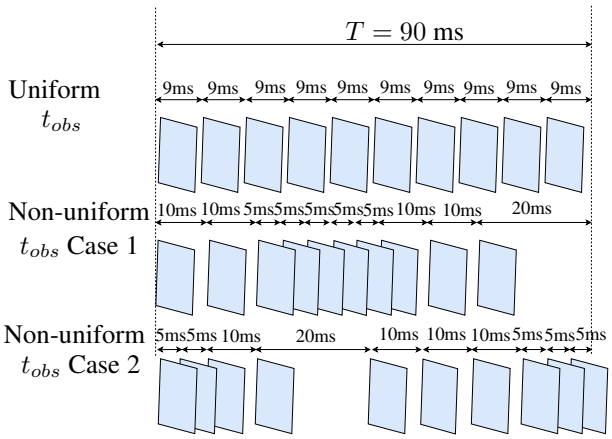

Figure 7 – *Subband generation rate. Case 1, corresponds to a subband generation rate with dense middle bandpass frames, corresponds to a subband generation rate with sparse middle bandpass frames.*

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 5. Extended studies in [7] have shown that the amount of information on the subbands produced by the OPL decomposition varies while time evolves. More specifically, according

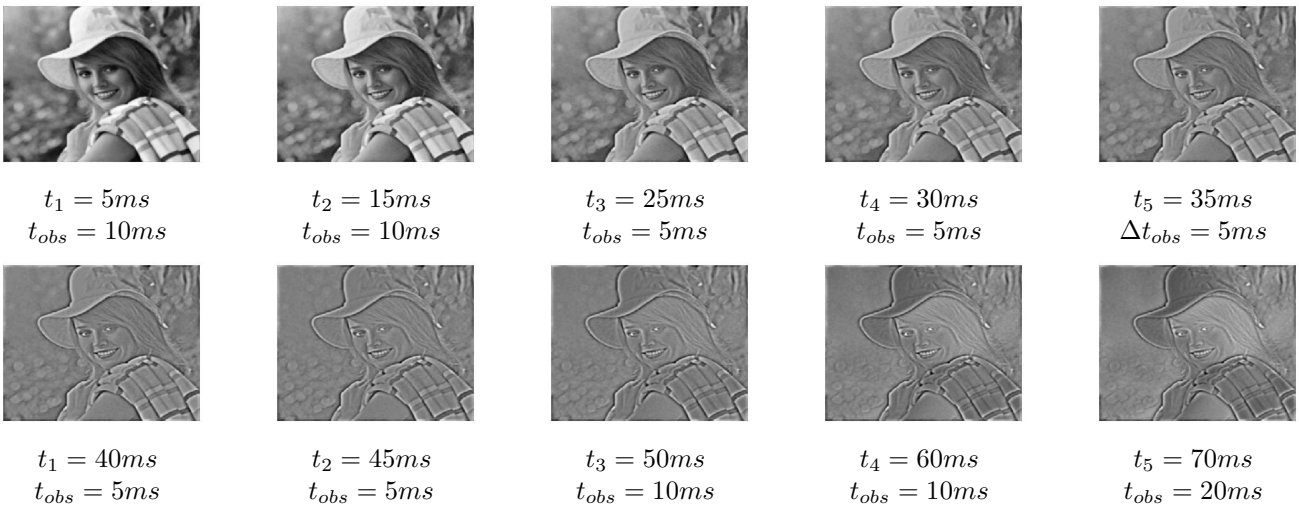

Figure 8 – *The non-uniform subband generation using the OPL filtering and the corresponding time of appearance* t_1, t_2, t_3, t_4, t_5 *and observation time* t_{obs} *respectively.*

to the bio-plausible filtering parameters given in [7], 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 achieve a sparse reconstruction and reduce the redundancy of the latest subbands, in our experiments 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 7.

Moving on, we experimented on the non-uniform case, trying the two different non-uniform schemes shown in Figure 7. The first one, corresponds to an attempt to keep most of the middle and most informative subbands in the bandpass range $25ms \leq t \leq 50ms$. The subbands produced by this scheme are visually presented in Figure 8. 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 7, 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 encoded layers and evaluate the quality of the reconstructed image compared to the original one as described in Figure 5. In Figure 9 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)[10].

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 9, 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

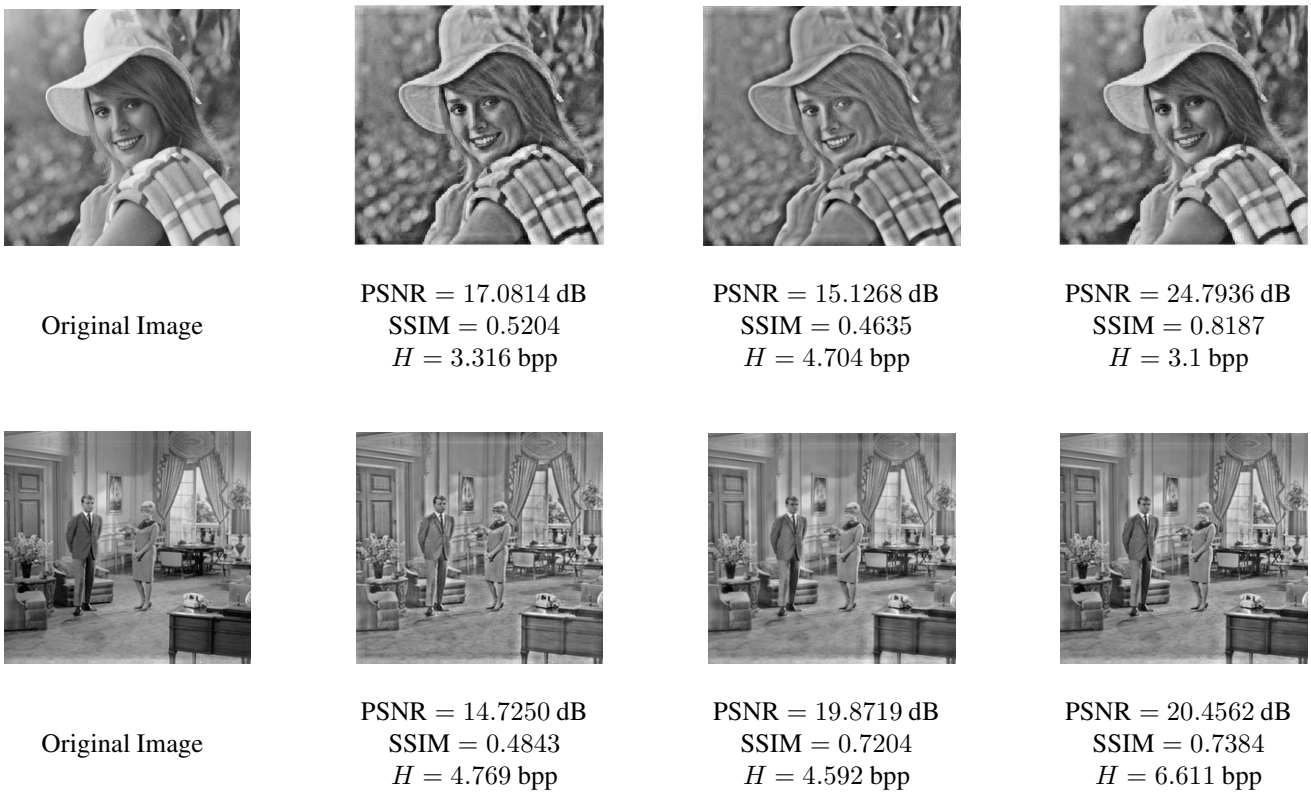

Figure 9 – *The comparison of the visual results and quality metrics of the PSNR, SSIM and Entropy for the a)original image (first image on the left) b)the uniform subband generation(second image from left to right) c) the non-Uniform subband generation with a sparser middle (second image from left to right) and d)the non-uniform scheme with a denser middle sumbbands (first image on the right)*

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.

Références

[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] H. Kolb, "How the Retina Works," *American Scientist the magazine of Sigma Xi, The Scientific Research Society*, vol. 91, pp. 28–35, 2004.

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

[8] J R Shewchuk, "An Introduction to the Conjugate Gradient Method Without the Agonizing Pain," *Science*, vol. 49, no. CS-94-125, pp. 64, 1994.

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

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