



HAL
open science

NEURO-INSPIRED QUANTIZATION

Effrosyni Doutsis, Lionel Fillatre, Marc Antonini, Julien Gaulmin

► **To cite this version:**

Effrosyni Doutsis, Lionel Fillatre, Marc Antonini, Julien Gaulmin. NEURO-INSPIRED QUANTIZATION. 25th IEEE International Conference on Image Processing (ICIP 2018), Oct 2018, Athènes, Greece. pp.689-693, 10.1109/ICIP.2018.8451793 . hal-02144872

HAL Id: hal-02144872

<https://hal.science/hal-02144872>

Submitted on 8 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.

NEURO-INSPIRED QUANTIZATION

Effrosyni Doutsis^{1,2}, Lionel Fillatre¹, Marc Antonini¹ and Julien Gaulmin²

¹Université Côte d’Azur, I3S, CNRS, UMR 7271, 06900, Sophia Antipolis, France.

²4G-TECHNOLOGY, 460 avenue de la Quiera 06370, Mouans Sartoux, France.

ABSTRACT

This paper presents a novel neuro-inspired quantization model which is the extension of the recently released perfect-Leaky Integrate and Fire (LIF) model. We propose that the LIF, which is a very efficient neuromathematical model that describes the spike generation neural mechanism, can lead to a groundbreaking and above all dynamic compression algorithm which is called LIF encoder/decoder. We also prove that under some assumptions, there is a link between the novel LIF encoder/decoder and the conventional Uniform Deadzone Quantizer (UDQ).

Index Terms— Visual system, Neuron, Leaky Integrate and Fire (LIF) model, Uniform Deadzone Quantization (UDQ)

1. INTRODUCTION

During the last decades, there have been a great progress on the decryption of the visual system. In the literature, a variety of models approximate the connectivity between neurons in order to communicate with each other. This communication happens with a very efficient and compact code which is generated in time through the visual pathway. There have been proposed several neuromathematical models which approximate the spike generation mechanism [1, 2, 3, 4, 5, 6]. Although each model is based on different assumptions, they all remain faithful to the dynamic properties of the neural processing.

In this paper, we proposed that mimicking the neural encoding process will be groundbreaking and useful in image processing society for many reasons. First of all, a neuro-inspired coder will enable to deal with a very high definition input taking for granted that the retinal sensors of human eyes can receive a very high definition signal $\approx 10^9$ bits per second [7] which is equivalent to the bitrate of a raw color video of HD resolution at 20Hz. In addition, through the visual pathway the transmission capacity is dramatically reduced leading to a very tightly packed code. Last but not least, the spike generation mechanism is dynamic, meaning that the neurons are able to process and encode a spatiotemporal input signal with respect to time.

There have been several attempts to build a bio-inspired codec in the past like the Rank Order Coder [8, 5, 9] or the Dithered Scalable Image Coder [10, 11]. However, both of them had roughly approximate the neural spiking generation mechanism. We have recently introduced the perfect-LIF method [12] which enables the sparse representation of an image according to the performance of the Leaky Integrate and Fire (LIF) model; a very well-known neural spiking mechanism which is based on the assumption that each neuron spikes individually [6]. The perfect-LIF is a simple thresholding function which is insufficient in terms of compression.

In this paper, we aim to extend the performance of the perfect-LIF and use spikes to encode a still-image. We propose that counting and encoding the number of spikes N , which are released within an observation window T , leads to an alternative and above all dynamic compression. In addition, we prove that under the assumption that the visual stimulus remains constant in time, the LIF encoder/decoder, which is also called neuro-inspired quantizer, can be approximated by a uniform deadzone scalar quantizer. Sections 2 and 3 are a recall of the LIF model and the Perfect-LIF function. Section 4 introduces the LIF encoder/decoder while the experimental results are given in Section 5. Some concluding remarks are given in section 6.

2. LIF MODEL

According to [6], the LIF approximates the neural encoding process by an electrical circuit:

$$I(t) = I_R(t) + I_C(t) = \frac{V(t)}{R} + C \frac{dV}{dt}, \quad (1)$$

where $I(t)$ is the input current which is divided into the current $I_R(t)$ which runs the resistor and $I_C(t)$ which charges the capacitor, C is the membrane capacitor of a neuron which is in parallel with the resistor R and $V(t)$ is the voltage across the resistor. We assume that at time $t = 0$, when we inject the current $I(t)$, the membrane potential of the neuron is $V(t = 0) = 0$ mV.

If we multiply eq. (1) by R , we introduce the time constant $\tau_m = RC$ of the “leaky integrator”. This yields the standard form:

$$\tau_m \frac{dV}{dt} = -V(t) + RI(t). \quad (2)$$

Thanks to 4G-TECHNOLOGY and ANRT for funding.

Whenever, the membrane potential of a neuron crosses the threshold θ , where $\theta > 0$, the neuron spikes. The moment the neuron spikes is called firing time $t = t^f$. For a given threshold θ a neuron spikes according to the following threshold criterion:

$$t^f : V(t^f) = \theta. \quad (3)$$

Immediately after the emission of a spike the potential is reset to $V_r' < \theta$ which is the new initial condition. In addition, the neuron will remain silent for an absolute refractory period Δ^{abs} and the integration starts again at time $t^f + \Delta^{\text{abs}}$. Since spikes are stereotyped events, i.e. with nearly identical shapes, they are fully characterized by their firing time.

2.1. Constant input

Assumption 1. *Let's assume that a neuron is stimulated by a constant current I which is switched on during some time T as following:*

$$I(t) = I \mathbf{1}_{[0 \leq t \leq T]}(t), \quad (4)$$

where $I \in \mathbb{R}$ the input intensity and $\mathbf{1}$ is the indicator function which equals 1 if $0 \leq t \leq T$, and 0 otherwise.

It has been shown in [6] that under the Assumption 1 the exact delay each spike arrives is given by:

$$d = \begin{cases} +\infty, & \text{if } RI < \theta, \\ h(RI; \theta) = -\tau_m \ln \left[1 - \frac{\theta}{RI} \right], & \text{if } RI > \theta. \end{cases} \quad (5)$$

3. PERFECT-LIF

In [12], we introduced that if the delay d of the first spike is a priori knowledge for the decoder, using the $h^{-1}(d; \theta)$ function it is possible to perfectly reconstruct all the input intensity \tilde{I} which are above the threshold:

$$\tilde{I} = \begin{cases} 0, & \text{if } d = +\infty \\ I = h^{-1}(d; \theta), & \text{if } d < +\infty, \end{cases} \quad (6)$$

where

$$h^{-1}(d; \theta) = \frac{\theta}{R \left(1 - \exp \left(-\frac{d}{\tau_m} \right) \right)}. \quad (7)$$

In addition, we have proven in [12, 13] that based on the Assumption 1 of the temporally constant input, we impose a temporal constraint which neglects all the spikes which are emitted with delay $d > T$. This constraint results in a new threshold λ given by:

$$\lambda = h^{-1}(T; \theta) = \frac{\theta}{R \left(1 - \exp \left(-\frac{T}{\tau_m} \right) \right)}. \quad (8)$$

Finally, the reconstruction of the time-constrained perfect-LIF encoder/decoder is given by:

$$\tilde{I} = \begin{cases} 0, & \text{if } d > T \\ I = h^{-1}(d; \theta) & \text{if } d \leq T. \end{cases} \quad (9)$$

The basic drawback of the above perfect-LIF function is that for all the intensities $I > \lambda$, it requires to code the firing times which is a finite set of real values. This will dramatically increase the memory cost of the algorithm and the number of bits which are necessary to store the input signal. However, due to the Assumption 1, we ensure that the interspike interval for each input intensity I is constant. To this extent, in section 4, we propose that it is more efficient to count and encode the number of spikes N which have been emitted within the observation window T .

4. LIF ENCODER/DECODER

In this paper, we are interested in building a neuro-based quantizer which imitates better the neural encoding process than the perfect-LIF [12]. As we show later on, the LIF encoder/decoder is able to dynamically compress a temporally constant input signal into spikes which is novel and advantageous with regards to the static state-of-the-art quantization models.

4.1. LIF Encoder: Counting the Number of Spikes

Lets suppose that one “listens” to the response of a neuron after a very large observation window T . Then, each intensity will produce a spike train of a high density. The density is described by the number of spikes N which is given by:

$$N = \left\lfloor \frac{T}{d} \right\rfloor, \quad (10)$$

The number of spikes depends on three parameters: the intensity of the input signal I , the value of threshold λ and the observation time T . For given λ and T the higher the intensity I , the smaller the delay d . Moreover, the higher the value of λ , the less the number of the neurons which spike. Last but not least, the longer the observation time T , the more the spikes each neuron emits.

In this section, we explain that counting and encoding the number of spikes, N , which corresponds to each intensity, instead of the delays d , results in the reduction of the memory cost. This is obvious due to the fact that the number of spikes belongs to a finite set of natural numbers whereas the delays to a finite set of real numbers. Of course, we need to pay the price of precision because by counting the number of spikes, we will not be able to perfectly reconstruct the intensities anymore.

Figure 1 illustrates how the LIF-based encoder/decoder allows to do the mapping and assign a single output intensity to a group of input intensities by using the number of spikes.

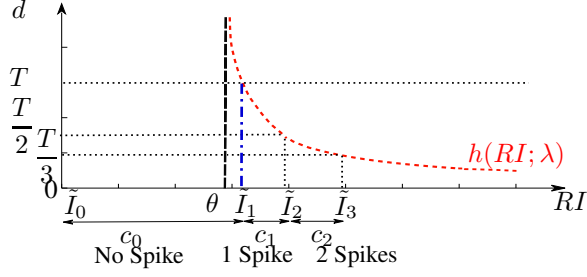


Fig. 1: LIF Decoder counting the number of spikes. The model provides different clusters which encode a group of input coefficients.

For any input intensity which is below the threshold $I < \lambda$, there will be emitted no spikes ($N = 0$), then \tilde{d} is:

$$\tilde{d} = \begin{cases} \infty, & \text{if } N = 0 \\ \frac{T}{N} & \text{if } N > 0. \end{cases} \quad (11)$$

Consequently, all the input values which belong to group $c_0 = \{I \mid \tilde{d} = +\infty\}$ will be recovered by a single output intensity $\tilde{I}_0 = 0$. Let us now suppose that only one spike arrives for the input signal, $N = 1$. According to eq.(11), when the first spike arrives the decoded delay equals $\tilde{d} = T$. All the input intensities which have caused the generation of a single spike will be reconstructed by the value \tilde{I}_1 (see eq. (13)). These input intensities belong to the group $c_1 = \left\{I \mid \frac{T}{2} < \tilde{d} \leq T\right\}$ because at time $\tilde{d} = \frac{T}{2}$ the second spike arrives. Then, all the input values which belong to the interval between $\frac{T}{2}$ and $\frac{T}{3}$ will be assigned to the group c_2 , etc. Finally, one could define the groups/clusters as following:

$$\begin{aligned} c_0 &= \left\{RI \mid \tilde{d} > T\right\} \\ c_k &= \left\{RI \mid \frac{T}{k+1} < \tilde{d} \leq \frac{T}{k}\right\}, \quad \forall k \in \mathbb{N}^+. \end{aligned} \quad (12)$$

An interesting remark is that the quality of reconstruction decreases while the number of spikes increases because the higher the number of spikes is, the more the input values RI which are going to be grouped together to the same cluster. Concerning the sign of the input intensity, we have decided in this paper to generate a function $\text{sgn}(I)$ which equals 0 if $I \geq 0$, and 1 otherwise.

4.2. LIF Decoder

If the decoder knows the number of spikes N , it is possible first to approximate the delay \tilde{d} and then, using the inverse function $h^{-1}(\tilde{d}; \theta)$, to reconstruct an approximation of the input intensity \tilde{I} . Based on eq. (11) the delay belongs to the set $\tilde{d} \in \left\{\infty, \frac{T}{1}, \dots, \frac{T}{k}, \dots\right\}$. Consequently, due to eq. (6) the

reconstructed value \tilde{I} in function of the delay is given by:

$$\begin{aligned} \tilde{I}_0 &= 0, & \text{if } \tilde{d} > T \\ \tilde{I}_k &= h^{-1}\left(\frac{T}{k}; \theta\right), & \text{if } \tilde{d} \leq T. \end{aligned} \quad (13)$$

We are able now to define also the length, l , of each cluster which is linked to the reconstructed values of eq. (13):

$$\begin{aligned} l_0 &= \lambda, & \text{for } \tilde{d} > T \\ l_k &= \tilde{I}_{k+1} - \tilde{I}_k & \text{for, } \tilde{d} \leq T. \end{aligned} \quad (14)$$

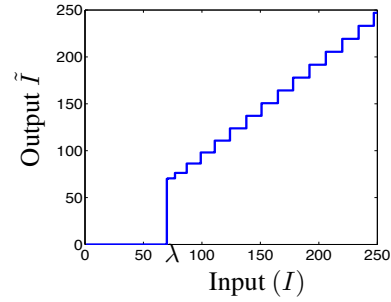


Fig. 2: A typical LIF encoder/decoder characteristic function computed by counting the number of spikes ($\lambda = 70$, $T = 150$ ms, $R=1 \Omega$, $C=1$ F).

The characteristic function of the proposed LIF encoder/decoder is presented in Fig. 2. As illustrated, the intensities $RI < \lambda$ will be discarded while the intensities $RI > \lambda$ are grouped into clusters of length l_k depending on the number of spikes they have emitted. In addition, we have noted that the size of the cluster which are closer to the threshold λ is smaller than the ones which are far from it. As a result, the reconstruction will be more accurate for the intensities around the threshold value.

4.3. Asymptotic Analysis

In this section, we are interested in studying the limiting behavior of the LIF-based encoder/decoder and find the link between the parameters of our novel neuro-inspired quantizer and the ones of conventional quantization methods which is currently used in image and video compression algorithms. As a result, we are going to analyze under which circumstances the LIF-based quantizer could be approximated by a Uniform Deadzone Quantizer (UDQ) which is the state-of-the-art in image compression algorithms [14, 15]:

$$Q_q(I) = \text{sgn}(I) \max\left(0, \left\lfloor \frac{|I| - \lambda}{q} + 1 \right\rfloor\right) \times q, \quad (15)$$

where I is the input value, λ the deadzone, q the quantization step and $\text{sgn}(I)$ a function which describes the sign of the input I .

Proposition 1. Assume $C \ll T \ll R$ and $\frac{T}{\tau_m} \ll 1$ then, the LIF-based encoder/decoder turns into a UDQ with quantization step $q = \frac{\theta C}{T}$.

Proof. The reconstructed value \tilde{I} of an input intensity which has emitted N spikes is computed according to the following function:

$$\tilde{I} = h^{-1}\left(\frac{T}{N}; \theta\right) = \frac{\theta}{R \left(1 - \exp\left(-\frac{T}{\tau_m N}\right)\right)}. \quad (16)$$

Under the assumption the $\frac{T}{\tau_m} \ll 1$ we are able to use the Taylor series to simplify the above function and compute the asymptotic value of \tilde{I} . The general formula for the Taylor series is given as following:

$$\frac{1}{1 - e^{-x}} = \frac{1}{x} + \frac{1}{2} + \varepsilon(x), \quad (17)$$

with $\varepsilon(x) \rightarrow 0$ when $x \rightarrow 0$. Hence, the reconstructed value can be computed by:

$$\tilde{I} = \frac{\theta C}{T} k + \frac{\theta}{2R} \quad \text{with} \quad q \approx \frac{\theta C}{T}, \quad (18)$$

the quantization step. \square

Based on Proposition 1, when C and θ remain constant, it follows that:

$$\lambda \approx \frac{\theta C}{T} = q \quad \text{and} \quad l \approx \frac{\theta C}{T} = q. \quad (19)$$

According to the definition above, when $T \rightarrow \infty$, $q = 0$.

Following the general definition of the UDQ (see eq. (15)) and taking under consideration the initial assumption that $\frac{T}{\tau_m} \ll 1$ and the fact that $\lambda \approx q$, the LIF-based quantizer can be described by the following relation:

$$Q_q(I) = \text{sgn}(I) \max\left(0, \left\lfloor \frac{|I|}{q} \right\rfloor\right) \times q. \quad (20)$$

5. EXPERIMENTAL RESULTS

This section is dedicated to the comparison of our novel LIF encoder/decoder and the conventional mid-thread UDQ which have been explicitly described in section 4.

We first applied the LIF encoder/decoder to an input image of size $n = 256 \times 256$ pixels. For each pixel intensity $I_i, i = 1 \dots, n$ we generate the spike trains. Then, we count the number of spikes N_i which are emitted during the observation window T . We estimate the number of bits required to store the number of spikes N_i with the Shannon entropy. Using this number of spikes we decode the delays \tilde{d}_i in order

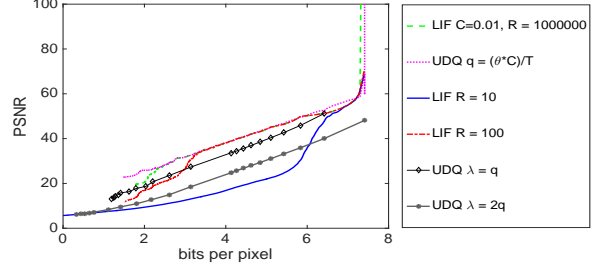


Fig. 3: Rate-Distortion comparison between LIF encoder/decoder and UDQ. The gray and black curves correspond to UDQ with the quantization step $q \in \{1, 10, 20, \dots, 90, 100\}$ with $\lambda = q$ and $\lambda = 2q$, respectively. The blue and red curves to the LIF encoder/decoder with $R = 10 \Omega$ and $R = 100 \Omega$ respectively. The green and pink curves illustrate that the LIF encoder/decoder can be approximated by UDQ according to Proposition 1 ($\theta \in \{1, \dots, 1000\}$, $C = 1 \text{ F}$, $T = 150 \text{ ms}$).

to reconstruct the input intensities $\tilde{I}_1, \dots, \tilde{I}_n$. Concerning the conventional UDQ quantization, we assign each input intensity to the correspondent quantization levels L . Based on the quantization levels we are able to compute the reconstructed value of each input intensity. Finally, we apply the Shannon entropy to estimate the number of bits required to encode the quantization levels.

Figure 3 shows the rate-distortion behavior when we compress lena image of size $n = 256 \times 256$ [16] using the LIF encoder/decoder and the UDQ, respectively. By tuning θ and R we are able to increase or decrease the number of spikes and, by extension, the number of bits which are required to store the input signal. The higher the rate is, the better the quality of the reconstructed signal. On the other hand, for the UDQ, the quantization step q links each input intensity to the correspondent quantization level L . The more the levels the higher the rate and vice versa. This figure also shows that when $C \ll T \ll R$ the LIF encoder/decoder can be approximated by a UDQ with a quantization step given by eq. (18).

6. CONCLUSION

In this paper we presented the extension of the perfect-LIF model which was recently released by our team. The novel LIF encoder/decoder is a neuro-inspired quantizer which mimics the dynamic spike generation neural mechanism. We introduced the link between the LIF encoder/decoder and the UDQ which results in the improvement of the performance of the latest for a given set of parameters. Our future goal is to be able to apply the LIF encoder/decoder to a time-varying input.

7. REFERENCES

- [1] AL L Hodgkin and AF F Huxley, "Currents carried by sodium and potassium ions through the membrane of the giant axon of Loligo," 1952.
- [2] Eugene M. Izhikevich, "Which model to use for cortical spiking neurons?," *IEEE Transactions on Neural Networks*, vol. 15, no. 5, pp. 1063–1070, 2004.
- [3] Z F Mainen and T J Sejnowski, "Reliability of spike timing in neocortical neurons.," *Science (New York, N.Y.)*, vol. 268, no. 5216, pp. 1503–6, 1995.
- [4] S J Thorpe, "Spike arrival times: A highly efficient coding scheme for neural networks," in *Parallel processing in neural systems and computers*, 1990, pp. 91–94.
- [5] R VanRullen and S J Thorpe, "Rate coding versus temporal order coding: what the retinal ganglion cells tell the visual cortex," *Neural computation*, vol. 13, no. 6, pp. 1255–83, 2001.
- [6] W. Gerstner and W. Kistler, *Spiking neuron models: Single Neurons, Populations, Plasticity*, Cambridge University Press, 2002.
- [7] D. H. Kelly, "Information Capacity of a Single Retinal Channel," *IRE Transactions on Information Theory*, vol. 8, no. 3, pp. 221–226, 1962.
- [8] J Gautrais and S J Thorpe, "Rate Coding vs Temporal Order Coding : a theoretical approach," *Biosystems*, vol. 48, pp. 57–65, 1998.
- [9] S J Thorpe, A Delorme, and R VanRullen, "Spike-based strategies for rapid processing," *Neural Networks*, vol. 14, no. 6-7, pp. 715–725, 2001.
- [10] K Masmoudi, M. Antonini, and P. Kornprobst, "Frames for exact inversion of the rank order coder," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, no. 2, pp. 353–359, 2012.
- [11] 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.
- [12] E. Doutsis, L. Fillatre, M. Antonini, and J. Gaulmin, "Bio-inspired Sparse Representation of Images," in *Gretsi*, 2017.
- [13] E. Doutsis, L. Fillatre, M. Antonini, and J. Gaulmin, "Retina-inspired Video Codec," in *Picture Coding Symposium (PCS)*, 2016.
- [14] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2325–2384, 1998.
- [15] Bernard Widrow, Istvan Kollar, and Ming Chang Liu, "Statistical Theory of Quantization," *IEEE Transactions on Instrumentation and Measurement*, vol. 45, no. 2, pp. 353–361, 1996.
- [16] A. Weber, "The USC-SIPI Image Database," 1977.