IMAGE SMOOTHING AND EDGE ENHANCEMENT AS NEUROMORPHIC DECODING BASED ON SAMPLING REPRESENTATION PARTITIONING

Viacheslav Antsiperov

Kotelnikov Institute of Radio engineering and Electronics,
Russian Academy of Sciences, Moscow, Russia

ABSTRACT

This work discusses the image neuromorphic encoding/decoding issues inspired by the mechanisms of visual perception encoding/decoding. The importance of this topic is directly related to the current problems of perceptual quality and perceptual reconstruction of images today. Therefore, to obtain reliable results in these directions, it was natural to turn to the most adequate mechanisms of perception. As a result, we propose a new approach to image processing, which uses the most realistic representation of the input data in the form of a stream of events or counts. Such events / counts simulate the firing of retinal receptors in response to the action of radiation recorded. The statistical model of the counts stream is chosen in the form of the two-dimensional inhomogeneous Poisson point processes, considered as a convenient representation of the input data. In the current paper such a representation is referred to as sampling representation. To adequately model the mechanisms of neural encoding of input data, we consistently use the concept of receptive fields. This general model implements well-known features of neural processing, including central/lateral inhibition. Decoding issues are considered in the context of Retinex paradigm of contrast detection. It is shown that the model of coupled ON-OFF receptive fields allows to restore sharp image details in the form of local edges. At the end of the work, we demonstrate an interpretation of synthesised encoding/decoding as classical smoothing and edge outlining of encoded images.

KEYWORDS

Neuromorphic Systems, Sampling Representation, Neural Encoding, Receptive Fields, Edges Detection

1. INTRODUCTION

The current time is often characterized as the time of Big Data. Over the past few decades, thanks to the explosive growth of the Internet and the development of its numerous services, the growth of the volumes of data surrounding us has increased unprecedentedly. However, this also brought a lot of specific problems that could be also called the “Big Problem”. This problem is big not only because it causes a lot of inconvenience for people, but also because it involves the need to store, process and exchange huge amounts of data between users, applications, and corporations. To a certain extent, the Big Problem is the other side of Big Data.

The Big Problem is most obvious in the case of visual content, for which the Internet has become a worldwide repository. It is estimated that the number of digital images will rise to 1.80 trillion in 2023 and it is expected to surpass two trillion in 2025. A massive portion of images are related

David C. Wyld et al. (Eds): SIPR, NCWC, BIBC, SOFEA, DSML, ARIA, NLP, CSEN -2023
pp. 157-168, 2023. CS & IT - CSCP 2023 DOI: 10.5121/csit.2023.131913

to the rapidly growing smartphone users. The development of CMOS video cameras and the increase in smartphone memory have led to the fact that in 2023 the average owner stores about 2100 photos on it. This in turn has led to the creation of specialized platforms on the Internet that support image exchange and give rise to a new unique media environment with a daily volume of data exchange of ~ 7 billion images [1].

The principal problems with video data lie in their usually significant binary volumes (even in comparison with audio data, not to mention texts). At the same time, the problems of storing large amounts of data associated with the required large memory resources are not so critical, but the problems of transmitting visual content over channels with physically limited capacity become fundamentally important. This problem is known as the "bottleneck problem" – an information systems problem where one component limits the performance of the entire system. The latter inevitably leads to increased transmission times, higher delays, and inconvenience for end users. At the same time, it cannot be said that the bottleneck problem is a part of the Big Problem and arose along with it. It always took place in a less acute form; moreover, it was, in fact, the main incentive for the creation of modern information theory by K. Shannon and his colleagues. Over the past decades, the theory has achieved impressive results. Let us note such a direction as the rate-distortion theory, directly related to the problem of the bottleneck [2]. Unfortunately, most results of this theory are of an asymptotic, recommendatory nature and in practical situations provide only potential (upper/lower) estimates of the performance of the system, but not specific methods for approaching them.

At the same time, fairly optimal practical solutions to the problem noted above exist and they were found by nature itself for living systems. This refers to the neurosensory systems and, first to the visual system of humans and higher vertebrates. Indeed, if we consider that the number of retinal receptors ("image pixels") reaches $\sim 10^8$, and the number of optic nerve axons (transmission channel capacity) is only 10^6 [3], then the degree of compression (distortion) of the input data encoding on the retina is ~ one hundred times, and this does not lead to a noticeable loss of information. In this regard, there is a strong opinion that a minimal deviation of reconstructed compressed images from the original ones does not in itself necessarily lead to good quality of perception. For example, it has been shown that the use of coding methods in generative adversarial networks can lead to a noticeable improvement in the quality of image perception, although the distortion of the original image may not be minimal [4]. In this regard, several attempts have recently been made to include additional elements in the theory of image coding that increase the resulting quality of their perception [5, 6]. New approaches have radically revised the classical methods for assessing image quality using distortion functions, which are usually defined as absolute or quadratic deviations of the restored version of the image from the original. It turned out that such functional metrics are poorly adequate to the peculiarities of human perception. In this regard, several attempts have been made to search for those non-traditional metrics that would objectively correspond to perception. Among the most well-known perceptual metrics are the structural similarity metric (SSIM) [7], the visual information-based metric (VIF) [8], and the metric based on spatial and temporal most apparent distortion (MAD) [9].

However, the greatest success in improving the quality of perception was achieved not by improving distortion metrics, but by revising the concept of distortion itself. We are talking about generative modelling, which is playing an increasingly important role in machine learning [10]. Generative models consider the entire set of (input/output) data as a set of random variables and, unlike discriminative models, are focused on their joint probability distributions rather than conditional ones. The practical success of generative models has not been fully clarified theoretically; however, one often hears the assertion that this may be due to more adequate modelling of perception features [11]. In this regard, special emphasis should be placed on the

role of data representations [12]. Modelling of retinal function and structure largely determines the representation of input-intermediate-output data and their relationships. Based on this, issues of choosing data representations should apparently play an important role in the development of new methods for images processing focused on high quality perception.

A detailed discussion of general issues on this topic can be found in our previous works, for example, in [13] (further references to the literature on the topic can also be found there). Current work is devoted to the presentation of new results obtained in the same direction – in the adaptation of neuro-encoding procedures inspired by the mechanisms of visual perception to the problems of optimal encoding (compression) and decoding of images. Thus, our main goal in this article is to present and discuss new results on the most rigorous statistical basis within the framework of adequate neural modelling.

2. STATISTICAL DESCRIPTION OF IMAGES IN THE FORM OF SAMPLING REPRESENTATIONS

A feature of our proposed approach is the special representation of images not in the form of continuous intensity $I(\vec{x}),\ \vec{x} \in \Omega$ of radiation incident on the retinal receptors, but in the form of a stream of random, discrete events at their output - in the form of the so-called (photo) counts. Note that from a physical point of view, also artificial imaging systems, such as CMOS cameras [14], form images through multiplication, accumulation, etc. samples of the initial photocurrent. A good discussion of the model of signals represented by streams of events and its statistical justification based on two-dimensional point Poisson processes can be found, for example, in [15]. A statistical description of this representation can also be obtained using the concept of an ideal recording device [13]. Namely, the result of registration by an ideal device of a flow of events from a photosensitive surface Ω is a set of samples $X = \{\vec{x}_i\}$, where each $\vec{x}_i,\ i = 1,\dots,N$ is the vector of coordinates of registered events - random points on Ω. Note that the number of recorded samples N is itself a random variable, its statistics is given by the Poisson distribution with the mean value $\overline{N} = \int_\Omega \alpha I(\vec{x})\, d\vec{x}$, where the coefficient $\alpha = \eta(h\bar{\nu})^{-1}$ is determined, among other things, by the quantum efficiency η of an ideal imaging device [13].

It is easy to show [15] that a set of random N samples $\{\vec{x}_i\}$ is also described by the probability distribution of a random number of points of some inhomogeneous point Poisson process with intensity function $\alpha I(\vec{x})$. Since the number of samples N is a random variable, this description is inconvenient for practical use (especially for large N). Therefore, we proposed to represent point processes by sets of random points, as in the original Poisson process, but with a fixed (controlled) number of points $N_s \ll \overline{N}$. It was shown [16] that the statistics of a fixed (non-random) size N_s sample $X_s = \{\vec{x}_j\}$ from $X = \{\vec{x}_i\}$ can be given by a single distribution density of the form:

$$\rho\big(X_s = \{\vec{x}_j\}, \mid I(\vec{x})\big) = \prod_{j=1}^{N_s} \rho\big(\vec{x}_j \mid I(\vec{x})\big), \quad \rho\big(\vec{x}_j \mid I(\vec{x})\big) = I\big(\vec{x}_j\big)/\int_\Omega I(\vec{x}) d\vec{x}. \tag{1}$$

In accordance with the above construction, the representation of images by counts $X_s = \{\vec{x}_j\}$, $j = 1,\dots,N_s$ was proposed to be called *a sampling representation*.

For illustrative purposes, as an example, we generated a sampling representation of image "butterfly-19" from the standard MPEG7 database [17], see Figure 1. The set $X_s = \{\vec{x}_j\}$ of $N_s = 10\ 000\ 000$ counts, for the GIF image "butterfly-19" of size 429×421 pixels, colour depth υ = 8

bits, was previously converted to PNG format with the same colour depth, but of size $s \times s$ =1000×1000 pixels. In this case, only two shades of gray 100 and 255 were used for the image. The counts were generated by the Monte Carlo rejection method with a uniform auxiliary distribution $u(\vec{x}) = s^{-2}$ and an auxiliary constant $M = 2^v$.

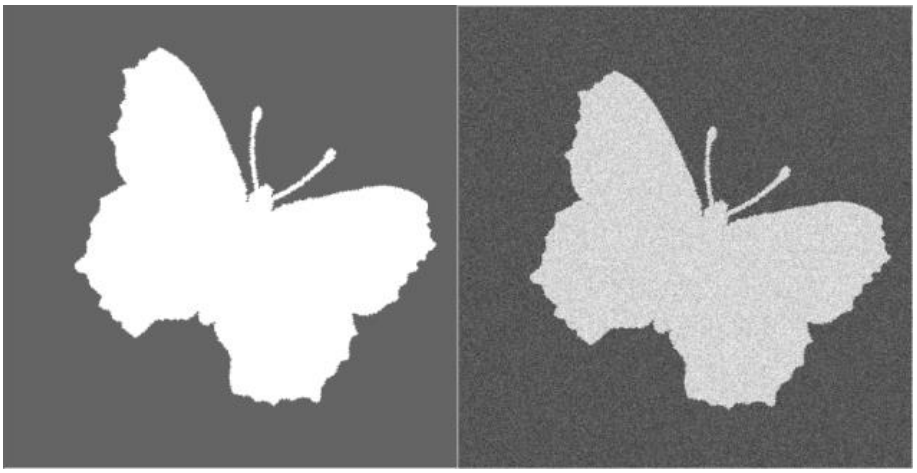

Figure 1. Representation of an image based on a sample of counts (sampling representation). On the left hand is the original image "butterfly-19" [17]. On the right hand is a sampling representation of 10 000 000 counts.

3. SAMPLING REPRESENTATION COMPRESSION BY A SYSTEM OF RECEPTIVE FIELDS

The sample representation $X_s = \{\vec{x}_j\}$ introduced above is most adequate to the data of receptors (rods/cones) in the outer layer of the retina when forming images based on the recorded intensity $I(\vec{x})$, $\vec{x} \in \Omega$. However, the impulses sent to the brain from neurons in the inner layer of the retina are different from the data directly recorded. They are formed based on the latter with the help of intermediate interneurons of the middle and inner layers. As a result, the output neurons of the retina, which form the input representation for subsequent parts of the visual cortex, aggregate samples from tens and sometimes thousands of receptors located in small areas of the retina, called receptive fields (RF). The systematic study of the RP system and the neural transformation of input data from receptors into a sequence of retinal output data is usually associated with the fundamental work of Hubel and Wiesel [18]. A modern presentation of the structure and functions of the RP can be found in [19].

The functions and sizes of various RPs are determined by the types of ganglion cells (retinal output neurons) associated with them. There are about ~20 types of ganglion cells and, accordingly, types of RP. In what follows, for simplicity, only the family of midget ganglion cells encoding the spatial distribution on Ω of intensity $I(\vec{x})$ is considered. Typical responses of midget cells to the nature of lighting/darkening of the corresponding RF are determined by its centre-antagonistic structure. Thus, the ON-cell is excited when the centre of the RF is stimulated and inhibited when the concentric periphery is stimulated. On the contrary, the OFF cell is activated upon stimulation of the RF periphery and inhibited upon stimulation of the centre [18]. The presence of two types of midget cells is due to the peculiarities of neural coding of positive/negative changes in stimuli (ON is activated when the stimulation of the centre exceeds the average stimulation in the field, OFF is the opposite).

As for the spatial location of the RFs, it was found that neighbouring ON and OFF cells have significantly overlapping fields, and the RF of cells of the same type practically do not overlap. In this case, non-overlapping RFs of each type are tightly adjacent to each other, forming a kind of mosaic that tightly covers the retina [20]. Thus, for the mathematical formalization of the structure of the RF system, it is sufficient to allow the intersection of ON fields only with neighbouring OFF fields and the absence of intersections with other ON fields, however, allowing for the contact of their boundaries. The same is true for the ON ↔ OFF exchange. A formal representation of the region Ω by square RFs with round centres, which will be used later for algorithmic purposes, is presented in Figure 2 (see details in [21]).

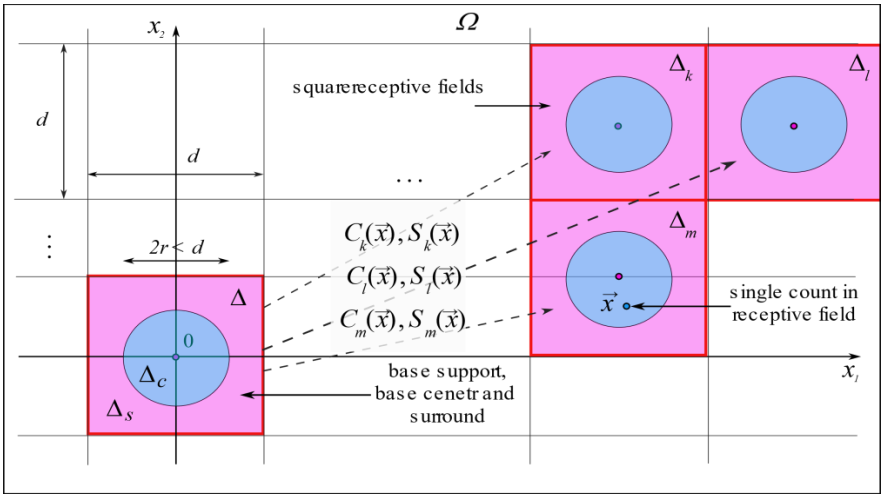

Figure 2. Partition of the image surface Ω by a system of receptive ON fields $\{C_k(\vec{x}), S_k(\vec{x})\}$ with square support $\Delta_k \cup \Delta_s$ located at the nodes of a regular square lattice.

In view of the noted symmetry in the arrangement of ON and OFF fields, we will consider the RF system for only one, for example, ON type of fields. Based on the previous brief review, we formalize the probability density model of the sample $\vec{x}_j$ on Ω – $\rho(\vec{x}_j \,|I(\vec{x}))$ (1) as a parametric family of probability densities $\mathbb{P} = \{\rho(\vec{x};\vec{\theta})|\vec{\theta} \in \Theta\}$, which are mixtures of K pairs of components $\{C_k(\vec{x}), S_k(\vec{x})\}$, $k = 1, \dots, K$:

$$\rho(\vec{x}\,;\vec{\theta}) = \sum_{k=1}^{K} w_k C_k(\vec{x}) + v_k S_k(\vec{x}) \tag{2}$$

where $\vec{\theta} = \{w_k, v_k\}$ are the positive weights of the mixture, the parameters of the model $\mathbb{P}$, and the mixture components $C_k(\vec{x})$ and $S_k(\vec{x})$ represent the centre and antagonistic surround of the k-th RF. The components $C_k(\vec{x})$ and $S_k(\vec{x})$ are specified by positive, normalized probability distribution densities having compact supports $\Delta_k^c = \{\vec{x}\,|C_k(\vec{x}) > 0\}$ and $\Delta_k^s = \{\vec{x}\,|S_k(\vec{x}) > 0\}$, which form the combined supports of the k-th RFs: $\Delta_k = \Delta_k^c \cup \Delta_k^s$ (see Figure 2):

$$\int_{\Delta_k^c} C_k(\vec{x})d\vec{x} = \int_{\Delta_k^s} S_k(\vec{x})d\vec{x} = 1. \tag{3}$$

The meaning of the introduced parameters $\vec{\theta} \in \Theta$ becomes transparent if we recall the identity of the density $\rho(\vec{x}_j \,|I(\vec{x}))$ of the normalized version of $I(\vec{x})$ fixed by (1). Essentially (2), up to a normalizing factor, specifies the expansion of intensity $I(\vec{x})$ over a system of local distributions, as is customary, for example, in wavelet or any other multi-resolution analysis [22]. Relations (3)

fix the normalization of the basis functions. The analogy can be extended even further if we assume that the supports of the centre and the antagonistic surround of the k-th RF Δ_k^c and Δ_k^s do not have common points $\Delta_k^c \cap \Delta_k^s = \emptyset$. Then to the normalization relations (3) we can also add relations of the orthogonality type:

$$\int_{\Delta_k^c} S_k(\vec{x})d\vec{x} = \int_{\Delta_k^s} C_k(\vec{x})d\vec{x} = 0. \tag{4}$$

Remembering, in addition, that the set of RF supports $\{\Delta_k\}$ constitutes an exact partition of the retinal surface, i.e., all of them do not intersect in pairs, but together they densely cover Ω, it is possible, just as in multi-resolution analysis, to express the model parameters $\vec{\theta} = \{w_k, v_k\}$ through the corresponding density $\rho(\vec{x}\,;\vec{\theta})$ (2) integrals over the corresponding RF supports:

$$\begin{aligned} w_k &= \int_{\Delta_k^c} \rho(\vec{x}\,;\vec{\theta})d\vec{x}\ , \\ v_k &= \int_{\Delta_k^s} \rho(\vec{x}\,;\vec{\theta})d\vec{x}\ . \end{aligned} \tag{5}$$

which leads to the interpretation of the parameters w_k, v_k also as the probabilities of a counts falling into the center Δ_k^c or into the surround $\Delta_k^s$of the k-th RF. An equivalent interpretation is to characterize w_k, v_k (5) as the average values of the characteristic functions $\Pi_k^c(\vec{x}) = 1,\ if\ \vec{x} \in \Delta_k^c, else\ 0$ and $\Pi_k^c(\vec{x}) = 1\ , if\ \vec{x} \in \Delta_k^s,\ else\ 0$ on the surface Ω .

Obviously, connections (5) cannot be used to find w_k, v_k, since the density $\rho(\vec{x}\,;\vec{\theta})$ is not known, only the sample $X_s = \{\vec{x}_j\}$ is known in relation to it. However, here you can use the standard technique trick, presented, for example, in [25]. Namely, taking into account the asymptotic of the law of large numbers and replacing the averages of $\Pi_k^c(\vec{x})$ and $\Pi_k^s(\vec{x})$ with their sample (empirical) averages, we can approximately write:

$$\begin{aligned} w_k &= \frac{1}{N_s}\sum_{j=1}^{N_s} \Pi_k^c(\vec{x}_j) = \frac{n_k^c}{N_s}\ , \\ v_k &= \frac{1}{N_s}\sum_{j=1}^{N_s} \Pi_k^s(\vec{x}_j) = \frac{n_k^s}{N_s}\ , \end{aligned} \tag{6}$$

where n_k^c and n_k^s are the numbers of counts in the centre and in the surround of the k-th RF.
Note that the approximate values of parameters (6) do not depend on the form of components $C_k(\vec{x})$ and $S_k(\vec{x})$, but only on the form of their supports Δ_k^c and Δ_k^s. It follows that for an approximate estimate of the probability density $\rho(\vec{x}\,;\vec{\theta})$ (3) only the numbers n_k^c and n_k^s of counts in the centres / surrounds of the receptive fields are sufficient. In other words, the sampling representation $X_s = \{\vec{x}_j\}$ of the image can be reduced (compressed) in the case under consideration to a "occupation number" representation $Y_s = \{n_k^c\ , n_k^s\}$, which in this context is sufficient statistics of the sampling representation.

4. THRESHOLD ENCODING OF COMPRESSED SAMPLING REPRESENTATIONS

Starting from the compressed representation of an image with occupation numbers $Y_s = \{n_k^c\ , n_k^s\}$, let's look in more detail at how this data can be encoded (compressed) when passing it

to subsequent processing / analysis stages. Let us assume for simplicity that the RF system is homogeneous – all its fields are identical in structure and function (see Figure 2). Let a typical field on Ω have area $\sigma = |\Delta|$, its centre has area $\sigma_c = |\Delta^c|$ and the antagonistic surround has area $\sigma_s = |\Delta^s|$, $\sigma = \sigma_c + \sigma_s$.

Further, let a typical RF has a simple set of functions – it can find the total number of counts n belonging to it, the number of counts in the centre n_c and their arbitrary linear combination $\alpha n + \beta n_\mathrm{c}$ (including $n_s = n - n_\mathrm{c}$). It is assumed that the coefficients α and β can be arbitrary both in magnitude and sign, which is determined by the excitatory / inhibitory nature of the corresponding RF regions. Due to the random nature of counts registration, the numbers n, n_c and n_s are also random, although dependent ($n = n_\mathrm{c} + n_s$). It is easy to show that for sampling representations $X_s = \{\vec{x}_j\}$ the typical field data n, n_c and n_s are Poisson random variables. Here it should be noted that since the centre and surround of the RF are non-overlapping areas, n_c and n_s are statistically independent. Accordingly, their probability distributions have the form:

$$\begin{aligned} P_\mathrm{c}(n_c|\,\lambda) &= \frac{(\sigma_c\lambda)^{n_c}}{n_c!}\exp\{-\sigma_c\lambda\}, \\ P_s(n_s|\,\mu) &= \frac{(\sigma_s\mu)^{n_s}}{n_s!}\exp\{-\sigma_s\mu\}, \end{aligned} \tag{7}$$

where λ and μ are the intensity values of the counts in the centre and in the surround of the RF:

$$\lambda = \frac{\alpha}{\sigma_c}\int_{\Delta^c} I(x)dx\,,\ \mu = \frac{\alpha}{\sigma_s}\int_{\Delta^s} I(x)dx \tag{8}$$

Note that the average values of distributions (7) are related to intensities (8) through $\bar{n}_\mathrm{c} = \sigma_c\lambda$ and $\bar{n}_s = \sigma_s\mu$. Therefore, the RF data n_c and n_s, being unbiased estimates of their averages $\bar{n}_\mathrm{c}$ and $\bar{n}_s$, also provide unbiased estimates n_c/σ_c and n_s/σ_s of λ and μ.

The joint distribution of n_c and n_s, due to their independence, is obtained by the usual multiplication of distributions (7). If we pass from the data n_c and n_s first to the numbers n_c and n, and then from them to $\delta = n_\mathrm{c} - (\sigma_c/\sigma)n$ and n, then after approximating the binomial distribution by Gaussian and a number of transformations and simplifications we will arrive at the following RF data model:

$$\begin{gathered} P(\delta, n|\,\lambda,\mu) = P(\delta|n,\lambda,\mu)P(n|\,\nu), \\ P(\delta|n,\lambda,\mu) = \frac{1}{\sqrt{2\pi\gamma^2}}\exp\left(-\frac{(\delta-\varepsilon)^2}{2\gamma^2}\right),\ \ \varepsilon = \frac{\sigma_c\sigma_s}{\sigma^2}\left(\frac{\lambda-\mu}{\nu}\right)n,\ \ \gamma^2 = \frac{\sigma_c\sigma_s}{\sigma_i^2}\frac{\lambda\mu}{\nu^2}n\,, \\ P(n|\,\nu) = \frac{(\sigma\nu)^n}{n!}\exp\{-\sigma\nu\},\ \ \nu = \frac{\sigma_c}{\sigma}\lambda + \frac{\sigma_s}{\sigma}\mu\,. \end{gathered} \tag{9}$$

For a complete statistical description of the (generative) RF model, it is necessary to specify an *a priori* joint distribution of intensities λ and μ. Let's choose it in the form:

$$\rho(\lambda\,|\,\mu) = \omega\delta(\lambda-\mu) + (1-\omega)\wp(\lambda) \tag{10}$$

where the weights ω and $1-\omega$ can be interpreted as the probability of the 0-hypothesis H_0 that λ and μ coincide and, accordingly, as the probability of the alternative $\bar{H}_0$ that λ and μ are

independent. From this interpretation it follows that $\wp(\lambda)$ represents the *a priori* unconditional probability distribution of each of the intensities λ and μ.

Basing on the generative RF model (9–10) and using standard statistical methods, it is possible to obtain *posterior* (depending on observed data δ and n) distributions λ and μ and their optimal (in the sense of the maximum *posterior* probability) estimates, to derive tests of H_0 vs $\overline{H}_0$, etc. Omitting intermediate calculations and approximations, we present only one of the final results. The first moments (average values) of the posterior distribution $\rho(\lambda, \mu\,|n_c\,, n)$, which can be chosen as MAP (maximum *posteriori* probability) estimates $\bar{\lambda}(\delta, n)$ and $\bar{\mu}(\delta, n)$ for the intensities λ and μ have the form:

$$\bar{\lambda}(\delta,n)=\begin{cases} n/\sigma, & |\delta|\le D\sqrt{n}\,; \\ n_c/\sigma_c\,, & |\delta|> D\sqrt{n}\,; \end{cases}\,, \qquad \bar{\mu}(\delta,n)=\begin{cases} n/\sigma, & |\delta|\le D\sqrt{n}\,; \\ n_s/\sigma_s\,, & |\delta|> D\sqrt{n}\,; \end{cases}\,. \tag{11}$$

where the threshold coefficient $D^2=2\sigma_c\sigma_s\ln\{\bar{\Lambda}_0\}/\sigma^2$ does not depend on the data, but depends only on the parameters of the problem, including the *a priori* likelihood $\bar{\Lambda}_0$, which, under reasonable assumptions, can be given by the following approximate expression:

$$\bar{\Lambda}_0=\frac{4\omega}{\pi(1-\omega)}\sqrt{\frac{\sigma_c\sigma_s\bar{n}}{\sigma^2}} \tag{12}$$

where $\bar{n}=\sigma\hat{v}$, $\hat{v}$ is the characteristic scale of the *a priori* probability distribution $\wp(v)$. Encoding RF data δ and n into threshold estimates $\bar{\lambda}(\delta, n)$ and $\bar{\mu}(\delta, n)$ for the Figure 1. image on a 50×50 RF lattice (see Figure 2) is illustrated in Figure 3.

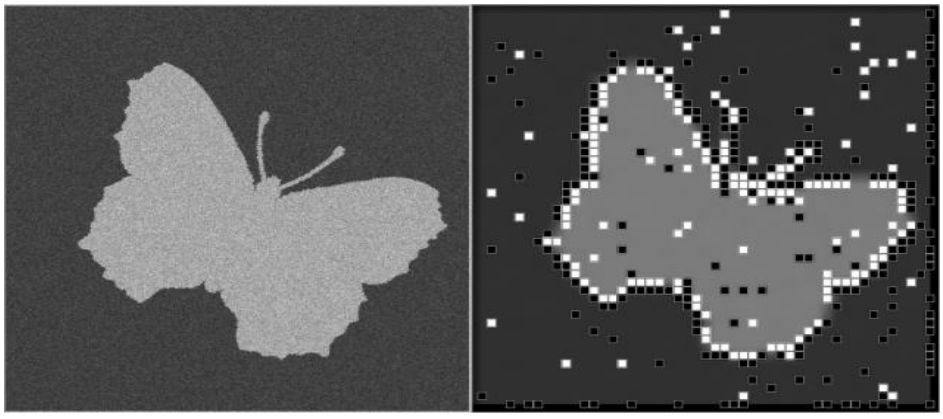

Figure 3. Illustration of the procedure (11) encoding results for a sampling representation of the "butterfly-19" image [17] Figure 1 by a 50×50 RF lattice. On the left hand is a sampling representation, on the right hand is the RF lattice with marked non-zero δ fields – ON responses ($\delta > 0$) in white, OFF responses ($\delta < 0$) in black.

Procedure (11) can be considered as a regression in data compression [23], if we consider n/σ as a predictor of both intensities λ and μ (provided that the 0-hypothesis H_0 holds), and values $\delta/\sigma_c = n_c/\sigma_c - n/\sigma\,, \delta/\sigma_s = n_s/\sigma_s - n/\sigma$ as residuals for such estimate. The only difference between (11) and the LASSO procedure (for the "smallest absolute shrinkage

operator") proposed in [23] is the dependence of the threshold $D\sqrt{n}$ on the total number of counts n on the RP. The latter is due to the correlation of the of Poisson noise intensity with the image signal power (i.e., with the intensity $I(\vec{x})$). In this regard, procedure (11) can be also defined as a noise dependent shrinkage encoder.

5. DECODING COMPRESSED SAMPLING REPRESENTATIONS USING SMOOTHING AND EDGE ENHANCEMENT

To assess the effectiveness of the proposed encoding procedure (11), it is desirable to evaluate the degree of distortion of images after their decoding (reconstruction). To do this, without resorting to formal statistical synthesis of decoders, it is possible, at least to a first approximation, to use well-known methods of interpolating images with low due to compression (coding) resolution. Considering that the number of RF counts $\{n_k\}$ is the output of a smoothing filter with a sliding window of area σ, we can immediately say that the distortions associated with this part of the code have the form of the original image blurring (see Figure 4 on the left). The filtering associated with the rest part of the code $\{\delta_k\}$ can be implemented by a piecewise-constant filter, having constant and positive value in the region of the centre of the RF, constant but negative at its surround and zero overall DC response. Similar filters are well-known in the field of digital image processing, in particular, the synthesized filter exactly matches the COSO (centre-ON-surround-OFF) filter proposed in [24]. The COSO filter was proposed to simulate the response of the LoG (Laplacian of Gaussian) filter used by Marr and Hildreth in the theory of edge detection [25]. Therefore, since the codes $\{\delta_k\}$ turn out to be closely related to the Laplace operator, the zeros of their linear interpolation along the edges of the RF lattice represent the points of intersection of the image edges (sharp contrast changes) with lattice (Marr's thesis). Therefore, numerous edge-enhancement methods can be used to reconstruct images that preserve meaningful details [26].

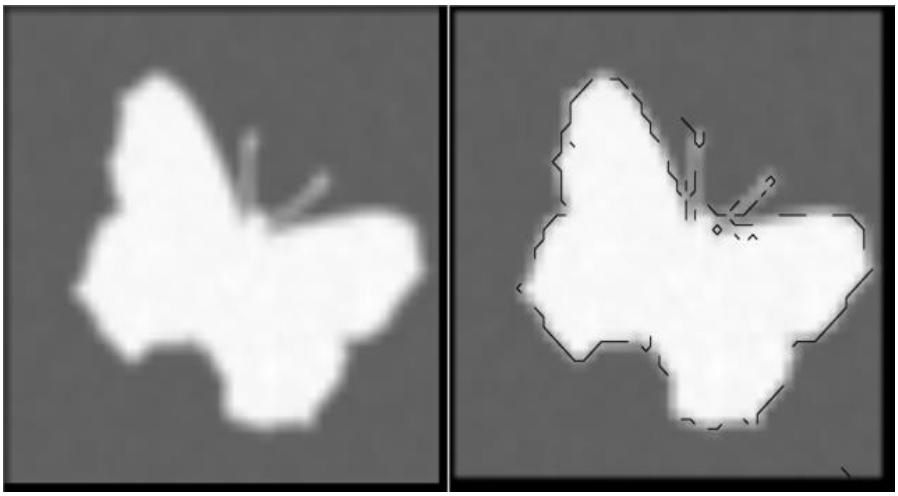

Figure 4. Example of restored (decoded) image "butterfly-19" [17] Figure. 1, defined by the compressed sampling representation $Y_s = \{n_k^c, n_k^s\}$ on a 50×50 lattice. On the left hand is a smoothed image decoded using only the "smooth" part $\{n_k\}$ of the code, on the right hand are edge contours, which were determined by $\{\delta_k\}$ added to the smoothed image.

To illustrate the reconstruction (decoding) of images, we used the encoded sampling representation of "butterfly-19" image, shown in Figure 3 on the right. The image area Ω was covered by a set of 2500 square receptive fields arranged at the nodes of a 50 × 50 square lattice

(see Figure 2 and details in [21]). At each k-th node, the values $\{\delta_k, n_k\}$ were calculated, after which δ_k was censored with a threshold $D\sqrt{n_k}$ (11). Classic bilinear interpolation, shown in Figure 4 on the left shows a reconstruction based only on a subset $\{n_k\}$ of the coded data, interpolating them first linearly along the vertical edges of the lattice, and then linearly across all rows of all lattice cells basing on the already interpolated values of the vertical edges. Obviously, the image obtained as a result of such an interpolation (a kind of smoothing) does not have very high visual quality, mainly due to the blurring of the boundaries between dark and light areas. To enhance edge contrast, we follow Marr's suggestions by detecting local edges of image and outlining them with a contrasting (black) color. Algorithmically, this procedure is implemented as follows. For each lattice cell – each receptive field, all four edges of its boundary are analyzed from the point of view of their intersections with the edges of objects in the image. A lattice edge is considered to intersect with some edge of the image if at its boundary node-points the values δ_i and δ_j are non-zero and if they have different signs. The middle of this lattice edge is taken as the intersection point. If the above condition is not met, the lattice edge is considered free of intersections. Once all edges of a cell boundary have been analyzed, it is determined whether the cell in question contains an image edge fragment. Namely, if it contains exactly two intersecting lattice edges, their intersection points are connected by a straight-line segment, which is then considered as a fragment of an image edge in this cell. Note that if two adjacent lattice cells have a common intersecting edge, then the corresponding fragments together form a section of the broken line of the contour of some image object. By constructing similar fragments for all cells, we, in a certain sense, decode the additional part of the code $\{\delta_k\}$ associated with the edges-details of the image. The results of the procedure described are presented on the right side of Figure 4 and give some insight into the quality of images reconstructed on the basis of the proposed neuromorphic encoding/decoding. These results (see Figure 4 right) appear to have better visual quality compared to classical bilinear interpolation (see Figure. 4 left).

6. CONCLUSIONS

In this work, in the context of modeling a number of functions and mechanisms of the periphery of the visual system (retina), it was possible to formulate a general approach to the synthesis of procedures for neuromorphic encoding/decoding of images. It is shown that on the basis of the proposed approach it is possible to substantiate a number of well-known features of visual perception, including those that form the basis of the popular image processing approach under the general name Retinex. Here, the first thing to note is the statistical substantiation of the importance of the central/lateral inhibition mechanism for enhancing image contrast.

Within the framework of the proposed approach, we were able to completely synthesize the encoding procedure and interpret it in terms of methods of optimal image processing and restoration known in DSP. By conveniently describing the input data in the form of a sampling representation, we succeeded in adequately modelling many neural coding mechanisms and we effectively incorporated them into the presented generative coding model. Of particular note here is the fact that the proposed model is strongly motivated by systems of ON-OFF coupled receptive fields –a universal principle underlying biological neural systems.

It should be stressed that the proposed generative approach to the synthesis of methods for perceptual coding of images, due to the possibility of transferring a part of the processing procedures directly to the periphery of image formation (by analogy with the retina-cortex system), opens up wide opportunities for research in the field of neuromorphic information systems. In this regard, basing on the results obtained and the numerical experiments performed, we can express the hope that the approach proposed in the work will find further theoretical development and effective use in applied problems. This is confirmed by the fact the proposed

approach has a natural extension to the field of parameters shrinkage methods, and the latter, as it turned out recently, has numerous, non-trivial connections with such areas of machine learning as anisotropic diffusion methods, wavelet approaches and variational techniques, which have shown themselves as the best tools in the field of CNN architecture [27].

ACKNOWLEDGEMENTS

The author would like to thank for the financial support of this work the Kotelnikov Institute of Radio Engineering and Electronics within the framework of the state contract "RELDIS".

REFERENCES

[1] Bull, D. R. & Zhang, F. (2021) *Intelligent image and video compression: communicating pictures.* 2nd ed. London: Acad. Press.

[2] Cover, T.M. & Thomas, J.A. (1991) *Elements of Information Theory*, New York: Wiley.

[3] Schiller, P.H. & Tehovnik, E.J. (2015) *Vision and the Visual System.* Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199936533.001.0001.

[4] Tschannen, M., Agustsson, E. & Lucic, M. (2018) "Deep Generative Models for Distribution-Preserving Lossy Compression" *Proc. of the 32nd International Conference on Neural Information Processing Systems (NIPS).* pp. 5933–5944.

[5] Blau, Y. & Michaeli, T. (2019) "Rethinking Lossy Compression: The Rate-Distortion-Perception Tradeoff" *Proc. of the 36th International Conference on Machine Learning, PMLR.*V. 97. pp. 675-685.

[6] Matsumoto, R. (2018) "Introducing the perception-distortion tradeoff into the rate-distortion theory of general information sources " *IEICE Communications Express*. V. 7(11). pp. 427–431. doi: 10.1587/comex.2018XBL0109.

[7] Wang, Z., Bovik, A., Sheikh, H., Simoncelli, E. (2004) "Image quality assessment: from error visibility to structural similarity" *IEEE Trans. Image Process.* V.13(4). pp. 600–612. doi: 10.1109/TIP.2003.819861.

[8] Sheikh, H., Bovik, A. & de Veciana, G. (2005) "An information fidelity criterion for image quality assessment using natural scene statistics: *IEEE transactions on image processing*. V. 14(12). P. 2117–2128. doi: 10.1109/TIP.2005.859389.

[9] Larson, E.C. & Chandler, D. M. (2010) "Most apparent distortion: full-reference image quality assessment and the role of strategy" *Journal of electronic imaging*, V.19(1), pp. 011006–011006. doi: 10.1117/1.3267105.

[10] Bishop, C.M. & Lasserre, J. (2007) "Generative or Discriminative? Getting the Best of Both Worlds" *Bayesian Statistics*. V. 8. London: Oxford University Press. pp. 3-24.

[11] Hassabis, D. & Kumaran, D., Summerfield, C., Botvinick, M. (2017) "Neuroscience-Inspired Artificial Intelligence" *Neuron*, V. 95(2). pp. 245–258. doi: 10.1016/j.neuron.2017.06.011.

[12] Bialek, W., Steveninck, R., & Tishby, N. (2006). "Efficient representation as a design principle for neural coding and computation" *2006 IEEE International Symposium on Information Theory*. doi:10.1109/isit.2006.261867.

[13] Antsiperov, V. & Kershner, V. (2023) "Retinotopic Image Encoding by Samples of Counts" De Marsico, M., Sanniti di Baja, G., Fred, A. (eds.) *Pattern Recognition Applications and Methods, ICPRAM 2021–2022. Lecture Notes in Computer Science*, Vol. 13822. Springer, Cham. doi: 10.1007/978-3-031-24538-1_3.

[14] Fossum, E. (2020) "The invention of CMOS image sensors: a camera in every pocket" *2020 Pan Pacific Microelectronic Symp*., pp. 1–6. doi: 10.23919/PanPacific 48324.2020.9059308.

[15] Streit, R. L. (2010) *Poisson Point Processes Imaging, Tracking, and Sensing*. Springer. doi: 10.1007/978-1-4419-6923-1.

[16] Antsiperov, V. (2022) "Generative Model for Autoencoders Learning by Image Sampling Representations" *Proceedings of the 11th Int. Conference on Pattern Recognition Applications and Methods – ICPRAM 2022*, pp. 354-361. DOI: 10.5220/0010836800003122

[17] Latecki, L.J., Lakamper, R. & Eckhardt, T. (2000) "Shape descriptors for non-rigid shapes with a single closed contour" *Proceedings IEEE Conference on Computer Vision and Pattern Recognition CVPR 2000* (Cat.No.PR00662), pp. 424–429. doi: 10.1109/CVPR.2000.855850.

[18] Hubel, D. H. & Wiesel, T.N. (2004) *Brain and Visual Perception: The Story of a 25-year Collaboration*. New York: Oxford University Press.
[19] Schiller, P.H. & Tehovnik, E.J. (2015) *Vision and the Visual System*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199936533.001.0001.
[20] Gauthier, J. L. & Field, G. D., et al (2009) "Receptive fields in primate retina are coordinated to sample visual space more uniformly" *PLoS Biol.* Vol. 7, No 4. p. e1000063. doi: 10.1371/journal.pbio.1000063.
[21] Antsiperov, V. (2023) "New Centre/Surround Retinex-like Method for Low-Count Image Reconstruction" *Proceedings of the 12th Int. Conf. on Pattern Recognition Applications and Methods (ICPRAM 2023)*, 2023, SCITEPRESS, pp. 517-528. doi: 10.5220/0011792800003411.
[22] Mallat ,S.G. (1989) "A theory for multiresolution signal decomposition: the wavelet representation" *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 11, No 7, pp. 674–693. doi: 10.1109/34.192463.
[23] Tibshirani, R. (1996) "Regression Shrinkage and Selection Via the Lasso" *Journal of the Royal Statistical Society: Series B (Methodological)*, Vol. 58. No 1, pp. 267–288. doi: 10.1111/j.2517-6161.1996.tb02080.x
[24] Allebach ,J. & Wong, P.W. (1996) "Edge-directed interpolation" *Proceeding of the 3rd IEEE Int. Conference on Image Processing*, Vol. 2, pp. 707–710. doi: 10.1109/icip.1996.560768.
[25] Marr, D. & Hildreth, E. (1980) "Theory of Edge Detection" *Proceedings of the Royal Society of London. Series B, Biological Sciences*, Vol. 207, No 1167, pp. 187–217.
[26] Ziou, D. & Tabbone, S. (1998) "Edge Detection Techniques-An Overview" *Pattern Recognition and Image Analysis: Advances in Mathematical Theory and Applications*, Vol. 8, No 4, pp.537-559.
[27] Alt, T., Weickert, J. & Pascal, P. (2020), "Translating Diffusion, Wavelets, and Regularisation into Residual Networks" *arXiv.org*. doi: 10.48550/arxiv.2002.02753.

AUTHOR

Viacheslav Antsiperov. Graduated from Moscow Institute of Physics and Technology in 1982. Received Candidate's degree (Physics and Mathematics) in 1986. At present, leading researcher at the Kotelnikov Institute of Radio Engineering and Electronics of the Russian Academy of Sciences. Scientific interests: information systems, processing and analysis of signals including image and speech recognition, biomedical informatics. Author of more than 140 papers.