Turkish Journal of Electrical Engineering & Computer Sciences

http://journals.tubitak.gov.tr/elektrik/

Research Article

Turk J Elec Eng & Comp Sci
(2014) 22: 223 – 241

doi:10.3906/elk-1206-61

Spatiotemporal realization of an artificial retina model and performance evaluation through ISI- and spike count-based image reconstruction methods

İrfan KARAGÖZ, Mustafa ÖZDEN*
Electrical and Electronics Department, Faculty of Engineering, Gazi University, Ankara, Turkey

Received: 21.06.2012 • **Accepted:** 04.09.2012 • **Published Online:** 20.12.2013 • **Printed:** 20.01.2014

Abstract: Development of an artificial retina model that can mimic the biologic retina is a highly challenging task and this task is an important step in the development of a visual prosthesis. The receptive field structure of the retina layer is usually modeled as a 2D difference of Gaussian (DOG) filter profile. In the present study, as a different approach, a retina model including a 3D 2-stage DOG filter (3D-ADOG) that has an adaptively changing bandwidth with respect to the local image statistic is developed. Using this modeling, the adaptive image processing of the retina can be realized. The contribution of the developed model in terms of the image quality is evaluated via simulation studies using test images. The first simulation results, including only the spike count-based reconstruction for a test video sequence, were previously published. In this study, in addition to the spike count-based reconstruction, the interspike interval measure is also used in the simulation study. The reconstruction results are compared using the statistical parameters of the mean squared error (MSE), universal quality index (UQI), and histogram similarity ratio (HSR), which characterize the image likelihood. To evaluate the performance of the model versus time, time-dependent changes in the MSE, HSR, and UQI parameters are obtained and compared to the standard model. From these results, it is concluded that the 3D-ADOG filter-based retina model preserves the spatial details of the image and produces a larger number of different gray tone levels, which are important for the visual perception of an image, in comparison with the well-known classical DOG filter-based retina model. The retina implant systems based on this model can provide better visual perception for implant recipients.

Key words: Image processing, retina modeling, artificial sight systems

1. Introduction

The visual system presents most of the environmental data to the brain. The absence of visual data greatly decreases quality of life. Although preventive treatments exist, there is no treatment method for complete vision loss caused by retinal diseases such as age-related macular degeneration, retinal detachment, and retinitis pigmentosa, which lead to photoreceptor loss [1–3]. In recent years, by overcoming this handicap, to create useful vision, many studies have been started to develop artificial vision systems (or visual prostheses) that electrically stimulate the remaining retinal ganglion cells [1,4–11].

In the visual system, the first step of visual processing starts in the eye. In the perception process of the image, the transformation of light into action potentials, called the 'phototransduction process', takes the first place in the process chain. The phototransduction process takes place through 5 different cell layers of the retina. These layers are the photoreceptor, horizontal, bipolar, amacrine, and ganglion cell layers, from

*Correspondence: mtozden@gazi.edu.tr

outside to inside, respectively. The photoreceptor cells introduce excitatory signals, whereas the horizontal cells introduce inhibitory signals. The horizontal cells make the lateral synapses with receptors along the whole retina. The synapses of the receptors and horizontal cells converge on the bipolar cells. Amacrine cells are somewhat similar to horizontal cells in terms of their functions and their synapses. The synapses of the bipolar and amacrine cells converge on the ganglion cells and the bipolar cells introduce excitatory signals, whereas amacrine cells introduce inhibitory signals. The last cell group, the ganglion cells, form the optic nerve with their axons. These cells convert the potentials from the bipolar and amacrine cells into action potentials and produce spike signals by a special coding mechanism. As the ganglion cells remain functional during vision loss due to photoreceptor loss, sight can be restored with the appropriate electrical stimulation signals.

In sight restoration studies, the main task is to convert the image data into encoded spike pulses that are meaningful for the visual cortex. After this step, these encoded pulses are sent to the targeted place of the visual pathway on which the microelectrode matrix placed. Based on electrode placement strategies, visual prostheses can be grouped into 3 main categories, as the visual cortex, optic nerve, and retinal implants [1,2]. Due to the ease of electrode placement and less surgical risks, many studies are focused on the retina implant system. In the retina implant system, the epiretinal approach uses an electrode array placed on the retina surface and aims to mimic the processing of the healthy retina layer. In the literature, aside from wavelet and Gabor filter-based approaches, the receptive field of the retina layer is implemented based on the spatiotemporal difference of Gaussian (DOG) filter and the spike generation of the retina is implemented based on a 'noisy leaky-integrate and fire' (nLIF) neuron model [5,9].

The nonlinear transduction and adaptation mechanism of the retina cannot be modeled based only on the DOG filter structure. However, it describes many of the biologic properties of the filtering behavior of the retina [12–14]. Furthermore, the image data are processed by the retina in both the spatial and temporal domains [5,15], and this concept is the basis of retina modeling studies [4,5,9].

Although standard DOG filter-based retina models are widely accepted in the literature, there are still some issues that should be improved. First, they do not include the local adaptation mechanism of the retina layer. However, there are some studies on modeling the adaptation mechanism that do not involve the DOG filter profile and deal with the adaptation problem as a distinct approach [4,9]. Second, some important parameters of the DOG filter-based retina model are determined by time-consuming trial and error techniques [6]. Finally, the spatial receptive field profile of the DOG filter cannot correctly model the contiguous structure of the receptive field of the retina [16]. By considering these issues, standard DOG filter-based retina models show user-dependent and relatively parametric characteristics.

In the present study, to overcome these disadvantages of the standard DOG filter-based retina models, a multistage DOG filter-based artificial retina model that changes the bandwidths of the DOG filters depending on the statistical characteristics of the local image data is proposed. The proposed model has a 2-stage spatiotemporal filtering structure that is based on the inner cell layers of the biologic retina, and the interactions of both the receptor-horizontal and bipolar-amacrine cells are modeled using an adaptive DOG filter [17,18]. The nLIF cell model, which is a good approximation for modeling the firing mechanism of ganglion cells, is used for the realization of the spike generation process in artificial retina models. The nLIF model is adopted to generate the outputs for the ON-center, OFF-center, and ON/OFF-center ganglion cells. Since there are both 2-stage spatial and temporal image filtering, the developed artificial retina model is referred to as a 3D 2-stage adaptive DOG filter-based retina model (3D-ADOG).

To evaluate the efficiency of the proposed retina model in terms of an image reconstruction perspective,

the obtained spike outputs are reconstructed as an image. This reconstruction process is performed with more conventional rate-based codes, which are based on the total number of spikes produced by each ganglion cell and measures of their interspike intervals (ISI) [14]. The ISI code is regarded as more robust to neural noise than other spike count- and spike latency-based coding alternatives [14]. For a comparative analysis, reconstructed images for both the proposed adaptive artificial retina model and the standard DOG model are compared using some statistical parameters. In a previous study [17], aside from the histogram similarity ratio (HSR) and universal quality index (UQI) parameters, the mean gray level values were used as a measure for comparison. In the comparisons of this study, aside from the HSR and UQI parameters, the mean squared error (MSE) measure, which is a well-known error comparison parameter, is used. Furthermore, different from the previous study, time-dependent changes in the MSE and UQI parameters are also evaluated for test images.

The rest of the paper is organized as follows. In Section 2, a detailed description of the proposed adaptive artificial retina model is given. Section 3 gives the simulation results and quantitative analysis with spike count- and ISI-based reconstructed images for the proposed and standard models. The discussion and conclusion of the study are presented in Section 4.

2. The proposed spatiotemporal artificial retina model (3D-ADOG)

In the DOG filter-based retina models, the receptive field of a ganglion cell is modeled by the differences of 2 Gaussians: a narrow Gaussian function of high amplitude describes the center, and a broader Gaussian function of lower amplitude describes the surround. The ON-center and OFF-center ganglion cell outputs are easily obtained by changing the polarities of the center and surround parts of the filter. The DOG filter-based receptive field model is shown in Eq. (1).

$$DOG(\sigma_c, \sigma_s) = A_c \cdot \frac{1}{2\pi\sigma_c^2} \cdot e^{\left(-\frac{x^2+y^2}{2\sigma_c^2}\right)} - A_s \cdot \frac{1}{2\pi\sigma_s^2} \cdot e^{\left(-\frac{x^2+y^2}{2\sigma_s^2}\right)} \tag{1}$$

Here, A_c, A_s and σ_c, σ_s are the spatial amplitude and standard deviation values of the center and surround components of the DOG filter, respectively. In this filter profile, for the ON-center receptive field, the maximum output values should be obtained only if the center part of the DOG filter is totally illuminated and the surround part is dark. The maximum output values for the OFF-center receptive field should be obtained only if the center part of the DOG filter is totally dark and the surround is totally illuminated. For uniform input data (or the illumination value), the output of the filter should be near zero. The gains of the DOG filters (A_c and A_s) in the proposed model are set to ensure these considerations. The block diagram of the model is given in Figure 1.

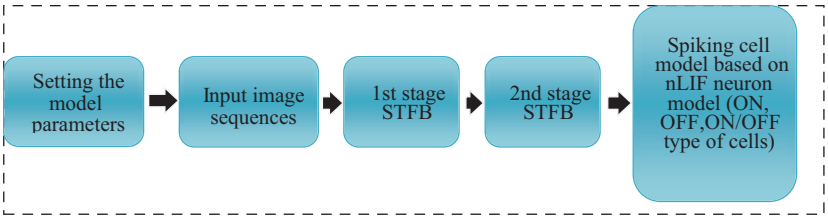

Figure 1. The proposed 2-stage adaptive DOG filter-based artificial retina model.

In the first step of the proposed artificial retina model, color image sequences are converted to their gray level intensity values (I). Consequently, the intensity image (I) is filtered by an adaptive spatiotemporal DOG

filter bank ($STFB_1$) to obtain an output resulting from the first process stage of the model, which contributes the interactions of the receptor, horizontal, and bipolar cells. The time delay between the center (τ_{c1}) and surround parts (τ_{s1}) of the DOG filter is conveniently selected as 0 ms and 5 ms, respectively [17,18]. The spatial center (A_{c1}) and surround (A_{s1}) amplitudes of the DOG filter are set to 3 and 1, respectively. The center bandwidth value of the DOG filter, σ_{c1}, is selected as the standard deviation of the normalized local image data (I_{norm}) and the surround bandwidth value, σ_{s1}, is defined as 6 times bigger than the center bandwidth value, according to Eq. (3). The normalization equation can be seen in Eq. (4).

$$STFB_1(\sigma_{c1},\sigma_{s1},\tau_{c1},\tau_{s1},t) = A_{c1}\cdot\frac{1}{2\pi\sigma_{c1}^2}\cdot e^{\left(-\frac{x^2+y^2}{2\sigma_{c1}^2}\right)}\cdot\delta(t-\tau_{c1}) - A_{s1}\cdot\frac{1}{2\pi\sigma_{s1}^2}\cdot e^{\left(-\frac{x^2+y^2}{2\sigma_{s1}^2}\right)}\cdot\delta(t-\tau_{s1}) \quad (2)$$

$$\sigma_{c1} = std(I_{norm}) \quad and \quad \sigma_{s1} = 6\cdot\sigma_{c1} \quad (3)$$

$$I_{norm} = \frac{I-\min(I)}{\max(I)-\min(I)} \quad (4)$$

Output values for the first stage, Out_{ST1}, are obtained by 2D convolution of the filter bank $STFB_1$ and image data I, according to Eq. (5).

$$Out_{ST1}(x,y,t) = I(x,y,t) * STFB_1(\sigma_{c1},\sigma_{s1},\tau_{c1},\tau_{s1},t) \quad (5)$$

The bandwidths of the DOG filters in the second stage are determined using the output values of the first stage (Eqs. (6) and (7)). Similar to the time delays of the first stage, the time delays of the central signals (τ_{c2}), due to bipolar cells, and the time delays of the surround signals (τ_{s2}), due to amacrine cells, are set to 10 ms and 15 ms, respectively. By considering the biological assumptions, the center and surround amplitudes in this stage, A_{c2} and A_{s2}, of the DOG filters are set to 2 and 1, respectively. The center bandwidth value σ_{c2} is selected as the standard deviation value of the normalized data (Out_{ST1_norm}) from the first stage and the surround bandwidth value σ_{s2} is defined as 6 times bigger than σ_{c2} (Eq. (7)). The normalization equation for the output of the first stage is given in Eq. (8).

$$STFB_2(\sigma_{c2},\sigma_{s2},\tau_{c2},\tau_{s2},t) = A_{c2}\cdot\frac{1}{2\pi\sigma_{c2}^2}\cdot e^{\left(-\frac{x'^2+y'^2}{2\sigma_{c2}^2}\right)}\cdot\delta(t-\tau_{c2}) - A_{s2}\cdot\frac{1}{2\pi\sigma_{s2}^2}\cdot e^{\left(-\frac{x'^2+y'^2}{2\sigma_{s2}^2}\right)}\cdot\delta(t-\tau_{s2}) \quad (6)$$

$$\sigma_{c2} = std(Out_{ST1\_norm}) \quad and \quad \sigma_{s2} = 6\cdot\sigma_{c2} \quad (7)$$

$$Out_{ST1\_norm} = \frac{Out_{ST1}-\min(Out_{ST1})}{\max(Out_{ST1})-\min(Out_{ST1})} \quad (8)$$

The output values of the second stage (Out_{ST2}) are obtained by 2D convolution of the $STBF_2$ filter bank and the Out_{ST1} signal, according to Eq. (9).

$$Out_{ST2}(x,y,t) = Out_{ST1}(x,y,t) * STFB_2(\sigma_{c2},\sigma_{s2},\tau_{c2},\tau_{s2},t) \quad (9)$$

The last step of the model is the calculation of the voltage levels of the ganglion cells (V_G). For the calculation, the values obtained from the second stage's output are regarded as exciting current values for the ganglion cells, and these values are used for the calculation of the voltage levels of the ganglion cells (V_G) for the spike generation process (Eq. (10)).

$$V_G(x,y,t) = nLIF(Out_{ST2}(x,y,t)) \quad (10)$$

The spike generation mechanism of the retinal ganglion cells is modeled using the nLIF model [19]. The nLIF model produces more realistic spike trains than the other well-known Poisson and Hodgkin–Huxley models.

$$\left\{ \begin{array}{ll} V(t) < spk_{thr} & \tau \frac{dV(t)}{dt} = -V(t) + R \cdot I(t) + n(t) \\ V(t) \geq spk_{thr} & V(t) = Vr,\ spike = 1 \end{array} \right\} \tag{11}$$

The basic formulation of the *nLIF* model is given in Eq. (11), where $I(t)$ is the input current, $V(t)$ is the membrane potential, τ is the time constant, and R is the typical resistance of the cell. A random noise source $n(t)$ with an amplitude of 0.1 mV is added to the integration process so that the neural noise can be considered. The typical cell resistance, time constant of the cell, and integration time constant are selected as 10 ohm, 5 ms, and 10 ms, respectively. Depending on the biological constants, the resting potential values (Vr) and the firing threshold values (spk_{thr}) of the cells are selected as –65 mV and –50 mV, respectively [17,18].

In the literature, several types of ON-center and OFF-center ganglion cells (sustained, direction selective, and transient) have been introduced. In this study, sustained types of these cells, which have a more important role for image forming in the visual cortex, ON-center, OFF-center, and ON/OFF-center ganglion cells, are studied.

The internal cell connections in the model are shown in Figure 2. The gray lines represent the inhibitory signals and the black lines represent the excitatory signals. There is a 5-ms time delay between the cell layers. For the OFF-center ganglion output, the bipolar ON-center signal is inverted and applied to the input of the OFF-center ganglion cell. The ON/OFF-center ganglion cell model is implemented as a hybrid cell model working with a dominant signal. If the ON-type input signal applied to this cell model is bigger than the OFF-type input signal, the cell works as an ON-center ganglion cell and vice versa [20,21]. The following section describes the simulation procedure and presents the simulation results.

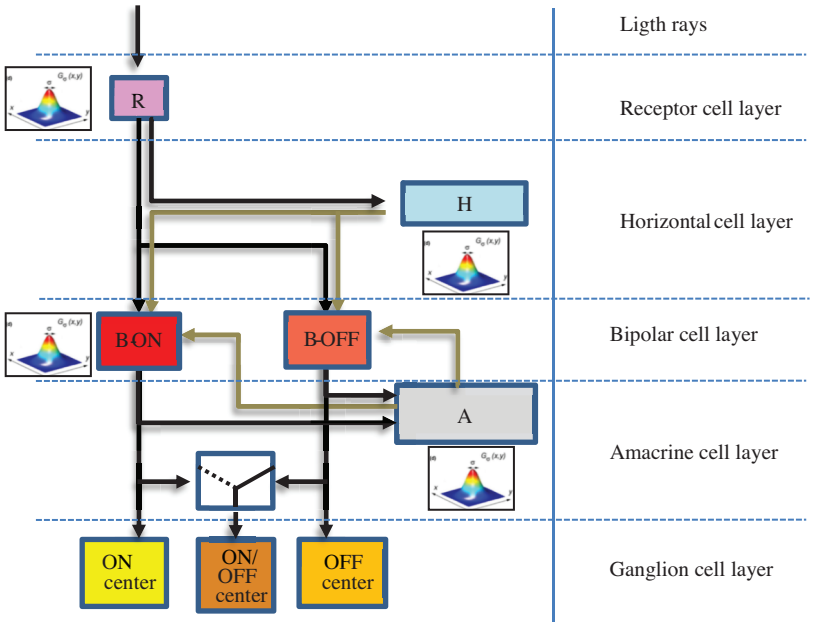

Figure 2. Signal flow between the cells in the model.

3. Simulation results

As an initial step before the simulation study on natural test images, the model is validated using special test patterns that can excite only ON-center and OFF-center ganglion cells. The test patterns used for this purpose are shown in Figure 3. For testing the ON-center ganglion cells, a synthetic pattern that includes white disks on a black background is generated. For OFF-center ganglion cells, an image containing black disks on a white background is used as the test pattern. These patterns are then processed by our artificial retina model, and ISI measure-based reconstructed images are obtained. The reconstruction results for the test patterns are given in Figure 4. From Figure 4, it is seen that the ON-center and OFF-center ganglion cells work in contrast to each other. The resulting reconstructed images are more similar to the original input test images. For the ON-center test pattern, the ON-center ganglion cells show a high firing rate and this corresponds to low ISI values, while the OFF-center ganglion cells show a low firing rate, corresponding to high ISI values (Figure 4a). For the OFF-center ganglion cells, the high firing rate (e.g., low ISI values) is obtained when the OFF-center test pattern is applied to the model (Figure 4b). Since the ISI values are reliable only if all of the ganglion cells produce at least one spike, in this test, all of the ganglion cells emitted at least 5 spikes in 20 ms with a simulation step time of 1 ms. This range of possible firing rates from the ganglion cells of 0 to 200 Hz is reasonable [14].

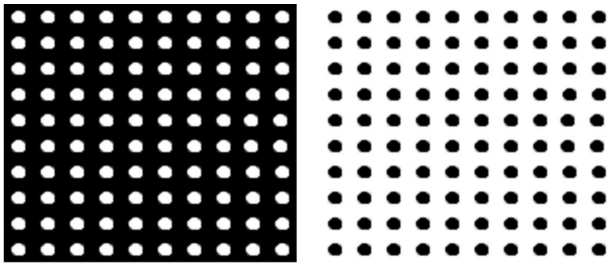

Figure 3. Test patterns for the ON-center (left) and OFF-center (right) ganglion cells.

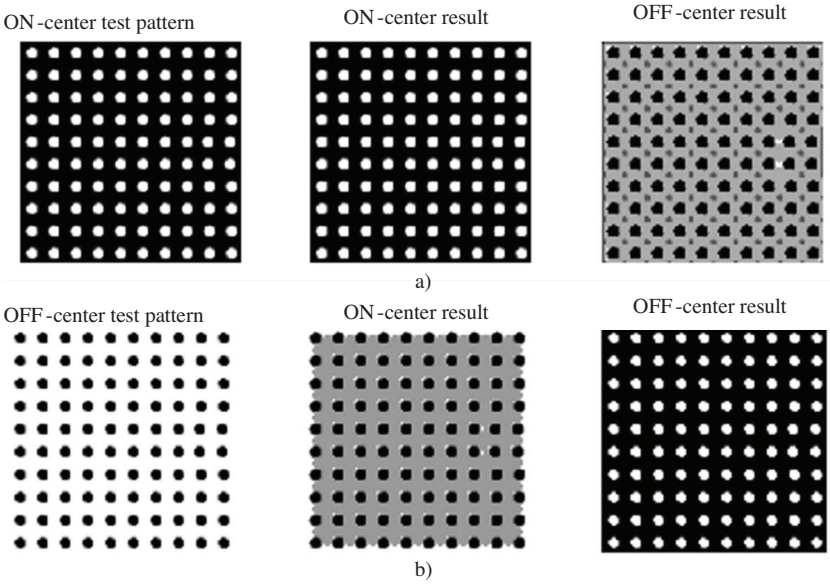

Figure 4. ISI measure-based reconstruction of the resulting spikes for the test patterns: a) ON-center test pattern and b) OFF-center test pattern.

After verification of the model with the test patterns, performance analyses of the model are performed using the natural test images, which have different texture and scene variations. The test images from the Berkeley database [22] used in the simulation study are shown in Figure 5. In these analyses, it is aimed to evaluate the efficiency of the adaptive stages in the model in terms of the image reconstruction quality. Next, in the simulation study, the images are processed with both the standard DOG filter and our proposed 3D-ADOG filter-based retina models. The obtained spike activities for both models are stored to be used in the reconstruction process and the reconstructed images are compared using statistical image comparison parameters. Since the OFF-center ganglion outputs cannot produce a meaningful image, in the comparisons, only ON-center and ON/OFF-center ganglion outputs are considered. In the image reconstruction process, the firing activities of the ganglion cells for the models are stored in N-dimensional matrix form, for a 50-ms time period (Figure 6). The spike counts and ISI measures of each ganglion cell are determined for the whole activity matrix, with a 1-ms time step (it then creates a 50 $\times$ M $\times$ N dimensional matrix, composed of 50 time frames for each test image, with the size of M $\times$ N pixels). Using these spike counts and ISI values, the output images are reconstructed as an image.

a) 175083.jpg b) 334025.jpg c) 48017.jpg

Figure 5. Original gray scale images used in the simulation study (from the Berkeley 'BSDS500' image database). The image dimensions are 241 $\times$ 161, 75 $\times$ 100, and 161 $\times$ 241 pixels, respectively.

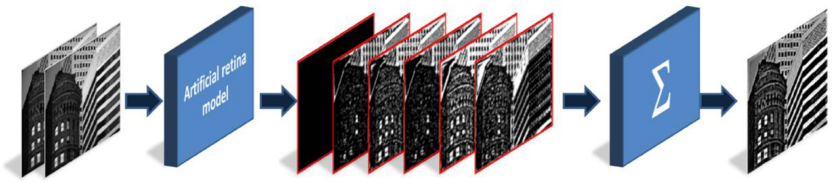

Figure 6. Image reconstruction from the spike trains: a) a frame of image sequence, b) obtained spike matrix sequence (or spike train), and c) the reconstructed image using the spike count- or ISI measure-based reconstruction methods.

In the spike count-based image reconstruction method, defining the total number of spikes of only one ganglion cell requires the counting of all spikes emitted from that ganglion cell during the period of firing activity. After determining the spike counts of all of the ganglion cells, gray tone values are assigned to each spike count and scaled to the original image's gray tone interval [0 $\sim$ 255]. Hence, in the reconstructed image, the lightest gray tone value represents the highest spike count, and the darkest gray tone value represents the lowest spike count. In the ISI measure-based image reconstruction, to measure the ISI value, the calculation of the time between the consecutive 2 spikes is sufficient. In the ISI-based reconstructed image, the lower ISI values are represented as lighter gray tone values and the higher ISI values are represented as darker gray tone

values. The gray tone intervals of the ISI-based reconstructed image are also normalized to the [0 ∼ 255] range. Hence, the image reconstruction process is done using the stored ganglion firing activity matrices [14,23]. The image reconstruction steps, including the input image/video sequence, N-dimensional firing activity matrix, and reconstructed image, are given in Figure 6.

Following the reconstruction step, in order to perform an objective evaluation, the reconstructed images are statistically compared with the original input images with respect to the MSE, UQI, and HSR parameters.

The MSE parameter basically produces the error values based on a pixel-by-pixel difference. For a pixel-frequency distribution-based comparison, the HSR parameter is used. For the HSR calculation, the Bhattacharyya coefficient is suitable. This coefficient is a measure of the distance between 2 normalized image histograms, and its value changes between 0 and 1 [24]. Finally, in the simulation study, the UQI parameter is used to measure the contrast distortion, luminance distortion, and correlation loss between the original and reconstructed images, and its range of output values is [–1, 1] [25].

In the simulation study, the effects of using adaptive or nonadaptive filtering stages in our retina model (3D-ADOG) are also analyzed. For this purpose, 2 versions of the 3D-ADOG model are derived: in the 3D-ADOG-1 model, only the first stage is adaptive and in the 3D-ADOG-2 model, only the second stage is adaptive. The simulation parameters for the artificial retina models are given in Table 1. The main difference between the models is whether the stages are adaptive or not. The other parameters are the same for each model. For some cases, the standard deviations of the smooth regions are nearly zero; the bandwidths of the filters are fixed as 0.5, similar to the standard DOG filter-based model. Using the parameters in Table 1, the artificial retina models are compared by considering the spike count-based and ISI measure-based reconstruction performances. In the following subsections, the simulation results are evaluated in terms of the spike count-based reconstruction performance, ISI measure-based reconstruction performance, and time-dependent performances for these 2 reconstruction methods, respectively.

Table 1. Artificial retina model parameters in the simulation study.

Artificial retina model parameters	3D-ADOG model	3D-ADOG-1 model	3D-ADOG-2 model	Standard DOG-based model
Spatial bandwidth of the DOG filters in the first stage	15 × 15 pixels	15 × 15 pixels	15 × 15 pixels	15 × 15 pixels
Spatial bandwidth of the DOG filters in the second stage	9 × 9 pixels	9 × 9 pixels	9 × 9 pixels	9 × 9 pixels
Center bandwidth of the DOG filters in the first stage	Adaptive (0 ∼ 1)	Adaptive (0 ∼ 1)	0.5 (not adaptive)	0.5 (not adaptive)
Center bandwidth of the DOG filters in the second stage	Adaptive (0 ∼ 1)	0.5 (not adaptive)	Adaptive (0 ∼ 1)	0.5 (not adaptive)
Simulation step time	1 ms	1 ms	1 ms	1 ms
Total simulation time	50 ms	50 ms	50 ms	50 ms
nLIF parameters	Same	Same	Same	Same

3.1. Spike count-based reconstruction performance

In the first phase of the simulation study, the spike trains obtained from the retina models are used in the spike count-based image reconstruction step. The reconstructed images, which are created using all of the spike data for each model, are shown in Figures 7–9. In these images, it is seen that when both stages are adaptive, the spatial details in the image are better described, whereas the other models present nearly the same blurry image quality. In addition, the resulting images for the 3D-ADOG-1 and 3D-ADOG-2 models include less blur distortion effect than the standard DOG filter-based model for both the ON and ON/OFF channels. However, some of the gray tone values in the reconstructed image for the 3D-ADOG model do not exactly match with some of the pixel values in the original image (especially for the white region in the '334025.jpg' image) due to adaptive stages, and the spatial details become more clear for visual perception for both the ON and ON/OFF channels. The spatial details in the '48017.jpg' image in Figure 7, the texture of the salamander and the spatial details of the leaf in the '175083.jpg' image in Figure 8, and the texture of the wall in the '334025.jpg' image in Figure 9 are obtained as closer to the original images than with the other models. Moreover, for all of the test images, only the 3D-ADOG model produces a meaningful edge-like image for the OFF channel.

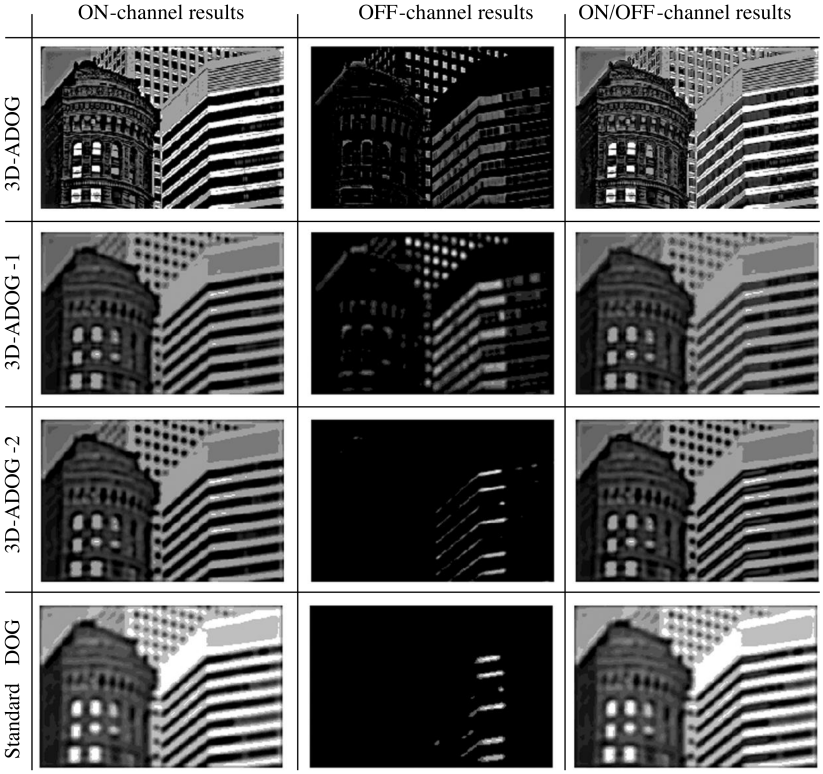

Figure 7. Spike count-based image reconstruction results for test image '48017.jpg'. Results for the 3D-ADOG, 3D-ADOG-1, 3D-ADOG-2, and standard DOG filter-based retina models (top to bottom). Reconstruction results for the ON channel, OFF channel, and ON/OFF channel, respectively (left to right).

The quantitative results for the models with respect to the 3 statistical parameters of the MSE, HSR, and UQI are presented in Table 2, where the values show the averaged values for all of the test images. As can be seen from Table 2, a significant improvement in terms of the MSE, HSR, and UQI parameters is obtained with the 3D-ADOG model for the ON channel (36.9% lower MSE value, 22.3% higher HSR, and 10.2% higher

UQI value) and for the ON/OFF channel (37.3% lower MSE value, 26.4% higher HSR, and 9.5% higher UQI value). Similarly, the other 2 versions of our model (3D-ADOG-1 and 3D-ADOG-2) yield slightly worse results for the HSR and UQI parameters and yield relatively better results for the MSE parameter for the ON and ON/OFF channels.

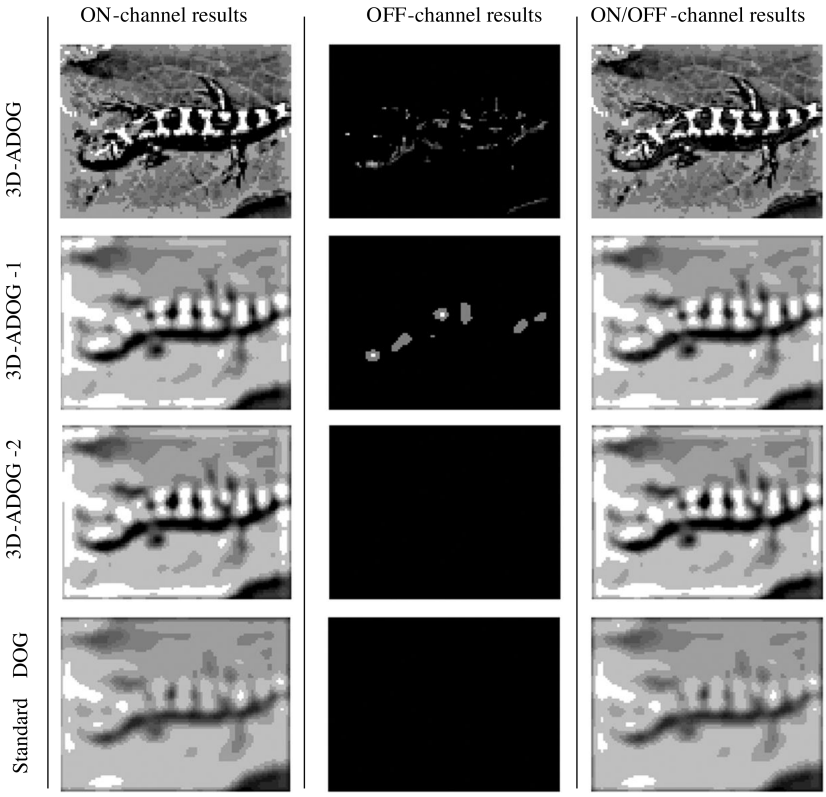

Figure 8. Spike count-based image reconstruction results for test image '175083.jpg'. Results for the 3D-ADOG, 3D-ADOG-1, 3D-ADOG-2, and standard DOG filter-based retina models (top to bottom). Reconstruction results for the ON channel, OFF channel, and ON/OFF channel, respectively (left to right).

Table 2. Quantitative comparison of the models based on the spike count-based reconstruction performance. The results are the average values over the results for all of the test images.

Artificial retina model	Comparison parameters for ON-center outputs			Comparison parameters for ON/OFF-center outputs		
	MSE	HSR	UQI	MSE	HSR	UQI
3D-ADOG	1440.0	0.722	0.653	1429.3	0.762	0.646
3D-ADOG-1	2191.7	0.549	0.492	2016.7	0.599	0.486
3D-ADOG-2	2269.2	0.547	0.482	2267.3	0.554	0.482
Standard DOG	2280.4	0.499	0.449	2278.9	0.499	0.456

3.2. ISI measure-based reconstruction performance

Aside from the spike count-based reconstruction, in the simulation study, the ISI measure-based reconstruction process is also performed. Because the interval between 2 spikes is an analog value, the ISI measure-based coding is potentially more accurate than the spike count-based image coding [14]. However, the time between

only 2 spikes is sufficient for calculating the interspike interval; similar to the count-based reconstruction, for each ganglion cell, the mean ISI value is calculated using the whole firing activity. Due to the characteristic of the ISI measure, it is expected that the gray tone values of the ISI-based reconstructed image may have a wider range. The ISI measure-based image reconstruction results for each retina model are given in Figures 10–12, where it can be seen that almost the same results in the image reconstruction are obtained for the ON, OFF, and ON/OFF channels. In these images, the gray tone levels are slightly lighter compared to the spike count-based reconstructed images. Similar to the previous results, in this analysis, more spatial details are preserved in the reconstructed images for the 3D-ADOG model.

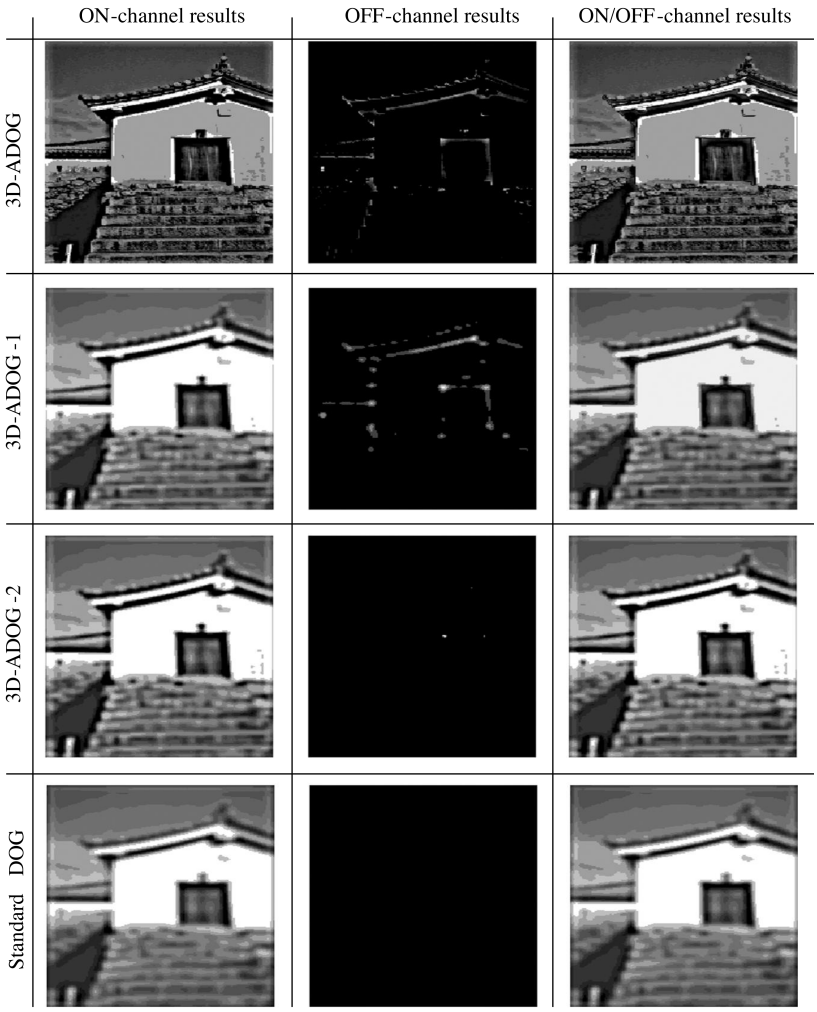

Figure 9. Spike count-based image reconstruction results for test image '334025.jpg'. Results for the 3D-ADOG, 3D-ADOG-1, 3D-ADOG-2, and standard DOG filter-based retina models (top to bottom). Reconstruction results for the ON channel, OFF channel, and ON/OFF channel, respectively (left to right).

The statistical evaluations of the models according to the MSE, UQI, and HSR parameters are presented in Table 3, which includes the averaged values for all of the test images. While the visual quality of the images is highly similar to the spike count-based reconstruction results, for the MSE parameter, the standard DOG filter-based model produces lower values for the ON and ON/OFF channels compared to the other models

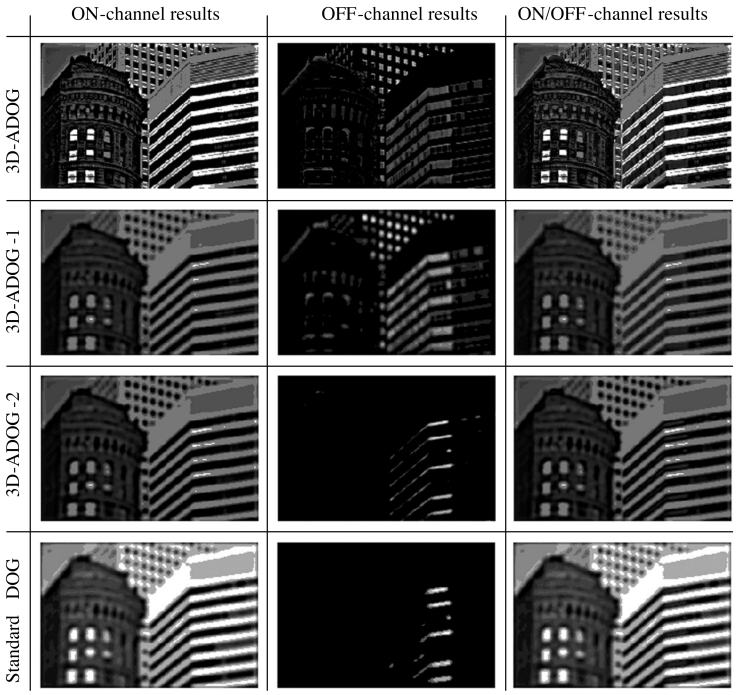

Figure 10. ISI measure-based image reconstruction results for test image '48017.jpg'. Results for the 3D-ADOG, 3D-ADOG-1, 3D-ADOG-2, and standard DOG filter-based retina models (top to bottom). Reconstruction results for the ON channel, OFF channel, and ON/OFF channel, respectively (left to right).

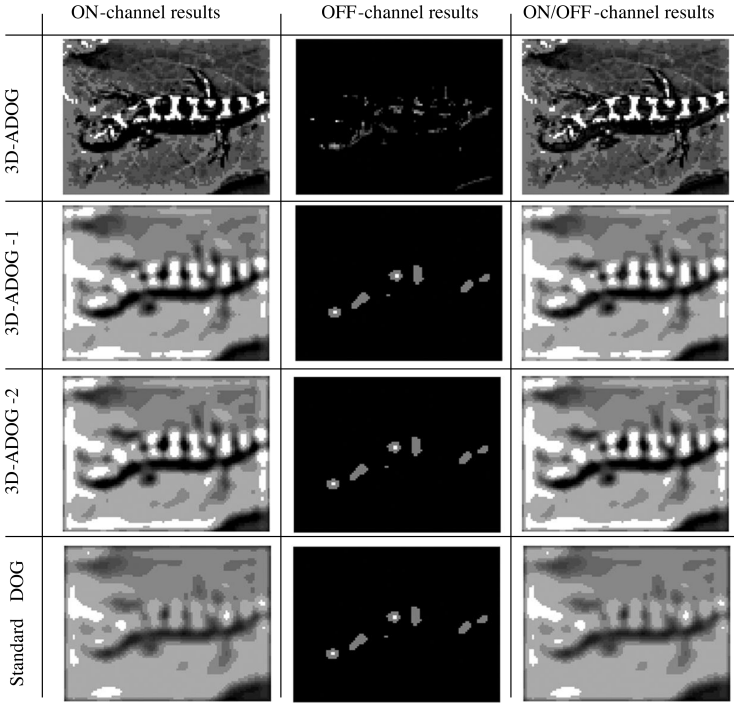

Figure 11. ISI measure-based image reconstruction results for test image '175083.jpg'. Results for the 3D-ADOG, 3D-ADOG-1, 3D-ADOG-2, and standard DOG filter-based retina models (top to bottom). Reconstruction results for the ON channel, OFF channel, and ON/OFF channel, respectively (left to right).

(29.0% and 21.4% lower MSE values for the ON and ON/OFF channels, respectively). This may be caused by slight changes in the gray level values in the images. Between the other models, the 3D-ADOG-1 model yields lower values for the ON and ON/OFF channels. For the HSR and UQI parameters, the 3D-ADOG model yields better results for the ON channel (10.0% higher HSR and 11.0% higher UQI value) and the ON/OFF channel (13.0% higher HSR and 10.5% higher UQI value).

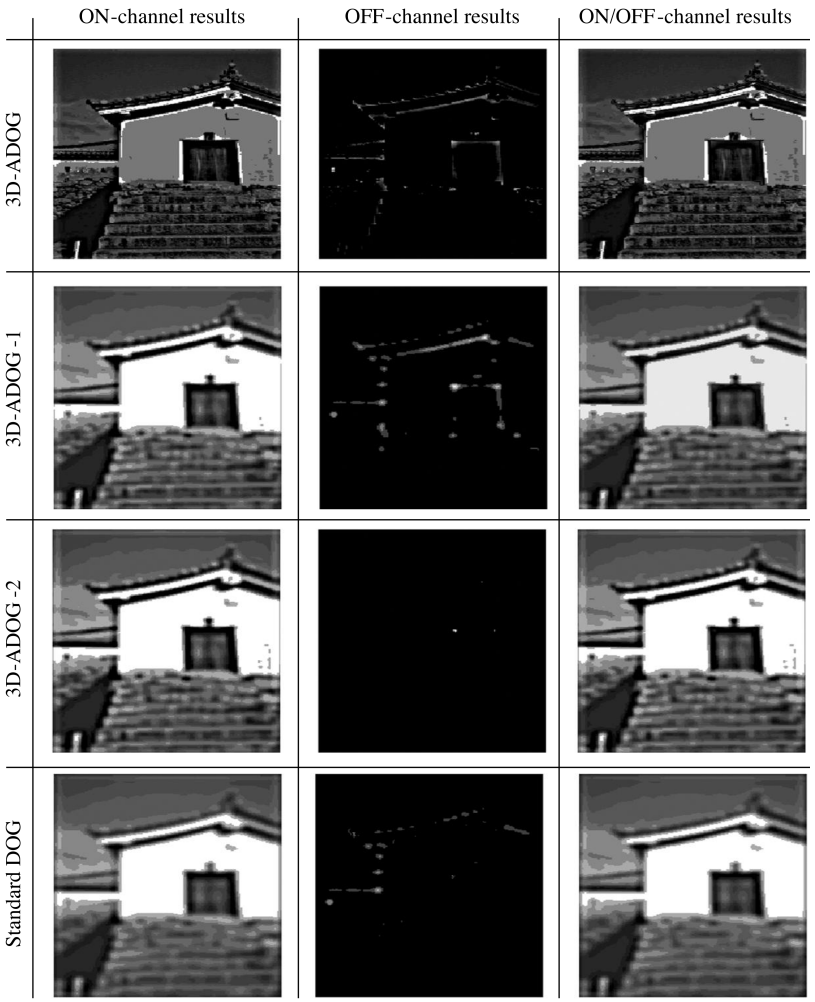

Figure 12. ISI measure-based image reconstruction results for test image '334025.jpg'. Results for the 3D-ADOG, 3D-ADOG-1, 3D-ADOG-2, and standard DOG filter-based retina models (top to bottom). Reconstruction results for the ON channel, OFF channel, and ON/OFF channel, respectively (left to right).

3.3. Reconstruction quality analysis by time

In order to evaluate the time-dependent reconstruction quality of the proposed model, the MSE, HSR, and UQI parameters are calculated at determined time intervals for both the spike count- and ISI measure-based reconstruction methods. For the reconstruction time periods, 10-ms time intervals are selected. Hence, parts of the whole spike activity at 10 ms, 20 ms, 30 ms, 40 ms, and 50 ms are used in the image reconstruction process. This process is shown in the Figure 13, where for the '48017.jpg' image, the image reconstruction results through the 10-ms time intervals are given for the 3D-ADOG model.

Table 3. Quantitative comparison of the models according to the ISI measure-based reconstruction performance. The results are the average values over the results for all of the test images.

Artificial retina model	Comparison parameters for ON-center outputs			Comparison parameters for ON/OFF-center outputs		
	MSE	HSR	UQI	MSE	HSR	UQI
3D-ADOG	2343.2	0.662	0.677	2311.2	0.696	0.669
3D-ADOG-1	1956.8	0.570	0.485	1780.5	0.614	0.483
3D-ADOG-2	2062.4	0.565	0.473	2060.5	0.570	0.473
Standard DOG	1399.0	0.563	0.460	1399.1	0.566	0.460

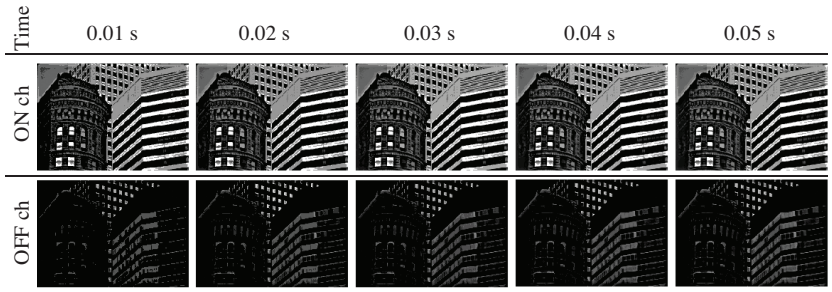

Figure 13. Time-dependent image reconstruction process at 10-ms time intervals. Top row: image reconstruction for the ON channel. Bottom row: image reconstruction for the OFF channel.

For the ON-center and ON/OFF-center ganglion outputs, a series of graphics that show the comparison results are given in Figures 14–16, where changes in the MSE, HSR, and UQI parameters according to time periods can be seen. The bar values in the graphics correspond to the average values of the MSE, HSR, and UQI parameters, which are obtained and averaged for the test images. The error bars also indicate the maximum and minimum values for the MSE, HSR, and UQI parameters.

The graphics including the time-dependent change of the MSE parameter for the ON and ON/OFF channels are given in Figure 14. Based on the spike count-based reconstruction (Figure 14a) for the ON and ON/OFF channels, as expected, the MSE values for the 3D-ADOG model decrease with time and the lowest MSE values are obtained for all of the time periods. The DOG model and the other 2 models (3D-ADOG-1 and 3D-ADOG-2) yield lower MSE values at 30 ms of reconstruction and become constant after this time period. In the ISI measure-based reconstruction results (Figure 14b), for the ON and ON/OFF channels, the 3D-ADOG model produces lower MSE values until the 40-ms time period and the other models produce low MSE values at the end of the reconstruction process, and the standard DOG filter-based model yields better results for this comparison. However, it should be noted that the error ranges of the 3D-ADOG model are very small compared to those of the standard DOG model.

The graphics including the analysis of the time-dependent change of the HSR parameter for the ON and ON/OFF channels are given in Figure 15. For the spike count-based reconstruction (Figure 15a), the 3D-ADOG model yields increasing HSR values until 40 ms of reconstruction for both the ON and ON/OFF channels. After that time, no significant increase is observed. As it can be seen from this graphic, similar to the 3D-ADOG model, the other models show an increasing trend with time. For the ISI measure-based reconstruction (Figure 15b), the HSR values for the 3D-ADOG model increase with time until 30 ms of reconstruction for both the ON

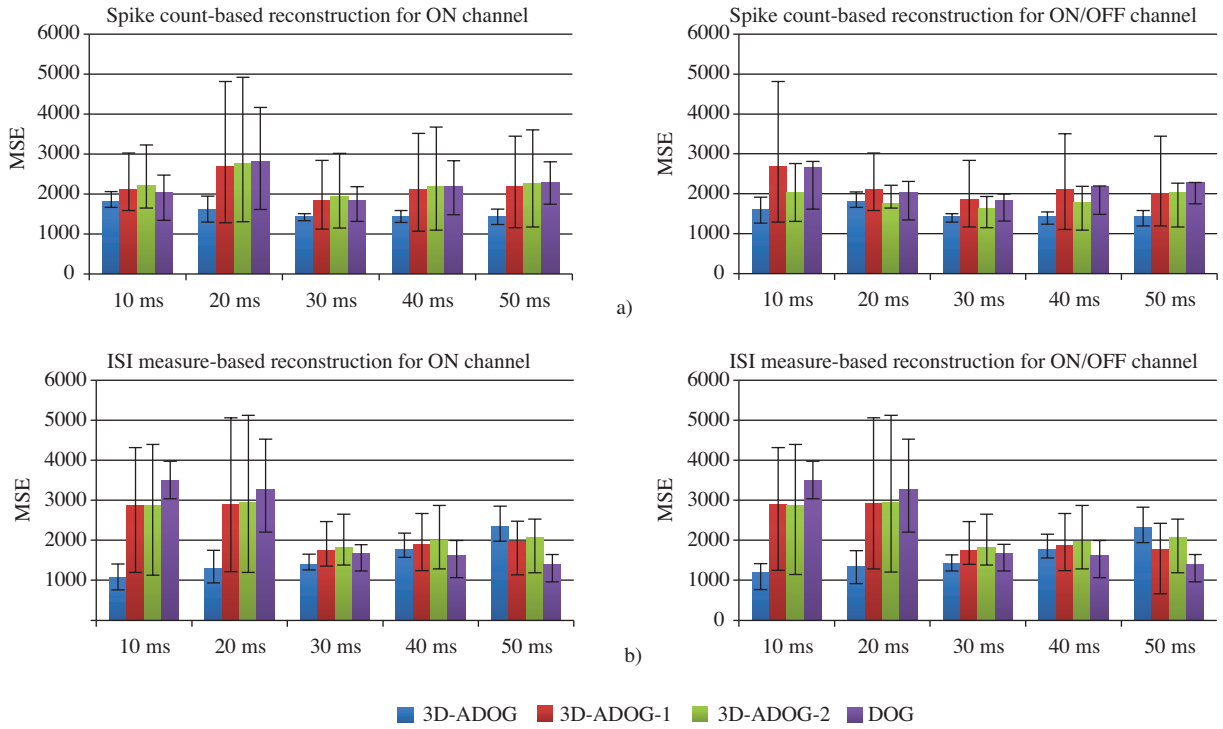

Figure 14. Time-dependent graphical analyses of the MSE values for the ON and ON/OFF channels: a) graphics for spike count-based reconstruction and b) graphics for the ISI measure-based reconstruction.

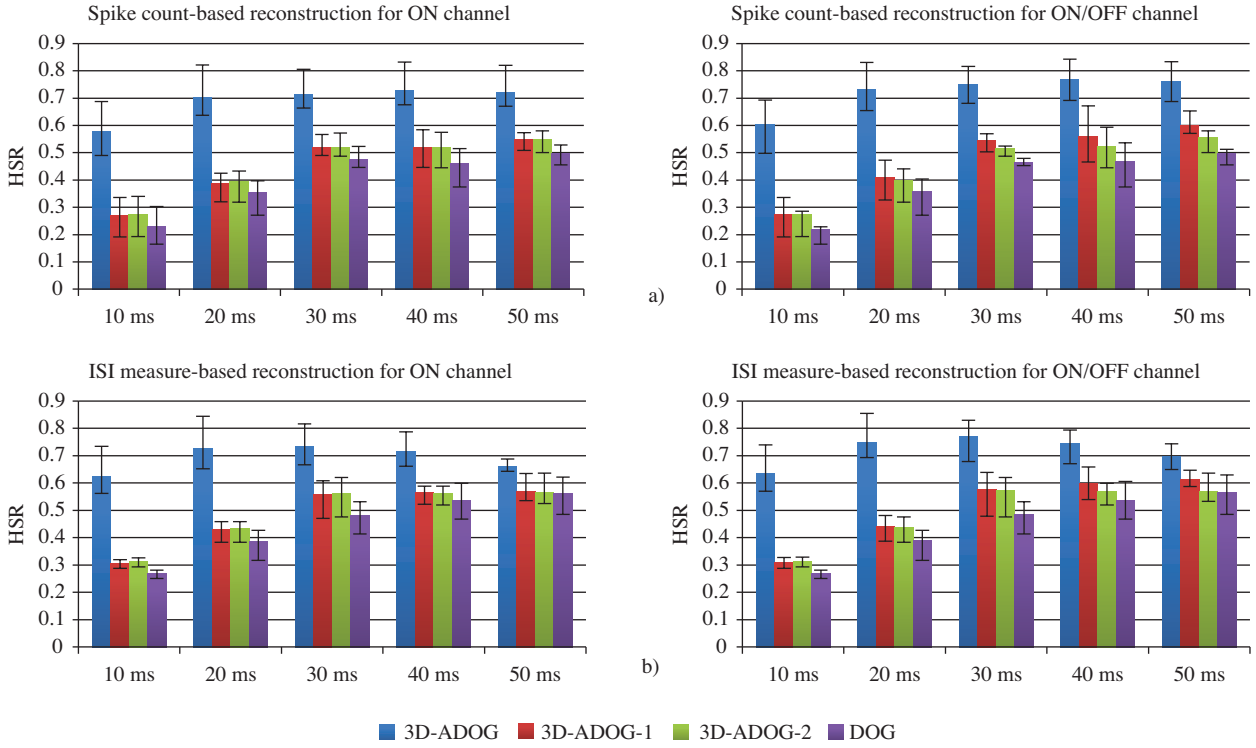

Figure 15. Time-dependent graphical analyses of the HSR values for the ON and ON/OFF channels: a) graphics for spike count-based reconstruction and b) graphics for the ISI measure-based reconstruction.

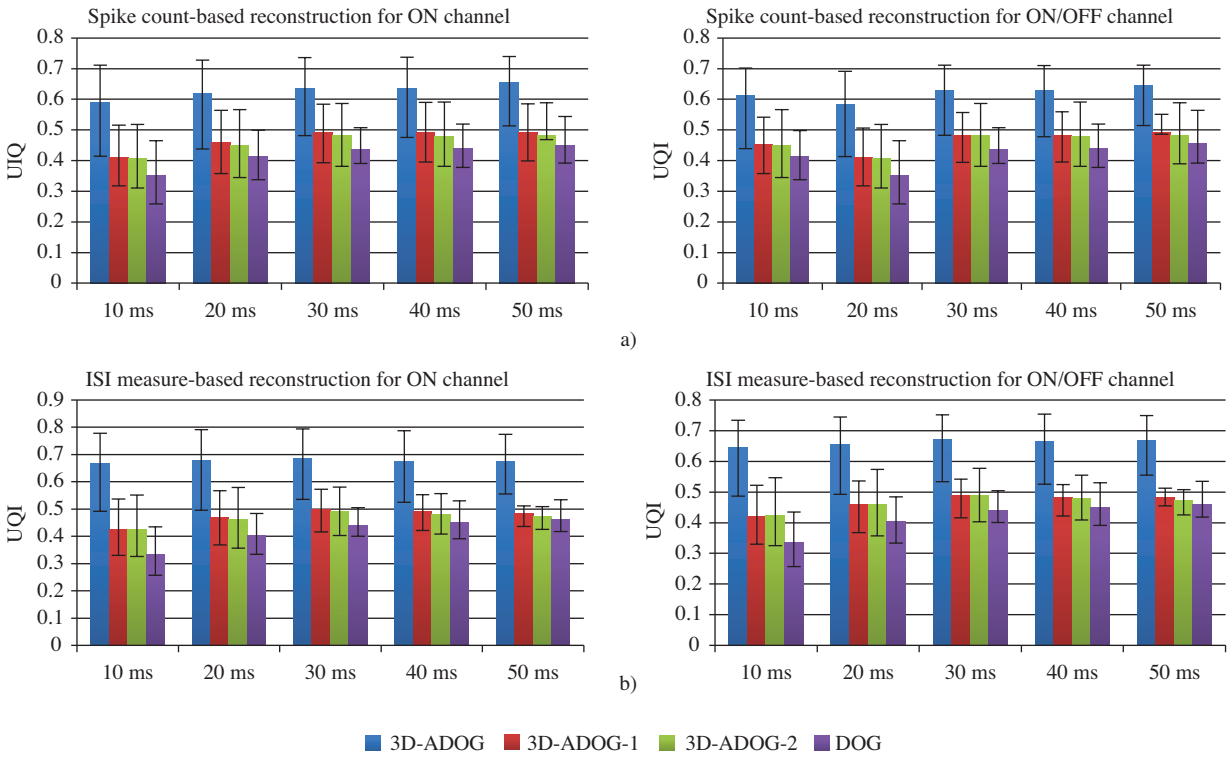

Figure 16. Time-dependent graphical analyses of the UQI values for the ON and ON/OFF channels: a) graphics for spike count-based reconstruction and b) graphics for the ISI measure-based reconstruction.

and ON/OFF channels. In some cases, although the maximum and minimum levels for the 3D-ADOG model are lower than the other models, the minimum HSR levels for this model are relatively higher than the HSR values of the other 3 models. However, the HSR values increase when many more spike trains are included in the reconstruction process, and the proposed model yields significantly high HSR values, even if less of the spike train is used (for example, results for the first 10 ms of reconstruction). The standard DOG filter-based model yields lower HSR values than the 3D-ADOG model, as can be seen from all of the graphics in Figure 15.

The time-dependent graphical analyses of the UQI parameter for the ON and ON/OFF channels are given in Figure 16. In this comparison, for both reconstruction methods, nearly the same UQI values are obtained in the simulation study (Figures 16a and 16b). The 3D-ADOG model yields better scores for all of the time periods of reconstruction, and lower UQI values are obtained for the standard DOG-based retina model.

As a result, the overall performance of the proposed 3D-ADOG model is much higher for the MSE parameter and is significantly high for the HSR and UQI parameters, even in the early times of reconstruction (10-ms, 20-ms, and 30-ms time periods). This means that the proposed model can show a better image reconstruction performance in a short time period. This also helps to increase the benefits of the visual prosthesis systems.

4. Discussion and conclusion

In this study, by considering a biologic model of the retina layer, a bioinspired 3D-ADOG was developed. This model is based on the 2-layered (inner and outer plexiform layers) structure and adaptive processing behavior of the biological retina. Both stages in the model include adaptive spatiotemporal filtering stages, which are

not included in the classical DOG filter-based retina model. In order to generate the proper spike activities of the retinal ganglion cells, the nLIF-based spiking cell model was used.

In this study, the aim was to achieve a higher visual perception performance for retina implant users via the developed model. The qualitative and quantitative performance of the model was evaluated by performing simulation studies. In the simulation studies, spike count- and ISI measure-based image reconstruction methods were used for the image reconstruction process. In the quantitative analyses, aside from the well-known classical MSE measure, the reconstructed images were compared using HSR and UQI parameters. Since the HSR parameter is important for analyzing the contrast range of the image, this parameter was used to measure the similarity of the gray tone levels in the reconstructed images. The UQI is a compact parameter for measuring contrast distortion, luminance distortion, and correlation loss of images. Using these 3 evaluation parameters, appearance-based, content-based, and compact statistical comparison-based performance analyses were performed.

When the spike count- and ISI-based reconstructed images were compared, it was seen that the ISI-based reconstructed images for all of the retina models include a larger number of different gray tone levels than the spike count-based reconstructed images. Thus, the HSR values were slightly lower for the ISI measure-based reconstruction than for the spike count-based reconstruction. Although it is possible to employ other encoding strategies (for example, rank order coding [14]) in the image reconstruction stage of this study, the well-known spike count- and ISI-based methods are regarded as reliable for verifying the relative comparisons of the models.

Simulation results based on the spike count showed that the proposed model yields better scores for the MSE, HSR, and UQI parameters for the ON and ON/OFF channels than the standard DOG filter-based retina model (Table 2). The ISI measure-based simulation results showed that our model yields higher values for the HSR and UQI (Table 3). Although the 3D-ADOG model preserves the spatial details in the images better, for the ISI-based reconstruction results for both the ON and ON/OFF channels, this model yielded relatively higher MSE values than the standard model. As the MSE measure depends on the difference of the pixel values in a $[0 \sim 255]$ interval, unimportant changes (e.g., small changes in the light conditions) for visual perception may cause important error values for this measure. Therefore, it can be considered that the HSR and UQI measures may coincide with visual perception because these measures have a relatively small variation in range.

It should be noted that the proposed model provides higher HSR and UQI values and a lower MSE value than the standard DOG-based model according to the time-dependent analysis of the MSE, HSR, and UQI values. This means that, using this model, it can be possible to create a meaningful visual image in a patient's mind with a short duration of electrical stimulation.

The proposed model shows a significant improvement compared to the standard model in the simulation results, but we should keep in mind that the real performance of the model can be evaluated only by clinical trials with real implant users. Qualitative and quantitative results in the study show that the adaptive behavior of the retina can be realized by implementing the 2-stage adaptive bandwidth-DOG filters and the quality of the obtained images from the retina implant devices can be improved using this adaptive retina model, which preserves the spatial image details well.

Acknowledgment

This work was supported by a grant from the Scientific and Technological Research Council of Turkey (TÜBİTAK, Grant 110E077).

References

[1] M. Matthaei, O. Zeitz, M. Keserü, L. Wagenfeld, R. Hornig, N. Post, G. Richard, "Progress in the development of vision prostheses", Ophthalmologica, Vol. 225, pp. 187–192, 2011.

[2] D. Weiland, M.S. Humayun, "Visual prosthesis", Proceedings of the IEEE, Vol. 96, pp. 1076–1084, 2008.

[3] E. Margalit, M. Maia, J.D. Weiland, R.J. Greenberg, M.D. Gildo, Y. Fujii, G. Torres, D.V. Piyathaisere, T.M. O'Hearn, W. Liu, G. Lazzi, G. Dagnelie, D.A. Scribner, E. de Juan Jr, M.S. Humayun, "Retinal prosthesis for the blind", Survey of Ophthalmology, Vol. 47, pp. 335–356, 2002.

[4] K.A. Zaghloul, K. Boahen, "Circuit designs that model the properties of the outer and inner retina", Ophthalmology Research: Visual Prosthesis and Ophthalmic Devices, pp. 135–159, 2007.

[5] C.A. Morillas, S.F. Romero, A. Martinez, F.J. Pelayo, E. Rosa, E. Fernandez, "A design framework to model retinas", BioSystems, Vol. 87, pp. 156–163, 2007.

[6] R. Eckmiller, D. Neumann, O. Baruth, "Tunable retina encoders for retina implants: why and how", Journal of Neural Engineering, Vol. 2, pp. 91–104, 2005.

[7] H. Wei, X. Guan, "The simulation of early vision in biological retina and analysis on its performance", Congress on Image and Signal Processing, Vol. 4, pp. 413–418, 2008.

[8] J. Liu, X. Gou, "Information processing model of artificial vision prosthesis", 2nd International Conference on Concrete Engineering and Technology, Vol. 2, pp. 551–555, 2010.

[9] A. Wohrer, P. Kornprobst, "Virtual retina: a biological retina model and simulator with contrast gain control", Journal of Computational Neuroscience, Vol. 26, pp. 219–249, 2009.

[10] M.S. Humayun, E. de Juan Jr, J.D. Weiland, G. Dagnelie, S. Katona, R. Greenberg, S. Suzuki, "Pattern electrical stimulation of the human retina", Vision Research, Vol. 39, pp. 2569–2576, 1999.

[11] D. Balya, I. Petras, T. Roska, "Implementing the multilayer retinal model on the complex-cell CNN-UM chip prototype", International Journal of Bifurcation and Chaos, Vol. 14, pp. 427–451, 2004.

[12] C.F. Cai, P.J. Liang, P.M. Zhang, "A simulation study on the encoding mechanism of retinal ganglion cell", Lecture Notes in Computer Science, Vol. 4689, pp. 470–479, 2007.

[13] L.J. Croner, E. Kaplan, "Receptive field of P and M ganglion cells across the primate retina", Vision Research, Vol. 35, pp. 7–24, 1995.

[14] R.V. Rullen, S.J. Thorpe, "Rate coding versus temporal order coding: what the retinal ganglion cells tell the visual cortex", Neural Computation, Vol. 13, pp. 1255–1283, 2001.

[15] J.W. Pillow, J. Shlens, L. Paninski, A. Sher, A.M. Litke, E.J. Chichilnisky, E.P. Simoncelli, "Spatio-temporal correlations and visual signaling in a complete neuronal population", Nature, Vol. 454, pp. 995–999, 2008.

[16] J.L. Gauthier, G.D. Field, A. Sher, M. Greschner, J. Shlens, A.M. Litke, "Receptive fields in primate retina are coordinated to sample visual space more uniformly", PLoS Biology, Vol. 7, pp. 1–9, 2009.

[17] I. Karagoz, M. Ozden, "Adaptive artificial retina model to improve perception quality of retina implant recipients", 4th International Conference on BioMedical Engineering and Informatics, pp. 91–95, 2011.

[18] I. Karagoz, M. Ozden, G. Sobaci, "Multi stage local adaptive DOG filter based retina model developed for visual prosthesis system and simulation results", Association for Research in Vision and Ophthalmology Annual Meeting, 2012.

[19] L.F. Abbott, "Lapique's introduction of the integrate-and-fire model neuron (1907)", Brain Research Bulletin, Vol. 50, pp. 303–304, 1999.

[20] A. Thiel, M. Greschner, J. Ammermuller, "The temporal structure of transient ON/OFF ganglion cell responses and its relation to intra-retinal processing", Journal of Computational Neuroscience, Vol. 21, pp. 131–151, 2006.

[21] M.N. Geffen, S.E.J. Vries, M. Meister, "Retinal ganglion cells can rapidly change polarity from OFF to ON", PLoS Biology, Vol. 5, pp. 640–651, 2007.

[22] Berkeley Computer Vision Group, "Contour detection and image segmentation resources Berkeley segmentation data set and benchmarks 500 (BSDS500)", available at http://www.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/BSR/BSR_bsds500.tgz, last accessed 20 July 2012.

[23] T. Gollisch, M. Meister, "Rapid neural coding in the retina with relative spike latencies", Science, Vol. 319, pp. 1108–1112, 2008.

[24] A. Bhattacharyya, "On a measure of divergence between two statistical populations defined by their probability distributions", Bulletin of the Calcutta Mathematical Society, Vol. 35, pp. 99–109, 1943.

[25] Z. Wang, A.C. Bovik, "A universal image quality index", IEEE Signal Processing Letters, Vol. 9, pp. 81–84, 2002.