

# Objects Movement Detection Based on Impulsed Neural Network of the Analogue of the Biological Retinal Model

Ihar Yeuseyenka<sup>1</sup>, & Ihar Melnikau<sup>2</sup>

<sup>1</sup> Research Center of Automotive Power Train and Electronic Control, School of Automotive Engineering, Hubei University of Automotive Technology, Shiyan, China

<sup>2</sup> Department “Asset Performance and Information Management,” AVEVA Solutions Ltd., Cambridge, the United Kingdom

Correspondence: Ihar Yeuseyenka, Research Center of Automotive Power Train and Electronic Control, School of Automotive Engineering, Hubei University of Automotive Technology, Shiyan, China. Tel: 1-777-892-9837

Received: July 20, 2022

Accepted: August 22, 2022

Online Published: August 24, 2022

doi:10.5539/cis.v15n4p46

URL: <https://doi.org/10.5539/cis.v15n4p46>

## Abstract

This paper describes a neural network model based on the biological analog of the impulse neural network of the retina, which makes it possible to identify movement objects in a video image and a motion detector based on the retinal operation. The proposed detector is an alternative to detectors based on deterministic methods and traditional neural networks. It requires less computational resources at the same video image processing speed.

**Keywords:** motion detector, impulse neural network, neuron model, movement object detection

## 1. Introduction

Automated selection and recognition of moving objects is a very promising area of research in the field of digital video processing and pattern recognition [Cheng, Y., et al., 2015] [Chakrabarti, I., Srinivasarao Batta, K. & Chatterjee, K., 2015] [Beyerer, J., Puente L. & Frese F., 2016]. It makes it possible, without human intervention, to control the technological process, record traffic rules violations, carry out automated information collection and analysis, etc. Research is also relevant because video surveillance systems are increasingly being used in many countries.

In large cities, for example, more and more video surveillance systems are being deployed on roads to monitor traffic. At the same time, there is a need for automated collection and analysis of information on traffic flows in major cities in order to respond quickly to changes in the road environment (e.g., changes in traffic light modes due to increased traffic density). In other words, there is a need for automated systems of estimating the traffic density and its composition [Sadhukhan, P. & Gazi, F., 2018] [Sudan, J., Changho, S., Eunmok, Y. & Gyanendra, P., 2021]. Such systems may use closed-circuit television as a source of information, but it is necessary to implement methods that allow, first, movement objects to be identified on video and, second, to be recognized as vehicles of a certain class, and do it quickly enough without requiring a lot of computing resources.

Before you can directly recognize moving objects, you must extract them from a static or dynamic background [Garcia-Garcia, B., Bouwmans T. & Silva A., 2020] [Sobral, A., 2014]. For this, there are a number of deterministic foreground extraction methods: background subtraction methods, time difference methods, optical flow methods, etc. The choice of method greatly affects the efficiency of the entire recognition system [Cheng, Y., et al., 2015] [Chakrabarti, I., Srinivasarao Batta, K. & Chatterjee, K., 2015] [Beyerer, J., Puente L. & Frese F., 2016] [Sudan, J., Changho, S., Eunmok, Y. & Gyanendra, P., 2021] [Menter, Z., Tee, W. & Dave, R., 2021] [Purohit, S., Singh Jat, D., Poonia, R., Kumar, S. & Hiranwal, S., 2021]. And the more efficient the method, the more complex it is, and requires more resources.

The purpose of this work is to present a fundamentally different approach to digital video processing, in particular to the separation of moving objects based on impulse neural networks. This will make it possible to create an effective motion detector that can be an alternative to detectors developed on the basis of deterministic methods and traditional neural networks [Yeuseyenka, I.A., Melnikau I.I. & Yemelyanov, I. A., 2022] [Sobral, A., 2014] [Mandal, M. & Vipparthi, S., 2021] [Ammar, S., Bouwmans, T., Zaghden, N., Neji, M., 2019] [Minematsu, T., Shimada, A. & Taniguchi, R., 2020] [Giraldo, J. & Bouwmans, T., 2020] [Giraldo, J., Javed, S. & Bouwmans,

T., 2022], including an improved way of parallel processing video segments [Yeuseyenko, I.A. & Melnikau, I.I., 2022].

## 2. Related Works

Early moving objects detection is based on the background subtraction methods and noise minimization on the image [Maddalena, L. & Petrosino, A., 2008/2012/2014]. P. Gil-Jimenez in his research [Gil-Jimenez, P., Maldonado-Bascon, R. & Gil-Pita, H., 2003] proposed using the classification capacity of a neural network for decreasing the false detection probability. Instead of thresholding the difference between the current frame and the reference, as it is the typical method, the system first classifies each zone of the image depending on its observed behavior, and so, perform the motion detection according to this classification. The main drawback of the proposed method is the high computational cost of the classification for every pixel. Moreover, the background behavior is supposed not to change frequently.

Z. Wang developed a new cooperative background model for multi-modal video surveillance based on a probability neural network [Wang, Z., Bao, H. & Zhang, L., 2009] [Wang, Z. & Bao, H., 2011]. Firstly, the probability of being foreground is estimated in the visible and infrared channel, and post-processed separately. Then, every pixel is classified into foreground, background. The change pixels by fusing this information, and foreground pixels are segmented into motion regions. Thirdly, the adaptive learning rate is computed for every frame and every pixel based on frame motion difference and pixel classification result, and the background model for every channel is updated.

In 2006 D. Culibrk proposed a novel neural network based on an approach to background modeling for motion-based object segmentation in video sequences [Culibrk, D., Marques, O., Socek, D., Kalva, H. & Furht, B., 2006]. The proposed approach is designed to enable efficient, highly-parallelized hardware implementation. Such a system can achieve real-time segmentation of high-resolution sequences. The basis of the approach is the employment of a novel neural network structure designed specifically to serve as a model of background in video sequences and a Bayesian classifier to be used for object segmentation. The new Background Modeling Neural Network is an unsupervised classifier.

L. Maddalena and A. Petrosino in their research [Maddalena, L. & Petrosino, A., 2008/2012/2014] developed SOBS (Self-Organizing Background Subtraction) algorithm 3dSOBS+. This SOBS algorithm implements an approach to moving object detection based on the neural background model automatically generated by a self-organizing method, without prior knowledge about the involved patterns. Such an adaptive model can handle scenes containing moving backgrounds, gradual illumination variations, and camouflage can include into the background model shadows cast by moving objects, and achieves robust detection for different types of videos taken with stationary cameras. Moreover, the introduction of spatial coherence into the background update procedure leads to the so-called SC-SOBS algorithm, which provides further robustness against false detections.

In [Ramirez-Alonso, G. & Chacon-Murguia, M., 2015] [Ramirez-Quintana, J., Chacon-Murguia, M. & Ramirez-Alonso, G., 2018] authors presented a video segmentation algorithm that takes advantage of using a background subtraction (BS) model with a low learning rate or a BS model with a high learning rate depending on the video scene dynamics. These BS models are based on neural network architecture, the self-organized map (SOM), and the algorithm termed temporal modular self-adaptive SOM, TMSA\_SOM. Depending on the type of scenario, the TMSA\_SOM automatically classifies and processes each video into one of four different specialized modules based on initial sequence analysis. This approach is convenient because, unlike state-of-the-art models, the proposed model solves different situations that may occur in the video scene with a specialized module. Furthermore, TMSA\_SOM automatically identifies whether the scene has drastically changed and automatically detects when the scene has become stable again and uses this information to update the background model in a fast way.

All methods and approaches considered above are aimed at finding the background, eliminating noise, and frame segmentation. The disadvantage of such approaches is the lack of information about the movement of the object in the image coming to the processing unit in real-time. The processing unit receives a pixel matrix of the image without any additional information and the entire load of the analysis of the current frame in real time is transferred to the computing unit/ controller (Figure 1). In other words, the computing unit can be represented as the human brain into which the frame image was inserted. The brain should analyze this image based on the history of previous frames and processing results. What's more, the brain doesn't know how this image is obtained: from the flow of a video camera at the moment, from watching a movie, from a fragment of a dream, or it's just a mirage or imagination. The brain gets a matrix of numbers. The computing center for finding

moving objects must extract the background from the frame. In addition, further processing of the frame is required to remove noise and interference (illumination changes, waving trees, cast shadows, moving background, camouflage, etc.) with the following regions segmentation.

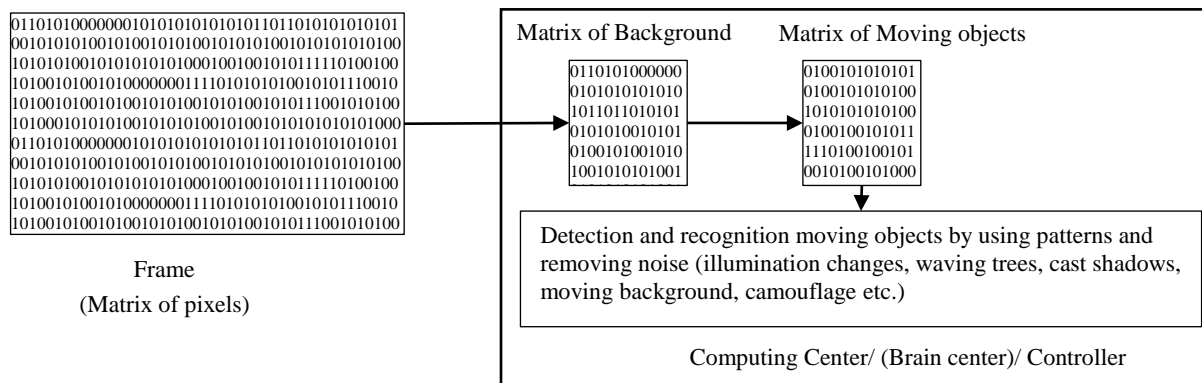


Figure 1. Traditional functional model of moving object detection

In this paper were proposed the principles of retinal operation based on the use of artificial impulse (spike) neural networks, which are the closest analogs of biological neural networks [Masland, R.H., 2001] [Wassle, H., 2004] [Olveczky, B., Baccus S. & Meister M., 2003] [Maass, W., 1997] [Wu, Q., McGinnity T., Maguire L. & Cai J., 2008] [Gerstner, W. & Kistler, W., 2002] (Figure 2).

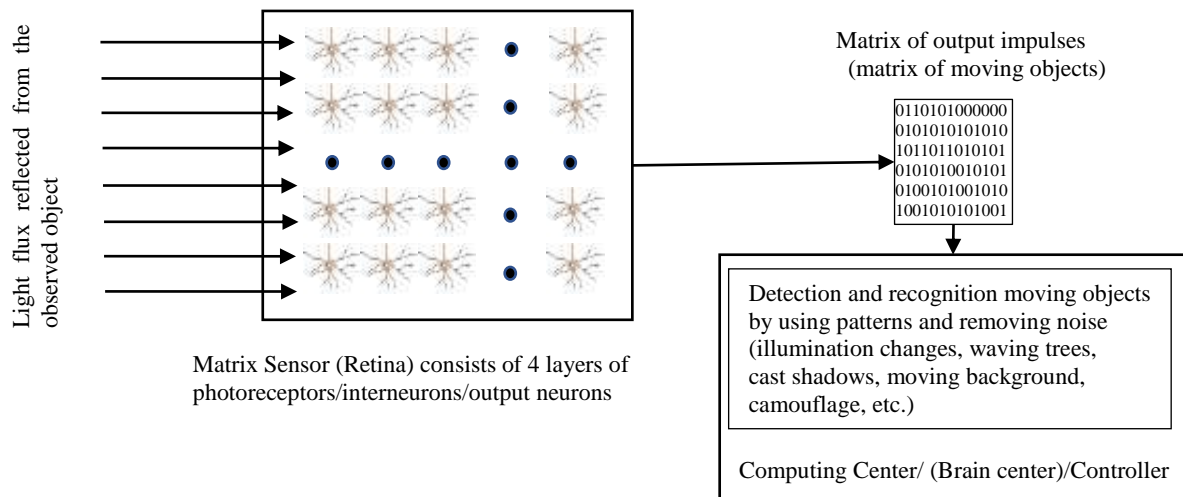


Figure 2. Functional model of moving object detection based on the network of impulse neurons

The light flux reflected from the objects falls on the retina receptors (described by the impulse neural network model), which record the energy/pulse change in the spike neurons (each of which corresponds to the pixel of the image frame) and generates an output matrix of impulses changes, i.e. the pixel matrix in which the energy of the impulse signal was changed (the matrix of movement objects). In this way, the computing center (brain center) obtains a matrix of all moving objects directly from the sensor retina (impulse neural network model) without computing the background and subtracting the background from the current image frame. The computing center in this case detects moving objects according to the patterns and removes the values of impulse signals caused by light changes, waving objects by wind, casting shadows, dynamic background, etc.

The originality of the proposed method (Functional scheme depicted in Figure 2) compared to previous works [Gil-Jimenez, P., Maldonado-Bascon, R. & Gil-Pita, H., 2003] [Wang, Z., Bao, H. & Zhang, L., 2009] [Wang, Z. & Bao, H., 2011] [Culibrk, D., Marques, O., Socek, D., Kalva, H. & Furht, B., 2006] [Maddalena, L. & Petrosino, A., 2008/2012/2014] [Ramirez-Alonso, G. & Chacon-Murguia, M., 2015] [Ramirez-Quintana, J., Chacon-Murguia, M. & Ramirez-Alonso, G., 2018] [Vaswani, N., 2018] [Mandal, M. & Vipparthi, S., 2021]

[Ammar, S., Bouwmans, T., Zaghden, N., Neji, M., 2019] [Minematsu, T., Shimada, A. & Taniguchi, R., 2020] [Giraldo, J. & Bouwmans, T., 2020] [Giraldo, J., Javed, S. & Bouwmans, T., 2022] is that the functional scheme operates on the principle of the impulsed neural network of the biological retinal model: the light reflected from the objects falls on the retina of the eye and any change in the luminosity of the reflected light is considered as motion.

Detection and recognition of the moving object, as well as the elimination of false alarms and removal of noise, will be carried out in the human brain (brain center, computing center, controller) based on the association of patterns of the shape of objects, paternal characteristics of the movement of objects and imagination derived from previous life experiences. The impulsed neural network (the retina of the eye) is designed only to detect movement by reacting the retina to the energy of light reflected from the objects around. An impulsed neural network does not detect a specific type of object or determine its shape. For this purpose, additional models of computation in the brain center are needed. For example, a sudden change in the lighting of a room (light switch on/off) will be identified as the movement of the whole environment around the observer. Therefore, the application of impulse neural networks requires the development of additional special methods and tools for the selection and detection of objects and the elimination of noise. The retina of the eye serves as a sensor-indicator of movement objects, color/brightness change, etc.

The advantage of the proposed approach is that it is not necessary to calculate the background matrix of the image and subtract it from the current image frame. In this work, the modeling of the impulsed neural network (retina) and the possibility of using this model for the indication of movement objects on the image have been described.

Both functional detection schemes (Figures 1, 2) require the elimination of noise in the computing center/controller. But the amount of computation in each scheme will be different. The use of impulse neural networks makes it possible to reduce the amount of computing in the controller. However, this approach requires the use of retina modeling (a special sensor based on the analog of the biological impulsed neural network of the retina).

This work focuses on motion indication and the development of motion sensors, but not on the detection of the type of moving object as such ones. The retina of the eye does not detect the type of moving object but sends a matrix of changes of the light luminosity reflected from the objects around.

Technologies and methods based on impulse neural networks modeling the retina are related to unsupervised methods. It is possible to adjust the sensitivity threshold of the impulse/energy change of the neuron and parameters of the models of impulsed neuron and impulsed neural network. All sensors are by nature designed to send the measured signal value to the controller and are built on unsupervised methods of functionality.

### **3. Indicating Movement Objects Using Impulse Neural Networks**

It is easy for a person to quickly isolate moving objects. But behind this skill lies a rather complex system of processing visual information - the retina of the eye [Masland, R.H., 2001] [Wassle, H., 2004]. The latter consists of complex chains of neurons, with photoreceptors (photosensitive cells) at the front of the neurons, which directly receive the optical signals and convert them into physiological excitations. The excitement from photoreceptors is transmitted via interneurons, or insertion neurons, which synaptically communicate with each other and bind photoreceptors to retinal ganglion cells that send signals through the optic nerve further into the brain. Different cell types are responsible for processing different image characteristics: brightness, color, movement of objects, etc. Under the influence of an external stimulus, whether a light signal or a signal from a neighboring neuron, the current neuron begins to emit pulses of the same amplitude. And the stronger the external impact, the more frequent the signals. Work [Olveczky, B., Baccus S. & Meister M., 2003] shows how a network of ganglion cells can instantly detect a moving object and even emit several such objects. The results of the research presented in the works [Masland, R.H., 2001] [Wassle, H., 2004] [Olveczky, B., Baccus S. & Meister M., 2003] [Maass, W., 1997] reveal the principles of retinal operation when separating moving objects from the point of view of physiology. But is it possible to use this information to create artificial neural networks that can isolate moving objects just as quickly and accurately?

The works of [Wu, Q., McGinnity T., Maguire L. & Cai J., 2008] [Dinu, A., Cirstea, M. & Cirstea, S., 2010] [Fujita, K., Okuno, S. & Kashimori, Y., 2018] [Miquel, J.R., Tolu, S., Schöller, F., & Galeazzi, R., (2021)] propose the use of artificial impulse (spike) neural networks, which are the closest analogs of biological neural networks. Unlike traditional neural networks [Sokic, E., & Konjicija, S., 2016] [Nielsen, M. A., 2015] [LeCun, Y., Bengio, Y. & Hinton, G., 2015], they do not transmit information in the form of average neural activity, but through a sequence of impulse signals [Maass, W., 1997] [Muder A., Alia A., Amer A., Saleh A. & Abdul, R.,

2020]. The output of a neuron is composed of short electrical impulses (also called potentials or spikes). The shape of the impulses does not change when transmitted by the axon. A chain of active potentials caused by a single neuron is called a pulse sequence, a series of identical events occurring at specific or random moments in time. Since all generated impulses are of roughly the same shape, the information is not contained in the form of impulses, but in their number and the exact timing of their occurrence.

#### 4. Model of an Impulse Neural Network for Detecting Moving Objects

The general structure of the impulse neural network used to isolate moving objects and used in the detector being developed is shown in Figure 3.

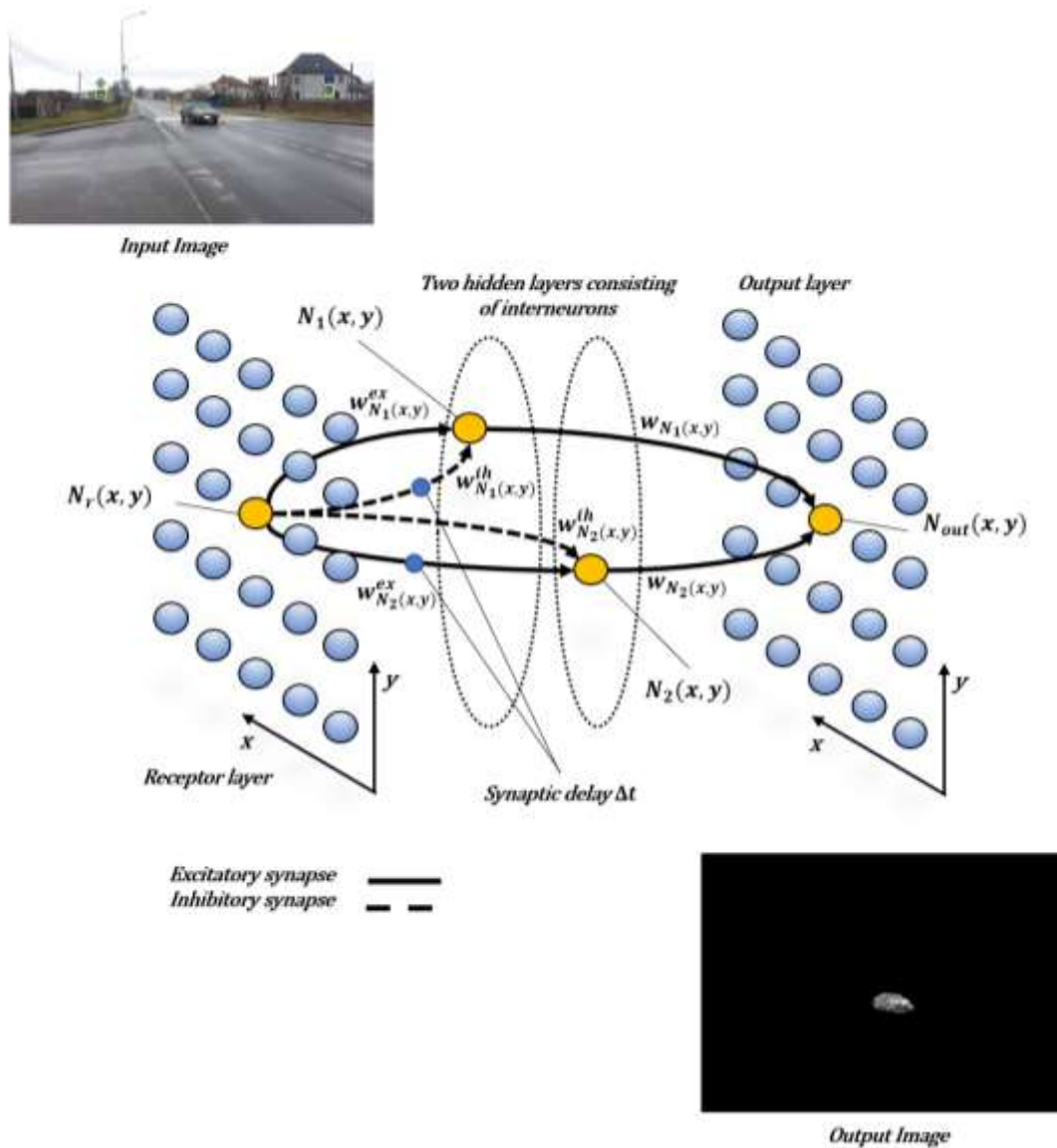


Figure 3. Motion detector based on the neural impulse network model

The input layer of neurons is an analog of the retinal photoreceptor layer, so what we're going to call the first layer neurons receptors. Each pixel  $(x, y)$  of the input frame of the video image corresponds to its receptor  $N_r(x, y)$ . The hidden layer is an analog of the retinal interlayer of the eye. It consists of two independent arrays of neurons  $N_1$  and  $N_2$ . They are the same size as the first layer and are connected by synaptic connections to both the input layer of neurons (receptors)  $N_r$  and the output layer of neurons  $N_{out}$ .

Each receptor  $N_r(x, y)$  is connected to the interneuron  $N_1(x, y)$  by means of an excitable synapse that transmits signals without delay and an inhibiting synapse that transmits signals with synaptic delay  $\Delta t$ . Each receptor  $N_r(x, y)$  is also connected to the interneuron  $N_2(x, y)$  by means of an excitatory synapse transmitting

signals with synaptic delay  $\Delta t$ , and an inhibiting synapse transmitting signals without delay. Let's suppose that  $I_{nr}(x, y, t)$  – the current flowing from the receptor  $N_r(x, y)$  at a time  $t$ . If the current flowing from the receptor  $N_r(x, y)$  is stable, i.e.  $I_{nr}(x, y, t) = I_{nr}(x, y, t - \Delta t)$ , and the excitatory and inhibitory inputs of the interneuron  $N_1(x, y)$  are balanced by adjusting the weights  $w_{N_1(x,y)}^{ex}$  and  $w_{N_1(x,y)}^{in}$  accordingly, the interneuron  $N_1(x, y)$  will be at rest. The situation is similar for interneuron  $N_2(x, y)$ . If the current of the receptor  $N_r(x, y)$  increases, i.e.  $I_{nr}(x, y, t) > I_{nr}(x, y, t - \Delta t)$ , the balance is disturbed since the signal coming from the excitatory synapse is stronger than the signal delayed for a period  $\Delta t$  coming from the inhibitory synapse. The interneuron  $N_1(x, y)$  begins to generate impulses (spikes). If the current of the receptor  $N_r(x, y)$  decreases, i.e.  $I_{nr}(x, y, t) < I_{nr}(x, y, t - \Delta t)$ , the interneuron  $N_1(x, y)$  does not react, but interneuron  $N_2(x, y)$  starts to generate impulses, since the delayed signal (for a period  $\Delta t$ ), flowing from the excitatory synapse is stronger than the signal from the inhibiting synapse. In other words, the neural network begins to react to the variation in pixel brightness that can be caused by the passage of objects moving over a static background.

The output layer of the  $N_{out}$  the neural network has the same dimensions as the input layer and the hidden layer. Each neuron of a given layer  $N_{out}(x, y)$  corresponds to each pixel  $(x, y)$  of the output frame of the video image. The interneurons  $N_1(x, y)$  and  $N_2(x, y)$  are connected by excitable synapses without delay to the output neuron  $N_{out}(x, y)$ . It produces signals only when it receives impulses from the interneuron  $N_1(x, y)$  or  $N_2(x, y)$ , otherwise, it is at rest. The magnitude of the grayscale of each pixel  $(x, y)$  of the output video frame is proportional to the frequency of impulse generation by the output neuron  $N_{out}(x, y)$  and has a value of 0 (black) if the output neuron  $N_{out}(x, y)$  does not generate any signals over a certain time period  $T$ . Otherwise, the brightness of the pixel  $(x, y)$  will be above 0 (Figure 4).



Figure 4. Selecting a Moving Object by Impulse Neural Network

The result of the simulation of the impulse neural network for the ideal case with a static background and without noise is shown in Figure 2. The result of this simulation demonstrates the capabilities of an impulsed neural network for educational purposes to explain the mechanism of impulsed neural networks in the detection/indication of movement objects.

### 5. Impulse Neuron Model

There are different models of [Gerstner, W. & Kistler, W., 2002] impulse neurons: Hodgkin-Huxley, «generalization-response» (Integrate-and-Fire), impulse response (Spike Response Model), etc. The most detailed and complex is the Hodgkin-Huxley model [Muder A., Alia A., Amer A., Saleh A. & Abdul, R., 2020]. It's based on an experimental study of a large number of squid neurons. The differential equation system of this model describes the exact reaction of the potential of the neuron membrane in response to different inputs. However, this realism results in high computational costs, and the model is not very suitable for experiments with neural networks composed of large numbers of neurons, as in this case.

Based on the work of [Wu, Q., McGinnity T., Maguire L. & Cai J., 2008], the motion detector being developed uses the model of the neuron «generalization-response» (integrate-and-fire, IaF), which is simpler in mathematical description and quite efficient.

The IaF model treats pulses as short-pulsed currents. Once the impulse arrives at the synapse, all associated

post-synaptic neurons are immediately charged. This voltage change is called post-synaptic potential. Once the potential of the neuron's membrane reaches the threshold value, it is reset and a new impulse is generated.

Let  $G_{x,y}(t)$  be the grayscale brightness for a single pixel  $(x,y)$  of the input image at time  $t$ ,  $q_{x,y}^{ex}(t)$  is the conductivity of the excitable synapse from the receptor  $N_r(x,y)$ ,  $q_{x,y}^{ih}(t)$  is the conductivity of the inhibiting synapse from the receptor  $N_r(x,y)$ , then the brightness transformation formulae of the grey scale will take the form of:

$$q_{x,y}^{ex}(t) = \alpha G_{x,y}(t), \quad q_{x,y}^{ih}(t) = \beta G_{x,y}(t), \quad (1)$$

where  $\alpha$  and  $\beta$  – some conversion coefficients. According to [Wu, Q., McGinnity T., Maguire L. & Cai J., 2008], the pulsed neuron  $N_1(x,y)$  can be characterized by the following equations:

$$\frac{dg_{N_1^{ex}(x,y)}(t)}{dt} = -\frac{1}{\tau_{ex}} g_{N_1^{ex}(x,y)}(t) + \alpha G_{x,y}(t), \quad (2)$$

$$\frac{dg_{N_1^{ih}(x,y)}(t)}{dt} = -\frac{1}{\tau_{ih}} g_{N_1^{ih}(x,y)}(t) + \beta G_{x,y}(t), \quad (3)$$

$$c_m \frac{dv_{N_1(x,y)}(t)}{dt} = g_m (E_m - v_{N_1(x,y)}(t)) + \frac{w_{N_1^{ex}(x,y)} g_{N_1^{ex}(x,y)}(t)}{A_{ex}} (E_{ex} - v_{N_1(x,y)}(t)) + \frac{w_{N_1^{ih}(x,y)} g_{N_1^{ih}(x,y)}(t-\Delta t)}{A_{ih}} (E_{ih} - v_{N_1(x,y)}(t-\Delta t)), \quad (4)$$

where  $g_{N_1^{ex}(x,y)}(t)$  and  $g_{N_1^{ih}(x,y)}(t)$  – the conductivity values of the membrane, respectively, of the excitatory and inhibiting synapses connecting the neurons  $N_r(x,y)$  and  $N_1(x,y)$ ,  $\tau_{ex}$  and  $\tau_{ih}$  – characteristic synaptic time of excitatory and inhibiting synapses respectively (usually 2 ms),  $\Delta t$  – synaptic delay in the transfer of impulse from neuron  $N_r(x,y)$  to neuron  $N_1(x,y)$ ,  $v_{N_1(x,y)}$  – neuron membrane potential  $N_1(x,y)$ ,  $E_m$  – equilibrium potential of the neuron membrane,  $g_m$  – neuron membrane conductivity value,  $E_{ex}$  и  $E_{ih}$  – equilibrium potentials values of excitatory and inhibiting synapses respectively,  $A_{ex}$  – the surface area of the neuron membrane  $N_1(x,y)$ , connected to the excitatory synapse,  $A_{ih}$  – the surface area of the neuron membrane  $N_1(x,y)$ , connected to the inhibiting synapse,  $c_m$  – specific capacity of the neuron membrane,  $w_{N_1^{ex}(x,y)}$  – the potency of the excitatory synaptic bond between receptor  $N_r(x,y)$  and interneuron  $N_1(x,y)$ ,  $w_{N_1^{ih}(x,y)}$  – the potency of inhibiting synaptic bond between receptor  $N_r(x,y)$  and interneuron  $N_1(x,y)$ . Coefficients  $w_{N_1^{ex}(x,y)}$  and  $w_{N_1^{ih}(x,y)}$  selected in such a way that the interneuron  $N_1(x,y)$  remained at rest during  $G_{x,y}(t) = G_{x,y}(t - \Delta t)$ . Similar equations are built for interneuron  $N_2(x,y)$ .

When the potential of an interneuron membrane  $N_1(x,y)$  or  $N_2(x,y)$  reaches the threshold  $v_{th}$ , it generates a pulse transmitted to the output neuron  $N_{out}(x,y)$ . Let's suppose that  $S_{N_1(x,y)}(t)$  reflects the sequence of impulses generated by the neuron  $N_1(x,y)$ , then

$$S_{N_1(x,y)}(t) = \begin{cases} 1, & \text{if interneuron } N_1(x,y) \text{ generated impulse at a time } t, \\ 0, & \text{if interneuron } N_1(x,y) \text{ did not generate impulse at the time } t. \end{cases} \quad (5)$$

The sequence of impulses generated by the neuron  $N_2(x,y)$ , denote  $S_{N_2(x,y)}(t)$ . Output neuron  $N_{out}(x,y)$  will be described by the following equations:

$$\frac{dg_{N_{out}(x,y)}}{dt} = -\frac{1}{\tau_{ex}} g_{N_{out}(x,y)}(t) + (w_{N_1(x,y)} S_{N_1(x,y)}(t) + w_{N_2(x,y)} S_{N_2(x,y)}(t)), \quad (6)$$

$$c_m \frac{dv_{N_{out}(x,y)}}{dt} = g_m (E_m - v_{N_{out}(x,y)}(t)) + \frac{g_{N_{out}(x,y)}(t)}{A_{ex}} (E_{ex} - v_{N_{out}(x,y)}(t)), \quad (7)$$

where  $g_{N_{out}(x,y)}$  – The conductivity value of each of the synapses connecting the interneurons  $N_1(x,y)$  and  $N_2(x,y)$  with the output neuron  $N_{out}(x,y)$ ,  $v_{N_{out}(x,y)}$  – neuron  $N_{out}(x,y)$  membrane potential,  $w_{N_1(x,y)}$  – the power of the excitatory synaptic bond between interneurons  $N_1(x,y)$  and the output neuron  $N_{out}(x,y)$ ,  $w_{N_2(x,y)}$  – the power of the excitatory synaptic bond between interneurons  $N_2(x,y)$  and output neuron  $N_{out}(x,y)$ .

Let's suppose that  $S_{N_{out}(x,y)}(t)$  – sequence of impulses generated by the output neuron  $N_{out}(x,y)$ . Then the frequency of impulses generated by the output neuron  $N_{out}(x,y)$ , shall be calculated as follows:

$$r(x,y,t) = \frac{1}{T} \sum_t^{t+T} S_{N_{out}(x,y)}(t), \quad (8)$$

where  $T$  – Measurement period of impulses generated by the output neuron  $N_{out}(x,y)$ .

Converting a magnitude  $r(x,y,t)$  for each pixel of the output image in the grayscale luminance, a lighter region can be obtained to describe the moving object (Figure 2).

## 6. Conclusion

The above approach for detecting and separating moving objects is an attempt to simulate the ability of the human eye to isolate moving objects quickly enough and surpass existing deterministic methods in terms of the selection speed of moving objects and economy computing resources. The motion detector has been developed on the basis of this approach, as a software module, which can be used in the field of digital video image processing (motion indication sensor is an impulsed neural network simulating the retina of the eye). It is intended to use this detector in automated traffic management systems as an alternative to existing detectors, even taking into account the possible improvement of the latter through the use of parallel computations for simultaneous processing of video segments and selection of moving objects within each of them [Sadhukhan, P. & Gazi, F., 2018] [Sudan, J., Changho, S., Eunmok, Y. & Gyanendra, P., 2021] [Yeuseyenka, I.A. & Melnikau, I.I., 2022]. Such an approach has already proven effective in quickly and accurately highlighting moving objects in video images [Miquel, J.R., Tolu, S., Schöller, F., & Galeazzi, R., (2021)] [Wu, Q., McGinnity T., Maguire L. & Cai J., 2008]. The main problem of using this type of detector in traffic management systems is to eliminate noise because any small change in the energy/brightness of the pixel of an image will be fixed as a moving object. Further research will be related to the optimization of parameters of the impulsed neuron model and the impulsed neural network to increase the accuracy and stability of the process of detecting the movement of target objects and prevent false alarms of the movement objects being created by changing lighting, shadows casting, dynamic background, weather conditions, disguise the object, etc. In addition, further research is planned on the optimal selection of the threshold of sensitivity of the brightness/energy of the motion detection impulse neuron to reduce motion false alarms and decrease the number of calculations, as well as the impact of this threshold of sensitivity on the quality and stability of the detection of moving objects.

The disadvantage of impulse-based sensors is the difficulty of detecting motion if the moving object and the background of the image have the same or similar color. In this case, the reflected light from moving objects and the background of the image will have approximately the same brightness, which creates serious problems for detecting the possible motion of objects. Increasing the threshold of sensitivity to the impulse neuron can lead to excessive noise in the image. To solve such problems, it is intended to develop an optimal segmentation of the image to regions by color/brightness and increase the sensitivity threshold in the area of the borders of these contours with the following experiments on the CDnet 2014 dataset (<http://changedetection.net/>) and on the BMC 2012 dataset (<http://backgroundmodelschallenge.eu/>) both quantitatively and qualitatively.

It should also be borne in mind that impulse neural network elements can be implemented in hardware [Dinu, A., Cirstea, M. & Cirstea, S., 2010], or software using modern parallel computing technologies based on graphics processors [Gonzalez, R. & Woods, R., 2018] [Appleyard, J., Kocisky, T. & Blunsom, P., 2016] [Weninger, F., Bergmann, J., & Schuller, B., 2015]. This can significantly accelerate the selection of moving objects in video images, although it may be costly to implement and pre-configure the motion detector.

## Acknowledgments

The authors acknowledge the support provided by the Director and all other staff members of the School of Automotive Technologies, Research Center of Automotive Power Train and Electronic Control, College of Automotive Engineering, Hubei University of Automotive Technology (Shiyan, China). The research was financially supported by the Hubei University of Automotive Technology within the framework of the Long-term



Innovation Project BK-202206 “Research on Intelligent Computer Integrated Manufacturing/Industrial control and monitoring. Methodology Development of Innovative Human-Machine Interfaces” and partly contributed by OYGJS-2021002 "Application research and development of a distributed photovoltaic power generation system" (School of Information Engineering, Xi'an Eurasia University, Xi'an, China) in order to develop innovative Human-Machine Interfaces, training equipment and its intelligent control system based on computer vision and to integrate them with IoT (Internet of Things) and IIoT (Industrial Internet of Things).

## References

- Ammar, S., Bouwmans, T., Zaghden, N., & Neji, M. (2019). *Moving Objects Segmentation Based on DeepSphere in Video Surveillance. Advances in Visual Computing*. ISVC 2019. Lecture Notes in Computer Science(), vol 11845. Springer, Cham. [https://doi.org/10.1007/978-3-030-33723-0\\_25](https://doi.org/10.1007/978-3-030-33723-0_25)
- Appleyard, J., Kocisky, T., & Blunsom, P. (2016). *Optimizing performance of recurrent neural networks on GPUs*. Retrieved from <https://arxiv.org/pdf/1604.01946.pdf>
- Beyerer, J., Puente L., & Frese F. (2016). *Machine Vision: Automated Visual Inspection: Theory, Practice and Applications*. Springer-Verlag, Berlin, New York. <http://dx.doi.org/10.1007/978-3-662-47794-6>. Retrieved from <https://publikationen.bibliothek.kit.edu/1000070905>
- Chakrabarti, I., Srinivasarao Batta, K., & Chatterjee, K. (2015). *Motion Estimation for Video Coding*. Springer Publishing, Switzerland. <https://doi.org/10.1007/978-3-319-14376-7>
- Cheng, Y., Hu, X., Wang, J., Wang, Y., & Tamura, S. (2015). Accurate Vessel Segmentation with Constrained B-Snake. *IEEE Trans. Image Process*, 24(8), 2440-2455. <https://doi.org/10.1109/tip.2015.2417683>
- Culibrk, D., Marques, O., Socek, D., Kalva, H., & Furht, B. (2006). *A Neural Network Approach to Bayesian Background Modeling for Video Object Segmentation*. International Conference on Computer Vision Theory and Applications, VISAPP 2006. Setubal. Portugal. February 2006. <https://doi.org/10.5220/0001374604740479>
- Dinu, A., Cirstea, M., & Cirstea, S. (2010). Direct Neural-Network Hardware-Implementation Algorithm. *IEEE Transactions on Industrial Electronics*, 57(5), 1845-1848. <https://doi.org/10.1109/TIE.2009.2033097>
- Fujita, K., Okuno, S. & Kashimori, Y. (2018). *Evaluation of the computational efficacy in GPU-accelerated simulations of spiking neurons*. Computing. pp. 907–926. <https://doi.org/10.1007/s00607-018-0590-0>
- Garcia-Garcia, B., Bouwmans T., & Silva A. (2020). Background Subtraction in Real Applications: Challenges, Current Models and Future Directions. *Computer Science Review*, 35. <https://doi.org/10.1016/j.cosrev.2019.100204>
- Gerstner, W., & Kistler, W. (2002). *Spiking Neuron Models. Single Neurons, Populations, Plasticity*. Cambridge University Press. 496 p. <https://doi.org/10.1017/CBO9780511815706>
- Gil-Jimenez, P., Maldonado-Bascon, R., & Gil-Pita, H. (2003). Background pixel classification for motion detection in video image sequences. *International Work Conference on Artificial and Natural Neural Network, IWANN 2003*, 2686, 718-725. [https://doi.org/10.1007/3-540-44868-3\\_91](https://doi.org/10.1007/3-540-44868-3_91)
- Giraldo, J., & Bouwmans, T. (2020). *Semi-Supervised Background Subtraction of Unseen Videos: Minimization of The Total Variation of Graph Signals*. IEEE ICIP 2020. pp. 3224-3228. Abu Dhabi. UAE. October 2020. <https://doi.org/10.1109/ICIP40778.2020.9190887>
- Giraldo, J., Javed, S., & Bouwmans, T. (2022). *Graph Moving Object Segmentation*. IEEE Transactions on Pattern Analysis and Machine Intelligence. <https://doi.org/10.1109/TPAMI.2020.3042093>
- Gonzalez, R., & Woods, R. (2018). *Digital Image Processing. Fourth Edition*. New York.: Pearson Education ISBN 978-0-13-335672-4. pp. 1022. Retrieved from <https://www.studocu.com/row/document/jamaa%D8%A9-aaman-alaarby%D8%A9/history-arabic/digital-image-processing-4th-edition-toc/7618609>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). *Deep learning. Nature*. pp. 436-444. <https://doi.org/10.1038/nature14539>
- Maass, W. (1997). Networks of Spiking Neurons: The Third Generation of Neural Network Models. *Neural Networks. Elsevier Science Ltd. No. 9(10)*, 1659-1671. <https://doi.org/10.1016/S0893-6080%2897%2900011-7>
- Maddalena, L., & Petrosino, A. (2008). A Self-Organizing Approach to Background Subtraction for Visual

- Surveillance Applications. *IEEE Transactions on Image Processing*, 17(7), 1168-1177. <https://doi.org/10.1109/TIP.2008.924285>
- Maddalena, L., & Petrosino, A. (2012). *The SOBS Algorithm: What Are the Limits? IEEE Workshop on Change Detection*, CVPR 2012, June 2012. <https://doi.org/10.1109/CVPRW.2012.6238922>
- Maddalena, L., & Petrosino, A. (2014). *The 3dSOBS+ algorithm for moving object detection*. *Computer Vision and Image Understanding*, CVIU 2014, May 2014. <https://doi.org/10.1016/j.cviu.2013.11.006>
- Mandal, M., & Vipparthi, S. (2021). *An Empirical Review of Deep Learning Frameworks for Change Detection: Model Design, Experimental Frameworks, Challenges, and Research Needs*. *IEEE Transactions on Intelligent Transportation Systems*. <https://doi.org/10.1109/TITS.2021.3077883>
- Masland, R. H. (2001). *The Fundamental Plan of the Retina*. *Nature Neuroscience*. *Nature Publishing Group*. No 4., 877-886. <https://doi.org/10.1038/nn0901-877>
- Menter, Z., Tee, W., & Dave, R. (2021). *A study of machine learning based pattern recognition in IoT devices*. In *Proceedings of the 3rd*. [https://doi.org/10.1007/978-981-16-3246-4\\_52](https://doi.org/10.1007/978-981-16-3246-4_52)
- Minematsu, T., Shimada, A., & Taniguchi, R. (2020). *Rethinking Background and Foreground in Deep Neural Network-Based Background Subtraction*. *IEEE International Conference on Image Processing*. ICIP 2020. Abu Dhabi. UAE. October 2020. <https://doi.org/10.1109/ICIP40778.2020.9191151>
- Miquel, J. R., Tolu, S., Schödl, F., & Galeazzi, R. (2021). *RetinaNet Object Detector Based on Analog-to-Spiking Neural Network Conversion*. 2021 8th International Conference on Soft Computing & Machine Intelligence (ISCMI), 201-205. <https://doi.org/10.1109/ISCMI53840.2021.9654818>
- Muder A., Alia A., Amer A., Saleh A., & Abdul, R. (2020). *Deep recurrent neural network for IoT intrusion detection system*. *Simul. Model. Pract. Theory*. <https://doi.org/10.1016/j.simpat.2019.102031>
- Nielsen, M. A. (2015). *Neural Networks and Deep Learning. Determination*. Press. Retrieved from <http://neuralnetworksanddeeplearning.com/index.html>
- Olveczky, B., Baccus S., & Meister M. (2003). *Segregation of Object and Background Motion in the Retina*. *Nature*. *Nature Publishing Group*. No. 423. pp. 401-408. <https://doi.org/10.1038/nature01652>
- Purohit, S., Singh Jat, D., Poonia, R., Kumar, S., & Hiranwal, S. (2021). *Proceedings of International Conference on Communication and Computational Technologies, Algorithms for Intelligent Systems*. Jaipur, India, 27–28 February 2021. <http://dx.doi.org/10.1007/978-981-15-5077-5>
- Ramirez-Alonso, G., & Chacon-Murguia, M. (2015). *Object detection in video sequences by a temporal modular self-adaptive SOM*. *Neural Computing and Applications*. <https://doi.org/10.1007/s00521-015-1859-2>
- Ramirez-Quintana, J., Chacon-Murguia, M., & Ramirez-Alonso, G. (2018). *Adaptive background modeling of complex scenarios based on pixel level learning modeled with a retinotopic self-organizing map and radial basis mapping*. *Applied Intelligence*. <https://doi.org/10.1007/s10489-018-1256-5>
- Sadhukhan, P., & Gazi, F. (2018). *An IoT based intelligent traffic congestion control system for road crossings*. 2018 International Conference on Communication, Computing and Internet of Things (IC3IoT), pp. 371–375, Chennai, India. <https://doi.org/10.1109/IC3IoT.2018.8668131>
- Sobral, A. (2014). *BGS Library: A Library Framework for Algorithm's Evaluation in Foreground/Background Segmentation*. Chapter 23. *Handbook on "Background Modeling and Foreground Detection for Video Surveillance"*. CRC Press, Taylor and Francis Group. July 2014. <https://doi.org/10.1201/b17223-29>
- Sokic, E., & Konjicija, S. (2016). *Phase-Preserving Fourier Descriptor for Shape-Based Image Retrieval*. *Signal Processing: Image Communication*, 40, 82-96. <https://doi.org/10.1016/j.image.2015.11.002>
- Sudan, J., Changho, S., Eunmok, Y., & Gyanendra, P. (2021). *Real-time object detection and tracking system for video surveillance system*. *Multimed Tools Appl* 80, 3981-3996. <https://doi.org/10.1007/s11042-020-09749-x>
- Vaswani, N. (2018). *Robust PCA and Robust Subspace Tracking: A Comparative Evaluation*. *IEEE Statistical Signal Processing Workshop*. SSP 2018. Freiburg, Germany. June 2018. <https://doi.org/10.1109/SSP.2018.8450718>
- Wang, Z., & Bao, H. (2011). *Cooperative Neural Network Background Model for Multi-Modal Video Surveillance*. *International Conference on Computational Intelligence and Security*, CIS 2011, pages 249-254, December 2011. Retrieved from <https://doi.ieeecomputersociety.org/10.1109/CIS.2011.63>

- Wang, Z., Bao, H., & Zhang, L. (2009). *PNN based motion detection with adaptive learning rate*. International Conference on Computational Intelligence and Security, CIS 2009, Beijing, December 2009. <https://doi.ieeecomputersociety.org/10.1109/CIS.2009.178>
- Wassle, H. (2004). *Parallel Processing in the Mammalian Retina*. Nature Reviews Neuroscience. Nature Publishing Group. No. 5. p. 747-757. <https://doi.org/10.1038/nrn1497>
- Weninger, F., Bergmann, J., & Schuller, B. (2015). Introducing CURRENNT: the Munich open-source CUDA recurrent neural network toolkit. *Journal of Machine Learning Research*, 16, 547-551. <https://doi.org/10.5555/2789272.2789289>
- Wu, Q., McGinnity T., Maguire L., & Cai, J. (2008). *Motion Detection Using Spiking Neural Network Model*. ICIC'08 Proceedings of the 4th international conference on Intelligent Computing. Berlin: Springer-Verlag. pp. 76-83. [https://doi.org/10.1007/978-3-540-85984-0\\_10](https://doi.org/10.1007/978-3-540-85984-0_10)
- Yeuseyenka, I. A., & Melnikau, I. I. (2022). Intelligent transport subsystem for estimating the traffic density and its composition based on artificial neural networks and parallel computing. *Journal of the Technical University of Gabrovo*, 64, 11-17. Retrieved from <http://umis.tugab.bg/prep/tomove/22/Vol64-2022-3-11-17.pdf/http://izvestia.tugab.bg/index.php?m=20&to m=22>
- Yeuseyenka, I. A., Melnikau I. I., & Yemelyanov, I. A. (2022). Detection and Selection of Moving Objects in Video Images Based on Impulse and Recurrent Neural Networks. *Journal of Data Analysis and Information Processing*, 10, 127-141. <https://doi.org/10.4236/jdaip.2022.102008>

### Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).