Tracking Dynamics Objects on Binocular Vision System

Rosen Spirov[1], Neli Grancharova[2] and Svilen Soyanov[3]

Abstract – **This paper presents the system to detect corresponding fragments of stereo pair and finally setting the parameters of dynamic objects, implemented in FPGA A software-based algorithm was independently developed and examined in MATLAB to evaluate its performance and verify its effectiveness.**

Keywords – **Image Processing, Filters, FPGA, StereoVideo.**

I. INTRODUCTION

The goal of this idea was to create an FPGA system to detect and track dynamics objects in real time, as in human vision.

Today in the scientific literature there are numerous models depicting anatomy, operation and interpretation of visual information in the brain as shown in Figure.1 [1].

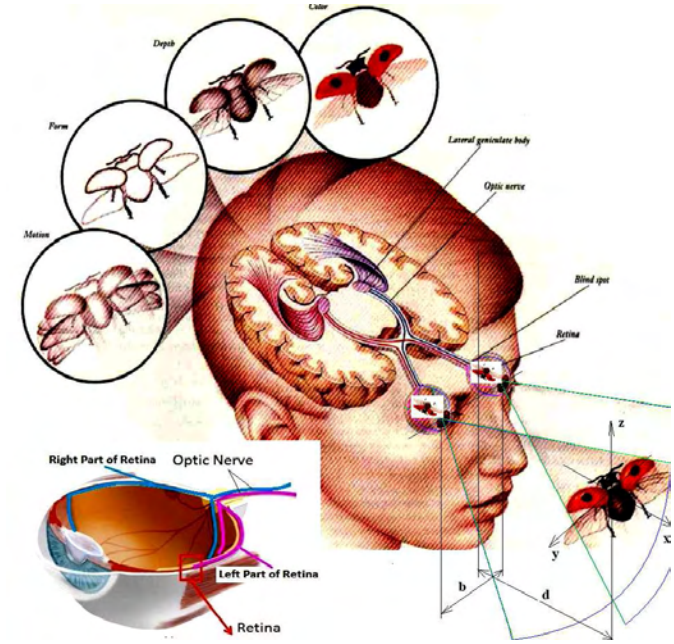

Fig.1. The visual information in the human brain

The medicine established laws of human perception, the characteristics of the optic tract, ways of forming mental images and their interpretation. Human perception is a function of a number of subconscious: processes for determining distances, edges, shapes, sizes, colors, placements; adaptation to light, heat, radiation; determination of changes plasticity, interference, other indicators for the objects and the environment.

In the human eye, the axons of the ganglion cells are attached to the bundle of optic nerve. The entire overview that can capture the eye of man is individual visual space. For each eye is left and right visual space. The visual information from the retina reach the area of the brain called the LGN. This is represented in Figure 2 [1].

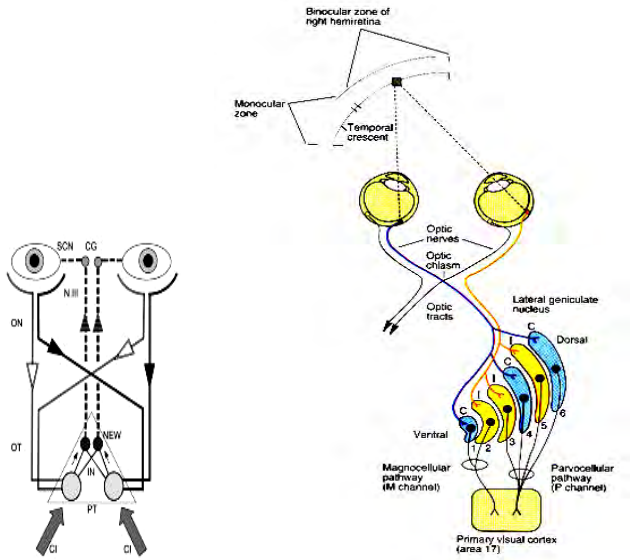

Fig.2. The eyes, visual tract and LGN area

Each of LGN consists of 6 levels- 2 inside levels that contain large cells -*magnocellular* and 4 outside levels with small cells - *parvocellular*. Between the layers there level - *interlaminar*, made up of other small cells - *koniocellular*. These layers receive input information from different types of ganglion cells. The process of entering information from each eye alternately flows to different types of layers. Visual tract continues to cortex lobe in the brain. The first cortex synapses of each neuron carries visual information in the middle part of the occipital lumbar part of the brain so on...in the *Brodman-card*. This area is located medially and a knurled surface with the external folds of the cortex. *Cyto-architecture* is perfectly symmetrical and dots - formed an increase in volume by surface corrugations of the cortex. Similarly, it is built and optic cortex share, as is shown in the Figure 3[1].

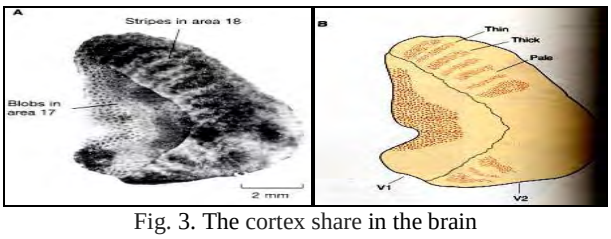

Fig. 3. The cortex share in the brain

[1]Rosen Spirov is with the Faculty of Electronics at Technical University of Varna, 1 Studentska Str, Varna 9010, Bulgaria, E-mail: rosexel@abv.bg.

[2]Neli Grancharova is with the Department of Termotechnics at Technical University of Varna, 1 Studentska Str, Varna 9010, Bulgaria, E-mail: nelly2000@abv.bg.

[3]Svilen Stoyanov is with the Technological College of Dobrich at Technical University of Varna, 1 Studentska Str, Varna 9010, Bulgaria, E-mail: svilenh@abv.bg.

This is consistent with the hypothesis, that there is carried out the pre-processing of visual space. Following anatomy work [1] on that partition can be represented with diagram:

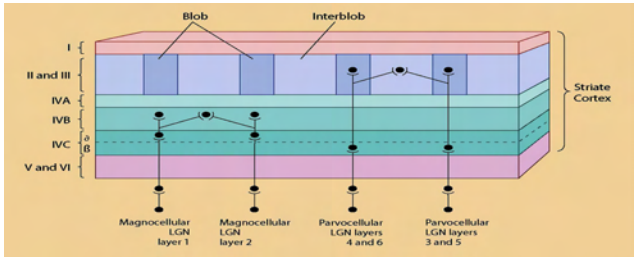

Fig. 4 . Operation of the cortex share

The separation of M and P ways are made in the cortex. Axons from both regions are terminated at layer 4 of the folds of the cortex, if (a) the terminal areas of axons in this layer have a branch each other- M - road, and (b) P - road drags second synapse information conveyed in the outer layers 2 and 3. The neurons equally selective ability are grouped in columns, according to the dominant eye and preferred orientation [1]. The orientation of the columns regularly presichat of blobs, as in figure 5.

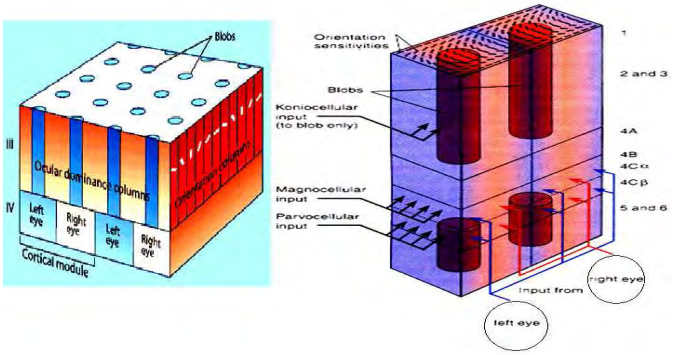

Fig.5. The processing in cortex share

Like the function of the human eye binocular video systems recreate a certain extent the human vision [3]. The own TV tracker detect and track dynamic objects and measure the parameters of the relative motion of the object set to TV system on Altera FPGA. They received images from both sensors differ in the location of objects on their screens and are determined by various observation points.

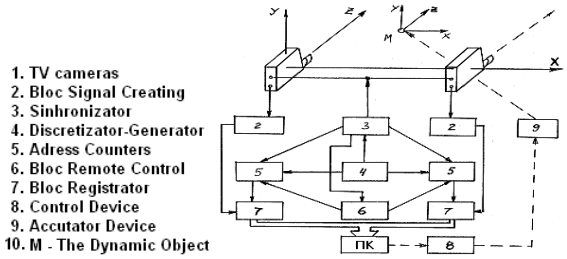

Fig. 6. TV trackers with automatically targeted

II. The Design and the Algorithm

The mathematical model is based on an algorithm to determine the parameters of dynamic objects in series stereo images [2], shown in figure 7. Based on the above algorithm is proposed structural diagram of the system for determining the parameters of dynamic objects, illustrated on figure 8.

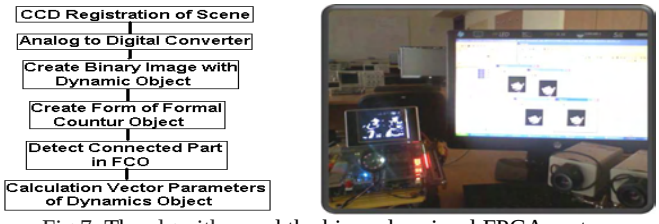

Fig.7. The algorithm and the binocular visual FPGA system

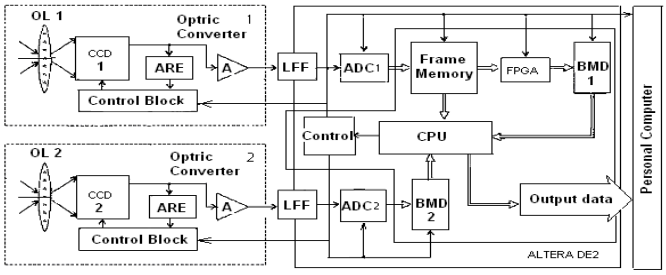

Fig. 8. The structural diagram

The description includes two optical converters (OC) with CCD [3]. The main blocks of OEП are: lens OL, CCD optical converter, control unit ARE and amplifier block (A). The output video signals of OC in digital form passed by low-pass filter (LFF), after that they enter the inputs of the 8- bits ADC. From output of ADC 2 digital signal is stored in a buffer memory device (BMD2) and from output of ADC1 digital signal is sent and saved in the block of frame memory. This block remembers N frames from the series images of the work scene. To achieve high-speed data processing is necessary to perform parallel calculations. The proposed implementation includes the structural organization of the block frame memory to design, as shown in Figure 9.

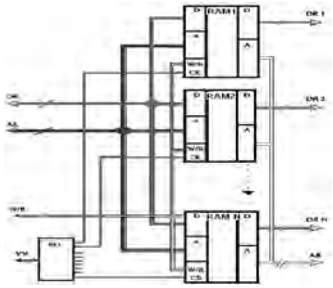

Fig. 9 The frame memory

Block of frame memory is a separate stack size N and type of organization FIFO, as the number of cells in the stack is equal to the number of frame. As a memory module used Dual-Port RAM (BMD). Both R and L ports allow independent applications to read and write to any cell array.

To achieve of parallel calculation is necessary every frame of image series to be stored in a separate of memory (BMD). The volumes of necessary memory blocks (BMD) using 8 bits quantization of size 640x480 pixels equivalence of 307.2 K bytes. Corresponding to input addresses AL, data DL, as well as the inputs W / R, the port L, are combined together.

Through port L recording is performed by the data outputs of the ADC and reading them from the microprocessor to determine related areas of the picture. The inputs to permit CE port L are connected to the outputs of the shift register, the outputs of which correspond to the number of used RAM (BMD). The management device (Control) controlled the work of register. Register carried move level logical unit to complete the recording of one frame of image. Thus, it is possible to work through the gate L only one module of RAM. Through a transferring registry is organized a stack FIFO. Through R port is carried out read data for forming the binary image, which contains a dynamic image areas S (k) by using FPGA. The corresponding input address AR, and also inputs W / R port R are merged together and entering into the control unit. Using Dual-Port RAM (BMD) realized bought of reading and record of information, without the need for waiting for the complete filling of the working RAM. This significantly shortens processing time. The device A formed signals of AL and AR, differing only by one bit, which in practice also simplifies the formation of the addresses. Thus, the output of the memory block of the frame memory is represented by the N- 1 - dimensional vector with values of pixels of the background image $X_{\min} = \left\lfloor x_{\min}^1, x_{\min}^2, ..., x_{\min}^{N-1} \right\rfloor$, which are formed on the basis of the N-1 past values $\left\{ x_{\min}^k, k = 0,1... \right\}$ for pixel in the process and current value of pixels in the image $x_{\min}^k$. For every component of the vector Xmn corresponding data bytes *DR1-DR N- 1* and the current value of the pixel corresponds byte data DR N from worked memory (ОЗУ). The task of checking the condition $\min \left| X_{\min} - x_{\min}^k \right| > \sigma^{ДОП}$, most often implemented using a FPGA [61]. In this case, for each component of the obtained vector with differences $X_{\min} - x_{\min}^k$ it is necessary to check the condition X>C, where X and C is 8 bits binary numbers *x7…x0* and *c7…c0*, thus *x0* and *c0* is a low bits. From the output signal Q=1 is removed by X>C or Q=0 by X$\leq$C. The output data for coding presents in terms of Boolean algebra. The condition X > C for three orders can be set as follows:

$$Q = x_2\bar{c}_2 + x_1\bar{c}_1(\overline{x_2 \oplus c_2}) + x_0\bar{c}_0(\overline{x_2 \oplus c_2})(\overline{x_1 \oplus c_1}) \quad (1)$$

The record (3.4) can be seen in the following form. The number X is larger than the number C, if one of three conditions:

- $x_2\bar{c}_2 = 1$, high bit of number *X* is equal to 1, and high bit of number *C* is equal to 0.
- $x_1\bar{c}_1(\overline{x_2 \oplus c_2}) = 1$, or the high bits is equals – its sum in modulo-2 of the inversion is equal to the 1 , in that $x_1=1$, *a* $c_1=0$.
- $x_0\bar{c}_0(\overline{x_2 \oplus c_2})(\overline{x_1 \oplus c_1}) = 1$, or the 2 high bits with codes X and *C* tally $x_0 = 1$, $c_0 = 0$. The condition (3.4) extends to any number of bits *n:*

$$Q = x_{n-1}\bar{c}_{n-1} + x_{n-2}\bar{c}_{n-2}(\overline{x_{n-1} \oplus c_{n-1}}) + x_{n-3}c_{n-3}(\overline{x_{n-1} \oplus c_{n-1}})(\overline{x_{n-2} \oplus c_{n-2}}) + ...$$

$$... + x_0c_0(\overline{x_{n-1} \oplus c_{n-1}})(\overline{x_{n-2} \oplus c_{n-2}})...(\overline{x_1 \oplus c_1}) \quad (2)$$

From equation (3.5), it follows that when $c_i=1$ *(i=0, 1, 2,…, n-1),* then in which as a multiplier enters Ci becomes zero. Thus, when set with the right side of eq.(2) removes all multiplications corresponding to the units of bits with binary representation C0 ... C7. In particular, when C = 1 from the right side of eq. (2) removes all multiplications and variable Q becomes identically zero (condition X> C in this case is impossible because C is the maximum representable number). Conversely, however, when C = 0 eq. (2) after simplification will take the form:

$$Q = x_{n-1} + x_{n-2} + x_{n-3} + ... + x_0 \quad (3)$$

From eq. 3.6 it is following that *Q=0* only by $x_{n-1} = x_{n-2} = x_{n-3} = ... = x_0 = 0$, but in another cases Q=*1*, or *X>C=0.* Thus, the number of the logic works in coding the FPGA is equal to the number of 0 in the binary number $C = \delta^{ДОП}$ For parallel data processing in an FPGA must be implemented N- 1 structures described by the eq. (2) , the outputs of which are connected to the inputs of the AND gate (AND). For this purpose simplify the task of checking the condition $\min \left| X_{\min} - x_{\min}^k \right| > \sigma^{ДОП}$ and in the next step is reduced to task for matching the values of the components of the vectors X_{mn} to etalons $X_{mn}{}^k$, for reporting limits of disperse $\delta^{ДОП}$ to compare be six senior diluted bytes of data DR. Such an organization is presented as follows:

$$Q = (DR1_5 \oplus DR2_5 \oplus ... \oplus DRN_5) \oplus (DR1_4 \oplus DR2_4 \oplus ... \oplus DRN_4) \oplus ...$$
$$... \oplus (DR1_0 \oplus DR2_0 \oplus ... \oplus DRN_0) \quad (4)$$

This equation is easily implemented with logic elements such as the functional organization of the simplified structure will look like as shown in figure 10.

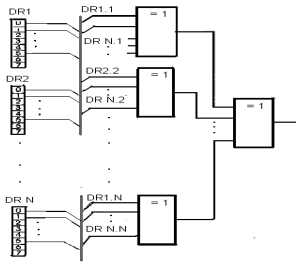

Fig. 9 The simplified structure

The FPGA handles data coming from the output of the block frame memory and stores them in buffer memory devices, the volume of which is 37.5 Kbyte. Microprocessor used to build FKO, make calculations by eq. (1) and eq. (2) of their parameters, such as processing the data. In addition to make calculations of matrix correction of Kalman filter [4]. This module is implemented with processors for embedded NIOS II on Altera FPGA. Operating time of the device T*y* is defined as follows:

$$T_y = T_{АЦП} + T_n + T_{FPGA} + T_1 + T_2, \quad (5)$$

Where $T_{АЦП}$ is the time of retention of the ADC; Tn - a delay of waiting at reading data from blocks of staffing memory is retention time processing with FPGA; is the time to create FCO in БЗУ1 including pre-filtration, filling gaps and

segmentation; is the time to detect their corresponding fragments of stereo pair and finally setting the parameters of dynamic objects. The time limit for the operation of the device in the presence of five sites in the frame constitutes 2.2ms, allowing processing of 25 frames per second.

III. Experimental Results

When a large number of observation accuracy of the estimate becomes sufficient to form, in practice, terminals for evaluation of the performance of compatibility, consistency and efficiency [5]. It should be noted that the width of the range of measurement of the brightness of the pixel determines the statistical difference between the minimum brightness levels of the background and a dynamic object set in the algorithm.

The reliability γ is assumed to be chosen greater than 0.95. Then the event, which range $(-\hat{x}^k_{mn}, +\hat{x}^k_{mn})$ roof parameter x^k_{mn} , will be faithful. After determining the: $\Phi(x_\gamma) = \frac{\gamma}{2}$, should find the corresponding tabular value tables for the function of Laplace . Thus, in $\gamma = 0{,}99$, то $\Phi(x_\gamma) = 0{,}495$ and in the next : $x_\gamma = 2{,}58$, and $\delta = 2{,}58\frac{\sigma_\chi}{\sqrt{N}}$.

Then the confidence interval will have limits :

$$(\bar{x}_B - 2{,}58\frac{\sigma_\chi}{\sqrt{N}}, \bar{x}_B + 2{,}58\frac{\sigma_\chi}{\sqrt{N}}) \qquad (6)$$

Therefore, with probability 0.99 we can be sure that the interval in eq. (6) covers parameters m_x, or in other words, with probability 0.99 magnitudes $\bar{x}_B$ gives a value for the parameter n_x accuracy:

$\delta = 2{,}58\frac{\sigma_\chi}{\sqrt{N}}$. From the experimental data when N = 50 is obtained $\max(\sigma_\chi) = 14$ and $\delta = 2{,}58\frac{14}{\sqrt{50}} \approx 5{,}1$. Rounded to obtain σ = 6 , so it is the minimum difference between the levels of background brightness and dynamic object in the algorithm are not less than 6 .

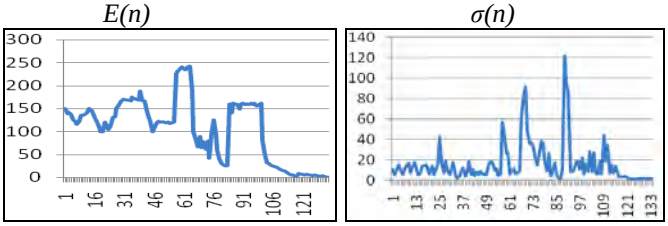

Fig.10. Values reported minimum and maximum luminance for each pixel E (n) and the values of dispersion σ (n) for each pixel in row.

Special attention should be paid to the pixels in the image located at the boundary of high contrast areas of the stage, the peculiarity of which lies in the sharp rise in values of dispersion. Such an increase of the variance explained by the error of quantization, i.e., quantization border pixel is set to tout for one, then the other area of the image. The minimum and maximum values of accounts brightness for each pixel along a row of image, processed statistically in the middle of MS Exel, containing contrast areas that are visually shown in Figure 11shows the values of the variance σ_x .

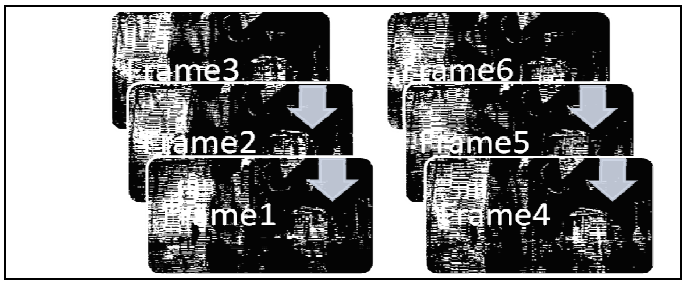

Fig. 11 The series stereo images containing dynamic object

Together, graphics dispersions of samples for every pixel lines in total have kind of figure 12.

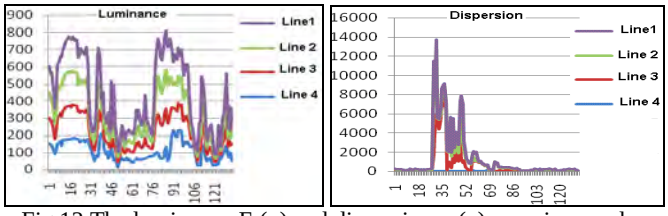

Fig.12 The luminance E (n) and dispersion σ (n) superimposed on each pixel for each row line of successive frames

On the right side are prominent peaks corresponding pixel in the contrast image area, but in contrast to the dynamic object are disposed in the same locality, and their position does not change in time, which can be taken into account in the subsequent separation of a dynamic object.

IV. Conclusion

The Image Processing Toolbox provided in MATLAB allowed the process of developing and testing the algorithm to be more efficient. Object detection and tracking has been an active research area for a long time because it is the initial important step in many different applications, such as video surveillance, face recognition, image enhancement, video coding, and energy conservation.

References

[1] Diaz J., E. Ros, F. Pelayo, *"Fpga-based real-time optical-flow system,"* Circuits and Systems for Video Technology, IEEE Transactions on, vol. 16, Feb. 2006

[2] Furi A., Hang H.M., *An efficient block-matching algorithm for motion compensated coding*, Proc. JSASSP, pp.1063-1066, 2007.

[3] Хуанга М. *Обработка изображении и цифровая фильтрация*, Под ред. Т:Мир 2009.

[4] Spirov R., *Practical Object Tracking System on FPGA*, 5-th Intern. Conference on Communications, Electromagnetic and Medical Application - CEMA2012, Athens, Greece, ISSN 1314-2100.

[5] Advanced Microcontroller Final Projects, or online at: http://people.ece.cornell.edu/land/courses/ece5760/FinalProjects