58th ILMENAU SCIENTIFIC COLLOQUIUM
Technische Universität Ilmenau, 08 – 12 September 2014
URN: urn:nbn:de:gbv:ilm1-2014iwk:3

EXTRACTION OF THE MOTION INDICATIONS IN THE SEQUENCE OF IMAGES

A. Mitsiukhin

Belarusian State University of Informatics & Radioelectronics,
Institute of Information Technology, Faculty of Computer Technology

ABSTRACT

The problem of extraction of the motion indications, i.e. coordinates and parameters of the motion in the sequence of images is considered. The reliability and precision of the extraction of the motion indications depends on the applied spatially-temporal algorithm of processing the images as well as on the noisiness of the images. This problem can be solved using the optimum or near-optimum processing algorithms. It is proposed to extract the moving objects on the images on the basis of the dyadic correlation approach. Detection and extraction of the dynamic object is implemented through intermediate calculations of the majority sequences describing the moving object. Here the scope of necessary calculation would be considerably reduces, should the distinctive structural properties of the applied transforms and signals be used.

Index Terms - extraction of the dynamic object, sequence of images, dyadic group, majority sequences, Walsh functions, dyadic-shift, noisy image, correlation function, Hamming distance.

1. INTRODUCTION

Let the video surveillance in the selected region of the space be performed at the moments of time t = 0,1, …, i, …, K–1. As a result, we have a sequence of K images $g_{y,x,t}$, where (y, x) are discrete spatial variables. The space is restricted by the image frame size $M \times N$, while the time – by the frame rate f and length of the sequence of images. Here the problem of extraction of the motion indications, i.e. estimation of the coordinates and parameters of the motion shall be solved with certain quality (probability of the correct extraction (detection) and measurement error).

2. THEORETICAL PRINCIPLES

The analysis of the images under consideration is based on the use of a cyclic group with the operation of dyadic shift on the finite intervals. In such a group, the additive modulo-two operation over the binary notations of decimal numbers is specified. The majority sequences on the dyadic group are formed with the help of the dyadic-shift operations [1]. The class of these sequences can be obtained by the majority addition of the Walsh functions $\{\mathrm{wal}_i(t)\}$, where the parameter i determines the function number, $i = 0,1,\dots,K-1.$ The discrete variable t determines the function count number and is to be written in the binary number system. To construct the majority sequences, it is necessary to select l functions from the set of Walsh functions $\{\mathrm{wal}_i(t)\}$, where l is an odd number. The selected functions will be designated as $\mathrm{wal}_{i_j}(t), j = 1,2,\dots,l.$. The Walsh functions with the index j have the period equal to

$$T = 2^{j-1}, j = 1,2,\dots,l. \quad (1)$$

The condition (1) is met by the Walsh function with one non-zero symbol in the binary code of value of the index j. Let the $\boldsymbol{b}_1, \boldsymbol{b}_2, \dots, \boldsymbol{b}_l$ designate the arbitrary vectors with the components $(b_1, b_2, \dots, b_l)$, $b_i \in \{1, -1\}$. The majority sequence is formed by summation of binary symbols according to the following rule [1]

$$a(t) = \operatorname{sign} \sum_{j=1}^{l} b_j \mathrm{wal}_{i_j}(t) = \begin{cases} 1, \sum_{j=1}^{l} b_j \mathrm{wal}_{i_j}(t) \geq 0; \\ -1, \sum_{j=1}^{l} b_j \mathrm{wal}_{i_j}(t) < 0. \end{cases} \quad (2)$$

Let the $\{a_i\} = \{a_0, \dots, a_{K-1}\}$ be a set of sequences of the cyclic groups, on which the dyadic-shift operation is determined. The whole set of sequences $\{a_i\}$ is obtained by specifying the coordinates of vectors $\boldsymbol{b} = (b_1, b_2, \dots, b_l)$. The number of different vectors $\boldsymbol{b}$ is 2^l. The dyadic shift of the sequence a_i transforms the latter into the sequence $a_i(t \oplus \tau) = a_\tau(t)$. The discrete variable τ determines the value of the dyadic shift of the function argument and is to be written in the binary code. The commutation of counts of the sequence a_i is actually performed on the interval consisting of $K = 2^l$ numbers. For example, the majority sequence (2) corresponds to three Walsh functions $(l = 3)$ with one non-zero symbol in the binary code, vector $\boldsymbol{b} = (b_1, b_2, b_3)$ with the coordinates $\boldsymbol{b} = (111)$:

$$a_0(t) = (111 - 11 - 1 - 1 - 1), \tau = 0 \quad (3)$$

Further, we shall use the following theorem.

Theorem [1]. When the vector $\boldsymbol{b}$ is fixed, the set of sequences $a(t)$ is invariant relatively to the dyadic shift of the argument. Then the whole set $\{a_i\}$ can be obtained by calculating the dyadic shifts of only one sequence corresponding to the fixed vector $\boldsymbol{b}$.

The set $\{a_i\}$ can be presented as a matrix $\boldsymbol{A}$. The rows of the matrix $\boldsymbol{A}$ are sequences $a_i(t \oplus \tau)$ arranged in the order of rising of magnitude of the dyadic shift $\tau = 0, 1, \dots, K - 1$. It is supposed to perform the extraction of the motion indication using the correlation approach on such set $\{a_i\}$.

We shall consider the analysis of the motion in the spatially-temporal images using the example with a single spatial coordinate and a coordinate reflecting the time of formation of frames (frame number), Fig. 1. The noisy image $g_{y,t}$ of the moving object with the spatial coordinate y is shown.

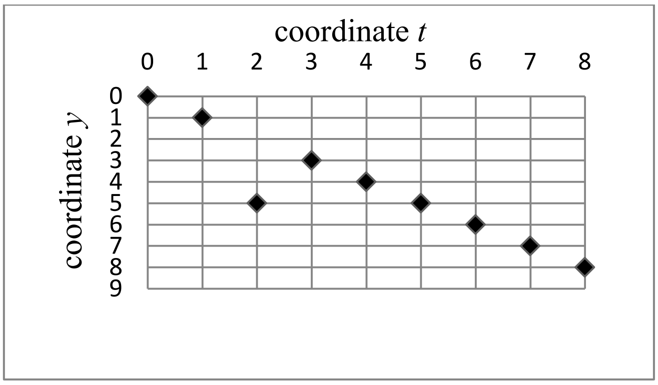

Figure 1. Spatially-temporal image of a moving object

Let $g_{t,y} = R$ be ratio on the $y \times t$. We shall consider the matrix presentation R of the form $M = [m_{ij}]$,

$$m_{ij} = \begin{cases} 1, if(y_i, t_j) \in R; \\ 0, if(y_i, t_j) \notin R. \end{cases}$$

To make the explanation more convenient, the object under observation is presented by a single pixel in each of the eight frames. The noisy binary image $g_{y,t}$ of the moving object with the spatial coordinate y is shown. The noise pixel of the image has the coordinates $(y,t) \rightarrow (5,2), m_{5,2} = 1$. Note that the horizontal (non-tilted) structure of brightness of the pixels corresponds to the motionless object.

Using the binary values m_{ij}, we shall form the sequence $y(t) = (y(0)y(1) \dots y(K-1))$. The values of samples of this sequence correspond to the coordinates of the mesh vectors of the frame images on the axis y. In the case under consideration, the mesh vector $\boldsymbol{y}$ describes the position of a single pixel of image of each frame on the axis y. The values of the vector components are equal to zero except for the component onto which the object is projected. Using the properties of the dyadic shift and dyadic ranking of the set $\{a_i\}$, the sequence K of images $g_{t,y}$ can be written in the form of projection onto the axis y can be written as follows:

$$\{y_i(t)\} = \{a_0(t \oplus \Delta\tau_{t,y})\}, i = 0,1, \dots, K-1. \tag{4}$$

where $\Delta\tau_t$ is the value of the dyadic shift for the time interval between the preceding frame and the subsequent one. For the time between the first frame and the second one (Fig. 1), the object is shifted by one pixel. So we obtain the value of sample $y_1(1)$ of the sequence $y_1(t)$. The shift of the object for the time interval between the two frames by j pixels along the axis y will correspond to the value of the sample $y_j(t)$. The j^{th} sequence of the set corresponds definitely to the constant value of the dyadic shift $\Delta\tau_t = j$. The value of the shift $\Delta\tau_t = j$ is proportional to the y-component of the object motion velocity. To calculate the y-component of the velocity, when the inter-pixel distance known, the value $\Delta\tau_t = \tau$ should be found. The dyadic delay τ (shift) can be determined by calculating the coefficients of the dyadic correlation function $r(t)_y$ of the sequence $y(t)$,

$$r(\tau)_y = \sum_{t=0}^{K-1} y(t)\, a(t \oplus \tau), \ \tau = 0,1, \dots, K-1. \tag{5}$$

The value of the coefficient of the correlation function depends on the position of the moving object on the image and shift τ. The determination of the shift comes to the comparison of the sequence $y(t)$ with each sequence of the set $\{a_i\}$ and selection of the nearest of them as regards the Hamming distance. The maximum value of the coefficient $r(t)_y$ determines the value of the argument τ_y. Should the frame rate f, inter-pixel distance d and value τ_y corresponding to the maximum coefficient correlation $r(t)_y$ be known, the motion direction and y-component of the velocity can be determined [2].

The matrix form

$$\boldsymbol{R}_y = \boldsymbol{A}\boldsymbol{y}, \tag{6}$$

where $\boldsymbol{y}$ is the vector notation of the sequence $y(t)$ corresponds to the expression (5).

Similar reasoning is also valid for obtaining the coefficients of the dyadic correlation function and calculations of the parameters of motion in the direction of the x axis. The correlation coefficients along the x axis are calculated from the formula:

$$r(\tau)_x = \sum_{t=0}^{K-1} y(t)\, a(t \oplus \tau), \ \tau = 0,1, \dots, K-1, \tag{7}$$

where $\{x_i(t)\} = \{a_0(t \oplus \Delta\tau_{t,x})\}, i = 0,1, \dots, K-1.$ (8)

$$\boldsymbol{R}_x = \boldsymbol{A}\boldsymbol{x}. \tag{9}$$

3. THE SIMULATION OF THE PROCESS OF FINDING THE MOVING OBJECT IN THE IMAGES DISTORTED BY THE NOISE

Let us illustrate the process of finding a moving object in the spatially-temporal image, Fig 1 on an example.

3.1. Exampl

Let $a_0(t)$ (3) be used as a sequence with zero dyadic shift, $\boldsymbol{b} = (111)$, $\tau = 0$. The whole set of majority sequences $\{a_i\}$ written in alphabetical order 1, –1 be mapped in the form of a square matrix $\boldsymbol{A}$ with the size $\times\, K$ (8×8), Fig. 2.

$$A = \begin{bmatrix} 1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 \\ 1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 \\ -1 & 1 & 1 & 1 & -1 & -1 & -1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 \\ -1 & 1 & -1 & -1 & 1 & 1 & -1 & 1 \\ -1 & -1 & 1 & -1 & 1 & -1 & 1 & 1 \\ -1 & -1 & -1 & 1 & -1 & 1 & 1 & 1 \end{bmatrix}.$$

Figure 2. Structure of majority sequences

The sequence (4) corresponds to the spatially-temporal image presented in Fig. 1

$$y_1(t) = a_0(t \oplus \Delta\tau_{t,y});\ y_1 = (1111-11-1-1).$$

The noise component of the image introduces an error into the sequence $y_1(t)$. Multiplying $\boldsymbol{A}$ by the vector $\boldsymbol{y}$ results in

$$\boldsymbol{R}_y = \boldsymbol{A}\boldsymbol{y} = (r_0 r_1 \dots r_{K-1})^T = (2, 6, 2, 2, -2, -2, -6, -2)^T. \quad (10)$$

Graph of the continuous correlation function (10) is shown in Fig. 3.

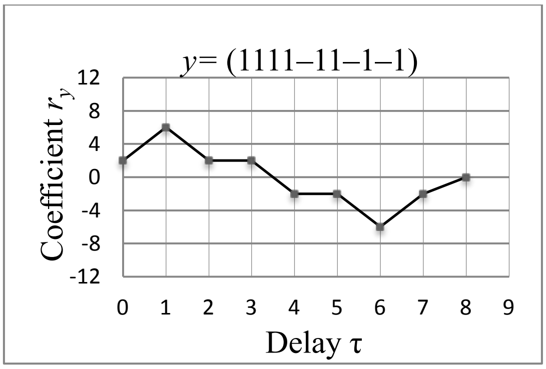

Figure 3. Correlation function of the distorted majority sequence

For comparison, Fig. 3 shows the view of correlation function of the non-distorted sequence reflecting the motion of the object along the y axis.

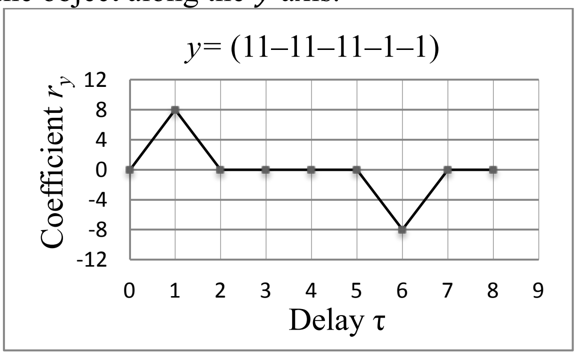

Figure 4. Correlation function of the majority sequence y_1

In the expression (10), the first component, i.e. $\tau_y = 1$ has the maximum value. The relationship between the correlation coefficient and velocity component in the direction of the y axis is established by the expression [2]

$$u_y = \tau_y df.$$

As seen from the example, the presence of noise on the image caused no effect on the correctness of extraction of the indication of the object motion. Because the minimum Hamming distance of the rows of the $\boldsymbol{A}$ is equal to $\mathrm{dist}_{\min}\boldsymbol{A} = \frac{K}{2}$, the $\lceil(\frac{K}{2} - 1)/2\rceil$ false components (distortions) in each coordinate direction can be removed.

3.2. Exampl

The experiment was performed on the interval of $K = 2^5$ numbers. The vector $\boldsymbol{b} = (11111)$ was chosen as a reference vector with the zero delay τ = 0. The problem of finding the dynamic object was solved when masking the latter with pulsed unipolar noise ("salt-and-pepper"). In Fig. 4, the object moves towards the top right corner of the image with the relative velocity of $u = \sqrt{5}$ pixels per frame.

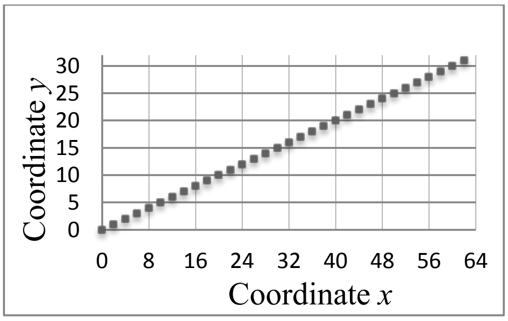

Figure 5. Superposed frames of the object motion images

The sequences $a_0(t), a_1(t)$ and $a_2(t)$ correspond to the image shown in Fig. 5. The following majority sequences correspond to the vectors:

$$\boldsymbol{b} = (11111) \to a_0(t) = (1111111 - 1111 - 11 - 1 - 1 - 1111 - 11 - 1 - 1 - 11 - 1 - 1 - 1 - 1 - 1 - 1 - 1),\ \tau = 0;$$

$$\boldsymbol{b} = (1111 - 1) \to a_1(t) = a_0(t \oplus 1) = (111111 - 1111 - 11 - 11 - 1 - 111 - 1\ 1 - 11 - 1 - 1 - 11 - 1 - 1 - 1 - 1 - 1 - 1),\ \tau = 1.$$

$$\boldsymbol{b} = (111 - 11) \to a_2(t) = a_0(t \oplus 2) = (11111 - 1111 - 111 - 1 - 11 - 11 - 1\ 11 - 1 - 11 - 1 - 1 - 11 - 1 - 1 - 1 - 1 - 1),\ \tau = 2.$$

Fig. 6 and 7 show the distorted half-tone images with the noise probability densities of $d = 0{,}25$ and $d = 0{,}45$, respectively. The distortions are caused by the presence of the additive noise $\eta(x, y, t)$,

$$a(x, y, t) = g(x, y, t) + \eta(x, y, t).$$

After the threshold conversion of the half-tone image shown in Fig. 7 into the binary one, the scene shown in Fig. 8 was obtained. The brightness threshold value of $T = 0{,}1$ was used. As seen, the images shown in Fig. 7 and 8 are heavily distorted by noise.

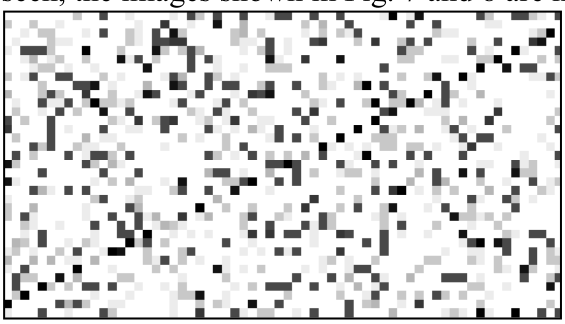

Figure 6. Image distorted by the pulsed noise, $d = 0{,}25$

Figure 7. Image distorted by the pulsed noise, $d = 0{,}45$.

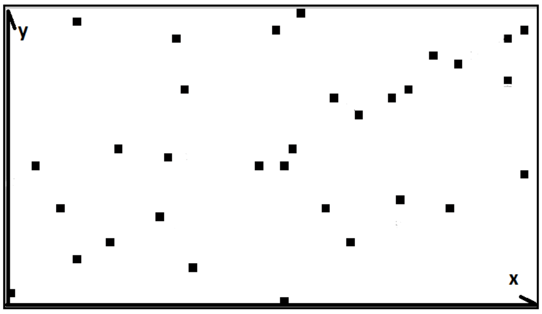

Figure 8. Result of the threshold conversion

Using the values of the coordinates of the axis y we shall create a one-dimensional sequence

$$y_1(t) = a_1(t) + \eta_y(t) = (-1111111111 - 1 - 1 - 11 - 11 - 11 - 1$$
$$-1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1).$$

Using the values of the coordinates of the axis x we shall create a one-dimensional sequence

$$y_2(t) = a_2(t) + \eta_x(t) = (11 - 111 - 1 - 111111 - 1 - 1 - 1 - 11 - 11$$
$$-1111 - 1 - 1 - 11 - 1 - 1 - 1 - 1 - 1).$$

The calculation of the correlation functions $\boldsymbol{R}_y$ and $\boldsymbol{R}_x$ makes it possible to determine definitely the numbers of correlation coefficients with the maximum values of

$$r_{\max,y} = r_{1,y} = 17; r_{\max,x} = r_{2,x} = 16.$$

The value of the object's relative velocity $u = \sqrt{5}$ is equal to the true value. It should be noted that the array of correlation coefficients along the x axis contains the components with the values near to $r_{2,x}$. It is evidence that the more reliable extraction of the motion indications requires the elongation of the observation time and the transition to longer sequences. In case of work with the longer sequences ($K \geq 2^7$), special structural properties of majority sequences make it possible to implement the fast algorithm of multiplying a vector by a matrix when calculating the correlation functions $\boldsymbol{R}_y$ and $\boldsymbol{R}_x$. The number of necessary computing operations decreases from $\sim K^2$ down to $\sim K \log_2 K$. The volume of computations can be considerably reduced.

CONCLUSION

The correlation method for extracting the motion indications in the sequence of images using the majority sequences determined on the dyadic group has been considered. The sufficiently reliable extraction of indications of the motion of hardly distinguishable objects can be provided by increasing the time of observation and length of the majority functions. Such an approach makes it possible to obtain the precision characteristics of the processing which are near to the optimum ones. In so doing, the acceptable computational and hardware-related complexity of processing is ensured.

REFERENCES:

[1] V.V. Losev, E.B. Brodskaya and V.I. Korzhik; edited by V.I. Korzhik, Search and Decoding of the Complex Discrete Signals, Radio and Communications, Moscow, 1988.
[2] A. Mitsiukhin, "Segmentation of Dynamical Images by Means of Discrete Hartley Transform", Proceedings 56th IWK, TU Ilmenau, – URN (Paper): urn: nbn: de: gbv: ilm1-2011iwk-011: 5, id 1100. – P. 1– 4.

CONTACT

Doz. A. Mitsiukhin mityuhin@bsuir.by