ENGINEERING TRANSACTIONS • Engng. Trans. • 39, 2, 139-161, 1991
Polish Academy of Sciences • Institute of Fundamental Technological Research

CHANGE DETECTION IN TWO IMAGES USING WALSH FUNCTIONS AND THE LIKELIHOOD TEST METHOD

M. NIENIEWSKI (WARSZAWA) and
P. K. PATHAK (ALBUQUERQUE)

The gray-value function in a local window is approximated by a series of the Walsh functions. Several terms of this series adequately model the gray-value function in the window, and the remaining variability can be attributed to noise. When comparing two images, one finds the coefficients of the Walsh functions for each consecutive position of the window in both images. Detection of the change between the corresponding windows is based on the maximum likelihood ratio testing, which consists in deciding between two hypotheses: H_0 – there is no motion, versus H_1 – there is motion. The results of change detection are presented for images of a natural scene, obtained by means of the CCD camera, as well as for simulated images subjected to the influence of Gaussian noise. The computer program described carries out the F-test for the maximum likelihood ratio. It is shown that such a program can be successfully used for change detection in image sequences.

1. Related work

Image sequences contain information about dynamic changes in a scene, which result, for example, from the relative motion of the camera and the objects in the scene. Before attempting any interpretation of the observed changes, it is necessary to carry out some preprocessing which determines if these changes can in fact be attributed to motion. This necessity becomes quite apparent if one considers that even taking two succesive pictures of the same scene usually does not result in exactly the same distribution of the gray values in the images.

Image sequence analysis has received considerable attention over the recent years, as illustrated by the references at the end of this paper, and includes diversified methods which are used in various situations. In the following, a short review of these methods, their applications and examples of images on which they were tested are given.

Most commonly, motion detection and estimation is used in dynamic scene analysis, moving object tracking, surveillance systems, robot navigation, obstacle avoidance, and in the coding of image sequences for efficient

transmission and storage of images. An introduction to motion detection is given in [1] and [6]. In particular, the author in [6] explains the concept of the optical flow, that is of the field of velocity vectors representing the motion of each point in the image. The work [1] extends this concept and considers also the problems of motion understanding. In the paper [7], an iterative procedure for calculation of the optical flow is given and illustrated by using several synthetic image sequences. In the paper [13], the optical flow for a natural scene of a ski jumper moving across the screen is calculated.

In the work [9] a technique of segmentation of a scene into stationary and nonstationary components based on using a moving camera is presented. The example of the scene contains both moving and nonmoving mechanical parts and tools in a laboratory environment.

In the paper [25], changes in images are detected in two steps. In the first step, a sample of the moving object is found by forming the difference image and finding a set of points enclosed by this image. In the second step this set is increased by means of a region growing technique so that the whole of the moving object is recovered. The method was proven to work well for polygons. Applications to simple laboratory scenes are presented. In the paper [10], a two-step procedure is given for processing image sequences in the case of translational motion between the camera and the objects. The first step, which is a feature extraction process, picks out small areas corresponding to distinguishing parts of the objects. In the second step, the direction of translational motion is found by a search for the image displacement paths along which a measure of feature mismatch is minimized. The examples given refer to roadsigns as read by a gyroscopically stabilized camera in a traveling car, as well as to the image sequence obtained with the camera held by a robot manipulator translating towards some industrial parts lying on the table.

In the work [20], a time-varying corner detector is described, which is based on the AND operation between the cornerness and the temporal derivative. The corner detector finds the measure of the cornerness in individual images, and the temporal derivative is calculated as a difference of the gray values at a point. Examples of image sequences include simulated images with random noise, and natural scenes depicting a moving toy car. Superposition of time-varying corners with the edges of the object renders the shape of the moving object. In the paper [2], the maximum likelihood method for the determination of moving edges is presented. A spatiotemporal edge in an image sequence is modeled as a surface patch in $3D$ spatiotemporal space. A likelihood test is developed, which enables one to detect the moving edges and to estimate their attributes. The method reduces to convolving the image with a set of masks. The experiments presented were carried out with sequences of urban scenes and of laboratory environment.

In the paper[11], a method and a circuit are described for extracting an

estimate of the velocity of a moving object from the television signal. The aim of the reported work was to improve efficiency of frame-to-frame coding by sending only those parts of a frame which have changed in comparison with the previous frame. The examples given refer to images of mannequins moved horizontally across the screen. The paper [3] extends the ideas of the work [11]. The algorithm presented in the paper [3] consists of two steps. The first step consists in the segmentation of the image into fixed and moving areas by means of dynamical programming. The second step, that is the estimation of the displacements, is based on the linear estimation technique.

In the paper [12], a method is presented for motion detection and velocity computation of large moving objects. The two-dimensional image sequence is projected on two spatial axes. This results in the generation of time sequences, and for both spatial axes the two-dimensional fast Fourier transforms are found. An estimate of the velocity of the moving object is calculated based on the analysis of the spectrum of both the spatial and temporal frequencies. The method is applicable to the tracking of multitargets in video data, of dust storms and clouds in weather forcasts, as well as to highway traffic monitoring, and to the control of robots and autonomous vehicles. In the work [16], a comparison of several algorithms for moving target identification is presented. These algorithms were tested on infrared images obtained from the sensor mounted on the geosynchronous orbit. In particular, some of these algorithms aim at removing the influence of the moving background, such as wind-driven clouds, while retaining the sensitivity with regard to detecting moving objects. In the paper [22], an algorithm is presented for the detection of small changes in radar images of the same terrain taken from approximately the same point of view at different times. The method removes local spatial distortions of one image with respect to the other. The two images are then brought into correspondence by matching the statistics of their gray-value functions. Subsequently, the gray values in one image are subtracted from the gray values in the other image. As an example, detection of a moving tugboat and a barge in a port is shown. In the paper [18], a method is proposed for the tracking algorithm. In this method, two Fourier transformations are used. The first one is a space to inverse-space transform, and the second one is a time to frequency transform, with separate transformations being performed for each of the two spatial axes. Numerical experiments presented indicate that the noise immunity can be improved by averaging frequency spectra. The method was developed for detecting a moving aircraft by means of a sensor mounted on a space platform.

2. Background

This paper presents the theory and a more comprehensive description

of the experiments, which were previously reported by the authors at the ICSP'90 ([15]). This paper extends and develops the concepts given by the authors of the paper [8] and [14]. In the paper [8],the change detection is based on the approximation of the gray-value function in a local window by a bivariate polynomial of the second degree. As an application of this approach, an example of a sequence of images depicting a street scene with a car moving across the screen is given in [8]. The estimation of the displacement vector based on the motion detection technique of [8] is discussed in the paper [14].

In the current paper the gray-value function of the image is approximated by the series of the Walsh functions of pixel coordinates. The advantage of using the Walsh functions rather than the bivariate polynomials is the simplification of calculations with equally satisfactory results, and the potential possibility of modifying the window size as well as the number of Walsh functions used. The coefficients of the Walsh functions are found for successive positions of the window in the image. It was established that one convenient size of the window is 4×4, in which the gray-value function is approximated by the series containing the first 3×3 terms. These nine terms adequately model the gray-value function, and the remaining variability of this function can be attributed to the noise generated by sensing and digitizing devices. When comparing two images, the coefficients of the Walsh functions for corresponding windows in both images are calculated. Detection of a change between the two windows is based on the test of the maximum likelihood ratio. This test is similar to the test for the local edge operator described in the work [24]. More specifically, the maximum likelihood test presented below decides between two hypotheses:

H_0, which means that the observed gray values in both windows are compatible and come from the same distribution, and hence there was no motion in the window.

H_1, which means that the gray values in both windows are incompatible and come from two different distributions, and hence there was a motion.

The maximum likelihood test described below is based on the assumption that the gray values are corrupted by additive noise with mean value zero and some unspecified variance, and that there is no correlation between noise at various pixels of the image.

3. Theoretical development of the method for detecting changes in images

The gray-value image function, denoted by $f(x, y)$, represents the brightness of the image at the point with coordinates x, y. These coordinates specify the row and the column in the image, and their values are natural numbers. The brightness is also expressed as an integer, in the range from

zero (black) to some maximum (white). The values of $f(x,y)$ are usually given by the matrix

$$\begin{bmatrix} f(0,0) & \dots & f(0,N-1) \\ \dots\dots & \dots & \dots\dots \\ f(N-1,0) & \dots & f(N-1,N-1) \end{bmatrix} . \tag{3.1}$$

Typically, it as assumed that the matrix (3.1) represents a square table of size which is some power of two: $N = 2^k$, where k – natural number. It is also assumed that the number of possible gray levels is a power of two: $G = 2^l$, where l – natural number. The $f(x,y)$ can be approximated by a linear expression of the form

$$\begin{aligned} f^*(x,y) = {} & c_{0,0}\varphi_{0,0}(x,y) + \dots + c_{0,n-1}\varphi_{0,n-1}(x,y) \\ & + \dots + c_{n-1,0}\varphi_{n-1,0}(x,y) + \dots + c_{n-1,n-1}\varphi_{n-1,n-1}(x,y) . \end{aligned} \tag{3.2}$$

In the above equation $f^*(x,y)$ denotes the approximating function; $\varphi_{0,0}(x,y), \dots, \varphi_{n-1,n-1}(x,y)$ denote the set of some known assumed functions; $c_{0,0}, \dots, c_{n-1,n-1}$ are the coefficients to be calculated, and $n \leq N$. The equation

$$f(x,y) = f^*(x,y) \tag{3.3}$$

should be satisfied with maximum accuracy at all pixels for which the value of $f(x,y)$ is given. Denoting the number of these pixels simply by p, one can write out p equations of the type (3.3). Usually, the number p is greater than the number $n \times n$ of unknowns, which results in an overdetermined system of equations. The solution to such a system is found by means of the least squares method ([4]). In the case when $\varphi_{0,0}, \dots, \varphi_{n-1,n-1}$ form an orthonormal set of functions, the calculations are simpler and the coefficients $c_{u,v}$ (where $0 \leq u, v \leq n-1$) are found from the equation

$$c_{u,v} = \sum_{i=0}^{p-1} f(x_i, y_i)\varphi_{u,v}(x_i, y_i) . \tag{3.4}$$

In principle the $\varphi_{u,v}$ could be arbitrary orthonormal functions, but in order to make calculations as simple as possible, it is assumed in the following that $\varphi_{u,v}$ are the Walsh functions $g(x,u;y,v)$, respectively, of order u, v,

$$\varphi_{u,v}(x,y) = g(x,u;y,v) , \tag{3.5}$$

and are defined by the equation ([5, 17, 19])

$$g(x,u;y,v) = \frac{1}{N} \prod_{i=0}^{m-1} (-1)^{b_i(x)b_{m-1-i}(u) + b_i(y)b_{m-1-i}(v)} , \tag{3.6}$$

where m – number of bits used in the binary representation of x (and y). The function $b_i(x)$ in the above equation, for a given argument x (and similarly $b_i(y)$), has the value of the i-th bit in the binary representation of x. For example, for $m = 3$ and $x = 6$ decimal (which is equivalent to 110 binary), one obtains $b_0(6) = 0$, $b_1(6) = 1$, $b_2(6) = 1$.

For any u and v, the values of $g(x, u; y, v)$ are equal to $(1/N)$ with a positive or negative sign. It can also be proved that the Walsh functions (3.6) form an orthonormal set of functions.

The following equations are derived for particular values $N = 4$ and $n = 3$. For any other N or n, the equations would be similar. The value of the gray-value function $f_1(x, y)$ at any pixel (x, y) of the first image can be expressed as

$$f_1(x,y) = \sum_{u=0}^{2}\sum_{v=0}^{2} c_{u,v} g(x,u;y,v) + e(x,y), \tag{3.7}$$

where the term $e(x, y)$ denotes a random component of $f_1(x, y)$. It is assumed in accordance with the papers [8] and [24] that the variable $e(x, y)$ has a normal distribution $N(0, \sigma_1^2)$ with mean value zero and variance σ_1^2. Although this assumption may not be satisfied in a general case, the examples of images of a natural scene in the papers [8] and [24] confirm that it leads to reasonable results in quite common situations. Under these circumstances, the probability density f_e of the variable $e(x, y)$ can now be written as

$$f_e\,[e(x,y)] = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left[-\frac{e^2(x,y)}{2\sigma_1^2}\right]. \tag{3.8}$$

It is further assumed ([8, 24]) that in many practical cases the random variables $e(x, y)$ at various pixels are independently distributed. It follows then that the joint probability density for several pixels is equal to the product of their individual probability densities as given by Eq.(3.8). Hence for $N = 4$, the likelihood function L, which is currently defined as the joint probability density for all the pixels in the window in image 1, can be represented as

$$L(f_1, c, \sigma_1) = \left[\frac{1}{2\pi\sigma_1^2}\right]^8 \times \exp\left\{-\frac{\sum_{x=0}^{3}\sum_{y=0}^{3}\left[f_1(x,y) - \sum_{u=0}^{2}\sum_{v=0}^{2} c_{u,v} g(x,u;y,v)\right]^2}{2\sigma_1^2}\right\}, \tag{3.9}$$

and depends on $f_1(x, y)$, the coefficients $c_{u,v}$, and σ_1. Using Eq.(3.9) directly would result in rather awkward calculations. Taking the natural logarithm of both sides of Eq.(3.9) leads to much simpler expressions

$$\ln L(f_1, c, \sigma_1) = -8\ln(2\pi) - 16\ln(\sigma_1) - \frac{1}{2\sigma_1^2}\left[\sum_{x=0}^{3}\sum_{y=0}^{3}\left[f_1(x,y) - \sum_{u=0}^{2}\sum_{v=0}^{2} c_{u,v}\, g(x,u;y,v)\right]^2\right]. \tag{3.10}$$

The coefficients $c_{u,v}$ and variance σ_1^2 are unknown, and their estimates are calculated by finding the maximum of the likelihood function for the given statistical sample, which in this case represents the structure of the gray values in the window. This maximum corresponds to the zero value of the derivatives of $\ln L$ with respect to $c_{u,v}$ and σ_1. Differentiating $\ln L$ with respect to $c_{u,v}$ and setting the derivative equal to zero gives

$$\frac{\delta \ln L}{\delta c_{u,v}} = \frac{1}{\sigma_1^2}\sum_{x=0}^{3}\sum_{y=0}^{3}\left[\left[f_1(x,y) - c_{u,v}\, g(x,u;y,v)\right] g(x,u;y,v)\right] = 0. \tag{3.11}$$

Performing multiplications as indicated in Eq.(3.11) and using the orthonormality property of the Walsh functions, one obtains the estimate of $c_{u,v}$ in the form

$$\hat{c}_{u,v} = \sum_{x=0}^{3}\sum_{y=0}^{3} f_1(x,y)\, g(x,u;y,v)\,, \tag{3.12}$$

which agrees with Eq.(3.4). Differentiating $\ln L$ with respect to σ_1 and setting the derivative equal to zero gives

$$\frac{\delta \ln L}{\delta \sigma_1} = -\frac{16}{\sigma_1} + \frac{1}{\sigma_1^3}\left[\sum_{x=0}^{3}\sum_{y=0}^{3}\left[f_1(x,y) - \sum_{u=0}^{2}\sum_{v=0}^{2} c_{u,v}\, g(x,u;y,v)\right]^2\right] = 0. \tag{3.13}$$

Raising the expression in the inner brackets in Eq.(3.13) to the second power, substituting from Eq.(3.12) into Eq.(3.13), and using the orthonormality property of the Walsh functions, one obtains the estimate of σ_1^2 in the form

$$\hat{\sigma}_1^2 = \frac{1}{16}\left[\sum_{x=0}^{3}\sum_{y=0}^{3} f_1^2(x,y) - \sum_{u=0}^{2}\sum_{v=0}^{2}\hat{c}_{u,v}^2\right]. \tag{3.14}$$

The following notation is subsequently used for the two hypotheses:

H_0 – the two samples come from the same distribution. The estimates of the coefficients of the Walsh series and of the variance are, respectively, $\hat{c}^*_{0,0}, \ldots, \hat{c}^*_{2,2}$, and $\hat{\sigma}_0^2$;

H_1 – the two samples come from two different distributions. The estimates for the first sample are $\hat{c}_{0,0}, ..., \hat{c}_{2,2}$, and $\hat{\sigma}_1^2$. Similarly the estimates for the second sample are $\hat{d}_{0,0}, ..., \hat{d}_{2,2}$, and $\hat{\sigma}_2^2$.

Summing up, assuming that the hypothesis H_1 is true, one finds the estimates $\hat{c}_{u,v}$ of the coefficients of the Walsh series for the first window by substitutions into Eq.(3.12), and the estimates $\hat{d}_{u,v}$ for the second window by substitutions into the equation

$$\hat{d}_{u,v} = \sum_{x=0}^{3}\sum_{y=0}^{3} f_2(x,y)g(x,u;y,v)\,, \tag{3.15}$$

where $f_2(x,y)$ – gray-level function in the second window.

Similarly, under H_1 one uses Eq.(3.14) for the estimate of σ_1^2 and

$$\hat{\sigma}_2^2 = \frac{1}{16}\left[\sum_{x=0}^{3}\sum_{y=0}^{3} f_2^2(x,y) - \sum_{u=0}^{2}\sum_{v=0}^{2} \hat{d}_{u,v}^{\,2}\right] \tag{3.16}$$

for the estimate of σ_2^2.

Assuming now that H_0 is true, one has to replace (3.9) by the following equation for the common likelihood function for both windows

$$L(f_1, f_2, c^*, \sigma_0) = \left[\frac{1}{2\pi\sigma_0^2}\right]^{16} \times \exp\left\{-\frac{\sum_{i=1}^{2}\sum_{x=0}^{3}\sum_{y=0}^{3}\left[f_i(x,y) - \sum_{u=0}^{2}\sum_{v=0}^{2} c_{u,v}^* g(x,u;y,v)\right]^2}{2\sigma_0^2}\right\}, \tag{3.17}$$

which reflects the fact that the current statistical sample consists of the gray values in two windows.

The estimates of $c_{u,v}^*$ and σ_0^2 are now found similarly as was the case when the hypothesis H_1 was assumed true. Differentiating the logarithm of Eq.(3.17) with respect to $c_{u,v}^*$ and σ_0 and setting the derivatives equal to zero, one finds after some algebraic manipulations the following estimates:

$$\hat{c}_{u,v}^* = \frac{1}{2}\sum_{x=0}^{3}\sum_{y=0}^{3}\left[f_1(x,y) + f_2(x,y)\right] g(x,u;y,v) = \frac{\hat{c}_{u,v} + \hat{d}_{u,v}}{2} \tag{3.18}$$

and

$$\hat{\sigma}_0^2 = \frac{1}{32}\left[\sum_{x=0}^{3}\sum_{y=0}^{3}\left[f_1^2(x,y) + f_2^2(x,y)\right] - 2\sum_{u=0}^{2}\sum_{v=0}^{2}\hat{c}_{u,v}^{*\,2}\right]. \tag{3.19}$$

Subsequently one has to find the maximum likelihood estimate of the structure of the gray values in the window given that H_0 is true, or that H_1 is true.

The case when H_1 is true will be considered first. Since the estimates $\hat{c}_{u,v}$ and $\hat{\sigma}_1^2$ are known, the joint probability density of getting the given values of $f_1(x,y)$ under H_1 can be written as

$$P_1 = \left[\frac{1}{2\pi\sigma_1^2}\right]^8 \exp\left\{-\frac{\sum_{x=0}^{3}\sum_{y=0}^{3}\left[f_1(x,y) - \sum_{u=0}^{2}\sum_{v=0}^{2}\hat{c}_{u,v}g(x,u;y,v)\right]^2}{2\hat{\sigma}_1^2}\right\} = \left[1/(2\pi)^8\right] e^{-8}\hat{\sigma}_1^{-16}, \tag{3.20}$$

where the right-most expression in Eq.(3.20) is obtained by substituting from Eq.(3.14) into the expression in the middle of Eq.(3.20).

Similarly, the joint probability density of getting the given values of $f_2(x,y)$ under H_1 is

$$P_2 = \left[1/(2\pi)^8\right] e^{-8}\hat{\sigma}_2^{-16}\,. \tag{3.21}$$

It follows then that the joint probability density of getting the given values of $f_1(x,y)$ and $f_2(x,y)$ under H_1 is equal to the product P_1P_2.

Assuming that H_0 is true, one arrives at a single equation for the joint probability density of getting the given values of $f_1(x,y)$ and $f_2(x,y)$ rather than two equations. By analogy with the case of H_1 being true, one obtains that the joint probability density for the sample consisting of the two windows can be expressed as

$$P_0 = \left[1/(2\pi)^{16}\right] e^{-16}\hat{\sigma}_0^{-32}\,. \tag{3.22}$$

In order to derive the statistical test for change detection, one defines the maximum likelihood ratio λ, which is the following function of P_0, P_1, and P_2 ([8, 24]):

$$\lambda = (P_1P_2)/P_0\,. \tag{3.23}$$

Substituting from Eqs.(3.20) – (3.22) into Eq.(3.23) would result in an awkward equation. However, this equation can be simplified by raising λ to the power of 1/8 so that

$$\lambda^{1/8} = \frac{\hat{\sigma}_0^4}{\hat{\sigma}_1^2\hat{\sigma}_2^2}\,. \tag{3.24}$$

When trying to evaluate Eq.(3.24), one has to find the most convenient expressions for $\hat{\sigma}_0^2$, $\hat{\sigma}_1^2$, and $\hat{\sigma}_2^2$. Substituting from Eqs.(3.14), (3.16) and (3.18) into Eq.(3.19), one obtains after some algebraic manipulations

$$\hat{\sigma}_0^2 = \frac{\hat{\sigma}_1^2 + \hat{\sigma}_2^2}{2} + \sum_{u=0}^{2}\sum_{v=0}^{2}\left[\frac{\hat{c}_{u,v} - \hat{d}_{u,v}}{2}\right]^2 . \tag{3.25}$$

Substituting from Eq.(3.25) into Eq.(3.24) gives

$$\lambda^{1/8} = \frac{\left[\frac{\hat{\sigma}_1^2 + \hat{\sigma}_2^2}{2} + \sum_{u=0}^{2}\sum_{v=0}^{2}\left[\frac{\hat{c}_{u,v} - \hat{d}_{u,v}}{2}\right]^2\right]^2}{\hat{\sigma}_1^2\,\hat{\sigma}_2^2} . \tag{3.26}$$

It will be assumed in further developments that the hypothesis H_0 is true.

For simplicity reasons the two estimates of variances, $\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$, in Eq.(3.26) can be replaced by one common estimate, equal to their average,

$$\hat{\sigma}_1^2 = \hat{\sigma}_2^2 = \frac{1}{32}\left[\sum_{x=0}^{3}\sum_{y=0}^{3}\left[f_1^2(x,y) + f_2^2(x,y)\right] - \sum_{u=0}^{2}\sum_{v=0}^{2}\left[\hat{c}_{u,v}^2 + \hat{d}_{u,v}^2\right]\right] . \tag{3.27}$$

Substituting from Eq.(3.27) into Eq.(3.26) results in

$$\lambda^{1/8} = 1 + F_1^2 , \tag{3.28}$$

where the variable F_1 is defined as

$$F_1 = \frac{\frac{1}{2}\sum_{u=0}^{2}\sum_{v=0}^{2}\left[\hat{c}_{u,v} - \hat{d}_{u,v}\right]^2}{\sum_{x=0}^{3}\sum_{y=0}^{3}\left[f_1^2(x,y) + f_2^2(x,y)\right] - \sum_{u=0}^{2}\sum_{v=0}^{2}\left[\hat{c}_{u,v}^2 + \hat{d}_{u,v}^2\right]} . \tag{3.29}$$

Since the variable λ is a monotonic function of the variable F_1, one can use F_1 instead of λ in statistical testing. For this purpose the distribution of the random variable F_1 is found provided that H_0 is true. Expressing F_1 as a ratio of its numerator and denominator

$$F_1 = N_1/D_1 , \tag{3.30}$$

one obtains the following results from the statistical analysis of N_1 and D_1.

Substituting from Eq.(3.7) into Eq.(3.12) leads to

$$\hat{c}_{u,v} = c_{u,v} + \sum_{x=0}^{3}\sum_{y=0}^{3} e(x,y)g(x,u;y,v) . \tag{3.31}$$

Since $e(x,y)$ has a distribution $N(0,\sigma_0^2)$, the above equation implies that the difference $c_{u,v} - \hat{c}_{u,v}$ has a distribution $N(0,\sigma_0^2)$, and a similar observation is

valid for $d_{u,v} - \hat{d}_{u,v}$. Hence, since it was assumed that H_0 is true, it follows that $c_{u,v} = d_{u,v}$, and the difference

$$\hat{c}_{u,v} - \hat{d}_{u,v} \tag{3.32}$$

has a distribution $N(0, 2\sigma_0^2)$. The auxiliary variable (N_1/σ_0^2), which represents the numerator of Eq.(3.30) in the appropriately normalized form, is distributed as χ^2 (chi square) with 9 degrees of freedom. By analogy with Eqs.(3.13) and (3.14), one can write for D_1

$$D_1 = \sum_{x=0}^{3}\sum_{y=0}^{3}\left[f_1(x,y) - \sum_{u=0}^{2}\sum_{v=0}^{2}\hat{c}_{u,v}g(x,u;y,v)\right]^2 + \sum_{x=0}^{3}\sum_{y=0}^{3}\left[f_2(x,y) - \sum_{u=0}^{2}\sum_{v=0}^{2}\hat{d}_{u,v}g(x,u;y,v)\right]^2 . \tag{3.33}$$

Combining Eq.(3.7) and an analogous expression for $f_2(x,y)$ together with Eq.(3.33) gives

$$D_1 = \sum_{x=0}^{3}\sum_{y=0}^{3}\left[e(x,y) + \sum_{u=0}^{2}\sum_{v=0}^{2}[c_{u,v} - \hat{c}_{u,v}]\,g(x,u;y,v)\right]^2 + \sum_{x=0}^{3}\sum_{y=0}^{3}\left[e(x,y) + \sum_{u=0}^{2}\sum_{v=0}^{2}\left[d_{u,v} - \hat{d}_{u,v}\right]g(x,u;y,v)\right]^2 . \tag{3.34}$$

Since $c_{u,v} - \hat{c}_{u,v}$ has a distribution $N(0, \sigma_0^2)$, it follows that the sum

$$\sum_{u=0}^{2}\sum_{v=0}^{2}[c_{u,v} - \hat{c}_{u,v}]\,g(x,u;y,v) \tag{3.35}$$

has a distribution $N\left[0, \frac{9}{16}\sigma_0^2\right]$, and that the expression

$$e(x,y) + \sum_{u=0}^{2}\sum_{v=0}^{2}[c_{u,v} - \hat{c}_{u,v}]\,g(x,u;y,v)\,, \tag{3.36}$$

which represents a typical term in the first double outer summation on the right-hand side of Eq.(3.34), has a distribution with mean zero and variance $\left[1 + \frac{9}{16}\right]\sigma_0^2 = \frac{25}{16}\sigma_0^2$. It is assumed, for simplicity reasons, that the expression (3.36) has a normal distribution with the same mean and variance. Similar

results are obtained for the second double outer summation on the right-hand side of Eq.(3.34). It follows then that $\left[\frac{16}{25} D_1/\sigma_0^2\right]$, which is a variable representing the denominator of Eq.(3.30) in an appropriately normalized form,is distributed as χ^2 with 14 degrees of freedom, where 14=32-18; 32 – number of terms in D_1; 18 – number of constrains, that is the number of equations for the estimates $\hat{c}_{u,v}$ and $\hat{d}_{u,v}$.

Defining another variable F, which is more convenient to use than F_1,

$$F = \frac{25}{16} \frac{N_1/9}{D_1/14} = \frac{175}{72} F_1 \tag{3.37}$$

one concludes that the variable F has an F distribution (named after R.A. Fisher) with (9, 14) degrees of freedom. The statistical test requires that a threshold F_{th} be chosen such that for $F \leq F_{\text{th}}$ the hypothesis H_0 would be accepted, and for $F > F_{\text{th}}$ the hypothesis H_0 would be rejected. The choice of F_{th} depends on the assumed level of significance of the test. A typical level is 1 or 5%, which means that the probability of not accepting H_0 when there was no motion is, respectively, 1 or 5%. As shown in the tables of the F distribution ([21]), for 1% significance level $F_{\text{th}} = 4.03$, and for 5% level $F_{\text{th}} = 2.65$.

4. Experimental results

The experiments were performed on both the simulated images and the images of a natural scene. The advantage of using simulated images is that one can consciously select the noise level so that it is easier to observe its influence. The experiments described were conducted on the IBM PC/AT computer with a program written in the Microsoft C language, version 5.1.

The results of the experiment with the simulated images are shown in Figs. 1 through 3. In particular, Fig. 1 shows the original 64 × 64 image similar to the one given in the paper [5]. The range of gray values in this image is 0 through 31. The gray values in Fig. 1 are shifted upwards in comparison with [5] to allow for their change both upwards and downwards. A random number generator ([23]) was used to generate a sample from a uniform distribution. By a suitable transformation this sample was converted into a sample from the normal distribution. Figure 2 shows the image of Fig. 1 corrupted by noise obtained from the distribution N(0, 1). Adding noise to the original gray values of Fig. 1 resulted in numbers with a fractional part. These numbers were subsequently rounded to the nearest integer in the range 0 through 31. Figure 3 shows a sequence of images obtained from the image of Fig. 1, in which all the characters, with the exception of the letter B and digit 3, are gradually displaced in various directions and the resulting images are corrupted by noise similarly as in Fig. 2. The images of Figs. 2 and 3 were used as an input to the program calculating the

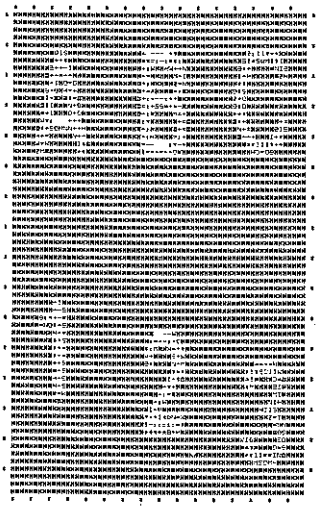

FIG. 1. Original noise-free image.

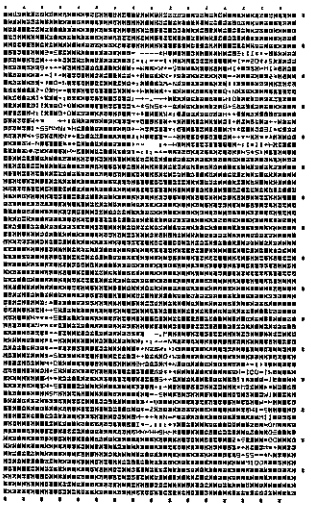

FIG. 2. Starting, reference image for the sequence of Fig. 3.

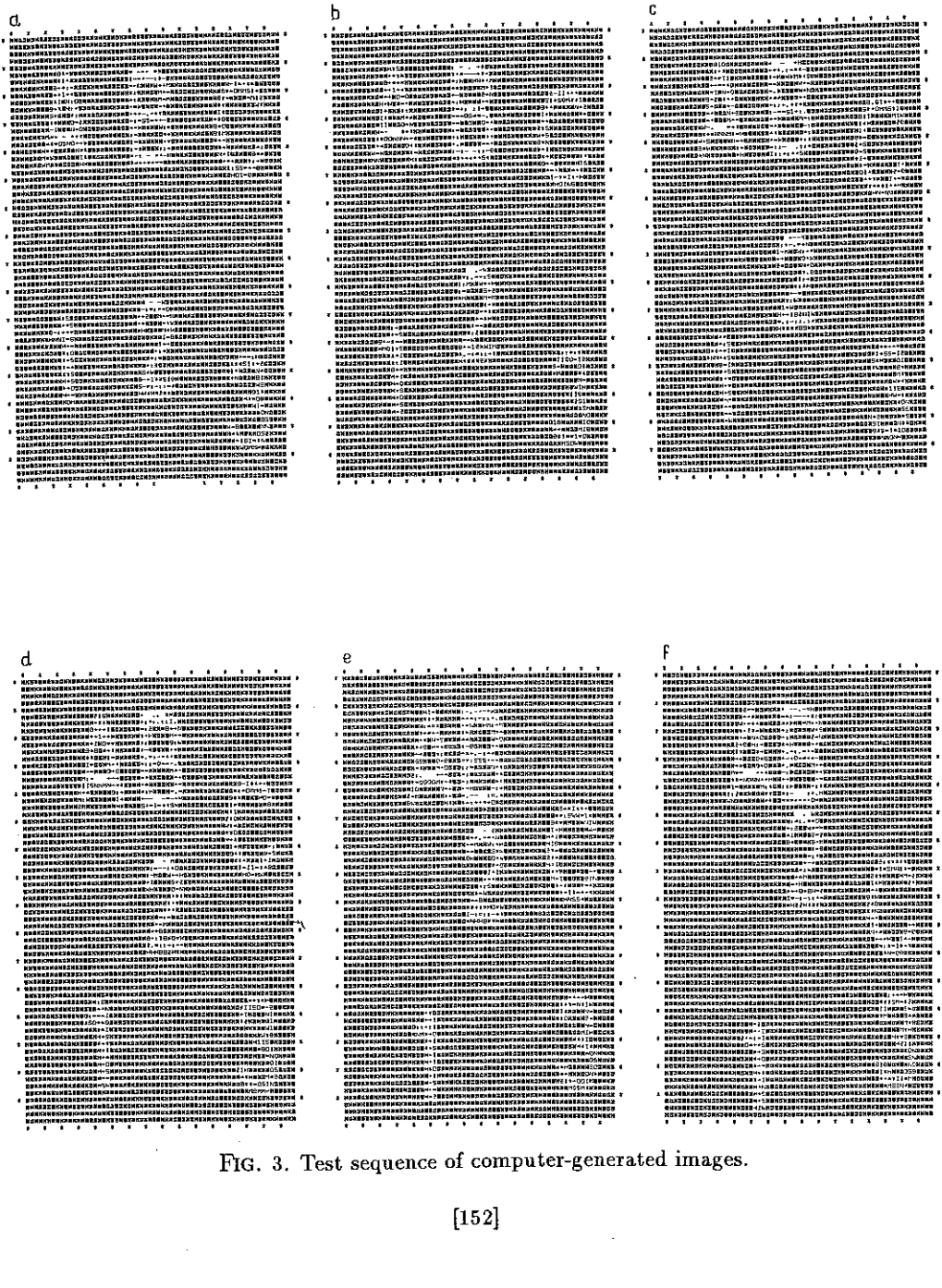

FIG. 3. Test sequence of computer-generated images.

Table 1. Likelihood test results for the sequence of Fig. 3 with $F_{th} = 4.03$.

a	b	c
0000000000000000	0000000000000000	0000000000000000
0011000000001000	0011000000001010	0011100000001010
0101000000000100	0011100000000000	0011100000000100
0101100000000010	0110100000000010	0101010000000010
1011100000000010	0111100000000010	0101010000000010
0000000000000000	0000000000001010	0000000000001110
0000000000000000	0000000000010000	0000000000000000
0000000000000000	0000000000000000	0000000110001000
0000000000000000	0000010110000000	0000000011000000
0000000110000000	0000000011000000	0000000010000000
0111000011000000	0110000010000000	0110000100000000
0011000001000000	0011000000100000	0011100011000000
0011000100010001	0011001010000000	0010100010000000
0000000110000000	1001000110000000	0001100110000000
0000000000000000	0000000000000000	0000000000000000
0000000000000000	0000000000000000	0000000000000000

d	e	f
0000000000000010	0000001100000010	0000000000000000
0010100000001010	0110110000001010	0010110000001010
0011110000000000	0011110000000000	0011110000000000
0111110000000110	0111110000000010	0111110000000010
0100010000000000	0101101000001010	0101101000001010
0000000000000000	0000000110001010	0000000011000000
0000000110001110	0000000011000000	0000000011000110
0000000011000000	0000000010001110	0000000110001110
1000000010000000	0000010110000100	0000000111001000
0100000110000000	0000000000000000	0100000000001110
0110000110000000	0110000110000000	0110000110000000
0010000011000000	0010000011000000	0110000011000000
0011101010000000	0010110010000000	0010010010000001
0001100110000000	0000010110000000	0000110110000000
0000100000000000	0000110000000000	0000010000000000
0000000000100000	0000000000000000	0000010000000001

Table 2. Likelihood test results for the sequence of Fig. 3 with $F_{th} = 2.65$.

a	b	c
0100000100000000	0000000000100000	0100000010000000
0011100000001000	0011100000001010	1111100000001010
0111000000100110	0111100000001010	0111100000001110
0101110000000010	0110100000000010	0110110000000110
1011100000000010	0111101000001110	1101010000001110
0000000000000100	0000000000011010	0000010000001110
0000000000010010	0000001010010000	0000100000010101
0000000000000000	0100000000000000	0000000110101000
1000010000000000	1000010110000000	1000010011000000
0010000110000000	0110001011000000	0110001011000000
0111010111000000	0111100010001000	0110000100001011
1011000011000000	0011000010100000	1011100011100000
0011000100010001	1011001011000000	0011100011010000
0001100111000000	1001000110000000	1001100111000000
0010000000000000	0100000011010000	0000100000010000
0000000000000000	0101000100000000	0100000000001011

d	e	f
0100000110100010	0000001100100010	0000101000100000
0110100000001010	1110111000011010	0010110000001010
0111110000001000	0111110000011000	0111110000001000
0111110000101110	0111111000001010	0111111000101010
0100011000000000	0101111000001010	0101111000011010
0000010000001000	0000000110001110	0000010111000000
0000000110011110	0000100011011000	0000000011000110
0000000011000100	0000000010001110	0010000111001110
1000010010000000	0000010110000100	1000001111001000
0110001110001001	0010001001000000	0110000000001110
1110010110010000	0110000110000000	1110010110001100
0010101111010000	0010000011010000	0110000011000000
0011101011000001	0010111010000000	1010010011100001
0001100111000000	1101010110000000	0000110111000000
0000100000110000	0000110001010000	0000010110000000
0100000000101000	0001010010001100	1001010010100011

coefficients of the Walsh functions and performing the maximum likelihood test. The results of this program are given in Tables 1 and 2 which show the values of 1 and 0, respectively, for each 4×4 window, depending upon whether there was a statistically significant change or not. There are six images a) through f) in Fig.3, and Tables 1 and 2 give the results of the comparison of each of these images with respect to the reference image of Fig.2. Tables 1 and 2 confirm the fact that the change detection program correctly detected the motion of the characters as well as the direction of the displacement. By using a relatively low value of $F_{\text{th}} = 2.65$, as shown in Table 2, one extracts a bigger part of the moving object, but at the same time there are more windows in which the value of 1 is obtained because of the noise.

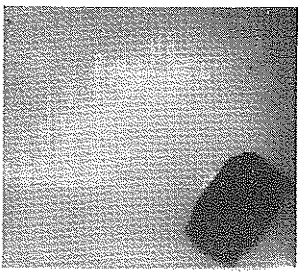

FIG. 4. Starting, reference image for the sequence of Fig.5.

The presented method of motion detection was also tested on real images obtained by means of the MC9000 Series Modular Camera, manufactured by RETICON. This camera has 256×256 pixels with 64 gray levels. The results of the experiment with images of the natural scene are shown in Figs. 4 through 6. In particular, Fig. 4 shows the reference image with a cube in the lower right corner. This cube was then moved as shown in Fig.5 a) through h). The images of Fig. 5 were tested by the computer program using Fig. 4 as a reference image. Figures 6 a) through h) show the corresponding results, with a cross placed for each 4×4 window in which statistically significant changes were detected. It was found in this case that a relatively high value $F_{\text{th}} = 30.0$ allows one to recover the image of the cube and does not result in too many errors in the background pixels. The influence of the threshold level is illustrated by Fig.7, which shows the results obtained for the image of Fig. 5 a) with F_{th} changing in the range from 20.0 down to 2.65. It is apparent from the examples given that the threshold level may have to be adjusted to the current operating conditions of the camera.

A problem related to detecting a moving object is that of extracting the image of the object ([1, 8]). One possible technique of doing this is the following ([1]). The comparison is made between two images in which the

object moves so far that there is no overlapping (compare the motion of the

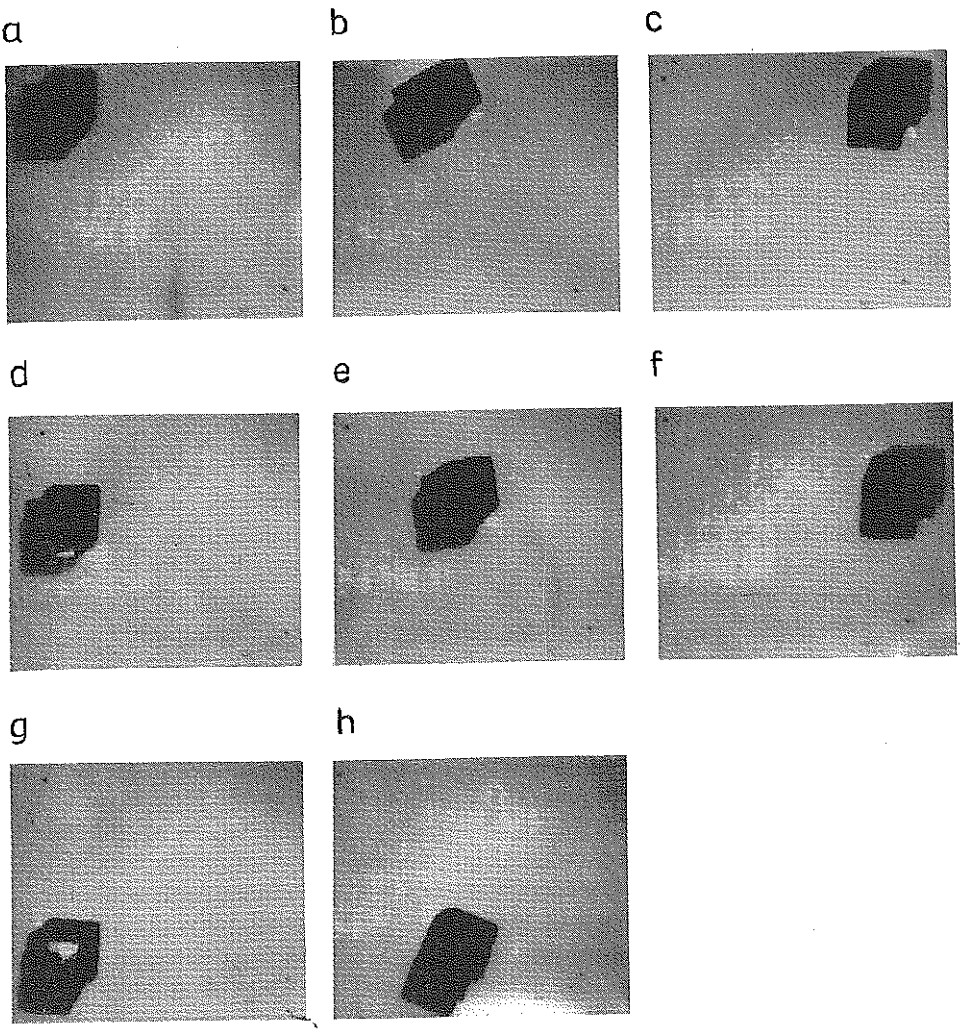

FIG. 5. Test sequence of images of a natural scene.

letter B and digits 1 and 2 between Figs.2 and 3 f)). For one of the images considered, all the windows indicating the motion of the object are merged into a single image of the object by means of a suitable neighbourhood analysis. In another technique ([8]), a sequence of, say, three images is tested, in which there is only a slight motion of the object. The second image in the sequence is used as a reference in the likelihood test for the first and the third image. The results from the two tests are logically OR-ed,

independently for each window in the image. This technique allows one to extract effectively most of the image of the object, but since there may be a considerable noise influence, noise filtering is necessary before the logical addition. For example, in the paper [8] all object images consisting of ten or less 4-connected neighbours were rejected.

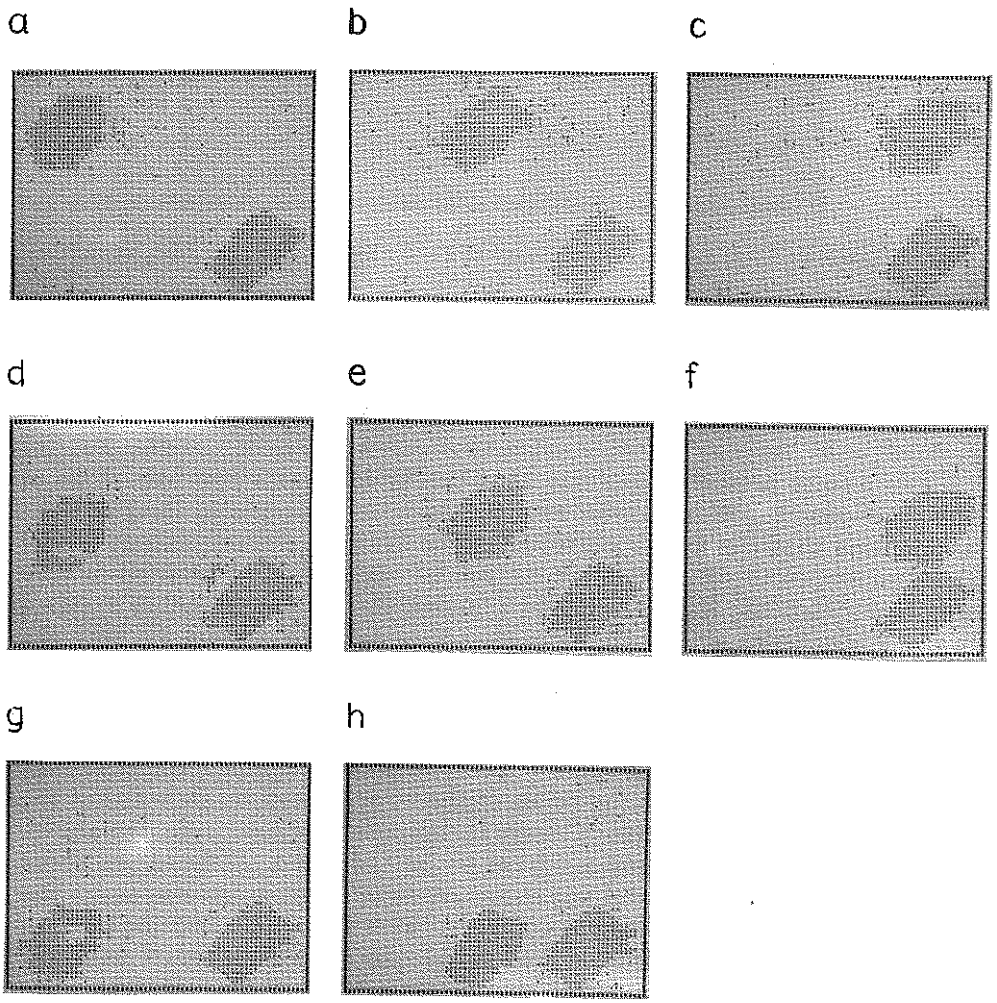

FIG. 6. Likelihood test results for the sequence of Fig. 5 with F_{th} =30.0.

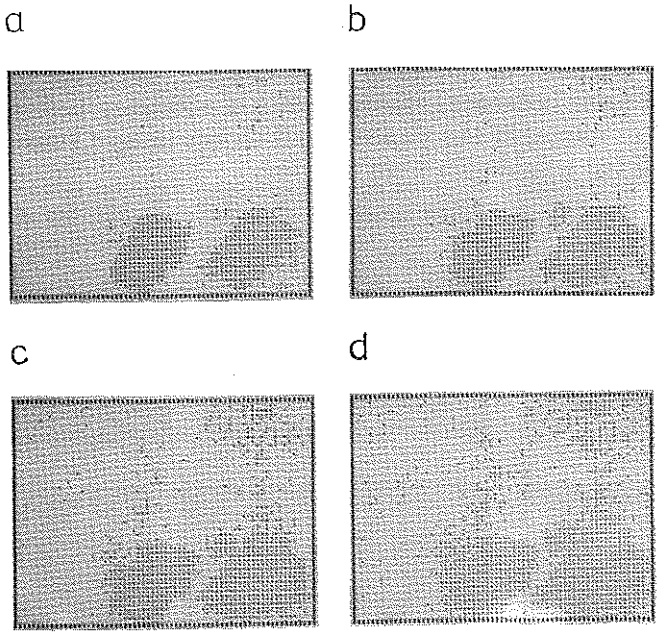

FIG. 7. Likelihood test results for the image of Fig.5 a) with: a) $F_{th} = 20.0$, b) 10.0, c) 4.03, d) 2.65.

5. CONCLUSIONS

The results presented indicate that the Walsh series approximation of the gray-value function can be used for detecting changes in images. The assumption that the first 3×3 terms of the Walsh series adequately model this function in the 4×4 window proved to work well and resulted in very simple calculations. However, bigger windows and a greater number of terms in the Walsh series might also be used as long as there is room for both the Walsh series approximation and the inclusion of the influence of noise. It is also possible to approximate by the Walsh series the difference image which is obtained by the subtraction of the respective gray levels from two images. In the case when there is no motion in the window, the non-zero gray levels in the difference image are the result of the noise. It can be shown that the likelihood test for the difference image is similar to the one presented above, although most of the equations have to be rederived. The experiments conducted by the authors did not show any significant improvement for the case with the difference image as compared with the approximation of both images separately, which is described in this paper. The calculations based on the Walsh series approximation and the maximum likelihood test are very simple and may be carried out in real time image analysis. One of the possible uses is finding the areas of interest in an image. Calculation

of the Walsh function coefficients can take the form of convolution of the image with local masks. The subsequent maximum likelihood test reduces to the calculation of the value of F or F_1 based on the definition (3.29). The detected areas of interest can then be processed by other methods.

REFERENCES

1. D.H. BALLARD and C.M. BROWN, *Computer vision*, Prentice-Hall, Englewood Cliffs, NJ, 1982.
2. P. BOUTHEMY, *A maximum likelihood framework for determining moving edges*, IEEE Trans. Pattern Anal. Machine Intell., PAMI-11, 499-511, 1989.
3. C. CAFFORIO and F. ROCCA, *Methods for measuring small displacements of television images*, IEEE Trans. Inf. Theory, IT-22, 573-79, 1976.
4. G. DAHLQUIST and A. BJÖRCK, *Numerical methods*, Prentice-Hall, Englewood Cliffs, NJ, 1974.
5. R.C. GONZALEZ and P. WINTZ, *Digital image processing*, Addison-Wesley, Reading, MA, 1977.
6. B.K.P. HORN, *Robot vision*, The MIT Press, Cambridge, MA, 1986.
7. B.K.P. HORN and B.G. SCHUNK, *Determining optical flow*, Artificial Intell., **17**, 185-203, 1981.
8. Y.Z. HSU, H.H. NAGEL, G. REKERS, *New likelihood test methods for change detection in image sequences*, Comput. Vision Graphics Image Processing, **26**, 73-106, 1984.
9. R. JAIN, *Segmentation of moving observer frame sequences*, Pattern Recogn. Letters, 1, 115-20, 1982.
10. D.T. LAWTON, *Processing translational motion sequences*, Comput. Vision Graphics Image Processing, **22**, 116-44, 1983.
11. J.O. LIMB and J.A. MURPHY, *Estimating the velocity of moving images in television signals*, Comput. Graphics Image Processing, 4, 311-27, 1975.
12. S.A. MAHMOUD, M.S. AFIFI, R.J. GREEN, *Recognition and velocity computation of large moving objects in images*, IEEE Trans. Acoust. Speech Signal Proc., ASSP-36, 1790-91, 1988.
13. W. MICHAELIS, *Möglichkeiten der Verschiebungsbestimmung bei der digitalen Auswertung von Bildfolgen*, Wiss. Zeitschrift, TH Ilmenau, **34**, H. 2, 95-106, 1988.
14. H.H. NAGEL, *On change detection and displacement vector estimation in image sequences*, Pattern Recogn. Letters, **1**, 55-59, 1982.
15. M. NIENIEWSKI and P.K. PATHAK, *Change detection in image sequences using Walsh functions and the likelihood test method*, Proceedings of Internat. Conference on Signal Processing, 995-98, Beijing, October 22-26, 1990.
16. T.J. PATTERSON, D.M. CHABRIES, R.W. CHRISTIANSEN, *Detection algorithms for image sequence analysis*, IEEE Trans. Acoust. Speech Signal Proc., ASSP-37, 1454-58, 1989.
17. W.K. PRATT, *Digital image processing*, John Wiley, New York, NY, 1978.
18. S.A. RAJALA, A.N. RIDDLE, W.E. SNYDER, *Application of the one-dimensional Fourier transform for tracking moving objects in noisy environments*, Comput. Vision Graphics Image Processing, **21**, 280-93, 1983.

19. A. ROSENFELD and A.C. KAK, *Digital picture processing*, Academic Press, Orlando, FL, 1982.

20. M.A. SHAH and R. JAIN, *Detecting time-varying corners*, Comput. Vision Graphics Image Processing, **28**, 345-55, 1984.

21. M.R. SPIEGEL, *Probability and statistics*, McGraw-Hill, New York, NY, 1975.

22. M.S. ULSTAD, *An algorithm for estimating small scale differences between two digital images*, Pattern Recogn., **5**, 323-33, 1973.

23. B. WICHMANN and D. HILL, *Building a random-number generator*, Byte, **12**, 127-128, March 1987.

24. Y. YAKIMOVSKY, *Boundary and object detection in real world images*, J.Assoc. Comput. Machinery, **23**, 599-618, 1976.

25. S. YALAMANCHILI, W.N. MARTIN, J.K. AGGARWAL, *Extraction of moving object descriptions via differencing*, Comput. Vision Graphics Image Processing, **18**, 188-201, 1982.

STRESZCZENIE

DETEKCJA ZMIAN W DWÓCH OBRAZACH PRZY UŻYCIU FUNKCJI WALSHA I METODY NAJWIĘKSZEJ WIARYGODNOŚCI

Funkcja szarości w oknie lokalnym jest aproksymowana za pomocą szeregu funkcji Walsha. Kilka wyrazów takiego szeregu przedstawia w zadowalający sposób funkcję szarości w oknie, a pozostałą zmienność tej funkcji można przypisać szumowi. Przy porównaniu dwóch obrazów znajduje się współczynniki funkcji Walsha dla kolejnych położeń okna w obu obrazach. Detekcja zmiany pomiędzy dwoma odpowiadającymi sobie oknami opiera się na teście stosunku maksymalnych wiarygodności, polegającym na wyborze między dwiema hipotezami: H_0 – ruch nie występuje, oraz H_1 – ruch występuje. Przedstawiono wyniki detekcji ruchu dla obrazów sceny naturalnej otrzymanych przy użyciu kamery CCD, a także dla obrazów symulowanych, poddanych oddziaływaniu szumu gaussowskiego. Opisany program komputerowy przeprowadza F-test dla stosunku maksymalnych wiarygodności. Wykazano, że program taki może być z powodzeniem użyty do detekcji zmian w sekwencjach obrazów.

РЕЗЮМЕ

ДЕТЕКЦИЯ ИЗМЕНЕНИЙ В ДВУХ ИЗОБРАЖЕНИЯХ ПРИ ПОМОЩИ ФУНКЦИЙ УОЛША И МЕТОДА МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ

Функция яркости в местным окне аппроксимирована рядом функций Уолша. Несколько членов такого ряда удовлетворительно моделирует функцию яркости в окне, а остаточная вариантность этой функции может быть объяснена влиянием шума. При сравнивании двух окон находят коэффициенты функций Уолша

для всех очередных положений окна в обоих изображениях. Детекция изменения между двумя соответствующими окнами основана на проверке соотношения максимальных правдоподобий и сводится к выборе одной из двух гипотез: H_0 – в окне нет движения, и H_1 – в окне есть движение. Представлены результаты детекции изменений для изображений натуральной сцены полученных при помощи оптоэлектронного сенсора типа CCD, а также для симулированных изображений подвергнутых действию гауссовского шума. Описываемая компьютерная программа производит F- тест для соотношения максимальных прадоподобий. Показано что такая программа может быть успешно применена для детекции изменений в секвенции изображений.

POLISH ACADEMY OF SCIENCES
ELECTROTECHNICAL INSTITUTE
INSTITUTE OF FUNDAMENTAL TECHNOLOGICAL RESEARCH.
and
DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF NEW MEXICO, ALBUQUERQUE, NM, USA.

Received December 20, 1989.