EURASIP Journal on Applied Signal Processing 2005:4, 588–599

Properties of Orthogonal Gaussian-Hermite Moments and Their Applications

Youfu Wu

EGID Institut, Université Michele de Montaigne Bordeaux 3, 1 Allée Daguin, Domaine Universitaire, 33607 Pessac Cedex, France
Email: youfu_wu_64@yahoo.com.cn

Jun Shen

EGID Institut, Université Michele de Montaigne Bordeaux 3, 1 Allée Daguin, Domaine Universitaire, 33607 Pessac Cedex, France

Received 7 May 2004; Revised 5 September 2004; Recommended for Publication by Moon Gi Kang

Moments are widely used in pattern recognition, image processing, and computer vision and multiresolution analysis. In this paper, we first point out some properties of the orthogonal Gaussian-Hermite moments, and propose a new method to detect the moving objects by using the orthogonal Gaussian-Hermite moments. The experiment results are reported, which show the good performance of our method.

Keywords and phrases: orthogonal Gaussian-Hermite moments, detecting moving objects, object segmentation, Gaussian filter, localization errors.

1. INTRODUCTION

Moments are widely used in pattern recognition, image processing, and computer vision and multiresolution analysis [1, 2, 3, 4, 5, 6, 7, 8, 9]. We present in this paper a study on orthogonal Gaussian-Hermite moments (OGHMs), their calculation, properties, application and so forth. We at first analyze their properties in spatial domain. Our analysis shows orthogonal moment's base functions of different orders having different number of zero crossings and very different shapes, therefore they can better separate image features based on different modes, which is very interesting for pattern analysis, shape classification, and detection of the moving objects. Moreover, the base functions of OGHMs are much more smoothed; are thus less sensitive to noise and avoid the artefacts introduced by window function's discontinuity [1, 5, 10].

Since the Gaussian-Hermite moments are much smoother than other moments [5], and much less sensitive to noise, OGHMs could facilitate the detection of moving objects in noisy image sequences. Compared with other differential methods (DMs), experiments show that much better results can be obtained by using the OGHMs for the moving objects detection.

Traffic management and information systems rely on some sensors for estimating traffic parameters. Vision-based video monitoring systems offer a number of advantages. The first task for automatic surveillance is to detect the moving objects in a visible range of the video camera [2, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22]. The objects can be persons, vehicles, animals, etc. [2, 12, 13, 15, 21, 23, 24].

In general, we can classify the methods of detecting the moving objects in an image sequence into three principal categories: methods based on the background subtraction (BS) [2, 12, 13, 18], methods based on the temporal variation in the successive images [1, 2, 25], and methods based on stochastic estimation of activities [11].

To extract the background image, one simple method is to take the temporal average of the image sequence; another is to take the median of the image sequence [2]. However, these methods are likely to be ineffective to solve the problems of the lighting condition change between the frames and the slow moving objects. For example, the mean method leaves the trail of the slow moving object in the background image, which may lead to the wrong detecting results.

In order to obtain the background image almost on real time, the adaptive background subtraction (ABS) method, proposed by Stauffer and Grimson [12, 13], can be adopted. In this method, a mixture of K Gaussian distributions adaptively models each pixel of intensity. The distributions are evaluated to determine which are more likely to result from a background process. This method can deal with the long-term change in lighting conditions and scene changes. However, it cannot deal with sudden movements of the uninteresting objects, such as the flag waving or winds blowing through trees for a short burst of time [11].

A sudden lighting change will then cause the complete frame to be regarded as foreground, if such a condition arises. The algorithm needs to be reinitialized [11, 12, 13]. It again demands a certain quantity of accumulated images.

Our paper is organized as follows. Section 2 presents OGHMs and their properties; Section 3 presents the detection of the moving objects by using OGHMs; Section 4 gives the experimental results and the performance comparison with other methods of detecting the moving objects; some conclusions and discussions are presented in Section 5.

2. OGHMS AND THEIR PROPERTIES

2.1. Hermite moments [5, 6]

Hermite polynomial is one family of the orthogonal polynomials as follows:

$$P_n = H_n\left(\frac{t}{\sigma}\right), \tag{1}$$

where $H_n(t) = (-1)^n \exp(t^2)(d^n/dt^n)\exp(-t^2)$, σ is the standard deviation of the Gaussian function.

The 1D nth-order Hermite moment $M_n(x, s(x))$ of a signal $s(s)$ can therefore be defined as follows:

$$\begin{aligned} M_n(x, s(x)) &= \int_{-\infty}^{\infty} s(x+t)P_n(t)dt \\ &= \langle P_n(t), s(x+t)\rangle \quad (n = 0, 1, 2, \dots). \end{aligned} \tag{2}$$

In the 2D case, the 2D (p, q)-order Hermite moment is defined as

$$\begin{aligned} &M_{p,q}(x, y, I(x, y)) \\ &\quad = \int_{-\infty}^{\infty}\int_{\infty}^{\infty} I(x+u, y+v)H_{p,q}\left(\frac{u}{\sigma}, \frac{v}{\sigma}\right)dudv, \end{aligned} \tag{3}$$

where $I(x, y)$ is an image and $H_{p,q}(u/\sigma, v/\delta) = H_p(u/\sigma) \times H_q(v/\sigma)$.

2.2. Orthogonal Gaussian-Hermite moments

The OGHMs was proposed by Shen et al. [5, 6]. The OGHMs of a signal $s(x)$ is defined as

$$M_n(x, s(x)) = \int_{-\infty}^{\infty} s(x+t)B_n(t)dt = \langle B_n(t), s(x+t)\rangle, \tag{4}$$

where $B_n(t) = g(t, \sigma)P_n(t)$, $g(x, \sigma) = (1/\sqrt{2\pi}\sigma)\exp(-x^2/2\sigma^2)$ and $P_n(t)$ is a Hermite polynomial function.

For calculating the OGHMs, we can use the following recursive algorithm:

$$\begin{aligned} M_n(x, s^{(m)}(x)) &= 2(n-1)M_{n-2}(x, s^{(m)}(x)) \\ &\quad + 2\sigma M_{n-1}(x, s^{(m+1)}(x)) \quad (n \geq 2), \end{aligned} \tag{5}$$

where $s^{(m)}(x) = (d^m/dx^m)s(x)$, $s^{(0)}(x) = s(x)$.

In particular,

$$\begin{aligned} M_0(x, s(x)) &= g(x, \sigma) * s(x), \\ M_1(x, s(x)) &= 2\sigma\left(\frac{d}{dx}\right)(g(x, \sigma)) * s(x), \end{aligned} \tag{6}$$

where "$*$" represents the operation of convolution.

In the 2D case, the OGHMs of order (p, q) of an input image $I(x, y)$ can be defined similarly as

$$\begin{aligned} &M_{p,q}(x, y, I(x, y)) \\ &\quad = \int_{-\infty}^{\infty}\int_{\infty}^{\infty} g(u, v, \sigma)H_{p,q}\left(\frac{u}{\sigma}, \frac{v}{\sigma}\right)I(x+u, y+v)dudv, \end{aligned} \tag{7}$$

where $H_{p,q}(u/\sigma, v/\sigma) = H_p(u/\sigma)H_q(v/\sigma)$, $g(u, v, \sigma) = (1/2\pi\sigma^2)\exp(-(u^2/2\sigma^2 + v^2/2\sigma^2))$.

Obviously, the 2D OGHMs are separable, so the calculation of the 2D OGHMs can be decomposed into the cascade of two steps of the 1D OGHMs calculation:

$$\begin{aligned} &M_{p,q}(x, y, I(x, y)) \\ &\quad = \int_{-\infty}^{\infty}\left(\int_{\infty}^{\infty} g(u, \delta)H_p\left(\frac{u}{\sigma}\right)I(x+u, y+v)du\right) \\ &\qquad \times g(v, \sigma)H_p\left(\frac{v}{\sigma}\right)dv. \end{aligned} \tag{8}$$

In order to detect moving objects in image sequences by using OGHMs, if a video image sequence $\{f(x, y, t)\}_{t=0,1,2,\dots}$ is given, for each spatial position (x, y) on the images, we define the temporal OGHMs as follows:

$$M_n(t, f(x, y, t)) = \int_{-\infty}^{\infty} f(x, y, t+v)B_n(v)dv. \tag{9}$$

Its recursive algorithm can be rewritten as follows:

$$\begin{aligned} &M_n(t, f(x, y, t)) \\ &\quad = 2(n-1)M_{n-2}(t, f(x, y, t)) + 2\sigma M_{n-1}(t, f^{(1)}(x, y, t)). \end{aligned} \tag{10}$$

In particular, we use only the moments of the odd orders up to 5.

2.3. The properties of the OGHMs

First of all, the Gaussian filter has a property as follows:

$$\left(\frac{d^n}{dt^n}\right)(g(t, \sigma) * f(x, y, t)) = g^{(n)}(t, \sigma) * f(x, y, t). \tag{11}$$

According to the recursive algorithm, the OGHMs have the following properties.

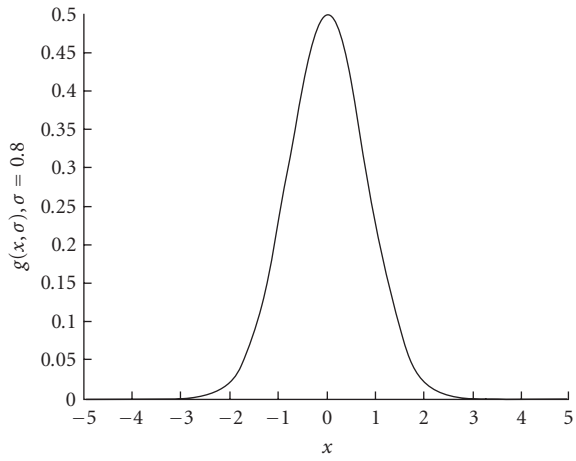

Figure 1: The mask of the 1D OGHMs of order 0.

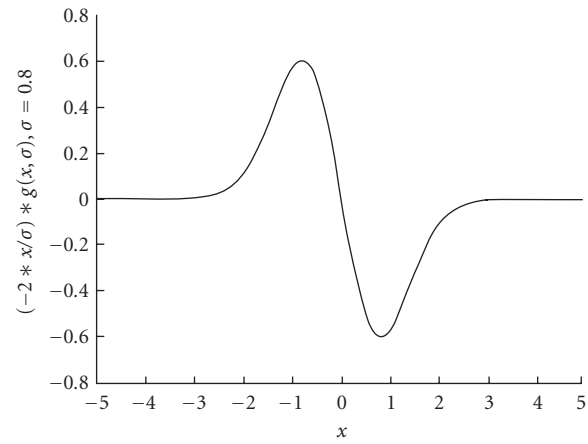

Figure 2: The mask of the 1D OGHMs of order 1.

Property 1. Given a Gaussian function $g(t,\sigma)$ and an image $f(x,y,t)$ of the image sequence $\{f(x,y,t)\}_{t=0,1,2,\ldots}$, we have

$$\begin{aligned} M_n(t,f(x,y,t)) &= \sum_{i=0}^{n} a_i\left(\frac{d^i}{dt^i}\right)(g(t,\sigma) * f(x,y,t)) \\ &= \left(\sum_{i=0}^{n} a_i\left(\frac{d^i}{dt^i}\right)g(t,\sigma)\right) * f(x,y,t), \end{aligned} \tag{12}$$

where a_i depends on σ only.

This property shows that the mask $\sum_{i=0}^{n} a_i(d^i/dt^i)g(t,\sigma)$ of the nth moment is the linear combination of the Gaussian function and its derivatives of different orders.

Property 2. Given a Gaussian function $g(t,\sigma)$ and an image sequence $\{f(x,y,t)\}_{t=0,1,2,\ldots}$, we have

$$M_n(t,f(x,y,t)) = \sum_{i=0}^{k} a_{2i}\left(\frac{d^{2i}}{dt^{2i}}\right)(g(t,\sigma) * f(x,y,t)) \quad \text{for } n = 2k \ (n \text{ is even}), \tag{13}$$

$$M_n(t,f(x,y,t)) = \sum_{i=0}^{k} a_{2i+1}\left(\frac{d^{2i+1}}{dt^{2i+1}}\right)(g(t,\sigma) * f(x,y,t)), \quad \text{for } n = 2k+1 \ (n \text{ is odd}), \tag{14}$$

where a_i depends on σ only.

This property shows that the mask of the OGHMs of odd order is the linear combination of the derivatives of odd orders of the Gaussian function, and the mask of the OGHMs of even order is the linear combination of the derivatives of even orders of the Gaussian function.

Property 3. Given a Gaussian function $g(t,\sigma)$ and an image sequence $\{f(x,y,t)\}_{t=0,1,2,\ldots}$, $M_n(t,f(x,y,t)) = \sum_{i=0}^{n} a_i(d^i/dt^i)(g(t,\sigma) * f(x,y,t))$, we note that $F(t,\delta) = \sum_{i=0}^{n} a_i(d^i/dt^i)g(t,\sigma)$; then $F(t,\sigma) = 0$ has n different real roots in the interval $(-\infty,\infty)$.

$F(x,\sigma)$ is called the base function of the OGHMs (also called the mask of the OGHMs). This property shows that the mask of the nth moment has n different zero crossings.

2.4. Some conclusions

From the properties of OGHMs, we see that these moments are in fact linear combinations of the derivatives of the filtered signal by a Gaussian filter. As it is well known, derivatives are important features widely used in signal and image processing. Because differential operations are sensitive to random noise, a smoothing is in general necessary. The Gaussian-Hermite moments just meet this demand because of the Gaussian smoothing included. In image processing, one often needs the derivatives of different orders to effectively characterize the images, but how to combine them is still a difficult problem. The OGHMs show a way to construct orthogonal features from different derivatives.

For facilely understanding the OGHMs, in Figures 1, 2, 3, 4, and 5, we give out some characteristics charts of the base function of OGHMs.

In the spatial domain, because the base function of the nth-order OGHMs will change its sign n times, OGHMs can well characterize different spatial modes as other orthogonal moments. As to the frequency domain behavior, because the base function of the nth-order OGHMs consists of more oscillations when the order n is increased, they will thus contain more and more frequency. Table 1 shows that the frequency windows' quality factor Q = (center/effective bandwidth) of OGHMs base function is large than that of other moment base function; therefore OGHMs separate different bands more efficiently. Moreover, from the properties of OGHMs, we see that these moments are in fact linear combinations of the derivatives of the signal filtered by a Gaussian filter, therefore, realizing differential operation and removing random noise.

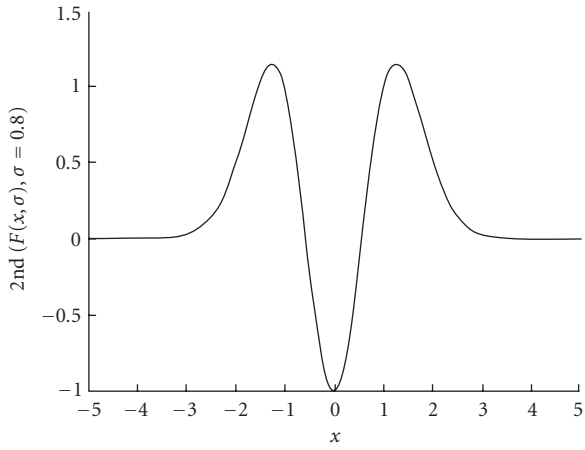

Figure 3: The mask of the 1D OGHMs of order 2.

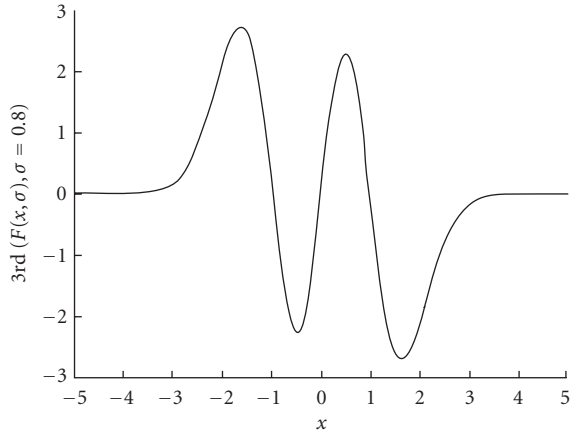

Figure 4: The mask of the 1D OGHMs of order 3.

From the viewpoint of frequency analysis, it seems that the OGHMs characterize images more efficiently. With the help of the technique representing frequency characteristics by the band center ω_0 and effective bandwidth B_e, we can better see the differences between the moment base functions. In Table 1, these characteristics for orders from 0 through 10 are shown. A similar conclusion holds in 2D cases. In general, one uses the max order up to 10.

3. DETECTING MOVING OBJECTS USING OGHMS

3.1. Calculating the OGHMs images

According to (12), (13), and (14), we can see that all the OGHMs are actually the linear combinations of the different order derivatives of the image filtered by a temporal Gaussian filter. As it is well known, the temporal derivatives can be used to detect the moving objects in image sequences. The OGHMs of odd orders are in fact the combinations of these derivatives, it is therefore reasonable to use them to detect the moving objects.

According to (12), M_n is equal to the temporal derivative of the image filtered by a Gaussian filter, and the OGHMs

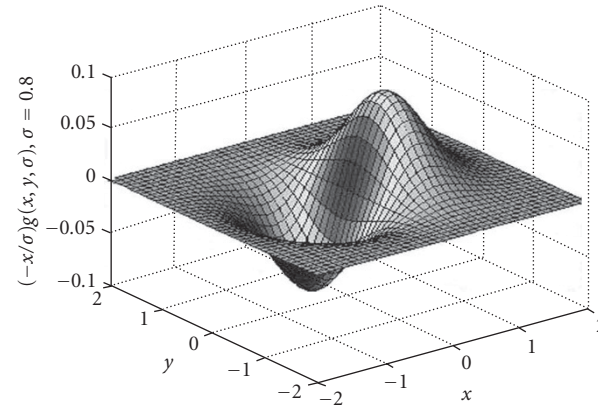

Figure 5: The mask of the 2D OGHMs of order $(1,0)$.

are much smoother than other moments, therefore much less sensitive to noise, which could facilitate the detection of the moving objects in noisy image sequences.

To detect the moving objects by using the OGHMs, we first calculate the temporal moments of an image sequence. According to (14), calculating the temporal OGHMs of an image sequence is equal to calculating the convolution of the mask $F(t,\sigma)$ with $f(x,y,t)$. In order to approach the mask $F(t,\sigma)$, according to the inequality $\int_{|t-Et|\geq\varepsilon} g(t)dt \leq (1/\varepsilon^2)\int (t-Et)^2 g(t)dt = \sigma^2/\varepsilon^2$, if we take $\varepsilon = 5\sigma$, then $\int_{|t-Et|\geq\varepsilon} g(t)dt \leq 1/25 = 4\%$. Hence the Gaussian function as the convolution kernel is common to choose $N_o = 10\sigma + 1$, namely the masks of size $2L+1$ with $L = 5\sigma$ is used. For practical computation reasons, we use only the moments of orders up to 5. Hence, in order to detect the moving objects using the OGHMs, the temporal OGHMs of an image sequence is calculated first.

Since both the positive values and negative values of the moments correspond to the moving objects (containing the noise), we take the absolute value of the moments instead of their original values. For example, for Figure 6, from Figure 7 to Figure 11, we present the OGHMs images, visualized by linearly transforming the absolute value of the OGHMs to the gray value ranging from 0 to 255. It can be seen that four moving objects (2 persons, one car, and one cyclist) are well enhanced in the moment images.

Thus M_3 contains more information than M_1 for detecting the moving objects. Therefore, we can use the third moment to detect the mobile objects. Figures 8 and 9 show the experimental results in this case.

By comparing the results obtained in the case of $\sigma = 0.3$ and $\sigma = 0.8$, it can be seen when a larger σ is used, the results of the detection of the moving objects are less sensitive to the noises, but the detected objects have larger sizes than their real sizes. This phenomenon can be explained in Section 3.4.

The M_5 is a weighted sum of the first-, third-, and fifth-order derivatives of image filtered by a temporal Gaussian filter. Therefore, M_5 still contains more information than these M_1 and M_3 for the detection of the moving objects. Figures 10 and 11 show the experiment results.

Table 1: Frequency characteristics of the base functions of geometric, Hermite, Legendre moments, and OGHMs.

Moment	Geometric moment			Hermite moment			Legendre moment			OGHMs		
order	ω_0	B_e	ω_0/B_e	ω_0	B_e	ω_0/B_e	ω_0	B_e	ω_0/B_e	ω_0	B_e	ω_0/B_e
0	0.40	2.11	0.1896	0.40	2.11	0.1896	0.40	2.11	0.1896	0.53	0.44	1.2045
1	1.25	3.87	0.3230	1.25	3.87	0.3230	1.25	3.87	0.3230	1.13	0.48	2.3542
2	1.59	4.71	0.3376	1.64	4.77	0.3438	1.94	4.90	0.3959	1.36	0.75	1.8133
3	2.32	5.74	0.4042	2.42	5.86	0.4130	2.60	5.74	0.4530	1.69	0.80	2.1125
4	2.61	6.23	0.4189	2.79	6.45	0.4326	3.25	6.46	0.5031	1.86	0.97	1.9175
5	3.28	7.00	0.4686	3.56	7.30	0.4877	3.92	7.15	0.5483	2.12	1.01	2.0990
6	3.52	7.36	0.4783	3.93	7.78	0.5051	4.59	7.77	0.5907	2.24	1.15	1.9478
7	4.15	7.98	0.5201	4.69	8.50	0.5518	5.29	8.41	0.6290	2.47	1.19	2.0756
8	4.36	8.27	0.5272	5.10	8.94	0.5705	6.01	9.01	0.6670	2.59	1.33	1.9474
9	4.95	8.79	0.5631	5.89	9.59	0.6142	6.77	9.64	0.7023	2.81	1.41	1.9929
10	5.14	9.03	0.5692	6.36	10.03	0.6341	7.55	10.23	0.7380	2.97	1.66	1.7892

3.2. Detecting the moving objects by integrating the first, third, and fifth moments

It is known that the third and fifth moments contain more information than the first moment, so we can integrate the first, third, and fifth moments.

Because the first, third, and fifth moments are orthogonal, one can consider that the first moment is the projection of image $f(x, y, t)$ on axis 1; the third moment is the projection of image $f(x, y, t)$ on axis 3; the fifth moment is the projection of image $f(x, y, t)$ on axis 5; and the axes 1, 3, and 5 are orthogonal. For getting the perfect real moving objects using the first, third, and fifth moments, we may use the vector module of the 3D space to regain its actual measure, namely $M(x, y, t) = \sqrt{M_1^2 + M_3^2 + M_5^2}$.

Henceforward, we principally adopt the $M(x, y, t)$ as OGHMs images (OGHMIs). We notice that the OGHMIs contain more information than a single derivative image or single OGHMs.

3.3. Segmenting the motion objects

We have noticed that OGHMI (OGHM) is also equal to the image transformation. This transformation can suppress the background, and enhance the motion information. We also know, for the noise image, we cannot remove all noise by a Gaussian filter. For obtaining the true region of the moving objects, having calculated the OGHMIs, we then detect the moving objects by the use of the segmentation of such images. To do this, a threshold for each OGHMI should be determined. One of the well-known methods for the threshold determination for image segmentation is the invariable moments method (IMM) [2, 3, 25]. But this method has not taken into account the spatial and temporal relations between moving pixels, which are important for image sequence analysis. To improve the method, instead of using a binary segmentation based on the threshold thus determined, we use a fuzzy relaxation for the segmentation of the moment images by taking into account these relations with the help of a nonsymmetric π fuzzy membership function [2]:

$$\pi\left(M(x,y); T, M_{\min}(x,y)\right) = \begin{cases} 1 - \dfrac{2\left(T - M(x,y)\right)^2}{\left(T - M_{\min}(x,y)\right)^2} & \\ \quad \text{if } 0 < \dfrac{\left(T - M(x,y)\right)}{\left(T - M_{\min}(x,y)\right)} \le 0.5, & \\ 2\left[\dfrac{\left(T - M(x,y)\right)}{\left(T - M_{\min}(x,y)\right)} - 1\right]^2 & \\ \quad \text{if } 0.5 < \dfrac{\left(T - M(x,y)\right)}{\left(T - M_{\min}(x,y)\right)} \le 1, & \\ 1 \quad \text{if } M(x,y) \ge T, & \\ 0 \quad \text{otherwise}, & \end{cases} \tag{15}$$

where $M(x, y)$ is the gray level of the OGHMI at the point (x, y), T is the segmentation threshold of the moment image $M(x, y)$, which is obtained by using the (IMM) [2], $M_{\min}(x, y)$ is the minimum of $M(x, y)$. $\pi(M(x, y); T, M_{\min}(x, y))$ is still noted as $\pi(x, y)$.

For each point in the moment images, the membership function, which gives a measure of the "mobility" of each pixel in each moment image, is first determined. We then apply a fuzzy relaxation to the membership function images, which gives the final results of the moving pixel detection [2].

3.4. Analyzing the localization errors

According to the theory of Gaussian filtering, a Gaussian filter with larger standard deviation is less sensitive to the noises, but brings larger localization errors [26]. How much is the localization error of detecting the moving objects? Now we give the discussion.

Let $F(x, y, t) = (f(x, y, t) + n(x, y, t))$ be the input image; $n(x, y, t)$ is the Gaussian white noise image; $f(x, y, t)$ is the true image at time t. Given the image $F(x, y, t)$, the Gaussian

Figure 6: Initial image.

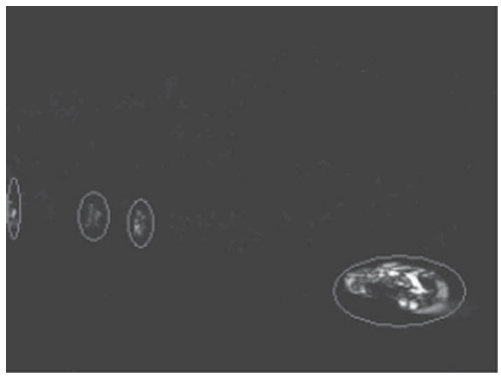

Figure 7: First OGHMs image.

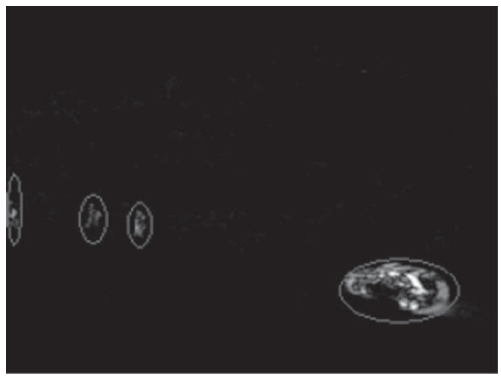

Figure 8: Third OGHMs image ($\sigma = 0.3$).

Figure 9: Third OGHMs image ($\sigma = 0.8$).

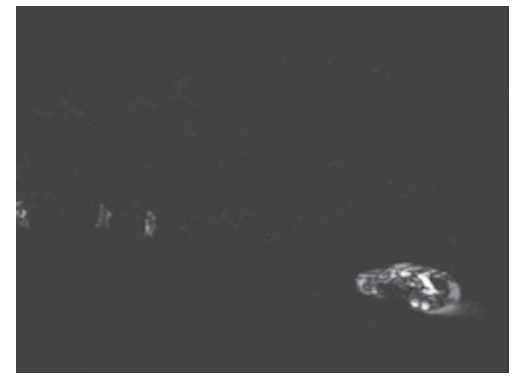

Figure 10: Fifth OGHMs image ($\sigma = 0.3$).

Figure 11: Fifth OGHMs image ($\sigma = 0.8$).

filter (finite impulse response) output is as follows:

$$H_F(x, y, t_0) = f(x, y, t) * g(t) + n(x, y, t) * g(t), \quad (16)$$

so we have

$$H_f(x, y, 0) = f(x, y, t) * g(t) = \int_{-w}^{w} f(x, y, -t)g(t)dt, \quad (17)$$

where ω is the length of integrating windows; in general, $\omega = 5\sigma$.

Supposing the noise variance is $n_0^2(x, y)$, then the noise output satisfies

$$E(H_n) = \left[E\left|n(x, y, t) * g(t)\right|^2\right]^{1/2} = n_0\left(\int_{-w}^{w} g^2(t)dt\right)^{1/2}, \quad (18)$$

where $H_f(x, y, t)$ and $H_n(x, y, t)$ are filter responses to image and noise, respectively. Let true edge point (we consider that the boundary of the background and the moving object point belong to the step-type edge, so its second derivative is

zero at the edge point [27]) at $t = 0$; and let t_0 be the edge point of the total response $H_F(x, y, t)$ (in fact, localization error point). Then,

$$H_F''(x, y, t_0) = H_f''(x, y, t_0) + H_n''(x, y, t_0) = 0, \tag{19}$$

where "′" represents the derivative of a function on t. Taylor expansion of $H_f''(x, y, t_0)$ at $t = 0$ is as follows:

$$H_f''(x, y, t_0) = H_f''(x, y, 0) + t_0 H_f'''(x, y, 0) + O(t_0^2), \tag{20}$$

where $H_f''(x, y, 0) = 0$ is assumed (the true edge point). We ignore the higher-order terms. Combining (19) and (20), we have

$$t_0 H_f'''(x, y, 0) \approx -H_n''(x, y, t_0). \tag{21}$$

From a derivative similar to that for noise, we have

$$E\left|H_n''(x, y, t_0)\right|^2 = n_0^2 \int_{-w}^{-w} \left(g''(t)\right)^2 dt. \tag{22}$$

Differentiating $H_f(x, y, t) = f(x, y, t) * g(t)$, and evaluating at $t = 0$ ($t \in \mathbb{R}$), we have

$$H_f'''(x, y, 0) = \int_{-w}^{w} f(x, y, -t) g'''(t) dt. \tag{23}$$

Combining (21), (22), and (23), the localization error is defined as follows:

$$\begin{aligned} T_0 &= \frac{1}{\left(E\left[(t_0)^2\right]\right)^{1/2}} \\ &= \frac{\left|\int_{-w}^{w} f(x, y, -t) g'''(t) dt\right|}{n_0 \left(\int_{-w}^{w} g''^2(t) dx\right)^{1/2}}. \end{aligned} \tag{24}$$

We must point out that the localization error T_0 is concerned with only the temporal domain (frame number). For obtaining the space localization error, we must estimate the motion speed of the moving object, then the space localization error equals to multiple the motion speed with $1/T_0$.

According to (24), we can obtain the following conclusions (1) If σ is bigger, the space localization error will be bigger. (2) If the speed of the moving object is faster, the space localization error will be bigger.

4. COMPARING THE EXPERIMENTAL RESULTS

In this section, we give some experimental results and the performance compared with other methods for detecting the moving objects. For these experiments, we choose the image sequence of the Reading University offered for test. In this image sequence, it contains 548 frame images (0953–1500). The size of each frame image is 768×576 and the gray level of pixel is 256. The images were acquired by a fixed camera and fixed parameters. Some original images are shown in Figures 12, 13, 14, and 15.

Figure 12: Initial image 0975.

Figure 13: Initial image 1105.

4.1. Comparing our method with a DM

For the original image 1105 (Figure 13), our experiments show that one cannot well detect the moving objects by using classical temporal DM because of the high-level noise. In Figure 16, we notice that the moving objects are very fuzzy and that there exist lots of noises. So this method is very sensitive to the noises.

However Figure 17 shows that the moving objects are detected by using OGHMs with $\sigma = 0.8$, and segmented by FRM [2]. We see that the moving objects are very well detected. The experiment result shows the good performance of our method.

4.2. Comparing our method with the BS method

In order to further test the performance of our method, the method of the BS is employed. To get the test image sequence, we artificially change the illumination condition of 10 frames (0980–0989) among 548 images (0953–1500). For this sequence, we first take the average of these 548 frames to generate a background image, and employ the method of the BS to detect the moving objects. Unfortunately, it fails for the 10 frames as the illumination was changed; an experimental result is shown in Figure 18; the background and the moving objects mix together. However, using the OGHMs, we succeeded to detect the moving objects in all frames of the image sequence, except 4 frames (0979, 0980, 0989, 0990);

Figure 14: Initial image 0981.

Figure 15: Initial image 1140.

Figure 16: Detecting the moving objects using DM, segmented by FRM (1105).

Figure 17: Detecting the moving objects using OGHMs, segmented by FRM ($\sigma = 0.8$).

Figure 18: Illumination abrupt change; background and moving objects are mixed by BS (0981).

Figure 19: Illumination change; detecting the moving objects by third OGHMs (0981).

an experimental result is shown in Figure 19. We can see that the moving objects are very well detected.

4.3. Comparing the real objects with the detected moving objects

Figure 20 shows an experiment result of detecting the moving objects by using OGHMs and segmenting the moving object using the 3D MRM [2]. Figures 21 and 22 show the superposition results of the original images with the moving objects that were detected by using OGHMs and the moving objects were segmented by 3D MRM. We can see that these detected moving objects conform to the real moving objects.

4.4. Comparison with ABS

To improve the performance of BS methods for motion detection, one can use ABS methods to update the background

Figure 20: Detecting the moving objects by OGHMs, segmented by 3D MRM (1105).

Figure 21: Superposition of the moving objects with initial image (0975).

Figure 22: Superposition of the moving objects with initial image (1105).

image at each instant so that it can take into account the environment changes such as the illumination change. One of the problems of such adaptive methods is the choice of value of the "learning rate" parameter for background updating, which in fact depends on the velocity of the moving objects. Unfortunately, in general, the velocity of the moving objects in a dynamic scene can change from time to time and at the same instant, it can change from object to object.

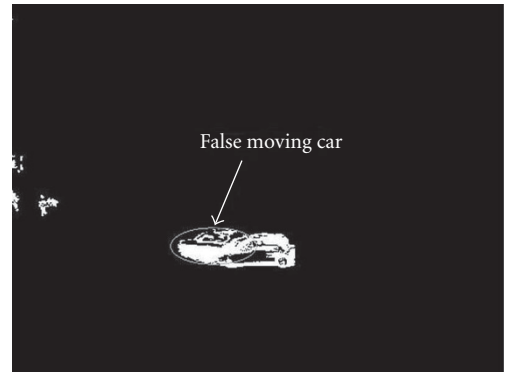

Figure 23: Detecting the moving objects using the adaptive background method (1140).

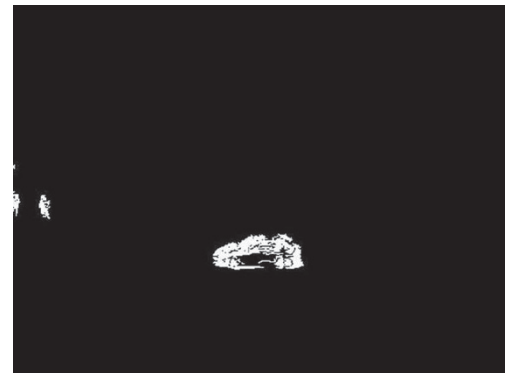

Figure 24: Detecting the moving objects using the OGHMs (1140).

Moreover, such methods can correct the illumination changes just efficiently in background updating only after the accumulation of a sufficiently large number of frames. Another problem of such methods is that when there exist slowly moving objects in the sequence, these moving objects would be considered as static objects in the background during the background updating process. So, when one detects the moving objects by BS, there are some risks to detect false moving objects that are in reality the trace of slowly moving objects at the preceding instants. By the use of OGHMs, because the background images are not at all used, such problems can be solved much easily. An experimental result is shown in Figure 23. From Figure 23, we can see that by the ABS method, two moving cars are detected, of which only one is the real moving car at the instant and another is in fact a false one that in reality does not exist at the instant. Figure 24 shows an experiment result of using our method. We can see that the moving objects are very well detected.

4.5. Comparing the integration performance with other methods

Detecting the moving objects can be divided into two categories: online detection and the offline detection. The online detection applies principally to real-time surveillance and so

Table 2: Comparing the performances of different method.

Method name	Computation of each pixel	Real time/online/offline	Advantages	Shortages
Temporal differential	Addition: 1 time	Approximation of real time; online, offline	Simple computing	Sensitive to the noise Existing localization errors
Temporal differential (Gaussian filtered)	Multiplication: $10\sigma + 1$ times Addition: $10\sigma + 1$ times	Approximation of real time; online, offline	Antinoises	Existing localization errors
Simple BS	Addition: average 2 times	Offline	Precise localization	No real time; sensitive to the noise
ABS K mixed Gaussian models	Multiplication: $20K$ times Addition: $7K$ times Exponential: K times Square root: K times	Approximation of real time; online	Precise localization	Sensitive to the noise; computation complexity; demand images accumulation
OGHMs	Multiplication: $10\sigma + 1$ times Addition: $10\sigma + 1$ times	Approximation of real time; online, offline	Antinoises	Existing localization errors

forth. The offline detection applies principally to the analysis of traffic accident and so forth. In general, the online detection treats with the long image sequence, for the past images, if without special demand, it does not save the images. And the offline detection treats with the short image sequence.

The online detection does not demand the high precise localization of the moving objects. It only demand to detect the moving objects in view range, for example, the automatic safe door opens and closes and so forth. However the offline detection demand high precise localization of the moving objects.

In evaluating the advantages or disadvantages of a detector of moving objects, we must pay attention to its applied situation, in addition to some convenience criterions. For example, the ABS for the offline detection is not significative, because obtaining a good background model must demand a lot of images.

In this section, some typical methods' performances advantages and disadvantages are shown in Table 2. It shows that our method has these advantages as follows.

(1) It can be used for the online and real-time detection. We know that for obtaining the moving objects at T time. We demand only the past 5σ frame image and later 5σ frame image. When σ is not very large, such as 1, 5 frame images are demanded after T time. In general, the CCD has the frame ratio 25/s. So delaying 0.2 second is completely acceptable. If we adopt the single-direction expansion technique, namely $f(x, y, T) = f(x, y, T + 1) = \cdots = f(x, y, T + 5\sigma)$, then it can apply to the online and real-time system.

(2) It does not need the image accumulations, except the simple DM; the other methods need image accumulations.

(3) It has the stronger antinoise ability.

The disadvantage is the existence of the localization errors.

Although the OGHMs method and the temporal DM belong to the same category, they are different since the OGHMs are not simple differential. However, it is reasonable weight of the differentials of different order.

4.6. Simple statistic comparison of SNR

In the domain of detecting the moving objects, the SNR is different from traditional SNR. Here the SNR is defined as follows: $\text{SNR} = M/(N_1 + N_2)$, where M is the total number of motion pixel in the believable (true) region of moving objects; N_1 is the total number of undetected motion pixels in the believable region of the moving objects and N_2 is the total number of the detected motion pixels in the nonbelievable region of the moving objects.

Practically, it is difficult to obtain the believable region of the moving objects, because the obtained believable region of the moving objects is the same as that of the detected. To our understanding, no comparable reports concerning the SNR of the motion detection were published up till now.

Here, we only can simply compare the SNR using the artificial believable region of the moving objects. In our experiment, 100 successive images are employed. For each image, the believable region of the moving objects is extracted artificially. Let D represent the total pixel number of the detected motion in the believable region of the moving objects. The experiment results are shown in Table 3.

4.7. Experiment of localization errors

Theoretically, according to (24), we can obtain the localization errors of the detected moving objects, and correct the detection results. However, in practice, because the discrete data can lead to error; sometimes, the experiment results are not accordant with the true case.

Table 3: The statistics showing the experiment results of SNR based on 100 images.

The name of the method	D (average)	N_1 (average)	N_2 (average)	SNR (average)
Temporal differential (unfiltered)	3547	2905	3262	1.0462
BS (average method, unfiltered)	5391	1061	957	3.1972
ABS	5169	1283	975	2.8574
OGHMs	5320	1132	1543	2.4120

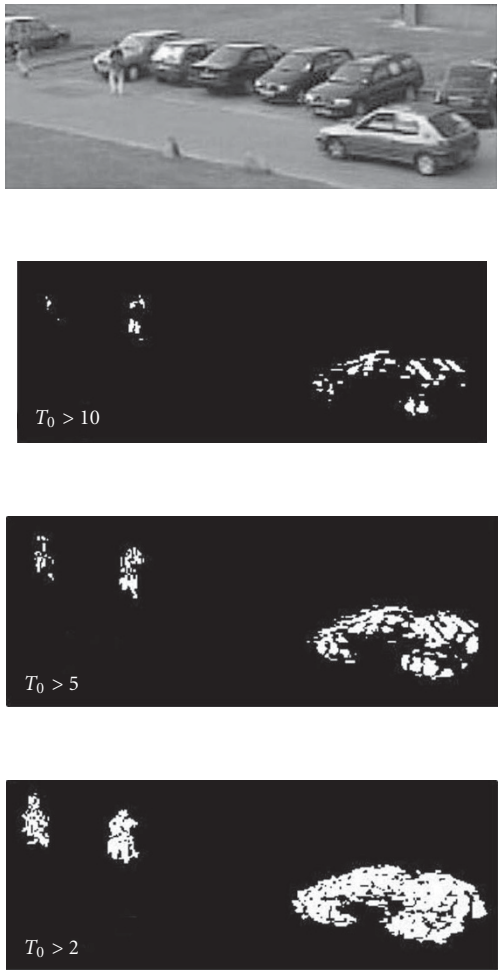

Figure 25: The temporal localization errors of the moving objects and their sketch maps.

To observe the localization errors, the points of having localization errors are represented by white. For example, $T_0 > 10$ represent the motion points that the temporal localization errors are less than 1/10 frame; $T_0 \geq 2$ represent the motion points that the temporal localization errors are less than 0.5 frame (see Figure 25).

5. CONCLUSIONS

In this paper, we have analyzed some properties of the OGHMs and proposed a new method for motion detection using the OGHMs. The experiment results are also reported, which show good performance of our method.

The main contribution of this paper is as follows. (1) Pointing out some properties of OGHMs. (2) Analyzing the meaning of each order moment. (3) Proposing a new method of detecting the motion objects using OGHMs. (4) Comparing the experiment results with other methods.

As for the application of OGHMs, we only make a try. The obtained results are simple; we still have a lot of research work to be completed; for example, (1) the antinoise ability (concrete quantification) of the OGHMs is still open; (2) equation (24) is a formula for estimating the localization error. However, because of the discrete data can lead to error, how much error is arisen by using discrete data is still open.

ACKNOWLEDGMENTS

This paper is supported by the advanced research plan of France and China (PRA SI 01-03), Chinese National Ministry of Education Science Foundation (2000–2003), and Institut National de Recherche Informatique et en Automatique (INRIA), France (2002–2003). We would like to thank Professor Mo Dai for his many beneficial suggestions.

REFERENCES

[1] J. Shen, W. Shen, D.-F. Shen, and Y. Wu, "Orthogonal moments and their application to motion detection in image sequences," *International Journal of Information Acquisition (IJIA)*, vol. 1, no. 1, pp. 77–87, 2004.

[2] Y. Wu and J. Shen, "Moving object detection using orthogonal Gaussian-Hermite moments," in *Visual Communications and Image Processing*, vol. 5308 of *Proceedings of SPIE*, pp. 841–849, San Jose, Calif, USA, January 2004.

[3] Y. Wu and J. Shen, "Detection the moving objects using orthogonal moment and analyze the action of moving objects," *Guizhou Science*, vol. 22, no. 3, pp. 20–28, 2004.

[4] J. Shen, Y. Wu, et al., "Motion detection and orthogonal moments," in *Proceeding of Conference on Science and Technology of Information Acquisition and the Application*, pp. 182–190, Auhui, China, December 2003.

[5] J. Shen, W. Shen, and D.-F. Shen, "On geometric and orthogonal moments," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 14, no. 7, pp. 875–894, 2000.

[6] J. Shen, "Orthogonal Gaussian-Hermite moments for image characterization," in *Intelligent Robots and Computer Vision XVI: Algorithms, Techniques, Active Vision, and Materials Handling*, vol. 3208 of *Proceedings of SPIE*, pp. 224–233, Pittsburgh, Pa, USA, October 1997.

[7] J. Shen and D.-F. Shen, "Orthogonal Legendre moments and their calculation," in *Proc. 13th IEEE International Conference on Pattern Recognition (ICPR '96)*, vol. 2, pp. 241–245, Vienna, Austria, August 1996.

[8] J. Shen and D.-F. Shen, "Image characterization by fast calculation of Legendre moments," in *Image and Signal Processing for Remote Sensory III*, vol. 2955 of *Proceedings of SPIE*, pp. 295–306, San Jose, Calif, USA, December 1996.
[9] B. C. Li and J. Shen, "Two-dimensional local moment, surface fitting and their fast computation," *Pattern Recognition*, vol. 27, no. 6, pp. 785–790, 1994.
[10] P. J. Green, "On use of the EM algorithm for penalized likelihood estimation," *Journal of the Royal Statistical Society: Series B*, vol. 52, no. 3, pp. 443–452, 1990.
[11] I. Haritaoglu, D. Harwood, and L. S. Davis, "W^4: real-time surveillance of people and their activities," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 809–830, 2000.
[12] C. Stauffer and W. E. L. Grimson, "Learning patterns of activity using real-time tracking," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 747–757, 2000.
[13] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR '99)*, vol. 2, pp. 246–252, Fort Collins, Colo, USA, June 1999.
[14] C. Eveland, K. Konolige, and R. C. Bolles, "Background modeling for segmentation of video-rate stereo sequences," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR '98)*, pp. 266–271, Santa Barbara, Calif, USA, June 1998.
[15] H. Fujiyoshi and A. J. Lipton, "Real-time human motion analysis by image skeletonization," in *Proc. IEEE Workshop on Application of Computer Vision (WACV '98)*, pp. 15–21, Princeton, NJ, USA, October 1998.
[16] C. Bregler, "Learning and recognizing human dynamics in video sequences," in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '97)*, pp. 568–574, San Juan, Puerto Rico, USA, June 1997.
[17] A. Blake, M. Isard, and D. Reynard, "Learning to track the visual motion of contours," *Artificial Intelligence*, vol. 78, no. 1-2, pp. 179–212, 1995.
[18] C. Ridder, O. Munkelt, and H. Kirchner, "Adaptive background estimation and foreground detection using Kalman-filtering," in *Proc. International Conference on Recent Advances in Mechatronics (ICRAM '95)*, pp. 193–199, Istanbul, Turkey, August 1995.
[19] S. A. Niyogi and E. H. Adelson, "Analyzing and recognizing walking figures in XYT," in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '94)*, pp. 469–474, Seattle, Wash, USA, June 1994.
[20] R. Polana and R. Nelson, "Low-level recognition of human motion," in *Proc. IEEE Workshop Nonrigid and Articulated Motion*, pp. 77–82, Austin, Tex, USA, November 1994.
[21] R. Polana and R. Nelson, "Detecting activities," in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '93)*, pp. 2–7, New York, NY, USA, June 1993.
[22] J. Yamato, J. Ohya, and K. Ishii, "Recognizing human action in time-sequential images using hidden Markov model," in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '92)*, pp. 379–385, Champaign, Ill, USA, June 1992.
[23] A.-R. Mansouri, "Region tracking via level set PDEs without motion computation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 947–961, 2002.
[24] R. Nelson, "Qualitative detection of motion by a moving observer," *International Journal of Computer Vision*, vol. 7, no. 1, pp. 33–46, 1991.
[25] R. Wang, "Comprehension of image," *Science Publishing Company*, pp. 95–101 and pp. 235–290, Hulan, 1994.
[26] S. Castan and J. Shen, "Box filtering for Gaussian-type filters by use of the B-Spline functions," in *Proc. 4th Scandinavian Conference on Image Analysis*, pp. 235–243, Trondheim, Norway, June 1985.
[27] J. Shen and S. Castan, "An optimal linear operator for step edge detection," *Computer Vision, Graphics, and Image Processing*, vol. 54, no. 2, pp. 112–133, 1992.

Youfu Wu is an Associate Professor and a Director of Education of Minority National Programs, University of Guizhou, China. Actually he is a Ph.D. student at Bordeaux-3 University, France. He is the author/coauthor of more than 20 publications in image processing and computer vision. Principal research interests include pattern recognition, image processing, and computer vision.

Jun Shen is a Professor and the Head of Image Laboratory at EGID Institute, Bordeaux-3 University, France. He received the "Doctorat d'État" degree from Paul Sabatier University, Toulouse, France, in 1986. He was awarded the "Outstanding Paper Honourable Mention" from the IEEE Computer Society in 1986. He is the author/coauthor of more than 130 publications in image processing and computer vision. Unfortunately, he died.