ZOOM-FACTOR AND ROTATION ANGLE ESTIMATION USING ZERNIKE MOMENTS

Lj.Panovski, Z.Ivanovski

Institute of Electronics,
Faculty of Electrical Engineering,
University "SS. Cyril and Methodius" of Skopje,
Karpos II bb,
Skopje, Republic of Macedonia.

ABSTRACT

A two-step procedure for zoom-factor and rotation angle estimation, based on the use of Zernike moments, is proposed in this paper, according to which the zoom-factor is estimated in the first step as a precondition for the estimation of the rotation angle in the second step. The experimental results show that the proposed method is robust in respect of the variety of images and to additive noise and has potential applicability to real-life video sequences involving camera rotation and camera zoom.

1. INTRODUCTION

Because of the insensibility of their magnitudes to rotation, the use of Zernike moments in pattern recognition has been studied extensively in the last decades [1], [2], [3]. It has recently been shown that this property of Zernike moments could be effectively used to reliably estimate the rotation angle of objects in some robot applications [4], i.e. when a robot arm has to manipulate objects on a conveyor belt. In such cases, the robot must first recognize the object, regardless of its orientation angle, and then to estimate its rotation angle for manipulation.

However, as it is shown in this paper, Zernike moments could be used for the estimation of the zoom-factor as well. An estimator of the rotation angle and the zoom-factor can be effectively applied in machine vision tasks when objects having different orientations and being viewed from different distances should be picked up by a robot, or in motion estimation tasks when scenes with both zoom and rotation are considered.

A situation with involving both zoom and rotation could be treated as follows: the zoom-factor is estimated in the first step, and then the estimation of the rotation angle is done in the second step. Since a convenient procedure for the estimation of the rotation angle is already given in [4] , the emphasis in this paper is put on the estimation of the zoom-factor. Once the zoom-factor is estimated, in order to estimate the rotation angle one can almost straightforwardly apply the estimation procedure given in [4].

2. ZERNIKE MOMENTS

The Zernike moments of order n and repetition m, $Z_{nm}(r,q)$, of a grey-scale image, $f(r,q)$, are defined inside the unit circle of the (r,q) polar coordinate system, as follows:

$$Z_{nm} = \frac{n+1}{p} \sum_{over\ unit\ disk} \sum V^*_{nm}(r,q)\, f(r,q) \qquad (1)$$

where:

$$V^*_{nm}(r,q) = R_{nm}(r) e^{-jmq}, \quad r \le 1 \qquad (2)$$

are the complex-conjugates of the Zernike basis images, and $R_{nm}(r)$ are the radial polynomials, defined as:

$$R_{nm}(r) = \sum_{s=0}^{(n-m)/2} (-1)^s \frac{(n-s)!}{s!\left(\frac{n+m}{2}-s\right)!\left(\frac{n-m}{2}-s\right)!} r^{n-2s} \qquad (3)$$

In (3), n and m are integers, with $n \ge m \ge 0$ and $n - m$ is even.

A. Zernike Moments of Zoomed and Rotated Images

Let $f_1(r,q)$ be an image known over the unit disk. Let $f_2(r,q)$, be its shrunken and rotated version, known over a disk with radius $1/a$, $a \ge 1$. Then the following equation is valid for $r \le 1/a$:

$$f_1(ar, q + a) = f_2(r,q) \qquad (4)$$

Now, let us denote the Zernike moments of order n and repetition m of the image $f_1(r,q)$ with $Z^{(1)}_{nm}$:

$$Z^{(1)}_{nm} = \frac{n+1}{p} \sum_{over\ unit\ disk} \sum R_{nm}(r)\, f_1(r,q) e^{-jmq} \qquad (5)$$

and let us introduce the following notation regarding the image $f_2(r,q)$:

$$Z^{(2,a)}_{nm} = \frac{n+1}{p} \sum_{r=0}^{1/a} \sum_{q=0}^{2p} R_{nm}(r)\, f_2(r,q) e^{-jmq} \qquad (6)$$

Then, from (4)-(6), we have:

$$Z^{(2,a)}_{nm} = \frac{e^{jma}}{a^n} \frac{n+1}{p} \sum_{over\ unit\ disk} \sum R^{(a)}_{nm}(r)\, f_1(r,q) e^{-jmq} \qquad (7)$$

where:

$$R^{(a)}_{nm}(r) = \sum_{s=0}^{(n-m)/2} (-a^2)^s \frac{(n-s)!}{s!\left(\frac{n+m}{2}-s\right)!\left(\frac{n-m}{2}-s\right)!} r^{n-2s} \qquad (8)$$

In case of rotation only, i.e. for $a = 1$, from (5) to (8) the following property of the Zernike moments can be easily shown:

$$Z^{(2)}_{nm} = Z^{(1)}_{nm} e^{jma} \; ; n \ge m \ge 0 \qquad (9)$$

However, when zooming is also involved, in order to have computationally simple relations, we have to restrict the analysis to Zernike moments having the same order as repetition, i.e., to the cases when $n = m$. For $n = m$, from (4)-(8), we have the following property of the Zernike moments:

$$Z_{nn}^{(2,a)} = \frac{1}{a^n} Z_{nn}^{(1)} e^{jna} \tag{10}$$

and

$$\frac{\left|Z_{nn}^{(1)}\right|}{\left|Z_{nn}^{(2,a)}\right|} = a^n \tag{11}$$

It should be noted here that in (10) and (11) $Z_{nn}^{(1)}$ and $Z_{nn}^{(2,a)}$ are to be calculated over different sampling grids, the one used for $Z_{nn}^{(2,a)}$ being the finer of the two.

From a computational point of view it is much better to apply the sampling grid of $Z_{nn}^{(2,a)}$ for the calculations of $Z_{nn}^{(1)}$, too. When $Z_{nn}^{(1)}$ and $Z_{nn}^{(2,a)}$ are calculated using the sampling grid of $Z_{nn}^{(2,a)}$, relation (11) has to be rewritten in the following manner:

$$\frac{\left|Z_{nn}^{(1)}\right|}{\left|Z_{nn}^{(2,a)}\right|} = a^{n+2} \tag{12}$$

We shall use relation (12) for the estimation of the zoom-factor a.

3. ZOOM-FACTOR ESTIMATION

Having two known images, $f_1(r,q)$ and $f_2(r,q)$, related according to (4) by an unknown zoom-factor a, and an unknown rotation angle a, we want to obtain a good estimate, $\hat{a}$, for the unknown value of a, and then to use $\hat{a}$ to obtain a good estimate $\hat{a}$ for the unknown rotation angle a.

We shall restrain the estimation procedure within a reasonable range of expected zoom-factors, $a_{min} \le a \le a_{max}$, and will then proceed as follows.

For a given value of n, we calculate $Z_{nn}^{(1)}$. Then, using the same sampling grid as for $Z_{nn}^{(1)}$, we calculate $Z_{nn}^{(2,k)}$ for the whole range of expected zoom-factors and construct the following quotient function:

$$q_n(k) = \frac{1}{k^{n+2}} \frac{\left|Z_{nn}^{(1)}\right|}{\left|Z_{nn}^{(2,k)}\right|}, \; a_{min} \le k \le a_{max} \tag{13}$$

Figure 1 shows four such quotient functions of an image obtained for four different values of n.

Before we proceed with the explanation of the proposed procedure, there is a need for a brief discussion on the properties of $q_n(k)$.

Let k_r denote the value of k for which $q_n(k_r) = 1$. Theoretically, provided the range of expected zoom-factors is correctly set, by inspecting the point at which $q_n(k_r) = 1$, one could expect to obtain the exact value of a, because relations (12) and (13) imply $a = k_r$. As an example, see $q_3(k)$ in Fig. 1.

However, it is possible to have $q_n(k_i) = q_n(k_j) = q_n(k_k) = \ldots = 1$ for $k_i \ne k_j \ne k_k \ldots$ The main cause for this is that, for a given n, it is possible to have:

$$\frac{\left|Z_{nn}^{(1)}\right|}{\left|Z_{nn}^{(2,k)}\right|} = k^{n+2} \tag{14}$$

for values of $k \ne a$, even under ideal conditions, because it is quite possible for two different images to have identical projections onto a common basis image. Namely, $f_1(r,q)$ for $r \le 1$ is in general different than $f_2(r,q)$ for $r \le 1/k \ne 1/a$. Thus this situation is inherent to the very problem and can not be prevented neither theoretically, nor practically. Of course, one of the solutions of (14), in case of multiple solutions, should correspond to $k_r \cong a$. Let us denote the solution $k_r \cong a$ as the true solution, and all other solutions as false ones. Experiments have shown that the false solutions, when they exist, are randomly spread along the k-axis, and that they occur at different values of k for different values of n. See, for example, $q_2(k)$ and $q_{14}(k)$ in Figure 1.

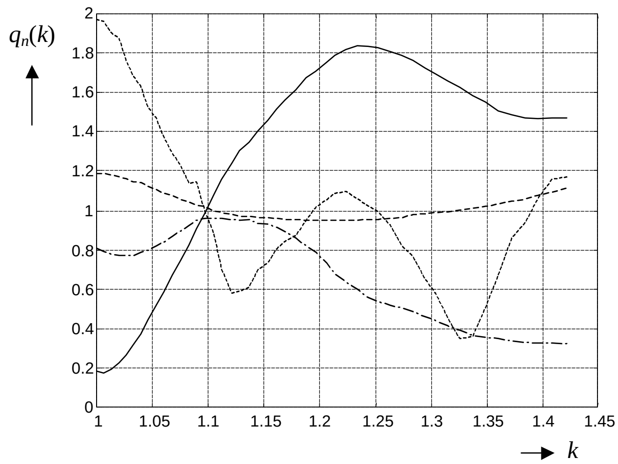

Fig.1 Four quotient functions for an image:
$q_2(k)$ - - - - - $q_6(k)$ –·–·–··
$q_3(k)$ ——— $q_{14}(k)$ – – –

Because of round-up errors and noise, and in cases when the sampling grid, although appropriately chosen with respect to the spread of the spatial spectrum of $f_1(r,q)$, is too coarse with respect to the spread of the spatial spectrum of $f_2(r,q)$, the value of the true solution could be slightly altered. We assume that the distribution for different values of n of the values for the true solution is a Gaussian one with zero mean and unknown variance.

The reasons causing variations of the true solution could in certain cases also cause the case when $q_n(k) \ne 1$ for $a_{min} \le k \le a_{max}$, even though the range of expected zoom-factors was correctly set. See, for example, $q_6(k)$ in Figure 1. Experiments have shown, though, that the occurrence of this situation depends in a random manner on the value of n.

Having in mind the aforesaid considerations, in the proposed procedure for the estimation of the zoom-factor, we calculate a set of moments with increasing values of n and create a family of quotient functions. We then determine all possible solutions of $q_n(k) = 1$ for all n. For each such solution a Gaussian probability density function is then constructed, with the value of the solution as its mean. All probability density functions could have the same variance. Then a total probability density function is constructed as a sum of all individual pd-functions. Finally, the value k_M, for which the maximum of the total pdf occurs, is taken to be the estimate of the zoom-factor a, $\hat{a} = k_M$.

One might think that computing $Z_{nn}^{(2,k)}$ for all k within the range $a_{min} \le k \le a_{max}$ would be computationally a rather complex task. Fortunately, this is not the case. We start by computing the necessary set of Zernike moments $Z_{nn}^{(2,k)}$ for the highest value of $k = a_{max}$. Having the values of the set $Z_{nn}^{(2,a_{\max})}$, all we need in order to obtain the set $Z_{nn}^{(2,k)}$ for the next lower value of k is to perform the calculations only for the pixels that belong to the outer ring on the (r,q)-plane having an inner radius equal to $1/a_{max}$ and a thickness of 1 pixel, and to add then the results to $Z_{nn}^{(2,a_{\max})}$. This procedure is repeated for all other lower values of k until $k = a_{min}$ is reached. Thus, the overall computational complexity is equivalent to the calculation of the sets $Z_{nn}^{(1)}$ and $Z_{nn}^{(2,a_{\min})}$.

4. ROTATION ANGLE ESTIMATION

When the zoom-factor is estimated, the estimation of the rotation angle is done following the procedure proposed in [4], applied on:

$$a = \frac{1}{n}\arg\left(\hat{a}^{n+2}\frac{Z_{nn}^{(1)}}{Z_{nn}^{(2,\hat{a})}}\right) \qquad (15)$$

5. EXPERIMENTAL RESULTS

Simulation experiments were performed with 87 random chosen 8 bits/pixel grey-scale images with a resolution of 640×480 pixels. These were down-scaled, using bilinear interpolation, with four different zoom-factors: a = 1.05, 1.1, 1.15 and 1.2. This image set was then used to test the proposed procedure for zoom-factor estimation. The corresponding algorithm was implemented in Matlab. Gaussian probability density functions with different variances were tried, but best results were achieved with $s^2 = 0.001$ and those are the ones that are presented in this paper.

Figure 2 illustrates how the obtained average errors fluctuate depending on the size of the set of used Zernike moments, S_n. Figure 3 shows the dependence of the errors' variances on the size of the set of Zernike moments.

In Figures 2 and 3, $S_n = 0$ denotes that just the Zernike moments of order and repetition 0 were used, whereas, for instance, $S_n = 5$ means that a set containing all Zernike moments of order and repetition 0 up to order and repetition 5 was used. The four shown curves correspond to the four zoom-factor values used in the tests.

As it could be seen, the amplitudes of the fluctuations of both the average error and the error variance are significant only for the lowest set sizes S_n, and then drop to quite lower values for $S_n \ge 8$.

Table 1 gives the obtained average error, error variance, maximal error and maximal error on the unit circle expressed in pixels for the whole image test set in the case when $S_n = 20$ was used, for all tested values of a.

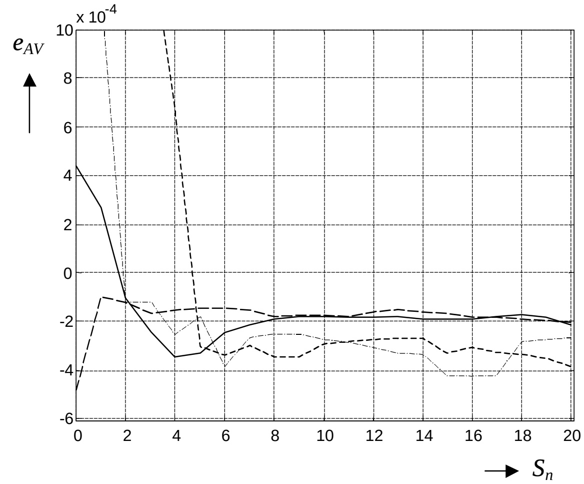

Fig.2 Fluctuations of the average error e_{AV} with the set size S_n for four different zoom-factors a

a = 1.05 ——— a = 1.15 - - - - -
a = 1.10 — — — a = 1.20 —·—··

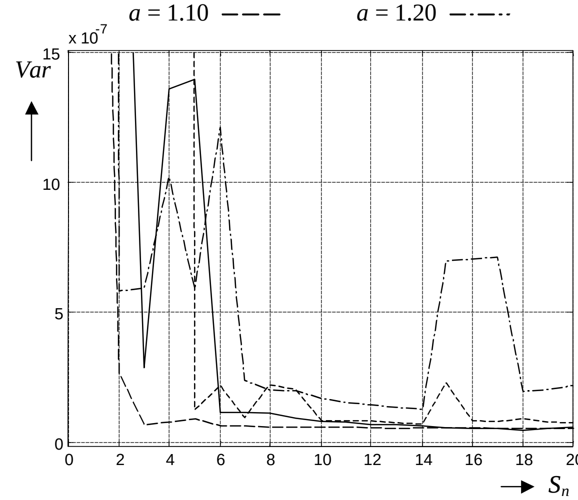

Fig.3 Dependance of the error variance on the set size S_n for four different zoom-factors a

a = 1.05 ——— a = 1.15 - - - - -
a = 1.10 — — — a = 1.20 —·—··

	a = 1.05	a = 1.1	a = 1.15	a = 1.2
Average error	-2.1×10^{-4}	-2×10^{-4}	-3.9×10^{-4}	-2.7×10^{-4}
Error variance	5.6×10^{-8}	5×10^{-8}	7.5×10^{-8}	2.2×10^{-7}
Maximal error	5×10^{-4}	5×10^{-4}	5×10^{-4}	2.4×10^{-3}
Maximal error in pixels	0.12	0.12	0.12	0.58

Table 1 The average error, error variance , maximal error and maximal error in pixels on the unit circle for four different values of zoom-factor

Experiments were also performed with Gaussian zero mean noise added to the zoomed images, but with an image set size of 30 and for just one value of the zoom-factor.

Table 2 shows the results for different SNR levels.

SNR [dB]	30	20	10
Average error	-2.1×10^{-4}	-2.5×10^{-4}	-1.9×10^{-4}
Error variance	3.1×10^{-8}	5.2×10^{-7}	4.1×10^{-5}
Maximal error	2×10^{-4}	9×10^{-4}	0.0114
Maximal error in pixels	0.048	0.216	2.736

Table 2 The average error, error variance , maximal error and maximal error in pixels on the unit circle for three different SNR levels in the case of a = 1.1

6. CONCLUSION

In this paper we propose a new and robust method of estimating the zoom-factor of an image using the magnitude information of Zernike moments whose orders are equal to their repetitions.

A number of simulation experiments were performed over a set of randomly chosen images and very good results were obtained. The results of the simulations on noise-free images exhibit an average error of the order of 10^{-4} when set size S_n was chosen large enough (over 8). This is equivalent to a maximal error at the top or the bottom of the image frame of 0.24 pixels. In the worst individual case the maximal error reached 0.58 pixels, a value that can still be considered as very good.

The simulation experiments that were performed on noisy images showed that the method gives good values for the error variance and the maximal error when the SNR level is equal or higher than 20 dB. Although the error variance and maximal error exhibit an increase for two orders of magnitude when the SNR level is 10 dB, the average error remains at the same order of magnitude.

The fact that the proposed method performed well on a randomly chosen set of noise-free and noisy images indicated its robustness and applicability to real-world situations. Indeed, the proposed method was tried on a quite small set of consecutive frames from video sequences containing true camera zoom action. On the grounds of the results obtained in these tries we can conclude that in order to obtain good results in real-life video sequences, the essential thing is to compensate all translational motion prior to the application of the proposed zoom-factor estimation method.

REFERENCES:

[1] L.Z.Wang, G.Healy, "Using Zernike Moments for the Illumination and Geometry Invariant Classification of Multispectral Texture", *IEEE Trans. Image Processing*, vol. 7, No.2, pp. 196-203, 1998

[2] P.J.G. Lisboa, S.J. Perantonis, "Invariant Character Recognition: Comparison of Zernike moments and High-Order Neural Networks", *Proc. Int. Joint Conf. on Neural Networks (IEEE&INNS)*, pp. 1421-1425, Singapore, Nov.1991.

[3] S.X. Liao, M. Pawlak, "On the Accuracy of Zernike Moments for Image Analysis", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20, No.12, pp. 1358-1364, Dec. 1998.

[4] Whoi-Yul Kim, Young-Sung Kim, “Robust Rotation Angle Estimator” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 21, No. 8, pp. 768-773, August 1999.