Vision based System for Optical Number Recognition

Anand B. Deshpande
Banashankri 3rd Stage
Bangalore. Karnataka.
India

ABSTRACT

Vision based system for Optical Number recognition (ONR) deals with the recognition of processed numbers rather than magnetically processed ones. ONR is a process of automatic recognition of numbers by computers in images and digitized pages of text. ONR is one of the most fascinating and challenging areas of pattern recognition with various practical applications. It can contribute immensely to the advancement of an automation process and can improve the interface between man and machine in many applications. Moments and functions of moments have been extensively employed as invariant global features of images in pattern recognition. This paper shows the implementation and analysis of ONR, regardless of orientation, size and position, feature vectors are computed with the help of statistical moments.

Keywords

Pattern recognition, Optical Number Recognition, Moments, Resolution.

1. INTRODUCTION

Optical Number Recognition (ONR) is a software technology aimed to enable computers that recognizes numbers which are in the form of digital images, without human intervention. ONR has been extensively studied in the last half century and progressed to a level, sufficient to produce technology driven applications. The applications of ONR have been found in automated guided vehicles (AGV), object recognitions, digital libraries, packaging industries etc. ONR includes essential problems of pattern recognition, which are common to other related topics such as 2D object recognition and image retrieval systems.

This paper explains algorithm to recognize numbers printed on credit card, mobile SIM (Subscriber Identity Module) or in any packaging industries in a scene regardless of its position, size, orientation, etc.

The section 2 gives the detailed information of the algorithm used to implement the ONR. Section 3 shows the obtained results and analyses the result. Section 4 gives the conclusion.

2. ALGORITHM

The Optical Number Recognition is carried out in mainly two stages. They are Training and Testing. Training stores the features of the trained images in a database. And Testing deals with the comparison of the test image features with trained image features to recognize the correct image. The below is the block diagram of the ONR engine.

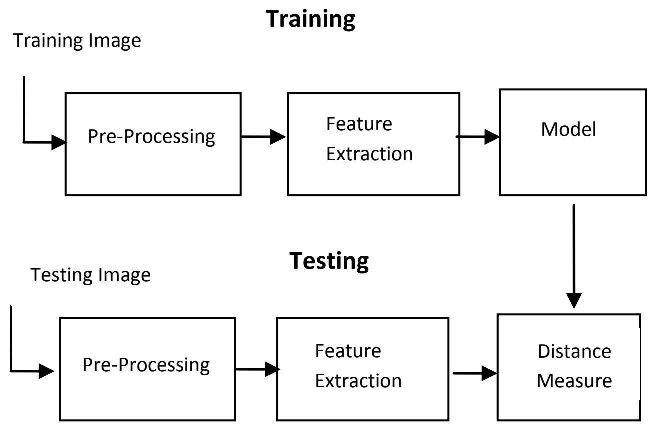

Fig. 1. Block diagram of ONR

2.1 Pre-processing

The preprocessing stage aims to make the image be suitable for different feature extraction algorithms. The binarization is performed using thresholding, a technique to partition the colors in the image into two sets using Niblack's algorithm [1]. Niblack's algorithm is a local thresholding method based on the calculation of the local mean and of local standard deviation. The threshold is decided by the formula:

$$T(x,y) = \mu(x,y) * \left[1 + k * \left(1 - \frac{\sigma}{C}\right)\right] \qquad (1)$$

Where T(x,y)- threshold, $\mu(x,y)$ -Mean, k & c- constant, and σ -standard deviation. The value of k is used to adjust how much of the total print object boundary is taken as a part of the given object. The constant k & care used to reduce its sensitivity to noise. The edges of binary images are obtained using Canny edge filter, which can estimate horizontal, vertical, and diagonal edges by detecting local maxima across the image's intensity gradient in these directions.

2.2 Feature Extraction:

The problem in ONR is the automatic recognition of a number in a scene regardless of its position, size, orientation, etc. In order to recognize different variations of the same number, image features which are invariant to certain transformations need to be used. Image invariants are features which have approximately the same values for samples of the same image which are, for instance, translated, scaled, rotated, skewed, blurred, or noise affected.

Shape-based image invariants are the most commonly used image features [2][3][4]. Image recognition based on these invariants includes three major issues: shape representation, shape similarity measure and shape indexing. Among these issues, shape representation is the most important issue. Various shape representation methods and shape descriptors exist in literatures. These methods can be classified into two categories: boundary-based invariants and region-based invariants. In boundary-based invariants, only the contour information of the shape is explored. The common boundary-based invariants include chain code and Fourier descriptors. Boundary-based invariants explore only the contour information; they cannot capture the interior content of the shape. On the other hand, these methods cannot deal with disjoint shapes where single closed boundary may not be available; therefore, they have limited applications. In region-based invariants, all the pixels within a shape are taken into account to obtain the mathematical representation. The most popular region-based methods include various moment-based invariants is Hu's Three Moment invariants.

2.3 Moment invariants

An essential issue in the field of pattern analysis is the recognition of objects and characters regardless of their position, size and orientation as illustrated in figure 1. The idea of using moments in shape recognition gained prominence when Hu (1962), derived a set of invariants using algebraic invariants [5][6][7]. Two-dimensional moments of a digitally sampled M X M image that has gray function f(x,y) where (x,y)=0…..M-1 is given by

$$m_{pq} = \sum_{x=0}^{x=M-1} \sum_{y=0}^{y=M-1} (x)^{p} . (y)^{q} f(x,y) \qquad (2)$$

Where p,q=0,1,2,…

The moment f(x, y) translated by an account (a, b)

$$m_{pq} = \sum_{x=0}^{x=M-1} \sum_{y=0}^{y=M-1} (x+a)^{p} . (y+b)^{q} f(x,y) \qquad (3)$$

The central moments m'_{pq} or μ_{pq} can be computed from eq.2 on substituting $a = -\bar{x}$ and $b = -\bar{y}$ as

$$\mu_{pq} = \sum_{x=0}^{x=M-1} \sum_{y=0}^{y=M-1} (x-\bar{x})^{p} . (y-\bar{y})^{q} f(x,y) \qquad (4)$$

Where $\bar{x} = \frac{m_{10}}{m_{00}}$ and $\bar{y} = \frac{m_{01}}{m_{00}}$

After applying scaling normalization the central moments changes as,

$$n_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}} \qquad (5)$$

Where $\gamma = \left[{p+q}/{2} \right] + 1$

Hu defines seven values, computed by normalizing central moments through order three, that are invariant to object scale, position, and orientation. In terms of the central moments, the three moments are considered and used this is work. The moments are as below,

$$M_1 = (n_{20} + n_{02}) \qquad (6)$$

$$M_1 = (n_{20} - n_{02})^2 + 4{n_{11}}^2 \qquad (7)$$

$$M_1 = (n_{30} - 3n_{12})^2 + (3n_{21} - n_{03})^2 \qquad (8)$$

Hu moment defined in equation above can be explained as followings [2];

M1- The sum of horizontal and vertical directed variance, more distributed towards horizontal and vertical axes and the values are enlarged.

M2 - The co-variance value of vertical and horizontal axes when the variance intensity of vertical axis and horizontal axis is similar.

M3 - the result emphasizing the values inclined to left or right and upper or lower axis.

The moments will be same for digit 6 and 9. To recognize these numbers, divide the image in two equal parts. Find the Mean for these parts. If Mean is more in first half, then the number is 9 else the number is 6.

2.4 Distance Measure

The distance between features of Training images and Test Image is measured by Euclidean distance measure given by

$$d(m,n) = \sqrt{\sum_{i=1}^{n} (m_i - n_i)^2} \qquad (9)$$

As the 4th, 5th, 6th and 7th moments result are not calculated, the mean of Hu 3 moments are calculated for each number for better match. The mean is consider here as the access point for recognition number. All moments for particular number is matched with the test image only if the mean value trained is matching with the mean values of the test image.

3. RESULT

The image is captured using a digital camera under controlled environment. The program executed on AMD Athlon 64 processor having 2.2GHz speed and 2GB RAM.

Table 1: Results of the numbers

Number\ HU	1	2	3	Mean
0	0.413	0.0245	0.0035	0.147
1	0.21	0.0036	0.0006	0.0714
2	0.32	0.0216	0.014	0.118
3	0.37	0.01	0.0009	0.126
4	0.38	0.017	0.0008	0.1326
5	0.35	0.01	0.0003	0.120
6	0.405	0.0383	0.0390	0.160
7	0.389	0.0785	0.0095	0.159
8	0.537	0.0155	0.0213	0.191
9	0.415	0.0387	0.0398	0.164

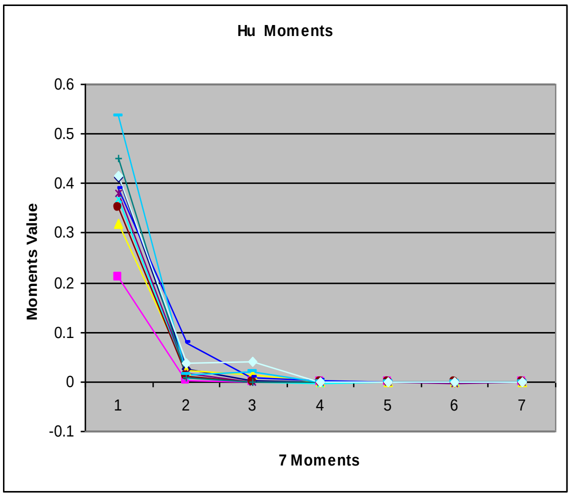

Fig. 2: Graphical representation of seven moments

The values of the computed moment invariants shown in table 1 and figure 2 are small. And also it has been found that most significant values concentrate at the first three moment invariants, and the values from forth to seventh moment invariants are close or equal to zero. The result of Hu seven moments for number 0 is shown below for different resolution.

Table 2: Comparison of three moments of the number 0 for different resolution

Resolution\Moments	1	2	3
512X512	0.413	0.0245	0.0035
256X256	0.4	0.0200	0.0030
32X32	0.34	0.0115	0.003

Table 3: Recognition Rate for different resolution

Resolution	Recognition Rate
512X512	100%
256X256	100%
32X32	72%

Table 3 lists the recognition result with images of different resolutions. It has been observed that for images with resolutions from 512×512 and 256×256, the recognition rate is 100%. It drops to 72% at resolution of 32×32 which verifies the spatial resolution invariance analysis shown in above table. It has been observed that, the moments for digit 6 and 9 are almost same but mean value is little bit different. At the time of packaging credit card or mobile SIM, there will maximum of 30 degree rotation.

4. CONCLUSION

This paper evaluated the performance of the Hu three moment's image invariants in ONR from the perspective of their invariance property to image transformations including scale, translation, rotation, different spatial resolutions. It has been observed and concluded that the Hu three moments can be used for Optical Number Recognition for images having resolution up to 256×256 under controlled environment

1. REFERENCES

[1] Graham Leedham, Chen Yan, Kalyan Takru, Joie Hadi Nata Tan and Li Mian, "Comparison of Some Thresholding Algorithms for Text/Background Segmentation in Difficult Document Images", Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003)

[2] M. K. Hu, "Visual pattern recognition by moment invariants," IRE Trans. Information Theory, Vol.IT-8, pp.179-187, February 1962.

[3] S. Pakchalakis and P. Lee, "Pattern recognition in gray level images using moment based invariant features," Image Processing and its Applications, IEE Conference Publication No.465, pp.245-249, 1999.

[4] R. Sivaramakrishna and N. S. Shashidhar, "Hu's moment invariants: how invariant are they under skew and perspective transformations," IEEE Conference on Communications, Power and Computing, pp.292-295, 1997.

[5] S. O. Belkasim, M. Shridhar, and M. Ahmadi, "Pattern recognition with moment invariants: a comparative study and new results," Pattern Recognition, vol. 24, pp. 1117–1138, 1991

[6] Muharrem Mercimek, Kayhan, "Real object recognition using moment invariants", S¯adhan¯a Vol. 30, Part 6, December 2005, pp. 765–775

[7] Dudani, Sahibsingh "Aircraft Identification by Moment Invariants", IEEE Conference, pp. 39-46, 2009

[8] M. Gruber and K. Y. Hsu, "Moment-based image mormalization with high noise-tolerance," Pattern Recognition, vol. 19, pp. 136–139, 1997.

[9] Jan Flusser, "Moment Invariants in Image Analysis", Proceedings of World academy of science engineering and technology, volume 11, 2006