Invariant Moments for 2D Object Categorization

Kateřina Nováková [1], *Jaromír Kukal* [2], *Dana Majerová* [2]

[1] CTU Prague, Faculty of Nuclear Sciences and Physical Engineering
[2] ICT Prague, Department of Computing and Control Engineering

Abstract: The recognition of 2D binary objects would be invariant to the translation, scaling and rotation. This problem can be solved by the TSR invariant moments. The calculation of them begins with general moments, over central and normalized moments and then seven invariant moments are produced. Thus every 2D binary object is represented as a pattern from $\mathbb{R}^7$ and after the second normalization as a pattern from $[0,1]^7$. It enables to use the artificial neural networks for the final TSR invariant categorization. The self-organization technique was used to learn Kohonen SOM artificial neural network.
Keywords: Invariant moments, 2D binary object, SOM, ANN, Matlab.

1 Introduction

The recognition of 2D objects is a traditional task for the artificial neural network (ANN). As in many other applications, the preprocessing and its small details have the significant role and influence to the quality of 2D object recognition. The paper is not oriented to the optical character recognition (OCR). That is why the invariance to the translation, scaling and rotation (TSR) of given object is necessary. But alas the TSR invariance is not a general property of ANN. The process of TSR invariance learning from examples would be too expensive in sense of ANN structure complexity and the time complexity of learning. The TSR invariant system based on invariant moments was introduced by Hu [3]. A very good illustrative example of TSR invariant recognition was given by Schalkoff [4]. The invariant moments varies in order in this example. Thus the preprocessing includes binary object separation, the calculation of general, central, normalized and invariant moments and the normalization of invariant moments. The traditional SOM with Kohonen learning was used for the final recognition.

2 Task description

Having a set of 2D binary images of various types, shifts, sizes and angles of rotation, there is necessary to categorize them using SOM network. We used ten classes of 2D objects. Every class is represented by sixteen objects in the training set (TRS) and using two objects in the testing (verification) set (TSS).

3 Invariant preprocessing

Any general binary image can be represented as a function $f : \mathbb{R}^2 \to \{0; 1\}$. Three types of moments are defined in the literature [2].

3.1 General moments

General moment of degree $p+q$ is defined as

$$m_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^p y^q f(x,y)dxdy$$

where $p, q \in \mathbb{N}_0$. But general moments are not invariant to translation, scaling and rotation. The first disadvantage is eliminated by using central moments.

3.2 Central moments

Let $x_t = m_{10}/m_{00}$, $y_t = m_{01}/m_{00}$. Then

$$\mu_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-x_t)^p (y-y_t)^q f(x,y)dxdy$$

is called central moment of $p+q$ order. The invariance to scaling is the main aim of moment standardization.

3.3 Standardized central moments

Standardized central moment of $p+q$ order is defined as

$$\nu_{pq} = \frac{\mu_{pq}}{(\mu_{00})^{\omega}},$$

where

$$\omega = \frac{p+q}{2} + 1.$$

The rotation invariance is the most difficult of all aims and can be solved individually.

3.4 TSR invariant moments

Hu [3] developed the system of seven TSR invariant moments:

$$\begin{aligned}
\varphi_1 &= \nu_{20} + \nu_{02} \\
\varphi_2 &= (\nu_{20} - \nu_{02})^2 + 4\nu_{11}^2 \\
\varphi_3 &= (\nu_{30} - 3\nu_{12})^2 + (3\nu_{21} - \nu_{03})^2 \\
\varphi_4 &= (\nu_{30} + \nu_{12})^2 + (\nu_{21} + \nu_{03})^2 \\
\varphi_5 &= (\nu_{30} - 3\nu_{12}) \cdot (\nu_{30} + \nu_{12}) \cdot [(\nu_{30} + \nu_{12})^2 - 3(\nu_{21} + \nu_{03})^2] + \\
&+ (3\nu_{21} - \nu_{03}) \cdot (\nu_{21} + \nu_{03}) \cdot [3(\nu_{30} + \nu_{12})^2 - (\nu_{21} + \nu_{03})^2] \\
\varphi_6 &= (\nu_{20} - \nu_{02}) \cdot [(\nu_{30} + \nu_{12})^2 - (\nu_{21} + \nu_{03})^2] + 4\nu_{11}(\nu_{30} + \nu_{12})(\nu_{21} + \nu_{03}) \\
\varphi_7 &= (3\nu_{21} - \nu_{03}) \cdot (\nu_{30} + \nu_{12}) \cdot [(\nu_{30} + \nu_{12})^2 - 3(\nu_{21} + \nu_{03})^2] - \\
&- (\nu_{30} - 3\nu_{12}) \cdot (\nu_{21} + \nu_{03}) \cdot [3(\nu_{30} + \nu_{12})^2 - (\nu_{21} + \nu_{03})^2]
\end{aligned}$$

4 Invariant pattern set

Every 2D binary object is than converted to the vector $\vec{\varphi} = (\varphi_1, \ldots, \varphi_7) \in \mathbb{R}^7$ which represents the object in 7^{th} dimensional vector space. Aplying the previous principle to $m \in \mathbb{N}$ objects we obtained the matrix representation of given pattern set

$$\mathbf{\Phi} = \begin{bmatrix} \Phi_{11}, & \ldots, & \Phi_{17} \\ \vdots & & \\ \Phi_{m1}, & \ldots, & \Phi_{m7} \end{bmatrix}$$

where Φ_{ij} is the value of j^{th} invariant for the i^{th} object. Both average values of invariant moments and their variances have various sizes. That is why the column normalization is necessary. Let $a_j = \min_k(\Phi_{kj})$, $b_j = \max_k(\Phi_{kj}) > a_j$. Than elements of normalized matrix $\mathbf{\Phi}^{\oplus}$ are

$$\Phi_{ij}^{\oplus} = \frac{\Phi_{ij} - a_j}{b_j - a_j} \in [0; 1]$$

Thus the vectors $\vec{a}, \vec{b} \in \mathbb{R}^7$ are parameters of signal processing which are obtained from training set TRS. The function called `tsr` was created in Matlab environment to realize the process from the set of binary objects to the matrix $\mathbf{\Phi}^{\oplus}$.

5 SOM as classificator

There is a time enough to finalize the process of classification using Kohonen [1] SOM ANN for the cluster forming and relationship visualization. Let $N \in \mathbb{N}$ be a number of SOM output neurons. Then the SOM realizes a function $SOM^* : \mathbb{R}^7 \to \{1, \ldots, N\}$. The input domain of SOM^* is the space of invariant moments and the output domain is the space of SOM neuron indices. The function SOM^* enables to decide, which of the output neurons is the winner using the 7-tuple of invariant moments. The standard hexagonal topology was used for $N = 48$. The output neurons are organized in six rows with eight elements inside. The SOM learning began with 1000 epochs for $\alpha = 0.5, R = 8$ and finished with 10000 epochs for $\alpha = 0.05, R = 3$. The standard software [7] was used for the realization of learning process. The quality of categorization can be measured using several simple criteria. Let C_1 be number of classes from training set which are represented by more than one neuron of SOM. Let C_2 be number of SOM neurons which are occupied by more than one class from the training set. Let C_3 be number of classes from testing set which are represented by more than one neuron of SOM. Let C_4 be number of SOM neurons which are occupied by more than one class from the testing set. Let C_5 be number of miss classified objects from the testing set (they are located in the neurons which corresponds with another class from the training set). Let C_6 be number of unidentified objects from the testing set (they are located in the neuron which is not occupied by any object from the training set). In our case the classification of 2D binary objects seems to be successful because of $C_1 = 4, C_2 = 1, C_3 = 0, C_4 = 1, C_5 = 0, C_6 = 0$. Because of C_2 and C_4 values we used the revisited SOM for similar objects where three most different objects were removed. The SOM learning (with the same map size) began with 1000 epochs for $\alpha = 0.5, R = 8$ and finished with 10000 epochs for $\alpha = 0.02, R = 5$. It comes to criteria values $C_1 = 3, C_2 = 0, C_3 = 0, C_4 = 0, C_5 = 0, C_6 = 0$.

6 Conclusion

The experiments prove the possibility of using Kohonen's SOM for the classification of grayscale objects transformed by translation, scaling and rotation. Seven features based on moments were used for description of objects, the projection of invariant vectors using PCA proved the object separability. The results are surprisingly optimistic. There are independent clusters of objects from given classes, which were produced by SOM network. Only two classes of ten were fusing. It is possible to eliminate this effect by using the revisited SOM for similar objects.

References

[1] Šnorek M.: Neuronové sítě a neuropočítače, Vydavatelství ČVUT, Praha, 2002

[2] Hlaváč V., Šonka M.: Počítačové vidění, Grada, Praha, 1992

[3] Hu M. K.: Pattern Recognition by Moment Invariance, Proceedings of the Institute of Radio Engineers, 49, 1428; 1961

[4] Schalkoff R. J.: Pattern Recognition, John Wiley & Sons, New York, 1992

[5] Sammon J. W.: A nonlinear mapping for structure analysis. IEEE Transactions and Computers, C-18(5), 1969

[6] Kohonen T.: Self-Organizing Maps, volume 30 of Springer Series in Information Sciences, Springer, Berlin, Heidelberg, 1995

[7] Kohonen T., Hynninen J., Kangas J., Laaksonen J.: The self-Organizing Map Program Package, volume A31 of Technical Report, Helsinki University of Technology, Laboratory of Computer and Information Science, Rakentajanaukio 2 C, SF-02150 Espoo, 1995, elektronická verze

Contacts

Kateřina Nováková
Czech Technical University in Prague
Faculty of Nuclear Sciences and Physical Engineering
Zikova 4, 166 35 Prague 6, Czech Republic
E-mail: katka.novakova@dc.fjfi.cvut.cz

Jaromír Kukal, Dana Majerová
Prague Institute of Chemical Technology
Faculty of Chemical Engineering
Department of Computing and Control Engineering
Technická 1905, 166 28 Prague 6, Czech Republic
Phone: +420-2-24354212, fax: +420-2-24355053
E-mails: {jaromir.kukal, dana.majerova}@vscht.cz