Proc. VIIth Digital Image Computing: Techniques and Applications, Sun C., Talbot H., Ourselin S. and Adriaansen T. (Eds.), 10-12 Dec. 2003, Sydney

Deformable Grid in Image Recognition

Yan Jikun[1], Xi Jianmin[1], and Zheng hui[2]

[1] The South-West Institute of Electronics & Telecommunication Technology
P.O. box 666, Chengdu, 610041, China
{yan_jk, xi_jm}@sina.com
[2] The University of Electronic Science &Technology of China
Chengdu, 610054, China

Abstract. Deformable template has been shown to significantly improve the performance of image recognition for difficult tasks such as character recognition, digit recognition, and trademark recognition etc. However, it is often time-consuming. Grid feature is a popular feature extraction scheme in image recognition. Here we propose deformable grid for image recognition. Unlike deformable template where deformation is applied to image, deformation is applied to grid in deformable grid. Because the number of grid is much less than that of image, our method is very timesaving comparing to deformable template. The approximate equality of deformable template and deformable gird is also shown. We tested our method in two image recognition experiments, namely, trademark recognition and off-line Chinese character recognition. We obtained improvement in recognition rate by 6.0% in first experiment, and 5.8% in second one.

1 Introduction

Selection of feature extraction scheme is probably the most important issue in achieving high performance in image recognition. Many feature extraction methods have been proposed in the literature. An important feature extraction method is based on grid[1,2]. A $n \times m$ grid is superimposed on image, and for each of the $n \times m$ zones, some feature is extracted, thus a feature vector of length $n \times m \times k$ (supposing the length of feature in each zone is *k*) is obtained. Examples of the features include average gray level, histogram of line segments from contour, count of black pixel[1,2,3] etc. We call this type of feature grid feature. In addition, some researchers observed that when near zone borders, small variations in the image could lead to large variation in the extracted feature. Therefore, they tried to compensate for this by using fuzzy borders.

Deformable template is an effective way to improve performance in many image recognition tasks, especially for such difficult ones as character recognition, digital recognition, and trademark recognition etc[4,5,6]. The key idea of deformable template lies in the process of reversing an input image back to one of its standard forms, which consists of applying a set of predefined inverse-deforming to the input image (See Fig.1). The inverse- deforming is independent of the input images and is expected to include the true deformation that actually makes the input image different from its standard pattern. ì Exhaustive searchingî is usually adopted if the way that input image differing from its standard pattern is unknown. Furthermore, a weighted-voting scheme should be employed to determine the most probable class of the input image. And the voting weight depends on the degree of deformation. That is, the larger deformation, the smaller weight and the less deformation, the larger weight. It is noted that deformable template is inherently not feature extraction method.

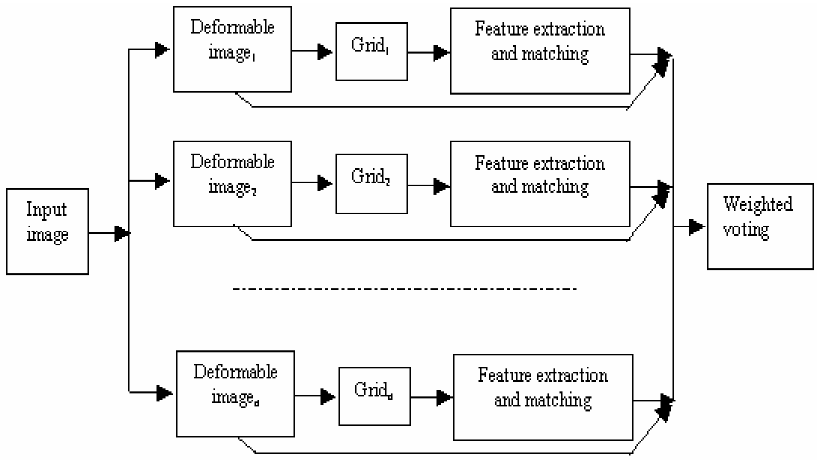


Fig. 1. Recognition scheme based on deformable template

Wile deformable template is effective in many tests, its critical shortcoming is time-consuming because of large quantities of operation on input image. For example, in ref.[5] 25 deformable images are created from an input image, thus lead to approximate tenfold more computation than recognition without deformable template. In this paper we propose deformable grid and investigate its application in trademark and off-line Chinese character recognition. Deformation is applied to grid superimposed on image but not to image itself. Because the number of grid is much less than that of image, our approach is very faster than deformable template. And because the grid feature extracted via deformable template and deformable grid is approximately equal, the two approaches possess approximately equal performance.

2 Deformable Grid

We first explain the key idea of deformable grid by an example, and then attempt to illuminate its effectiveness by two experiments. A Chinese character image named ì pingî is shown in the left part of Fig 2. Because of writing habit the character squeezes to the top. So we need to push downwards the image in order to obtain standard template of ì pingî, then certain grid would be superimposed on the character and feature could be extracted based on the grid where deformable template is employed. But another way would produce approximate feature: create grid firstly, and then push the gird upwards, extract feature via the deformable grid. Although we have taken the pushing operation as example, it is clear that the argumentation is applicable to other deformation such as rotation, slanting etc.

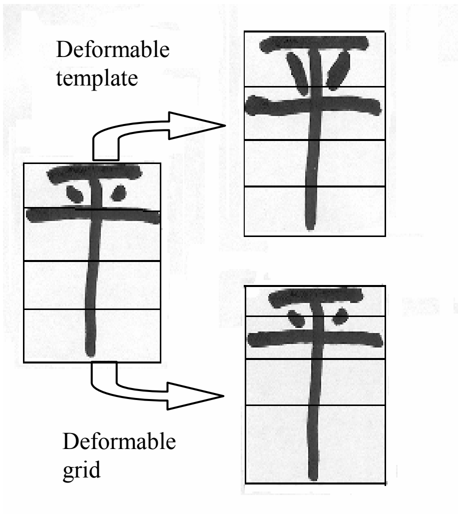


Fig. 2. Comparison between deformable template and deformable grid

We will further reveal the inherent relationship between deformable grid and deformable template. Suppose P the image to be recognized,

$$P = \begin{pmatrix} p_{11}, & p_{12}, & \cdots, & p_{1N} \\ p_{21}, & p_{22}, & \cdots, & p_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ p_{M1}, & p_{M2}, & \cdots, & p_{MN} \end{pmatrix} \qquad (1)$$

Suppose C the function to cut image to grid,

$$\mathrm{C}(P) = \begin{pmatrix} P_{11}, & P_{12}, & \cdots, & P_{1n} \\ P_{21}, & P_{22}, & \cdots, & P_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ P_{m1}, & P_{m2}, & \cdots, & P_{mn} \end{pmatrix} \qquad (2)$$

where $P_{ij} (1 \le i \le m, 1 \le j \le n)$ is the submatrix consisted of those elements locating in same zone of grid.

Suppose E the function to extract feature whose parameter is $C(P)$. E arrays those sub-features from P_{ij} s to compose feature F, that is:

$$F = E[C(P)] \tag{3}$$

Suppose D the function to create deformation from image whose parameter is sub-matrix P_{ij} ,

$$P_{ij}' = D(P_{ij}) \tag{4}$$

where P_{ij}' is the deformed image.

The process of extracting feature from deformed image could be described as (priority of operator is from right to left):

$$E \bullet C(P') = E \bullet C \bullet D(P) \tag{5}$$

As observed from Fig. 2, the two functions C and D satisfy exchange law when deformation is small. That is:

$$E \bullet C \bullet D(P) \approx E \bullet D \bullet C(P) \tag{6}$$

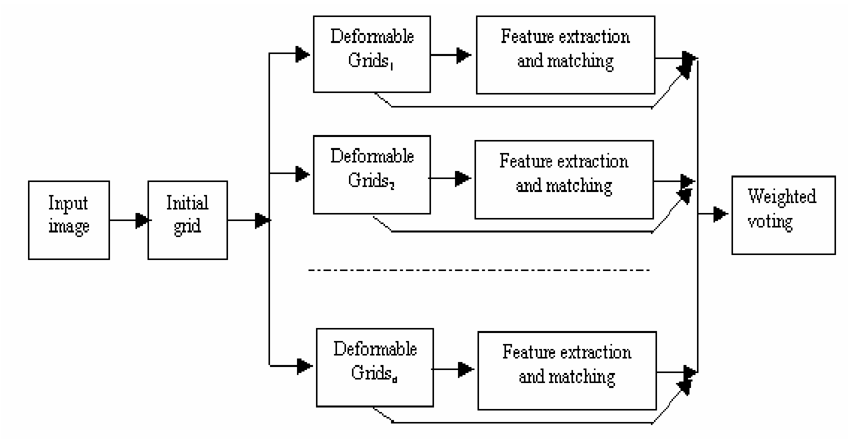


Fig. 3. Recognition scheme based on deformable grid

The right of equation (6) means cut image to grid firstly, and then produce deformation of the grid, finally extract feature via deformed grid. It is clear that the right side stands for deformable grid and the left side deformable template. The approach of deformable grid reduces computation greatly by exploiting the fact that C and D satisfy exchange law when deformation is small. Fig.3 shows how deformable grid works in practice: first create initial grid for input image, secondly deform initial grid to produce some deformable grids, thirdly extract grid features based on every deformable grids, and finally obtain result by weighted voting.

In addition, a control factor *a* is needed to control the degree of deformation, that is:

$$P' = D(P, a) \tag{7}$$

Suppose that the legal space of *a* is *A*. In order to catch the true deformation type and degree *a*, we can try to cover the overall space of control factor by sampling *a* with small sampling interval and obtaining those finitely number of values, i.e. $a_1, \quad a_2, \quad \cdots, \quad a_n$.

In order to avoid that an initial grid is deformed into certain degree so that we extract a completely different feature without being penalized, a confidential factor is defined for each a_i. The larger deformation, the smaller w_i, and contrariwise the less deformation, the larger w_i. Clearly w_i is the weight of weighted-voting scheme.

We perform two experiments, namely, trademark recognition and off-line handwritten Chinese character recognition to demonstrate the application of deformable grid.

3 Trademark recognition

Trademark recognition is an important research issue since the increasing number of registered trademarks puts a heavy burden on manual examiners. Hence, it is imperative to develop trademark recognition system.

The grid feature we used in experiment represents local direction of trademark contour. We think that, unlike characters, contour of trademark fails to represent all information. Therefore, inner black pixels (contrast to contour) are used, too.

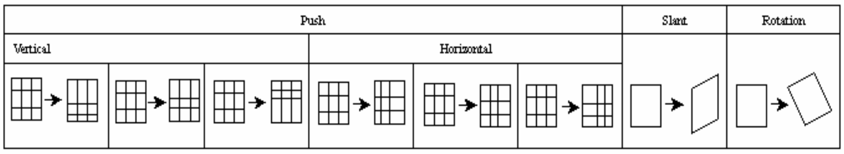
The initial grid is called gravity grid obtained by partitioning the trademark image hierarchically in both horizontal and vertical directions in order to balance the number of pixels inside up and down part (for horizontal partition) or inside left and right part (for vertical partition). The feature of image is 1280-dimension vector due to the fact that the gravity grid is 16×16 and a 5-dimension sub-vector is extracted from each zone. Euclid distance is employed to measure the similarity between two features.

The following four types characterize the deformation of trademark image[5] (See table 1):

(1) Pushing: the pixels is pushed to somewhere, i.e. left, right, or middle in vertical direction and top, down, or middle in horizontal direction;
(2) Slant: left part lifts and right part descends or left part descends and right part lifts
(3) Rotation of image.

To ìcapturingî those deformations, we adopt eight types deformations depicted in table 1.

Table 1. Eight types of deformation



3.1 Pushing deformation

Since pushing deformations in horizontal direction is similar to those in vertical direction, we take vertical direction as an example to explain our method. Three-order polynomial function $f(x)=ax^3+bx^2+cx+d(0>x>1)$ is used as deformation function with constraints below:

$$f(0)=0 \;\; f(1)=1 \qquad \textbf{(8)}$$

$f'(x)=3ax^2+2bx+c$ $f''(x)=6ax+2b$ The inflexion of $f(x)$ is $x_c= -b/3a$.

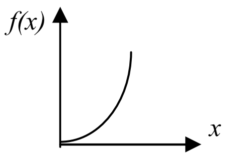


Fig. 4. Deformation function for pushing downward

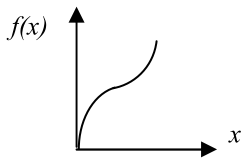


Fig. 5. Deformation function for pushing to middle part

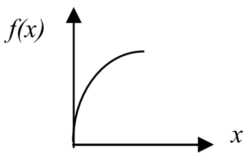


Fig. 6. Deformation function for pushing upward

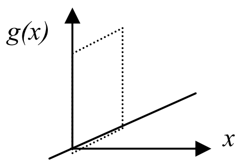


Fig. 7. Deformation function for slanting

Let $x_c =0,$ taking in count (8) we obtain $f(x)=a(x^3-x)+x$ with its figure depicted in Fig.4. We are able to push x downwards via $f(x)$ and the closer of x to 0 , the larger of deformation.

Let $x_c =0.5,$ we obtain $f(x)=a(x^3-1.5x^2+0.5x)+x$ with its figure depicted in Fig.5. We are able to push x to middle via $f(x)$ and the closer of x to 0.5, the larger of deformation.

Let $x_c = 1$, we obtain $a(x^3-3x^2+2x)+x$ with its figure depicted in Fig.6. We are able to push x upwards via $f(x)$ and the closer of x to 1, the larger of deformation.

3.2 Slant deformation

The function of $f(x)=g(x)+x$ is used where $g(x)=ax$ and its figure is depicted in Fig.7. We can lift right part of grid on condition that $a>0$, and descend on condition that $a<0$.

3.3 Rotation deformation

When employing function $x'=xCosa-ySina$, $y'=xSina+yCosa$ we are able to rotate grid clockwise on condition that $a<0$, or rotate grid anti-clockwise on condition that $a>0$.

Four control factors are used for each of those eight types of deformation showed in table 1. Therefore, thirty-three dissimilar grids could be obtained based via gravity grid (including the gravity grid), thus lead to thirty-three features.

3.4 Experiment evaluation of trademark recognition

A collection of 106 trademarks is provided by University of Maryland (ftp://ftp.cfar.umd.edu/pub/documents/contrib/databases/UMDlogo_database.tar). Beside those trademarks we have collected others from books, newspapers and faxs etc to construct our own collection that consists of 406 trademarks (106 from UMD). Those are treated as training samples. And test samples are produced by manually bedaubing, adding noise, pushing, slanting, or rotating each trademark in training samples to create four deformable images. Although unable to cover all possible types of deformation, we attempt to account for a wide variety of deformation. We perform both deformable template-based and deformable grid-based recognition. The two recognition methods employ the same feature extraction scheme, as well as deformation function and control factor. The 16×16 gravity grid is also used in both experiments. We report our result in table 2, where ìwithoutî (deformable template or deformable grid) means the recognition process with neither one. The usage of deformable template improves recognition rate by 7.5% and deformable grid 6.0%.

Table 2. Experiment result of trademark recognition

Scheme	without	with	Improvement
Deformable template based	85.0%	92.5%	7.5%
Deformable grid based	85.0%	91.0%	6.0%

4 Off-line Chinese character recognition

Another experiment used to assess deformable grid scheme is off-line Chinese character recognition, which is a challenging image recognition task because of large quantities of different writing habits and styles.

We use the grid feature of ìfuzzy line elementî proposed in [7] where gravity grid is 8×8. The experiment based on deformable template is performed as well.

Institute of Automation, Chinese Academy of Sciences provides handwritten Chinese character database including 3755 frequently used characters with 100 samples per character. Fifty samples of each character are used as training samples and remnant as test samples. The experiment result is reported in table 3. Deformable template gains improvement in recognition rate by 6.1%, and deformable grid 5.8%.

Table 3. Experiment result of off-line Chinese character recognition

Scheme	without	with	Improvement
Deformable template based	84.3%	90.4%	6.1%
Deformable grid based	84.3%	90.1%	5.8%

5 Conclusion

A new approach named deformable grid is proposed. We attempt to compensate the shortcoming of deformable template by the approach of deformable grid. It has been observed that these two methods are approximately equal on condition that the deformation is little. Furthermore, two experiments, namely, trademark recognition and off-line Chinese character recognition, are performed respectively to assess the performance of deformable grid. In both experiments, the performance of deformable template is slightly better than that of deformable grid. The result is comprehensive considering the approximate equality between these two methods and larger deviation would occur when meeting large deformation template. In addition, deformable template in trademark recognition works better than in off-line Chinese character recognition. So does deformable grid. It should be noticed that trademarks have less deformation than handwritten character, so same deformation function and control factor are able to ìcaptureî more types of deformation. This means that deformation model must be analyzed carefully in order to devise the most suitable deformation function and control factor. Although deformable template shows slightly better performance than deformable grid in both experiments, our method is competitive because it cut computation to several percent of what is needed by deformable template. And the more types of deformation are adopted, the more time could be saved. On the other hand, our method consistently improves the recognition rate, be trademark recognition or off-line Chinese character recognition, and by using different grid feature types. Therefore, we believe deformable grid can be adopted to improve other grid-feature-based recognition. But it should be stressed out that deformable grid works merely for grid feature-based scheme, and will fail when other popular feature is used, such as invariant moment, Fourier descriptor, K-L transformation etc. Clearly this demand constrains the application of deformable grid in some tasks.

References

1. Qivind D.Trier, Anil K.Jain, orfinn Taxt, :Feature Extraction Methods for Character Recognitonó A Survey, Pattern Recognition, 1996,29(4) 641-662
2. Dengsheng Zhang, Guojun Lu, In: A Comparative Study of Three Region Shape Descriptorsî, Digital Image Computing Techniques and Applications(DICTA2002), Melbourne, Australia, 2002,21-22
3. Ji-Kun Yan, Hui Zheng, Jian-Min Xi: Logo Recognition in Low Quality Document Imagesî, Proceedings of International Conference on Intelligent Information Technology (ICIIT2002), Beijing, China, 2002,185-189
4. Anil K.Jain ,Douglas Zongker, :Representation and Recognition of Handwritten digits using Deformable Templates ,IEEE Tran. On Pattern Analysis and Machine Intelligence, 19(12),1997
5. Thien M.Ha, Horst Bunke, Off-Line, Handwritten Numeral Recognition by Perturbation Method, IEEE Tran. On Pattern Analysis and Machine Intelligence,19(5),1997,535-539
6. Anil K.Jain, Aditya Vailaya,: Shape-based Retrieval: A Case Study with Trademark Image Databases, Pattern Recognition, 31(9), 1998,1369-1390
7. Ji-Kun Yan, Hui Zheng, Zhen Zhang,: Off-line Handwritten Chinese Character Recognition Based on Grid Distortion, Journal of Computer Research and Development(in Chinese),38(Suppl.),2001,127-131