한국컴퓨터정보학회 동계학술대회 논문집 제19권 제1호 (2011. 1)

A Matching Strategy to Recognize Occluded Number

Thi Thuong Pham○, Hyung-Il Choi**, Gye-Young Kim*
○*Dept. of Computing, Soongsil University
**Dept. of Media, Soongsil University

폐색된 숫자를 인식하는 매칭 방법

Thi Thuong Pham○, 최형일**, 김계영*
○*숭실대학교 컴퓨터학과
**숭실대학교 미디어학과
e-mail: thuongmm@yahoo.com, hic@ssu.ac.kr, gykim11@ssu.ac.kr

● 요 약 ●

This paper proposes a method of occluded number recognition by matching interest points. Interest points of input pattern are found via SURF features extracting and matched to interest points of clusters in database following three steps: SURF matching, coordinate matching and SURF matching on coordinate matched points. Then the satisfied interest points are counted to compute matching rate of each cluster. The input pattern will be assigned to cluster having highest matching rate. We have experimented our method to different numerical fonts and got encouraging results.

키워드: Interest point, Matching

I. Introduction

Nowadays number recognition is a common task in digital technology. A lot of endeavored researches have been done in literature and got encouraging results. However, number recognition in general still gets into difficulty because of occlusion, illumination... that they can effect to performance of recognizing systems.

Lanlan [2] addressed the text occlusion problem by restoring the occluded text before applying OCR engine. The text is visible watermark characters which are occluded by foreground contents. This method fills the occluded region patch-by-patch in the order decided by the complexity factor inside each patch. The proposed restoration technique produces qualitative restoration resulting in improving of recognition performance. In such a recognizing systems, restoring step always requires additional computation cost and effects to whole system speed.

This paper proposes a two-phase method to recognize occluded digit without restoring step (see Fig. 1). Both off-line training phase and recognizing phase extract SURF features to find out the interest points. These interest points are SURF matched to find alike points called origins that are used to implement coordinate matching. Number of coordinate and SURF matched points of each cluster is counted to calculate matching rate. The cluster having highest matching rate will be assigned to the input pattern.

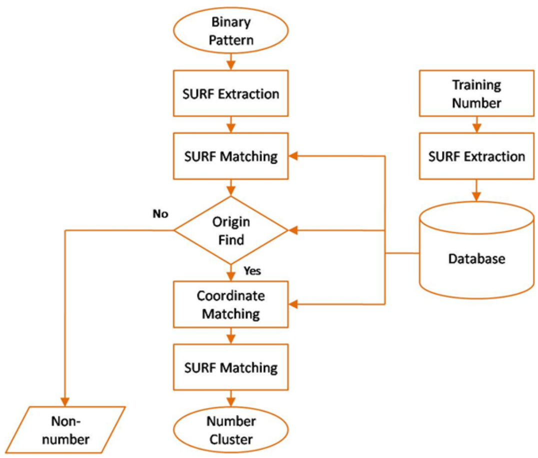

Fig. 1. Flowchart of proposed method

This method takes full advantage of features which have already extracted without requiring any restoring step in advance.

II. SURF Features Extraction

The progress of SURF features [1] extraction can be divided into two main steps: Interest point detection and Interest point description. The interest point detection step detects interest points at distinctive location in the image. The interest point description step describes local character of each interest point and represents it by a feature vector.

1. Interest Point Detection

For fast detection of interest points, Gaussian second order derivatives are approximated by box filters. The Hessian matrix determinant is calculated and interpolated in image space to find local maximum. The detected local maximum position refer to interest point location.

2. Interest Description

A circular window is constructed around every detected interest point and orientation is estimated using Haar Wavelet responses to have invariance to rotation. SURF descriptions are obtained by taking a rectangular window around every detected interest point in the direction of orientation. The windows are split into 4×4 sub regions. For each sub region Haar wavelet responses are extracted for equally spaced sample points. Finally the wavelet response in horizontal (dx) and vertical (dy) direction are summed up for each sub region. The absolute values of wavelet responses (|dx|and|dy|) are summed up to find the polarity of image intensity changes. Hence feature vector for each interest point is given by

$$fv = \left(\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y|\right) \tag{1}$$

3. SURF Features Extraction

We uses rotation variant version of SURF (U- SURF) in our system because rotation invariant is not necessary for recognizing digit. U-SURF is faster to compute and can increase distinctiveness, while maintain a robustness to rotation of about +/- 15 degree. Every interest points are described by its coordinate and a 64-dimension vector. In our experience, 64-dimension vector is distinctive enough among different interest points.

Binary image is used to extract SURF features in both training phase and recognizing phase. The number characters on training image are not fixed on size but fixed on Arial font. The training images are collected and classified under clusters from 0 to 9. We extract SURF features from training images of each numerical character and store the features produced on the image which amount of features is the most. This will avoid mismatching caused by interest points in input image might not appear in training image and vice versa.

III. Interest Points Matching

1. SURF Matching

SURF matching in our system is implemented by Nearest Neighbor Search in which the distance between two SURF vectors of two interest points is calculated. Two points are considered SURF matching if their distance smaller than a ThresSURFMatch value which experimentally chosen to be 0.3. SURF matching results in one interest point in input image maybe match with several interest points in training image (Fig. 2)

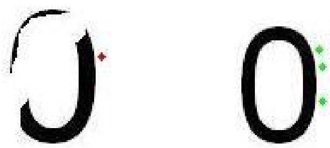

Fig. 2. An example of SURF matching result: input image (left) and training image (right)

2. Coordinate Matching

Coordinate matching refers to training interest point's coordinate is close enough to input interest point's coordinate. If P(X_P, Y_P) is an interest point in training image, Q (X_Q, Y_Q) is an interest point in input image. Q and P is coordinate matching if their coordinate satisfy

$$\begin{cases} X_Q - X_P < ThresCoorMatch \\ Y_Q - Y_P < ThresCoorMatch \end{cases} \tag{2}$$

2.1 Origin Finding

Origin of input image is an interest point which has unique corresponding matching point in training image. Correspondingly, the unique matching point in training image is called origin of training image. The matching here is in terms of both SURF features and coordinates.

In order to find origin, we choose an interest point named A in input image that has at least one SURF matched point in training image. We also choose the second interest point named B in input image that separates from A and has at least one SURF matched point in training image. Two interest points in input image produce one line that its direction is unique. In training image we consider the direction of every line between a matched point of A and a matched point of B and take out the line that its direction is similar to direction of AB in input image (see Fig. 3).

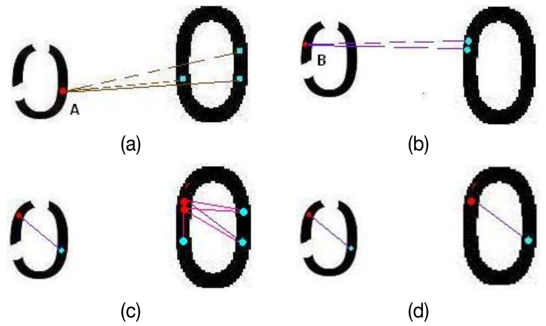

Fig. 3. Steps of origin finding (a), (b), (c) and (d)

In our system, two line directions are similar if $\tan(\theta_1) - \tan(\theta_2) < ThresTangent$ where $\tan(\theta_1) = \frac{d_y}{d_x}$, $\tan(\theta_2) = \frac{D_y}{D_x}$ (Fig. 4) and ThresTangent = 0.08.

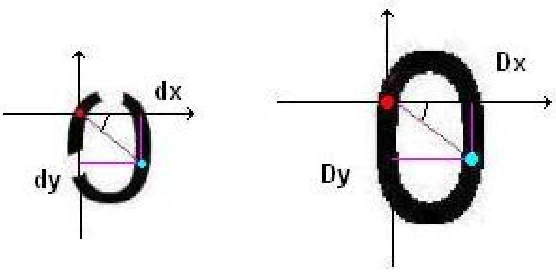

Fig. 4. Two lines passing origin of input image (left) and training image (right) having same direction

If two line directions are similar, either of two points will be chosen as origin of input image, and its SURF matching point is origin of training image. Additional, the proportion of training image dimension to input image dimension is also calculated called TrainInRate. TrainInRate lets us know how many times input pattern is scaled up or down comparing to training pattern. In Fig. 4 if the chosen origin is red pixel then TrainInRate is calculated as

$$TrainInRate = \frac{D_x}{d_x} = \frac{D_y}{d_y} \tag{3}$$

where (Dx, Dy) and (dx, dy) are the translated coordinates of the non-origin point in training image and input image, respectively.

If we correctly find out origin, the TrainInRate is unique for each pair of input and training image. Therefore TrainInRate can be used for identify position of interest points in training image when its position in input image was known.

2.2 Coordinate Matching

We translate all interest points' coordinate to new coordinate system where origin is a chosen interest point using formular

$$\begin{cases} x = X - X_0 \\ y = Y_0 - Y \end{cases} \tag{4}$$

Where (X_o, Y_o) is the new origin's coordinate in old coordinate system.

Each interest point at (dx, dy) in input image, its corresponding coordinate (d'x, d'y) in training image will be

$$\begin{cases} d'x = dx \times TrainInRate \\ d'y = dy \times TrainInRate \end{cases} \tag{5}$$

We search training image for interest points (Dx, Dy) satisfy condition

$$\begin{cases} Dx - d'x < ThresCoorMatch \\ Dy - d'y < ThresCoorMatch \end{cases} \tag{6}$$

where ThresCoorMatch = 2. Any point satisfied (6) is coordinate matching with interest point (dx, dy) in input image.

3. SURF Matching on Coordinate Matched Points

We need to implement SURF matching again to eliminate the coordinate matched points that are not SURF matching. In fact, we do not need any additional computation during this step except for choosing the pair of points such that its SURF distance smaller than ThresSURFMatch. Since all the SURF distance between two points have been already calculated in the section III.1.

All interest points in each cluster that both SURF matching and coordinate matching will be counted to calculate matching rate. Matching rate of each cluster is the proportion of number of matched interest points to the amount of cluster's interest points. The input pattern will be assigned to cluster having highest matching rate.

Ⅳ. Results

We prepared a set of testing images including 200 images of different sizes and fully comprise digits from 0 to 9. The fonts contain Symbol, Calibri, Franklin Gothic Medium, Arial and In-Out gate number extracted in subway station. Occluded numbers are produced by erasing accidentally any parts of unoccluded number patterns. Sample of testing images are shown in Fig. 5.

Fig. 5. Samples of testing occluded digits

Our proposed method gets the recognition accuracy of 82%. This is an encouraging result because the occluded patterns in our testing sets diverse in sizes, fonts and types of occlusion.

Ⅴ. Conclusions

This paper presents a method of occluded number recognition without restoration in advance. The core module of recognizing process is implementation of coordinate matching after SURF matching on interest points extracted from input image. Experimental results with different font-size numbers demonstrate the encouraging result of our method with the recognition rate of 82%. Based on these results, we plan to expand the test image font and speed up system for a real time application.

References

[1] Herbert Bay, Andreas Ess, Tinne Tuylelaars, Luc Van Gool, “Speed-Up Robust Features (SURF)”, Computer Vision and Image Understading 110, 346~359, 2008.

[2] Lanlan Chang, Jun Sun, Misako Suwa, Hiroaki Takebe, Yuan He, Satoshi Naoi, “Occluded Text Restoration and recognition”, Proceeding of the 9th IAPR International Workshop on Document Analysis Systems, 2010.