CITY UNIVERSITY OF HONG KONG
香港城市大學

Approaches to Symbol Recognition and Spotting
符號識別和定位算法

Submitted to
Department of Computer Science
電腦科學系
in Partial Fulfillment of the Requirements
for the degree of Doctor of Philosophy
哲學博士學位

by

ZHANG Wan
張琬

September 2008
二零零八年九月
Abstract of thesis entitled:

Approaches to Symbol Recognition and Spotting
Submitted by ZHANG Wan
for the degree of Doctor of Philosophy
at City University of Hong Kong in September 2008

The problem of symbol recognition and symbol spotting have been addressed in this thesis. First of all we reviewed the literature of the symbol recognition problem from three aspects: 1) how to represent a symbol, 2) how to matching the symbols based on their representations, and 3) how to solve the symbol alignment problem during symbol matching. As the main contributions, we then propose two different approaches for symbol recognition and symbol spotting.

The first contribution is a statistical approach to symbol recognition. In the approach, a symbol is measured and represented in 2-dimensional kernel densities around the sampling points on its skeleton. The Kullback-Leibler (KL) divergence is then exploited to measure the similarity between the densities of two symbols. Compared to the results of the same test sets reported in (Yang, 2005), the best overall results in the 2003 GREC Contest, our method achieves the better overall performance except in two subsets with combination of deformation and degradation. On the hand-drawn drawings generated with the 50 models also from the 2003 GREC Contest, our recognition results are 96% and 94% for the two generated sets respectively. It is
much better than those of Su’s method, which are 72% and 74% correspondingly. We propose two methods to eliminate the rotation effect. One is to adjust the rotation angel by minimizing the KL divergence between the test symbol and the model symbol with the gradient-based method. The other is to exploit the independent component analysis (ICA) technique, which considers the higher-order statistical information of the data and whose outputs are generally invariant to any invertible linear transformation. The first method gives quite accurate results, as shown in the experiment section, and the second method reduces the computation greatly with the performance degraded just around 5%. By introducing ICA, it only takes about 80ms to match a pair of symbols, which is only $\frac{1}{30}$ of that for the gradient-based algorithm.

The second approach is a hybrid approach to symbol recognition or spotting of symbols in vectorial forms. We calculate all primitive-pair relationships in a symbol and generate the signature of the symbol with those relationships. Finally, the symbol is represented as a feature set containing all the primitive-pair relationships. Matching between two symbols/drawings is then reduced to the problem of matching two feature sets. We apply the approach on the GREC2003 test sets with symbols in vectorial forms. Since all the best matching models for the test symbols are correctly retrieved, the recognition accuracies are all 100%. Note that, the retrieved most similar models for a test symbol may not be unique. If the recognition standard is that the recognition result must be unique, the accuracy of the approach drops. However, all of the recognition accuracy are still above 92%. For this database, at most two best matching models are retrieved for each test symbol. Under such situation, the approach can be applied to filter out
the impossible similar symbols and further accurate recognition can be performed based on those results. The approach preserves high recognition accuracy, while speeding up the matching procedure. Since all the models can be processed in advance, the important factor to reveal the efficiency of the approach in the symbol recognition task is the average time to create a signature for a test symbol and the time for matching a pair of symbol. Averagely, a signature of a symbol in the proposed approach can be created in 0.025 seconds. Moreover, matching a pair of symbols only costs 0.01 seconds. Compared with the time cost used in the proposed statistical method, which costs about 2.4 seconds for matching a pair of symbols with rotation estimation, the proposed approach improves the matching speed by two hundred times.

Moreover, we generate a general framework for the performance evaluation of certain recognition approaches. All the preprocessing steps are the same for the to-be-evaluated approaches. Both symbol databases and shape databases are applied to do evaluation. In symbol databases, a symbol is unique and people target to find the exact matching symbol for the testing symbol. However, shape databases usually include shapes in various categories or classes and people aim at finding the characteristics of a class and retrieving the top similar shapes for the testing symbol. Different traditional measures, e.g. recognition accuracy and top matching rates, are applied to evaluate the performance of the methods. Generally, the symbol recognition approaches perform well on the symbol databases and have strong competitive performance in retrieving the top similar shapes. Furthermore, new measures, namely homogeneity and separability, are proposed to explore more characteristics of the methods so that they can be better
understood. High homogeneity means that a descriptor can represent the symbols in the same class in a highly similar way. High separability means that a descriptor can distinct symbols in different classes well. These two measures are expected to evaluate how well distributed are the symbols in the space of representation provided by the symbol descriptor. Experimental results show that the proposed statistical method in Chapter 3 describes a relatively sparse inner-class structure, that if we regard the representations of the shapes in one class as points in one class, sparse inner-class structure means that their distribution is relatively sparse. It is also an explanation to the question why the proposed statistical method is robust to noisy and distorted symbols.
## Contents

Abstract  i

Acknowledgement  v

1 Introduction  1
  1.1 Motivation of This Work  1
  1.2 Contribution of This Work  7
  1.3 Organization of the Thesis  11

2 Literature Review  12
  2.1 Solutions to Symbol Alignment  12
  2.2 Symbol Representation Forms  16
    2.2.1 Structural Representation  16
    2.2.2 Statistical Representation  20
    2.2.3 Hybrid Representation  24
  2.3 Symbol Matching  25
  2.4 Summary  28

3 Kernel Density for Symbol Matching  30
  3.1 Related Work in Other Fields  31
  3.2 Basic Introduction  32
  3.3 KL Divergence and Maximum Likelihood  33
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4 Kernel Density Estimator</td>
<td>34</td>
</tr>
<tr>
<td>3.5 Preprocessing</td>
<td>36</td>
</tr>
<tr>
<td>3.5.1 Denoising</td>
<td>36</td>
</tr>
<tr>
<td>3.5.2 Vectorization</td>
<td>40</td>
</tr>
<tr>
<td>3.6 Solutions to the Rotational Alignment Problem</td>
<td>43</td>
</tr>
<tr>
<td>3.6.1 Angle Searching Algorithm</td>
<td>44</td>
</tr>
<tr>
<td>3.6.2 Matching Axes’ Selection by Independent Component Analysis</td>
<td>45</td>
</tr>
<tr>
<td>3.7 Summary</td>
<td>47</td>
</tr>
<tr>
<td>4 Vectorial Signature</td>
<td>49</td>
</tr>
<tr>
<td>4.1 Basic Framework</td>
<td>50</td>
</tr>
<tr>
<td>4.2 Brief Description</td>
<td>51</td>
</tr>
<tr>
<td>4.3 Relationship between Line Segments</td>
<td>53</td>
</tr>
<tr>
<td>4.4 Relationship between Line Segment and Arc</td>
<td>54</td>
</tr>
<tr>
<td>4.5 Relationship between Line Segment and Circle</td>
<td>56</td>
</tr>
<tr>
<td>4.6 Relationship between Arcs</td>
<td>57</td>
</tr>
<tr>
<td>4.7 Signature Generation and Optimization</td>
<td>57</td>
</tr>
<tr>
<td>4.8 Similarity Evaluation between Two Segmented Symbols</td>
<td>59</td>
</tr>
<tr>
<td>4.9 Symbol Spotting in a Large Image</td>
<td>61</td>
</tr>
<tr>
<td>4.10 Summary</td>
<td>62</td>
</tr>
<tr>
<td>5 Experiments</td>
<td>64</td>
</tr>
<tr>
<td>5.1 Performance of the Proposed KDE Approach</td>
<td>65</td>
</tr>
<tr>
<td>5.1.1 Symbol Databases</td>
<td>66</td>
</tr>
<tr>
<td>5.1.2 Sharvit Shape Database</td>
<td>76</td>
</tr>
<tr>
<td>5.1.3 Selected Shapes from Mpeg-7 Database</td>
<td>78</td>
</tr>
<tr>
<td>5.2 Performance of the Proposed Vectorial Signature</td>
<td>80</td>
</tr>
<tr>
<td>5.3 Summary</td>
<td>87</td>
</tr>
</tbody>
</table>
List of Figures

1.1 Framework of the symbol recognition/spotting task . . 3

2.1 Various transformations of symbol . . . . . . . . . . . . . . 13
2.2 A simple example of SRG . . . . . . . . . . . . . . . . . . 19
2.3 An example of SDT, reprinted from (Liu et al., 2004) . 19
2.4 An example of Network, reprinted from (Ah-Soon & Tombre, 2001) . . . . . . . . . . . . . . . . . . . . . 20

3.1 On the upper row from left to right: a symbol, its sampling points, its density; On the lower row there are the distorted symbol, the corresponding sampling points and density . . . . . . . . . . . . . . . . . . . . . 33
3.2 Example symbols in engineering drawing . . . . . 37
3.3 Comparison between original and denoised images . . 39
3.4 The flowchart of ANR . . . . . . . . . . . . . . . . . . . 41
3.5 The first row corresponds to the model symbol, and the second row is a corresponding hand-drawn symbol. The symbols on the first row are the model symbols before and after sampling, and the one after whitening; The second row gives the hand-drawn symbols. . . . . . . . 43
3.6 The axes produced by ICA of a symbol before and after rotation. The ICA outputs, or ICs, are obtained by projecting the data points onto these axes.

4.1 Angle feature to describe the relation of a line-line pair.

4.2 Angle feature to describe the relation of a line-arc pair; $O$ is the center of the arc, Point $s$, $e$, and $m$ are the start point, end point, and middle point of the arc, respectively.

4.3 Three sections to represent the space relationship of the line-arc pair.

4.4 Three relationships between a line segment and a circle.

4.5 Angle to describe the relation of arc-arc pair; $O_1$, $O_2$ are the centers of the arcs $\hat{AB}$, $\hat{CD}$, respectively.

5.1 Results of 20 and 50 models tests with rotation, scaling and combination under noise; On $x$-axis, $s$ means scaling, $r$ is rotation and 20/50 is the number of model symbols.

5.2 A regular symbol from the Grec’03 contest.

5.3 Noisy versions of the symbol in Fig 5.2 at different noise levels.

5.4 Average recognition results of test sets with noise at different levels.

5.5 Image pair; the Left image is the hand-drawn one and right image is the model.

5.6 Some unrecognizable samples in GREC 2005. The first two symbols belong to the set of degradation 5; The last two come from the set of degradation 6.

5.7 Nine Models Selected from SIID.
5.8 Retrieved similar shapes for several query shapes by KDE method ........................................ 79
5.9 Model pairs sharing the same signature in the database ......................................................... 82
5.10 Top 5 similar models retrieved for the query model; values under symbols are the similarities between the retrieved model and the query one ......................................................... 83
5.11 An engineering drawing example ................................................................. 84
5.12 Filtering results for searching symbol 'arrow' in an engineering drawing example. Notice, that the definition of the y axis in the result is different from the original drawing ......................................................... 85
5.13 An architectural drawing example ................................................................. 86
5.14 Another engineering drawing example before and after primitives spotting ...................... 87
6.1 Sample model symbols ................................................................. 92
6.2 Homogeneity and separability of the various descriptors on symbol database ...................... 93
6.3 Top-10 most similar shapes of several query shapes retrieved by SIHA method .................... 95
List of Tables

5.1 Accuracy on Test Sets with Rotation, Scaling and Combination of them (%) .......................... 66
5.2 Accuracy on Test Sets with Degradation (%) .................. 67
5.3 Accuracy on Test Sets with Deformation (%) .................. 67
5.4 Accuracy on Test Sets with Deformation and Degradation(%) 68
5.5 Accuracy on Test Sets with Degradation in GREC'05(%) 73
5.6 Accuracy on Test Sets with Degradation and Rotation Transformation in GREC'05(%) .................. 73
5.7 Accuracy on Test Sets with Degradation and Global Transformation in GREC'05(%) .................. 74
5.8 Accuracy on Test Sets with Degradation and Scaling Transformation in GREC'05(%) .................. 74
5.9 Accuracy on Test Sets with Global Transformations in GREC2003 with the algorithm using ICA(%) ........ 76
5.10 Recognition accuracy of the top 10 similar shapes retrieved from Sharvit data set (%) .................. 78
5.11 Recognition accuracy of the top 11 similar shapes retrieved from the 216 shapes (%) .................. 79
5.12 Accuracy on test sets with rotation, scaling, and combination of them (%); See the text for explanation ........ 81
5.13 Average candidate list size of test sets .................. 82
5.14 Original primitive set size and those after filtering the engineering drawing ................................... 85

6.1 Accuracy on selective test sets in GREC2003 (%) ................................................................. 94

6.2 Recognition accuracy of the top 10 similar shapes retrieved from Sharvit data set by several algorithms (%) 96

6.3 Homogeneity and separability summarization on Sharvit data set by several algorithms (%) ....................... 96

6.4 Recognition accuracy of the top 11 similar shapes retrieved from the 216 shapes(%) ...................................... 97

6.5 Homogeneity and separability on selected shapes from Mpeg-7 database ............................................. 98