

**Hybrid Segmentation on Slant & Skewed
Deformation Text in Natural Scene Images**

by

Fei XiaoLei

Master of Science in Software Engineering



2010



**Faculty of Science and Technology
University of Macau**

Hybrid Segmentation on Slant & Skewed Deformation Text in
Natural Scene Images

by

Fei Xiao-Lei

A thesis submitted in partial fulfillment of the
requirements for the degree of

Master of Science in Software Engineering

Faculty of Science and Technology
University of Macau

2010



Approved by _____
Supervisor

Date _____

In presenting this thesis in partial fulfillment of the requirements for a Master's degree at the University of Macau, I agree that the Library and the Faculty of Science and Technology shall make its copies freely available for inspection. However, reproduction of this thesis for any purposes or by any means shall not be allowed without my written permission. Authorization is sought by contacting the author at

Address: N/A

Telephone: +(853)-66657624

Fax: N/A

E-mail: shkirin001@gmail.com



Signature _____

Date _____

University of Macau

Abstract

Hybrid Segmentation on Slant & Skewed Deformation Text in
Natural Scene Images

by Fei XiaoLei

Thesis Supervisor: Prof. Dong Ming-Chui
Department of Electrical and Electronics Engineering

Master of Science in Software Engineering

The works presented in this dissertation originate from our years of research on vision-to-text machine translation, in which the prototyping system applied on camera-based images has been developed by our group. However, it is found in our practice that text detection in complex background scene image is far more challenging than expected. The biggest challenge comes from the fact that text line varies in size, grey, shape and color. Furthermore, traditional learning-based classification method on image block is unsuitable for text detection that embeds in complex background scene images. Consequently, the conventional grey-value based OCR (Optical Character Recognition) software has encountered insufficiency on recognizing the text located with colorful background in scene images.

While in this thesis, a general framework is proposed for text detection based on a novel hybrid text line detection paradigm. The text detection process can be divided into two stages. Firstly, image pixels are grouped into regions through Vector Quantization method so that candidate text lines are obtained via layout analysis of regions with same color label. Secondly, wavelet histogram features of the candidate text lines are extracted to discriminate text pattern and non-text pattern by employing an SVM (Support Vector Machine) classifier. Here the classified text pattern is testified using the OCR technique on mentioned text to see whether it has textural feature or not. Then in the refinement procedure, a homography operation is utilized

to project the distort text from image plane to the real world plane. Domain knowledge is used to find corresponding points between the text in image plane and in real world plane.

The performance of proposed text detection technique is evaluated by using an international benchmark database ICDAR 2003. As to the testing of text detection applied on scene images in the aforementioned database, the proposed method achieved the precision of 0.60 and recall of 0.63, which is superior to those published methods. Coming to the testing of time consuming, our method can not reach the best performance but still acceptable, which was affected by the two-stage hybrid detection strategy. In summary, the proposed method is robust against complex background in scene images so it will make great contribution to the vision-to-text machine translation research and application.

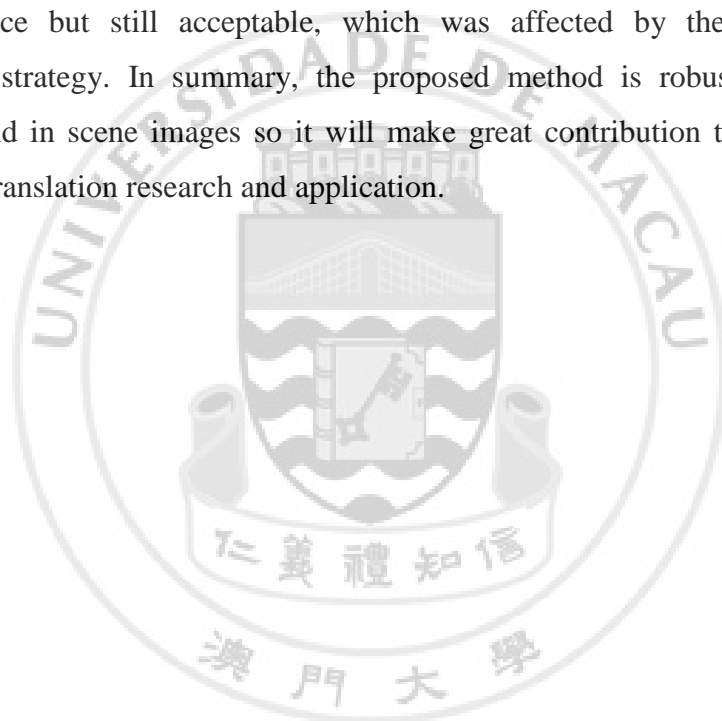


TABLE OF CONTENTS

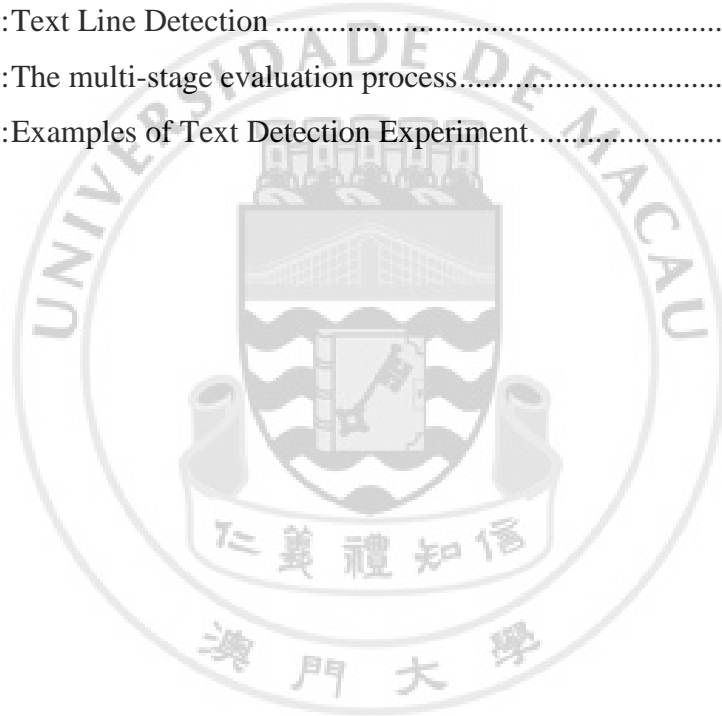
List of FIGURES	iii
List of TABLES	iv
List of ABBREVIATIONS	v
Chapter 1: Introduction	1
1.1 Research background of Text Detection.....	1
1.2 Text Pattern and Text Detection	2
1.3 Challenges of Text Detection.....	5
1.4 Current State of The Art	7
1.5 Bottleneck Problems and Research Goals	15
1.6 Statement of Contributions	16
Chapter 2: Image Processing and Image Segmentation	19
2.1 Analysis of Object Detection Technology.....	19
2.2 Image Wavelet Decomposition.....	20
2.3 Support Vector Machine Classification Method.....	24
2.4 Generalized Learning Vector Quantization	28
Chapter 3: Hybrid Segmentation of Text in Complex Background	30
3.1 Peer Group Filtering	31
3.2 Color Image Vector Quantization.....	34
3.3 Text Line Location.....	36
3.4 Text Segmentation from Complex Background	39
3.5 OCR Result Based Feedback	44
Chapter 4: Homography Based Text Recovery	45
4.1 Homography Based Recovery	45
4.2 Skewed Text Detect	47
4.3 Skewed Text Estimate.....	47
Chapter 5: Experiments	49
5.1 Preparation of Benchmark Database.....	49

5.2 Evaluation Method.....	50
5.3 Experiment Results and Discussion.....	52
Chapter 6: Conclusion and Future Works	56
6.1 Dissertation Summary.....	56
6.2 Future Work	58
BIBLIOGRAPHY	59
APPENDIX: Papers and Patent	67



LIST OF FIGURES

Figure 1-1:Scene Image Vision-to-Text Translation System	2
Figure 1-2:Comparison of Human Face Pattern & Text Line Pattern.....	3
Figure 1-3:Text Line Detection	4
Figure 1-4:Process of Text Line Detection.....	4
Figure 2-1:Time-Frequency analysis of WFT and Wavelet	21
Figure 3-1:Text Line Detection	31
Figure 5-1:The multi-stage evaluation process.....	50
Figure 5-2:Examples of Text Detection Experiment.....	54



LIST OF TABLES

Table 1-1: Summary of Text Detection Literature Showing Test Data, and Results.....	7
Table 5-1: Performace comparison with other algorithms on ICDAR2003	55
Table 5-2: Performace comparison with other algorithms on ICDAR2005	55

