

Comparative Study For Character Recognition Using Various Segmentation Techniques

BITS ZG628T: Dissertation

by

SHEEVANGI RELAN

2015HT12731

Dissertation work carried out at

SAMSUNG INDIA ELECTRONICS PVT LTD, Noida



**BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE
PILANI (RAJASTHAN)**

April 2017

Comparative Study For Character Recognition Using Various Segmentation Techniques

BITS ZG628T: Dissertation

by

SHEEVANGI RELAN

2015HT12731

Dissertation work carried out at

SAMSUNG INDIA ELECTRONICS PVT LTD, Noida

Submitted in partial fulfillment of M.Tech. Software Systems degree
programme

Under the Supervision of

PANKAJ GARG ,

LEAD ENGINEER

SAMSUNG INDIA ELECTRONICS PVT LTD, Noida



**BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE
PILANI (RAJASTHAN)**

April,

CERTIFICATE

This is to certify that the Dissertation entitled **Building an Auto-Filling Database using Machine Learning** and submitted by **Paras Santra** having ID-No:**2014HT13006** for the partial fulfillment of the requirements of M.Tech. Software Systems degree of BITS, embodies the bonafide work done by him/her under my supervision.

Signature of the Supervisor

Place :

Date :

MihiraKantaRath
Senior Development Manager,
Samsung Research Institute,
Noida

Birla Institute of Technology & Science, Pilani

Work-Integrated Learning Programmes Division

Second Semester 2016-2017

BITS ZG628T: Dissertation

ABSTRACT

BITS ID No. : **2015HT12731**

NAME OF THE STUDENT : **SHEEVANGI RELAN**

EMAIL ADDRESS : **sheevangirelan10@gmail.com**

STUDENT'S EMPLOYING ORGANIZATION & LOCATION : **Samsung India Electronics Pvt Ltd , Noida**

SUPERVISOR'S NAME : **PANKAJ GARG**

SUPERVISOR'S EMPLOYING ORGANIZATION & LOCATION : **Samsung India Electronics Pvt Ltd , Noida**

SUPERVISOR'S EMAIL ADDRESS: **pankaj786garg@gmail.com**

**DISSERTATION TITLE : Comparative Study For Character Recognition
Using Various Segmentation Techniques**

ABSTRACT : (Should be neatly word processed; should not exceed one page)

Broad Academic Area of Work: (Specify one from the areas listed in the guidelines)

Key words (Specify the technical keywords of the Project in alphabetical order)

Sheerangi Relan
Signature of the Student

Name: Sheerangi Relan

Pankaj
Signature of the Supervisor

Name: PANKAJ GARG

Date: 26-02-2017

Place: Noida

Date : 26-02-2017

Place: Noida

ACKNOWLEDGEMENTS

The satisfaction and euphoria that accompanies the successful completion of any task would be incomplete without mentoring the people who made it possible, because success is the epitome of hard work, perseverance, undeterred missionary zeal, steadfast determination and most of all "Encouraging Guidance".

I express my gratitude to my supervisor Mr. PANKAJ GARG for providing me a means of attaining my most cherished goals.

I record my heart full of thanks and gratitude to my additional examiner Mr. UDIT GUPTA for providing me an opportunity to carry this project, along with purposeful guidance and moral support extended to me throughout the duration of the project work.

SHEEVANGI RELAN

Contents

Introduction	1
Literature Review	2
OVERVIEW OF ALGORITHM	3
DataBase Training	5
Text Segmentation	6
IMAGE RECOGNITION	8
SCHEDULE AND RESULT	9
Summary	10
References	11

List of Figures

1. **Figure 1. Input Output Example..... 01**

2. **Figure 2. DataBase Training05**

3. **Figure 3. Text Segmentation06**

4. **Figure 4: Heuristic Tests07**

5. **Figure 5: Image Recognition..... 08**

Introduction

Imagine a world in which any person can take a picture of a signboard , billboard ,label with a mobile phone, letting to know what the message on the board is about , the ratings of the store , , and review the brands products that is of interest to the person.

Character Recognition techniques have been discussed in detail a number of times , But it still remains a very challenging field as numerous parameters affect it like clutter noise, orientation , text font , illumination , different color , non-standard text , and many more.

Text recognition has a lot of applications like automatic indexing , information retrieval , car plate recognition , content based image retrieval

We would input an image and the output would be a recognized text "STARBUCKS COFFEE"



Figure 1 : Input output Example

Literature Review

Text Segmentation

It is a technique to crop image in such a way that only the text portion of image is visible.

The project uses two different segmentation techniques

- **MSER (Maximally Stable External Region)**

This technique was proposed by Matas et al. The algorithm is a feature detector , that extracts from an image a number of co-variant regions, called MSERs.

- **Multi Scale Morphology**

It's an iterative collection of non-linear operations related to the shape or morphology of features in an image. They mostly rely on relative ordering of pixel values.

Recognition

It is the process of identifying and detecting an object or a feature in a digital **image**

The project uses two different recognition techniques

- **SIFT(Scale Invariant Feature Transform)**

In computer vision it's an algorithm that describes and detects local features .It has many applications like robotic mapping, Image stitching.

- **Traditional Recognition**

Using mat lab's built in functions it removes the unwanted spaces

OVERVIEW OF ALGORITHM

The algorithm is split into three phases.

Phase 1 is the **Database Training Phase**,

It creates a database required by the recognition phase . This phase is only executed once at the beginning to initialize the algorithm.

Phase 2 is the **Text Segmentation Phase**, which crops image so that only the portion of the image containing the text part is kept and is input to recognition phase.

There are two segmentation techniques: **MSER** and **Multi-Scale Morphology**.

Phase 3 is the **Recognition Phase**, makes use of the segmented image to convert the image into text.

There are two recognition techniques used

SIFT features matching and **traditional recognition technique**.

We make use of two segmentation techniques as well as two recognition techniques , resulting as following.

K=Threshold

DatabaseTraining() // Run once in starting

//Using MSER Segmentation and SIFT

```

MSER_Image = MSER_Segmented(image)

MSER_SIFT_Matches = SIFT_Recognition(MSER_Image)

if (MSER_SIFT_Matches > K)

return MSER_SIFT_Matches


//Using Multi Scale Morphology Segmentation and SIFT


MORPH_Image = MORPH_Segmented(image)

MORPH_SIFT_Matches = SIFT_Recognition(MORPH_Image)


if (MORPH_SIFT_Matches > K)

return MORPH_SIFT_Matches


//Using Traditional Recognition and MSER


MserMatches = TraditionalRecognition(MSER_Image)

if (MserMatches is valid)

return MserMatches


//Using Traditional Recognition and Morphology


MorphMatches = TraditionalRecognition(MORPH_Image)

if (MorphMatches is valid)

return MorphMatches

```

DataBase Training



Figure 2 : Data Base Training

The flow for Phase 1 is shown in Fig. 1.

A code book is formed in data base training phase. The following steps are followed

- Gathering of N manually segmented, images.
- Extraction of 128-dimension SIFT feature images and developing a Code sample list using K-Means algorithm. Using K-Means as per [1], K centroids get generated that are used as the K codes and entered into form the Code sample list.
- Next each feature is taken for a particular training image and we observe which code sample it quantizes to.
- A histogram is formed.

Text Segmentation

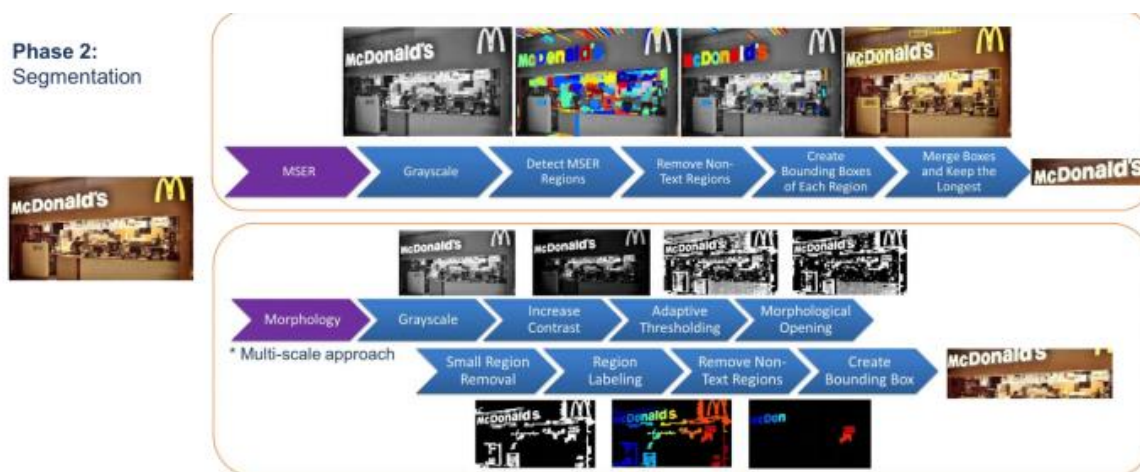


Figure 3 : Text Segmentation

Text Segmentation is essential for both SIFT and traditional recognition techniques. For Phase 2, two text segmentation techniques are used. Both flows are shown in Fig. 2.

A. Maximally Stable Extremal Region (MSER)

- It refers to a set of external regions that are present in a grayscale image[2] .
- An external region is characterized as a connected region with pixel intensities above (or below) a threshold.
- For maximum stability, the size of external regions must be nearly constant .
- MSER is mostly used to detect characters in natural figures as the nearly constant color and high contrast of texts results to stable intensity .
- Overlapping bounding boxes are combined to form one bounding box.

B. Multi-Scale Morphology

- The Multi-Scale Morphology refers to iterative morphology considering multiple assumptions of text size in the process.
- The initial step is to perform threshold.
- Convert colored image to grayscale.

- Secondly increase the contrast of the grayscale image.
- Now apply adaptive threshold on the image.
- Finally we segment the text using morphology.
- Firstly assuming a large scale white region to be the foreground text , morphological opening operation is performed and then small region removal is done .
- Next , the regions are labeled and we conduct heuristic tests shown in below Figure 4 to crop the background area.

Test	Description
1	The aspect ratio of letters are somewhat close to 1.
2	A letter has to be close horizontally to other letters.
3	A letter cannot take up too much space compared to the entire image.
4	Letters are probably of similar height.
5	Letters are probably located at similar y-coordinate region.

Figure 4 Heuristic tests

IMAGE RECOGNITION

For Phase 3, two recognition techniques are used. Both flows are shown in Fig. 4.

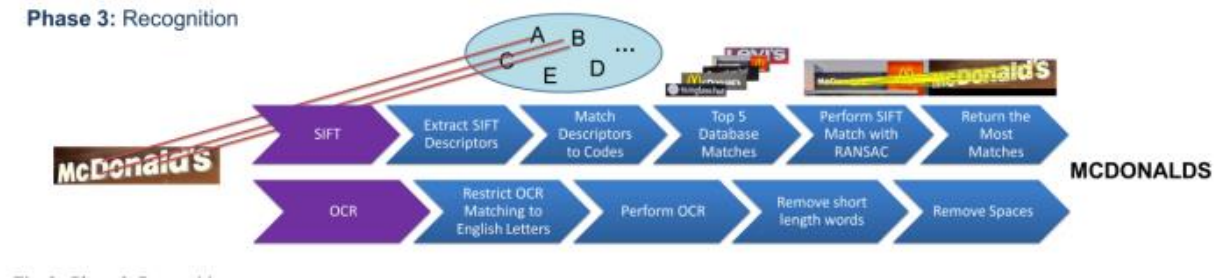


Figure 5 : Image Recognition

A. SIFT Feature Matching

- Extract the SIFT descriptors of the cropped testing image.
- Now quantize each feature to a particular code in the K-Means Code sample list.
- A histogram is formed of cropped testing image.
- This above prepared histogram is then compared with the histogram of the already prepared training images
- pair wise SIFT feature matching is performed between chosen image (training image) and test image[3]

SCHEDULE AND RESULT

SNo	Description of Work	Start Date	End Date	Status
1	System Study ,Software installations Study Segmentation MSER, Multi scale morphology, train database, Image recognition techniques	16-jan-2017	30-jan-2017	Done
2	Design and Analysis Design the algorithm Acquire test data	1-feb-2017	18-feb-2017	Done
3	Development Implementation of the DATABASE TRAINING, TEXT SEGMENTATION, IMAGE RECOGNITION Verify it against the test data	19-feb-2017	17-Mar-2017	Ongoing
4	Testing Recognition Prepare final report	18-Mar-2017	27-Mar-2017	Not started

Summary

The report discusses in detail about the algorithmic techniques among which the comparison will be performed .It discusses about the MSER ,Multi Scale morphological techniques .

Also discusses about the recognition techniques that will be used SIFT and Traditional recognition techniques using mat lab in built function.

The various stages of the algorithm have been explained .

The aim of the project is as follows

- To study "**MSER versus Morphology**" segmentation techniques , on basis of success rate of only using MSER or Morphology , also what are the results when both techniques are used as a combination to identify text of an image.
- To study "**SIFT versus Traditional Recognition**" recognition techniques , on basis of success rate of only using SIFT or Traditional Recognition.

References

- [1]Sivic, J.; Zisserman, A., "Video Google: a text retrieval approach to object matching in videos," in Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on , vol., no., pp.1470-1477 vol.2, 13-16 Oct. 2003
- [2] J. Matas, O. Chum, M. Urban, and T. Pajdla. Robust wide baseline stereo from maximally stable extremal regions. In Proceedings of British Machine
- [3] VLFEAT, Available: [http:// http://www.vlfeat.org/](http://www.vlfeat.org/)