

Features extraction and image clustering

Lorenzo Cioni, Francesco Santoni

lore.cioni@gmail.com, francesco.santoni@gmail.com

19 January 2015

Abstract

In this paper we discuss a method for extracting features from images containing handwritten US states to cluster them.

Contents

1	Introduction	1
2	The project	2
2.1	Troubles	2
3	Clustering	2
3.1	Features	2
3.1.1	Whitespace	2
3.1.2	Loop	2
3.1.3	Dot	2
3.1.4	Diagonal line	2
3.1.5	Horizontal line	2
3.1.6	Vertical Line	2
3.1.7	Cross	2
3.2	LCS	2
3.3	Affinity Propagation	2
4	Results	2
5	Conclusion	2

1 Introduction

Segmentation and clustering of big data is one of the main purpose of Artificial Intelligence at now. The aim of our work is to improve the performance of the clustering method implementend prevoiusly in an algorithm.

2 The project

2.1 Troubles

3 Clustering

3.1 Features

3.1.1 Whitespace

3.1.2 Loop

3.1.3 Dot

3.1.4 Diagonal line

3.1.5 Horizontal line

3.1.6 Vertical Line

3.1.7 Cross

3.2 LCS

In order to compare the generated strings each others we use the *Longest Common Subsequence* (LCS) algorithm for evaluating distances.

3.3 Affinity Propagation

4 Results

5 Conclusion