

Towards Semantic Clustering – A Brief Overview

Phei-Chin Lim

Faculty of Computer Science and Information Technology
Universiti Malaysia Sarawak
94300 Kota Samarahan, Sarawak, Malaysia

pclim@fit.unimas.my

Narayanan Kulathuramaiyer

Faculty of Computer Science and Information Technology
Universiti Malaysia Sarawak
94300 Kota Samarahan, Sarawak, Malaysia

nara@fit.unimas.my

Dayang NurFatimah Awg. Iskandar

Faculty of Computer Science and Information Technology
Universiti Malaysia Sarawak
94300 Kota Samarahan, Sarawak, Malaysia

dnfaiz@fit.unimas.my

Abstract

Image clustering is an important technology which helps users to get hold of the large amount of online visual information, especially after the rapid growth of the Web. This paper focuses on image clustering methods and their application in image collection or online image repository. Current progress of image clustering related to image retrieval and image annotation are summarized and some open problems are discussed. Related works are summarized based on the problems addressed, which are image segmentation, compact representation of image set, search space reduction, and semantic gap. Issues are also identified in current progress and semantic clustering is conjectured to be the potential trend. Our framework of semantic clustering as well as the main abstraction levels involved is briefly discussed.

Keywords: Image Clustering, Semantic Gap, Semantic Clustering, Concept Description, Symbolic Description.

1. INTRODUCTION

Advancement in digital imaging devices, technology and cost-decrease in storage devices contributed to the creation of large-scale digital images in various domains. The question now is how to effectively extract semantically meaningful information (knowledge) from these image collections. One of the fundamental of understanding and learning is to discover the natural groupings of images based on similarity of multiple characteristics or latent aspects of meaning. In this paper, the former referred to image clustering while the latter leads to semantic clustering.

One of the reasons for writing this paper is that we hardly find any work in reviewing image clustering methods by highlighting the challenges and provide association between image clustering and semantic clustering. Semantic clustering has attracted many research efforts after the year 2000 in terms of papers published. A simple experiment is conducted where we searched for publications containing the phrases *image clustering* and *semantic clustering* using