Unsupervised Emotional Scene Detection from Lifelog Videos Using Cluster Ensembles

Hiroki Nomiya, Department of Information Science, Kyoto Institute of Technology, Kyoto, Japan
Atsushi Morikuni, Department of Information Science, Kyoto Institute of Technology, Kyoto, Japan
Teruhisa Hochin, Department of Information Science, Kyoto Institute of Technology, Kyoto, Japan

ABSTRACT

An emotional scene detection method is proposed in order to retrieve impressive scenes from lifelog videos. The proposed method is based on facial expression recognition considering that a wide variety of facial expression could be observed in impressive scenes. Conventional facial expression techniques, which focus on discriminating typical facial expressions, will be inadequate for lifelog video retrieval because of the diversity of facial expressions. The authors thus propose a more flexible and efficient emotional scene detection method using an unsupervised facial expression recognition based on cluster ensembles. The authors’ approach does not need to predefine facial expressions and is able to detect emotional scenes containing a wide variety of facial expressions. The detection performance of the proposed method is evaluated through some emotional scene detection experiments.

Keywords: Ensemble Clustering, Facial Expression Recognition, Lifelog Video, Machine Learning, Video Retrieval

INTRODUCTION

Lifelog means a person’s record of life and has recently attracted attention (Aizawa et al., 2004; Gemmell et al., 2002). Although lifelog can be recorded as various types of data such as text, images and videos, the authors focus on lifelog videos (Datchakorn et al., 2005). Due to the improved performance and the minimization of recently developed video cameras, anyone can easily record his/her own lifelog videos. Therefore, lifelog videos will play an important role in our future life.
role in recording one’s life. However, lifelog videos have a serious problem in that it is difficult to efficiently retrieve useful scenes from the enormous amount of video data. As a result, valuable lifelog videos often remain unused.

In order to solve the issue, the authors have proposed an impressive scene detection method for lifelog video retrieval (Nomiya et al., 2013). Impressive scenes are considered to be useful since they contain recordings of important events that need to be retrieved during the lifelog video retrieval. The authors attempt to obtain impressive scenes by detecting emotional scenes based on facial expression recognition. It is important for human communication to utilize nonverbal communications including facial expressions (Mehrabian, 1971). Therefore, the emotions and the facial expressions of the persons in a video will be changed in impressive scenes.

Facial expression recognition has been extensively studied and can be applied to video-scene detection (Datcu & Rothkrantz, 2007; Fanelli et al., 2010; Littlewort et al., 2004). However, most of the existing approaches focus on the recognition of typical facial expressions such as six basic facial expressions including anger, disgust, fear, happiness, sadness, and surprise (Ekman & Friesen, 1975; Ekman & Friesen, 1978). In lifelog videos, however, more complicated and/or subtle facial expressions such as subtle smiles, broad smiles and wry smiles can be observed. This makes it difficult to detect useful scenes having a wide variety of facial expressions from lifelog videos because most of existing methods need to predefine the facial expressions.

Moreover, most of facial expression techniques are based on supervised learning. The supervised learning approaches generally require a large amount of training data to construct a good facial expression recognition model. Since it is quite difficult to automatically collect training data, preparing sufficient training data requires considerable human effort.

In the authors’ approach, a facial expression recognition model is constructed on the basis of an unsupervised learning using a clustering technique. For the purpose of improving the recognition accuracy, the authors introduce an ensemble learning approach (Dietterich, 2000) called cluster ensemble (Strehl & Ghosh, 2003), which integrates multiple clusters into a more discriminative cluster. Because the proposed method is unsupervised, it neither requires learning data nor the predefinition of facial expressions. The authors show the flexibility and efficiency of the proposed method through some emotional scene detection experiments.

The remainder of this paper is organized as follows. The next section presents some related works. The section following that describes the facial features used to recognize facial expressions. The next section then elaborates the facial expression recognition model based on the cluster ensemble. The section after explains the emotional scene detection method. The section after shows the emotional scene detection experiments using lifelog videos and a facial expression video database that has various emotions. Finally, the last section concludes this study.

**RELATED WORKS**

The facial expression recognition plays the most important role in the emotional scene retrieval. In order to precisely and efficiently recognize facial expressions, various kinds of facial expression recognition techniques have been proposed. The performance of the facial expression recognition can be greatly influenced by the facial features. Currently, there are two major types of facial features, appearance features and geometric features (Tian et al., 2011).

The appearance features are based on the skin texture of a face and can describe the appearance changes of a face such as wrinkles and furrows. The appearance feature can be obtained
Fuzzy Rule-Based Vulnerability Assessment Framework for Web Applications
[www.igi-global.com/article/fuzzy-rule-based-vulnerability-assessment-framework-for-web-applications/152244?camid=4v1a](www.igi-global.com/article/fuzzy-rule-based-vulnerability-assessment-framework-for-web-applications/152244?camid=4v1a)

Comparison Between Internal and External DSLs via RubyTL and Gra2MoL
[www.igi-global.com/chapter/comparison-between-internal-external-dsls/71818?camid=4v1a](www.igi-global.com/chapter/comparison-between-internal-external-dsls/71818?camid=4v1a)