Feature coding for image classification based on saliency detection and fuzzy reasoning and its application in elevator videos

Xiao Lv* , Dingdong Zou, Lei Zhang and Shangyuan Jia
Chongqing special equipment inspection and research institute
No.5 Furongyuan Road Northern New Chongqing
PEOPLE'S REPUBLIC OF CHINA
lvxiao87@126.com, cq_zdd@126.com, zl_816@163.com, 1458018035@qq.com

Abstract: - Feature coding is an fundamental step in bag-of-words based model for image classification and have drawn increasing attention in recent works. However, there still exits ambiguity problem, and it is also sensitiveness to unusual features. To improve the stability and robustness, we introduce saliency detection and fuzzy reasoning rules to propose an novel coding scheme. In detail, saliency maps generated by saliency detection are first used to divide each image into salient and non-salient region, then a structured dictionary is obtained by combing two separated codebooks in them. Secondly, fuzzy reasoning rules are introduced to choose the most salient and stable codewords to encode. Finally, saliency maps are incorporated into pooling operation named saliency based spatial pooling to introduce spatial information. Experiments on several datasets demonstrate our approach outperforms all other coding methods in image classification. Furthermore, we also apply it into elevator video event classification, which shows the potential application in intelligent elevator video surveillance, such as overload detection, violence detection, video summarization.

Key-Words: - Image classification, feature coding, saliency detection, fuzzy reasoning, elevator video event

1 Introduction

Automatic image classification is one of the most fundamental problems in computer vision and pattern recognition, whose aim is to assign one or more category labels to an image. It has drawn increasing attention from the researchers around the world due to its widespread prospects in a wide range of applications, e.g., image retrieval [1, 36], video retrieval [2], video surveillance [3], human-computer interaction [4], web content analysis [5], and biomedical [6, 37]. There are many approaches proposed for image classification in the literatures. Among them, the bag-of-words (BOW) model [7] and its extensions [8] achieve the state-of-the-art performance in several famous databases, such as Caltech 101 [9], Scenes 15 [10], Caltech 256 [11], and PASCAL VOC [12].

The BOW quantizes local descriptors into discrete visual codewords and counts their occurrence frequencies in the entire image. Then the resulting histogram is used as the image representation. Fig. 1 shows the general framework of the BOW model. It usually comprises of the following common steps: (1) Feature extraction. It extracts images' local features by detectors or dense sampling and then calculates their descriptors, such as Harris detector [13], affine invariant salient region detector [14], SIFT (Scale-Invariant Feature Transform) [15] descriptor, HOG (Histogram of Oriented Gradient) [16] descriptor. (2) Codebook generation. After obtained local descriptors, a codebook is usually needed to represent them. It is typically generated by clustering (e.g., K-means [17]) over a subset of descriptors, which is randomly sampled from all descriptors in database in real application for computational efficiency. (3) Feature coding and pooling. In this step, each local descriptor first activates a number of codewords, and generate a coding vector. Then, all responses on each codeword are integrated into one value by feature pooling. Various coding and pooling strategies will be described in detail in Section 2. The output of this step is a vector whose length is equal to the size of the codebook, namely the final image representation. (4) Classification. Finally, the image representation vectors are sent to a classifier, such as SVM (Support Vector Machine) [18-19] for classification.

Among these steps, feature coding and pooling is the fundamental component, which will greatly influences image classification in terms of both

* Corresponding author.