Descriptor Optimization for Semantic Concept Detection Using Visual Content

Mohamed Hamroun, LabRI- University of Bordeaux, Tunisia, France

Sonia Lajmi, MIRACL, University of Sfax, Tunisia, France

Henri Nicolas, LabRI- University of Bordeaux, Tunisia, France

Ikram Amous, MIRACL, University of Sfax, Tunisia, France

ABSTRACT

Concept detection has been considered a difficult problem and has attracted the interest of the content-based multimedia retrieval community. This detection implies an association between the concept and the visual content. In other words, the visual characteristics extracted from the video. This includes taking knowledge about the concept itself and its context. This work focuses on the problem of concept detection. For that, several stages are elaborated: first, a method of extraction and semi-automatic annotation of the video plans for the training set is proposed. This new method is based on the genetic algorithm. Then, a preliminary concept detection is carried out to generate the visual dictionary (BoVS). This second step is improved thanks to a noise reduction mechanism. This article's contribution has proven its effectiveness by testing it on a large dataset.

KEYWORDS

Bag Of Visual Features, Concepts Detection, Descriptor Optimization, Multimedia Indexing, Semi-Automatic Annotation

1. INTRODUCTION

Following the rise of digital technologies, the quantity of multimedia documents exposes. This process made their indexation very expensive and manually impossible. As a result, the need for indexing systems that can analyses, store and retrieve multimedia documents automatically, based on their content, has been felt in many application areas. However, current indexing techniques still face problems of feasibility or quality. Their performance remains very limited and depends on several factors such as the variability and the amount of data to be processed. The recognition concepts (as bike, chair, etc.) approaches, proposed until now, face the problem of variability of shapes and positions. These approaches suffer from scaling problems.

Concept detection is a very crucial problem in the field of image and video processing. It is essential for different applications in the analysis and automatic video indexing. For example, in the field of automatic car driving, it is important to distinguish between the concept rock and bag, which have a very similar appearance once captured from the camera sensor. This sensor records and analyzes in real-time the scene in front of the car. Concept detection systems are expected to find key-frames containing class or category (or concepts) of the considered object. Therefore, the resolution of the localization problem requires a system, not only able to recognize the concepts, but also to indicate their location, with precision, in the key-frames using an enclosing frame, a polygon or a pixel. Locating concepts in key-frames is a very difficult task because of the variability of video content.

Existing approaches tackle the detecting concept problems according to different levels of supervision. We can distinguish between highly supervised and weakly supervised detection techniques. Highly supervised techniques require a set of positive and negative annotated images. For this kind of techniques, the location of the objects must be well indicated for the learning phase. As for the weakly supervised techniques, they aim to perform the same tasks but without any indication of the object's location in the key-frames. Such concepts can appear several times and arbitrarily in positive videos, the location task becomes easier with a high-level supervision.

The simplest solution for concept detection would be to manually annotate the concepts of all training data (highly supervised learning). Recently, interesting work on small dataset such as Caltech04 (Zhang & Chen, 2010; Nguyen et al., 2009; Opelt & Pinz, 2005) or Weizmann (Winn et al., 2005) or just in few categories of PASCAL-VOC (Zhang & Chen, 2010; Pandey & Lazebnik, 2011) have shown that supervised learning is a very promising approach for the concept detection. Those approaches are effective for limited sized dataset or for well-defined and easily selected concepts. Indeed, they suffer from scaling problems: they are not applicable to large corpus (such as TRECVID). Therefore, weakly supervised learning is required as a greatest alternative to reduce costs (Galleguillos et al., 2008; Fergus et al., 2007; Zhang & Chen, 2010; Prest et al., 2012; Crandall & Huttenlocher, 2006; Nguyen, et al., 2009; Opelt & Pinz, 2005).

The following questions are the focus of this paper: What is the amplitude of weakly supervised methods in concepts detection applied for large video collections like TRECVID?

As an attempt to answer this question, we proceed as follows:

- We begin by proposing a framework based on weakly supervised learning techniques for the detection concepts in large video collections.
- Next, we evaluate the proposed method on the TRECVID 2015 dataset and compare our performance to that of other concept detection approaches.
  The innovative aspects of the proposed approach are:
- A semi-automatic annotation method based on the genetic algorithm in order to reduce the manual annotation in the indexing process.
- An automatic concept detection method based on our own optimized PMC descriptor (Hamroun et al., 2018).

18 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/article/descriptor-optimization-for-semantic-concept-detection-using-visual-content/237145

Related Content

Information Technology Process Improvement Decision-Making: An Exploratory Study from the Perspective of Process Owners and Process Manager
Sandy A. Lamp, Kathleen M. Hargissand Caroline Howard (2012). *International Journal of Strategic Information Technology and Applications (pp. 18-35).*
www.irma-international.org/article/information-technology-process-improvement-decision/67348

Strategic Management of International Subcontracting: A Transaction Cost Perspective
Yue Wang (2010). *Strategic Information Systems: Concepts, Methodologies, Tools, and Applications (pp. 1219-1229).*
www.irma-international.org/chapter/strategic-management-international-subcontracting/36753

Supply Chain Management and Portal Technology
Scott Paquette (2010). *Strategic Information Systems: Concepts, Methodologies, Tools, and Applications (pp. 1134-1141).*
www.irma-international.org/chapter/supply-chain-management-portal-technology/36748

New Interpretation of Some Forgotten Problems
Dan Ciulin (2016). *International Journal of Strategic Information Technology and Applications (pp. 1-63).*
www.irma-international.org/article/new-interpretation-of-some-forgotten-problems/186762

A Proposed Architecture for Autonomous Mobile Agent Intrusion Prevention and Malware Defense in Heterogeneous Networks
Robert O. Sealeand Kathleen M. Hargiss (2011). *International Journal of Strategic Information Technology and Applications (pp. 44-54).*
www.irma-international.org/article/proposed-architecture-autonomous-mobile-agent/60144