

Survey on Semantic Concept Detection

Julin Rose Jacob

PG student

Computer Science Department

Karunya University

Coimbatore, India

Abstract

Concept detection plays an important role in digital image processing. This paper is based on the survey of various techniques used in the semantic concept detection. One of the main opportunities to improve the performance in concept detection is feature extraction and matching image. In semantic concept detection explains the concept detection in the video shots. For getting the accurate video shots, there are different types of techniques used. Most of the papers are explained about various techniques that used in concept detection. Each technique has got its own advantages as well as disadvantages.

1. Introduction

Semantic concept detection defines detecting the concepts that presence in the video shots. One of the technique that can be used for the powerful retrieval or filtering systems for multimedia is semantic concept detection. Semantic concept detection also provides a semantic filter to help analysis and retrieve a multimedia content. It also determines whether the element or video shot is relevant to a given semantic concept. Concept detection means detecting the visual concepts like people, object or any locations, given a predefined concept lexicons and sufficient annotated examples. Semantic concept detection based on the image retrieval or by any other methods. For developing the models for annotated concepts we can use annotated training datasets. In the semantic concept detection there is an existing technique. The existing techniques of semantic concept can be classified as:

1. Support vector machine
2. Extreme learning machine

In support vector machine (SVM) is concept classifier is based on generic visual feature that manually annotated with the video shots. The concept classifier indicates the probability of the target concept that present in the given video shots. In SVM enhances the practical performances of content based video retrieval to come extent. By

using the technique SVM has good accuracy and get efficient value for the concept detection.

In semantic concept detection framework [1] includes four main stages. The four main stages are

1. Normalization and feature extractions.
2. Data splitting.
3. Rule generation.
4. Rule selection and classifications.

In normalization and feature extraction explains that extracting audio-visual features that set based on the shot boundaries of each video and the numerical features used in extracted and normalized. Then the data is splitting into training and testing data sets. The training data set is discretized and after discretization a rule generation occurs. The rule is generated based on the feature process. After the rule generation the selection and classification occurs. One of the main ideas for semantic concept detection is a statically learning problem. For this each video shot can be detected and associated with the visual features. Semantic concept detection is detecting the presence or absence of the semantic concepts in video shots. One of the main task is to detect efficiently and analyzing the semantic representation of video shots.

By evaluating the semantic conceptions based on the TRECVID datasets. TRECVID datasets are used for promoting the content based analysis and retrieval from digital video. TRECVID used various test data sets from broadcast videos, TV program producers etc. The following of this paper explains various techniques that used in the semantic concept detection and also explained the comparison of the techniques used for semantic concept detection and summarizing this paper.

2. Methodology

2.1. Association Rule Mining

Association rule mining (ARM) [2][3][10] is used for detecting the concepts from the video shots. It is one of the techniques that used for the semantic concept detection. In ARM has been adopted for bridging the semantic gap between the

low-level features and the high-level features. In semantic concept detection for video to perform high efficiency and good performance, here we use classification name association rule mining. In association generation rule that implemented by three categories. They are objective measures, subjective measures and semantic measures. The objective measures are calculated based on the statistics, distances etc. In the subjective measures includes both data and users domain knowledge about the data. The third category is semantic measures explains the feature value pairs. These are evaluated based on the TRECVID datasets. The proposed framework is shown in fig.1

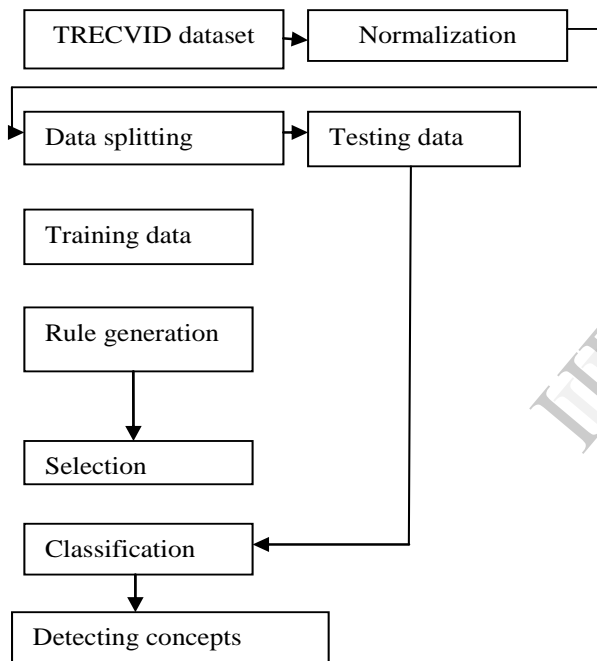


Fig.1. Block diagram for proposed framework of ARM.

In ARM consists of two main phases. The first phase explains the association rule generation and the second phase explains the rule selection. According to mining process generating and rank the rules with reduction of time and sparse costs. To improve the efficiency here we use algorithm is traditional ARM algorithm. It's also evaluating the criteria of video shots.

2.2. Semantic indexing of multimedia content

For large video libraries contains certain tools for representing, searching and retrieving the video content. In this paper [4] we present a learning based approach for semantic indexing. To

exploiting the various features like visual, audio and text cues. For this here we use Gaussian Mixtures Models (GMM), Hidden Markov Models (HMM) and Support Vector Machine (SVM). In HMM models generalized the sound recognition and the GMM models [5] explain the maximum posterior classifiers and neighbors of classification of speech and sound. In HMM [6] is based on the early fusion of the context for audio visual recognition. In this method mainly focus on the audio, visual and textual modalities for semantic modeling video. In this technique also explain two main modeling. One is probabilistic modeling for semantic concepts and discriminant technique. SVM approach is also used for late fusion approach

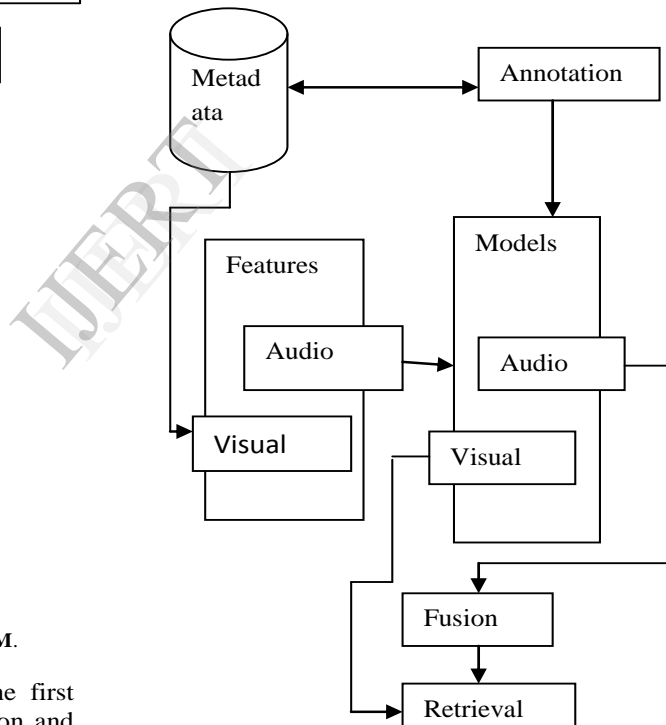


Fig.2. Block diagram for semantic concept analysis.

For analyzing the semantic concept and retrieving the contents, here we use three main components. First component is defining a lexicon of semantic concepts and annotating examples. Second is the learning the representation and the third is the data retrieval. The entire framework for

semantic labelling to the data retrieval is shown in Fig.2.

2.3. Reranking Approach

A reranking approach [7] is used for context based concept fusion in video indexing and retrieval. In reranking approach automatically discovers related concepts from the video and also incorporate the detection to refer the results for search. Reranking is rooted in pseudo-relevance feedback (PRFB) [8] for text search. Reranking is unique and takes a ranked list of results from initial search result of the semantics of the target. In initial search lists includes related concepts. In pseudo-relevance feedback for text search includes pseudo-positive. In pseudo-positive means the documents are assumed to be true. In pseudo negative are obtained from video collections. Both pseudo-positive and pseudo-negative are extracted from the test search results. Here we assume that pseudo-positive for top ranked results and pseudo-negative results from the lower ranked results.

Consider the target concept as S and the concept detector scores as C. The main objective is to improve the performance by target pseudo labels S and concept detection scores C. here we have to find the subset of concept lexicon by measuring the mutual information between the two [7]:

$$I(S; C) = \sum_S \sum_C P(S, C) \log P(S, C) / P(S)P(C)$$

Here P(S,C),P(S) and P(C) are estimated counting features in sampled video shots.

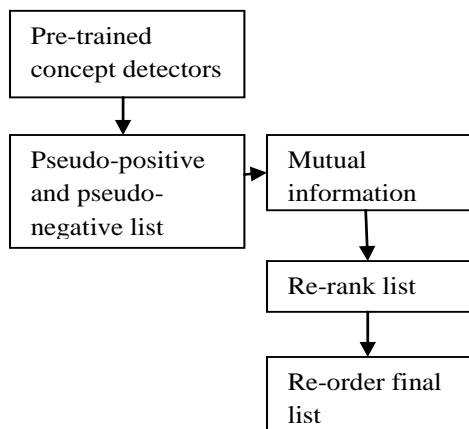


Fig. 3. Block diagram of the Reranking approach

In figure 3 represents the block diagram explaining the reranking based context fusion method. In this figure explains that we need to target a concept by incoming query. Then by initial

ranked list divides it into pseudo-positive and pseudo-negative list. Giving this both ranked list to the related concepts contains large set of pre-trained concept detectors. After this reranked list obtained and the final output will be as re-ordered final lists.

2.4. Extreme learning machine

In Extreme Learning Machine (ELM) [9] explains that it is a single hidden layer feedforward networks that focused on input sets and training samples. ELM generates the hidden layer output matrix. For past decades the application of neural networks are slower than the other techniques. It is slower because of two main reasons. The one reason behind this is slow gradient-learning algorithm and the second reason behind this is parameters are tuned by these algorithms. By avoiding these here we propose the algorithm is extreme learning machine for single hidden layer feedforward neural networks. By using this ELM algorithm can produce good performance and can learn faster than other algorithms. In single hidden layer feedforward network contain N hidden neurons. By using ELM we can determine the hidden nodes and find the output matrix.

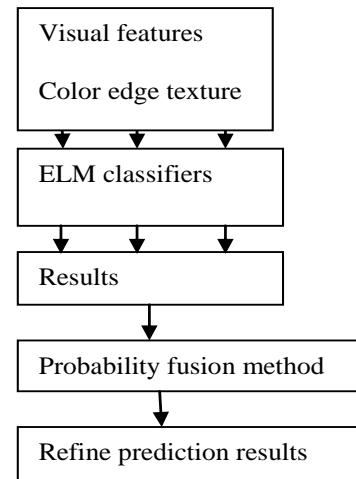


Fig.4. Block diagram for ELM for multicategories framework.

ELM is also good for multi-categories problem. To improve the accuracy of the semantic concept detection ELM is used based on the multi-modality classifier combination frameworks. It explains three different steps. First step explains that extracting the visual features like color, edge

and texture. Besides the classical ELM methods ELM classifier is used to extracting each feature for training the datasets. In this step also used One-Against-All (OAA) [14] method is used. In second step explains prediction results and the probability fusion methods. The final results explain the probability of the prediction results in the concept of the video shots. By evaluating the ELM by using the mean average precision (MAP). In ELM the MAP ranges from 0.3 to 0.5. The above block diagram represents the ELM framework in figure 4.

2.5. Support Vector Machine

Semantic concept detection is an important in concept based semantic video retrieval. To analyzing the semantic concept detection support vector machine (SVM) [11] is used. SVM is traditional algorithm that used for semantic concept detection. SVM is the concept classifiers. In concept classifier is based on the visual features that manually labelled video shots and also indicates the probability of a target concept that present in the video shots. SVM are supervised learning models that are associated with learning algorithms.

For improving the semantic concept detection the SVM detector is used. SVM detector is used for finding the prior probability of concepts in the video or the image. For the semantic level fusion SVM classifier is used. By using the SVM classifier we can used to train each SVM classifiers and train each features. There are different steps used for improving the detection score. First step is the detecting the detection score from the each individual SVM detectors. Then incorporates the concept and employ a probability prediction rule to estimate the detection. Finally apply a weighted linear combination to aggregate probabilities into final detection score.

To produce final detection score, V_i by combining all the estimates [11].

$$V_i = \sum_{j=1}^n W_{ij} P_{E_j} (Y_i = 1)$$

Here W_{ij} denotes the weight of the estimate of the j^{th} detector for E_i . By using the TRECVID [13] data sets we can evaluate the SVM detectors.

3. Comparison

By comparing various techniques used in the semantic concept detection consider the mean average precision of each technique. The following table represents the MAP values that used in each technique. Reranking approach [7], SVM [11] and ELM [9] are the techniques used for the comparisons in semantic concept detection. In the following table also explains which techniques performance is good.

TABLE1: COMPARISONS OF VARIOUS TECHNIQUE USED IN SEMANTIC CONCEPT DETECTION.

Ref no	Technique used	MAP
1	Reranking approach	0.153
2	SVM	0.167
3	ELM	0.179

4. Conclusion

Semantic concept detection is one of the concept detection that can be used in the image processing. Semantic concept detection defines detecting the presence or absence of the concepts in video shots. There are various techniques that are used for the effective and accurate concept detection. In this paper various techniques are analyzed. And also explains various advantages and disadvantages of various techniques used in the semantic concept detection.

5. References

- [1] L.Lin.,M.L.Shyn.,and C.C.Chen.Correlation based interestingness measure for video semantic concept detection: A survey ACM Computing Surveys, September 2006.
- [2] L. Lin, G. Ravitz, M.-L. Shyu, and S.-C.Chen. Video semantic concept discovery using multimodal-based association Classification .In *IEEE International*

- Conference on Multimedia and Expo*, pages 859–862, July 2007.
- [3] B. Liu, W. Hsu, and Y. Ma. Integrating classification and association rule mining. In *International Conference on Knowledge Discovery and Data Mining (KDD98)*, pages 80–86, August 1998.
- [4] W.H.Adams.,G.Iyengar.,Ching-yung Lin.Semantic Indexing Of Multimedia content using visual,audio and text cues.EURASIP journal on applied signal processing 2003:2,1-16.
- [5] E. Scheirer and M. Slaney, “Construction and evaluation of a robust multifeature speech/music discriminator,” in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, vol. 2, pp. 1331–1334, IEEE, Munich, Germany, April 1997.
- [6] M. A. Casey, “Reduced-rank spectra and minimum-entropy priors as consistent and reliable cues for generalized sound recognition,” in *Proc. Eurospeech*, Aalborg, Denmark, September 2001.
- [7] Lyndon S.Kennedy, Shih-Fu Chang. AReranking Approach for context-based concept fusion in video indexing and retrieval,july 9-11,2007.
- [8] J. Carbonell, Y. Yang, R. Frederking, R. Brown, Y. Geng, and D. Lee. Translingual information retrieval: A comparative evaluation. Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence, pages 708–715, 1997.
- [9] Bo Lu, Guoren Wang, Ye Yuan, Dong Han.semantic concept detection for video based on extreme learning machine,july 2012
- [10] N. Zhao, S.-C. Chen, and S. H. Rubin. Video semantic concept discovery using multimodal-based association classification. In *IEEE International Conference on Information Reuse and Integration (IRI07)*, pages 373–378, August 2007.
- [11] Yusuf Aytar, O.Bilal Orhan and Mubarak Shah, ”improving semantic concept detection and retrieval using contextual estimates”
- [12] G.Iyengar, H. Nock, and C. Neti, “Discriminative model fusion for semantic concept detection and annotation in video,” *ACM Multimedia*, pp. 255-258, Berkeley, USA, 2003.
- [13] A. Yanagawa, W. Hsu, S.-F. Chang, Brief Descriptions of Visual Features for Baseline TRECVID Concept Detectors. Columbia University ADVENT Technical Report 219-2006-5, 2006.
- [14] H.J. Rong, G.B. Huang, Y.-S. Ong, Extreme learning machine for multi- categories classification applications, in: Proceedings of IEEE International Joint Conference on Neural Networks, 2008, pp. 1709–1713.