Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 48 (2015) 101 – 106

Procedia Computer Science

International Conference on Intelligent Computing, Communication & Convergence

(ICCC-2015)

Conference Organized by Interscience Institute of Management and Technology,

Bhubaneswar, Odisha, India

Fuzzy Ontology Based Multi-Modal Semantic Information Retrieval

G.Nagarajan[a], R.I.Minu[b]

[a]*Sathyabama University,Chennai-42,India*
[b]*Jerusalem College of Engineering,Chennai-100,India*

Abstract

Lack of today's technology of information retrieval system is the semantic between the information resources. The semantic can be expressed for a resource using the underlying technology of ontology. So the main objective of this paper is to design an information retrieval system for both text and image data using the concept of ontology. The main focus of this paper is to improve information retrieval for sports events using Ontologies. As it is complex problem, it is addressed into the following sub-problems in this paper. (1) Integrating domain knowledge and images using fuzzy ontology and Retrieving the required Muti-modal information using fuzzy rule set. (2) Providing image semantic by constructing visual codebook for affine covariant-Semantic segmented patches. (3) Analysing the discriminative power of each visual word using the probabilistic latent semantic and quantizing them using the Chi-Square test. In this work, the domain of Sports is considered for creating both Low-level visual ontology for certain sport event images and also for building a high-level domain ontology from the information on web. These two Ontologies are integrated using Fuzzy concepts.

Peer-review under responsibility of scientific committee of International Conference on Computer, Communication and Convergence (ICCC 2015)

*Keywords:*Ontology; Knowledge representation; Semantic Analysis; OWL; Machine Learning; Computer Vision;

1. Introduction

Images are the mental and visual treat for human beings. The amount of different strata of images is increasing exponentially in web and also in research repositories. The concept of image classification, image recognition, and

Peer-review under responsibility of scientific committee of International Conference on Computer, Communication and Convergence (ICCC 2015)
doi:10.1016/j.procs.2015.04.157

image retrieval has become more[1] labyrinthine than ever before due to the availability of enormous amount of images. The complication in analyzing the images with respect to their context is due to the semantic gap[2] between the image features and the human linguistic about the image. If an application can study and analyze an image as human does, the semantic gap would be reduced relatively. The ultimate objective of this work is to create a system which would visualize and understand the images. The mere existence of natural living thing can be studied and analyzed efficiently by ontology[3], in which every existing resources are considered as an entity and they are schematized hierarchically via their relationship. The motivating concept of this work is to create a multimodal ontology for sport images, so that it would be easier to analyze and study the image automatically by a machine. Thus, a machine can visualize an image as human does. This would provide a way to develop a Multimodal Image Retrieval System.

Nomenclature

m	Number of row pixels
n	Number of column pixels
k	Number of key point generated

2. Related Works

Some of the most widely used image retrieval systems are text based image retrieval, content based image retrieval, and ontology based image retrieval techniques that are reviewed here. An image can be expressed in text by annotating it in relation to the object around it. There are certain types of annotation like manual annotation, semi-automatic annotation and automatic annotation, Linguistically the images are all indexed in this annotation concept. . For web images Schroff . F, et. al. [4] designed a framework. Here the image and the non-image were identified and the unwanted images were removed from the search. Then from the text around the image in the web pages was used to index the image. The authors also used probabilistic methods for image annotation. Su et. al.[5] tried to reduce the semantic gap using fuzzy membership function. From the web document both visual and textual data were extracted, and model was created. For image analysis, the algorithm used was not discussed; just the creation of model was specified. From the text, the keywords were all extracted. Both the models were fused with fuzzy set. This concept was implemented for the retrieval process. For categorization of the images, the probability based pLSA approach was used as highlighted by the authors Kim.S and Daijin[6] and Zhang. J., et.al [7]. The comparative study of this work with the other content-based image retrieval system provided satisfactory results. In the work of Farah.I.Ret. al.[8] for each satellite image, they had created three types of ontology: the scene model ontology sensor model ontology, and spatial relation model ontology. Then these ontologies were all merged and used for semantic image retrieval system. For all the created ontology, the image based keyword and certain metric were used. In text based image retrieval system, they had incorporated the idea of ontology to it. The concepts to fill the semantic gap were not elaborated in this technique. In the work of Yildirimet.al.[9] the content of the basketball game video was extracted. The author concentrated only on high level semantic of the video's description. The action activity of the basketball game was organized and a complete ontology for the videos was created. From the created ontology the exact action and activity of the player or the referee was extracted by involving the fuzzy concept.

3. Technical Approach

The main objective of this research work is to enhance information retrieval system using Ontologies. This objective gets multiplied into several aspirations, as follows: As the visual contents are uncertain data, we require a probabilistic model to handle the uncertain situation. In the domain of basketball for dunk action image, the ball may be near the hoop or far from it. Therefore, it requires a fuzzy set to guide the classification module with higher degree of confidence. The information gathered from web regarding basketball with respect to player name, team name, and different action is to be integrated with the action images using the generated codebook. Due to the disambiguate information from image we requires a fuzzy set rule to integrate this information in a created domain ontology. To represent the semantics of visual content a hybrid technique of Bag-of-visual word based image indexing was proposed. The feature used to index the images has to preserve the image semantic signature. The major weakness of existing frameworks is that visual features are disambiguated to index in database. This is more of a statistical calculation that does not represent the actual semantic relations properly. The overall proposed framework is shown in Fig.1

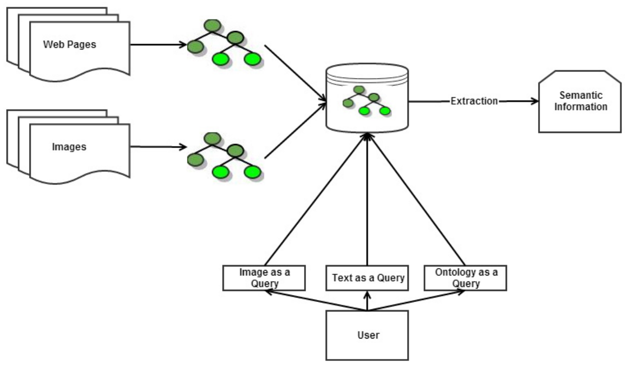
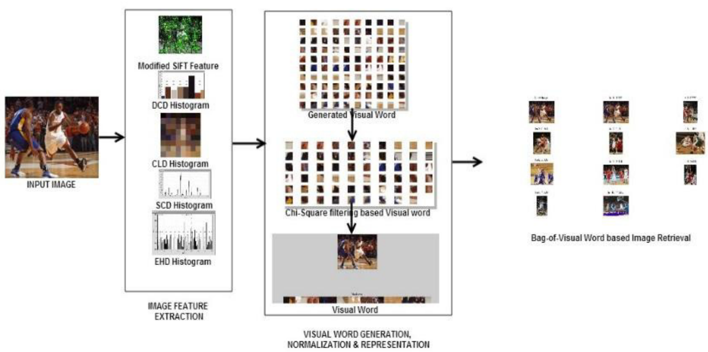


Fig. 1. Proposed System Framework

1.1. Affine covariant based visual codebook generation

1.2. The basic idea behind bag-of-visual word generation is to first quantize local descriptors into visual words and represent each image as a vector of words like a document. Therefore, by analyzing the image, a visual analogue of a work is used in the BoW model, which is based on the vector quantization process by clustering low-level feature of the image. In this paper, from the sport event image datasets the features extracted are: modified SIFT and MPEG 7 based dominant color, color layout, scalable color and edge histograms. From these vector values after quantization and analysis the visual vocabulary would be created. This created vocabulary dictionary is integrated in the domain specific visual ontology for semantic understanding of the image through ontology.



The overall framework for the creation of visual word is shown in Fig.2

Fig.2Methodology used to create a BoW

Thus from the input sports image, the interesting points, with respect to its Affine transformation are identified. From the location vector of those interesting points image patches are generated. The algorithm used for the extraction of image patches is explained in Algorithm 1.

Algorithm 1: Image Patch Generation Input: Image Output: Image Patches
Step 1: Convert the input image of size (mxnx3)into grayscale image of size (mxn) Step 2: Implement Harris-Affine transformation on image and identify the location vector of the interesting point (kx4), where k is the number of interesting point , k < m & n Step 3: For each identified interesting point a patch of size 10 x 10 is extracted. Step 4: Thus K number of image patches are Generated for the input image of (mxn) size

1.3. Fuzzy Ontology integration

1. Visual ontology is used to provide semantic to the generated visual word. In this work the domain of Basket-ball game images was considered. An ontology with respect to the image feature and domain knowledge was created. For the creation of ontology, Fuzzy concept with respect to cognitive spatial relationship was also incorporated to identify action on the image. Fig.2 shows the creation of fuzzy basketball ontology. As the creation of such ontology requires data from image and documents, it is said to be multimodal ontology.

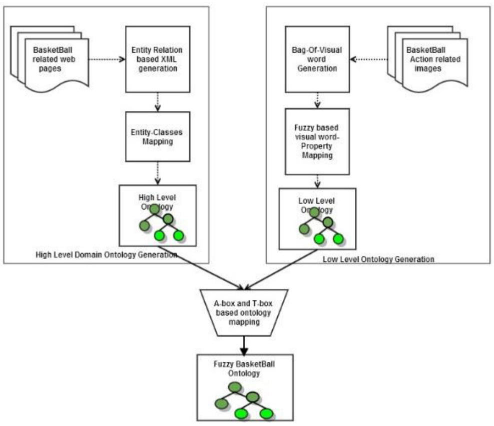


Fig. 2 Creation of Fuzzy basketball ontology

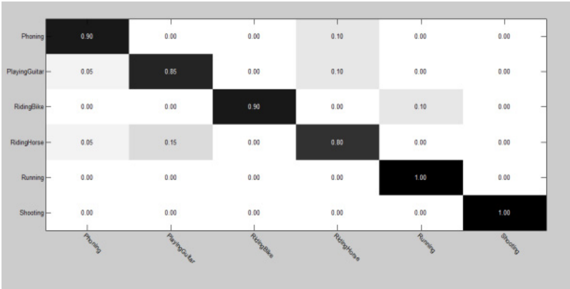
Mainly the created ontology consists of three main classes Item, Action and Cognitive spatial relationship. In item class the general living and non-living object such as player name, referee name, team name, ball, basket hoop, free throw line etc. are listed. In action class the general action of basket-ball games such as Assist, Dunk, Freethrow, Jumpball, Pass, Rebound and. In this ontology the concept of cognitive spatial relationship is used to find the relationship between ball, body, hand, basket and free throw line in the image. To measure the distance and to identify the cognitive spatial relationship fuzzy membership function where used . The ontology entities are selected generically so that these classes can also be modified for any indoor and outdoor sport events. It requires modification only on action attributes. The subclasses for such cognitive_spatial_relation classes. Thus the core component of this thesis is the creation of domain specific sports event ontology and mapping this ontology with the extracted low-level and high-level knowledge base. The created ontology can be used as the back-end system for the semantic search engine framework. In this work other than adding the low-level visual word feature to the respective action image, the type of action is also identified using the concept of cognitive spatial relationship. In ontology the semantic is provided by the property entities. There are two type of property object property and data property. The relationship between two classes is said to be the object property. Thus these property are used to specify the action of the given image. To represent the distance between the balls with respect to other notable item, the data-property balldistace where used, which has the sub-property ballhead, ballhoop and ballplayer. Using the distance value the action is identified. The four main property are balldistance, playername, teamname and Bagofvisualword. To identify the action occurred in the image the distance between the ball and other notable item is the important criteria. The hasTouchSpatial, hasdisjointSpatial and hasOverlapSpatial properties are used to link the relationship between the action and the cognitive_spatial_relation class. The main objective of this work is to identify the action of the basketball by giving the image as input. Let us consider the image given as input and it is shown in Fig.3. Using the modified SIFT and bag-of-visual word concept the object on the images is identified.

Fig.3The class, data property and object property relation for the Dunk action image

The three main cognitive spatial relationships are considered for identifying the action on it. The used relation are touch, disjoint and overlap. From the input images the ball and the person are overlapped; then the ball and hoop are disjoint. This relation between the object are identified using ground truth based fuzzy system. Thus the identified class with respect to images are Person, Ball and Hoop of item classes. The identified cognitive_Spatial_relation are BallAwayFromHoop and PlayerPartiallyInsideBall. The relation between ball and person is ‘hastouchSpatial’ and the semantic relation between ball and hoop are ‘hasDisjointSpatial’.

4. Results and Implementation

2. The entire procedure was conceived as an interface using Matlab and implemented on Stanford University sport event datasets. Fig.4(a) shows the satisfactory confusion matrix for this approach and the output is also shown in the preceding Fig.4(b). The overall accuracy in identifying the image with respect to other systemsis shown in. Fig.4(c).



(a)

(b)

(b)

Fig.4(a)Confusion matrix on identifying the games using pLSA (b) Overall Image retrieval using pLSA (c) The overall accuracy of this system

References

1. Bernal. J (2010), “Feature detectors and feature descriptors: Where we are now”, Technical Report 154, Computer Vision Center& Computer Science Department UniversitatAutònoma de Barcelona.
2. Hao.M.J., Zhu, J., Lyu, M. T., and King, I (2010),“Bridging the Semantic gap between Image contents and Tags”,IEEE Transaction on Multimedia, Vol.12, No.5, pp.462-470.
3. Hyvönen. E., Saarela, S., Styrman, A., and Viljanen, K (2003),“Ontology-Based Image Retrieval”, Proceedings of WWW2003, Budapset, Hungary.
4. Schroff. F., Criminisi, A., and Zisserman, A (2011), “Harvesting Image Databases from the Web”, IEEE Transactions on pattern analysis and Machine Intelligence, Vol. 33, No.4, pp.754-766.
5. Su.J.H, Wang, B. W., Hsu, T. Y., Chou, C. L., and Tseng, V. S (2010),“Multi modal image retrieval by integrating web image annotation, concept matching and fuzzy ranking techniques”,International Journal of Fuzzy Systems, Vol.12, No. 2, pp.136-149.
6. Kim . S and Daijin.K (2009),“Scene classification using pLSA with visterm spatial location”,Proceedings of 1st ACM International Workshop on Interactive Multimedia for Consumer Electronics, New York, pp. 57–66.
7. Zhang. J., Li, D., Hu, W., Chen, Z., and Yuan, Y.(2014), “ Multilabel Image Annotation Based on Double-Layer PLSA Model”, The Scientific World Journal, 2014.
8. Farah.I.R., Messaoudi, W., Ettabâa, K. S., and Solaiman, B (2008),“Satellite Image Retrieval based on Ontology merging”,ICGST-GVIP Journal, Vol.8 ,No.2, pp. 1-6.
9. YakupYildirim ,Yazici, A., and Yilmaz, T (2013),“Automatic semantic content extraction in videos using a fuzzy ontology and rule based Method”, IEEE Transactions on knowledge and data engineering,Vol.25,No.1.