A Comparative Study of Multimedia Retrieval Using Ontology for Semantic Web

P.Kannan[1], P.Shanthi Bala[2], Asst.Professor and Dr.G.Aghila[3], Professor
School of Engineering and Technology, Department of Computer Science
Pondicherry University, Pondicherry, India
p.rajkannan@gmail.com[1], shanthibala.cs@gmail.com[2]

***Abstract*— Multimedia plays significant role in today's IT world. The revolution of multimedia makes it more familiar to the users because of its expressiveness. Multimedia has a wide range of application in many fields like e-learning, teleconferencing, online medical transcription, etc. Semantic web is an emerging technology to fulfill the user's needs. Over the past decade, lots of research is going on for retrieval of multimedia content in semantic web. This paper discusses various retrieval techniques of multimedia content using ontology for the semantic web. In addition, it discusses the advantages of Text, Image, Video and Audio based retrieval systems. It also presents an analysis based on its merits.**

***Index Terms*— Multimedia, Ontology, Semantic Web, Semantic-based search.**

I. INTRODUCTION

Use of multimedia makes the user's perception in more realistic manner with less effort. Generally, multimedia is the combination of text, image, video, audio, graphics and interactive contents. The number of users searching multimedia over the web is increasing. Multimedia enhances the applications with rich content and use of media tools making the communication more effective. So, Multimedia retrieval is desirable for most of the web users. Even effective retrieval of multimedia in "World Wide Web" is still average. The end user can retrieve the multimedia data very effectively based on the depth knowledge of the multimedia content representation with well-defined structure.

In Semantic Web, the retrieved information has a well-defined meaning. It enables people to create data stores on the web, build vocabularies and write rules for handling data. Semantic web aims to present web data that it is understood by machines to do aggregating and searching the information in web without human operator. It adds meta-data to the existing documents for extending those documents into well defined information. This extension enables the web to be processed automatically by machines and used by humans.

Ontology is a key enabling technology for semantic web. It is used as a main component for representing the knowledge in semantic web. Ontology is defined as "a formal explicit specification of shared conceptualization" [22]. It consists of concepts, relationship between concepts, properties of concepts and instances. Some of the ontology languages are RDF, OWL, SPARQL and SKOS.

With the use of proper representation user can effectively retrieve the multimedia content for semantic web. But, it needs additional mechanisms to reduce the semantic gap between the objects. Lexical libraries like WordNet, ConceptNet, FrameNet, VerbNet, Propbanksup, etc. can be used for expanding the semantics of descriptions associated with the objects. For visual objects, low level features such as size, colour, texture, file types, etc. can be used for finding the appropriate objects based on the user requirements. In addition, various techniques are used for retrieving the multimedia content. This paper provides the detailed information of the existing retrieval techniques of text, image, audio and video.

Section 2 describes about text, audio, video and audio retrieval system. Section 3 compares all the multimedia retrieval system. Finally, section 4 concludes the work with future enhancements.

II.MULTIMEDIA RETRIEVAL

A. Text Retrieval

The traditional keyword-based search uses annotations of content to find the relevance's. The classic technique Vector-Space Model (VSM) is a combination of annotation weighting and ranking algorithm [6]. It uses tf-idf model, term frequency (frequency of a term in document) and inverse document frequency (frequency of a term across the collection of documents) for giving weight to the terms. Based on the relevance of information to the user queries, the results are ranked. The problem with KeyWord based search is complexity between words and concepts due to synonymy (different words denote same content) or homonymy (same word denote different concepts).

,

ISBN: 978-81-909042-2-3 ©2012 IEEE

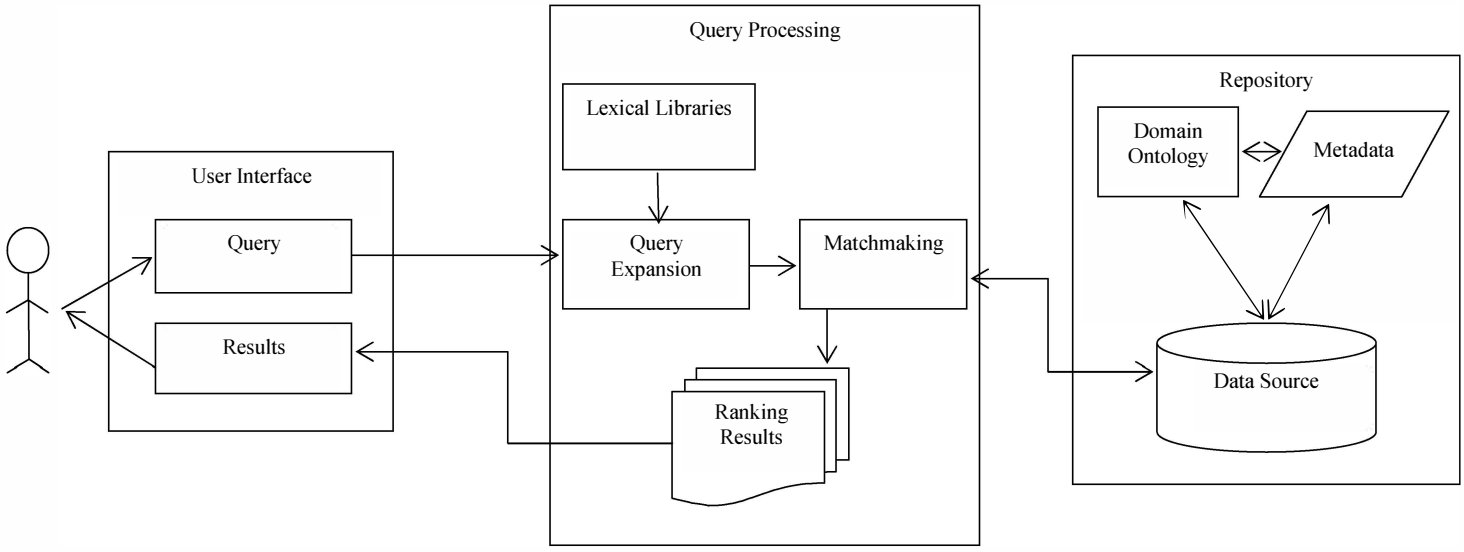

Fig.1 Text Retrieval System

But in semantic web, the search is not only based on the keyword and also based on the meaning of keywords (semantics). For semantic search, Castells, Miriam, and Vallet proposed an adaptation of vector-space model for ontology based information retrieval model, where ontology based knowledge base constructed from domain concept, topic and document. The additional information such as type, structure, relations, classification, and rules about the concepts which are referenced in the documents, are represented in ontology based knowledge base which returns the instances against the user queries. The documents are semi-automatically annotated with the instances. They assigned to a weight that reflects how relevant the instance is considered for the document semantic. The weights are computed automatically by an adaptation of tf-idf algorithm based on the occurrence of instances in documents [6]. Fig.1 shows the text retrieval system based on semantics.

Mustafa, Khan, and Latif proposed a method of searching in existing metadata by matching the RDF triples (subject, object and predicate) instead of keywords in [1]. Predicate describes the relationship between subject and object. For retrieving the data related to user queries, the relevance between document metadata and RDF triples measured by concepts similarity, relations similarity and RDF triples similarity. Jun Zhai and Kaitao Zhou presented a retrieval system using SPARQL query language [2]. The SPARQL query language for RDF used to retrieve data according to the relations of synonym, inheritance and part-of relations between the concepts from the large data sets. A new structure of ontology which has 6-tuples (atomic term, complex terms, instances, instances description, attribute assignment and axioms) has been proposed for semantic information retrieval in [3]. The authors expressed the structure of the ontology using Description Logic (DL).

The fuzzy ontologies are efficient in text and multimedia object representation and retrieval [5]. By combining fuzzy domain ontology with fuzzy linguistic domain variable ontology, a fuzzy ontology framework has been developed in [4]. This framework is the extension of RDF data model. It includes concepts, properties of concepts and values of properties. For facilitating information retrieval at semantic level, the relations between fuzzy concepts defined in fuzzy linguistic domain variable ontology.

B. Image Retrieval

Image retrieval goes through three kinds of retrieval techniques including text-based retrieval, content-based retrieval and semantic based retrieval [7] [13]. Most of the text-based image retrieval system requires the images to be annotated manually [8].

In text-based retrieval, images retrieved through matching the user queries and annotated keywords. The content-based image retrieval system uses low-level visual features like color, texture, shape and location, etc. for retrieving the images. These low level features are extracted from images automatically. The semantic-based image retrieval technique is the combination of low level features and semantic of the images. Initially, Yang, Ming Dong and Farshad Fotouhi proposed image retrieval system based on Natural language Processing (NLP) approach [9].

This processing model integrates WordNet, online lexical information system and low level visual features. Lisa Fan and Botang Li proposed a hybrid model which comprises ontology reasoner and Bayesian Network (BN) [10]. The ontology reasoner retrieves all the neighboring information of the input keyword and the BN model used to calculate the relevance between input keyword and neighboring information. So this model not only retrieves the target image but also retrieves the images having neighboring information about the keyword.

ISBN: 978-81-909042-2-3 ©2012 IEEE

Fig.2 shows the image retrieval with fuzzy set.

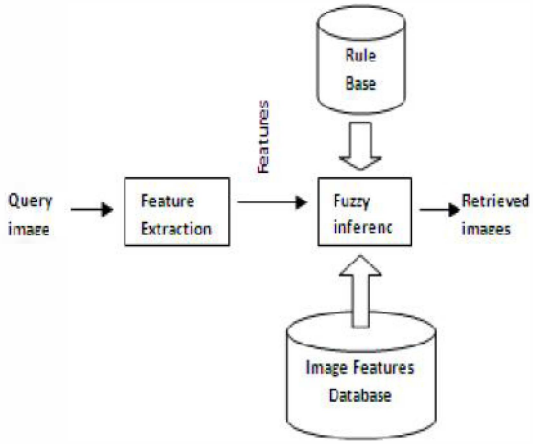

Fig.2 Fuzzy System for Image Retrieval (as given by [23])

With the use of multi-modality ontology a semantic retrieval system has been developed by Wang, Chia and Song Liu in [11]. Three sub-ontologies such as Domain Ontology, Textual Description Ontology and Visual Description Ontology constitute multi-modality ontology. Later, the new ranking mechanism based on semantic similarity between the concepts used to improve the results. Chang and Huang proposed the method of partitioning each image in the different major categories into suitable regions with the help of novel fuzzy segmentation algorithm [8]. Based on the Discrete Cosine Transform (DCT) coefficients images classified into different major categories. The regions derived from images in each individual major category are classified into different sub-categories based on low level features. Then the image search carried through these sub-categories. Fuzzy Domain Ontology (FDO) introduced in [12]. FDO has been constructed by adding fuzzy membership to the concepts and relationship between them in domain ontology, which is used to describe the high-level semantic features. In addition, the Classifier Ontology maps the low-level visual features to high-level semantic features and semantic ontology describes semantics of images in the image database.

Liua, Shaob and Liua [7] presented an image retrieval technique which accepts image as input and will construct the image ontology with the help of SIFT (Scale Invariant Feature Transform) features. SIFT used to extract the low level features of the image. The retrieval technique will calculate the feature similarities between input image and images in ontology. It also considers the similarity of semantic features and concepts.

C. Video Retrieval

Mostly video files annotated by a large number of tags in archives. So, in text-based video retrieval users may face difficulties to find the exact video. Content-based video retrieval is used to find the videos efficiently for users [14]. In this method, low-level visual feature extraction, shot detection and object recognition are the most important phases [16].

Pictorially enriched ontologies are based both on visual and linguistic concepts as described in [15], where visual features extracted from MPEG videos. Annotation is performed automatically on selected video clips associating occurrences of events or entities to high level concepts, by checking their similarity to visual concepts that are hierarchically linked to higher level semantics. Based on four dimensions such as Action, Location, Time and Shooting technique events are modeled as in [14]. Combination of low-level visual features called as semantic patterns, which is used to characterize concepts in the above four dimensions. Video is transformed into a 20-Dimensional multi stream by sequentially aggregating low-level visual features from all shots. The video ontology further refined by extracting new semantic patterns from subspaces of videos which cannot be retrieved by previously extracted patterns. The authors also noted complex events can be retrieved by combining semantic patterns in different dimensions.

Ying Dai introduced semantic tolerance relation model in [17]. In this model, each image/video is represented by its semantics and low level features. It defines the semantic categories of image/key-frame which represent the segment of video, together with the tolerance degrees between them. In this model, the semantics of each image/key-frame is represented by the class-weighted vector. These weights assigned based on classification result to the defined classes. The Bayesian Classifier is used to classify the images in this model. Fig.3 shows video retrieval system based on Shot Selection from archives and Object Recognition.

ISBN: 978-81-909042-2-3 ©2012 IEEE

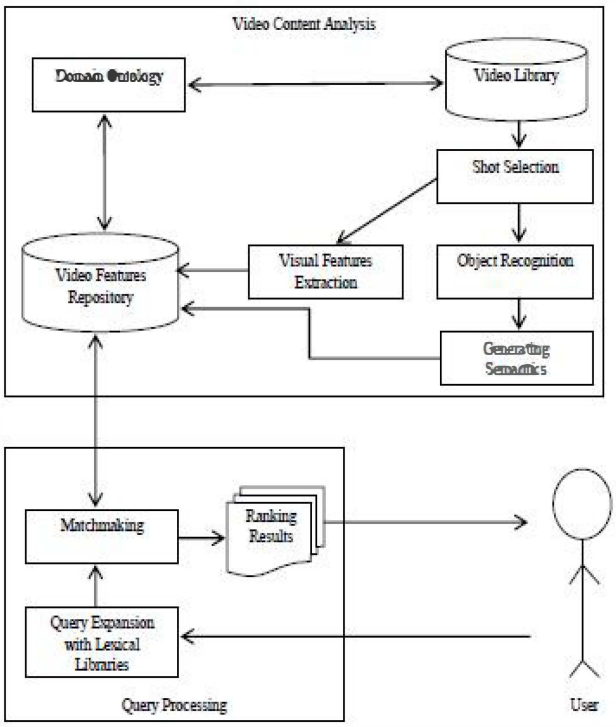

Fig.3 Video Retrieval System

Lamberto Ballan, Marco Bertini, Alberto Del Bimbo and Giuseppe Serra proposed a retrieval system based on ontologies and semantic concept classifiers [18]. Using WordNet, the semantic linguistic relations between the concepts is determined for defining the ontology schema. Then the concept detectors are linked to the corresponding concepts in ontology. The rules defined in Semantic Web Rules Language (SWRL) automatically, learn the knowledge embedded in the ontology.

D.Audio Retrieval

In this digital media era, audio plays significant role among the users because of its expressiveness. Most of the audio information retrieval system compares the audio domain with the features extracted from the audio signals.

Latifur Khan and Dennis McLeod proposed a customized selection/retrieval of audio information with the use of ontologies in [19]. This model consists of several tasks such as Segmentation, Meta data acquisition, Selection, Scheduling, etc. The segmentation phase used to identify the boundaries of audio objects, which is defined by five tuple such as Identifier, Start time, End time, Description and Audio data. The accuracy of audio information retrieval is based on the generation of metadata. Both fully automated content extraction and selected content extraction approaches have been used for metadata acquisition. Metadata has the description of audio with boundary. Based on the user's query the audio information will be retrieved.

Query-by-example audio retrieval system based on semantic similarity proposed in [20]. The user provides an audio example instead of semantic description and the system returns audio content that is similar to the query. Unordered set of feature vectors are extracted from the audio signals to represent the audio tracks. There are two kinds of approaches used in this audio retrieval system. First one is query-by-acoustic example (QBAE) which retrieves the audio acoustically similar to the query. Second approach is query-by-semantic example which retrieves the audio based on the semantic information. An integrated system introduced in [21], which can be used for text based retrieval, content-based query-by-example and automatic annotation. It uses hybrid network that linking the sounds through a measure of perceptual similarity and linking the semantic tags through user-provided weights or lexical libraries which helpful for improving the retrieval of audio information.

III. COMPARISON OF RETRIEVAL SYSTEM

For text retrieval, vector space model adopted with semi automatic annotation and instead of using the keyword-indices, the ontology knowledge-base used. It enhances precision and recall rate in retrieval. RDF Triples approach mainly focus on concept and semantic similarities. It is mainly used for unstructured data. The SPARQL query language used to retrieve the data based on the semantic relationship (synonym, part of, kind of) between the concepts in ontology. New ontology structure which has six tuple having the word association at semantic level focused on enhancing the efficiency of a retrieval system. The fuzzy ontology approach ensures the reusability concept of ontologies. Table 1 summarizes all the retrieval approaches of multimedia.

For image retrieval, content based approach with semantic features widely used. The natural language processing approach integrated with lexical libraries and low-level features used for the retrieval of images. It provides the semantic base for ontology creation of the image databases and improves retrieval accuracy. The hybrid model which combines ontology and Bayesian Network aim to improve the quality of image retrieval. But in this model, the reasoning and ranking of the concepts has to be improved for better recall rate. Multi-Modality ontology approach provides scalability of retrieval system and ranking mechanism has to be improved for large domain concepts. The region-based approach narrows the semantic gap when retrieving the images and overcomes the shortcomings of image-image matching.

Image semantics and SIFT features are used to improve the accuracy of retrieval. The fuzzy domain ontology for image narrows the search range and reduces the useless information retrieval. This approach improves retrieval system efficiency and accuracy with the use of fuzzy set.

In video retrieval, shot selection and object recognition plays the most important role. Automatic shot selection and annotation used in pictorially enhanced ontology to perform automatic annotation of video for retrieval. In video ontology approach, complex events can be detected by combining the semantic patterns in different dimensions. Semantic-based tolerance image representation improves the effectiveness of

ISBN: 978-81-909042-2-3 ©2012 IEEE

retrieval through automatic annotation of image/video according to its semantics. In rule learning approach, rule based annotation provides better precision and recall.

Audio retrieval can be achieved in a high recall and precision with the use of ontologies. Metadata plays major role in retrieving the audio effectively. Though Query-by-semantic example system having low precision and recall it shows better result than Query-by-acoustic example system. The hybrid model linking sounds through perceptual similarity and linking semantics through lexical libraries and improves the effectiveness of the audio retrieval system.

IV. CONCLUSION

This paper discussed about various retrieval techniques used for multimedia using ontology for the semantic web. It is focused mainly towards text, image, audio and video retrieval. It is observed that semantic-based search improves the performance of the retrieval of multimedia in all the ways. This paper includes important techniques of multimedia retrieval among more techniques. However, it provides retrieval mechanisms from the past to present which are crucial. Our research focuses mainly on the effective video retrieval using ontology for the semantic web. It will emphatically enhance existing video retrieval mechanisms.

Table. I Various Retrieval Approaches of Multimedia

Category	Approaches	Comments
Text	1.Semi Automatic Annotation with the Adaptation of VSM	Enhances precision and recall when VSM used with ontology knowledge-base
	2.Measurement of RDF Triples	It can retrieve unstructured data
	3.Using Description Logic (DL)	Enhances efficiency of the retrieval system
	4.Fuzzy Domain Ontology	Ensures reusability concept of ontologies
Image	1.Natural Language Processing (NLP) Approach	Improves retrieval accuracy
	2.Hybrid Network Model	Effective with small image library
	3.Multi-Modality Ontology	Provides scalability
	4.Based on Discrete Cosine Transform Coefficients	Used in image-image matching
	5.Fuzzy Domain Ontology	Fuzzy set improves the image retrieval accuracy
	6.Use of Scale Invariant Transform Features (SIFT)	Improves retrieval accuracy
Video	1.Pictorially Enriched Ontologies	Automatic video annotation with higher level concepts
	2.Semantic Patterns	Complex events in video can be retrieved
	3.Semantic Tolerance Relation Model	Automatic representation of image/video based on semantics
	4.Rule Learning	Annotation of video based on rules
Audio	1. Audio Structuring/Personalized Retrieval	Better precision and recall compared to keyword-based search
	2.Using Semantic Similarity	Improves audio retrieval both qualitatively and quantitatively
	3.Unifying Semantic and Content-based approach	Hybrid model ensures the effectiveness of the retrieval system

ISBN: 978-81-909042-2-3 ©2012 IEEE

ACKNOWLEDGMENT

We thank every contributor of this comparative study work. We thank all the authors who provided valuable information through their research papers, other publications and the university which made available the required data.

REFERENCES

[1] Jibran Mustafa, Sharifullah Khan and Khalid Latif, "Ontology Based Semantic Information Retrieval," *4th International IEEE Conference on Intelligent Systems, IEEE*, 2008.

[2] Jun Zhai and Kaitao Zhou, "Semantic Retrieval for Sports Information Based on Ontology and SPARQL," *International Conference of Information Science and Management Engineering, IEEE*, 2010.

[3] Wei Wang, Hongwei Wang and James N. K. Liu, "A New Structure of Ontology Model for Semantic Information Retrieval," IEEE, 2009.

[4] Jun Zhai, Yan Cao, and Yan Chen, "Semantic Information Retrieval Based on Fuzzy Ontology for Intelligent Transportation Systems," *IEEE*, 2008.

[5] Parry D, "A Fuzzy Ontology for Medical Document Retrieval," *Proceedings of the Second Workshop on Australasian Information Security, Data Mining and Web Intelligence, and Software Internationalization*, Dunedin, New Zealand, 2004, pp. 121-126.

[6] Pablo Castells, Miriam Ferna´ndez, and David Vallet, "An Adaptation of the Vector-Space Model for Ontology-Based Information Retrieval," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 19, No. 2, February 2007.

[7] Xuejun Liua, Zhenfeng Shaob and Jun Liua, "Ontology-based image retrieval with SIFT features," First International Conference on Pervasive Computing, Signal Processing and Applications, 2010.

[8] Tsun-Wei Chang and Yo-Ping Huang, "An Ontology Oriented Region-Based Image Retrieval Strategy," *IEEE*, 2008.

[9] Changbo Yang, Ming Dong and Farshad Fotouhi, "Learning the semantics in image retrieval – A Natural Language Processing Approach," *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW'04)*.

[10] Lisa Fan and Botang Li, "A Hybrid Model of Image Retrieval Based on Ontology Technology and Probabilistic Ranking," *Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence (WI 2006 Main Conference Proceedings) (WI'06)*.

[11] Huan Wang, Liang-Tien Chia and Song Liu, "Semantic Retrieval with Enhanced Matchmaking and Multi-Modality Ontology," *IEEE*, 2007.

[12] Yuemei Ren and Xianyi Cheng, "Semantic-based Image Retrieval using Fuzzy Domain Ontology," *Second International Symposium on Intelligent Information Technology Application, IEEE*, 2008.

[13] Yihun Alemu, Jong-bin Koh, Muhammed Ikram and Dong-Kyoo Kim, "Image Retrieval in Multimedia Databases: A Survey," *2009 Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IEEE*, 2009.

[14] Kimiaki Shirahama, Kazuyuki Otaka and Kuniaki Uehara, "Content-Based Video Retrieval Using Video Ontology," *Ninth IEEE International Symposium on Multimedia 2007-Workshops, IEEE*, 2007.

[15] Marco Bertini, Alberto Del Bimbo and Carlo Torniai, "Enhanced Ontologies for Video Annotation and Retrieval," *ACM*, 2005.

[16] Liang Bai, Songyang Lao, Gareth J. F. Jones and Alan F. Smeaton, "Video Semantic Content Analysis based on Ontology," *International Machine Vision and Image Processing Conference, IEEE*, 2007.

[17] Ying Dai, "Semantic Tolerance-based Image Representation for Large image/video Retrieval," *Third International IEEE Conference on Signal-Image Technologies and Internet-Based System, IEEE*, 2008.

[18] Lamberto Ballan, Marco Bertini, Alberto Del Bimbo and Giuseppe Serra, "Video Annotation and Retrieval Using Ontologies and Rule Learning," *IEEE*, 2010.

[19] Latifur Khan and Dennis McLeod, "Audio Structuring and Personalized Retrieval Using Ontologies," *IEEE*, 2000.

[20] Luke Barrington, Antoni Chan, Douglas Turnbull and Gert Lanckriet, "Audio Information Retrieval using Semantic Similarity," *IEEE*, 2007.

[21] Gordon Wichern, Harvey Thornburg and Andreas Spanias, "Unifying Semantic and Content-based Approaches for Retrieval of Environmental Sounds," *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics October 18-21, 2009.*

[22] R. Studer, Richard Benjamins and Dieter Fensel, "Knowledge Engineering: Principles and Methods," *Data and Knowledge Engineering, vol. 25, no. 1-2, pp. 161-197, 1998.*

[23] Abolfazl L., M. Shahram Moin, and Kambiz Badie: "Semantic–Based Image Retrieval: A Fuzzy Modelling Approach", *IEEE* (2008).

ISBN: 978-81-909042-2-3 ©2012 IEEE