Learning Context Determination based on Relevance Feedback for Memory Recall

Bela Joglekar
Ph.D. Research Scholar
Bharti Vidyapeeth University
Pune, India

Parag Kulkarni
EKALAT Research Lab
Pune, India

ABSTRACT
The brain continuously receives information through various stimuli and processes this information through cognition. Patients suffering from short term memory loss (e.g. in Parkinson's disease, seizures or epileptic attacks), lose a short episode of memory. The originality of the research lies in retrieving back any lost cognition due to damage/disease by presenting context-specific sequence of images to the subject under study. The approach proposes mapping the lost memory episode to a corresponding set of stored ranked images which can help regain memory loss. A framework is presented for implementation of context determination through relevance feedback. A comprehensive overview and analysis of existing techniques is also presented for context based retrieval of images.

General Terms
Memory loss, Memory recall, Image Retrieval, Machine Learning

Keywords
Context, Content, Relevance Feedback, Image Ranking

1. INTRODUCTION
The brain continuously receives information through various stimuli. It perceives this information. This information is converted into an electrical signal by encoding it in real time. This encoding is done by the neuron. The novelty of the study involves testing whether the original signal that captured some images/situation, helps in retrieving back the lost cognitive function due to damage/disease (e.g. Parkinson's disease, seizures or epileptic attacks) when it is fed at a different time instant, in the form of sequential ranked images.

Here reproduction of image is done on the basis of context. A set of images that faithfully represent a real time event will be identified. These images will be stored in a database. They are representative real time images which will help a person to regain his memory loss. A scenario/episode lost from memory can be built up in front of a person by deriving context from these representative images. Thus memory recall can be achieved by mapping images of the lost event to the corresponding most likely images retrieved from the database. The event will be build up progressively by sequential relevance feedback of images representing the real time scenario. The environment or context is used as relevance feedback for restoring any lost cognitive function. Accuracy is achieved by eliminating redundancies through feature selection process in the most likely images. Many state-of-the-art approaches are reviewed.

The paper is organized as follows: Section 1 gives the motivation and introduction for proposed study. A comprehensive overview of existing techniques is reviewed in section 2. Section 3 gives a detailed framework of the various modules and processes involved for learning context determination through relevance feedback. The discussions and future work is discussed in section 4.

2. CONTEXT DETERMINATION FOR IMAGE RETRIEVAL
Context can be used in pattern recognition [1] where any entity can be viewed differently in different contexts. The properties of the entity change in accordance with the context. Context can help in perceiving what is actually not there. Literature survey on context with respect to image classification and text recognition uses suitable approximations for probability distribution. Techniques for using context in text recognition are classified into dictionary look-up methods, Markov and probability distribution approximation methods, and hybrid methods [1]. Content-Based Image Retrieval (CBIR) computes numerical values for the features using image processing algorithms. Features may be obtained from a single extractor or several feature extractors. Optimization of selection plays an important role as features may be obtained from combination of several extractors also. This may cause irrelevancy and redundancy problem. This problem can be solved by sorting done according to a similarity criteria employed. So the output is a set of images ranked by their similarity to the current case. Evaluation is done iteratively by retrieving a large number of relevant images in the initial positions of the output. In order-based ranking evaluation, a score is assigned to each element (images in this case) according to its relevance and position in the ranking [2]. Ranking of images is done on the basis of single valued function based on genetic algorithm approach. The ranking uses fitness functions for graphical curves and it is shown that the genetic feature selection method improves ranking quality of image retrieval and performs better than wrapper and filter approach. Dimensionality reduction of the images is done by two methods, namely, feature extraction and feature selection. Feature extraction transforms the image to a lower feature space which makes the image highly uncorrelated with the original image. This may lose original representation of the image when its dimensions are changed. Feature selection methods outperform feature extraction methods as they give the most relevant feature subset corresponding to the original image according to a user defined criterion. Hence, for context determination, feature selection algorithms are preferred for image retrieval. Major advantages of image retrieval based on content include finding images that have not been categorized and using automated cataloguing [3]. Multiple features may be extracted from multiple images and may form components of the original image. Either same feature from different query images or different features from different query images are considered. A multihistogram intersection technique is employed for multiple query images [3]. For each query

image, a corresponding color and texture histogram is formed. Intersection in multihistograms for the queries gives a similarity ranking measure between the query image and corresponding image in the database, which improves retrieval precision.

Constrained Similarity Measure (CSM) learns similarity between images for content based image retrieval (CBIR) [4]. Boundary clustering between images is done by two techniques of machine learning, namely support vector machine (SVM) and AdaBoost. Ranking of images inside the boundary is done by measuring Euclidean distance to the query. The similarity measure is achieved through learning a boundary that separates positive and negative examples. Positive examples correspond to the most likely dimensions which help in image retrieval by relevance feedback. Negative examples correspond to the minimum likely dimensions in the process of image retrieval. Different ways of learning boundary are discussed. One way to learn boundary is by using Bayesian Classifiers to estimate model parameters but which requires a large number of input positive and negative examples. This disadvantage is overcome by using large margin classifiers such as AdaBoost and Support vector Machine (SVM) which are nonparametric and require comparatively less number of positive and negative examples in learning the boundary. The system can be provided with preselected positive and negative examples or adaptive relevance feedback can be used. In this, the learned boundary is adapted to each query and the system builds up the boundary iteratively through the user's relevance feedback [4]. Relevance feedback for interactive image retrieval adjusts weights of different levels according to a model representing certain prominent image contents.

A novel probabilistic framework is proposed to process multiple sample queries in content based image retrieval (CBIR). A comprehensive review is presented for existing relevance feedback methods [5]. The feature spaces of similar images can be combined after examining the images through probabilistic distance normalization (PN) and gaussian normalization (GN). Here the results presented clearly show that the proposed probabilistic algorithm outperforms other relevance feedback models by giving a higher retrieval precision. The MindReader system proposed in [6], computes the parameters (weights) of a given distance function. These weights will compute the distance from the query thus representing dissimilarity, and will iteratively give best query point when user feedback is given as input at each iteration. In the approach stated in [7], each image is considered as a vector of weights. Each weight represents the ranking/significance of a feature within the vector as well as its significance with respect to the entire database. The queries are updated by removing the redundancies and keeping the most relevant weights, while ignoring the irrelevant ones.

G. Ciocca, R. Schettini proposed a simple algorithm that computes a new query point based on standard deviation from a set of relevant features used. The query point is computed based on minimum deviation which corresponds to maximum correlation. This new query point represents the set of relevant images of interest to the user [8]. But the disadvantage in this approach includes treating the feature space globally instead of locally and ignoring the dependencies between features.

One of the tools for relevance feedback mechanism includes using Support Vector Machine (SVM) which gives a clear discrimination between relevant and non-relevant examples [9]. SVM approach has a major disadvantage which includes critical parameter setting according to different databases and different descriptors. Another disadvantage of SVM is small sample size. But this issue has been addressed in [9]. Another approach to CBIR presented in [10] includes self-organizing map (SOM) method, in which each SOM classifies each image into one or two dimensional grid of cells. After parallel iteration of all SOM's, a score is assigned to each cell resulting in similar descriptor values appearing close to the cell, and dissimilar values appearing distant to the map. The resulting maps are low pass filtered to propagate this information to the neighboring cells. So this method yields a new image ranking strategy, where score assigned to each image is the sum of scores assigned to the cell it belongs to in each of the SOM's.

The cluster-based unsupervised learning strategy method [11] proposes applying a clustering algorithm to a collection of images close to the query. The images having minimum distance to the query are included as the most relevant ones, and clustered together thus providing a similarity measure.

In [12], an approach based on nearest neighbor paradigm is proposed, where image ranking is done according to a relevance score depending on the nearest-neighbor distances. This helps in representing images in dissimilar spaces, as the images far to the neighbor are low in similarity. Here images are ranked according to their dissimilarity to a set of relevant images.

A relevance feedback algorithm based on fuzzy sets [13] has defined a fuzzy set such that, degree of similarity of each image in the database helps in determining positive and negative selected outputs. At each iteration, a set of parameters is modified making the process progressively selective.

Content Based Image Retrieval (CBIR) can be made more efficient by integrating region based image retrieval (RBIR) with relevance feedback [14]. Here a region in an image is considered as the most similar region when a weighting scheme assigns maximum weight to the pixels closer to the query. Group Biased Discriminant Analysis (GBDA) is used as a measure for re-estimating both similarities and corresponding feature weights between images. Positive feedbacks are grouped by providing heuristic pre-clusters as a reference measure. The future work talks about improving retrieval performance for studying importance of region on the basis of user feedback.

Rough set theory is a technique that helps to reach upto a reduct of an image or data from a set of images or data which reduces redundancies and hence helps in reaching an accurate inference of the data set under study. Information system of Rough Set is an extension of set theory for the study of intelligent systems that are characterized by incomplete information to classify imprecise, uncertain or incomplete information, or knowledge that is expressed in terms of data. Rough set theory discovers data dependencies and helps to reach to a reduct out of a large amount of data. Feature selection algorithms like RSReduct and MRSReduct are discussed in [23] [24], which implement rough set theory. The reduct outcome of rough set theory discovers data dependencies and removes any irrelevancy and redundancies leading to a more accurate and reduced dataset [25].

Reduction of attributes is achieved such that the reduced set provides the same degree of information prediction of the decision attribute as the original. So prediction error by using rough set theory is minimum as decision attribute is a true reduct from all the original features of the data/image under consideration.

Multi-sensor data fusion is a technique that integrates complementary information from multiple images. It combines data from multiple sensors and related information from associated databases to achieve improved accuracies and more specific inferences instead of using a single sensor. Applications of sensor fusion involve fuzzy-based multi-sensor data fusion classifiers, which has the advantage of flexibility of integrating multi-sensor/contextual and prior information. Adaptive weighted estimate fusion algorithm is tested for fault tolerance [26]. Wavelet based image fusion hybrid architecture (pixel based for high image contrast and region based for high edge /boundary details) is implemented [27]. Contextual information is studied in terms of phase level and feature level context. Incomplete data fusion due to fluctuant sensors is dealt with by deleting information from redundant sensors using rough set theory [28].

The fundamental idea behind relevance feedback is to present the user with candidate images and allow the user to decide relevant and irrelevant images by modifying either the feature space, parameter space or classification space to separate positive and negative examples [29]. Representative test set of images include personal photos, COREL database of images, web images or videos. The query point is moved iteratively towards the positive examples and away from the negative examples. A disadvantage in relevance feedback is the size of the sample, as this technique requires a small training set. Optimization based parameter updating method gives higher accuracy than heuristic method [29]. For video retrieval, keyframe based retrieval is used which splits the video in a small set of frames. Boosting and k-means method of learning is compared with the vector-space model. It is shown that k-means performs best in handling variety of queries, whereas, boosting performed well as a similarity measure for finding similarity compositions in various video frames [29].

3. PROPOSED FRAMEWORK

In this paper, we have proposed a framework that can be used to build context. The study will deal with classification of sequence of images based on context. A set of images that can best describe a real time event will be identified. These images will be stored in a database. They are the most likely images which will build a scenario corresponding to the lost real time episode. They will be presented to the subject under observation in a sequence. So short term memory loss can be overcome by deriving context from the above most likely images. Thus, mapping the lost memory episode to a corresponding set of stored ranked images can help regain memory loss. The environment or context is used for feedback and restoring any lost cognitive function. Accuracy is achieved by eliminating redundancies through feature selection process in the most likely images. The overall process involved can be decomposed into two phases, first is training/supervised phase and another is test phase, as shown in figure 1 below.

First phase involves classification of images into clusters. Each cluster will consist of the most correlated images. The clusters are formed using one of the various learning algorithms like AdaBoost and SVM [4] [16]. Then the training set is formed out of various input images. This will act as an input to the feature extraction stage. For feature extraction, local features, global features, texture will be considered. In order to make this process more efficient, dimensions of the images is reduced in next step so as to reduce the size of the data to be handled. The most relevant feature subset is then obtained using feature selection process.

Second phase is the test phase. It involves processing of input query and providing output to the user. The concept of relevance feedback [4] [17] is used here for refining the query as per user's perspective. This is an iterative process that will keep on refining the query [7] [18]. The final query is obtained through a similarity measure that maps the query features with the image features stored in the database of training set. Mapped features are ranked using one of the ranking algorithms so as to build a story/episode out of the selected features through the ranked images. This will correspond to the short event lost from memory.

3.1 Feature Extraction

For extraction of relevant visual content, visual codebook [15] is used. Clustering of images in visual codebook can be done by

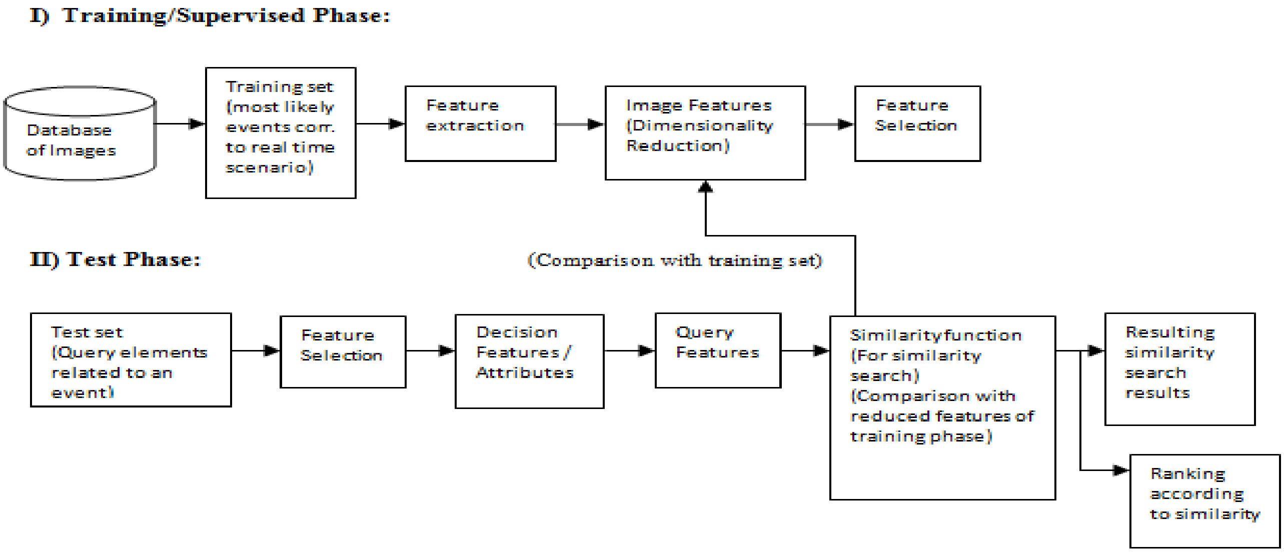

Figure 1. Block diagram of proposed framework

static clustering and dynamic clustering. In static clustering, the index and a similarity for comparing images are directly provided. Dynamic clustering involves adaptive k-means clustering which allows user to adapt to the image database. This process is scalable to huge datasets also. The figure 2 given below represents the steps involved in construction of training set. Kernel function 'k' is used for similarity measure between histogram [3]. This is done by calculating Euclidian distance.

3.2 Query Processing

Typically, input query will comprise of frames or information related to the event that is to be recalled. This can be in the form of features or an image itself. If it is in the form of image, then the features are extracted using feature extraction algorithm [3] [19].

The user can keep on refining his query until optimum match to the image in the data bank is achieved. The query point after this will be modified and next match will be searched iteratively. Algorithms like SVM and AdaBoost are used to readjust the parameters provided in the query.

3.3 Image Ranking

There are various kinds of ranking algorithms like unsupervised ranking algorithm, ranking algorithm for implicit feedback, ranking using SVM technique, ranking algorithm using dynamic clustering, etc. [2] [20] [21]. Here images will be ranked in descending order from the most likely matches scaling down to the least ones. The most optimum matches will be added to the database for matching with any further similar queries. So image ranking will provide the user with maximum positive examples in the first few iterations.

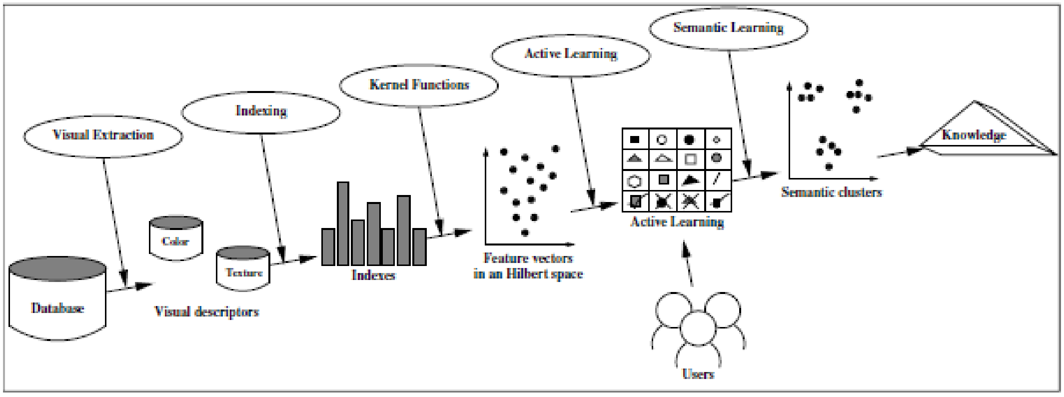

Figure 2. Training set construction

3.4 Visual features

Visual features are characterized by primitive, logical and abstract features. Primitive features define colour or shape, logical features define the identity of objects (e.g. source from where taken) and abstract features define the significance of image presented [4]. Feature representation of an image causes loss of information as the image is represented only as a set of prominent features which define it. This is called as semantic gap between the original image and its feature representation. Sensory gap describes gap between the actual structure and its digitally represented image.

3.4.1 Color

Red, green, blue (RGB) color space is used for images taken with the same surrounding condition each time. So this color space is limited in use. Hue, saturation, value (HSV) spaces are more frequently used. So there is a difference between RGB color space and the hue, saturation, pixel values of each color. This means that differences in the color space are similar to the differences between colors that humans perceive.

3.4.2 Texture

Texture measures convey greater detail than color measures. Wavelets and Gabor filters are some of the common measures for capturing texture of images. Gabor filters perform better according to the properties of human visual cortex for edge detection. These texture measures catch the characteristics of an image or parts of an image with respect to changes in directions and scale.

4. CONCLUSION

A framework is proposed for context determination based on relevance feedback. The approach involves mapping a lost memory episode to a set of most likely stored ranked images which build up the lost event and thus help regain memory lost due to damage in cognition (e.g. Parkinson's disease, seizures or epileptic attacks). This is done by context determination through relevance feedback which will progressively build up the lost event in the form of sequential image presentation to the affected user. Image ranking will assist in evaluating the results within a short time span. For retrieval and ranking precision, and to reach an optimum relevance feedback algorithm, many approaches are reviewed. The process implementation will involve a combination of the approaches stated which will result in maximum relevance and minimum redundancy to regain lost cognition. Positive and negative examples can be effectively classified for incremental region building by including more number of sample ranked images. The selection of nearest similar neighbor can be done by taking a heuristic pre-clustered region reference which represents real context. Testing with generic samples first and then updating data with real time samples will lead to more feasibility in integrating relevance feedback with memory recall. The outcome of the proposed framework is a novel application in learning context determination through relevance feedback which can compensate for any short term memory loss.

5. REFERENCES

[1] Godfried T. Toussaint, 'The use of Context in Pattern Recognition', Pattern Recognition, Vol. 10, pp 189-204

[2] Sergio Francisco da Silva, Marcela Xavier Ribeiro, Joao do E.S.Batista Neto, Caetano Traina-Jr, Agma J.M.Traina, 'Improving the ranking quality of medical image retrieval using a genetic feature selection method', Decision Support Syst., Journal of Science Direct, Elsevier 2011 10.1016/j.dss.2011.01.015

[3] Jinshan Tang and Scott Acton, Department of Electrical and Computer Engineering, University of Visginia, Charlottesville, VA 22904-4743, USA 'An Image Retrieval Algorithm using Multiple Query Images', 0-7803-7946-2/03/$17.00 02003 IEEE

[4] Guo-Dong Guo, Anil K. Jain, *Fellow*, *IEEE*, Wei-Ying Ma, *Member*, *IEEE*, and Hong-Jiang Zhang, *Senior Member*, *IEEE* 'Learning Similarity Measure for Natural Image Retrieval With Relevance Feedback', IEEE Transactions On Neural Networks, Vol. 13, No. 4, July 2002

[5] Miguel Arevalillo-Herráeza, FrancescJ.Ferria, JuanDomingob, ‘A naive relevance feedback model for content-based image retrieval using multiple similarity measures’, Journal of Science Direct, Elsevier 10.1016/j.patcog.2009.08.010 pp: 619 – 629

[6] Y. Ishikawa, R. Subramanya, C. Faloutsos, ‘Mindreader: querying databases through multiple examples’, in: Proceedings of the 24th International Conference on Very Large Data Bases, VLDB, New York, USA, 1998, pp. 433–438.

[7] Y. Rui, S. Huang, M. Ortega, S. Mehrotra, ‘Relevance feedback: a power tool for interactive content-based image retrieval’, IEEE Transaction on Circuits and Video Technology 8 (5) (1998) 644–655

[8] G. Ciocca, R. Schettini, ‘A relevance feedback mechanism for content-based image retrieval’, Information Processing and Management 35 (1) (1999) 605–632.

[9] S. Tong, E. Chang, ‘Support vector machine active learning for image retrieval’, in: ACM Multimedia Conference, ACM Press, Ottawa, Canada, 2001, pp. 107–118.

[10] J. Laaksonen, M. Koskela, E. Oja, PicSOM: ‘self-organizing image retrieval with MPEG-7 content descriptors’, IEEE Transactions on Neural Networks 13 (4) (2002) 841–853.

[11] Y. Chen, J. Wang, R. Krovetz, Clue: ‘cluster-based retrieval of images by unsupervised learning’, IEEE Transactions on Image Processing 14 (8) (2005) 1187–1201.

[12] G. Giacinto, ‘A nearest-neighbor approach to relevance feedback in content based image retrieval’, in: Proceedings of the 6th ACM International Conference on Image and video retrieval (CIVR'07), ACM Press, Amsterdam, The Netherlands, 2007, pp. 456–463.

[13] M. Arevalillo-Herráez, M. Zacarés, X. Benavent, E. de Ves, ‘A relevance feedback CBIR algorithm based on fuzzy sets’, Signal Processing: Image Communication 23 (7) (2008) 490–504.

[14] Wan-Ting Su, Ju-Chin Chen, Jenn-Jier James Lien, ‘Region-based image retrieval system with heuristic pre-clustering relevance feedback’, Journal of Science Direct, Elsevier 10.1016/j.eswa.2009.12.015 pp: 4984–4998

[15] Philippe Henri Gosselin, Matthien Cord, Sylvie Phylipp-Folignet, 2007. “Combining visual dictionary kernel based similarity and learning strategy for image category retrieval”, computer vision and image understanding 110(2008) 403-417.

[16] Tieu, K. and Viola, P. 2004. “Boosting Image Retrieval”, International Journal of Computer Vision 56(1), 17-36.

[17] Y. Rui *et al.*, “A relevance feedback architecture in content-based multimedia information retrieval systems,” in Proc. IEEE Workshop Content- Based Access of Image and Video Libraries, 1997.

[18] Yin, P.Y., Bhanu, B., Chang, K.C., and Dong, A. 2005. ‘Integrating Relevance Feedback Techniques for Image Retrieval Using Reinforcement Learning’, IEEE Transactions on Pattern Analysis and Machine Intelligence 27(10), 1536-1551.

[19] N.H. Bergboer , E.O. Postma, H.J. van den Herik. 2006. “Content based object detection in still images”. Image and vision computing, 24 (2006) 987–1000.

[20] Md. Farooque. “Image indexing and retrieval”. Documentation Research and Training Centre Indian Statistical Institute Bangalore-560059, 2003

[21] Christos Faloutsos . “ImageMap: An Image Indexing Method Based on Spatial Similarity”. Dept. of Computer Science Carnegie Mellon University, 2001

[22] Ji Zhu, “ Multiclass AdaBoost”. Department of Statistics, University of Michigan, Ann Arbor, MI 48109, 2006.

[23] Z. Pawlak, ‘Rough Sets’, International Journal of Computer and Information Sciences, vol 11, pp. 341-356, 1982

[24] Yailé Caballero,Delia Álvarez ,Rafael Bello and María M. García, ‘Feature Selection Algorithms using Rough Set Theory’, Seventh International Conference on Intelligent Systems Design and Applications

[25] Yan Huang, Shulin Chen, ‘An Algorithm of Attribute Reduction Based on Rough Sets’, 2008 International Conference on Computer Science and Software Engineering

[26] Li Shu-qing, Zhang Sheng-xiu, ‘A Congeneric Multi-Sensor Data Fusion Algorithm and Its Fault Tolerance’, IEEE ICCASM 2010

[27] Susmitha Vekkot, Pancham Shukla, ‘A novel architecture for Wavelet Based Image Fusion’ World Academy of Science, Engineering and Technology57 2009

[28] Chen Wu, Xiaohua Hu, Enbin Wang, University of Science and Technology, ‘Combination of Granules, Rough Sets with Evidence Theory and Its Application in Incomplete Data Fusion for Belief Estimation’

[29] Michael S. Lew, Nicu Sebe, Chabane Djearba Lifl, Ramesh Jain, ‘Content-Based Multimedia Information Retrieval: State of the Art and Challenges’ACM Transactions on Multimedia Computing, Communications and Applications, Vol. 2, No. 1, February 2006, Pages 1–19.