ISSN 2395-695X (Print)
ISSN 2395-695X (Online)

International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)

Vol. 2, Special Issue 10, March 2016

AN APPRAISAL OF CONTENT BASED IMAGE RETRIEVAL BY MEANS OF UNLABELLED IMAGES

T. Dharani#1, Dr.I. Laurence Aroquiaraj*2
*Assistant Professor
*Department of Computer Science, Periyar University Salem - 636011, Tamil Nadu, India
#Ph.D Research Scholar
#Department of Computer Science, Periyar University Salem - 636011, Tamil Nadu, India
1dharanimca28@gmail.com
2 laurence.raj@gmail.com

***Abstract*— Creature easily recognizes the information more images than the text. A depiction is worth a thousand words. In our fast life depiction is more useful than the text one. Depictions are the most useful for creature identifying the information. In contemporary epoch CBIR system having diverse image database. That's why we go for content based image retrieval system for image assessing in a proficient way. In CBIR system having some demerits that is annotation of image, feature extraction, dimensionality reduction, unlabelled images, etc. In this effort an appraisal of Content Based Image Retrieval systems based on how treated unlabelled images in various manner. CBIR is a technique which uses satisfied features of image such as color, shape, texture, salient features, semantic, edge, region, etc…to search user required image from titanic image database according to user's interaction in the form of a query image. We working with Content Based Image Retrieval system viz. relevant, irrelevant and unlabelled images for analyzing capable image for different image retrieval process viz. ICA, Semantic, MRF, Ontology etc. To analyzing the efficient imaging for Content Based Image Retrieval, We prepared an appraisal by using principles of Content Based Image Retrieval based on unlabelled images. And also give some future scopes and suggestions for improve the CBIR system using unlabelled images.**

***Keywords*- MRF, ICA, BMMA, CBIR, Unlabelled images.**

I. INTRODUCTION

Pictures are the most common and convenient means of conveying or transmitting information. A picture is worth a thousand words. Pictures concisely convey information about positions, sizes and inter-relationships between objects [1]. An image is an array, or a matrix, of square pixels (picture elements) arranged in columns and rows [2]. Feature means characteristics of object that is feature is a significant piece of information extracted from an image which provides more detailed understanding of the image. Feature extraction is refers that dimensionality reduction of that object. It plays an important role in image processing. Features are classified into three types in image processing, that is low, middle and high. *Low level* features are color, texture and *middle level* feature is shape and *high level* feature is semantic gap of objects [3]. An image retrieval system is a computer system for browsing, searching and retrieving images from a large database of digital images. Most traditional and common methods of image retrieval utilize some method of adding metadata such as captioning, keywords, or descriptions to the images so that retrieval can be performed over the annotation words. But current era of image retrieval method is adding image, video, text and image, ROI of image as a searching query. Various possible applications for CBIR technology have been identified. Some of these are Investigations, Shapes identification, Medical diagnosis, Journalism, advertising Media, Fashion and graphic design, Remote sensing, Trademark databases, Art galleries, museums and archaeology, Architectural and engineering designs, Cartography, Digital Forensics, Radar engineering [5] etc. Image retrieval is classifying mainly two types; they are Text Based Image Retrieval (TBIR) and Content Based Image Retrieval (CBIR). Initially it working based on text of image. Text Based Image Retrieval is having demerits of efficiency, lose of information, more expensive task and time consuming. Overcome these problems by using Content Based Image Retrieval (CBIR) system for image retrieval. CBIR working by using features of the content of the image is known as Content-Based Image Retrieval (CBIR). Nowadays it is classifying many types; they are color based image retrieval, texture based image retrieval, shape based image retrieval, semantic based image retrieval. Thus the above types of image retrievals are come under CBIR types of retrievals. The image retrieval system acts as a classifier to divide the images in the image database into three classes are relevant, irrelevant and unlabelled images. Image annotation means the name or label of an image. This label is more useful for image searching purpose. In this sense, an annotated image can be represented by a feature vector **x**, e.g. a set of image features or Eigen

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)

International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)

Vol. 2, Special Issue 10, March 2016

features, and its label y that neither is relevant or irrelevant nor unlabelled images. It seems that many supervised learning approaches could be employed to approach to this classification problem.

Unfortunately, they are confronted by three main challenges. The first one is that the annotated or labelled training samples are too limited. Generally, the labels are provided by queries and relevance feedbacks, which will not be many. Limited training data would only result in weak classification. The second challenge is the dimensionality of learning, since high dimensional visual data would pose practical difficulties for feature weighting, classification, clustering, and selecting and dimensionality reduction. Limited training data would also prevent effective dimensionality reduction schemes. Therefore, a new learning scheme needed for such a challenging scenario. Considering there are a large number of unlabelled images in the given database, we may use them to boost the weak classifier learned from the limited labelled data, since unlabelled data contain information about the joint distribution over features. The interesting images to the user are only a very small portion of the large image database, in which most images remain unlabelled. Much work regards the problem as a strict three classes classification problem, with equal treatments on positive, negative and unlabelled examples. It is reasonable to assume positive examples to cluster in certain way, but negative examples usually do not cluster since they can belong to any class but unlabelled examples are belong either positive or negative examples [4]. But in this fast life CBIR system have many problems are feature extraction and classification, indexing and fast searching, limitation in selecting images with finer changes, search mechanism, data management, time consuming and occupies too much memory, effective similarity measure, transportation problem, human attention principles in the process of image contents identification, understanding visual content, dimensionality in pattern recognition, computer vision and pattern recognition, human visual perception in distinguishing images etc

It is almost impossible to estimate the real distribution of unlabelled images in the database based on the relevant feedback. The spreading activation theory is used to give annotation for the untagged images. First it is identifying the candidate annotations and then refining the candidate annotations. Milestone of CBIR system is low level feature extraction. Feature extraction may be done from region or an entire image. The Bayes point machines approach used to give the soft annotation to the unlabelled image. The novel system for content based image retrieval in large unlabelled database called PicSOM, and it is based on tree structured self-organizing maps (TS-SOMs). The TS-SOMs working as hierarchical levels. It is presents the result image based on the query image, resembles each other [4]. The section II of this paper presenting research background of image classification, feature clustering, feature selection with unlabelled image. Section III of this paper we are give some recommendations for improve image retrieval system with unlabelled image. Section IV is conclusion of this paper.

II. RESEARCH BACKGROUND

A. Basic idea of Content Based Image Retrieval system

Content Based Image Retrieval (CBIR) is a technique used for extracting similar images from an image database. Content Based means search analysis based on the contents of an image. Content refers to the color, texture, shape or any other information. The feature descriptors of the content can be obtained from the image itself. All CBIR systems perform search and retrieval operations in image database using image features like color, texture and shape. CBIR system is working with heterogeneous image database. All databases are having three types of images like relevant, irrelevant and unlabelled.

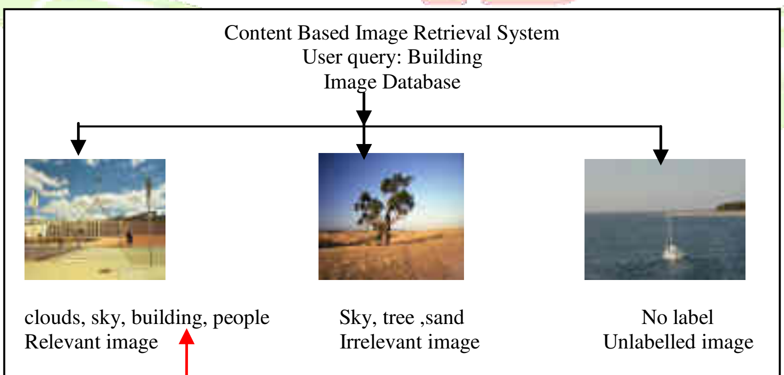

Fig. 1 Example of CBIR system image classification

B. Basic diagram of CBIR system

Basically CBIR system having some components they are, preprocessing, feature extraction, similarity measures, relevance feedback. These are the main pillar of CBIR system.

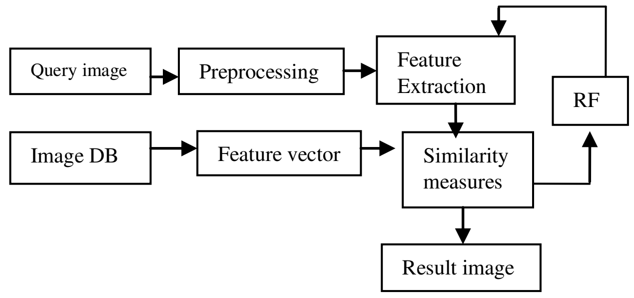

Fig. 2 Basic diagram of CBIR system

C. Feature Extraction Types

It is used to identify the meaningful information from image to user based on features like color, texture and shape etc.

Independent Component Analysis (ICA) It is a computational method to get hidden values of random variables. ICA basically designed for multivariate data. It helps in extraction rich features of an image [6]. It is working with image features are average value, min value and max value. ICA is working with dimension and data points of input data for calculating mean value. In ICA input data has to be centre. *Heuristic Features Extraction* Heuristic features for each image are stored in the database. 'Water', 'Sand', 'Sky', 'Brick', 'Bark', 'Cloud', 'Grass' etc. are some of the mostly used heuristic features. For color images color based recognition is used. Hence to reduce the space and time complexity correlogram is calculated is for 8 colors and 8 distances. Heuristic features are can be efficiently recognized by shape based [11] recognition method. *Semantic Color Extraction* The storage is dependent upon accuracy of segmentation algorithm because segmentation algorithm created initial segments on the basis [11] of semantic colors only. Hue, saturation and intensity values will be converted to semantic color names by using HSI space algorithm of each image. *Ontology Based Approach* The images returned by search engine may not be relevant to the user query; therefore the content of each image needs to be verified. This module is responsible to check the images for the compliance with the input query. The objects in each image are detected using shape based feature extraction and these objects are passed to Color based Feature Extraction technique which uses MTH algorithm to calculate the pixel color and color value of the objects and texture classification technique to identify texture / pattern (if any) in the objects. In the next step, the low level features extracted in the previous step are converted into high level ontology concepts; afterwards SPARQL query is generated using these concepts and executed on the ontology knowledgebase. If the result class(es) matches user search query, the image is included in the resultant set otherwise it is discarded. As a result only the related images remains and the non-relevant images are discarded in this step [16]. *Active Contour Selection* Many image Retrieval systems take a simple approach by using typically normal-based distances on the extracted feature set as similarity measurement function. The main purpose behind these proposed CBIR systems is that given a good set of features extracted from the images in the dataset (here the images are stored.) then for two images to be similar, their extracted features vectors have to be close to each other. The techniques or algorithms used in our proposed CBIR are Feature Extraction algorithm in this Active Contour Selection processes the image and compute feature vector. This feature is used for the extraction of information from Wang dataset. Features can be human vision related, low-level features, middle-level features, high-level features and semantic related. Most commonly used features for retrieval are color, shape, texture. *Object Detection Method: Histogram of Gradient with Entropy* Histogram of Gradient (HOG)12 is an edge based descriptors used in computer vision and image processing for the purpose of object detection. The image is divided into small regions called cells, then the local appearance and shape of the object is obtained by distribution of intensity gradient and edge directions in HOG descriptors. Since the HOG descriptor operates on localized cells, the method upholds invariance to translation and photometric transformations [49].

D. Feature Classification Algorithms

Local binary patterns (LBPs) Local binary patterns (LBPs) have been extensively used in image classification. The most common usage of LBPs is to calculate their histogram and use them as features. Here, for sake of comparison, LBPs are extracted and recorded as barcodes through concatenation of binary vectors around each pixel. As supplementary information, barcodes could enhance the results of existing "bag-of-features" methods, which are generally designed to capture the global appearance of the scene without much attention to the local details of scene objects (e.g. shape of a tumor in an MR scan) [12]. *Feed-forward back propagation neural network* Neural network has been inspired from the brain that consists of network of "neuron like" units called nodes. This network can be used in many fields like classification, optimization, and control theory and for solving regression problems. Mainly NNs are helpful, or we can say that can perform effectively where classification is required on the basis of training and testing. NN are preferred over other networks because it is dynamic in nature. Dynamic in nature means weights of input can be adjusted according to the desired output. At this weight adjustment is known as learning process [13]. *Image classification* Image classification and clustering are the supervised and unsupervised classification of images into groups. In supervised classification, we are given a collection of labelled (pre-classified) images, and the problem is to label a newly encountered, yet unlabeled images. Typically, the given labelled (training) images are used to do the machine

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 10, March 2016

learning of the class description which [15] in turn are use to label a new image. *Biased Maximum Margin Analysis* In BMMA, labelled positive feedbacks are mapped as close as possible, whereas labelled negative feedbacks are separated from labelled positive feedbacks by a maximum margin in the reduced subspace. The SemiBMMA can incorporate the information of unlabeled samples into the RF and effectively alleviate the over fitting problem caused by the small size of labelled training samples [18]. *Genetic programming* The framework described in uses only the information provided by the images labelled as relevant. One natural extension would be to incorporate non-relevant images/regions to the learning and ranking processes. The objective in this case is to find database images which are near to relevant and far from non-relevant images. Two components of the described framework were adapted to cope with such an extension: the training set redefinition and the similarity function employed to rank regions. Redefinition of the training set the new training set is composed by relevant, non-labelled, and non-relevant regions. Let IRR be the set of regions labelled as non-relevant over all iterations. Fitness computation process is assign the highest fitness values are assigned to those individuals that rank relevant regions at the first positions and the non-relevant ones at the last positions [20]. Here t_i is the training set of image region 1; if ti is relevant:, _1; if ti is non-relevant:, 0; if ti is unlabeled:. *Modified K-Nearest Neighbour Algorithm* Inspired the traditional KNN algorithm, the main idea is classifying the test samples according to their neighbour tags. This method is a kind of weighted KNN so that these weights are determined using a different procedure. The procedure computes the fraction of the same labelled neighbours to the total number of neighbours [25].

E. Feature Clustering Algorithms

Image clustering In unsupervised classification (or image clustering), the problem is to group a given collection of unlabeled images into meaningful clusters according to the image content without a priori knowledge. The fundamental objective for carrying out image classification or clustering in image mining is to acquire content information the users are interested in from the image group label associated with the image. Intelligently classifying image by content is an important way to mine valuable [15] information from large image collection. *Fuzzy C- Means Clustering* In this clustering obtaining feedback from user interaction, obtain the feedback from the users on prior results. Feedback is in the form of relevant or irrelevant to request during searching time in CBIR system. Results found to be not satisfied means again learn the system through a feedback algorithm and hence results are refined till the results are satisfied [19].

F. Feature Segmentation Method

Region-based Active Contour Model A region-based active contour model uses statistical information of regions both inside and outside the curve for contour evolution, for example, the Chan-Vese (C-V) model. This model is based on the assumption that the pixel regions of the image are statistically homogenous. It deals well with noisy images, blur images, and images that have multiple holes, disconnected regions etc. In MRI brain image analysis the region based active contour model since considers global properties of images such as contour lengths and MRI image pixel regions as against local properties such as gradients [50].

G. Similarity Measures

Supervised Hashing Hashing methods minimize the hamming distance of "neighbouring" image pairs. "Neighbouring" is defined by its semantic meaning, i.e., whether the two images belong to same category or not. Therefore, supervised information can be naturally encoded as similar and dissimilar pairs. Feature vectors are randomly selected from to build the label matrix. Note that we need to provide labels for only a small number of image pairs. Therefore, labelled data are explicitly constrained by both semantic information and visual similarities, whereas unlabeled data are mainly constrained by visual similarities and implicitly affected by labelled data [14]. *Hash value* Similarity measurement algorithm performs matching which compares between extracted features of images to find whether they are similar or not and up to what level they are matching. Also compare two images, distance of zero shows that it is very similar picture , distance of 5 means a few things may be different if 10 means fully different picture. Hashing algorithm proposed advocates to generate the hash value for each image in the dataset. The hash value of query image is used to calculate the hamming distance. Finally the retrieval of images from the dataset is of same class and cross class is checked to calculate the ranking of retrieval [17]. *Canonical Correlation Analysis* Given image features $x_1, ..., x_n$ and text features $y_1, ..., y_n$, matrices A and B found by CCA maximize the correlation between canonical variables $s_i = A^t x_i$ and $t_i = B^t y_i$. We define the metric for the combined latent space denoted by z. Common CCA generates two canonical spaces, denoted by s_i and t_i. An obvious approach is to search for similar [23] images with their Euclidean distance within the canonical space generated from the image features.

H. Relevance Feedback

Graph based image re-ranking We denote by $\phi i \in \{0, 1\}$, $i \in V$, the feedback that the user would give to the ith image, namely $\phi i = 1$ if the image is relevant and $\phi i = 0$ if it is non-relevant. Clearly, we always assume the query image itself to be relevant, i.e., $\phi 0 = 1$. At each feedback iteration $t > 0$, the user is given the opportunity to indicate which images among the provided ones are relevant for the query. By doing so, the system gathers more and more information about the relevancy of the images in the database [24].

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 10, March 2016

Author	Month & Year	Image features	Problems	Based on survey
Krishna N. Parekh, et al., [26]	December 2014	Color-edge; color-shape-texture; color-texture; shape-texture	Extract the image features	Different combinations of low level feature extraction
Ashwin G. Parmar, et al., [27]	November, 2014	Color, texture and shape using saliency map.	Retrieve revelent images based on their contents	Feature extracting methods are discussed, analyzed and compared
Rikin K Thakkar, et al., [28]	December 2013	Color, texture and shape	Identify unlabelled image.	CBIR with some image retrieval methods based on unlabelled data
Mr. Pranjul mishra, et al., [29]	March-April, 2014	Color, texture	Efficient and more accurate of image retrieval	Clustering techniques

I. *Image Matching*

Image matching system usually provides a user interface for communicating with the user. It collects the required information from the user and displays the matching results to him. However, as the images are matched based on low-level description features, the target or the similar images may be far away from the query in the feature space, and they are not returned in the limited number of retrieved images of the first display. Therefore, in some matching systems, there is a relevance input from the user, where human and computer can interact to increase matching performance. Similarity computation is the system computes the similarity between the query image and the database images graphically represents the computed mean and standard deviation which represents principal color and the variation of pixel colors respectively. Matching: The system retrieves and presents a sequence of images ranked in decreasing order of similarity. As a result, the user is able to find relevant images by getting the top-ranked images first. These images are ranked based on the similarity. Under each image, a ranking percentage value is attached so that the user can see which images are relevant or irrelevant. The amount of percentage represents the degree of relevance. Higher percentage values higher the relevance [21]. *Chain Code Based Exact Match (CCEM) Algorithm* The CCEM algorithm assumes a preprocessing transformation phase which takes a main matrix (MM) and a sub-matrix (SM) and transforms them into vectors TMM and TSM respectively. The CCEM algorithm searches for an exact match of image SM inside a larger image MM, using their corresponding transformed vectors TSM and TMM respectively. The search process starts by detecting the first pixel match and continues comparing in the current row pixel by pixel. When it reaches the end of the current row, it determines the offset of the column in the next row. Search is [22] terminated when one of the following two cases occurs: If the last element of TSM is reached, then a sub-matrix match is found. If the value of TMM is less than the corresponding one in TSM, then a match is not found. *Chain Code Based Approximate Match (CCAM) Algorithm* The CCAM algorithm in which incorporates a percentage of mismatch or error value in the search process which gives CCAM the flexibility over CCEM for [22] handling mismatch pixels. Similar to CCEM, CCAM performs the transformation preprocessing phase to gain the advantage of size reduction by 50% on average and thereby reduce the number of comparisons. The search phase starts by comparing the pixels in SM and MM, pixel by pixel.

TABLE II
Summary of Previous image features, problems and future scopes in CBIR

Ashwani Kr. Yadav, et al., [30]	2014	Color, texture and shape	No universally accepted feature extraction	CBIR systems their behaviour, texture analysis and various feature extraction with representation

TABLE III
Summary of features, techniques and CBIR

Previous image problems, future scopes in

Author	Month & Year	Image features	Problems	Techniques	Future scope
Sapanjot Kaur, et al., [31]	September 2015	Average value, min value and max value.	Feature extraction and classification	Independent Component Analysis (ICA), NN (neural network).	Deep learning neural network, integrating with Fuzzy C-means clustering algorithm
H.R.Tizhoosh, [32]	September 2015	Projection angles	Annotate medical images	Radon barcodes (RBC), local binary patterns (LBP) and local Radon binary patterns (LRBP)	ROI barcodes for intra-class retrieval
Harjeet Kaur1, et al., [33]	August 2014	Color	Classification	Feed-forward back propagation neural network	Fuzzy C-means clustering algorithm
Gurpreet Kaur, et al., [34]	2013	Color, texture shape	Indexing and fast searching of images based on visual features. Dimension reduction	Back-propagation neural network and K-Mean algorithm	Deep learning neural network
Shruti S. Betageri, et al., [35]	July 2014	Color and texture	Limitation in selecting images with finer changes	Color histogram, wavelet and nearest neighbour	Improve the retrieval performance
MS. PRAGATI ASHOK DEOLE, et al., [36]	May 2014	Color, shape and texture	Classification of image	K-nearest neighbours Algorithm (KNN). Relative standard derivation	Improve the retrieval performance

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)

International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)

Vol. 2, Special Issue 10, March 2016

Cheruvu Mahammad Ibraheem, et al., [37]	October 2015	Color, shape and texture	Faster and accurate of CBIR behaviour	Feature extraction methodologies	Various databases and to be applied to video retrieval
Maryam Shahabi lotfabadi, et al., [38]	December 2014	Color, shape and texture	Dealing with incomplete and uncertain data	Rough Set and a 1-v-1 Support Vector Machine (SVM)	1-v-r (one-versus-rest) (SVM), and rough set
S.Sasikala, et al., [39]	March 2015	Color, shape and texture	Search mechanism , data management and limitation in selecting images with finer changes	Color histogram and wavelet nearest neighbour classification	Color appearances
Maryam Shahabi Lotfabadi, et al., [40]		Color, shape and texture	Feature selection	Fuzzy Rough Set feature selection	Rules extracted from the selecting features of the Fuzzy Rough Set feature selection
Chengyou Wang, et al., [41]	January 2015	Color, texture and shape	Time consuming and occupies too much memory	Grading retrieval algorithms based on DCT compressed domain and DWT compressed domain	Improve the retrieval performance
N. Gnaneswara Rao, et al., [42]	2014	Color and texture	Effective Similarity Measure, Transportation problem	Optimal Cost Region Matching (OCRM) similarity measure for region based image retrieval	Improve the retrieval performance
Ni Li1, et al., [43]	February 2016	Color and texture	Human performance prediction	Improved KNN	Improve prediction capability and applicability.
Juan Villegas-Cortez, et al., [44]	2014	Color	Fast retrieval	K-means, Algoritmo Gen_etico	Improved the retrieval
Mohammad Al-Azawi, [45]	January 2015	Color, texture, shape and saliency	Human attention principles in the process of image contents identification	Irregularity of Regions	Tree or binary search, support vector machine (SVM) and Swarm-based neural networks
Orod Razeghi, [46]	July 2015	Color, texture and shape	Understanding visual content, dimensionality in pattern recognition	Object Recognition	Object classification and pattern recognition applications
Javier Espinosa, [47]	2014	Array of pixels	Computer vision and pattern recognition	Historical Document Recognition	Follow different ways
Eng. Ahmed K. Mikhraq, [48]	2013	Color, texture and shape	Human visual perception in distinguishing images	Clustering and Genetic Algorithm	Feedback from the user, Parallel computing

III. RECOMMENDATIONS

Generally image database is classified into relevant, irrelevant and unlabelled image. This is also classified an accurate manner, that is relevant, irrelevant of labelled image and also relevant, irrelevant of unlabelled image. In research background of labelled images are retrieved in an efficient manner with various image visual features using as user query during image retrieval. But unlabelled images are still inefficient one for retrieval process. The unlabelled images are more useful for image classification. In this paper we are mention some recommendations for improve image retrieval by using unlabelled images from World Wide Web. The Crisp set using for image retrieval process working with only labelled images from image database. In most Content Based Image Retrieval system is allow to user for giving label for unlabelled image during retrieval time. It is not consider unlabelled image for retrieval process.

A. Big Data in CBIR system

Big data is focused on deriving knowledge from multiple data formats using intelligent analytics techniques. Medical analytics is a typical example, which encompasses data in multiple formats available as text, images and the data in the databases. In general, Apache Hadoop processes very large scale data files containing large number of sequences of input record, which can be processed parallel. This large data file is split automatically, along with distribution of file system block boundaries. Because of granularity of these input splits in distributed blocks usually contain many data records. Increasingly large and complex algorithms make processing image data more challenging.

B. *Adaptive similarity*

In this approach, the query image is first categorized into one of the predefined adaptive weight categories, such as "portrait" and "scenery." Inside each category, a specific pretrained weight schema is used to combine visual features adapting to this kind of images to better re-rank the text-based search result. This correspondence between the query image and its proper similarity measurement reflects the user intention. *Keyword expansion* In our approach, query keywords are expanded to capture users' search intention, inferred from the visual content of query images, which are not considered in traditional keyword expansion approaches. *Image pool expansion* The image pool retrieved by text-based search accommodates images with a large variety of semantic meanings and the number of images related to the query image is small. In this case, re-ranking images in the pool is not very effective. Thus, more accurate query by keywords is needed to narrow the intention and retrieve more relevant images.

C. Markov random field

Markov network or undirected graphical model is a set of random variables having a Markov property described by an undirected graph. In other words, a random field is said to be Markov random field if it satisfies Markov properties are pairwise Markov property: Any two non-adjacent variables are conditionally independent given all other variables, Local Markov property: A variable is conditionally independent of all other variables given its neighbours, Global Markov property: Any two subsets of variables are conditionally independent given a separating subset. A Markov network or MRF is similar to a Bayesian network in its representation of dependencies; the differences being that Bayesian networks are directed and acyclic, whereas Markov networks are undirected and may be cyclic.

D. Artificial Neural Network Levenberg-Marquardt (ANNLM)

Artificial Neural Network Levenberg-Marquardt (ANNLM) to categorize inputs into the collection of target categories (relevant, irrelevant or unlabelled images) in accordance with the feature extraction parameters. The Levenberg-Marquardt (LM) scheme is a higher-order adaptive approach and considerably decreases the Mean Square Error of a neural network. In this work attempts to apply optimized LM scheme in order to diminish the errors during the process of classifying the tumour in brain images.

E. Semantic modelling

The step aims to learn a model f(x) for each semantic concept that can be implemented for automatic semantic annotation on unlabeled images. Since the images from several platforms do not have corresponding tags, it will cost great human efforts for manual annotation for supervised model learning. Adapt the one-to-one adaptive support vector machine for collaborative semantic learning. $D^F = D^F_l \cup D^F_u$ denote a partially labeled primary dataset from one unlabeled platform, where D^F_l is the manually labeled part and D^F_u is the unlabeled part.

F. Deep learning

It has been developed to learn feature hierarchies with higher-level features formed by the composition of lower-level ones. This aims to construct multiple levels of feature representations where higher layers characterise more abstract features. Deep learning mainly offers the three advantages, The first is the expressive power where the combination of (distributed) features at each layer can define the exponential order of higher-level features, and this order is further exponentially increased by passing through layers. The second advantage is the invariance property where more abstract features are generally invariant to subtle changes in visual appearances. The last one is the explanatory factor that the learnt feature hierarchy can capture valuable patterns or structures underlying raw images or videos. Finally, a classifier for detecting a concept is created by using the learnt hierarchy as initialisation of a multi-layer neural network, or building a supervised classifier by constructing the feature vector of each example based on the hierarchy.

Re-Ranking via Visual Feature Verification The relevant image grouping approach can find more relevant images for the query sketch, some irrelevant images may appear in the top *N* results. If we re-rank the top *N* results by measuring their similarities in the visual feature space, then the refined search results will be more satisfactory. Our aim is to filter out irrelevant images using content matching or spatial constraints, which are often used in retrieval result verifications. *Shape-adaptive discrete wavelet transform (SA-DWT)* The idea was to firstly perform a 1D DWT decomposition on region rows, and then apply the same DWT on resulting region colomns. This resulting in pixels values outside the segment related with the pixels values inside the segment. Since we use an odd biorthogonal wavelet filters, the segment extension used is the symmetrical one. The image mask may be defined as image texture within the edge. The mask permits ROI pixels to be distinguished from their surroundings.

IV. CONCLUSION

In this section we surveyed the field of content-based image retrieval, by providing an overview of the most important aspects characterizing that kind of images. The image databases are classified into the labelled image as relevant, irrelevant and also unlabelled image. The use of keywords is also listed. In this paper, we have presented a survey on content based image retrieval based on unlabelled images. We compare the many image database classification processes based on unlabelled images and also give some basic components of CBIR systems like feature extraction, classification, clustering, segmentation, similarity measures, relevance feedback and image matching. The image classifying processes are more useful to the next generation of the content based image retrieval system with unlabelled images. We have presented the recommendations to improve image classification for the future directions of the current content based image retrieval systems.

ACKNOWLEDGMENT

The second Author of this paper acknowledges UGC-Minor Research Project (No.F.41-1361/2012 (SR)) for the financial support.

REFERENCES

[1] Minakshi Kumar,"DIGITAL IMAGE PROCESSING", *Photogrammetry and Remote Sensing Division, Indian Institute of Remote Sensing, Dehra Dun, Satellite Remote Sensing and GIS Applications in Agricultural Meteorology ,pp. 81-102.*
[2] https://www.spacetelescope.org/static/projects/fits_liberator/image_processing.pdf.
[3] T. Dharani, I. Laurence Aroquiaraj, "CONTENT BASED IMAGE RETRIEVAL SYSTEM with MODIFIED KNN ALGORITHM", *International Journal of Innovations in Scientific and Engineering Research(IJISER)*, Vol. 1 Issue 1 Jan 2014.
[4] T. Dharani, I. Laurence Aroquiaraj, ―A Survey on Content Based Image Retrieval,‖ *IEEE-2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering (PRIME)*, PRIME 2013, 978-1-4673-5845-3/13/2013.
[5] Ashwani Kr. Yadav, R. Roy, Vaishali and Archek Praveen Kumar, "Survey on Content-based Image Retrieval and Texture Analysis with Applications", *International Journal of Signal Processing, Image Processing and Pattern Recognition* Vol. 7, No. 6 2014.
[6] Sapanjot Kaur Gaganpreet Kaur, "A Novel Content Based Image Retrieval using ICA and Neural Network", *International Journal of Emerging Technology and Innovative Engineering* Volume 1, Issue 9, September 2015.
[7] H.R.Tizhoosh, "BARCODE ANNOTATIONS FOR MEDICAL IMAGE RETRIEVAL: A PRELIMINARY INVESTIGATION", *IEEE International Conference on Image Processing (ICIP)*, September 27-30, 2015.
[8] Harjeet Kaur, Jagbir Singh Gill, "CONTENT BASED IMAGE RETRIEVAL SYSTEM BASED ON NEURAL NETWORKS", *International Journal of Application or Innovation in Engineering & Management (IJAIEM)* Volume 3, Issue 8, August 2014.
[9] Desai Devanshi Manojbhai, R.Rajamenakshi, "Large Scale Image Feature Extraction from Medical Image Analysis", *International Journal of Advanced Engineering Research and Science (IJAERS)* Vol-3, Issue-1 , Jan- 2016.
[10] Naresh M.Bhagat, Dr. M.A.Pund, "A Study Paper on "Required Image Retrieval by Capturing User Intention Using Image Pool Clustering", *International Journal of Computer Science and Mobile Computing*, Vol.5 Issue.1, pg. 25-30, January- 2016.
[11] Dipankar Hazra, Debnath Bhattacharyya, Kil-hwan Shin, " Heuristic and Intermediate FeaturesBased", *Contemporary Engineering Sciences*, Vol. 7, 2014.
[12] H.R.Tizhoosh, "BARCODE ANNOTATIONS FOR MEDICAL IMAGE RETRIEVAL: A PRELIMINARY INVESTIGATION", *IEEE International Conference on Image Processing (ICIP),* September 27-30, 2015.
[13] Harjeet Kaur, Jagbir Singh Gill, "CONTENT BASED IMAGE RETRIEVAL SYSTEM BASED ON NEURAL NETWORKS", *International Journal of Application or Innovation in Engineering & Management (IJAIEM)* Volume 3, Issue 8, August 2014.
[14] Xiaofan Zhang, Wei Liu, Wei Liu, Sunil Badve, Shaoting Zhang, "Towards Large-Scale Histopathological Image Analysis: Hashing-Based Image Retrieval", *IEEE TRANSACTIONS ON MEDICAL IMAGING*, VOL. 34, NO. 2, FEBRUARY 2015.
[15] Aboli W. Hole, Prabhakar L. Ramteke, "Design and Implementation of Content Based Image Retrieval Using Data Mining and Image Processing Techniques", *International Journal of Advance Research in Computer Science and Management Studies,* Volume 3, Issue 3, March 2015.
[16] Umar Manzoor, Mohammed A. Balubaid, Bassam Zafar, Hafsa Umar, M. Shoaib Khan, "Semantic Image Retrieval: An Ontology Based Approach", *International Journal of Advanced Research in Artificial Intelligence,* Vol. 4, No.4, 2015.
[17] MAITHIILI.K, ELAKKIYA.P, GAYATHRI.G, SANGEETHA.As "CONTENT BASED IMAGE RETRIEVAL WITH HASH CODES", *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)* Volume 4 Issue 4, April 2015.
[18] Satish kumar, Suresh.M.B, "Biased Maximum Margin Analysis for Content based Interactive Image Retrieval", *International Journal of Computer Science and Network Security*, VOL.15 No.8, August 2015.
[19] Jasmine Samraj, NazreenBee . M, "Content Based Medical Image Retrieval Using Fuzzy C- Means Clustering With RF", *International Journal of Innovative Science, Engineering & Technology*, Vol. 2 Issue 1, January 2015.
[20] J.A. dos Santos, C.D. Ferreira, R. da S. Torres, M.A. Gonçalves, R.A.C. Lamparelli, "A relevance feedback method based on genetic programming for classification of remote sensing images", *Information Sciences* 181, 2011.
[21] Mrigank Sharma, Amrita Priyam, "IMAGE MATCHING ALGORITHM BASED ON HUMAN PERCEPTION", *International Journal of Innovative Research in Science, Engineering and Technology* Vol. 2, Issue 5, May 2013.
[22] Azzam Sleit, Heba Saadeh, Ibraheem Al-Dhamari, Afaf Tareef, "An Enhanced Sub image Matching Algorithm for Binary Images", *RECENT ADVANCES in APPLIED MATHEMATICS,* ISSN: 1790-2769, ISBN: 978-960-474-150-2.
[23] Yoshitaka Ushiku, Tatsuya Harada, Yasuo Kuniyoshi, "IMPROVING IMAGE SIMILARITY MEASURES FOR IMAGE BROWSING AND RETRIEVAL THROUGH LATENT SPACE LEARNING BETWEEN IMAGES AND LONG TEXTS", *Proceedings of 2010 IEEE 17th International Conference on Image Processing* September 26-29, 2010.
[24] Malay Kumar Kundu, Manish Chowdhury, Samuel Rota Bul`o, "A Graph-Based Relevance Feedback Mechanism in Content-Based Image Retrieval", *KNOWLEDGE-BASED SYSTEMS* · JANUARY 2015.
[25] T. Dharani, I. Laurence Aroquiaraj, "Content Based Image Retrieval System using Feature Classification with Modified KNN Algorithm", sep, 2013.

[26] Krishna N. Parekh, Mr. Narendra Limbad, "A Survey on different approaches of CBIR and their combinations", *International Journal Of Engineering And Computer Science,* Volume 3 Issue 12 December 2014.
[27] Ashwin G. Parmar1, Mukti S. Pathak2, "A Survey On Flower Image Retrieval Based On Saliency Map And Feature Extraction", *International Journal Of Engineering And Computer Science* Volume 3 Issue 11 November, 2014.
[28] Rikin K Thakkar, Ms Ompriya Kale, "Content Based Image Retrieval for Unlabelled Images", *IOSR Journal of Computer Engineering*, Volume 15, Issue 4 (Nov. - Dec. 2013).
[29] Mr. Pranjul mishra [1], Ms. Sonam[2], Mrs. S. Vijayalakshmi[3], "Content Based Image Retrieval Using Clustering Technique: A Survey", *INTERNATIONAL JOURNAL OF RESEARCH IN COMPUTER ENGINEERING AND ELECTRONICS.* VOL : 3 ISSUE :2 (MARCH-APRIL'14).
[30] Ashwani Kr. Yadav, R. Roy, Vaishali and Archek Praveen Kumar, "Survey on Content-based Image Retrieval and Texture Analysis with Applications", *International Journal of Signal Processing, Image Processing and Pattern Recognition,* Vol. 7, No. 6 (2014).
[31] Sapanjot Kaur ,Gaganpreet Kaur, "A Novel Content Based Image Retrieval using ICA and Neural Network", *International Journal of Emerging Technology and Innovative Engineering* Volume 1, Issue 9, September 2015.
[32] H.R.Tizhoosh, "BARCODEANNOTATIONSFORMEDICALIMAGERETRIEVAL: APRELIMINARYINVESTIGATION", *International Conference on Image Processing*, September 27-30, 2015.
[33] Harjeet Kaur1, Jagbir Singh Gill2, "CONTENT BASED IMAGE RETRIEVAL SYSTEM BASED ON NEURAL NETWORKS", *International Journal of Application or Innovation in Engineering & Management*, Volume 3, Issue 8, August 2014.
[34] Gurpreet Kaur1, Mnupreet Kaur2, "Content Based Image Retrieval Using BPNN and K-Mean Algorithm", *International Journal of Science and Research*, 2013.
[35] Shruti S. Betageri1, V. S. Malemath2, Shivanand M. Patil3, "Content Based Image Retrieval Using Color and Textural Features", *International Journal of Advanced Research in Computer and Communication Engineering*, Vol. 3, Issue 7, July 2014.
[36] MS. PRAGATI ASHOK DEOLE1, PROF. RUSHI LONGADGE2, "Content Based Image Retrieval using Color Feature Extraction with KNN Classification", *International Journal of Computer Science and Mobile Computing*, Vol.3 Issue.5, May- 2014.
[37] Cheruvu Mahammad Ibraheem1, Dr. G. Umamaheswara Reddy2, "Content Based Image Retrieval using HSV Color, Shape and GLCM Texture", *International Journal of Advanced Research in Computer and Communication Engineering,* Vol. 4, Issue 10, October 2015.
[38] Maryam Shahabi lotfabadi, Mohd Fairuz Shiratuddin, Kok Wai Wong, "Content Based Image Retrieval system with a combination of Rough Set and Support Vector Machine", *International Joint Conferences on Computer, Information, Systems Sciences, & Engineering*, 12 - 14 December.
[39] S. Sasikala R. Soniya Gandhi, "Efficient Content Based Image Retrieval System with Metadata Processing", *International Journal for Innovative Research in Science & Technology*, Volume 1 Issue 10 March 2015.
[40] Maryam Shahabi Lotfabadi, Mohd Fairuz Shiratuddin, Kok Wai Wong, "Evaluation of Fuzzy Rough Set Feature Selection for Content Based Image Retrieval System with Noisy Images".
[41] Chengyou Wang, Xinyue Zhang, Rongyang Shan, and Xiao Zhou, "Grading Image Retrieval Based on DCT and DWT Compressed Domains Using Low-Level Features", *Journal of Communications,* Vol. 10, No. 1, January 2015.
[42] N. Gnaneswara Rao1 T. Sravani2 and V. Vijaya Kumar3 , "OCRM: Optimal Cost Region Matching Similarity Measure for Region Based Image Retrieval", *International Journal of Multimedia and Ubiquitous Engineering.* Vol.9, No.4 (2014).
[43] Ni Li1 & Haipeng Kong1 & Yaofei Ma1 & Guanghong Gong1 & Wenqing Huai1, "Human performance modeling for manufacturing based on an improved KNN algorithm", Springer, February 2016.
[44] Grigori Sidorov, "Modelado computacional de habilidades lingüísticas y visuals", *Research in Computing Science 74 (2014).*
[45] Mohammad Al-Azawi, "A New Approach to Automatic Saliency Identification in Images Based on Irregularity of Regions", January – 2015.
[46] y Orod Razeghi, "An Investigation of a Human in the Loop Approach to Object Recognition", July 2015.
[47] Javier Espinosa, "Clustering of Image Search Results to Support Historical Document Recognition", Thesis no: MGCS-2014-06.
[48] Eng. Ahmed K. Mikhraq , "Content-Based Image Retrieval (CBIR)System Based on the Clustering and Genetic Algorithm", 1434H 2013.
[49] R. Aarthi1*, K. P. Anjana1 and J. Amudha2," Sketch based Image Retrieval using Information Content of Orientation", *Indian Journal of Science and Technology*, Vol 9(1), January 2016.
[50] A. Shenbagarajan1*, V. Ramalingam1, C. Balasubramanian2 and S. Palanivel1, "Tumor Diagnosis in MRI Brain Image using ACM Segmentation and ANN-LM Classification Techniques", Indian Journal of Science and Technology, Vol 9(1), January 2016.