Content based Image Retrieval Survey of Feature Selection and Matching Techniques

Monika Soni
Dept. of Computer Science & Engineering
M.Tech Student, DCRUST Murthal, Sonepat, India

Dr. Parvinder Singh
Dept. of Computer Science & Engineering
Associate Professor, DCRUST Murthal, Sonepat, India

Abstract—Content Instituted Picture Retrieval (CBIR) is a method that uses discernible contents such as color, form, sense, and supplementary picture features to reclaim comparable pictures from a colossal repository opposing a given query image. This has come to be an alert scutiny span alongside the advent of the digital mass media in all most all applications. Sense as a main picture feature has been learned for countless years in the area of picture analysis. Countless sense descriptors are counseled for CBIR and picture analysis. But, discovering a solitary best representation of sense pictures that can distinct disparate sense classes alongside well-defined borders in the feature space always stayed as an unsolved problem. This paper reviews a little of most accepted CBIR Feature extraction and matching methods.

Keywords*:* Content-based image retrieval (CBIR), feature extraction, textures, wavelets

I. INTRODUCTION

With the present outpouring of multimedia-enabled arrangements, the demand for multimedia retrieval has increased by leaps and bounds. Due to the intricacy of multimedia contents, picture understanding is a difficult-albeit-interesting case of scutiny, inside the area of multimedia retrieval. Obtaining priceless vision from a large-scale multimedia repository, normally denoted to as "multimedia mining", has presently caught up as a area of attention amongst researchers. Typically, in the progress of an picture requisition arrangement, semantic picture retrieval relies deeply on the connected captions, e.g., filenames, groups, annotated key-words, and supplementary manual descriptions. Hunting of pictures is predominantly established on associated metadata such as keywords, text, etc. The word CBIR [1] describes the procedure of reclaiming wanted pictures from the colossal collection of database on the basis of features that can be automatically removed from the images. The ultimate aim of a CBIR arrangement is to circumvent the use of textual descriptions in the hunt for an picture by the user. Unfortunately, this kind of a textual-based picture retrieval arrangement usually suffers from these problems: high-priced manual annotation and inaccurate and inconsistent automated annotation. On one hand, the price associated alongside manual annotation is prohibitive alongside regards to a large-scale data set. On the supplementary hand, improper automated annotation yields distorted result for semantic picture retrieval. As a consequence, a number of influential picture retrieval algorithms have been counseled to deal alongside such setbacks above the past insufficient years. CBIR is the mainstay of present picture retrieval systems.

In CBIR, retrieval of picture is established on similarities in their contents, i.e., textures, colors, forms etc.[2], that are believed the lower level features of an image. These standard ways for picture retrieval are established on the computation of the similarity amid the users query and images. In CBIR every single picture stored in the database, has its features removed and contrasted to the features of the query image. Thus, mainly, it involves two procedures, viz, feature extraction and feature matching.
Feature extraction involves the picture features to a distinguishable extent. Average RGB, Color Moments, Concurrence, Native Color Histogram, Globe Color Histogram and Geometric Moments are utilized to remove features from the examination image. Feature matching, on the supplementary hand, involves matching the removed features to yield result those display discernible similarities.

Feature vectors are computed for the given image. The Euclidean distance [3] is utilized as default implementation for contrasting two feature vectors. If the distance amid feature vectors of the query picture and pictures in the database is tiny plenty, the corresponding picture in the database is to be believed as a match to the query. The find is normally established on similarity rather than on precise match and the retrieval result are next ranked accordingly to a similarity index.

Figure 1 displays the block diagram of a frank CBIR arrangement.

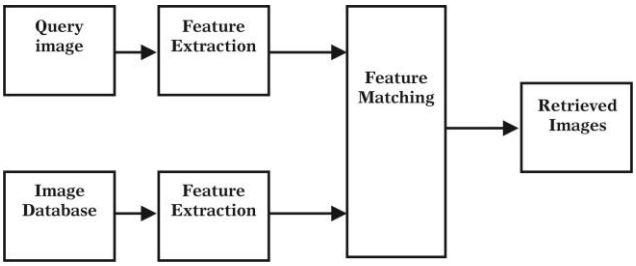

Figure 1: Block diagram of a basic CBIR system

1.1 Keyword Based Image Retrieval

the Keyword based Image Retrieval system [4] used keywords as descriptors to index an image. General Framework of keyword based image retrieval is shown in Fig.2

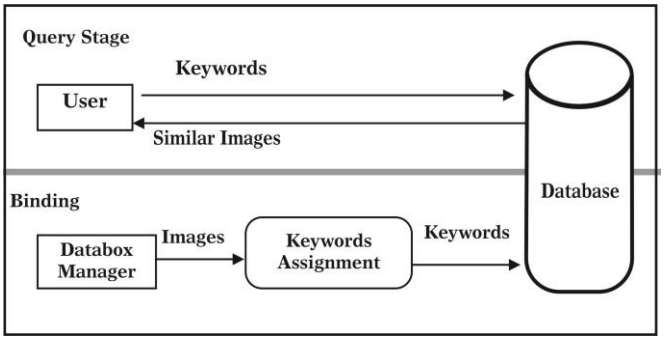

Figure 2: General Framework of Keyword Based Image Retrieval

Before pictures are being stored in the database, they are examined manually and allocated keywords that are most appropriate to represent their contents. These key- words are stored as portion of the qualities associated to the image. Across query period, the picture retrieval arrangement will accord from the user one or countless keywords that contain the find criteria. A keyword matching procedure is next gave to reclaim pictures associated alongside the keywords that match the find criteria.

“A picture is worth a thousand words"; this acquainted proverb emphasizes that discernible data is inherently unclear and semantically rich. The content of an im-period is far richer than what each set of keywords can express, just retaining text to delineate the content of the picture that frequently reasons ambiguity and inadequacy in giving an picture database find and query processing. This setback is due to the difficulty in enumerating precise words and phrases in delineating the content of pictures as the content of an picture is far richer than what each set of keywords can express. As the textual annotations are established on speech, variations in annotation will pose trials to picture retrieval. Comprehensive surveys of main text-based picture retrieval methods can be discovered.

1.2 Content Based Image Retrieval

Content-based picture retrieval (CBIR) [5] next has been utilized as an alternative to text established picture retrieval. IBM was the early, who seize an initiative by counseling query-by picture content (QBIC). QBIC industrialized at the IBM Scutiny Center is an open framework and progress technology. Unlike keywords established arrangement, discernible features for contents-based arrangement are removed from the picture itself. CBIR can be categorized established on the kind of features utilized for retrieval that might be whichever

low level or elevated level features. At main years, low level features contain colour (distribution of color intensity across image), sense (Homogeneity of discernible patterns), form (boundaries, or the interiors of objects delineated in the image), spatial relations (the connection or arrangement of low level features in space) or combination of above features were used. Finished Framework of Content established Picture Retrieval is shown in Fig. All pictures will experience the low level feature extraction procedure beforehand being added to the pictures database. In feature extraction period, features such as colour, form or sense are removed from the image. User provides a example picture and the similarity measurement engine is accountable in approximating the similarity amid the query picture and database pictures and next ranking them according to their similarity to the given query picture

Spatial feature is proved functional and competent in grating alongside supplementary low level features such as colour, form and sense to more rise the assurance in picture understanding. Although there are countless urbane algorithms to delineate color, form, sense and spatial features ways, these algorithms do not gratified and comfort to human perception.

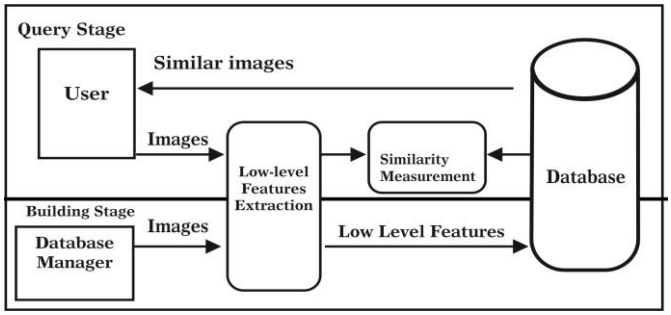

Figure 3: General Framework of Content Based Image Retrieval

This is generally due to the unavailability of low level picture features in delineating elevated level thoughts in the users' mind such as find an picture of a baby is weeping loudly. The merely method a contraption is able to present automatic extraction is by removing the low level features that embodied by the color, sense, form and spatial from pictures alongside a good degree of efficiency.

1.3 Semantic Based Image Retrieval

Neither a solitary features nor a combination of several discernible features might fully arrest elevated level believed of images. Besides, due to the presentation of Picture retrieval established on low level features are not satisfactory, there is a demand for the mainstream of the scutiny converges to retrieval established on semantic meaning by trying to remove the cognitive believed of a human to chart the low level picture features to elevated level believed (semantic gap). In supplement, representing picture content alongside semantic words permits users to admission pictures across text query that is extra intuitive, easier and favored by the front conclude users to express their mind difference alongside employing pictures [6]. For example, users' queries could be 'Find an picture of sunset rather than 'find me an picture encompasses red and yellow colors'. Finished Framework of Semantic established Picture Retrieval is shown in Fig. 4.

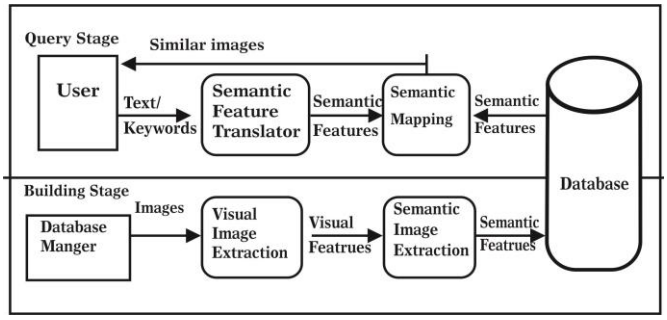

Figure 4: General Framework of Semantic Based Image Retrieval

All pictures demand to go across the discernible feature extraction procedure whereas it needs to remove the low level features of pictures to recognize meaningful and interesting regions/objects established on the comparable characteristics of the discernible features. Next, the object/region features will go into Semantic Picture

Extraction procedure to become the semantics description of pictures to be stored in database. Picture retrieval can be queried established on elevated level concept. User can have query established on a set of textual words. Then, it will go into semantic features translator to become the semantic features from the user query. The semantic mapping procedure is utilized to find the best believed to delineate the segmented region/objects established on the discernible features amid the query picture and database images. This mapping normally will be completed across oversee or unsupervised discovering instruments to associate the low level features alongside object believed and will be annotated alongside textual word across picture annotation procedure.

1.4 Applications

The police uphold picture database of convicts, offense scenes, and stolen items. In the health occupation, X-rays and scanned picture database are retained for diagnosis, monitoring, and scutiny purposes. In architectural and engineering design, picture database exists for design undertakings, finished undertakings, and contraption parts. In publishing and publicizing, journalists craft picture databases for assorted events and hobbies such as sports, constructions, personalities, nationwide and global events, and product advertisements. In past scutiny, picture databases are crafted for records in spans that contain arts, sociology, and medicine. In a tiny collection of pictures, easy browsing can recognize an image. This is not the case for colossal and varied collection of pictures, whereas the user encounters the picture retrieval problem. An picture retrieval setback is the setback encountered after hunting and reclaiming pictures that are relevant to a user's appeal from a database. To resolve this setback, text-based and content-based are the two methods adopted for find and retrieval in an picture database.

II. CHALLENGES

These are few Challenges that are still unsolved and researchers are trying to resolve these issues.

1. It is a challenging task, as understanding an picture and elucidating semantics is a tough thing for both humans and computers. Most comparable pictures are returned to the user as arrangement response. But this way has two main flaws due to that CBIR arrangements are incapable to produce good results. They locale the retrieved pictures on the center of distance or similarity alongside query picture and produce the results; in this stare they do not confirm their output. The setback alongside this way is that countless pictures could materialize as the reply pictures as they are not relevant at all
2. Secondly they don't consider the neighbors for the finalization of obtained results, Since simply color, texture and shape features cannot sufficiently delineate image semantics, semantic- based image retrieval is still an open problem.

Dimensionality It is a well-documented fact that the accuracy of retrieval result does not vitally rise as the dimensionality of a feature vector increases. Hughes has clarified that, alongside a finite database size, the accuracy of the result will plummet afterward a maximum top but next plummet afterward an optimal level is reached. This phenomenon contradicting one's intuitive assumption is recognized as the Plague of Dimensionality [7]. Presentation does not always enhance as the number of variables is arbitrarily increased. It could even deteriorate. It is vital to note that they have experimented alongside fluctuating scope and that at higher scope, the result of dimension reduction is extra prominent.

Scope: On the supplementary hand, the results of fluctuating scope have not consented as far attention. This is generally due to the nature of most normal CBIR arrangements that merely display retrieved pictures to the user. The last can plainly reclaim pictures till his expectations are met. Though, retrieved pictures can additionally be utilized for consecutive tasks. This includes arbitrating the meaning across RF and supplementary computationally intensive tasks. The redundant data from those irrelevant pictures is unusable and not desirable. It is therefore vital to illuminate a system's presentation below those assorted conditions.

Database Size The results of a fluctuating database size are reliant on the system's contraption discovering skills and the nature of the samples. In a contraption discovering nature, the CBIR accuracy increases alongside rising database size even though of the supplementary samples. On the supplementary hand, lacking contraption

discovering, the nature of the added examples is important. Reclaiming pictures on the basis of their pixel content stays mainly unsolved.

1.5 Research Issues

Remarkable observations in the review of related works are as follows:

(1) Despite the fact that Gabor filters are a widely acclaimed natural and excellent tool in segmentation , texture feature classification and extraction, only few CBIR systems utilize Gabor filters for texture feature extraction.

(2) In the existed CBIR systems, the texture features are obtained during segmentation from pixels or small blocks. Such features do not appropriately represent the properties of an entire region; thus it is required to study texture feature extraction from the whole region after segmentation.

(3) Even though RBIR systems enhanced the retrieval accuracy, they require high complex computations to calculate similarity, since these systems need to consider each sector in the database images, resulting in high retrieval response time. Thus, we requried a solution to decrease the number of database regions included in the similarity computation.

(4) The existing CBIR systems use either global features, or region based features to represent the content of an image. Each type of these features can be relevant in representing images with certain semantics. For example, global features are useful for retrieving textured images that doesn't have a specific regions in accordance to the user, such as natural scenes used as backgrounds. Thus, utilizing an integration of both types of features can improve the performance of the retrieval system.

III. CBIR METHODS

Color based Methods

Color is one of the most vital features of images. Color features are described subject to a particular color space or model. Every single pixel of the picture can be embodied as a point in a 3D color space. A number of color spaces have been utilized in works, such as RGB, LUV, HSV. There is no accord on that is the best. Though, one of the desirable characteristics of an appropriate color space for picture retrieval is its uniformity. Uniformity way that two color pairs that are equal in similarity distance in a color space are observed as equal by viewers. In supplementary words, the measured proximity amid the colors have to be undeviatingly connected to the psychological similarity amid them. RGB space is a extensively utilized color space for picture display. It is composed of three color constituent Red, Green and Blue. In HSV space is extensively utilized in computer graphics and is a extra intuitive method of delineating color. The three color constituents are hue, saturation (lightness) and worth (brightness). The hue is invariant to the adjustments in illumination and camera association and hence extra suited to object retrieval Once the color space is enumerated, color feature can be removed from pictures or regions. A number of vital color features have been counseled in the literatures, encompassing color histogram, color moments (CM), color coherence vector (CCV) and color correlogram, etc. Amid them, CM is one of the simplest yet extremely competent features.

Table 1 provides a summary of different color methods excerpted from the literature including their strengths and weaknesses

Technique	Description	Advantages	Disadvantages
Color Moment	Color moments have been prosperously utilized in countless retrieval arrangements, exceptionally after the picture encompasses just the object. The early order (mean), the subsequent (variance) and the third order (skewness) color moments are computed in this.	1. Effectual and competent in representing color allocations of images. 2. Very compact Representation as difference to supplementary feature extraction. 3. narrow down the find space	1. Not enough to describe all colors, no spatial info.
Color Histogram	The color histogram serves as an competent representation of the color content of an picture if the color outline is exceptional contrasted alongside the rest of the data set. Competent in describing both the globe and innate allocation of colors in an image. The allocation of the number of pixels for every single quantized bin, can be described for every single constituent in the color space.	1. It is robust to translation and rotation about the view axis. 2. Easy to compute.	1. Histogram alongside a colossal number of bins will rise the computational price (and additionally come to be improper for constructing effectual indexes for picture databases.) 2. Sensitive to noise.
CCV	It is a disparate method of incorporating spatial data into the color histogram, color coherence vector was proposed. Every single histogram bin is partitioned into two kinds, i.e., consistent, if it belongs to a colossal uniformly-colored span, or incoherent, if it does not. Picture is described as the vector.	1. Due to its additional spatial information, it has been shown better retrieval results than the color histogram.	1. High dimension. 2. High computation cost.
Gabor Filter	Gabor filter has been extensively utilized to remove picture features, exceptionally sense features. It is optimal in words of minimizing the combined uncertainty in space and frequency, and is frequently utilized as an orientation and scale tunable frontier and line (bar) detector. The picture is filtered alongside a bank of Gabor filters or Gabor wavelets of disparate favored spatial frequencies and orientations. Every single wavelet arrests power at a specific frequency and association that furnish a localized frequency as a feature vector.	1.Rotationally symmetric(perform the same in all directions) 2.One of the great advantage are the invariant properties of the extracted features.	1. It requires large space for storage.

Texture based Methods

Texture is one more vital property of images. Assorted sense representations have been investigated in outline credit and computer vision. Texture is a extremely functional characterization for a expansive scope of image. It is usually trusted that human discernible arrangements use sense for credit and interpretation. In finish, color is normally a pixel property as sense can merely be measured from a cluster of pixels. A colossal number of

methods have been counseled to remove sense features. Instituted on the area from that the sense feature is removed, they can be mainly categorized into spatial sense feature extraction methods and spectral sense Feature extraction methods. For the preceding way, sense features are removed by computing the pixel statistics or discovering the innate pixel constructions in early picture area, whereas the last transforms an picture into frequency area and Next computes feature from the transformed image. As the most public method for sense feature extraction, Fourier manipulation spectra, co-occurrence matrices, and multi-resolution filtering methods such as Gabor and wavelet change and wavelet transform. Table 1 summarizes their pros. and cons.

SHAPE based Methods

Shape features of objects or spans have been utilized in countless content-based picture retrieval systems. Contrasted alongside color and sense features, form features are normally delineated afterward pictures have been segmented into spans or objects. As robust and precise picture segmentation is tough to accomplish, the use of form features for picture retrieval has been manipulated to distinct requests whereas objects or spans are effortlessly available. The state-of-art methods for form description can be categorized into whichever boundary-based (rectilinear forms, polygonal approximation, finite agent models, and Fourier-based form descriptors) or region-based methods (statistical moments). A good form representation feature for an object ought to be invariant to translation, rotation and scaling. In this serving, we briefly delineate a little of these form features that have been usually utilized in picture retrieval requests.

Related work

Zhijie Zhao, et al in “Content Based Image Retrieval Scheme using Color, Texture and Shape Features”, 2016 [8]In this paper, proposed scheme is based on three noticeable algorithms: color distribution entropy(CDE), color level co-occurrence(CLCM) and invariant moments. The correlation of the color spatial distribution in an image is considered in CDE. CLCM matrix shows the texture feature of an image, which is a new proposed descriptor that is grounded on co-occurrence matrix to seize the modification of the texture. Hu invariant moments are frequently used owing to its invariance under translation, changes in scale, and also rotation. This proposed scheme achieves a modest retrieval result by utilizing these diverse and primitive image descriptors, at the similar time, the retrieval result is better when use the texture feature alone which they proposed than use gray level co-occurrence. The similarity measure matrix is based upon Euclidean distance.

J.K. Dash, et al in "Content-based image retrieval using fuzzy class membership and rules based on classifier confidence" 2015 [9] In this paper a novel texture image retrieval scheme is proposed. In this scheme more effort are spent on reducing the overall search time of newly proposed approach called Class membership-based Retrieval (CMR). Although classifier is used in CMR prior to the retrieval, it never exploits the classification confidence of the classifier to limit the search space to few of the output classes. This will identify the confidence in classification and use this information to limit the search space to single output class based on the confidence label and thereby reduces the overall search time.

Xiaofan Zhang, et al in “Towards Large-Scale Histopathological Image Analysis: Hashing-Based Image Retrieval”2015 [10] In this paper, they focus on developing scalable image-retrieval techniques to cope intelligently with massive histopathological images. Specifically, they represent a supervised kernel hashing technique which leverages a small amount of supervised information in learning to compress a 10,000-dimensional feature vector of image into only tens of binary bits with the informative signatures preserved. And then these binary codes are indexed into a hash table that enables real-time retrieval of images in a large size database. Critically, the supervised information is employed to bridge the semantic gap between low-level and high-level image features diagnostic information. In their future work, they will examine more types of features, especially those features stemming from segmentation and architectures. Furthermore, they will incorporate appropriate feature-fusion techniques to design a hybrid hashing method such that multiple types of features can be systematically fused to boost image-retrieval accuracy. They will also evaluate our framework in more applications in histopathological image analysis.

E. G. Karakasis, et al in “Image Moment Invariants as Local Features for Content Based Image Retrieval using the Bag-of-Visual-Words Model” 2015 [11] In this paper an image retrieval framework is presents that uses affine image moment invariants as descriptors of local image areas. In this a detailed feature vectors are generated by providing the construced moments into a Bag-of-Visual-Words representation. The image moment invariants selection has been done on the bases of compact representation of image areas as well as due to the ability to remain static under affine image transformations. In order to evaluate and discuss the overall approach three different setups were examined. The retrieval results are promising as compared with other extensively used local descriptors, allowing the proposed framework to present it as a reference point for future image moment local descriptors applied to the general task of content based image retrieval. they plan to extend the experiments, by replacing the SURF detector with different detectors and benchmark the descriptor with respect to other databases used in image retrieval research (e.g., Holidays Database (Jegou et al., 2008)). Furthermore, a different approach used in clustering and the study of alternative moment invariants constitute a new future research directions.

VUPPALA SWATHI, et al in “Feature Based Image Retrieval System using DWT and SVD” 2015 [12] In the past decade, more and more information has been published in computer readable formats. In the meanwhile, most of the information in journals, older books, and newspapers has been digitized and made computer readable. Big archives of music films, satellite pictures, images, newspapers, books and magazines have been made accessible for computer users. Internet facilitates this for the human to access this huge amount of informative data. The greatest challenge of the World Wide Web is that the huge amount of information is available about a given topic, because of that it becomes difficult to locate accurate and relevant information. Most users know what information they need, but are unsure where to find that information. Search engines can facilitate the ability of users to locate such relevant information.

IV. Conclusion and Future Scope

Image retrieval setbacks were early challenged alongside algorithms that endeavored to remove the discernible properties of a portrayal in a globe manner, pursuing the human intuition of assessing an image’s content. The experimentation alongside retrieval arrangements and the evaluation of their result employing both verbose pictures as well as pictures alongside occluded objects, displayed that the extraction of pictures salient spans lead to enhanced accurately consented result in the difference of employing the finished depiction. Thus, a representation of the picture by its points of attention proved to be a extra robust solution.

The competent content-based picture retrieval (CBIR) needs effectual extraction of low level features like color, sense and forms for indexing and fast query picture matching alongside indexed pictures for the retrieval of comparable images. Features are removed from pictures in pixel and compressed areas.

References

[1] Veltkamp, Remco, Hans Burkhardt, and Hans-Peter Kriegel, eds. “State-of-the-art in content-based image and video retrieval”. Vol. 22. Springer Science & Business Media, 2013.

[2] Rashedi, Esmat, Hossein Nezamabadi-Pour, and SaeidSaryazdi. "A simultaneous feature adaptation and feature selection method for content-based image retrieval systems." Knowledge-Based Systems 39 (2013): 85-94.0

[3] Liberti, Leo, CarlileLavor, Nelson Maculan, and Antonio Mucherino. "Euclidean distance geometry and applications." Siam Review 56, no. 1 (2014): 3-69.

[4] Rajam, I. Felci, and S. Valli. "A survey on content based image retrieval." Life Science Journal 10, no. 2 (2013): 2475-2487.

[5] Liu, Guang-Hai, and Jing-Yu Yang. "Content-based image retrieval using color difference histogram." Pattern Recognition 46, no. 1 (2013): 188-198.

[6] Hui, Wang Hui. "Performance and Analysis of the Automated Semantic Object and Spatial Relationships Extraction in Traffic Images." Advances in Computer Science: an International Journal 2, no. 4 (2013): 37-42.

[7] Har-Peled, Sariel, Piotr Indyk, and Rajeev Motwani. "Approximate Nearest Neighbor: Towards Removing the Curse of Dimensionality." Theory of computing 8, no. 1 (2012): 321-350.

[8] Wu, Lei, RongJin, and Anubhav K. Jain. "Tag completion for image retrieval." Pattern Analysis and Machine Intelligence, IEEE Transactions on 35, no. 3 (2013).

[9] Ezekiel, Soundararajan, Mark G. Alford, David Ferris, Eric Jones, Adnan Bubalo, Mark Gorniak, and Erik Blasch. "Multi-Scale Decomposition Tool for Content Based Image Retrieval." In Applied Imagery Pattern Recognition Workshop (AIPR): Sensing for Control and Augmentation, 2013 IEEE, pp. 1-5. IEEE, 2013.

[10] Kumar, Ashnil, Jinman Kim, WeidongCai, Michael Fulham, and Dagan Feng. "Content-based medical image retrieval: a survey of applications to multidimensional and multimodality data." Journal of digital imaging 26, no. 6 (2013).

[11] Desai, Padmashree, Jagadeesh Pujari, and Kamakshi Prasad. "Content Based Image Retrieval using Morphological Operators and BTC." (2013).

[12] Malik, Fazal, and BaharumBaharudin. "Analysis of distance metrics in content-based image retrieval using statistical quantized histogram texture features in the DCT domain." Journal of King Saud University-Computer and Information Sciences 25, no. 2 (2013).