

ISSN 1112-9867

Special Issue

Available online at <http://www.jfas.info>

AN ANALYSIS OF A DISCRETE COSINE TRANSFORM (DCT) COMPRESSION
TECHNIQUE ON LOW LEVEL FEATURES OF IMAGE FOR IMAGE RETRIEVAL
SYSTEM

M. A. M. Shukran^{*}, N. Abdullah, A. M. A. Zaidi, M. R. M. Isa and M. N. Ismail

Department of Computer Science, Faculty of Defense Science and Technology, National
Defence University of Malaysia, Sungai Besi Camp, 57000 Kuala Lumpur, Malaysia

Published online: 10 September 2017

ABSTRACT

A digital image is basically formed together with metadata which describes the image such as information about the file (e.g. location captured, author, date, time and etc.), tags and keywords. These properties allow images to be stored accordingly to its categories in a database. Those images later can be retrieved from the system by running a text-based query which explained the characteristics of an image. Although retrieving by text-based query method is a common and widely used to find images in a very large database, this method however does not guarantee that the results returned is as described in searching query due to the fact that it is totally dependent on the accuracy of stated image properties. In order to improve the accuracy of the text-based retrieval method, a Content-Based Image Retrieval (CBIR) has been proposed as a complement to the standard method.

Keywords: image compression; discrete cosine transformation (DCT); image processing; content-based image retrieval (CBIR) system.

Author Correspondence, e-mail: afizi@upnm.edu.my

doi: <http://dx.doi.org/10.4314/jfas.v9i3s.2>



1. INTRODUCTION

A single digital image may have contained vast number of information which might be relevant to the image or the other way around. Such information which was derived from metadata of an image file stores the details about the image such as file name, document type, date created, resolution and image dimension. This information is vital in retrieval processes using text-based approach as it would help to determine the images which was indexed by the search engine in a database. The search engine will have matched a search query with the indexed metadata such as keywords or tags that identifies lists of images with sorted relevancy.

This approach, however does not always returns relevant images since the results retrieved from the similarity matching of the search query are dependent on the accuracy of the image's metadata which might be generated automatically or by a human operator. Therefore, the text-based approach should be integrated together with a content-based method to establish an image retrieval system with high reliability and less overhead costs. When dealing with a content-based method, this absolutely means that the features of image which must be derived and used in complement with a searching keyword. Colour, texture and shape are a few features which can be extracted from an image and exhibits a very useful property that will be used in a similarity-matching process.

Colour is one of the major features in an image. This is why many image retrieval systems utilize colour feature extraction for retrieval purposes. Colour can be utilized as a feature by employing a colour space such as Red, Green and Blue (RGB) model which is then divided into n cells and creating histogram with n bins where each bin contains the number of pixels in each cell. A system will compare the query image's histogram with each model image's histogram based on a similarity metric function. Since digital images are sometimes equipped with redundant information and also made up of various colours, therefore storing those images into archive as well as processing a colour-rich image through feature extraction can be a time-consuming task. Hence, a digital image should pass through an image compression process to minimize unnecessary or any repeated information which cause the file to increase in size without degrading the quality of the image. There are two common types of image

compression which are known as lossy and lossless.

A digital image which has not been compressed is known as 'raw' image file. This raw format as the name suggest is a file which has not been processed and commonly produced by image input devices such as digital camera and image scanner. It is a modern form of 'photo negatives', whereby it contained all the information to produce a 'positive' image by adjusting the colour space to reduce its file size and usually resulted in a compressed form [1].

Lossy compression method like Joint Photographic Experts Group (abbreviated as JPEG) has been widely used on the Internet as a common image format, due to its capabilities of transforming a digital raw image's file size with significant reduction without degrading the image quality when the compression was first applied. The JPEG lossy compression can be performed by using a mathematical operation of the Discrete Cosine Transform (DCT), where the domain is transformed from a two-dimensional (2D) spatial domain to a frequency domain. The next sections of this paper will discuss in details some of the available image compression methods, followed by the application of the DCT in image processing and how both of these implementations have become the crucial factors in developing an image retrieval system.

2. METHODOLOGY

2.1. Feature Extraction Techniques-Colour Histogram

Colour histogram is a common method used to extract colour information from an image. The colours found in the image can be represented in a graphical way such as histogram. The histogram is used to show distributions of colour variables, in which it exhibits amount of pixels for each colour ranges with respects to the three-dimensional colour spaces such as RGB and HSV (hue, saturation and value). Likewise, the feature information of a greyscale image can also be represented using the same techniques. However, the pixels' amount is computed based on the intensities ranging from white to black rather than using the colour space. This is known as intensity histogram, which has about 256 different possible intensities for an 8-bit greyscale image.

Extracting colour or greyscale information using the histogram has some good points where it

has no effect to rotation, translation and scaling due to the pixel-wise properties of a digital image. This means that each individual pixel in an image will be processed on its own position at a given time. A colour histogram can be constructed by counting the number of pixel blocks with similar colours in an image. Since there might be a wide range of colours in an image which would take a long time to extract the pixel values, thus reducing the colour space into fewer number of bins through quantization will simplified the process [2].

An illustration of a greyscale image with extracted pixel values mapped onto a resampled image is shown in Fig. 1. The image was first pixelated into a 12x16 pixel blocks with the purpose on making it easier to portray the pixels that build up a digital image. Each pixel's intensities are represented by a single 8-bit number, whose range is from 0 (black) to 255 (white) [3].

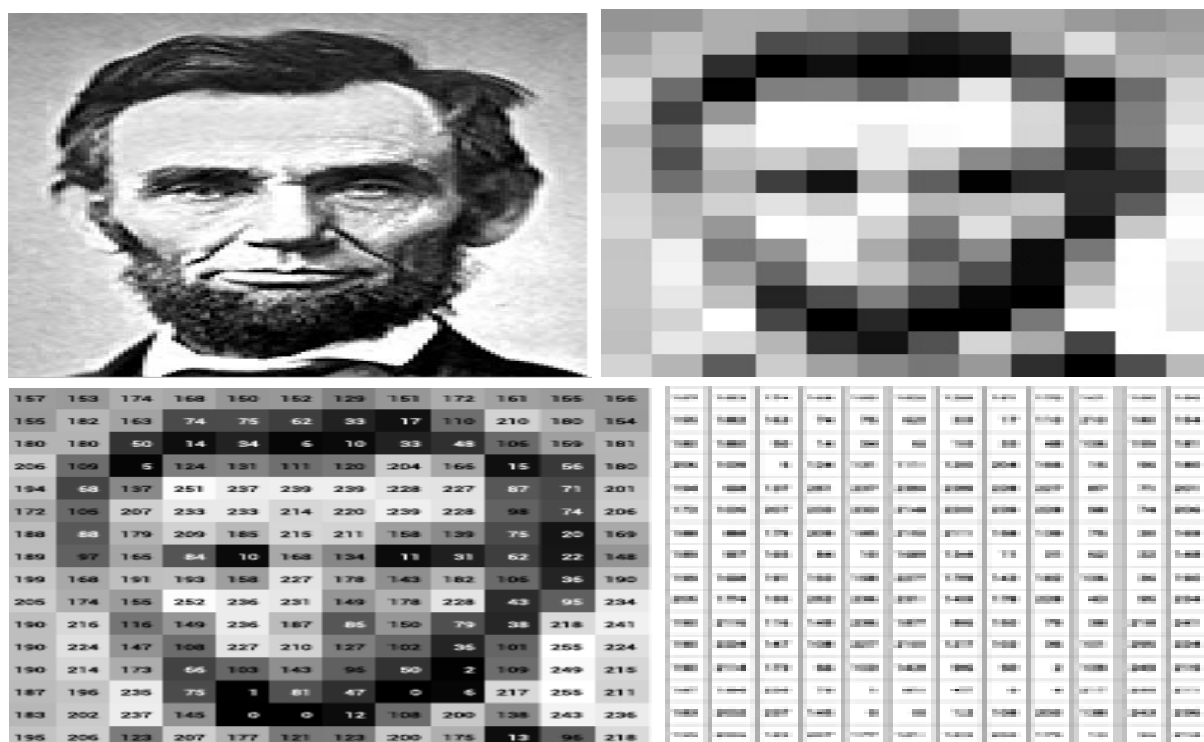


Fig.1. Greyscale image of Abraham Lincoln with mapped pixel values ranging from 0 to 255 based on 256 grey levels

As for the colour image, it is quite different from the greyscale since the image is formed by the three colour channels. In this case, the RGB channel with the range from 0 to 255 for each red, green and blue colours of an 8-bit image. The values gained from the extraction of the image below proves that each pixels are represented by the red, green and blue colours with

varying intensities. The combinations of those RGB colours will absolutely produce the exact colours of the following image.

Considering the original image with a resolution of 512x512 pixels, which is about 262,144 blocks of pixels available to be portrayed. Therefore, it is impractical to extract the values for all available pixels of the original image. With regard to this matter, the image was first reduced into an 8x8 pixel grids to make it clearer about the colour extraction in a colour image. Although this process will remove all the details from the original image and appears to be non-identical to the first image, this approach however will greatly explain how various intensities of the three colour channels would have resulted in broad array of colours in respective pixels' blocks

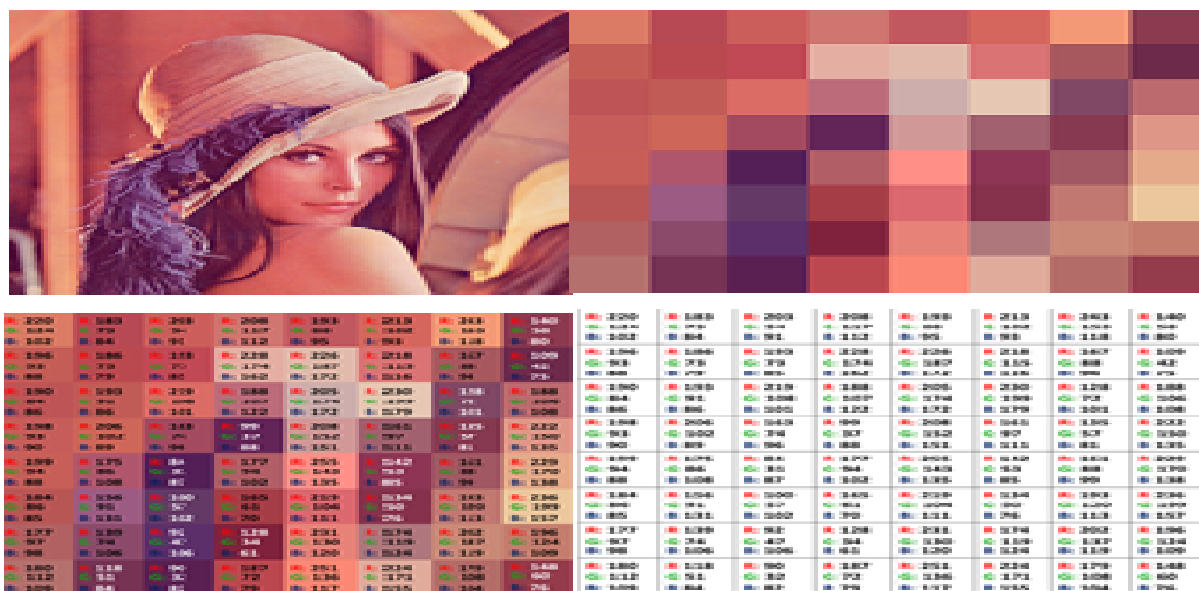


Fig.2. A colour image of Lena Söderberg with a varying values of three colour channels in each pixel blocks based on RGB model

2.2. Colour Quantization

Colour quantization refers to the process of reducing the number of colours available in an image into a smaller range. This process allows a new image to be produced with selected colour palette, hence ensuring that the image could be display despite of the colour depth of various devices such as desktop monitor, television and smartphone screen and also keeps the quantized image as identical as the original image. Based on the images shown in Fig. 2, the image was first undergoing a pre-processing techniques known as data binning process that

decrease the number of pixels from actual image into an 8x8 pixels blocks image. The data binning also combines several numbers of pixels to form a single and big pixel while keeping the actual dimensions of the actual image.

The histogram of the actual and quantized images was compared and obviously can be seen that the number of bins and the amount of pixels representing each colours in the image was greatly decreased in Fig. 3, as a result of data binning and colour quantization. Implementing these processes into a CBIR system will absolutely speed up the retrieval activity, but somehow it will negatively impact the accuracy of the search results as the details of that particular image was minimized significantly.

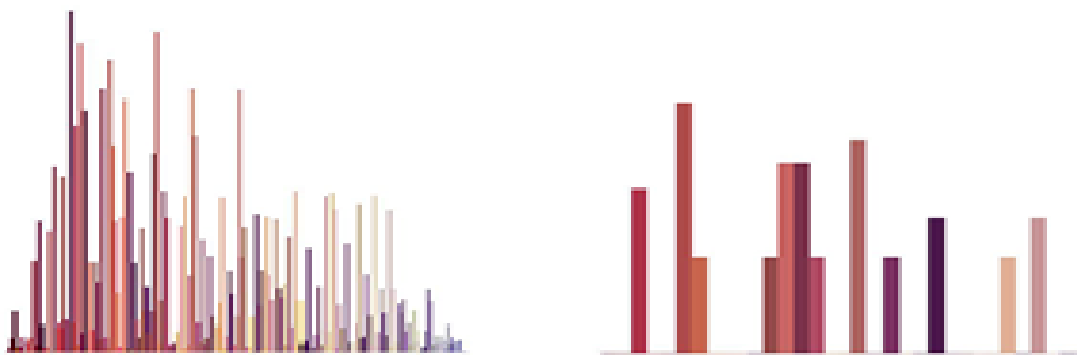


Fig.3. Colour distribution histograms for image in Fig. 2 before and after the quantization process

The colour histogram for each red, green and blue colour channels was combined into a single graph to show diverse formation of colours as a consequence of varying intensities from each RGB channels rather than representing it as separate individual channel. The greyscale images in Fig. 1 as opposed to the colour histogram was extracted to obtain the pixel values and visualized into an intensity histogram. There is not much different between the intensities of the actual image and the quantized image whereas the pixels' amount were highly reduced in the quantized image. The second histogram in Fig. 4 shows less number of pixels available for each available greyscale bins while maintaining the same intensity range for the sample image.



Fig.4. Greyscale intensities histograms for image in Fig. 1 with some pixels' number reduced in certain bins resulted from quantization process

3. RESULTS AND IDSCUSSION

3.1. Image Compression

A very large image database consists of millions collection of images, which are stored for various purposes. In medical application, images scanned from imaging modalities such as the Magnetic Resonance Imaging (MRI), ultrasound, tomography and radiography are archived in databases for clinical analysis and diagnosis. Meanwhile, in aerospace field, images are collected through imaging satellites for environmental observation, defence and security [17-18] and the earth observation [19]. A massive amount of data can be obtained from the image collections where valuable information could also be generated from that data, thus it is important to keep all the images in the database. The number of digital images in databases are increasing from day to day and there is an urgency to provide mechanisms that can handle this matter effectively. According to [4], an estimation of 1 Exabyte of data will be generated by the hospitals for the purpose of diagnostic imaging. The advancement of imaging technology in hospitals has led to a massive image archiving problems. The data storage capacity has always been a major concern since loads of images are stored into the database every day. As time passed, the capacity of the database storage eventually reached the maximum. Thus, more storage is required since the image grows bigger and there is a need to store all the images efficiently while keeping the costs lower.

Those problems pose a new set of challenge which includes an increase in time for retrieval process as well as the relevancy of image results with respect to the query.

The images stored in the database have to be compressed in order to allow more images to be

stored in image database. Therefore, providing appropriate image compression techniques would overcome those problems. As for now, there are two types of image compression techniques namely the lossless and lossy compressions. A lossless compression normally preserves every single bits of data from the original image file when it was decompressed and usually end up with a larger file size. Lossy compression on the other hand offers a smaller image size since some data was removed from the original image file [5]. An image compression system well-known as DCT-Based encoder comprises of three main blocks as illustrated in Fig. 5 [6]. The first block is known as transform block, where an image will be break into a non-overlapping 8x8 blocks preceded by a DCT coefficient blocks of the same size. The process continues with the next block, namely a quantization block where a large set of input values in a given range of an image are mapped to the same output that transforms the image into a smaller set of data. Finally, a lossless (entropy) encoder block was applied to the quantized image data whereby frequently occurring patterns are represented with a less number of bits while a rarely occurring patterns most likely are represented with more bits [6].

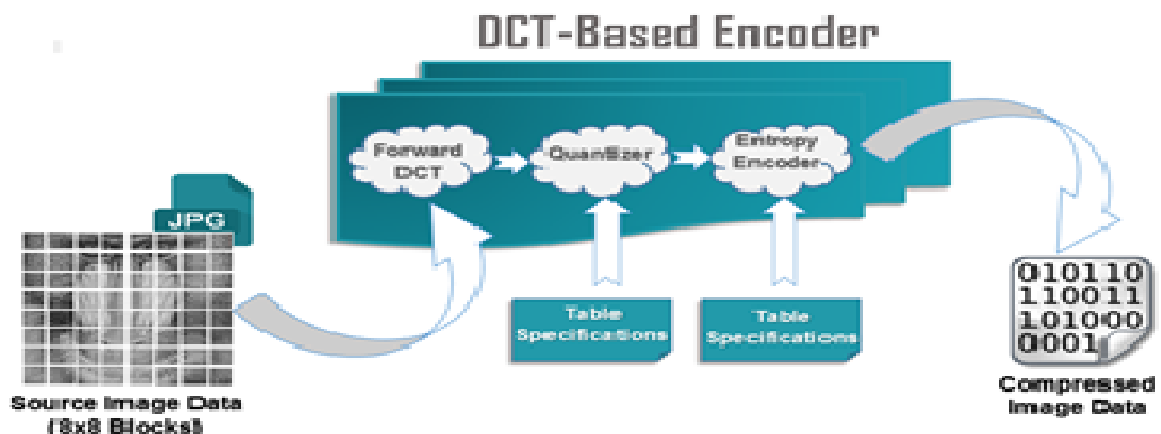


Fig.5. A DCT-Based encoder model

The JPEG, which originated from the name of the committee who proposed the standard, the Joint Photographic Experts Group is a form of lossy compression that was extensively used to compress images and the most image file format that is available on the World Wide Web. A mathematical algorithm which is related to the Fourier transform known as the Discrete Cosine Transform (DCT) was used in the JPEG lossy compression algorithm. The DCT compressed analog signals by discarding redundant information from the raw image file

carefully through quantization process to make a compact image with smaller size in terms of bytes [7]. The bits' number from an original image was reduced during the quantization process, in other words the number of colours used in an image was scaled-down so that the image can be view on limited colours display devices. There would always be a trade-off between image quality and file size when the JPEG method was used in image compression.

Similarly, the Portable Network Graphics (PNG) is also one of the most used image file on the internet, however the PNG used a lossless data compression algorithm compares to the JPEG image. The advantage of using a lossless compression over the lossy compression is that it is capable of reforming the original image from the compressed image file. The PNG image file was first pre-compressed using a prediction method called single filter method. The 2-stage PNG compression method was then followed by a DEFLATE compression algorithm, that was originated from LZ77 algorithm. During the PNG compression, the 'method 0' filter predicts the value of each pixel based on the values of previous neighbouring pixels and subtracts the predicted colours of the pixel from the actual value. While the JPEG file suffers from a digital generation loss due to the multiple decompression and compression processes, the PNG image file however does not degrade the quality of the original image even though the image was saved multiple times, exhibiting the properties of a lossless image format. Nevertheless, the lossy JPEG image file would always be smaller in size (in terms bytes) which is contrary to the PNG image size despite of the quality.

3.2. Discrete Cosine Transform (DCT)

Discrete Cosine Transform (DCT) as used in JPEG compression is a technique that convert data from a spatial domain into a frequency domain and very similar to Discrete Fourier Transform (DFT), but only use real numbers instead of complex numbers. The DCT also make use of Cosine functions and real coefficients rather than using both Sine and Cosine functions which are usually applied in the DFT. This technique was broadly used to compress image, audio and video file to make the size smaller by discarding high-frequency signals from a raw or uncompressed digital file. In image processing area, feature extraction of low-level features such as colour and texture can be performed on an image using the DCT. It has been proven that the DCT was able to de-correlate and compacting the energy of image

data very well [8]. Energy compaction is referring to the potential of compression techniques to package the energy of the spatial frequency possibly into lowest coefficients. A high compaction energy allows only a few coefficients to be transmitted rather than the entire set of pixels of an image. Due to the enormous amount of visual data stored in a compressed JPEG and MPEG format, this has induced research and application on image retrieval [9].

Since the DCT technique can be applied on a colour space with a single channel, thus any input that comes as colour image (usually with red, green and blue channels) should be converted into a grey-level image [7]. A grey-level digital image can be defined to be a function of two variables, $f(x, y)$ where x and y are spatial coordinates, and the amplitude f at a given pair of coordinates is called the intensity of the image at that point [9]. An example of DCT and DFT signals represented as amplitude spectrum for a sample image in Fig. 6 are as shown in the Fig. 7 and Fig. 8 respectively.



Fig.6. An image of a Dandelion flower with three RGB channels (left) and a converted grey-scale image (right)

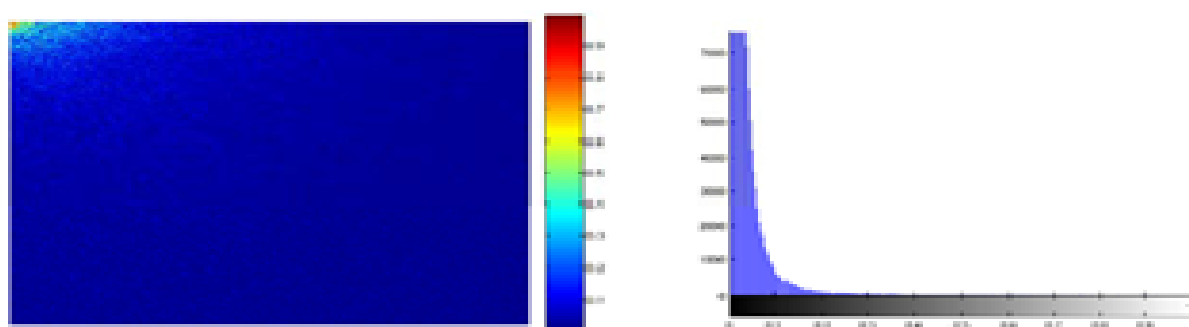


Fig.7. The amplitude spectrum of a grey-scaled Dandelion image in Fig. 6 using the DCT

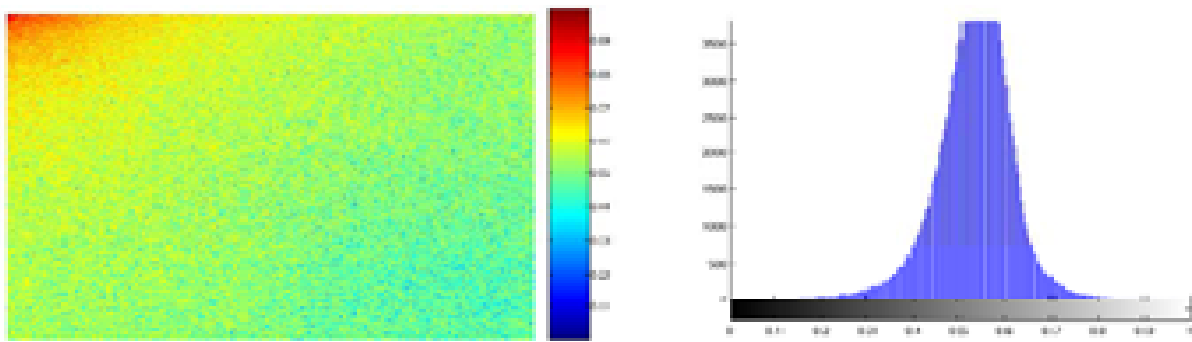


Fig.8. The amplitude spectrum of a grey-scaled Dandelion image in Fig. 6 using the DFT. It is clearly seen from Fig. 7 and Fig. 8 that the energy compaction is better in the DCT by comparing the concentrated histogram formed in both graphs. Hence, it can be concluded that the DCT has better energy compaction in comparison with the DFT because more coefficients are eliminated, which resulted in a greater number of bit savings for the same loss. The inverse DCT can be applied on compressed images, where the sequence of data can be reconstructed to obtain the original images.

3.3. Image Retrieval System

The applications on image databases have grown bigger in the past few decades, which give rise to a new problem on the retrieval process. The prime factors of the CBIR system development are to enhance the accuracy, speed, and the stability of the image retrieval system [11]. The accuracy of a CBIR system can be measured based on the similarity of a query image with the retrieved images from the database that are relevant to the search [7].

Meanwhile, the stability of a retrieval system is dependent on the capability of the CBIR system to perform queries on various image inputs. The speed not to be missed can be determined by evaluating the time required for the CBIR system to produce relevant results in real time. Often both color and shape features were combined to analyze the similarity of images in the CBIR system. Similarities of a query image with the images in a database are decided from the metric between two images.

The color content in images has become the basis when developing a CBIR system. This is because the pixels' value found in a digital image can be very useful when it comes to the computation process. It can be done swiftly and much simpler as compared to the shape and texture features. Color is also an intuitive feature and vital in the similarity matching process. It

is compulsory to have knowledge of the colour properties and its representation in digital images since the reliability of a colour extraction is dependent on those concept [12]. Developing a CBIR system based on colours can be a lengthy task at the beginning. The main reason why it does takes a lot of effort and time is because there is a need for constructing a database of images which stores extracted colours information from a very large collection of images.

This is an essential part of development process since the extracted features information of the image collection will be compared to the query, whereby in this case the query image will be extracted first before the similarity matching phase. During the similarity matching process, the Euclidean distance of the query image will be computed as well as the collection of images in the database. This metric helps to determine possible sets of images, which are completely or almost identical to the query image. The query results are determined by the shortest Euclidean distance found between the query image and the images in the database. If the images in a database have a shorter Euclidean distance, this means those images might share commonalities at some point with respect to the query image. Therefore, the resulting images can be considered as relevant.

The purpose of similarity measures through Euclidean distance in a process known as similarity matching as executed in CBIR system is to elicit the degree of relevance of images in a searching results based on the query image.

Similarity matching in a CBIR model as depicted in Fig. 9 is a key component of a content-based image retrieval (CBIR) system because retrieving a collection of images, which are nearly identical to the query image is its main objective. In an ideal world, image indexing and retrieval systems would not only be able to identify and index relevant features of an image but also be able to retrieve only relevant images matching a query by a human user [13]

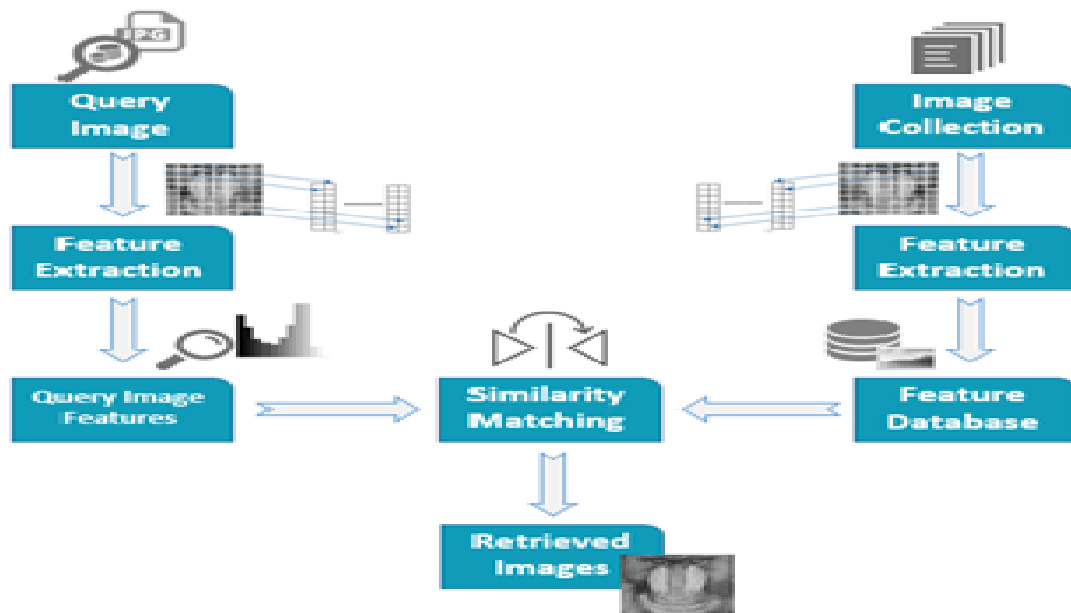


Fig.9. A Content-Based Image Retrieval (CBIR) model

Feature extraction using the DCT coefficients produced an efficient searching results that matched with the query. When using the DCT, a JPEG image will be divided into an 8x8 blocks of pixels which is then transformed into a frequency domain. The individual block of the image is then represented by 64 components known as DCT coefficients $\{DC, AC1, \dots, AC63\}$ [11].

During the compression, the upper coefficients which contained detailed information of the image will be eliminated whereas all essential and general information known as lower DCT coefficients remained in the compressed image [10]. Improvement over the CBIR system can be done with the integration of the semantic-based image retrieval method. Since there are underlying properties of image such as tags and keywords describing the characteristics about the image, thus it can be used as a complement to the similarity matching phase while retrieving images from the database. This will definitely help to increase the accuracy of the query results, as well as shortening the time required for retrieval and also reducing the number of irrelevant images shown by the system. Text-based search provides results with semantic similarity, while content-based search provides results with visual similarity [14]. During the similarity matching phase, low-level image features (which is the colour in this context) will be used to match with any similar or exact image from the database followed by looking on the metadata of resulted images.

The idea behind the introduction of text-based searching in a CBIR system arise because of the human perception when perceiving images from a large collection. The CBIR system, which is normally run by an automated computer program is using the low-level features to match the query image, where the process is not visible to the human. The application of digital image processing in image retrieval has made a broad gap between a human perception and computer process. This is called as semantic gap where the high-level perceptions of human might be conflicting with the low-level feature known to the computer [15]. The distinction between the contextual knowledge of a human and the computational representation of a computer was shown in Fig. 10 [16]. Hence, it can be deduced that the Text-Based Image Retrieval (TBIR) method would be the best companion when developing a Content-Based Image Retrieval (CBIR) system.

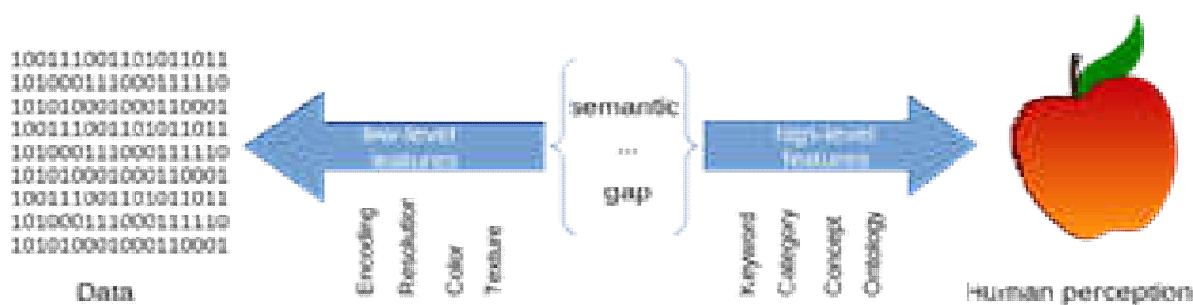


Fig.10. An illustration of semantic gap between human perception and computer interpretation

4. CONCLUSION

The evolution of image retrieval system embarks as there is an increase on the numbers of applications which use a very large image database for specific purposes. Considering that there is an increase on demand for a Content-Based Image Retrieval rather the conventional Text-Based Image Retrieval which might be bias, therefore a reliable method for feature extraction must be applied upon the development of the system. Since the DCT is one of the technique which has better compression result which removes all redundant information from a raw image, therefore the approach must be implemented to the CBIR system which would help to further improves the retrieval process. Feature extraction from digital images can be

performed by using a colour histogram technique. An image has a rich content of pixels' value which formed the colours and structures of a digital image. These pixel values are very meaningful as it will be used to make a similarity matching in a retrieval process.

Retrieving images from a very large scale database has become a big challenge in information technology (IT) world. The emergence of automated retrieval system such as a CBIR system has improved the image retrieval process whereby the low-level features or simply known as contents of an image was used instead of a text-based approach. However, another issue arises when it comes to the human perception throughout the retrieval process as there seems to be a barrier between human and computer when perceiving the image [20]. Therefore, a text-based method should be a complementary to the content-based method as to ensure that the CBIR system could produce query results with high accuracy in an acceptable period of time.

5. REFERENCES

- [1] Verhoeven G J. It's all about the format-Unleashing the power of RAW aerial photography. *International Journal of Remote Sensing*, 2010, 31(8):2009-2042
- [2] Katira C, Vora N, Wali K, Medhekar A. Advanced content based image retrieval using multiple feature extraction. *International Journal of Innovative Research in Science, Engineering and Technology*, 2015, 4(10):9805-9812
- [3] Nixon M. S., Aguado A. S. *Feature extraction and image processing for computer vision*. Massachusetts: Academic Press, 2012
- [4] Frost and Sullivan. *U.S. medical imaging informatics industry reconnects with growth in the enterprise image archiving market.* : Frost and Sullivan, 2012
- [5] Vijayvargiya G, Silakari S, Pandey R. A survey: various techniques of image compression. *International Journal of Computer Science and Information Security*, 2013, 11(10):11-19
- [6] Abomhara M, Khalifa O O, Zakaria O, Zaidan A A, Zaidan B B, Rame A. Video compression techniques: An overview. *Journal of Applied Sciences*, 2010, 10:1834-1840
- [7] Shukran M A, Chung Y Y, Yeh W C, Wahid N, Zaidi A M. Artificial bee colony based data mining algorithms for classification tasks. *Modern Applied Science*, 2011, 5(4):217-231
- [8] Xiong Z, Ramchandran K, Orchard M T, Zhang Y Q. A comparative study of DCT-and

wavelet-based image coding. *IEEE Transactions on Circuits and Systems for Video Technology*, 1999, 9(5):692-695

[9] Tsai T, Huang Y P, Chiang T W. Dominant feature extraction in block-DCT domain. In *IEEE International Conference on Systems, Man and Cybernetics*, 2006, pp. 3623-3628

[10] Shukran M A, Chung Y Y, Yeh W C, Wahid N, Zaidi A M. Image classification technique using modified particle swarm optimization. *Modern Applied Science*, 2011, 5(5):150-164

[11] Lin C H, Chen A L, Kuo T C, Tsay C Y. An efficient approach on content-based image retrieval with DCT coefficient indexes. In *Symposium on Advanced Database Systems for Integration of Media and User Environments*, 1999

[12] Singha M, Hemachandran K. Content based image retrieval using color and texture. *Signal and Image Processing*, 2012, 3(1):39-57

[13] Rorissa A. Image retrieval: Benchmarking visual information indexing and retrieval systems. *Bulletin of the Association for Information Science and Technology*, 2007, 33(3):15-17

[14] Barrios J M, Díaz-Espinoza D, Bustos B. Text-based and content-based image retrieval on Flickr. In *IEEE 2nd International Workshop on Similarity Search and Applications*, 2009, pp. 156-157

[15] Alzu'bi A, Amira A, Ramzan N. Semantic content-based image retrieval: A comprehensive study. *Journal of Visual Communication and Image Representation*, 2015, 32:20-54

[16] Wikipedia, Semantic gap. 2017. https://en.wikipedia.org/wiki/Semantic_gap

[17] Zaman F H, Sulaiman A A, Yassin I M, Tahir N M, Rizman Z I. Development of mobile face verification based on locally normalized Gabor wavelets. *International Journal on Advanced Science, Engineering and Information Technology*, 2017, 7(4)

[18] Zaman F H, Ali M H, Shafie A A, Rizman Z I. Efficient human motion detection with adaptive background for vision-based security system. *International Journal on Advanced Science, Engineering and Information Technology*, 2017, 7(3):1026-1031

[19] Nur A Z, Mohamad H J, Siti Z A Z, Zairi I R, Development of space weather monitoring platform for space and earth's electromagnetism observation. *ARNP Journal of Engineering*

and Applied Sciences, 2017, 12(10):3308-3311

[20] Norizan M, Zairi I R, Mazidah P, Nur H R H, Defining image attributes for frames filtering rule. Journal of Basic and Applied Scientific Research, 2013, 3(10):381-389

How to cite this article:

Shukran MAM, Abdullah N, Zaidi AMA, Isa MRM, Ismail MN. An analysis of a discrete cosine transform (dct) compression technique on low level features of image for image retrieval system. J. Fundam. Appl. Sci., 2017, 9(3S), 11-27.