ISSN:2229-6093



# Image Retrieval Based on Combined features of DCT and Shape Descriptor

*Arvind Sharma*
*M.Tech(CTA), SOIT'*
*Rajiv Gandhi Technical,*
*University, Bhopal ,(M.P.)*
arvindks82@gmail.com

*Dr. Sanjeev Sharma,*
*HOD,School of Information Technology,*
RGPV ,Bhopal (M.P.)
sanjeev@rgtu.net

*Nishchol Mishra*
*Asst. Professor, SOIT,*
*Rajiv Gandhi Technical*
*University, Bhopal(M.P.)*
nishchol@rgtu.net

**Abstract:** *In the image retrieval technology, the aim is to retrieve an image based on combined features of DCT color coefficients and shape descriptor. There were many algorithms based on single feature and combined features, presented in past. Those algorithms have lack of speed and require more computational capability. In this algorithm by combining two features of image, most similar images are retrieved. The idea of DCT was to decouples the color component of image using YCbCr and transform these into DCT coefficients and measure the similarity moreover extract the region based shape descriptor to calculate central moment of an image with the help of edge information and morphological operations to find normalized feature vectors. Now, calculate similarity with combined features of quantized DCT color coefficients and normalized feature vectors to retrieve an image.*

**Key Words:** Discrete Cosine Transformation (DCT), Edge detection, Morphological operations, Shape analysis, Central moment.

## I. INTRODUCTION

Content Based Image Retrieval is a set of techniques for retrieving relevant images from an image database based on automatically derived image features. One of the main tasks for CBIR systems is similarity comparison, extracting feature signatures of every image based on its pixel values and defining rules for comparing images [3]. These features become the image representation for measuring similarity with other images in the database [1]. As processors become increasingly powerful, and memories become increasingly cheaper, the deployment of large image databases for a variety of applications have now become realizable for example, geography, medicine, architecture, advertising, design, fashion, publishing. CBIR is the retrieval of image based on visual features such as color, texture and shape. It involves two steps [2]:

(1)Feature Extraction: In this, extract image features to a distinguishable extent.

(2)Matching: This step involves matching these features to yield a result that is visually similar.

Such a set of feature attributes are called feature vector or feature descriptor. In this paper, we have presented an elegant and effective system for content-based image retrieval. The system exploits the global and regional features of the images. The images are extracted using DCT and central moment. Combine the feature vectors of both descriptors to calculate similarity and retrieval of images. The Fig.1 shows a block diagram of the proposed system.

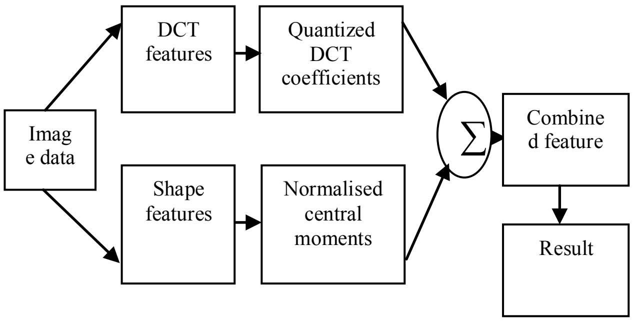


***Figure1 Image retrieval system with combined features***

## II. EXTRACTION OF DCT COEFFICIENTS

Color information of image transformation into DCT coefficients has been proved to be a quick and effective method for image matching [4]. It transforms the input data into a format to reduce inter-pixel redundancies in the input image. A color space conversion from RGB to YcbCr is done and linear mathematical transformation is applied to map the pixel values onto a set of coefficients, which are then quantized.

*A. Conversion from RGB to YCbCr color space*

Medical research proved that the human eyes have different sensitivity to color and brightness. The eye is more sensitive to changes in brightness than changes in color [4]. YcbCr represents color as brightness and two color difference signals. In YCbCr, the Y is the brightness (luma), Cb is blue minus luma (B-Y) and Cr is red minus luma (R-Y). The transformation of RGB color model to YCbCr color model can be derived as a linear transformation of the RGB color space [6]:

$$Y = 0.299R + 0.587G + 0.114B$$

$$Cb = -0.1687R - 0.3313G + 0.5B$$

$$Cr = 0.5R - 0.4187G - 0.0813B$$

The conversion is done in order to reduce the computational load and it decouples intensity and color informations.

*B. DCT Coefficient Transformation*

The discrete cosine transform is a fast transform. It has excellent compaction for highly correlated data and gives good compromise between information packing ability and computational complexity. In this paper by using DCT, minimize the number of bits required to represent the information in an image, by removing the redundancy between neighbouring pixel values [5]. DCT transforms each color component into DCT coefficients by using following [6]:

$$C(u,v) = \alpha(u)\alpha(v)\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\cos\left[\frac{\pi(2x+1)u}{2N}\right]\cos\left[\frac{\pi(2y+1)v}{2N}\right]$$

$$\alpha(u) = \begin{cases} \sqrt{\frac{1}{N}} & for \quad u = 0 \\ \sqrt{\frac{2}{N}} & for \quad u \neq 0. \end{cases}$$

Where,

- u is the horizontal spatial frequency,
- v is the vertical spatial frequency,
- f(x, y) is the pixel value at coordinates (x, y),
- C(u, v)is the DCT coefficient at coordinates(u, v).

The result of the DCT transformed image is quantized by reducing the high frequency pixels simply by dividing it by a constant quantization matrix and then rounding it. A lot of the values will become too small [6]. All the quantized DCT coefficients are stored in an array of feature vector.

## III. EXTRACTION OF SHAPE FEATURE

Shape representations can be generally divided into two categories [7]:

- Boundary-based, and
- Region-based.

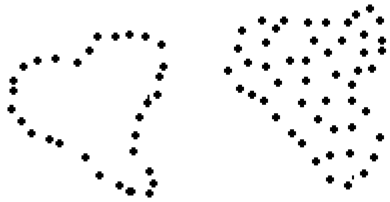

***Figure2: Boundary-based & Region- based***

Boundary-based shape representation only uses the outer boundary of the shape, i.e., the pixels along the object boundary. Region-based shape representation uses the entire shape region by describing the considered region using its internal characteristics; i.e., the pixels contained in that region [8, 9]. The most successful representations for shape categories are Fourier Descriptor and moment invariant. The main idea of central moment is to use region-based moments of centroid of image, which are invariant to transformations [10]. In this paper, we use a shape descriptor in terms of central moment, computed on the region of binary image, which can be extracted by applying edge detection method and morphological operations. Then calculate the features of the image in terms of normalized central moment.

*A. Extraction of edge image*

For the method parameter, we select SOBEL edge detection. The block can perform a threshold operation on the gradient magnitudes and output a binary image, which is a matrix of boolean values. If a pixel value is 1, it is an edge. In order to suppress noise, a certain weight is correspondingly increased on the centre point, and its digital gradient approximation equations may describe as follows [11, 12]:

Generally, the size of its gradient:

$$G(x, y) = (Gx^2 + Gy^2)^{1/2}$$

Can also adopt similar to that:

$$G(x, y) = |Gx^2| + |Gy^2|$$

The total gradient value G may get the edge by threshold method.

ISSN:2229-6093



*B. Morphological operations*

Morphological image processing (morphology) describes a range of image processing techniques that deal with the shape of features in an image. Morphological operations such as erosion, dilation, opening, and closing, often, you need to use a combination of these to perform your morphological image analysis. Dilation and erosion are basic morphological operations. First, define A as the reference image and B is the structure image used to process A[13]. Dilation is defined by the equation:

$$A \oplus B = \{z \mid [(\hat{B})_Z \cap A] \subseteq A\}$$

Where $B\hat{}$ is B rotated about the origin. Dilation has many uses but a major one is bridging gaps in an image due to the fact that B is expanding the features of A. Erosion on the other hand can be considered a narrowing of features on an image. Again defining A as the reference image and B as the structure image:

$$A \ominus B \ \{z \mid (B)_Z \subseteq A\}$$

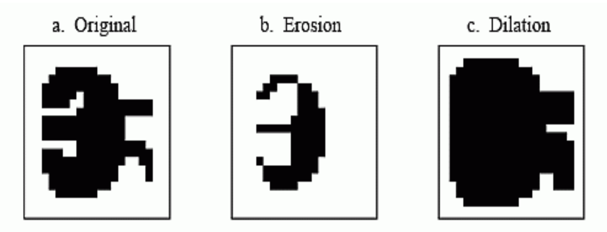


**Fig 3 Erosion (b) and dilation (c) of image (a)**

*C. Central moment*

Moments have been frequently used as features for image processing, shape recognition and classification. It was Hu, who first set out the mathematical foundation for two-dimensional moment invariants and demonstrated their applications to shape recognition [7]. Hu defines seven set of feature vectors computed from central moments through order three. Translation invariance is achieved by computing moments that are normalized with respect to the centre of gravity so that the centre of mass of the distribution is at the origin (central moments). From the second and third order values of the normalized central moments a set of seven invariant moments can be computed which are independent of rotation. The moments used to construct the moment invariants are defined in the continuous but for practical implementation they are computed in the discrete form. Given a function *f(x, y)*, these regular moments are defined by [7, 10]:

$$M_{pq} = \iint x^p y^q f(x, y)dxdy \qquad \ldots\ldots\ldots (1)$$

$M_{pq}$ is the two-dimensional moment of function *f(x,y)*. The order of the moment is (p + q), where p and q are both natural numbers. For implementation in digital from this becomes:

$$M_{pq} = \sum_x \sum_y x^p y^q \ f(x, y) \qquad \ldots\ldots\ldots (2)$$

To normalize for translation in the image plane, the image centroid is used to define the central moments. The co-ordinates of the centre of gravity of image are calculated using equation (2) and are given by:

$$\bar{x} = M_{10} / M_{00} \qquad \bar{y} = M_{01} / M_{00} \qquad \ldots\ldots\ldots (3)$$

The central moments can then be defined in their discrete representation as:

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q \qquad \ldots\ldots\ldots (4)$$

The moments are further normalised for the effects of change of scale using the following formula:

$$\eta_{pq} = \mu_{pq} / \mu_{00}{}^{\gamma} \qquad \ldots\ldots\ldots (5)$$

Where the normalization factor: $\gamma = (p + q/2) + 1$. From the normalized central moments a set of seven values can be calculated and are defined by:

$$\Phi_1 = \eta_{20} + \eta_{02}$$

$$\Phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta^2_{11}$$

$$\Phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$

$$\Phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{03} + \eta_{21})^2$$

$$\Phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})^2 [(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \times [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

$$\Phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$$

$$\Phi_7 = (3\eta_{21} - \eta_{03})^2 (\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{30})^2 (\eta_{21} + \eta_{03}) \times [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{30})^2]$$

Where $\Phi_i$, $1 \leq i \leq 7$, set of feature vectors.

ISSN:2229-6093



## IV. PROPOSED WORK

***Step:1)*** Input image from the database.

***Step:2)*** Convert this true color image into YCbCr color space by a linear transformation of RGB color space.

***Step:3)*** Partition an image into 8x8 blocks of pixel and level shifting is done to work in the range of DCT.

***Step:4)*** Calculate the DCT coefficients for each color component of YCbCr image, luma ,chroma (red, blue) by using 8x8 DCT equation.

$$C(u,v) = \alpha(u)\alpha(v)\sum_{x=0}^{7}\sum_{y=0}^{7} f(x,y)\cos\left[\frac{\pi(2x+1)u}{16}\right]\cos\left[\frac{\pi(2y+1)v}{16}\right]$$

$$\alpha(u) = \begin{cases} \sqrt{\frac{1}{8}} & for \quad u = 0 \\ \sqrt{\frac{2}{8}} & for \quad u \neq 0. \end{cases}$$

***Step: 5)*** Quantization is applied simply by dividing with a constant quantization matrix for luma and chroma components.

***Step:6)*** Measure the similarity between query image and database image based on quantized DCT coefficients and store them in array of vector to further combine with the feature vectors of shape descriptor.

***Step: 7)*** Convert again query image into gray scale to reduce complexity by reducing hue and saturation color information.

***Step: 8)*** Apply SOBEL edge detection method to find the edges where the gradient of an image is maximum. Generally, the size of its gradient:

$$G(x, y) = (Gx^2 + Gy^2)1/2$$

***Step: 9)*** Perform morphological operations:
(1) Dilation - To remove gaps in edges,
(2) Erosion - To remove the noise and smooth the image object,
(3) Filling - To fill the holes in the region.

***Step:10)*** Calculate central moment, described above section, by equations (1-5) .

***Step:11)*** Measure the similarity between query image and database image based on shape descriptor.

***Step:12)*** Finally, combine these feature vectors with the quantized DCT coefficients for matching. So we have common image, found by both feature.

## V. RESULT AND ANALYSIS

**Figure: Image by combining both features**

| Images | DCT values | | |
|---|---|---|---|
| | luma | blue | Red |
| Fishgirl.jpeg | 53 | 6 | 23 |
| Basketball.jpeg | 53 | 6 | 23 |
| Flower1.jpeg | 53 | 5 | 29 |
| Apple2.jpeg | 53 | 5 | 17 |
| Panathin.jpeg | 53 | 6 | 21 |
| Bee.jpeg | 53 | 6 | 27 |

**Table 1 Quantized DCT values**

| Normalized feature vectors | Image1 | Image2 |
|---|---|---|
| | Basketball .jpeg | Panathin.jpeg |
| $\Phi_1$ | 1.8301 | 1.8353 |
| $\Phi_2$ | 12.7972 | 11.2705 |
| $\Phi_3$ | 13.7791 | 13.3348 |
| $\Phi_4$ | 20.9780 | 21.9495 |
| $\Phi_5$ | 38.4985 | 40.2267 |
| $\Phi_6$ | 27.7140 | 28.0587 |
| $\Phi_7$ | 39.1819 | 39.7565 |

**Table 2 Feature vectors by central moments of image**

**Analysis:**

Values are shown in table 1, the query image have the values (53, 6, 23). It searches the database images based on these values. The images, those have nearest values of query image, will be similar images. As we have shown in table: fishgirl.jpeg, flower1.jpeg, apple2.jpeg, panathin.jpeg and bee.jpeg, all have the nearest values (53, 6, 23), (53, 5, 29), (53, 5, 17), (53, 6, 21), (53, 6, 27) respectively. In table2 all 7 feature vectors are calculated corresponding to query image and database image that has small difference in these feature vectors. In table we can see the query image basketball.jpeg have the values (1.8301, 12.7972, 13.7791, 20.9780, 38.4985, 27.7140, 39.1819) and its similar image Panathin.jpeg have the values (1.8353, 11.2705, 13.3348, 21.9495, 40.2267, 28.0587, 39.7565) that has small difference.

Finally, Combine both the DCT based and shape based feature vectors. Sometimes in DCT based matching we have the same maximum value for each matrices for luminance and both color value as shown in table basketball.jpeg and fishgirl.jpeg have the same coefficients (53, 6, 23). That is why, shape based matching is applied to refine the result, obtained from the DCT based matching. It is hardly possible that the two images have the same feature vectors by normalizing the central moment of images but, there are some chances for very large image database, so we combine both the descriptors to similarity measure.

## IV. CONCLUSION AND FUTURE WORK

In this paper, according to the DCT coefficients of an image and normalized central moments combined algorithm are proposed. To some extent, we have shown in the experiments that the performance of the proposed system is better than those based on only color and only on shape features respectively. Although the retrieval result based on the combined low level features is better than only based on single feature, to increase the robustness of the proposed system, a more sophisticated matching algorithm is currently under development. We are also constantly adding new images to the database in attempt to better evaluate the system's performance. It is widely recognized that most current content-based image retrieval systems work with low level features, such as color, texture and shape, and that next generation systems should operate at a higher semantic level. Although this is difficult in general, it should be feasible for specific domains.

## REFERENCES

[1] "Content-Based Image Retrieval Using Color and Edge Direction Features" Jianlin Zhang School of Software Nanchang University Nanchang, Wensheng Zou Institute of Computer Technology Engineering Nanchang University Nanchang, 978-1-4244- 5848-6/10 , 2010 IEEE.

[2] "A Review of Content-Based Image Retrieval", G. Rafiee, S.S. Dlay, and W.L. Woo School of Electrical, Electronic and Computer Engineering, Newcastle University, England, United Kingdom, 2010 IEEE.

[3]. Huiyu Zhou*, Abdul H. Sadka, Mohammad R. Swash, Jawid Azizi and Abubakar S.Umar, School of Engineering and Design, Brunel University, "Content Based Image Retrieval and Clustering: A Brief Survey" Recent Patents on Electrical Engineering 2009, Bentham Science Publishers Ltd 2, 187-199 187.

[4] THE 2D-DCT COEFFICIENT STATISTICAL BEHAVIOUR: A COMPARATIVE ANALYSIS ON DIFFERENT TYPES OF IMAGE SEQUENCES M. Mitreaa,b*, F.Prêteuxb, A. Vlada,c, C. Fetitab aFaculty of Electronics and Telecommunications, "Politehnica" University, Bucharest–Romania bARTEMIS Project Unit, GET / INT, Evry–France cThe Research Institute for Artificial Intelligence, Romanian, Academy Journal of Optoelectronics and Advanced Materials Vol. 6, No. 1, March 2004, p. 95 – 102.

[5] "Image Retrieval Using Local Compact DCT-based Representation" ˇ Stˇep´an Obdrˇz´alek1 and Jiˇr´ı Matas1,2 1 Center for Machine Perception, Czech Technical University, Prague, CZ 2 Centre for Vision Speech and Signal Processing, University of Surrey, Guildford, UKDAGM'03, 25th Pattern Recognition Symposium September 10-12, 2003, Magdeburg, Germany Springer-Verlag.

[6] "Enhancement of Color Images by Scaling the DCT Coefficients", Jayanta Mukherjee, Senior Member, IEEE, and Sanjit K. Mitra, Life Fellow, IEEE, IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 17, NO. 10, OCTOBER 2008.

[7] "Review of shape representation and description techniques", Dengsheng Zhang∗, Guojun Lu Gippsland School of Computing and Info. Tech., Monash University, Churchill, Vic 3842, Australia, 16 July 2003.

[8] "Shape Classification Using the Inner-Distance" Haibin Ling, Student Member, IEEE, and David W. Jacobs, Member, IEEE, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 29, NO. 2, FEBRUARY 2007.

[9] Zyad Shaaban and Thawar Arif, Faculty of Information Technology, Applied Science University, Amman , "Shape Representation and Retrieval Using Radial Basis Probabilistic Neural Networks", Jordan IJCSNS International Journal of Computer Science and Network Security, VOL.8 No.12, December 2008..

[10] "Shape-based Image Retrieval" RYSZARD S. CHORA´S University of Technology & Life Sciences Faculty of Telecommunications & Electrical Engineering S. Kaliskiego 7, 85-796 Bydgoszcz POLAND Proceedings of the 7th WSEAS International Conference on SIGNAL PROCESSING (SIP'08), Istanbul, Turkey, May 27-30, 2008.

[11] "A Descriptive Algorithm for Sobel Image Edge Detection", O. R. Vincent, Clausthal, O. Folorunso, University of Technology, Germany and University of Agriculture, Abeokuta, Nigeria, proceedings of Informing Science & IT Education Conference (InSITE) 2009.

[12] SEGMENTATION AND OBJECT RECOGNITION USING EDGE DETECTION TECHNIQUES, Y.Ramadevi, T.Sridevi, B.Poornima, B.Kalyani Department of CSE , Chaitanya Bharathi Institute of Technology Gandipet, Hyderabad, International Journal of Computer Science & Information Technology (IJCSIT), Vol 2, No 6, December 2010.

[13] "Morphological Background Detection and Enhancement of Images With Poor Lighting", Angélica R. Jiménez-Sánchez, Jorge D. Mendiola-Santibañez, Iván R. Terol-Villalobos, Gilberto Herrera-Ruíz, Damián Vargas-Vázquez, Juan J. García-Escalante, and Alberto Lara-Guevara IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 18, NO. 3, MARCH 2009 .

[14] "Application for Morphological Image Processing", Author: Dan Campbell UW-Madison Computer Engineering/Computer Science,the university of Wisconsin Madison, Course: ECE533,2006.