An Improved Distance Metric for Image Retrieval using Texton Features

Dr. Renuka Devi S M
Department of Electronics & Communication Engineering
G.Narayanamma Institute of Technology & Science Hyderabad, India

***Abstract*—** *Texture is an important characteristics in expressing the image details during content based image retrieval. This paper compares the retrieval performance of two texture features i.e., Texton co-occurance matrix (TCM) and Multi-texton histogram (MTH) using various distance measures. In this paper have proposed a modified distance metric and have shown the performance of this metric to be better over the recent proposed distance metric. Our proposed distance is similar to l_1 norm and hence has advantage of reduced computation cost. We also show that our proposed distance measure follows metric properties. Experiments on 1000 images Wang database have shown the effectiveness of our proposed metric and is found to result in 4-5% improvement in precision.*

Keywords— Texton, GLCM, Texton Co-Occurrence Matric, MTH

I. INTRODUCTION

With the advent of high-speed general-purpose digital computers it is becoming possible to perform mathematical or algorithmic processes pictorial data from images of photographic quality. In small collection of images simple browsing can identify an image. This is not the case for large and varied collection of images, where the user en-counters image retrieval problem. Image retrieval problem is the problem of searching and retrieving images that are relevant to a user's request from a database. To facilitate the retrieval of image data, many image retrieval techniques have been developed. Enormous research work exists in proposing various descriptors based on color, shape and texture features for image [11], [12]. The features used in content based image retrieval mainly depend on the type of the images. For example texture images are better described by texture features, images find better description by color, shape, and texture features.

In this paper, we are concerned with the task of developing a new distance metric that yields a better image retrieval based on texture and edge information of an image. In a Search for meaningful features for describing pictorial information, .it is only natural to look toward the types of features which human beings use in interpreting pictorial in-formation. Texture information and edge information are the fundamental elements used in human interpretation. Image Texture gives us information about the spatial arrangement of color or intensities in an image or selected region of an image. Edge information gives the outlines of an object and boundaries between objects and the background in the image. With the texture and edge information, we tried to explore for a better distance metric which a modified distance metrics is proposed by G. H. Liu et al [2]. In our experiment, we found that city block distance gives better results over Euclidean distance. This motivated our proposed distance metrics to be variation of l1 normal. Our results indicate that an improvement in 3% and 5% in precison and recall is obtained over the distance proposed by G. H. Liu et al.

II. RELATED WORK

In the usual texture analysis approach, GLCM is obtained from the original gray image itself. The features extracted from this matrix like contrast, entropy describe the texture content of an image. As an improvement to this, TCM was proposed wherein GLCM is obtained by texton extracted image. Later MTH [6] was proposed wherein both the texture information and edge information is embedded in feature generation of histogram.

A. Gray Level Co-occurance Matrix(GLCM)

The gray co-occurrence matrix is a traditional statistical method for texture analysis. The co-occurrence matrices characterize the relationship between the values of neigh-boring pixels [1]. For a coarse texture these matrices tend to have high values near the main diagonal, whereas for a fine texture the values are scattered. We denote the values of a gray image f as w_N. N={0, 1,.., 255} with f(P)= w. The pixel position is P = (x, y), let $P_1 = (x_1, y_1)$, P2 = (x_2, y_2), f $(P_1) = w_0$, $f(P_2) = w_2$. The probability Pr of two values (w_0 w_1) with pixel positions related by d, defines the cell entry (w_0 w_1) of the GLCM, usually the range of values: 0, 45, 90 and 135 degrees are considered. Haralick et. al [1] extracted a set of 13 features based on the co-occurrence matrix, such as energy, homogeneity, contrast, entropy etc.

1) Texton Co-occurance Matrix(TCM):

The concept of texton is a very useful tool in texture analysis. In general, textons are defined as a set of blobs or emergent patterns sharing a common property all over the image. Example tex- ton used to locate the textons in the given image is shown in figure 1. TCM features of an image [4] represent the spatial correlation of textons which discriminate color, texture and shape features simultaneously. The block diagram to obtain TCM features for an image is shown in figure 3.

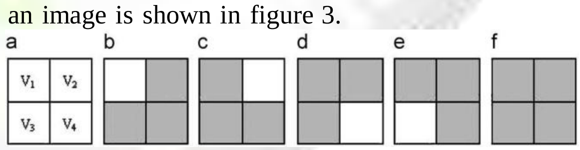

Fig. 1: Five types of textons used in TCM: (a) 2x2 grid (b)T1 (c)T2 (d)T3 (e)T4 and (f)T5

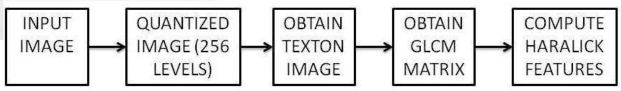

Fig. 2: The block diagram to obtain TCM features

The original color image is quantized into 256 colors in RGB color space, denoted by C(x,y) [2]. Five special types of texton templates are used to detect the textons, shown in Figure 1. The flow chart of texton detection is illustrated in Fig. 2. In an image, we move the 2x2 grid from left-to-right and top-to-bottom throughout the image to detect textons with one pixel as the step-length. If the pixel values that fall in the texton template are the same, those pixels will form a texton, and their values are kept as the original values. Otherwise they will be set to zero. Each texton template will lead to a texton image, an example of texton detection result is shown. The five texton templates will lead to five texton images. These are combined to obtain a final texton image (as shown in Fig. 2(b)). The GLCM is obtained from texton image. Finally thirteen haralick features [1] are obtaining GLCM.

III. MULTI –TEXTON HISTOGRAM(MTH)

The block diagram to obtain MTH feature of an image is shown in figure 4. Based on the texton theory, texture can be decomposed into elementary units, the texton classes of colors, elongated blobs of specific widths, orientation and aspect ratios, and the terminators of these elongated blobs. In order to extract color information and simplify the manipulation, in this work the RGB color space. Each color space is quantized into four levels and finally a 64-level grey image is obtained. Denote by C (x, y) the quantized image, where x, y = {0, 1, ..., N − 1}. Then each value of C (x, y) is a 6-bits binary code, ranging from 0 to 63.

The texton templates defined in MTH are different from those in TCM. Four special texton types are defined on a 22 grid, as shown in Fig.5. Denote the four pixels as V_1, V_2, V_3 and V_4. If the two pixels highlighted in gray color have the same value, the grid will form a texton. Those 4 texton types are denoted as T_1, T_2, T_3 and T_4, respectively. The working mechanism of texton detection is illustrated in Fig. 6. Given an color image C (x, y), we move the 22 block from left-to-right and top-to-bottom throughout the image to detect textons with 2 pixels as the step-length. If a texton is detected, the original pixel values in the 22 grids are kept unchanged. Otherwise it will have zero value. Finally, the obtained a texton image is denoted by T (x, y). The four texton types used in MTH contain richer information than those in TCM because the co-occurring probability of two same valued pixels is bigger than that of three or four same-valued pixels in a 22 grid. As for the texton detection procedure, MTH is also faster than TCM. In the texton detection of TCM, the 22 grid moves throughout the image with one pixel as the step length, and the detected textons in a neighbourhood overlap. The final texton image needs to be fused by the overlapped components of textons, and this will increase the computational complexity. While the step-length is set to two pixels in MTH reduces the computational cost.

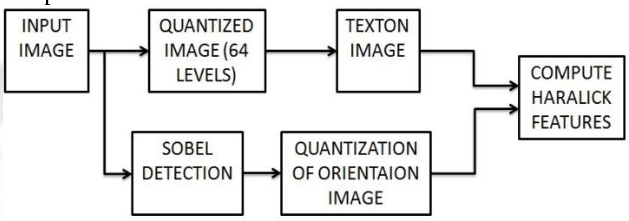

Fig. 4: The block diagram to obtain MTH features

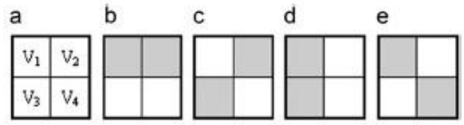

Fig. 5: Five types of textons used in MTH: (a) 2 2 grid (b) T_1 (c) T_2 (d) T_3; (e) T_4 and (f) T_5

Texture orientation analysis plays an important role in computer vision and pattern recognition. For instance, orien- tation is used in pre-attentive vision to characterize textons [3-5]. Orientation of texture images has a strong influence on humans perception of a texture image. Texture orientation can also be used to estimate the shape of textured images. The orientation map in an image represents the object boundaries and texture structures, and it provides most of the semantic information in the image. By applying some gradient operator, such as the Sobel operator, to a gray level image along horizontal and vertical directions, we can have two gradient images, denoted by gx and gy. The reason that we use the Sobel operator is that it is less sensitive to noise than other gradient operators or edge detectors while being very efficient. A gradient map g(x,y) can be obtained, with the gradient magnitude and orientation defined as $g(x,y)=gx^2+gy^2$ and angle(x,y)=arctan(gy/gx).

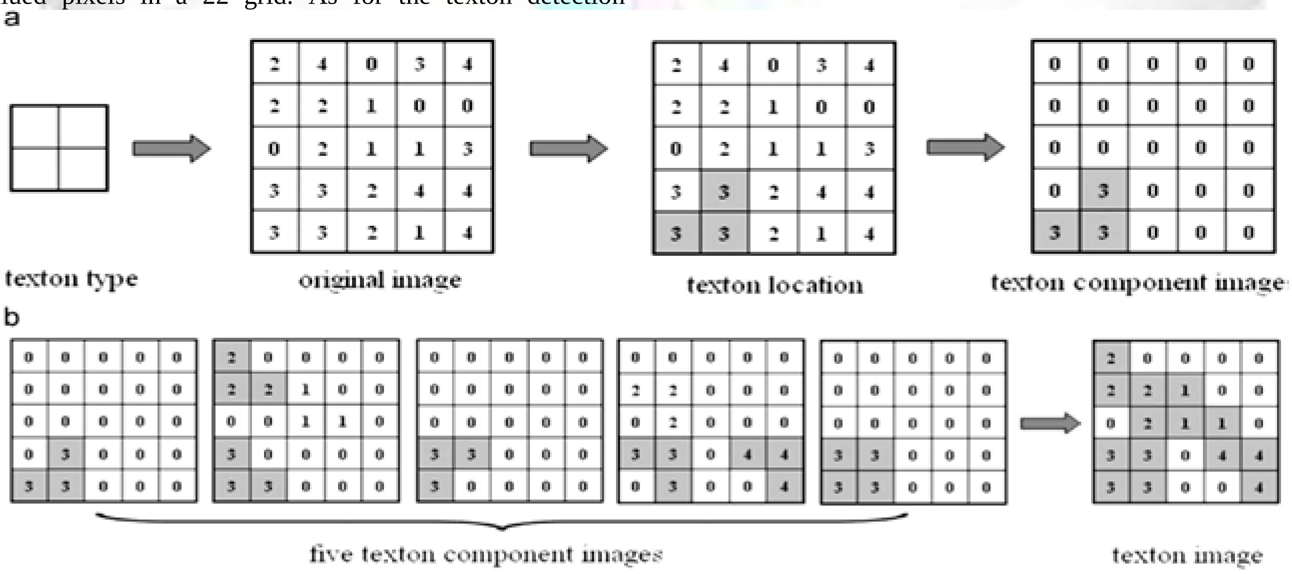

Fig. 3: The flow chart of texton detection in TCM (a)An example of texton detection (b)the five detected texton images and the final texton image.

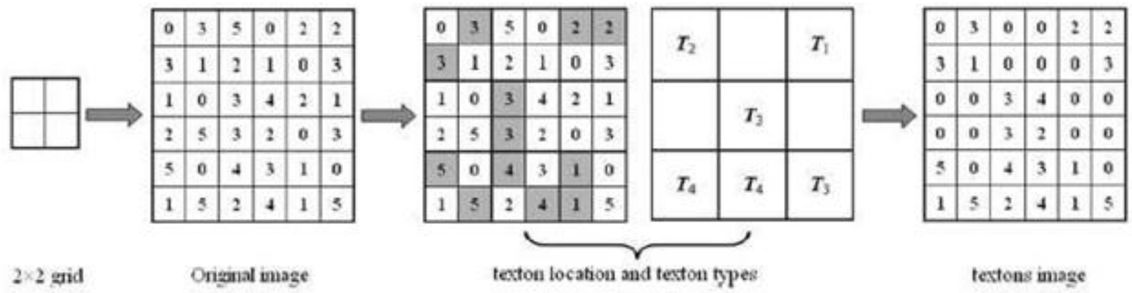

Fig. 6: Illustration of the texton detection process

As for full color images, there are red, green and blue channels. If we convert the full color image into a gray image, and then detect the gradient magnitude and orientation from the gray image, much chromatic information will be lost. In order to detect the edges caused by chromatic changes, the following method is followed. In the Cartesian space, let a = (x_1, y_1, z_1) and b = (x_2, y_2, z_2). Then dot product is defined as a.b = $x_1 x_2 + y_1 y_2 + z_1 z_2$

$$\cos(\widehat{a,b}) = \frac{a.b}{|a||b|} = \frac{x_1x_2 + y_1y_2 + z_1z_2}{\sqrt{x_1{}^2 + y_1{}^2 + z_1{}^2}.\sqrt{x_2{}^2 + y_2{}^2 + z_2{}^2}}$$

In color images consider the two vectors to be $a(R_X, G_X, B_X)$ and $b(R_Y, G_Y, B_Y)$, then

The angle between a and b is then

$$\cos(\widehat{a,b}) = \frac{a.b}{|a|.|b|}$$

Therefore texture orientation is given by,

$$\theta = \arccos[\cos(\widehat{a,b})] = \mathrm{arc}\cos[\frac{a.b}{|a|.|b|}]$$

After the texture orientation θ of each pixel is computed, quantize it uniformly into 18 orientations with 10 degrees as the step-length. The co-occurrence matrix characterizes the relationship between the values of neighboring pixels, while the histogram- based techniques have high indexing performance and are simple to compute. If we use the co-occurrence matrix to represent image features directly, the dimension will be high and the performance can be decreased. If we use histogram only to represent image features, the spatial information will be lost. In order to combine the advantages of co-occurrence matrix and his- togram, the MTH descriptor is proposed. The values of a texton image T are denoted as wn, n={0, 1, ..., W − 1}. Denote by P_1= (x_1, y_1) and P_2 = (x_2, y_2) two neighboring pixels, and their values are T (P_1) = w1 and T (P2) = w2. In the texture orientation image θ(x, y), the angles at P_1 and P_2 are denoted by by θ(P_1) =v_1 and θ(P_2)=v2. In texton image T, two different texture orientations may have the same color, while in texture orientation image θ(x, y) two different colors may have the same texture orientation. Denote by N the co-occurring number of two values v_1 and v_2, and by N the co-occurring number of two values w_1 and w2. With two neighboring pixels whose distance is D, MTH is defined as follows:

$$H(T(P_1)) = \begin{cases} N\{\theta(P_1) = v_1 \cap \theta(P_2) = v_2, |P_1 - P_2| = D\} \\ \quad \text{where } \theta(P_1) = \theta(P_2) = v_1 = v_2 \end{cases}$$

$$H(\theta(P_1)) = \begin{cases} \bar{N}\{T(P_1) = w_1 \cap T(P_2) = w_2, |P_1 - P_2| = D\} \\ \quad where\ T(P_1) = T(P_2) = w_1 = w_2 \end{cases}$$

$$\text{Where } T(P_1) = T(P_2) = w_1 - w_2$$

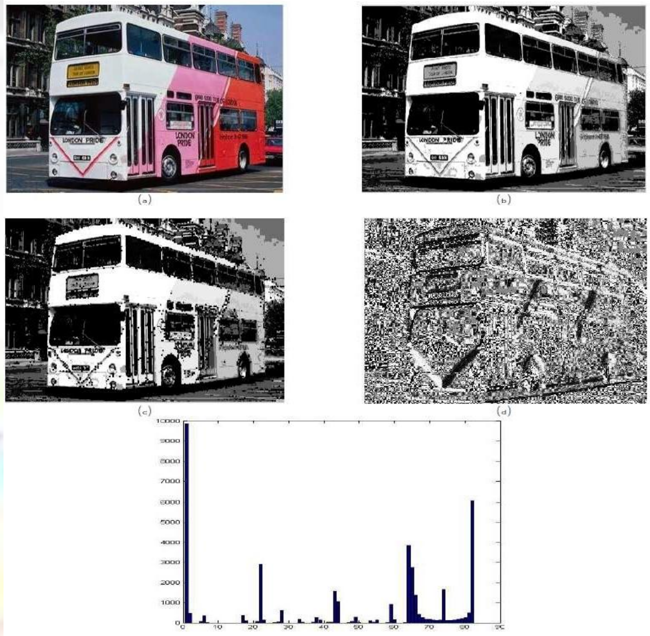

Fig. 7: MTH detection process (a) Query image (b) 64 level quantized image (c) texton image (d) Texture oriented image (e) Multi-texton Histogram

The proposed algorithm analyzes the spatial correlation between neighboring color and edge orientation based on four special texton types, and then forms the textons co-occurrence matrix and describes the attribute of texton co occurrence matrix using histogram.

This is why we call it multi-texton histogram (MTH). HTP1 can represent the spatial correlation between neighboring texture orientation by using color information, leading to a 64 dimensional vector. HP1 can represent the spatial correlation between neighboring colors by using the texture orientation information, leading to a 18 dimensional vector. Thus in total MTH uses a 64+18=82 dimensional vector as the final image features in image retrieval.

Where k is any positive integer constant.

IV. DISTANCE METIC

A. *Distance Metric Proposed G. H. Liu Et. Al.*

For each template image in the dataset, an M-dimensional feature vector T=[T_1, T_2….T_M] will be extracted and stored in the database. Let Q = [Q_1, Q_2…..Q_M] be the feature vector of a query image, the distance metric between them is simply calculated as

$$D = \sum_{i=1}^{M} \frac{|T_i - Q_i|}{1 + T_i + Q_i}$$

The above formula is as simple as the L_1 distance, which has no square or square root operations. Hence the computational cost and time is very less and is more suitable for large scale image datasets. Actually, it can be considered as a weighted L_1 distance with the factor $\frac{1}{1+T_i+Q_i}$ being the weight. For the proposed MTH, average distance over 100 query images is considered.

B. Our Proposed Distance Metric

Our proposed distance measure is a modification over the above distance metric. The reason for the proposed distance metric to be similar to L_1 norm is two-folds. One, our experiments using Euclidean distance and L_1 norm have shown that performance is better over the later than former.

Two, the computational cost of L_1 norm is much less than Euclidean. Our proposed distance metric is

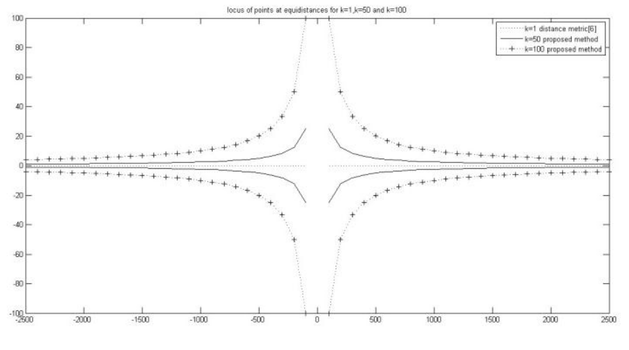

Fig. 8: Locus of distance metric for k=1, 50, 100

$$D = \sum_{i=1}^{M} \frac{|Ti - Qi|}{k + |Ti| + |Qi|}$$

The plot of the locus of equidistance points is shown in figure 8, which is done as follows. Let us consider T, Q be two dimensional vectors. To plot the locus of points at equidistances, consider Q to be origin (0,0) and T be (x, y) and distance D to be 1. By replacing T = (x, y), Q = (0, 0) and D = 1; we can rewrite above equation as

$$1 = \frac{|x|}{k + |x|} + \frac{|y|}{k + |y|}$$

$$|y| = \frac{k^2}{|x|}$$

The above is a hyperbolic equation. The plot of curves for k = 50, 100 is shown in figure 8. As k value increases, the curves moves away from origin. Our proposed distance metric follows the distance metric properties [5]. A metric on a set X is a function d. for all x,y,z in X, this function is required to satisfy the following conditions:

1) d(x, y) >= 0 (non-negativity, or separation axiom) this means that the distance should be always greater than or equal to zero.
2) d(x, y) = 0 if and only if x=y (identity of indis- cernibles) i.e., entities x and y are identical if every predicate possessed by x is also possessed by y and vice versa.
3) d(x, y) = d(y, x) (symmetry) i.e., the distance from points x to y is same as distance from points y to x.
4) d(x, z)d(x, y) + d(y, z) (subadditivity) i.e., distance between two points x and z is less than or equal to the sum of distances of each point with some other point say y.

V. Experimental Results

In this section, we demonstrate the performance of our proposed distance metric using Corel dataset. We show that our proposed distance metric performance better over the l1-norm, Euclidean and the recent distance metric proposed by G.H.Liu. We experimented using two texton features, TCM and MTH. In both the cases, proposed distance metric is found to give an improved performance of 4-5% for various scopes. In the experiments, we selected randomly 20 images from every category as query image, and results shown in table aare the average precison of 200 images of queries.

A. Datasets

The Corel image dataset is the most commonly used dataset to test image retrieval performance is widely used. The dataset is Corel 1000 dataset. It contains 10 categories. There are 100 images from diverse contents such as sunset, beach, flower, horses, buses, food, etc [9,10].

B. Performance Measure

Precision and recall are used to evaluate the performance of the proposed approach. Precision is the number of the retrieved relevant images over the total number of retrieved images, and recall is the number of the retrieved relevant images over the total number of relevant images in the database. To calculate precision and recall, only those re-trieved images from the same semantic category as the query are counted as relevant. The number of images returned to the user is called scope.

Distance metric	Precison for various scopes				
	s=20	S=40	S=60	S=80	S=10
Our Proposed Distanc	73.05	64.85	59.12	54.29	49.5
Distance measure by	69.1	62.05	56.8	46.66	46.66
City Block	61.4	54.53	49.72	46.18	42.75
Euclidea	55.5	47.8	43.1	39.5	36.94
Distance metric	Precison for various scopes				
	s=20	S=40	S=60	S=80	S=10
Our Proposed Distanc	14.7	25.94	35.47	43.43	49.5
Distance measure by	13.82	24.82	34.08	41.06	46.66
City Block	12.28	21.81	29.83	36.94	42.75
Euclidea	11.1	19.1	25.9	31.6	36.94

Table 1: Performance Measure

VI. Conclusion

In this paper we have proposed a new distance metric over the distance proposed in [2] by G.H.Lui et. al for image retrieval based on Texton. We also showed that this measure satisfies the properties required for a metric. In our experiments we used the benchmark Wang database, to compare the performance of proposed distance measure other metrics. An improvement in 4-5% improvement is observed using both the texton features i.e., TCM and MTH over the recent proposed distance.

REFERENCE

[1] R.M. Haralick, K. Shangmugam, I. Dinstein, "Textural feature for image classification", IEEE Trans. Syst. Man Cybern SMC-3 (6) (1973) 610621.

[2] G.-H. Liu, J.-Y. Yang, "Image retrieval based on the texton co-occurrence matrix", Pattern Recognition 41 (12) (2008) 35213527.

[3] B. Julesz, "Textons, the elements of texture perception and their interactions", Nature 290 (5802) (1981) 9197.

[4] B.Julesz, "Texton gradients: the texton theory revisited", Biological Cybernetics, 54 (1986) 245251.

[5] R.C.Gonzalez,R.E. Woods, "Digital Image Processing", third ed, Prentice Hall, 2007.

[6] Guang-Hai Liu, Lei Zhang, Ying-Kun Hou, Zuo-Yong Li, Jing- Yu Yang, "Image retrieval based on multi-texton histogram", 2010.

[7] Y. Liu, D. Zhang, G. Lu, W.-Y. Ma, "A survey of content-based image retrieval with high-level semantics", Pattern Recognition, 40 (11) (2007) 262282.

[8] N. Alajlan, M.S. Kamel, G.H. Freeman, "Geometry-based image retrieval in binary image databases", IEEE Transactions on Pattern Analysis and Machine Intelligence 30 (6) (2008) 100311013.

[9] S. M. Renuka Devi and Chakravarthy Bhagvati, "Connected component method to find components of gmm in image retrieval", International Conference on Computational Intelligence and Communication Networks (CICN 10), vol. 0, pp.5054, 2010.

[10] S. M. Renuka Devi and Chakravarthy Bhagvati, "A Novel Method for Long Term Learning by Cluster Merging, 5th In- dian International Conference on Artificial Intelligence (IICAI11), 427–440.

[11] Niblack, Carlton W., et al. "QBIC project: querying images by content, using color, texture, and shape." IS&T/SPIE's Symposium on Electronic Imaging: Science and Technology. International Society for Optics and Photonics, 1993.

[12] Hiremath, P. S., and Jagadeesh Pujari. "Content based image retrieval using color, texture and shape features." Advanced Computing and Communications, 2007. ADCOM 2007. International Conference on. IEEE, 2007.