Analysis of Distance Measures in Content Based Image Retrieval

Anjali Batra[1] and Dr. Meenakshi Sharma[2]

[1] HCTM TECHNICAL CAMPUS, KAITHAL, HARYANA, INDIA (136027)

Abstract

Content predicated image retrieval (CBIR) provides an efficacious way to probe the images from the databases. The feature extraction and homogeneous attribute measures are the two key parameters for retrieval performance. A homogeneous attribute measure plays a paramount role in image retrieval. This paper compares six different distance metrics such as Euclidean, Manhattan, Canberra, Bray-Curtis, Square chord, Square chi-squared distances to find the best kindred attribute measure for image retrieval. Utilizing pyramid structured wavelet decomposition, energy levels are calculated. These energy levels are compared by calculating distance between query image and database images utilizing above mentioned seven different kindred attribute metrics. A sizably voluminous image database from Brodatz album is utilized for retrieval purport. Experimental results shows the preponderating of Canberra, Bray-Curtis, Square chord, and Square Chi-squared distances over the conventional Euclidean and Manhattan distances.

1 Introduction

ontent-based image retrieval (CBIR), additionally kenned as query by image content (QBIC) and content-based visual information retrieval (CBVIR) is the application of computer vision techniques to the image retrieval quandary, that is, the quandary of probing for digital images in immensely colossal databases (optically discern this survey for a recent scientific overview of the CBIR field). Content-predicated image retrieval is opposed to traditional conceptpredicated approaches (optically discern Concept predicated image indexing).

"Content-based" designates that the search analyzes the contents of the image rather than the metadata such as keywords, tags, or descriptions associated with the image. The term "content" in this context might refer to colors, shapes, textures, or any other information that can be derived from the image itself. CBIR is desirable because searches that rely pristinely on metadata are dependent on annotation quality and broadness. Having humans manually annotate images by entering keywords or metadata in an astronomically immense database can be time consuming and may not capture the keywords desired to describe the image. The evaluation of the efficacy of Author ? ?: H.C.T.M., Technical Campus, Kaithal, Haryana, -mails: minnyk@gmail.com, anjalibatra18@gmail.com keyword image search is subjective and has not been well-defined. In the same regard, CBIR systems have homogeneous challenges in defining success.

II.

2 Related Literature

Due to exponential increase of size of soi-disant multimedia files in recent years because of the substantial increase of affordable recollection storage on one hand and the wide spread of World Wide Web (www) on the other hand, the desideratum for the efficient implement to retrieve the images from the immensely colossal data base becomes crucial. This motivates the extensive research into image retrieval systems. From the historical perspective, the

earlier image retrieval systems are rather text-predicated with the thrust from database management community since the images are required to be annotated and indexed accordingly. However with the substantial increase of the size of images as well as size of image database, the task of utilizer-predicated annotation becomes very cumbersome and at some extent subjective and thereby, incomplete as the text often fails to convey the affluent structure of images. In the early 1990s, to surmount these difficulties this motivates the research into what is referred as content based image retrieval (CBIR) where retrieval is predicated on the automating matching of feature of query image with that of image database through some image-image kindred attribute evaluation. Therefore images will be indexed according to their own visual content such as color, texture, shape or any other feature or a coalescence of set of visual features. The advances in this research direction are mainly contributed by the computer vision community.

3   III.

4   Proposed Work

We apply different distance metrics and input a query image based on similarity features of which we can retrieve the output images. These distance measures or metrics have been illustrated as follows: The distance between two points in a grid based on a strictly horizontal and/or vertical path (that is, along the grid lines), as opposed to the diagonal or "as the crow flies" distance. The Manhattan distance is the simple sum of the horizontal and vertical components, whereas the diagonal distance might be computed by applying the Pythagorean Theorem. Distance measures such as the Euclidean, Manhattan and Standard Euclidean distance have been used to determine the similarity of feature vectors. In this CBIR system Euclidean distance, Standard Euclidean distance and also Manhattan distance is used to commonly to compare the similarity between the images. Distance between two images is used to find the similarities between query image and the images in the database.a) Euclidean distance
EU (u, v) = ?$(x1-x2)$ 2 + $(y1-y2)$ 2**(1)**

5   d) Mahalanobis distance

The Mahalanobis distance is a measure of the distance between a point P and a distribution D, introduced by P. Mahalanobis in 1936. [1] It is a multidimensional generalization of the idea of measuring how many standard deviations away P is from the mean of D. This distance is zero if P is at the mean of D, and grows as P moves away from the mean: Along each principal component axis, it measures the number of standard deviations from P to the mean of D. If each of these axes is rescaled to have unit variance, then Mahalanobis distance corresponds to standard Euclidean distance in the transformed space. Mahalanobis distance is thus unit less and scale-invariant, and takes into account the correlations of the data set. The Mahalanobis distance of an observation from a group of observations with mean and covariance matrix Sis defined as: [2] Mahalanobis distance (or "generalized squared inter point distance" for its squared value [3]) can also be defined as a dissimilarity measure between two random vectors and of the same distribution with the covariance matrix S:

If the covariance matrix is the identity matrix, the Mahalanobis distance reduces to the Euclidean distance. If the covariance matrix is diagonal, then the resulting distance measure is called a normalized Euclidean distance: where s i is the standard deviation of the x_i and y_i over the sample set.

Mahalanobis distance is preserved under fullrank linear transformations of the space spanned by the data. This means that if the data has a nontrivial null space, Mahalanobis distance can be computed after projecting the data (non-degenerately) down onto any space of the appropriate dimension for the data.

6   e) Chebyshev distance

The Chebyshev distance between two vectors or points p and q, with standard coordinates and , respectively, is This equals the limit of the L p metrics: hence it is also known as the L ? metric. Mathematically, the Chebyshev distance is a metric induced by the supremum norm or uniform norm. It is an example of an injective metric. In two dimensions, i.e. plane geometry, if the points p and q have Cartesian Under this metric, a circle of radius r, which is the set of points with Chebyshev distance r from a center point, is a square whose sides have the length 2r and are parallel to the coordinate axes. On a chess board, where one is using a discrete Chebyshev distance, rather than a continuous one, the circle of radius r is a square of side lengths 2r, measuring from the centers of squares, and thus each side contains 2r+1 squares; for example, the circle of radius 1 on a chess board is a 3×3 square.

In one dimension, all Lp metrics are equalthey are just the absolute value of the difference. The two dimensional Manhattan distance also has circles in the form of squares, with sides of length ?2r, oriented at an angle of ?/4 (45°) to the coordinate axes, so the planar Chebyshev distance can be viewed as equivalent by rotation and scaling to the planar Manhattan distance. However, this equivalence between L 1 and L ? metrics does not generalize to higher dimensions. A sphere formed using the Chebyshev distance as a metric is a cube with each face perpendicular to one of the coordinate axes, but a sphere formed using Manhattan distance is an octahedron: these are dual polyhedra, but among cubes, only the square (and 1dimensional line segment) are self-dual polytopes. The Chebyshev distance is sometimes used in warehouse logistics, [4] as it effectively measures

the time an overhead crane takes to move an object (as the crane can move on the x and y axes at the same time).

On a grid (such as a chessboard), the points at a Chebyshev distance of 1 of a point are the Moore neighborhood of that point.

7 IV.

8 Experiments on Matlab

L1 (numOfReturnedImages, queryImageFeatureVector, dataset) function L1(numOfReturnedImages, queryImageFeatureVector, dataset) % input: % numOfReturnedImages: num of images returned by query % queryImageFeatureVector: query image in the form of a feature vector % dataset: the whole dataset of images transformed in a matrix of % features % % output: % plot: plot images returned by query % extract image fname from queryImage and dataset query_image_name = queryImageFeatureVector (:, end); dataset_image_names = dataset (:, end); queryImageFeatureVector (:, end) = [];

9 Confusion Matrix

Confusion matrix is used to compare the performance of the CBIR system using different distance metrics. To evaluate the overall performance of the CBIR system and compare the different distance metrics for retrieval accuracy, confusion matrix is calculated. A confusion matrix represents the actual classifications compared with the number of correct and incorrect prediction. The confusion matrix is n-by-n matrix, where n is the number of classes from the dataset. Each row represents the number of instances in actual class. Each column represents the number of instances in predicted class. Table **??** shows a confusion matrix for a 3 classes classification model. In this confusion matrix, of the actual 5 A instances, the system predicted that the 5 instances were A, and of the 5 B instances, it predicted that 1 was A, 3 were B and 1 was C. All correct predictions are located in the diagonal of the table, so the other positions except the diagonal are errors. Accuracy (AC) is the most intuitive assessment from the confusion matrix. It is the correct classifications divided by all classifications. In the confusion matrix, the overall accuracy is calculated as the sum of the diagonal numbers divided by the sum of all the numbers in the matrix. For example, the accuracy of the example in Table **??** is: $(5+3+1)/ (5+0+0+1+3+1+2+2+1) = 0.6$ Table 1 : Confusion Matrix Predicted Actual A B C A 5 0 0 B 1 3 1 C 2 2 1

10 Feature Extraction

When the input data to an algorithm is too sizably voluminous to be processed and it is suspected to be notoriously redundant (much data, but not much information) then the input data will be transformed into a reduced representation set of features Transferring the input data into the set of features is called feature extraction. The features provide the characteristics of the input type to the classifier by considering the description of the pertinent properties of the image into a feature space. If the extracted features are meticulously culled, it is expected that they will extract the pertinent information from the input data in order to perform the desired task utilizing this reduced representation in lieu of the full size input. Feature extraction is simplifying the amount of data required to describe an immensely colossal set of data accurately. When performing analysis of hard data one of the major quandaries stems from the number of data's involved. Analysis with an astronomically immense number of data's generally requires a substantial amount of recollection and computation power or a relegation algorithm which over fit's the training sample and generalizes poorly to incipient samples. Feature extraction can be utilized in the area of image processing which involves utilizing algorithms to detect and isolate sundry desired portions or shapes (features) of a digitized image or video stream. Another paramount feature processing stage is feature cull. However, when immensely colossal and G can be found that can accurately relegate the training data, but do not pertain to unseen test data. Feature cull is partially up to the designer to cull a felicitous feature set, but automatic methods can withal be utilized. In culling features, it is consequential to consider whether features will avail in discriminating unseen data, and how perplexed the interactions between the features are liable to be in order for them to be utilized in discrimination.

11 GLCM (Gray level Co-occurrence matrix)

A gray level co-occurrence matrix (GLCM) contains information about the positions of pixels having similar gray level values. A co-occurrence matrix is a two-dimensional array, P in which both the rows and the columns represent a set of possible image values. A GLCM Pd [i, j] is defined by first specifying a displacement vector d = (dx, dy) and counting all pairs of pixels separated by d having gray levels i and j. The GLCM is defined by Where nij is the number of occurrences of the pixel values (i, j) lying at distance d in the image. The co-occurrence matrix Pd has dimension n × n, where n is the number of gray levels in the image. For example, if d= (1, 1). The efficacy of the image retrieval is predicated on the performance of the feature extraction and kindred attribute quantification. In this section we describe the performance metrics which have been adopted not only to evaluate the efficacy of image retrieval but withal to ascertain of the stability of the results. In order to evaluate the retrieval performance of CBIR, three quantifications are utilized: precision, and F-Score.

12   Figure 3 : Confusion Matrix

The precision in image retrieval can be defined as: precision is the measurement of the retrieved relevant images to the query of the total retrieved images. The recall in image retrieval can be defined as: Recall is the measurement of the retrieved relevant images to the total database images.

V.

13   Conclusion

Query image is given as input and using different similarity metrics, we can retrieve the required number of output images. [1]

Figure 1:

$$x = (x_1, x_2, x_3, \ldots, x_N)^T$$

Figure 2:

$$_1\mu = (\mu_1, \mu_2, \mu_3, \ldots, \mu_N)^T$$

Figure 3: Figure 1 :

$$D_M(x) = \sqrt{(x - \mu)^T S^{-1} (x - \mu)}.$$

Figure 4:

$$2.\vec{x}$$

Figure 5: Figure 2 :

$$\vec{y}$$

Figure 6:

13 CONCLUSION

[164] [Liu et al. ()] 'A survey of Content-based image retrieval with high level semantics'. Y Liu , D Zhang , G Lu , W Ma . *Pattern Recognition* 2007. 40 (1) p. .

[166] [Selvarajah and Kodituwakku ()] 'Analysis and Comparison of Texture Features for Content Based Image Retrieval'. S Selvarajah , S R Kodituwakku . *International Journal of Latest Trends in Computing* 2045-5364. 2011. 2 (1) p. .

[169] [Bae and Jung ()] Hee-Jung Bae , Sung-Hwan Jung . *Image retrieval Using texture based on dct. International conference on Information, communications and signal processing ICICS 97*, (Singapore) 1997. p. .

[171] [Hiremath and Pujari ()] 'Based Image Retrieval Using Color'. P S Hiremath , J Pujari . *Texture and Shape Features, presented at the Proceedings of the 15th International Conference on Advanced Computing and*, (Content) 2007.

[174] [Carson et al. (1999)] 'Blobworld: A System for Region-Based Image Indexing and Retrieval'. C Carson , M Thomas , S Belongie , J M Hellerstein , J Malik . *Proc. Visual Information Systems*, (Visual Information Systems) June 1999. p. .

[177] [Bai ()] 'Color Textured Image Retrieval By Combining Texture and Color Features'. C Bai . *European Signal Processing Conference*, (Bucharest: Romania) 2012.

[179] [Smeulders et al. (2000)] 'Content-Based Image Retrieval at the End of the Early Years'. A W M Smeulders , M Worring , S Santini , A Gupta , R Jain . *IEEE Transactions on Pattern Analysis and Machine Intelligence* December 2000. 22 (12) p. .

[182] [Hafner et al. ()] 'Efficient Color Histogram Indexing for Quadratic Form Distance Functions'. J Hafner , H Sawhney , W Equitz , M Flickner , W Niblack . *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1995. 17 (7) p. .

[185] [Suresh et al. ()] 'Feature Extraction in Compressed Domain for Content Based Image Retrieval'. P Suresh , R M D Sundaram , A Arumugam . *IEEE International Conference on Advanced Computer Theory and Engineering (ICACTE)*, 2008. p. .

[188] [Fuh et al. (2000)] 'Hierarchical Color Image Region Segmentation for Content-Based Image Retrieval System'. C S Fuh , S W Cho , K Essig . *IEEE Transactions on Image Processing* Jan. 2000. 9 (1) p. .

[190] [Jain and Vailaya ()] 'Image Retrieval using Color and Shape'. A Jain , A Vailaya . *Pattern Recognition* 1996. 29 (8) p. .

[192] [Deshpande and Borse ()] 'Image Retrieval with the use of different color spaces and the texture feature'. Gauri Deshpande , Megha Borse . *International Conference on Software and Computer Applications*, 2011. 9 p. .

[194] [Majumdar and Sural ()] 'Performance Comparison of distance metrics in content-based Image Retrieval applications'. A Majumdar , A K Sural , S . *Proc. of Internat. Conf. on Information Technology*, (of Internat. Conf. on Information TechnologyBhubaneswar, India) 2003. p. .

[197] [Niblack et al. ()] 'The QBIC Project: Querying Images by Content using Color Texture and Shape'. W Niblack , R Barber , W Equitz , M Flickner , E Glasman , D Pektovic , P Yanker , C Faloutsos , G Taubin . *Proc. SPIE Int. Soc. Opt. Eng., in Storage and Retrieval for Image and Video Databases*, (SPIE Int. Soc. Opt. Eng., in Storage and Retrieval for Image and Video Databases) 1993. 1908 p. .

[201] [Smith and Chang (1996)] 'Visualseek: A Fully Automated Content-Based Image Query System'. J Smith , S Chang . *Proceedings of the 4 th ACM International conference on Multimedia table of Contents*, (the 4 th ACM International conference on Multimedia table of ContentsBoston, Massachusetts, United States) Nov. 1996. p. .

[205] [Flickner et al. ()] 'WALRUS: A Similarity Retrieval Algorithm for Image Databases'. H Flickner , W Sawhney , J Niblack , Q Ashley , B Huang , Dom . *IEEE Computer* 1995. 1999. 28 (9) p. . (SIGMOD Record)