iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 2 **Issue:** VI **Month of publication:** June 2014

DOI:

ISSN: 2321-9653

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

Automatic Image Annotation and Retrieval Using Group Sparsity

Miss. Jayachitra1, Mrs. M. Jagadheeswari2

1Research scholar, 2Lecturer in computer science,

Department of computer science,

Sri Krishna Arts & science,

Coimbatore-641008

Abstract: Automatically assigning relevant text keywords to images is an important problem. Many algorithms have been proposed in the past decade and achieved good performance. Efforts have focused upon model representations of keywords, whereas properties of features have not been well investigated. In most cases, a group of features is preselected, yet important feature properties are not well used to select features. In this paper, we introduce a regularization-based feature selection algorithm to leverage both the sparsity and clustering properties of features, and incorporate it into the image annotation task. Using this group-sparsity-based method, the whole group of features [e.g., red green blue (RGB) or hue, saturation, and value (HSV)] is either selected or removed. Thus, we do not need to extract this group of features when new data comes. A novel approach is also proposed to iteratively obtain similar and dissimilar pairs from both the keyword similarity and the relevance feedback. Thus, keyword similarity is modeled in the annotation framework. We also show that our framework can be employed in image retrieval tasks by selecting different image pairs. Extensive experiments are designed to compare the performance between features, feature combinations, and regularization-based feature selection methods applied on the image annotation task, which gives insight into the properties of features in the image annotation task. The experimental results demonstrate that the group-sparsity-based method is more accurate and stable than others.

LITERATURE SURVEY

2.1 AUTOMATIC THUMBNAIL CROPPING AND ITS EFFECTIVENESS

B.Suh, H.Ling and David W. Jacobs [2] proposed a new method that gives the effectiveness of saliency based cropping methods for preserving the recognisability of important objects in thumbnails. The method is a general cropping method based on the saliency map based on a model of

ISSN: 2321-9653

human visual attention. A saliency map of a given image describes the importance of each position in the image. In this method, used the saliency map directly as an indication of how much information each position in images contains. The merit of this method is that the saliency map is built up from low-level features only, so it can be applied to general images. Then select the portion of the image of maximal in formativeness.

Although this saliency based method is useful, it does not consider semantic information in images. The semantic information can be used to further improve thumbnail cropping, using automatic face detection. This domain is chosen because a great many pictures of interest show human faces, and also because face detection methods have begun to achieve high accuracy and efficiency.

2.2 TEXTURE BASED IMAGE INDEXING AND RETRIEVAL

N.Rao, Dr.Vijaya Kumar and V.Venkatesh Krishna proposed this system [3], as there is a prominent increment in computing power, rapidly reducing storage cost and worldwide access to the Internet, digital acquisition of information has become increasingly popular in recent years. Digital information is preferable to analog formats because of convenient sharing and distribution properties. This trend has motivated research in image databases, which were nearly ignored by traditional computer systems due to the enormous amount of data necessary to represent images and the difficulty of automatically analyzing images. Currently, storage is less of an issue since huge storage capacity is available at low cost. However, effective indexing and searching of large-scale image databases remains as a challenge for computer systems. The image retrieval is a system, which retrieves the images from an image collection where the retrieval is based on a query, which is specified by content and not by index or address. The query image is an image in which a user is interested and wants to find similar images from the image collection. The image retrieval system retrieves relevant images from an image collection based on automatic derived features. The derived features include primitive features like texture, color, and shape. The features may also be logical features like identity of objects shown, abstract features like significance of some scene-depicted etc. The present method implemented by three steps. First, for each image in the image collection, a feature vector of size ten, characterizing texture of the image is computed based on the Wavelet transformation method. The Wavelet transformations are used because they capture the local level texture features quite efficiently, where feature vectors are stored in a feature database.

Second, using clustering algorithm to construct indexed image database based on the texture feature vectors obtained from wavelet transformation, and finally, given a query image, its feature vector is computed and compared to the feature vectors in the feature database, and relevant images to the query image from the image database returned to the user. Every care has been taken to ensure that the features and the similarity measure used to compare two feature vectors are efficient enough to match similar images and to discriminate dissimilar ones. The main aim of this approach is that not even a single relevant image should be missed in the output as well as to minimize the number of irrelevant images.

The steps involved in the methodology are:

Vol. 2 Issue VI, June 2014

ISSN: 2321-9653

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

1. Haar Wavelet transformation is used for feature extr action.
2. Precomputing the texture feature vectors for all the images in the database using Haar wavelet Transformation.
3. Clustering the images based on feature vectors using modified ROCK clustering algorithm.
4. Computing the feature vector of the query image as and when presented.
5. Comparing query image with indexed database, identifying the closest cluster for the query image and retrieves those images.
6. Presenting the result as the thumbnail set of images.

The Extraction of feature vector is the most crucial step in the whole image retrieval system. This is because these feature vectors are used in all the subsequent modules of the system. It is to be realized that the image itself plays no part in the following modules. It is the feature vectors that are dealt with. The quality of the output drastically improves as the feature vectors that are used are made more effective in representing the image. Haar Wavelets are useful for hierarchically decomposing functions in ways that are both efficient and theoretically sound. Broadly speaking, a wavelet representation of a function consists of a coarse overall approximation together with detail coefficients that impudence the function at various scaled.

The wavelet transform has excellent energy compaction and de-correlation properties, which can be used to effectively generate compact representations that exploit the structure of data.

By using wavelet sub band decomposition, and storing only the most important sub bands (that is, the top coefficients), the fixed-size low-dimensional feature vectors independent of resolution, image size and dithering effects are computed. In addition, wavelets are robust with respect to color intensity shifts, and can capture both texture and shape information efficiently. Furthermore, wavelet transforms can be computed in linear time, thus allowing for very fast algorithms.

Another important issue in content-based image retrieval is effective indexing and fast searching of images based on visual features. Because the feature vectors of images tend to have high dimensionality and therefore are not well suited to traditional indexing structures, dimension reduction is usually used before setting up an efficient indexing scheme. The basis of the clustering method in indexed image database is that, the images belonging to the same cluster are similar or relevant to each other when compared to images belonging to different clusters. The images are clustered using modified ROCK. The modified ROCK allows us to minimize the undesirable results of the ROCK algorithm. The feature vector of each image is a vector of size 10.The Euclidean distance measure is used to measure the similarity between feature vectors of query image and indexed database image. In the present method, calculated representative Feature vector of Cluster (FC) as the minimum Euclidean distance, this resulted in good cluster-matching results.

The representative feature vector of cluster (FC) is computed from the following equation.

ISSN: 2321-9653

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

$$F_{ci} = min|F_i - F_j|$$

(2.1)

Where, j =1,2,…..n and i=1,2,..n and $j \neq i$. F_{ci} denotes representative features of cluster i and F_i, F_j represents feature vector of the given cluster. Query by example allows the user to formulate a query by providing an example image. The system converts the example image into an internal representation of features. Images stored in the database with similar features are then searched. It can be further classified into query by external image example, if the query image is not in the database, and query by internal image example, if otherwise. For query by internal image, all relationships between images can be pre-computed. The main advantage of query by example is that the user is not required to provide an explicit description of the target, which is instead computed by the system. It is suitable for applications where the target is an image of the same object or set of objects under different viewing conditions. Most of the current systems provide this form of querying.

2.3 CONTENT-BASED IMAGE INDEXING AND SEARCHING

J.Z.Wang, G.Wiederhold, proposed a method for content based image indexing [4]. Every day, large numbers of people are using the Internet for searching and browsing through different multimedia databases. To make such searching practical, effective image coding and searching based on image semantics is becoming increasingly important. In current real-world image databases, the prevalent retrieval techniques involve human-supplied text annotations to describe image semantics. These text annotations are then used as the basis for searching, using mature text search algorithms that are available as free-ware. However, there are many problems in using this approach.

For example, different people may supply different textual annotations for the same image. This makes it extremely difficult to reliably answer user queries. Furthermore, entering textual annotations manually is excessively expensive for large-scale image databases. Image feature vector indexing has been developed and implemented in several multimedia database systems such as the IBM QBIC System developed at the IBMAlmaden Research Center, the Virage System developed by Virage, Inc., and the Photobook System developed by the MIT Media Lab.

For each image inserted into the database, a feature vector on the order of 500 elements is generated to accurately represent the content of the image. This vector is much smaller in size than the original image. The difficult part of the problem is to construct a vector that both pre-serves the image content and yet is efficient for searching. Once the feature vectors are generated, they are then stored in permanent storage.

To answer a query, the image search engine scans through the previously computed vector indexes to select those with shortest distances to the image query vector. The distance is computed by a measure such as the vector distance in Euclidean space. For partial sketch queries, usually a mask is computed and applied to the feature vector. They developed a new algorithm to make semantically-meaningful comparisons of images efficient and accurate. To accurately encode semantic

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

features of images, the wavelets based on continuous functions was employed, as described by Daubechies wavelets.

Many color image formats are currently in use, e.g., GIF, JPEG, PPM and TIFF are the most widely used formats. Because images in an image database can have different formats and different sizes, the data must be normalized first.

For our test database of relatively small images, a rescaled thumbnail consisting of 128 X 128 pixels in Red-Green-Blue (RGB) color space is adequate for the purpose of computing the feature vectors. Bilinear interpolation is used for the rescaling process. This method resample's the input image by overlaying the input image with a grid with 128 X 128 points.

This gives one grid point for each pixel in the output image. The input image is then sampled at each grid point to deter-mine the pixel colors of the output image. When grid points lie between input pixel centers, the color values of the grid point are determined by linearly interpolating between adjacent pixel colors (both vertically and horizontally).This rescaling process is more effective than a Haar-like rescaling, i.e., averaging several pixels to obtain a single pixel to decrease image size, and replicating pixels to increase image size, especially when the image to be rescaled has frequent sharp changes such as local texture. It is necessary to point out, however, that the rescaling process is in general not important for the indexing phase when the size of the images in the database is close to the size to be rescaled. The sole purpose for the rescaling is to make it possible to use the wavelet transforms and to normalize the feature vectors.

Here, the assumption the images in the database to have sizes close to 128 X 128. In fact, images may be rescaled to any other size as long as each side length is a power of two. Therefore, to obtain a better performance for a database of mostly very large images, a bilinear interpolation is suggested to rescale to a large common size, with side lengths being powers of two, and then apply more levels of Daubechies' wavelets in the indexing phase. Since color distances in RGB color space do not reflect the actual human perceptual color distance, the images are converted and stored in a component color space with intensity and perceived contrasts.

2.4 COLOR IMAGE RETRIEVAL TECHNIQUE BASED ON COLOR FEATURES AND IMAGE BITMAP

T.C.Lu and C.C.Chang, [5] proposed image retrieval technique based on color features and image bitmap. The field of color image retrieval has been an important research area for several decades. For the purpose of effectively retrieving more similar images from the digital image databases, this paper uses the color distributions, the mean value and the standard deviation, to represent the global characteristics of the image. Moreover, the image bitmap is used to represent the local characteristics of the image for increasing the accuracy of the retrieval system. As the experimental results indicated, the proposed technique indeed outperforms other schemes in terms of retrieval accuracy and category retrieval ability. Furthermore, the total memory space for saving the image features of the proposed method is less than Chan and Liu's method. Image retrieval has been a very active research topic since the 1970s. Most convenient image retrieval schemes are annotated-based that annotate each image in an image database by using keywords for similar image retrieval. However, the annotated-based retrieval method has two major problems that make it impractical. The first problem is that the size of the image database has become larger and larger such that to create keywords for each image is time-consuming. The second problem is that different people may give the same image different keywords. Afterwards, the content-based image retrieval (CBIR) techniques were proposed to solve the problems of the annotated-based image retrieval methods. In a CBIR system, images are automatically indexed by summarizing their visual contents through automatically extracted primitive features, such as shape, texture, color, size, and so on. Performance is poor for a CBIR system, which only uses a color feature to search for similar images from a huge database. Thus, some researchers analyzed the color distribution of the image to increase the retrieval accuracy. For example, in 1996, Gong et al. proposed an image indexing and retrieval scheme. In their scheme, an image is split into nine equal sub-areas. They presented each sub-area by using a color histogram to model the color spatial information. In 1997, Stricker and Dimai split an image into an oval central region and four corner sub-regions for image indexing Gagliardi and Schettini described the image in the CIELAB color space with two palettes and integrated different color information descriptions and similarity measurements to enhance the effectiveness of a CBIR system. Kou used the mean value, the standard deviation, and the skewness of pixels from each bin in a color histogram as the image features to search similar images. In 2003, Chan and Liu proposed a CBIR system based on color differences on edges in spiral scan order. With a view to increasing the retrieval accuracy, Chan and Liu combined the color feature with color differences among adjacent pixels for image retrieval. The color differences can be viewed as the local features of an image.

Each image in an image database may be different from all the others, but at the same time all images may share certain common characteristics. Hence, we need the statistical description of images to capture these common characteristics and use them to represent an image with fewer bits. The statistical descriptions used in this paper are the mean values and the standard deviations of images. In the proposed scheme, each pixel of a color image is represented by a vector

$$P_i = \begin{bmatrix} R_i \\ G_i \\ B_i \end{bmatrix} \quad (2.2)$$

Where Pi is the ith pixel of the image, $1 \leq I \leq M$, the size of the image is M and the components of Pi are the RGB components of the color image. The mean value (l) and the standard deviation (r) of the color image are determined as follows:

The mean value (l) and the standard deviation (r) of the color image are determined as follows:

$$\mu = \frac{1}{M}\sum_{i=1}^{M} Pi \quad (2.3)$$

$$\sigma = \left[\frac{1}{M-1}\sum_{i=1}^{M}(Pi-\mu)^2\right] \quad (2.4)$$

Vol. 2 Issue VI, June 2014

ISSN: 2321-9653

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

The above Eq.2.3 and 2.4 is the mean and standard deviation values of an image. The mean value and the stand deviation are the global feature of the image that depicts the global characteristics of images. For the purpose of enhancing the retrieval accuracy, this paper adopts image bitmap as the local feature to describe the local characteristics of the image.

In the first step to generate the image bitmap, the scheme divides the image into several non-overlapping blocks. Let Bj = {b1,b2. . .bk} be the jth block of the image, where $1 \leq j \leq m$. The symbol k is the total number of pixels in the block, and m is the total number of blocks in the image. In the second step, the scheme computes the mean value for each block. Let μBj be the mean value of the block Bj that is computed using the expression

$$\mu = \frac{1}{k}\sum_{x=1}^{k} b_x$$

(2.5)

The above Eq.2.5 is mean value of image blocks. This paper uses two different measurements for the global features and the local feature to evaluate the similarity between the two images. For the global features, μ and σ, the scheme uses Euclidean distance to calculate the similarity. On the other hand, for the local feature, the scheme uses hamming distance to evaluate the distance between the two bitmaps. Afterwards, the overall similarity is obtained by linearly combining of these two similarity values. However, the linear combination will become meaningless because the magnitude similarity value may dominate the others. We need to range the two similarity values on the same scale. This paper uses the Gaussian normalization to normalize the features into the same criterion.

$$d(A,B) = \frac{H(T^A,T^B)}{3 \times m} + \sqrt{\sum_{z\in\{R,G,B\}}\left(\mu_z'^A - \mu_z'^B\right)^2 + \sum_{z\in}}$$

(2.6)

Where $\mu_z'^A$, $\sigma_z'^A$, $\mu_z'^B$, $\sigma_z'^B$ represent the normalized mean value and standard deviation of the image A and B in z color space . The scheme computes the distance between a user query image and each image in the database by using Eq. 2.6. The image with the smallest distance is the image the most similar to the query image.

2.5 COLOR IMAGE RETRIEVAL USING M-BAND WAVELET TRANSFORM BASED COLOR – TEXTURE FEATURE

M.Acharya and M.K.Kundu [6] proposed an adaptive approach to unsupervised texture segmentation using M-band wavelet transforms. The advent of internet, there has been an explosion in the amount of visual data available to us. Over the years it has become increasingly difficult to manage ever increasing database of multimedia. Due to the difficulty of manual annotation of images, it is imperative to index images automatically, for efficient retrieval depending on its content. Several solutions to this effect have been proposed to the problem and it is still an active area of research. Several color, texture, shape etc. based approaches have been proposed independently .Texture of an image is a very useful

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

characteristic for image retrieval, so Texture based retrieval using wavelet transform is an active topic of research, owing to many beautiful properties of the wavelet transform. The ability of wavelet transform to capture textural characteristics is comparable to Gabor Transforms, and the availability of fast algorithms for computation of wavelet transform has facilitated to its increasing use over other methods.

Orthogonal M-band wavelet transform is a direct generalization of dyadic orthogonal wavelet transform. Dyadic wavelet transform is not suitable for analysis of high frequency signals, as it decomposes the frequency channel logarithmically but M-band wavelet transform divides the time-scale space both logarithmically as well as linearly thereby giving better resolution at high frequencies. The M-channel filters are used for decomposing the time-scale space into M×M sub bands. Human eye shows varying sensitivity response to different spatial frequencies. A Human Visual system divides an image into several bands, than actually visualizing the complete image as a whole. This fact motivated us to use the M-band filters which are essentially frequency and direction oriented band pass filters. Using a 1-D, 16 tap 4 band orthogonal filters with linear phase and perfect reconstruction for the multi-resolution analysis. This ensures that the textural characterization of the image is independent of the color characterization. Wavelet decomposition over the intensity plane characterizes the texture information, while the wavelet decomposition over chromaticity planes characterizes color. An over-complete decomposition resulting in the same size of the sub-bands as the image is important here, to obtain the features for each pixel of the image, to be clustered further. The 16 sub-bands coefficients obtained are used as the primitive features. Natural images exhibit spatial variation of the texture. As a result, texture based retrieval of images cannot assume the textures to be homogeneous.

A localized characterization of textures thus becomes necessary. Since the estimation of local energy for each of the 16 sub-band images. The energy values, for each sub-band and for each plane of the color image are used as the feature for a pixel and clustered using Fuzzy C-Means (FCM). It has been shown that Earth Mover’s Distance(EMD)is a very is very useful distance metric while measuring perceptual distance between two color texture the Earth Mover’s Distance(EMD) is used as the metric for similarity matching. Earth Mover’s distance uses a signature over traditional histogram for similarity matching has successfully reported the use of EMD as an efficient metric for content based image retrieval with several advantages over other similarity and dissimilarity measures. The Earth Mover’s Distance is formally discussed in the next section. The feature vector comprising of the cluster centers of the energy measurement over sub-bands, with the number of pixels of the image in each cluster comprises the image signature. To keep the computations minimum, Fuzzy C-Means (FCM) was preferred keeping the number of clusters as 3.MPEG-7 or Multimedia Content Description Interface is an ISO standard focusing on multimedia retrieval. It includes descriptors and description schemes for efficient retrieval of images, videos, audio and graphic files based on the content. Several low level feature extraction algorithms using color, texture, motion, and shape are facilitated for image and video retrieval and benchmarking of new schemes. MPEG-7 provides Scalable Color Descriptor (SCD), Color Structure Descriptor

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

(CSD), Dominant Color Descriptor and Color Layout Descriptor for color based retrieval and Texture Browsing Descriptor, Homogeneous Texture Descriptor (HTD) and Edge Histogram Descriptor (EHD) for texture based retrieval. Texture Browsing Descriptor is a compact descriptor based on Gabor Wavelets. It characterizes a texture's regularity, directionality and coarseness.HTD characterizes image texture by filtering the image with a bank of scale and orientation sensitive Gabor filters. It computes the energy and standard deviation of the energy of the output of the filter banks in the frequency bands as the features. The underlying algorithm for Texture Browsing Descriptor and HTD assumes that the images comprises of homogeneous textures. EHD on the other captures spatial distribution of edges which gives a better texture measurement even if it is not homogeneous.

2.6 SUPERVISED LEARNING OF SEMANTIC CLASSES FOR IMAGE ANNOTATION AND RETRIEVAL

N.Apostol, H.Alexander and T.Jelena, [7] proposed Semantic concept based query expansion and reranking for multimedia retrieval. A probabilistic formulation for semantic image annotation and retrieval is proposed. Annotation and retrieval are posed as classification problems where each class is defined as the group of database images labeled with a common semantic label. It is shown that, by establishing this one-to-one correspondence between semantic labels and semantic classes, a minimum probability of error annotation and retrieval are feasible with algorithms that are 1) conceptually simple, 2) computationally efficient, and 3) do not require prior semantic segmentation of training images. In particular, images are represented as bags of localized feature vectors, a mixture density estimated for each image, and the mixtures associated with all images annotated with a common semantic label pooled into a density estimate for the corresponding semantic class. This pooling is justified by a multiple instance learning argument and performed efficiently with a hierarchical extension of expectation-maximization. The benefits of the supervised formulation over the more complex, and currently popular, joint modeling of semantic label and visual feature distributions are illustrated through theoretical arguments and extensive experiments. The supervised formulation is shown to achieve higher accuracy than various previously published methods at a fraction of their computational cost. Finally, the proposed method is shown to be fairly robust to parameter tuning. From an implementation point of view, SML requires answers to two open questions. The first is how do we learn the probability distribution of a semantic class from images that are only weakly labeled with respect to that class? That is, images labeled as containing the semantic concept of interest, but without indication of which image regions are observations of that concept. We rely on a multiple instance learning type of argument to show that the segmentation problem does not have to be solved a priori: It suffices to estimate densities from all local appearance descriptors extracted from the images labeled with the concept.

The second is how do we learn these distributions in a computationally efficient manner, while accounting for all data available from each class? We show that this can be done with recourse to a hierarchical density model proposed for image indexing purposes. In particular, it is shown that this model enables the learning of semantic class densities with a

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

complexity equivalent to that of the unsupervised formulation, while 1) obtaining more reliable semantic density estimates, and 2) leading to significantly more efficient image annotation. Overall, the proposed implementation of SML leads to optimal (in a minimum probability of error sense) annotation and retrieval, and can be implemented with algorithms that are conceptually simple, computationally efficient, and do not require prior semantic segmentation of training images. Images are simply represented as bags of localized feature vectors, a mixture density estimated for each image, and the mixtures (associated with all images annotated) with a common semantic label pooled into a density estimate for the corresponding semantic class. Semantic annotation and retrieval are then implemented with a minimum probability of error rule, based on these class densities.

The overall SML procedure is illustrated its efficiency and accuracy are demonstrated through an extensive experimental evaluation, involving large-scale databases and a number of state-of-the-art semantic image labeling and retrieval methods. It is shown that SML outperforms existing approaches by a significant margin, not only in terms of annotation and retrieval accuracy, but also in terms of efficiency. This large-scale experimental evaluation also establishes a common framework for the comparison of various methods that had previously only been evaluated under disjoint experimental protocols. This will hopefully simplify the design of future semantic annotation and retrieval systems, by establishing a set of common benchmarks against which new algorithms can be easily tested. Finally, it is shown that SML algorithms are quite robust with respect to the tuning of their main parameters.

2.7 GENETIC ALGORITHM BASED IMAGE RETRIEVAL

S.F.da Silva, M.A.Batista, and C.A.Z.Barcelos, [8] proposed an Adaptive image retrieval through the use of a genetic algorithm. Considered an image database where all the images of the search universe of the system are stored. The image database is linked to the module of feature extraction. The output data of the module of feature extraction is a structure containing the identification-code, and the features vectors of color, shape and texture, for each image of the database. This data (identification-code/features) is stored in the feature database.

A Genetic Algorithm works with a population of individuals, also known as chromosomes, which represent the possible solutions to a given problem. These are usually randomly generated, however, if there is some knowledge available concerning the problem domain (heuristic), it can be incorporated into a fraction of the initial set of potential solutions. The individuals evolve in successive iterations known as generations, by means of genetic operators of crossover, where the offspring inherit the genetic features of the parents, and through mutation, that is a small random alteration in the individual's features. The evolution process halts when the system no longer improves, or when a preset maximum number of generations is reached. The output of the genetic algorithm is usually the best individual of the end population. For each problem to be solved, one has to supply a fitness function, and indeed its choice is crucial to the good performance of the GA. Given an individual (or chromosome), the fitness function must return a numerical value that represents the individual's utility. This score will be used in the selection process of the parents

Vol. 2 Issue VI, June 2014

ISSN: 2321-9653

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

and in the survival selection process for the next generation, so that the best adapted individuals will have the greatest likelihood of being chosen. The fitness function must, therefore, be appropriate for the problem being dealt with, because the GA's effectiveness will, to a large degree be determined by how faithfully the fitness function characterizes the function that is to be optimized When the user carries out a search, feature vectors of color, shape and texture are extracted from the query image by the feature extraction module and compared, through similarity measures, found in the image's feature vectors from the range of images found in the collection, which is stored in the database. The similarity measure module returns a similarity value for each image of the collection, in relation to the query image. After, the images are sorted in decreasing order of similarity (ranking) and the *n* first are shown to the user not satisfied with the result of the search, the user can carry out feedback to the system, indicating the relevant images for the search process, according to his point of view.

Based on the user's feedback the relevance feedback mechanism will adjust the similarity measure to the user's criteria through feature weights (of color, shape and texture) and region weights. The relevance feedback mechanism has as its basis a real-code genetic algorithm designed for the inference of weights that maximize the retrieval accuracy in accordance to the user's requirements expressed in the user's feedback.

The retrieval process itself is based on the local similarity pattern as where the image areas are uniformly partitioned into rectangular regions and the similarity between images is measured by corresponding region similarities. Similarity between regions, and therefore between images is computed through three features: color, shape and texture, represented by color moment's edge-direction histogram and texture neighborhood feature vectors, respectively. Distance between pairs of color feature vectors is computed by Euclidean distance, while distances between pairs of shape and texture feature vectors are computed by city-block distance.

The image similarity model is defined in Eq.2.7, where q is the query image, i an image belonging to the database, r an image region, f an image feature, I the whole image database and R = {r1, r2, . . . , r16}. $S_F(q, i, r, f)$ is the similarity between q and i, in relation to f in the region r, $\omega_F(r,f)$ weighting with real values in range [−1, 1] the importance of f in the region r, $\omega_R(r)$ weighting with real values in range [−1, 1] the importance of the region r, and finally $S_I(q,i)$ defines the image similarity between q and i.

$$S_I(q,i) = \sum_{r\in R}(\omega_R(r)\sum_{f\in F}(\omega_F(r,f)S_F$$
(2.7)

The task of the GA consist of finding the $\omega_R(r)$ and $\omega_F(r,f)$ that maximizes the retrieval accuracy according to the context defined by the query image and the set of relevant images chosen by the user. The use of negative and positive weights for the features into the GA allows one to express, in a continuous way, the concepts of relevance, irrelevance and undesirability in the similarity model used

ISSN: 2321-9653

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

Existing System

In the previous systems, image content is stored in visual features which can be divided into four classes according to the properties they describe. The classes are color, texture, shape, and structure. Color and texture contain important information but, for instance, two images with similar color histograms can represent very different things. Therefore the use of shape-describing features is essential in an efficient content-based image retrieval system. Although shape description has been intensively researched, there exists no direct answer as to which kind of shape features should be incorporated into such a system. A major problem in automatic feature extraction is segmentation. Even if it were known that there is a single object in the image, it is in general a non-trivial problem to locate it. On the other hand, when there are no specific objects the result of segmentation is probably an irrelevant part of the original image. For the use of a general database of images, such as the World Wide Web, it might then be reasonable to use some statistical shape features for the whole image instead. Another basic concept to be considered in selecting an appropriate shape description technique is whether some invariant properties such as transformation, rotation, and scaling invariances are needed. The use of these is not always beneficial because they reduce the discrimination power of the features.

PROPOSED SYSTEM

The aim of this work was to design shape features for a content-based image retrieval system which is going to be used as an image search engine for large-scale databases like the World Wide Web. The system is based on the Tree Structured Self Organizing Maps which are used as the indexing structure of the images. Ideally, images that are similar to each other with respect to a particular feature extraction method, should cluster together on the corresponding map. The responses on the different maps are combined in such a way that the most relevant features are automatically weighted as the query proceeds. The incorporated relevance feedback mechanism thus adapts the system to the user's preferences until he/she finds the preferred images from the database.

1)Using the group sparsity, a feature selection framework is introduced to solve the image annotation problem, which improves annotation performance. Group sparsity enables either selecting or removing a whole group of features. If this group of features is removed in the training data, these features also do not need to be extracted in the testing stage.

2) Using keyword similarity and relevance feedback, a novel method is proposed to obtain iteratively similar and dissimilar pairs for training purposes. The positive and negative samples from our method are more separated than those using previous techniques, which reduce noise and benefit the annotation performance.

3) By selecting similar and dissimilar image pairs with category information instead of keyword information, this framework can also be used to solve image retrieval problems.

4) We give some insight into the feature properties and provide in-depth comparisons between regularization methods in this image annotation application.

Advantages of Proposed

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

Relevance feedback (RF) has been demonstrated to be a powerful tool which involves the user in the loop to enhance the performance of CBIR. Popular RF schemes can exhibit some general limitations of over sensitivity to subjective labelling by users and the inability to accumulate knowledge over different

sessions and users. The conventional process of RF is as follows:

1) from the retrieved images, the user labels a number of relevant samples as positive feedbacks, and a number of irrelevant samples as negative feedbacks;

2) the CBIR system then refines its retrieval procedure based on these labelled feedback samples to improve retrieval performance

These two steps can

be carried out iteratively. As a result, the performance of the system can be enhanced by gradually learning the user's preferences.

1. Subspace learning based methods either find a low-dimensional subspace of the feature space, such that the positive and negative samples are well separated after projection to this subspace.
2. Support vector machine (SVM) based methods either estimate the density of positive instances or regard RF as a classification problem with the positive and negative samples as training sets.

DRAWBACKS OF EXISTING

To give text annotations to all images manually is tedious and impractical . In addition, automatic image annotation is generally beyond current techniques. Finally, a picture says more than a thousand words.

It is based on representing images by using low-level visual features, which can be automatically extracted

from images, to reflect the color , texture , and shape

METHODOLOGY

Image Acquisition

IA is a significant super-set of the support for digital still imaging drivers that was provided by the Still Image Architecture (STI) in Windows. Whereas STI only provided a low-level interface for doing basic transfers of data to and from the device (as well as the invocation of an image scan process on the Windows machine through the external device), IA provides a framework through which a device can present its unique capabilities to the operating system, and applications can programmatically take advantage of those features. According to Microsoft, IA drivers are made up of a user interface (UI) component and a driver core component, loaded into two different process spaces: UI in the application space and the driver core in the IA service space.

Information Extraction

Image Information extraction (IIE) is a type of image information retrieval whose goal is to automatically extract structured pixel information, i.e. categorized and contextually

ISSN: 2321-9653

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

and semantically well-defined data from a certain domain, from unstructured machine-readable images

The design of a system to extract information automatically from paper-based maps and answer queries related to spatial features and structure of object and color data is considered. The foundation of such a system is a set of image-analysis algorithms for extracting spatial features. Efficient algorithms to detect symbols, identify and track various types of lines, follow closed contours, compute distances, find shortest paths, etc. from simplified map images have been developed. A query processor analyzes the queries presented by the user in a predefined syntax, controls the operation of the image processing algorithms, and interacts with the user.

Image Querying

The growth of the World Wide Web over the past decade or so, vast amounts of information is available to anyone in possession of a personal computer with a modem and an Internet connection. Tasks such as finding a favorite picture have been made easy by search engines like Google. One can simply type in a few lines about the picture, and then it’s just a matter of sorting through a few top matches before one has the entire collection of similar pictures on the screen by specifying the picture itself.

These are displayed as small thumbnails in the upper right part of the window. If one of these was the desired image, the user simply clicks on that thumbnail to retrieve the full-sized image. Otherwise, if the desired image doesn't appear among the thumbnails, the user may modify the query and ``match" again. Because the queries may be matched against the database very quickly, we have implemented an ``interactive" mode for the application. In this mode, the application re-evaluates the painted query in the database as each stroke is drawn. Every time the user pauses from painting for a moment, the application retrieves a new set of thumbnails. Median time to retrieve the target image appears to be around twenty seconds for a database images.

Similarity Identification

Similarity Identification and cluster analysis indicated that subjects could precisely perceive the similarity between every two images according to the salience of form features. It has been found that even though there are multiple and complicated combinations of form features in objects, it is possible for people to make decisions for the feature matching behavior

For the polar image the Fourier transform and decimation are performed similarly as with the Fourier features and a 128-dimensional feature vector is obtained. The method is invariant to translation in the polar plane, and therefore rotation invariant with respect to the center of the image and translation invariant along the radius from the center.

we consider the analysis of thousands of unorganized, low resolution images of an object. With very low resolution images, standard computer vision techniques of finding corresponding points and solving for image warping parameters or 2D geometry may fail. Two recent techniques in statistical pattern recognition, locally linear embedding (LLE) and Isomap,

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

give a mechanism for finding the structure underlying point sets for which comparisons or distances are only meaningful between nearby points.

Image Retrieval

An image retrieval system is a computer system for browsing, searching and retrieving images from a large database of digital images. Most traditional and common methods of image retrieval utilize some method of adding metadata such as captioning, keywords, or descriptions to the images so that retrieval can be performed over the annotation words. Manual image annotation is time-consuming, laborious and expensive; to address this, there has been a large amount of research done on automatic image annotation. Additionally, the increase in social web applications and the semantic web have inspired the development of several web-based image annotation tools

All the Fourier-based features presented here are sensitive to occlusion: the direct use of the Fourier transform may lead to very different magnitude spectra for occluded images. In addition, if some parts of an image are missing, the calculation of the centroid will go wrong and significantly differing log–polar images will result.

Feedback and Ranking

The proposed model assumes that the user expects the best possible retrieval results after each Feedback iteration, the search engine is required to return the most semantically relevant images based on the previous feedback samples. The system likes to label a large number of images during Feedback and does not limit the number of feedback iterations for image labelling. With the proposed system, we can embed various Feedback. when a query (image) is given, the low-level visual features are extracted. Then, all images in the database are sorted based on the Euclidean measure. If the user is satisfied with the results, the retrieval process is ended. However CBIR assumes that the user expects the best possible retrieval results after each Feedback iteration, the search engine is required to return the most semantically relevant images based on the previous feedback samples.

Algorithm

PSO

Initialize feature weightings randomly, then use the variances of the positive and negative feedback samples' features as study principle, utilize particle swarm optimization (PSO) algorithm to optimize weightings according to user's retrieval requirement, and obtain retrieval results at last. Experiments show that the proposed algorithm is validity.

Initial Population

The IPSO requires a population of potential solutions to be initialized at the beginning of the PSO process. These potential solutions are encoded as binary values and that are called as chromosomes. Chromosomes are represented from the image features (i.e., color, texture, and edge) in an image. Usually, the initialization process varies with the applications; here, we adopt the first query results of as ample image as initial candidate images.

Fitness Function

ISSN: 2321-9653

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

The fitness function is employed to evaluate the quality of the chromosomes in the population. The use of IPSO allows the fusion of human and computer efforts for problem solving. Since the objective of this system is to retrieve the images that are most satisfied to the users' need, the evaluation might simultaneously incorporate users' subjective evaluation and intrinsic characteristics of the images. Hence, in our approach, the quality of a chromosome C with relation to the query q is defined as

$$F(q,C) = \omega_1 . sim(q,C) + \omega_2 . \delta$$

(4.1)

where $sim(q,C)$ represents the similarity measure betweenimages, δ indicates the impact factor of human's judgment,the coefficients w_1 and w_2 determine the relative importance of them to calculate the fitness, and Σ wi =1.Inthis paper, they are both set to 0.5. The similarity measure between images is defined as

$$sim(q,C) = \sqrt{\sum_{t\in\{H,S,V\}} \left(\mu_t'^q - \mu_t'^C\right)^2 + \sum_{t\in\{H,S,V\}} \left(\sigma_t'^q - \sigma_t'^C\right)^2}$$

$$+\frac{H(BM^q,BM^C)}{3\times M}$$

$$+|E^q - E^C| + \frac{|EHD^q - EHD^C|}{5\times 80}$$

(4.2)

Where μ_t' and σ_t' represent the normalized mean value and standard deviation of the image I in t color space BMI means the image bitmap feature of the image I. EI and EHDI represent the entropy and Edge Histogram Descriptor of the image I. For two images, the hamming distance used to evaluate the image bitmap similarity is defined by

$$H(BM^q,BM^C) = \sum_{j=1}^{m} \left(IH_j^q - IH_j^C\right)$$

$$+\sum_{j=1}^{m}\left(IS_j^q - IS_j^C\right) + \sum_{j=1}^{m}\left(IV_j^q - IV_j^C\right)$$

(4.3)

A user's preference is included in the fitness evaluated by the user. We use an impact factor to indicate the human's judgment or preferences, and the values of the impact factor are carried out with constant range from 0.0 to 1.0with an interval of 0.1.

CONCLUSION

The image retrieval is done by considering color, Texture, and Edge features in the proposed technique. The color and bitmap method involves extracting only the local and global features such as mean and standard deviation. But in the proposed technique, color, texture, and Edge features are extracted and then Interactive Genetic Algorithm is applied on these image features.

The estimated parameters in the proposed technique include the Precision and Recall values. The precision value improves by nearly 30% and the Recall

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

value improves 5% when compared to color and image bitmap method.

FUTURE ENHANCEMENT

The visual features and annotation methods are combined may increase the efficiency of retrieval. The query techniques can be explored in particular domain-specific using this system. Image based video retrieval can be implemented for efficient retrieval of videos.

BIBLOGRAPHY

[1] M.Antonelli, S.G.Dellepiane, and M.Goccia, "Design and implementation of Web based systems for image segmentation and CBIR," IEEE Trans. Instrum. Meas., vol. 55, no. 6, pp. 1869–1877, Dec. 2006.

[2] B.Suh, H.Ling, B.B.Bederson, and D.W.Jacobs, "Automatic thumbnail cropping and its effectiveness," in Proc. ACM Symp. User Interface Software and Technology, 2003, pp. 95–104.

[3] N.Gnaneswara Rao, Dr.Vijaya Kumar, V.Venkatesh Krishna, "Texture Based Image Indexing and Retrieval", IJCSNS International journal of Computer Science and Network Security, Vol.9 No.5, May 2009

[4] J.Z.Wang, G.Wiederhold, O.Firschein and S.X.Wei. "Content Based image indexing and searching using Daubechie's wavelets"- Digital libraries, pp.311-328, 1998

[5] T.C.Lu and C.C.Chang, "Color image retrieval technique based oncolor features and image bitmap," Inf. Process. Manage., vol. 43, no. 2,pp. 461–472, Mar. 2007.

[6] M.Acharya and M.K.Kundu,"An adaptive approach to unsupervised texture segmentation using M-band wavelet transforms, Signal processing 81, pp.1337-1356, 2001.

[7] N.Apostol, H.Alexander and T. Jelena, "Semantic Concept Based Query Expansion and Reranking for Multimedia Retrieval",proc.ACM Int'l Conf.Multimedia(Multimedia' 07), Sept.2007

[8] S.F.da Silva, M.A.Batista, and C.A.Z.Barcelos, "Adaptive image retrieval through the use of a genetic algorithm," in Proc. 19th IEEE Int.Conf. Tools with Artif. Intell., 2007, pp. 557–564.

[9] E. J. Delp and O. R. Mitchell, "Image coding using block truncation coding," IEEE Trans. Commun., vol. COM-27, no. 9, pp. 1335–1342, Sep. 1979.

[10] R.M.Haralick and L.G.Shapiro, Computer and Robot Vision: Volume I. Reading, MA: Addison-Wesley, 1992.

[11] T.Sikora, "The MPEG-7 visual standard for content description—An overview," IEEE Trans. Circuits Syst. Video Technol., vol. 11, no. 6, pp. 696–702, Jun. 2001.

[12] G.Beligiannis, L.Skarlas, and S.Likothanassis, "A generic applied evolutionary hybrid technique for adaptive system modeling and information mining," IEEE Signal Process. Mag.Special Issue on "Signal

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

Processing for Mining Information", vol. 21, no. 3, pp. 28–38, May 2004.

[13] G.N.Beligiannis, L.V.Skarlas, S.D.Likothanassis, and K. G. Perdikouri, "Nonlinear model structure identification of complex biomedical data using a genetic-programming-based technique," IEEE Trans. Instrum. Meas., vol. 54, no. 6, pp. 2184–2190, Nov. 2005.

[14] C.-Y. Chang and D.-R. Chen, "Active noise cancellation without sec-ondary path identification by using an adaptive genetic algorithm," IEEE Trans. Instrum. Meas., vol. 59, no. 9, pp. 2315–2327, Sep. 2010.

[15] G. Paravati, A. Sanna, B. Pralio, and F. Lamberti, "A genetic algorithm for target tracking in FLIR video sequences using intensity variation function," IEEE Trans. Instrum. Meas., vol. 58, no. 10, pp. 3457–3467,Oct. 2009.

[16] Z.Steji, Y.Takama, and K.Hirota, "Genetic algorithm-based relevance feedback for image retrieval using local similarity patterns," Inf. Process. Manage., vol. 39, no. 1, pp. 1–23, Jan. 2003.