Content Based Image Retrieval Genetic Algorithm for Relevance Feedback

Nikita R. Hatwar[1], Rajendra F. Hatwar[2]
AP, Department of Information Technology, Priyadarshini College of Engineering, Nagpur, India[1]
Scientist Grade B, Department of Information Technology, National Informatics Centre, Nagpur, India[2]

Abstract

Image retrieval using content has become a hot topic in the field of Digital Image Processing and Computer Vision. In this proposed work, we focus on the problem of image retrieval using genetic algorithm for relevance feedback. Texture and Color are the important features in CBIR. Texture and Color is extracted using Multispectral Simultaneous Autoregressive Model (MSAR). The color is represented by ratios of sample color means. The features are extracted by segmenting the image into regions of uniform texture/color using an unsupervised histogram clustering approach that utilizes MSAR and color features. The principle of Genetic Algorithm is based on "survival of the fittest". After extracting the features Genetic Algorithm (GA) will be applied to retrieve the images from the database. The fittest image will be retrieved.

Keywords

Color, Texture, Multispectral Random Field Models, Color Texture Segmentation.

1. Introduction

With the advances in the computer technologies and the advent of the World Wide Web, there has been an explosion in the amount and complexity of digital data being generated, stored, transmitted, analyzed, and accessed. Much of this information is multimedia in nature, including digital images, video, audio, graphics, and text data. In order to make use of this vast amount of data, efficient and effective techniques to retrieve multimedia information based on its content need to be developed. Among the various media types, images are of prime importance. Not only is it the most widely used. media type besides text, but it is also the basis for representing and retrieving videos and other multimedia information. This paper deals with the retrieval of images based on their contents, even though the approach is readily generalizable to other media types. Keyword annotation is the traditional image retrieval paradigm. In this approach, the images are first annotated manually by keywords. They can then be retrieved by their corresponding annotations. However, there are three main difficulties with this approach, i.e. the large amount of manual effort required in developing the annotations, the differences in interpretation of image contents, and inconsistency of the keyword assignments among different indexers [1], [2], [3].

As the size of image repositories increases, keyword annotation approach becomes infeasible. This paper describes a color texture-based retrieval system that finds images similar to a query image in a large database. The similarity criterion is based on the notion of containing same/similar color texture regions. The approach involves characterizing color texture using features derived from a class of multispectral random field models. These features are then used in an unsupervised histogram clustering-based segmentation algorithm to find regions of uniform texture in the query image.

The retrieval process involves identifying those images in the database that contain the same kind of texture found in the query image. The process of retrieval is done with the help of GA as it is the most efficient optimized searching algorithm.

2. Texture Characterization with Multispectral simultaneous Autoregressive Model

The texture of the color images is characterized using a class of Multispectral Random Field Model called the Multispectral Simultaneous Autoregressive (MSAR) Model [5], [6]. The MSAR model is shown to be effective for color texture synthesis and classification [5], [6]. For mathematical simplicity, the model is formulated using a toroidal lattice assumption. A location within a two-dimensional M by M lattice is denoted by $\mathbf{s}$. The value of an image observation at location $\mathbf{s}$ is denoted by the vector value $\mathbf{y}(\mathbf{s})$, and the image observations are assumed to have zero mean. The MSAR model relates each lattice position to its neighbouring pixels, both within

and between image planes, according to the following model equation:

$$y_i(s) = \sum_{j=1}^{p} \sum_{r \in N_{ij}} \theta_{ij}(r) \, y_j(s \ominus r) + \sqrt{\rho_i} \, w_i(s), \quad i = 1 .. \, P$$

$y_i(s)$ = i^{th} element of a zero mean observation vector
s and r = two dimensional lattices
P = number of image planes (in this case, P = 3, for the three image color planes: Red, Green, and Blue)
Nij = neighbour set relating pixels in plane i to neighbour in plane j (only interplane neighbour sets, i.e. Nij, i=j, may include the (0, 0) neighbour)
θ_{ij} = coefficients which define the dependence of $y_i(s)$ on the pixels in the its neighbour set Nij
ρ_i = noise variance of image plane i
$w_i(s)$ = i.i.d random variables with zero mean and unit variance
$\ominus$ denotes modulo M addition in each
The parameters associated with the MSAR model are *q* and *r* vectors which collectively characterize the spatial interaction between neighbouring pixels. These vectors are taken as features representing the underlying color texture present in the M by M image.

3. Color Characterizations

In addition to modelling color texture, the color content of the image by itself is also important. Additional features focusing on the color alone are also considered. This is done using the sample mean of the three color components in the red, green, and blue (RG B) space. The defined feature vector is:

$$\mathbf{f}_R = \left\{ \frac{\hat{\mu}_r}{\hat{\mu}_g}, \frac{\hat{\mu}_r}{\hat{\mu}_b} \right\}$$

with $\hat{\mu}_i s$ being the sample mean of the respective color component. This is a two-dimensional feature vector containing ratios of the color means. It is assumed that the observed value at each pixel is a product of illumination and spectral reflectance. Under this assumption, the ratios of the color means are invariant to uniform changes in illumination intensity, i.e. the power of the illumination source changes uniformly across the spectrum. Such a change in illumination would result in changing each $\hat{\mu}_i s$ by a scale factor making the ratios invariant to illumination changes. This property makes the color features more robust.

4. Segmentation with a Histogram Based Algorithm

The first step in the retrieval process is to segment the query image into regions of uniform color texture. It should be noted that since the ultimate goal is to retrieve images similar to the one presented to the system, it suffices to find dominant regions of texture in the image and locating very exact boundaries of regions is not as important.

The segmentation algorithm used in this work relies on scanning the image with a sliding window and extracting texture and color features from each window. These features are then clustered using an unsupervised histogram-based algorithm. Mapping the identified clusters back into the image domain results in the desired segmentation. This neighbour set results in a 20-dimensional MSAR feature vector. Therefore together i^{th} the color feature set, a 22-dimensional feature vector is used to characterize each window.
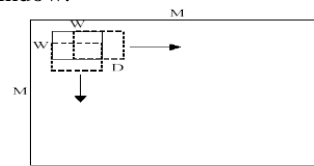

Figure 1: Feature Vector Extraction with Sliding Window

The windowing operation consists of sliding a window from left to right and top to bottom across the image. M is the size of the image in pixels, W is the size of the window in pixels, and D is the size of the sliding step in pixels. D is set to 4 pixels for this work. To find the optimum window size for each case, the size of the window W varies from W1 = 4 pixels to W2 = 28 pixels in increments of 4 pixels. In later sections, it is explained how the best window size is decided. The texture bounded by each window is characterized using the MSAR and color features. The neighbourhood used for the MSAR model is a set that contains neighbors above, below, to the left, and to the right of the pixel as illustrated in Figure 2. This neighbour set is used for both inter and intraplanes of the model.

<pre>
    X
  X O X
    X
</pre>

Figure 2: Neighbour Set used with the MSAR Model

4.1 Clustering Algorithm

Once all 22-dimensional features are extracted from the sliding window, they are clustered in the feature space using an unsupervised histogram-based peak climbing algorithm [7], [8], [9]. The 22-dimensional histogram is generated by quantizing each dimension and dividing the space into hyper boxes. Next, the number of feature vectors falling in each hyper box is counted and this count is associated with the respective hyper- box creating the required histogram.

After the histogram is generated in the feature space, a peak climbing clustering approach is utilized to group the features into distinct clusters. This is done by locating the peaks of the histogram. In Figure 3 this peak climbing approach is illustrated for a two-dimensional space.
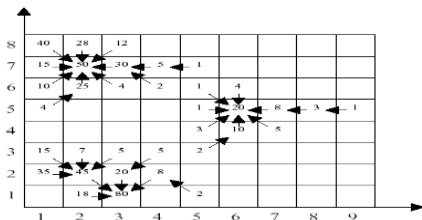

Figure 3: Illustration of the Peak Climbing Approach for a Two-dimensional Space

The number in each cell (hyper box) represents a hypothetical count for the feature vectors captured by that cell. By examining the counts of the 8-neighbors of a particular cell, a link is established between that cell and the closest cell having the largest count in the neighbourhood. At the end of the link assignment, each cell is linked to one parent cell, but can be parent of more than one cell. A peak is defined as being a cell with the largest density in the neighbourhood, i.e. a cell with no parent. A peak and all the cells that are linked to it are taken as a distinct cluster representing a mode in the histogram. Once the clusters are found, the windows associated with features grouped in the same cluster are tagged as belonging to the same category.

A major component of this algorithm is the number of quantization levels associated with each dimension. To decide this parameter, the total number of non-empty cells and the percentage of them capturing only one vector for each selection of quantization levels are examined. The best number of quantization levels is selected as the largest one that maximizes the measure below [9].

$$M_i = (N_{ci} - N_{ui}) * Nui$$

N_{ci} = number of non-empty cells
N_{ui} = number of cells capturing only one sample
The algorithm also includes a spatial domain cluster validation step. This step involves constructing a matrix B for each cluster m as:

$$B_m (i, j) = 1 \text{ if sample } \in \text{ cluster m}$$
$$B_m (i, j) = 0 \text{ otherwise}$$

The (i, j) index corresponds to the location of a window used to scan the image. A cluster is considered compact if only a very small number of its elements (1's) have a 0 neighbour, i.e. a cluster is considered valid (compact) if only a very small number of its elements have neighbouring elements that do not belong to that cluster. A cluster that does not pass this test is merged with a valid cluster that has the closest centroid to it.

During the segmentation process, the best window size to scan the image with is chosen in an unsupervised fashion. The optimum window size is obtained by sweeping the image with varying window sizes, and choosing the smallest one out of at least two consecutive window sizes that produce the same number of clusters. The parameters for this operation are those described in Section 4.1.

5. Genetic Algorithm

Genetic Algorithms are the search and optimization procedures that are motivated by the principles of natural selection and natural genetics. Some fundamental ideas of genetics (for representation) and natural evolutionary process (for algorithm) are borrowed and used artificially to construct search algorithm that are robust and required minimum problem information. It provides good optimization capabilities when search space is large

Various features of GA are: 1) They are very robust technique. 2) Easy to program. 3) Operate on several solutions simultaneously.

5.1 Algorithmic Phases

1. Initialize the Population
2. Select individuals for the mating pool
3. Perform Crossover
4. Perform Mutation
5. Insert offspring into the population
6. Repeat step 2 until satisfying result is obtained.

First we will initialize the population in this case the population will be the set of images in the database. Individual images will be selected from the population, perform crossover and mutation operation. Add the new offspring to the mating pool.

5.2 Genotype and Phenotype

Genes are the basic "instructions" for building an organism. A chromosome is a sequence of genes. Biologists distinguish between an organism's genotype (the genes and chromosomes) and its phenotype (what the organism actually is like) Example: You might have genes to be tall, but never grow to be tall for other reasons (such as poor diet).

5.3 Types of GA

1) Binary coded GA: Parameters are coded into bits (0's &1's) Close to natural genetics. The search space is discrete
2) Real-coded GA (RCGA): Parameters are used as it is and are close to natural formulation of problem and search space continuous

For this the RCGA will be used that is whatever is the value of the selected features will be taken as it is. Chromosome represents a solution and population is a collection of such solutions.

5.4 Operators in GA

There are Reproduction or Selection operators and Variation (Crossover or Mutation) operators. In GA the role of selection and recombination (or crossover) operators are very well defined. Selection operator controls the direction of a search and recombination operator generates new regions for a search. The performance of a GA on a particular problem strongly depends on the degree of exploration and exploitation of search space.

Reproduction or Selection Operator

It creates a mating pool (or intermediate population) by performing the following tasks:

Identify good (usually above-average) solutions in a population. Make multiple copies of good solutions. Eliminate bad solutions from the population

Some common methods (operators) are:

Tournament selection, Proportionate selection and Ranking selection. The selection operator cannot create any new solutions in the population. It only makes more copies of good solutions at the expense of not-so-good solutions. For this purpose ranking selection will be used.

Crossover Operator:

The recombination (or crossover) operation is a method for sharing information among chromosomes. The crossover operator is the main search operator in the GA. In crossover solutions are picked from mating pool at random and some portion of the solutions is altering to create new solutions. Crossover operator in Binary GA:

Single point crossover operator, Multi point crossover operator and Uniform crossover operator Crossover operators in Real Coded GA:

They are neighbourhood based and use uniform, polynomial, triangular or lognormal probability distribution. The role of mutation in GA is to restore lost or unexpected genetic material into population to prevent the premature convergence of GA to sub-optimal solutions; it ensures that the probability of reaching any search space is not zero.

6. Relevance Feedback

In order to help the users retrieve the correct images they seek, relevance feedback techniques have been developed. This involves allowing users to make further selections from the initial lot of images, presented for a query. The users can keep on refining the search from the results of the previous search until they get the desired images or closest to what they desire. Issues regarding relevance feedback have been presented where the linear and kernel-based biased discriminate analysis, Bias Map is proposed to fit the unique nature of relevance feedback as a small sample biased classification problem. Also a word association via relevance feedback (WARF) formula is presented and tested for erasing the gap between low-level visual features and high-level semantic annotations during the process of relevance feedback [12].

7. Image Retrieval Procedure

The retrieval system is based on the described segmentation approach and the associated features. All the images in the database are first segmented. The texture associated with each region of each image is then characterized by a single 22-dimensional feature vector. This is done by fitting the maximum fitting square to each region and extracting features from this square. Of course this operation could be done in an off-line manner to speed up the process during the query phase. When a query image is presented, a similar procedure is performed. For instance, the image may be segmented into three regions. A feature vector associated with each of the regions is then extracted from the corresponding maximum fitting square. Next, these three feature vectors are presented to the space containing the feature vectors of the database one at a time. For each case, those images that contain a feature vector similar to the one being presented are found. Each 22-dimensional feature vector are stored as one

chromosome. If three regions are produced than the length of the chromosome will be 66. The images with the minimum Euclidean distance will be retrieved from the database. It should be noted that the current approach only retrieves images that contain a textured region similar to one of the regions in the query image.

8. Conclusion

This paper presents a color image retrieval system that is based on modelling color textures in the image. The texture is characterized with a Multispectral Random Field Model. Features associated with this model along with color-only features are used to segment the image. A histogram-based clustering algorithm is used for this purpose. Once the image is segmented, features associated with each of its regions are used to find images similar to it in a database. The technique used to retrieve images from database is MSAR. For relevance feedback GA is used.

References

[1] R. Barber, W. Niblack, and et al., "The QBIC project: Querying images by content using color, texture and shape," in Proc. SPIE Conf. on Vis. Commun. and Image Proc., Feb 1994.

[2] B. Dom, D. Lee, , D. Petkovic, D. Steele, H. Sawhney, J. Ashley, J. Hafne, M. Flickner, M. Gorkani, P. Yanker, Q. Huang, W. Niblack, , and, "Query by image and video content: The QBIC system," IEEE Computer, 1995.

[3] C. Faloutsos, D. Petkovic, M. Flickner, R. Barber, W. Equate and W. Niblack, "Efficient and effective querying by image content," tech. rep., IBM Research Report, 1993.

[4] B. S. Manjunath and W. Y. Ma, "Texture Features for Browsing and Retrieval of Image Data," *IEEE Trans. On Pattern Analysis and Machine Intelligence*, vol. 18, no.8, August 1996, pp. 837-842.

[5] A. Khotanzad and J. W. Bennett, "Multispectral Random Field Models for Synthesis and Analysis of Color Images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, no. 3, March 1998, pp. 327-332.

[6] J. W. Bennett, "Modelling and Analysis of Gray Tone, Color, and Multispectral Texture Images by Random Field Models and Their Generalizations," *Southern Methodist University*, Dissertation, May 1997.

[7] A. Khotanzad and J. Y. Chen, "Unsupervised Segmentation of Textured Images by Edge Detection in Multidimensional Features," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 11, no. 4, April 1989, pp. 414-421.

[8] J. Y. Chen, "Segmentation of Textured Images by Edge Detection in Multidimensional Features," *Southern Methodist University*, Dissertation, September 1987.

[9] A. Bouarfa and Khotanzad, "Image Segmentation by a Parallel, Non-Parametric Histogram Based Clustering Algorithm," *Pattern Recognition*, vol. 23, no. 9, September 1990, pp. 961-963.

[10] Khotanzad, Alireza, and Orlando J. Hernandez. "Color Image Retrieval Using Multispectral Random Field Texture Models." IEEE Signal Processing Society Ninth DSP (DSP 2000) Workshop. 2000.

[11] Bennett, Jesse, and Alireza Khotanzad. "Multispectral random field models for synthesis and analysis of color images." Pattern Analysis and Machine Intelligence, IEEE Transactions on 20.3 (1998): 327-332.

[12] Zhou, X.S., Huang, T.S. (2002). Relevance Feedback in Content-Based Image Retrieval: Some Recent Advances, *Proceedings of 6th Joint Conference on Information Sciences*, March 8-13, 2002, Research Triangle Park, North Carolina, USA, 15-18.

Books

[13] Jain, Anil K. Fundamentals of digital image processing. Vol. 3. Englewood Cliffs: prentice-Hall, 1989.

[14] Rafael C. Gonzalez and Richard E. Woods, "Digital Image Processing" Addison-Wesley, 1992.

[15] Golberg, David E. "Genetic algorithms in search, optimization, and machine learning." Addion wesley 1989 (1989).

[16] Deb, Kalyanmoy. Multi-objective optimization using evolutionary algorithms. Vol. 16. John Wiley & Sons, 2001.

The author was born in Nagpur (MS) and her date of birth is 21st June 1975 She has done her B.E.(CSE) from VNIT, Nagpur and M. Tech (CSE) from GHRCE Nagpur. She is life member of ISTE and CSIAST. She has several papers to her credit.

The author was born in \\Kamptee, Nagpur and his date of birth is 20th Oct., 1971. \He has done his M. Sc. (Computer Science) from R. T. M., Nagpur University.