Automatic Image Annotation and Retrieval: A Survey

M.Sangeetha[1],K.Anandakumar[2],A.Bharathi [3]

[1]Research Scholar, Bharathiar University, Coimbatore - 641 046, Tamil Nadu
[2]Department of MCA, BannariAmman Institute of Technology, Erode - 638401, Tamil Nadu, India
[3]Department of IT, Bannariamman Institute of Technology, Erode - 638401, Tamil Nadu, India
---***--

Abstract - *Image Annotation and Retrieval (IR) is one of the most exciting and fastest growing research areas in the field of multimedia technology. Manual image annotation is time-consuming, laborious and expensive. Automatic Image annotation and retrieval is a dominant research area in computer science. It is concerned with the storage of images and assigning meaningful keywords to it. There are several methods developed for efficient automatic image annotation which uses various optimization techniques. Image retrieval is a technique for searching images in large databases. It is the study concerned with the technique of searching and retrieving digital images from a collection of image databases. To perform image retrieval, user gives query as images, text as a keyword(s), and image links, then the system will retrieve images "similar" to the query. The comparison used for searching criteria could be meta tags, color spreading in images, region or shape attributes. In image search, searching of images is carried out based on combined metadata such as keywords and text that are annotated with each image. The purpose of this paper is to show the survey study done on the optimization techniques for Image annotation and retrieval .*

Key Words: Image Annotation, Retrieval, Optimization, Content based image retrieval, Feature Extraction

1.INTRODUCTION

With the advent of various multimedia devices such as digital cameras, mobile phones etc.. The number of images has increased dramatically. Thus, Effective methods are required for organizing, searching and browsing these images. Many search engines use text-based searching methods for retrieving images. Indexing images based on its semantic content will improve the image search quality. However, as it is impossible to manually annotate all images, Automatic image annotation (AIA) might be a promising solution. The goal of Automatic image annotation is to assign meaningful keywords to an image by automatically checking the semantics of the image. Automatic image annotation is essential to label a huge collection of unlabeled photos. Traditional methods of image annotation are not adequate as the amount of images to be indexed is huge, which makes it impractical and error prone. Thus to enhance annotation performance, optimization techniques are used. Optimization is a commonly encountered mathematical problem in all engineering disciplines. It means finding the best possible solution. Feature weighting is a technique used to approximate the optimal degree of influence of individual features. In this survey various methods to annotate images are discussed. Also, various approaches for optimum feature selection are interpreted and their limitations are discussed.

A common ground in most of current IR systems is to exploit low-level features such as color, texture and shape, which can be extracted by a machine automatically. While semantic level retrieval would be more desirable for users [12], given the current state of technology in image understanding, this is still very difficult to achieve. This is especially true when one has to deal with a heterogeneous and unpredictable image collection such as from the WWW. Current research fights to bridge the gap between low-level,

statistical, descriptions and high-level semantic content. Thus methods inspired by artificial intelligence [13], textual retrieval [14, 15], and psychology & human-computer interaction [16, 17], are starting to influence the research. Synthetically, image retrieval starts off by the design of a robust, meaningful and flexible feature set to characterize all plausible images in the collection. Then clever manipulation of the features tries to uncover some higherlevel similarity between the query and the database candidates. An interactive, iterative, and user-oriented query process

Typical image retrieval process. Early IR systems [18, 19, 20, 21] mainly relied on a global feature set extracted from images .For instance, color features are commonly represented by a global histogram. This provides a very simple and efficient representation of images for the retrieval purpose. However, the main drawback with this type of systems is that they have neglected spatial information. More recent systems have addressed this problem. Spatial information is either expressed explicitly by the segmented image regions [6, 7, 8] or implicitly via dominant wavelet coefficients [10, 11]. Most systems use the query by example approach, where the user selects one or several images, and the system returns the ones judged similar. An alternative way of querying the image database based on content, is by allowing the user to sketch the desired image's color/texture layout, thus abstracting himself, the objects searched for [22,23, 18]. Other more targeted systems allow the user to specify spatial constraints on the dominant objects. All of these methods suffer somewhat from the drawback that the system relies on the users abilities and does not adapt to his/her needs.

2. FEATURE DESCRIPTORS FOR IMAGE ANNOTATION AND RETRIEVAL

A feature is a significant piece of information extracted from an image which is used to interpret an image. Image feature extraction is a type of dimensionality reduction that detect and isolate various desired portions of a digitized image. An image has several visual attributes such as color, texture, shape etc.

Color is one of low-level visual features, which is commonly used for image processing. It is invariant in image scaling, rotation and translation and has the characteristics of easy calculation. The comparison of color histograms[3][5][6] is one of the most widely used techniques for image annotation. A histogram is the distribution of the number of pixels of an image. The number of elements in a histogram relates to the number of bits in each pixel of an image. RGB[4] color model (RGB: Red, Green, Blue) is the most commonly used color model. It is an additive color model: the colors red, green, and blue are combined to generate other colors. It is not perceptually uniform this means that the measure of the variation perceived by a human is different from the mathematical distance. HSV (HSL, HSB)[1] models are much closer to human eye perception of color The components of these models are: Hue, Saturation, and Value (lightness or brightness). The hue represents the chromatic component in this model. Saturation refers to the predominance of a particular hue in a color.

Texture can be a very useful feature For browsing, searching and retrieval of images,Texture descriptor provides measures of the properties such as smoothness, coarseness, and regularity etc. Texture is usually Denoted by the values of energy, entropy, contrast and homogeneity. gray-level co-occurrence matrix (GLCM)[11] examines texture based on the spatial relationship of pixels. Discrete Wavelet Transform (DWT)[3][6][9] are used for analysis of textures recorded With different resolution. Discrete Cosine Transform (DCT) [7][9] is a powerful transform to extract features from face images. Histogram of Oriented Gradient (HOG)[1] descriptor technique counts occurrences of gradient orientation in localized portions of an image - detection window, or region of interest. SIFT[10] has been proven to be the most robust local invariant feature descriptor. It is designed mainly for gray images. However, color provides valuable information in object description and matching tasks.

Shape descriptor is a set of numbers that are used to describe a given shape. Shape features can be obtained by gradient vector flow (GVF), chain code, edge histogram descriptor (EHD), and Zernike moments, contour co-occurance matrix(CCM)[4] and edge co-occurance Matrix (ECM) are some histogram based shape descriptors.

3. CLASSIFICATION METHODS

Image classification analyzes the numerical properties of various image features and organizes data into categories. Classification algorithms typically employ two phases of processing: training and testing. In the initial training phase, characteristic properties of typical image features are isolated and, based on these, a unique description of each classification category, i.e. training class, is created. In the subsequent testing phase, these feature-space partitions are used to classify image features.

3.1. Support vector machine (SVM)

A Support Vector Machine (SVM)[4] is a discriminative classifier formally defined by a separating hyper plane. It is a vector space based machine learning method where the goal is to find a decision boundary between two classes that is maximally far from any point in the training data. Optimization techniques can be used along with SVM classifiers to enhance the retrieval performance.

3.2. Minimum distance classifier

The minimum distance classifier[10] is used to classify unknown image data into classes which minimize the distance between the image data and the class in multi-feature space. The distance is defined as an index of similarity so that the minimum distance is identical to the maximum similarity. Euclidian distance, Normalized Euclidian distance, Mahalanobis distance, Manhattan distance etc are some of the distance measurements that are often used by minimum distance classifiers.

3.3. K-nearest neighbour classifier

k-Nearest Neighbors[4][11] is a classification method that stores all available cases and classifies new cases based on a similarity measure (e.g., distance functions). Euclidian distance, Manhattan distance, minkowski distance are some of the distance measurements used. The input consists of the k closest training examples in the feature space and the output is a class membership. An object is classified by a majority vote of its neighbours, with the object being assigned to the class most common among its k nearest neighbours.

4. OPTIMIZATION

Optimization is a frequently encountered mathematical problem in all engineering disciplines. It means finding the best possible solution. Optimization of image annotation techniques are an important research area in computer science. Optimization can be achieved in different phases of annotation such as feature extraction, feature selection and weighted feature selection. In feature extraction phase optimum features are extracted from an image using an optimization algorithm along with feature extraction techniques. In [4] feature extraction process is done using Particle swam optimization with SVM classifier. In feature selection, optimization algorithms such as particle swam optimization, genetic algorithm, Firefly algorithm etc are employed. In [9] and [10] a PSO-based feature selection algorithm is utilized to search the feature space for the optimal feature subset. Instead of assigning equal weights to different features, weights are appropriately assigned using optimization algorithms such as Genetic algorithm, Particle Swam Optimization etc which gives an optimized feature vector of each image.

5. STATE OF THE ART IN OPTIMIZATION FOR IMAGE ANNOTATION

Sreekumar. K and Ajimi Ameer[1] proposed a method for Efficient Automatic Image Annotation using Optimized weighted Complementary Feature Fusion using Genetic Algorithm. First, SURF, HSV and HOG features are obtained from the training set images and an average feature descriptor representing each class is obtained. Feature vectors are clustered into k clusters using K means clustering. Instead of assigning equal weights to different features, weights are appropriately assigned using Genetic algorithm which gives an optimized feature vector of each image..

Dong Yang and Ping Guo[2] focused on Image modelling with combined optimization techniques for image semantic annotation. In this framework, low-level image features are extracted from sub-blocks of a given image. Affinity propagation algorithm is applied to estimate the image feature distribution. Then a Bayesian classifier is built using Gaussian mixture model for image semantic annotation.

Darsana B and G. Jagajothi[3] proposed a novel approach of, distributed retrieval of images using particle swarm optimization and Hadoop file systems. It deals with problems of semantic gap and delayed response time in content based image retrieval by coalescing automatic relevance feedback, a stochastic algorithm and distribution of image retrieval. Particle swam optimization is used to assign weights to features.

Lei Wang and Latifur Khan[5] proposed a method for Automatic Image Annotation and Retrieval Using Weighted Feature Selection. For a given cluster, relevant features determined based on histogram analysis and assign greater weight to relevant features as compared to less relevant features.An adaptive content based image retrieval (CBIR) approach based on relevance feedback and Firefly algorithm is proposed by T.Kanimozhi and K.Latha[6]. In addition to the color descriptor, wavelet-based texture descriptor is considered to improve the retrieval performance.

C.Ramesh babu durai, V.Duraisamy and C.Vinothkumar[7] focused on Improved Content Based Image Retrieval Using Neural Network Optimization with Genetic Algorithm. Here feature extraction is done using Discrete cosine Transform. Gaussian Fuzzy Feed Forward Neural Network algorithm is used for classification and optimize the momentum and learning rate using Genetic algorithm. The classification accuracy obtained is 96.29%.

S. Bahrami and M. Saniee Abadeh[8] tried to propose an automated annotation based method to solve image annotation problem. Here Genetic algorithm is utilized for feature selection ,Multi-Label KNN algorithm to weight neighbours and to generate a novel weighted matrix.

Rabab M. Ramadan and Rehab F. Abdel – Kader[9] proposed a method for Face Recognition Using Particle Swarm Optimization-Based Selected Features. The algorithm is applied to coefficients extracted by two feature extraction techniques: the discrete cosine transforms (DCT) and the discrete wavelet transform (DWT). The performance of the proposed algorithm is compared to the performance of a GA-based feature selection algorithm and was found to yield better results with less number of selected features.

Bae-Muu Chang, Hung-Hsu Tsai and Wen-Ling Chou presents a content-based image retrieval method using three kinds of visual features and 12 distance measurements, which is optimized by particle swarm optimization (PSO) algorithm. Firstly, three kinds of features such as color, texture, and ape are extracted from the image. To calculate similarities between query image and images in the database, appropriate distance measurements for each kind of features are used.

PSO algorithm is utilized for finding the approximately optimal weights for three similarities with respect to three kinds of features.

TABLE 1 Comparison of different optimization techniques

Author	Publication Year	Feature extraction method	Optimization method	Classification method	Database	Advantages/Disadvantages
SreeKumar.K, Ajimi Ameer	2015	SURF,HOG, HSV	Genetic Algorithm	K-Means Clustering	Corel 1000 dataset	As the number of iterations increases in genetic algorithm, weights obtained becomes much more accurate and yields better result.
Lei Wang, Latifur Khan	2004	Histogram analysis	Weighted feature selection	K-Means Clustering	Duygulu02	Better annotation performance and correspondence accuracy
Darsana B G. Jagajothi	2013	Color histogram bins wavelet texture energy values.	Particle swam optimization	Similarity evaluation	Corel image database, SIMPLicity, web images.	Dynamically modifies the feature space by feeding automatic relevance feedback and clustering relevant images using metaheuristics.
Vinay Kumar Lowanshi, Shweta	2014	RGB histogram CCM/CCH	Particle swam optimization	SVM and kNN	Corel 1000 dataset	Improved accuracy than existing approach
T.Kanimozhi, K.Latha	2013	Color histogram, Color moment, edge direction histogram, wavelet texture feature	Firefly algorithm	Euclidian distance	Corel database	Highly efficient, robust and highly rapid for image accuracy based application.
C.Ramesh babu durai, V.Duraisamy , C.Vinothkumar	2012	Discrete Cosine Transform	Genetic algorithm	Fuzzy Feed Forward Neural Network	MRI medical image	Efficient, robust
Rabab M. Ramadan , Rehab F. Abdel – Kader	2009	Discrete cosine transforms, Discrete wavelet transform	Particle swam optimization	Euclidean distance	Face images	Generating excellent recognition accuracy with the minimal set of selected features.

6. CONCLUSION

This paper, attempts to provide a comprehensive survey on the latest developments of Automatic Image Annotation techniques with a special emphasis on Optimization. Optimization can be achieved in different phases of annotation such as feature extraction, feature selection and weighted feature selection. From this survey a method for efficient Automatic image annotation is proposed. An image has several dominant characteristics like color texture, shape etc. These different descriptors for the images can form a combined feature vector. However, in order to have optimum performance particle swam optimization algorithm (PSO) based feature selection may be used.

REFERENCES

[1] Ajimi Ameer and SreeKumar.K, Efficient Automatic Image Annotation using Optimized weighted Complementary Feature Fusion using Genetic Algorithm, Second International Symposium on Computer Vision and the Internet (VisionNet'15)

[2] Dong Yang and Ping Guo, Image modeling with combined optimization techniques for image semantic annotation, Springer-Verlag London Limited 2010

[3] Darsana B and G. Jagajothi, Distributed Retrieval of Images using Particle Swarm Optimization and Hadoop, *international Journal of Computer Applications (0975 – 8887) Volume 71– No.8, May 2013*

[4] Vinay Kumar Lowanshi, Shweta Shrivastava and Vineet Richhariya, An Efficient Approach for Content based Image Retrieval using SVM, KNN-GA as Multilayer Classifier, [5] Lei Wang and Latifur Khan, Automatic Image Annotation and Retrieval Using Weighted Feature Selection , Multimedia Software Engineering, 2004. Proceedings. IEEE Sixth International Symposium

[6] T.Kanimozhi and K.Latha, A Meta-Heuristic Optimization Approach for Content Based Image Retrieval using Relevance Feedback Method, Proceedings of the World Congress on Engineering 2013 Vol II, WCE 2013, July 3 - 5, 2013, London, U.K. [7] C.Ramesh babu durai, V.Duraisamy and C.Vinothkumar, Improved Content Based Image

Retrieval Using Neural Network Optimization with Genetic Algorithm, International Journal of Emerging Technology and Advanced Engineering (ISSN 2250-2459, Volume 2, Issue 7, July 2012)

[8] S. Bahrami and M. Saniee Abadeh, Automatic image annotation using an evolutionary algorithm, Telecommunications (IST), 2014 7th International Symposium

[9] Rabab M. Ramadan and Rehab F. Abdel – Kader, Face Recognition Using Particle Swarm Optimization-Based Selected Features, International Journal of Signal Processing, Image Processing and Pattern Recognition Vol. 2, No. 2, June 2009
[10] A. Finkelstein C.E. Jacobs and D.H. Salesin. Fast multiresolution image querying. In Computer graphics proceeding of SIGGRAPH, pages 278–280, Los Angeles, 1995.

[11] K.-C. Liang and C.-C. Jay Kuo. WaveGuide: A joint wavelet image description and representation system. 1998. to appear.

[12] J.P. Eakins. Automatic image content retrieval - are we getting anywhere? In Proc. of Third International Conference on Electronic Library and Visual Information Research, pages 123–135, May 1996.

[13] T. Minka. An image database browser that learns from user interaction. Master's thesis, MIT Media Laboratory, 1995.

[14] S. I. Gallant and M. F. Johnston. Image retrieval using image context vectors: first results. In Storage and Retrieval for Image and Video Databases III, volume 2420, pages 82–94, 1995.

[15] Z. Pecenovic. Intelligent image retrieval using Latent Semantic Indexing. Master's thesis, Swiss Federal Institute of Technology, Lausanne, Vaud, April 1997.

[16] D. McG. Squire and T. Pun. Assessing agreement between human and machine clusterings of image databases. Pattern Recognition, accepted, to be published 1998.

[17] M. Richeldi and P. L. Lanzi. ADHOC: A tool for performing effective feature selection. In Proceedings of the International Conference on Tools with Artificial Intelligence, pages 102–105, 1996.

[18] M. Flickner et al. Query by image and video content: The QBIC system. Computer, pages 23–32, September 1995.

[19] R.W. Piccard A. Pentland and S. Sclaroff. Photobook: Content-based manipulation of image databases. International Journal of Computer Vision, 18(3):233–254, 1996.

[20] J.R. Bach et al. The Virage image search engine: An open framework for image management. In Storage and Retrieval for Image and Video Databases III, volume 2420 of SPIE, pages 76–87, 1995.

[21] V. E. Ogle and M. Stonebraker. Chabot: Retrieval from a relational database of images. Computer, pages 40–48, September 1995.

[22] K. Hirata and T. Kato. Query by visual example. In EDBT'92, pages 56–71, 1992.

[23] M. Egenhofer. Spatial-query-by-sketch. In IEEE Symposium on Visual Languages, pages 60–67, 1996.