International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

Reducing Semantic Gap in Image Retrieval by Integrating High Level Query and Low Level Facial Features

Anusree B S

M.Tech Student, Department of Computer Science and Engineering, Sarabhai Institute of Science and Technology
Vellanad, Trivandrum, India

Abstract—As a result of rapid increase in the digital images, annotation of human faces in images and retrieval of images is one of the desirable needs in the current world. For face annotation, face detection and recognition are two important tasks to be performed and is still challenging due to the wide variety of faces and the complexity of noises and image backgrounds. Automatic face annotation facilitates improved retrieval and organization of digital images. In the proposed work annotation of human faces is done using PCA and MLP. Two scenarios are for naming people in an image finding all faces, and assigning names to all faces. For naming, free text type of annotation is used. This makes annotation task easier, but more difficult to use the annotation later for image retrieval. Efficient image retrieval of an annotated face can be done using the combination of name label and Euclidean distance measurement mechanism.

Keywords: Annotation based image retrieval, CBIR,PCA,MLP,Rprob

1. Introduction

Due to the increase in digital images, the need to efficiently store and retrieve these images from a large collection of image databases arises. Many image retrieval systems have been developed to browse, search and retrieve images from large databases. Currently image retrieval has two approaches: content-based image retrieval (CBIR) and annotation based image retrieval (ABIR).Mainly, they differ in the way the query is formulated. CBIR systems search images using low level features such as color, texture, shape, spatial layout etc. which can be automatically extracted and used to index images. But, without the explicit knowledge on what users want to search, CBIR systems have limited success. Humans tend to associate images with keywords rather than query image. The requirement of CBIR systems is to provide a similar query image to the retrieval system. The CBIR systems fail to meet user expectations, since these systems are unable to index images according to the high level features as perceived by the user. Main challenge in Content Based Image Retrieval is the two gaps namely semantic gap and sensory gap [1].

A good way to reduce the semantic gap is image annotation and is the task of associating text to the semantic content of images. Image annotation can be used as an intermediate step to image retrieval process. It enables users to retrieve images by text queries and can provide semantically better results than CBIR.

The process of assigning a set of keywords to an image is called as image annotation. Image annotation systems attempt to reduce the semantic gap. The task of automatically assigning semantic labels to images is automatic image annotation or linguistic indexing [9]. The main idea of automatic image annotation is to automatically learn semantic descriptors from a collection of image samples, and use it to label new images. If the images are annotated with semantic labels, images can be retrieved using those semantic labels or keywords. Human beings perceive facial images and comparison is by measuring the similarities using high-level features. These high-level semantic concepts cannot be compared directly with low-level features. Generally facial images are different from other images, because facial images are complex, multidimensional, and similar in overall configuration.

Annotation Based Image Retrieval (ABIR) system retrieves images using annotations or labels in images and is an attempt to incorporate more efficient semantic content into both text based queries and image captions. Consequently, textual information plays an important role in image retrieval. However, CBIR has been researched far more than ABIR [2].

The proposed work is a facial image retrieval system for finding out all similar facial images from the pool of digital images by using face recognition technique and to retrieval all those facial images by integrating the semantic label of the facial image and the similarity between query facial image and recognized facial images. The main aim of the system is to reduce the semantic gap between high level query requirement and low level facial features of the human face image.

2. Related Work

In [3] free text annotation, keyword annotation and annotation based on ontologies are reviewed. Different keyword vocabularies and annotation standards can be used for individual databases, and user need not necessarily know the exact vocabulary.

In [4] intensity based Haar like features are extracted and is integrated with supervised manifold learning method. Manifold learning method is employed for constructing an efficient face manifold. Neighborhood graphs are constructed. To classify the new input images nearest neighbor method is used.

A relevance feedback system [5] for retrieving a face images from a large image database based on the face in user's mind. A series of queries and answers are made between a user and the system. System displays a set of images from the pool of images and user can provide feedback to the system. The purpose is to retrieve the user target image in his mind from the image database. Disadvantage of this method is the difference between mental matching and feature-based matching.

In [6] Latent semantic analysis is used for the purpose of image annotation and retrieval. In this work for annotation labels are used and annotation is treated as unsupervised learning process. For retrieval labels are used retrieval is treated as supervised learning process. Retrieved images were ranked based on its likelihood values.

In [7] DaidiZhong and IrekDefee propose a unified framework for pattern retrieval that uses structural and statistical information of pattern information. The method is purely based on local feature sets and the framework is not limited to the face regions

3. Proposed Work

For face annotation, face detection and recognition are two important tasks to be performed. Automatic face annotation facilitates improved retrieval and organization of digital images. In the proposed work annotation of human faces is done using PCA and MLP. Also the retrieval of similar images of a query image face can be obtained by employing Euclidean distance measurement mechanism and a supervised label refinement can also be employed.

A. Image Preprocessing

A supervised training system will be supplied with digital images and database for further processing will be prepared by training the system with those image samples. For training, face regions in the image sample can be identified and detected using Viola-Jones face detector algorithm. The detected face regions are extracted and for adjusting the variation in light, gray scale conversion is carried out. Histogram equalization can be performed to adjust the contrast. Through this adjustment, the intensities can be better distributed on the histogram.

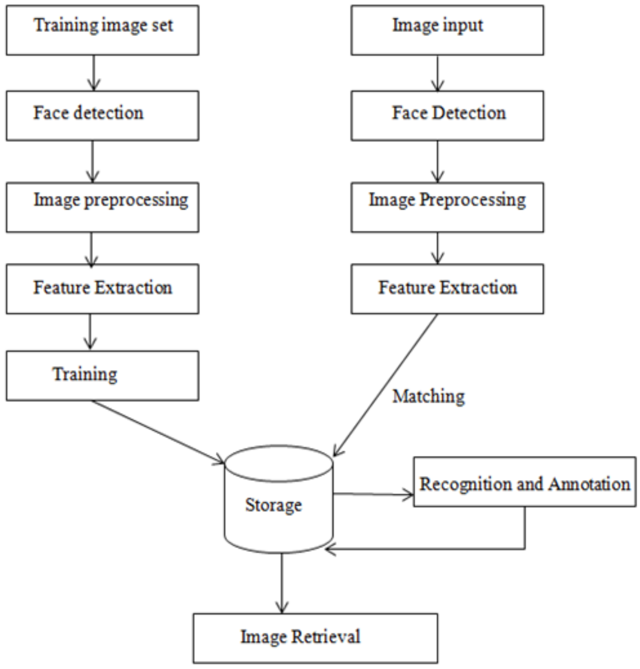

Fig: Proposed System Architecture

B. Performing PCA and MLPs

Principal Component Analysis can be performed on images to reduce dimensionality and to extract feature vector. Feature vector of an image describes image feature. This feature vector can be fed into an artificial neural network for training of images. A resulting weighted value obtained will be unique for the images and a label is given to each value which can be used further for annotation purpose.

- Principal Component Analysis

  The PCA is performed by computing the eigenvectors and eigenvalues of the covariance matrix. The covariance is determined by the tendency to two random variables that vary together. Covariance, cov(X, Y) = E[E[X] - X] . E[E[Y] - Y] where E[X] and E[Y] denote the expected value of X and Y respectively.

  Algorithm

  STEP 1: Prepare the Data

  $S = \{\Gamma_1, \Gamma_2, \Gamma_3, \ldots \ldots \Gamma_M\}$

  STEP 2: Obtain the Mean, ψ as

  $\psi = \frac{1}{M}\sum_{n=1}^{M}\Gamma_n$

  STEP 3: Mean is subtracted from original image as

  $\phi_i = \Gamma_i - \psi$

  STEP 4: Calculate the Covariance Matrix, C

  $C = \frac{1}{M}\sum_{n=1}^{M}\phi_n\phi_n^T$

  $= AA^T$

  $A = \{\phi_{1,}\phi_{2,}\phi_3, \ldots, \phi_n\}$

  STEP 5: Eigenvectors and Eigenvalues of the Covariance Matrix are calculated and principal components are selected.

  In this step, the eigenvectors and the corresponding eigenvalues will be calculated. From M eigenvectors only M' will be chosen, which have the highest eigenvalues.

- MultiLayer Perceptron

  A multilayer perceptron (MLP) is a feed forward artificial neural network model that maps sets of input data on to a set of appropriate outputs. MLP consist multiple layers of nodes as a complete directed graph. A supervised learning technique called backpropagation is performed for training the network. The training algorithm used is resilient backpropagation (Rprob) [8].

- Resilient backpropagation algorithm
  Rprob algorithm [8] performs a direct adaptation of the weight step based on local gradient information and the effort of adaptation is not blurred by gradient behavior. In Rprob the number of learning steps is significantly reduced when compared to the original gradient-descent procedure as well as other adaptive procedures. for each weight its individual update-value Δ_{ij}, which solely determines the size of the weight-update. This adaptive update value which is evolved during the learning process is based on its local sight on the error function **E,** according to the learning-rule as,

$$\Delta_{ij}^{(t)} = \begin{cases} \eta^{+} * \Delta_{ij}^{(t-1)}, if \frac{\partial E^{(t-1)}}{\partial w_{ij}} * \frac{\partial E^{(t)}}{\partial w_{ij}} > 0 \\ \eta^{-} * \Delta_{ij}^{(t-1)}, if \frac{\partial E^{(t-1)}}{\partial w_{ij}} * \frac{\partial E^{(t)}}{\partial w_{ij}} < 0 \\ \Delta_{ij}^{(t-1)}, else \end{cases}$$

where $0 < \eta^{-} < 1 < \eta^{+}$

Once the update-value for each weight is adapted, weight update will be as

$$\Delta w_{ij}^{(t)} = \begin{cases} -\Delta_{ij}^{(t)}, if \frac{\partial E^{(t)}}{\partial w_{ij}} > 0 \\ +\Delta_{ij}^{(t)}, if \frac{\partial E^{(t)}}{\partial w_{ij}} < 0 \\ 0, else \end{cases}$$

$$w_{ij}^{(t+1)} = w_{ij}^{(t)} + \Delta w_{ij}^{(t)}$$

However, there is one exception: If the partial derivative changes sign, i.e. the previous step was too large and the minimum was missed, the previous weight-update is reverted.

$$\Delta w_{ij}^{(t)} = -\Delta w_{ij}^{(t-1)}, if \frac{\partial E^{(t-1)}}{\partial w_{ij}} * \frac{\partial E^{(t)}}{\partial w_{ij}} < 0$$

C. Face Annotation

Once images are trained and labeled with the neural network system, an image query given to the system can be annotated. The image query can contain faces and those faces can be detected and preprocessed. The feature vector of each face in the query image is computed and is compared with the feature vector of training samples stored in the database. Each feature vector is indexed by a label and these labels are used for the annotating images associated to that feature vectors.

D. Similar Image Retrieval

Once all the detected faces in the query images were annoted similar images of a particular detected face can be retrieved. Filtering of the images can be done with the aid of semantic information associated with the query face image i.e. the labels associated with the query face image is used for finding out feature vectors of all the faces that has the same label. For similar face areas the system will extract the vector feature of the query face image and will use Euclidian Distance measurement methodology to compute the distance between this query features vector Q and the resultant features vectors D.

FaceSimilarityDistance (Q,D)=$\sqrt{\sum_{i=1}^{n}(Qi-Dj)^2}$

The most identical feature vectors will be retrieved and display on the top. The actual images that points to the feature vectors will contain the query face and will be the result.

4. Conclusion

Face annotation is formulated as an extended face recognition problem in which faces are trained from a collection of well-controlled labeled facial images. In this work face recognition and face annotation is performed using PCA and MLP. Usually query for image retrieval will be either image or text. In this work image retrieval is implemented by combining bothtext query and image. First retrieval of feature vectors having same label is retrieved and later comparison with query face is done using Euclidean distance measurement. The images that correspond to the similar feature vector that shows a match will be the result of image retrieval.

5. Acknowledgment

This work was supported in part by the Department of Computer Science & Engineering, SIST, Trivandrum. I would like to show my gratitude to Dr. C G Sukumaran Nair (HOD), Associate Professors, Sudha SK and Resmy V.R, for their valuable guidance.

References

[1] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain.Content-based image retrieval at the end of the early years.IEEE Trans. Pattern Anal. Mach. Intell., 22(12):1349–1380, 2000.
[2] M. Inoue. On the need for annotation-based image retrieval.Workshop on Information Retrieval in Context, 2004.
[3] Allan Hanbury," A survey of methods for image annotation", http://www.sciencedirect.com/science/article/pii/S1045926X08000037
[4] Bau-Cheng, S., C. Chu-Song, and H. Hui-Huang,Face Image Retrieval by Using Haar Features, in19th International Conference on Pattern Recognition (ICPR 2008) 2008: Tampa, Florida. p. 1-4.
[5] Fang, Y. and D. Geman, Experiments in Mental Face Retrieval, in Audio- and Video-Based Biometric Person Authentication, T. Kanade, A. Jain, and N. Ratha, Editors. 2005, Springer Berlin / Heidelberg.p.637-646.
[6] T.-T. Pham, N.E. Maillot, J.-H.Lim, and J.-P. Chevallet, "Latent Semantic Fusion Model for Image Retrieval and Annotation," Proc.16th ACM Conf. Information and Knowledge Management(CIKM),2007.
[7] Daidi, Z. and D. Irek, Face Retrieval Based on RobustLocal Features and Statistical-Structural LearningApproach. EURASIP Journal on Advances in SignalProcessing, 2008.2008.

[8] Martin Riedmiller and Heinrich Braun, A Direct Adaptive Method for Faster Backpropagation Learning:The RPROP Algorithm,Proc IEEE Int Conf Neural Networks, 1993
[9] R. Datta, D. Joshi, J. Li, J.Z. Wang, Image retrieval: ideas,influences and trends of the new age, ACM Computing Surveys 40(2) (April 2008).
[10] R. Zhao and W. Grosky. Narrowing the semantic gap - improved text-based web document retrieval using visual features, IEEE Trans. on Multimedia, 4(2):189–200, 2002.
[11] HimaliChaudhari and D.D.Patil, A Survey on Automatic Annotation and Annotation Based Image Retrieval, International Journal of Computer Science and Information Technologies, Vol. 5 (2) , 2014, 1368-1371