ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 1.852

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

Interactive Image search and Retrieval using Artificial Neural Network

S.Dilli Babu [*1], Gobiga Jegatheesan [2], Mrs.S.Anusuya [3]
Undergraduate Student [*1], Undergraduate Student [2], Associate Professor [3]
Department of Computer Science Engineering, Saveetha School of Engineering, Saveetha University
Chennai – 600089, India
dillibabu101093@yahoo.com

Abstract

Human Computer Interaction is an immense field of computer studies to improve human interactions with machine with the influence of knowledge drawn from human and computers. An interactive method to search and retrieve images from databases through voice recognition system is proposed in this paper. Initially, the images are stored in the databases with voice samples and ensured that each image is associated with its own samples. These samples could of any keywords describing features of the image. During retrieval, the input sample is checked against with the samples stored in the database and corresponding image is displayed for view. ANN is applied for voice recognition to train the samples and to classify it. The experimental section of this paper shows the image retrieval using voice recognition in simplest manner with good accuracy.

Keywords: Image Search, Image Retrieval, Voice recognition, ANN, BackPropagation.

Introduction

A class of chemical compounds called Development of interactive software applications is a state of the art today and it occupies major portions of IT business. The study says general attitude of computer users turned towards interactive solutions and success of interactive applications depend user experiences with it. The revolution of human computer interactions (HCI) made computers more usable even to physically challenged people. Particularly, speech based interfaces is becoming rapidly maturing part of HCI to communicate with the machine through voice recognition system. The information retrieval generally can be classified as text based, content based and interactive based retrieval. Many approaches such as content based image retrieval (CBIR), Query By Image Content (QBIC), visual seek and virage have been proposed for image retrieval particularly for face and finger print recognition. But they are time consuming methods and more user intervention is needed to locate desired image. Most of the methods of image retrieval are keywords and content based image retrieval. The popular search engines such as Google, Imagenet, and Flicker follow keyword based retrieval system. The main purpose of image processing concerns with the tasks such as visualization, Image sharpening and restoration, Image retrieval and Image Recognition. The manual and text based image search and retrieval is not sufficient to meet the challenges since the scale of digital images increasing exponentially. Interactive image retrieval system is rapidly evolving approach and is widely used to retrieve images from databases of internet.

Related work

The process of interactive image retrieval proposed in [1] using JIGSAW algorithm where images are segmented and query is generated for the segmented image. The search process checks the query; accordingly images are retrieved in fast and accurate manner. Retrieval based on image description for mobile apps has been proposed in [2] [3] [4] [6]. The images are segmented into multiple regions and keywords describing object features such as, size of the objects, color of the objects and resolution of the image are assigned to these regions and are numbered. The image search is performed based on the numeric value assigned to these objects. Content based Image retrieval [4] retrieve images from the data bases based on visual content defined by a set of low level features of an image. But the interpretation of images would have different perception among people. The same images would be

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 1.852

annotated differently by different people. Thus the semantic gap is a big challenge to CBIR system. Locality-sensitive hashing (LSH) and support vector machine (SVM) have been introduced [7] to reduce the limitations of CBIR. The main disadvantage of semantic gap is reduced by SVM feedback method while LSH introduced for indexing. Interactive image query system for content-based image retrieval [8] divides an image into multiple regions and desired region is selected based on input query. The texture features are obtained with wavelet transform and target region is matched against region defined in query. Interactive image recognition for human-machine interfacing [9] developed to recognize the images using image metamorphosis and intelligent system. Leaning and recognition of images is implemented with neural network. Images are converted to feature vectors and compared with training set. To improve classification accuracy of image retrieval system Back propagation algorithm has been applied to remote sensing images [10]. Probabilistic neural network is used to classify vowel signals of different individuals [11]. A framework combining Principal component analysis and back propagation is discussed to classify emotional information [12]. In this paper we propose a simple and an interactive method to retrieve images interactively and efficiently than traditional system.

Interactive image search and retrieval using Artificial Neural Network

Artificial neural network is a biological neural network composed of interconnected group of neurons, performs computations and process information based on human brain system. Many types of neural networks such as feed forward neural networks, radial basis function (RBF) networks, Kohonen self-organizing networks, recurrent networks, stochastic neural networks, modular neural networks, dynamic neural networks, cascading neural networks and neuro-fuzzy networks have been applied for classification and categorization, prediction and pattern recognition.

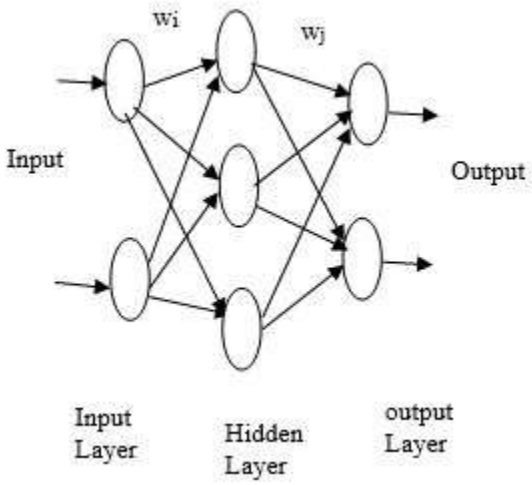

Figure 1. A Multi layer back propagation network

Most of the tasks of image processing such as pattern recognition, feature extraction, segmentation of objects are solved using neural networks. Back Propagation neural network is a kind of layered feed-forward network structure; forward nature of network propagates activation to produce an output and backward propagates error to determine weight changes. The back propagation algorithm is used to learn the weights of multi-layer neural network with a fixed architecture and it performs gradient descent to try to minimize the sum squared error between networks output values and the given target values. In this work we employ back propagation algorithm, a well-known supervised training algorithm to train digitized samples. A multilayer back propagation network is shown in Figure 1. A set of inputs are fed into the network, multiplied with weights and summed up. The modification function is applied to the result to find the output. The error is calculated by comparing actual output with desired output and propagated to the network; accordingly weights are adjusted to reduce the error.

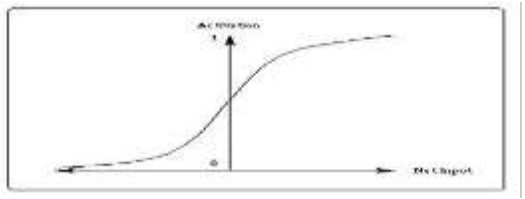

Figure 2. Sigmoid function

The back propagation learning algorithm uses Sigmoid activation function shown in Figure 2. The

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 1.852

activation function is defined as in Eq. (1) and it from 0 to1.

$$f(x) = \frac{1}{1+e^{-v}} \quad (1)$$

v is the weighted neurons of input layer. The
The total error in a network is calculated as

$$E = \frac{1}{2}\sum_{k}(t_k - a_k)^2 \quad (2)$$

Where t_k is target output and a_k is the actual output. So the error E depends on the output and it is calculated as

$$a_k = \sum_{k=1}^{n} w_k x_k \quad (3)$$

Where n is the number of input units to the neuron, wk is the k^{th} weight and x_k is the k^{th} input to the neuron.

The derivative of the error with respect to activation is written as

$$\frac{\partial E}{\partial a_k} = \frac{\partial(1/2(t_k - a_k)^2}{\partial a_k} = -(t_k - a_k) \quad (4)$$

The weight is updated using gradient descent as

$$\Delta w_k = \alpha(t_k - a_k)x_k \quad (5)$$

Figures 3.a and 3.b depict the overall process of proposed work. The relative paths of the images to be stored are stored along with digitized voice samples. These samples in turn are learned and recognized using back propagation algorithm. In the process of image retrieval, the analog input is converted into digitized form and then checked against in the databases. If matches, the associated image has been retrieved once image path has been read.

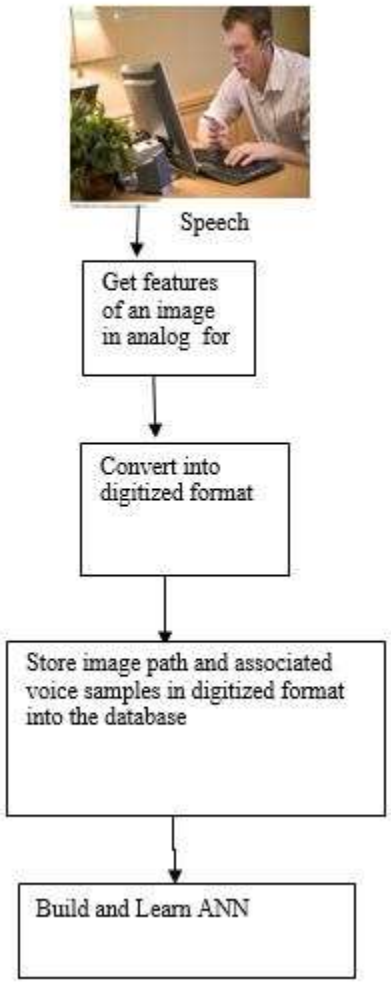

Figure 3.a Part 1 of proposed process

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 1.852

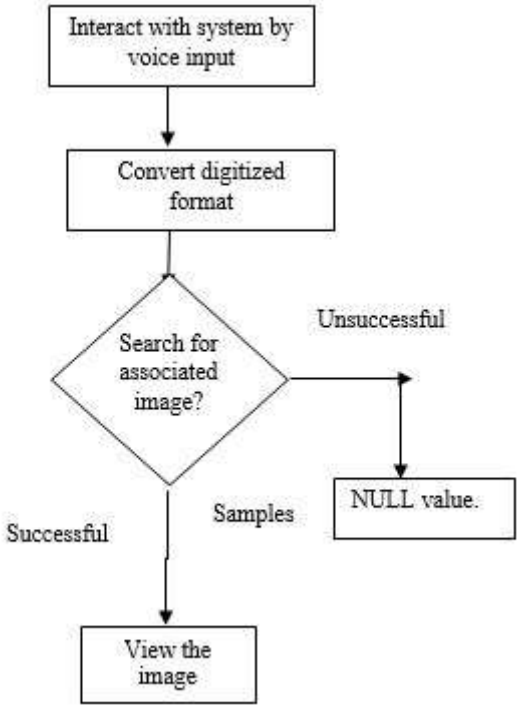

Figure 3.b Part 2 of proposed process

In this experiment, the target data set is generated with the audio signals of 50 speakers and each image is accompanied with four voice samples describing features of an image. ANN is designed with three hidden layers and each layer consists of 10 neurons.

Figure 4 shows the snapshot of sample records present in the database where each record contains the path of the image and its associated digitized samples. The experiment is also demonstrated on the images with same theme. Figure 5 shows some images retrieved with two common words and with two different words pertaining to different images.

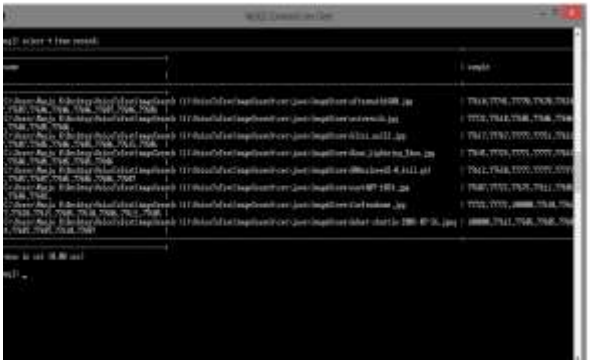

Figure 4.Sample records stored in database

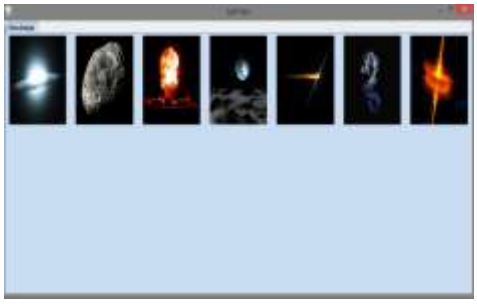

Figure 5 Retrieval of images with same theme

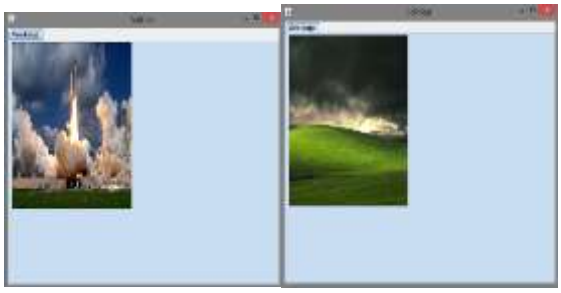

Figure 6 Images retrieved with proposed approach

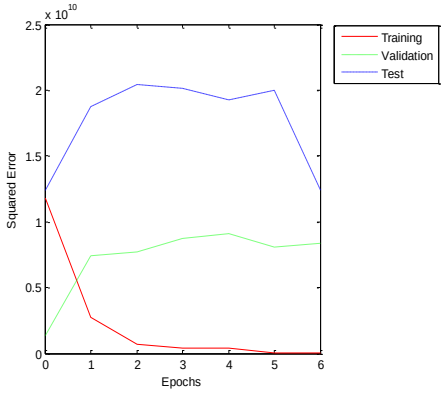

Figure 7. Squared Error Vs Epochs
Figure 7 shows the activation function converges at 6th epochs

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 1.852

Conclusion

The purpose of this work is to retrieve images from the databases through speech recognition system. The voice samples are given as input to the searching mechanism and checked against with samples stored in the database. If it matches the associated image is retrieved otherwise the system produce an error message. The proposed approach would be useful for the disabled people to view images without the use of mouse and keyboard. The existing standard methods are not reliable for all interactive environments and it takes much time to retrieve the image. The accuracy of proposed work is good compared to other text based methods. In future, we extend this work to support different kinds of images.

Table 1. Sample inputs and recognition rate

User audio input	Number of Relevant images	Number of images retrieved	Recognition rate
"Planets","night", "black"	10	7	70%
"greeneries","cloudy","raining"	10	8	80%
"rocket", "space", "launching"	10	8	80%

References

1. *Yang Wang , Tao Mei , Jingdong Wang , Houqiang Li , Shipeng Li,JIGSAW: Interactive Mobile Visual Search with Multimodal Queries. University of Science and Technology of China, Hefei 230027, P. R. China Microsoft Research Asia, Beijing 100080.*
2. *Y. Cao, H. Wang, C. Wang, Z. Li, L. Zhang, and L. Zhang, "MindFinder: interactive sketch-based image search on millions of images," in Proc. of ACM Multimedia, 2010, pp. 1605–1608.*
3. *V. Chandrasekhar, D. M. Chen, A. Lin, G. Takacs, S. S. Tsai, N. M. Cheung, Y. Reznik, R. Grzeszczuk, and B. Girod, "Comparison of local feature descriptors for mobile visual search," in Proc. IEEE International Conf. Image Processing, 2010, pp. 3885–3888*
4. *Ig-Jae Kim ; Je-Ho Lee ; Yong-Moo Kwon ; Sang-Hui Park , "Content-based image retrieval method using color and shape features" Information, Communications and Signal Processing, 1997. ICICS., Proceedings of 1997 International Conference on Digital Object Identifier: 10.1109/ ICICS.1997.652119 Publication Year: 1997 , Page(s): 948 - 952 vol.2*
5. *N. Ben-Haim, B. Babendo, and S. Belongie, "Improving web-based image search via content based clustering," in Proc. of IEEE Conf. Computer Vision and Pattern Recognition Workshop, 2006, p. 106.*
6. *Houqiang Li, Member, Yang Wang, Tao Mei *, Senior Member, Jingdong Wang, Member, and Shipeng Li, Fellow, "Interactive Multimodal Visual Search on Mobile Device". 2011.*
7. *Tingting Dong , Zhicheng Zhao, A novel interactive image retrieval method based on LSH AND SVM, Proceedings of IC-NIDC,2012.*
8. *Fumihiho Kawanobe, Shigeru Takano, and Yoshihiro Okada, Towards interactive image query system for content-based image retrieval, WISE '00 Proceedings of the First International Conference on Web Information Systems Engineering (WISE'00)-Volume 1 - Volume 1 Page 419,2000*
9. *Md. Tajmilur Rahman, M. A. Jobayer Bin Bakkre, Al-Amin *Ahsan Raja Chowdhury, Md. Al-Amin Bhuiyan, On Interactive Image Recognition for Human-Machine Interfacing, International Conference on Convergence Information Technology, 2007*
10. *Jiefeng Jiang ; Jing Zhang ; Gege Yang ; Dapeng Zhang ; Lianjun Zhang, Application of back propagation neural network in the classification of high resolution remote sensing image: Take remote sensing image of beijing for instance , 18th International Conference on Geoinformatics, 2010*
11. *Chee Peng Lim , Siew Chan Woo, Aun Sim Loh , Rohaizan Osman, Speech Recognition Using Artificial Neural Networks, Proceeding Fourth International Workshop on Semantic Media Adaptation and Personalization,2009*
12. *Sheguo Wang,Xuxiong Ling,Fuliang Zhang,Jianing Tong,Speech Emotion Recognition Based on Principal Component Analysis and Back Propagation Neural Network, International Conference on Measuring Technology and Mechatronics Automation 2010.*