International Journal of Engineering Research and Development
***e**-ISSN: 2278-067X, **p**-ISSN: 2278-800X, www.ijerd.com*
Volume 3, Issue 10 (September 2012), PP. 08-14

Content Based Image Retrieval through Fundamental and Visual Feature Description Techniques

Venkata Naga Srinivas Palaparthi[1], Chittibabu Dondapati[2], Subhakar Rao Golla[3], J.N.S.S Janardhana Naidu[4]
[1]*Assoc Professor, HOD in Dept of I.T, Swarnandhra College of Engineering & Technology*
[2,3,4]*Asst Professor in Dept of I.T, Swarnandhra College of Engineering & Technology*

Abstract—Content-Based Image Retrieval (CBIR) uses the visual contents of an image such as color, shape, texture and special layout to represent and index the image. The feature vectors of images in the database form a feature database. To retrieve images, users provide the retrieval system with example images or sketched figures. The system then changed these examples into its internal representation of feature vectors. The similarity between the characteristics vectors of the query example or sketch and those of the images in the database are then calculated and retrieval is performed with the assist of an indexing scheme. The indexing proposal provides an efficient way to search for the image database. Recent retrieval systems have included users Relevance Feedback (RF) to modify the retrieval process in order to generate perceptually and semantically more meaningful retrieval results.

Keywords—Content Based Image Retrieval, Artificial Neural Network, Relevance Feedback, Visual Features.

I. INTRODUCTION

The explosive growth of digital image collections on the web sites is calling for an efficient and intelligent method of browsing, searching and retrieving images. However, multimedia database contains huge number of images in the web. Due to the complexity of multimedia contents, image sympathetic is difficult but interesting issue in this field. Content-based image retrieval [3], a technique which uses visual contents to search images from large scale image databases according to user's interest. Earlier approaches to the content –based multimedia retrieval do not adapt the query and retrieval model based on the user's perception of the visual resemblance [1]. CBIR is one of the current image retrieval systems. The relevance feedback from the image retrieval domain is used in the content based image retrieval [8]. In this project, an artificial neural network (ANN)-based approach is proposed to explore a promising solution to the Web image retrieval (IR). Compared with other image retrieval methods, this new approach uses the Content-Based features [5] to improve retrieval performance.

II. EXISTING SY STEM

Early techniques were not generally based on visual features but the textual annotation of images. Text-based image retrieval uses traditional database techniques to manage images. Through text description, images can be organized by topical or semantic hierarchies to facilitate easy navigation and browsing based on standard Boolean queries.
However, since automatically generating descriptive texts for a wide spectrum of images is not feasible, most text-based image retrieval systems require only manual annotation of images. Obviously, annotating images manually is a cumbersome and expensive task for large image databases, and is often subjective, context-sensitive and incomplete. As a result, it is difficult for the traditional text-based methods to support a variety of task-dependent queries.

III. STUDIES ON PROPOSED SYSTEM

3.1. Image Content Descriptors

Generally speaking, image content may include both visual and semantic content. Visual content can be very general or domain specific. **General visual content** include color, texture, shape, spatial relationship, etc. Domain specific visual content, like human faces, is application dependent and may involve domain knowledge. Semantic content is obtained either by textual annotation or by complex inference procedures based on visual content. Our proposed system is based on General visual content which includes

- **a.** Color
- **b.** Texture
- **c.** Shape

3.1.1. Color

3.1.1.1. Color Histogram

The color histogram is a method for describing the color content of an image, it counts the number of occurrences of each color in an image. The color histogram of an image is rotation, translation, and scale-invariant; therefore, it is very suitable for color-based CBIR: content-based image retrieval using solely global color features of images. However, the

main drawback of using the color histogram for CBIR is that it only uses color information, texture and shape-properties are not taken into account.
This may lead to unexpected errors; for example, a CBIR engine using the color histogram as feature is not able to distinguish between a red cup, a red plate, a red flower, and a red car.

Figure 3.1.1.1: Two distinct images are shown. However, when represented by their color histograms, they

Are judged as identical.
Two distinct images are shown. However, when represented by their color histograms, they are judged as identical. [Color tuple histograms, color coherent vectors, color correlograms, local color regions, and blobs. These methods are concerned with optimizing color matching techniques on a spatial level, but disregard the basic issue of intuitive color coding.
In other words, the way the engine is processing color, is not related to human color processing. In our opinion, the issue of color coding (or categorization) should be stressed prior to exploring these techniques. Therefore, we will focus on color histogram based image retrieval. [6].

3.1.1.2. Color Quantization

In order to produce color histograms, color quantization has to be applied. Color quantization is the process of reducing the number of colors used to represent an image. A quantization scheme is determined by the color space and the segmentation (i.e., split up) of the color space used. A color space is the representation of color in a three dimensional space.

Figure 3.1.1.2: From top to bottom: The original image using 2563 colors, quantized in 8 bins, and quantized in 64, bins using RGB color space**.**

In applying a standard quantization scheme on a color space, each axis is divided into a number of parts. When axis are divided in k, l, and m parts, the number of colors (n) used to represent an image will be n=k.l.m. A quantization of color space in n colors is often referred to as an n-bins quantization scheme

3.1.1.3. Color space

A color space specifies colors as tuples of (typically three) numbers, according to certain specifications. Colors spaces lend themselves to (in principle) reproducible representations of color, particularly in digital representations, such as digital printing or digital electronic display. The purpose of a color space is to facilitate the specification of colors in some standard, generally accepted way.

One can describe color spaces using the notion: perceptual uniformity. Perceptually uniform that two colors that are equally distant in the color space are perceptually equally distant. Perceptual uniformity is a very important notion when a color space is quantized. When a color space is perceptually uniform, there is less chance that the difference in color value due to the quantization will be noticeable on a display or on a hard copy. In the remainder of this section, several color spaces with their quantization schemes will be described. In addition, the conversion of color images to gray-scale images, using the specific color space, will be described. The quantization of color images transformed into gray-scale images is independent of the color spaces: the gray-scale axis is divided in the number of bins needed for the specific quantization scheme. In this thesis, gray-scale images were quantized in 8, 16, 32, 64, and 128 bins.

3.1.1.4. The RGB color space

The RGB color space is the most used color space for computer graphics. Note that R, G, and B stand here for intensities of the Red, Green, and Blue guns in a CRT, not for primaries as meant in the CIE RGB space. It is an additive color space: red, green, and blue light are combined to create other colors. It is not perceptually uniform. Each color-axis(R, G, and B) is equally important. Therefore, each axis should be quantized with the same precision. So, when the RGB color space is quantized, the number of bins should always be a cube of an integer. In this thesis, 8 (23), 64 (43), 216(63), 512 (83), and 4096 (163) bins are used in quantizing the RGB color space. The conversion from a RGB image to a gray value image simply takes the sum of the R, G, and B values and divides the result by three.

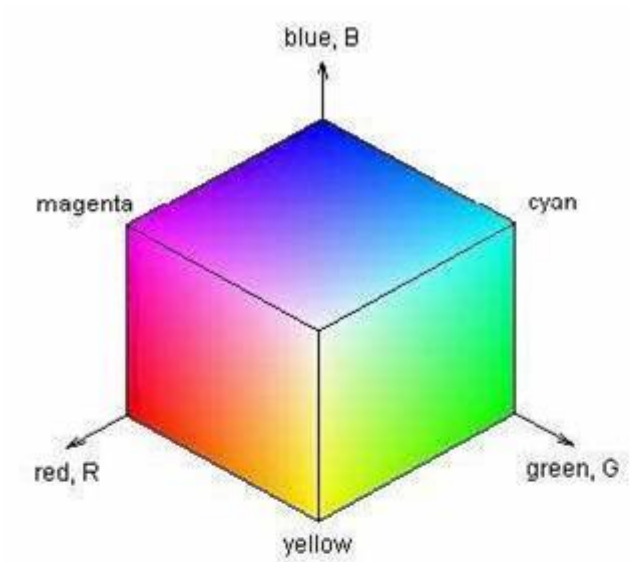


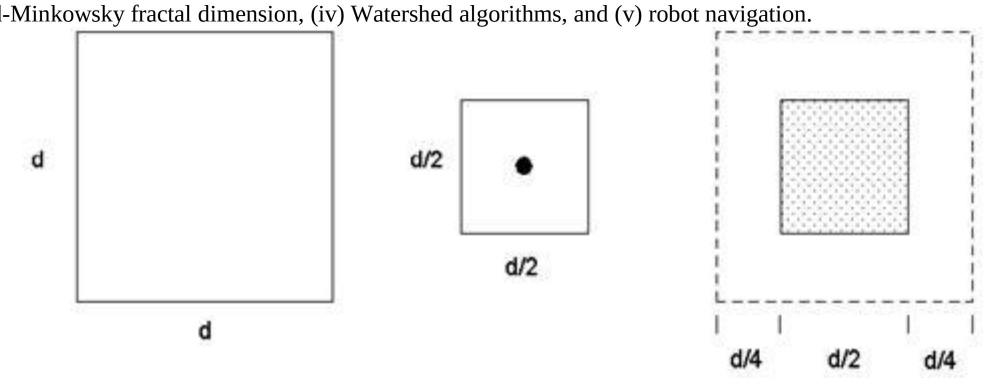
Figure 3.1.1.4. : The RGB color space visualized as a cube

Euclidean Distance transformation (EDT)

Region growing algorithms can be applied to obtain distance transformations. A distance transformation creates an image in which the value of each pixel is its distance to the set of object pixels O in the original image:

D(p) = min { dist(p,q), q € 0}

The Euclidean distance transform (EDT) has been extensively used in computer vision and pattern [2] recognition, either by itself or as an important intermediate or ancillary method in applications ranging from trajectory planning to neuromorphometry. Examples of methods possibly involving the EDT are: (i) skeletonization; (ii) Voronoi tessellations; (iii) Bouligand-Minkowsky fractal dimension, (iv) Watershed algorithms, and (v) robot navigation.



The process of erosion illustrated. The left figure is the original shape A. The square in the middle is the erosion marker B (dot is the center). The middle of the marker runs over the boundary of the result of erosion of A by B (A□B) is given by the solid shape on the right, in which the outer (dotted) square projects the original object A.

```
For k:=1 to R
    For every pixel p in the boundary
        If NOT [p is a vertex ) AND (k modulo 5=0)
            AND (k modulo 410!=0)]
              Grow p as 8-n
                  Otherwise
                    Grow p as 4-n
```

Algorithm for hexadec number f iterations

Several methods for calculation of the EDT have been described in the literature, both for sequential and parallel machines. However, most of these methods do not produce exact distances, but only approximations. Chamfer distance transformation using two raster scans on the image, which produces a coarse approximation of the exact EDT can be used. To get a result that is exact on most points but can produce small errors on some points, four raster scans can be used.

3.2.2. Texture

Texture is another important property of images. Various texture representations have been investigated in pattern recognition and computer vision. Basically, Texture representation methods can be classified into two categories: structural and statistical. Structural methods, including morphological operator and adjacency graph, describe texture by identifying structural primitives and their placement rules. They tend to be most effective when applied to textures that very regular. Statistical methods including Fourier power spectra, co-occurrence matrices, shift-invariant principal component analysis (SPCA), Tamura feature, old decomposition, Markov random field, fractal model, and multi-resolution filtering techniques such as Gabor and wavelet transform, characterize texture by the statistical distribution of the image intensity. The structural techniques deal with the arrangement of image primitives, such as the description of texture based on regularly spaced, parallel lines. The co-occurrence matrix was used to perform texture analysis because it is an important gray-scale texture analysis method.

The co-occurrence matrix:

The co-occurrence matrix is constructed from an image by estimating the pair wise statistics of pixel intensity. In order to (i) provide perceptual intuitive results and (ii) tackle the computational burden, intensity was quantized into an arbitrary number of clusters of intensity values, which we will name: gray values.

The co-occurrence matrix Cd (I, j) counts the co-occurrence of pixels with

Gray values I and j at a given distance d. the distance d is defined in polar coordinates (d,a),with discrete length and orientation. In practice, a takes the values 0º,45º,90º,135º,225º,270º,and 315º. The co-occurrence matrix Cd (I, j) can now be defined as follows:

$C_d(i,j)=pr(I(p_1))=i$^$I(p_2)=j||p_1-p_2|=d)$

Where Pr is probability, and p_1 and p_2 are positions in the grapy-scale image I.

The algorithm yields a symmetric matrix, which has advantage that only angles up to 180º need to be considered. A single co-occurrence matrix can be defined for each distance (d) by averaging four co-occurrences matrices of different angles (i.e., 0º, 45º, 90º, and 135º).

Let N be the number of gray-values in the image, then the dimension of the co-occurrence matrix $C_d(i,j)$ well be N×N. So, the computational complexity of the co-occurrence matrix depends quadratic ally on the number of gray-scales used for quantization.

Because of the high dimensionality of the matrix, the individual elements of the co-occurrence matrix are rarely used directly for texture analysis./Instead, a large number of textural features can be

Derived from the matrix such as: energy, entropy, correlation, inverse difference moment, intertia, Haralick's correlation, cluster shade, and cluster prominence.

3.2.3. Shape:

The shape extraction phase is divided in three stages (i) coarse image segmentation,(ii) pixel wise classification, and(iii) smoothing. The coarse image classification phase, only color information is used because the regions are too small for our texture descriptor to be informative.

There are a plenty of shape descriptors available that can be divided into two main categories: region-based and contour-based methods use only the information present in the contour of an object.

There are some techniques, for example, Fourier transforms and moments, that can be applied using both approaches with only small changes in algorithms.

Using only contour information in shape analysis can be beneficial:

- Information inside the object's contour is low when dealing only with the contour
  (Whether this is an advantage or a disadvantage, depends on the application)
  It takes less space to store different objects (data compression)

- Shape description are faster to calculate because there are less image pixels to process (although the overhead that comes from contour tracking must be included in total computation time)
- Variations in a contour are more easily detected

IV. PROPOSED SYSTEM

Content-based image retrieval uses the visual contents of an image such as color, shape, and texture to represent and index the image. To retrieve images, users provide the retrieval system with example images or sketched figures. The system then changes these examples into its internal representation of feature vectors. The similarities/distances between the feature vectors of the query example or sketch and those of the images in the database are then calculated and retrieval is performed.

V. ROLE OF NEURAL NETWORK

Fig: 5.1 an artificial neural network (ANN) or commonly just neural network (NN) [9] is an interconnected group of artificial neurons that uses a mathematical model or computational model for information processing based on a connectionist approach to computation. In most cases ANN is an adaptive system that changes its structure based on external or internal information that flows through the network.

An artificial neuron is a device with many inputs and one output. The neuron has two modes of operation; the training mode and the using mode. In the training mode, the neuron can be trained to fire(or not),for particular input patterns. In the mode, when a taught input pattern is detected at the input, its associated output becomes the current output. If the input pattern does not belong in the taught list of input patterns, the firing rule is used to determine whether to fire or not.

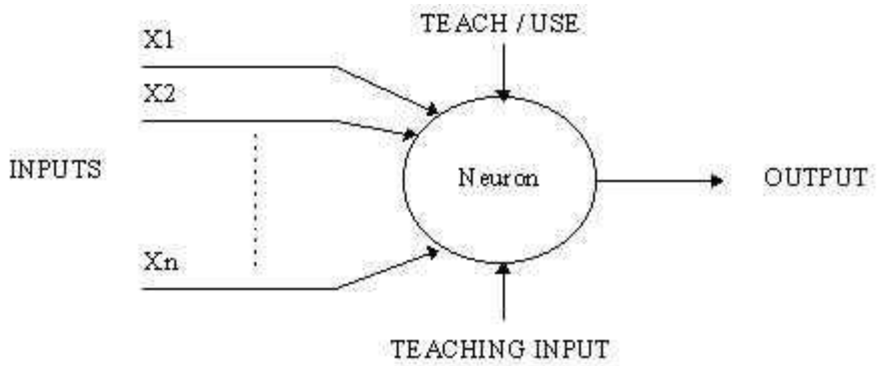


Figure 5.1 Simple Neural Networks

An important application of neural network is pattern recognition [10]. Pattern recognition can be implemented by using a feed-forward neural network [11] that has been trained accordingly. During training, the network is trained to associate outputs with input patterns. When the network is used, it identifies the input pattern and tries to output the associated output pattern. The power of neural networks comes to life when a pattern that has no output associated with it; it is given as an input. In this case, the network gives the output that corresponds to a taught input pattern that is least different from the given pattern.

The commonest type of artificial neural network [8] consists of three groups, or layer, of
"Input" in its is connected to a layer of "hidden" units, which is connected to a layer of "output" units.

- The activity of the units represents the raw information that is fed into the network.
- The activity of each hidden unit is determined by the activities of the input units and the weights on the connections between the input and the hidden units.
- The behavior of the output units depends on the activity of the hidden units and weights between the hidden and output units.

This simple type of network Fig: 5.2 is interesting because the hidden units are free to constructs their own representations of the input. The weights between the input and hidden units determine when each hidden unit is active, and s0 by modifying these weights, a hidden unit can choose what it represents

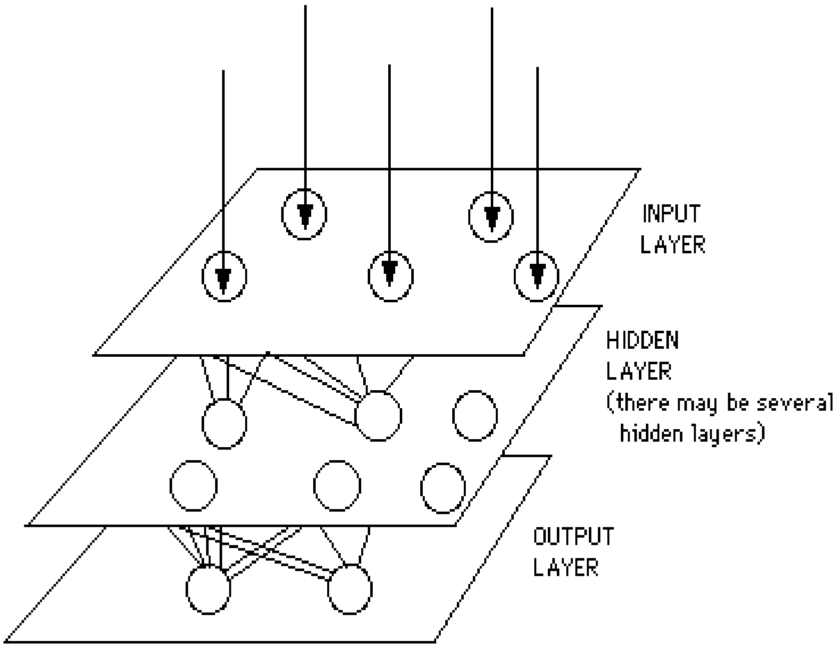


Figure 5.2 Layered Neural Networks

The computing world has a lot to gain from neural networks. Their ability to learn by example makes them very flexible and powerful. Furthermore there is no need to devise an algorithm in order to perform a specific task; i.e. there is no need to understand the internal mechanism of that task. They are also very well suited for real time systems because of their fast response and computational times which are due to their parallel architecture. Neural networks also contribute to other areas of research such as neurology and psychology. They are regularly used to model parts of living organisms and to investigate the internal mechanisms of the brain.

VI. DESCRRIPTION

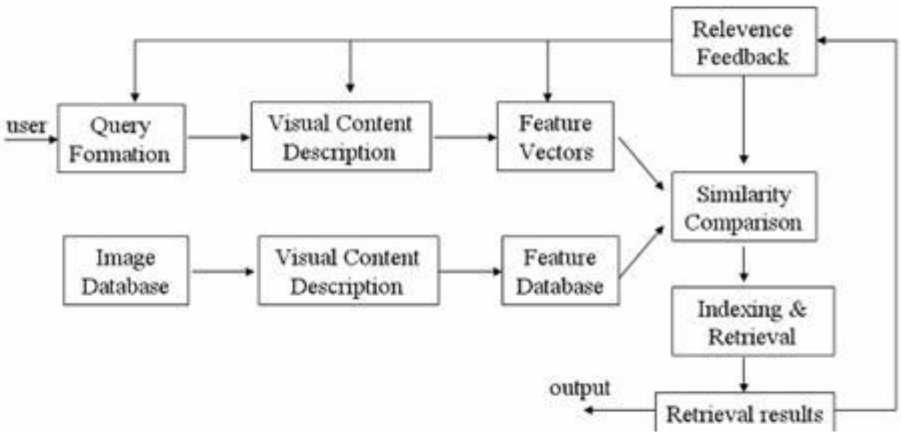


Figure 6.1 Content-Based Image Retrieval System

For content-based image retrieval Fig 6.1, user interaction with the retrieval system is crucial since flexible formation and modification of queries can only be obtained by involving the user in the retrieval procedure. User interfaces in image retrieval [4] systems typically consist of a query formulation part and result presentation part.

6.1 Query Specification

Query by concept is to retrieve images according to the conceptual description associated with each image in the database. Query by sketch and query by example is to draw a sketch or provide an example image from which images with similar visual features will be extracted from the database. The first two types of queries are related to the semantic description of images which will be introduced in the following chapters.

Query by sketch allows user to draw a sketch of an image with graphic editing tool provided either by the retrieval system or by some other software. Queries may be formed by drawing several objects with certain properties like color, texture, proof shape, sizes and locations. In most cases, a coarse sketch is sufficient, as the query can be refined based on retrieval results. 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. Query by example can be further classified into query by external image example, if the query image is not in the database, and query by internal image.

It is suitable for applications where the target is an image of the same object or set of objects under different viewing conditions Query by group example allows user to select multiple images. The system will then find the images that the best match the common characteristics of the group of examples. In this way, a target can be defined more precisely by specifying the relevant feature variations and removing irrelevant variations in the query. In addition, only group properties can be refined by adding negative examples. Many recently developed systems provided both queries by positive and negative examples.

VII. CONCLUSION

In this project, some fundamental techniques for content-based image retrieval, including visual content description, similarity/distance measures, indexing scheme, user interaction and system performance evaluation are discussed. The main emphasis is on visual feature description techniques. General visual features are most widely used in content-based image retrieval are color, texture, shape, and spatial information. Most CBIR research focuses on the utilization of advanced algorithms. An important constraint for CBIR is the complexity of the algorithm chosen. Each of the three features used for CBIR [4] (i.e., color, texture, and shape), was developed from a human-centered perspective, where in parallel improvements to algorithms were an issue. From Neural Network the ability to learn by example makes very flexible and powerful and there is no need to understand the internal mechanisms of the tasks. They are well suited for real time systems because of their fast response and computational times which are due to their parallel architecture.

VIII. FUTURE ENHANCEMENT

The CBIR techniques, as presented in this thesis, should be plugged into a General Frame work from which an online CBIR System can be developed. Next to images, video material can be based on its content for this purpose, frequently, so called key frames are selected, which can be analyzed as images. Hence for the CBIR techniques can be utilized. Note that where with CBIR only one images is present, in content-based video retrieval a set of images describes a part of the video.

REFERENCES

[1]. D.H. Kim and Chung,"QCluster8: Relevance Feedback Using Adaptive Clustering for Content Based Image Retrieval," Proc.ACM SIGMOD, pp.599-610, 2003

[2]. S.K. Chang,Q.Y.Shi, and C.Y. Yan,"Iconic indexing by 2-D Strings," IEEE Trans. On pattern Anal.Machine Intell. Vol.9, No.3,pp.413-428 May 1987

[3]. Y.Rui,T.Huang, and S.Mehrotra," Content Based Image Retrieval with Relevance Feedback in MARS," Proc IEEEInt cpnf. Image Processing, pp.815-818, oct.1997

[4]. J.Liu,Z.Li,M.Li,H.Lu, and S.Ma," Humane Behaviour Consistent Relevance Feedback Model for Image Retrieval," Proc. 15th Int'l Conf. Multimedia, pp. 269-272, Sept 2007.

[5]. Y. Rui, T. Huang, and S. Mehrotra, "Content-Based Image Retrieval with Relevance Feedback in MARS," Proc. IEEE Int'l Conf. Image Processing, pp. 815-818, Oct. 1997.

[6]. X. Jin and J.C. French, "Improving Image Retrieval Effectiveness via Multiple Queries," Multimedia Tools and Applications, vol. 26, pp. 221-245, June 2005.

[7]. R.Burnelli and O.Mich, Image Retrieval by examples, IEEE Transactions on Multimedia, 2(3):164-171, September 2000.

[8]. K.Al-Ghoneim and B.V.K.V.Kumar," Learning ranks with Neural networks," in Proc. SPIE Applications Science Artificial Neural Network Vol.2492,1995,PP.446-464

[9]. P.F.Baldi and K.Homik," Learning in linear neural networks: A Survey," IEEE Trans. Neural Networks, Vol.6,PP.837-858

[10]. D.E.Brown V.Corruble and C.L. Pittard,"A Comparision of Decision Tree Classifier with Back Propagation Neural Networks for multimodal classification problems," Pattern Recognition, Vol.26,PP.953-961,

[11]. D. S. Lee, S. N. Sriharia, and R. Gaborski, "Bayesian and neural-network pattern recognition: A theoretical connection and empirical results with handwritten characters," in Artificial Neural Networks and Statistical Pattern Recognition, I. K. Sethi and A. K. Jain, Eds. New York: Elsevier, 1991, pp. 89–108