ESTIMATION OF AGE USING FACIAL IMAGE

[1]K.BasavaRaju, [2]Dr.Y.RamaDevi, [3]Dr.P.V.Kumar.
[1]Research Scholar,JNTUK, [2]HOD,CSE Dept.CBIT,[3]Professor,CSE Dept., O.U. Telangana.
Email:[1]basavark@gmail.com, [2]yrdcse.cbit@gmail.com, [3]pvkumar58@gmail.com

Abstract—Image mining is a vital technique which is used to mine knowledge straightforwardly from image. Image segmentation is the primary phase in image mining. Image mining is simply an expansion of data mining in the field of image processing. Image mining handles with the hidden knowledge extraction, image data association and additional patterns which are not clearly accumulated
in the imagesThe face of a human being conveys a lot of information about identity and emotional state of the person. Face recognition is an interesting and challenging problem, and impacts important applications. In this work a novel and computational fast algorithm is proposed for predicting age of humans with PeanoCountTree (P-Tree). The predicting system was developed and tested based on texture features extracted local gradient patterns (LGP) and gray level co-occurrence matrix (GLMC) to give better and more predicting accuracy with a range of time period. In our research work, we empirically evaluate face recognition which considers both shape and texture information to represent face images based on Local Binary Patterns for person independent face recognition. The face area is first divided into small regions from which Local Binary Patterns (LBP) extracted and concatenated into a single feature vector. This feature vector forms an efficient representation of the face and is used to measure similarities between images.

Keywords— Local binary pattern (LBP), feature extraction, classification, pattern recognitin, feature vector, Decision tree.

I. INTRODUCTION

Knowledge discovery systems are constrained by three main limited resources: time, memory and sample size. Recognition of human skin is an important task for both computer applications. In computer graphics, facial animation is an important problem which necessitates reliable skin texture recognition. In addition to computer vision and graphics, skin recognition is useful in dermatology and several industrial fields.Many skin segmentation methods depend on skin color [1] [2] which has many difficulties. For the above reasons combining the texture features of skin with its color feature will increase the accuracy of skin recognition. Skin is a complex landscape that is difficult to model for many reasons. The skin texture features depends on many variables such as body location (knuckle vs. torso), subject parameters (age/gender/health) and imaging parameters (lighting and camera). Also as with many real world surfaces, skin appearance is strongly affected by the age with folds on the face as shown in the below figure1.

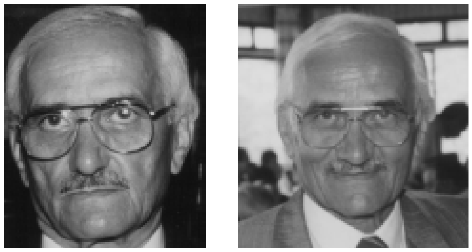


Figure 1: wrinkles changes based on the age on the face.

In many commercial areas, very large quantities of data are generated and collected every day, such as international airports, security surveillance systems, and Verification and authentication in different places. These data arrive too fast to be analyzed or mined in time. Such kinds of data are called "data streams" [3]. Classifying open-ended data streams brings challenges and opportunities since traditional techniques often cannot complete the work as quickly as the data is arriving in the stream [4]. Such data sets can be very large and are often archived in deep storage before valuable information can be obtained from them. An objective of spatial data stream mining is to mine such data in near real time prior to deep storage archiving. Classification is one of the important areas of data mining. In classification task, a training set (or called learning set) is identified for the construction of a classifier. Each record in the learning set has several attributes, one of which, the goal or class label attribute, indicates theclass to which each record belongs. The classifier, once built and tested, is used to predict theclass label of new records that do not yet have a class label attribute value.A test set is used to test the accuracy of the classifier. The classifier, once certified, is usedto predict the class label of future unclassified data. Different models have been proposed forclassification such as decision trees, neural networks, Bayesian belief networks, fuzzy sets, andgeneric models. Among these models, decision trees are widely used for classification. We focuson decision tree induction in this paper. ID3 (and its variants such as C4.5) [5, 6] and CART [7]are among the best known classifiers that use decision trees. Other decision tree classifiersinclude Interval Classifier [8] and SPRINT [9] which concentrate on making it possible to minedatabases that do not fit in main memory by only requiring sequential scans of the data.Classification has been applied in many fields, such as retail target marketing, customer retention,fraud detection and medical diagnosis . Spatial data is a promising area for classification. Inthis paper, we propose a decision tree based model to perform classification on spatial datastreams.A new data structure, the Peano Count Tree (P-tree) [10] is used to build the decision treeclassifier. P-trees represent spatial data bit-by-bit in a recursive quadrant-by-quadrant arrangement. With the information in P-trees, we can rapidly build the decision tree. Each new component in a spatial data stream is converted to P-trees and then added to the training set assoon as possible. Typically, a window of data components from the stream is used to build (or rebuild) the classifier. There are many ways to define the window, depending on the data andapplication.

In this paper, the focus is on feature extraction and building fast classifieralgorithm. We are implementing our work in four stages.

I. Image mining, Ptree and decision tree
2. Local Binary Patterns.
3. Feature extraction.
4. Application part (Lighting condition, Pose, size and Surgery problems).

Here we finished three things and images are correctly recognised. Last stage is still in progress.

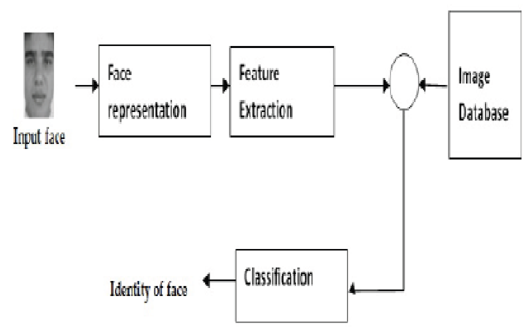


Figure2: General structure of steps in classification.

Most existing skin segmentation techniques involve the Classification of individual image pixels into skin and non-skin categories on the basis of pixel color. Lots of relative studies of skin color pixel classification have been reported. In [11] Nidhal K. Al abbadi proposed a method for skin texture recognition using neural network. They proposed a skin recognition system. This system is using skin color feature and texture feature. In [12] authors proposed a texture recognition system based on Grey Level Co-occurrence Matrix (GLCM) for automatic recognizing the texture. Based on the differences in texture appearance skin texture is categorized into 3 different disease classes. Features extracted from GLCM are contrast, homogeneity, mean and variance. Based on the differences in texture appearance skin texture is categorized into 3 different disease classes. Brand and Mason [13] compared three different techniques on the Compaq database: thresholding the red-green ratio, color space mapping with 1D pointer and RGB skin probability map. In [14] authors implemented a classifier using MLP neural network for face detection. Face can be detected through different features similar to shape skin texture and skin color. They are interested by the design of ANN not by the features. In [15] authors proposed a technique for image segmentation using texture content. Texture features are extracted from spatial blocks using quad tree decomposition. All feature sets are computed from Quadrature Mirror Filter (QMF) wavelet representation. Texture feature extraction can be done based on block-based features, wavelet sub band features.Numbers of comparative studies of skin color pixel classification have been reported.LDP is a general framework to encode directional pattern featuresbased on local derivative variations [16].The n^{th} -order LDP isproposed to encode the $(n-1)^{th}$-order local derivative direction variations, which can capture more detailed information thanthe first-order local pattern used in local binary pattern (LBP).Different from LBP encoding the relationship between the centralpoint and its neighbors, the LDP templates extract high-orderlocal information by encoding various distinctive spatial relationshipscontained in a given local region.

II. HUMAN SKIN COLOR MODEL

A human skin color model is used to decide if a color is a skin or nonskin color. Major requirements of a skin color model are listed below:

• ***Very low false rejection rate at low false detection rate.***

Skin color detection is first step in skin segmentation; therefore it is imperative that almost all skin colors are
detected while keeping the false detection rate low. False detections can be handled later when more *a priori* knowledge about the object of interest (ie. face, hand) is available.

• ***Detection of different skin color types.*** There are many
skin color types, ranging from whitish and yellowish to blackish and brownish, which must be all classified in one class, skin color.

• ***Handling of ambiguity between skin and nonskin colors.*** There are many objects in the environment that have the same color as skin. In these instances, even a human obser- ver cannot determine if a particular color is from a skin or nonskin region without taking into account contextual info r mation. An effective skin colormodel should address this ambiguity between skin and nonskin colors.

• ***Robustness to variation in lighting conditions.*** Skin color can appear markedly different under different
lighting. It is impractical to construct a skin color model that
works under all possible lighting conditions.However, a good skin color model should exhibit some sort of robust- ness to variations in lighting conditions. Inour work, we aim to create a skin color model for typical office lighting and daylight conditions.

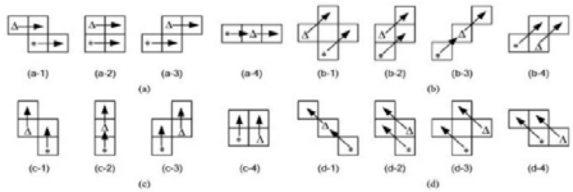
III. FEATURE EXTRACTION

A. Local gradient patterns

To investigate the feasi- bility and effectiveness of using high-order local patterns for face representation. An Local gradient patterns is proposed, in which the first-order derivative direction variations based on a binary coding function. In this scheme, LBP is conceptually regarded as the nondirectional first-order local pattern operator, becauseLBP encodes all-direction first-order derivative binary resultwhile LDP encodes the higher-order derivative informationwhich contains more detailed discriminative features that thefirst-order local pattern (LBP) can not obtain from an image.Given an image $I(Z)$, we calculate first-order derivatives along 0°, 45°, 90° and 135° directions, which is denoted as $I'_\alpha(Z)$ where α= 0°, 45°, 90° and 135°.

$$(a) I'_{0°}(Z_0) = I(Z_0) - I(Z_4)$$
$$(b) I'_{45°}(Z_0) = I(Z_0) - I(Z_3)$$
$$(c) I'_{90°}(Z_0) = I(Z_0) - I(Z_2)$$
$$(d) I'_{135°}(Z_0) = I(Z_0) - I(Z_1)$$

If Z_0 is one point in $I(Z)$,neighboring point around Z_0 (see Fig. below). So the four first-order derivatives at $Z=Z_0$



Once the first order gradient details of the image is obtained. The energy of the local gradient patterns is computed from the LDP image by the following equation.

$$Energy = \sum_{i,j}^{M,N} LDP(i,j)^2 \qquad (1)$$

B.Gray Level Co-occurrence Matrix (GLCM)

Texture analysis has been an active area of research in pattern recognition. A variety of techniques have been used for measuring textural similarity. In 1973, Haralick et al. proposed co-occurrence matrix (GLCM) representation of texture features to mathematically represent gray level spatial dependence of texture in an image [17]. In this method the co-occurrence matrix is constructed based on the orientation and distance between image pixels. Meaningful statistics are extracted from this co-occurrence matrix, as the representation of texture. Since basic texture patterns are governed by periodic occurrence of certain gray levels, co-occurrence of gray levels at predefined relative positions can be a reasonable measure of the presence of texture and periodicity of the patterns. Several texture features such as entropy, energy, contrast, and homogeneity, can be extracted from the co-occurrence matrix of gray levels of an image. The gray level co-occurrence matrix $C(i, j)$ is defined by first specifying a displacement vector $d_{x,y} = (\delta x, \delta y)$(where δx, δy are the displacements in the x and y directions respectively) and then counting all pairs of Pixels separated by displacement $d_{x,y}$and having gray levels i and j . The matrix $C(i, j)$ is normalized by dividing each element in the matrix by the total number of pixel pairs.Using this co-occurrence matrix, the texture featuresMetrics are computed as follows.

$$Entropy = \sum_{i,j}^{M,N} C(i,j)\log(C(i,j)) \qquad (2)$$

$$Contrast = \sum_{i,j}^{M,N} (i-j)^2 C(i,j) \qquad (3)$$

$$Homogeneity = \sum_{i,j}^{M,N} \frac{C(i,j)}{1+|i-j|} \qquad (4)$$

C.Feature dataset Organization using PeanoCount Tree

All the values have been scaled to values between 0 and 255 for simplicity. The pixel coordinates in raster order constitute the key attribute. One can view such data as table in relational form where each pixel is a tuple and each band is an attribute.

Figure3 : preprocessed image divided into 64 regions.

There are several formats used for spatial data, such as Band Sequential (BSQ), Band Interleaved by Line (BIL) and Band Interleaved by Pixel (BIP). The new format which was called as bit Sequential Organization (bSQ) is introduced. Since each intensity value ranges from 0 to 255, which can be represented as a byte, we try to split each bit in one band intoa separate file, called a bSQ file. For example, for a TIFF image with three bands, we have 24bSQ files.

This example has been taken from [18]. The following relation table contains 4 features of 4-bit data values.

LABEL	**Attributes**			
Record	**B1**	**B2**	**B3**	**B4**
0	0011	0111	1000	1011
1	0011	0011	1000	1111
2	0111	0011	0100	1011
3	0111	0010	0101	1011
4	0011	0111	1000	1011
5	0011	0011	1000	1011
6	0111	0011	0100	1011
7	0111	0010	0101	1011
8	0010	1011	1000	1111
9	0010	1011	1000	1111
10	1010	1010	0100	1011
11	1111	1010	0100	1011

12| 0010| 1011| 1000| 1111
13| 1010| 1011| 1000| 1111
14| 1111| 1010| 0100| 1011
15| 1111| 1010| 0100| 1011

Table 1. Feature sets in relational table

This dataset is converted in bSQ format. The feature-1bit-bands would be like below if we display the bSQformat in 2- dimension

B11	B12	B13	B14
0000	0011	1111	1111
0000	0011	1111	1111
0011	0001	1111	0001
0111	0011	1111	0011

There is a constraint on bSQ formation. P-tree requires the number of rows and columns in bSQ file be multipleof four. In case, the number of tuples in database is not aMultiple of four we need to pad zero vectors at the end;so that it is.

Based on this idea, to identify a person, the recognition
System works as follows:

- Basic p-trees of the enrolled feature sets are
  Constructed as preprocessing. This is a one-time job.
- Feature vector of the person to be recognized isextracted.
- A tuple p-tree of this feature vector is generatedbased on the basic p-trees.
- If the root count of this tuple p-tree are classified by Peano-Tree classifier.

These tuple p-tree values are subjected through decision tree will different age groups will have different gradient energy.

D. P-Tree based Decision Classifier

Classification is a data mining technique that typically involves three phases, a learningphase, a testing phase and an application phase. A learning model or classifier is built during thelearning phase. It may be in the form of classification rules, a decision tree, or a mathematicalformula. Since the class label of each training sample is provided, this approach is known assupervised learning. In unsupervised learning (clustering), the class labels are not known inadvance.In the testing phase test data are used to assess the accuracy of classifier. If the classifier passes the test phase, it is used for the classification of new, unclassified data tuples. This is theapplication phase. The classifier *predicts* the class label for these new data samples.In this paper, we consider the classification of spatial data in which the resulting classifier isa decision tree (decision tree induction). Our contributions include

(1)A set of classification-ready data structures called Peano Count trees, which are compact, rich in information and facilitate classification;

(2)A data structure for organizing the inputs to decision tree induction, the Peano count cube.

(3)A fast decision tree induction algorithm, which employs these structures. We point out the classifier is precisely the classifier built by the ID3 decision tree induction algorithm.

The point of the work is to reduce the time it takes to build and rebuild the classifier as new data continue to arrive. This is very important for performing classification on data streams.

A Decision Tree is a flowchart-like structure in which each node denotes a test on an attribute. Each branch represents an outcome of the test and the leaf nodes represent class's orclass distributions. Unknown samples can be classified by testing attributes against the tree. Thepath traced from root to leaf holds the class prediction for that sample. The basic algorithm forinducing a decision tree from the learning or training sample set is as follows [19]:

- Initially the decision tree is a single node representing the entire training set.
- If all samples are in the same class, this node becomes a leaf and is labeled with that class label.
- Otherwise, an entropy-based measure, "information gain", is used as a heuristic for selecting the attribute which best separates the samples into individual classes (the "decision" attribute).
- A branch is created for each value of the test attribute and samples are partitioned accordingly.

- The algorithm advances recursively to form the decision tree for the sub-sample set at each partition. Once an attribute has been used, it is not considered in descendent nodes.
- The algorithm stops when all samples for a given node belong to the same class or when there are no remaining attributes (or some other stopping condition)

The attribute selected at each decision tree level is the one with the highest information gain. The information gain of an attribute is computed by using the following algorithm. Assume B[0] is the class attribute; the others are non-class attributes. We store the decision path for each node. For example, in the decision tree below (Figure 3), the decision path for nodeN09 is "Band2, value 0011, Band3, value 1000". We use RC to denote the root count of a P-tree, given node *N*'s decision path B[1], V[1], B[2], V[2], … , B[t], V[t], let P-tree P=PB[1],v[1]^PB[2],v[2]^…^P B[t],v[t] , we can calculate node *N*'s information I(P) through

$$I(p) = -\sum_{i=1}^{n} p(i)\log 2(p(i))$$

Wherepi=*RC (P^PB[0], V0[i])/RC(P),* here V0[1]..V0 [n] are possible B[0] values if classified by B[0]at node *N*. If *N* is the root node, then P is the *full P-tree* (root count is the total number of transactions).Now if we want to evaluate the information gain of attribute A at node *N*, we can use the formula: Gain (A) =I (P)-E (A), where entropy

$$E(A) = \sum_{i=1}^{n} I(P\wedge PA, VA[i]) * \frac{RC(P^{PA}, VA[i])}{RC(P)}$$

WhereVA [1], VA[n] are possible A values if classified by attribute A.Thus, the B1 basic P-trees are as follows (tree pointers are omitted).

P1, 1P1, 2 P1, 3P1, 4
5 7 16 11
0014 0403 4403
0001 0111 0111

Then we generate basic P-trees and value P-treessimilarly to F2, F3 and F4.Start with A = F2. Because the node currentlydealing is the root node, P is the full P-tree. So pi canbe 3/16, 1/4, 1/4, 1/8, 3/16, thus we can calculate

*I(P) = 3/16*log2(3/16) + 4/16*log2(4/16) + 4/16*log2(4/16) + 2/16*log2(2/16) + 3/16*log2(3/16)*
I(P) = 2.281

To calculate E (B2), first P^PA, VA[i] should be allthe value P-trees of B2. Then I (P^PA, VA[i]) can becalculated by ANDing all the B2 value P-trees andB1 value P-trees. Finally we get E (B2) =0.656 andGain (B2) =1.625.Likewise, the Gains of B3 and B4 are computed:Gain (B3) = 1.084,Gain (B4) = 0.568. Thus, B2 isselected as the first level decision attribute.Branches are created for each value of F2 andsamples are partitioned accordingly.

B2=0010 → Sample_Set_1
B2=0011 → Sample_Set_2
B2=0111 → Sample_Set_3
B2=1010 → Sample_Set_4
B2=1011 → Sample_Set_5

Advancing the algorithm recursively to each subsampleset, it is unnecessary to rescan the learning set to form these sub-sample sets, since the P-trees forthose samples have been computed.The algorithm will terminate with the decision tree:

B2=0010 → B1=0111
B2=0011 → B3=0100 → B1=0111
→ B3=1000 → B1=0011
B2=0111 → B1=0011
B2=1010 → B1=1111
B2=1011 → B1=0010

Based on the above describe entropy and gain values root nodes are selected. The branch will give the path of particular tuple in the data set with class label as shown in the following decision tree.

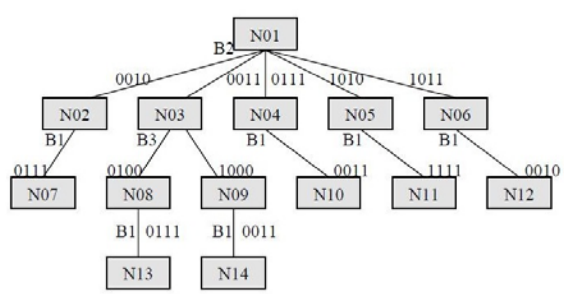


Figure 4: Decision Tree on example dataset

IV.EXPERIMENTAL RESULTS

A spatial image can be viewed as a 2-dimensional array of pixels. Associated with each person to computed energy of local gradient pattern image and Gray level Co-occurrence metric (GLCM) from the data base. The energy of LGP is taken as class label andglobal texture features as various descriptive attributes, called “features”. Based on idea of texture analysis, it is understood with the increase age of particular person the wrinkles on the face will appear. Which change the properties of skin texture in terms of more number of changes will happen. This can be identified through first gradient operators.

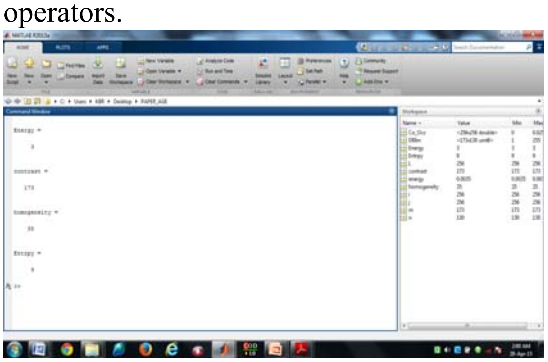


Figure 5:Energy, contrast, homegeneity, entropy are calculated.

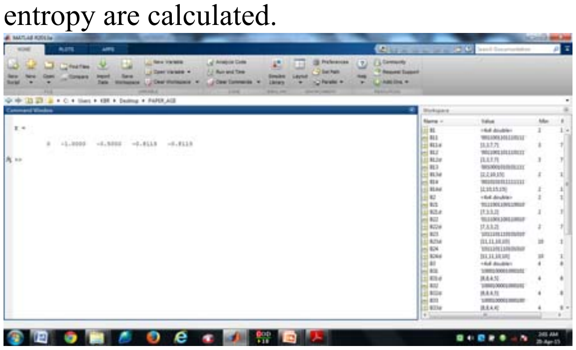


Figure 6. Entropy calcuated as per the procedure

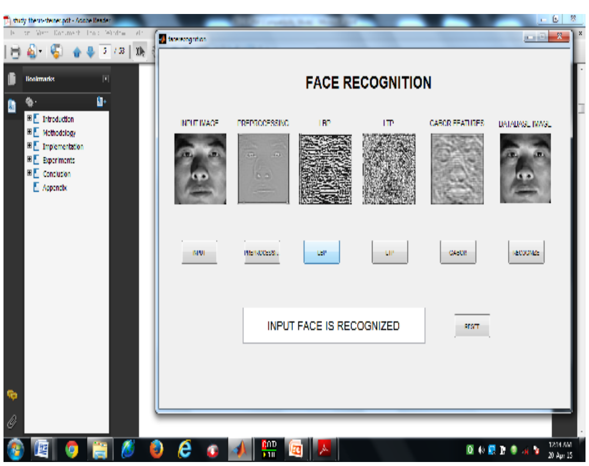


Figure 6 : Image processing based on above procedure and recognition of a human face.

V. Conclusion

This system is an application of image mining. Here knowledge is obtained based on wrinkles and by using these wrinkles it is easy to guess age of a person with the help of assumption rules in bio- informatics. This system is based on texture analysis. Here after preprocessing we performed LBP, LTP, Gabor filter entropy calculations for a given image. Here first order gradient operator and GLCM is taken as a method of feature extraction.For analyzing texture, symmetrical normalized GLCM is computed along four directions. From this matrix texture features are calculated and averaged over 4 directions. The gradient energy is calculated along with contrast, homogeneity and entropy. These features are very useful to recognize texture of skin and these features are classified with a novel P-tree classifier. This classifier uses decision tree induction that is especially usefulfor the classification of spatial data streams. We use the data organization, bit Sequential organization (bSQ) and a lossless, data-mining ready data structure, the Peano Count Tree (P-tree), to represent the information needed for classification in an efficient and ready-to-use form. This makes classification of open-ended streaming datasets feasible in near real time.

REFERENCES

[1] Hwei-Jen Lin, Shu-Yi Wang, Shwu-Huey, and Yang –Ta-Kao “Face Detection Basedon Skin Color Segmentation and NeuralNetwork” IEEE Transactions on, Volume: 2,pp1144-1149, ISBN: 0-7803-9422-4, 2005.

[2] Son Lam Phung, AbdesselamBouzerdoum,and Douglas Chai “Skin Segmentation UsingColor and Edge Information” IEEE ISSPAISBN: 0-7803-7946-2 2003.

[3] Domingos, P. and Hulten, G., “Mining high-speed data streams”, Proceedings of ACMSIGKDD 2000.

[4] Domingos, P., &Hulten, G., “Catching Up with the Data: Research Issues in Mining DataStreams”, DMKD 2001.

[5] J. R. Quinlan andR. L. Riverst, “Inferring decision trees using the minimum descriptionlength principle”, Information and Computation, 80, 227-248, 1989.

[6] Quinlan, J. R., “C4.5: Programs for Machine Learning”, Morgan Kaufmann, 1993.

[7] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, “Classfication and RegressionTrees”, Wadsworth, Belmont, 1984.

[8] R. Agrawal, S. Ghosh, T. Imielinski, B. Iyer, and A. Swami. “An interval classifier fordatabase mining applications”, VLDB 1992.

[9] J. Shafer, R. Agrawal, and M. Mehta, “SPRINT: A scalable parallel classifier for datamining”, VLDB 96.

[10] William Perrizo, Qin Ding, Qiang Ding, Amlendu Roy, “On Mining Satellite and OtherRemotely Sensed Images”, DMKD 2001.

[11] Al abbadi, N.K. et. al., “Skin texture recognition using neural network,” In Proc. of the International Arab Conference on Information Technology, Tunisia, 2008.

[12] Parekh, R., Mukherjee, A., “Advances in Telemedicine: A Multimedia-Based Texture Recognition Diagnostic System”, In Proc. of the Business and Health Administration Association (BHAA ‟09) International Conference, Chicago, Illinois, USA, pp. 88-97, March 18-20, 2009.

[13] J. Brand and J. Mason, “A Comparative Assessment of Three Approaches to Pixel- Level Human Skin Detection,” Proc. IEEE Int’l Conf. Pattern Recognition, vol. 1, pp. 1056-1059, Sept. 2000.

[14] Smach, F. ET. al., “Design of a neural networks classifier for face detection”, Science Publication, Journal of Computer Science, Vol. 2, issue 3, pp. 257-260, 2006.

[15] Smith, J.R. et.al., “Quad tree segmentation for texture based image query”, In Proc. of the second ACM international conference on Multimedia, San Francisco, California, United States, pp. 279 - 286, 1994.

[16] Baochang Zhang, YongshengGao, Sanqiang Zhao, and JianzhuangLiuLocal Derivative Pattern Versus Local BinaryPattern: Face Recognition With High-OrderLocal Pattern Descriptor.IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 19, NO. 2, FEBRUARY 2010.

[17] TinkuAcharya “Image Processing Principlesand Applications” A JOHN WILEY & SONS,MC. 2005.

[18] Qiang ding, William Perrizo, “Cluster analysis ofspatial data using peano count tree”, Proceedings ofCATA2002, San Francisco, USA, April 4-6, 2002.

[19] Jiawei Han, MichelineKamber, “Data Mining: Concepts and Techniques”, MorganKaufmann, 2001.

[20] “Association rule mining on remotely sensedimages using peano count trees,” Pacific-AsiaConference on Knowledge Discovery and Data Mining,pp. 66-79, 2002