Increasing the Reliability of Fuzzy Angle Oriented Cluster Using Peer-to-Peer

Remani Naga Venkata Jagan Mohan[1,*], Vegi Srinivas[2], and Kurra Rajasekhara Rao[3]

[1] CSE, Swarnandra College of Engg&Tech., Narsapuram-534275, India
[2] CSE, Dadi Institute of Engg. & Tech., Anakapalli-531002, India
[3] Dept. of CSE, K.L. University, Vaddeswaram, Guntur-522502, A.P., India
{mohanrnvj,srini.vegi}@gmail.com, krr_it@yahoo.co.in

Abstract. The vast volume of data is collected and it needs to be analyzed rapidly for Quality of Data and Quality of Service (QoS), both are used for verification of sharing the information not only for web applications, but also used for many user applications over a network. In this paper, we proposed MapReduce (i.e., Parallelized and Distributed) process used for improving the performance of peer to peer communication on angle oriented clusters in Big Data. To study of this paper, the data set classified into two types namely, Clock wise and Anti-clock wise rotations using Fuzzy cluster classification. The data is extracted by using the angle oriented DCT (Discrete Cosine Transform) that invokes certain normalization techniques. Also, matching the data is compared with the technique of similarity based approach using Tanimatto distance. A high recognition rate is observed using Nelson model for this approach, and it is proved by giving an example.

Keywords: Angle Oriented, Fuzzy Cluster, MapReduce, Peer-to-Peer, Quality of Service.

1 Introduction

Reliability can be classified into two major concepts; the first one is simply the continuity of correct service delivery. On other hand, the probability of failure-free operation of a computer program for a specified period of time in a specified environment. To increasing the reliability in every software process, we can follow the two aspects to measure namely, Quality of Data and the other one is Quality of Service. Quality of Data is important for recognition of the system. Data Quality aspects are accuracy, completeness, update status, relevance consistency, reliability and accessibility. The purpose of data quality is fitness to serve in a particular perspective. It is important for operational, transactional and reliability processes. It is affected by inserting, storing the data and managing the data. The quality of data assurance is the process of verifying the reliability and effectiveness of data. To maintain the quality that needs data, going through periodical and scrubbing it. Quality of Service is a set of characteristics that are related to the performance of the

* Corresponding author.

K. Singh, A.K. Awasthi, and R. Mishra (Eds.): QSHINE 2013, LNICST 115, pp. 928–938, 2013.

connection. For every application software supports highly adaptable and be capable of tolerating a wide range of reconfigurations and extensions whereas guarantee to conduct the meeting their Quality of Service (QoS). Earlier works on these lines were proposed by various authors, like *Improving QoS for Peer-to-Peer Applications through Adaptation* is suggested by Daniel Hughes et al., 2003[3].

In our present study of this approach, quality of service is relies on complexity of software for providing high quality services to their internal and/or external clients. While the process is going on to meet the requirements such as high availability, dynamic resource allocations, and ease of management in the manufacture environment. As in the new technologies such as Hadoop and MapReduce continue their perceptions into mainstream market, more applications will be developed and moved into production. Quality of Service will certainly develop a critical consideration to move to the Big Data.

As mentioned in the above concept, is applicable on the angle oriented recognition system and is important for law-enforcement security and video-surveillance systems. In this system, we use to compare the image database and input image by peer-to-peer communication i.e., communicate directly with one-to-one or one-to-many. It is a common way to share resources on a peer-to-peer is by modifying the file sharing controls through the operating system. The angle orientation is broadly classified into two categories i.e., Clock wise and Anti-Clock wise and is suggested by Jagan Mohan et al., 2011-12 [4, 5, and 6] discussed about how to increase the reliability of angle oriented face recognition using DCT [17]. It is used to recognize the feature images of the faces even though they are in angle oriented. If the angle of input image is not equal to 90^0, rotate the image into 90^0 and then apply normalization technique such as geometric and illumination technique. Recognition of an image, by using rotational axis is easy to achieve or recognize the face. If input image rotates from horizontal axis to vertical axis the face rotates anti-clock wise and the face appears in which it is the same as the database pose, then the object is recognized. Similarly, if input image rotates from vertical axis to horizontal axis the face rotates clock wise and the face appears in which it is the same as the database pose, then the object is recognized. Therefore, if input image is angle oriented, the pose is changed or angle is altered using rotational axis and then only comparison is done.

By careful observation and comprehensive reference of above concepts, the organization of this paper as follows; In Section II, Classification of Fuzzy angle oriented cluster approaches are introduced. Section III, deals with Distance Classifier Method. Linear Regression technique is introduced in Section IV. Angle oriented with MapReduce described in section V. Cluster Reliability classifier method is discussed in Section VI. Finally, the experimental results on angle oriented cluster database are provided in section VII, and we present our Conclusion and Future Perspectives in Section VIII.

2 Nomenclature of Fuzzy Cluster

Earlier works, on Fuzzy Classification were proposed by various authors stated Fuzzy Classifications. Kuok et. al., designated fuzzy association rules such as, Mining Fuzzy Association Rule is the discovery of association rule using fuzzy set concepts such that

the quantitative attribute can be handled, 2007 [10]. Jen Chianga.I et al., have given a Fuzzy classification trees for data analysis, 2002 [8][9], proposed Fuzzy classification trees are a model that integrates the fuzzy classifiers with decision trees that can work well in classifying the data with noise. Instead of defining a single class for any given instance, fuzzy classification predicts the degree of possibility for every class.

In our present discussion, I = {0^0, 30^0, 45^0, 60^0, 90^0, 180^0, 360^0} represents all the attributes appearing that are angles in Input Training Database is C = {0^0,30^0,60^0,90^0} contains all the possible items of a database C. Jagan Mohan et. al., discussed about the classification of clustering database images using decision tree given an Efficient K-Means Cluster Reliability on Ternary Face Recognition using Angle Oriented Approach, 2012 [12]. Whole database images denoted by C which is the root node i.e., [0^0-90^0], divided into two groups namely, Cluster C_1 and C_2 representing internal nodes i.e., Clock wise and Anti Clock wise. It tells that, the similar image object groups being rotated in the clock wise belong to the internal node, cluster C_1. On the other hand, C_2 which has a group of similar image objects rotates in Anti clock wise direction. Again the Cluster C_1 is Re-grouped into three terminal nodes C_{11} [0^0-30^0], C_{12} [31^0-60^0] and C_{13} [61^0-90^0], called nested clusters. On the other hand, the cluster C_2 is re-grouped into three terminal nodes C_{21} [0^0-30^0], C_{22} [31^0-60^0], C_{23} [61^0-90^0], called nested clusters of C_2. The θ is the angle of rotation in each cluster.

Fuzzy sets and their corresponding membership functions have to be defined by domain experts each of the fuzzy sets can be beheld as [0, 1] valued attribute, called fuzzy attribute.

Note: A membership function f (x) is characterized by the following mapping: f: x→ [0, 1], x ε X where x is a real number describing an object or its attribute. Let X= [0^0, 90^0]. Define f(x) = Sin x.

A fuzzy rule as follows

$$F = \sum_{\theta=0^\circ}^{90^\circ} \sin\theta \tag{1}$$

In this regard, we propose the variable θ is broadly classified into three fuzzy relations. The fuzzy relation can be classified as follows:

$$\sum_{\theta=0^\circ}^{30^\circ} Sin\theta \ , \sum_{\theta=31^\circ}^{60^\circ} Sin\theta \ , \sum_{\theta=61^\circ}^{90^\circ} Sin\theta \tag{2}$$

Finally, the fuzzy relation can be represented by

$$\mathrm{F} = \sum_{\theta=0^\circ}^{30^\circ} Sin\theta \ + \sum_{\theta=31^\circ}^{60^\circ} Sin\theta + \sum_{\theta=61^\circ}^{90^\circ} Sin\theta \tag{3}$$

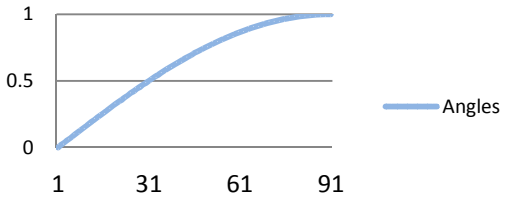

Fig. 1. Fuzzy Cluster Graph

The Θ values vary between 0^0-90^0 as shown in the above graph. By this we recognize fuzzy cluster values always in between 0 and 1.

3 Distance Classifier Method

There are few transformation techniques like DCT, DWT, KLT, SVD and Gabber methods are used for calculate the extraction of image feature vectors have certain limitations like poor discriminatory power and ability to handle large computational load. While studying of this section, we discussed, *Tanimatto distance classifier method for image matching* as in [12, 15].

The classification system analyzes the numerical properties of various image features and organizes the data into categories. This classification includes a wide range of theoretic-decision approaches for the identification of images for each image debits one or more features and each of these features belongs to one of several distinct classes. In practice, the minimum distance classifier works well, provided the distance between means is large when compare to the randomness of each class with respect to its mean. The minimum distance classifier is used to categorize into an unknown image data classes and therefore, the minimum distance between the image data and the classes in multi feature space exists. The distance is defined as an index of similarity and in consequence the minimum distance is identical to the maximum similarity. And now, this paper describes Tanimatto distance is as follows.

3.1 Tanimatto Distance Measure

The Tanimatto distance classifier is a comprehensive of Jaccard Coefficient and can be used for document data but reduce the Jaccard coefficient in the case of binary attributes. This is represented by T.

$$\text{Tanimatto Distance} = \frac{\boldsymbol{X}.\boldsymbol{Y}}{||\boldsymbol{X}||^2 + ||\boldsymbol{Y}||^2 - \boldsymbol{X}.\boldsymbol{Y}} \tag{4}$$

If a "similarity ratio" is given over bitmaps, where each bit of a fixed-size array represents the presence or absence of a character being modeled [2]. The definition of the ratio is the number of common bits divided by the number of bits set in either sample. The same calculation is expressed in terms of vector product and magnitude. This representation relies on the fact that, for a bit vector (where the value of each dimension is either 0 or 1) then

$$X.Y = \sum_{i} (X_i \Lambda Y_i) \ and \ |X|2 = \sum_{i} (X_i) \tag{5}$$

4 Regression

In many applications, it is necessary to estimate or predict the values of a numeric attribute that has a continuous range. The numeric attribute to be estimated may depend on a multiple attributes called as multivariate regression [14]. Regression in which the predictor variables are measured with error, regression with more predictor variables than observations, and casual inference with regression. These problems are usually solved by assuming some form of an equation that relates the input variables to dependent variable. In this paper, we proposed linear regression technique used in reduce function.

4.1 Linear Regression

In linear regression, the equation takes the form of a line and hyper plane in higher dimensions. If y is the dependent variable and x_1, x_2, x_3… x_n are the input variables, then the relation between them is of the form

$$Y = w_1x_1 + w_2x_2 + \ldots + w_nx_n \tag{6}$$

The training data contains multiple records giving the values of y for specific values of the input variables. The problem thus reduces to computing the values of w_1, w_2, w_3, …, w_n that best fits the data. Once a regression model has been fit to a group of data, examination of the residuals allows the modeler to investigate the validity of the assumption that a linear relationship exists.

In the multivariate regression, the time complexity can be calculated for the number of parameters and the sample size is O (kp^2), where k is the sample size, and p is the number of parameters to be estimated. Also it deals with the popular regression methods like; robust regression, decision tree regression, and support vector regression.

5 Map Reduce Query Processing

The notion of angle oriented cluster based recognition query processing, using nearest neighbor classifier with the help of Tanimatto distance and MapReduce programming model is proposed. MapReduce has been working on query processing for optimization technique [1]. MapReduce is a functional programing model that implements for processing and generating large data sets as in[7][11][13]. It is well-developed technology for distributing and parallelized environments and also working on large cluster data sets. Initially, the user takes on input image pairs i.e., Clock wise clusters or Anti-Clock wise clusters. For each cluster is segmented into three nested clusters, to classifies in fuzzy cluster classifications as we already discussed in the above section. In each cluster produces a set of intermediate key value pairs (images) with the help of map functions written by users.

Now, first we calculate the mean value for each image i.e., intermediate key value with the help of image normalization process and feature extraction. Different authors are, Mathew Turk and Alex Pentland, 1991 [16] expanded the idea of face recognition. Jagan Mohan R.N.V. and Subbarao R, 2011 [4] given the idea to increase the Reliability of Angle Oriented Face Recognition using DCT. Each cluster consisting of intermediate key values associated with the same intermediate key and passes them to the Reduce function with the help of MapReduce library. In the same way remaining clusters are also having the same process.

Finally, the user of the reduce function is accepts an intermediate key and a set of values for that key for each cluster. We compare the mean value with the target image of database with the help of face matching using similarity based distance measures are calculated i.e., Tanimatto distance with the help of linear regression technique.

Note: Def. V: DXI→ [-1, 0]

$$(x, y) = xy = \min(x, y) - \max(x, y) \tag{7}$$

$$\forall \ (x, y) \in D \times I \tag{8}$$

The user request to produce the output value whichever is minimum is required database image. The remaining values are denied.

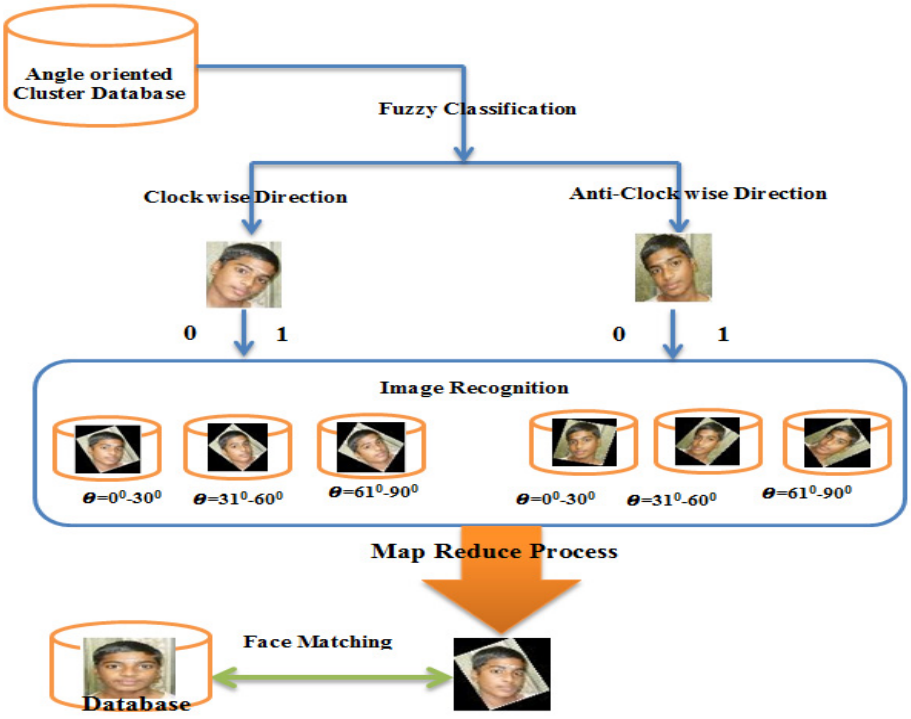

Fig. 2. MapReduce with Angle Orientation

6 Cluster Reliability

The Cluster Reliability for an angle oriented recognition system is more efficient, as shown in the above figure. This process is analyzed through experimental results.

6.1 Nelson Model

To find the cluster reliability we used distance classifier method i.e., a software reliability model is called Nelson Model [12].In this model, the software reliability is assessed by,

$$R = 1 - \frac{n_f}{n} \qquad (9)$$

Where n is total no.of runs and f is no. of failure runs.

7 Experimental Results

7.1 Anti-clock Wise Rotations

The experimental results are observed for both clock wise and anti-clockwise rotations for various clusters. The method for Tanimatto Distance is used for face

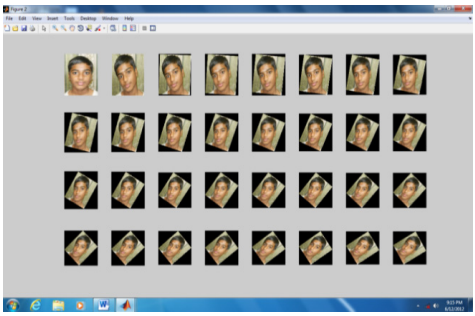

Fig. 3. Anti-Clock Wise Rotation Images

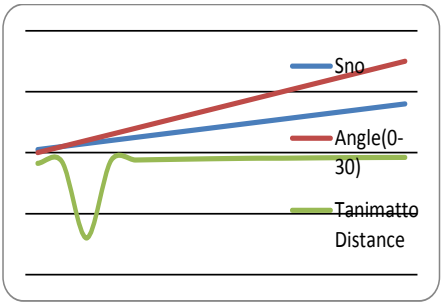

Fig. 4. C11 (0^0-30^0)

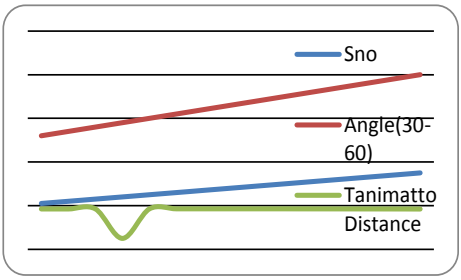

Fig. 5. C12 (30^0-60^0)

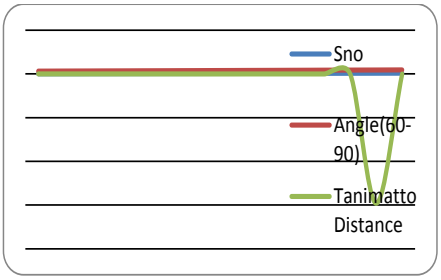

Fig. 6. C13 (60^0-90^0)

matching. The Nelson model of software reliability is used for efficiency of the face recognition in all types of clusters. The mean recognition values of the above method for all the three clusters c_{11}, c_{12} and c_{13} are calculated and the corresponding graphs for each of the three clusters are depicted in figures 4, 5 and 6 respectively.

7.2 Clock-Wise Rotations

As in the Clock wise rotation, the experimental results are calculated in anti-clock wise direction also. The same method along with Nelson model for reliability are adopted for face matching, and the mean recognition values using the algorithm given in section 1 for the three clusters c_{21}, c_{22} and c_{23} are calculated and their corresponding results are given in figures 8, 9 and 10 respectively.

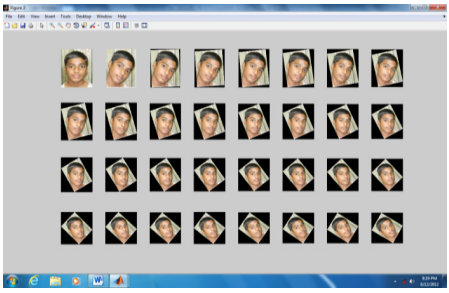

Fig. 7. Clock Wise Rotation Images

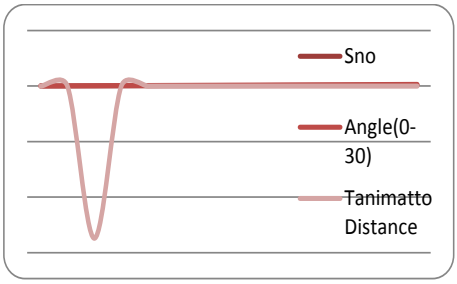

Fig. 8. C21 (0^0-30^0)

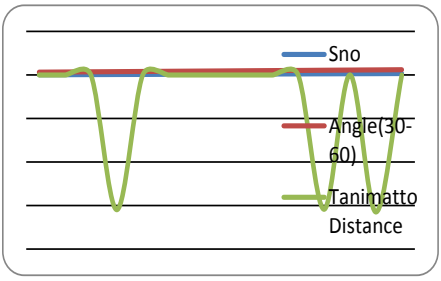

Fig. 9. C22 (30^0-60^0)

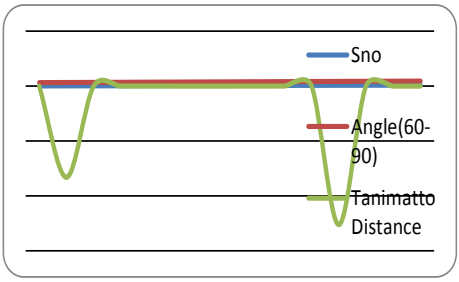

Fig. 10. C23 (60^0-90^0)

7.3 Comparison between DCT and KLT

The Tanimatto distance method is used in DCT and KLT techniques are experimented under standard execution environment by considering the created data by the students of Sri Vasista Educational Society. The phenomenal growth of DCT reliability is observed when compared with the KLT. The graph clearly shows that the reliability performance of DCT is constantly increasing with respect to KLT, while the number of records is increased.

Table 1. Performance Records in DCT and KLT

S. No.	No. of records	Performance in DCT	Performance in KLT
1	1000	91.46	65.42
2	2000	92.01	64.23
3	3000	93.25	52.15
4	4000	94.12	54.12
5	5000	94.62	53.10
6	6000	96.5	52.63
7	7000	97.23	51.71
8	8000	97.56	50.46
9	9000	98.46	50.04
10	10000	98.89	54.13

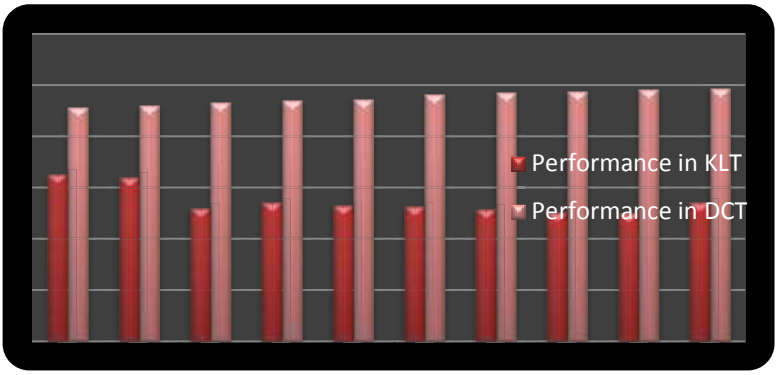

Fig. 11. Bar Chart for Performance in DCT and KLT

8 Conclusion and Future Percepective

In this paper, the Angle oriented approaches on Face Recognition is proposed and it is also used for MapReduce programming model. This model has been successfully used at Google for many different purposes. This model is easy to use, even for programmers without experience with parallel and distributed systems, since it hides the details of parallelization, fault tolerance, locality optimization, and load balanceing. MapReduce is used for a large variety of problems are easily solvable. MapReduce can implement in large clusters of datasets including thousands of clusters. The concept of face recognition is studied through Fuzzy cluster classification on the basis of rotation of images both clock-wise and anti-clock wise directions. Then, each cluster is segmented with respect to angles for face recognition using distance classifier method, Tanimatto Distance. Nelson model of reliability

approach is considered for comparing cluster reliability. It is proved through the experimental results that the Tanimatto method has high probability recognition rate. This approach has several applications in Cloud Computing, Spatial Mining and new technologies like Biometric Systems etc.

References

1. Atnafu, S., Brunie, L., Kosch, H.: Similarity-based operators and query optimization for multimedia database systems. In: Proc. of Int. Database Eng. and App. Symposium (IDEAS), pp. 346–355. IEEE computer Society, Grenoble (July 2001)
2. Atnafu, S., Brunie, L., Kosch, H.: Similarity-based operators in image database systems. In: Wang, X.S., Yu, G., Lu, H. (eds.) WAIM 2001. LNCS, vol. 2118, p. 14. Springer, Heidelberg (2001)
3. Hughes, D., Warren, I., Coulson, G.: AGnuS: The Altruistic Gnutella Server. In: proceedings of the Third International Conference on Peer-to-Peer Computing, Linköping Sweden (2003)
4. Jagan Mohan, R.N.V., Subbarao. R.: Increasing the Reliability of Angle Oriented Face Recognition using DCT. In: Published In the Proceedings of International Congress on Productivity, Quality, Reliability, Optimization and Modeling, Allied Publishers (2011)
5. Jagan Mohan, R.N.V., Subba Rao, R., Raja Sekhara Rao, K.: Similarity of Inference Face Matching on Angle Oriented Face Recognition. Journal of Information Engineering and Applications 1(1) (2011) (print) ISSN2224-5758, ISSN 2224-896X
6. Jagan Mohan, R.N.V., Subba Rao, R., Raja Sekhara Rao, K.: Similarity of Inference Face Matching on Angle Oriented Face Recognition. Intelligent Systems from Published in Journal of Computer Engineering and 3(2) (2012) (Paper), http://www.iiste.org ISSN 2222-1719, ISSN 2222-2863
7. Dean, J., Ghemawat, S.: MapReduce: Simplified Data Processing on Large Cluster. In: OSDI 2004: Sixth Symposium on Operating System Design and Implementation, San Francisco, CA (December 2004)
8. Jen Chianga, I., Yung, J., Hsu, J.: have given a Fuzzy classification trees for data analysis. Fuzzy Sets and Systems 130, 87–99 (2002), http://www.elsevier.com
9. Makkithaya, K., Subba Reddy, N.V., Dinesh Acharya, U.: Intrusion Detection System using Modified C-Fuzzy Detection Tree Classifier. Published in the International Journal of Computer Science and Network Security 8(11) (November 2008)
10. Kuok, C.M., Fu, A., Wong, M.H.: Mining Fuzzy Association Rules in Databases. SIGMOD Record 27 (1998)
11. Ramakrishna, V.M., Nepal, S.: Similarity Query Processing in Image Databases. Technical Report, University Melbourne (2000)
12. Subbarao, R., Raja Sekhara Rao, K., Jagan Mohan, R.N.V.: Efficient K-Means Cluster Reliability on Ternary Face Recognition using Angle Oriented Approach. Published In the Proceedings of International Conference on Advances in Communication, Navigation and Signal Processing Technically Co-Sponsored by IEEE, Hyderabad Section, March 17-18, Dept of ECE, Andhra University College of Engineering (2012)
13. Babu, S.: Towards Automatic Optimization of MapReduce Programs. In: The1st ACM Symposium on Cloud Computing (2010)

14. Schmerr, L.W., Prabhu, G.M., Forouraghi, B.: Fuzzy Multiobjective Optimization with Multivariate Regression Trees. In: IEEE World Congress on Computation Intelligence, vol. 2, pp. 1400–1405 (1994)
15. Tanimoto, T.T., Rogers, D.J.: A Computer Program for Classifying Plants. Science 132, 1115–1118 (1960)
16. Turk, M., Pentland, A.: Eigen faces for recognition. Journal of Cognitive Neuroscience 3(1), 71–86 (1991)
17. Ziad Hafed, M., Levine, M.: Face Recognition using Discrete Cosine Transform. International Journal of Computer Vision 43(3), 167–188 (2001)