Target Inference on Evaluation of Angle Oriented Cluster

R.N.V.Jagan Mohan1,*, K.Raja Sekhara Rao2

1Dept of CSE, Swarnandhra College of Engg & Tech, Narsapur-534280
2Sri Prakash College of Engineering, Raju Peta, Tuni-533401
*Corresponding Author: mohanrnvj@gmail.com

Abstract  In general, any field consists of unnecessary data. Several algorithms exist to remove unwanted data because it cannot seal to this processes. Research Scholars are still studying to complete this work. For Instance, face recognition system suffers in-depth pose verification problem over the last few decades. To solve this problem we used angle orientation technique. It consists of various angles of input images (same person with different direction) to compare with the database image. To remove needless data i.e., unsupervised image is the best solution to recognize a target inference. So with this idea we are attempting a small approach for this kind of applications. In this paper, we introduced a ternary cluster relation on angle oriented images. Again, various angles of images form into three nested clusters in Clock wise and/or Anti-clock wise directions. In this, we used multivariate analysis technique to improve the quality of cluster with the help of evaluation of cluster and also statistical approaches of tackle outlier detection methodology and bootstrapping technique to find the target inference. The experimental results are produced on angle oriented cluster images to increase the performance using analysis of variance test.

Keywords  Angle Oriented, Bootstrapping, Cluster Reliability, Multivariate Analysis, Tackle Outlier, Ternary Cluster

1. Introduction

In any engineering approach, verification and validation technique is more important before introducing any reliable detection system. At this moment, in many reliable systems, loading of object may be duplicated and is a common problem in many applications like recognition system. For example, in angle orientation system, object can be in various directions of angles or poses, so it leads to data redundancy. In 2011, Jagan Mohan et al found the novel approach based on angle orientation method in Increasing the Reliability of Angle Oriented Recognition System [7, 8]. In this system various directions (i.e., clockwise and anti-clock wise) in the angles of input images (different directions of same person) are compared with database pose of same person. The difficulty of this system is to compare the database pose with training set images which involves more iteration of comparisons between images of various angles. To increase the performance of the system we need to apply quality metrics for reducing the iterations and comparisons. To solve this problem we have to use Evaluation of clustering techniques (Image and cluster quality metrics) and optimizing the iterations to recognize the target pose (redundancy). In the selection of the best suitable target pose by comparing with the database pose requires preprocessing techniques.

The Evaluation of image quality is necessary in image coding for minimizing the miss representation. There are two approaches namely, subjective and objective evaluation. The subjective evaluation is divided into two categories rating scale and comparison methods. In the rating scale method, the subjective views a sequence of images under comfortable conditions and assigns each image to one of several methods. The subject evaluation method may be used in impairment scale. In the comparison method, the scale is based on a comparison with a set of images and used in a comparison scale. Several novelists found various subject evaluation tests which are complex and time consuming and there is no constructive method for performance improvement, and it is difficult to use them as part of the design process. So, we have attempted to find objective measure to evaluate the quality of the processed images. Earlier works on these lines have discussed the image quality. Some authors, Ismail et al proposed the idea of image quality metrics in statistical evaluation of image quality measures in journal of electronic imaging, 2002[5].

This paper is organized as follows: Section I deals with Ternary approach for angle orientation. The Evaluation of cluster is presented in section II. In section III, we deal with taxonomy of multivariate analysis for poses of images. Section IV is discussed about Map Reducing Algorithm. Experimental results are shown in section V. Conclusion is in section IV.

2. Ternary Approach

In mathematics, ternary approach deals with three parts
which includes relationship between them. Three parts may also be referred to as three dimensions and it is finitary relation. Cluster is a collection of datum’s that are alike to one another within the same cluster and/or unlike to the objects in other clusters. Clusters are formed based on the data. No predefined class (cluster) is formed with an unsupervised image to form it into cluster. There are three major methods in clustering namely Partition, Hierarchical and Model Based cluster.

1. Partitioning Cluster: Form arbitrary partitions and then iteratively refine them by some condition.
2. Hierarchical cluster: creating ordered breakdown with the set of data (or objects) using some condition.
3. Model-based cluster: A model is offered for each of the clusters for finding the best fit of that model to each other. For example: Fuzzy Decision Tree.

To overwhelm the above said three clusters, we introduced a novel method namely ternary cluster with the idea of ternary relation; it is the amalgamation of the above three cluster properties for angle orientation clusters. It is shown in figure A. In Ternary Cluster, Entire training image is denoted as root node i.e., I and it is divided into two groups of major similar objects called clusters I₁ and I₂ (used partitioning clustering) which are internal nodes.

Jagan Mohan et. al., discussed about the classification of clustered database images using decision tree given an Efficient K-Means Cluster Reliability on Ternary Face Recognition using Angle Oriented Approach, 2012 [9]. The similar image object groups which are rotated in the clock wise direction, belongs to the internal node, i.e., cluster I₁. Similarly that of I₂ containing the group of similar image objects which are rotated in anti-clock wise direction. Again, the cluster I₁ is grouped into three terminal nodes (segments), I₁₁, I₁₂ and I₁₃, called nested clusters (using hierarchy cluster). Likewise I₂₁, I₂₂ and I₂₃ defined as nested clusters which are three terminal nodes of cluster I₂. If \( \theta \) is the angle for rotation in each cluster, the nested cluster I₁₁ consists of images with angle 0° to 30°, I₁₂ comprises the images in between 31° to 60° and I₁₃ bears 61° to 90°. The same type of segmentation followed in internal node of I₂, the images of angle with an interval of 30° starts with 0° and ends at 90° in anti-clock wise direction (used fuzzy logic). Cluster images based on their visual content. Now we have to form ternary cluster with unsupervised images. Fuzzy sets and their corresponding membership functions have been defined by domain experts. Each of the fuzzy sets can be held as [0, 1] valued attribute, called fuzzy attribute.

3. Evaluation of Clusters

Evaluation of cluster improves the quality of clusters in view of satisfying the clustering objective. According to this objective, instances in the same cluster must be closer to each other than to instances across clusters. The most challenging issue in the cluster is to check whether it is good or not. In 1964, Bonner has suggested the notion of evaluation criteria. There are two types of evaluation criteria.
namely Internal and External [4, 10].

3.1. Internal Criteria of a Cluster

By using similarity measures, we can measure the internal quality of the cluster. Here we used the Mean Square Error [5] and Sum of the Squared Error [6]. This is used to check whether the various angles of pose images in a cluster fit to one input image or not.

Mean Square Error:

\[ \text{MSE} = \frac{1}{MN} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [x(i,j) - \hat{x}(i,j)]^2 \]

Sum of Squared Error (SSE):

\[ \text{SSE} = \sum_{k=1}^{K} \sum_{\forall x_{i} \in C_{k}} ||x_i - \mu_k||^2 \]

3.2. External Criteria of a Cluster

It is used to examine, whether the forming of clusters are matched with the given instance or not. Here we used one of the simple and transparent measure called purity measure for testing whether structure of cluster matched to the given input pose or not.

\[ \text{purity}(\Omega, C) = \frac{1}{N} \sum_{k=1}^{\max_k} |\omega_k - c_j| \]

Where \( \Omega = \{\omega_1, \omega_2, \ldots \} \) is set of clusters and \( C = \{ C_1, C_2, \ldots \} \) is set of classes.

3.3. Bootstrap Principle

After examining the cluster quality, now it is time to take cluster for finding target inference so here we used the statistical approach of bootstrap principle which is suggested by Efron et al, An Introduction to the Bootstrap, Davison and Chapman & Hall 1993, 1997[1, 3]. It is useful when size of the population is high. A Bootstrap data set is one which is created by randomly selecting n points from the training set D, with replacement. We used this technique for taking the sample data set which consists of various angles of pose images. By using this principle we randomly pick any one of the clusters to get target inference.

4. Taxonomy of Multivariate Analysis

To understand and find a target inference, it is very important to examine more than one face simultaneously from a set of poses of an image. So it is needed to consider more than one variable i.e., more than one image from various angles of images. Multivariate analysis is a graphical tool mainly used to find patterns and relationships between several variables simultaneously. Kappor et al have discussed the Multivariate analysis technique that allows the analysis of more than two variables at once [14].

4.1. Multivariate Normal Distribution

Multivariate normal distribution is a probability distribution in a multivariate analysis. Multivariate normal distribution has a mean \( \mu \) and variance-covariance matrix \( \Sigma \) of random n-vector \( X \) and is denoted as \( X \in N(\mu, \Sigma) \) and its density is given by

\[ f(x) = \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} \exp\left(\frac{-1}{2}(y - \mu)\Sigma^{-1}(y - \mu)\right) \]

The following is the very special property of multivariate normal distribution, which is used to test the independency of the random variable.

4.2. Independence

Let \( X \) is a normal random vector. The components are independent iff they are uncorrelated, i.e., \( \text{Cov}(X_i, X_j) = 0 \) then they are uncorrelated so the two components \( X_i \) and \( X_j \) are independent.

In this paper we used this extraordinary property in the following two cases:

1. We have to compare all angles of images and check whether all are belonging to one input image or not. In this case if they are not uncorrelated then all poses of images belongs to one particular image. i.e., \( \text{Cov}(X_i, X_j) \neq 0 \) and \( X_i, X_j \in C \) (\( X_i, X_j \) are from poses of images) which means that they are not independent and there is relation between these angles or poses. After succeeded from step 1, among all angles of images we have to test which angle or pose is a match with the input images. In this case we have to test the independency property between database image and the angles or poses of images. i.e., \( \text{Cov}(X_i, X_j) \neq 0 \). Here if we find any one of the angle or pose is not independent to input image, which means it is the target inference for the input image.

After step 2, there is a possibility of getting more than one image as a target inference so here we have to use the tackle outlier. This is explained in the following section.

4.3. Tackle Outlier Detection in Multivariate Analysis

In pose orientation analysis, object can be viewed in many directions. So to analyze such data for finding target object from various angles of image data recorded or sampled one of the steps is obtaining target object by observation or detection. It is very difficult to find out the best output pose results from various images. So it takes the image which is nearly matched to the database pose using tackle outlier. Tackle outlier is mainly useful when there is more than one as result and it chooses best of them which matches to the target object is introduced by Michael Thompson, 2009[15].
5. Map Reduction

Jagan Mohan et al proposed this method in similarity based query optimization on map reduce using Euler angle oriented approach, 2012[2]. It uses the two interfaces Map and Reduce, which is used for processing and generating clusters. In map phase, Researcher have processed the cluster in order to analyze clusters which consist of various angles or poses of images. Mapping is used to load the images and transform it into key-value pair to store all these angles of images [11, 12, and 13]. In the above sections 2, 3 we mentioned that there is a possibility of getting more than one as target inference. So here it is needed to use reduce phase. After mapping is finished the key-value files are shuffled and for getting the new key-value pair using Tanimatto distance with the extension of Jaccard methods. This is shown in below figure:

The hypotheses for the comparison of independent groups are
H₀: Mean values of all groups are not equal.
H₁: Mean values of two or more groups are equal.

6. Experimental Results

For angle orientation images we performed ANOVA test for Quality of the cluster, Independence test of Multivariate analysis for finding target inference and finally we tested the software reliability to increase the performance of Angle orientation images. All these are shown in below sections:

6.1. Analysis of Variance

It is a statistical tool which is used for measuring the quality. It is used to test whether the hypotheses is acceptable or not. Here various angles of images taken as the hypotheses. And checking whether all the poses of images are acceptable to one object or not.

The Table 6.1.1 value of F at 1% level with (2,42) D.F. is 5.18. We see that the calculated value of F i.e., 0.137525099 is less than the tabulated value 5.18 at 1% level. Hence, we accept the Null Hypothesis H₀ at 1% level and conclude that all the means values of three clusters are not same.
Table 6.1.2. Anova Table for Anti-Clock wise

| Source of Variation (SV) | Sum of Squares (SS) | Degrees of Freedom (d.f.) | Mean Squares (MS) | Test Statistic (F-ration of Variance) |
|--------------------------|---------------------|---------------------------|------------------|---------------------------------------|
| Between Samples          | 384929.21           | 2                         | 192464.6         |                                       |
| Within Samples           | 8166541             | 42                        | 189919.6         | \(F = \frac{192464.6}{189919.6} = 1.013400663\) |
| Total                    | 8551470.2           | 44                        | -                |                                       |

6.2. Testing the Target Inference

The following is the result after doing above all procedure which is used in section 3, 4.

![Figure 6.2.1](image-url) 22⁰ of angle of input image i.e., anti-clockwise, database image, G 29⁰ of angle of input image i.e., Clockwise.

The first figure is the target object among all the angles of images i.e., 22⁰ of 1st cluster in anti-clockwise and third figure 29⁰ of 1st cluster in clockwise compared with data base pose shown in figure G.

6.3. Software Reliability Test on Angle Oriented Images

To conduct cluster reliability test on angle oriented images like anti-clock wise and clock wise images. The Probability of failure is calculated by angle oriented images, testing a random sample of all available input images.

\[
\text{Probability} = \frac{\text{Number of failing case}}{\text{Total Number of cases under considerations}}
\]

7. Conclusion

This paper mainly deals with finding the target inference among various angles of poses. Here we group the images into cluster based on ternary relation and introduced the ternary cluster and then the evaluation of the cluster. We have various poses of images. So we need to examine more than one image. For this we used multivariate analysis. For finding target inference we have used bootstrapping principle which chooses the cluster and tests the independence between various angles of images. We used the independency of multivariate normal distribution. If we found more than one object matched to the input image then we used tackle outlier and map reduction algorithm. Experimental results are produced with the help of ANOVA.

REFERENCES

[1] Davison, A. C., Hinkley, D. V., Bootstrap Methods and their Application, Cambridge University Press 1997.
[2] Dileep Kumar, K and Jagan Mohan, R.N.V. and Vamsidhar, Y., Similarity based Query Optimization on Map Reduce using Euler Angle Oriented Approach, International Journal of Scientific & Engineering Research, Volume 3, Issue 8, and August-2012 ISSN 2229-5518.
[3] Efron, B. and Tibshirani R. J.: An Introduction to the Bootstrap, Chapman & Hall 1993.
[4] Guha, S. et al., “CURE: An Efficient Clustering Algorithm for large Databases,” In Proceedings of the ACM SIGMOD Conference, 1998.
[5] Ismail et al proposed the idea of image quality metrics in statistical evaluation of image quality measures in journal of electronic imaging, 2002.
[6] LiorRokach and Oded Maimon in the Data Mining and Knowledge Discovery Handbook.
[7] Jagan Mohan R.N.V. and Subbarao. R., Increasing the Reliability of Angle Oriented Face Recognition using DCT, published In the Proceedings of International Congress on Productivity, Quality, Reliability, Optimization and Modeling, Allied Publishers, 2011.
[8] Jagan Mohan, R.N.V and Subba Rao, R and Raja Sekhara Rao, K., Similarity of Inference Face Matching On Angle Oriented Face Recognition, Computer Engineering and Intelligent Systems ISSN 2222-1719 (Paper) ISSN 2222-2863 (Online),Vol 3, No.2, 2012, www.iiste.org.
[9] Jagan Mohan, R.N.V and Subba Rao, R and Raja Sekhara Rao, K.,Efficient K-Means Cluster Reliability on Ternary Face Recognition Using Angle Oriented Approach, International Journal of Informatics and Communication Technology (IJ-ICT)Vol,2, No.1, January 2013, pp. 38–45 ISSN: 2252-8776.
[10] Koutroubas, K. and Theodoridis, S, Pattern Recognition, Academic Press, 1999.
[11] Rajasekhar B. and Sunil Kumar B. and Rajesh Vibhudi and Rama Krishna B.V., Quality of Cluster Index based on Study of Decision Tree, published in International Journal of Research in Computer Science eISSN 2249-8265, Volume 2 Issue 1, pp. 39-43, 2011.
[12] Michael J. Swain and Dana H. Ballard., Color Indexing published in International Journal of Computer Vision, 7, 1, 11-32, 1991.
[13] Nezamabadi-pour, H. and Sarayazdi, S., Object Based Image Indexing and Retrieval in DCT Domain using Clustering Techniques, Transactions on Engineering, Computing and Technology, 3, 2004.
[14] V.K.Kapoor and S.C.Gupta, Fundamentals of Mathematical Statistics, Sultan Chand & Sons.
[15] Thompson Michael, Rogues and Suspects: How to Tackle Outliers, Analytical Methods Committee, ISSN 1757-5958,AMCTB No 39, April 2009.