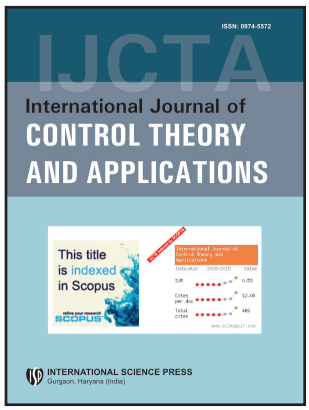

International Journal of Control Theory and Applications

ISSN : 0974-5572

Volume 10 • Number 16 • 2017

Observation of a New Solution for Handling Biometric Data by Experimenting with Distance Similarity Measures

R. Ramya[a] and T. Sasikala[b]

[a]*Research Scholar, Department of Computer Science and Engineering, Sathyabama University, Chennai, India. Email: ramya.rav@gmail.com*
[b]*Research Supervisor, Principal, JEPPIAAR SRR Engineering College, Chennai, India. Email: sasi_madhu2k2@yahoo.co.in*

Abstract: Effective handling of biometric data becomes a mandatory part in any biometric processing system. Distance Similarity measures play a major role in solving pattern matching problems. Comparing their efficiencies in almost similar images has been attempted. Though different distance similarity measures are available, few of them take their importance according to their specific area of need. This paper has made an attempt to analyze the various similarity distance measures for comparing biometric images. It has been observed that Manhattan Distance method shows a huge difference and outperforming with the other well known methods like hamming and Euclidean. The efficiency can be seen in the results where we have taken almost similar biometric images. Manhattan distance shows high distance showing that it was able to differentiate even able almost similar images than the other methods which were not able to do it.

Keywords: Similarity Distance Measures, Biometrics, Manhattan.

1. INTRODUCTION

Biometrics considers programmed methods of identifying a person on the basis of computable biological and behavioural characteristics. The features of human beings taken into account are fingerprint, face, iris, retinal, hand geometry, signature and voice. Fingerprint recognition [1, 2, 3, 4] takes into account the patterns found on the tip of finger. Traditional policy method of matching minutiae is used in some, while others make use of pattern-matching devices. There are a wide range of fingerprint devices than for any other biometric features. Due to the fall in prices of fingerprint devices fingerprint verification has gained more acceptance of usage. Face identification scrutinize the characteristics of the face. A digital camera is required to build an image of the face of the user for identification. Since the facial scanning requires additional peripheral device it gives a little discomfort for the users to comfortably afford it. However, few companies have provided the facial database for ease of use.

Biometrics based on iris, analyses features that are found in the tissue that surrounds the pupil. Iris scanning, however, uses a camera element avoids the close connection amid the user and the reader. Iris biometrics works well in recognition mode. The issues of system integration and easy way of usage comparatively with fingerprint

acts as a setback but improvements can be expected when new products emerge in the market. Retina based biometric analyses the blood vessels located at the rear of the eyeball. Retinal scanning requires the user to glance into a container and give full attention on a given point. It becomes inconvenient when people wear glasses or contact lenses. Despite of its accuracy the way of use is unfortunately not welcomed by the users.

Hand Geometry involves the evaluation of the shape of hand. Precision can be achieved by tuning and configuring the performance that can provide a wide variety of applications. Some of the areas where it is used are in recording attendance and time. Since it can be easily coupled into other systems and processes it has been used successfully in many biometric projects.

Signature verification [5] analyzes the manner in which a person signs his/her name. The factors such as the speed in which the person signs, the velocity, and the pressure used are considered as important as manner in which it is signed. It can be used in transactions and other areas where there is a need for high security. The devices used for signature verification provide reasonable accuracy.

Voice Recognition [6, 7] is the recognition of an individual from their voice. It is also called Speaker recognition. Recognizing the speaker simplifies the work of interpreting the speech in systems that have been taught on particular person's voices. It can be used to verify or validate the personality of a speaker as a component of a security process.

All these biometrics types share similar fundamentals. The features are scanned and the key features are determined. It is then changed into a digital string, and added into the database. If a match is found, the system indicates the likelihood of the genuinity of the person who he/she is claimed to be.

This technology has become the groundwork for a wide array of extremely safe recognition and individual authentication solutions. It has taken its giant leap due to the high increase of fraud in transactions and breaches in security. Applications areas where it has its huge demand are federal, central and state government, law enforcement, counter-terrorism, access control, driver licenses, airport, child recovery, smart cards and social networks. Emphasis for different applications depends upon the below criteria which determines what kind of biometrics should be used for a particular application:

- precision
- swiftness
- intrusiveness
- ecological tolerance
- liability to being deceived
- price

A successful working of a biometric system highly relies on the performance of the similarity measure function [8]. Similarity is the evaluation of like objects to each other. Choosing the best distance function among all the various similarity measure functions becomes an essential part when applying them over the set of data to be used. An algorithm applied using the most suitable similarity functions gains an overall good performance than selecting a bad similarity measure function. Similarity echoes the potency of correlation between two data items while dissimilarity measures the deviation between two data items [9][10].

1.1. Similarity Distance Measure Functions

To measure how close a sample data resembles a template data a fresh sample has to be taken. A characteristic based on statistics of the data distribution is taken into account by a good similarity measure function [11].

Some of the well established measures are, Euclidean Distance, Mahalanoblis Distance, and Manhattan Distance.

1.1.1. Euclidean Distance

It is a known and a famous similarity distance function. It is the sum of the squared distances of two vector values (x_i, y_i)[12].

$$d_{\mathrm{E}} = \sqrt{\sum_{j=1}^{m} (x_i - y_i)^2} \tag{1}$$

Euclidean Distance is a variant to the dimensionality of the vectors. [13]

1.1.2. Mahalanobis Distance

It can be given as:

$$\mathrm{D_M}(x) = \sqrt{(x-\mu)\mathrm{T}\sum^{-1}(x-\mu)} \tag{2}$$

with $\mu = (\mu_1, \mu_2, \mu_3, ..., \mu_p)$ representing mean and Σ for covariance matrix for a multivariate vector $x = (x_1, x_2, x_3, ... x_p)$ Mahalanobis distance is also a measure of dissimilarity between $\vec{x}$ and $\vec{y}$ two random vectors of the same distribution with the covariance matrix Σ:

$$d(\vec{x}, \vec{y}) = \sqrt{(\vec{x}-\vec{y})\mathrm{T}\sum^{-1}(\vec{x}-\vec{y})} \tag{3}$$

If the covariance matrix is the identity matrix, then it is the same as Euclidean distance. If the covariance matrix is diagonal, then it is called normalized Euclidean distance:

$$d(\vec{x}, \vec{y}) = \sqrt{\sum_{i=1}^{p} \frac{(x_i - y_i)^2}{\sigma_i^2}} \tag{4}$$

where, σ_i is the standard deviation of the x_i over the sample set. Mahalanobis distance is not dependent on the scale of measurements.[14]

1.1.3. Manhattan Distance

It is the sum of the lengths of the projections of the line segment between the points onto the coordinate axes. Also, Manhattan distance is the sum of the absolute differences of the two vector values (x_i, y_i)

$$d_{\mathrm{M}} = \sum_{i=1}^{n} |x_i - y_i| \tag{5}$$

2. LITERATURE REVIEW

Pagnin et. al., [15] who used a search at the center and have tried to show that when the process of matching is done using distances like existing distance Hamming and Euclidean, then the stored data(the template to be referred) is leaked and favors the attacker. It finally helps the attacker to recuperate the biometric template.

Erkin et. al., [16] for the first time have taken into account the trouble of preserving the privacy in biometric identification. A privacy preserving face recognition system was proposed. It is based eigenface approach which was introduced by Turk et. al., [17, 18]. They calculated the Euclidean distance between face image feature vector from client and server's face image database. In [19, 20], the key idea is to get the closest match for

an input biometric data based on the Euclidean distance. Here, the input biometric data is encrypted using the public key published by the client and late sent to the server. On the server side each biometric data uses an additive homomorphic encryption using the same public key. It then uses the Euclidean distances to find the closest match.

Blanton and Gasti [21] have proposed a protocol for iris codes. It proved to be secure since it was based on additive homomorphic encryption and garbled circuits. Further the similarity between iris codes is measured by Hamming distance. In case of Finger codes they used the Euclidean distance for recognition of fingerprint.

Osadchy et. al., [22, 23] introduced a privacy-preserving face identification system. The method is based on additive homomorphic encryption and unaware transfer. The similarity of binary feature vectors are measured using Hamming distances.

In [24], the authors introduced a new measure of distance between projected vertex sets of intrinsic graphs to mitigate the effect of the differences between views and preserve the intrinsic graphs. This distance is defined as the weighted sum of squared Euclidean distances between every cross-view data pair in two graph embedding models. Having sets of multi-view data, MiLDA aims to find a common subspace of higher discriminability between classes.

The transformed feature vectors in the common subspace are classified using a nearest neighbor classifier. Most recently, deep learning has attracted the interest of many researchers, such that a deep network has become the model of choice for unconstrained face recognition. Various deep learning architectures such as deep neural networks, convolutional deep neural networks, deep belief networks and recurrent neural networks have been applied to fields like computer vision, automatic speech recognition, natural language processing, audio recognition and bioinformatics.

3. SYSTEM MODEL

To evaluate the binary vector similarity measures, a fingerprint database is considered. Daugman has proposed the degrees of freedom based on mismatch of a biometric trait.[25].The biometric verification is the one that considers two randomly selected biometric data and checks whether the two samples is of the same person or of two different persons. Figure 1 depicts the Fingerprint verification Model. Initially features are extracted from the fingerprint biometric image x and y. In [25] Daughman has used the Hamming distance method, but in our proposed method Manhattan Distance measure has been used.

The distributions are assumed to be normal and the FAR (False Acceptance Rate) and the FRR (False Rejection Rate) are minimized. The result which was taken with many different samples of fingerprint images shows that Manhattan Distance measure gives the most accurate similarity.

4. RESULT AND DISCUSSION

The analysis has been done considering the distances like Hamming, Euclidean, Weighted Euclidean, Mahalanobis and Manhattan. In the below table it can be seen that Manhattan distance method shows the highest distance comparatively with other methods. The more the distance the less is the similarity. The database which we used had the maximum of dissimilar biometric fingerprint samples. Instead of identifying the other biometric samples as not same, the widely used methods like Hamming, Euclidean, Weighted Euclidean show that they are similar in most of the cases. This can be seen in the Table 1 given below. In none of the samples Manhattan is showing the FAR and it outperforms than the other similarity distance measures.

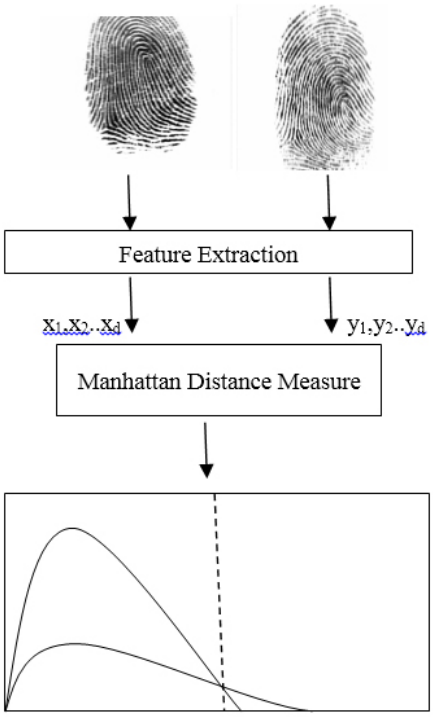

Figure 1: Fingerprint Verification Model

Table 1
Comparison of Distance Measure Functions with Biometric Sample Dataset

Distance Method	*Biometric Sample 1*	*Biometric Sample 2*	*Biometric Sample 3*	*Biometric Sample 4*	*Biometric Sample 5*
Manhattan	22502	9416	25884	13144	16770
Hamming	2.540039e+02	2.552930e+02	2.540273e+02	2.554453e+02	2.554258e+02
Euclidean	1.115745e+04	2.882039e+03	8.293131e+03	3.260512e+03	5.708954e+03
Weighted Euclidean	7.889508e+03	2.037909e+03	5.864129e+03	2.305530e+03	4.036840e+03
Mahalanobis	1.411291e+02	3.769138e+01	1.189785e+02	4.544620e+01	7.778393e+01

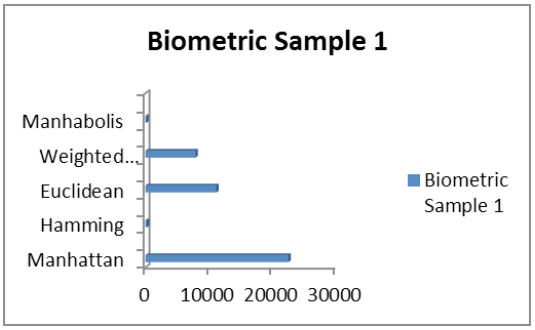

Figure 2: Distance Similarity chart for Biometric Sample 1

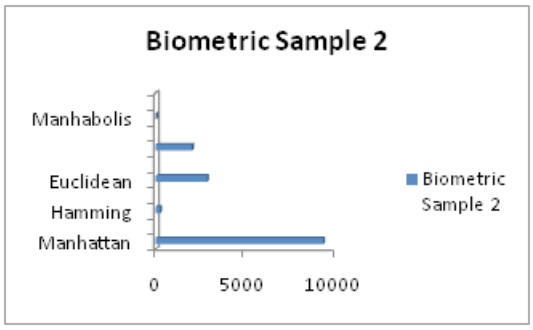

Figure 3: Distance Similarity chart for Biometric Sample 2

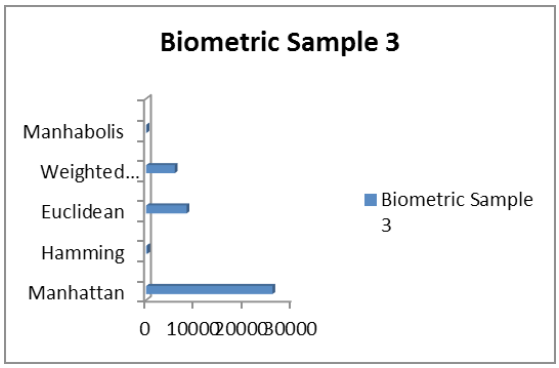

Figure 4: Distance Similarity chart for Biometric Sample 3

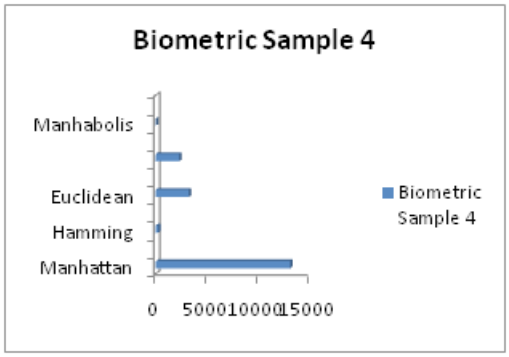

Figure 5: Distance Similarity chart for Biometric Sample 4

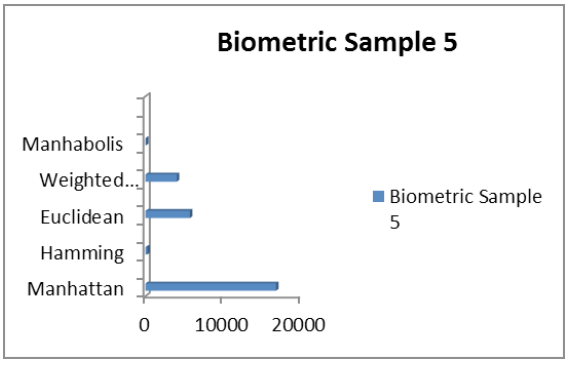

Figure 6: Distance Similarity chart for Biometric Sample 5

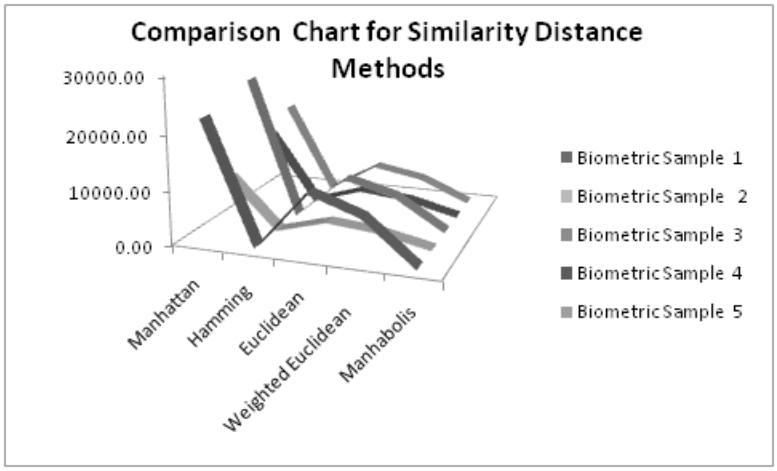

Figure 7: Comparison Chart of Similarity Distance Measures

Figure 7 shows the elevated levels of the Manhattan Distance similarity method. Though currently Hamming has been widely used, this paper has proved that Manhattan method performs more precisely than the other methods considered.

CONCLUSION

Choosing the similarity measure plays an extremely important part since distance measures are used to retrieve similar images from the database for a given query. This paper has experimented with the distance similarity functions like Hamming, Euclidean, Weighted Euclidean, Mahalanobis tested with the fingerprint database. It has been shown that Manhattan is the best similarity measure showing the dissimilar biometric sample comparatively well enough than the widely used other similarity measures. In all the samples experimented, Manhattan has proved itself as the best distance similarity measure. These promising results encourages to the use of Manhattan method rather than the other similarity measures.

REFERENCES

[1] Yampolskiy, Roman V and Govindaraju, Venu,"Use of behavioral biometrics in intrusion detection and online gaming", Defense and Security Symposium. International Society for Optics and Photonics, 2006, pp. 62020U-62020U

[2] M. Neuhaus and H. Bunke, "An error-tolerant approximate matching algorithm for attributed planar graphs and its application to fingerprint classification", Proc. Joint IAPR Int. Workshops Structural, Syntactic, and Statistical Pattern Recognition, 2004, pp. 180 -189

[3] Barral, Claude, Jean-Sébastien Coron, and David Naccache. "Externalized fingerprint matching." Biometric Authentication. Springer Berlin Heidelberg, 2004,pp. 309-315

[4] Schwaighofer, Anton. "Sorting it out: Machine learning and fingerprints."Special Issue on Foundations of Information Processing of TELEMATIK 1 (2002),pp.18-20.

[5] H. Lei, S. Palla and V. Govindaraju, "ER2:An Intuitive Similarity Measure for On-Line Signature Verification", IWFHR '04: Proceedings of the Ninth International Workshop on Frontiers in Handwriting Recognition (IWFHR'04), IEEE Computer Society, 2004, pp. 191--195.

[6] T. Kinnunen and I. ainen, "Class-discriminative weighted distortion measure for VQ-based speaker identification: In Proc. Joint IAPR International Workshop on Statistical Pattern Recognition, Windsor, Canada, August 6-9, 2002, pp. 681-688.

[7] O. Mut and M. Göktürk, "Improved Weighted Matching for Speaker Recognition", The Third World Enformatika Conference, WEC'05, Istanbul, Turkey, April 27-29, 2005, pp. 229-231.

[8] Yampolskiy, Roman V., and Venu Govindaraju. "Similarity measure functions for strategy-based biometrics." International Conference on Signal Processing (ICSP 2006), Vienna, Austria. 2006.

[9] Huang, Anna. "Similarity measures for text document clustering."Proceedings of the sixth new zealand computer science research student conference (NZCSRSC2008), Christchurch, New Zealand. 2008,pp 49-56

[10] Taghva, Kazem, and Rushikesh Veni. "Effects of similarity metrics on document clustering." Information Technology: New Generations (ITNG), 2010 Seventh International Conference on. IEEE, 2010,pp.222-226

[11] Lee, Kwanyong, and Hyeyoung Park. "A new similarity measure based on intraclass statistics for biometric systems." ETRI journal 25.5 (2003): 401-406.

[12] Sturn A. Cluster analysis for large scale gene expression studies (Doctoral dissertation, Graz University of Technology).

[13] S. Yang and I. Verbauwhede, A Secure Fingerprint Matching Technique, In Proc. ACM Workshop on Biometrics Methods and Applications, 2003, pp. 89-94.

[14] Mahalanobis distance http://en.wikipedia.org/wiki/Mahalanobis_distance Date accessed: 25/07/2015.

[15] E. Pagnin, C. Dimitrakakis, A. Abidin, and A. Mitrokotsa, "On the leakage of information in biometric authentication," in Progress in Cryptology– INDOCRYPT 2014. Springer, 2014, pp. 265–280.

[16] Z. Erkin, M. Franz, J. Guajardo, S. Katzenbeisser, I. Lagendijk, and T. Toft, "Privacy-preserving face recognition," in Privacy Enhancing Technologies. Springer, 2009, pp. 235–253.

[17] Turk, Matthew, and Alex Pentland. "Eigenfaces for recognition." Journal of cognitive neuroscience 3.1 (1991): 71-86

[18] M. Turk and A. Pentland, "Face recognition using eigenfaces," in IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1991, pp. 586–591

[19] Sadeghi, Ahmad-Reza, Thomas Schneider, and Immo Wehrenberg. "Efficient privacy-preserving face recognition." International Conference on Information Security and Cryptology. Springer Berlin Heidelberg, 2009, pp. 229-244

[20] Evans, David, et. al., "Efficient privacy-preserving biometric identification."Proceedings of the 17th conference Network and Distributed System Security Symposium, NDSS. 2011.

[21] M. Blanton and P. Gasti, "Secure and efficient protocols for iris and fingerprint identification," in Computer Security–ESORICS. Springer, 2011, pp. 190–209.

[22] M. Osadchy, B. Pinkas, A. Jarrous, and B. Moskovich, "SCiFI - a system for secure face identification," in IEEE Symposium on Security and Privacy (SP), 2010, pp. 239–254.

[23] Osadchy, Margarita, et. al., "System for secure face identification (SCIFI) and methods useful in conjunction therewith." U.S. Patent No. 8,542,886. 24 Sep. 2013.

[24] Yan, Shuicheng, et. al., "Graph embedding and extensions: A general framework for dimensionality reduction." IEEE transactions on pattern analysis and machine intelligence 29.1 (2007).

[25] Daugman, John. "Evolving methods in iris recognition." IEEE International Conference on Biometrics: Theory, Applications, and Systems,(BTAS07),(online). http://www. cse. nd. edu/BTAS_07/John_Daugman_BTAS. pdf. Accessed Sept. 2016.