The International Journal of Engineering and Science (IJES)
|| Volume || 7 || Issue || 7 Ver. II|| Pages || PP 08-16 || 2018 ||
ISSN (e): 2319 – 1813 ISSN (p): 23-19 – 1805

Implementation On Face Recognition Techniques

Ali Mohammed Dige, M. Abdulkadir
Department of Electrical and electronics Engineering, University of Maiduguri, Nigeria.
Department of Electrical and electronics Engineering, University of Maiduguri, Nigeria.
Corresponding Author: Ali Mohammed Dige

ABSTRACT
The term biometrics comes from the combination of the Greek words 'bios', which Means life, and 'metric', which in its turn means measuring. Biometric technologies aim primarily at identifying a person's unique features, be those physiological or behavioural. Failure rate is a critical consideration in the configuration and day to day running of a biometric system. Two types of failure rate must be considered; false acceptance rates and false rejection rates. PCA (principal component analysis) model system was used for the face recognition. The system was successfully designed using MATLAB. The face recognition was implemented and the system performance was tested using ORL database (face sets) validation, generalization and rejection testing methods. The system performance was good and the efficiency of the system was proven by plotting a graph. The system was design and was successfully tested and it has proven to be efficiently, although under bad condition the system does its best to provide the right top five ranking image.
KEYWORDS: *face recognition techniques and big data.*

Date of Submission: 04-07-2018 Date of acceptance: 19-07-2018

I. INTRODUCTION

Biometrics is a method refers to as a means of authentication techniques that is deeply based on the measurable physical characteristics, humans have used biological traits or physical characteristics such as face, gait (a manner of running, stepping or walking) and voice for thousands of years to identify one another. Each individual has its own biological traits that are uniquely different from each other including twins, examples of these physical trait that are used in biometric system are; fingerprint, retina, DNA, voice, iris and face etc. In Biometric identification, fingerprint is used for thumbprint application, voice for a voice identification application and face for a facial recognition application. Also biometric identification refers to identifying an individual based on his/her distinguishing physiological and/or behavioral characteristics. As these characteristics are distinctive to each and every person, biometric application is more reliable and capable than the traditional token based and knowledge based technologies differentiating between an authorized and a fraudulent person (A.K. Jain et al, 2003).

Big data is a term that refers to data sets or combination of data sets whose size (volume), complexity (variability) and rate of growth (velocity) made them difficult to be captured, processed or analyzed by conventional technologies and tools. By volume refers to amount of data stored. Can be from megabytes to petabytes. Variety means different type of data and source of data. And velocity refers to the speed of data processing.

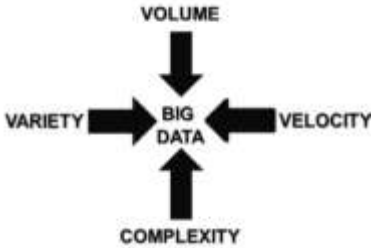


Fig. 1 Big data problem

Physical challenged non registerble users: during the rollout and use of a system based on biometrics, it is inevitable that some users will be found who cannot register for a given system due to their physical characteristics. Thus, a secondary authentication mechanism is always needed for biometric systems. Circumvention: when rolling out biometric system, the view of a system as a chain is maintained. Failure rate: failure rate is a critical consideration in the configuration and day to day running of a biometric system. Two types of failure rate must be considered; false acceptance rates and false rejection rates.

How to overcome the challenges These are as follows; Liveness detection : This refers to the ability of the system to distinguish between a sample feature provided by a live human and a copy of a feature provided by an artifact. Steganographic and watermarking techniques: Steganographic means secret communication. Its involve hiding critical information is unsuspected carrier data. The technique can be suitable for transferring critical biometric information from a client to sever.

Research on biometric method has gained renewed attention in recent years brought on by an increase in security concerns. The recent world attitude towards terrorism has influence people and their government to take action and be more proactive in security issues. Many biometric techniques have been developed and are improve with the most successful being applied in everyday law enforcement and security applications. In 1903 it was discovered that some people share the same measurement and physical characteristic (Canton, 2003).
Chengjum Liu and Henry Wechsler (2002) presented new coding scheme linear discriminant models for indexing and retrieval from large image databases. Joss Beveridge et al (2003) provide the PCA and LDA algorithms for face recognition. Kyeungnam kim (1998) has proposed PCA to reduce the large dimensionality of the data space. Face recognition using principal component analysis (PCA) the original face is reconstructed with some error, since the dimensionality of the image space is much larger than that of face space.

Jun-Ying et al (2005) have combined the characteristics of PCA with LDA. This improved method is based on normalization of within-class average face image, which has the advantages of enlarging classification distance between different-class samples. El-Bakery (2007) has proposed a new PCA implementation for last face detection based on the cross-correlation in the frequency domain between the input image and eigenvectors.

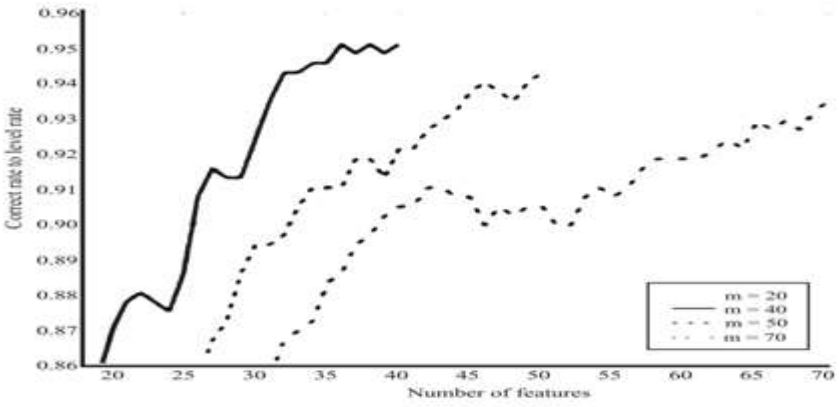
Wangmeng Zuo et al (2006) have described in combination of two novel LDA-based methods for face recognition. The combination of the two methods which performed LDA on distinctly different subspace and this may be effective in further improving the recognition performance. Moshe Butman and Jacob Goldberger (2008) have introduced a face recognition algorithm in "face recognition using classification-based linear projections" based on linear subspace projection.

II. MATERIAL AND METHOD

Biometrics refers to measurements of properties of the human body and behavior. These properties are identifying of individuals and are therefore fit for identification. Biometrics may also be used for medical and forensic purposes, but this paper focuses on applications within identity management. This paper looks at the latter application of biometrics. Compared to other forms of authentication ("something you know", "something you have"), biometrics has important advantages. A biometric characteristic is generally more difficult to spoof or copy, most individuals already have suitable characteristics and they carry it always with them. Identifying biometric characteristics are also referred to as modes and may be hard (high degree of identification) or soft (to a lesser extent identifying), as well as physiologically (measured on the body) or behavioral (measured on behavior). Multimodal biometrics combine measurements of more than one attribute in biometric applications (Sanjekar, P. S et al, 2013).

Independent component analysis

ICA is a statistical technique that is used in unfolding factors that are hidden in an underlie collections of signal or measurements and random variables. ICA at first was an effective technique for blind source separation (P.Comon et al, 1994) it was Bartlett and Sejnowski that came up with a technique to used ICA for face recognitions (M. Bartleet and T. Sejnowski, 1997).



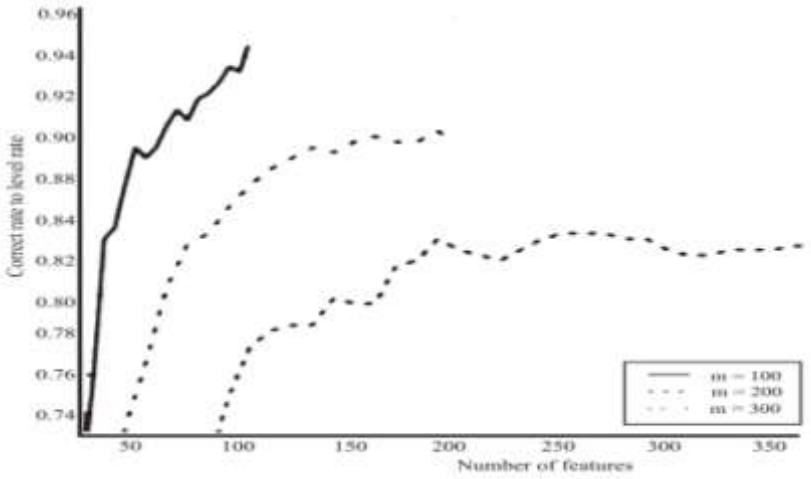


Fig. 2 Comparative performance of ICA as a discriminant criterion

Linear discriminant analysis of principal component

These types of facial recognition system involves in combining two different types of facial recognition algorithm to make a better one, these facial recognition algorithm are PCA and LDA. The PCA is used to reduce the dimension of the data while retaining the quality of the data set and the LDA technique to obtain the linear classifier. The notion behind adding PCA and LDA is to increase the efficiency capacity of LDA when involves in few sample per class. Both PCA and LDA have been used on their own for face recognition (P. Comon, 1994).

Fig.3 PCA+LDA bases

Principal component analysis

Principal component analysis (PCA) is a statistical method used in a dimensionality reduction of data while retaining quality of the data set. As for PCA because is a statistical method for handling and analysis data, a PCA-based face recognition algorithm requires its own set of algorithm structure that support it. Implementing the correct support structure depend on a lot critical design decision, with each decision on its own has the potential of affecting overall performance of the face recognition algorithm (internet 2016). PCA-based face recognition system consists of normalization, project and recognition modules.

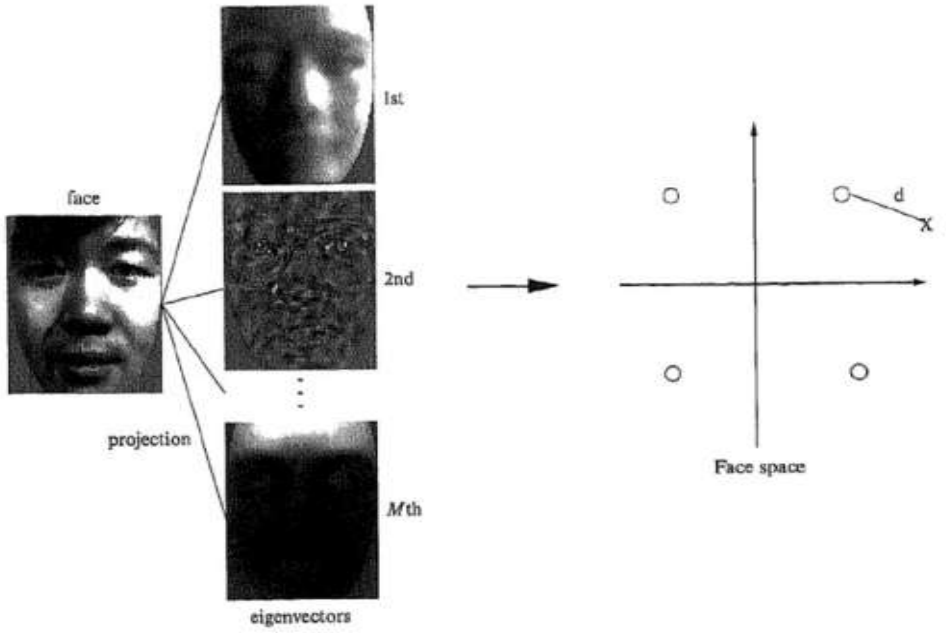


Fig.4 PCA face space

The face recognition application was chosen due to circumstance that the biometric application is based on Big Data and for the Big Data analysis, Hadoop system was used. PCA method was use for the face recognition because it's one on the most preferred holistic colour-based techniques in the world and it also answered the question that people are always asking; what features of the face is the system using to recognised face. PCA Eigenfaces state that these features are nose, eyes and mouth etc. There are lots of ways to implement PCA system and in this project, the PCA system was design using MATLAB software. The code for the PCA system was initially obtains in an openCV. OpenCV is an online library designed for computational efficiency and with a strong focus on real time application, is released under a BSD (Berkeley software distribution) license. BSD licenses are a family of permissive software licences.

III. RESULTS AND DISCUSSION

Generalisation Testing is a testing method where a probe in the probes set is compute against the gallery sets; in other words, this is a normal face recognition that is occurring between the query sets and gallery sets. A generalisation test was carried out between the probes sets (1-5 to 40-1) and gallery sets (1-6 to 40-10) and the Euclidean distances can be seen below in figure 2, the yellow highlighted Euclidean distances was identify as the correct match for the query image which was as expected, due to factors of environment, the system sometimes cannot perform efficiently. In figure 1 is an example of generalisation testing method that was carried.

Rank Number	TAR%	FAR%	FRR%
1	86.5	67.5	13.5
2	89	55	11
3	90.5	47.5	9.5
4	91.5	42.5	9.5
5	94.5	27.5	5.5
6	95	25	5
7	95.5	22.5	4.5
8	96.5	17.5	3.5
9	96.5	17.5	3.5
10	96.5	17.5	3.5

Table 1

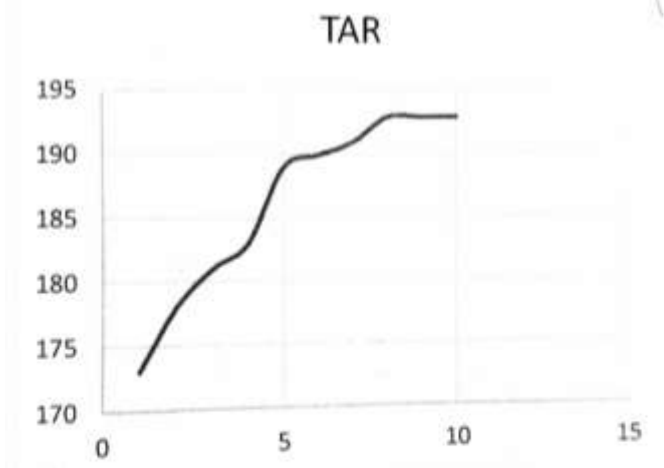


Fig.1 TAR

		Subject 1					Subject 2					Subject 3				
		1-1	1-2	1-3	1-4	1-5	2-1	2-2	2-3	2-4	2-5	3-1	3-2	3-3	3-4	3-5
		1.0e+16 *	1.0e+16 *	1.0e+16 *	1.0e+16 *	1.0e+16 *	1.0e+16 *	1.0e+16 *	1.0e+16 *	1.0e+16 *	1.0e+16 *	1.0e+16 *	1.0e+16 *	1.0e+16 *	1.0e+16 *	1.0e+16 *
subject 1	1-6	0.1995	0.2838	0.1328	0.3984	0.3810	0.2078	0.1924	0.2102	0.2257	0.1611	0.3460	0.3909	0.3360	0.3186	0.3699
	1-7	0.0758	0.1230	0.0676	0.1854	0.1587	0.2942	0.3075	0.2902	0.3386	0.2454	0.5003	0.5392	0.4681	0.3836	0.3958
	1-8	0.2290	0.2955	0.1664	0.1379	0.1049	0.3770	0.4490	0.3763	0.4701	0.3900	0.5917	0.5969	0.4929	0.5190	0.5424
	1-9	0.1726	0.2517	0.2564	0.0731	0.0789	0.5369	0.6172	0.5553	0.6471	0.5245	0.6611	0.6273	0.5414	0.6341	0.6667
	1-10	0.1638	0.1980	0.1347	0.1412	0.0972	0.3458	0.3966	0.3480	0.4247	0.3399	0.4716	0.4736	0.3877	0.4289	0.4691
subject 2	2-6	0.5511	0.5878	0.3715	0.8789	0.8218	0.0562	0.0174	0.0528	0.0188	0.0358	0.2428	0.3030	0.3006	0.2140	0.2417
	2-7	0.3887	0.4298	0.2258	0.5537	0.5133	0.0362	0.0667	0.0235	0.0647	0.0586	0.2732	0.3281	0.2896	0.1796	0.1800
	2-8	0.4249	0.4759	0.2471	0.6291	0.5839	0.0144	0.0250	0.0035	0.0249	0.0271	0.2518	0.3075	0.2756	0.1796	0.1923
	2-9	0.6031	0.5834	0.3612	0.8979	0.8029	0.0774	0.0482	0.0582	0.0370	0.0704	0.2948	0.3776	0.3653	0.1985	0.1985
	2-10	0.3529	0.3874	0.1984	0.5545	0.5069	0.0285	0.0399	0.0141	0.0430	0.0309	0.2650	0.3230	0.2888	0.1711	0.1781
	3-6	0.5099	0.6223	0.5019	0.6558	0.6969	0.2408	0.2716	0.2619	0.2775	0.2541	0.0249	0.0126	0.0000	0.1412	0.2063
	3-7	0.4344	0.4695	0.3453	0.6605	0.6442	0.1864	0.1910	0.1725	0.1968	0.1788	0.0935	0.1611	0.1412	0.0000	0.0181
[illegible]	3-8	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Fig. 2 Euclidean distances

0.0887	0.1397	0.1551	0.6980	0.6253	0.1544	0.4302	0.2351	0.2542	0.1920	0.5784	0.4915	0.5094	0.4726
0.3674	0.2865	0.2879	0.9833	0.9014	0.2853	0.4065	0.2985	0.1947	0.3097	0.5603	0.4396	0.3819	0.4517
0.1586	0.2218	0.2698	1.0809	0.9853	0.2379	0.3615	0.2637	0.1889	0.2565	0.8040	0.6960	0.7214	0.6940
0.2980	0.2642	0.2711	0.9276	0.8453	0.2515	0.4132	0.2822	0.2126	0.2798	0.5613	0.4426	0.4062	0.4452
0.4380	0.3277	0.3399	1.1272	1.0356	0.3212	0.3380	0.2931	0.1685	0.3305	0.6062	0.4874	0.4486	0.5201
0.0465	0.1220	0.1490	0.5870	0.5178	0.1377	0.5011	0.2320	0.3652	0.1663	0.5789	0.5056	0.5914	0.4695
0.3036	0.0352	0.0411	0.5180	0.4460	0.1154	0.3385	0.1239	0.2996	0.1082	0.2642	0.2576	0.3491	0.2836
0.2508	0.0851	0.0573	0.4943	0.4370	0.1669	0.4617	0.2430	0.3323	0.2015	0.3475	0.2978	0.2606	0.2909
[illegible]	0.0622	0.0440	[illegible]	[illegible]	0.1607	0.4843	0.2357	0.3910	0.1843	0.2935	0.2636	0.2898	0.2551
0.3145	0.0548	0.0819	0.8619	0.7667	0.1401	0.1934	0.0950	0.1041	0.1163	0.4626	0.4362	0.4793	0.4792
0.7238	0.4075	0.4949	1.6515	1.5145	0.4710	0.0338	0.2739	0.1583	0.3988	0.8766	0.8574	0.9305	0.9746
0.7170	0.3664	0.4435	1.5102	1.3818	0.4223	0.0107	0.2373	0.1860	0.3572	0.7572	0.7451	0.8191	0.8598
0.3840	0.1827	0.2183	1.0746	0.9745	[illegible]	0.1850	[illegible]	0.0066	[illegible]	0.5749	0.5147	0.5396	0.5648
0.7131	0.3544	0.4240	1.4840	1.3585	0.3955	0.0100	0.2076	0.1534	0.3288	0.7344	0.7225	0.8034	0.8370
0.7159	0.3581	0.4401	1.5283	1.3967	0.4057	0.0257	0.2062	0.1412	0.3272	0.7688	0.7617	0.8660	0.8783
0.7992	0.3451	0.2656	0.3364	0.3078	0.3404	0.7152	0.3968	0.6521	0.3651	0.0676	0.0580	0.0328	0.0672
0.7729	0.3329	0.2622	0.2608	0.2300	0.3240	0.7638	0.3897	0.7094	0.3432	0.0291	0.0332	0.0768	[illegible]
0.7681	0.3660	0.2762	0.3758	0.3510	0.3504	0.7300	0.4157	0.6276	0.3865	0.1029	0.0635	0.0065	0.0678
0.8195	0.3252	0.2667	0.3762	0.3351	0.3027	0.6297	0.3187	0.6104	0.3034	0.0218	0.0300	0.0908	0.0600
0.9097	0.3667	0.2926	0.5146	0.4734	0.3780	0.5892	0.3669	0.5696	0.3784	0.0865	0.0845	0.0612	0.1297

Fig.3 Euclidean distances

DOI:10.9790/1813-0707020816

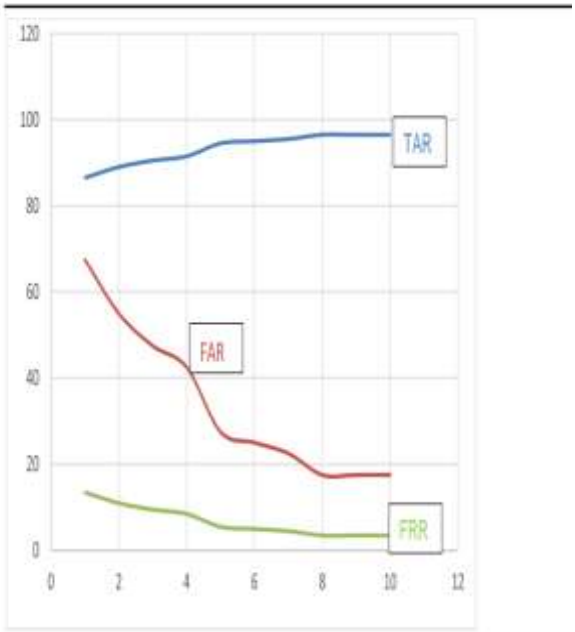


Figure 4 TAR FRR FAR

Validation testing is a testing method that occurs between the same sets of gallery or query sets. This is a gallery set (1-40) computing against gallery set (1-40) or query set (1-40) against query set (1-40). In the validation testing, the top match is always the right match. Figure 4 is an example of validation testing which was carried out. Table 2 is an example of the system performances depending on the ranking number and the Graph in figure 5 is the plot of the system performance (TAR represents the system performance). It can be seen that in Validation testing, the system performance is 100 percent accurate because the top one match is always a right match.

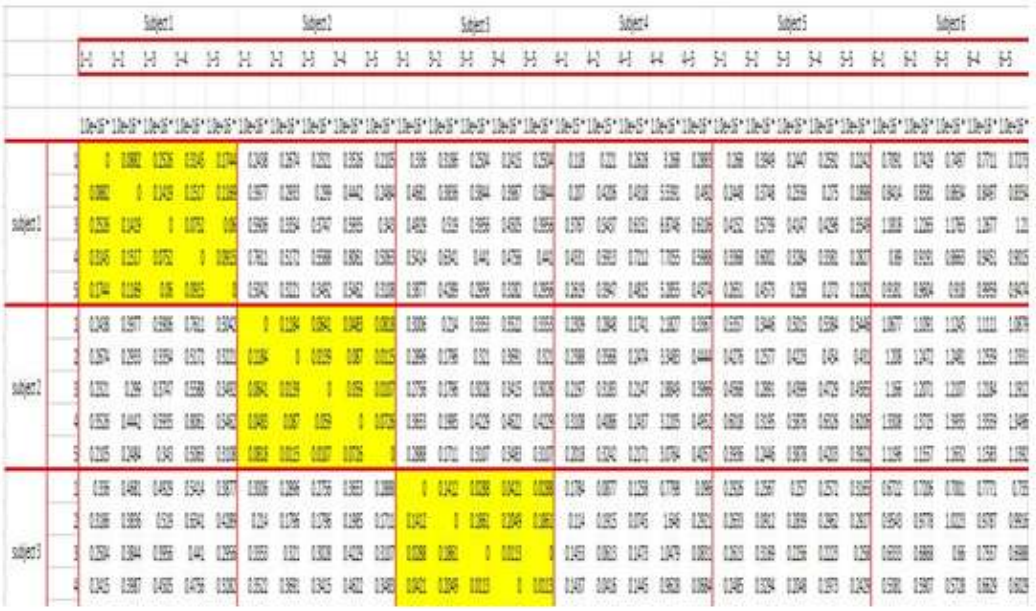


FIGURE 5 Euclidean distances

Rejection testing

Rejection testing is a testing method that is carried out between the same sets of images, rejection testing was tested against the gallery set.

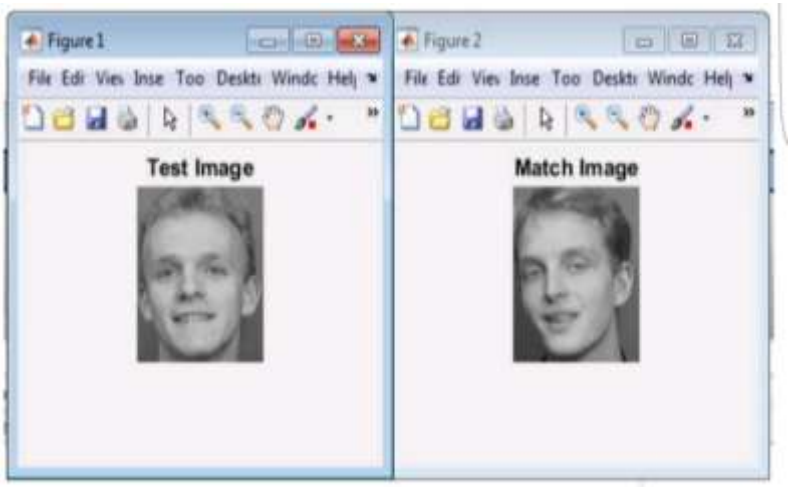


Figure 6 system error

The MapReduce system was tested on a local PC by running face recognition, the system recognition performance was still the same but there was a big difference in the execution time. When a face recognition was run without the MapReduce, the system execution time was 17.280, while when the system was implemented using the MapReduce function, the execution time for the same face recognition in figure 7 was 8,910,11 and 12

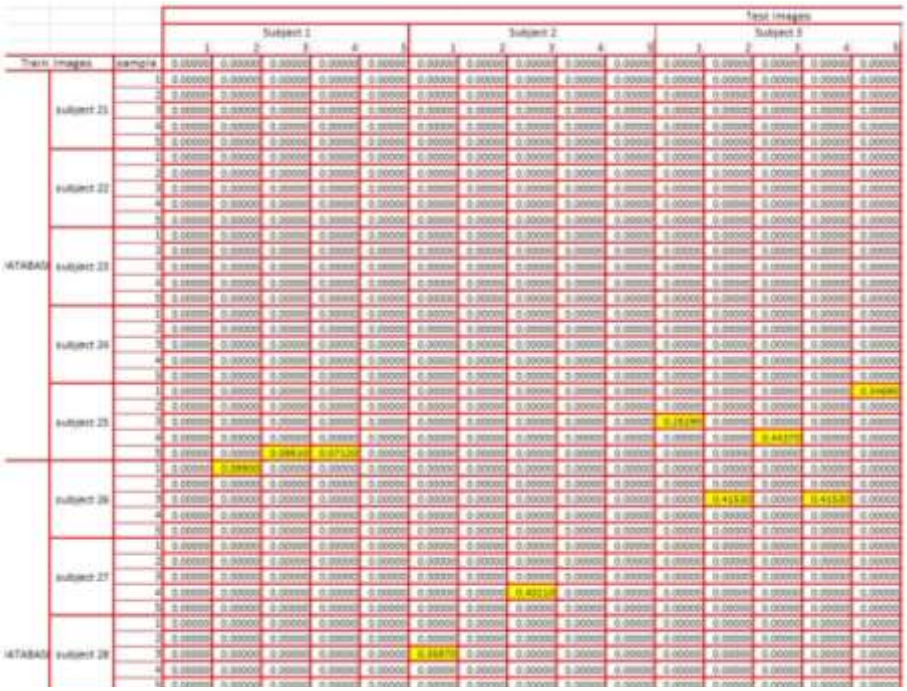


Fig.7 Threshold

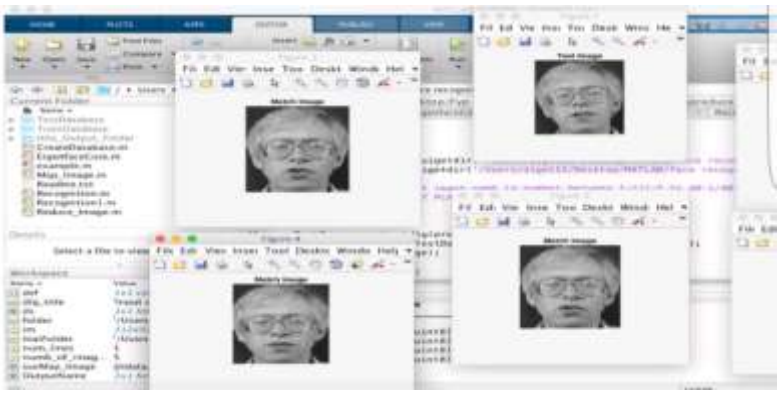


Figure8 match images

The execution time without MapReduce and with MapReduce it was 7.49. After the face recognition was computed, a stem plot was plotted (it can be seen below in figure 4.8 and figure 4.9) to show the Euclidian distances.

Figure 9 match image from the training set

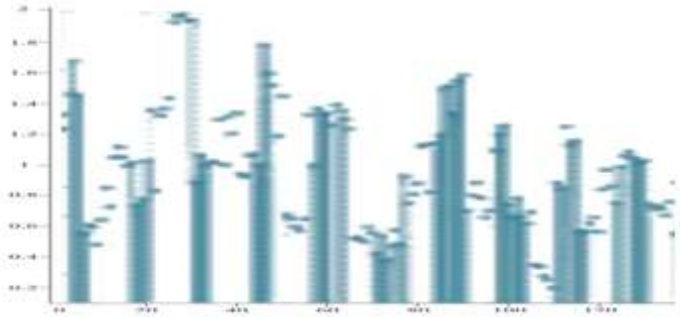


Figure 10 steam plot before sorting the Euclidean distances in ascending order

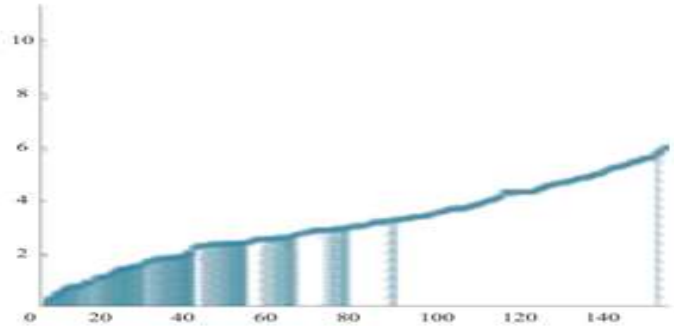


Figure 11 steam plot after sorting the Euclidean distances in ascending order

IV. CONLUSION

The purpose of this paper explains how implementation of Biometric application based on Big Data was done. The main objective is to implement a face recognition application based on Hadoop (Big Data tool). The system was successfully designed using MATLAB. This paper work involves the use of PCA technique for the face recognition and MapReduce for the analysis and management of the Big Data. Based on the literature review and the practical testing that was carried out in this project, it has been concluded that PCA is easier to implement compare to the rest of the face recognition technique, PCA should also be the preferred holistic colour based face recognition method and the level of performance of PCA based algorithm is reasonable. Face recognition based on Big Data has also been reported that is being used by government intelligence for security purpose and it have actually help improved the level of security. It's been concluded that, a big or multiple processor or a big RAM (memory) are required in order to help fasten the analysis of Big Data despite having Hadoop. The system was design and was successfully tested, it was tested using validation testing method, generalisation testing method and rejection testing method and it has proven to be efficiently, although under bad condition the system does its best to provide the right top five ranking image.

REFERENCES

[1]. Anil K. Jain, Arun Ross, and Salil Prabhakar, "An Introduction to Biometric Recognition", IEEE Transactions on Circuits and Systems for Video technology, Vol. 14, No.1, Jan. 2004, pp:4-20, URL: An Introduction to Biometric Recognition, last accessed 14 Apr.2016.

[2]. Xiaoguang Lu" Image Analysis for Face Recognition" published by Dept. of Computer Science & Engineering Michigan State University

[3]. S.T.Gandhe, K.T.Talele, A.G.Keskar "Intelligent face recognition techniques: A comparative study" published in IAENG International Journal of Computer Science

[4]. Anil Kumar Sao and B. Yegnanaarayana "Templatematching Approach for Pose Problem in Face Verification Speech and Vision Laboratory" Department of Computer Science and Engineering, Indian Institute of Technology Madras

[5]. X. Jin, B. W.Wah, X. Cheng and Y. Wang, Significance and challengesof big data research, Big Data Research, 2(2) (2015), pp.59-64.

[6]. R. Kitchin, Big Data, new epistemologies and paradigm shifts, BigData Society, 1(1) (2014), pp.1-12.

[7]. C. L. Philip, Q. Chen and C. Y. Zhang, Data-intensive applications challenges, techniques and technologies: A survey on big data, Information Sciences, 275 (2014), pp.314-347.

[8]. K. Kambatla, G. Kollias, V. Kumar and A. Gram, Trends in big dataanalytics, Journal of Parallel and Distributed Computing, 74(7) (2014),pp.2561-2573.

[9]. S. Del. Rio, V. Lopez, J. M. Bentez and F. Herrera, On the use of map reduce for imbalanced big data using random forest, Information Sciences, 285 (2014), pp.112-137.

[10]. http://www.itl.nist.gov/div898/software/dataplot/refman2/auxillar/falpos.htm

[11]. M. Bartlett, H. Martin, y T.J. Sejnowski. Independent component representations for face recognition. In Proc de la SPIE Symposium on Electric Imaging, 1998.

[12]. P. Comon. Independent component analysis, a new concept? Signal Processing, 36:287–314, 1994.

[13]. D. L. Swets and J. J. Weng, "Using discriminant eigenfeatures for image retrieval", IEEE Trans. PAMI., vol. 18, No. 8, 831-836, 1996.

[14]. C.Magesh Kumar, R.Thiyagarajan , S.P.Natarajan, S.Arulselvi,G.Sainarayanan,‖ Gabor features and LDA based Face Recognition with ANN classifier‖,Procedings Of ICETECT 2011

[15]. Onsen TOYGAR Adnan ACAN ,‖Face recognition using PCA,LDA and ICA approaches on colored images‖,Journal Of Electrical and Electronics Engineering, vol-13,2003

[16]. Y. Cheng, C.L. Wang, Z.Y. Li, Y.K. Hou and C.X. Zhao,‖ Multiscale principal contour direction for varying lighting face recognition‖,Proceedings of IEEE 2010

[17]. Marian Stewart Bartlett, Javier R. Movellan, Terrence J. Sejonowski, ―Face Recognition by Independent Component Analysis‖, IEEE Transactions on Neural Networks, vol-13, No- 6,November 2002, PP 1450-1464.

[18]. Sanjekar, P. S., & Patil, J. B., An Overview of Multimodal Biometrics, Signal & Image Processing: An International Journal (SIPIJ) Vol. 4, 2013

Ali Mohammed Dige"Implementation On Face Recognition Techniques." The International Journal of Engineering and Science (IJES) 7.7 (2018): 08-16