International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9, Issue-7, May 2020

Face Recognition by Reconstructing A 2.5 D Face using Photometric Stereo

Bhavnesh Jaint, Shivam Maheshwari, Sushant Agarwal, Sampada Singh

Abstract*: *Face recognition is an important application of image analysis and it has received a lot of interest in the last decade. There is a critical need for a reliable identification system. As of now, face recognition is not reliable enough in the majority of security applications, therefore a low cost, accurate, and viable identification method are required for face recognition. Two dimensional (2D) face recognition systems that are already existing are often not reliable. Three dimensional (3D) face recognition systems produce more accurate and robust than 2D systems but they are very costly due to large scanning and coded light and also consume a lot of time in the recognition process. This paper aims to produce a low-cost 3D face recognition system (2.5D) using photometric stereo which is less explored in face recognition systems. The capabilities of photometric stereo for use in face recognition are evaluated using a number of experiments conducted using the photometric stereo system and it is implemented and shown to be better than our traditional 2D systems. This system is aimed to solve a number of issues we see in face recognition systems like illumination, distance from the camera and pose and thus, it could be a useful application for biometric authentications in homes, governmental organizations and financial institutions

Keywords: Biometric Control, Cosine Similarity Face Recognition, Feature Extraction, Intensity Matrix, Photometric Stereo, Recognition Rate, Similarity Score, Surface Normal

I. INTRODUCTION

Biometric Access control has picked up over the years because of the limited success of physical checks in protecting vital assets [1]. This technique has been utilized by government and private offices to their benefits against burglary and terrorist activities. Restrictive access through biometric security has become even more significant after the expanding frequencies of psychological militant exercises in the cutting edge world.

Revised Manuscript Received on May 30, 2020.
***** Correspondence Author

Bhavnesh Jaint*, Department of Electrical Engineering, Delhi Technological University (Formerly Delhi College of Engineering), New Delhi, India. Email: bhavneshmk@gmail.com

Shivam Maheshwari, Department of Electrical Engineering, Delhi Technological University (Formerly Delhi College of Engineering), New Delhi, India. Email: shivammaheshwari_bt2k16@dtu.ac.in

Sushant Agarwal, Department of Electrical Engineering, Delhi Technological University (Formerly Delhi College of Engineering), New Delhi, India Email: sushantagarwal8@gmail.com

Sampada Singh, Department of Electrical Engineering, Delhi Technological University (Formerly Delhi College of Engineering), New Delhi, India Email: sampadasingh01@gmail.com

A portion of the frameworks is the Fingerprint Access Control and Face recognition. Biometric Access Control is a security framework that gives restrictive access in the wake of checking to one's unique physical characteristics. The system filters the face, irises, fingers, and veins for unique traits, changes over the information into digital format, and afterward utilizes an intricate calculation to make a match. There is almost no opportunity of it being misused, making it a valuable application for financial institutions, government, and organizations. The present system of passwords and pin numbers in administrations has drawn a great deal of criticism because of the expanding incidents of hacking. Biometric control can likewise be utilized to improve attendance in organizations. The data can be confirmed with the entries in the manual registers kept up by each office. Face recognition can be easily and routinely followed in day to day lives. The wide accessibility of desktop and embedded systems have made an interest in automatic processing of images and videos in various applications, including biometric authentication. Thus, there is a need for research in automatic face recognition. Research in face recognition is required not only due to the difficulties faced by recognition problems but also in various applications where human identification is needed. Face recognition, as one of the essential biometric technologies, turned out to be increasingly more significant from quick advances in technologies like the Internet and increased demand for security systems. Face recognition is natural and easy to use, thus having several advantages over the other traditional biometric technologies A face identification system is expected to automatically distinguish faces present in images and videos. It can work in either or both of two modes: face verification (or validation), and face identification (or recognition). Face verification compares a query face image to a template face image (one-to-one match). Face identification is a one-to-many match process where the query face image is recognized by comparing it with all the stored images in the database. Despite the fact that progress in face recognition has been empowering, it turns out to be a difficult task, particularly for unconstrained tasks where viewpoint, illumination, expression, accessories, and so on can change extensively [2]. We will give a brief review of technical advances and analyze technical challenges in the following sections.

II. ISSUES IN FACE RECOGNITION

A. Pose Variation

Head's movements, which can be depicted by the egocentric rotational angles, that is, pitch, roll, and yaw, or camera changing point of view could prompt changes in facial appearance and creates variations in the face, making automated face identification across a pose a tedious task.

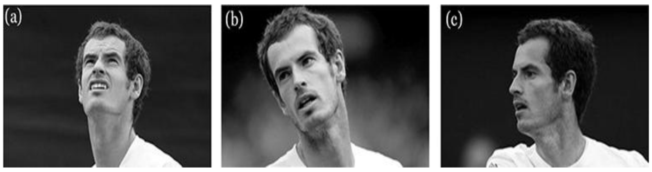

Fig. 1. (a) pitch (rotation around x-axis), (b) roll (rotation around z axis) and (c) yaw (rotation around y-axis).

B. Presence or absence of structuring elements (occlusions)

The presence of components such as beard and/or mustache (Figure 2b), cap (Figure 2c), sunglasses (Figure 2d), etc. or occlusions of the face (Figure 2e) by background objects or its presence can also result in the diversity of the images.

Fig. 2. (a) without occlusion, (b) facial hairs, (c) cap, (d) glasses (e) obstruction by hand (partial occlusion).

Therefore, face's pictures taken in an unconstrained environment require effective recognition of faces altered by occlusions and accessories, as handled by appropriate approaches like texture-based algorithms.

C. Facial Expression Changes

A person's emotional state can also be a cause of the change in the facial expression and thus could result in variability in the face. Thus, it is required that the different facial expressions are automatically and efficiently recognized for the evaluation of emotional states and the automated face recognition. Humans can express their emotions in form of macro-expressions, which could possibly express anger, fear, happiness, sadness, disgust, surprise, and various other facial patters with all of these expressions producing non-rigid motion of the face.

Fig. 3. Different facial expressions reflecting emotions such as (a) resentment, (b) disgust, (c) grief or (d) cheerfulness.

D. Aging of the face

Fig. 4. Illustration of the human aging process, where the same person has been photographed (a) at a younger age and (b) at an older age, respectively.

Face appearance could also be changed due to the aging of the human face and could have an impact on the process of face recognition if there is a significant gap between each image capture.

E. Variability in the lighting condition

The performance of face recognition system can be degraded with large variations in illuminations. Low level of lighting in the background or foreground creates problems in the recognition process as the shadows appearing on the face makes it difficult to detect the facial pattern. A high level of lighting would result in overexposure of the face and undetectable facial patterns. Automatic face recognition and detection with largely varying levels of lighting, be it low level or high level applies to image processing techniques like illumination normalization, for example, histogram equalization or various machine learning methods.

Fig. 5. (a) more exposure to light, (b) large shadows

A lot of effort has been made in 3D methods of face recognition [3]. In this paper, we aim to review one such method for 3D data capture which is photometric stereo (PS). In this technique, multiple 2D images of an object are captured each of the images with different light source directions. The adjustments in pixel intensities at each point are utilized to derive surface orientation. The directions are represented by surface normal vectors situated at points on the target object which is oriented orthogonally to the surface at that point. Photometric stereo is used for the purpose of estimating the surface normal of the objects under unstable lighting conditions. It depends on the way that the amount of light reflected by a surface is dependent on the orientation of the surface relating to the observer and the light source. The space of possible surface orientations is limited by estimating the amount of light reflected in the camera. Given enough light sources from various angles, the surface orientation might be constrained to a single orientation or even over-constrained.

III. PROCESS FLOW

There is a process flow common in all face recognition systems [4]

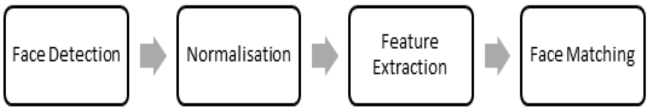

Fig. 6. Process Flow

A. Face Detection

At this step, our aim is to find whether there exists a face in an image or not and if there exists an image then our task is to find the location and extent of it in the given image because all other things present in the image are not useful in the recognition process and are considered as noise.

B. Normalization

It is a very important step; the accuracy of matching depends a lot on the feature extracted from the face and for correct feature extraction image should be normalized both geometrically and photometrically.

International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9, Issue-7, May 2020

C. Feature Extraction

At this point, we quantify our image with the help of certain parameters. Broadly they are classified into 2 categories, 1st is on the basis of the geometry of the face and another one is on the basis of the texture of the skin. It could be a vector representation of the image where elements of the vector are the pixels of the image or it could be a geometrical representation of certain key ratios of face geometry.

D. Feature Matching

It's done with the help of a classifier like a neural network, support vector machine, classification based on Euclidean distance with a certain threshold.

IV. PROPOSED METHODOLOGY

Face detection is a crucial component in a face recognition system. It is a very well researched domain; several methods have already been proposed for finding a human face in an image and then localizing them. It's not our area of interest, our objective here is to deal with the problem of face recognition under variable illumination which significantly affects the performance of a face recognition system. It is a well-established fact that face recognition executed on 3D data performs better as compared to 2D data [5]. But the capture of 3D data is an expensive process, it requires sophisticated instruments such as laser or coded light to capture the 3D image of a person. These devices take a considerable amount of time to scan a single face (more than 30 seconds in case of laser light), which is not desirable. There are some less expensive options such as Stereo which works by taking 2 images of a person from different angles and uses triangulation to find the surface gradient and depth of the face. But this process of triangulation creates holes in the 3D model and it's also computationally very demanding. All these problems can be handled by using photometric stereo. It's a prominent concept in computer vision for 3D modeling of objects. This technique can be used to created 3D models of the face from 2D images taken under different lighting. These 3D Models are not exactly similar to as taken by a 3D capture device but they provide good enough accuracy when compared with the cost involved in the process of image capture and computational efforts required. Also, the images used for creating the model contains only the face and model generated after photometric stereo contains depth information for the face only, not the entire head. So, the technique here is known as 2.5D, not 3D.

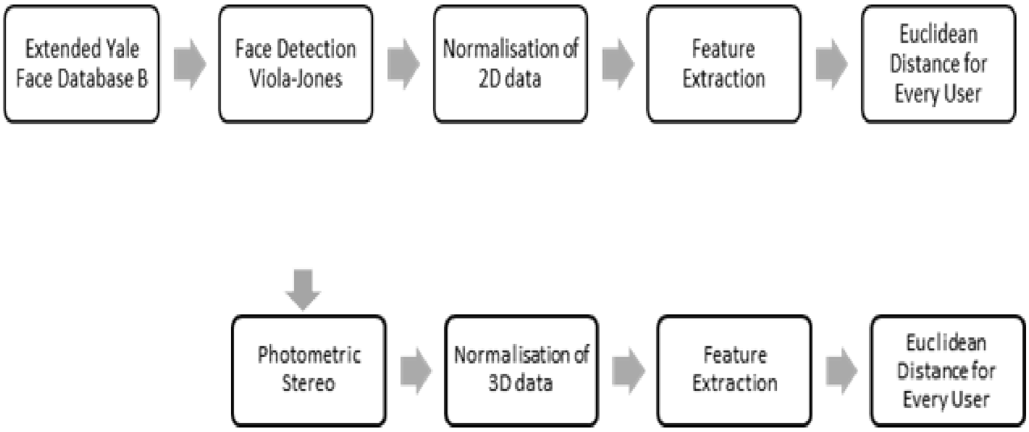

Fig. 7. Steps followed in the recognition process

A. Photometric Stereo

Photometric stereo works by creating a 3D model of an object by taking images under variable lighting conditions. It's based on SFS (Shape from shading) method which tries to recover the shape of an object from a single 2D image by using the gradient of the image in a certain direction [6], it gives an idea about the slope or depth of an object. But the problem with SFS method is that the image gradient can also be due to color or texture information which is difficult to separate by using a single image. Photometric stereo is an enhanced version of the famous SFS method which works by using several images of one person under different lighting from the known direction. It is used to find surface normal for each pixel which is then integrated to create a heightmap. This technique of 3D modeling is well researched and documented but its use in face recognition is comparatively less researched. The surface normal for each pixel is calculated as follows [7]

$$X(x,y) = \frac{I2 - I3}{d * 0.867} \quad \text{....(1)}$$

$$Y(x,y) = -\left(\frac{(2 * I1) - I2 - I3}{d * 0.867}\right) \quad \text{....(2)}$$

$$Z(x,y) = \frac{I1 + I2 + I3}{(3 * h)} \quad \text{....(3)}$$

Where I1, I2, and I3 are the intensity values of the three images at the pixel coordinate (x, y). x, y are the coordinates of a given pixel in the three images, d & h respectively are the distance between light sources and between the light source and the person whose image is being captured. Since between 0 - 255. These 3 coordinates are created for all the pixels in the image by using equation (1) & equation (2), which constitutes a matrix for surface normal. Then an integration algorithm is applied on the surface normal matrix *S* which gives us the height data. Frank & Chellappa Algorithm is used for the reconstruction of the face and obtaining the depth data [8]. It's an effective method known for its accurate results and high tolerance to noise. It uses the Fourier transform for the purpose of integrating the surface normal and reconstructing the object's surface.

Retrieval Number: G5447059720/2020©BEIESP
DOI: 10.35940/ijitee.G5447.059720
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

V. FEATURE EXTRACTION

The success of a face recognition system relies heavily on what features are extracted for the purpose of quantifying the human face.

Features are extracted in such a manner so that it becomes easier to uniquely identify a human face. Most often we extract more than one features and together the set of these features is used in the recognition process [9].

In this paper, we compared the accuracy of recognition on both the 2D data and 2.5D data (generated from the photometric stereo model).

Features that are extracted in the case of 2D data are: -

- Nose length
- Distance between the right eye and right ear
- Distance between left eye and left ear
- Distance between 2 ears
- Distance between 2 eyes
- Distance between both eyes and nose tip
- Distance between chin to both ears

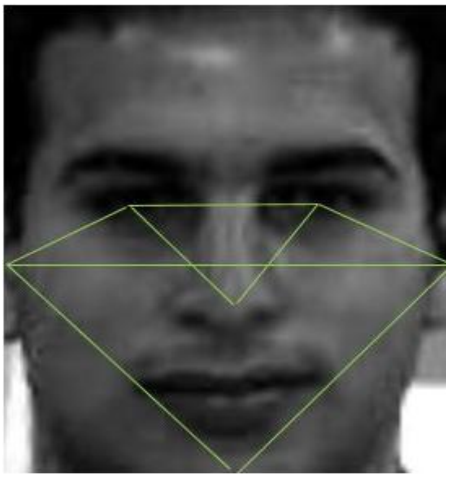

Fig. 8 Parameters in 2D image

All these parameters together are used in the process of face matching. In the case of 2D data, we used three images for each person to find these parameters, and then the average of these parameters for all the three images is considered for creating the feature vector.

Features that are extracted in the case of 2.5D data are: -

- Nose length
- Nose height at the tip
- Nose height in the middle
- Distance between the right eye and right ear
- Distance between left eye and left ear
- Distance between 2 ears
- Distance between 2 eyes
- Distance between both eyes and nose tip
- Distance between chin to both ears

In the case of 2.5D data, two more important features can be calculated such as nose height at the tip and in the middle of the nose bridge. Also, other common parameters will not be exactly the same because of the new structure created. These features are localized by running a certain algorithm on the surface normal and intensity matrix. The nose is the first feature that is being localized here based on the heuristic that the nose is an extended part on a human face and it sticks out the most in comparison to other parts of the human face [10]. We'll localize the nose tip as a point having the most height difference near a surface. A horizontal profile is created for

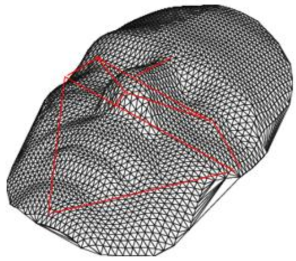

Fig. 9. Parameters in 2.5D image

each pixel using the heightmap. Then, the highest value in the heightmap is found using a simple filter algorithm and the corresponding x and y coordinates are then located. These x and y coordinates define the location of the nose.

VI. FACE DATABASE USED

There are various publicly available databases available for researchers having different scopes and sizes. To get the best results, it is advised to use a standard face recognition database. In this project, Extended Yale faces dataset B has been used. It contains 16,128 images captured from 28 subjects under nine poses and 64 illumination variations.

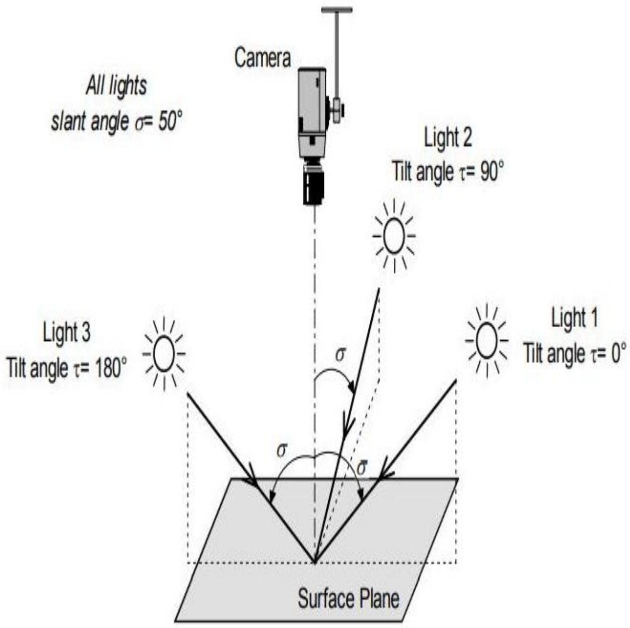

Fig. 10. Camera & light source set up for photometric stereo

For the purpose of our experiment, the database is divided into three sets, training set, gallery setting, and probe set. Out of 28 subjects, 20 subjects are being used for storing in the system as known identities. 11,520 images of these 20 subjects are divided between the training set and gallery set on the basis of the angle between the light source direction and the optical axis of the camera. Each subject of this database has 576 sample images, each image has a file name in a format that describes the subject number, its pose, angle 'α' between light source direction and the optical axis of the camera and the angle of elevation of the light source with respect to the camera.

For building a 2.5D face model using photometric stereo, we used three images per person, one with -50 degrees of elevation of the light source and 'α' as zero degree, other two with zero degrees of light source elevation and +50 degree and -50 degree of angle 'α' respectively as shown in figure 10. Only the frontal face images with this configuration are selected which constitutes the Training set, the rest of the frontal images that are not used for building the photometric stereo model are used for the purpose of testing. Gallery set contains frontal images for each of the 20 known faces under the remaining 61 illumination condition and the last set, probe set contains the images for the rest of the 8 subjects (unknown to the system) under all 64-illumination condition. Samples images for one subject from the Extended Yale database B are shown in Figure 11.

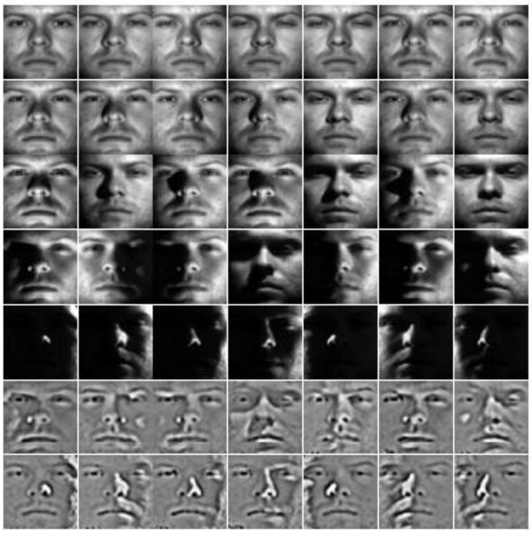

Fig. 11. Extended Yale Face Dataset B Sample Images

VII. RESULT AND DISCUSSION

Three metrics are calculated to check the performance of the system:

- Recognition Rate (RR): - Percentage of correctly identified persons from the gallery set, which are known to the system.
- False Acceptance Rate (FAR): - Percentage of unknown faces from the probe set accepted by the system
- False Rejection Rate (FRR): - Percentage of known faces from the gallery set rejected by the system

The similarity score '*S'* is being created and a threshold of 97% is set for accepting or rejecting a person by the system. It's obtained by using the concept of cosine similarity between feature vectors [11]. Let us suppose the feature vector for the training image is denoted by *'x'* and the feature vector of the test image is denoted using *'y'*. Then S is calculated as follows:

$$S = \frac{x.y}{|x||y|} \qquad \text{....(3)}$$

Here $|x|$ and $|y|$ represents the magnitude of the two feature vectors x and y respectively. The value of S lies between 0 and 1. The results obtained from the experiment are mentioned in Table I.

Table-I: Experimental Results

Approach	RR	FAR	FRR
2D	65.24%	4.29%	3.52%
2.5D	74.59%	2.53%	2.34%
Approach	RR	FAR	FRR
2D	65.24%	4.29%	3.52%
2.5D	74.59%	2.53%	2.34%

All these metrics are calculated using a threshold of 97% in similarity score. A low threshold results in high FAR and low FRR. For security systems, FAR is a very important metric and its high value signifies more chances of an intruder getting through the security system. Increasing the value of threshold lowers FAR value but it simultaneously increases the FRR which means more known persons getting rejected by the system. So there exists a tradeoff between FAR and FRR and it's adjusted as per the user requirement by changing the value of the threshold.

VIII. CONCLUSION

Face Recognition is a popular biometrics technique used nowadays. Most of the face recognition systems available are designed for constrained situations such as standard lighting conditions, neutral facial expressions, frontal view face, and no occlusions. All the other conditions which are not constrained deteriorates the performance of the face recognition system. Several types of research have been done that prove that 3D recognition techniques perform better under unconstrained conditions especially variable illumination and different poses. But these face recognition techniques require 3-dimensional data, and 3D image capturing devices are very expensive & takes considerable time to capture a single human face. The method proposed here uses 2-dimensional data to find surface normal and then height data after integrating the surface normal using photometric stereo. This proposed methodology clearly outperforms the traditional 2D recognition techniques in recognizing a person using the images taken under unstable lighting conditions. The improvement of 9.35% is achieved in the recognition rate, also the false acceptance rate and false rejection rate are decreased by significant numbers which makes the overall system better. These improvements in important metrics are achieved at a low cost by increasing the accuracy of extracted features and adding another dimension to the 2D images using photometric stereo without using any expensive 3D image capturing device.

REFERENCES

1. Pithadiya, Kunal. (2016). An Analytical Survey on Biometric Access Control for Physical Security. 10.13140/RG.2.1.2811.6081.
2. Kortli, Yassin & Jridi, Maher & Falou, & Atri, Mohamed. (2020). Face Recognition Systems: A Survey. Sensors. 20. 342. 10.3390/s20020342.
3. Gokberk, Berk & Salah, Albert & Akarun, Lale & Etheve, Remy & Riccio, Daniel & Dugelay, Jean-Luc. (2009). 3D Face Recognition. 10.1007/978-1-84800-292-0_9.
4. Zhao, Wen-Yi & Chellappa, Rama & Phillips, P. Jonathon & Rosenfeld, Azriel. (2003). Face Recognition: A Literature Survey. ACM Comput. Surv.. 35. 399-458. 10.1145/954339.954342.
5. Abate, Andrea & Nappi, Michele & Riccio, Daniel & Sabatino, Gabriele. (2007). 2D and 3D Face Recognition: A Survey. Pattern Recognition Letters. 28. 1885-1906. 10.1016/j.patrec.2006.12.018.

6. Herbort, Steffen & Wöhler, Christian. (2011). An introduction to image-based 3D surface reconstruction and a survey of photometric stereo methods. 3D Res. 2. 1-17. 10.1007/3DRes.03(2011)4.
7. Lenglet, C., 2003. 3D surface reconstruction by shape from shading and photometric stereo methods. Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV2003), pp 1-20.
8. School of computer science and software engineering at the University of Western Australia, 2006. Frankot-Chellappa surface integration algorithm implemented in Matlab [online]. Available from: http://www.csse.uwa.edu.au/~pk/research/matlabfns
9. S. R. Benedict and J. S. Kumar, "Geometric shaped facial feature extraction for face recognition," 2016 IEEE International Conference on Advances in Computer Applications (ICACA), Coimbatore, 2016, pp. 275-278.
10. Zhou, Song & Xiao, Sheng. (2018). 3D face recognition: a survey. Human-centric Computing and Information Sciences. 8. 10.1186/s13673-018-0157-2.
11. Rahutomo, Faisal & Kitasuka, Teruaki & Aritsugi, Masayoshi. (2012). Semantic Cosine Similarity.

AUTHORS PROFILE

Bhavnesh Jaint received her degree in B.E. in Electronics and Telecommunication Engineering from Government College of Engineering, Jabalpur, M.P India, and M.E. degree in Electrical Engineering from Jabalpur Engineering College, Rajiv Gandhi Proudhyogiki Vishwavidyalaya Bhopal, M.P., India. During her M.E., she has completed a dissertation in the area of Fuzzy logic. She has started her career as Lecturer of Electronics and Communication at BBITRC, Jhangribad affiliated to UPTU. She has worked as Assistant Professor at ABES Engineering College, Ghaziabad. Presently, she is an Assistant Professor in the Department of Electrical Engineering at Delhi Technological University, Delhi, India. She is a member of the different statutory committees of the university. She is also the life member of professional societies Indian Unit for Pattern Recognition and Artificial Intelligence. Also, a member of the IEEE. Her research interests include wireless sensor networks, Embedded System IOT, Control System.

Shivam Maheshwari, undergraduate student, Department of Electrical Engineering, Delhi Technological University (Formerly Delhi College of Engineering), Delhi, India. His research areas include machine learning, computer vision, predictive modeling.

Sushant Agarwal, undergraduate student, Department of Electrical Engineering, Delhi Technological University (Formerly Delhi College of Engineering), Delhi, India.

Sampada Singh, undergraduate student, Department of Electrical Engineering, Delhi Technological University (Formerly Delhi College of Engineering), Delhi, India.