International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-8 Issue-11, September 2019

A Real Time Virtual Dressing Room Application using Opencv

Rshami S. Shinkar, Nagaraju Bogiri

***Abstract**: The real time clothes trying having long waiting queues for trial rooms is a problem faced at many places and store. In this system, by using real time system cloth processing is done by data drive approach which is worn by human. To give exact look to skeleton, by using height, skin color, it starts creation of clone person which is prior to real-time simulation. We use hardware sensors like motion, light, camera sensors which are controlled by using GUI (Graphical User Interface) software. Hardware uses latest depth camera of Kinect Sensing element along Unity SDK. We use operating system which can communicate with the user friendly software which should manage the controllers. This proposed approach can offer good GUI which is user friendly to end user and also to retailer. Because of this user friendly system, it should increase the level of marketing more than current system. Proposed system should provide not only good solution for dressing but also can solve the issues which are related to retailer and end user.*

Keywords: Depth Image, Kinect Sensor, Augmented Reality, Unity SDK, Skeleton

I. INTRODUCTION

In stores usually it takes more time to change the clothes or trying clothes. In online shopping, we cannot try clothes. In proposed system we will try to upgrade use of cloth in less time [8]. For this approach we are going to create an environment which should be the effective room and which should be virtual [1]. For this system main issue should occurs regarding to adjustment according to model of cloth with their ideal situation, plate, turning, spinning and ordering [2]. The most important thing of this problem is removing user and their body parts [4]. In literature, for detection of parts of user body, skeletal tracking and pose approximation there were many ways proposed. For online shopping users, by using web camera it is going to be simple.

To make platform independent approach and portable approach we should implement this by using OpenCV [5].
Our approach is as follows:

- By using depth and user label data, by video stream user should be extracted.
- According to the skeletal tracker, Positioning of the 2D cloth models.
- From the sensor distance should be scaled between body joints and user.
- To prevent blockage of body parts which is unnecessary skin color detection is used.
- On user stratified of the model.

Manuscript published on 30 September 2019.
***Correspondence Author(s)
Rshami S. Shinkar, Computer Engineering Department, K. J.College of Engineering & Management, SavitribaiPhule Pune University, Pune, India.
NagarajuBogiri, Computer Engineering Department, K. J.College of Engineering & Management, SavitribaiPhule Pune University, Pune, India.

For checking the practical performance according to user interface a simple application should be designed. The interface permits the user to settle on a dress by suggests that of hand movements. An extreme way of interacting with machines is action recognition using the Kinect technology. For human body, by using Kinect by Microsoft, analyzation of depth mapping and skeletal joint information is done and also we should recover three dimensional information scenes [9]. According to that we can detect position of human body such as stand, walk, seat, wave etc. Kinect has large no of application areas related with computer, medical, electronics, robotics and many other. By using own body movements, people can play games. In medical purposes, by using Kinect remote location doctor can operate patient. To interact with human's, Kinect is used applications related to the software and medicines. Using Kinect is great thing in computer vision. The computer vision is based on Human Computer Interaction i.e. HCI. The Kinect sensor senses the environment and produces depth map for it. By using the mean shift algorithm, the human body is tracked using skeletal tracking. In skeletal tracking, 24 joints of human body should be identified by using Kinect sensor. Kinect find out the action by the human body and gestures according to 3D joint information, [3] after that according to the input given to machine, machine should replies.

A. KINECT SENSOR:

Kinect Sensor is manufactured by Microsoft. In Kinect Sensor, depth sensor, four-microphone array and a color camera were present. The depth sensor includes IR (Infrared) senor. IR Camera is a monochrome sensor which is present in IR sensor. IR camera is usually related with idea regarding to structured lights. IR laser moved from broadcast grating and turned into IR dots set is known as IR projector. Projector, Camera and Dot pattern regarding with IR have a known respective geometry. Dot in the image and dot in the projector pattern become same at that time it should be represented in three dimensions.

Depth map for IR should be created by using the Kinect sensor. The darker pixel is near camera point because depth values should be encoded with gray values. If there are no depth values present, at that time it should consider than point is too near or too long to be calculated and no representation can be made.

Fig.1: Kinect Sensor.

Specifications of Kinect cameras are as follows:

RGB Image	resolution	1920*1080
	fps	30fps
Depth sensor	resolution	512*424
	fps	30fps
Depth measurement range		500~4500mm
View angle	RGB Image	Horizontal: 84.1 deg Vertical: 53.8 deg
	Depth sensor	Horizontal: 70 deg Vertical: 60 deg

Fig.2: Specification of Kinect Sensor.

II. REVIEW OF LITERATURE

The current methodology of on-line searching doesn't guarantee the right size of the wear. This result in a number of products being returned and the time taking to replace it with the correct sized one is long. This is a major setback for the online shopping industry.

The various approaches to obtain the desired results are as follows,

Srinivasan K. and Vivek S. [1], in proposed system they should discussed about the growth in online shopping. A person wants to purchase the dress with full satisfaction and also enjoy maximum utilization and justifies the need to develop an algorithm which virtually dresses people with the selected dress. Human silhouette with variable background and noisier environment. Which is the more challenging task in still image using image processing?

Pros: The wide choice of online availability with the assistance of feeling how clothes would look.

Cons: No provision for 3D viewing and sensitive to light conditions.

Ari Kusumaningsih and Eko Mulyanto Yuniarno [2],A virtual dressing room for Madura batik dress has been successfully developed. The proposed system has a purpose tom make dressing room specialized for Madura batik clothes supposed to create attention from customer and should contributes in improving sales performance and promote Madura's heritages as also. Efficient and fast computation methods needed to process numerous 3D models. So that, we don't have to use high performing computer for implementing this virtual dressing room.

Pros: The distance of objects from the Kinect and to compose a "depth map" of the image.

Cons: Lighting conditions affected depth map.

Ting Liu and Ling Zhi Li [3], work uses user extraction from Kinect video stream and avatar system for skeletal tracking to align the clothes' models with users. And a virtual dressing software prototype is developed allowing clothes' 3D models to overlay users and were convenient to view in front, side and back perspectives. Furthermore, improving clothes modeling approaches that achieve rapid reconstruction based on real clothes is also of great use.

Pros: The user can view the real-time collocation effect with the change of hats' textures and clothes 'models.

Cons: Only alignment of clothes according to body is used and dynamic movement is not considered.

Stephen Karungaru and Kenji Terada [4], in this Project, they propose a method to acquire human body length / perimeter easily using Kinect. Experimental results confirmed that human data can be acquired from Kinect sensor. We also confirmed problems in case of error in acquired data. Future issues include improving the accuracy of acquisition of person's data and the CG.

Pros: Most work is focused on acquiring human data.

Cons: No interactive activities are focused and as they say accuracy is low. Human data is acquired and not used

Dr. Anthony L. Brooks and Dr. Eva Petersson Brooks [5], the open-structured surveys received wide-ranging input from the public attending the live demonstrations at Malls and Messe events.13 wheelchair-bound individuals gave direct input as well as others who were either friends or associated with a wheelchair-bound person that they considered would benefit from a dedicated adaptation of the product. Yet that distance had to be shut enough to permit the person associate degree operable read of the interface management detail.

Pros: This technique/camera as the core of the VDR allows a person to be scanned and identified from the background for the superimposing of the apparel layer over the mirrored self

Cons: Complexity is more as VDR system is used.

Reizo NAKAMURA and Masaki IZUTSU [6], this paper show processes that estimate of body suites size. First, person recognition be got by Kinect. And, person area in the image be extracted using person recognition data. Next, user's mark points are extracted using contour tracing. The size of the body suites was presumed using it.

Pros: Size estimated using the distance data of the number of frames that can be retrieved from two Kinect. Reasons to use two Kinect, this is for improving the accuracy by using the information obtained from the other.

Cons: As use of multiple Kinect's improves accuracy but increases cost also.

Poonpong Boonbrahma and Charlee Kaewrat [7], Using the physical parameter from our experiment, the appearance of the fabrics under simulation can be predicted. The simulation results will tell the distinction among customers sporting jean, satin, silk or cotton, which will be very useful for setting up the virtual fitting room.

Pros: Simulation is done in different environments.

Cons: Need more precision and actual experiments than just simulation in different environments are needed.

Umut Gültepe and Uğur Güdükbay [8], we propose a novel virtual fitting room using depth sensor data. The framework yields a realistic fitting experience for standard body types with customized motion filters, body measurement and physical simulation.

Retrieval Number: H7396068819/19©BEIESP
DOI: 10.35940/ijitee.H7396.0981119
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

The proposed scaling method adjusts the avatar's body size parameters and determines a suitable apparel size, and prepares the collision mesh and the physics simulation. In future work, we'd wish to improve the standard of the measurements and visual scaling by exploitation information from an RGB sensing element likewise, as a result of it provides extra information. We would prefer to increase the amount of collision spheres for higher collision detection

Pros: Depth sensor data is used to filtering room. Adjust the size.

Cons: Quality of measurement is low. Provides additional data of room which is of no use.

Ayushi Gahlot and Purvi Agarwal [9], this system is based on the Kinect technology which contains the features such as action reorganization. This system is concentrate on capturing of the three dimensions information of scene and also find outs user actions by recovering the depth image info and real- time skeletal pursuit. The Kinect technology has revolutionized the way humans interrelate with the machines. It has a large number of applications areas. This system also covers proposed approaches to skeletal related action recognition using Kinect.

Pros: The Kinect Technology interacts with the Human Body for pose estimation.

Cons: Action predictions are main purpose. Pose estimitation is done but not used for any purpose

III. PROPOSED METHODOLOGY

Main features of the proposed methodology are as follows,

1. User Extraction/Detection from the Input Video

By separating user area user allows creating environment and background from the video stream and layered it onto a virtual manner in the user interface. Also, for skin detection it is also useful and also useful to find out region of interest. The user id and depth image is given by the Kinect SDK. At the time of working device, depth image separate background from user. The background is removed by combining the RGBA image with the separated depth image for each pixel by setting the alpha channel to zero if the pixel does not lie on the user.

2. Human Body Detection:

The main function of this step is to determine whether human body appears in a given image, and where these are located at. The expected outputs of this step are patches containing each body parts in the input image. In order to create more organic structure recognition system a lot of study and simple to style, body alignment is performed to justify the scales and orientations of the patches

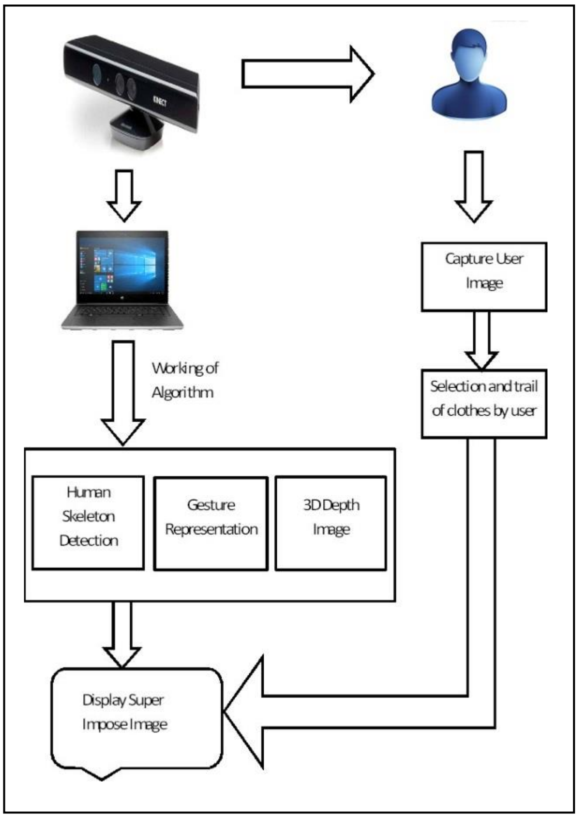

Fig.3: System Architecture

Overall working of the proposed methodology is as shown in Fig.3. A Kinect camera/sensor is connected to windows PC further more we can extend it to large display. Kinect continuously records the depth data and sends the image frames to the PC. On windows system application is running using Unity3D and visual C# along with OpenCV libraries. Unity uses Kinect plug in to get data from sensor. It then processes each frame along with libraries to predict the human pose and action. Once pose is stable system imposes predefined 3D models of clothes on human body and displays it. If action is predicted then the gesture is calculated and action is taken accordingly; for example if swipe right gesture is calculated then clothes gets change.

3. Kinect Skeletal tracking:

Large number of dimensions is described in the skeletal tracking. Skeletal tracking dimensions describe unique individuals, shapes, size, hair and clothing etc.

In this, body of human is showed by the body part joints such as head, neck, hair and arms. All body part joints were shown according to their 3D coordinates. Because of this, we are treating the sectional is a auction process of the depth image according to per-pixel classification task. If we separate each pixel individually, it will avoid combinative search.

For training, Realistic images of humans with different sizes and shapes and variety of poses are generated from large database. By using mean sift algorithm, in three dimensional joint geographical nodes of the per-pixel distribution obtained are figure out. Boost execution of this algorithm is consider as 5ms per frame on XBOX 360 Graphical Processing unit i.e. GPU. There for, the whole idea of pipeline is as follows:

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Stage 1- classification of body parts per pixel.

Stage 2- with the help of mean shift algorithm, for finding global centroids of local modes of density hypothesize of body joints is done.

Stage 3- by using temporal continuousness mapping of hypothesized joints to the skeletal joints and suite a skeleton is done.

4. Use of OpenCV in project:

A. User Extraction:

Pertained models such as haar cascades and hog cascades which are xml format files and are made by training number of images in which human is present and not present, are available in OpenCV distribution folder. One of the cascades are upper body, lower body and full body. If we use full body cascade it searches for the full human body in the camera, which can be used for user extraction.

Skin Segmentation

HSV and YCbCr color spaces are mostly used for skin color partition. We preferred YCbCr color space and the RGB images are change into YCbCr color space by using following equations:

Y = 0:299R + 0:587G + 0:114B

Cb = 128 + 0:169R + 0:332G + 0:5B

Cr = 128 + 0:5R + 0:419G + 0:081B

According to internet report that the most representative color ranges of human skin on YCbCr color space. A threshold is applied to the color element of the image within the following ranges to segment human skin.

$77 < Cb < 127$

$133 < Cr < 173$

$Y < 70$

YCbCr Color System:

YCbCr is also same like RGB. Another color space. But bit different spaces are taken.

Here Y is the luma component of the color. Luma component is the brightness of the color. That means the light intensity of the color. The human eye is more sensitive to this component.Cb and Cr is the blue component and red component related to the chroma component. That means “Cb is the blue component relative to the green component. Cr is the red component relative to the green component.”These components are less sensitive to the human eyes.Since the Y component is more sensitive to the human eye, it needs to be more correct and Cb and Cr is less sensitive to the human eye. Therefore it needs not to be more accurate. When in JPEG compression, it uses these sensitivities of the human eye and eliminates the unnecessary details of the image.

IV. ALGORITHM

1. Pose estimation:

In action prediction process, basically need of pose estimation method should be considered. This method used to make the procedure constant to dissimilarity in aspect, body structure and different clarifications of similar actions. For Post estimation situation of three joints should be considered: shoulder on left side, shoulder on right side and right hip. The subjects pose is given perfectly by using this. By having poses of users we can further modify the clothing model in real time according to poses and hence actual experience of clothing can be generated.

2. Mean Shift:

Mean shift is easy and interactive algorithm which shifts particular data point to the mean of data points in its neighborhood is generalized and examined. In its special cases, k-means clustering algorithms were generated. It is shown that mean shift is a mode-seeking process on the surface constructed with a "shadow" kernel. For Gaussian kernels, mean shift is a gradient mapping. Convergence is studied for mean shift iterations. Cluster analysis if treated as a deterministic problem of finding a fixed point of mean shift that characterizes the data. Applications in clustering and Hough transform are demonstrated. Mean shift is also considered as an advance algorithm that performs multi start global optimization. Shifts of clothes models can be calculated using mean shift algorithm.

5. Mathematical Model:

The mathematical model of any software system can be presented using set theory as elaborated below,

Let S be the closed system defined as, S = {Ip, Op, A, Ss, Su, Fi}

Where, Ip=Set of Input, Op=Set of Output, Su= Success State, Fi= Failure State and A= Set of actions, Ss= Set of user’s states.

- Set of input=Ip={username, password}
- Set of actions =A={F1,F2,F3,F4,F5,F6}

Where,

- F1 = Authentication of system
- F2 = Fetching Kinect values
- F3 = User display in mirror
- F4 = Slide one by one image in mirror
- F5 = If user selects particular products then system show to size of this products
- F6 = Gesture controls
- Set of user’s states=Ss={initialization state, gesture state, selection of products, check size, stop}
- Set of output=Op={Show user image in mirror with clothes}
- Su=Success state={initialization Success, gesture control Success, Show super impose image}
- Fi=Failure State={Kinect Failure, Power Failure}
- Set of Exceptions= Ex ={ Null Values Exception while showing state, Bad light error while recording user}.

V. PROPOSED IMPROVEMENTS

Project is the one dealing with depth camera. Therefore this method will use three types of technologies that are:

a) Markers

Markers are plotting human body points in an image.

b) Optical tracking

Illusion and motion tracking of human bodies

c) Depth cameras

3D representation for real time user experience

Retrieval Number: H7396068819/19©BEIESP
DOI: 10.35940/ijitee.H7396.0981119
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-8 Issue-11, September 2019

d) Gesture Controls

To use system and toggle through clothes gesture controls. Minimized set of gestures will be used with gesture recognition

MODIFICATION IN EXISTING SYSTEM

In many of the existing techniques most of the focus is on just acquiring human data [2][4] like pose, tracking points and not focused on the use of that data acquired from Kinect. Also some techniques work on human interaction but either they are costly by using multiple Kinect's [6] or not efficient. We in proposed methodology combine human data and interaction of human with the system and apply it on a virtual dressing experience where clothes are also 3D modeled and are dynamic in sizing. Some of the major modifications are as follows,

- Track and record the area position data of the user. The creation of beautiful data visualizations or Info graphics this data will be used later.
- Depending on automatic changes in body cloth size should be adjusted.
- Detection of gestures and recognize it for changing clothes or start or stop system. Minimal set of gestures is used to interact and maximum automation approach is used.
- No need for calibration to start detecting human body.
- Automatic rescaling and movement of clothes along with human body change and movement.

Capturing the motion, recognition of face, recognition of voice, etc. These are some of the many features Microsoft made accessible to everyone with their lease of the Kinect and the Kinect SDKs. In this project, other than Kinect devices are not considered, being that likely lots of the research will be done based on its RBG camera. But the reality is that Kinect is able to combine all these different technologies in one small device, making it a very strong tool.

Hence, According to all features and possibilities Kinect offers this device is suitable to achieve the proposed goal. We also think that if the combination of Kinect with the "Virtual Dressing" project is successful, other than a satisfactory device for our experiment, it will be a device that, due to its characteristics, will allow researchers to go much further in their examination, allowing then to explore new and different areas.

VI. RESULT AND DISCUSSION

One of the best features of Microsoft Kinect for Windows SDK is Skeletal Tracking. By using the Infrared (IR) camera, the sensor can track people and follow body actions. The Skeletal Tracking method can recognize up to six persons present in the field view. Out of six persons the method can track up to two people in detail. In the proposed system we will track only one human at a time all body data is plotted as shown in fig.2 and is used for pose estimation. Once pose is confirmed the 3D models of clothes are imposed on human body and are viewed on screen. Also interaction is done by getting gestures from human data.

Comparative results of existing and proposed system is as follow,

Parameter	Existing System	Proposed System
Human Body Tracking	Yes	Yes
3D Depth Image	Somewhat	Yes
Gesture Controls	No	Yes
Real Time Simulation	No	Yes
Dynamic Cloth Fitting	No	Yes

Table1*:* Comparative Results

With reference to Table it is clear that we overcome various problems in existing system and our approach works efficiently.

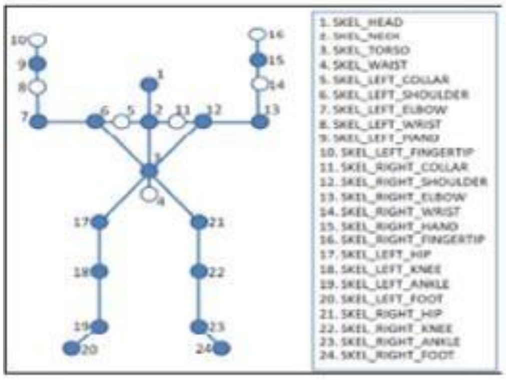

Fig.4: Skeleton Tracking with Points Mapping

As shown in Fig4. The user skeleton is estimated and tracked with body points. Kinect provides the data about change in motion and system will move the skeleton accordingly. This skeleton tracking is used in pose estimation and gesture calculation and further clothes are imposed accordingly on human body and viewed.

Fig.5: Human Body Detection in Various Light Conditions

As shown in Fig.5 the light conditions plays important role in human body detection and estimations. If lights are low or dark then this may affect system.

Following Fig.6 shows the time analysis for superimposing clothes images on various persons.

We tested it with five different persons having different body shapes and measured time taken in milliseconds by system to detect human and process it to superimpose clothes on human body accordingly. All experimentation was carried in normal light conditions.

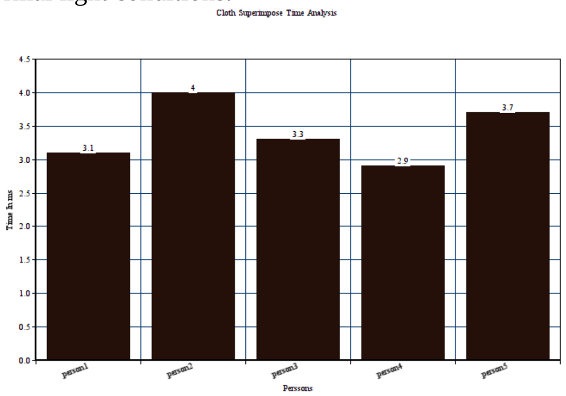

Fig.6: Cloth Superimpose Time Analysis

VII. CONCLUSION

Here only a front image is required in the virtual dressing room application. For every product to layer it on user and two dimensional graphics of given product appear to be comparatively acceptable and real for multiple uses. Given approach is used for range the models with user and to test the system in many states. This system covered all drawbacks and demonstrations have resulted with allowable presentation rates for general positions. According to the model used for fitting there are many possible implementations. This should attainable for use a similar variation to the pictures instead of straightforward scale-rotate ability for matching many joints altogether though it'd necessary additional estimation. Many pictures at different angels should be used because of it; it is feasible for creating more practical video stream. Anyone should bring a same effect using three dimension models and also provide them according to ongoing angle and position. To go along with the model another point of view should make it easy and feasible for implementing physics engine.

REFERENCES

1. Srinivasan K. and Vivek S., "Implementation Of Virtual Fitting Room Using Image Processing" IEEE2017
2. Ari Kusumaningsihand EkoMulyantoYuniarno, "User Experience Measurement On Virtual Dressing Room Of Madura Batik Clothes", IEEE2017
3. Ting Liu and LingZhi Li, "Real-time 3D Virtual Dressing Based on Users", IEEE2017
4. Naoyuki Yoshino, Stephen Karungaru "Body Physical Measurement using Kinect for Vitual Dressing Room"2017 6thIIAI
5. Dr. Anthony L. Brooks and Dr.EvaPetersson Brooks "Towards an Inclusive Virtual Dressing Room for Wheelchair-Bound Customers"978-1-4799-5158-1/14/$31.00 ©2014 IEEE
6. Masaki IZUTSU, and Shosiro HATAKEYAMA "Estimation Method of Clothes Size for Virtual Fitting Room with Kinect Sensor" 978-1-4799-0652-9/13 $31.00 © 2013IEEE
7. PoonpongBoonbrahma and CharleeKaewrat "Realistic Simulation in Virtual Fitting Room Using Physical Properties of Fabrics."
8. UmutGultepe and UgurGudukbay. "Real time virtual fitting with body measurement and motion smoothing."
9. AyushiGahlot and PurviAgarwal. "Skeleton based Human Action Recognition using Kinect.

AUTHORS PROFILE

Ms. Rshami S. Shinkar, ME(Computer Engineering) K. J. College of Engineering, Pune, India

Mr. Nagaraju Bogiri, ME(Computer Engineering) K. J. College of Engineering, Pune, India

Retrieval Number: H7396068819/19©BEIESP
DOI: 10.35940/ijitee.H7396.0981119
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)