International Journal of Emerging Engineering Research and Technology
Volume 2, Issue 7, October 2014, PP 119-126
ISSN 2349-4395 (Print) & ISSN 2349-4409 (Online)

Advance Human Computer Interaction Using Air Signature

Mandar Jangam[1], Jayashree Suryavanshi[1], Aman Gupta[1], Nida Shaikh[1]

[1]Computer Engineering, RMD Sinhgad School of Engineering, Pune, India

Abstract: *Human Computer Interaction in the field of input and output techniques has developed a lot of new techniques over the last few years. With the recently released full multitouch tablets and notebooks the way how people interact with the computer is coming to a new dimension. As humans are used to handle things with their hands the technology of multi-touch displays or touchpad's brought much more convenience for use in daily life. But for sure the usage of human speech recognition will also play an important part in the future of human computer interaction. In this paper we are introducing several promising directions toward achieving multimodal HCI by integrating Hand gestures, Voice recognition and In-Air Signatures. Thereby the gesture and speech recognition take an important role as these are the main communication methods between humans and how they could disrupt the keyboard or mouse as we know it today. This research can benefit for many disparate fields of study that increases our understanding of different human communication modalities and their potential role in HCI.*

Keywords: *Mobile Phones, Accelerometer, Gestures, Hand Writing, Voice Detection, Air Signature*

1. Introduction

With the ever increasing role of computers in society, HCI has become an increasingly important part of our daily lives. It is widely believed that as the computing, communication, and display technologies progress even further, the existing HCI techniques may become a bottleneck in the effective utilization of the available information flow. For example, the most popular mode of HCI still relies on the keyboard and mouse. These devices have grown to be familiar but tend to restrict the information and command flow between the user and the computer system. This limitation has become even more apparent with the emergence of novel display technology such as virtual reality [1],[3],[5] and wearable computers [7], [9]. Thus, in recent years, there has been a tremendous interest in introducing new modalities into HCI that will potentially resolve this interaction bottleneck.

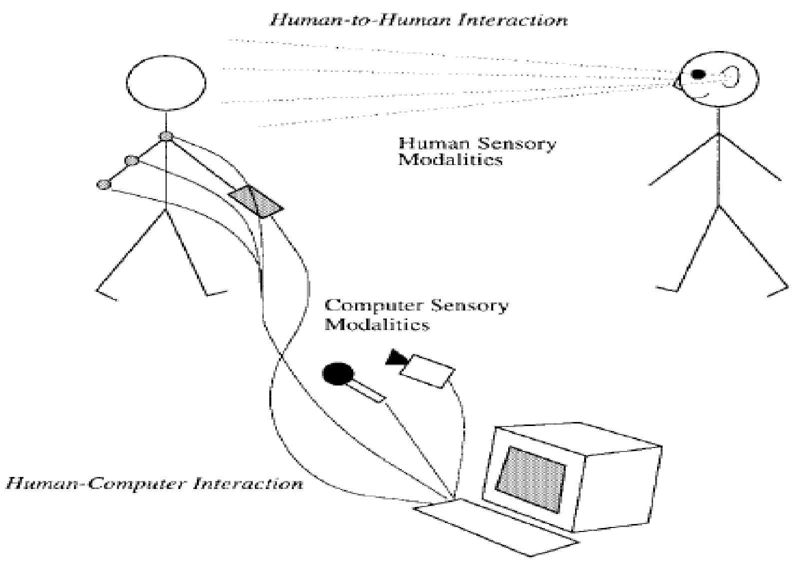


Fig1. *Human-to-human interaction and Human-to-computer interaction*

One long-term goal in HCI has been to migrate the "natural" means that humans employ to communicate with each other into HCI (Fig. 1). With this motivation, ASR has been a topic of research for decades [10]. Some other techniques like automatic gesture recognition, analysis of facial expressions, eye tracking, force sensing, or EEG have only recently gained more interest as potential modalities for HCI. Though studies have been conducted to establish the feasibility of these novel modalities using appropriate sensing and interpretation techniques, their role in HCI is still being explored. A limiting feature of modern interfaces that has also become increasingly evident is their reliance on a single mode of interaction—a mouse movement, key press, speech input, or hand motion. Even though it may be adequate in many cases, the use of a single interaction mode proves to be inept in HCI. For example, in manipulating a [three-dimensional (3-D)] virtual object, a user may employ [two-dimensional (2-D)] mouse motion to select the object, then point with the mouse at a control panel to change the object's color. On the other hand, in a more natural setup, the same user would point at the object with his hand and say: "Make it green." Almost any natural communication among humans involves *multiple, concurrent modes of communication*. Surely, any HCI system that aspires to have the same naturalness should be multimodal. Indeed, studies have shown that people prefer to interact multimodally with computers, since among other things, such interaction eases the need for specialized training [5], [11]. The integration of multimodal input for HCI can also be seen from the perspective of *multisensor data fusion* [12]. Different sensors can, in that case, be related to different communication modalities. It is well known that multiple types of sensors may increase the accuracy with which a quantity can be measured by reducing the uncertainty in decision making [12], [13].

Why Multi Modalities in HCI?

The interaction of humans with their environment (including other humans) is naturally multimodal. We speak about, point at, and look at objects all at the same time. We also listen to the tone of a person's voice and look at a person's face and arm movements to find clues about his feelings. To get a better idea about what is going on around us, we look, listen, touch, and smell. When it comes to HCI, however, we usually use only one interface device at a time—typing, clicking the mouse button, speaking, or pointing with a magnetic wand. The "ease" with which this unimodal interaction allows us to convey our intent to the computer is far from satisfactory. An example of a situation when these limitations become evident is when we press the wrong key or when we have to navigate through a series of menus just to change an object's color. We next discuss the practical, biological, and mathematical rationales that may lead one to consider the use of *multimodal interaction* in HCI

1.1. Aim

Most of the previous work had only one or two modalities integrated in a single system. Observing the need of the people we are integrating three modalities together to make an Advanced Multi Model System i.e. In-Air Signatures, Hand gestures and Voice recognition.

1.2. Related Work

Most of the previous work consisted of one or two integrated system i.e. either Multi touch and gesture, gesture and voice, only gesture etc. The Systems such as Microsoft Surface, iPad, 10/GUI, Display Multi touch Technology and Skinput had one or two integrated systems. We present in this paper a multi model system integrated with Hand Gesture, Voice recognition and In-Air Signature, a step forward in the way of current existing systems.

1.3. Drawback of Existing System

1. Existing systems had only one or two integrated features.
2. The In-Air signature was used for Authentication purpose which was not safe.
3. Mobile was not used as a interface for mouse navigation.
4. External hardwares are required.

1.4. Proposed System

In this article we are focusing on integrating In-Air Signatures, Voice Recognition and Hand Gestures into a single system, which will give us a multi model system. The system is based on dynamic programming, as a method to find the distance between two points and give us a proper result. In this system we are using our mobile phone as a medium of interaction with the system. We connect our mobile phone to the laptop by using WI-FI and within the WI-FI range we can control the mouse operation and game operations using air signature and media Operations using voice input. We are using EUCLIDEAN ALGORITHM to calculate the distance between the samples.

1.1.1. Advantage of Proposed System

1. This paper will tell you about how to integrate different modalities together and make our model efficient to work.
2. We are using WI-FI as the medium of connectivity as the range of WI-FI is more than bluetooth, which is being used as a way of connecting two different devices together from a long time.
3. We are using mobile phone as the medium for gestures because it is readily available with every human being today.
4. No other external hardware is required except a smartphone with our designed application.

2. EXPLANATION OF IMPLEMENTED METHODS

2.1. Testing

Here we are testing our devices i.e. smartphone and the implemented system for errors pertaining in them. We first test our mobile phone for the left, right, up and down values and see whether our system detects the correct values and changes according to it. If any other action is given by the system other than the expected ones signifies the presence of error.

2.2. Callibration

Here we calibrate our system for proper values of each direction i.e. left, right, up and down. We will take the values of each axis i.e. X, Y and Z and match those values with the calibrated values in the system and according to those given values the system will decide on which side the gesture is and which side to move.

2.3. Mouse Control

Here we will check the working of our phone for proper mouse gestures. Here we are trying to control mouse movements and its operations like single click and double click using voice commands and the actions are generated by movement of phone in air with proper hand gestures.

2.4. Game Control

Here we are trying to generate key strokes like forward, backward, and left, right using our mobile phone.

2.5. Media Control/Ppt Control

Here we are controlling media player options like play, pause, next track, previous track etc. by using hand gestures and voice commands. Here we have added a very useful feature i.e. PPT Control which can be useful in our professional activities.

3. ALGORITHM

3.1. Euclidean Algorithm

The immediate consequence of this is that the squared length of a vector **x** = [*x*1 *x*2] is the sum of the squares of its coordinates (see triangle OPA in Figure 2, or triangle OPB – |OP|2 denotes the squared length of **x**, that is the distance between point O and P); and the

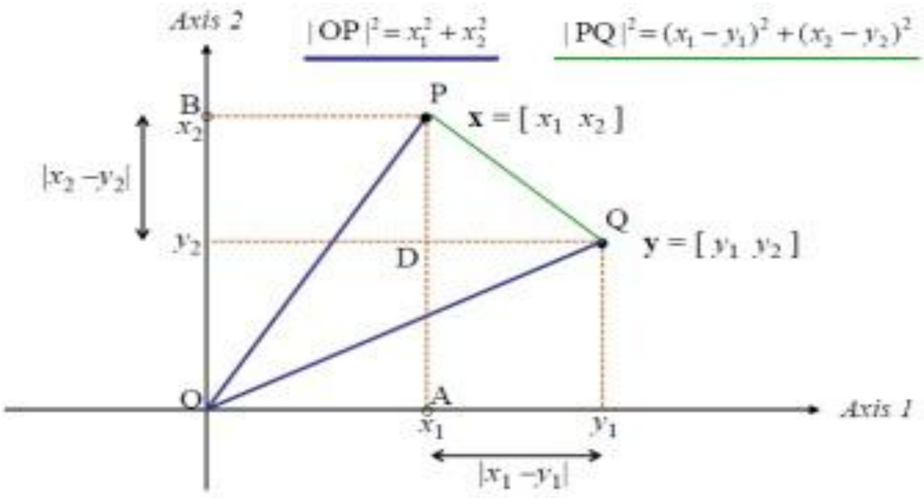


Fig2. *Pythagoras theorem applied to distances in two dimension space*

squared distance between two vectors **x** = [*x*1 *x*2] and **y** = [*y*1 *y*2] is the sum of squared differences in their coordinates (see triangle PQD in Exhibit 4.2; |PQ|2 denotes the squared distance between points P and Q). To denote the distance between vectors **x** and **y** we can use the notation *d* **x,y** so that this last result can be written as:

$$d^2_{\mathbf{x},\mathbf{y}} = (x_1 - y_1)^2 + (x_2 - y_2)^2 \quad (1)$$

That is, the distance itself is the square root

$$d_{\mathbf{x},\mathbf{y}} = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \quad (2)$$

What we called the squared length of **x**, the distance between points P and O in Figure 2, is the distance between the vector **x** = [*x*1 *x*2] and the zero vector **0** = [0 0] with coordinates all zero:

$$d_{\mathbf{x},0} = \sqrt{x_1^2 + x_2^2} \quad (3)$$

Which we could just denote by *d***x.** The zero vector is called the *origin* of the space.

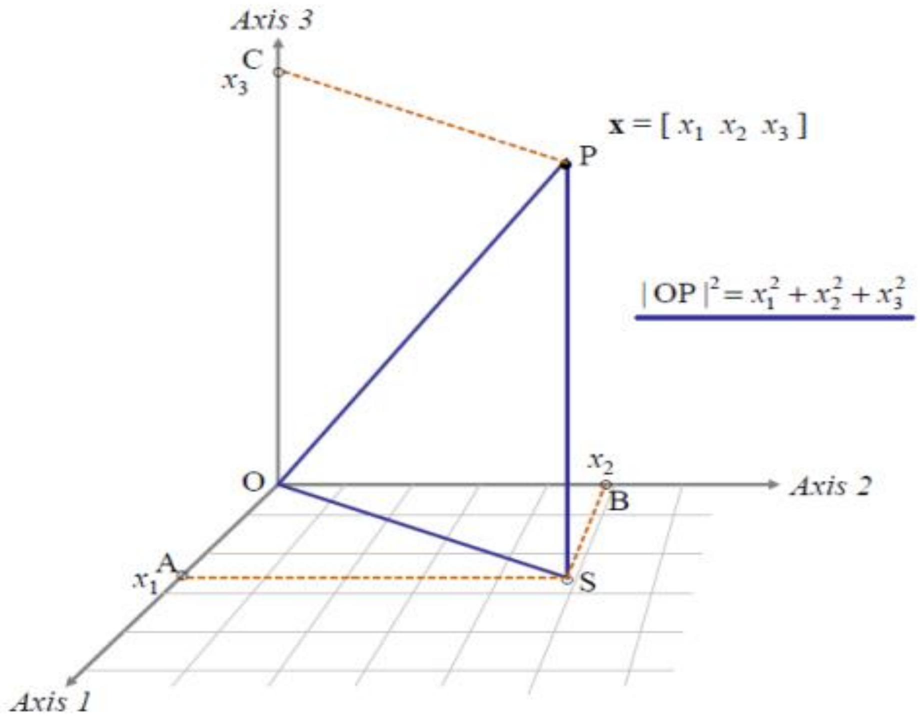


Fig3. *Pythagoras theorem extended in three dimensional space*

We move immediately to a three-dimensional point **x** = [*x*1 *x*2 *x*3], shown in Figure3. This figure has to be imagined in a room where the origin O is at the corner – to reinforce this idea 'floor tiles' have been drawn on the plane of axes 1 and 2, which is the 'floor' of the room. The three coordinates are at points A, B and C along the axes, and the angles AOB, AOC and COB are all 90° as well as the angle OSP at S, where the point P (depicting vector **x**) is projected onto the 'floor'. Using Pythagoras' theorem twice we have:

|OP|2 = |OS|2 + |PS|2 (because of right-angle at S)

|OS|2 = |OA|2 + |AS|2 (because of right-angle at A)

And so

|OP|2 = |OA|2 + |AS|2 + |PS|2

That is, the squared length of x is the sum of its three squared coordinates and so

$$d_{\mathbf{x}} = \sqrt{x_1^2 + x_2^2 + x_3^2} \tag{4}$$

It is also clear that placing a point Q in Figure 3 to depict another vector **y** and going through the motions to calculate the distance between **x** and **y** will lead to

$$d_{\mathbf{x},\mathbf{y}} = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2} \tag{5}$$

Furthermore, we can carry on like this into 4 or more dimensions, in general *J* dimensions, where *J* is the number of variables. Although we cannot draw the geometry any more, we can express the distance between two *J*-dimensional vectors **x** and **y** as:

$$d_{\mathbf{x},\mathbf{y}} = \sqrt{\sum_{j=1}^{J} (x_j - y_j)^2} \quad (6)$$

This well-known distance measure, which generalizes our notion of physical distance in two- or three-dimensional space to multidimensional space, is called the *Euclidean distance* (but often referred to as the ‘Pythagorean distance’ as well).

4. SYSTEM ARCHITECTURE

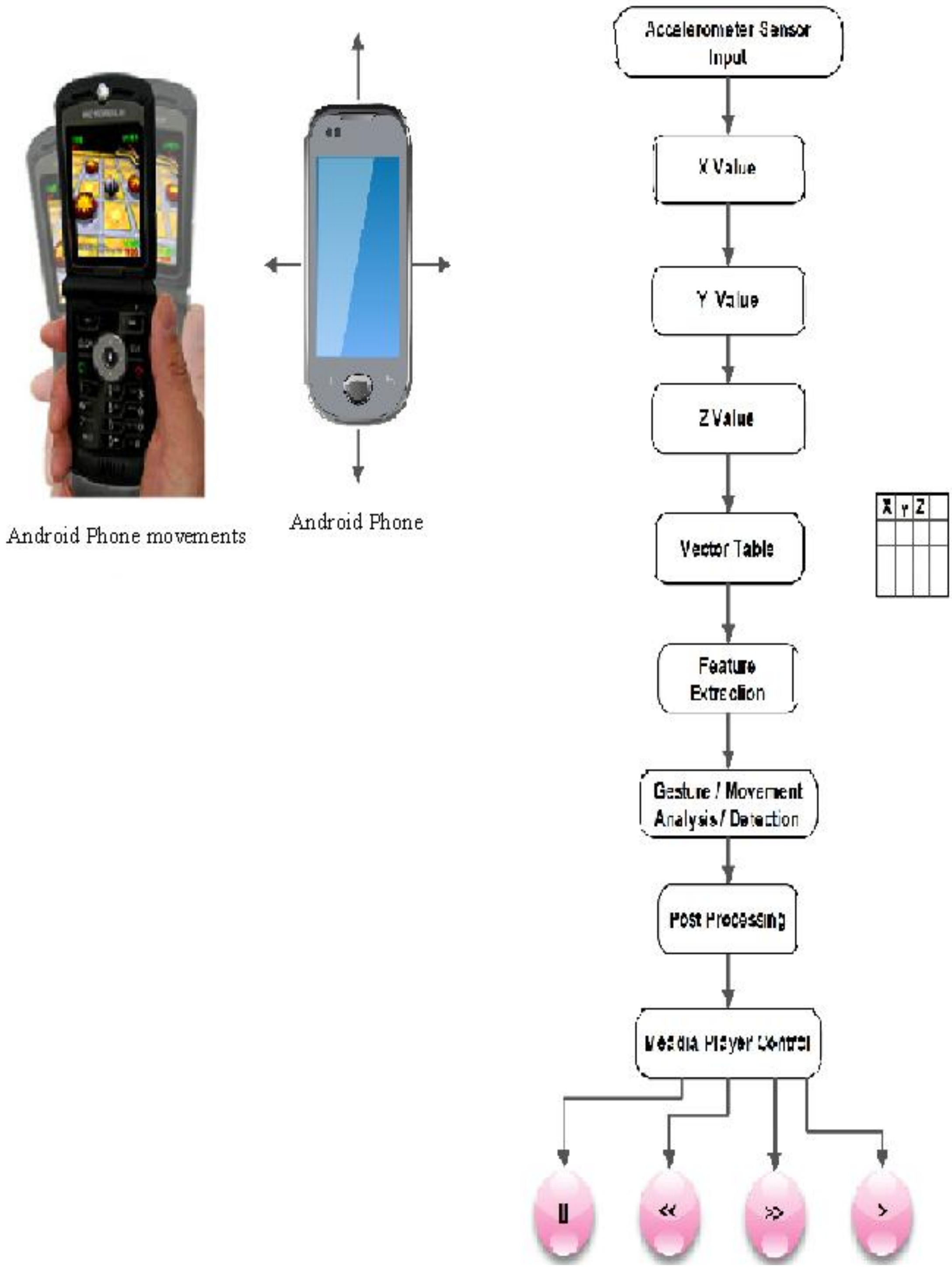


Fig4. *Architecture of Android Application*

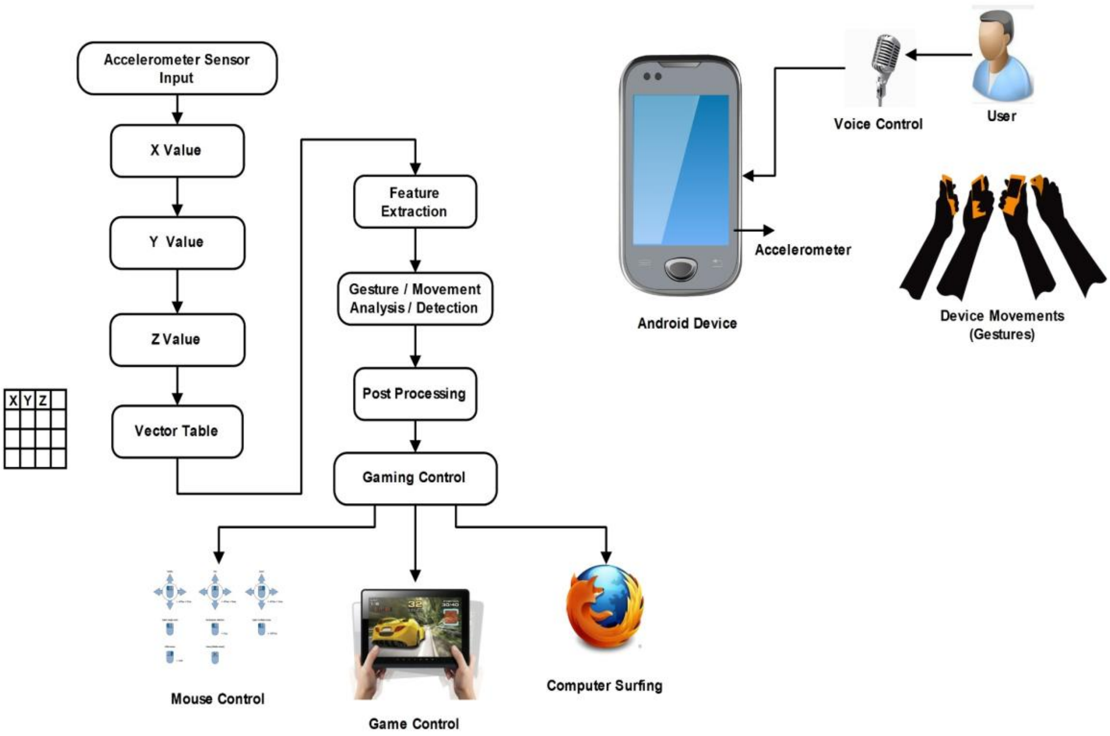


Fig5. *System Architecture*

5. Conclusion

Our system introduces a lot of approaches of new future Human Computer Interaction methods and also devices or prototypes in which these techniques are already in use. We have made an effort to disrupt the use of Mouse and Keyboard. Many new methods are going into the sector of using human hand gestures and even multi modal methods to interact with computer. Many devices in market are using these techniques and many sophisticated methods will push in the market soon. As we are used to act with our hand movements and communicate with our voice, these parts will play a major role in our interaction with computer.

References

[1] J. A. Adam, "Virtual reality," IEEE Spectrum, vol. 30, no. 10, pp. 22–29, 1993.

[2] Time series distances measures to analyze in-air signatures to authenticate users on mobile phones

[3] H. Rheingold, Virtual Reality. New York: Summit Books, 1991.

[4] Toward multimodal human-computer interface

[5] A. G. Hauptmann and P. McAvinney, "Gesture with speech for graphics manipulation," Int. J. Man-Machine Studies, vol. 38, pp. 231–249, Feb. 1993.

[6] Using Mobile Phones to Write in Air

[7] S. Mann, "Wearable computing: A first step toward personal imaging," IEEE Computer Mag., vol. 30, pp. 25–32, Feb. 1997.

[8] Review on Surround Sense Hand Gestures for Mobile Devices

[9] R. W. Pickard and J. Healey, "Affective wearables," in Proc. Int. Symp. Wearable Computing, Cambridge, MA, Oct. 1997.

[10] L. R. Rabiner and B. Juang, Fundamentals of Speech Recognition. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[11] S. Oviatt, A. DeAngeli, and K. Kuhn, "Integration and synchronization of input modes during multimodal human–computer interaction," in Proc. Conf. Human Factors in Computing Systems (CHI'97), Atlanta, GA, pp. 415–422.

[12] D. L. Hall and J. Llinas, "An introduction to multisensor data fusion," Proc. IEEE, vol. 85, pp. 6–23, Jan. 1997.

[13] R. R. Murphy, "Biological and cognitive foundations of intelligent data fusion," IEEE Trans. Syst., Man, Cybern., vol. 26, pp. 42–51, Jan. 1996.

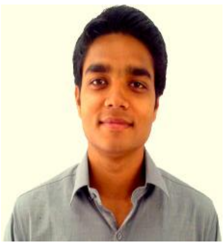
[14] N. Clarke and S. Furnell, "Authenticating mobile phone users using keystroke analysis," International Journal of Information Security,vol. 6, pp. 1–14, 2007.

[15] J. Liu, Z. Wang, L. Zhong, J. Wickramasuriya, and V. Vasudevan, "uWave: Accelerometer-based personalized gesture recognition and its applications," in IEEE PerCom, 2009.

AUTHORS' BIOGRAPHY

Mandar Jangam is pursuing Bachelors of Engineering Degree from RMD Sinhgad School of Engineering, Warje, Pune-58. He is responsible for giving the module ideas.

Jayashree Suryavanshi is Pursuing Bachelors of Engineering Degree for RMD Sinhgad School of Engineering, Warje, Pune – 58. She is responsible for Design analysis and Algorithm of the System.



Aman Gupta is pursuing Bachelors of Engineering Degree from RMD Sinhgad School of Engineering, Warje, Pune-58. He is responsible giving the overall idea of the project.

Nida Shaikh is pursuing Bachelors of Engineering Degree from RMD Sinhgad School of Engineering, Warje, Pune-58. she is responsible Design Analysis and Algorithm of the system.