

# Hand Gesture Recognition Using Radial Length Metric

Warsha M.Choudhari<sup>1</sup>, Pratibha Mishra<sup>2</sup>, Rinku Rajankar<sup>3</sup>, Mausami Sawarkar<sup>4</sup>

<sup>1</sup>Professor, Information Technology, Datta Meghe Institute of Engineering, Technology & Research, Wardha, India

<sup>2</sup>Professor, Information Technology, J.L.Chaturvedi Collage of Engineering, Nagpur, India

<sup>3</sup>Professor, Information Technology G.H.Raisoni Institute of Engineering, Technology for Women, Nagpur, India

<sup>4</sup>Professor, Wireless System & Computing, Tulsiram Gaikwad Patil Collage of Engineering, Nagpur, India

**Abstract:** *A gesture recognition system is such a system that recognizes and differentiates between 'gestures'. These gestures can be any type of facial or body gestures. Various facial expressions constitute facial gestures. Similarly the various gestures that can be made using our hand, or the palm to be more specific are called 'HAND GESTURES' Sign languages are the most raw and natural form of languages could be dated back to as early as the advent of the human civilization, when the first theories of sign languages appeared in history. It has started even before the emergence of spoken languages. Since then the sign language has evolved and been adopted as an integral part of our day to day communication process. Now, sign languages are being used extensively in international sign use of deaf and dumb, in the world of sports, for religious practices and also at work places. Gestures are one of the first forms of communication when a child learns to express its need for food, warmth and comfort.*

**Keywords:** computer vision, gesture recognition, radial length, stereoscopic displays, tactile information

## 1. Introduction

The keyboard and mouse are currently the main interfaces between man and computer. In other areas where 3D information is required, such as computer games, robotics and design, other mechanical devices such as roller-balls, joysticks and data-gloves are used. Humans communicate mainly by vision and sound, therefore, a man-machine interface would be more intuitive if it made greater use of vision and audio recognition. Another advantage is that the user not only can communicate from a distance, but need have no physical contact with the computer. However, unlike audio commands, a visual system would be preferable in noisy environments or in situations where sound would cause a disturbance. The visual system chosen was the recognition of hand gestures. The amount of computation required to process hand gestures is much greater than that of the mechanical devices.

Hand gesture recognition using computer vision — a viable proposition.

- **Man-machine interface:** using hand gestures to control the computer mouse and/or keyboard functions. Eg.controls various keyboard and mouse functions using gestures alone.
- **3D animation:** Rapid and simple conversion of hand movements into 3D computer space for the purposes of computer animation.
- **Visualization:** Just as objects can be visually examined by rotating them with the hand, so it would be advantageous if virtual 3D objects (displayed on the computer screen) could be manipulated by rotating the hand in space.
- **Computer games:** Using the hand to interact with computer games would be more natural for many applications.
- **Control of mechanical systems (such as robotics):** Using the hand to remotely control a manipulator.

## 2. Optical Data Collection Method

Since the hand is by nature a three dimensional object the first optical data collection method considered was a stereographic multiple camera system. Alternatively, using prior information about the anatomy of the hand it would be possible to garner the same gesture information using either a single camera or multiple two dimensional views provided by several cameras. These three options are considered below:

- **Stereographic system:** The stereographic system would provide pixilated depth information for any point in the fields of view of the cameras. This would provide a great deal of information about the hand. Features that would otherwise be hard to distinguish using a 2D system, such as a finger against a background of skin, would be differentiable since the finger would be closer to camera than the background. However the 3D data would require a great deal of processor time to calculate and reliable real-time stereo algorithms are not easily obtained or implemented.
- **Multiple two dimensional view system:** This system would provide less information than the stereographic system and if the number of cameras used was not great, would also use less processor time. With this system two or more 2D views of the same hand, provided by separate cameras, could be combined after gesture recognition. Although each view would suffer from similar problems to that of the "finger" the combined views of enough cameras would reveal sufficient data to approximate any gesture.
- **Single camera system:** This system would provide considerably less information about the hand. Some would be very hard to distinguish since no depth information would be recoverable. Essentially only "silhouette" information could be accurately extracted.

The silhouette data would be relatively noise free (given a background sufficiently distinguishable from the hand) and would require considerably less processor time to compute than either multiple camera system. It is possible to detect a large subset of gestures using silhouette information alone and the single camera system is less noisy, expensive and processor hungry. Although the system exhibits more ambiguity than either of the other systems.

### 3. Hardware Setup

- **Lighting:** The task of differentiating the skin pixels from those of the background and markers is made considerably easier by a careful choice of lighting. If the lighting is constant across the view of the camera then the effects of self-shadowing can be reduced to a minimum. The intensity should also be set to provide sufficient light for the CCD in the camera. However, since this system is intended to be used by the consumer it would be a disadvantage if special lighting equipment was required. It was decided to attempt to extract the hand and marker information using standard room lighting. This would permit the system to be used in a non-specialist environment.

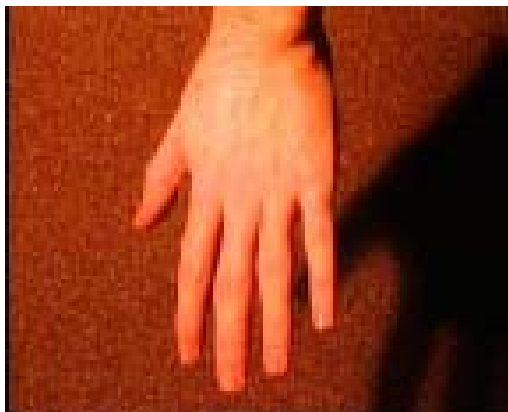


Figure 3.1: The effect of self shadowing

- **Camera orientation:** It is important to carefully choose the direction in which the camera points to permit an easy choice of background. The two realistic options are to point the camera towards a wall or towards the floor (or desktop). However since the lighting was a single overhead bulb, light intensity would be higher and shadowing effects least if the camera was pointed downwards.

- **Background:** In order to maximize differentiation it is important that the color of the background differs as much as possible from that of the skin. The floor color in the project room was a dull brown. It was decided that this color would suffice initially.

### 4. Gesture Based Applications

Gesture based applications are broadly classified into two groups on the basis of their purpose: multidirectional control and a symbolic language.

**3D Design:** CAD (computer aided design) is an HCI which provides a platform for interpretation and manipulation of 3-Dimensional inputs which can be the gestures. Manipulating 3D inputs with a mouse is a time consuming task as the task involves a complicated process of decomposing a six degree freedom task into at least three sequential two degree tasks has come up with the 3DRAW technology that uses a pen embedded in polhemus device to track the pen position and orientation in 3D. A 3space sensor is embedded in a flat palette, representing the plane in which the objects rest. The CAD model is moved synchronously with the users gesture movements and objects can thus be rotated and translated in order to view them from all sides as they are being created and altered.

**Tele presence:** There may raise the need of manual operations in some cases such as system failure or emergency hostile conditions or inaccessible remote areas. Often it is impossible for human operators to be physically present near the machines. Tele presence is that area of technical intelligence which aims to provide physical operation support that maps the operator arm to the robotic arm to carry out the necessary task, for instance the real time ROBOGEST system constructed at University of California, San Diego presents a natural way of controlling an outdoor autonomous vehicle by use of a language of hand gestures [1]. The prospects of tele presence includes space, undersea mission, medicine manufacturing and in maintenance of nuclear power reactors.

**Virtual reality:** Virtual reality is applied to computer-simulated environments that can simulate physical presence in places in the real world, as well as in imaginary worlds. Most current virtual reality environments are primarily visual experiences, displayed either on a computer screen or through special stereoscopic displays. There are also some simulations include additional sensory information, such as sound through speakers or headphones. Some advanced, haptic systems now include tactile information, generally known as force feedback, in medical and gaming applications.

**Sign Language:** Sign languages are the most raw and natural form of languages could be dated back to as early as the advent of the human civilization, when the first theories of sign languages appeared in history. It has started even before the emergence of spoken languages. Since then the sign language has evolved and been adopted as an integral part of our day to day communication process. Now, sign languages are being used extensively in international sign use of deaf and dumb, in the world of sports, for religious

practices and also at work places. Gestures are one of the first forms of communication when a child learns to express its need for food, warmth and comfort. It enhances the emphasis of spoken language and helps in expressing thoughts and feelings effectively [2].

A simple gesture with one hand has the same meaning all over the world and means either 'hi' or 'goodbye'. Many people travel to foreign countries without knowing the official language of the visited country and still manage to perform communication using gestures and sign language. These examples show that gestures can be considered international and used almost all over the world. In a number of jobs around the world gestures are means of communication.

In airports, a predefined set of gestures makes people on the ground able to communicate with the pilots and thereby give directions to the pilots of how to get off and on the run-way and the referee in almost any sport uses gestures to communicate his decisions. In the world of sports gestures are common. The pitcher in baseball receives a series of gestures from the coach to help him in deciding the type of throw he is about to give. Hearing impaired people have over the years developed a gestural language where all defined gestures have an assigned meaning. The language allows them to communicate with each other and the world they live in

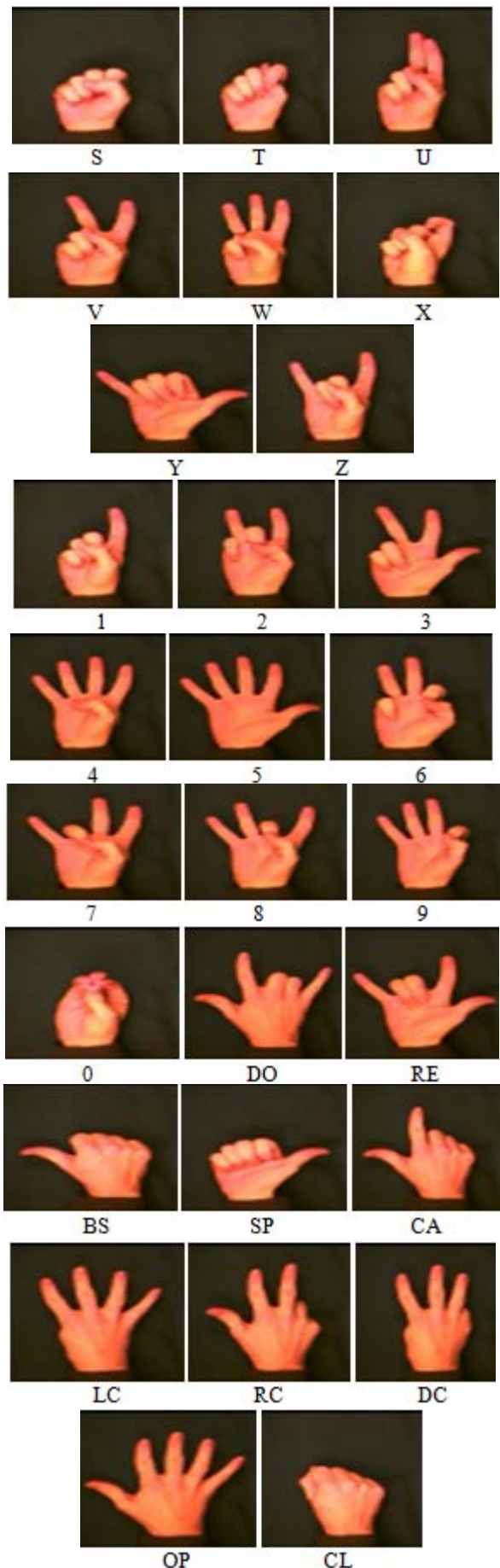
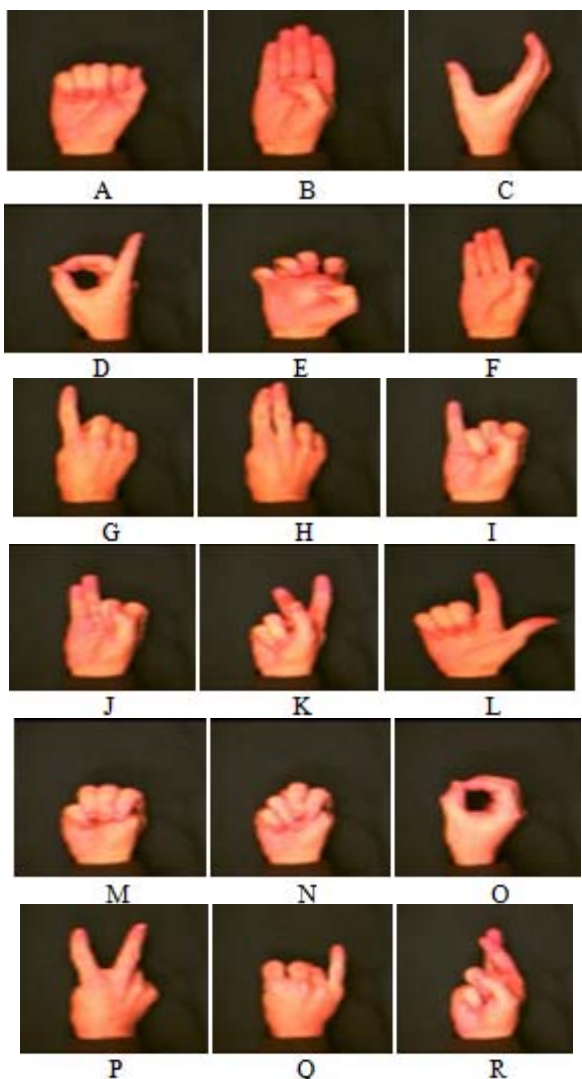


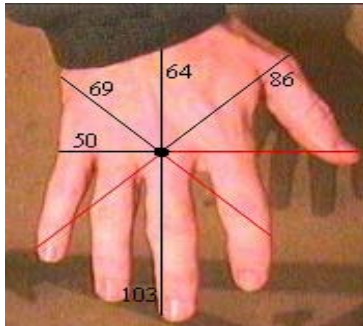
Figure 4.1: Natural form of Sign languages



## 5. Recognition Method

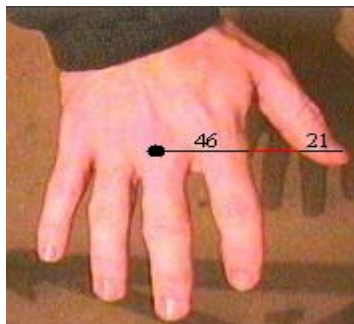
### 5.1 Re-evaluation of radial length metric

A simple method to assess the gesture would be to measure the distance from the hand centroid to the edges of the hand along a number of radials equally spaced around a circle. This would provide information on the general “shape” of the gesture that could be easily rotated to account for hand yaw (since any radial could be used as datum).



**Figure 5.1:** Gesture with radials marked. The black radial lengths can easily be measured (length in pixels shown). However, the red radials present a problem in that they either cross between fingers or palm and finger

However, a problem is how to measure when the radial crosses a gap between fingers or between the palm and a finger. To remedy this it was decided to count the total number of skin pixels along a given radial.



**Figure 5.2:** One of the problem radials with outlined solution. If only the skin pixels along any given radial are counted then the sum is the “effective” length of that radial. In this case the radial length is  $46 + 21 = 67$ .

All of the radial measurements could then be scaled so that the longest radial was of constant length. By doing this, any alteration in the hand camera distance would not affect the radial length signature generated.

### 6. Evaluation of Radial Length Metric

To evaluate this method a program was written to calculate the radial length signature of a given gesture and display it in the form of a histogram.

The measurement is not affected by hand-to-camera distance. The measurement is affected by the yaw of the hand, but this only shifts the readings to the left or right and does not affect their shape.

A formal description of radial length calculation is as follows:

Examine a typical radial at angle the score for that radial is:

$$radscore(\theta) = \sum_{x \in R_\theta} S(x)$$

where  $S()$  is the skin pixel predicate defined earlier and where

$$R_\theta = \left\{ (x, y) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} c_x \\ c_y \end{pmatrix} + r \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} \forall r > 0 \right\}$$

The signature for a given gesture  $g$  could then be calculated as:

$$g = \left\{ \frac{radscore(\theta)}{\max_\alpha (radscore(\alpha))} \mid 0 \leq \theta < 2\pi \right\}$$



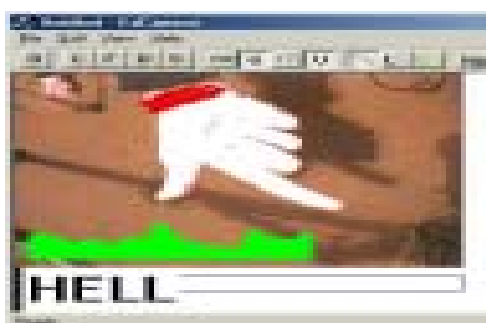
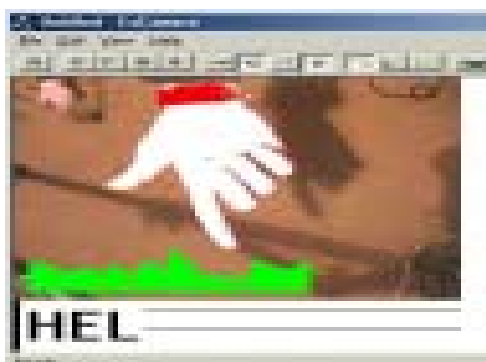
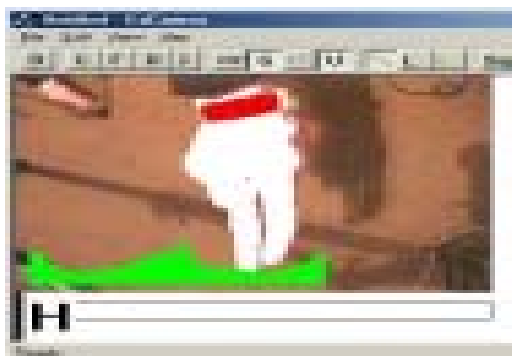
**Figure 6.1:** Images showing the histogram for two different gestures. The two histograms are sufficiently different to permit differentiation

### 7. Removing the Hand Yaw Degree of Freedom

In order to counter the shifting effect of hand yaw, a wrist marker was used. The angle between the centroid of this marker and the centroid of the hand was then used as the initial radial direction. This, along with the maximum radial length scaling makes the system robust against changes in hand position, yaw and distance from camera.

The radial measurements are very similar no matter how the hand is positioned. Using this improved system the sign language letters ‘a’ through to ‘o’ were taught to the system.

This enabled a very limited sign language “word processor” to be made



## 8. Conclusion

Radial Length Metric method involved the comparison of radial length signatures. This was more suitable, but it was found that the amount of information provided about individual fingers was dependent on the relative angle of the radial and the long axis of the finger, making some gestures hard to differentiate.

## References

- [1] G. Charvat, L. Kempel, E. Rothwell, C. Coleman, and E. Mokole. A Through-dielectric Radar Imaging System. In Trans. Antennas and Propagation, 2010.
- [2] R. Bowden, D. Windridge, T. Kadir, A. Zisserman, and M. Brady, “A linguistic feature vector for the visual interpretation of sign language,” in Proc. 8th Eur. Conf. Comput. Vis., New York: Springer-Verlag, 2004, pp. 391–401.
- [3] Bauer & Hienz, Relevant feature for video- based continuous sign language recognition. Department of Technical Computer Science, Aachen University of Technology, Aachen, Germany, 2000.
- [4] R. Block. Toshiba Qosmio G55 features SpursEngine, Visual Gesture Controls.