**SIXTH SENSE TECHNOLOGY**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Pranav Kondalkar (Ex 6th sem)

Sudhakar Tembhare(ExTc 6th sem)

Akansha Lingayat (ExTc 6th sem)

**Abstract:**

***“Invent is an attempt to make programming more like thinking.”***

*In brief, Invent=‘Imagine...Explore...&...Learn'.*

*We've evolved over millions of years to sense the world around us. When we encounter*

*something, someone or some place, we use our five natural senses to perceive information*

*about it; that information helps us make decisions and chose the right actions to take. But*

*arguably the most useful information that can help us make the right decision is not naturally*

*perceivable with our five senses, namely the data, information and knowledge that mankind has*

*accumulated about everything and which is increasingly all available online. Although the*

*miniaturization of computing devices allows us to carry computers in our pockets, keeping us*

*continually connected to the digital world, there is no link between our digital devices and our*

*interactions with the physical world. Information is confined traditionally on paper or digitally*

*on a screen. Sixth Sense bridges this gap, bringing intangible, digital information out into the*

*tangible world, and allowing us to interact with this information via natural hand gestures.*

*‘Sixth Sense’ frees information from its confines by seamlessly integrating it with reality, and*

*thus making the entire world your computer.*

*'Sixth Sense' is a wearable gestural interface that augments the physical world around*

*us with digital information and lets us use natural hand gestures to interact with that*

*information. The Sixth Sense prototype is comprised of a pocket projector, a mirror and a*

*camera. The hardware components are coupled in a pendant like mobile wearable device. Both*

*the projector and the camera are connected to the mobile computing device in the user’s pocket.*

*The projector projects visual information enabling surfaces, walls and physical objects around*

*us to be used as interfaces; while the camera recognizes and tracks user's hand gestures and*

*physical objects using computer-vision based techniques. The software program processes the*

*video stream data captured by the camera and tracks the locations of the colored markers*

*(visual tracking fiducials) at the tip of the user’s fingers using simple computer-vision*

*techniques. The movements and arrangements of these fiducials are interpreted into gestures*

*that act as interaction instructions for the projected application interfaces. The maximum*

*number of tracked fingers is only constrained by the number of unique fiducials, thus Sixth*

*Sense also supports multi-touch and multi-user interaction.*

*Why sixth sense? It simply culls data based on information- which is still limited- our*

*inherent nature knows more than information, as information is limited to words and*

*descriptions and senses go beyond these.*

**Keywords**

Marker caps, Camera, Battery powered Micro projector, Mirror, Cell Phone, Image

recognization, Marker technology, Fluid Interference Group.

**Introduction:**

We have the basic five senses in

our human body that constitute of the nose

(smell), ear (hear), tongue (taste), eye

(vision), skin (feel).This sensing device in

our body sense what is around us but

suppose in a case we come to condition

were the thing we want to sense is not

familiar to us for that very case we need a

sense that has the huge amount of data

stored in it. In today’s world that much

huge data is available only at one place

that is internet. Suppose we have a sense in

our body whose data constitute the internet

than consider how useful it would be that

is what the 6th sense is.

Although the miniaturization of

computing devices allows us to carry

computers in our pockets, keeping us

continually connected to the digital world,

there is no link between our digital devices

and our interactions with the physical

world. Information is confined

traditionally on paper or digitally on a

screen. Sixth Sense bridges this gap,

bringing intangible, digital information out

into the tangible world, and allowing us to

interact with this information via natural

hand gestures. ‘Sixth Sense’ frees

information from its confines by

seamlessly integrating it with reality, and

thus making the entire world your

computer. This technology is also called

by its developer as WEAR YOUR

WORLD (WYW).

This technology constitute the

basic components that are Marker caps,

Camera, Battery powered Micro projector,

Mirror, Cell Phone. The hardware

components (Camera, Battery powered

Micro projector, Mirror) are coupled in a

pendant like mobile wearable device. Both

the projector and the camera are connected

to the mobile computing device in the

user’s pocket. The projector projects visual

information enabling surfaces, walls and

physical objects around us to be used as

interfaces; while the camera recognizes

and tracks user's hand gestures and

physical objects using computer-vision

based techniques. The software program

processes the video stream data captured

by the camera and tracks the locations of

the colored markers (visual tracking

fiducials) at the tip of the user’s fingers

using simple computer-vision techniques.

The movements and arrangements of these

fiducials are interpreted into gestures that

act as interaction instructions for the

projected application interfaces. The

maximum number of tracked fingers is

only constrained by the number of unique

fiducials, thus Sixth Sense also supports

multi-touch and multi-user interaction.

**About the Components**

**Marker Caps**:

Marker caps are specially designed colored

finger caps that are designed to wear in

fingers. They come in different color

formats are used to identify one from

another for this technology we can also

color our finger in different colors. The

main motto of using colored finger process

is during the process of image processing

of what is in front of the camera it must

identify all the finger instruction that are

given by the user such as to draw some

thing or to make a rectangular frame for

taking the picture of a place and so many

instructions . Due to different color in

different finger the camera can recognize

what the instructions are by identifying the

pixels that are on the finger than to the

instruction given by the finger is identified

as if we draw a wrist watch on our hand

than by identifying the pixels movement

identifies that the thing drawn is a circular

object and as is was drawn at the wrist

position hence it is a watch. Again as in

front the camera someone makes a

snapshot icon due to the variation a pixels

color the camera identifies the marker caps

fingers and finally the camera takes the

snapshot of what is between the frames.

Even with the help of marker caps the

camera can recognize minimize and

maximizing finger instructions. The

software program processes the video

stream data captured by the camera and

tracks the locations of the colored markers

(visual tracking fiducials) at the tip of the

user’s fingers using simple computervision

techniques. The movements and

arrangements of these fiducials are

interpreted into gestures that act as

interaction instructions for the projected

application interfaces. The maximum

number of tracked fingers is only

constrained by the number of unique

fiducials, thus Sixth Sense also supports

multi-touch and multi-user interaction.

**Camera**

This is the most vital component in the

technology that has evolved or is under

discussion. There is a high quality camera

with the computer vision features and a

very quick image processing and image

recogniziation capability. It identifies

everything that is kept in front of it. It is

connected to the mobile computing device

in the user’s pocket. It recognizes and

tracks user's hand gestures and physical

objects using computer-vision based

techniques.

**Battery Powered Micro Projector**

The projector projects visual information

enabling surfaces, walls and physical

objects around us to be used as interfaces**.**

The physical surface can be anything that

is it can be a wall, a table, stone anything

and if we didn’t get anything we can

project the information on our hand.

**Mirror**

It is used to project at user desired location

with the help of the projector.

**Cell phone**

Cell phone or a mobile computing device

that is connected with internet and the

identified object by the camera is than

search by the mobile computing device

over the internet and useful result are

passed to the projector which project the

information for the user the giving him the

detailed view what it is.

**Technologies Involved**

**Computer-Vision Techniques:**

**“Processing of large data sets such as**

**images”**

Computer vision is the science and

technology of machines that see. As a

scientific discipline, computer vision is

concerned with the theory for building

artificial systems that obtain information

from images. The image data can take

many forms, such as a video sequence,

views from multiple cameras, or multidimensional

data from a medical scanner.

As a technological discipline,

computer vision seeks to apply the theories

and models of computer vision to the

construction of computer vision systems.

Examples of applications of computer

vision systems include systems for:

Controlling processes (e.g. an

industrial robot or an autonomous

vehicle).

Detecting events (e.g. for visual

surveillance or people counting).

Organizing information (e.g. for

indexing databases of images and

image sequences).

Modeling objects or environments

(e.g. industrial inspection, medical

image analysis or topographical

modeling).

Interaction (e.g. as the input to a

device for computer-human

interaction).

Computer vision can also be described as a

complement (but not necessarily the

opposite) of biological vision. In

biological vision, the visual perception of

humans and various animals are studied,

resulting in models of how these systems

operate in terms of physiological

processes. Computer vision, on the other

hand, studies and describes artificial vision

system that is implemented in software

and/or hardware. Interdisciplinary

exchange between biological and computer

vision has proven increasingly fruitful for

both fields.

Sub-domains of computer vision include

scene reconstruction, event detection,

tracking, object recognition, learning,

indexing, motion estimation, and image

restoration**.**

Typical tasks of computer vision

Recognition

The classical problem in computer vision,

image processing and machine vision is

that of determining whether or not the

image data contains some specific object,

feature, or activity. This task can normally

be solved robustly and without effort by a

human, but is still not satisfactorily solved

in computer vision for the general case:

arbitrary objects in arbitrary situations.

The existing methods for dealing with this

problem can at best solve it only for

specific objects, such as simple geometric

objects (e.g., polyhedrons), human faces,

printed or hand-written characters, or

vehicles, and in specific situations,

typically described in terms of welldefined

illumination, background, and

pose of the object relative to the camera.

Different varieties of the recognition

problem are

*Object recognition*: one or several prespecified

or learned objects or object

classes can be recognized, usually together

with their 2D positions in the image or 3D

poses in the scene.

*Identification*: An individual instance of an

object is recognized. Examples:

identification of a specific person's face or

fingerprint, or identification of a specific

vehicle.

*Detection*: the image data is scanned for a

specific condition. Examples: detection of

possible abnormal cells or tissues in

medical images or detection of a vehicle in

an automatic road toll system. Detection

based on relatively simple and fast

computations is sometimes used for

finding smaller regions of interesting

image data which can be further analyzed

by more computationally demanding

techniques to produce a correct

interpretation.

*Content-based image retrieval*: finding all

images in a larger set of images which

have a specific content. The content can be

specified in different ways, for example in

terms of similarity relative a target image

(give me all images similar to image X), or

in terms of high-level search criteria given

as text input (give me all images which

contains many houses, are taken during

winter, and have no cars in them).

*Pose estimation*: estimating the position or

orientation of a specific object relative to

the camera. An example application for

this technique would be assisting a robot

arm in retrieving objects from a conveyor

belt in an assembly line situation.

*Optical character recognition* (or OCR):

identifying characters in images of printed

or handwritten text, usually with a view to

encoding the text in a format more

amenable to editing or indexing (e.g.

ASCII).

Motion

Several tasks relate to motion estimation,

in which an image sequence is processed

to produce an estimate of the velocity

either at each point in the image or in the

3D scene. Examples of such tasks are:

*Egomotion*: determining the 3D rigid

motion of the camera.

*Tracking*: following the movements of

objects (e.g. vehicles or humans).

Scene reconstruction

Given one or (typically) more images of a

scene, or a video, scene reconstruction

aims at computing a 3D model of the

scene. In the simplest case the model can

be a set of 3D points. More sophisticated

methods produce a complete 3D surface

model.

Image restoration

The aim of image restoration is the

removal of noise (sensor noise, motion

blur, etc.) from images. The simplest

possible approach for noise removal is

various types of filters such as low-pass

filters or median filters. More sophisticated

methods assume a model of how the local

image structures look like, a model which

distinguishes them from the noise. By first

analyzing the image data in terms of the

local image structures, such as lines or

edges, and then controlling the filtering

based on local information from the

analysis step, a better level of noise

removal is usually obtained compared to

the simpler approaches.

Systems

The organization of a computer vision

system is highly application dependent.

Some systems are stand-alone applications

which solve a specific measurement or

detection problem, while other constitute a

sub-system of a larger design which, for

example, also contains sub-systems for

control of mechanical actuators, planning,

information databases, man-machine

interfaces, etc. The specific

implementation of a computer vision

system also depends on if its functionality

is pre-specified or if some part of it can be

learned or modified during operation.

There are, however, typical functions

which are found in many computer vision

systems.

*Image acquisition*: A digital image is

produced by one or several image sensors,

which, besides various types of lightsensitive

cameras, include range sensors,

tomography devices, radar, ultra-sonic

cameras, etc. Depending on the type of

sensor, the resulting image data is an

ordinary 2D image, a 3D volume, or an

image sequence. The pixel values typically

correspond to light intensity in one or

several spectral bands (gray images or

color images), but can also be related to

various physical measures, such as depth,

absorption or reflectance of sonic or

electromagnetic waves, or nuclear

magnetic resonance.

*Pre-processing*: Before a computer vision

method can be applied to image data in

order to extract some specific piece of

information, it is usually necessary to

process the data in order to assure that it

satisfies certain assumptions implied by

the method. Examples are

Re-sampling in order to assure that

the image coordinate system is

correct.

Noise reduction in order to assure

that sensor noise does not introduce

false information.

Contrast enhancement to assure

that relevant information can be

detected.

Scale-space representation to

enhance image structures at locally

appropriate scales.

*Feature extraction*: Image features at

various levels of complexity are extracted

from the image data. Typical examples of

such features are

Lines, edges and ridges.

Localized interest points such as

corners, blobs or points.

.

*Detection/Segmentation*: At some point in

the processing a decision is made about

which image points or regions of the

image are relevant for further processing.

Examples are

Selection of a specific set of

interest points

Segmentation of one or multiple

image regions which contain a

specific object of interest.

*High-level processing*: At this step the

input is typically a small set of data, for

example a set of points or an image region

which is assumed to contain a specific

object. The remaining processing deals

with, for example:

Verification that the data satisfy

model-based and application

specific assumptions.

Estimation of application specific

parameters, such as object pose or

object size.

Classifying a detected object into

different categories

**Visual Tracking Fiducials**

Fiducials is an object used in the

field of view of an imaging system which

appears in the image produced, for use as a

point of reference or a measure.

**Applications**

Drawing anything on a physical

object without touching it.

Making projected object move on

finger instruction.

Capturing image by just making

snapshot icon in front of what we

want to capture.

Maximizing and minimizing an

image.

Collecting the entire captured

image a rearrange them.

Went to a shopping mall and want

to know about the quality of

product.

Meet a distance friend and don’t

recall about him. You can get all

his information that he has posted

on the internet as his blog.

Want to know experts rating about

a book.

Want to know the status of your

flight.

Check out your mail on a wall or

any physical object.

Make hand as your cell phone

keypad and use it to dial some one

Want to see time in a virtual watch.

**Conclusion**

A technology developed by the fluid

interference group at the MIT, California

.It was launched on Feb 2009 in California

by the director of the MIT. The device that

works on the above technology cost only

$350. The main credit for this innovation

goes to a Indian PhD student name Pranav

Mistry who is currently doing his PhD in

design from MIT. The computational

power of the brain. The brain is the most

efficient device as far as Image processing

and pattern recognition is concerned. But

our brain contains only a limited set of

data but if the data base for image

processing and pattern recognition is the

internet than our efficiency can be

unlimited to the utilization of the internet

data. That is how efficiently we can utilize

the internet information the more efficient

device would be.

There are many more chapter

regarding to this can be uncovered

regarding to this and this will come in to

light as we start using it. The most

important features of the device is it

handiness that is it is very handy.

**References**

Sixthsense. Pranav Mistry.

CNet News: MIT's 6th Sense device could trump Apple's multitouch

New York Times - At TED, Virtual Worlds Collide With Reality

TED Talks - Pattie Maes & Pranav Mistry: Unveiling the "Sixth Sense," gamechanging

wearable tech

Keith Price's Annotated Computer Vision Bibliography and the Official Mirror Site

Keith Price's Annotated Computer Vision Bibliography

USC Iris computer vision conference list

Pedram Azad, Tilo Gockel, Rüdiger Dillmann (2008). Computer Vision - Principles

and Practice. Elektor International Media BV. ISBN 0905705718.

Dana H. Ballard and Christopher M. Brown (1982). Computer Vision. Prentice Hall.

ISBN 0131653164.

Fay Huang, Reinhard Klette and Karsten Scheibe (2008). Panoramic Imaging -

Sensor-Line Cameras and Laser Range-Finders. Wiley. ISBN 978-0-470-06065-0.

Bernd Jähne and Horst Haußecker (2000). Computer Vision and Applications, A

Guide for Students and Practitioners. Academic Press. ISBN 0-13-085198-1.

Bernd Jähne (2002). Digital Image Processing. Springer. ISBN 3-540-67754-2.

Reinhard Klette, Karsten Schluens and Andreas Koschan (1998). Computer Vision -

Three-Dimensional Data from Images. Springer, Singapore. ISBN 981-3083-71-9.

Tony Lindeberg (1994). Scale-Space Theory in Computer Vision. Springer. ISBN 0-

7923-9418-6.