VISION BASED MAN-MACHINE INTERACTION

Aykut & Erkut ERDEM

Presentation Overview

- Introduction
- Current State of Art and Example Projects
- Projects
  - Vision-Based Single-Stroke Character Recognition for Wearable Computing
  - Computer Vision-Based Mouse
- Future Projects

Introduction

- As the role of computers in our daily life increases, it is expected that many computer systems will be embedded into our environment. These systems must provide new types of human-computer-interaction with interfaces that are more natural and easy to use.
- Today, the keyboard, the mouse and the remote control are used as the main interfaces for transferring information and commands to computerized equipment. In some applications involving three-dimensional information, such as visualization, computer games and control of robots, other interfaces based on trackballs, joysticks and datagloves are being used.

Introduction (Cont'd.)

In examining how effective an input device is, we should take into consideration the following properties:

1. **Input speed** The rate at which characters can be typed, usually given in units of characters per second or minute, or words per minute.
2. **Error rate** The number of errors, usually given in terms of errors per hundred characters.
3. **Learning rate** The speed at which one can learn to use a specific input device.
4. **Fatigue** How tired a user becomes during a session of typing
5. **Muscle strain** How much and where the strain is put on the muscles. Repetitive motions injuries can also be classified here.

Introduction (Cont'd.)

6. **<u>Portability</u>** Taking into account mass, ease of carrying, and use in a public environment, how small or uncumbersome a device can get and still be useable.
7. **<u>User preferences</u>** How readily will users give up their old interface and use a new one?

Note that all the devices listed before suffers from at least one of these properties.

Introduction (Cont'd.)

- Besides in our daily life, we humans use our vision and hearing as main sources of information about our environment. Therefore, one may ask to what extent it would be possible to develop computerized equipment able to communicate with humans in a similar way, by understanding visual and auditive input.

- So, nowadays there is a trend for interacting with computers without need for special external equipments like the keyboard and the mouse.

Introduction (Cont'd.)

- An alternative way of such human-computer interaction systems are vision based man-machine communication systems. These methods provide both conventional and unconventional means for entering data into computers. Application areas of vision based text and data entry systems include regular computers as well as wearable computing in which flexible and versatile man-machine communication systems other than the ordinary tools of keyboard and mouse may be needed.

- Computer vision based man-machine communication systems can be developed by taking advantage of the character recognition systems developed in document analysis and image analysis methods.

Current State of Art

- Current researches are mainly concentrated on visual hand/face gesture recognition and tracking systems. The outcome these projects are new vision-based user interfaces and new type of input and control devices.

Example Projects

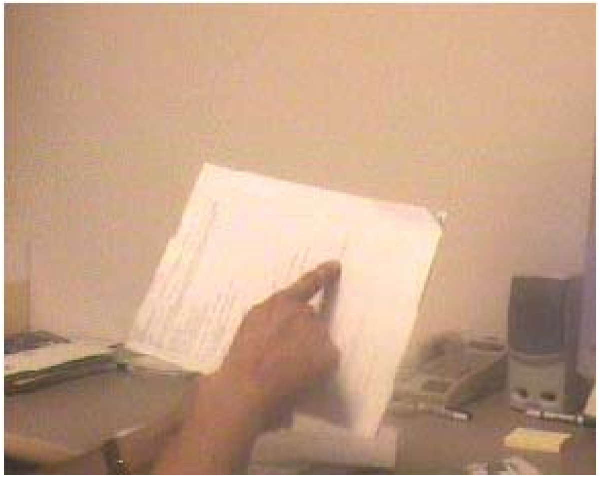
- **Visual Panel**, *Microsoft Research*
- **Visual Gesture Research**, *University of Illinois at Urbana-Champaign*
- **A Prototype System for Computer Vision Based Human Computer Interaction**, *KTH (Royal Institute of Technology), Sweden*
- **The Perceptual Browser (PBrowser)**, *IIHM Group, CLIPS-IMAG Lab, France*
- **FingerPaint**, *IIHM Group, CLIPS-IMAG Lab, France*

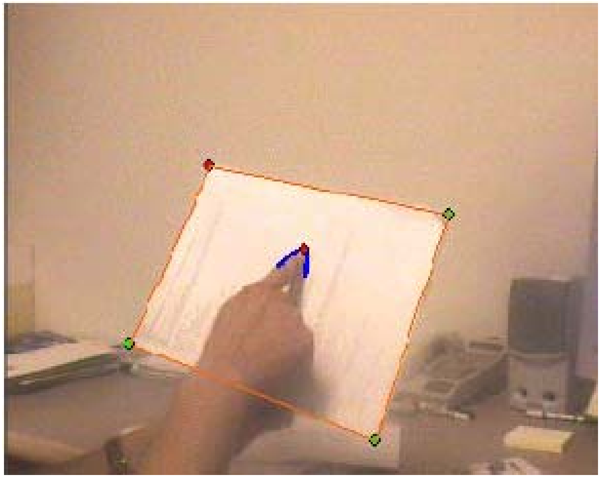
Visual Panel

- *VisualPanel* employs an arbitrary quadrangle-shape panel and a tip pointer like fingertip as an intuitive input device. Taking advantage of the panel, the system can fulfill many UI tasks such as controlling a remote and large display, and simulating a physical keyboard.

- Users can naturally use their fingers or other tip pointers to issue commands and type texts. The system is facilitated by accurately and reliably tracking the panel and the tip pointer and detecting the clicking and dragging actions.

Visual Panel (Cont'd.)





Tracking the panel and the fingertip

Visual Panel (Cont'd.)

Finger Painting

Visual Panel (Cont’d.)

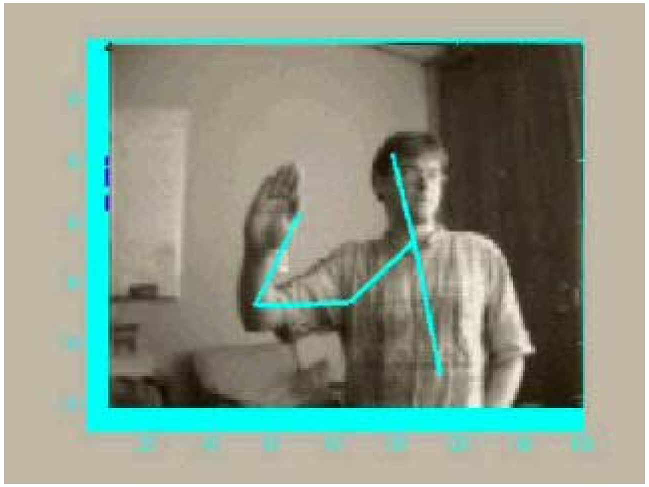
Virtual Keyboard

Visual Gesture Research

- The goal of visual gesture research (VGR) is to study the problems involved in implementing an immersive visual gesture interface in order to achieve a more natural human computer interaction. Several topics are explored including the study of hand modeling, various tracking algorithms, and analysis and synthesis of gestures.

Visual Gesture Research (Cont'd.)

Visual Gesture Research (Cont'd.)



A Prototype System for Computer Vision Based Human Computer Interaction

The Perceptual Browser

- The Perceptual Browser (PBrowser) is a prototype software demonstrating the introduction of Computer Vision in classical Graphical User Interfaces. Interacting with PBrowser, users control the scrolling in a window with head motion

FingerPaint

- FingerPaint is a demonstration system in which computer vision is used to allow people to interact with the system using several devices, such as a pen, an eraser, or even bare fingers

Projects

- **Vision-Based Single-Stroke Character Recognition for Wearable Computing** ,

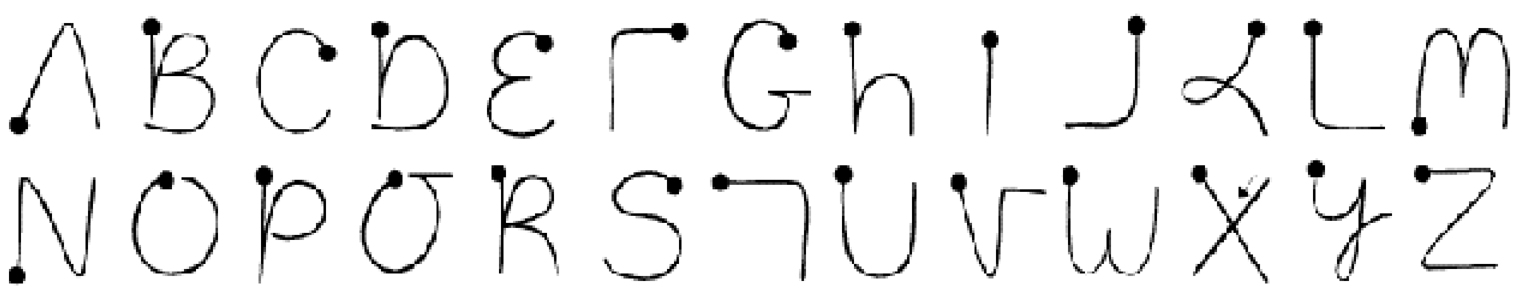
  An approach for data entry using a head-mounted digital camera to record characters drawn by hand gestures or by a pointer

- **Computer Vision Based Mouse** ,

  A computer vision based mouse which can control and command the cursor of a computer or a computerized system using a camera without a physical connection

Vision-Based Single-Stroke Character Recognition for Wearable Computing

- People want increasingly flexible and mobile ways to communicate using computers.
- A new approach for data entry using a digital camera to record characters drawn by hand gestures or by a pointer.
- Each character is drawn as a single, isolated stroke—the same technique used for characters in the Graffiti alphabet.

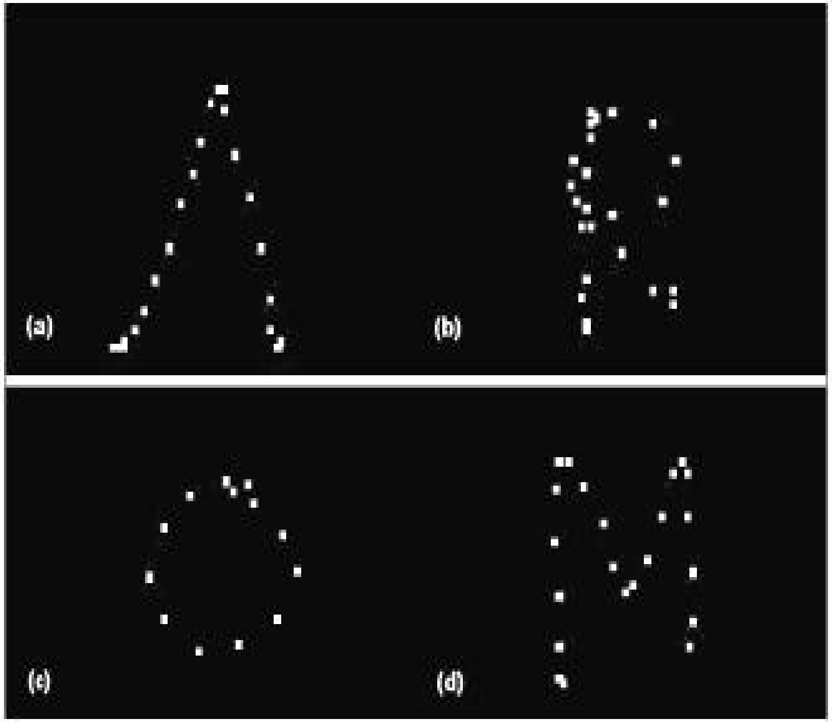


Vision-Based Single-Stroke Character Recognition for Wearable Computing (Cont'd.)

- Potential applications in mobile communication and computing devices such as
    - mobile phones,
    - laptop computers,
    - handheld computers, and
    - personal data assistants.

Vision-Based Single-Stroke Character Recognition for Wearable Computing (Cont'd.)

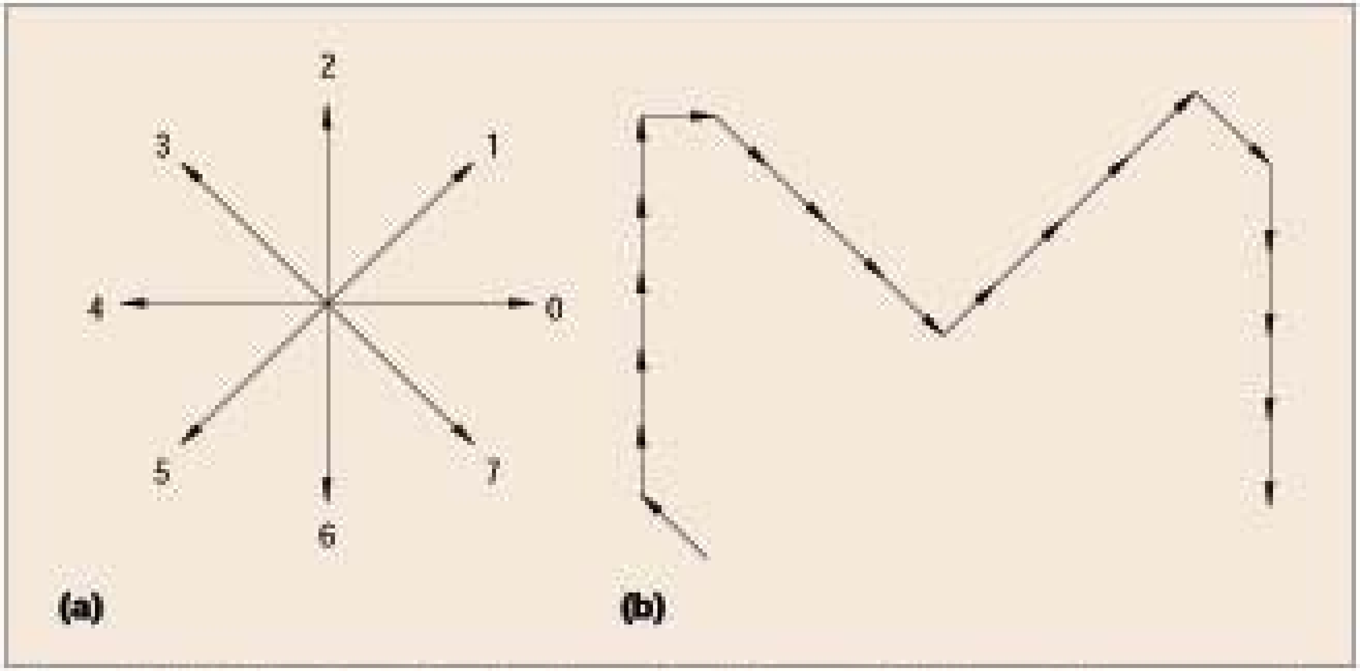
- A user draws unistroke, isolated characters with a laser pointer or a stylus on their forearm or a table.
- A camera on their forehead records the drawn characters and captures each character in sequence.
- The image sequence starts when the user turns the pointer on and ends when they turn it off.



Laser beam traces generated by image sequences

Vision-Based Single-Stroke Character Recognition for Wearable Computing (Cont'd.)

- A chain code describes the unistroke characters drawn.



a) Chain code values for the angles, b) a sample chain coded representation of char."M"=32222207777111176666

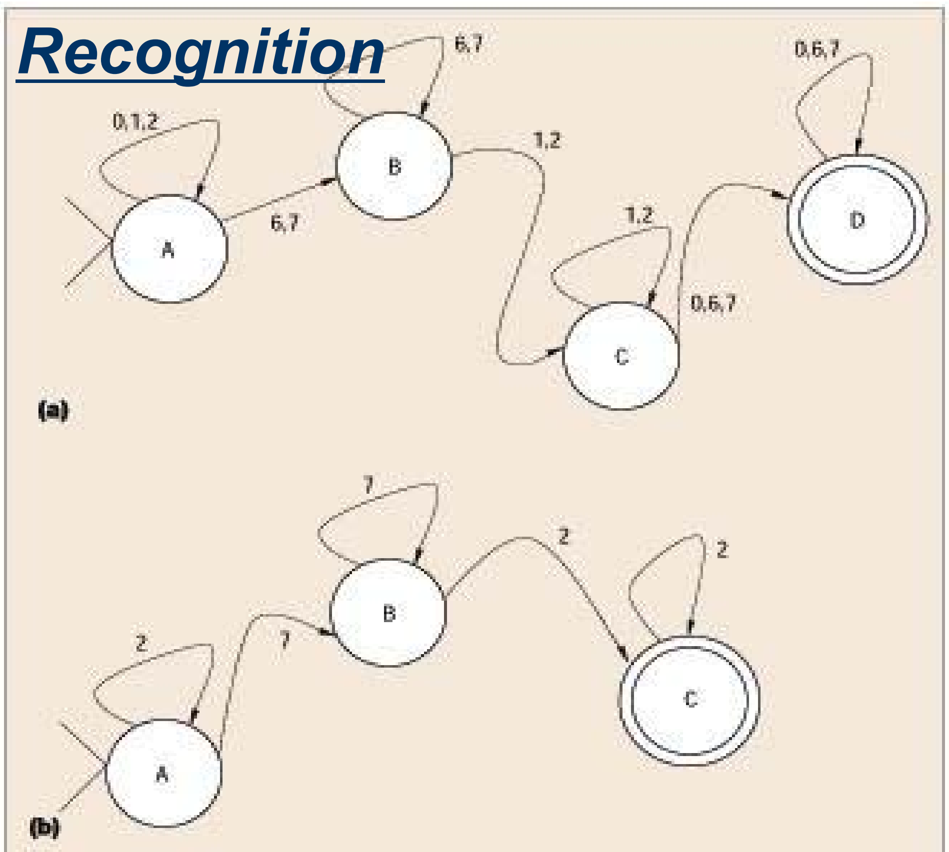
Vision-Based Single-Stroke Character Recognition for Wearable Computing (Cont’d.)

Recognition

- Chain code is the input for the recognition system.
- Recognition system consists of *finite state machines* corresponding to individual characters.
- FSMs generating the minimum error identify the recognized character.
- Certain characters such as *Q* and *G* might be confused: system also considers the beginning and end strokes.
- Weighted sum of the error from a finite state machine and the beginning and end point error.

Vision-Based Single-Stroke Character Recognition for Wearable Computing (Cont'd.)

Recognition



Finite state machines for the characters a) "M" and b) "N".

Vision-Based Single-Stroke Character Recognition for Wearable Computing (Cont'd.)

Video Processing

- To extract chain code from the video, marker positions corresponding to a character are processed.
- If the marker is in the initial frame, it is tracked in the consecutive images.
- A red laser pointer is used to write the characters.
- Images are decomposed into red, green, and blue components.
- Thresholding followed by a connected component analysis identifies the red mark.

Vision-Based Single-Stroke Character Recognition for Wearable Computing (Cont'd.)

Overall Algorithm

- Step 1: extraction of chain code
- Step 2: analysis using finite state machines
- Step 3: accounting for errors due to beginning and end points
- Step 4: determining characters

Both the time and space complexity of the recognition algorithm are $O(n)$, where n is the number of elements in the chain code.
To prevent noisy state changes, look-ahead tokens can act as a smoothing filter on the chain code.

Vision-Based Single-Stroke Character Recognition for Wearable Computing (Cont'd.)

Experimental Results

- A red laser pointer, black background fabric, and a Web cam
- Intel Celeron 600 processor with 64 Mbytes of memory
- 160 $\times$ 120 pixel color images at 13.3 frames per second
- Recognition rate of 97% at a writing speed of about 10 wpm
- Writer-independent and requires little training
- Perspective distortion up to about 45 degrees does not affect character recognition

Published in *IEEE Intelligent Systems and Applications*
May-June 2001

Computer Vision Based Mouse

- A ***computer vision based*** mouse can control and command the cursor of a computer or a computerized system using a camera.
- In order to move the cursor on the computer screen the user simply moves the mouse placed on a surface within the viewing area of the camera. The computer analyzes the video generated by the camera using computer vision techniques and moves the cursor according to mouse movements.
- In this system there is no need to have a cable connection between the computer and mouse nor a wireless transmitter-receiver pair as the mouse movements are transferred to the computer by the camera.

Computer Vision Based Mouse (Cont'd.)

- The concept of vision based finger mouse is first proposed by *Quek et al* in which the user controls the cursor by moving his fingers in the three-dimensional space.
- In our mouse system the approach is simpler than that in the sense that we place specific reference points on the mouse and the mouse moves on a two-dimensional surface. We track the reference points and the location of the cursor is updated accordingly.
- This approach is computationally more efficient than finding the finger tip in a cluttered background and tracking it. In addition to the above features our mouse has well defined regions corresponding to buttons to implement clicking. To click a button the user simply covers a button region for some time by his or her finger or a pointer.

Computer Vision Based Mouse (Cont'd.)

- The resulting image processing system can be also used in mobile communication and computing devices such as mobile phones, laptop computers, handheld computers, and PDAs. The advantages of our computer vision based mouse system compared to earlier systems are the following:
  - The background can be controlled by of the user. A mousepad which is clearly distinguishable from the background can be used. This simplifies the image analysis process for extracting the mouse boundaries or to track the reference point on the mouse.
  - Reference marks can be placed on the mouse. These marks are easier to find and track by the vision system of the computer.

Computer Vision Based Mouse (Cont'd.)

- Well defined regions corresponding to mouse buttons can be placed on the mouse. To click a button the user covers one of these regions momentarily by his or her finger or pointer.
- There is no need to place an electronic circuit inside the mouse which is essentially a passive device. The mouse can be made of any hard material.

Computer Vision Based Mouse (Cont'd.)

Video Analysis and Recognition System

- The video transmitted by the camera is analyzed image by image in real time. Let It be the image at time instant t extracted from the video generated by the camera. The main image processing problem that we encouter is to find the mouse in the image It. This can be done in varios ways. We are going to describe three different methods:
  - By edge detection
  - By color analysis *(We used this method)*
  - By motion analysis

Computer Vision Based Mouse (Cont'd.)

By Edge Detection

- Let us assume that It(nt, mt) be the pixel corresponding to one of the corners of the mouse or the center of an edge of the mouse or the location of a specific reference mark on the mouse etc. The pixel It(nt, mt) is obtained by performing edge detection and after some simple image processing operations such as thresholding etc.

- Whenever the next image It+1 is available the same edge detection operation is repeated over the new image It+1 and the reference mark is extracted.

- Let It+1(nt+1,mt+1) be the pixel corresponding to the reference mark on the mouse. If (nt,mt) and (nt+1,mt+1) are the same and the pixel values are close to each other then it is assumed that the mouse has not moved and the cursor remains wherever it is on the screen.

Computer Vision Based Mouse (Cont'd.)

By Edge Detection

- If (nt,mt) and (nt+1,mt+1) are different from each other and the pixel values are close to each other then the cursor is moved in the direction of the vector (nt,mt) - (nt+1,mt+1).

- The length of the cursor movement is also proportional to the difference vector. This can be adjusted by the user and according to the viewing area of the camera.

- In the computer vision based mouse there are specific regions corresponding to mouse buttons. To click a button the user covers a region by his or her finger or a pointer for some time corresponding to a button. The edges of these button regions are also detected during image analysis. If a change is detected inside one of these regions it is assumed that it is assumed that it is pressed.

Computer Vision Based Mouse (Cont'd.)

By Color Analysis

- In this approach a small reference mark which has a different color from the mouse, the mousepad and the background is placed on the mouse. This mark is known by the computer a priori. This mark is used as the reference point of the mouse.

- Whenever the mouse is moved by the user the location of the reference point changes. The detection of the reference point can be carried out by adaptive thresholding. By tracking the reference point the cursor is moved by the computer as described above.

- This color coding approach is easier and more robust then the edge detection method as the viewing area of the camera is more or less known by the image analysis system.

Computer Vision Based Mouse (Cont'd.)

By Color Analysis

- Similarly the mouse buttons can be color coded, too. For example the left mouse button may be a green region and the right mouse button may be a blue region etc.
- Whenever the green regions is covered by a finger or a pointer momentarily it is assumed that the left mouse button is clicked.
- The user should not cover the reference point and the regions corresponding to mouse buttons during the normal mode of operation.

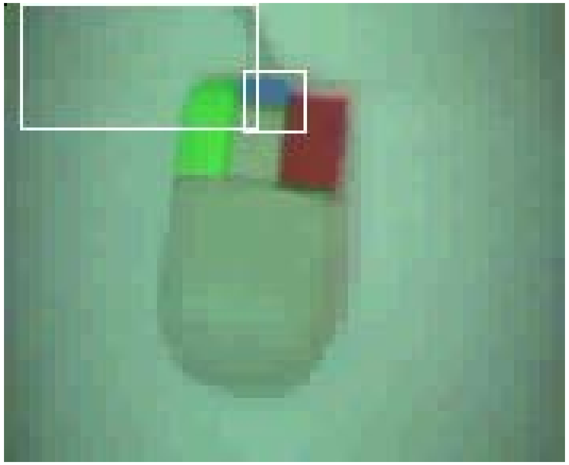
Computer Vision Based Mouse (Cont'd.)

By Motion Analysis

- In this case two consequitive images It and It+1 are processed together.
- The difference image It - It+1 contains only the moving regions in the viewing area of the camera (assuming that the camera is fixed during time segment t+1- t). If the reference point is fixed it does not appear on the difference image. If it moves it appears on the difference image and the cursor is updated accordingly.

Computer Vision Based Mouse (Cont'd.)

Current Implementation



Future Projects

- ***Vision Based Keyboard Systems***
  Computer Vision Based Weightless and Wearable Keyboards for Mobile Computing Devices and Future Cellular Phones.

- ***Vision Based Recognition System for Continuous Writing Based on Graffiti Alphabet***
  This project is a generalized version of previous character recognition project. This time we plan to recognize continuous writing.

References

- *Vision for man machine interaction,* J.L. Crowley, J.Coutaz
- *VISUAL PANEL: Toward A Vision-Based Mobile Input Interface For Anywhere,* Ying Wu, Ying Shan, Zhengyou Zhang, Steven Shafer
- *Visual Input for Pen-Based Computers,* Mario E. Munich, Pietro Perona