Computer Vision for Wearable Visual Interface

Walterio Mayol
Department of Computer Science
Faculty of Engineering
University of Bristol

Takeshi Kurata
National Institute of Advanced Industrial Science and Technology (AIST)
Table of Contents

- Introduction: Why visual sensors?
- Fundamentals of camera models
- Where to wear an extra eye?
- Kinds of cameras
- Devising real-time wearable visual systems
- Lunch
- Relevant computer vision techniques
- Applications
- Discussion of social issues with wearable cameras and conclusions
Introduction: Why visual sensors?

“These days, cameras are cheap, small and ubiquitous, and mobile computing power is improving constantly”
Why visual sensors?

Advantage of visual sensors:

They can recover several properties of the world:

- Identity of objects.
- Ambient properties such as colours, motion, etc.
- 3D structure (multiple cameras or when in motion).
- Provide a signal compatible with the pinnacle of human senses.
- Detect people and their activities.
- Among others…
Cameras help wearables

Some key applications of wearable computers:
- Memory enhancement.
- Ability enhancement.
- Data collection.
E.g. Casual Capture of Events

Images, panoramas and videos automatically detected offline from camera motion.
Summary of Phil’s family holidays in the Dolomites:

Phil Cheatle. Media Content and Type Selection from Always-on Wearable Video. pp 979-982. ICPR 2004.
E.g. Game assistant

Fundamentals of camera models

- Camera Model
  - Perspective (Projective) or Perspective approximations (Affine)?
  - You need Euclidean shape/motion?

- 2-D transform (Affine, Homography, etc) may be enough.
  - Object: Planer or not?
  - Motion: Just rotation or involving translation?
Perspective Camera

- So-called a Pinhole camera
- Non-linear

\[
\begin{pmatrix}
X \\
Y \\
Z \\
1
\end{pmatrix}
= \begin{pmatrix}
f & 0 & 0 & 0 \\
0 & f & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
x \\
y \\
1
\end{pmatrix}
\]
Affine Camera: Orthogonal

- The simplest affine approximation
- Ignore the distance and position effects

\[ \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = f \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \]
Affine Camera: Weak Perspective

- Or scaled orthogonal
- Zero-order approximation
- Distance effect

\[
\begin{bmatrix}
X \\
Y \\
Z_c
\end{bmatrix} = 
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix} \begin{bmatrix}
x \\
y \\
1
\end{bmatrix}
\]
Affine Camera: Paraperspective

- First-order approximation
- Distance and position effect
- Still linear
- Need to know the focal length

\[
\begin{pmatrix}
X \\
Y \\
Z
\end{pmatrix} = \begin{pmatrix}
1 & 0 & -\frac{X_c}{Z_c} & X_c \\
0 & 1 & -\frac{Y_c}{Z_c} & Y_c \\
0 & 0 & 1 & Z_c
\end{pmatrix} \begin{pmatrix}
X \\
Y \\
Z \end{pmatrix}
\]

Diagram showing the virtual image plane, focal point, average depth plane, centroid of the object, and world origin.
Shape-from-motion & Wearable

Affine factorization method

- Robust estimation (M-estimator) to obtain dominant 2D motion
- SVD to decompose affine shape and motion with the LMedS (Least Median of Squares) method
- Metric constraints for each specific camera model
2-D Transform & Wearable: Panorama-based Annotation

- Pre-registered panoramic images are used to estimate the position and direction of users with 2-D image transform.
Image mosaicing with wearables

- 2-D Affine or Homography
- Real World OCR (Optical Character Reader)
  - Make a panoramic image by stitching (mosaicing)
  - Texture analysis to detect text regions

Where to wear an extra eye?

Where to wear an extra eye?

- Head
  - Mann

- Chest
  - Starner et al.
  - Vardy et al.

- Wrist
  - Hujimoto et al.

- Shoulder
  - Kurata
  - Mayol

- Foot
  - Kemp et al.
Simulating camera placement

Articulated Humanoid Model with 1800 polygons. Virtual Cameras Around Shape to Evaluate Sensor Performance.

Matlab model at: http://www.robots.ox.ac.uk/~wmayol/3D/nancy_matlab.html

Mayol et al. 2001
Simulating camera placement
Simulating camera placement

Field of View (FOV) for camera ~4cm from body

Mayol, Tordoff & Murray 2001
How far away?

![Graph showing occlusion (%) vs. distance from surface (cm)]

- Shoulder
- Chest
- Sternum
- Temple
- Forehead
Purposive placement
Example of criteria fusion

FOV + Motion + Handling view
## Sensor placement

<table>
<thead>
<tr>
<th>Zone</th>
<th>FOV</th>
<th>Best suited to task</th>
<th>Body motion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head</td>
<td>Wide</td>
<td>Linked to user’s attention</td>
<td>Large</td>
</tr>
<tr>
<td>Shoulder</td>
<td>Wide</td>
<td>Most tasks</td>
<td>Small</td>
</tr>
<tr>
<td>Chest</td>
<td>Reduced</td>
<td>Linked to user’s workspace</td>
<td>Small</td>
</tr>
</tbody>
</table>
Kinds of Camera

- Narrow FOV
- Rotating Narrow FOV
- Fish-eye
- Catadioptric

Images from Nayar 1997
Some notes on hardware for prototypes

Current common interfaces:

- USB (currently cheap with interlaced sensors)
- Firewire (currently more expensive, better sensor & control)
- Analog (small units, needs capture card)
- Direct digital (directly interface the chip)

Firewire cameras w/progressive scan, full control of shutter speed, WB, frame rate, etc

Interlacing scanning

Progressive scanning
Single-lens Camera

- Most studied case.
- More suitable for high resolution of a particular region.
- Small volume.
- Wearable positions: head looking ahead, head looking to hands, chest, foot, etc.

Clarkson & Pentland MIT 2001
Catadioptric: mirror-lens

- Easy to obtain views round the user.
- Usually low resolution and large volume.
- Hard to decide where to wear it.

Trinocular and more

- 3D Visual Information

- Two way stereo (near/distant)
  - miniBEE by ViewPLUS
  - 150g, 2.0W, USB2.0
  - www.viewplus.co.jp
What kind of camera?

Compound eye with human-like resolution
A lesson from The Animal Kingdom?

**Salticidae** (Jumping) spiders.


- Wider FOV low-acuity eyes
- “Mobile” narrow FOV high-acuity eyes

8 eyes in total
Devising real-time wearable visual systems

- Computational/Software architectures
  - Device
  - Software Environments

- Active vision approach
  - Light control (IR Vision)
  - Camera head (pan/tilt/roll/zoom) control
Computational/Software architectures

- Mobile Device (Smartphone, PDA)
  - J2ME
  - QUALCOMM BREW & C++
  - Windows Mobile (for Smartphone)
  - Symbian OS & C++
  - Embedded Linux, ITRON, etc

- WearComp
  - Embedded Comp, Laptop (subnotebook) PC
  - Library, SDK, Toolkit…
  - IPP, OpenCV, DirectShow, ARToolkit, Georgia Tech Gesture Toolkit
Cell Phone as a platform for CV

Thomas Riisgaard Hansen
- http://www.daimi.au.dk/~thomasr/
- http://www.pervasive-interaction.org/Mixis/

- **J2ME**
  - **Pro:** Portability, easy to get started
  - **Con:** Limited functionality, slow

- **Symbian OS & C++**
  - **Pro:** Runs on a lot of phones, fast, programmers can control a lot of functions
  - **Con:** Limit compatibility, hard to program

- **Windows Mobile for SmartPhone**
  - **Pro:** It is Microsoft, resembles the programming you do on Windows PCs.
  - **Con:** Runs only on a limited number of devices
Running on a cell phone: MIXIS

- Tracking a fixed-point with a cell phone
  - Randomized Hough Circle Detection
    - Much faster than Non-Randomized Hough Algorithms
    - Hand-drawn black circle is tracked
  - Optimized version of the CamShift (Continuously Adaptive Mean Shift)
    - Face and colour tracking
    - The fixed-point is always with the user
    - C++ for Symbian OS 7.0s on a Nokia 7610
    - Image resolution: 160x120 or 320x240

http://www.pervasive-interaction.org/videos/Mixis.mov
Running on a cell phone: QR (Quick Response) Code

- Open Code (can be used for phishing)
  - QR Code (R) is registered trademarks of DENSO WAVE INCORPORATED in JAPAN and other countries.
  - http://www.qrcode.com/
- Standardization
  - AIM (Automatic Identification Manufacturers ) International, JIS, ISO
- Readable from any direction in 360°
- Dirt and Damage Resistant
  - QR Code has error correction capability. Data can be restored even if the symbol is partially dirty or damaged
- Getting Popular (Especially in Japan)
  - Newspaper, Magazine, Poster, and Business card
  - NTT DOCOMO, au by KDDI, Vodafone
Running on a cell phone: SyncR on BREW

- **SynchroReality by OLYMPUS**
  - Natural picture/icon as a marker

1. Download a set of templates
2. See some picture through the camera
3. Display the relevant info
4. Access the detailed info
Running on a cell phone:
OKAO Vision (OKAO = face in Japanese)

- Face Recognition Product by OMRON
  - http://www.omron.com/r_d/vision/01.html
- Symbian OS, BREW, embedded Linux, ITRON
- 0.2% FRR (False Reject Rate) with 1.0% FAR (False Acceptance Rate)
- Fast (1.0 sec. with ARM9core 200MHz)
- Previously employed Gabor-wavelet transform and graph matching, but the current method is customized for low computational power.

Low resolution images (such as in-camera images with 0.1Mpixel resolution) can be recognized.

Face measurably bigger than the screen
Small face
Rotated face
Face out of screen centre
28pixel
Running on a cell phone: Text Region extraction & OCR

- Work by Keiichi TAMAI et al., ORMON SOFTWARE
- Compensation for defocus, motion blur, and (lens) shading: LOG (Laplacian of Gaussian) filter, partial equalization
- Compensation for roll: Hough transform
- Performance: 1.48sec (ARM9core, 196MHz), 89.4% recall, about 50% precision
Other Related Works with a smartphone and PDA

- **PDA & ARToolkit**
  - Handheld AR Project (ARToolkit Plus)
    - [http://studierstube.org/handheld_ar/](http://studierstube.org/handheld_ar/)
    - [http://graphics.tudelft.nl/~wouter/](http://graphics.tudelft.nl/~wouter/)

- **SmartPhone & {Marker-based AR, Object Recognition}**
  - A. Henrysson, M. Billinghurst, M. Ollila, “Face to Face Collaborative AR on Mobile Phones”, ISMAR 2005 (7fps on Nokia6630, ARM9Core 210MHz, Symbian OS).
    - [http://www.itn.liu.se/~andhe/UMAR/](http://www.itn.liu.se/~andhe/UMAR/)
    - [http://www.uni-weimar.de/~bimber/](http://www.uni-weimar.de/~bimber/)
Marker-based AR on a mobile device

- See-through Information Display
  - PDA + HMD
  - SynchroReality by OLYMPUS
Embedded Comp: ViewRanger by SGI Japan

- [ ] http://www.sgi.co.jp/solutions/bbu/viewranger/
- [ ] SH-4 240MHz, Net BSD, 130g, 2.5W
- [ ] In: NTSC, AUDIO, RS232Cx2, Out: AUDIO
- [ ] CF (WiFi, PHS, FOMA), 100Base-TX
- [ ] Live video streaming (MPEG) and Live audio communication via Apache
- [ ] Local logging of video and audio (MPEG4)
- [ ] Motion Tracking, Edge Detection, etc
- [ ] Programmable (g++)
Toolkit: IPP, Open CV

- IPP (Intel Performance Primitives)
  - Signal processing (1-D), Image processing (2-D), Matrix computing

- OpenCV
  - [http://sourceforge.net/projects/opencvlibrary/](http://sourceforge.net/projects/opencvlibrary/)
  - Basic operations
  - Statistics and moments computing
  - Image processing and analysis
  - Structural analysis
  - Motion analysis and object tracking
  - Pattern recognition
  - 3-D reconstruction and camera calibration
  - Graphic interface and acquisition
Toolkit: ARToolkit

Popular toolkit for Marker-based tracking

- Platforms: from SGI, Linux, Windows to PDA, Smartphone

Overview

Input Image $\Rightarrow$ Thresholding Image $\Rightarrow$ Marker Detection $\Rightarrow$ Virtual Image Overlay $\Rightarrow$ Pose and Position Estimation
Coordinate Systems

Camera Coordinates

Observed Screen Coordinates

Image Distortion Function

Ideal Screen Coordinates

(xc, yc)

(xd, yd)

Marker Coordinates

Marker

Zm is directed to upside
Toolkit:

Georgia Tech Gesture Toolkit (GT²K)

- Support Experiments in HMM-based Gesture Recognition
  - http://www.gt2k.cc.gatech.edu
  - Platforms: Linux, Unix, Windows via CygWin (no promises)

- GT²K Provides suite of tools for gesture recognition
  - data preparation, training, validation, recognition

- User provides
  - data, feature vector, model topology and grammar, results interpreter

- When to Use GT²K
  - Applications using symbolic/iconic gestures
    - sign language, handwriting
  - Not good for tracking applications (e.g. finger mouse)
Application:
American Sign Language translator

- one-way, mobile translation, recognizes basic vocabulary, grammar


- Hat-mounted camera, wrist-mounted accelerometers
GT²K Background

- Cambridge University's HMM (Hidden Markov Model) Toolkit (HTK)
  - [htk.eng.cam.ac.uk/](http://htk.eng.cam.ac.uk/)
  - can be difficult for non-expert users
  - intended for speech recognition
  - users must understand how to relate speech to gesture

- GT²K uses HTK for modeling and recognition
  - abstracts lower level details of HMMs
  - abstracts speech-specific details of HTK
  - allows quick prototyping of gestural interfaces
GT²K Functionality Overview

- **Data Preparation**
  - design gesture models
  - specify gestures
  - GT²K generates gesture grammar
  - label data for training
    - one gesture per file (automatic labeling)
    - multiple gestures per file (assisted labeling)

- **Training**
  - determines which examples to use for training
    - Cross-validation (2/3 training, 1/3 validation)
    - Leave-one-out (one validation, rest training)
    - User-defined
GT²K Functionality Overview

- **Validation**
  - Automatically validates model based on chosen method (e.g., cross-validation)
  - Reports accuracy
  - Produces confusion matrix

- **Recognition**
  - Returns name of most likely gesture

- **Uses classified gesture to issue command to devices**
  - Hand moving up ➤ volume up
  - Hand moving down ➤ volume down
The Active Vision Approach

Active Vision

“Efficient use of available optical and computational resources”

Examples:
1) A face tracking camera.
2) A camera that emits IR illumination only when a new observation is needed.
Active vision approach: IR (Infrared) Vision

- The most successful vision system
  - measures only what you want to measure
- IR Vision system: Simplify image processing
  - Segmentation, Detection, Tracking
  - Can use Invisible targets

- Drawback
  - Sunlight, heat source
  - Additional camera to capture normal images
    - could be solved by synchronizing the light flashing and shutter speed
IR Vision:
Invisible Visual Marker for AR

Y. NAKAZATO, M. KANBARA, N. YOKOYA, “Discreet markers for user localization”, ISWC2004

- Easy infrastructures
- No impairing the scenery

×Visual Marker

✓Invisible Visual Marker
System Overview

Translucent Retro-reflective Marker

Reflection Infrared Light

Wearable Computer (MP-XP7310 [JVC])

Video Capture Unit (NV-UT200 [NOVAC])

IR Camera and IR-LEDs

RS-232C

USB
IR Vision: Gesture Pendant

- Device for control of home automation systems via hand gestures (HMM-based)
- The Gesture Pendant can also analyze their movements for pathological tremors.
IR Vision: Low Vision Aid

  - [http://www.hitl.washington.edu/research/wlva/](http://www.hitl.washington.edu/research/wlva/)
- Help visually impaired people avoid obstacles in their path
- The closest objects appear bright
- Warning icon projected into the eye with a laser light
- Laser light through a vibrating optical fiber directly into the user’s retina

Source: Seattle Post-Intelligencer, 2004
Active vision approach: Camera motion control

- **Wearable Visual Robot**
  
  Mayol, Tordoff, Murray. ISWC 2000

- **Wearable Active Camera/Laser (WACL)**
  
  Sakata, Kurata, Kato, Kourogi, Kuzuoka. ISWC 2003
Max velocity with highest detail

\[ V_s = \frac{\Delta \phi}{\Delta t} \]

(Angle per pixel)

(Shutter speed)

E.g. Webcam 640x480 @ 30Hz & FOV of 42 degrees horizontal:
- Max time to collect light is \(~30\text{ms}\)
- Detail starts to blur at just 2 deg/s

Person walking 2m away moves at \(~50\text{ deg/s}\) on the image.
Why an Active Solution?

Passive Camera
Slow shutter       Fast shutter

Active Camera
Slow shutter       Fast shutter
WACL: Wearable Active Camera/Laser

- Camera/Laser head on a small pan/tilt platform
  - The laser pointer is almost parallel to the camera
  - The laser spot is close to the center of the video images
  - Easy to miniaturize and reduce weight because of the simple mechanism
- Stabilization capability
  - Based on image registration and inertial sensors
- The WACL interface allows the remote collaborators
  - not only to independently set their viewpoint into the wearer’s task space
  - but also to point to real objects in the task space with the laser spot.

[Image: Expert’s view and Worker’s view]

[Link: http://www.is.aist.go.jp/weavy/movie/2004/WACLHMC-secC-VS.mpg]
Experiences of camera wearability

Study of Collar Structures
A tip for building robust prototypes

“Friendly Plastic” by www.amaco.com re-usable polyethylene thermoplastic that softens in warm water and becomes hard in cold water.
Relevant computer vision techniques

- Integration with inertial sensors
- Gesture identification
- SLAM: Simultaneous Localization and Mapping
Integration with inertial sensors: Personal Positioning (PP)

- M. Kourogi, T. Kurata, “Personal Positioning based on Walking Locomotion Analysis with Self-Contained Sensors and a Wearable Camera”, ISMAR03
- Detect a cycle of walking locomotion and walking direction by analyzing the acceleration/angular velocity/magnetic vectors
- Dip angle: uniquely determined by the latitude and longitude of the location to handle disturbances of the magnetic field
Personal Positioning (PP): Recognize human locomotion behavior

Walking on a flat floor

The forward direction obtained by PCA

Walking while in walking

speed estimation

Going up stairs

Going down stairs

Taking an elevator
Fast Image Registration: Pseudo-motion vector

- Approximation of the actual 2-D motion vector

\[
\begin{align*}
\frac{\partial I}{\partial x} u_p + \frac{\partial I}{\partial y} v_p + \frac{\partial I}{\partial t} &= 0 \\
\frac{\partial I}{\partial x} u_p + \frac{\partial I}{\partial t} &= 0 \\
\frac{\partial I}{\partial y} v_p + \frac{\partial I}{\partial t} &= 0
\end{align*}
\]

- Block-wise matching
- Pixel-wise matching

Image brightness vs time

- 1-D image brightness
- temporal gradient
- spatial gradient

Pixel-wise matching

The position of the pixel
PP Demo Video (Dead Reckoning only)

In 2003 International Robot Exhibition at Tokyo Big sight

Demo Video (3x speed play)

http://www.is.aist.go.jp/weavy/movie/2003/robot_exhibition-1VI.MPG
http://www.is.aist.go.jp/weavy/movie/2003/robot_exhibition-2VI.MPG
Indoor Navigation

- Dead-Reckoning and Image Registration
- Position adjusted by image registration
- Position adjusted by map-matching with locomotion behavior
Outdoor Navigation

- Dead-Reckoning and GPS
Integration with inertial sensors: AirGrabber

- [http://chihara.naist.jp/research/research_UC.html](http://chihara.naist.jp/research/research_UC.html)
Gesture identification

- Gestures can be explicit or not.
- Most work for wearables has perhaps not surprisingly concentrated on hands.
Skin detection

\[ p(\text{class}_i \mid \text{color}) = \frac{p(\text{color} \mid \text{class}_i) p(\text{class}_i)}{\sum_{j=1}^{N} p(\text{color} \mid \text{class}_j) p(\text{class}_j)} \]

Probabilistic Method

Bounding Box Method

Note: Other colour spaces such as RGB, HSV etc are also in use.
Skin detection

Colour Image  Skin segmentation  Image filtered with 5x5 “averaging mask”
ASL Recognition

Skin images converted to vector of characteristics:
(e.g. hand position in the image, dimension and orientation of bounding ellipse)

~97% accuracy over a 40 word vocabulary using grammar as constraint.
**ASL Recognition**

<table>
<thead>
<tr>
<th>part of speech</th>
<th>vocabulary</th>
</tr>
</thead>
<tbody>
<tr>
<td>pronoun</td>
<td>I, you, he, we, you(pl), they</td>
</tr>
<tr>
<td>verb</td>
<td>want, like, lose, dontwant, dontlike, love, pack, hit, loan</td>
</tr>
<tr>
<td>noun</td>
<td>box, car, book, table, paper, pants, bicycle, bottle, can, wristwatch, umbrella, coat, pencil, shoes, food, magazine, fish, mouse, pill, bowl</td>
</tr>
<tr>
<td>adjective</td>
<td>red, brown, black, gray, yellow</td>
</tr>
</tbody>
</table>

**ASL Test Lexicon**

The four state HMM used for recognition.

Starner & Pentland ISCV95
SLAM: Simultaneous Localization and Mapping

“How to know in Real-Time where sensor AND objects are without any positioning prior”

• Key problem in robots and wearables.
• Refinement of Structure From Motion:
  • Demands causality (no bundle adjustment).
SLAM: *Simultaneous* Localization and Mapping

Key aim: not just *account* for uncertainty but *reduce* it

See Work by Andrew Davison
See: http://www.doc.ic.ac.uk/~ajd
Davison’s Algorithm for monocular SLAM

Sensor’s state: 3D position, 3D Orientation (quaternion), 3D velocity, 3D angular velocity.

\[ \mathbf{x}_v = \begin{pmatrix} \mathbf{r} \\ \mathbf{q} \\ \mathbf{v} \\ \mathbf{\omega} \end{pmatrix} \]

Feature’s state: 3D position.

\[ \mathbf{y}_i = \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} \]

“Full covariance” EKF

\[ \mathbf{x} = \begin{pmatrix} x_v \\ y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \mathbf{P} = \begin{bmatrix} P_{xx} & P_{xy1} & P_{xy2} & \cdots & P_{xyn} \\ P_{yx1} & P_{yx1} & P_{yx2} & \cdots & P_{yxyn} \\ P_{y1x} & P_{y1y1} & P_{y1y2} & \cdots & P_{y1yn} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P_{ynx} & P_{yny1} & P_{yny2} & \cdots & P_{ynyn} \end{bmatrix} \]

Davison’s Algorithm for monocular SLAM

Visualisation of a constant velocity motion model.

Update Model

\[
\begin{pmatrix}
    r_{\text{new}} \\
    \dot{q}_{\text{new}} \\
    v_{\text{new}} \\
    \omega_{\text{new}}
\end{pmatrix}
=
\begin{pmatrix}
    r + (v + V)\Delta t \\
    \dot{q} \times \dot{p}((\omega + \Omega)\Delta t) \\
    v + V \\
    \omega + \Omega
\end{pmatrix}
\]

Davison’s Algorithm for monocular SLAM

Use a calibration object to get initial scale
Davison’s Algorithm for monocular SLAM

Particle filter to estimate depth

Shi-Tomasi “corner” operator to detect subsequent features

Images from Davison ICCV, 2003
SLAM applied to the Wearable Robot

Davison, Mayol, Murray, ISMAR 2003
Applications

- Hand gesture-based visual interface
- Wearable Augmented Reality
- Wearable Augmented Memory
- Remote collaboration with AR and 3D map building
- Real World Coupled Visual Interface Development Kit
Hand gesture-based interface: WVT: Wearable Virtual Tablet


- Basic Scheme: Fingertip drawing on a plane using an infrared camera

1. Input-plane extraction (Hough transform)

2. Fingertip detection (Arc [semi-circle] detection)

3. Discrimination between input/non-input (distance, shade of the fingertip)
Wearable Augmented Reality: Outdoor See-Through Vision

  - [http://www.image.esys.tsukuba.ac.jp/Eindex.html](http://www.image.esys.tsukuba.ac.jp/Eindex.html)
- Initial estimation employing a GPS and compass
- Tracking pose changes by inertia sensor
- Pose adjustment by referring landmarks (natural feature detection)
Wearable Augmented Reality: AR Manual Authoring

T. Okuma, T. Kurata, K. Sakaue, “Fiducial-less 3-D Object Tracking in AR Systems Based on the Integration of Top-down and Bottom-up Approaches and Automatic Database Addition”, ISMAR03

Wearable Augmented Memory: Lifelog

- “Lifelog” Project proposed by DARPA, 2003
- Lifelog research by Aizawa-lab, Univ of Tokyo
  - http://www.hal.t.u-tokyo.ac.jp/
  - 70 years recording will be available later on!?  
- 70 x 365 x 16 x 60 x 60 sec x 1Mbps = 184 TB

But, how to watch? retrieval, indexing, summarization
Remote collaboration with AR and 3D map building
Remote collaboration with AR and 3D map building

Davison, Mayol, Murray, ISMAR 2003
Demo Video
http://www.doc.ic.ac.uk/~ajd/Movies/wearableslam2.mpg
AIST has developed a commercial RWCVI-DK (Real World Coupled Visual Interface Development Kit) in cooperation with Media Drive Corporation that includes:

- Personal Positioning (PP) with self-contained sensors
- Hand-Gesture (HG) Interface
- Real-World OCR (RWOCR)

**Weavy: Wearable Visual Interface Project**
- [http://www.is.aist.go.jp/](http://www.is.aist.go.jp/)

**Media Drive**
Hand gesture-based interface on the Weavy


- The Distributed Monte Carlo (DMC) Tracking Method
  - An extension of the Condensation algorithm for heterogeneous distributed architecture
  - Different numbers of samples and different dynamical models on the wearable and infrastructure side respectively

- The Functionally-Distributed (FD) Hand Tracking Method
  - Combination of the DMC Tracking Method and hand color modeling with a GMM
Applications of Hand gesture-based interface

Secure PIN Input with HG
http://www.is.aist.go.jp/weavy/movie/2002/passwdV1.MPG

Real World OCR with HG
Social Issues with wearable cameras

- Cameras have been around for some time and acceptability is evolving.
- Different countries (and sometimes states) have different legislations, mainly related to pictures of people.
- The problem are not the cameras themselves but how you network them [Pentland 2000].
- Any comments on this?
Thank You, Gracias, Arigato!!

Contact Info

- Walterio Mayol
  - wmayol@cs.bris.ac.uk
  - http://www.cs.bris.ac.uk/~wmayol/

- Takeshi Kurata
  - kurata@ieee.org
  - http://www.is.aist.go.jp/kurata/
  - http://unit.aist.go.jp/itri/itri-rwig/