A Panorama-based Method of Personal Positioning and Orientation and Its Real-time Applications for Wearable Computers

National Institute of Advanced Industrial Science and Technology (AIST)
Masakatsu Kourogi, Takeshi Kurata, Katsuhiko Sakaue
Purposes

- Personal **positioning and orientation** of wearable users which can be used for:
  - Personal navigation
  - Annotation overlay, smart reminder

Actual output of our system:

- Head-worn display
- CCD camera
Previous works

- Artificial markers (fiducials)
  - Problems:
    - Difficulty to apply large-scale applications
    - Fiducials need to be captured closely

- Dedicated sensors such as inertial sensors, GPS, IR beacon, etc.
  - Problems:
    - Applicable environments restricted
    - Sensor precision limited
Our approach: panorama-based method

- A set of panoramic images are used as a database of real-world environment.

From the user’s camera

Input frames

- Located
- Orientated

Position at which a panorama is taken

Environmental map

Direction

Position

5 [deg]

235 [deg]
The problems

- We use *affine transform* that has difficulty to align input frames with panoramic images *if*...
  - User’s view direction is apart from the horizontal line.
- Change of *lighting condition* severely affects the performance of registration.
The problem

- Input frames cannot be aligned with a panorama that are mapped onto a cylindrical surface using **affine transform**.

An input frame

A panorama
The solution

- To re-project the input frames onto a cylindrical surface and align the projected frame with the panorama.
The solution

- **Thorough search**: initial estimates are given to cover the whole part of the panorama.

- **Partial search**: initial estimates are given around the previous result to save computational costs.
Omni-directional camera

- We use an *omni-directional* camera, HyperOmni Vision, to acquire panoramas.
  - Easy to generate
  - Geometrically correct
The problem

- Image registration is prone to change of lighting conditions.
  - A panoramic image was captured at different time and with a different camera system.

Different lighting condition gives totally different brightness.
The method: motion vector

- Pseudo motion vector is computed and selected or discarded at each pixel.

\[
\begin{align*}
\text{Pseudo motion vector} \quad & \quad u_p = -(1 - w) \frac{\partial I}{\partial t} + c_w \frac{\partial I'}{\partial t} \\
& \quad v_p = -(1 - w) \frac{\partial I}{\partial x} + c_w \frac{\partial I'}{\partial y}
\end{align*}
\]

\(c = 2.0, \quad w = 0.25\)
The method: confidence value

- Weighted sum of absolute difference (wSAD) of brightness and gradient is used as a evaluation measure to be minimized.

\[
wSAD = (1 - w) \sum_{x,y} | I_p(x+u, y+v) - I_f(x, y)| + cw \sum_{x,y} | I'_p(x+u, y+v) - I'_f(x, y)|
\]

We use \( c = 2.0, w = 0.25 \)
Experiments

- The proposed method can robustly handle most of difficult situations, empirically.

The correct parameters

New

Old

The correct parameters
False parameters

True match
False match

Confidence value

Width of panorama
Height
Each task is running in distributed and parallel manners.

Input/output image is compressed and sent via wireless network.

The VizWear system overview

IEEE 802.11b Wireless LAN (11Mbps)

Compressed image, data

Access point

Gigabit ethernet

PC cluster

PC: No.1

PC: No.2

PC: No.3

Reg

C/D

Reg

C/D

Reg

C/D

Image registration task (thread)

Compress/decompress (thread)

UI generation (thread)
The VizWear: outlook

- Head-set
  - Camera
  - Inertial sensors
- Wearable display
  - MicroOptical Clip-On display
- Wristwatch display
- Mobile PC (B5-size)
  - Video capture card
  - Wireless LAN card

1.53kg
The VizWear: hardware

- Computers and wired/wireless network
  - PC cluster consisting of 3 PCs (OS: Linux)
    - Dual Intel PentiumIII-1.2GHz, 850MHz, connected via Gigabit ethernet
    - 5 access points for roaming (Conformed to IEEE 802.11b, 11Mbps)
  - Wearable PC (OS: Windows Me): B5-size notebook PC
    - CPU: Intel Mobile PentiumIII-600MHz
    - Connected via Wireless LAN (Conformed to IEEE 802.11b, 11Mbps)

- Inertial sensor, CCD camera
  - InterSense, InterTrax2 (connected via USB)
  - Kanagawa Battery Co., Ltd, Very small Color CCD camera
The VizWear software

- The software is designed for parallel and distributed computing.
  - Each task is implemented as a thread.
    - Image registration task is further divided so that it can exploit data parallelism of the task.
  - Inter- and intra-PC communication is achieved by our open library. It internally uses:
    - PVM (Parallel Virtual Machine) library
    - POSIX or Win32 thread APIs
    - Berkley socket APIs
Experiments (off-line) : Registration

- Evaluate accuracy of registration between input frames and the panoramic image.

Manually select and find matching between input frames and the panorama.

100 frames are used in the experiments.

Compare accuracy of the previous and the improved method.
Results: accuracy of registration

- The proposed method improves accuracy of registration 3-10 pixeles
  - Imminent when inclination angle is apart from the horizontal line.
**Results: computational cost**

- The old method vs. the new method.

<table>
<thead>
<tr>
<th>The proposed method</th>
<th>50-100 msec/frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>Previous method (affine)</td>
<td>40-80 msec/frame</td>
</tr>
</tbody>
</table>

**Breakdown of cost**

<table>
<thead>
<tr>
<th>Task</th>
<th>Time (msec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Re-projection onto the surface</td>
<td>5-10</td>
</tr>
<tr>
<td>Image registration (affine)</td>
<td>40-80</td>
</tr>
<tr>
<td>Compute confidence value</td>
<td>5-10</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>50-100</strong></td>
</tr>
</tbody>
</table>

Computational cost increased by 10-20%. Real-time processing can still be achieved.
Experiments (on-line) : Positioning

- Panoramic images were captured at 45 points of the environment shown below:

Booth B

Booth A

Actual path of the user

Enterance

The user with VizWear is moving around the environment.

50cm intervals between the grids

Demo video #1

Panoramic images were captured at 45 points of the environment shown below:

The user with VizWear is moving around the environment.

50cm intervals between the grids

Demo video #1
Results: Positioning

- **100cm accuracy** of positioning achieved for a typical office environment.

![Graph showing measured and real positioning data]

- Annotation overlay is very stable.
- Positioning is accurate where motion parallax is large.
Demo video clip

Demo video #2
Exhibition held in Japan

Demo video #3
Wearer’s view
Conclusion

- We proposed an improved panorama-based method of personal positioning and orientation.
  - Input frames are re-projected onto panorama surface.
  - Gradient-based motion vector is used.
- Real-time applications are demonstrated.
  - The VizWear system is implemented.
  - Personal navigation and annotation overlay.