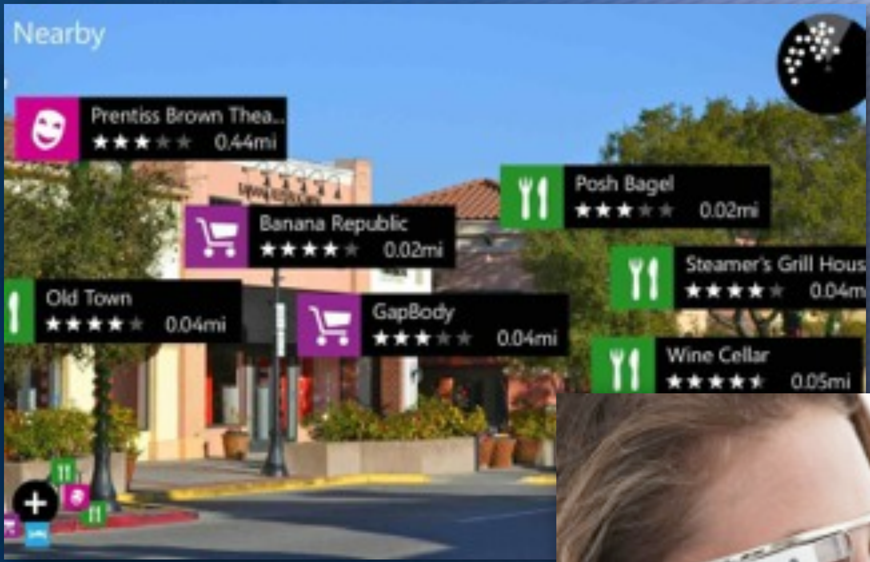
ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Distributed Systems Group | Institute for Pervasive Computing

# Handheld Augmented Reality

Reto Lindegger
lreto@ethz.ch

# AUGMENTED REALITY




# A Definition

Three important characteristics:

- Combines real and virtual environment
- Interactive in real-time
- Registered in 3D

Definition from:

A Survey of Augmented Reality

by Ronald T. Azuma

(1997)

# A Definition

What this excludes:

- 2D Overlay

- 3D Objects rendered in a movie

# A Definition

## What this includes:

Virtual objects in an rendered real-world reconstruction

Interaction between real-world objects and virtual objects

Information in a video feed

# Handheld Augmented Reality

- Devices which fit into a user‘s hand
- Portable and (ideally) not infrastructure dependent
- Smartphones as augmented reality displays
  - Commodity hardware
  - Widely used
- But:
  - Limited computation power
  - User has to hold device all the time

# 4 Problems and 4 Solutions

- Indoor Navigation
- One Handed Mobile Device Interaction
- Situated Visualization
- 3D Surface Reconstruction

# INDOOR NAVIGATION
## WITH ACTIVITY-BASED INSTRUCTIONS


# Problem Statement

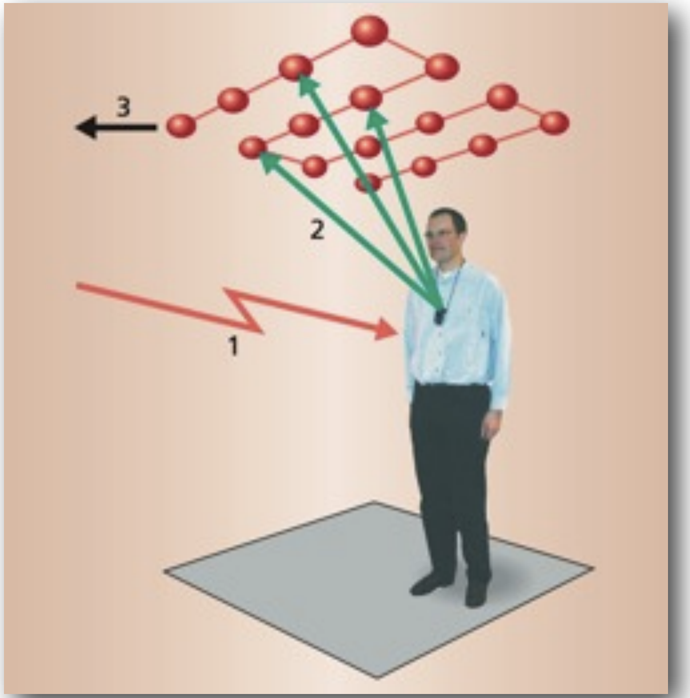
- Navigation in unknown, complex building
- No GPS
- Low effort for installation, low cost
- Sparse localization
- As accurate as possible

# Previous Work

Sensing Infrastructure

- Cyberguide Project, infrared
- BAT system, ultrasonic waves
- Chittaro and Nadalutti, RFID

➡ Require instrumentation of the environment



# Previous Work

Sparse infrastructure

- „You-are-here“ maps
- Information at checkpoints
- Way between checkpoints completely up to the user

➡ Finding next checkpoint not assisted

# Previous Work

Measuring movement

- Start point known
- Measure movement with sensors and camera
- Instruction based on movement and estimated location

➡ Inaccurate over time

➡ Phone movement indistinguishable from user movement

# Solution

Alessandro Mulloni et al. :

Handheld Augmented Reality Indoor Navigation with Activity-Based Instructions

- Combine egocentric and exocentric navigation
  - egocentric: turn-by-turn navigation
  - exocentric: „you-are-here“ maps
- Main requirements:
  - Robustness to user failures
  - Minimal instrumentation of the environment
  - Adaptiveness to localization accuracy
  - Interactive flow of activities with minimal user input

# Solution

Sparse localization: Info points

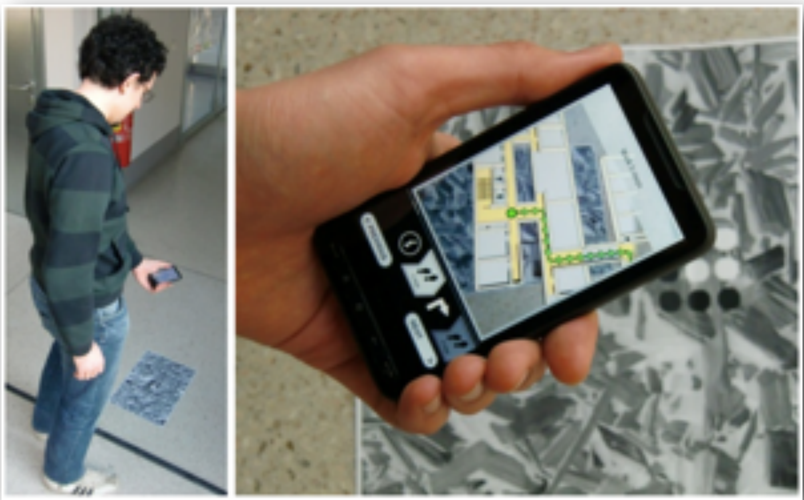
- floor-mounted posters
- Recognized by camera
- Act as checkpoint for user and software
- Change user interface: more information
- Recalculate path

# Solution

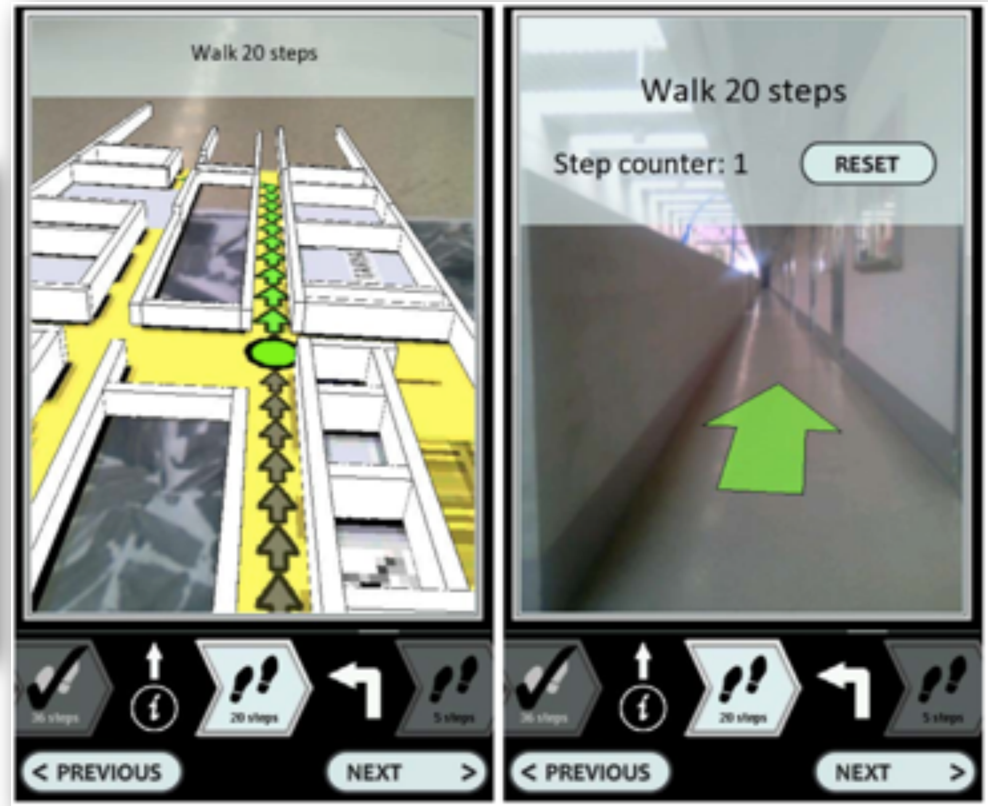
Between info points: Activities

- Navigation between info points activity-based
- Total route described in sequence of activities
  - „Go 5 steps“, „Turn left“, „Go 8 steps“
- Accelerometer used to count steps
- User can actively activate next activity
  - Scrolling through list and selecting any activity also possible

# Solution



Info point, recognized by camera




At Info points

Between Info points

# Evaluation

- User study: navigation in an unfamiliar building
- With and without information point
- Compare used time, navigation errors and user feedback

- Navigation works well, info points improve performance and usability
- Step counter just hint, but not used to switch to next activity

# Personal Opinion

- Application could be used in a museum or in an airport
  - Guide visitors through building
  - Find shortest way to an exhibit
- Useful for people with visual impairment
  - Extended with voice in- and output
  - Has to be quite accurate
- Not realistic for daily use as it is now

# LOOKING AT YOU

## FUSED GYRO AND FACE TRACKING FOR VIEWING LARGE IMAGERY ON MOBILE DEVICES

# Problem Statement

- Photos are getting bigger:
  - High resolution
  - Panoramic images
  - Multi-Perspective images
- Screens are getting smaller:
  - Smartphones
  - Tablets
- Many sensors on mobile devices available:
  - Accelerometer
  - Gyroscope
  - Compass
  - Camera

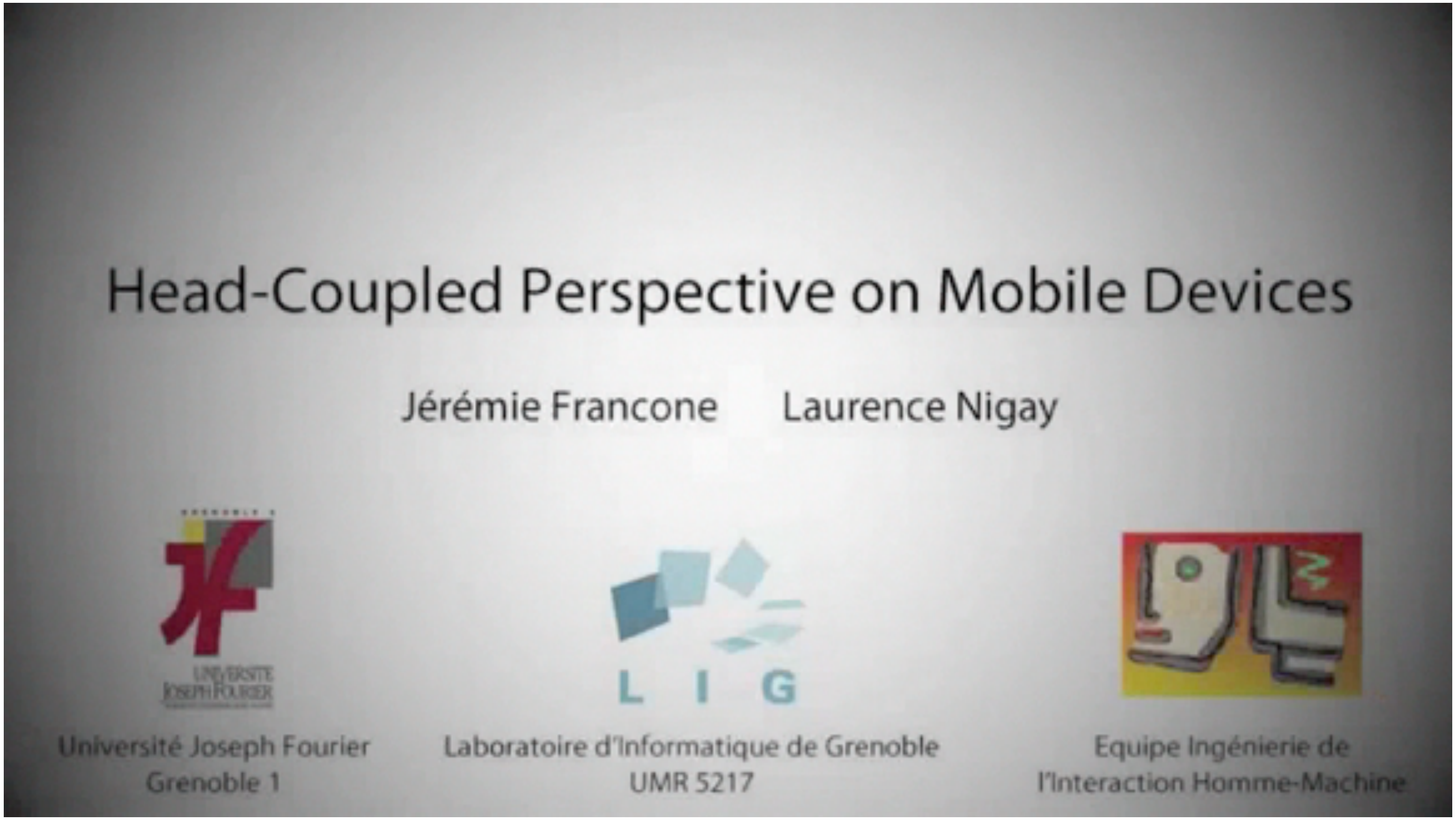
# Previous Work

- Scrolling speed coupled with zoom level
- Google Street View (compass and gyro)
- Tourwrist (360° panoramas)
- Gaze-enhanced scrolling techniques
- Glasses-free 3D displays (face tracking)

  http://iihm.imag.fr/en/demo/hcpmobile

# Previous Work

Glasses-free 3D displays (face tracking)



# Solution

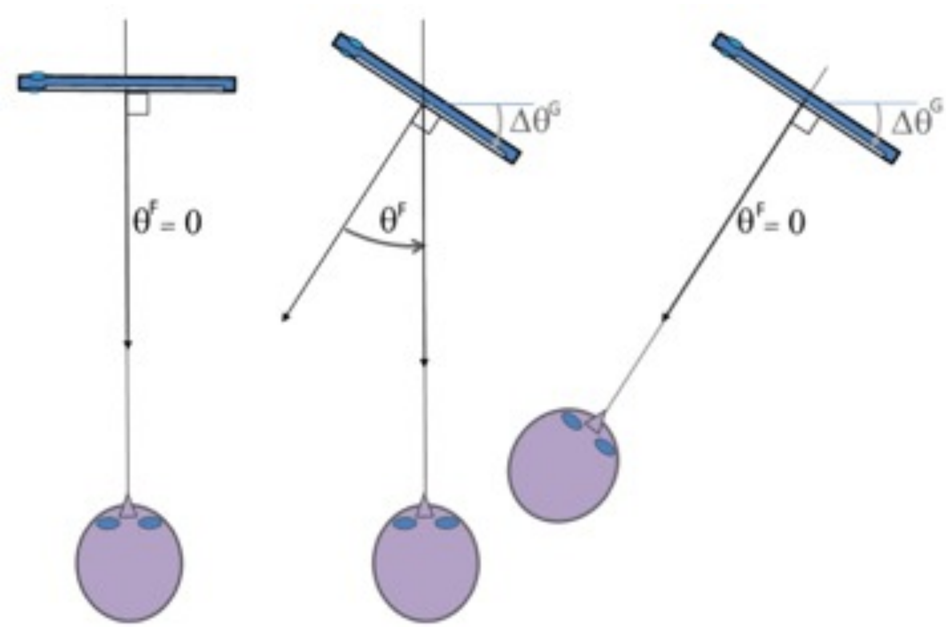
Neel Joshi et al. :

Fused Gyro and Face Tracking for Viewing Large Imagery on Mobie Devices

- Different applications
  - Large field of view
  - Wide multi-perspective panorama
  - Multi-views
  - Combinations
- Touch-based interaction has drawbacks
  - Hand obscures part of the picture
  - Difficult to distinguish between navigation and other interaction
  - Requires 2 hands

# Solution

Using only gyroscope data is not enough



# Solution

- In the real world we
  - move our gaze relatively to a scene
  - move an object relatively to our gaze
- In both cases, head moves relatively to the scene/object
- Exploiting the relative position of the head to the screen
- Face tracking can provide input in 3 dimensions
  - position, based on face location (x,y)
  - depth, based on face size (z)
  - but suffers from noise, high latency and limited field of view

# Solution

- Combine gyroscope data and face tracking!
- Features:
  - Navigation by moving head and/or device
  - Natural and smooth navigation
  - Don‘t have to spin 360° in place
  - Zooming by changing distance between face and device

# Solution

The application in action

# Evaluation

- User study: Find a mark placed somewhere on a large image
- Comparison of different input techniques
  - Finger based, combined gyro and face tracking
- 50% of the participants indicated that touch control is easier than gesture control
- Motion based methods not significantly faster than finger based

# Personal Opinion

- Very interesting technology for viewing multi-view images
  - 3D images look impressive
- Useful to view large panoramic images
  - Applications like StreetView
  - Get impression of the location where picture was taken
- Not useful for „flat“ images
  - Maps
  - Large high-resolution non-panoramic images

# SITE LENS

## SITUATED VISUALIZATION TECHNIQUES FOR URBAN SITE VISITS

# Problem Statement

- Urban planner, urban designer or architects visit a site before the design activity
- Interested in different information and facts about the location
  - Carbon monoxide measurement
  - demographics
  - traffic flows
  - congregation of people

➡ Much data from different sources, has to be processed and analyzed after the visit, with maps, videos and photos etc.

# Previous Work

- ArcGIS
- Vidente Project: visualization of subsurface features (pipelines, power cables) for utility field workers

# Solution

Sean White et al. :

SiteLens: Situated Visualization Techniques for Urban Site Visits

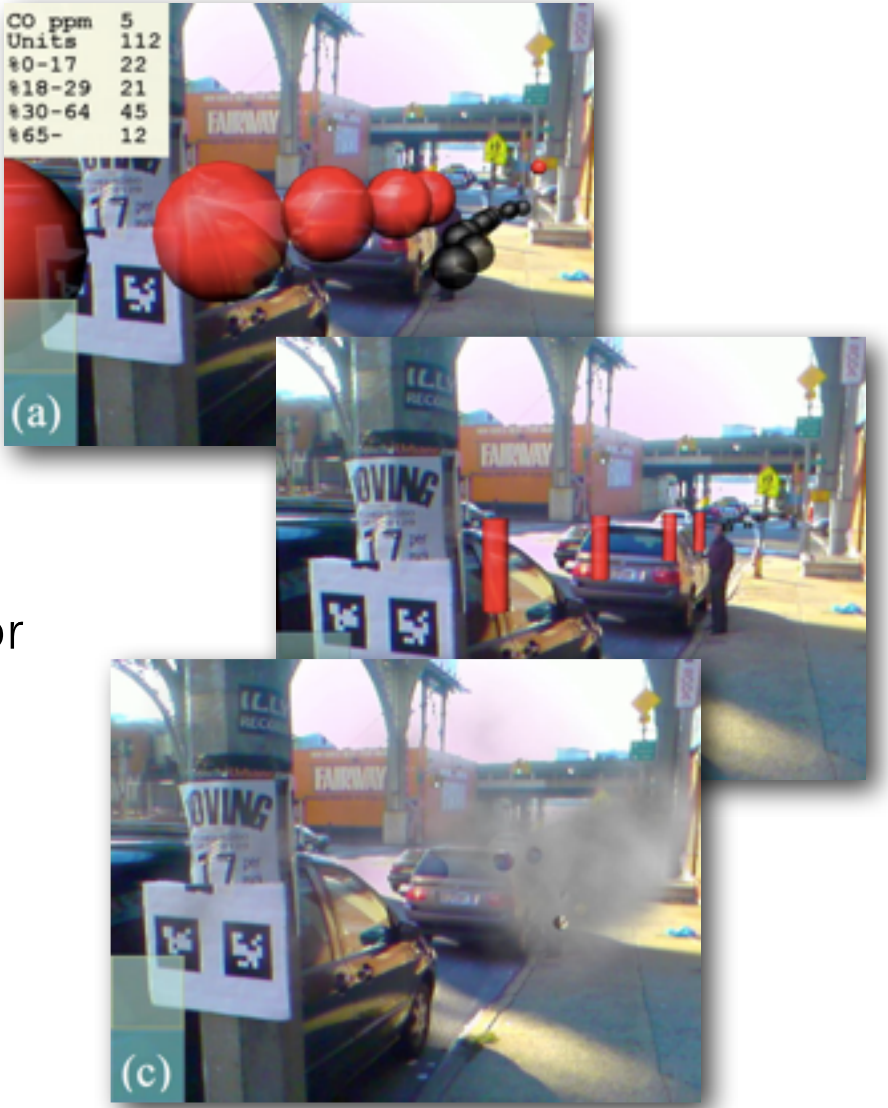
- Preparation: Acquire environmental data
  - collecting data by measurements
  - get data from EPA (Environmental Protection Agency)
- Three different modes for presenting data:
  - screen fixed display
  - world-fixed, augmented reality display

# Solution

## Visual Representation

- Spheres
  - Value is mapped to altitude and color
- Cylinders
  - Value is mapped to length and color
- Smoke
  - Value is mapped to density

- Position indicates, where the data was collected



# Evaluation

- Feedback from urban designers and planner
- Good to have measured data combined with visual inspection
  - Map data alone could not explain high CO level at the end of a street, visual inspection revealed that there were cars waiting to enter the highway
- Different preferences for representation:
  - Spheres vs cylinders
  - Clouds: phycological impact more important than location accuracy
- Live sensor data as improvement

# Personal Opinion

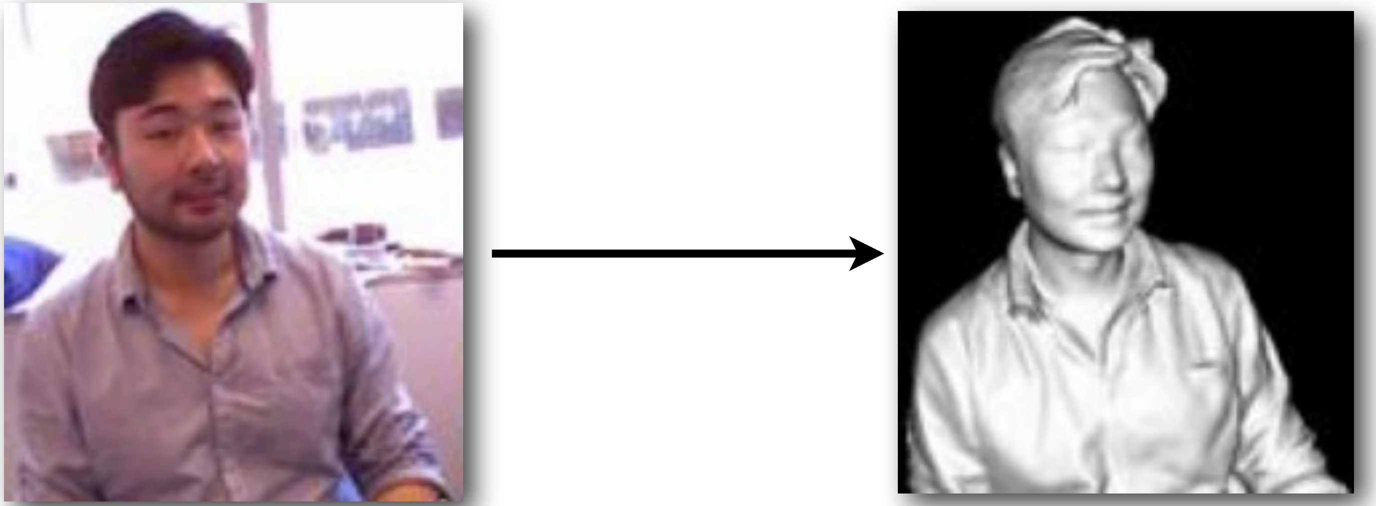
- Great idea to visualize measurements in the location where they were taken
- Importance/usefulness depends on need of urban planner/ designer
- Only few measurements available
  - User of the application could collect and share data
  - Accuracy of data might not be guaranteed

# KINECT FUSION

## REAL-TIME DYNAMIC 3D SURFACE RECONSTRUCTION AND INTERACTION

# Problem Statement

- Reconstructing an real-world object or scene
- Moving camera freely
- Process data in real time

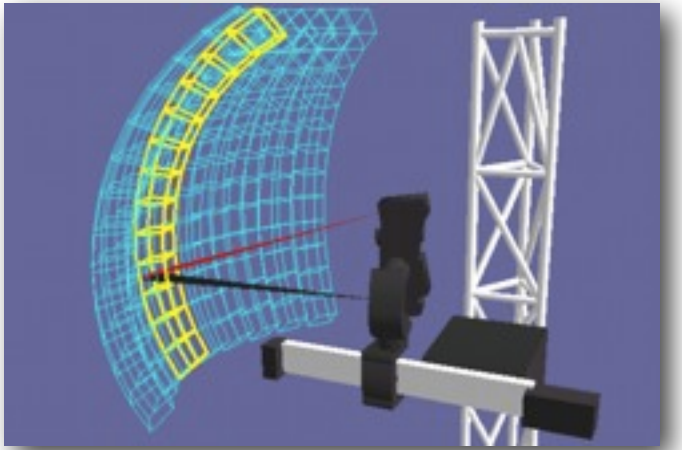


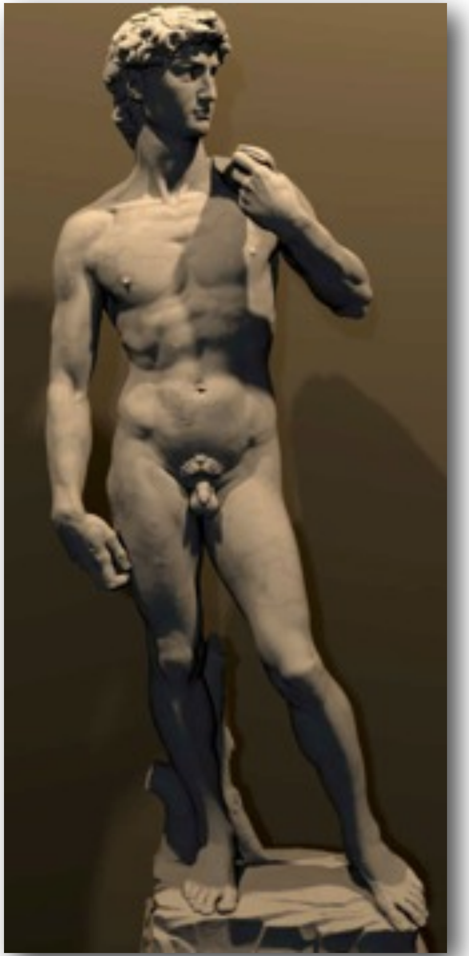
# Goals

- Real-time camera tracking and 3D reconstruction
- Capture detailed 3D models with exact geometry
- Dynamic interaction
- Infrastructure-less
- Support whole room construction and interaction

# Previous Work

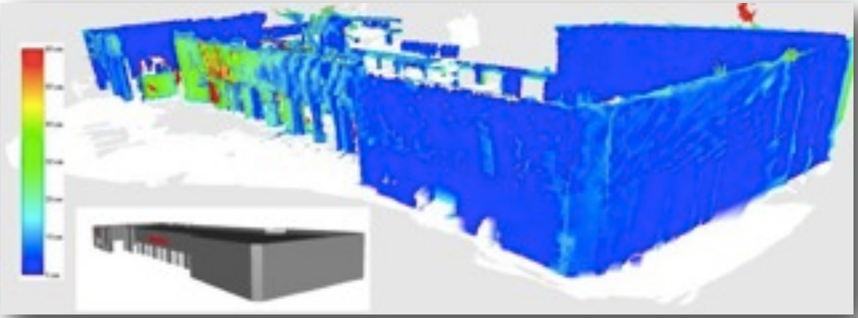
- The Digital Michelangelo Project [2000]
  - Laser rangefinder and cameras
  - Large and quite immobile



# Previous Work

- Real-Time Visibility-Based Fusion of Depth Maps [2007]
  - Reconstruction by combining multiple depth maps
  - Depth maps computed from images captured by moving camera



# Solution

- Generate 3D reconstruction in 4 phases
  - Get live depth map from Kinect, convert into 3D points an normals
  - Calculate the movement of the camera relative to the scene
  - Derive global camera position, update consistent 3D model
  - Render view of the volume and 3D surface with raytracing

# Solution

Shahram Izadi et al. :

KinectFusion: Real-Time Dynamic 3D Surface Reconstruction and Interaction

- Use commodity hardware: Kinect
  - Cheap, Portable
  - Provides depth map
  - Provides RGB camera

# Solution

- Moving the camera leads to
  - New viewpoints
  - More details of the scene
- Reconstruction becomes more complete over time

- RGB camera not used for 3D model, but can be used generate texture for the model

# Solution

- Object segmentation through direct interaction
  - Scan complete scene first
  - Move an object around
  - System monitors 3D reconstruction and observes rapid change in the model
  - Repositioned item is segmented from the background
- Augmented reality: interaction with the model
  - Place (virtual) objects in this 3D reconstruction
  - Shadow from virtual objects on real-world objects
  - Reflection of real-world objects on virtual objects
  - Collision detection

# Solution

- Static background, moving foreground
  - Algorithm can distinguish between moving foreground in the scene and camera movement
  - Background movement use for camera tracking
  - Foreground movement tracked separately
- Intersection between foreground and background can be determined
  - Intersection is recognized as touch
  - Allows multi-touch on every surface

# Solution

# Personal Opinion

- Great success, fascinating technique
- Opens new possibilities for applications
  - Robotic
  - Entertainment
  - Designing, planning
- Handheld camera, but visualization on screen
  - People in the scene not necessarily see the augmented reality
  - Combination with other techniques lead to new possibilities (projection)
- Good to see interesting and useful application of Kinect

# SUMMARY

Many applications for handheld augmented reality:

- Indoor Navigation with Activity-Based Instructions
- Fused Gyro & Face Tracking for Touch-free interaction
- Situated Visualization Techniques for Urban Site Visits
- Real-Time Dynamic 3D Surface Reconstruction and Interaction

Potential for improvement, but most techniques seem advanced