Augmented Reality –
Linking Real and Virtual Worlds

Suya You
Integrated Media Systems Center
Computer Science Department
University of Southern California
Outline

- Virtual Reality (VR) vs. Augmented Reality (AR)
- AR system and applications
- AR technical barriers
- Current research status & achievements
- Future direction
Virtual Reality (VR)

- Virtual
- Interactive
- Immersive

a computer generated, interactive, three-dimensional environment in which a person is immersed
Everything is Computer Generated (Virtual)
The Virtual World is Interactive
The User is Immersed
Virtual Reality

Advanced computer, networking, display, interaction, and graphics technologies allow us to create *photorealistic* 3D scene, and *view*, *hear*, *touch*, and even *smell* the created world,

However,

Everything is still *not* real. The user cannot sense the real world around him/her.
Augmented Reality

Augmented Reality (AR)
Supplements the real world with virtual (computer-generated) objects that appear to coexist in the same space as the real world.
Why AR?

AR offers an intuitive and natural means for people to navigate and work effectively in the real world - *links real and virtual worlds*
Augmented reality will further blur the line between what's real and what's computer-generated.

Reality-Virtuality Continuum (Milgram, P., H. Takemura, 1994)

- Virtual Environment
- Augmented Virtuality
- Augmented Reality
- Real Environment
AR Applications

Medical

UNC
MIT
Applications

Engineering

Columbia University

ECRC
Applications

Engineering

USC

Boeing
Applications

Engineering

Siemens Research

UW
Applications

*Consumer Design*

Tokyo University

USC
Applications

Consumer Design

ECRC

USC
Applications

Design and Inspection

USC
Applications

Commerce

Broadcasting and Advisement
Applications

Entertainment

MR Systems Lab

USC Golf

USC
Applications

**Entertainment**

*Daredevil*, *X-Men 2* (Rhythm & Hues /IMSC/USC)
Applications

Navigation

Robot navigation  Personal navigation
Applications

Military

Simulations & training
Information enhancement
Missions & operations
How Does AR Work?

The idea of mixture of virtual with real objects is not new. Hollywood and photography people have used it since very long before…
Augmented Reality (AR) - supplements the real world with virtual (computer-generated) objects that appear to coexist in the same space as the real world

- Combines real and virtual objects in a real environment
- Runs interactively, and in real time
- Registers real and virtual objects with each other
A Brief Look Back in Time

Ivan Sutherland’s vision of AR - User is “inside” the computer

- 1962 "Sketchpad“ dissertation
- 1965 "Ultimate Display"
  - Real and synthetic objects coexist
- 1968 “Head Mounted Display” AR System (MIT, Harvard, Univ. of Utah)
  - Graphics (vector mono)
  - HMD (optical see-through)
  - Tracker (mech. on head)


Sutherland’s 1968 System
35 Years Later...

- Moore’s law
  - Faster computers and graphics generators
- Display advances
  - HMD, LCD, OLED, laser,...
- Engineering, algorithms, and infrastructure
  - Tracking systems – GPS/RF, optical/IR, ultrasonic, magnetic, computer vision ...
- AR system components have made transition from laboratory to commercial domain
  - Customers want it!
...But AR is Still in the Laboratory

- Sutherland’s vision of AR is still “experimental”
  - AR System – components and classification
  - Main technical barriers to producing a practical AR - as envisioned in 1965
  - Current research status and achievements
  - Future research direction - consider strategies to overcome the barriers
AR System Components

- AR system classification
  - From display viewpoint
    - Optical see-through based
    - Video see-through based
    - Monitor based
  - From tracking viewpoint
    - Artificial feature
    - Natural feature
  - From application viewpoint
    - Indoor
    - Outdoor
Optical See-through AR

- Scene generator
- Head Tracker
- Monitors
- Graphic images
- Optical combiners
- Real world
- Head locations
Optical See-through AR (con.)
Video See-through AR

Video of real world

Scene generator

Graphic images

Video compositor

Combined video

Head tracker

Video cameras

Monitors

Real World

Head locations
Video See-through AR (con.)
Monitor Based AR
Artificial Landmark Based AR
Natural “Feature” Based AR

AR with a computer vision based tracking

Using GPS/INS for pose estimation
“Classic” Sutherland AR vision barriers
- See-Through HMD
- Tracker
- Computing
- HCI

Example systems
- Boeing
- Columbia University (MARS)
- ONR (BARS)
Technical Barriers - towards Sutherland’s AR vision

Boeing wire harness experiments (1997)

- **Application barriers**
  - Interface software
    - not optimized for task
  - Tracking/calibration
    - Robust, accurate, low latency
  - HMD - resolution acceptable
    - Weight, social acceptance
Technical Barriers - towards Sutherland’s AR vision

Columbia “MARS” & ONR “BARS” experiments in how humans use and interact with outdoor AR systems

- **Application barriers**
  - HMD – resolution and brightness
  - Added PDA displays for off-eye info
  - Tracking - GPS and tilt/compass
  - Dropouts, latency, and accuracy
  - System Integration – bulk, wires, power,...
Summary of “Classic” AR Barriers

“Classic” Sutherland AR vision barriers (See-Through HMD, Tracker, CPU, HCI)

- Tracking - primarily orientation and latency
  - Issues of *scale* – larger spaces, less infrastructure
- Display - brightness, resolution, and weight
  - Issues of *sight* – poor substitute/aid for human visual system
- Bulky AR gear and cabling – computation load
  - Issues of *speed* – smaller, cheaper, faster, lighter, low-power
- Interface design and software – human factors / task support
  - Issues of *design* - optimally applying technologies - hands-free, voice, gesture, multi-displays
“Classic” AR - Technologies on the Horizon

- Tracking - \textit{(scale)}
  - High precision gyros
  - Next generation GPS, G3 cell phones, Bluetooth
  - Computer Vision…
  - UWB (Ultra-Wideband) arrays

- Display - \textit{(sight)}
  - OLEDs (bright), LCD resolutions…

- Bulky AR gear and cabling \textit{(speed)}
  - Ubiquitous digital interfaces - IEEE1394 firewire
  - Moore’s Law…

- Interfaces and software \textit{(design)}
  - Hands-free, voice, gesture, multi-display …
  - More testbed development and applications research…
Beyond “Classic AR”

- Sutherland’s vision of classic AR is still limited ... *sight, speed, scale,* and *design*... Yet, research and prototypes abound and the community of people working in AR seem to be grower rapidly...

- Video see-through AR
  - Started (1992) at UNC
  - Appears to be the most favorable extension of AR vision
Technical Barriers – Video AR

AR using artificial landmarks

- Requiring pre-designed landmarks – tracking & recognition
- Assuming calibrated features are always in view - limited workspace

(Video)
Technical Barriers – Video AR

AR using natural features

- Not requiring any calibrated features – natural feature tracking & pose estimate
- Pure vision approach and computation – accuracy, robustness, speed

(Video)
Technical Barriers – Video AR

AR using hardware devices

- Not requiring any calibrated features – off-the-shelf products

(Video)
Technical Barriers – Video HMD & Tracking

- **Pros:** Self-contained – simplicity, passive
  - View and track through same camera – human model
    - Z-compositing, latency-match, ...

- **Barriers:** *sight, scale, and speed*
  - *Sight* – camera and display limits of resolution, field of view, dynamic range, and focus
  - *Scale* – tracking assumes calibrated features in view
    - usually indoors, but outdoor operation is possible
  - *Speed* – robust vision tracking could consume 2-3 orders of magnitude CPU speedup (or HW accel)

- Most barriers not present with see-though HMD
Camera Sight Barriers

- Camera resolution/frame rates lower than displays
- Sensors are *finally* free of analog (NTSC) interfaces, but computers and interfaces remain limited
  - Firewire interfaces – 1288x1032 @30Hz = ~320Mb/s (byte/pixel)
  - Tracking could use 60-100Hz ~1Gb/s
- Requires a new generation of sensors and interfaces
  - Perhaps spurred by consumer HD TV/camcorders
Display Sight Barriers

We are surrounded by visual data -- except when we work with computers...

- HMD FOV is limited – 40-60 degrees – tunnel vision
  - Lens system size, weight, cost
  - Display element resolution and aspect

90 degree FOV
FOV Implications

- **Informal task experiment - R/C car driving**
  - Indoor lab with 40/120 degree camera(s) and display
  - 40 deg - can not turn without hitting furniture/walls

- **Golf task experiment in “life size” AR**
  - Golf putt game through a video AR system using Sony Glasstron 800x600 HMD and matching camera (FOV ~30 deg)
  - Only *tall* people see ball and hole at same time!
GRIDS (Geospatial Registration of Information for Dismounted Soldiers)

- Outdoor AR tracking experiment
  - Single technology is inadequate
    - INS, GPS, Vision
    - One degree error = 11 pixels
- Hybrid approach can overcome many barriers
  - Robustness, Accuracy, Speed
CV pose tracking assumes calibrated features in view
  - Indoor pre-calibrated environments

Outdoor operation is feasible
  - Autocalibration – determine 3D parameters of scene points and lines (modeling while tracking)
  - Panoramic based pose tracking – wide-baseline panoramic images for pose without calibrated features
  - Model based pose tracking – 2D-3D correspondence
  - Hybrid approach – vision/inertial/GPS sensors

These and other CV methods depend on robust and fast 2D pixel-level computing (features or flow)...
Computing Speed Barriers

- **Speed** – robust CV tracking could consume 2-3 orders of magnitude CPU speedup (or HW accel.)
  - Natural feature tracking (2D) is slow – much more complex than BW square target detection/tracking

USC *Fastrack* – “is capable of tracking hundreds of features from one frame to another with sub-pixel accuracy in only a few seconds on a standard personal computer, effects artists can process roughly 40 percent of movie shots without having to provide extensive input to the computer.

Rhythm & Hues used this software with great success on several recent blockbuster films, including Daredevil and X-Men 2.
Computing Speed Barriers

- 2D tracking takes ~1-10 sec/frame depending on complexity of algorithms and these are “simple” algorithms - no recognition or abstraction
  - Speedup of 30-300 to achieve 30Hz frame rates
- Resolutions (256x256) need to increase to 1kx1k (or more)
  - Speedup of 3
- Frame rate increase from 30Hz to 90Hz (or more)
  - Factor of 3
- Total of 100-1000 speedup in computing consumed before we can really start thinking about what to do with excess cycles...
- Very few hardware efforts (Sarnoff – highly app specific)
  - Similar situation in graphics in 1980’s prior to SGI
Summary Video AR Barriers

- Tunnel vision in video HMDs is likely to remain a *sight* barrier for several years
- *Scale* and *speed* barriers mean that usable video AR is confined to small work areas with calibrated environments
- Toys, training, and medical AR applications seem most tolerant of small scale & tunnel vision
  - Highly structured environments with infrastructure
Beyond HMDs - New Frontiers for AR

- **Projective Mixed Reality (PMR)**
  - No HMD – unencumbered viewing
  - Virtual objects & textures + real world = PMR

- **Modeled Mixed Reality**
  - Virtual elements as 3D model
  - 3D Model + multiple projected textures = AR object

- **PAD AR**

- **Offline AR**
Projective Mixed Reality - FlatWorld

- Developed at ICT/IMSC/USC for military training
- Projected stereo “digital walls” + functional stage props
- Passive haptics – “the world works”
  - Doors/windows open, user’s move freely
  - Modular “flats” and “wild” units (optional props)
- Novel narrative training possibilities
- Other applications
  - Boeing wire harness guidance
    - projected info + real wires
    - No tracking, no HMD, no wires
    - Leverage projector technology gains in past 5 yrs - LCD/DLP
FlatWorld Concept

Paramount Studios Flats
FlatWorld Preview

FlatWorld door/window flat with projected imagery
FlatWorld Applications

Schematic of a multi-room digital flat simulation facility
Modeled Mixed Reality

- Merge virtual elements with real world elements represented as 3D models
  - 3D camera – pixel by pixel modeling
    - MR Labs – virtual sets
    - UNC – Tele-Immersion
    - CMU – Z-Key
    - USC – AVE
MODELED MIXED REALITY

- AVE merges 3D models with real world videos
PDA Augmented Reality

- Hand-held AR display
  - Computing power, display…
- Technology merger with great potential
  - Convergence: cell phone, PDA, ubiquitous computing, Bluetooth
  - Unobtrusive, social patterns exist
- USC Tourist Machine
  - PAD, Bluetooth GPS, Vision tracker
  - Whole campus 3D model, map, image, video, text…
Offline Augmented Reality

Offline AR – a sidestep *real time* speed bump

- Most commercially developed
- Allows/encourages manual intervention and interaction – limits demands on the machine
  - Siemens - pipe visualization/inspection
  - NASA/USC/Boeing – AR training video authoring
  - RealViz – “match move” for special effects
- Interaction and playback can be real time
Summary: AR Barriers

- **AR barriers identified in new context:**
  - **Scale** - Size of working area
  - **Speed** - Vision tracking, sensor/processing rates
  - **Sight** – Lightweight, wide FOV, hi-res display

- **Atoms and Bits**
  - **Speed** is primarily a “bits” barrier
    - Material science and Moor’s Law make predictable advances
    - Industry and market impediments
  - **Scale** and **Sight** barriers are more related to “atoms”
    - Innovation is needed...
    - Less predictable, more creative solutions, more difficult...
Research Status & Achievements

HMD Display

- VGA, SVGA - 640x480, 800x600
- Monochrome, 0.1, 0.2, 1.5 million pixels
- FOV: 30, 45, 51 degrees
- LCD, OLED, Retinal-scan
- Weight: < few grams
- Cost: <$$$

65
Concept: A universal visor display for all visual information
- Two eye viewing, stereo option, wide option (80 degree FOV w/o stereo)
- Imagery, maps, overlays

Benefit: User has single window for all sensors, including his own Night Vision
- No additional devices needed for other digital information
Research Status & Achievements

Tracking and Registration - Indoor

- Pre-calibrated environments (landmark, model…)
- Computer vision, Optical, Ultrasonic, RF, UWB, Magnetic, INS…
  - HIT Lab AR Toolkit, USC MKTracker libraries, UNC Hiball, Intersense Inc.
Research Status & Achievements

**Tracking and Registration - Outdoor**

- **Computer vision**
  - Natural feature, Model based, Affine/Projective vision, Autocalibration, Panoramic image...

- **Hardware**
  - Compass, Inertial gyro, accelerometer, Laser range finer, GPS...

- **Hybrid approach**
  - Magnetic/vision, Inertial/vision, Inertial/ultrasonic...

- INS tracker
Research Status & Achievements

Computation

Handheld, Wearable, Tablet PC
- CPU: 400Mhz (PDA), 1Ghz...
- Wireless: 802.11b/g, Bluetooth
- Interaction: voice, gesture ...

1GHz ultra-personal computer (OQO, Inc)

MIThril (MIT)
Research Status & Achievements

Human Factors and Perception

- Cognitive science
  - Latency - <10 milliseconds
  - Registration error – relative, absolute.
  - Depth perception – resolution, conflict
  - Fatigue and strain – monocular, binocular

- Visualization
  - Rendering – CG acceleration
  - Annotation – 2D, 3D, data density
Academic and Research

There is still long way to go. How do we get there…

- Most of the needed technologies require collaborations between several disciplines as well as good engineering
  - CS, EE, Physics, Human Factors, Psychology, Industrial Eng,…
  - Commercial Industry/University partnerships

- Requiring Proposals and Funding from Govt and Industry
  - ARvika, Europe
  - Mixed Reality Labs, Japan
  - GRIDS, BARS, USA
ARVIKA – A Joint European Project

- Using AR technologies to support *development, production,* and *servicing*
  - Automobile
  - Aero plane
  - Machine, Automation

- Sponsored by German Federal Ministry of Education and Research
  - 89 Consortiums of firms, institutes and universities
  - Project starts from 1999. Expected product release in 2003/04

- Work on wide fundamental and application issues
  - Tracking,
  - Interaction
  - Presentation
  - Mobile communication
  - Information management
Ongoing Research

Development

Crashtest 10/1/198BL1
50 km/h
Ongoing Research

Production
Ongoing Research

Production
Ongoing Research

Mixed Reality Systems Laboratory - Japan

- Sponsored by Japan Key-Technology Research Center
  - Funded in 01/1997 by Japanese government and Canon
  - Extended to 03/2001 with three Universities

- Research themes
  - AR technologies
  - 3D image display
  - Sponsored 1999 ISMR
  - Applications - games
Ongoing Research

**BARS - Battlefield Augmented Reality System (ONR)**
New Generation AR Directions

- Projective Mixed Realities
- Modeled Mixed Realities
- PDA Augmented Realities
- Offline Mixed Realities

- Traditional AR still remain actively

*With hope:* this community’s hard work will hasten the day we see AR “practical AR” dotting the landscape... arts, entertainment, education, commerce.
Acknowledgements

- Members of CGIT/IMSC/USC
- Funding supports – NSF, DARPA, ONR, ARL, NASA, NGA, and industrial partners
- Collaborators and cited materials from the web and publications