Interactive multimedia

Overview

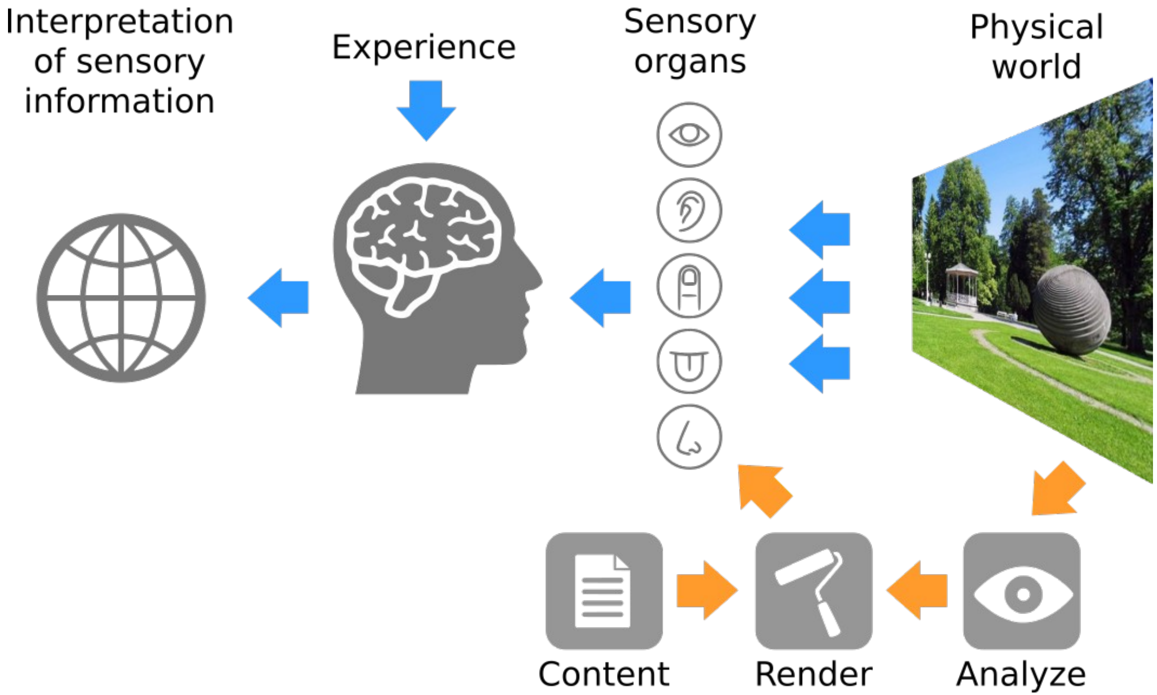
- Augmented reality and applications
  - Marker-based augmented reality
  - Marker-less augmented reality
  - Rendering virtual content
- Interactive surfaces
  - Multi-touch technology
  - Interactive tabletops

Reality–virtuality continuum

- Real world
- Augmented reality
- Mixed reality
- Virtual reality

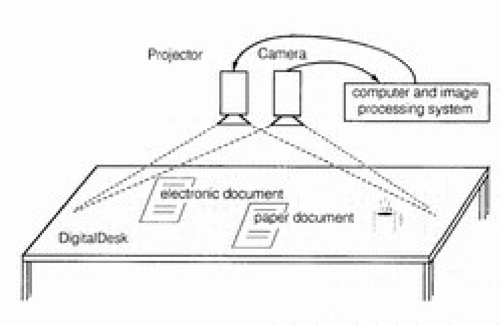
What is augmented reality?

- Reality is subjective
  - Sight
  - Hearing
  - Smell
  - Haptics
  - Balance
- Augmenting sensory information



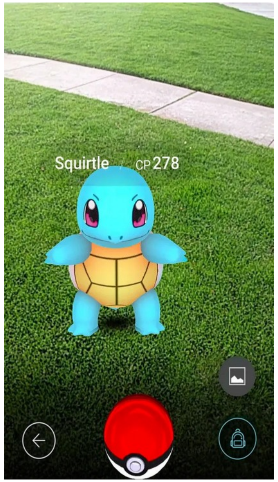
Augmenting visual information

- Superpositioning digital information on top of real imagery
- Who Framed Roger Rabbit (1988)
- Tom Caudell - Boeing (1990)
- DigitalDesk - Xerox & University of Toronto (1991)
- Virtual Fixtures (1992)



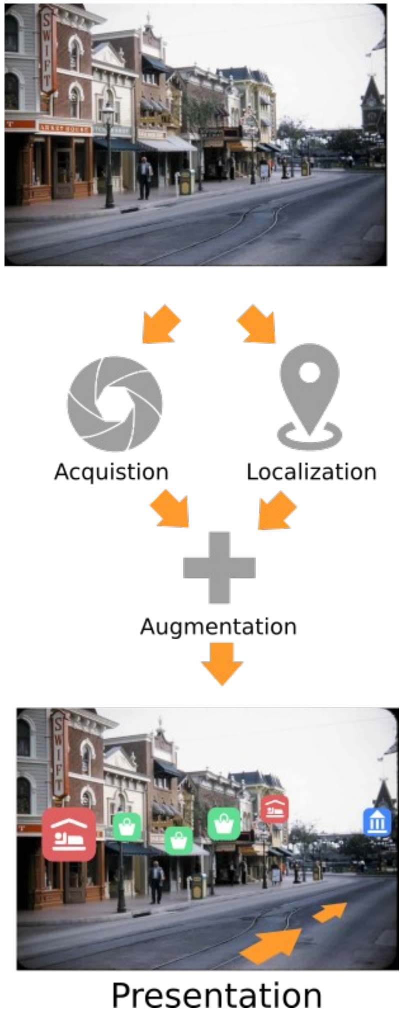
Types of augmented reality

- Anchor based
  - Global
  - Local
- Input based



Classical approach

- Visual information acquisition: camera
- Camera localization
  - Image: camera
  - Depth: depth camera
  - Other: GPS, WiFi, IMU
- Displaying augmented information: monitor, mobile phone, projector, smart glasses

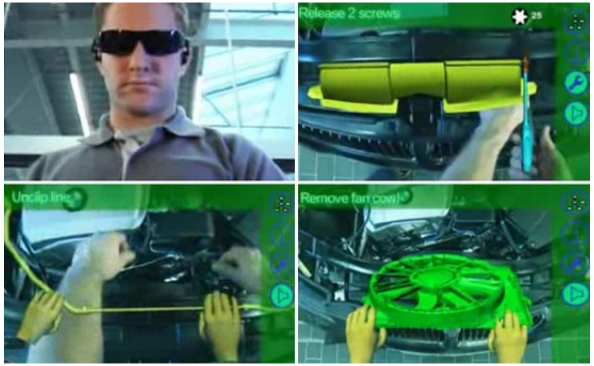


Application examples (TV)

- Olympic games 2004
  - Monitor/TV
  - Robotic camera
- USA elections 2008 – CNN
  - Hologram conference
  - 35 cameras
  - 20 computers

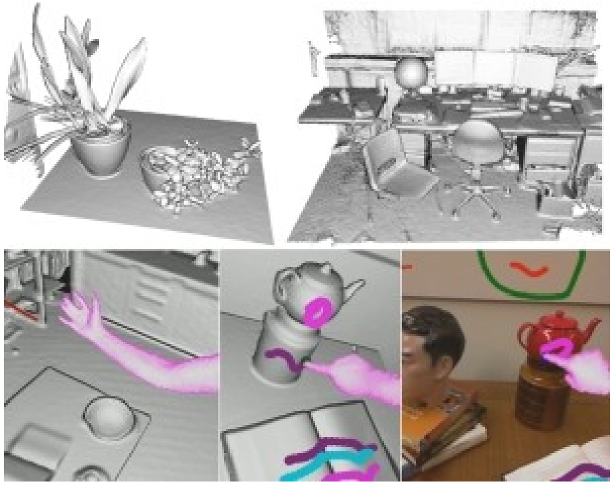
Application examples (Mobile)

- Mobile devices
  - Pokemon Go
  - Vuforia
  - ARKit, ARCore
  - Wikitude
  - BlippAR
- Wearables
  - BMW
  - Hololens



AR using depth information

- Depth cameras
  - Active (IR light)
  - Passive (Stereo systems)
- Automatic scene reconstruction
- Easier interaction with objects



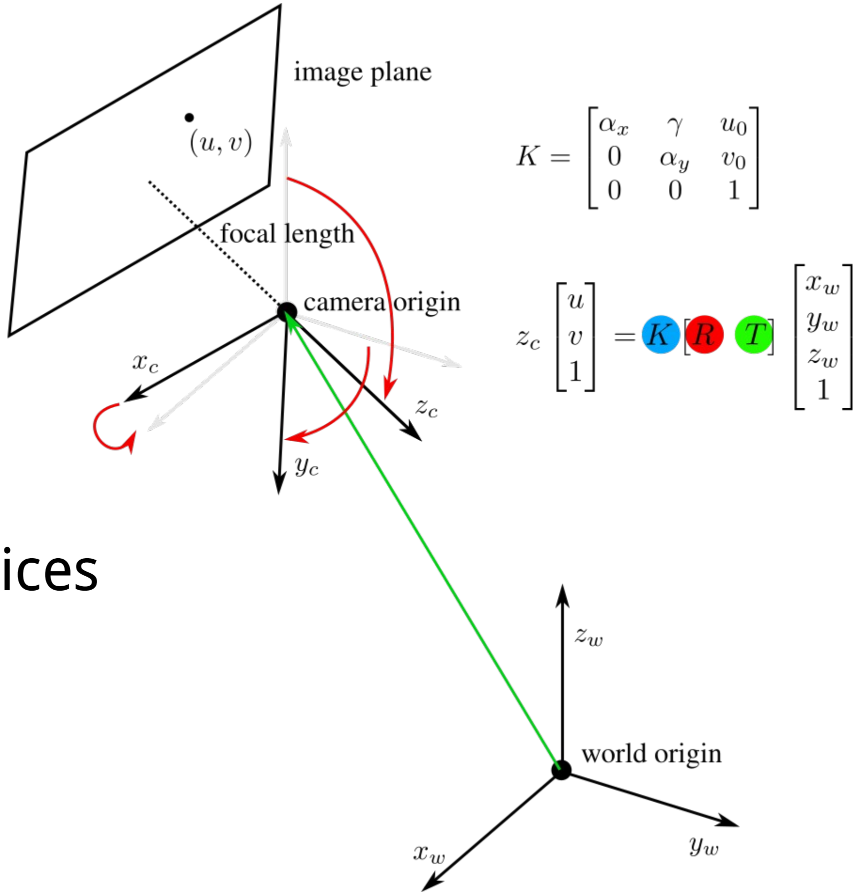
Izadi et al. "KinectFusion: Real-Time Dynamic 3D Surface Reconstruction and Interaction", SIGGRAPH, 2011

AR with visual anchor

- Localization with visual information
  - Detect key object in image
  - Determine relative position of the object to camera
  - Draw information with this relation

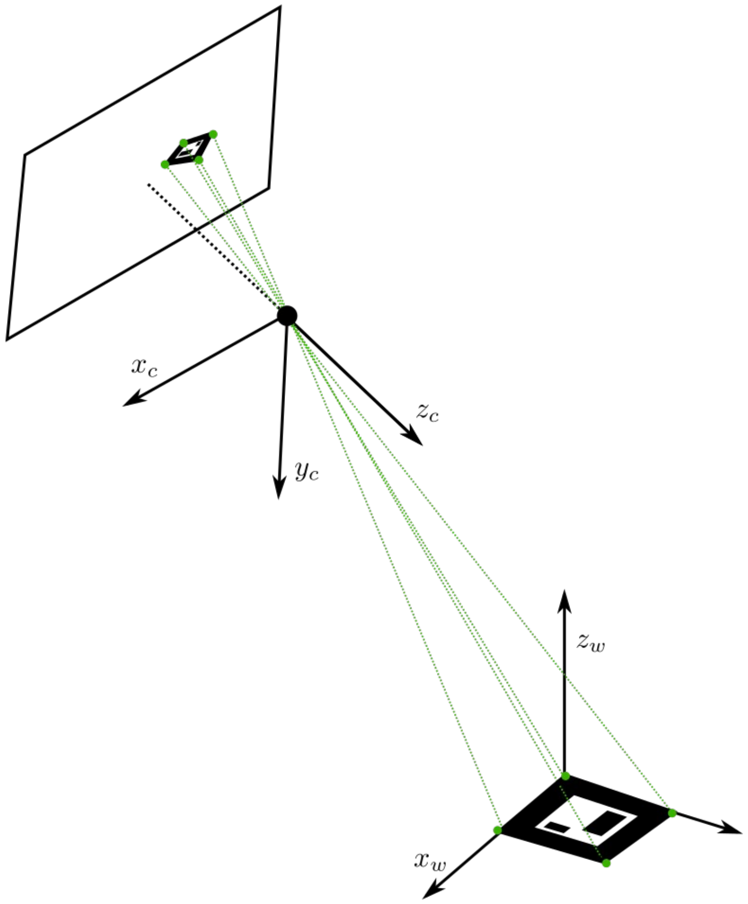
From point to pixel

- Transform 3D point to camera coordinate system (pixels)
- Required data
  - K ... camera calibration matrix (intrinsic parameters)
  - R,t ... rotation and translation matrices (extrinsic parameters)



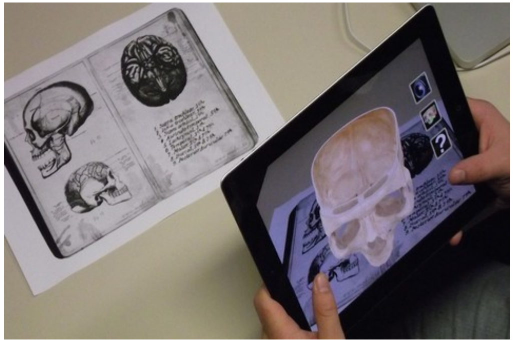
AR with binary marker

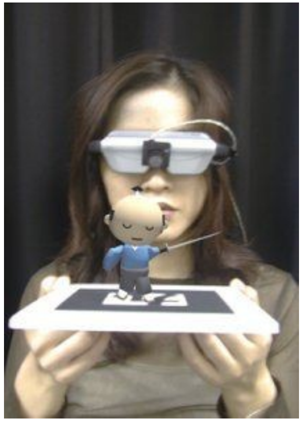
- Detect markers that are easy to detect and identify
  - Detect marker from edges
  - Identify marker with correlation
- Known marker size
  - Compute relation to camera
  - Use corners of marker to compute relative position



Applications of marker AR

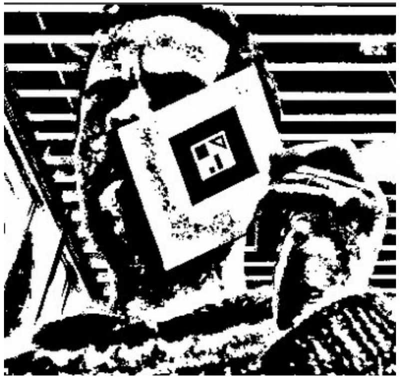
- Catalogs
- Books
- Tourism
- Gaming





Detecting a binary marker

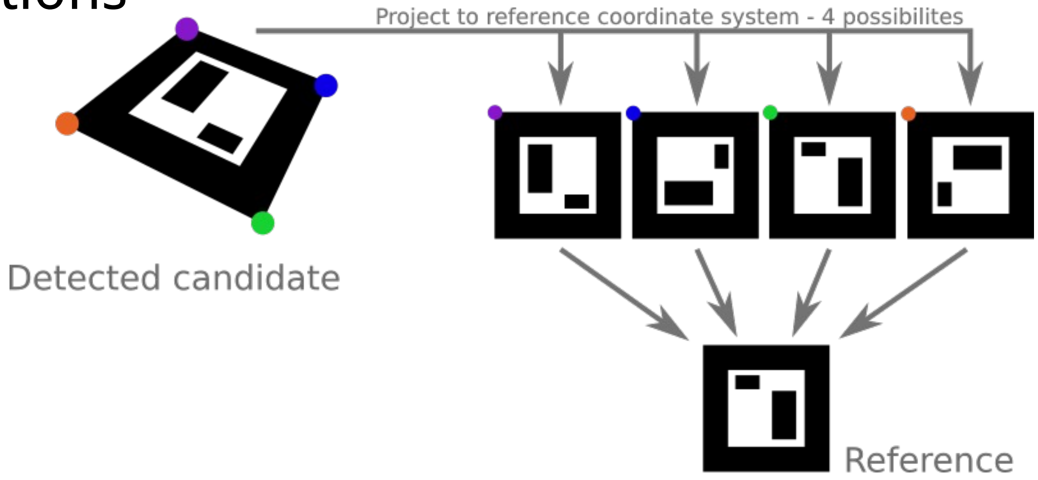
- Image processing approach
  - Speed
  - Robustness
- Finding possible candidates
  - Adaptive threshold
  - Trace contours
  - Estimate contours as polygons
  - Contours with 4 corners



Recognizing binary marker

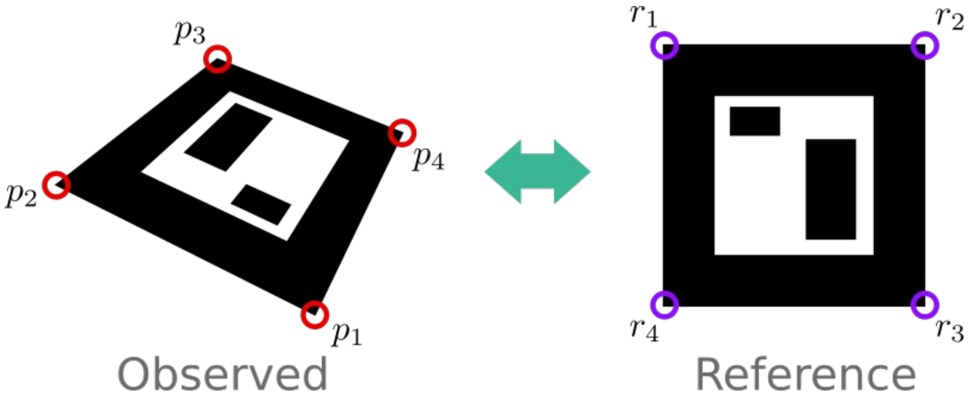
- Threshold on similarity
  - Project region to reference position
  - Normalized cross correlation
  - Orientation – test all four options

$$G(i,j) = \frac{(\mathbf{h}^T - \hat{h})(\mathbf{f}_{ij} - \hat{f})}{\sqrt{\mathbf{h}^T\mathbf{h}}\sqrt{\mathbf{f}_{ij}^T\mathbf{f}_{ij}}}$$

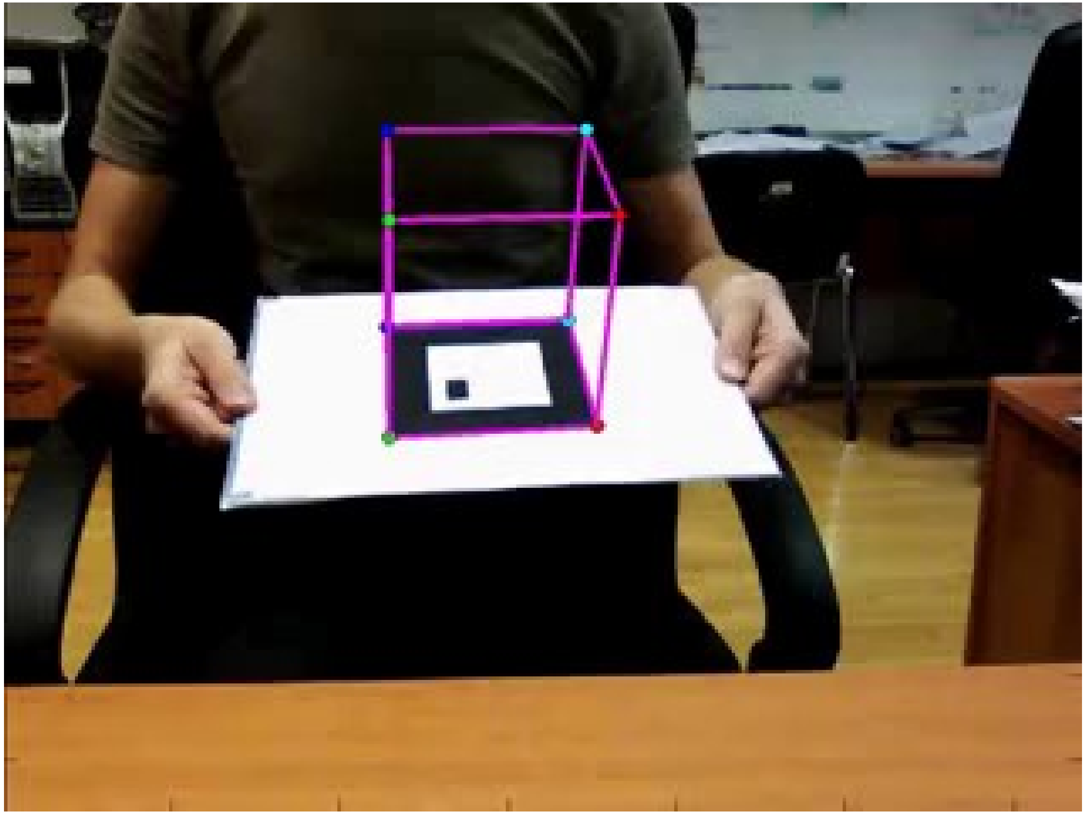


Generalizing transformation

- Determine camera extrinsic parameters using a detected marker
- Transformation between planes (homography)
  - Marker plane (reference)
  - Camera plane (observed)



Example video

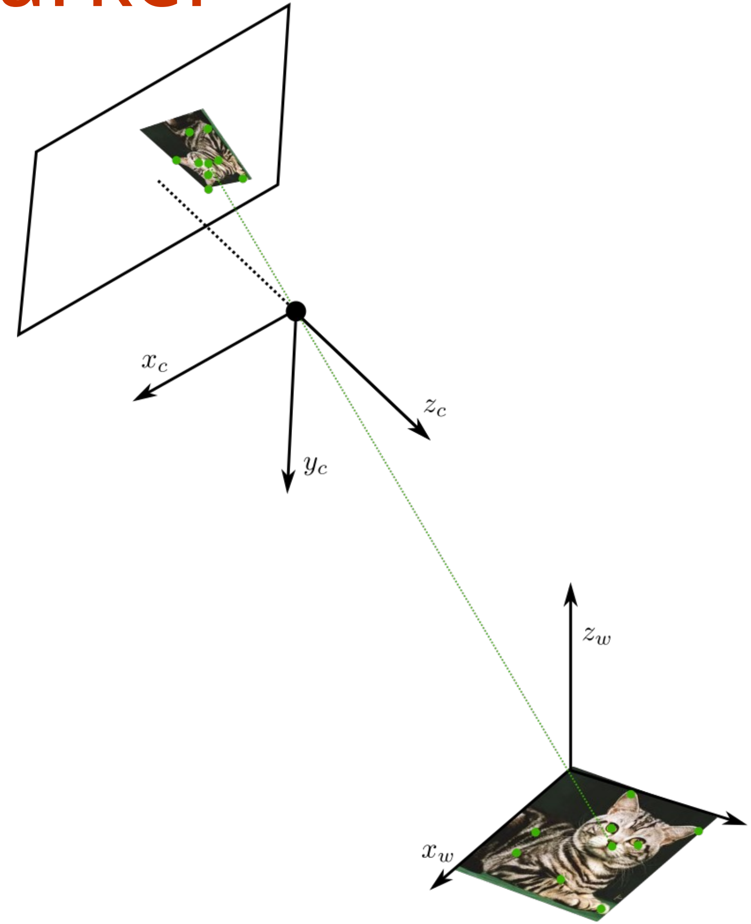


Problems with binary markers

- Artificial appearance
- Entire marker must be visible
  - No touching
  - No overlapping
  - Within the image
- Image quality
  - Contrast
  - Motion blur

AR with arbitrary planar marker

- Match an arbitrary surface
  - Describe local texture
  - Robust matching
- Less constrained can use existing textures from real world
  - Posters
  - Building facades

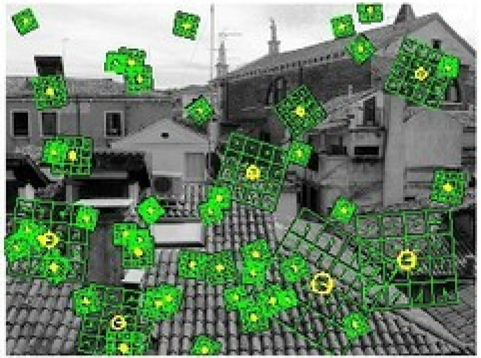


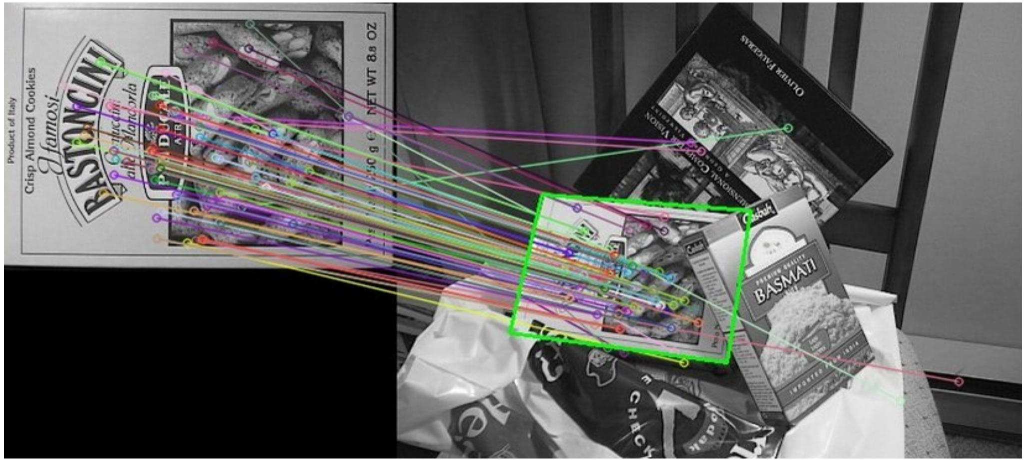
Augmented reality with planar marker

- Natural surfaces
  - Unable to detect corners robustly
  - Partial occlusions
  - Can detect feature-points
- Over-sample reference points
  - Not all points will match correctly
  - Robustly estimate homography

Matching keypoints

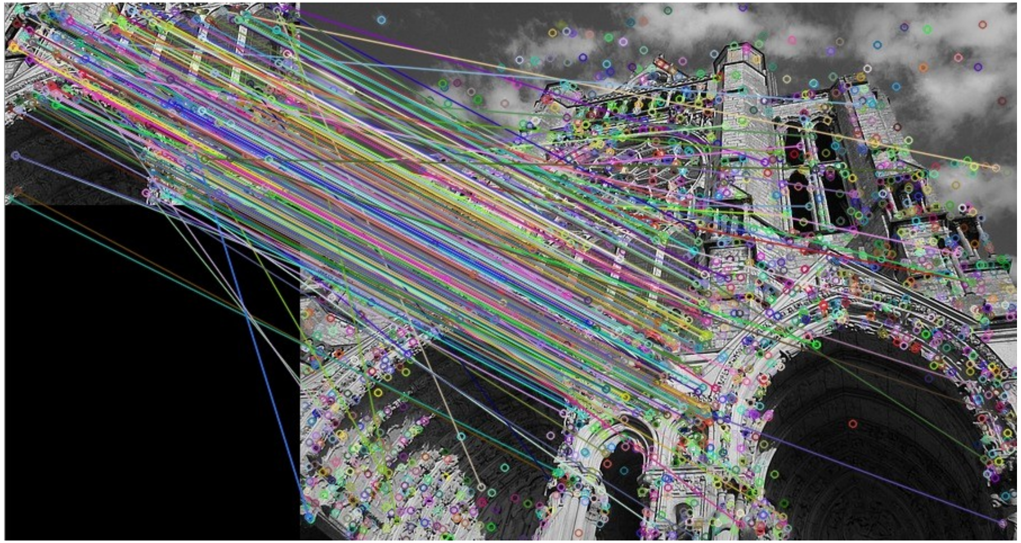
- Detector of keypoints
- Descriptor of regions
  - SIFT, SURF
  - BRIEF, ORB
- Matching descriptors
  - Distance function
  - Symmetric matches





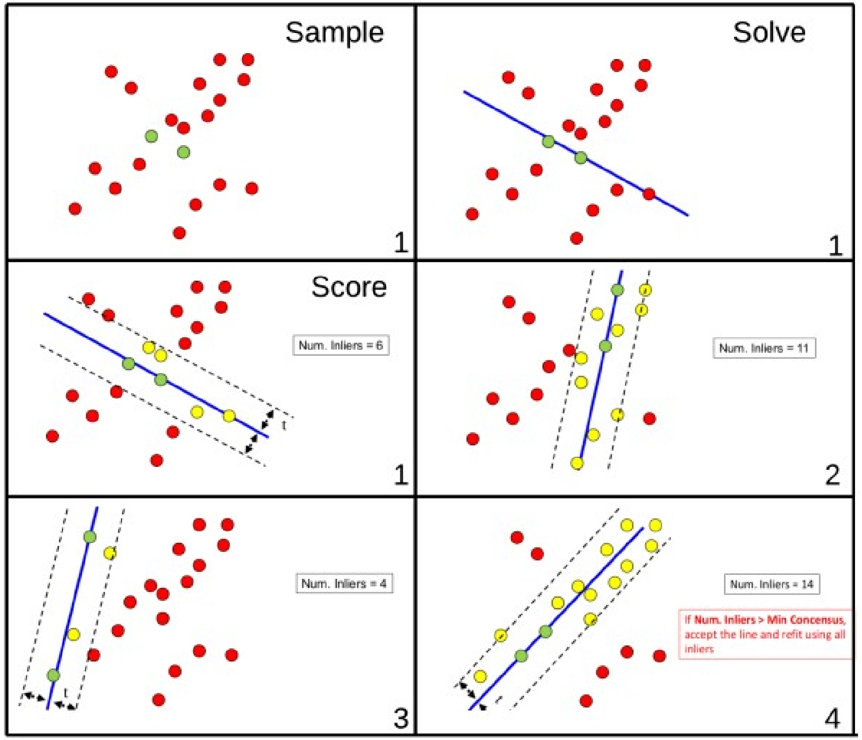
Robust estimation of homography

- Many correspondences
  - Over-determined system
  - Not all correspondences are correct
- Robust matching
  - Exclude outliers from calculation
  - Find sub-set of correspondences that agree on a model



RANSAC

- Random Sample Consensus
  - Meta algorithm (used for many tasks)
  - Probabilistic interpretation
- Repeat k times
  - Select random set of 4 correspondences
  - Estimate model – homography (DLT)
  - Look which other pairs agree with the model (projection from one plane to the other is small enough)
  - Take the model with largest support (inliers)



RANSAC for line fitting (source: F. Moreno)

Reference plane example

Disney AR coloring book

Urban AR

Beyond planar markers

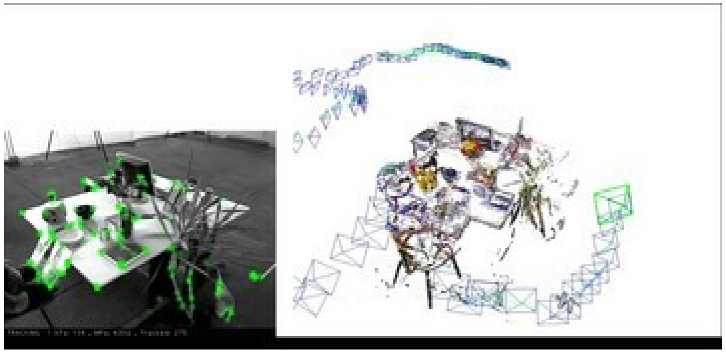
- Reference objects
- Deformable surfaces
- No reference

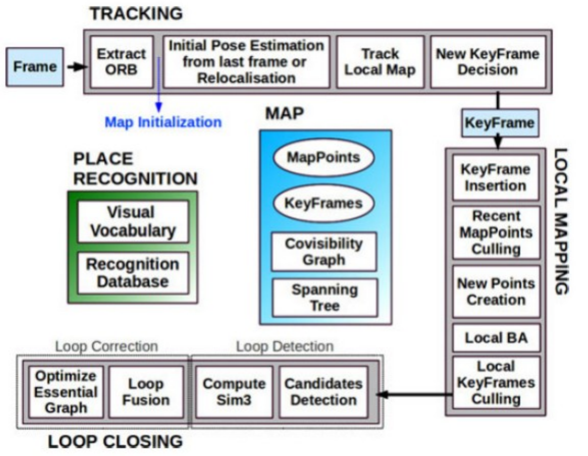
Positioning without an explicit anchor

- Entire world is an anchor
  - Detect points
  - Not planar
- Visual SLAM
  - Simultaneous localization and mapping
  - Mapping – building a 3D model of world
  - Lots of parameters - slow

ORB-SLAM

- ORB local features
  - Tracking
  - Localization
  - Loop closing
- Robust re-initialization
- Fast operation
  - Large environments
  - Co-visibility graph
  - Discarding redundant features

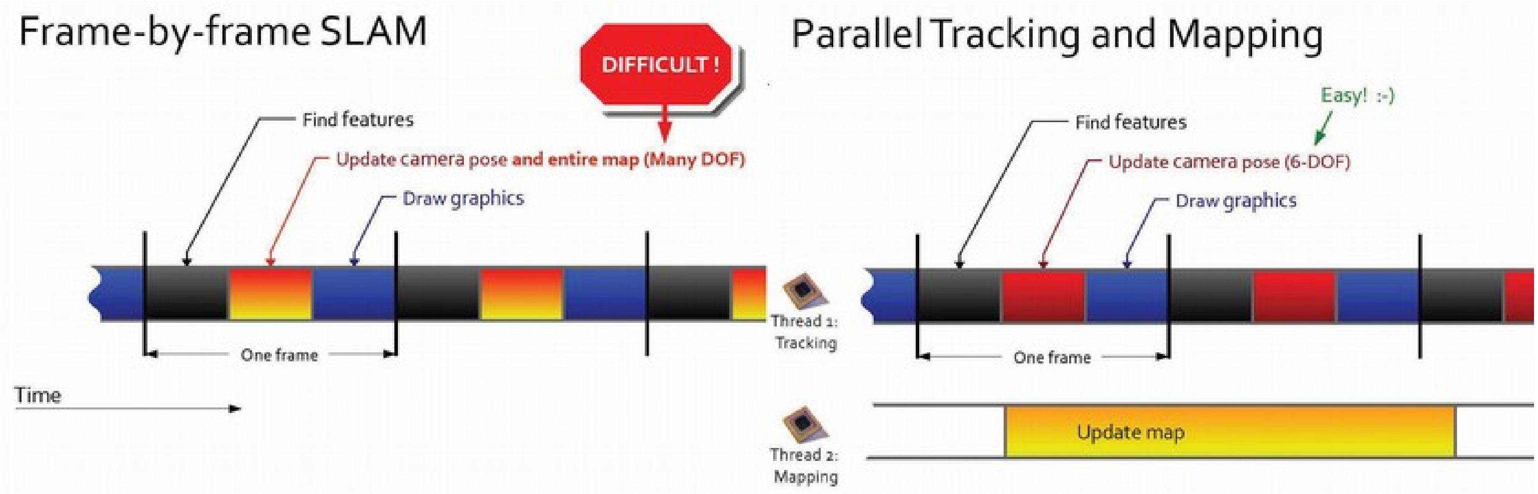




R. Mur-Artal, J. M. M. Montiel, and J. D. Tardos, "ORB-SLAM: A Versatile and Accurate Monocular SLAM System,"IEEE Trans. Robot. 2015.

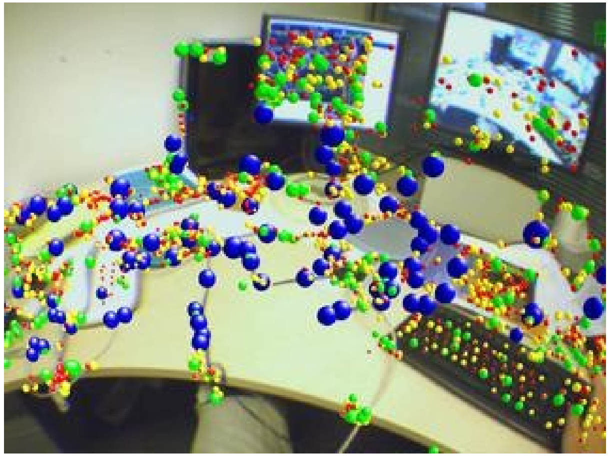
Parallel tracking and mapping (PTAM)

- Camera tracking
- Mapping of points



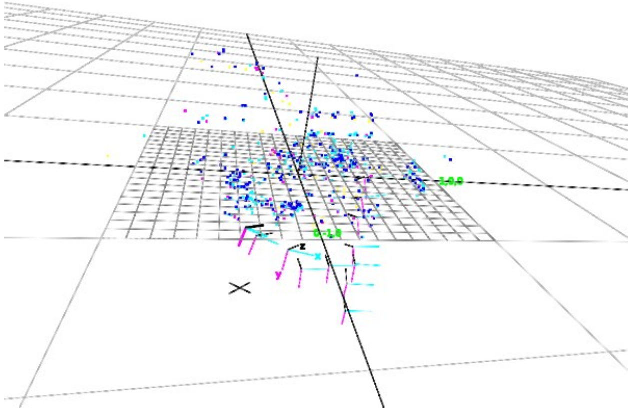
Camera tracking

- Motion model
  - Better starting position
- Re-project points from map to image
  - Find local matches
  - Correct camera position



Building a map

- Initialization
  - Stereo correspondences
  - Sideways camera panning
  - Possible initialization with a marker
- Updates
  - Key-frames (not every frame is processed)
  - Bundle adjustment



Initializing map building

- Internal
  - Correspondences (ORB-SLAM)
  - Known motion (PTAM)
- External
  - Sensor fusion (ARKit, ARCore)
  - Pre-built map (scene prior)

Initialization with scene prior

Realistic rendering

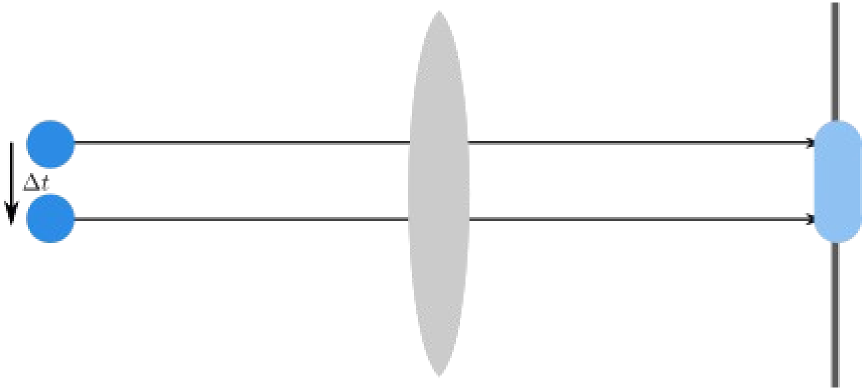
- Acquired images are degraded by various factors
- Augmented reality will be more immersive if these factors are replicated on virtual objects

A terrible image

Graphics don't match

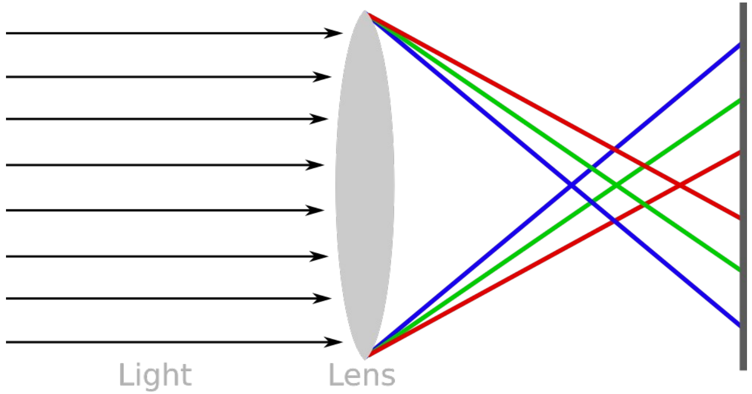
Motion blur

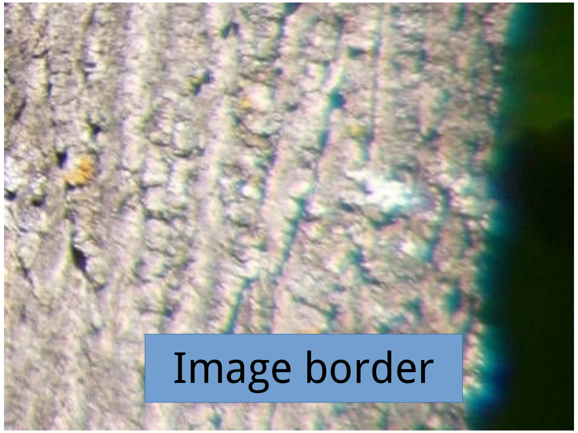
- Exposure time + rapid motion
- Simulated by smoothing image with directional filter



Chromatic aberration

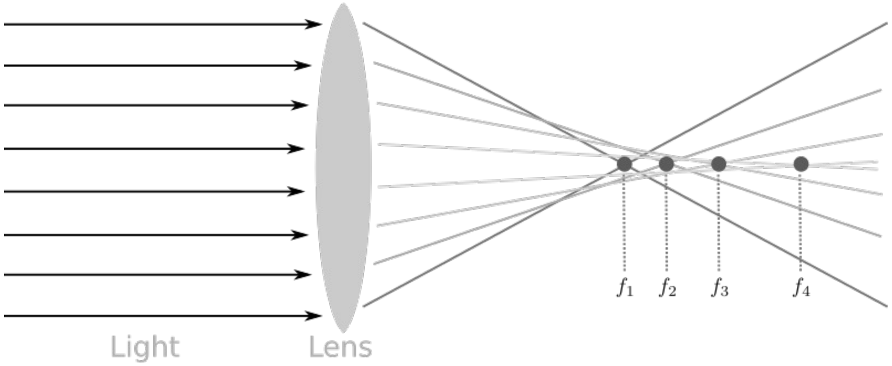
- Different wavelenghts bend under different angles
- Simulated by distorting individual color channels

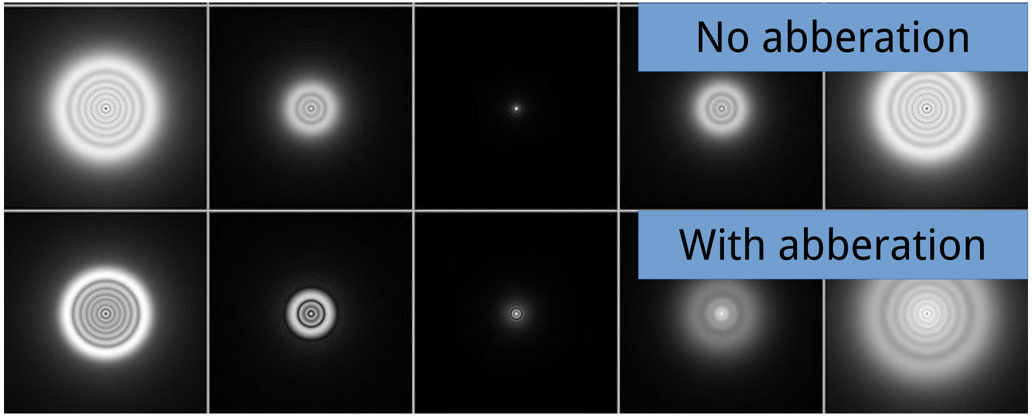




Spherical aberration

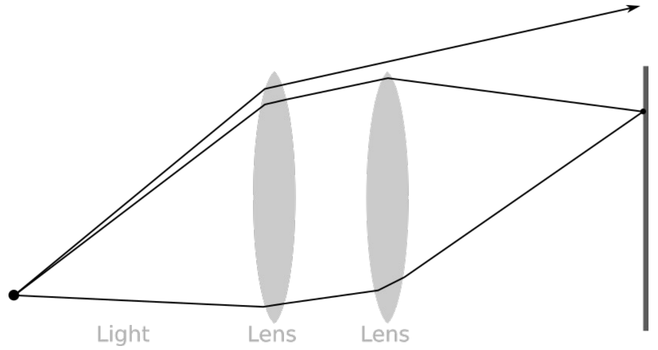
- Spherical lenses do not focus light perfectly
- Rays on the edge are focused closer
- Compensated by using multiple lenses





Vignetting

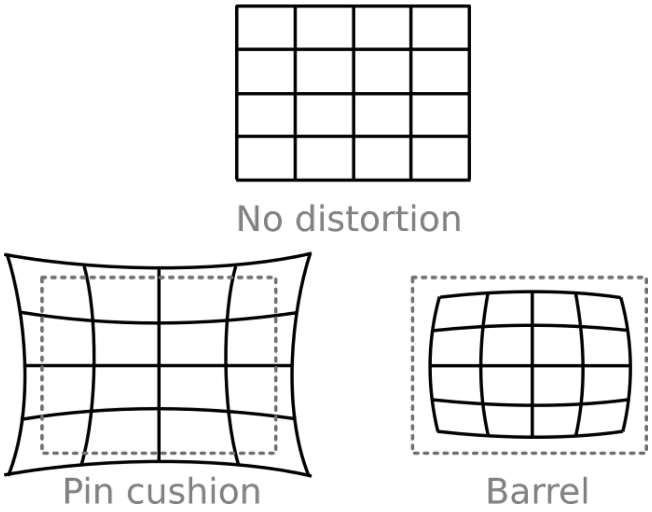
- Optical vignetting – multiple lenses

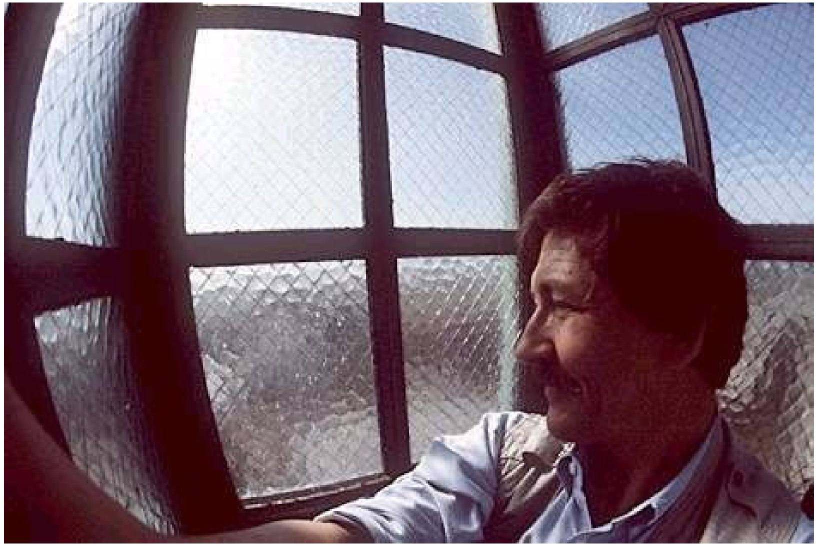


- Pixel vignetting
  - Digital sensors
  - Angle dependence
  - Software compensation

Radial distortion

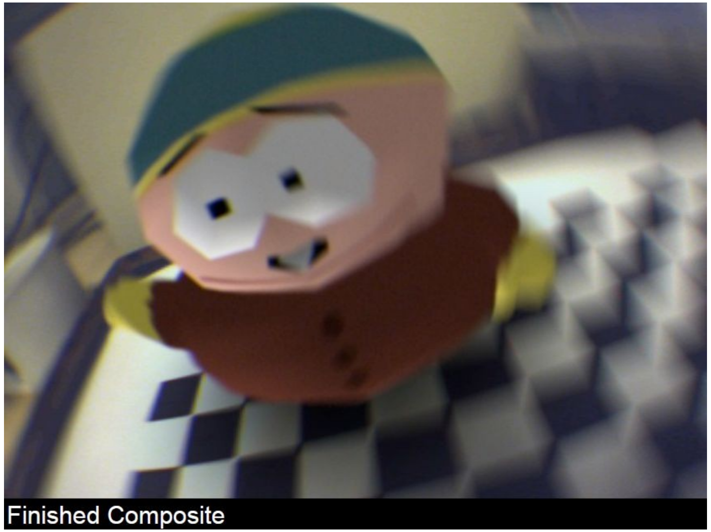
- Imperfect lenses
- Deviations are apparent near the edge of image

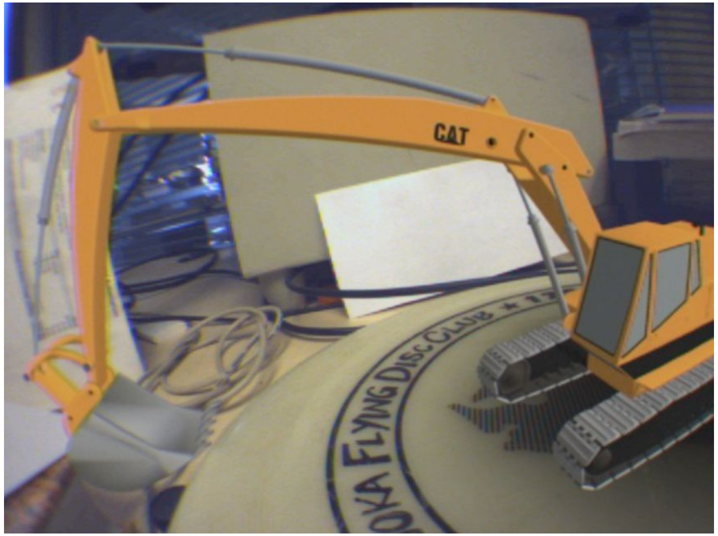




Illumination

- Matching lighting setup
  - Position and intensity
  - Ambient conditions
- Determine parameters
  - Number of lights
  - Position, strength
  - Environmental illumination





Gerhard Reitmayr, Axel Pinz, Visual Coherence in Augmented Reality

Augmented reality challenges

- Perception
  - Full understanding of the scene
  - Immersive experience
- Hardware requirements
  - Wearable computing
  - Battery
- Content
  - What are useful applications

Ethical considerations

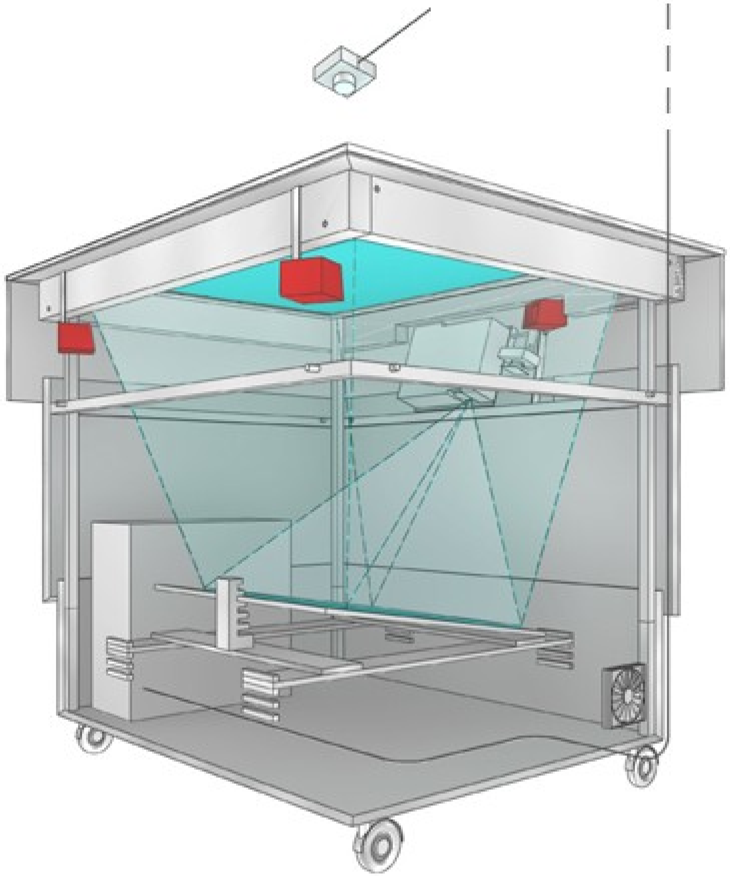
- Ownership
  - Private property
  - Data ownership
  - Anonymity
- Misleading information
  - Who provides the information
  - Deception

Interactive surfaces

- Small form
  - Smart phone
  - Tablets
- Multi-user Tabletops
  - Ergonomics
  - Dedicated purpose

Enabling technologies

- Touch sensor
  - Size, shape
  - Embedding / integration
  - Latency, multi-touch
- Display
  - Front (LCD, projector)
  - Back (projector)
- Software / application

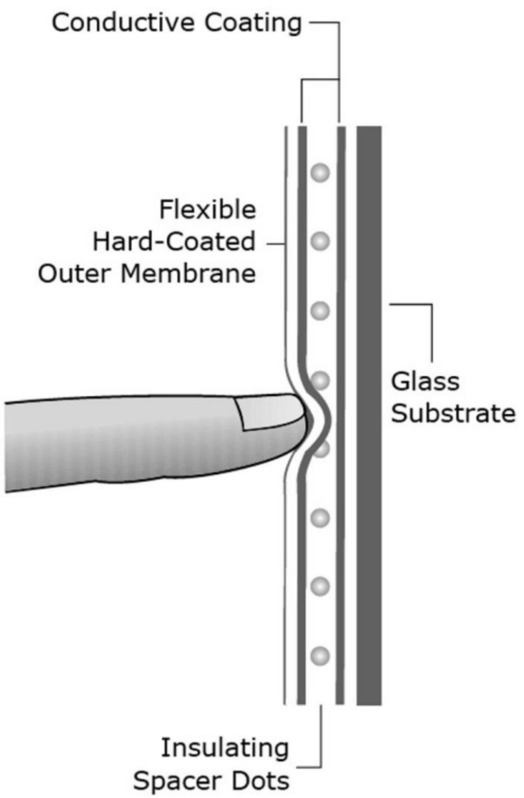


Touch technologies

- Electronics
  - Resistive panels
  - Capacitive panels
- Optical
  - FTIR
  - Diffused illumination
  - Depth camera
  - Laser
- Ultrasonic

Resistive sensors

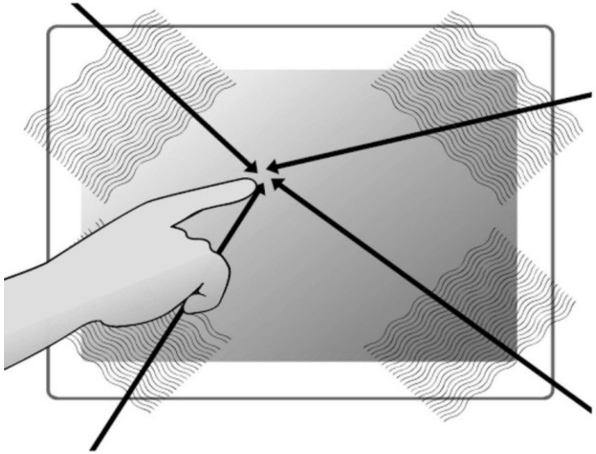
- Two conductive layers with insulation
- Detect position
  - Switching electrodes
  - Horizontal and vertical
- Low-cost, low-power
- Physical pressure
- Reduced display quality

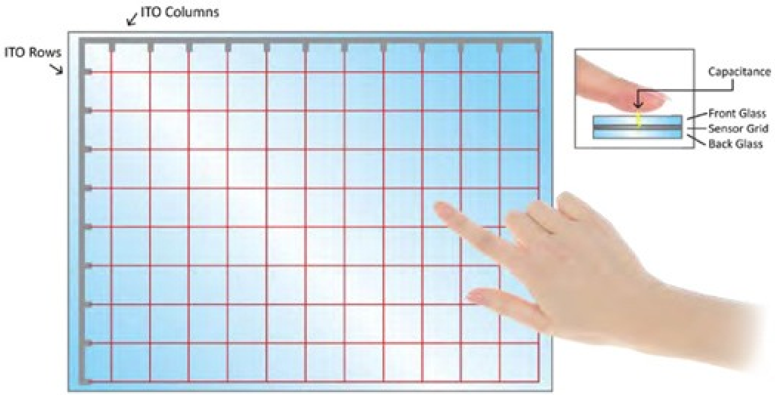


Schoning et al. 2008

Capacitive sensors

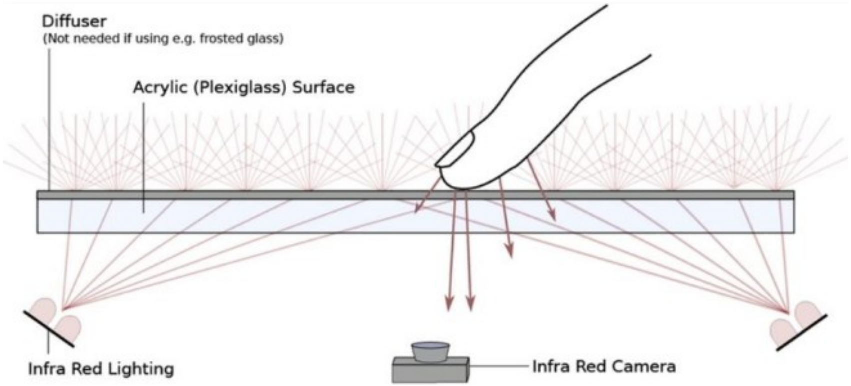
- Surface capacitive
  - Electrodes in corners
  - Hard to detect multi-touch
- Projected capacitive
  - Grid of sensors
  - Can detect multiple inputs
  - Used in mobile phones

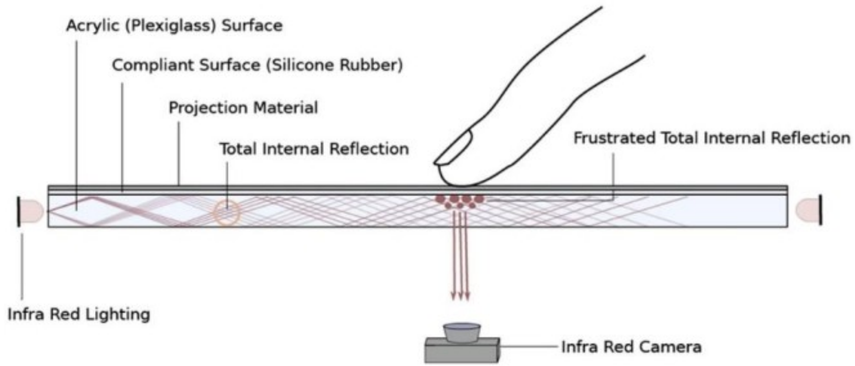




Infrared

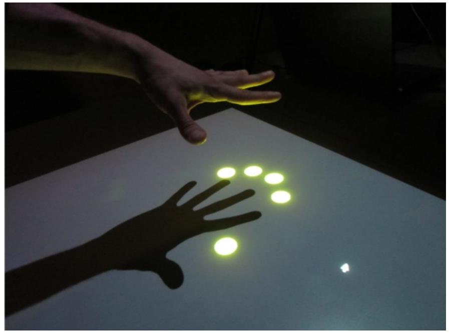
- IR light sources
- Detect reflected light
- Diffused Illumination
  - Wrong detections (hovering)
  - Easier to detect objects
- Frustrated Total Internal Reflection
  - Detects only touching objects

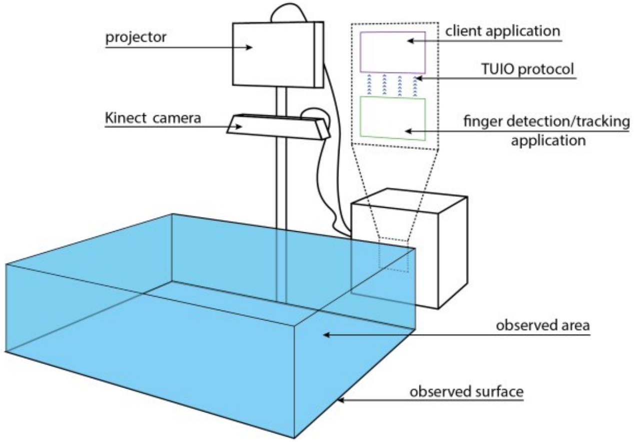




Depth camera

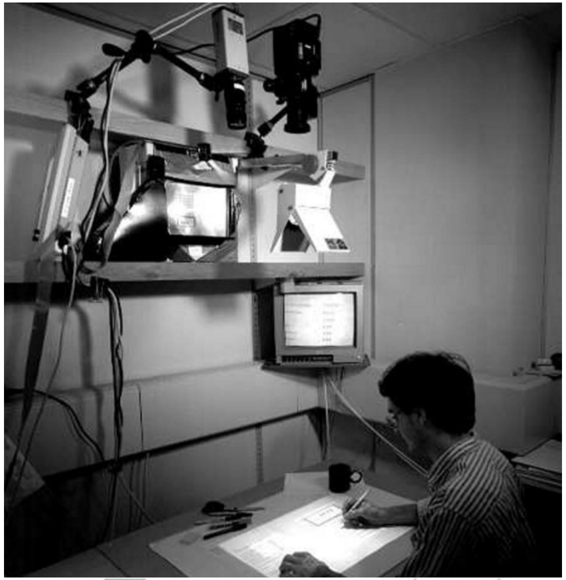
- Use 2.5D depth information
  - IR projector
  - Stereo
- Computationally intensive
  - Finger detection
  - Gesture recognition
  - Object recognition

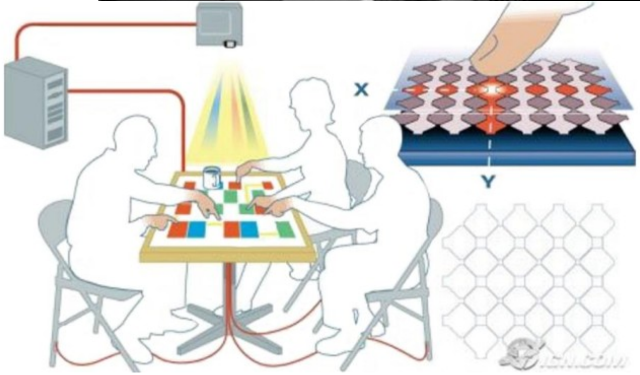




Tabletops - early prototypes

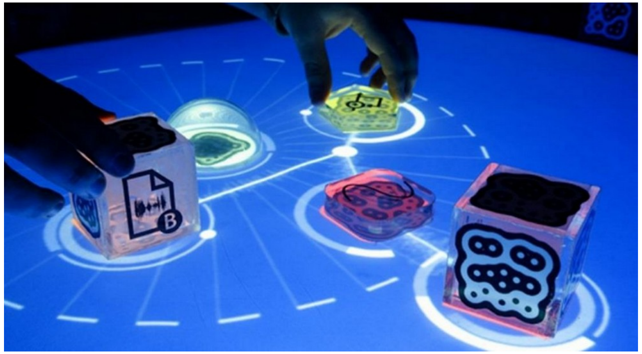
- DigitialDesk (Xerox)
  - Augmented desk prototype
- DiamonTouch (Mitsubishi)
  - Capacitive surface
  - Touch association





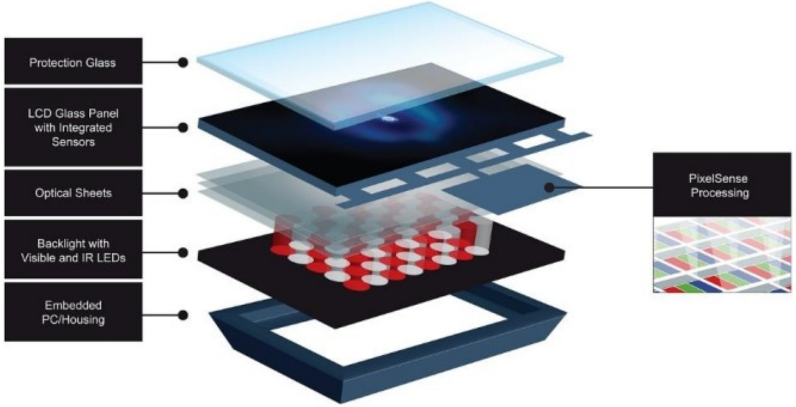
Tabletops - DIY

- Based on IR light
- Reactable (2007)
    - Diffused illumination
    - Fiducial markers
- Jeff Han (2008)
    - Popularization of FTIR



Tabletops - commercial

- Microsoft PixelSense
  - Embedded IR sensors
  - Samsung SUR40
  - Surface Studio
- HP Sprout
  - 3D scanning
  - Blended reality



Tabletop challenges

- Ergonomic issues
  - Neck muscle strain or back problems (gorilla arm)
  - Surface size and position
- Usability issues
  - Visibility and reachability in multi-user scenarios
  - Use-cases, added value