Interactive Environments

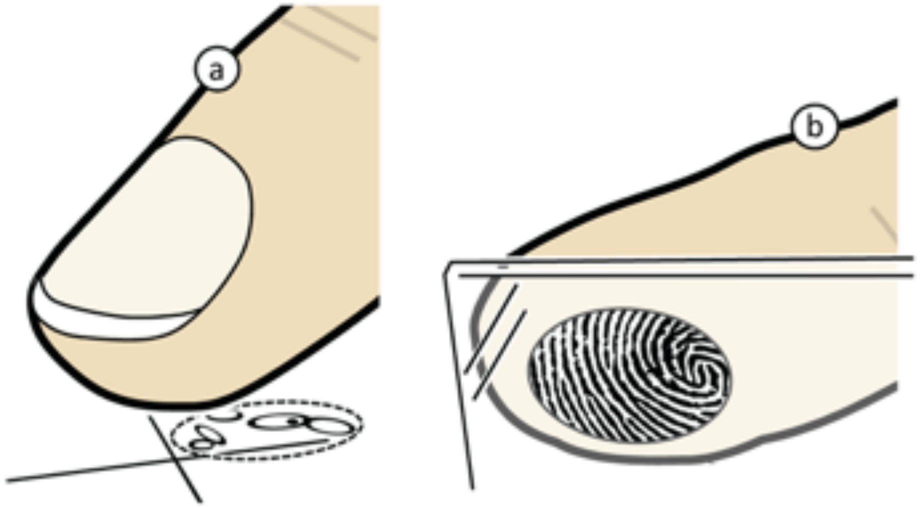
context and task

theory

interaction techniques

in/output technologies

Christmas lectures



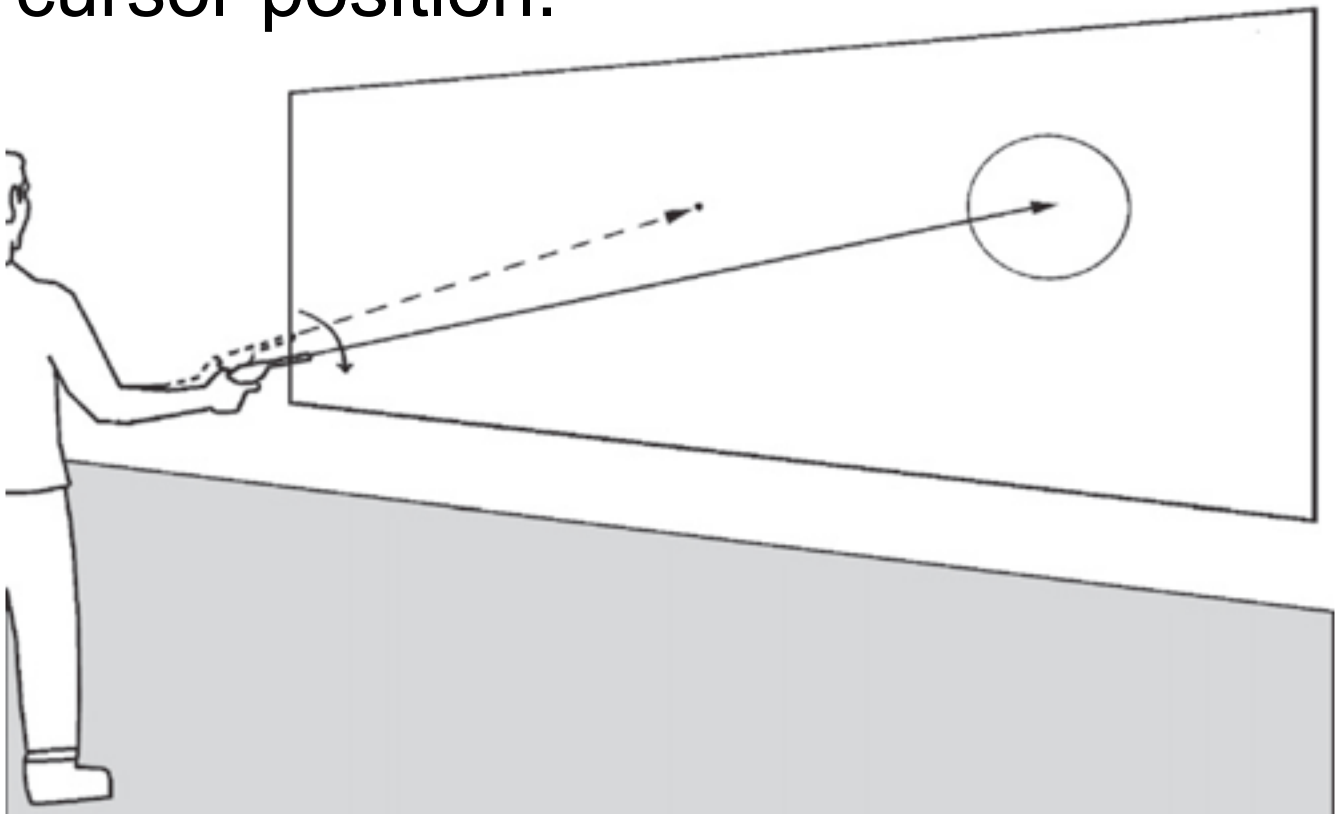
- 17.12. 10-12h MMI2: guest lecture by Christian Holz http://www.christianholz.net
- 16.12. 10-12h, B101, Infoviz: christmas lecture with optical illusions and visual fun
  - bring cookies
  - material won‘t be asked in the exam

Some Theory for Instrumented Env.

- Pointing (...again???, really??? ;-)
  - yes, because we finally move to 3D!
- Crowd Sourcing (huh?!?)
  - yes, because instr. env. are inhabited by **people**
- Spatial Augmented Reality (what???)
  - yes, because that looks like the perfect mixture of virtual and physical worlds...

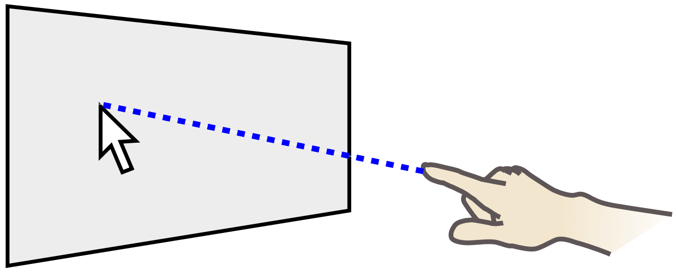
pointing in mid-air

- pointing in desktop or mobile environments
  - models in which users either touch a target directly or translates an input device to cause a proportional translation of a cursor
- Distal pointing makes use of different types of movement (e.g. wrist rotation.)
  - both position and orientation of input device determines the cursor position.



Literature: Kopper R. et al.: A human motor behavior model for distal pointing tasks, International Journal of Human-Computer Studies, Volume 68 Issue 10 (2010)

RayCasting



- place cursor at point where ray emanating from the index finger intersects the screen.
  - problems: jittery cursor movements due to natural hand tremors.
  - solution:
    - use of hand palm or forearm
      - reduces some of jittery with body-parts more proximal in the kinematic chain.
    - use filtering techniques
      - e.g. Kalman filter, two stage mean filter based on angular velocity, etc.

Literature: Vogel, D.: Distant Freehand Pointing and Clicking on Very Large, High Resolution Displays

Repetition

- human motor behavior model for pointing tasks?
  - Fitts' law
  - time to acquire a target is dependent on its size and on the amplitude of movement.
    - MT = a+b * ID
    - a,b, empirically determined constants
    - ID = index of difficulty of the task

Target 1

$$ID = \log_2\left(\frac{A}{W} + 1\right),$$

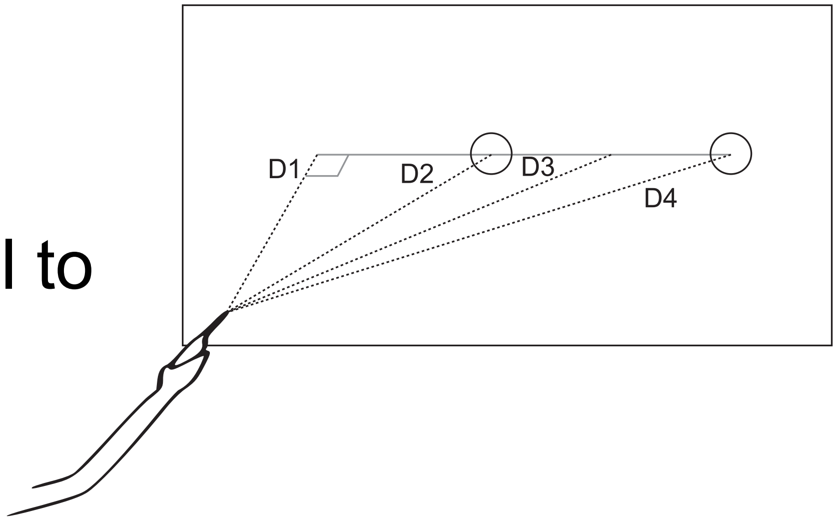
Target 2

Why do you think distal pointing is not well described using Fitts' law. What might be other factors that influence the pointing time?

Literature: Kopper R. et al.: A human motor behavior model for distal pointing tasks, International Journal of Human-Computer Studies, Volume 68 Issue 10 (2010)

Integrating D into Fitts' ID

- reason for W^2
  - decrease in performance as W gets smaller is approximately proportional to
    - decrease in performance as A gets larger
    - decrease in performance as D gets larger
- accounts for the users distance to the display (D)
  - problem: unclear which value should be used for D if distance to initial pointing location different from distance to final pointing location.
  - solution: resolve ambiguity by using angular measurements of target size and movement amplitude

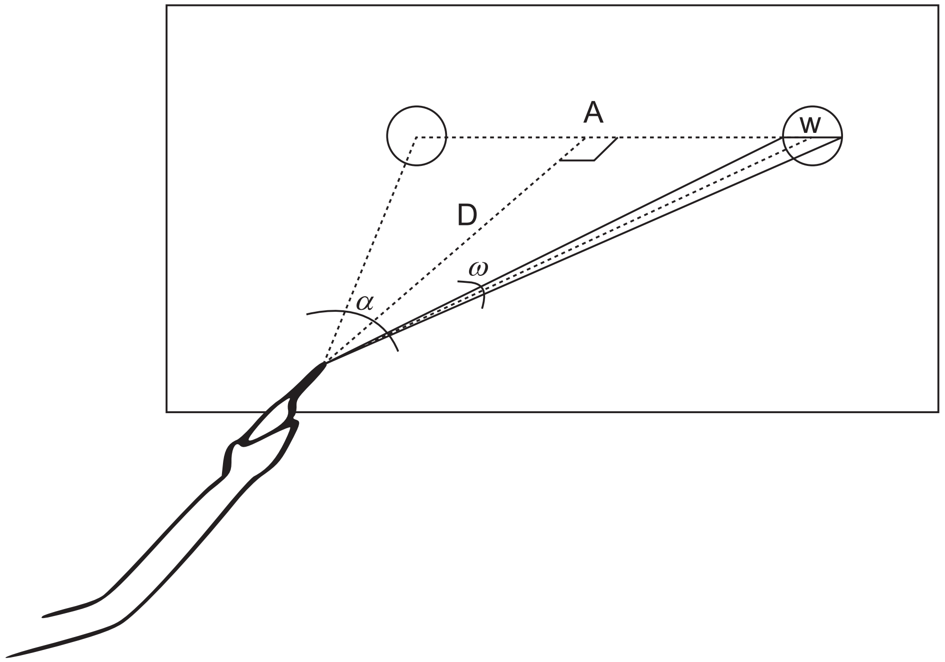


$$ID_{RAW} = \log_2\left(\frac{A \cdot D}{W^2} + 1\right)$$

Literature: Kopper R. et al.: A human motor behavior model for distal pointing tasks, International Journal of Human-Computer Studies, Volume 68 Issue 10 (2010)

Integrating angular measurements for ID

- the amplitude of user movement in a distal pointing task decreases as user moves away from display (arm/wrist rotation is smaller)
  - more appropriate parameters:
    - angular movement (α)
    - angular size of target (ω)



$$\alpha = 2\arctan\left(\frac{0.5A}{D}\right)$$

$$\omega = \arctan\left(\frac{0.5(A+W)}{D}\right) - \arctan\left(\frac{0.5(A-W)}{D}\right)$$

$$ID_{ANGULAR} = \log_2\left(\frac{\alpha}{\omega^k} + 1\right)$$

Literature: Kopper R. et al.: A human motor behavior model for distal pointing tasks, International Journal of Human-Computer Studies, Volume 68 Issue 10 (2010)

$$ID_{ANGULAR} = \log_2\left(\frac{\alpha}{\omega^k} + 1\right)$$

- k is a constant power factor determining the relative weights of ω and α.
  - not always a linear relationship
    - pointing consists of two phases:
      - ballistic phase: pointer moves very rapidly to point
      - correction phase: fine-grained adjustments to acquire target.
    - natural hand tremor
    - Heisenberg effect: unintentional movement of cursor when a button is pressed

Literature: Kopper R. et al.: A human motor behavior model for distal pointing tasks, International Journal of Human-Computer Studies, Volume 68 Issue 10 (2010)

Testing various possibilities for ID

- Regression analysis ID vs. ID_{raw} vs. $ID_{angular}$:
  - find the best model of human motor behavior
- ID: $R^2 = 0.686$
  - 30% of data points cannot be explained by the model.
  - take the users' distance to display into account!

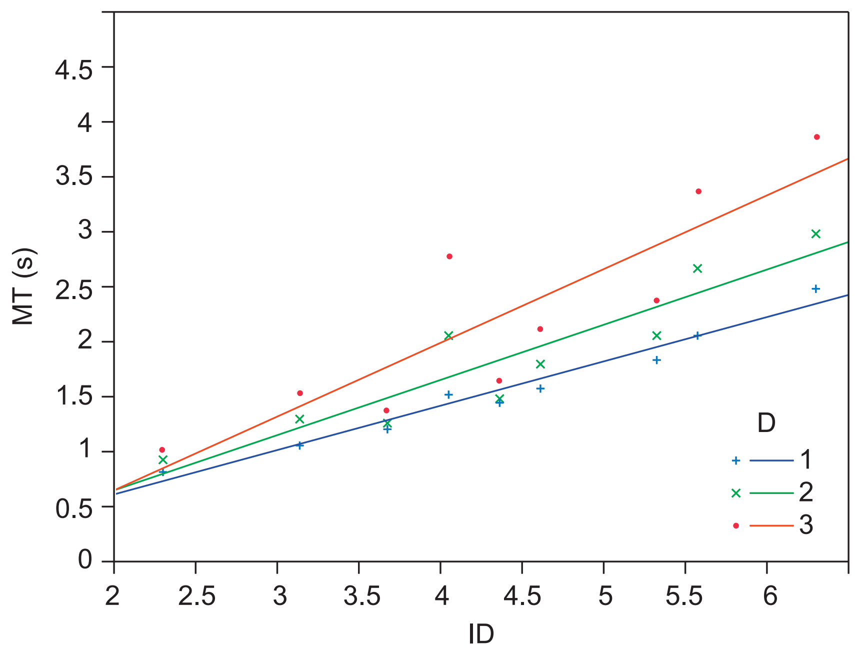


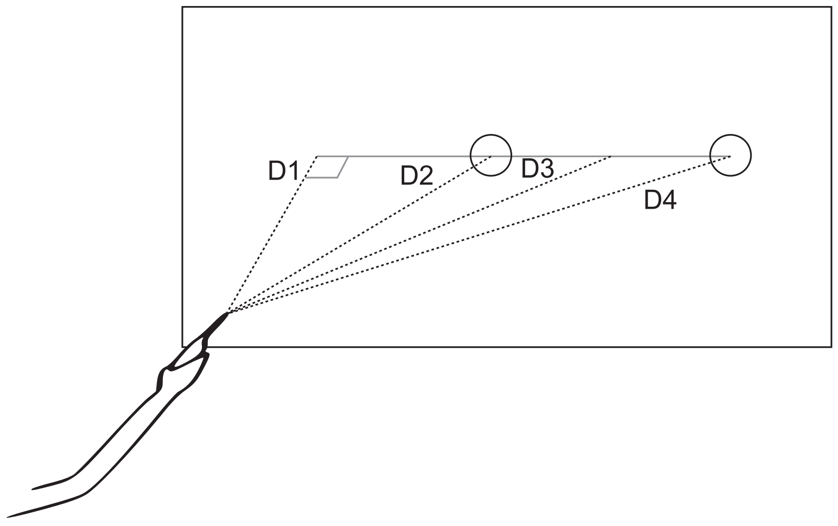
Table 1
Fit of Fitts' law for each distance to the display.

D (m)	*a*	*b*	RMS	R^2
1	−0.204	0.402	0.106	0.963
2	−0.362	0.502	0.267	0.864
3	−0.707	0.672	0.484	0.776

Literature: Kopper R. et al.: A human motor behavior model for distal pointing tasks, International Journal of Human-Computer Studies, Volume 68 Issue 10 (2010)

Testing various possibilities for ID

- Regression analysis ID vs. ID_{raw} vs. $ID_{angular}$:
  - find the best model of human motor behavior
- ID_{raw}: $R^2 = 0.928$
- users stood in the center of movement
  - less generic model
  - in the experimental setup people stood in the center of movement.

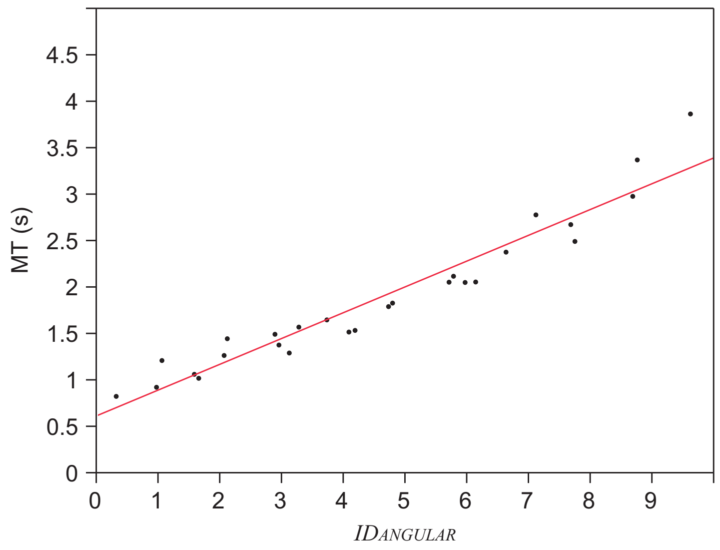


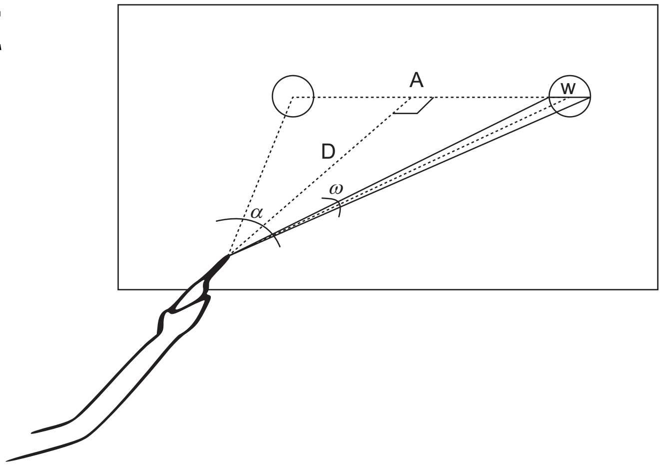
$$ID_{RAW} = \log_2\left(\frac{A \cdot D}{W^2} + 1\right)$$

Literature: Kopper R. et al.: A human motor behavior model for distal pointing tasks, International Journal of Human-Computer Studies, Volume 68 Issue 10 (2010)

Testing various possibilities for ID

- Regression analysis ID vs. ID_{raw} vs. $ID_{angular}$:
  - find the best model of human motor behavior
- $ID_{angular}$: $R^2 = 0.929$ (k=3)
- more generic and expressive
- outliers for high index of difficulty
  - as angular width gets extremely small, a linear increase in acquisition time is not adequate
    - hand tremor and Heisenberg effect





$$ID_{ANGULAR} = \log_2\left(\frac{\alpha}{\omega^k} + 1\right)$$

Literature: Kopper R. et al.: A human motor behavior model for distal pointing tasks, International Journal of Human-Computer Studies, Volume 68 Issue 10 (2010)

Proposing an improved model

- take into account imprecision in two dimensions

$$ID_{DP} = \left[\log_2\left(\frac{\alpha}{\omega^k} + 1\right)\right]^2$$

  - to avoid requiring two parameters to denote the size of target assume dimension of target parallel to direction of movement.
- ID_{DP}: $R^2 = 0.961$ (k=3)

Literature: Kopper R. et al.: A human motor behavior model for distal pointing tasks, International Journal of Human-Computer Studies, Volume 68 Issue 10 (2010)

finally...

- the predicted model of performance for distal pointing under their experimental condition and their input device

$$MT = \underline{1.091} + \underline{0.028} ID_{DP}$$

Literature: Kopper R. et al.: A human motor behavior model for distal pointing tasks, International Journal of Human-Computer Studies, Volume 68 Issue 10 (2010)

Design Guidelines

- angular measurements of target size and movement amplitude are the critical factors in distal pointing performance.
  - distance of the user from the target is significant.
  - targets that might be large when standing near the display might be hard to acquire when standing in a distance.
    - UI could dynamically adapt to user's distance
- angular target size has more influence on pointing difficulty of distal pointing tasks than angular amplitude.
  - increase target size (limited screen space, aesthetics considerations)
  - increase effective target size without increasing the scale of entire UI

Literature: Kopper R. et al.: A human motor behavior model for distal pointing tasks, International Journal of Human-Computer Studies, Volume 68 Issue 10 (2010)

Hybrid pointing techniques

- Absolute and Relative Mapping (ARM) a.k.a dual-mode pointing techniques
  - manual control of the CD-ratio allowing users to increase the effective angular width of targets as needed.
  - ARM uses absolute ray-casting technique as default (cursor appears at intersection of ray with the screen)
  - when pressing a button, users temporarily enter a “precision mode” (Quasimode) with a 10:1 CD-ratio
    - increases the effective angular width of nearby targets by a factor of 10

Literature: Kopper R. et al.: A human motor behavior model for distal pointing tasks, International Journal of Human-Computer Studies, Volume 68 Issue 10 (2010)

Hybrid pointing techniques

- Explicit mode switch: Dual-mode target acquisition techniques
  - Interactions using head tracking, gaze-tracking
    - object selection is often preceded by visual search for the target.
- Implicit mode switch : Adaptive Pointing
  - adapt mode switch dynamically to e.g. cursor speed

Crowdsourcing

- crowdsourcing paradigm: tasks are distributed to and completed by networked people.
  - company's production cost can be greatly reduced

watch:
https://www.youtube.com/watch?v=-Ht4qiDRZE8 (15min)
https://www.youtube.com/watch?v=tx082gDwGcM (50min.)

- history:
  - 2003: Luis von Ahn et al. pioneered concept of 'human computation', use human abilities for tasks which are difficult for computers.
  - 2006: Jeff Howe coined the term "crowdsourcing"

Yuen, M.-C. et al.: A Survey of Crowdsourcing systems, IEEE International Conference on Privacy, Security, Risk and Trust, 2011

Do you have example tasks which are hard to do by computers but trivial to humans?

Labeling Images with words

- women
- cooking
- street
- crowded
- hot food...

application: image search, accessibility for visually impaired.

Further example: locating objects in images
application: train computer vision algorithms

Using Humans Cleverly

- The ESP game
  - two strangers play a game over the web.
    - they see a common image
    - their goal is to type the same word as the other person
    - they need to agree on as many images as they can.
    - tabu words: related to the image, but people cannot agree on.
      - come from the game itself.
      - each time an image goes through another game, it results in a new world for the image
      - it's also making the game harder, more fun.

Dealing with “cheating”

- pair up and agree for a word which does not label the image.
- prevention:
  - probabilistic approach: random test images
    - label not corrupt given that subject labeled all test images correctly
  - repetition: store a label after n pairs have agreed on it.

crowdsourcing

- is a distributed problem-solving and business production model.
  - *"an idea of outsourcing a task that is traditionally performed by an employee to a large group of people in the form of an open call"* (Jeff Howe)
- crowdsourcing sites have 2 types of users
  - requesters and workers
  - workers are motivated through rewards, gain of credibility, fun or altruist
- Application areas:
  - voting system
  - information sharing system
  - game system
  - creative system
- e.g. Amazon Mechanical Turk

Voting System

- voting task: select an answer from a number of choices
  - the answer most people picked is considered to be correct.
  - voting tasks can evaluate correctness of voting tasks.
- some examples:
  - geometric reasoning tasks (difficult to reproduce algorithmically)
  - Named entity annotation (identify/categorize textual references to objects in the world)
  - Opinions (subjective)
  - Spam identification: Vipul's Razor anti-spam mechanism use human votes to determine if a given email is spam.

Information Sharing System

- share various types of information among the crowd.
  - monitor noise pollution
  - Wikipedia: online encyclopedias written by users; anyone can contribute.

Game System

- pioneered by Luis Von Ahn et al.
  - games with purpose: produce useful metadata as a by-product.
  - taking advantage of people's desire to be entertained to solve problems
- peekaboom: object location in images
- Squigl system: outlines of objects in images
- Matchin system: rank images based on appeal
- TagATune system: annotation for sounds and music
- CommonConsensus system: commonsense knowledge (reasoning)
- ...

Creative systems

- human creativity cannot be replaced by any advanced technologies
  - e.g. drawing, coding
- Foldit: game allowing players to assist in predicting protein structures
  - important area of biochemistry seeking for cures for diseases
  - taking advantage of human's puzzle-solving intuitions

Creative systems

- art: http://www.thejohnnycashproject.com
  - people contributed with frame images
  - resulting in hundreds of images per frame
  - each time you watch this video you see a unique image composition

Crowdsourcing: Algorithm

- model performance of a crowdsourcing system [1]
  - completion time as a stochastic process
  - statistical method for predicting the expected time for task completion on MTurk
    - found that time-independent variables of posted tasks affect completion time

[1]Wang et al.: Estimating the completion time of crowdsourced tasks using survival analysis models, CSDM 2011

Crowdsourcing: data sets

- crowdsourcing datasets are available for further research:
  - 100,000 images with English labels from ESP [1]
  - TagATune released their dataset as well: sound clips with human annotation [2]
  - Körner and Strohmaier: list of social tagging datasets made available for research [3]

[1] ESP Game dataset: http://server251.theory.cs.cmu.edu/ESPGame100k.tar.gz
[2] Tagatune Dataset website: http://tagatune.org/Magnatagatune.html
[3] C. Körner and M. Strohmaier. A call for social tagging datasets. SIGWEB Newsl., January 2010.

follow work of
- Hrvoje Benko
- Andrew Williams

Spatial augmented reality

- Virtual Reality:
  - *technology that makes diving into a completely synthetic, computer-generated world possible. Senses such as vision, hearing, haptics, smell etc., are controlled by a computer while our actions influence the produced stimuli.* [1]
- Augmented Reality
  - brings virtual elements to a real environment (or live video of real environment) through a display (hand-held, HMD)
- Spatial augmented reality
  - augments real world without using any display.
  - uses digital projectors to display on real world surfaces.

[1] Bimber and Raskar: Spatial augmented reality: Merging real and virtual worlds, AK Peters Ltd, 2005

The Vision

http://inventinginteractive.com/wp-content/uploads/2010/01/avatar_45.jpg

How to achieve Spatial Augmented Reality

- **Projectors and their working principles**
- Using projectors as shader lamps
- Combining two projectors
- Combining many projectors
- Steerable projectors
- Projection on structured surfaces
- Combining it all with today‘s technology

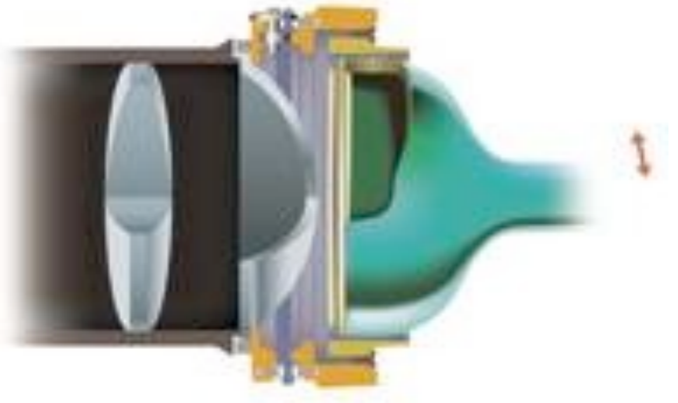
Projectors

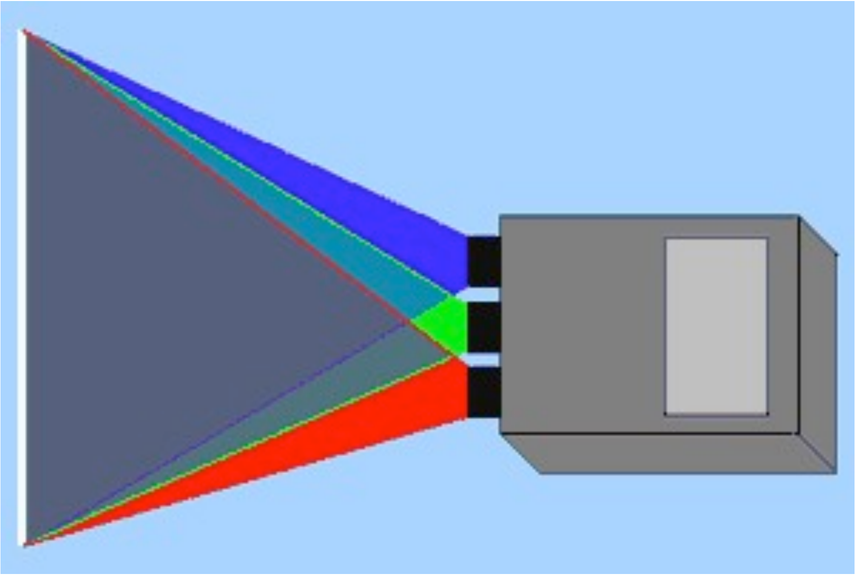
- Key Criteria
  - Resolution
  - Brightness
  - Weight
  - Noise
  - Lens
  - Image correction
  - Projection distance
  - Connections
  - Lamp life time

CRT projector

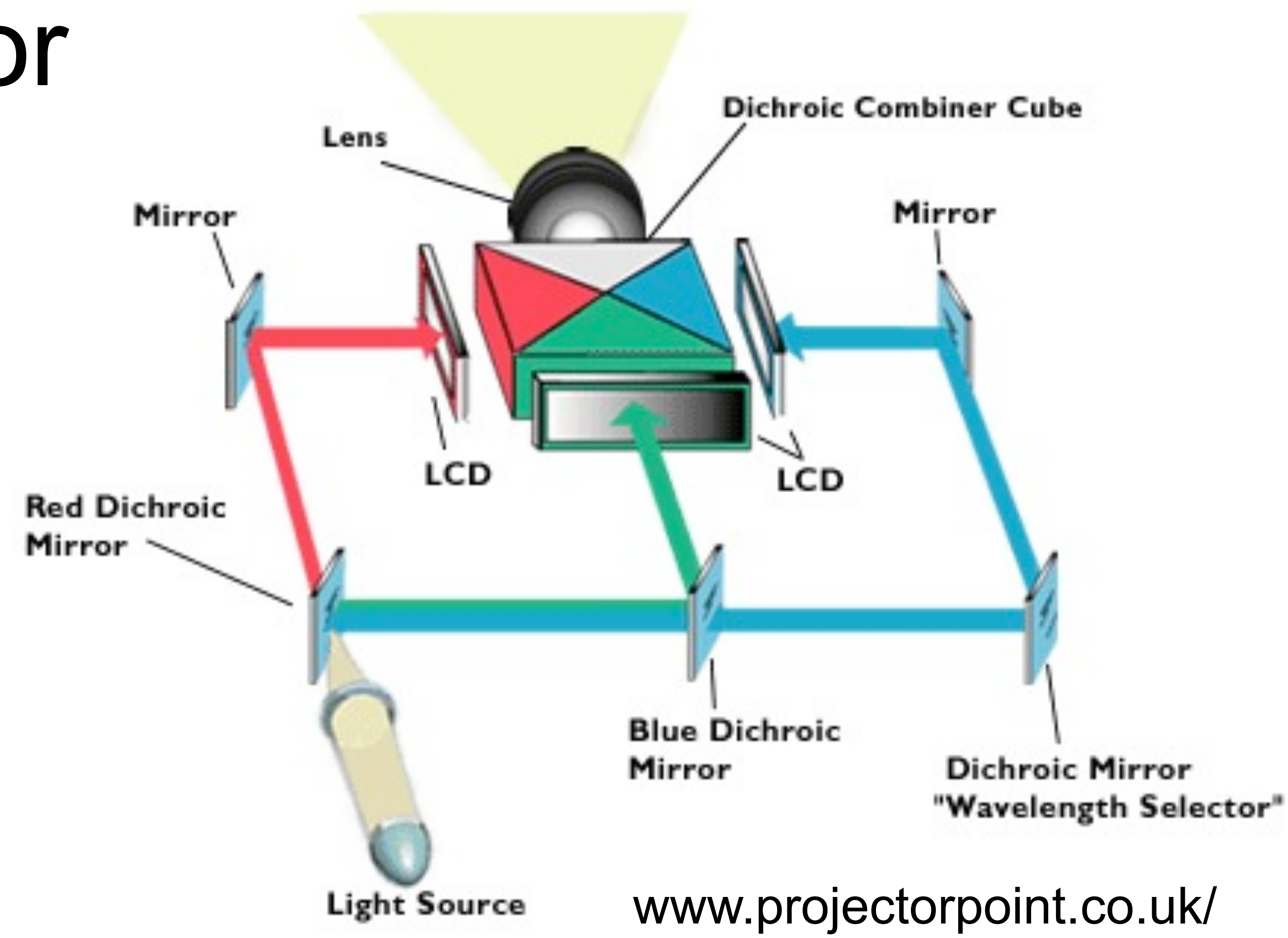
- Use R,G+B CRTs as light sources
- Good black areas
- Low brightness
- Fast
- Need to calibrate convergence!

www.projektoren-datenbank.com/rohre.htm

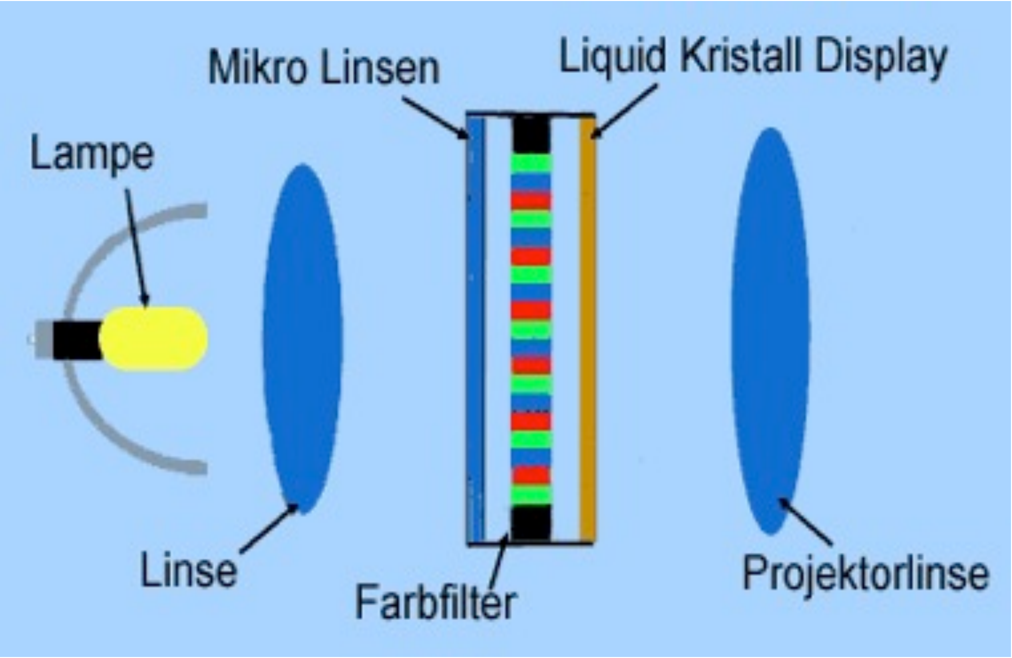




LCD projector

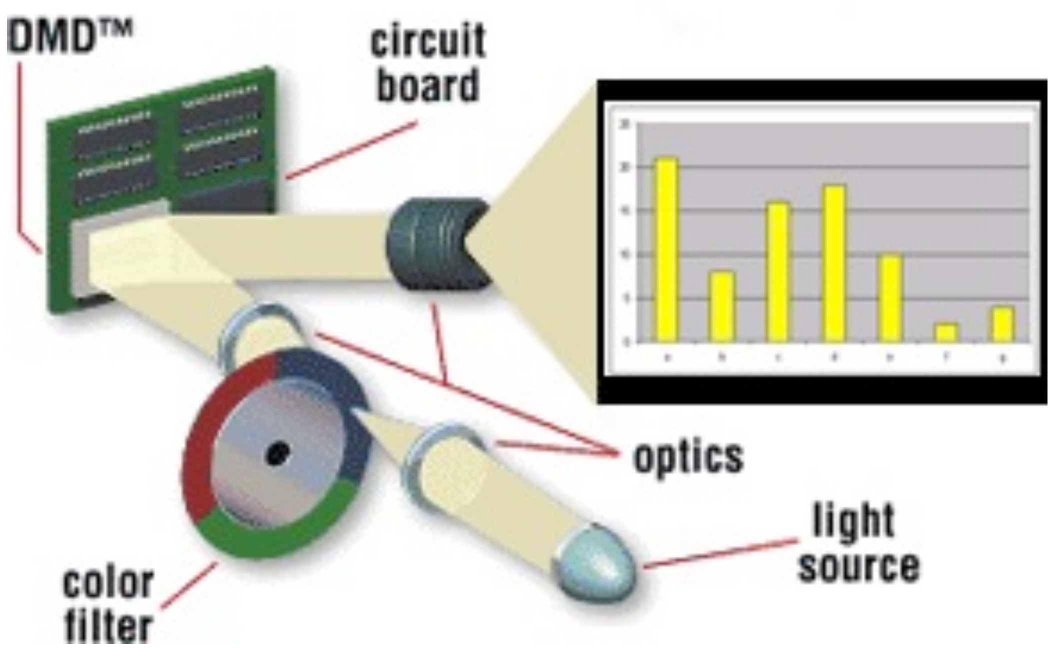


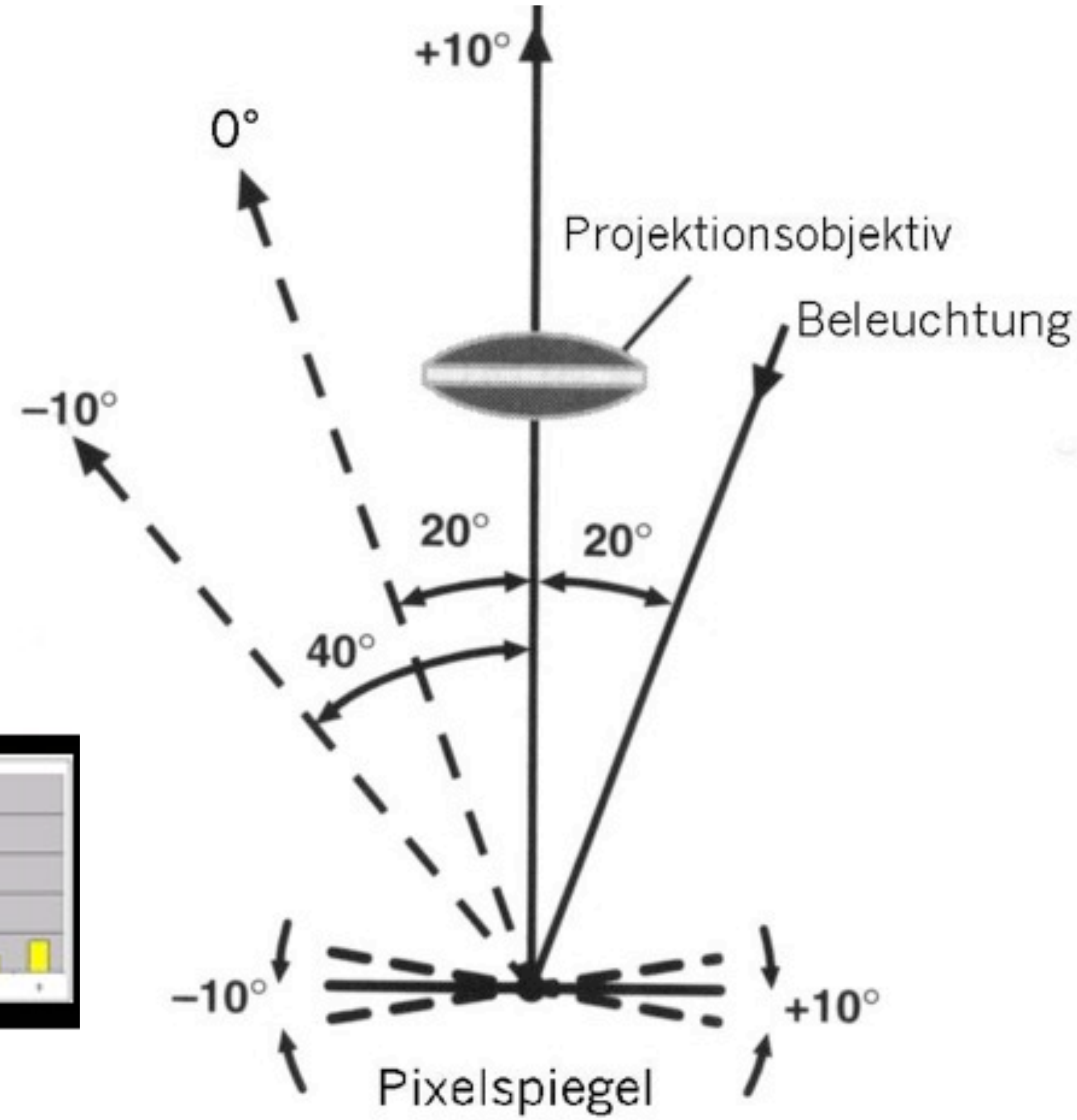
www.projectorpoint.co.uk/projectorLCDvsDLP.htm

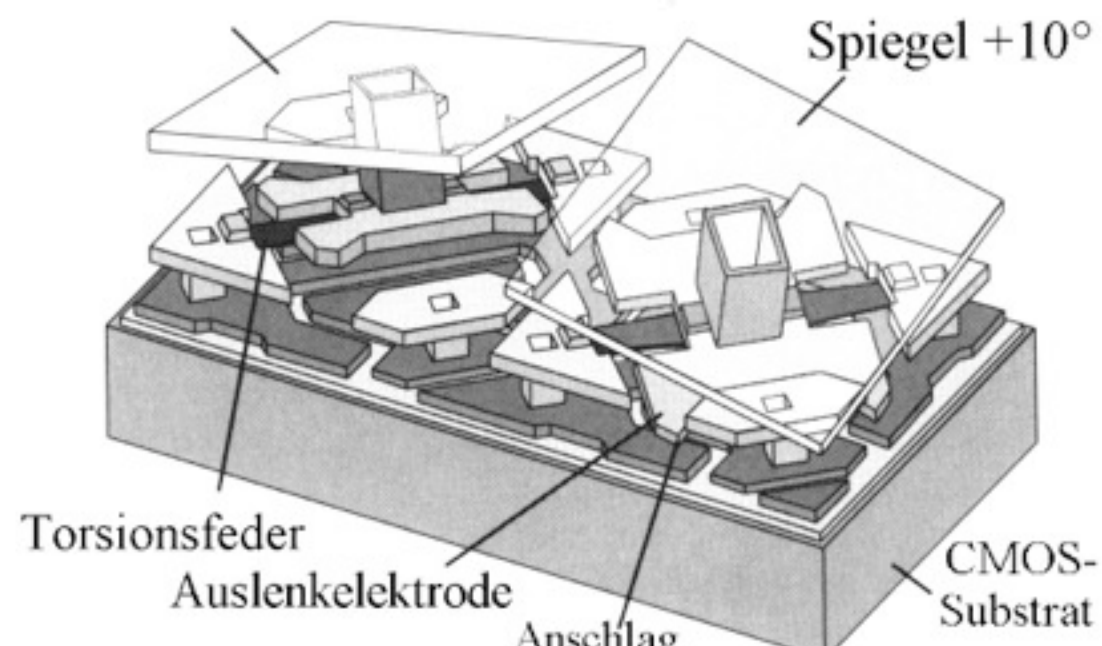


www.projektoren-datenbank.com/lcd.htm

DLP projector

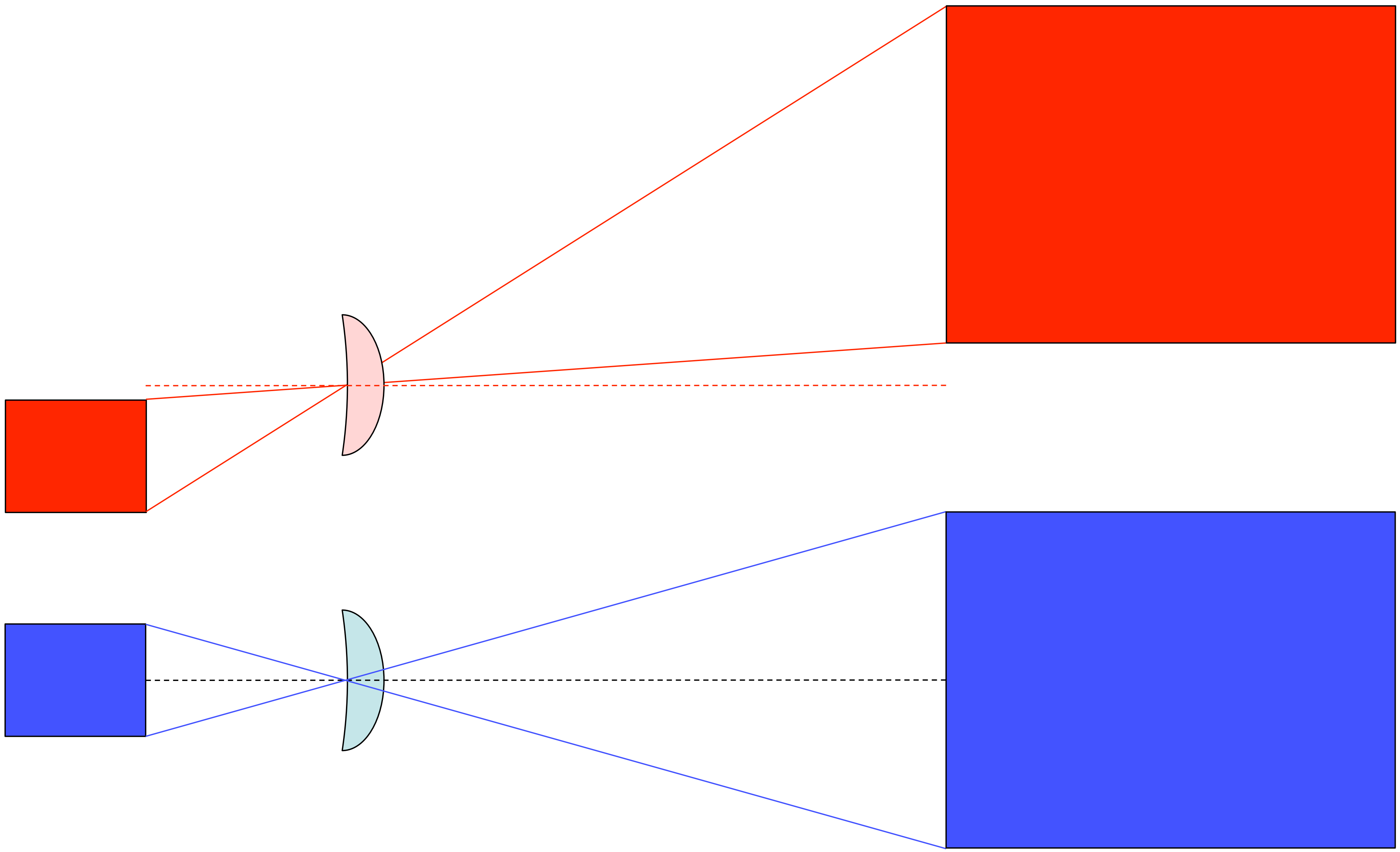






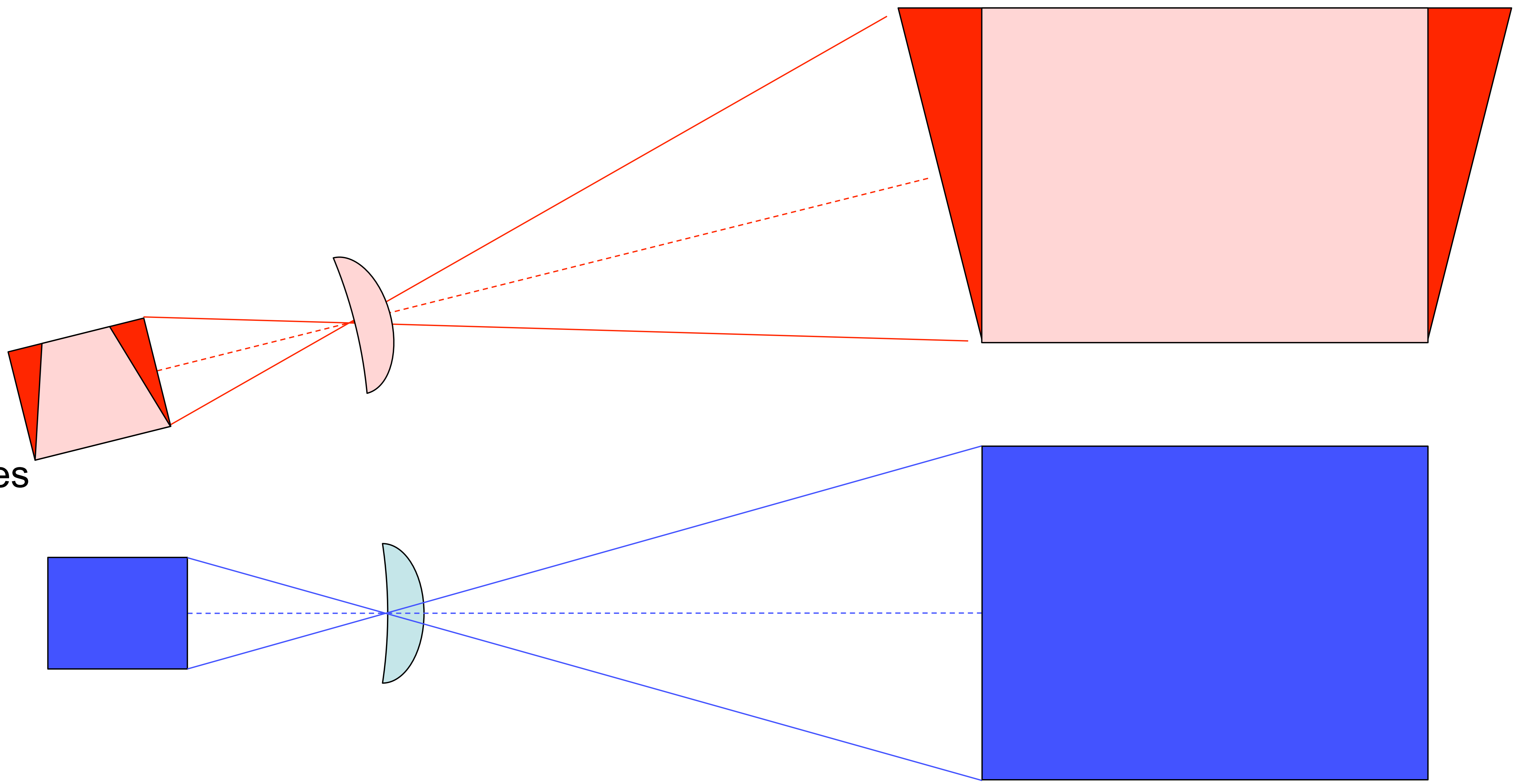
Lens shift

- Optical construction
- No loss of resolution



Keystone correction

- Computed correction
- Loss of resolution!



How to achieve Spatial Augmented Reality

- Projectors and their working principles
- **Using projectors as shader lamps**
- Combining two projectors
- Combining many projectors
- Steerable projectors
- Projection on structured surfaces
- Combining it all with today‘s technology

Shader Lamps: Basic Idea

Rearranging terms in optical path

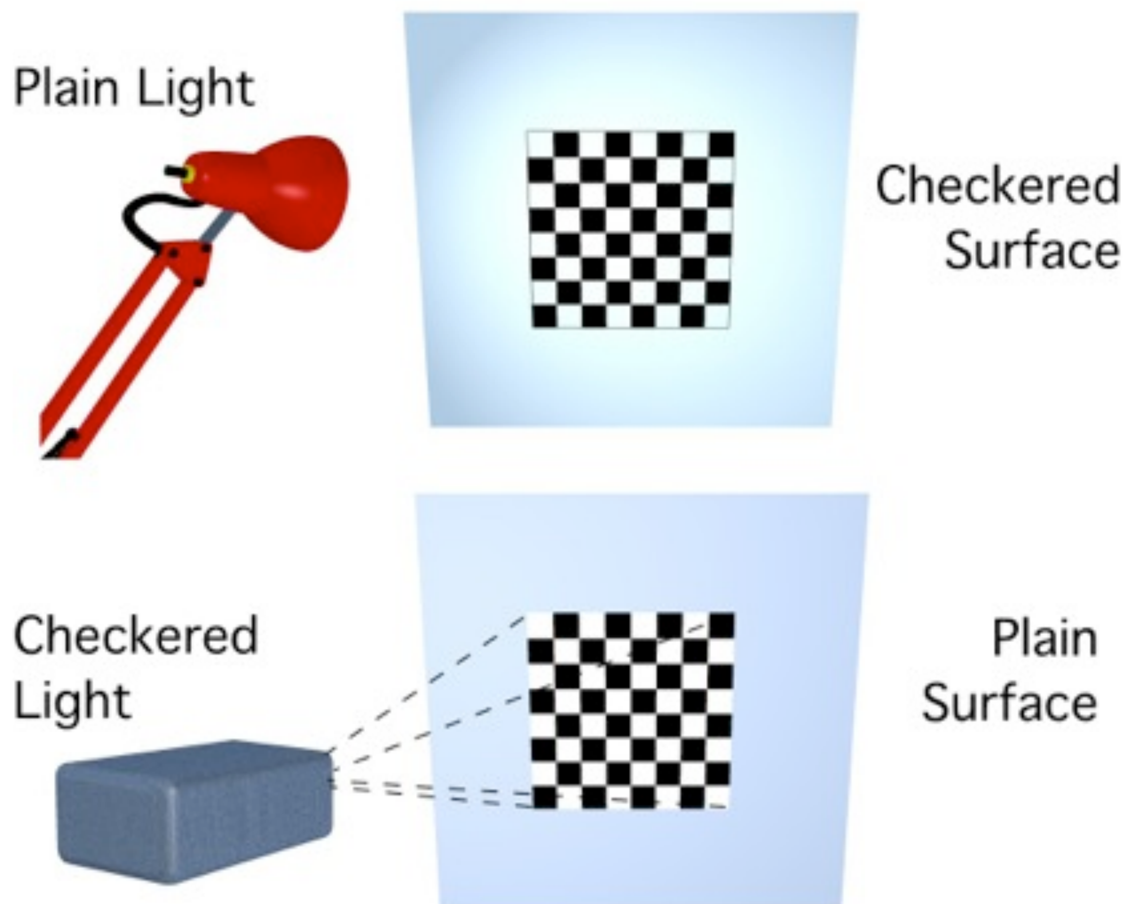
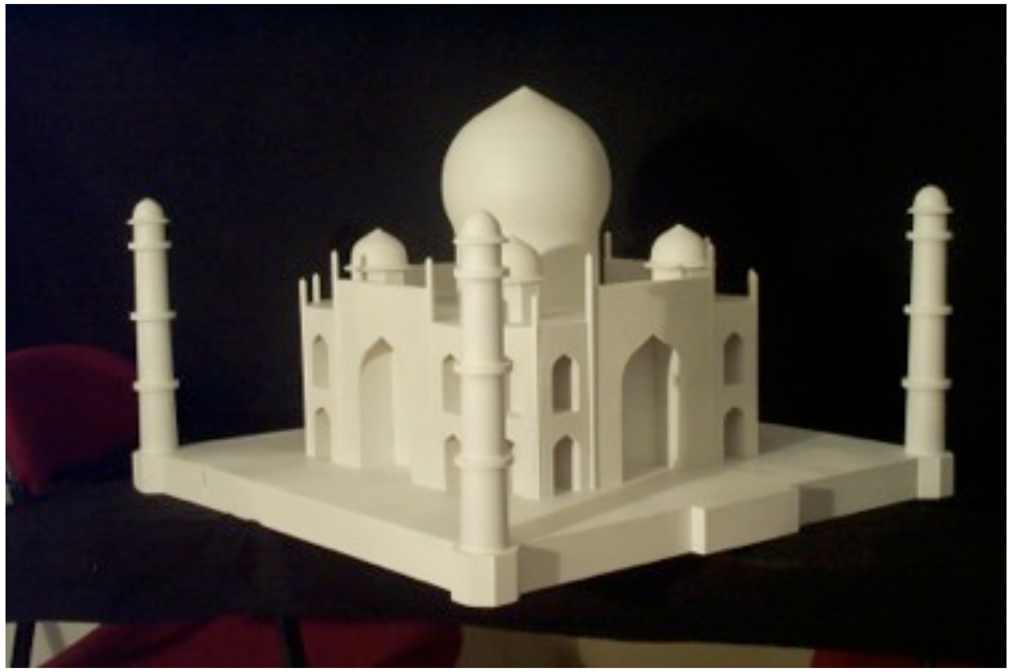
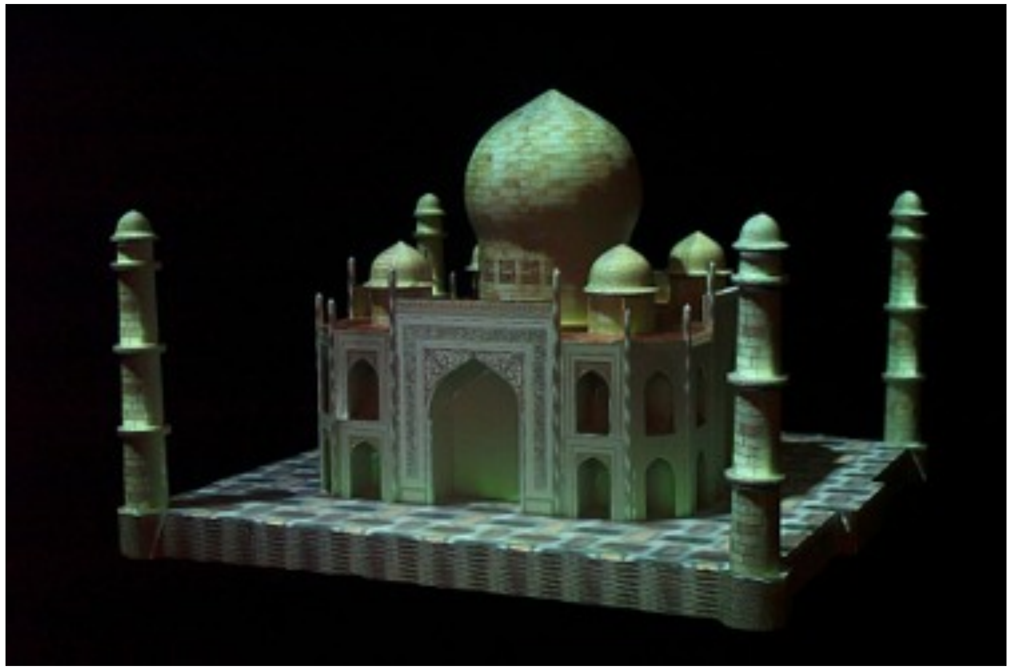


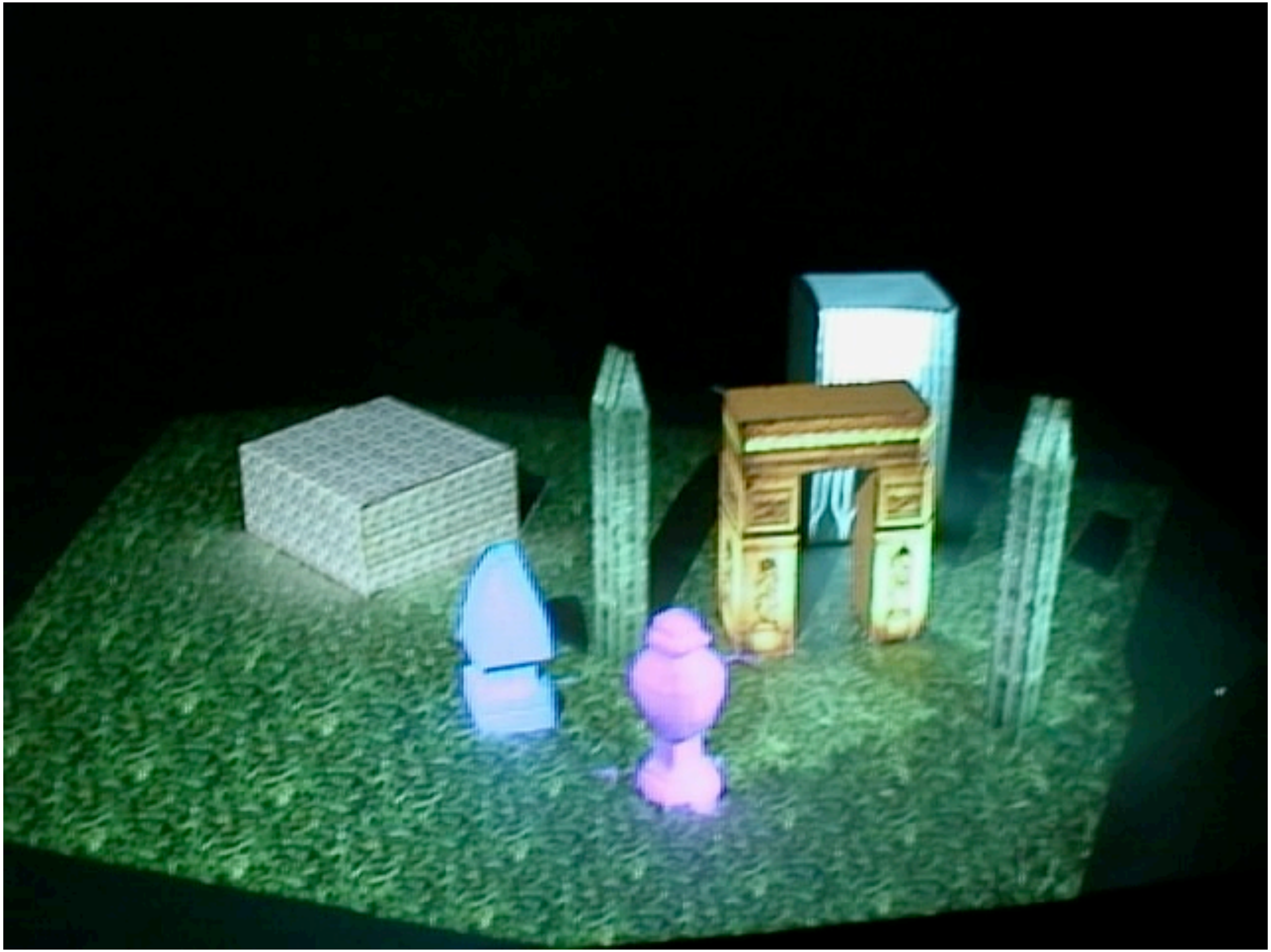
Image based Illumination

- Basic Idea
  - Render images and project on objects
  - Multiple projectors
  - View and object dependent color

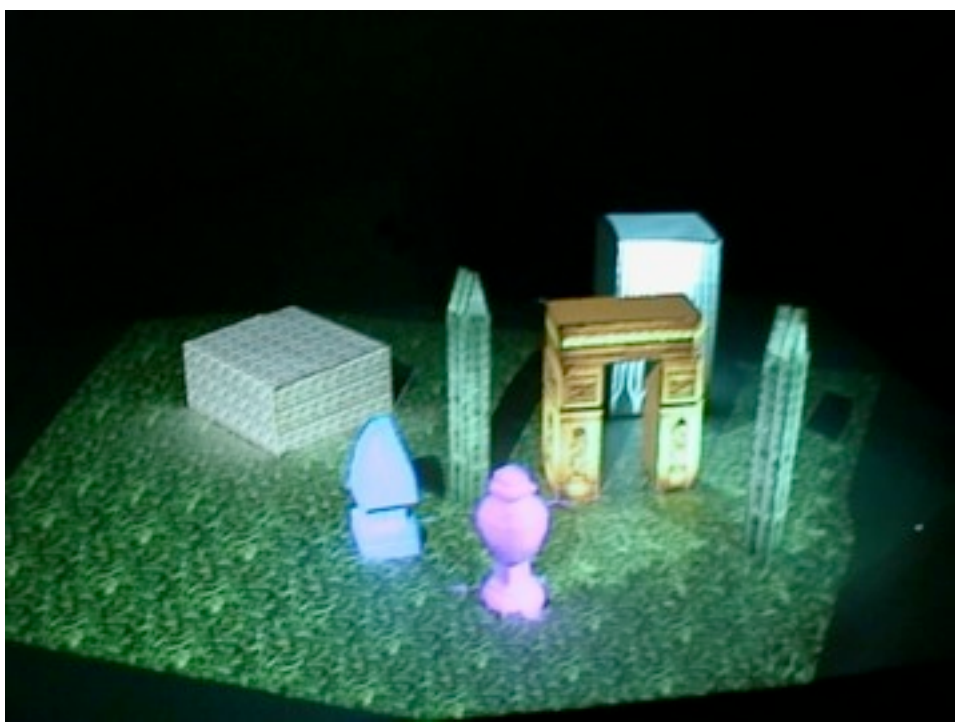




Shaderlamps: Example



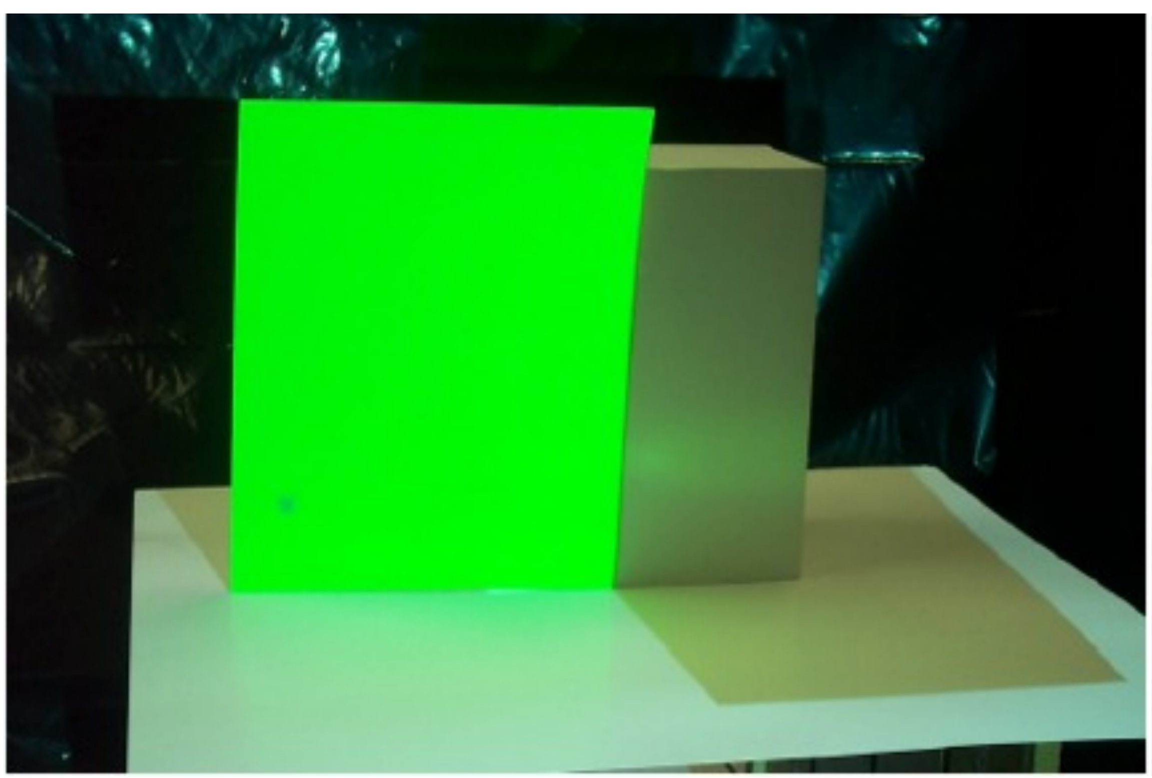
Problem: shadow areas
Solution: two projectors



Every visible surface must be illuminated by at least one lamp (projector)

Radiosity

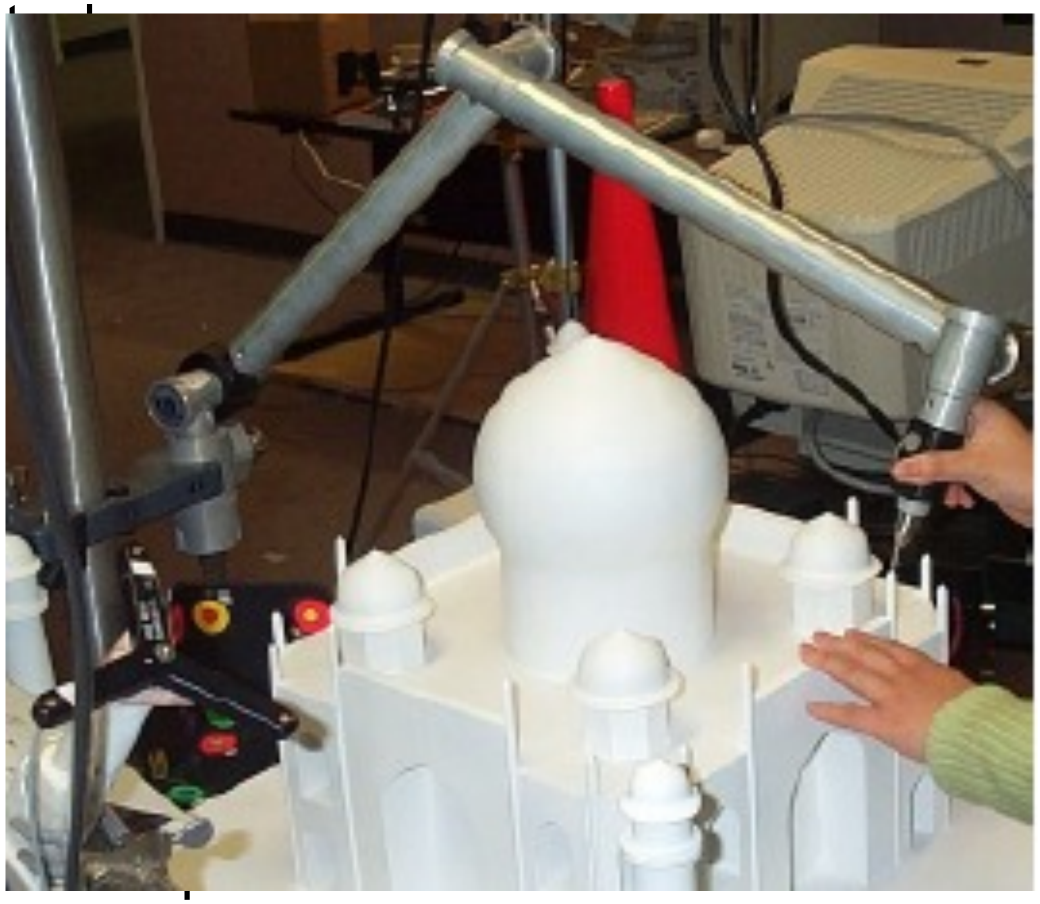
- Objects illuminated by direct and indirect light
- Parts of an object can scatter light onto other parts of object and other objects
- High computational effort to calculate correctly
- Often approximated by „ambient light“
- Comes for free with shaderlamps!

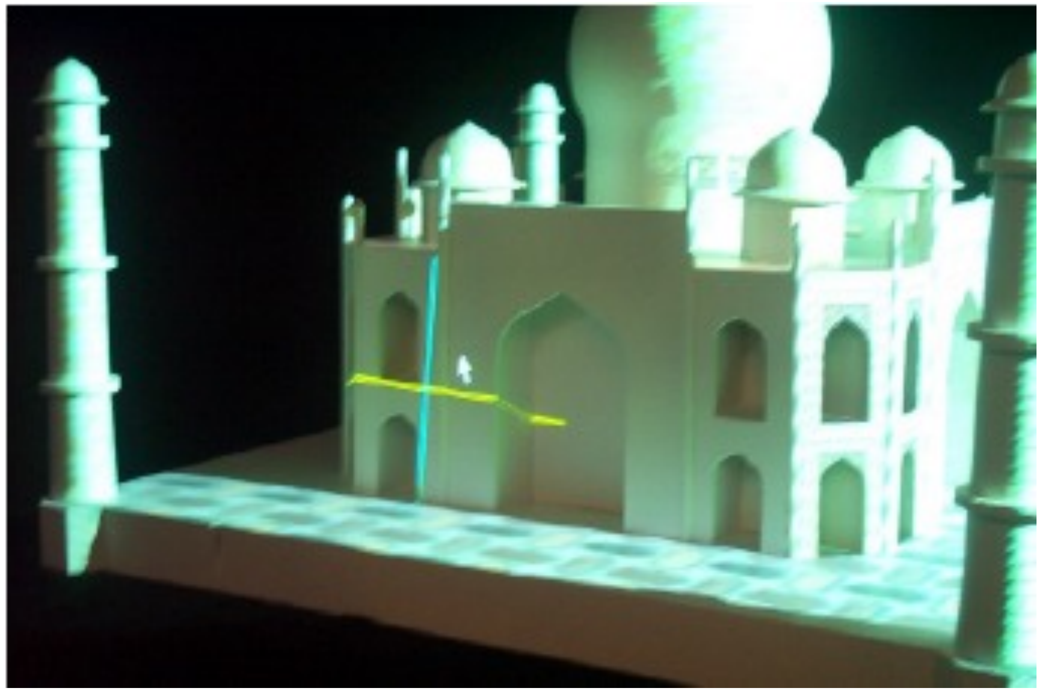


How to achieve Spatial Augmented Reality

- Projectors and their working principles
- Using projectors as shader lamps
- **Combining two projectors**
- Combining many projectors
- Steerable projectors
- Projection on structured surfaces
- Combining it all with today‘s technology

Manual Projector Alignment





- Position projector roughly
- Adapt to geometric relationships between physical objects
- Take fiducials on physical object and find corr. projector pixels
- Compute 3x4 projection matrix
- Decompose into intrinsic & extrinsic projector params

Occlusion and Overlaps

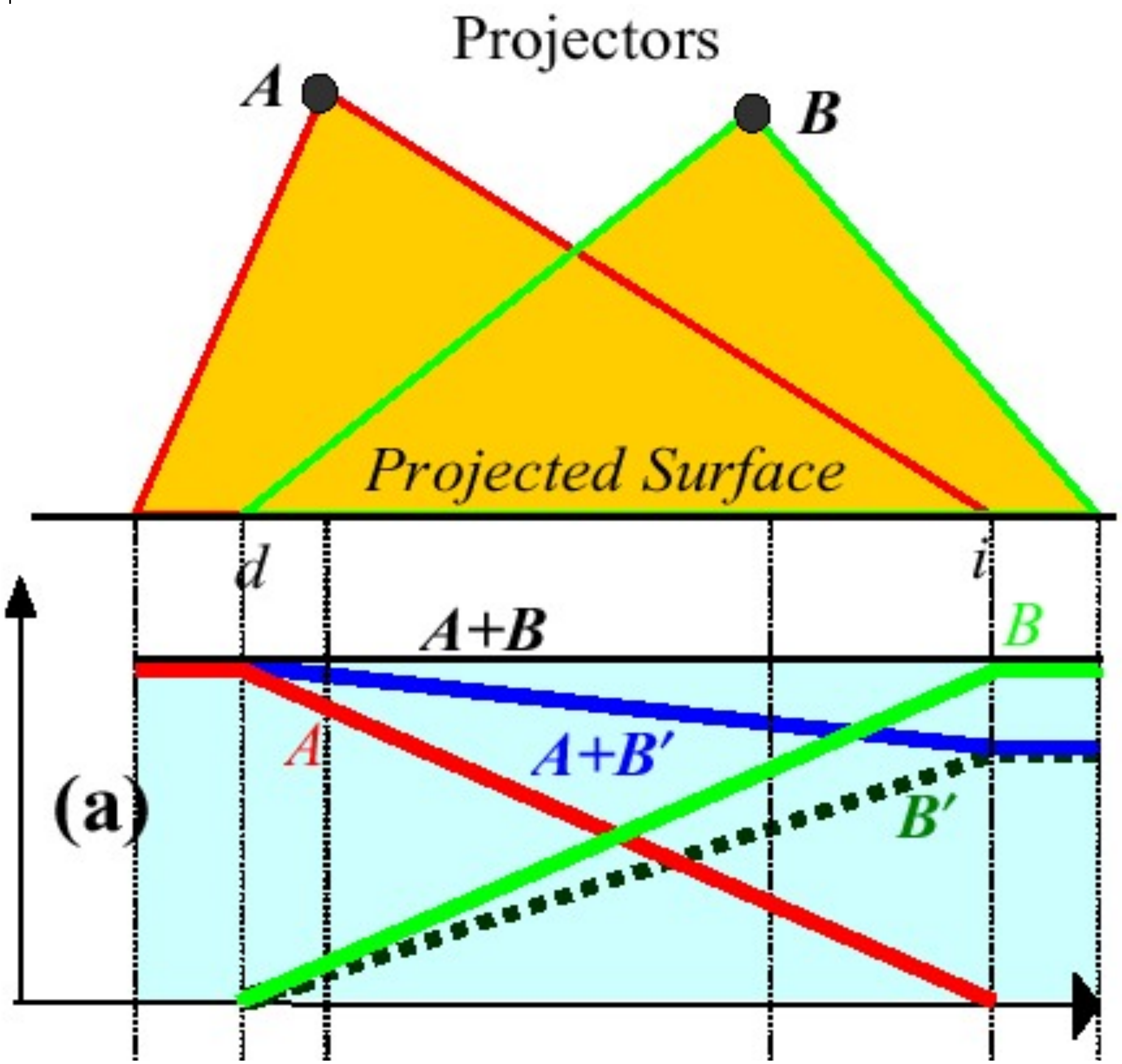
- Several problems:
  - No color equivalence between two projectors (manufacturing & temperature color drift)
  - Minimize sensitivity to small errors in calibration parameters or mechanical variations

- Relatively good solution: Feathering

Feathering

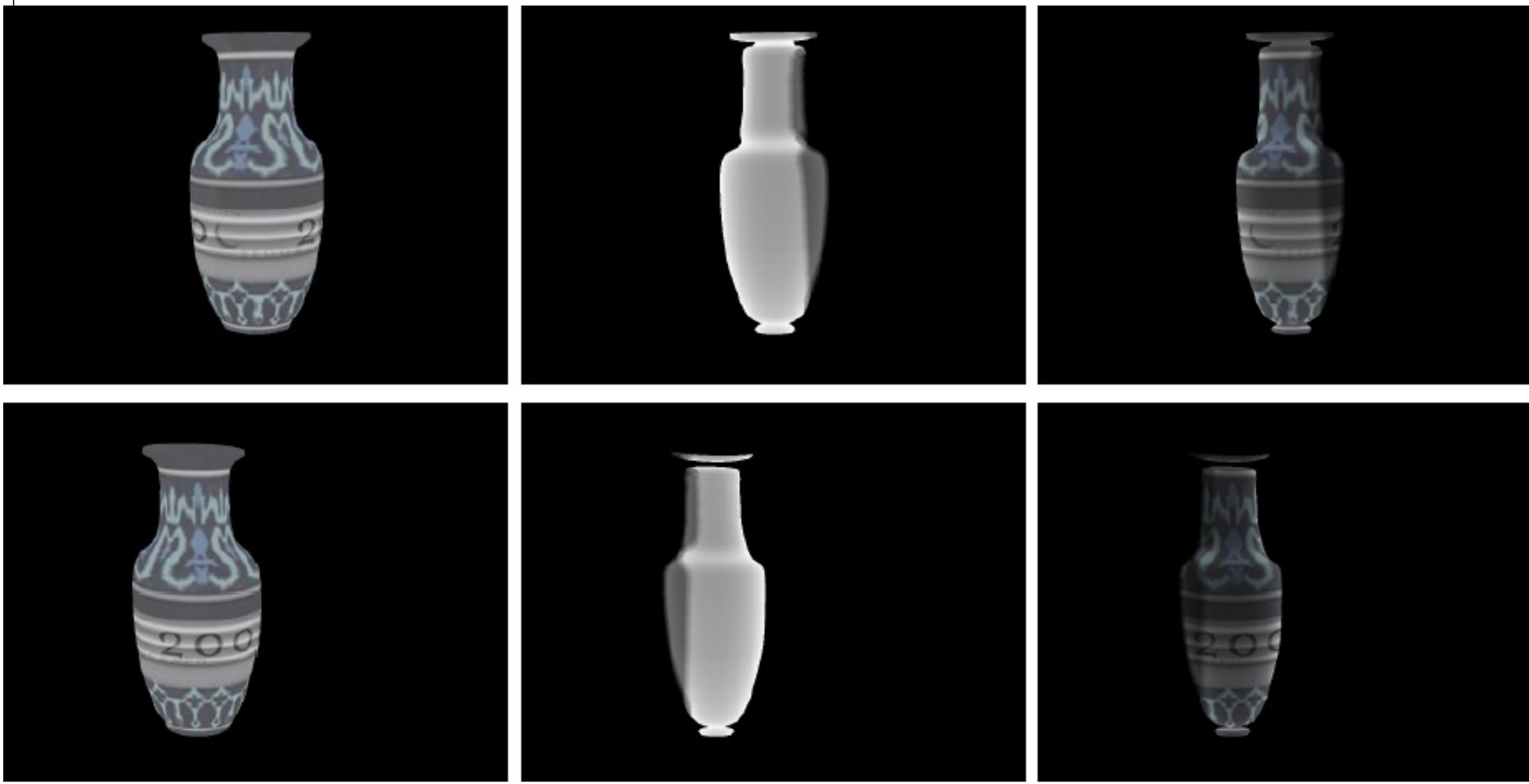
- Normally the overlap region is a well-defined contiguous region
- Intensity of every pixel weighted proportional to Euclidian distance to nearest boundary pixel of image
- Weights in range [0,1] multiplied with intensities in the final image

Feathering



- If both projectors produce the same color, A+B are at maximum and constant over surface
- If not, A+B´ produces a smooth transition

Examples



How to achieve Spatial Augmented Reality

- Projectors and their working principles
- Using projectors as shader lamps
- Combining two projectors
- **Combining many projectors**
- Steerable projectors
- Projection on structured surfaces
- Combining it all with today‘s technology

Luminance Attenuation Map

[Majumder & Stevens, VRST 2002]

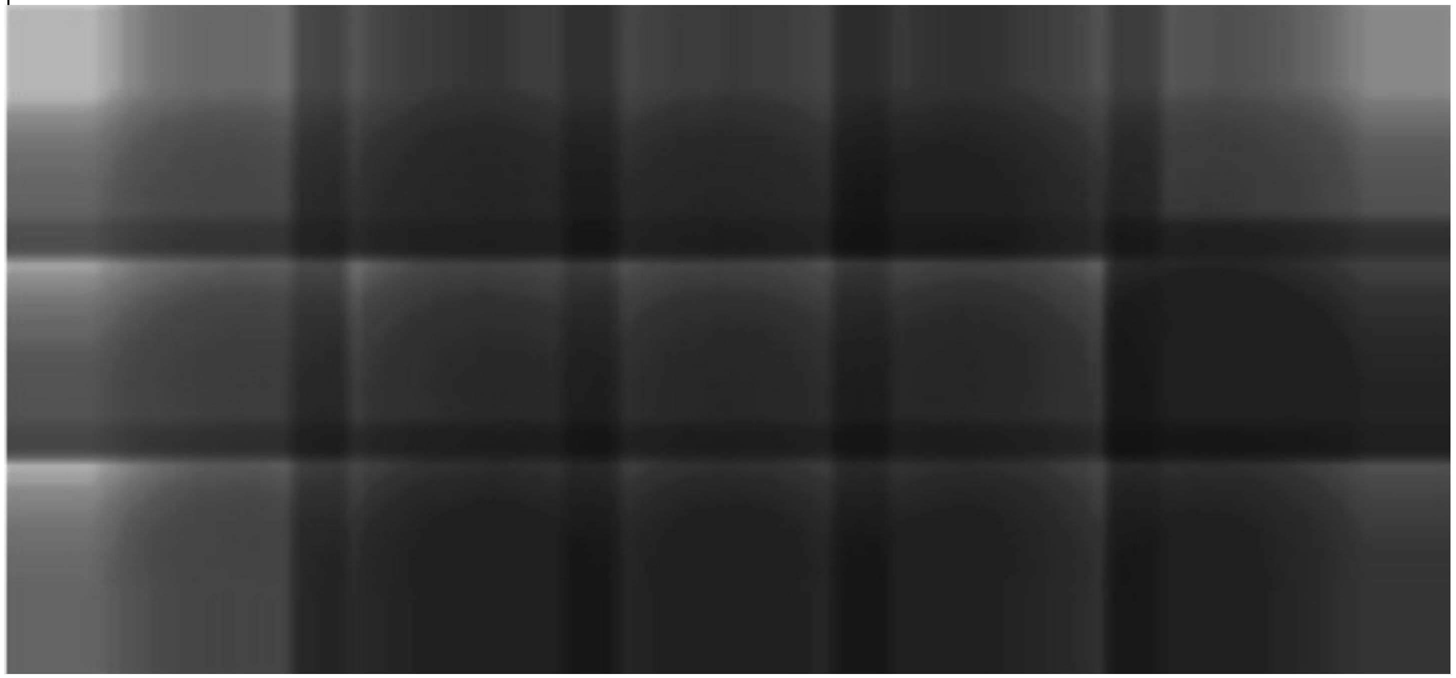
- Large display wall with 5x3 projectors
- Linear ramps (feathering) don‘t work perfectly
- Goal: get rid of the remaining unevenness
- Strategy: don‘t assume, but measure!

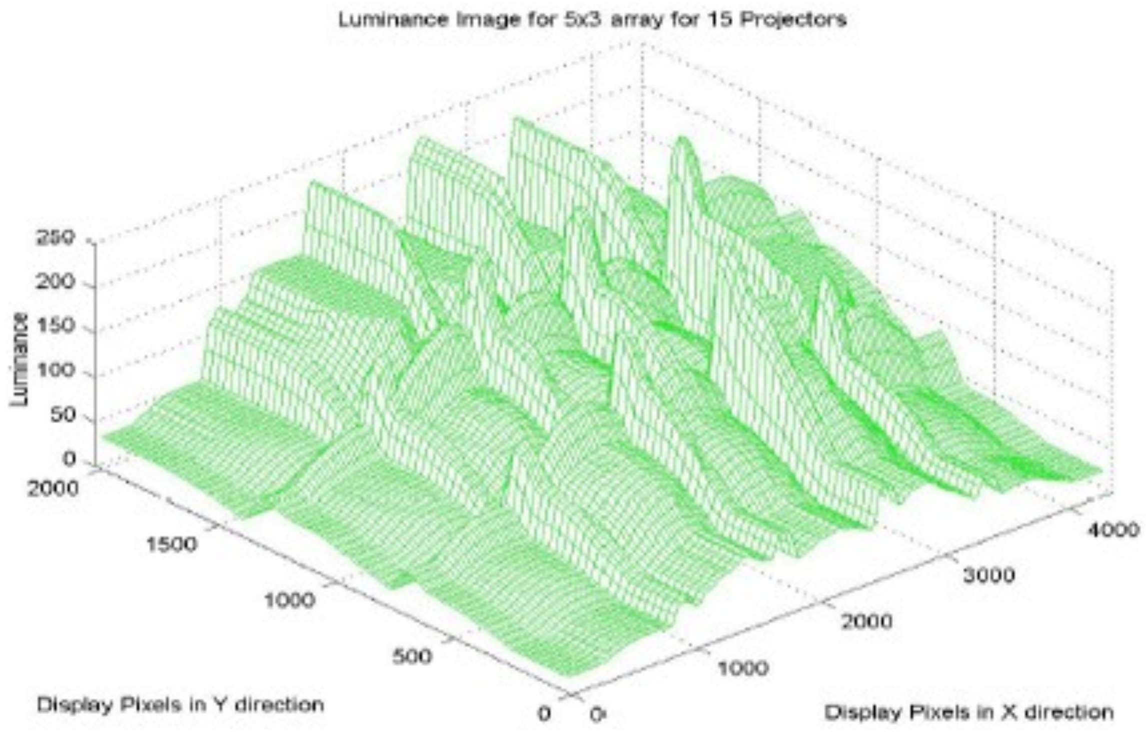
Calibration step

- **Measuring the Luminance Response:** The *luminance response* of any pixel is defined as the variation of luminance with input at that pixel. We measure the luminance response of every pixel of the display with a camera.
- **Finding the Common Achievable Response:** We find the common response that every pixel of the display is capable to achieving. The goal is to achieve this *common achievable response* at every pixel.
- **Generating the Luminance Attenuation Map:** We find a luminance attenuation function that transforms the measured luminance response at every pixel to the common achievable response.

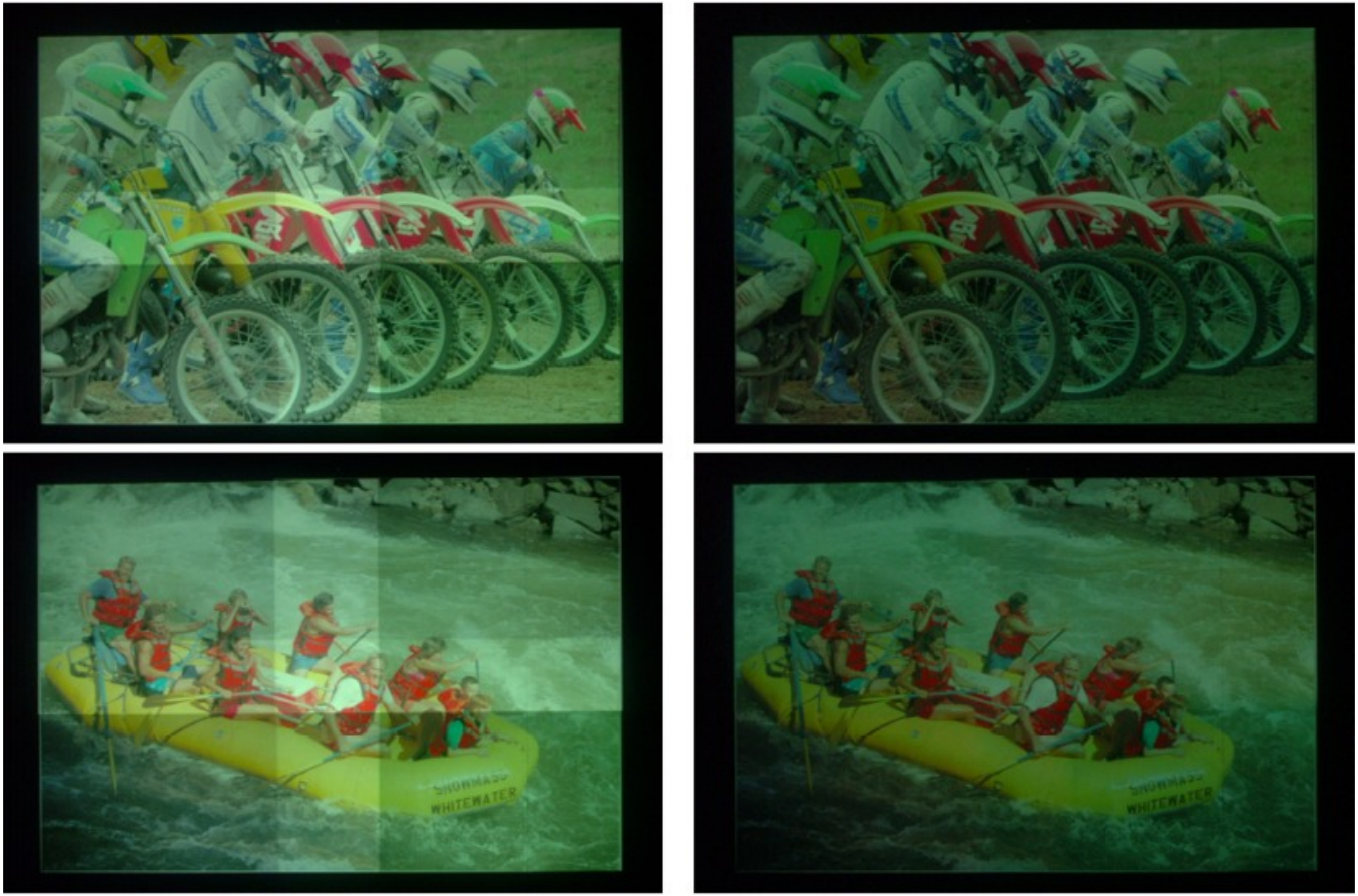
Measured luminance response

- Gives a factor for multiplication of the final images (just as in feathering)
- Can be done in graphics hardware via alpha channels





LAM: results

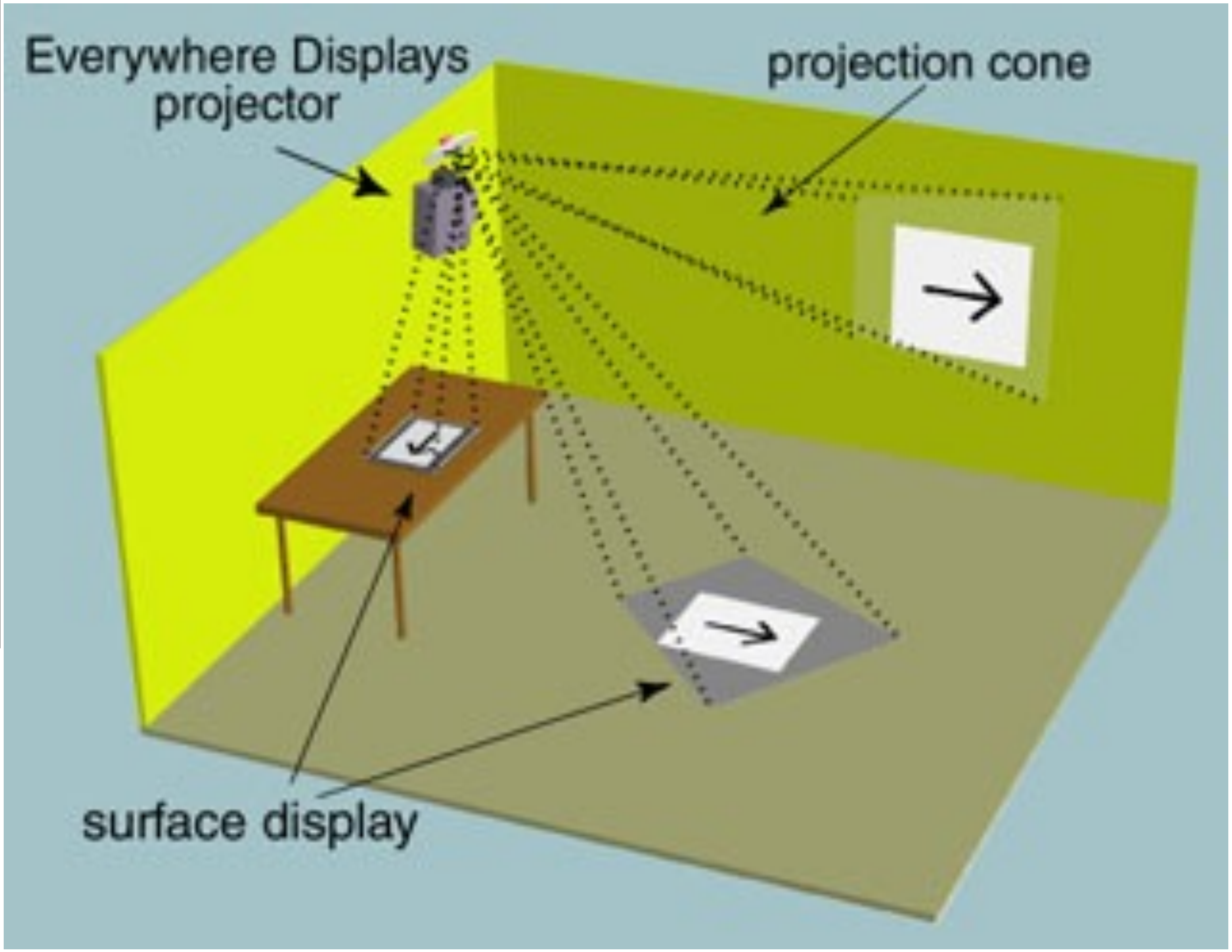


How to achieve Spatial Augmented Reality

- Projectors and their working principles
- Using projectors as shader lamps
- Combining two projectors
- Combining many projectors
- **Steerable projectors**
- Projection on structured surfaces
- Combining it all with today‘s technology

Everywhere Display Projector (IBM)

http://www.research.ibm.com/ed/



Claudio Pinhanez

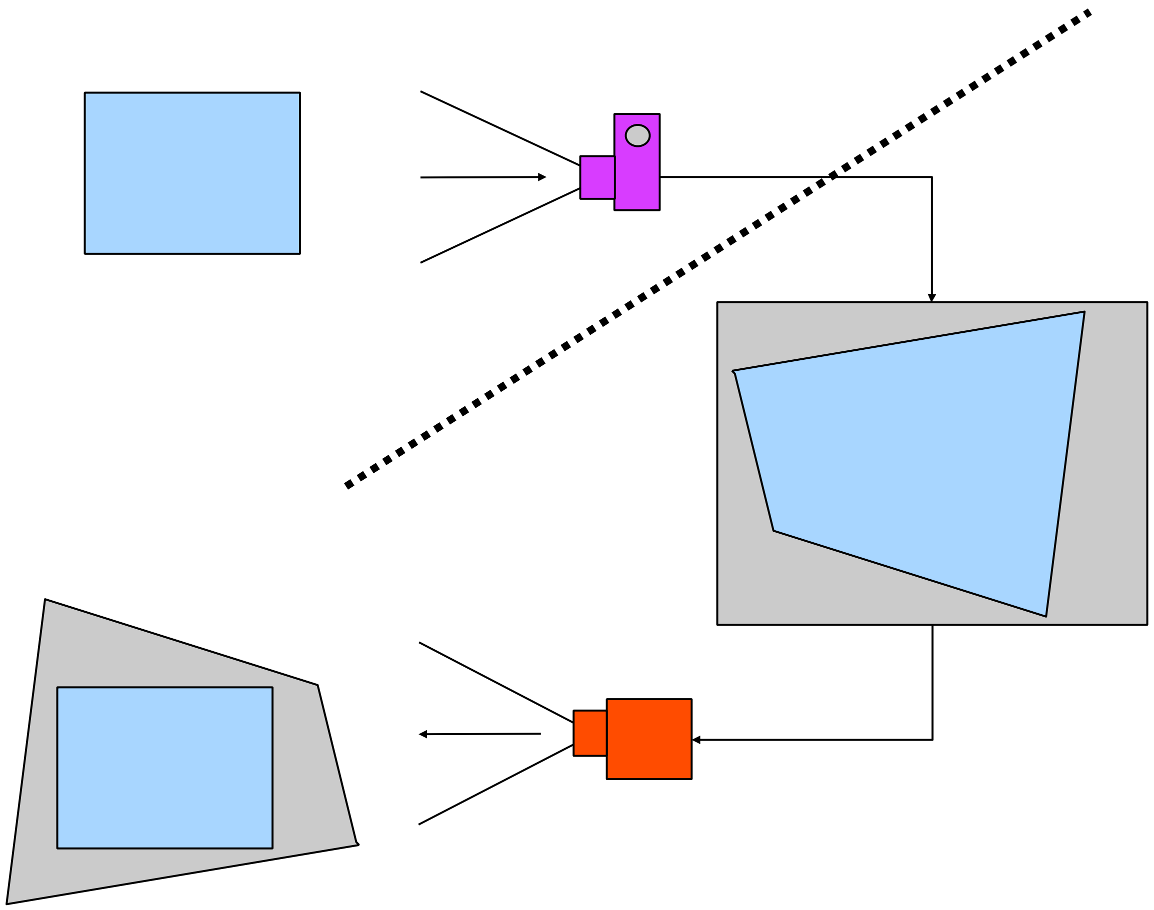
www.research.ibm.com/ed/

Everywhere display (cont.)

Output: a projector and a rotating mirror

Input: a camera for interaction, NOT for image rectification!

Undistorting the projected image

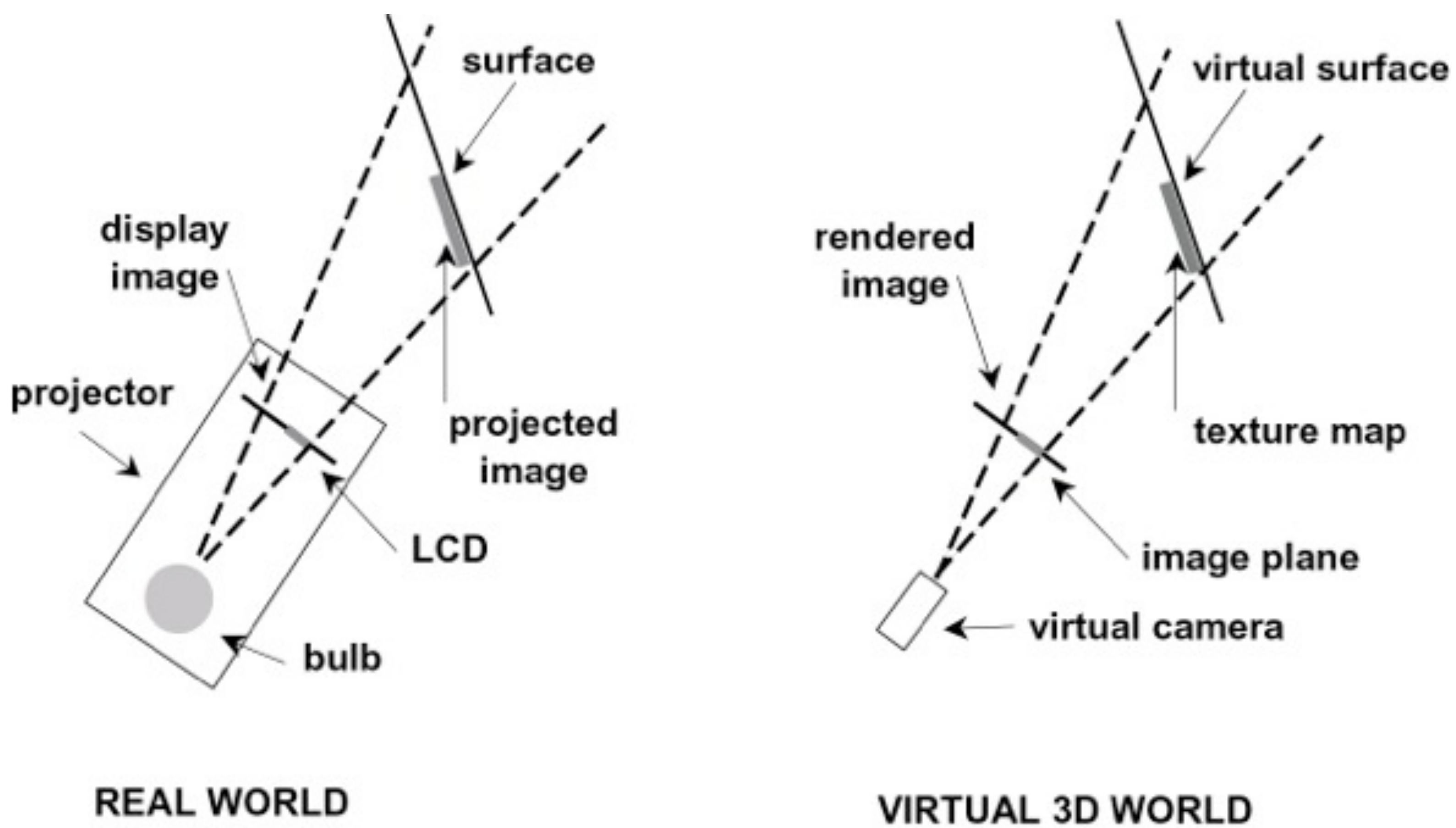


- Place original image in the 3D model
- Virtual camera image shows it distorted
- Project the distorterd image from 3D model with the Real projector into the real world

– Distortions cancel each other out IF virtual camera and real projector are in the same location

Everywhere display (cont.)

- Correct distortions
  - Use the fact that camera and projectors are geometrically the same (optically inverse)
- Use standard HW components
  - 3D-Graphics board and VRML-world



Everywhere display (cont.)

BLUESPACE office scenario

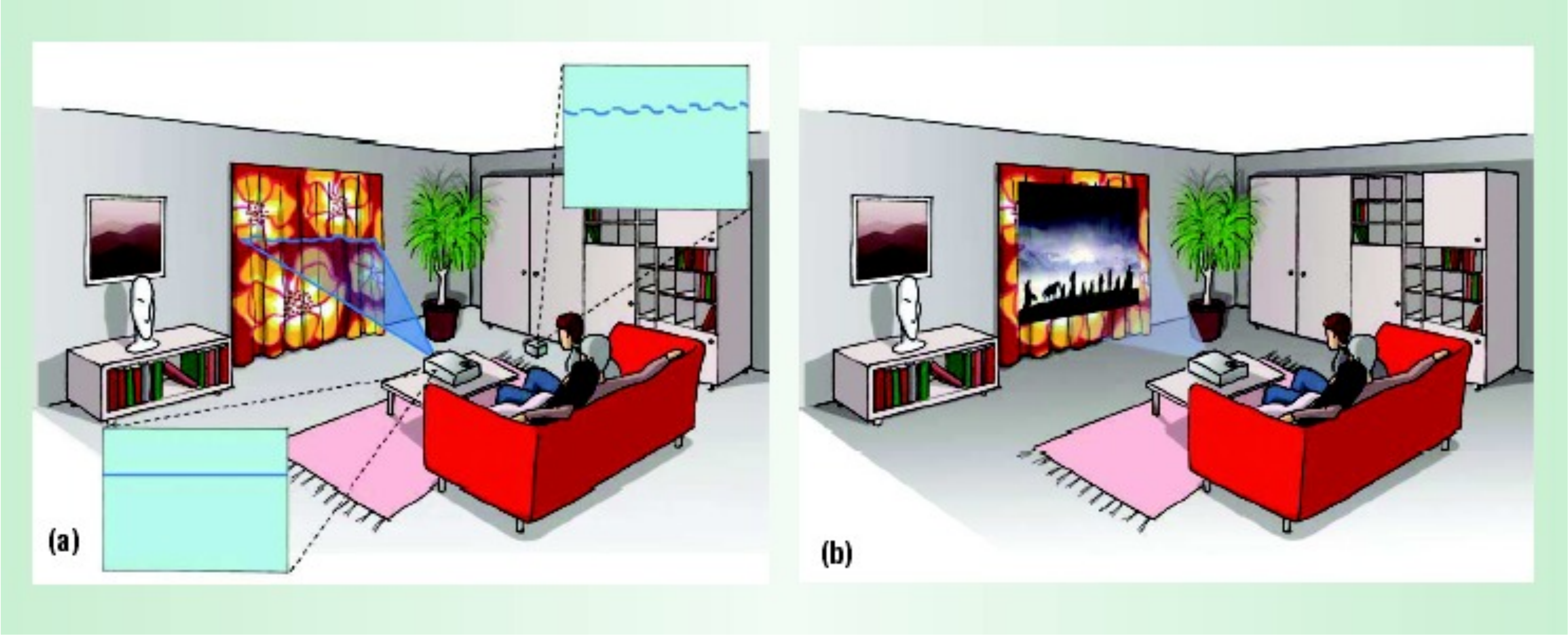
How to achieve Spatial Augmented Reality

- Projectors and their working principles
- Using projectors as shader lamps
- Combining two projectors
- Combining many projectors
- Steerable projectors
- **Projection on structured surfaces**
- Combining it all with today‘s technology

Smart Projectors

[Oliver Bimber et al., IEEE Computer, January 2005]

- Projection onto curved surfaces can be solved by 3D rectification, …but:
- What if the projection surface is not uniformly colored?
- See Video (scientific) or Video (TV)



context and task

theory

pointing

crowd

SAR

interaction techniques

in/output technologies

How to achieve Spatial Augmented Reality

- Projectors and their working principles
- Using projectors as shader lamps
- Combining two projectors
- Combining many projectors
- Steerable projectors
- Projection on structured surfaces
- **Combining it all with today‘s technology**

Examples

- IllumiRoom (see context and task chapter)
  - peripheral projected illusions.
- Mano-a-Mano

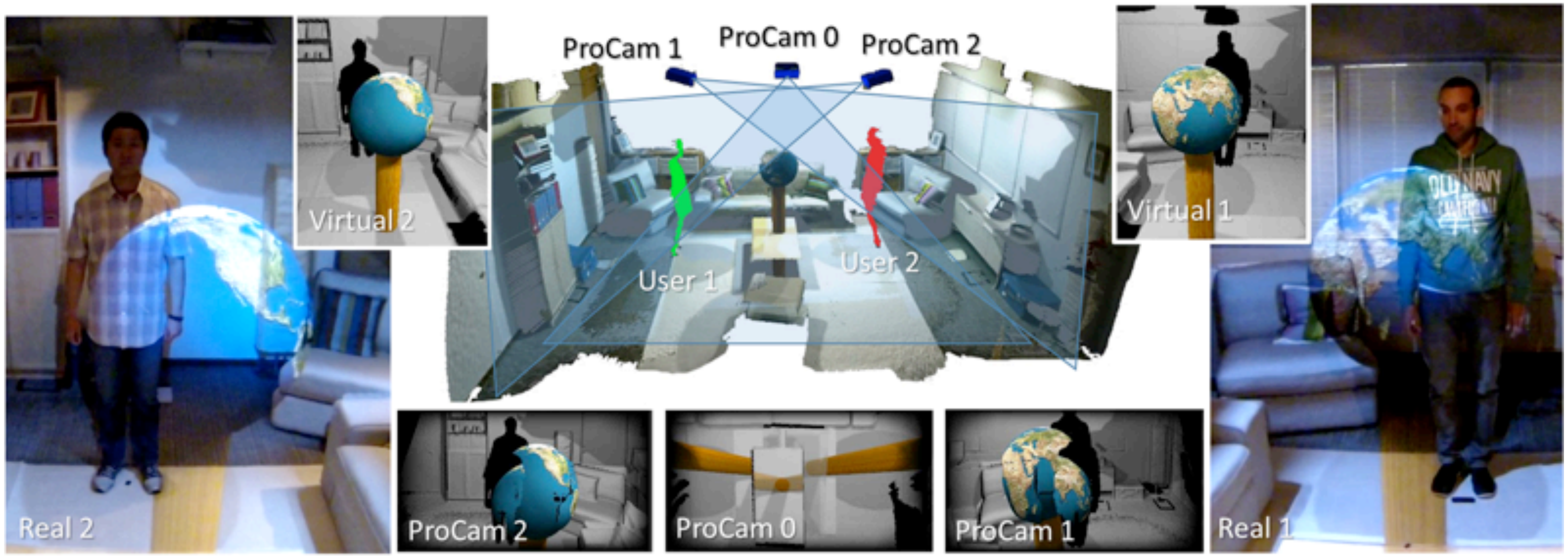
Literature: Benko, H. et al: Dyadic Projected Spatial Augmented Reality, UIST 14

Spatial Augmented Reality (SAR)

- can change surface appearance of objects
- requirement: How would you implement that? What technology to use?
  - knowledge about the users' head position
  - geometric model of physical environment
    - alter the projected graphics to account for distortion of projected image.
- SAR is view-dependent rendering
  - supports single view
  - Mano-a-Mano supports separate perspective views for two users when arranged face-to-face.

Hardware configuration

- 3 HD video projectors, each paired with a Kinect
- 1 PC driving all three projectors
- 3 PCs each running one Kinect (Kinect SDK can support only one camera per PC)
  - sending images to main PC via network
  - depth data is merged into single scene using Unity 3D



Literature: Benko, H. et al: Dyadic Projected Spatial Augmented Reality, UIST 14

Calibration

- Calibrate projector/Kinect pair
- Calibrate relative pose of each projector camera pair.
- get information about the physical environment

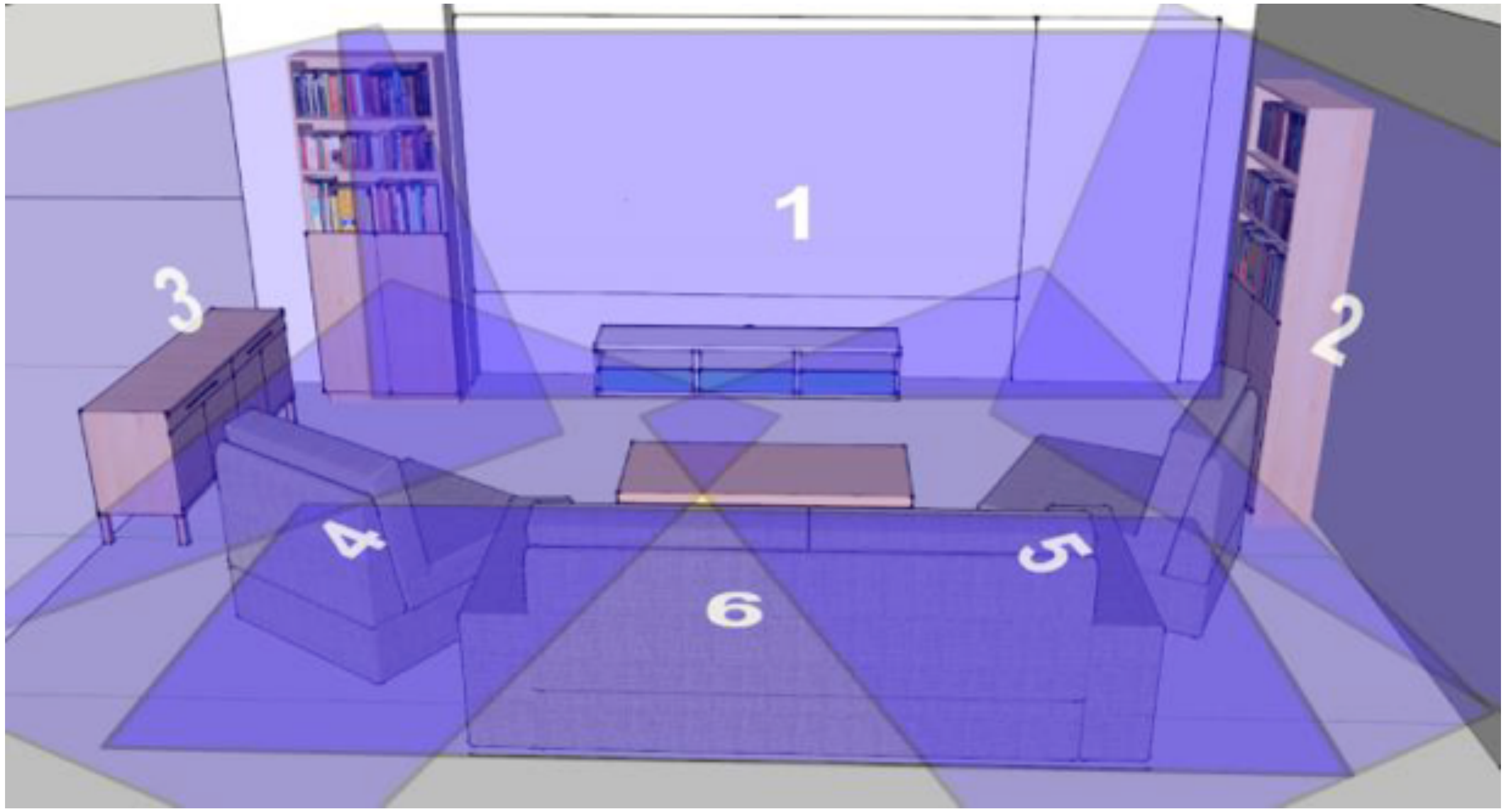
Calibration: projector/camera pair

- requirement: pose, focal length and optical center of each projector and Kinect camera.
- idea:each projector in turn displays a series of gray code patterns, these patterns are observed by the color camera of paired Kinect.
- result: precise mapping of 3D point between camera's coordinate frame to corresponding point in projectors' image.

Literature: Benko, H. et al: Dyadic Projected Spatial Augmented Reality, UIST 14
Literature: Jones, B. et al: RoomAlive: Magical Experiences Enabled by Scalable, Adaptive Projector-Camera Units, UIST'14

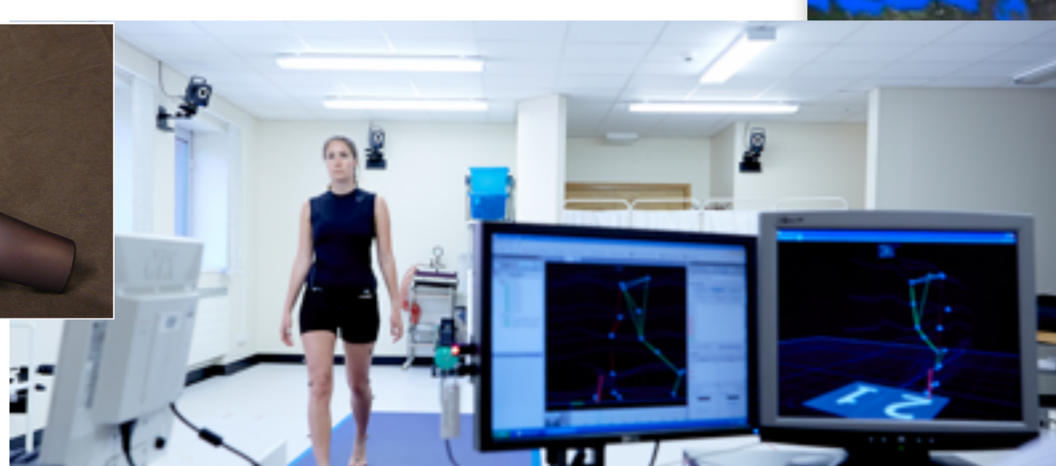
Calibration: relative pose of each pair

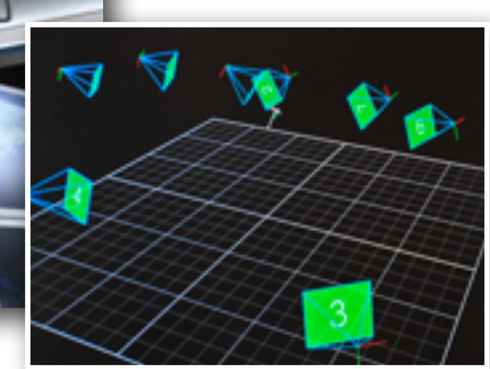
- have all Kinect color cameras observe the gray code patterns of all other projectors
  - look for regions where the other projectors overlap with the camera's own paired projector
- result: world coordinate system for all projectors and cameras.



Side Story: VICON Cameras

- VICON is not a depth camera!
  - yet very precise in tracking (precision in mm range)
  - requires passive markers
- manual calibration procedure uses a specific delivered object (wand) with mounted markers
  - distance between markers is defined
- swing the wand around the room
  - each camera registers which part of the wand is





http://www.vicon.com/content/images/other_vicon_software.jpg

Calibration: physical environment

- use depth camera to scan the environment.
- Kinect for Windows version 2 more precise than original Kinect
  - constant precision of depth (0.5m - 4.5m)
  - depth precision degrades with distance.

Summary

- mid-air pointing model
  - further development of Fitts' law prediction models
  - understanding what effects interaction performance leads to the development and improvements of techniques
- crowdsourcing
  - involving the inhabitants of an environment...
  - how it developed
  - applications and resulting data sets you can make use of
- spatial augmented reality
  - geometric projection concepts
  - multiple projectors
  - how to perfect the illusion