
Active Stereo Vision

COMP 4102A

Winter 2016

Gerhard Roth

Version 1

Why active sensors?

- Project our own texture using light (usually laser)
 - This simplifies correspondence problem (much easier)
- Pluses
 - Can handle different ambient lighting conditions
 - Can get 3d data when there is no natural texture (i.e. white wall)
- Minus
 - Need active source and a way to project it (laser dangerous?)
 - Need more complex hardware
- A number of different systems, but two principles
 - Triangulation (same as stereo but the light source replaces second camera) with camera and light source
 - Time of flight (produce a pulsed beam of light, measure distance by time light takes to return)

Pulsed Time of Flight

Basic idea: send out pulse of light (usually laser), time how long it takes to return

Advantages:

- Large working volume (up to from 20 to 1000 m.)

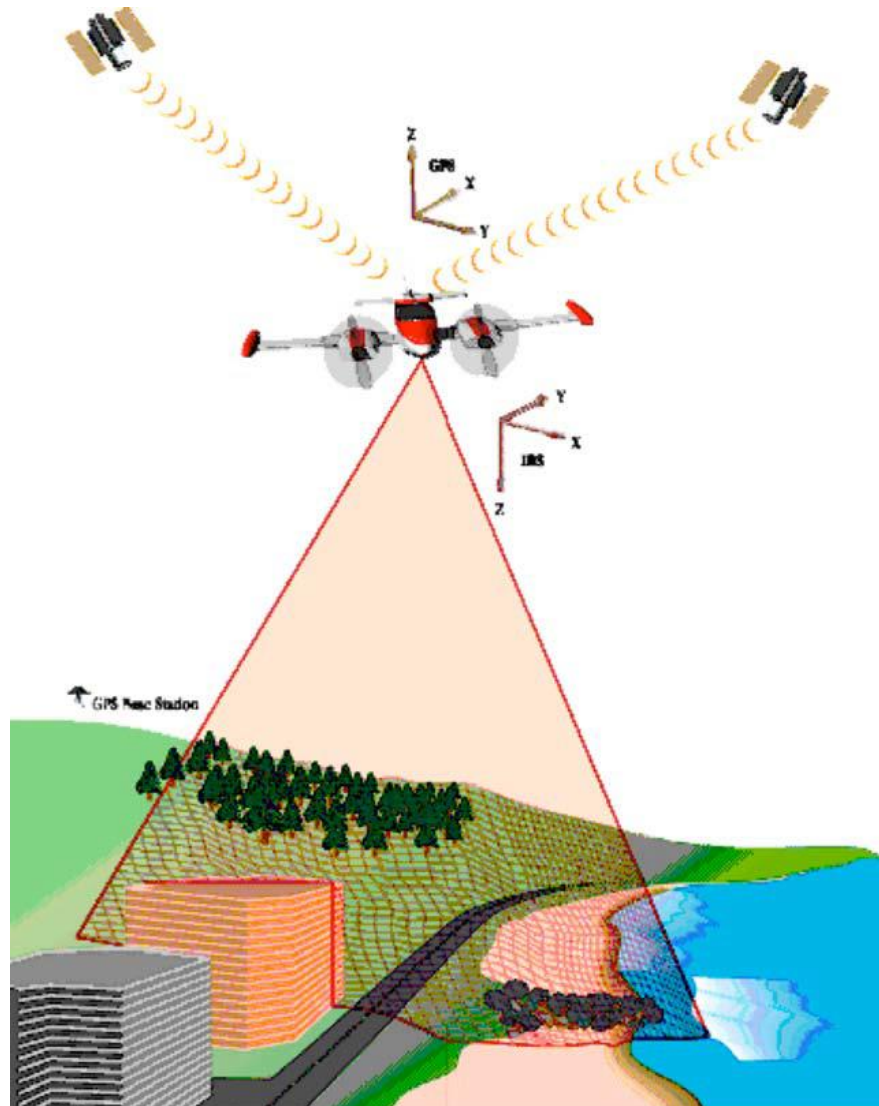
Disadvantages:

- Not-so-great accuracy (at best ~5 mm.)
 - Requires getting timing to ~30 picoseconds

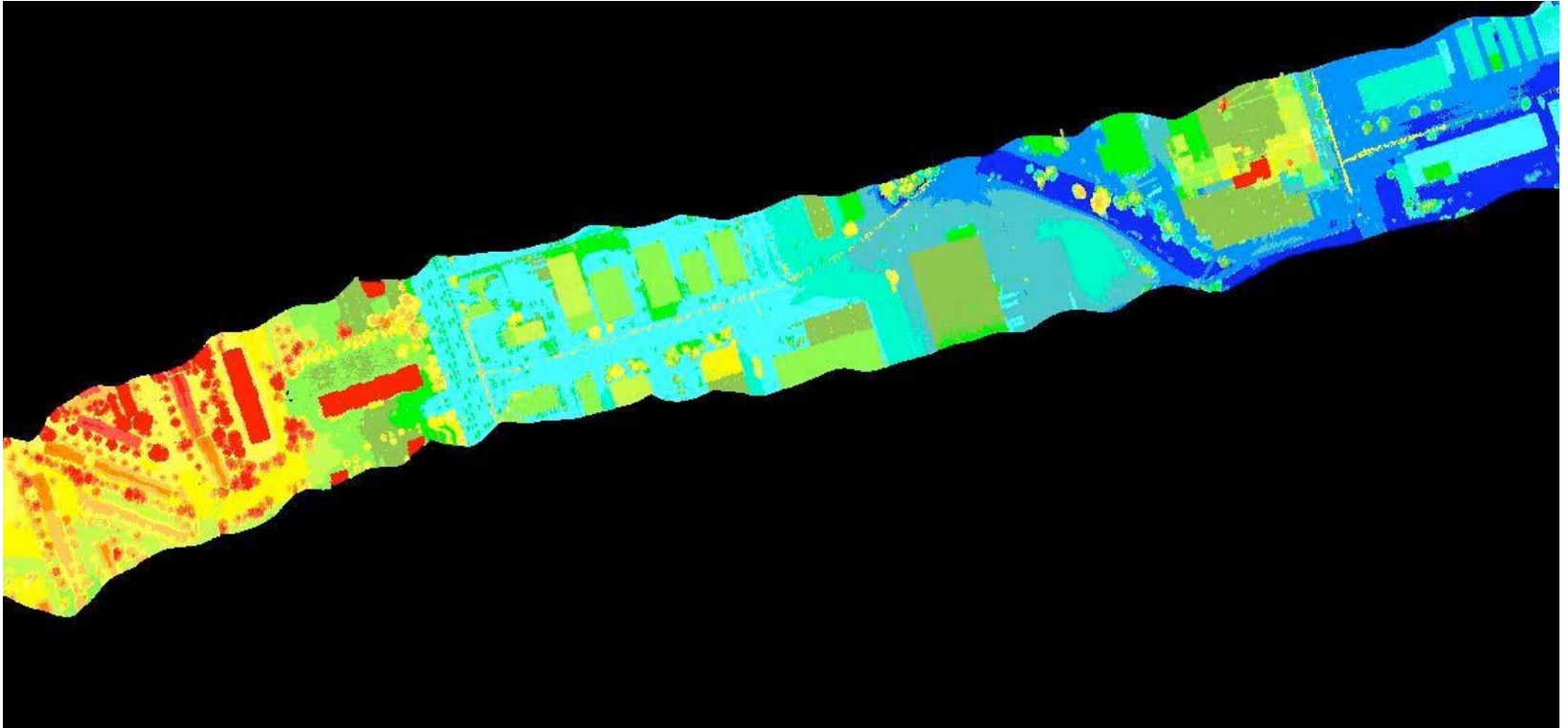
Often used for scanning buildings, rooms, archeological sites, etc.

The only practical long range measuring technology (triangulation fails over 20 meters)

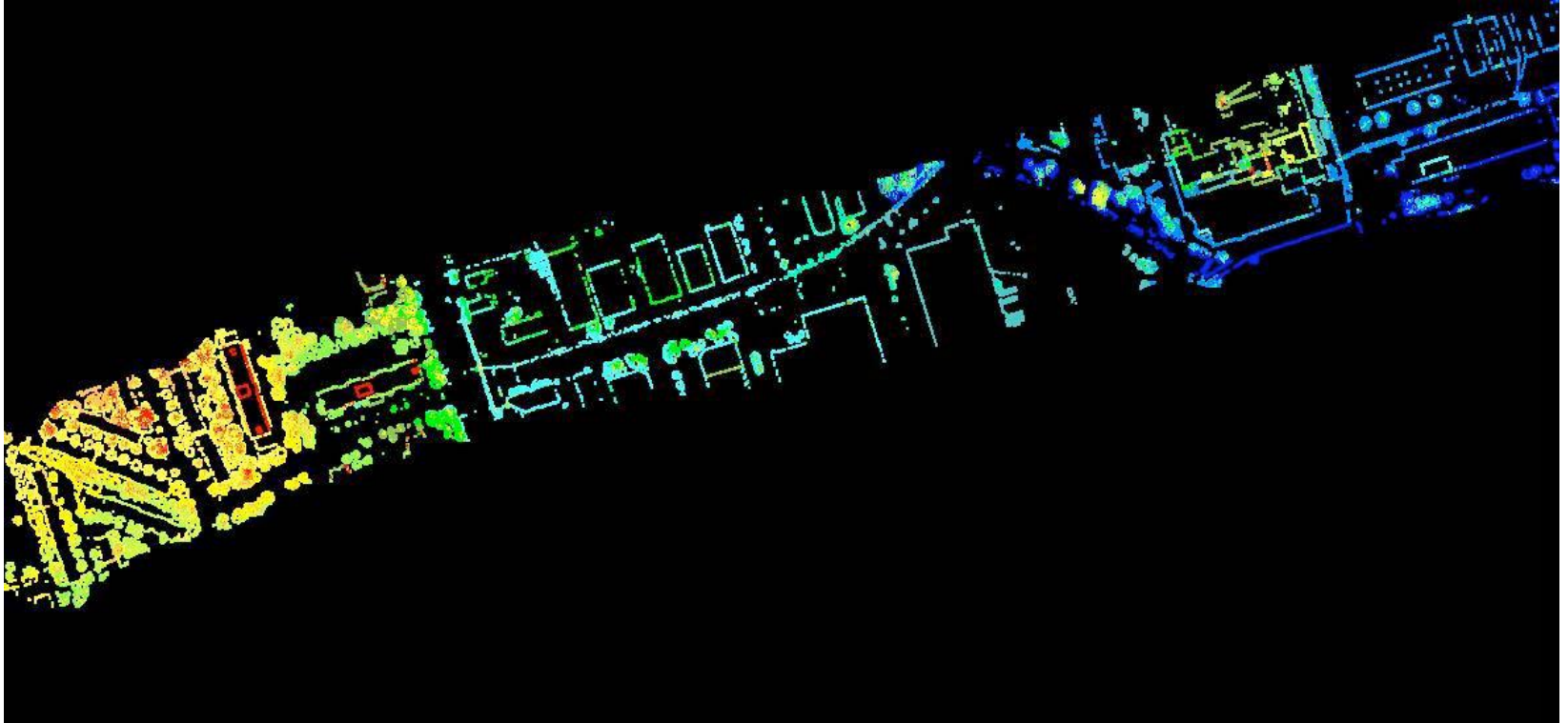
Optech – Airborne Laser Mapping



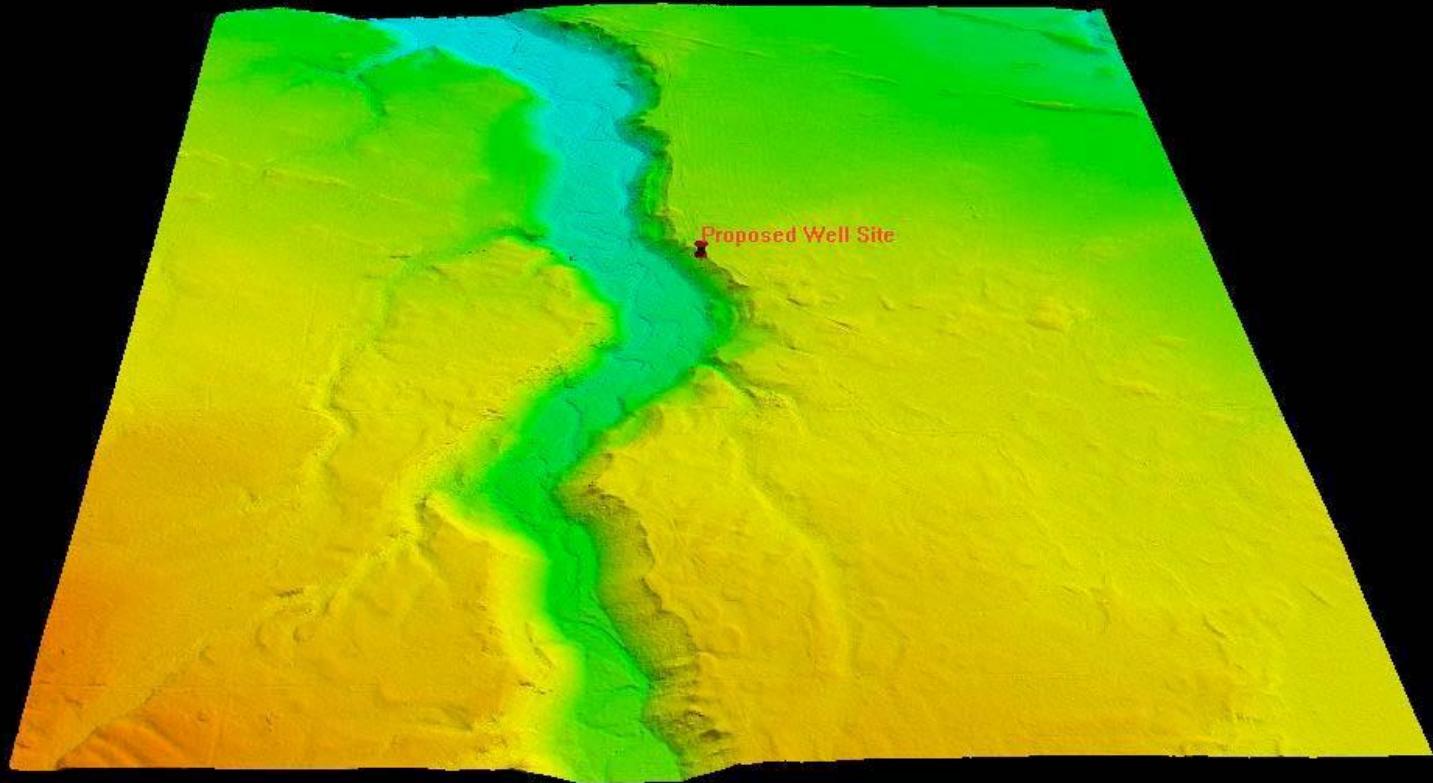
Raw Image – depth is colour coded



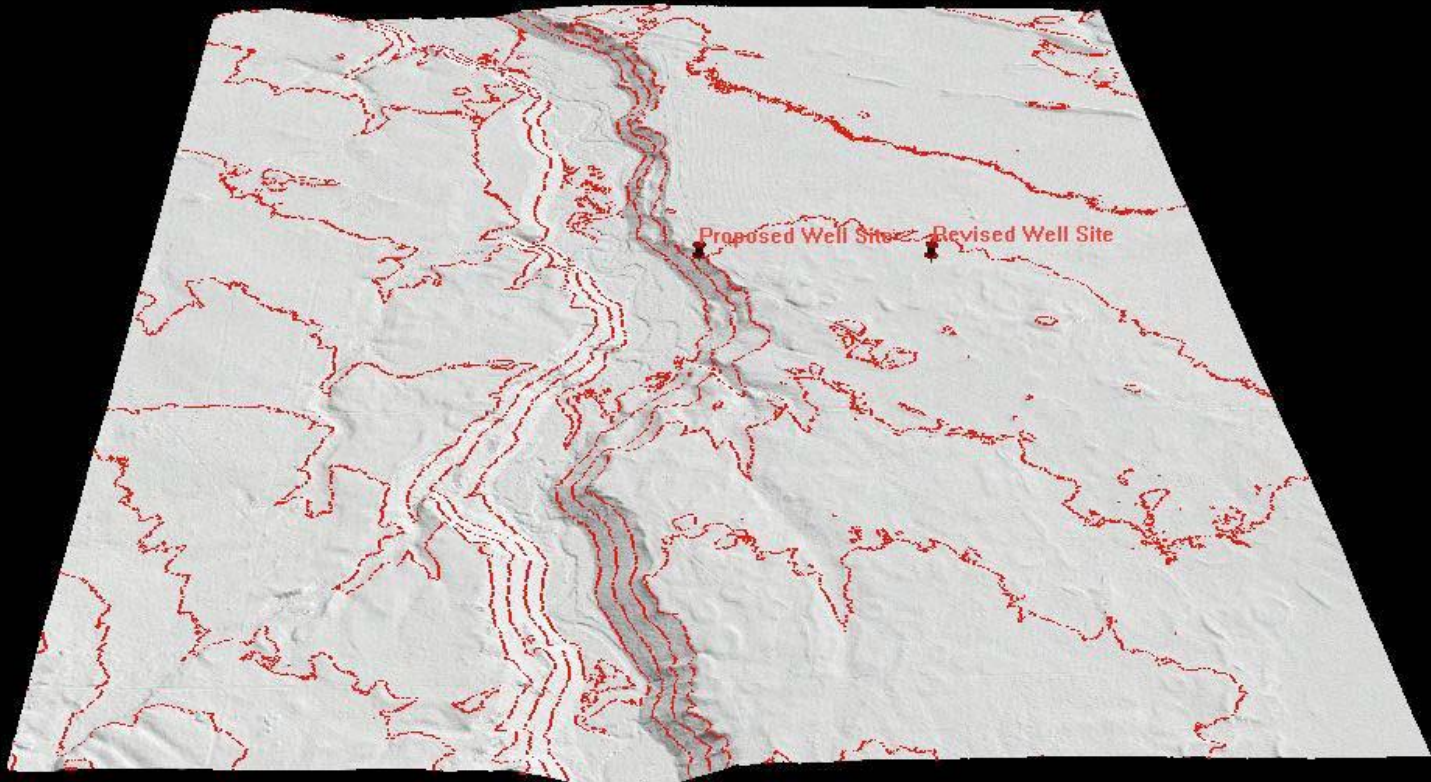
Building, outlines, trees and wires



Bare Earth Model

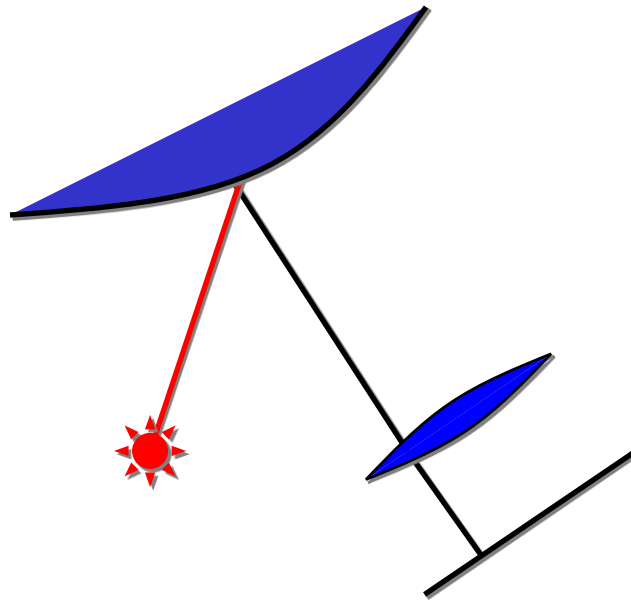


Removing the trees



Triangulation

- One or two cameras and a light source
- Many possible light sources and variations
- Still use triangulation to find the depth

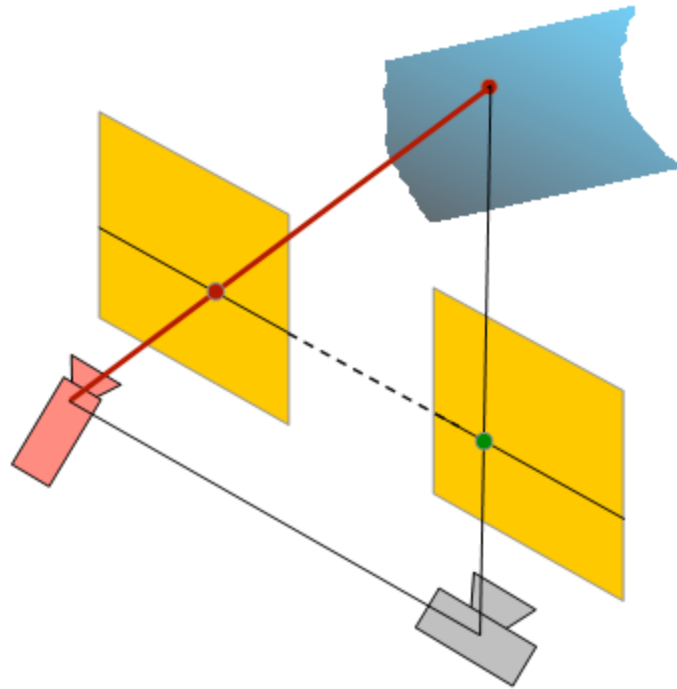


Simplest possible triangulation system?

- Take two calibrated stereo cameras
- Use a laser pointer to shine light on where we want the depth
- Find that laser spot in both images, this feature must correspond, so you get 3d
- This is easy because the laser spot is very bright compared to the rest of the world
- This works, but getting data is very slow since you must move around the laser spot
- Very easy to build, and to make it work!

Triangulation system with one camera?

- What if you have a single laser pointer and also a single camera looking at the spot?



Triangulation with one camera?

- Assume laser moved by a calibrated motor
 - Then you know direction in space of the laser beam
- Camera calibrated and you know baseline, so can find laser spot location in image
- Then you can still triangulate to find depth!
 - Even though you only have one camera
- Need a very accurate and high speed motor to move the laser spot around the scene
- This is complex hardware but is exactly what was done at NRC over about 30 years!

Triangulation can be very accurate

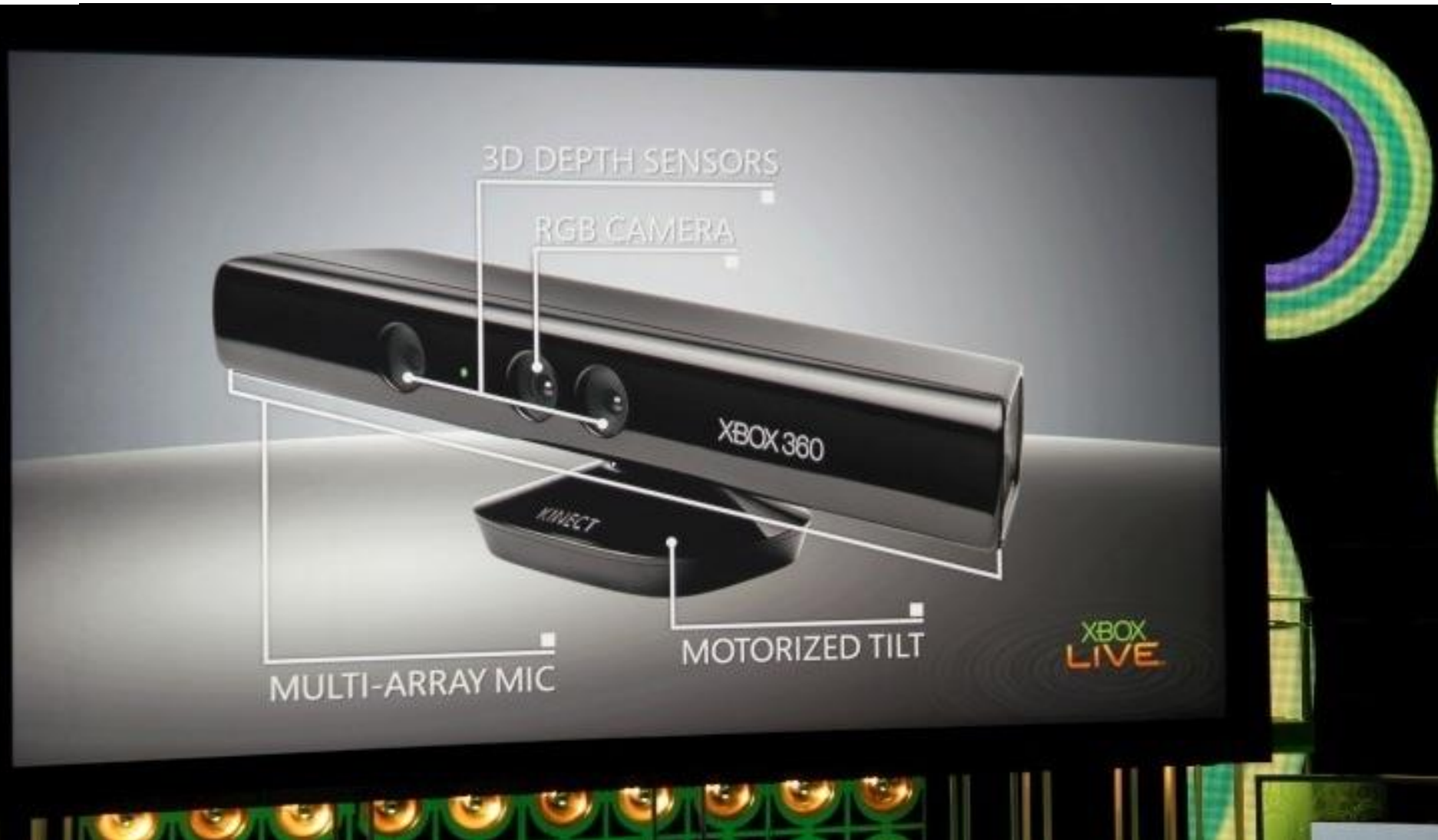
Can get accuracy down to 20 microns
(1/50th of a millimeter!)



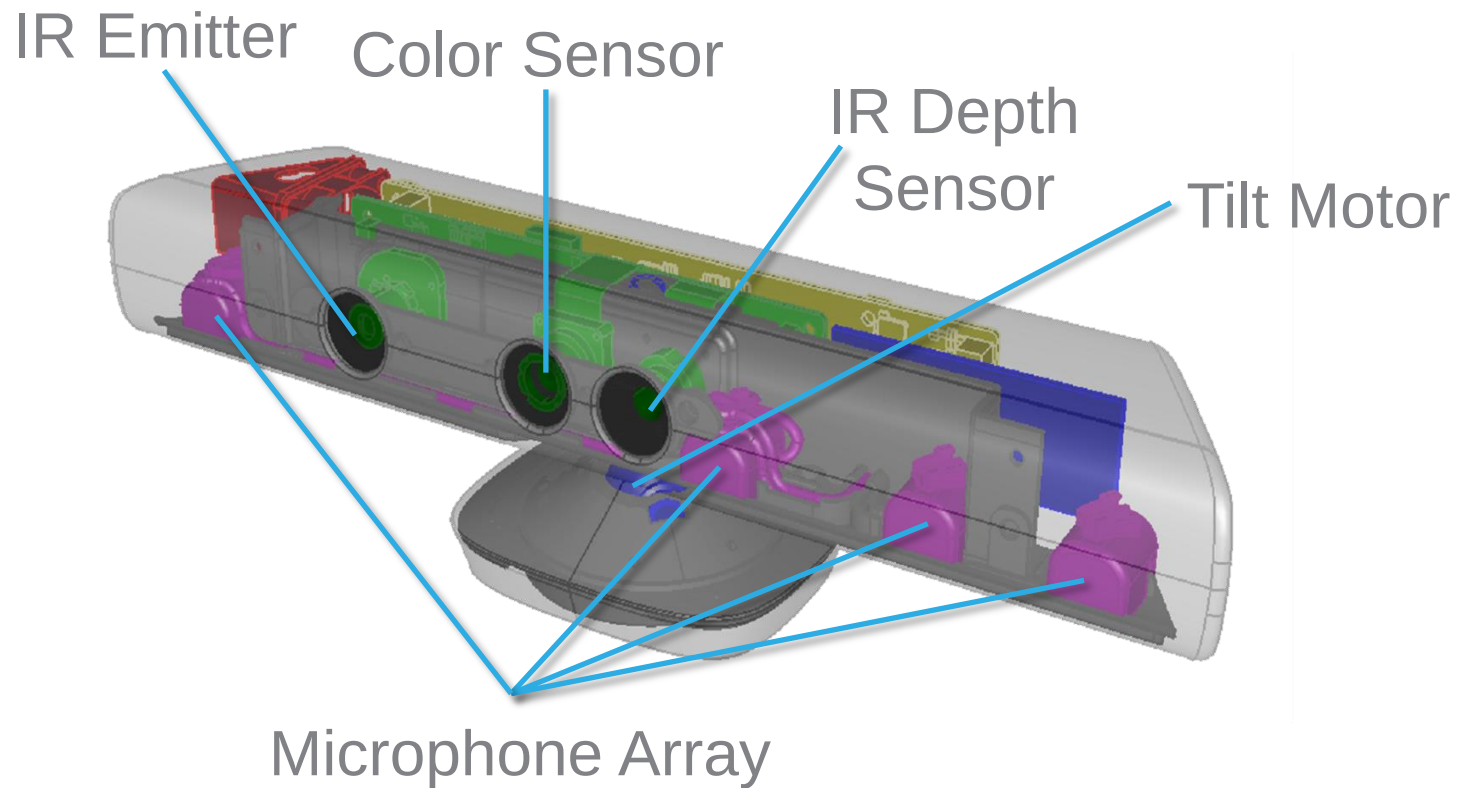
Microsoft Kinect

- Triangulation based system for finding depth
- Designed to interpret motions, not to build accurate 3d models or measure objects
- Frequency of infrared projector similar to sun
- So can not be used close to a window or be taken outdoors
- Still, for Human Computer Interaction, Kinect is a big breakthrough
- The first inexpensive and mass produced active sensor for consumers and researchers

Kinect Hardware



Kinect Hardware



Kinect Hardware



Sensors/Resolution of Kinect

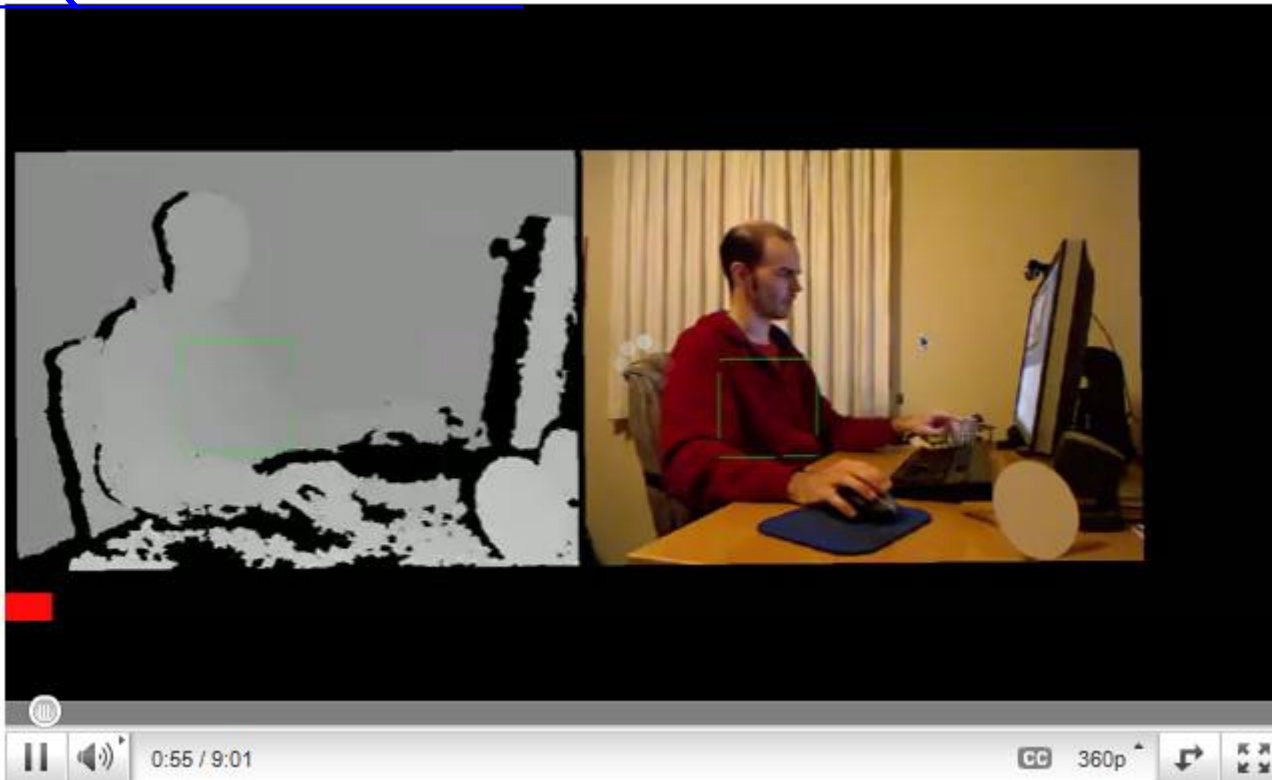
- Separate sensors for depth and colour
- Color
 - 12 FPS: 1280X960 RGB
 - 15 FPS: Raw YUV 640x480
 - 30 FPS: 640x480
- Depth
 - 30 FPS: 80x60, 320x240, 640x480
 - Not that accurate unless extra calibration is done
- Depth and colour registered so you can get the colour for each depth point

Depth and Intensity Images

- Depth image shown in “depth map” style, with brighter points closer to camera

<http://www.youtube.com/watch?v=inim0xWiR0u>

<http://www.youtube.com/watch?v=7TGF30-5KuQ&feature=related>



How does Kinect get depth

- Project “pseudo-random” dots on world



[http://www.youtube.com/watch?v=dTKlNGSH9Po
&feature=related](http://www.youtube.com/watch?v=dTKlNGSH9Po&feature=related)

Local patterns are almost unique

- What is the principle? Uses self identifying patterns of dots (like glyphs)
- What are glyphs?
- A local pattern that identifies itself uniquely

Qrcode

Augmented Reality Tags



Glyphs printed in paper (Dataglyphs)

CareCo

Personnel Services
1637 Bear Vista Blvd
Whittier, CA 94031
Tel: (408) 857-0921

January 12, 1996

J. Craig Erickson
78 Hirsch Road
Brenton, IL 32564

Dear Craig:

At our recent Open Enrollment Benefit Carnival, you elected to spend your 1996 Nifty-Flex Benefit Allowance on

Medical care	\$1525
Dental care	\$287
Childcare	\$122
Prepaid Donut Plan	\$45
Total	\$1979

This leaves a balance in your Nifty-Flex Benefit account of \$21. Please indicate if you would like this to be

☐ Donated to charity.

☐ Paid to you as ordinary income.

☐ Held in your Nifty-Flex Benefit account.

Please sign and return this form to us to confirm these choices. If you would like to make other changes, please visit your local benefit counselor and fill out a new choice form.

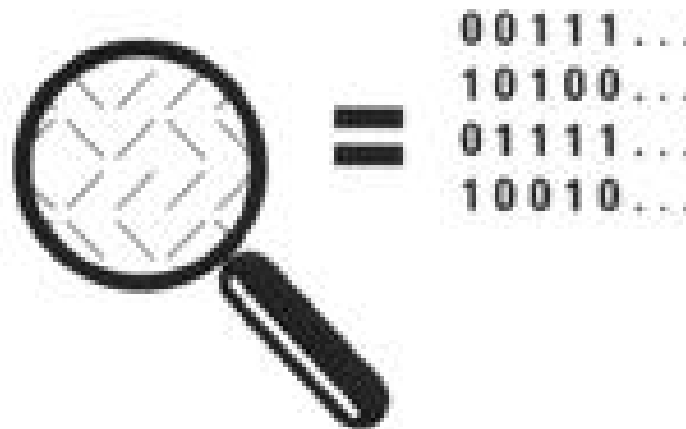
Cooper M Price

Cooper M. Price
Manager, Benefits Services

YES! This is what I want! _____

J. Craig Erickson

Old Xerox technology
A little pattern that is hard to
see but encodes a unique bit string

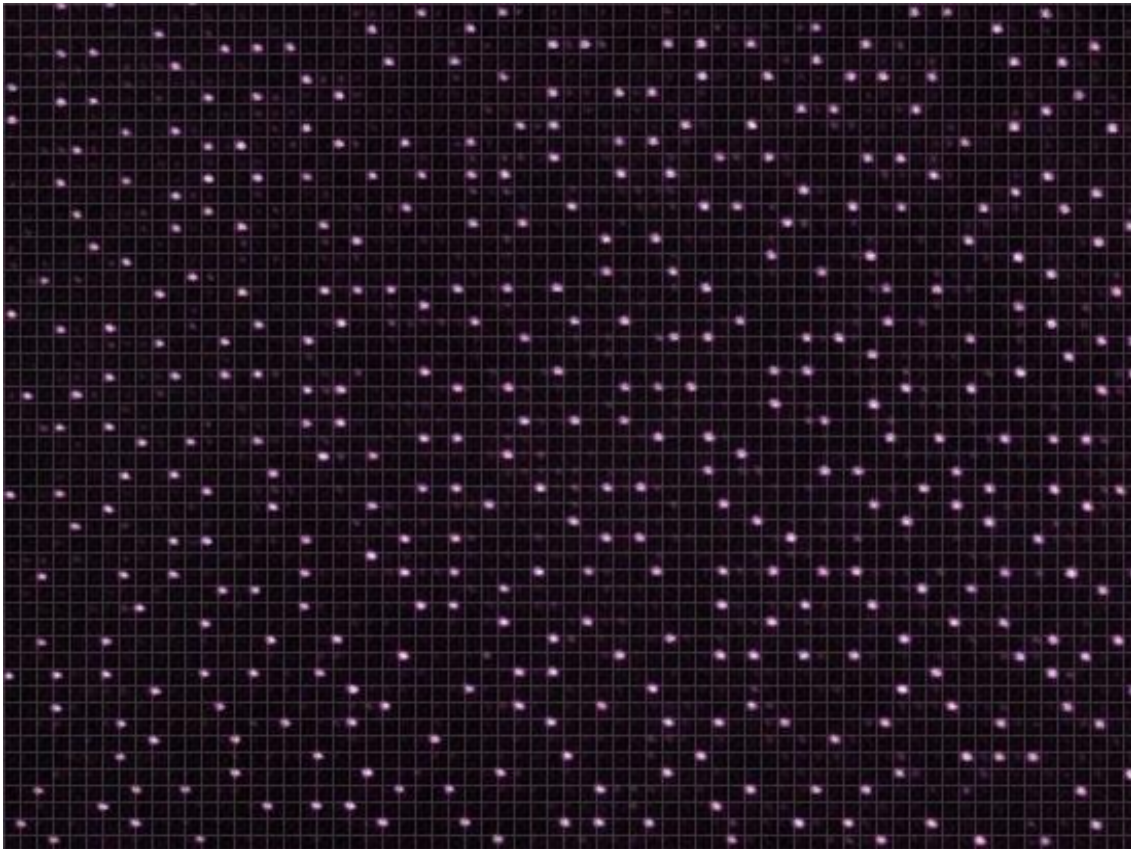


Kinect Projects dots which are glyphs



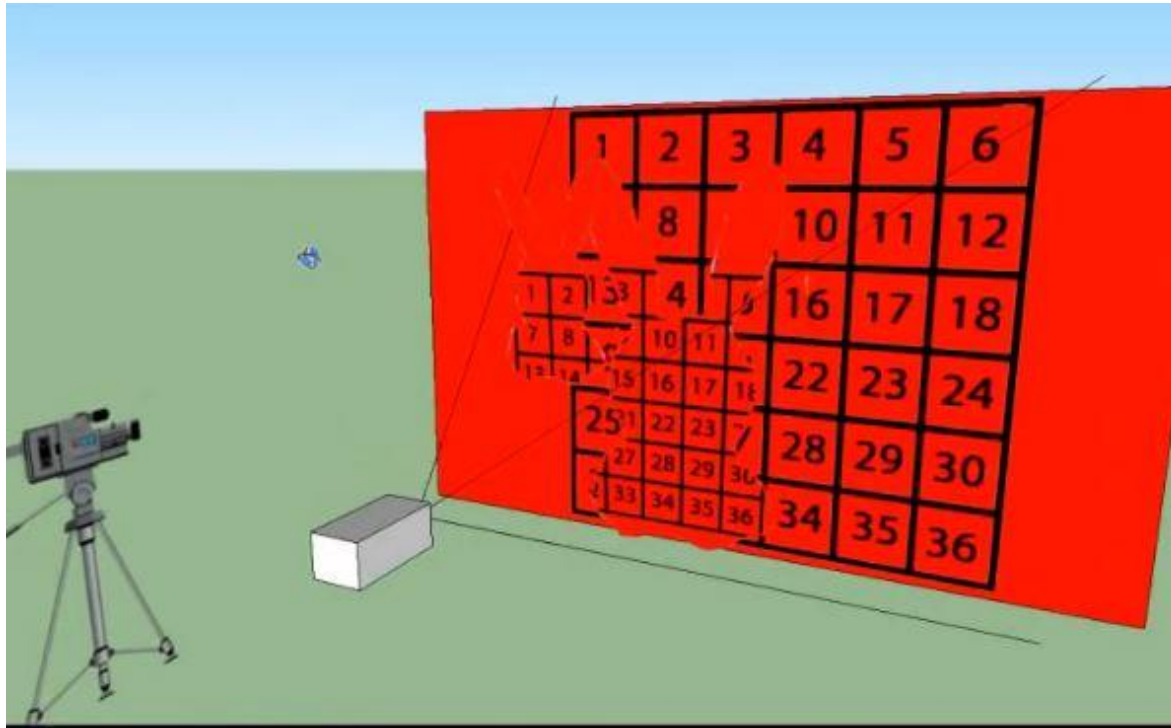
Kinect Glyphs – almost unique

- Local pattern identifies location of projection
- Find local identifier by looking in a small region around a given point => code



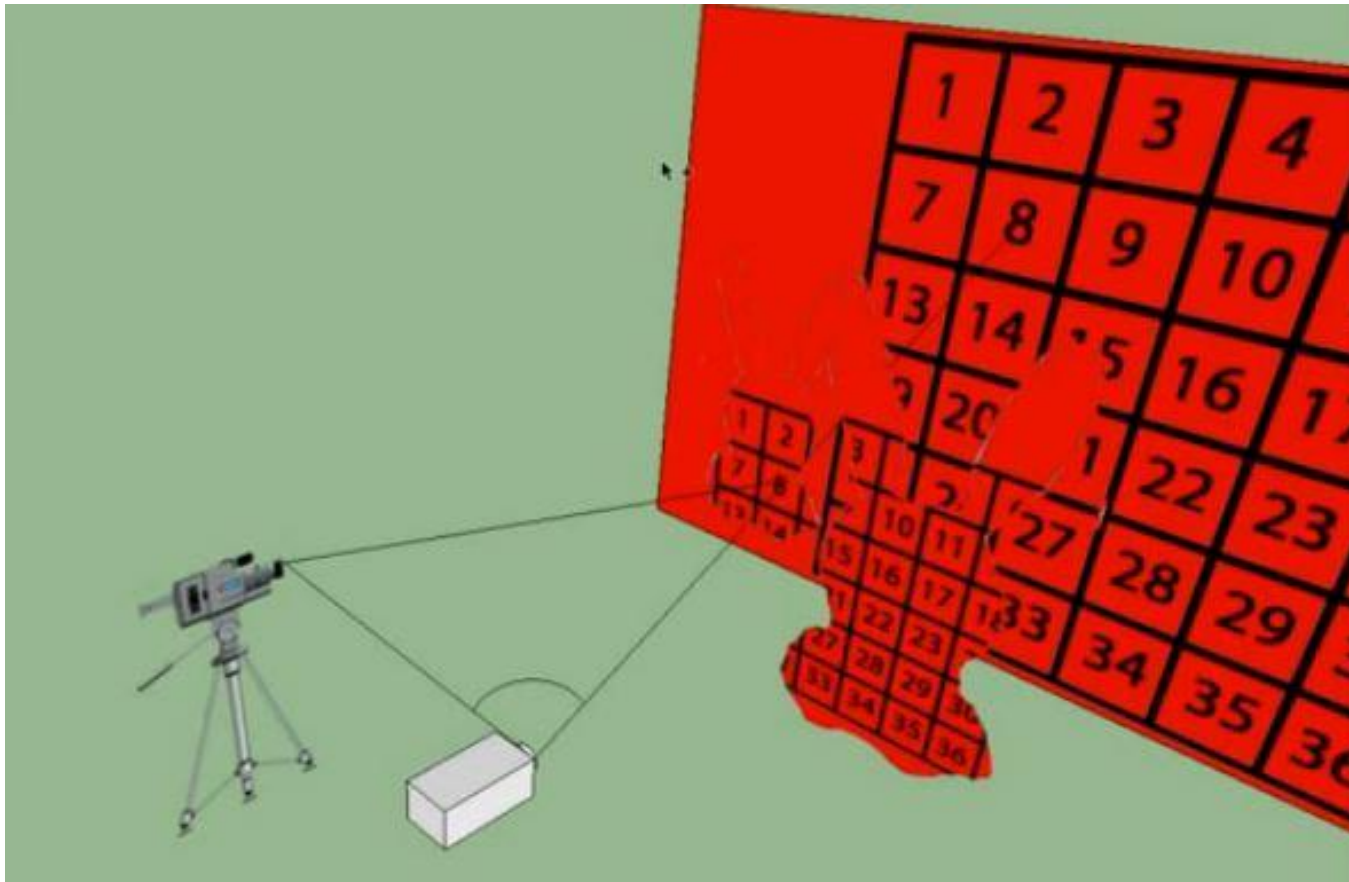
How do Pseudo-Random dots work?

- One you get the glyph, a prior calibration tells you the angle(s) and therefore the ray for that particular point
- So now you can triangulate to get depth!



How do Pseudo-Random dots work?

- Repeat this process for each small region in the dot image to get depth at that point



Kinect Depth Acquisition Summary

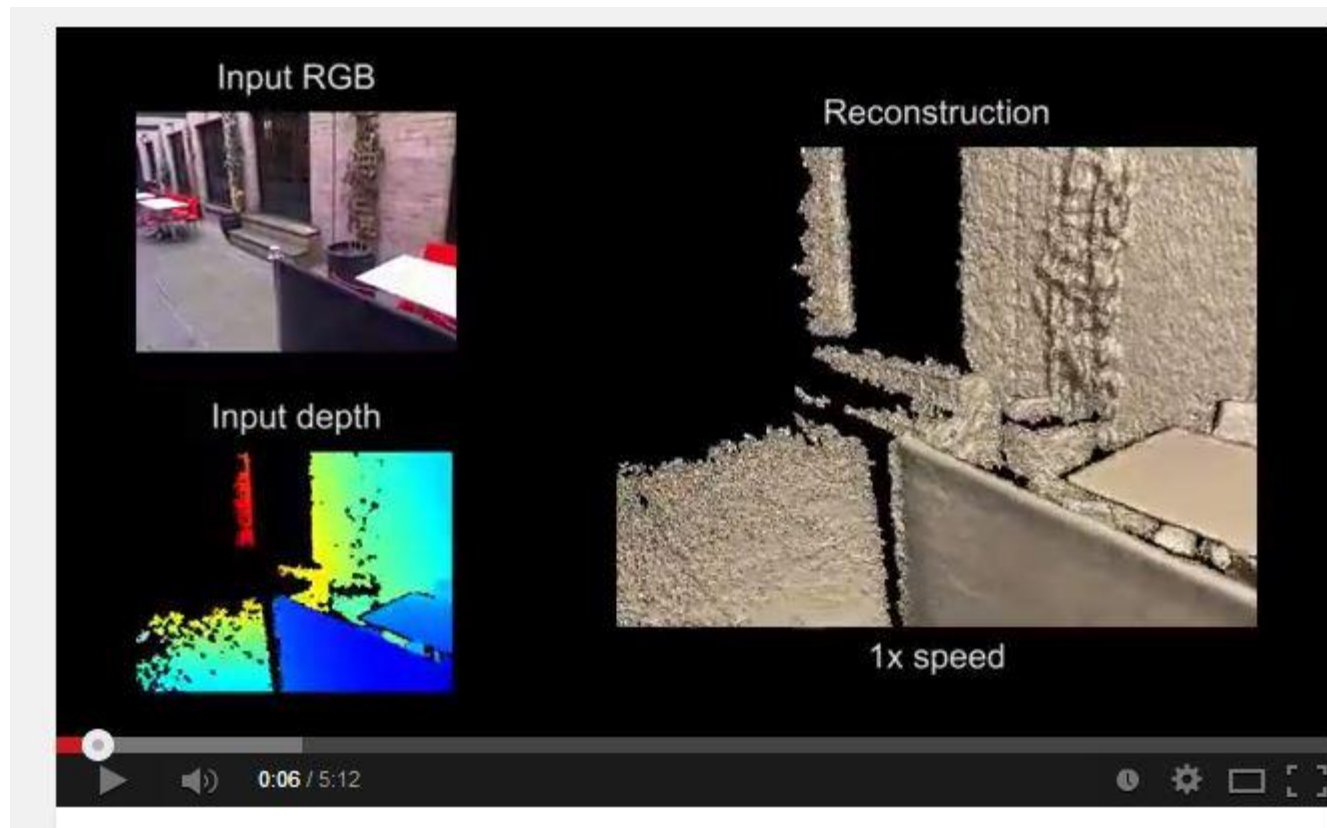
- There is a projector for the laser dots and a sensor just for these dots (infrared)
- We can recognize the glyph in the infrared image so can triangulate to find the depth
- This requires a prior calibration process so that we know the rays for the laser dots
- Still just ordinary triangulation process
- There is a another camera that produces a separate and distinct intensity image
- The Kinect returns both a depth map and the overlayed intensity image

Model Building with a Kinect

- Given a series of depth images (from Kinect) and overlaid intensity what can we do?
- A simple model building algorithm
 - Take overlapping depth images
 - In intensity image find some surf images
 - Each surf feature has range value in depth
 - Each surf feature has range value in depth
 - Align the overlapping depth images
- If you repeat this process enough times you get one big model

Kinect for model building

- <http://www.youtube.com/watch?v=NsrnmniEvO4s>



Depth sensor better than intensity?

- Is it easier to use a Kinect (depth sensor) or an ordinary digital camera to make models?
- Using a Kinect is much better because the depth accuracy from the Kinect does not change as you move camera
- Depth accuracy depends on baseline alone
- With an intensity image sequence the quality of any depth reconstruction process depends on the spacing between the images
- Can not rotate the intensity camera and get depth, but can rotate Kinect camera

Limitations of Kinect

- Not that accurate unless you do more complex calibrations
- It was designed to interpret motions, not to build accurate 3d models or measure objects
- Frequency of infrared projector similar to sun
- So Kinect can not be used close to a window or be taken outdoors in bright sunlight
- Multiple Kinects interfere with each other
- For Human Computer Interaction, Kinect is a big breakthrough; inexpensive and useful