

---

# Active Stereo Vision

COMP 4900D

Winter 2012

Gerhard Roth

# 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:

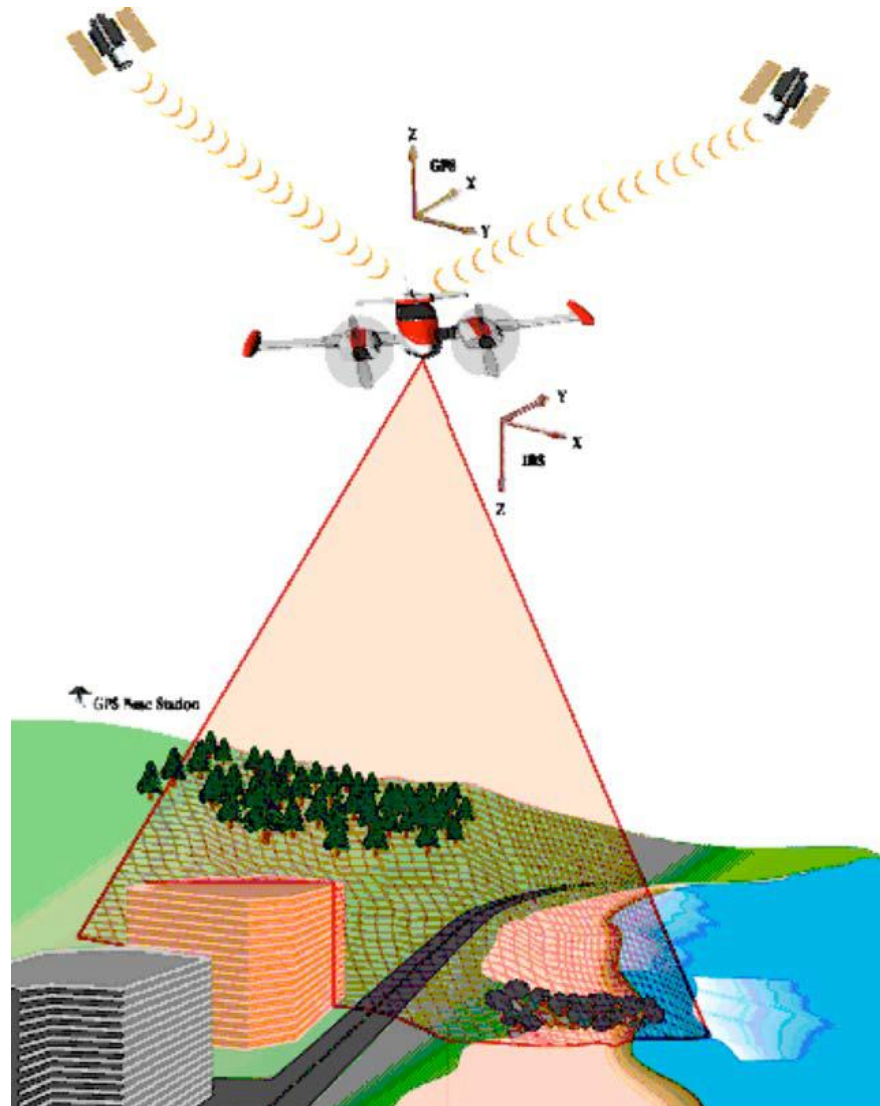
- Not-so-great accuracy (at best ~5 mm.)
  - Requires getting timing to ~30 picoseconds

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

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

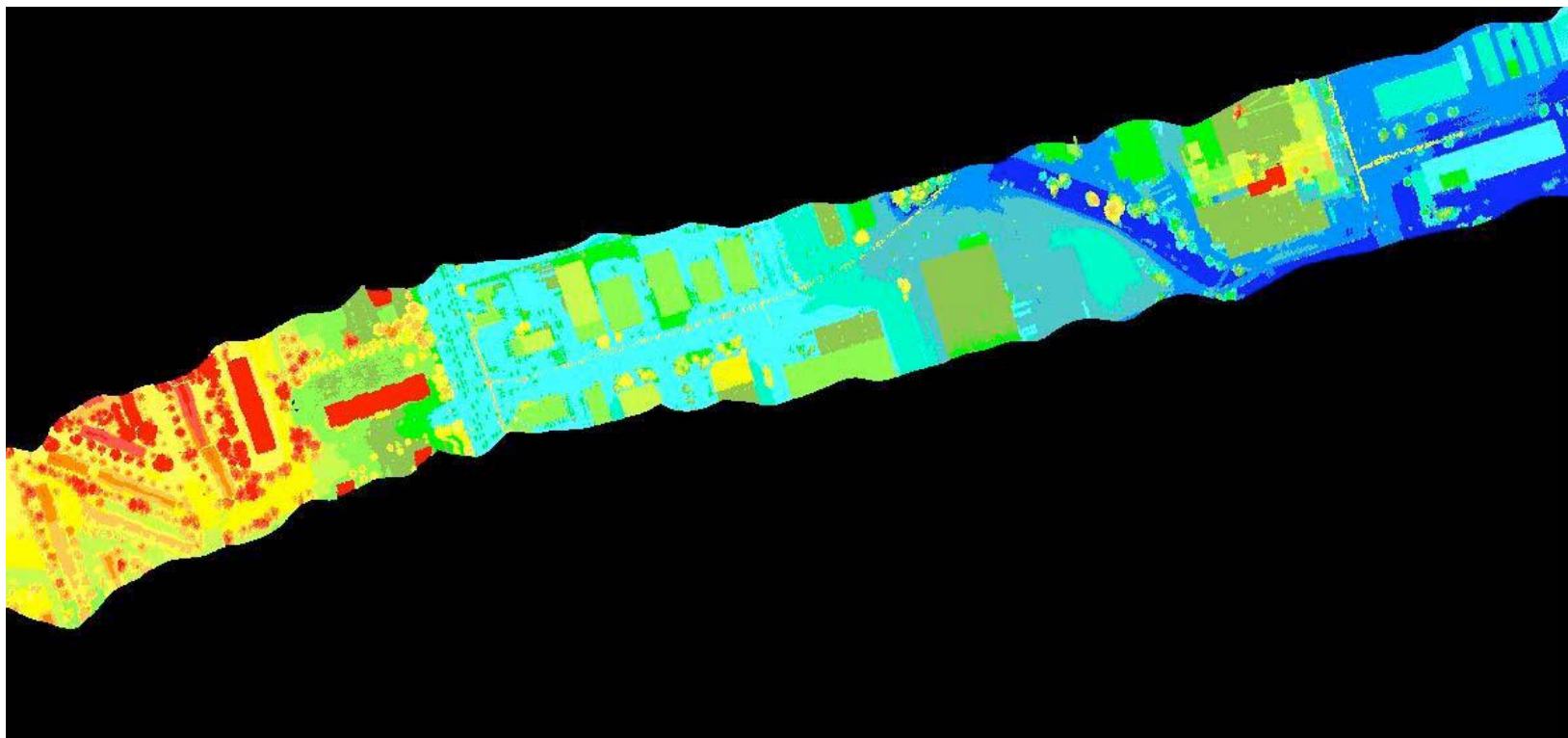
# Optech – Airborne Laser Mapping

---



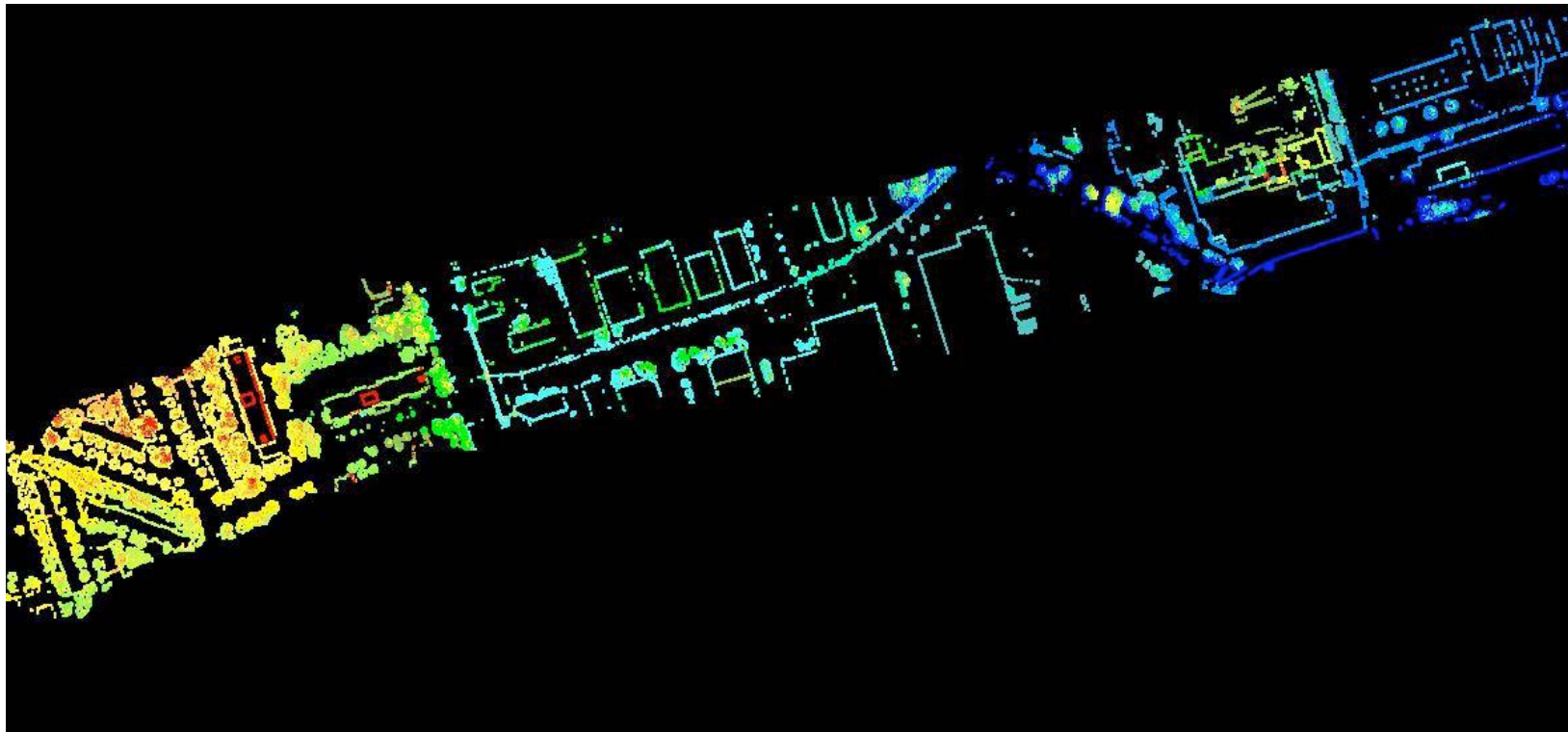
# Raw Image – depth is colour coded

---



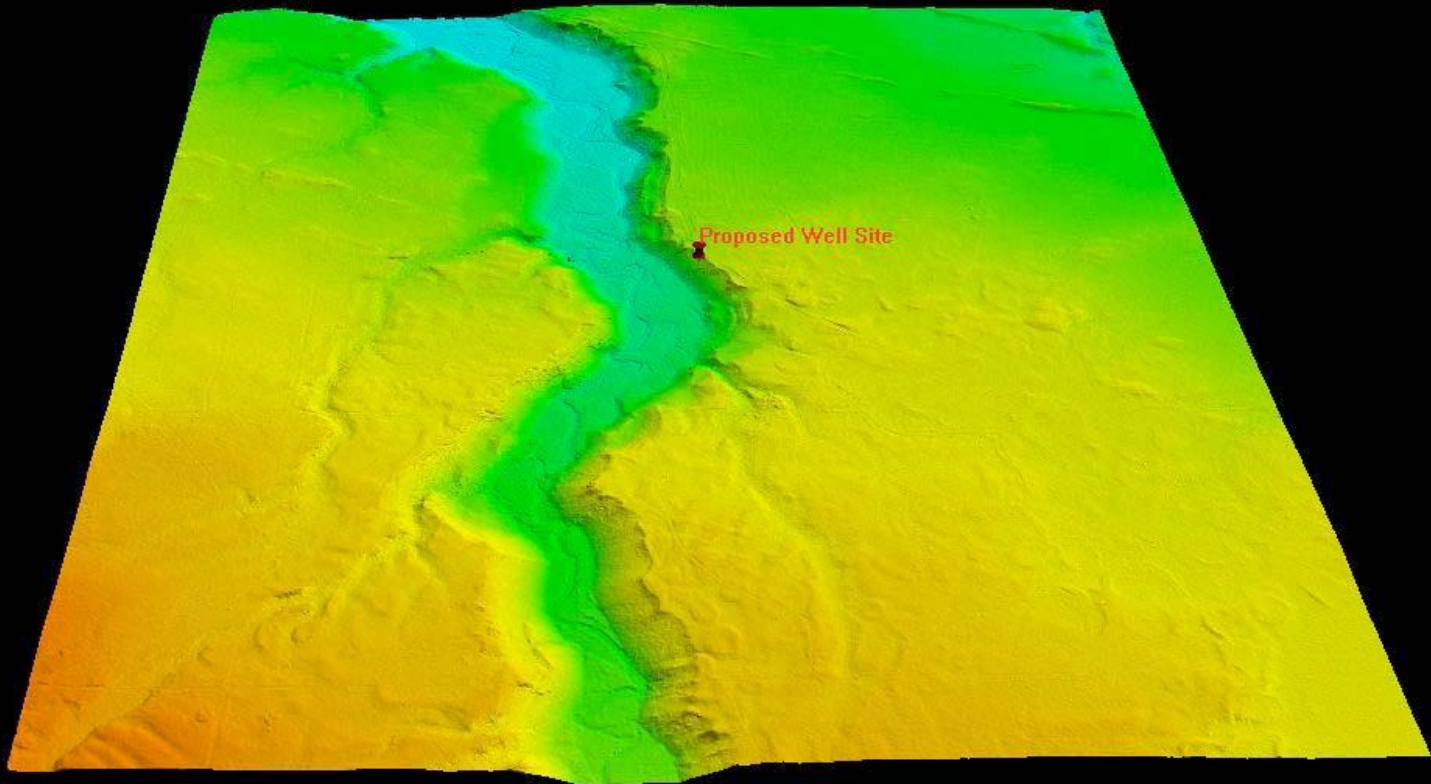
# Building, outlines, trees and wires

---



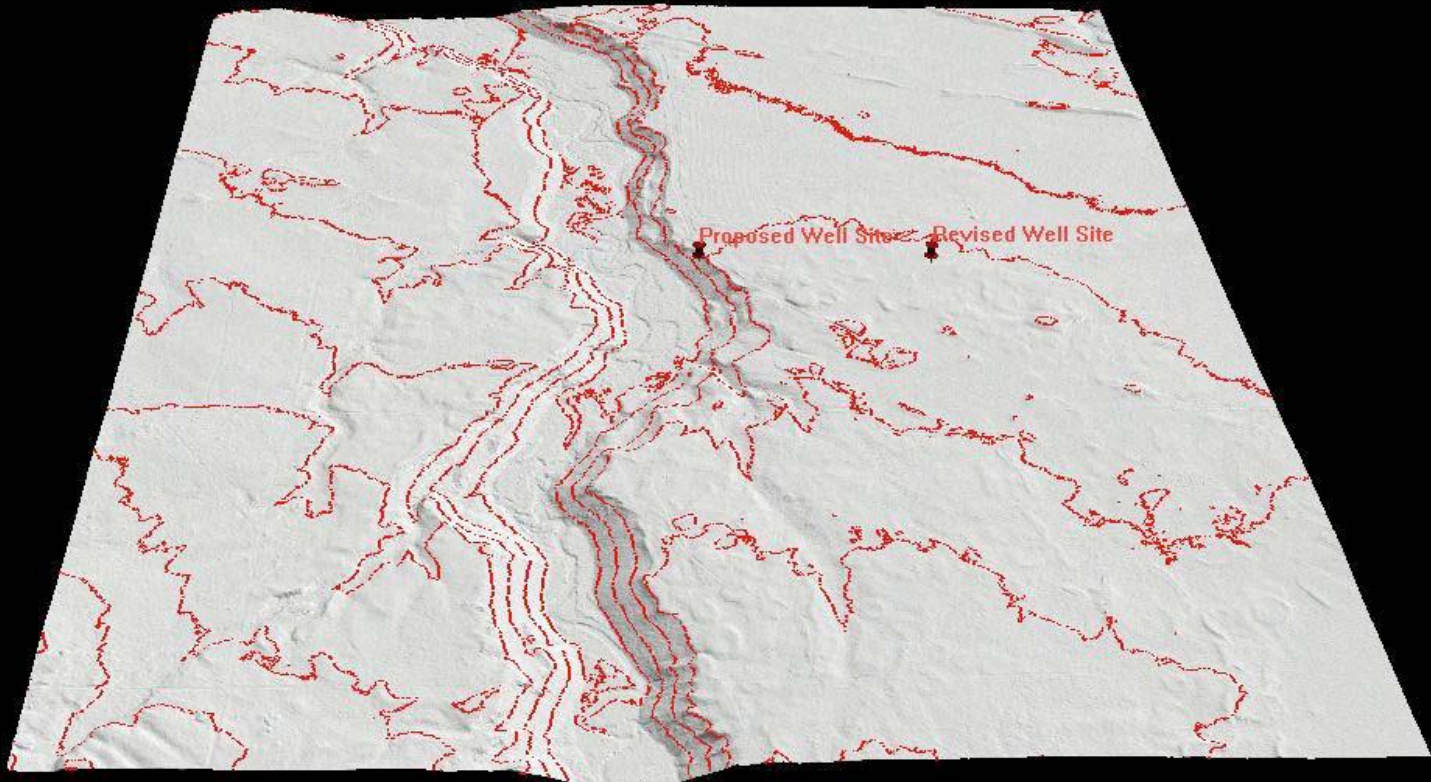
# Bare Earth Model

---



# Removing the trees

---





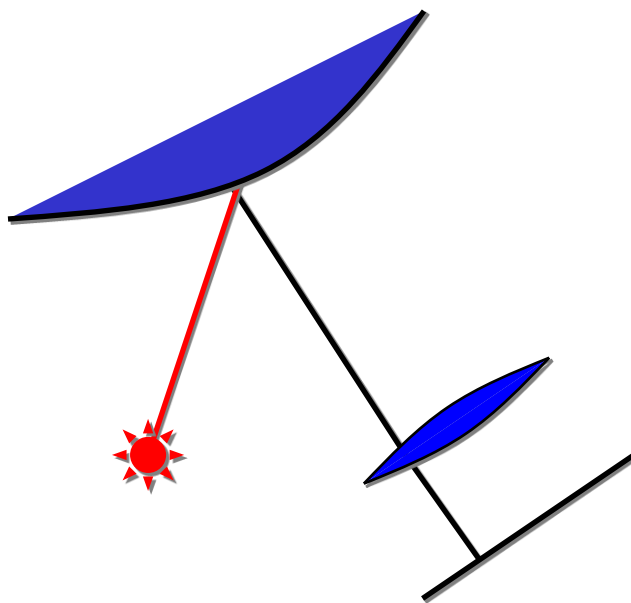
# Triangulation

---

Like stereo, but one camera 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 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

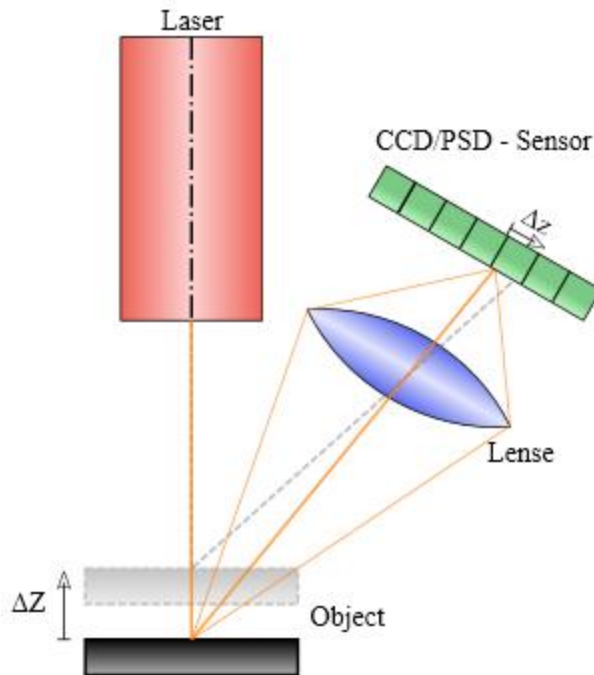
Very easy to build, and to make it work!

# Moving Laser Spot with Mirrors

---

Move the laser spot over the scene with a spinning mirror

Complex, but can gather thousands of points per second very accurately (NRC pioneers!)



# Triangulation can be very accurate

---

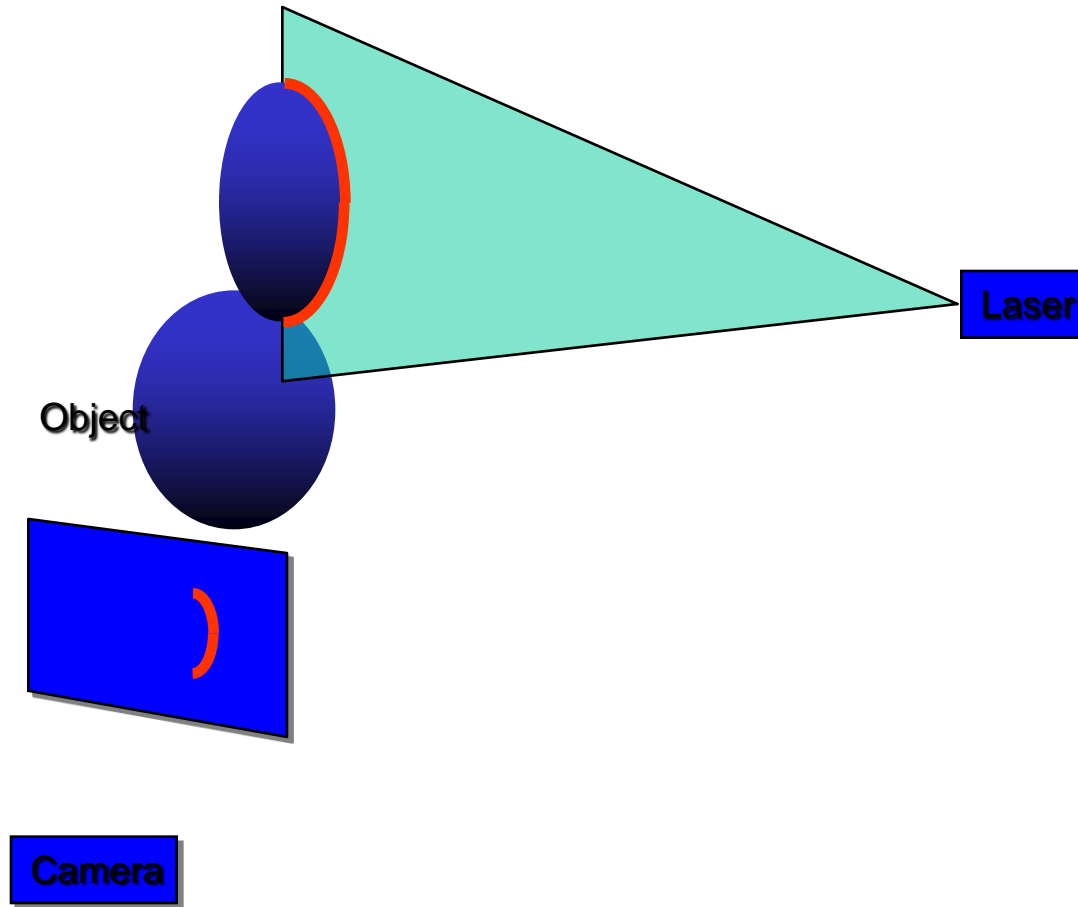
Can get accuracy down to 20 microns  
(1/50<sup>th</sup> of a millimeter!)



# Spreading out the spot to a line

---

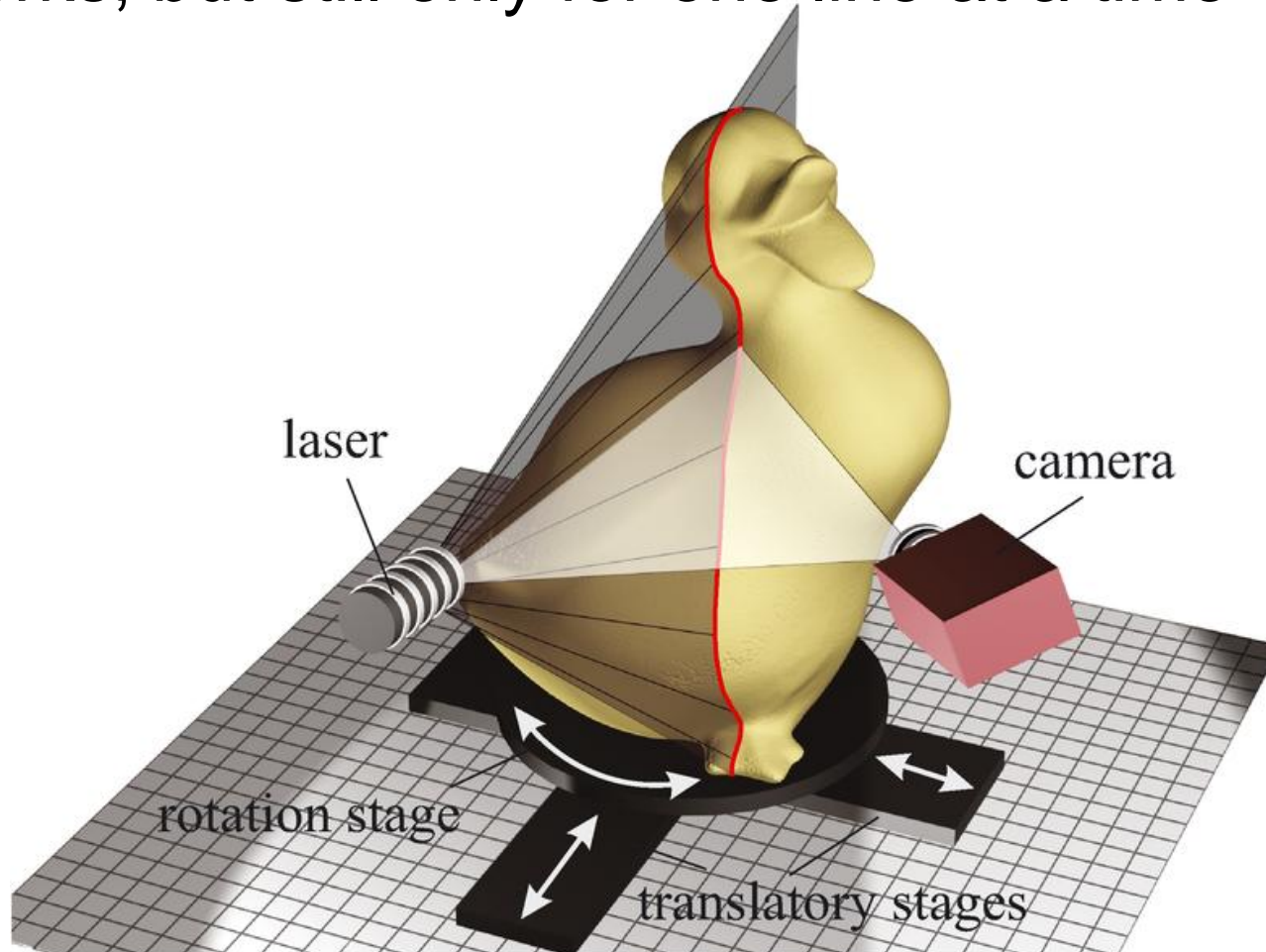
Project a stripe, not a dot – very easy and cheap



# Depth from a laser line (structured light)

---

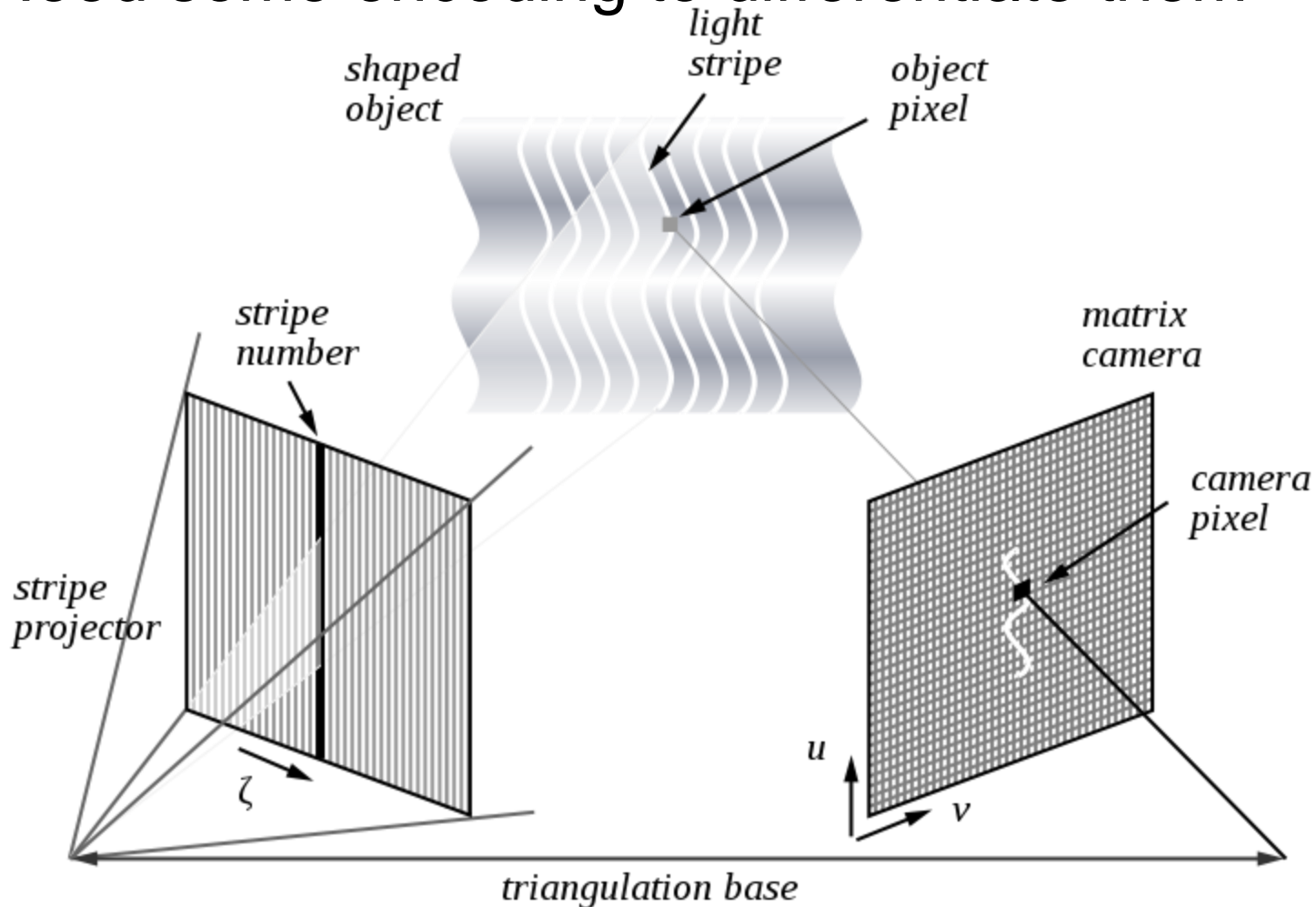
Process each horizontal line in the camera  
Works, but still only for one line at a time



# Depth from multi-stripe projector

Can use ordinary projector for multiple stripes

Need some encoding to differentiate them



# Multi-Stripe Triangulation

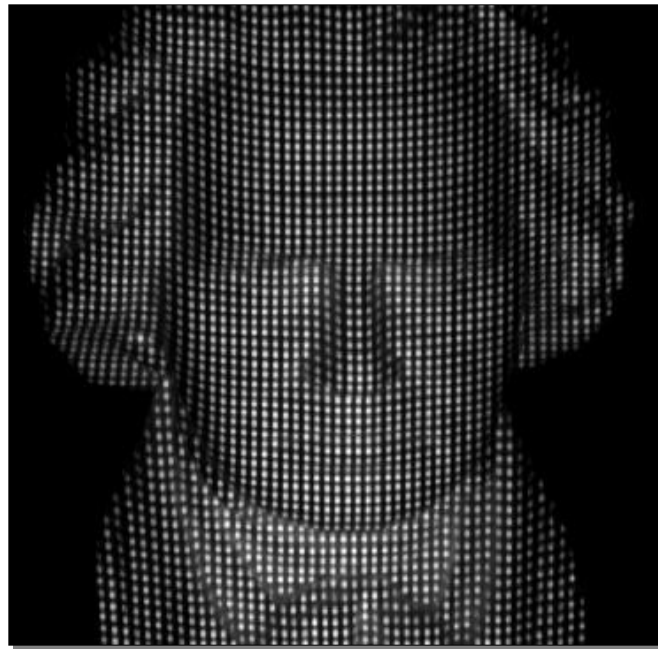
---

To go faster, project multiple stripes

Also commonly called structured light

But which stripe is which?

Answer #1: assume surface continuity





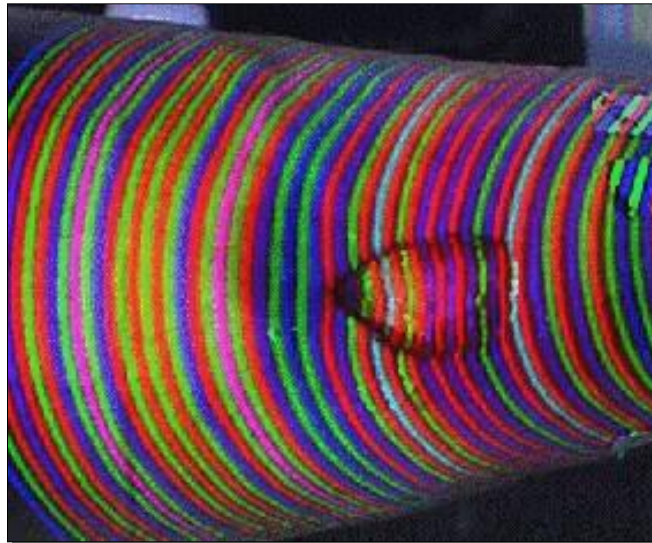
# Multi-Stripe Triangulation

---

To go faster, project multiple stripes

But which stripe is which?

Answer #2: colored stripes (or dots)



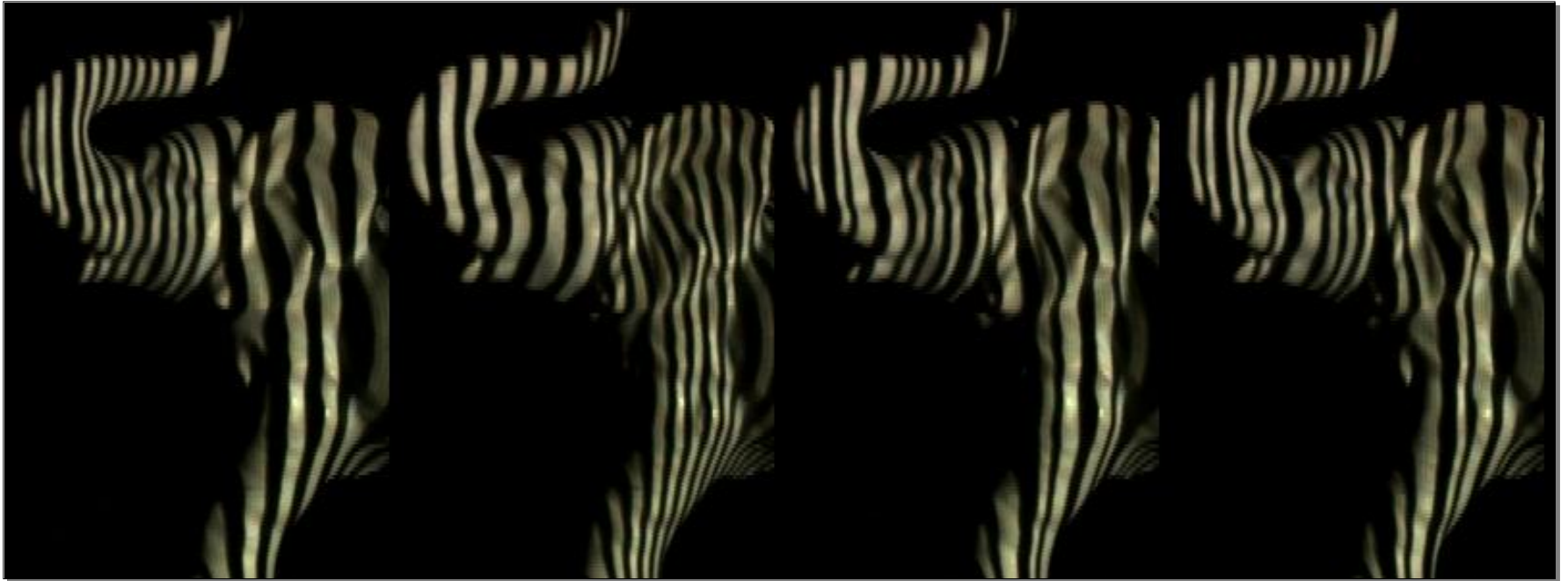
# Multi-Stripe Triangulation

---

To go faster, project multiple stripes

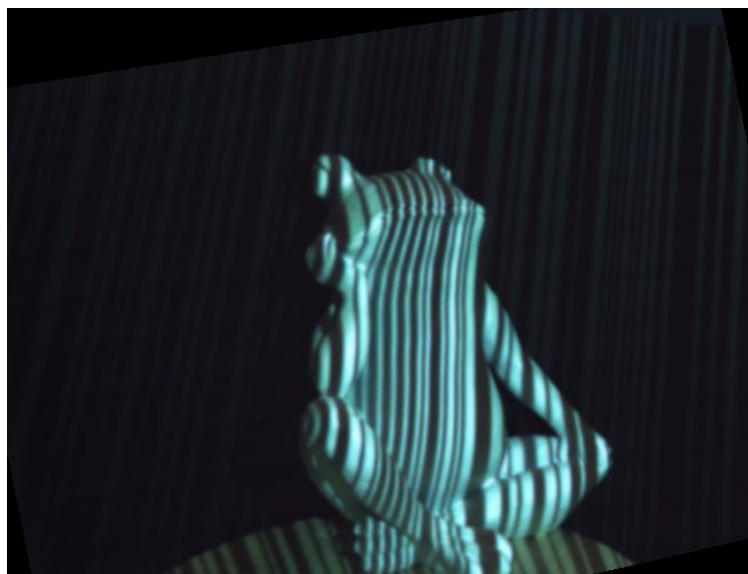
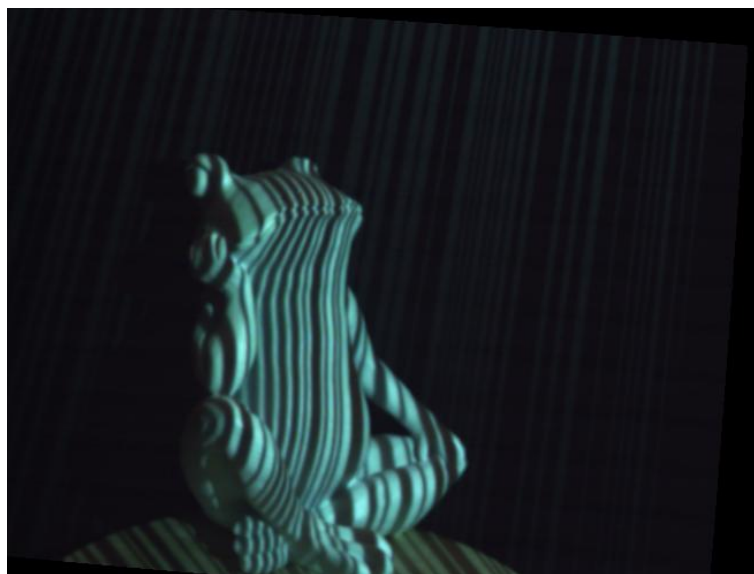
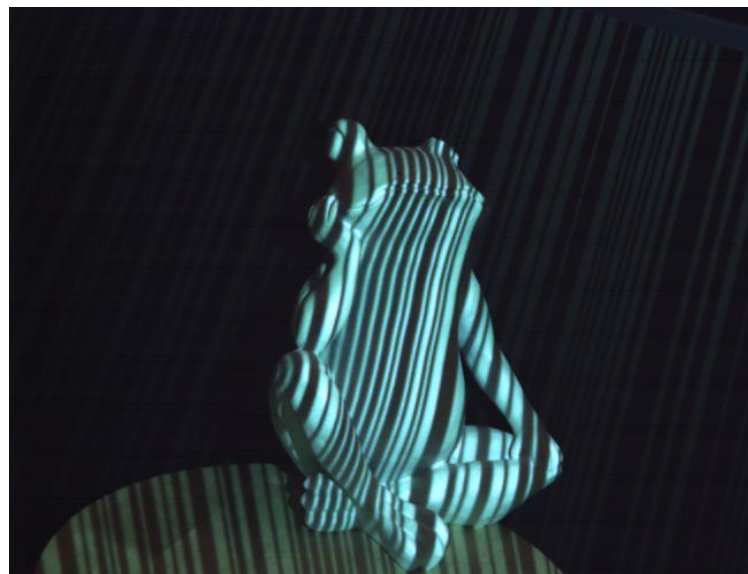
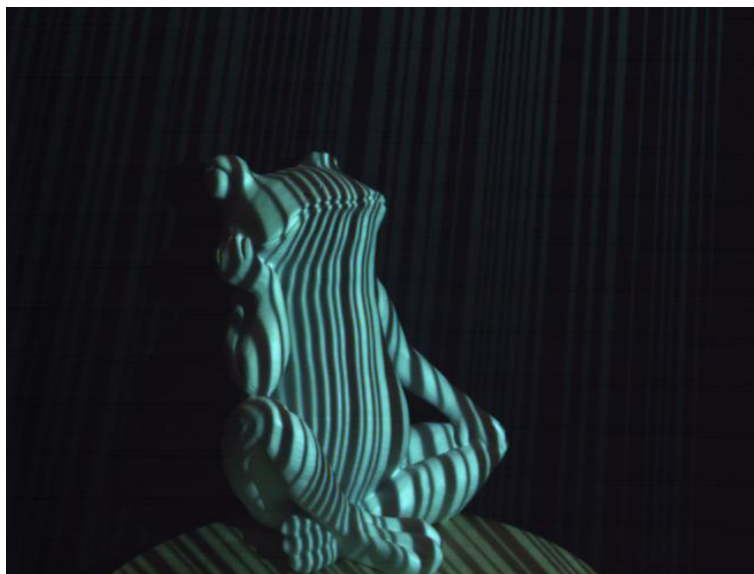
But which stripe is which?

Answer #3: time-coded stripes



# Active Stereo (Structured Light)

---

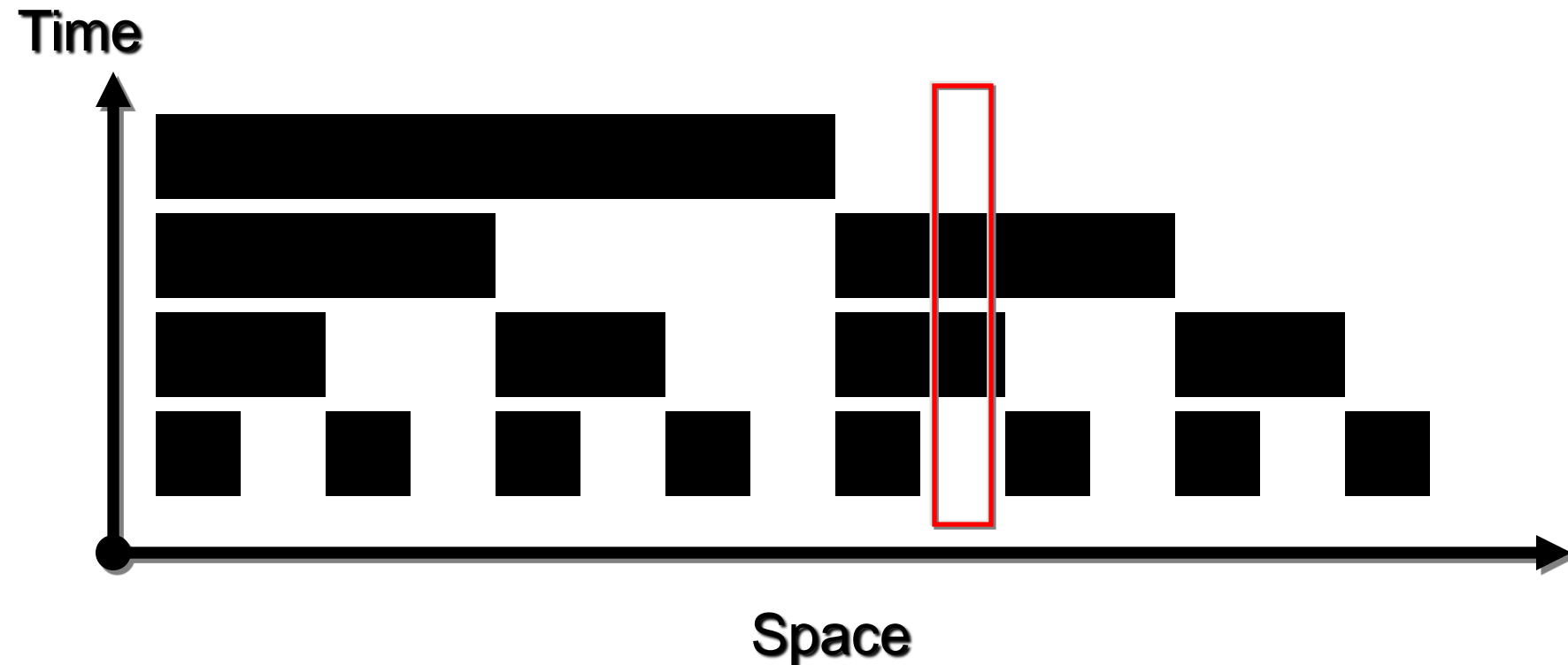


rectified

# Time-Coded Light Patterns

---

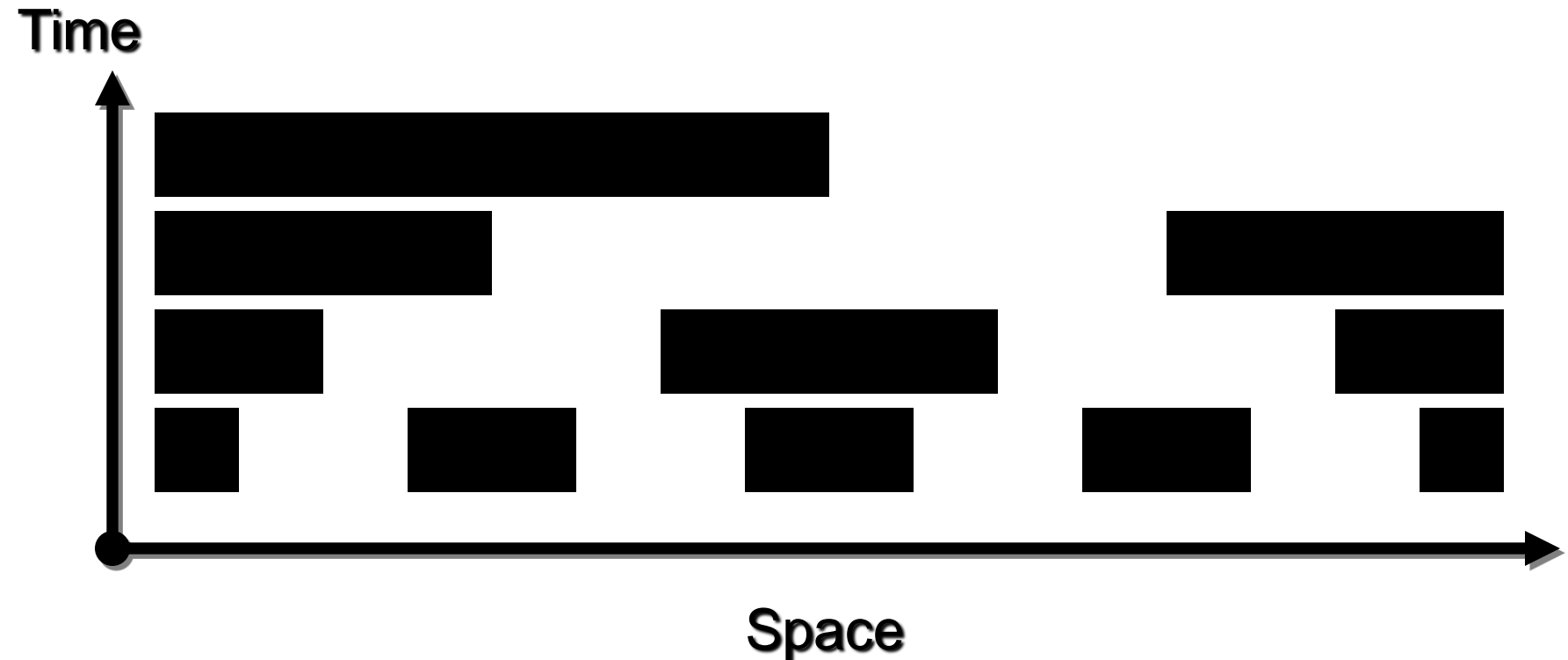
Assign each stripe a unique illumination code over time [Posdamer 82]



# Gray-Code Patterns

---

To minimize effects of quantization error:  
each point may be a boundary only once



# Microsoft Kinect

---

This camera is first mass produced 3d camera

What is the principle? Uses self identifying patterns of dots (like glyphs)

What are glyphs?

A local pattern that identifies itself uniquely


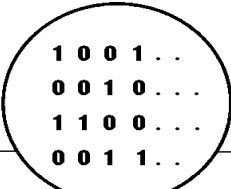
Qrcode



Augmented Reality Tags



# Glyphs printed in paper (Dataglyphs)



**CareCo**

*Personnel Services*  
1637 Bear Vista Blvd  
Wholesale, 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.

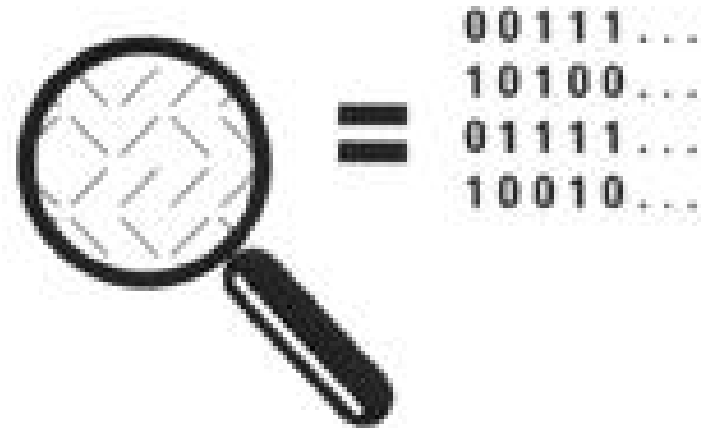
*Cooper M Price*

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 glyphs

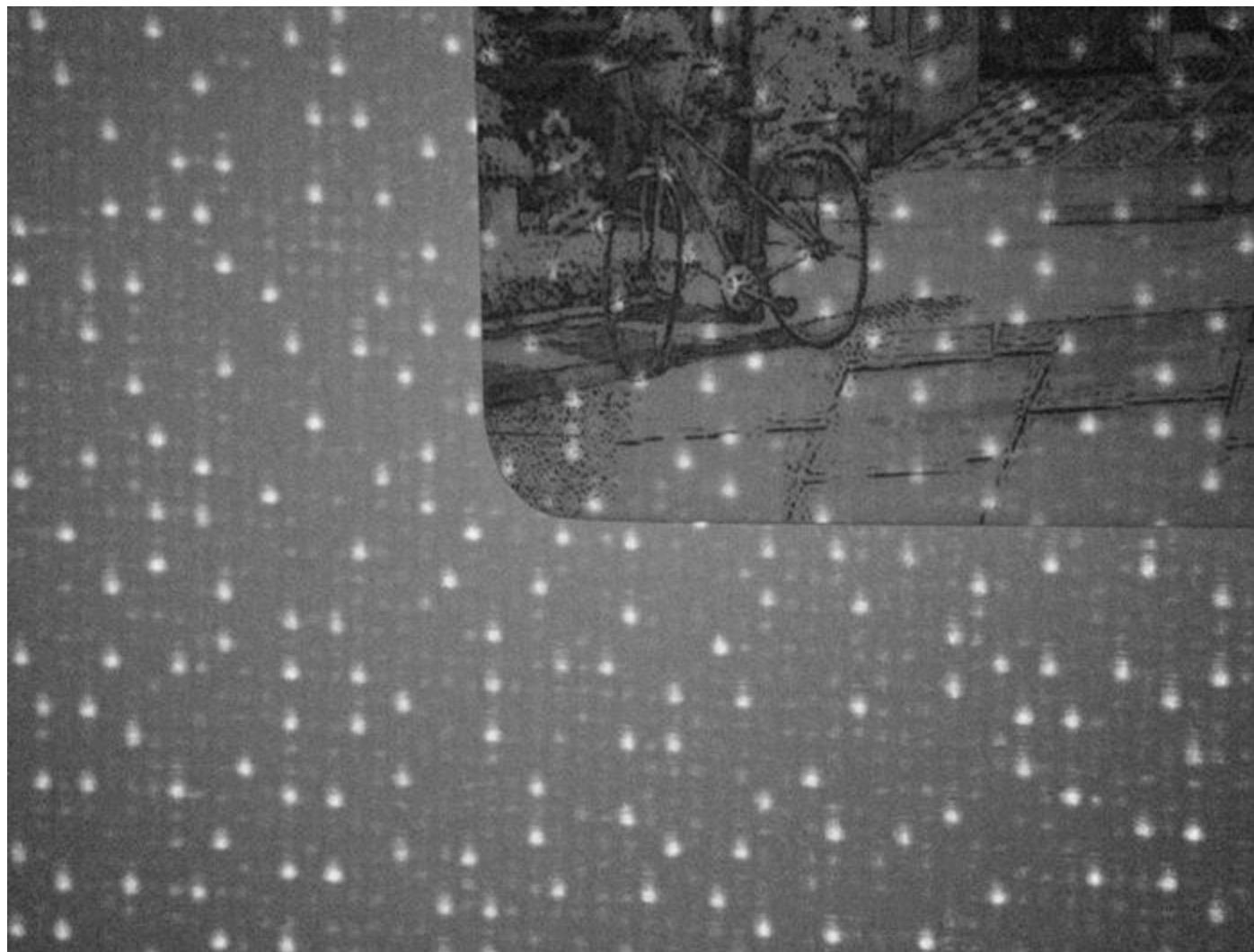
---





# Kinect Projects glyphs

---

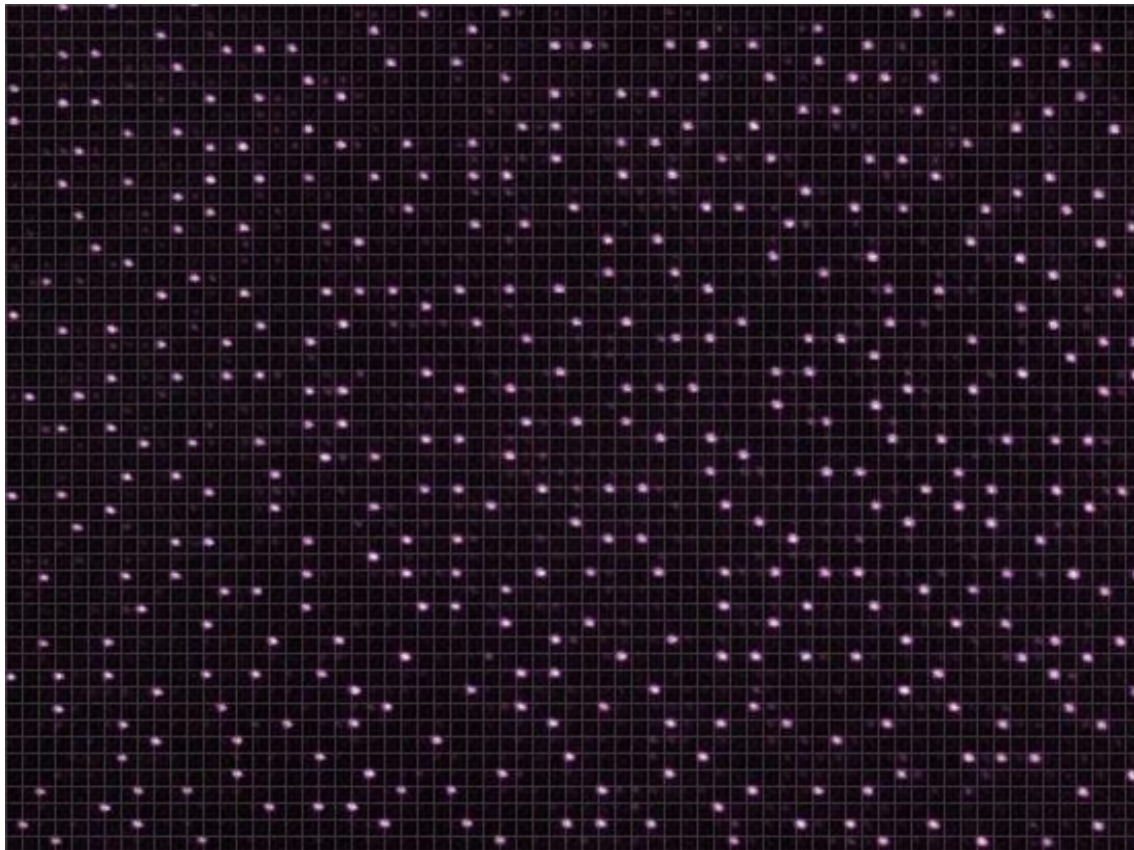


# Kinect Triangulation

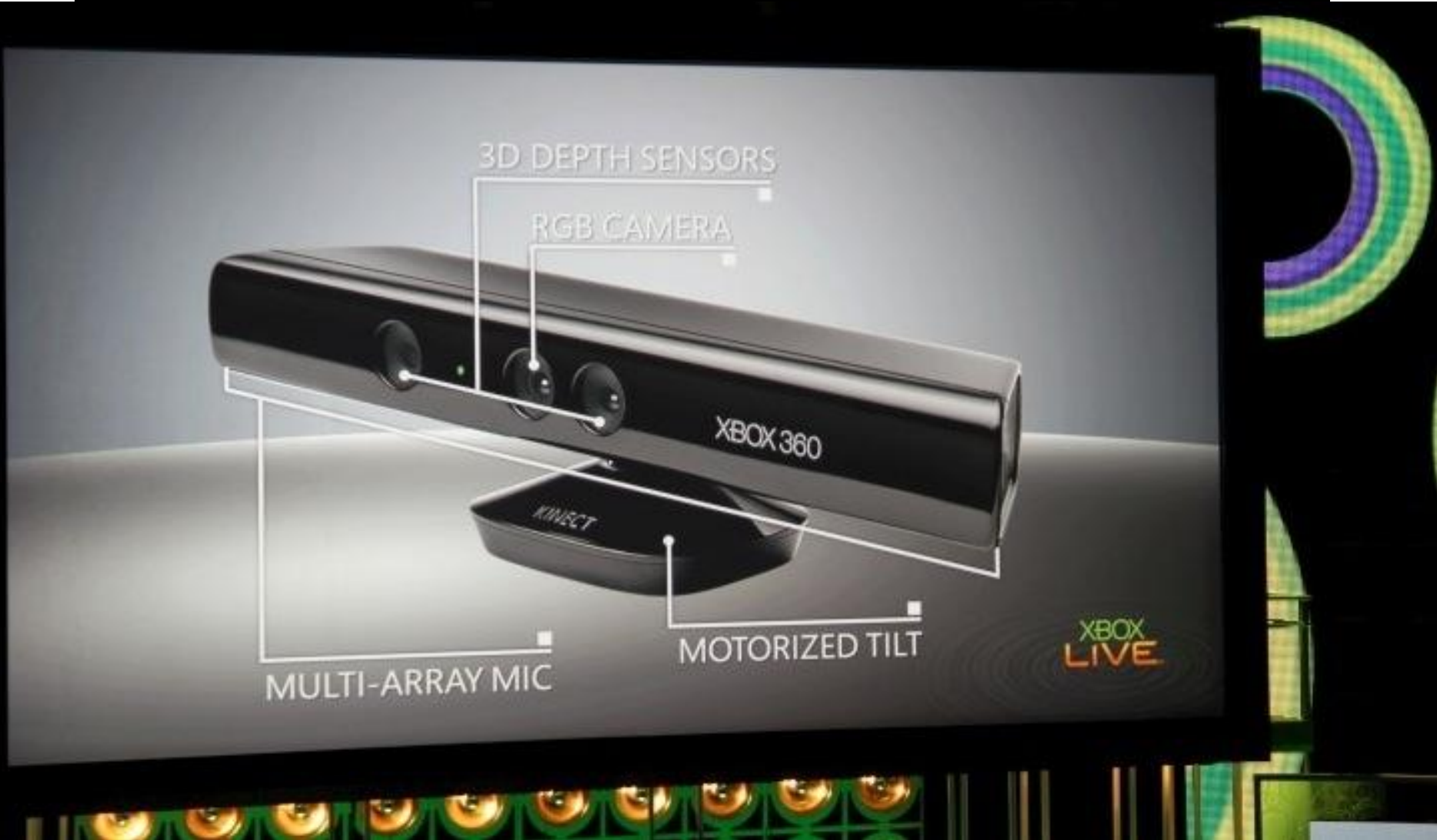
---

Local pattern identifies location of projection

By simple processing we find local identifier



# Kinect Hardware



# Kinect Hardware

---

There is a projector for the laser dots and a sensor just for these dots (infrared)

Since we can recognize the glyph in the infrared image we can triangulate to find the depth

This requires a prior calibration process so that we know the rays for the laser dots

But it is 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 overlaid intensity image

# Kinect Demo

---

There is a plethora of open source software where you can program a Kinect

In real-time display a depth image as you would get from a real-time stereo camera

Usually the depth is colour coded, or displayed in a way that colour relates to the depth value

I will now show you a demonstration!

# 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

Use this range value to register (align) the overlapping range images

Repeat this enough times you get one big model

# Depth better than intensity?

---

Is it easier to use a Kinect or an ordinary digital camera to make 3d models

Using a Kinect (or a stereo camera) is much easier because the depth accuracy from the Kinect does not change as you move camera

Depth accuracy depends on baseline alone

With an ordinary intensity image 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 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