

Autofocus (AF)

CS 178, Spring 2013

ⓧ
Begun 4/18/13, finished 4/25.



Marc Levoy
Computer Science Department
Stanford University

Outline

- ◆ viewfinders and manual focusing
 - ◆ view cameras and tilt-shift lenses
 - ◆ passive autofocus
 - phase detection
 - contrast detection
 - ◆ autofocus modes
 - ◆ lens actuators
-
- ◆ active autofocus
 - time-of-flight
 - triangulation

Single lens reflex (SLR) with autofocus



Nikon F4
(film camera)

Viewfinder coverage & magnification

- coverage is fraction of sensor image that is covered by the viewfinder, i.e.

$$\frac{h_1}{h_2} \quad \text{or sometimes} \quad \frac{\text{area}_1}{\text{area}_2}$$

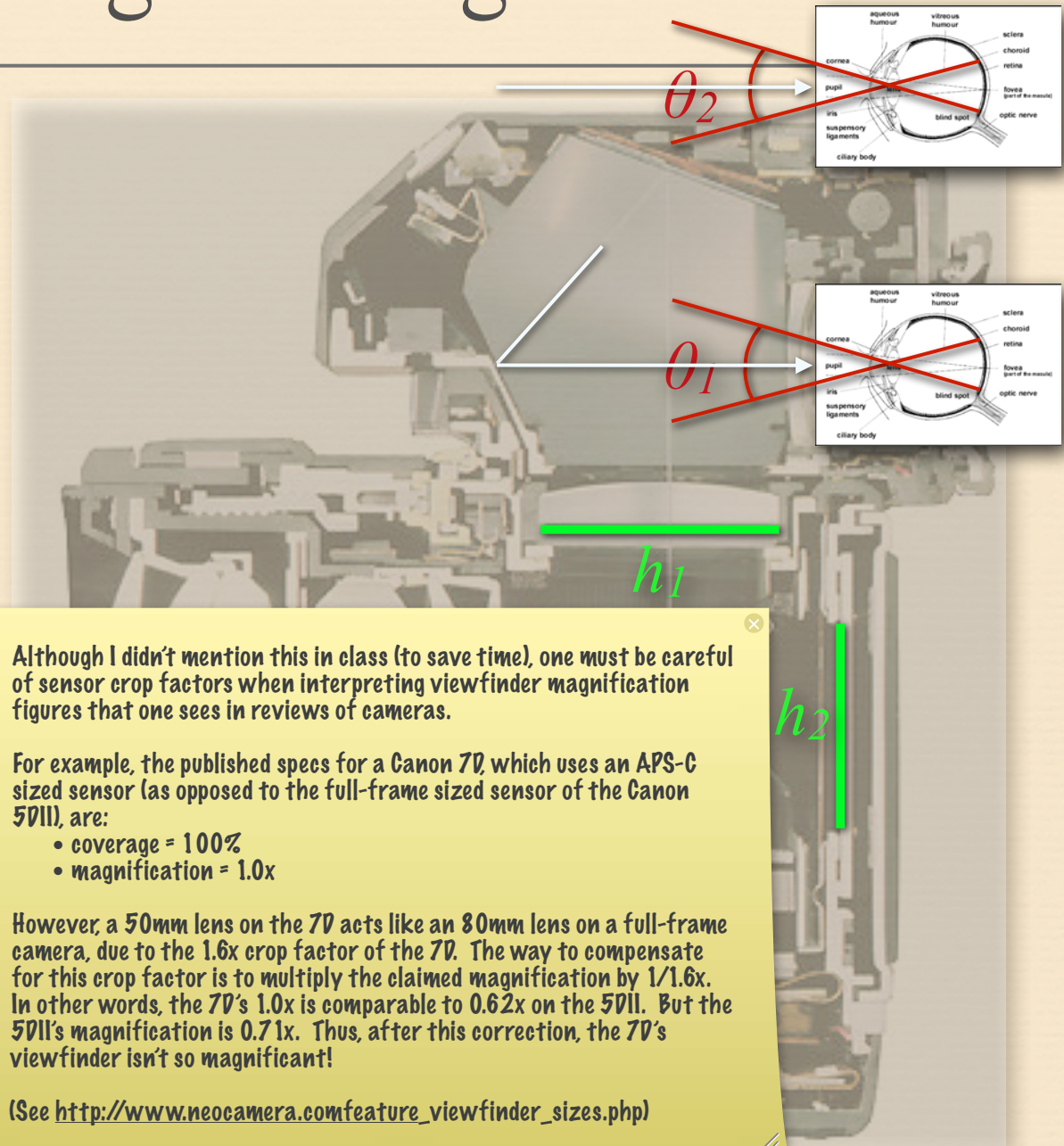
- magnification is apparent size of objects in viewfinder relative to unaided eye, i.e.

$$\frac{\theta_1}{\theta_2} \approx \frac{\tan \theta_1}{\tan \theta_2}$$

with a 50mm lens on camera

- example: Canon 5D II

- coverage = 98%
- magnification = 0.71x



Although I didn't mention this in class (to save time), one must be careful of sensor crop factors when interpreting viewfinder magnification figures that one sees in reviews of cameras.

For example, the published specs for a Canon 7D, which uses an APS-C sized sensor (as opposed to the full-frame sized sensor of the Canon 5DII), are:

- coverage = 100%
- magnification = 1.0x

However, a 50mm lens on the 7D acts like an 80mm lens on a full-frame camera, due to the 1.6x crop factor of the 7D. The way to compensate for this crop factor is to multiply the claimed magnification by 1/1.6x. In other words, the 7D's 1.0x is comparable to 0.62x on the 5DII. But the 5DII's magnification is 0.71x. Thus, after this correction, the 7D's viewfinder isn't so magnificent!

(See http://www.neocamera.com/feature_viewfinder_sizes.php)

Electronic viewfinders



point-and-shoot



SLR “live view”



electronic viewfinder

◆ pros

- same view as lens without need for reflex mirror
- can tone map to show effect of chosen exposure

◆ cons

- poor resolution and low dynamic range relative to optical
- Is the display being tone mapped? Will the shot look like this?

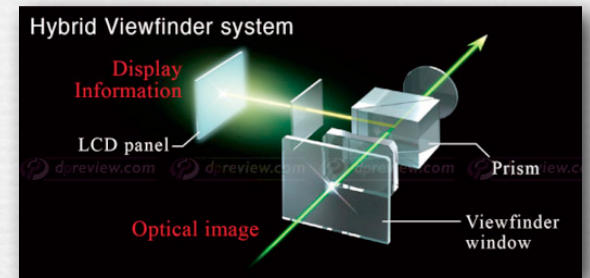
Non-thru-the-lens optical viewfinder



point-and-shoot



Fuji X100 (new) and its hybrid viewfinder



◆ pro

- infinite resolution and dynamic range, like SLR viewfinder
- retro (cool!)

◆ con

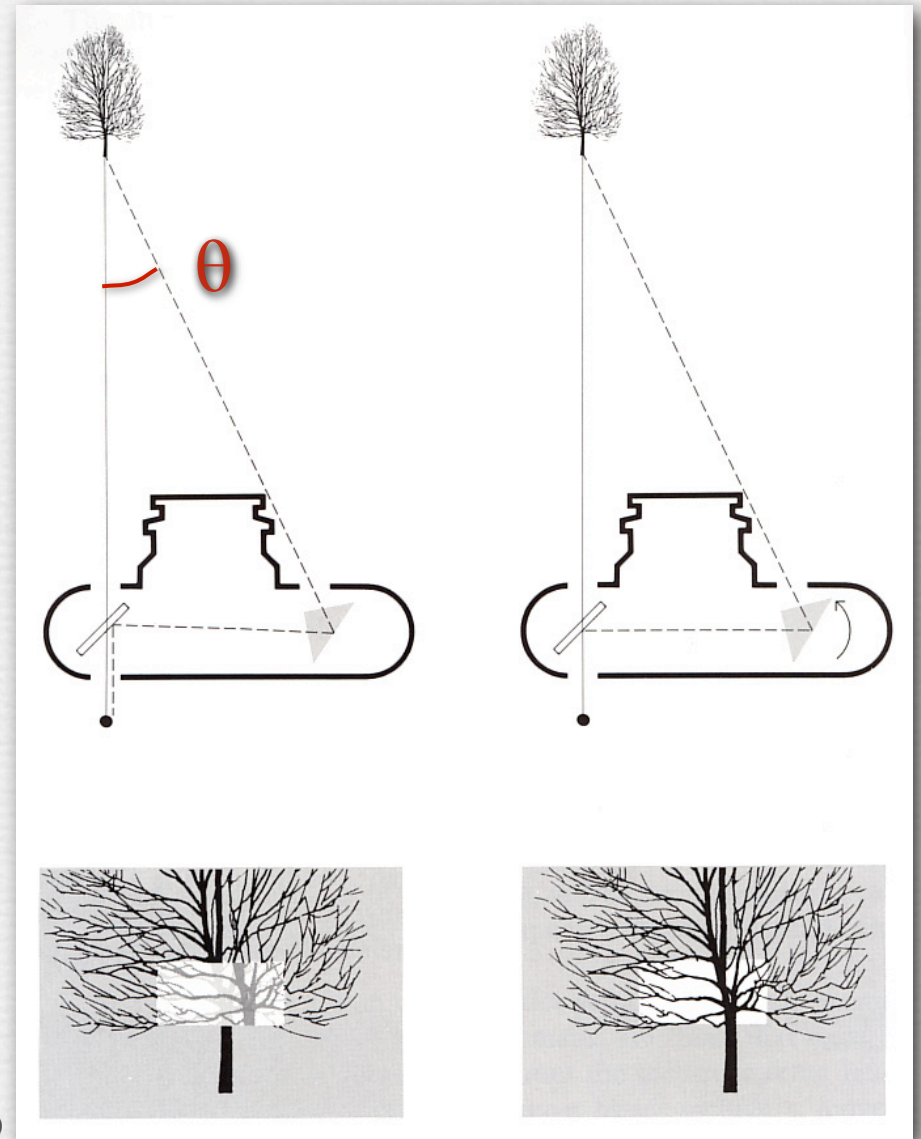
- different perspective view than main lens sees
- low magnification (appears small to eye)

Manual rangefinder

- ◆ accurate
- ◆ painstaking
- ◆ different perspective view than main lens sees
- ◆ triangulation concept widely applicable



Leica M9
(digital full-frame)



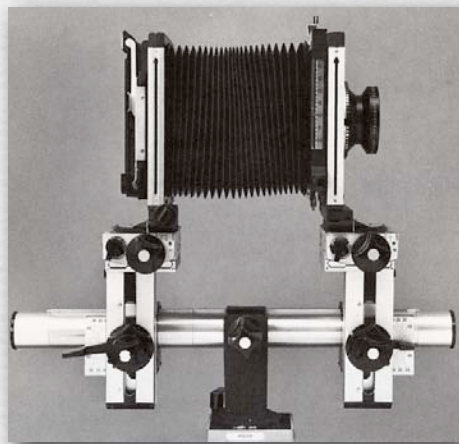
(Adams)

View camera with focusing screen

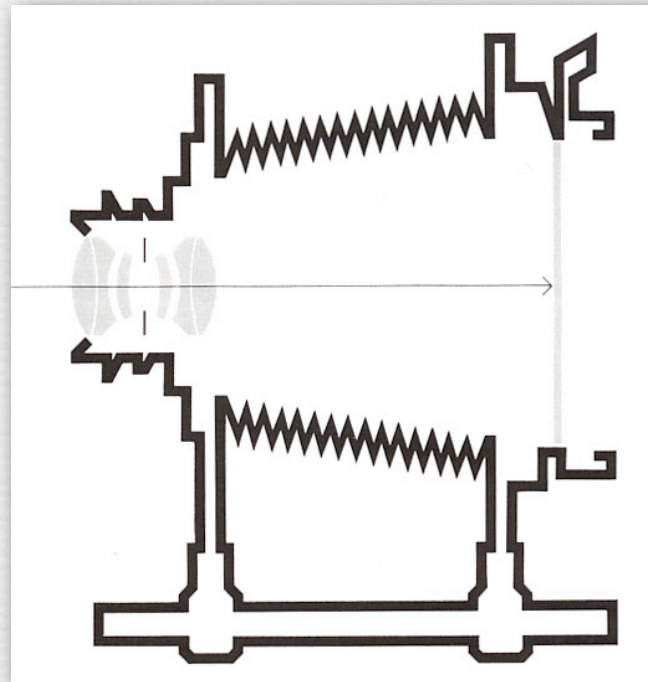
Sinar with digital back



- ◆ large format: 4×5" or 8×10"
 - film or scanned digital
- ◆ ground glass focusing screen
 - dim
 - hard to focus
 - inverted image



Sinar 4×5



(Adams)



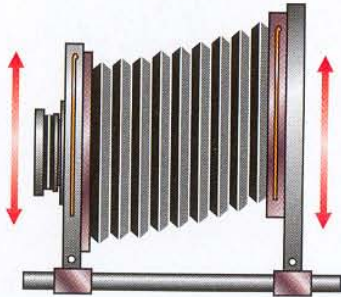
© QT Luong / terragalleria.com

#kica39809

View camera movements

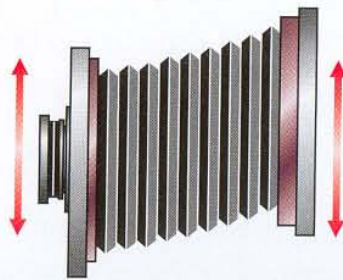
VIEW CAMERA MOVEMENTS

Side View



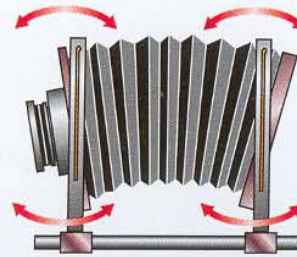
Rise and fall move the front or back of the camera in a flat plane, like opening or closing an ordinary window. Rise moves the front or back up; fall moves the front or back down.

Top View



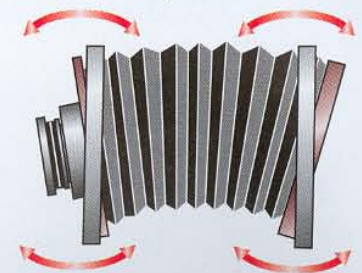
Shift (like rise and fall) also moves the front or back of the camera in a flat plane, but from side to side in a motion like moving a sliding door.

Side View



Tilt tips the front or back of the camera forward or backward around a horizontal axis. Nodding your head yes is a tilt of your face.

Top View



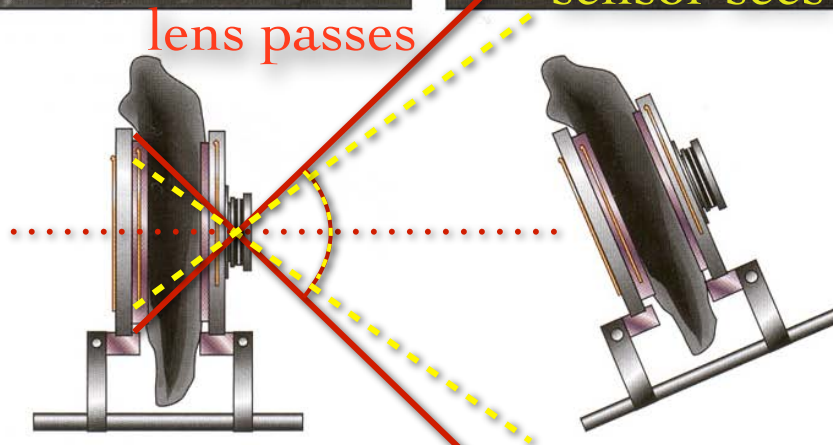
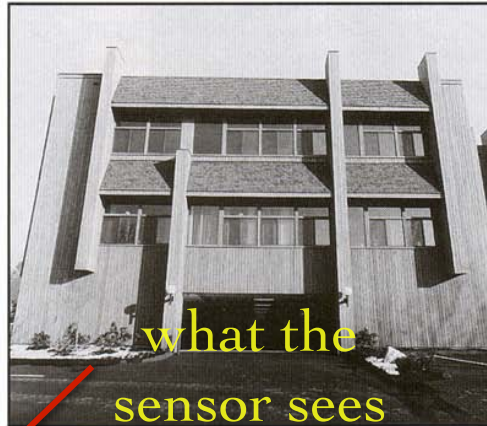
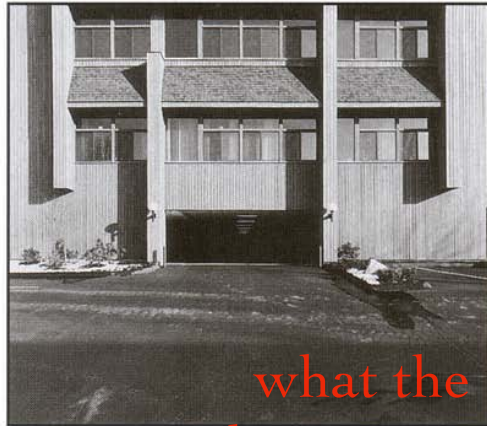
Swing twists the front or back of the camera around a vertical axis to the left or right. Shaking your head no is a swing of your face.

(London)

Off-axis perspective

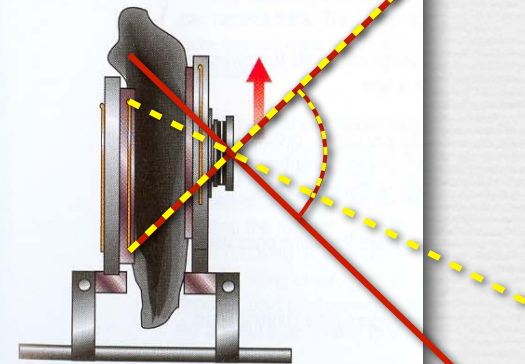
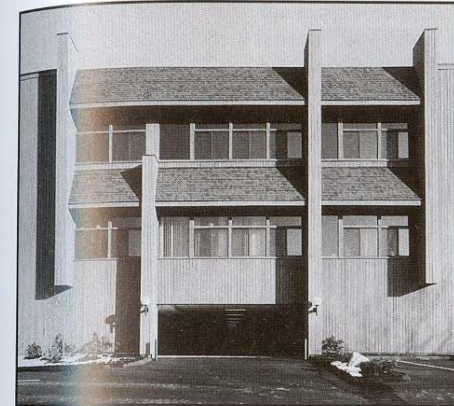
NOW AVAILABLE IN PHOTOSHOP !!

CONTROLLING CONVERGING LINES: THE KEYSTONE EFFECT



Standing at street level and shooting straight at a building produces too much street and too little building. Sometimes it is possible to move back far enough to show the entire building while keeping the camera level, but this adds even more foreground and usually something gets in the way.

Tilting the whole camera up shows the entire building but distorts its shape. Since the top is farther from the camera than the bottom, it appears smaller; the vertical lines of the building seem to be coming closer together, or converging, near the top. This is named the keystone effect, after the wedge-shaped stone at the top of an arch. This convergence gives the illusion that the building is falling backward—an effect particularly noticeable when only one side of the building is visible.

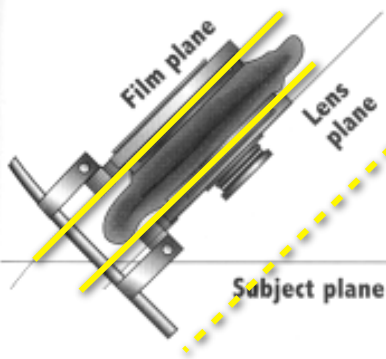


To straighten up the converging vertical lines, keep the camera back parallel to the face of the building. To keep the face of the building in focus, make sure the lens is parallel to the camera back. One way to do this is to level the camera and then use the rising front or falling back movements or both.

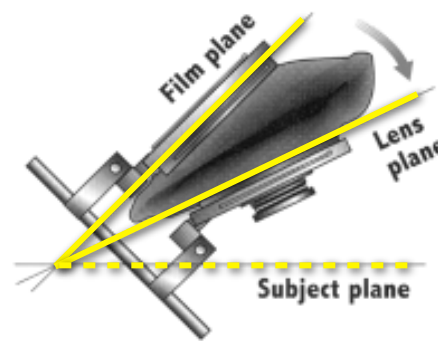
Another solution is to point the camera upward toward the top of the building, then use the tilting movements—first to tilt the back to a vertical position (which squares the shape of the building), then to tilt the lens so it is parallel to the camera back (which brings the face of the building into focus). The lens and film will end up in the same positions with both methods.

Tilted focal plane

ADJUSTING THE PLANE OF FOCUS TO MAKE THE ENTIRE SCENE SHARP



The book is partly out of focus because the lens plane and the film plane are not parallel to the subject plane. Instead of a regular accordion bellows, the diagrams show a bag bellows that can bring camera front and back closer together for use with a short focal-length lens.



Tilting the front of the camera forward brings the entire page into sharp focus. The camera diagram illustrates the Scheimpflug principle, explained at right.

- Scheimpflug condition

(London)

◆ cannot be done after the photograph is taken



Ansel Adams, Railroad Tracks



Ansel Adams, Monument Valley

Tilt-shift lenses



Canon TS-E
90mm lens



Tilt-shift lenses



Canon TS-E
90mm lens



The “miniature model” look



Canon TS-E
24mm II



- ◆ simulates a macro lens with a shallow depth of field, hence makes any scene look like a miniature model

The “miniature model” look



Canon TS-E
24mm II



- ◆ simulates a macro lens with a shallow depth of field, hence makes any scene look like a miniature model

Not a tilt-shift lens

(<http://www.tiltshiftphotography.net/>)

- ◆ gradient blur in Photoshop



Not a tilt-shift lens

(<http://www.tiltshiftphotography.net/>)

- ◆ gradient blur in Photoshop



original

Not a tilt-shift lens

(<http://www.tiltshiftphotography.net/>)

- ◆ gradient blur in Photoshop



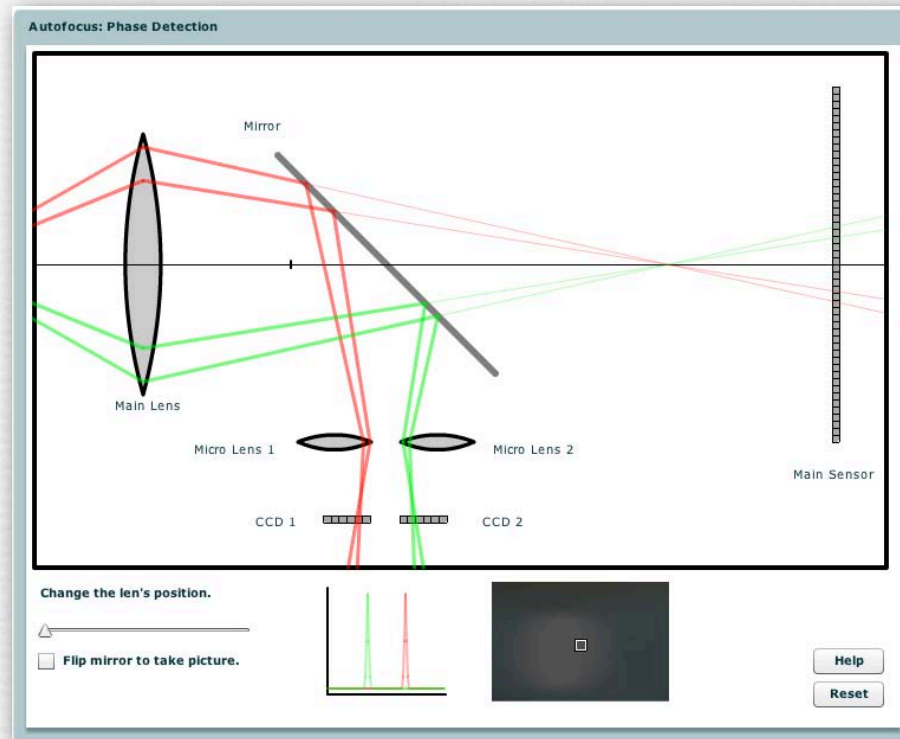
Q. Is this “fake” identical to the output of a real tilt-shift lens?

Recap

- ◆ the optical viewfinder in a single lens reflex camera (SLR)
 - gives you the same perspective as your final photograph
 - has high resolution and no limit on dynamic range
 - can't tone map to show effect of exposure
- ◆ view cameras let you eliminate vanishing points
 - this can alternatively be done in Photoshop
- ◆ view cameras also let you tilt the focal plane
 - this cannot be done in Photoshop, although you can fake it
- ◆ tilt-shift lenses provide both functions for SLRs
 - can be faked by blurring in Photoshop, but it looks different

Questions?

Passive autofocus: phase detection

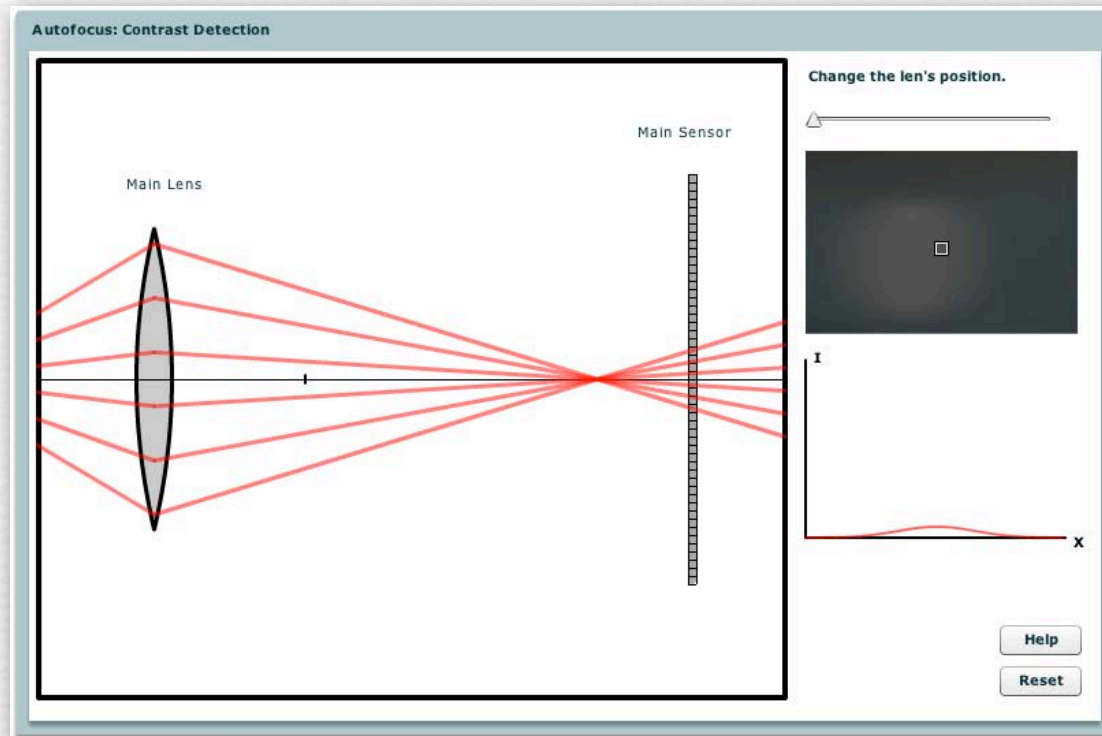


(FLASH DEMO)

<http://graphics.stanford.edu/courses/cs178/applets/autofocuspd.html>

- ◆ as the lens moves, ray bundles from an object converge to a different point in the camera and change in angle
- ◆ this change in angle causes them to refocus through two lenslets to different positions on a separate AF sensor
- ◆ a certain spacing (*disparity*) between these images is “in focus”

Passive autofocus: contrast detection

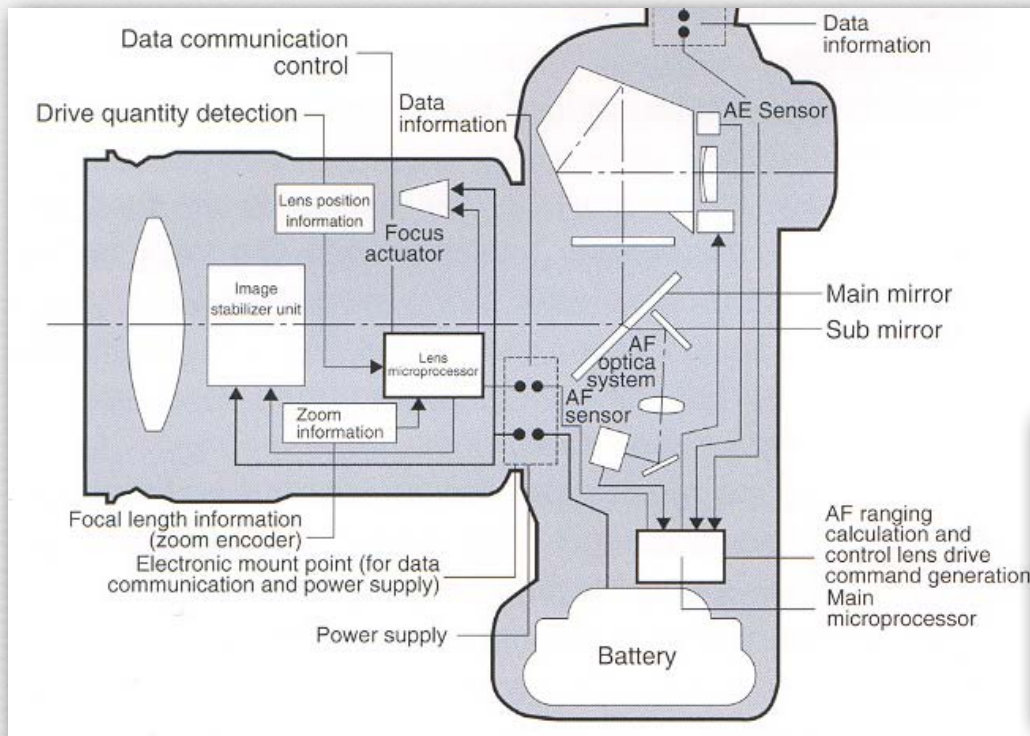


(FLASH DEMO)

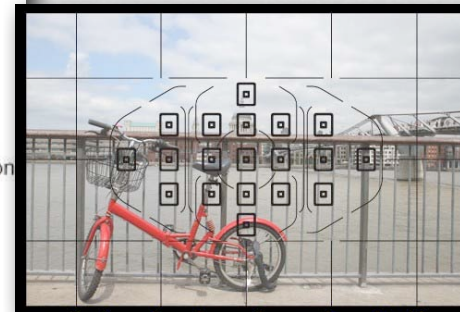
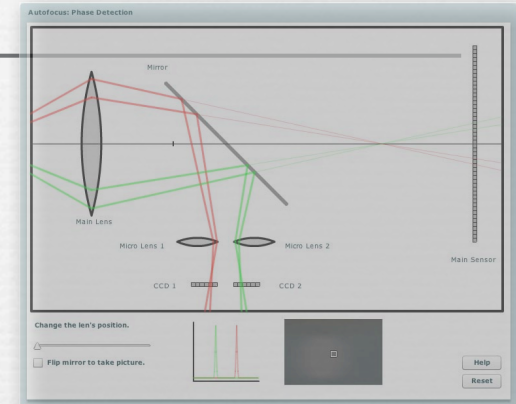
<http://graphics.stanford.edu/courses/cs178/applets/autofocuscd.html>

- ◆ an image sensor will see an object as contrasty if it's in focus, or of low contrast if it's not
- ◆ move the lens until the image falling on the sensor is contrasty
- ◆ compute contrasty-ness using local *gradient* of pixel values

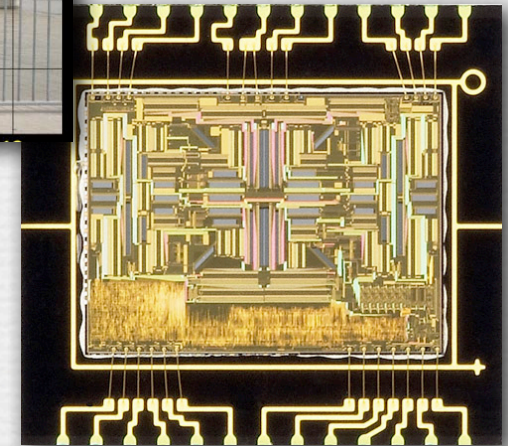
Most SLRs use phase detection



(Canon)

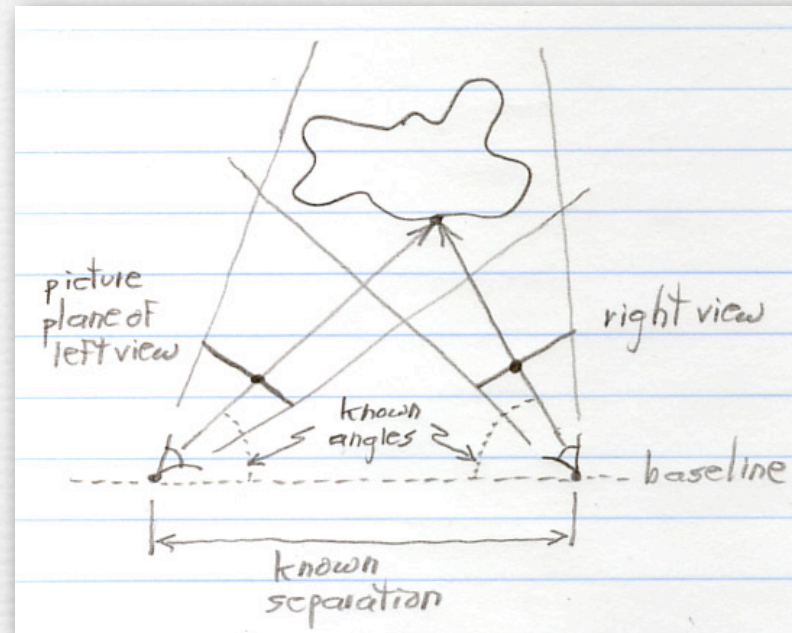


Canon 7D



- ◆ distance between subimages allows lens to jump directly into focus, without hunting
 - equivalent to *depth-from-stereo* in computer vision
- ◆ many AF points, manual or automatic way to choose among them
 - closest scene feature is often the best, but also consider position in FOV

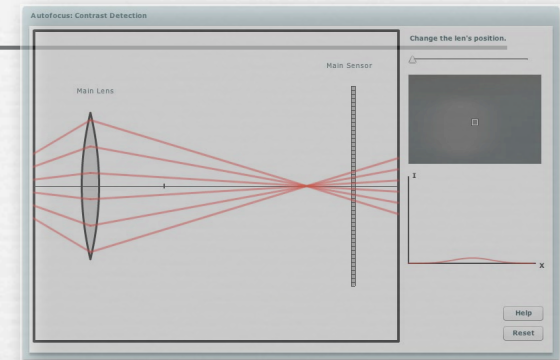
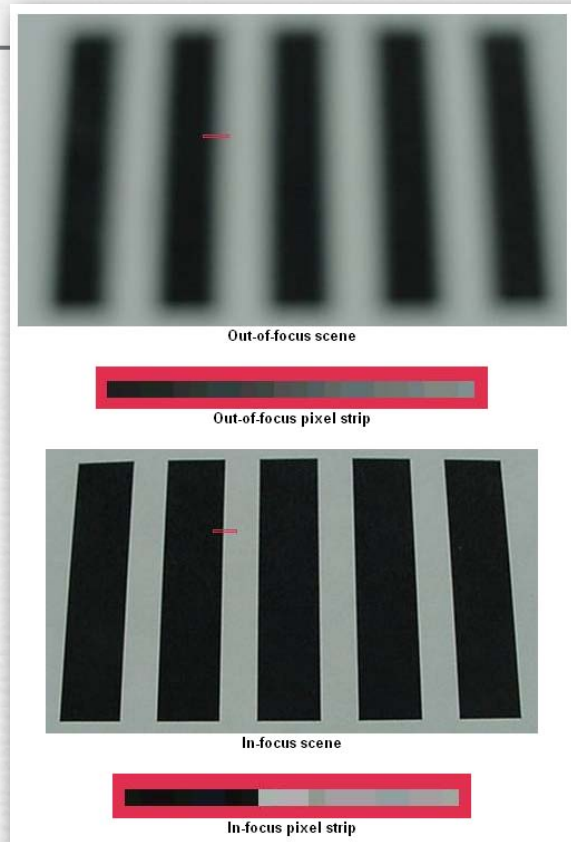
Phase detection is like depth from stereo (contents of whiteboard)



- ◆ start by finding corresponding features in two views of an object; express each correspondence as a pair of pixels (dots)
- ◆ knowing the separation and aim (angles) of the two cameras, the positions of these two pixels can be projected (arrows) until they intersect; this gives the 3D location of the feature
- ◆ the larger the separation, the more accurate the depth estimate

Most DSCs use contrast detection

(howstuffworks.com)



- ◆ requires repeated measurements as lens moves, measurements are captured using the main sensor
 - equivalent to *depth-from-focus* in computer vision
- ◆ slow, requires hunting, suffers from overshooting
 - it's ok if still cameras overshoot, but video cameras shouldn't

Autofocus modes

- ◆ AI servo (Canon) / Continuous servo (Nikon)
 - predictive tracking so focus doesn't lag axially moving objects
 - continues as long as shutter is pressed halfway

Servo focus can misfocus quickly



Canon 1D Mark III, 300mm f/2.8

Servo focus can misfocus quickly



Canon 1D Mark III, 300mm f/2.8

Servo focus can misfocus quickly



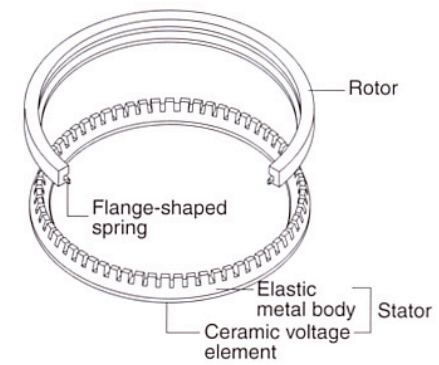
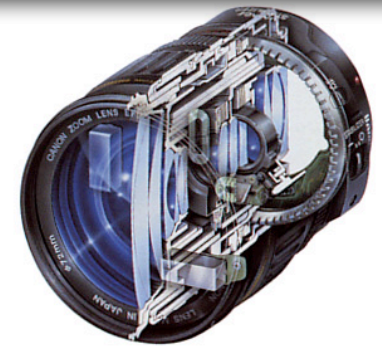
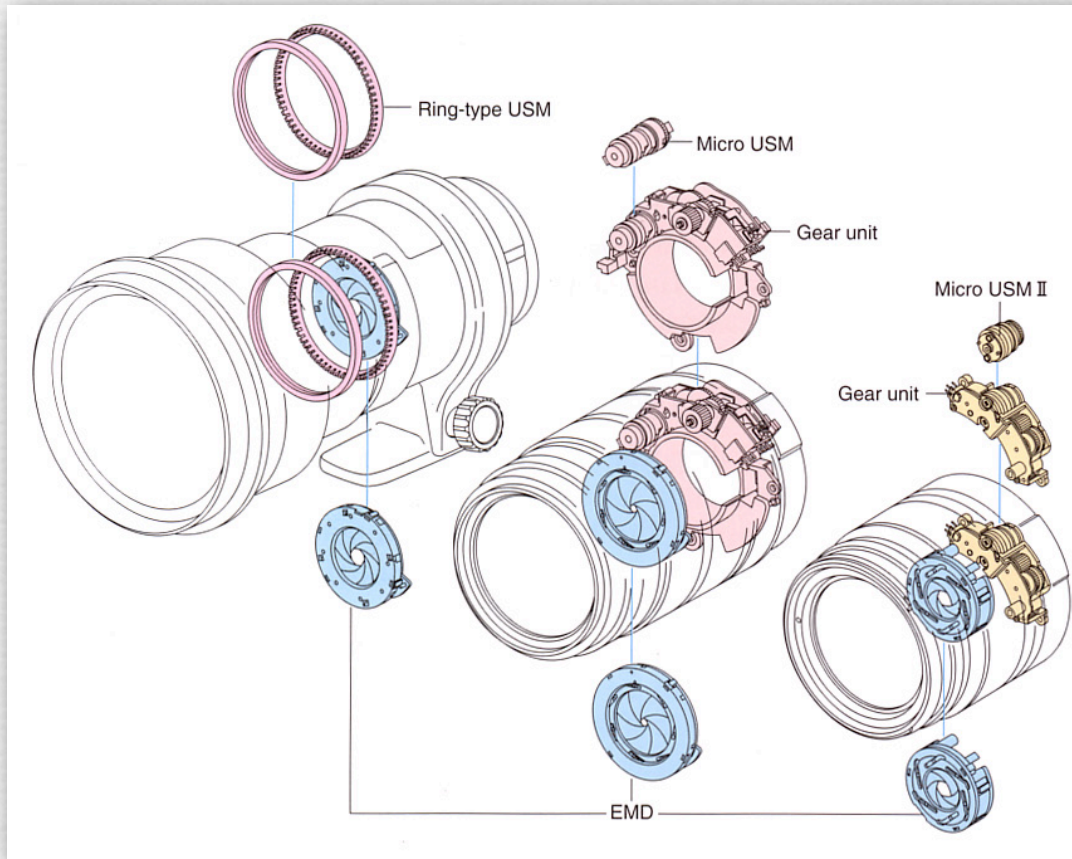
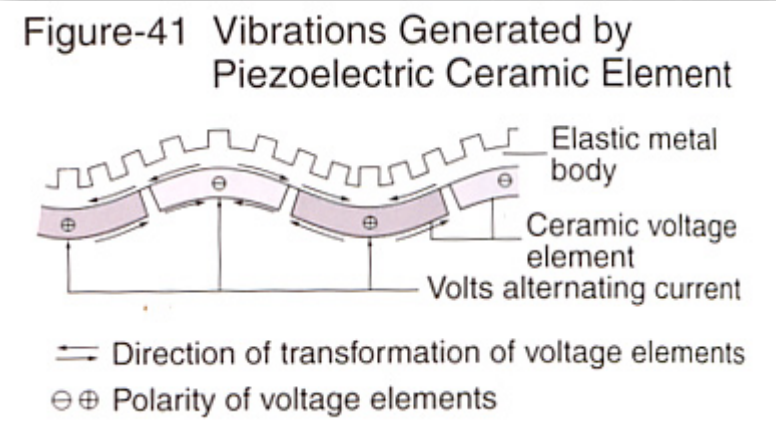
Canon 1D Mark III, 300mm f/2.8

Autofocus modes

- ◆ AI servo (Canon) / Continuous servo (Nikon)
 - predictive tracking so focus doesn't lag axially moving objects
 - continues as long as shutter is pressed halfway
- ◆ focusing versus metering
 - autofocus first, then meter on those points
- ◆ “trap focus”
 - trigger a shot if an object comes into focus (Nikon)
- ◆ depth of field focusing
 - find closest and furthest object; set focus and N accordingly
- ◆ overriding autofocus
 - manually triggered autofocus (AF-ON in Canon)
- ◆ all autofocus methods fail if object is textureless!

Lens actuators

◆ Canon ultrasonic motor (USM)



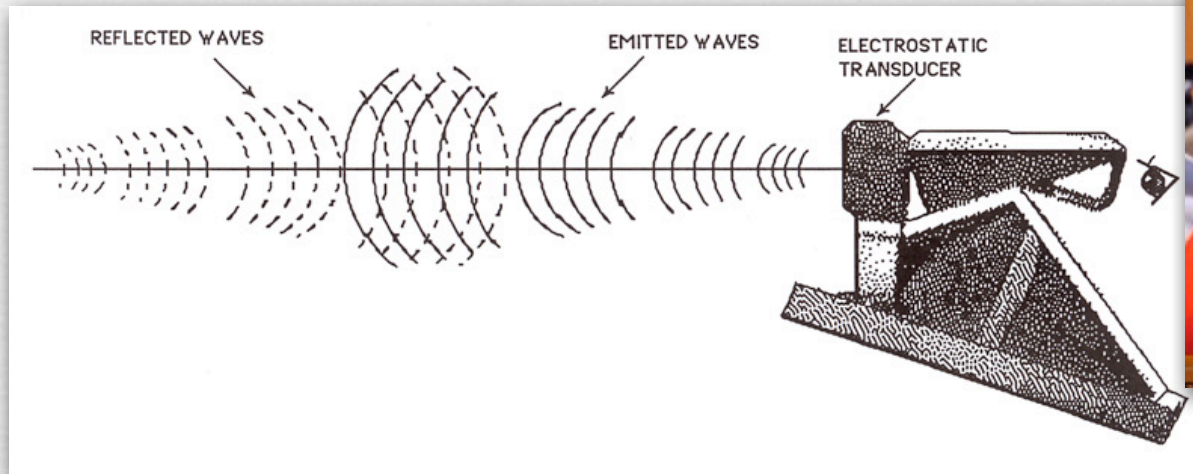
(Canon)

Recap

- ◆ *phase detection* uses *disparity* between views of a scene feature as seen through left and right sides of aperture to judge misfocus
 - most SLRs use phase detection
 - permits direct jump to in-focus position
 - like depth-from-stereo in computer vision
- ◆ *contrast detection* uses *gradient* of pixel values from main sensor's view of a scene feature to judge misfocus
 - most DSCs use contrast detection
 - requires repeated measurements (hunting) as lens moves
 - like depth-from-focus in computer vision

Questions?

Active autofocus: time-of-flight

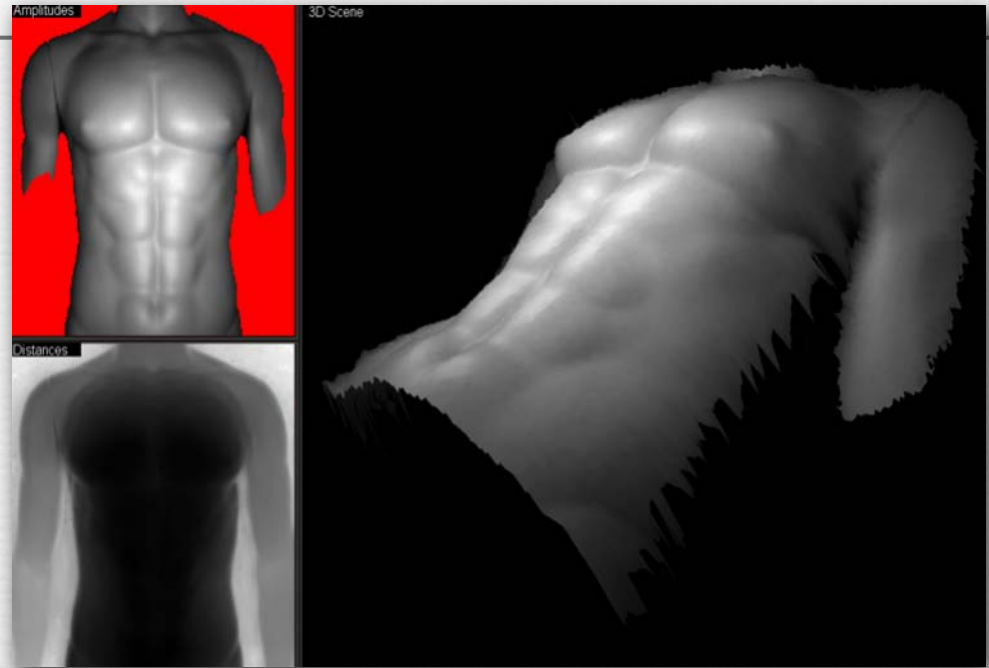


(Goldberg)



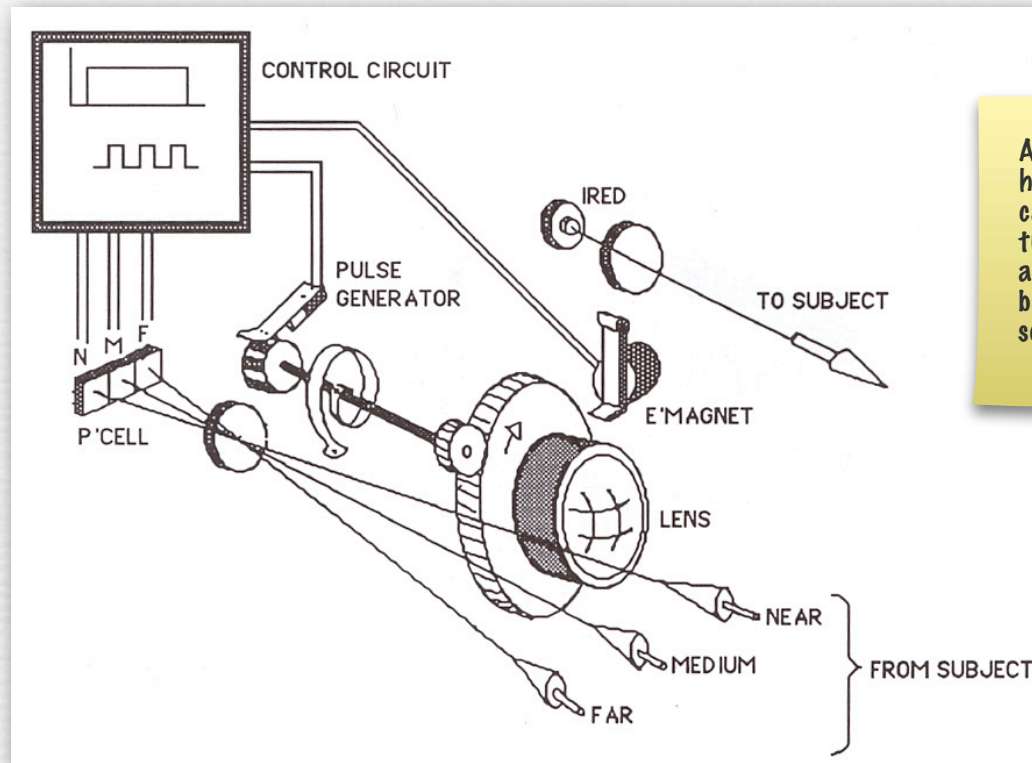
- ◆ SONAR = Sound Navigation and Ranging
- ◆ Polaroid system used ultrasound (50KHz)
 - well outside human hearing (20Hz - 20KHz)
- ◆ limited range, stopped by glass
- ◆ hardware salvaged and re-used in amateur robotics

Active autofocus: time-of-flight



- ◆ LIDAR = Light Detection and Ranging
- ◆ Swiss Ranger SR4000 uses infrared light
- ◆ light travels 1 foot per nanosecond, so accuracy requires fast circuitry ($\pm 5\text{mm}$ is typical)
- ◆ doesn't work in bright scenes

Active autofocus: triangulation



As I mentioned in class, the technology depicted here is not used on any current consumer camera. Don't confuse it with the LED that turns on to illuminate a dark scene and thereby assist many cameras with phase or contrast based passive autofocusing. That LED is sometimes called the "autofocus assist light".

(Goldberg)

- ◆ infrared (IR) LED flash reflects from subject
- ◆ angle of returned reflection depends on distance
- ◆ fails on dark or shiny objects

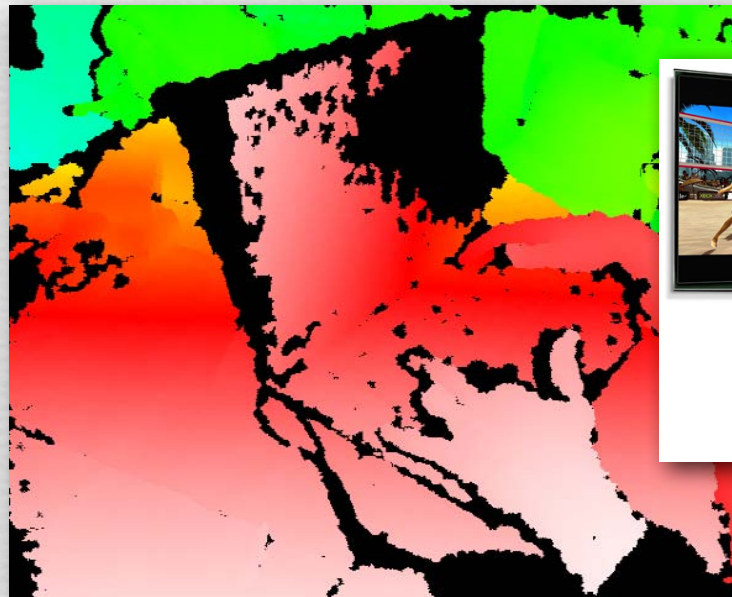
Triangulation rangefinding using texture-assisted stereo

IR video projector RGB camera IR depth camera

Kinect for XBox 360
developed by Prime Sense



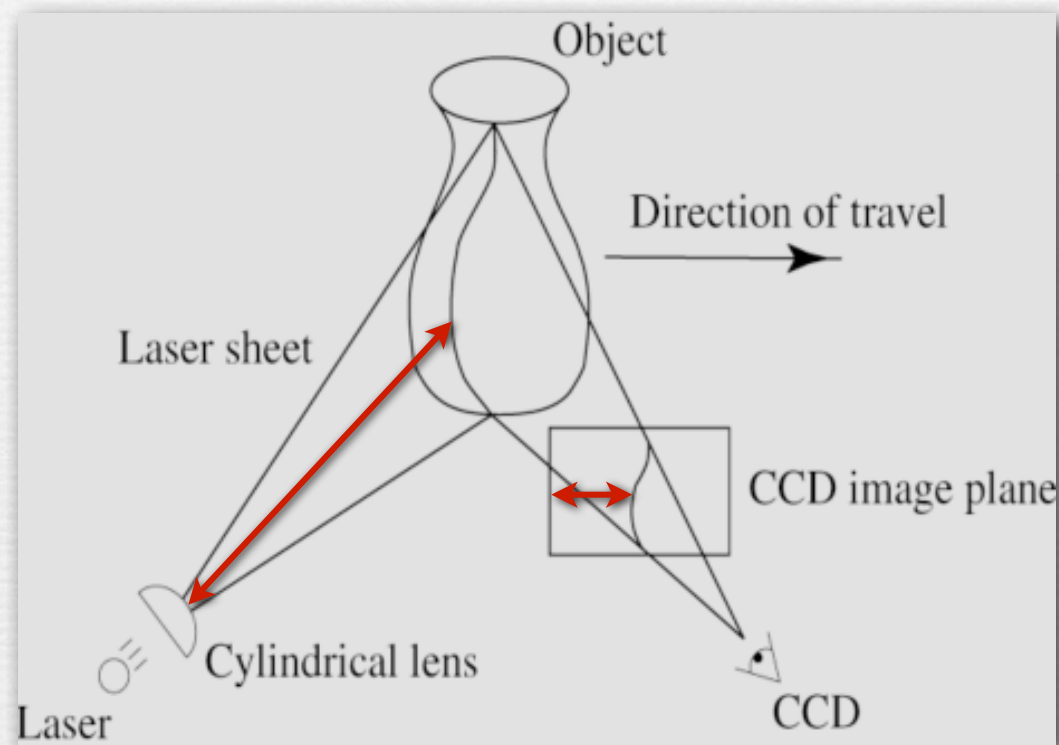
projected texture



depth map

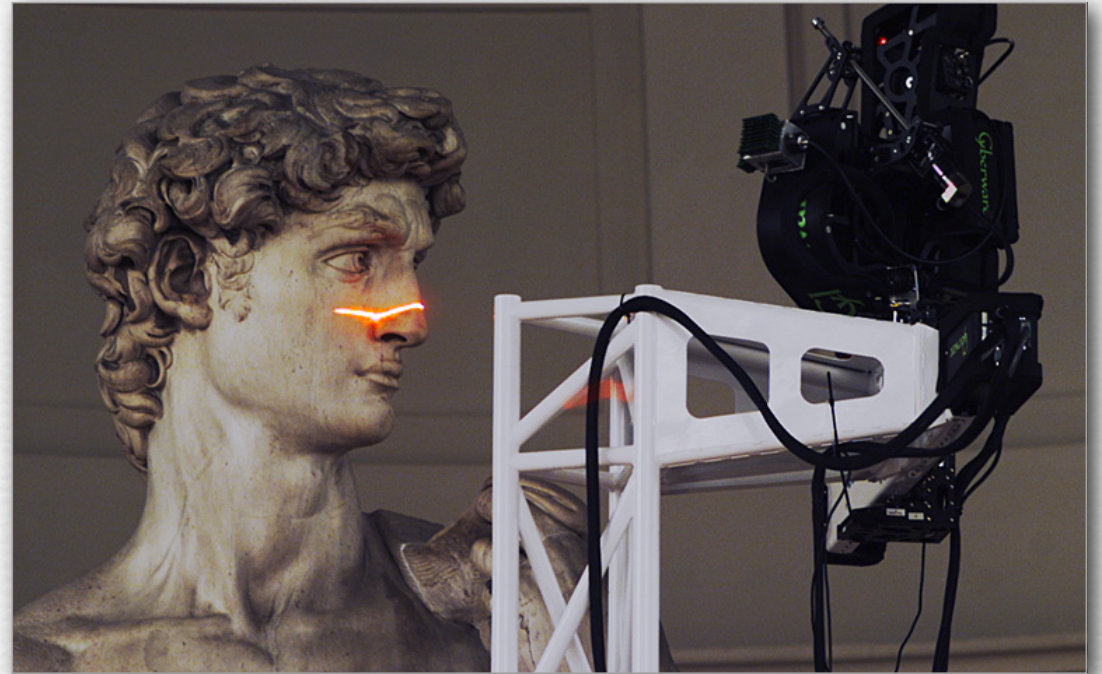


Triangulation rangefinding using laser stripe scanning



- ◆ laser sheet illuminates a curve on the object
- ◆ distance from left edge of image gives distance from laser
- ◆ move object or sweep laser to create *range image* $z(x,y)$

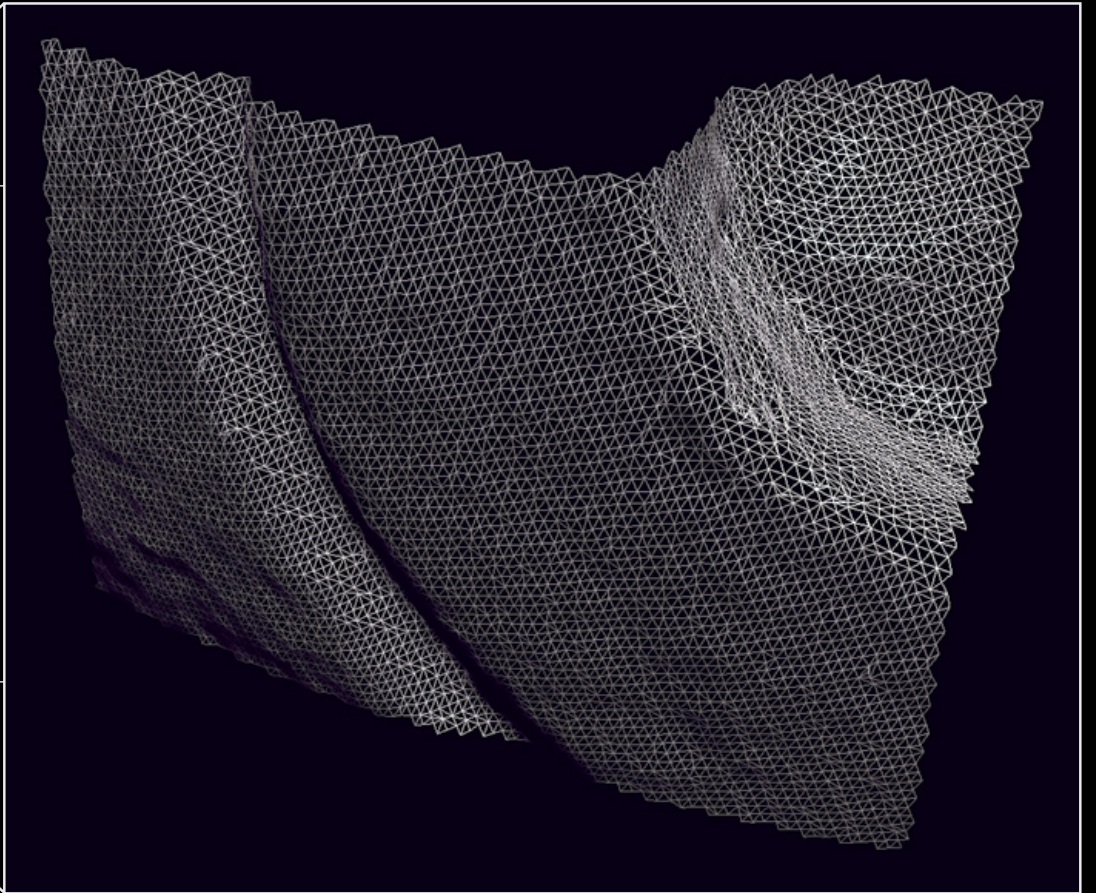
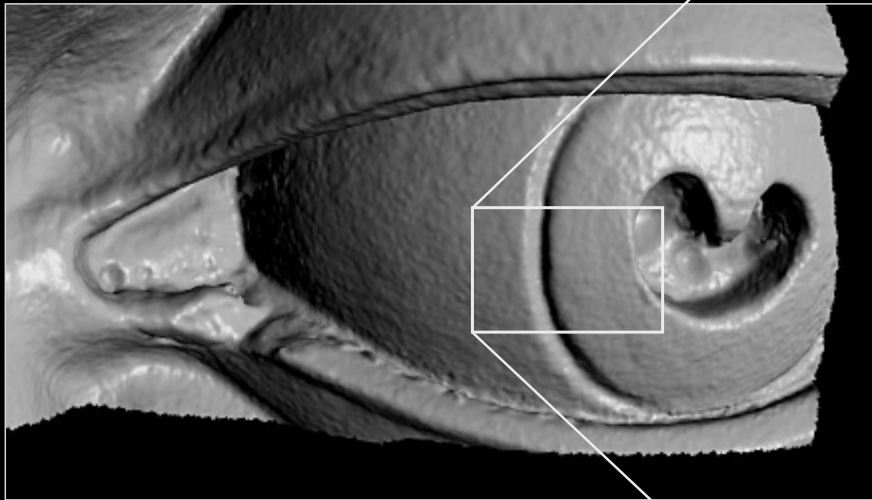
Scanning Michelangelo's David



- ◆ 480 range images
- ◆ 2 billion polygons
- ◆ 22 people × 30 nights

<http://graphics.stanford.edu/projects/mich/>





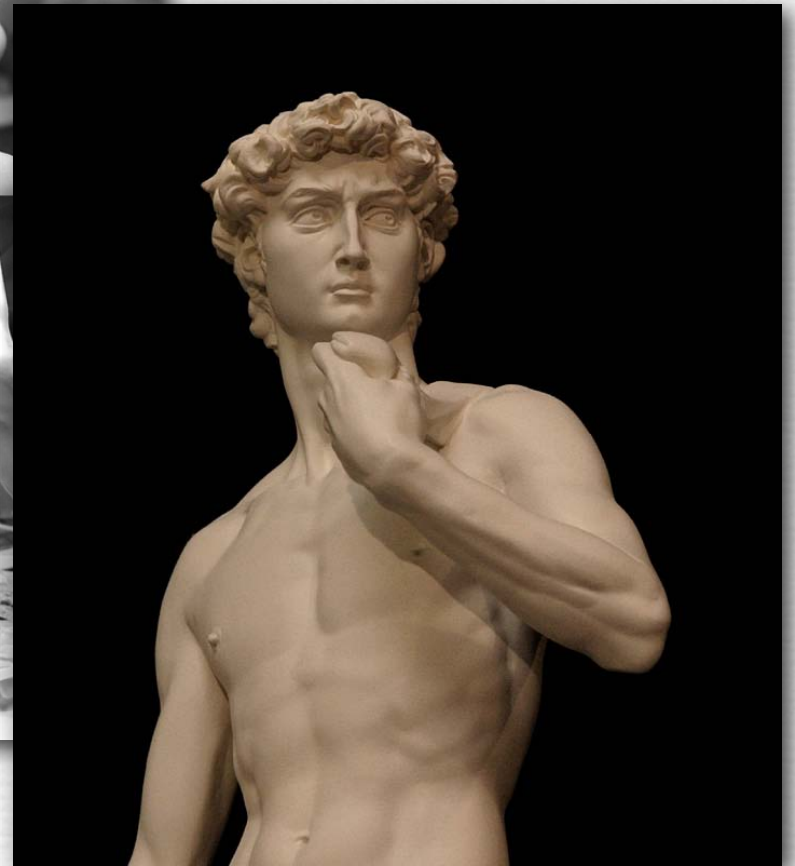
Uses of the 3D model



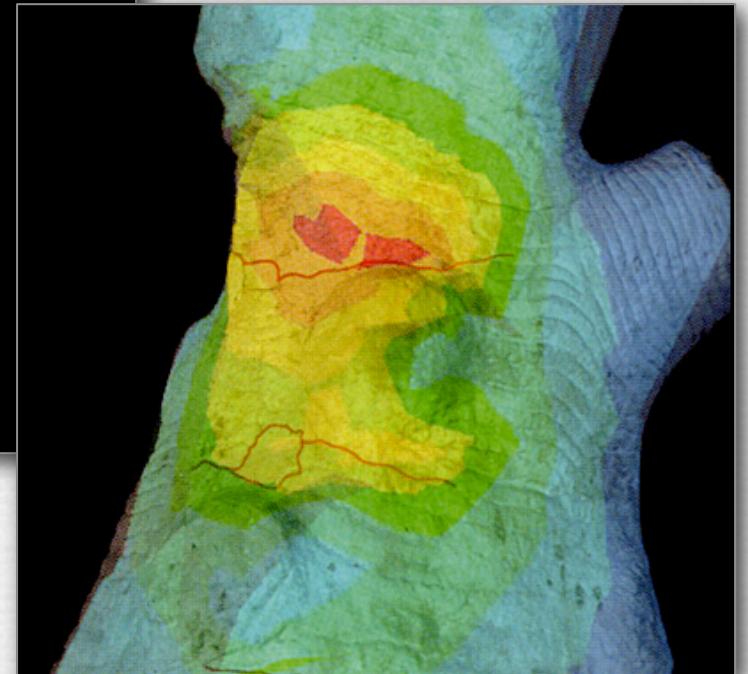
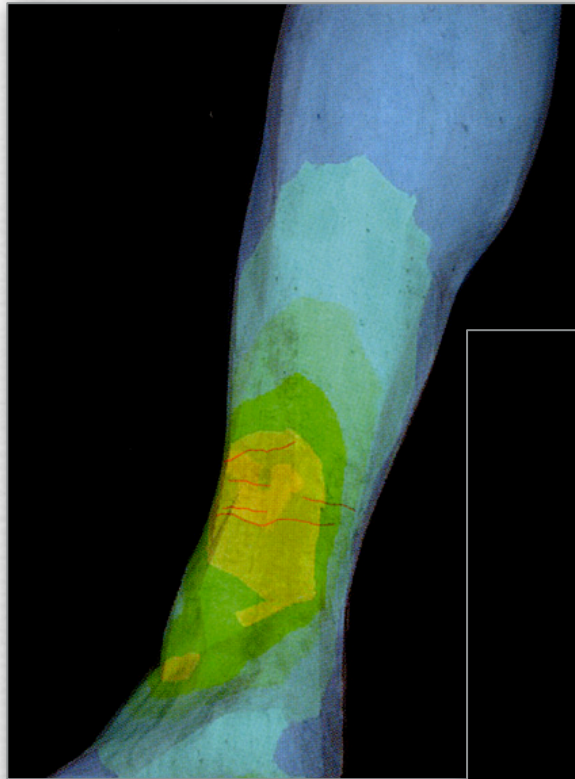
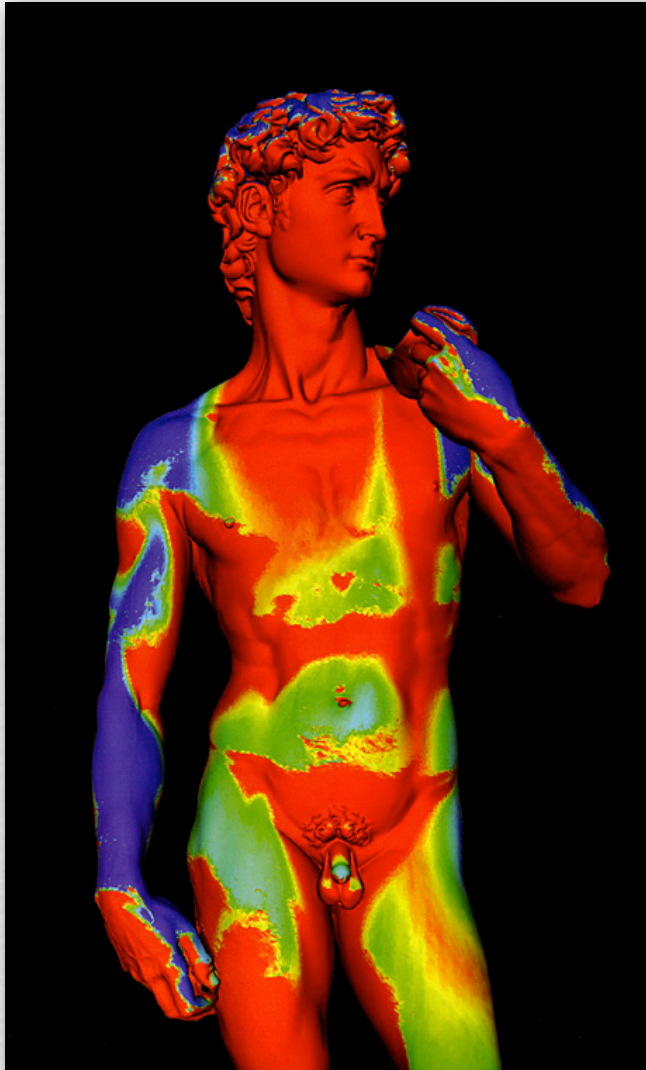
interactive kiosk



physical replica



Uses of the 3D model



scientific studies

cinefex

number 107
\$12.50
Canada - \$15.00

- ◆ commonly used to digitize physical maquettes to ease creation of 3D CG models



3D model



mocap



rendered



Slide credits

- ◆ Goldberg, N., *Camera Technology: The Dark Side of the Lens*, Academic Press, 1992.
- ◆ Canon, *EF Lens Work III: The Eyes of EOS*, Canon Inc., 2004.
- ◆ Adams, A., *The Camera*, Little, Brown and Co., 1980.
- ◆ Kerr, D.A., *Principle of the Split Image Focusing Aid and the Phase Comparison Autofocus Detector in Single Lens Reflect Cameras*.