Seeing things that are hard to see: (a) the big world of tiny motions and (b) accidental cameras

Bill Freeman

Massachusetts Institute of Technology

December 17, 2012

Joint work with: Fredo Durand, Antonio Torralba, Michael Rubinstein, Neal Wadhwa, Hao-Yu Wu, John Guttag.

Massachusetts Institute of Technology

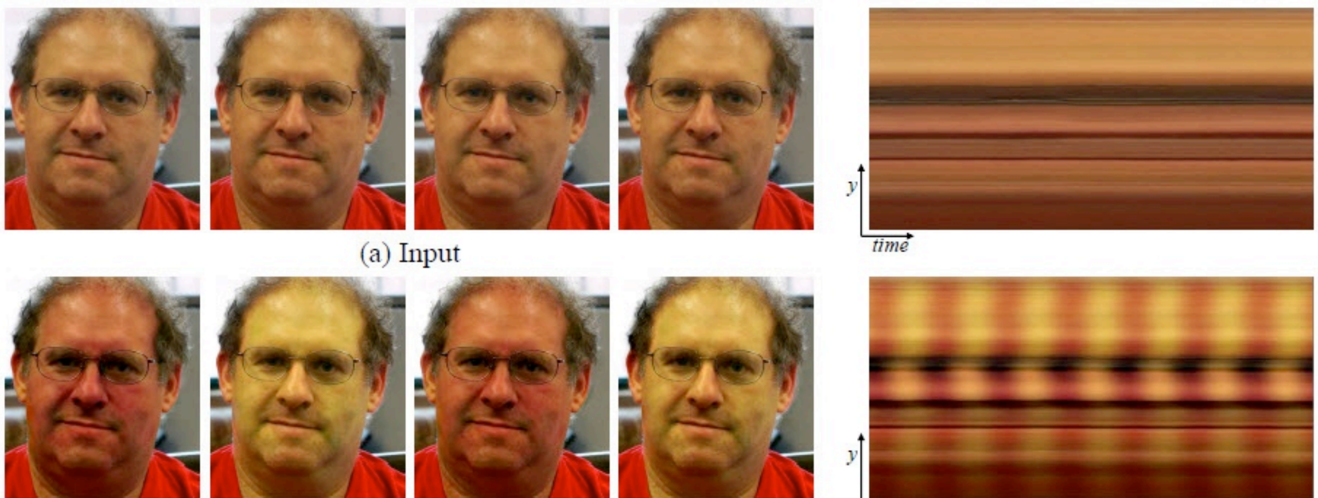
ICVGIP 2012

Massachusetts Institute of Technology

Quanta Research

Eulerian Video Magnification
for Revealing Subtle Changes in the World

with Hao-Yu Wu, Eugene Shih*, John Guttag, Fredo Durand and Bill Freeman
MIT CSAIL, and *Quanta Research Cambridge, Inc.



(a) Input
(b) Magnified
(c) Spatiotemporal *YT* slices

Imperceptible Changes in the World

Respiratory motion

Buildings swaying in wind

Pulse and Blood flow

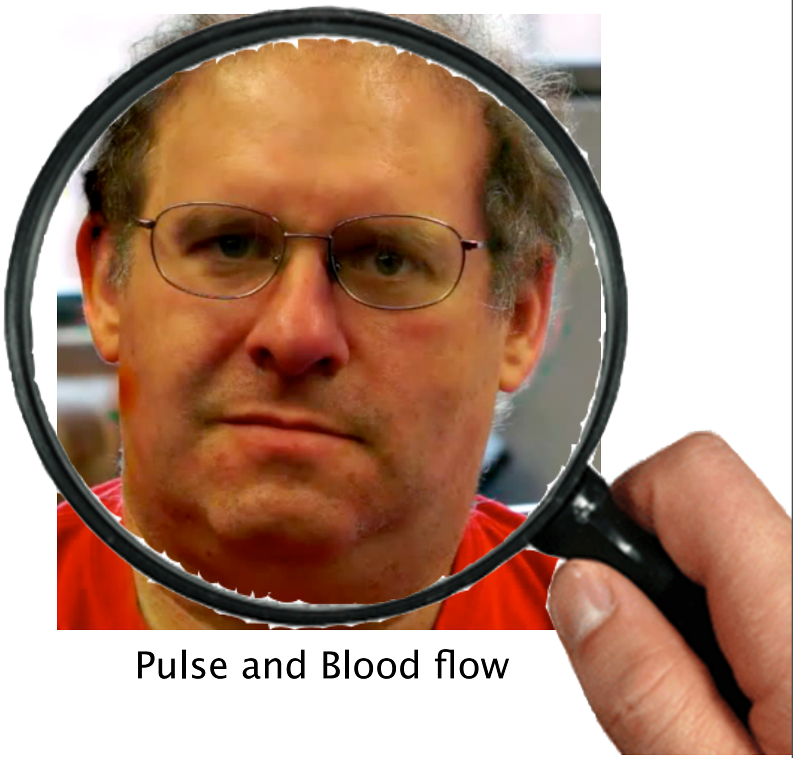
Magnifying Glass for Temporal Variations

Respiratory motion

Pulse and Blood flow

Magnifying Glass for Temporal Variations

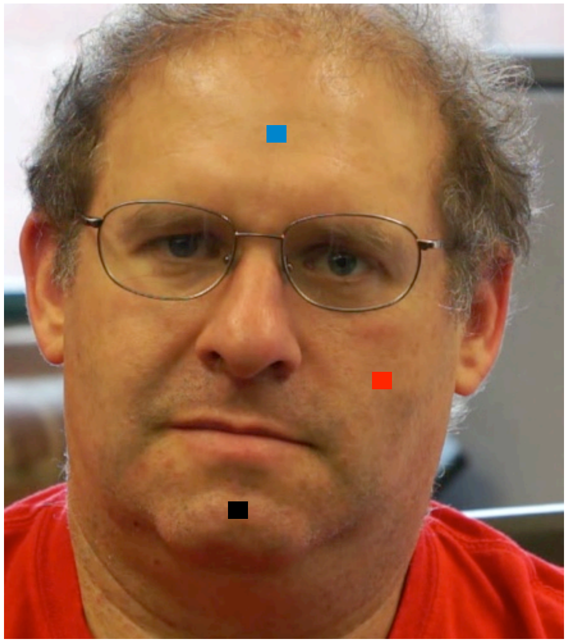
Respiratory motion

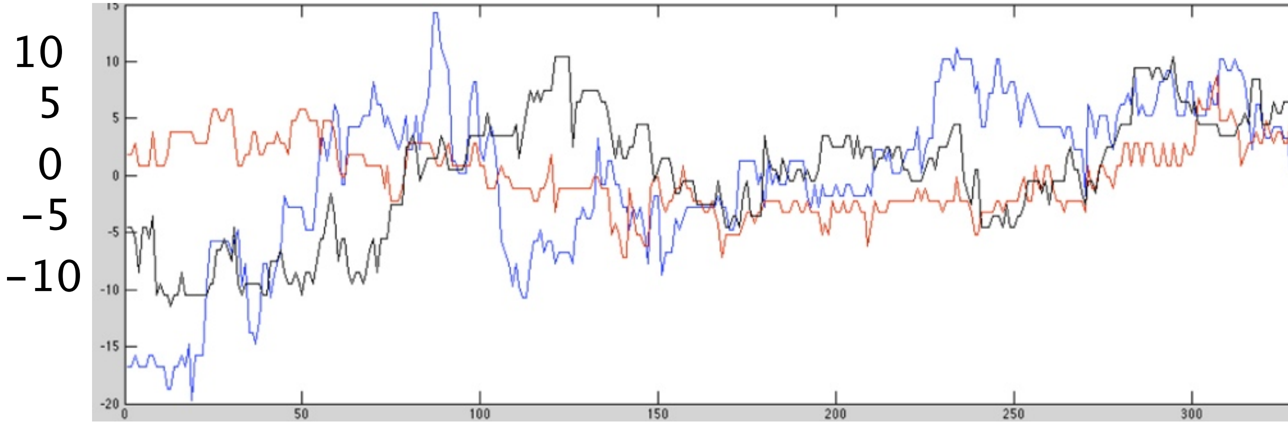


Pulse and Blood flow

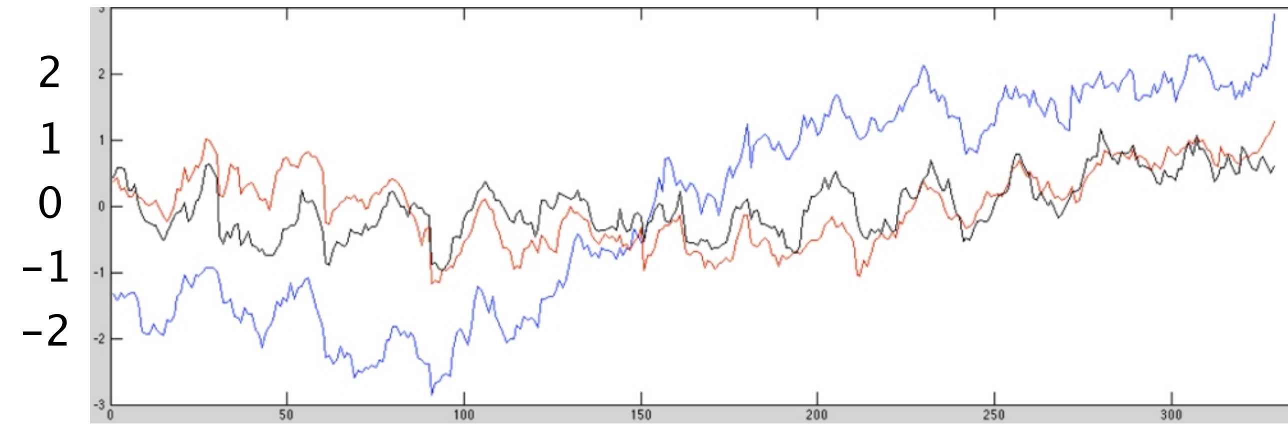
Amplifying Subtle Color Variations

1. Average spatially to overcome sensor and quantization noise





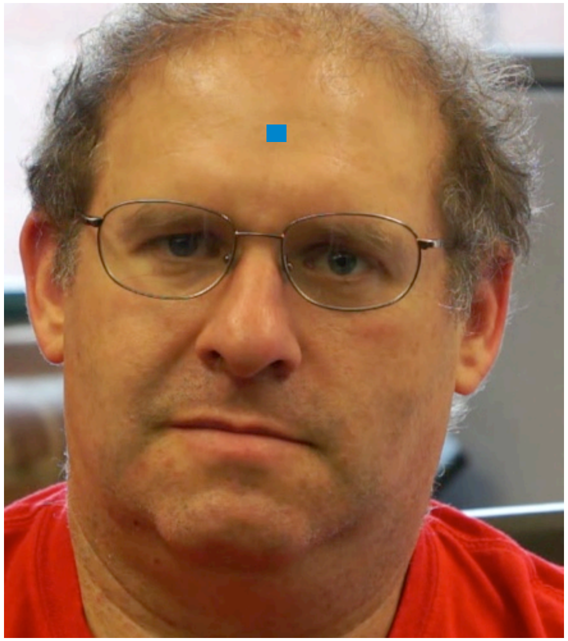
Luminance traces for 3 image locations (zero mean)

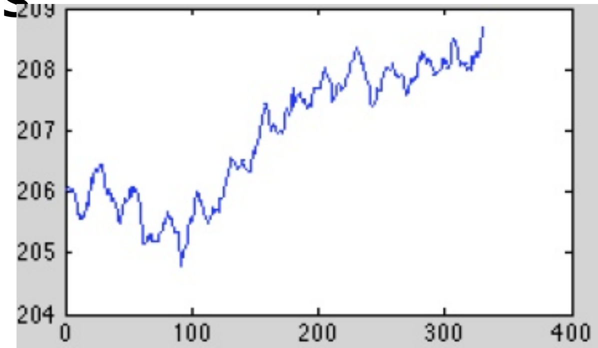


Spatial average in those locations

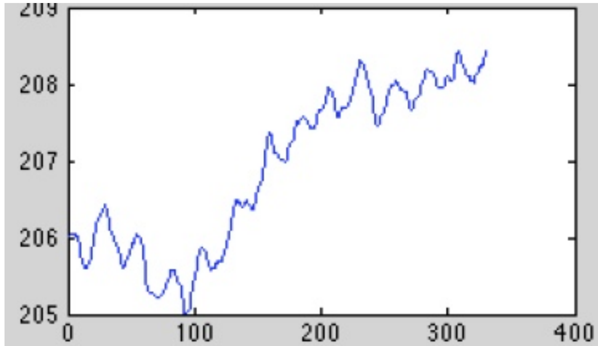
Amplifying Subtle Color Variations

2. Filter Temporally

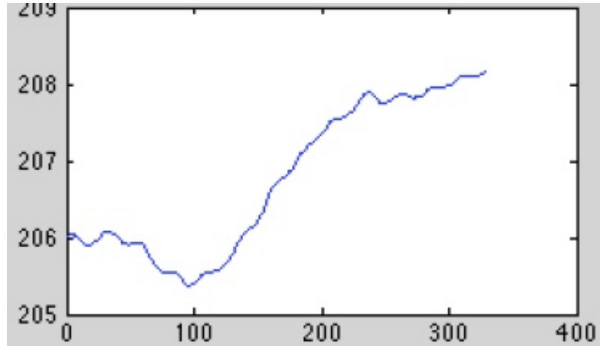




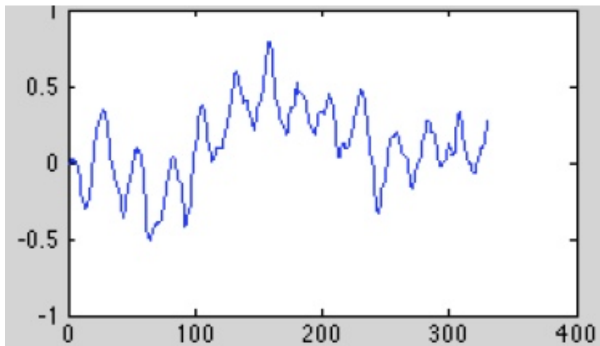
Luminance trace



IIR1 [.4,.6]



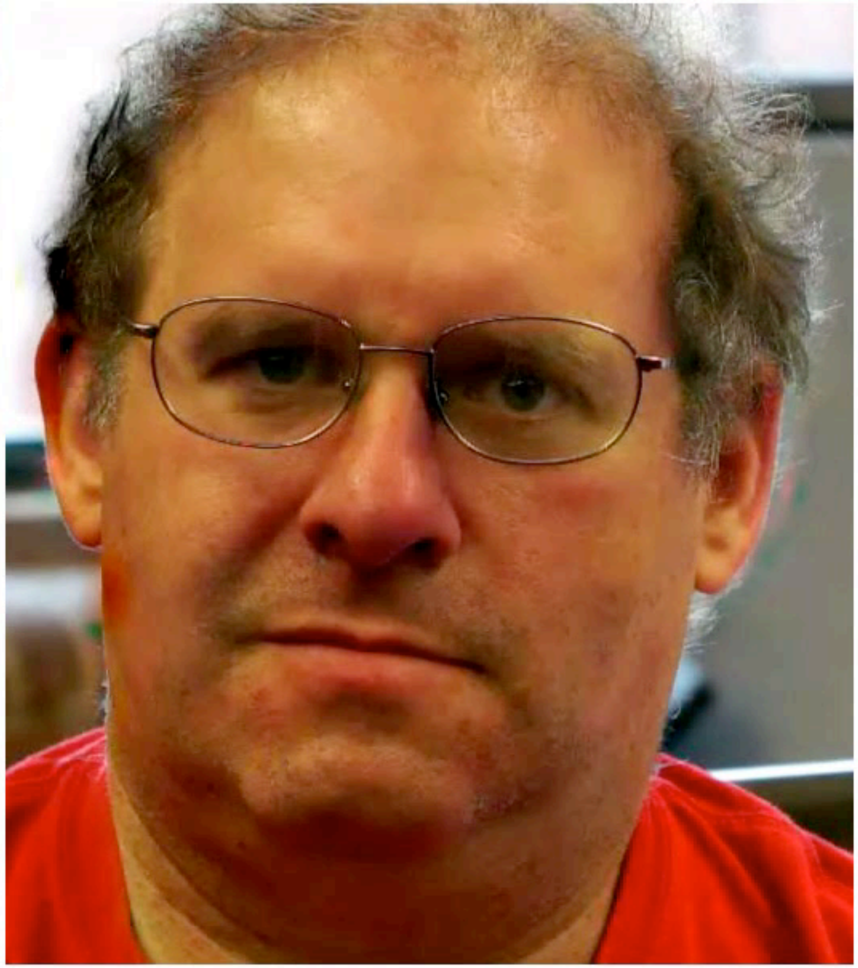
IIR2 [.05,.95]



Temporally bandpassed trace

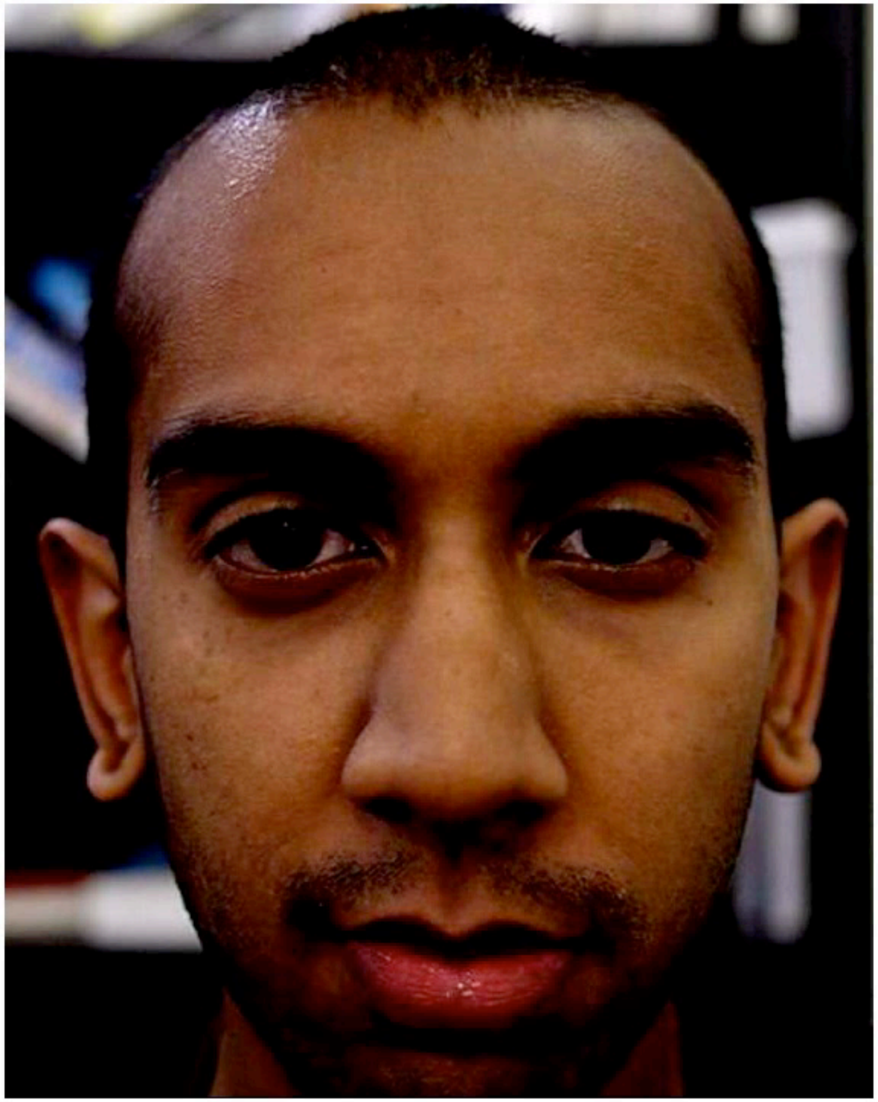
Color Amplification Results

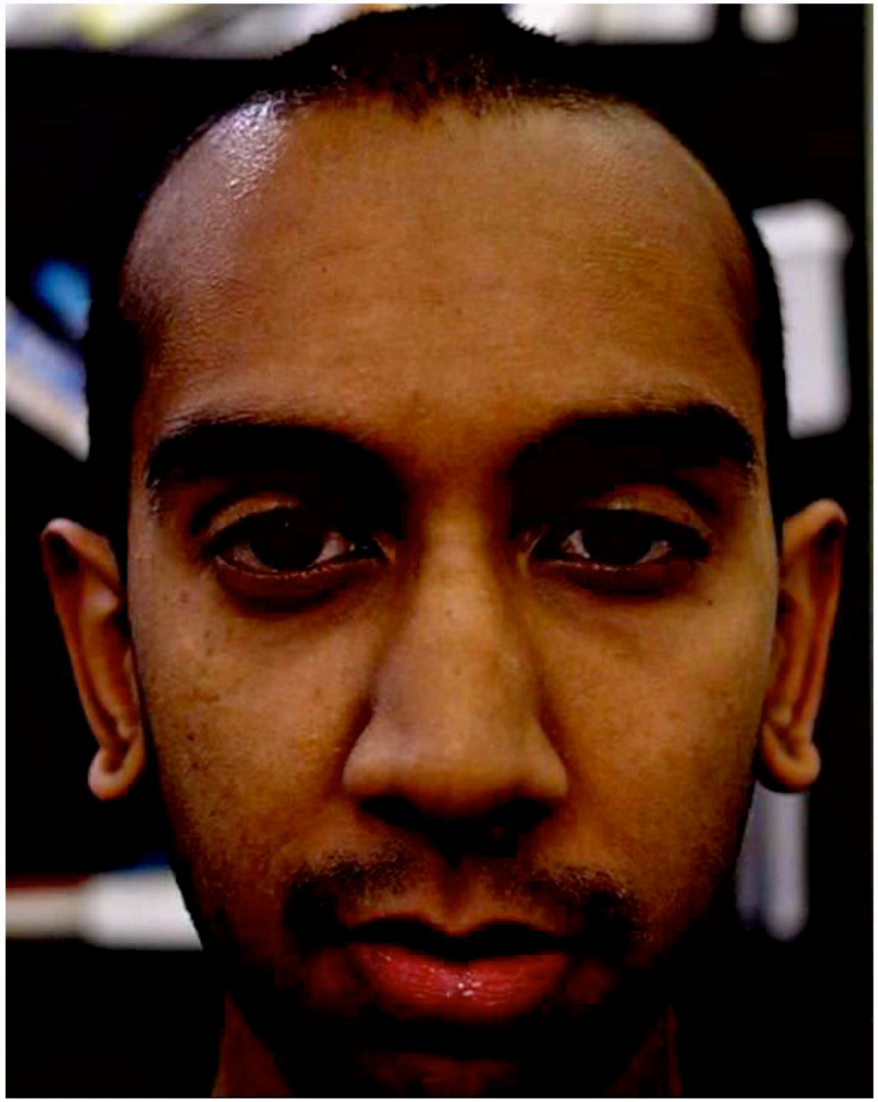
Source



Color-amplified (x100)
0.83-1 Hz (50-60 bpm)
(**ideal filtering**)

Color Amplification Results

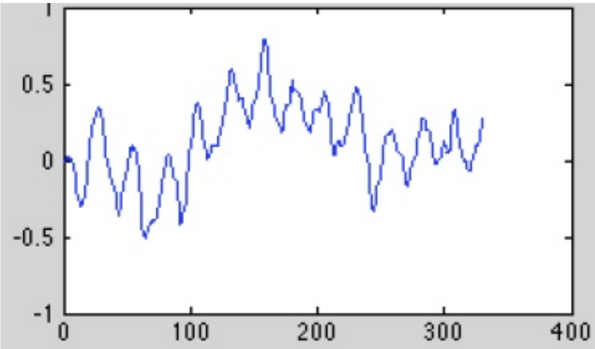




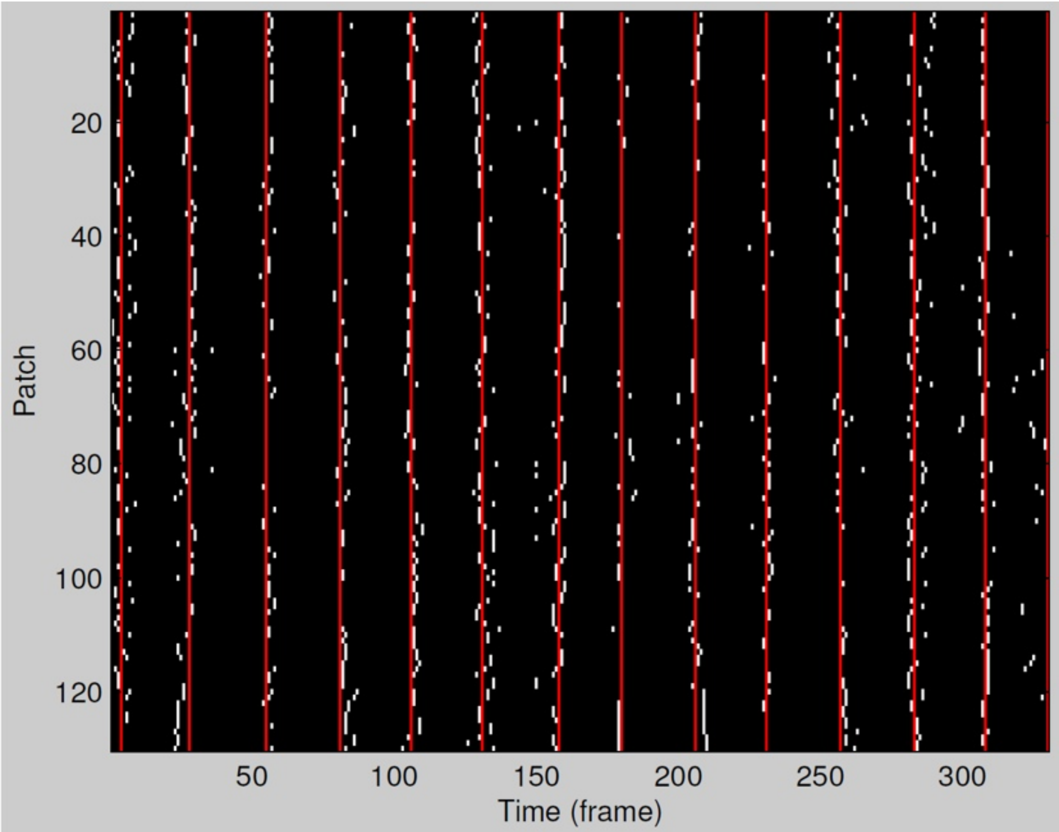
Source

Color-amplified (x120)
0.83-1 Hz (50-60 bpm)

Heart Rate Extraction

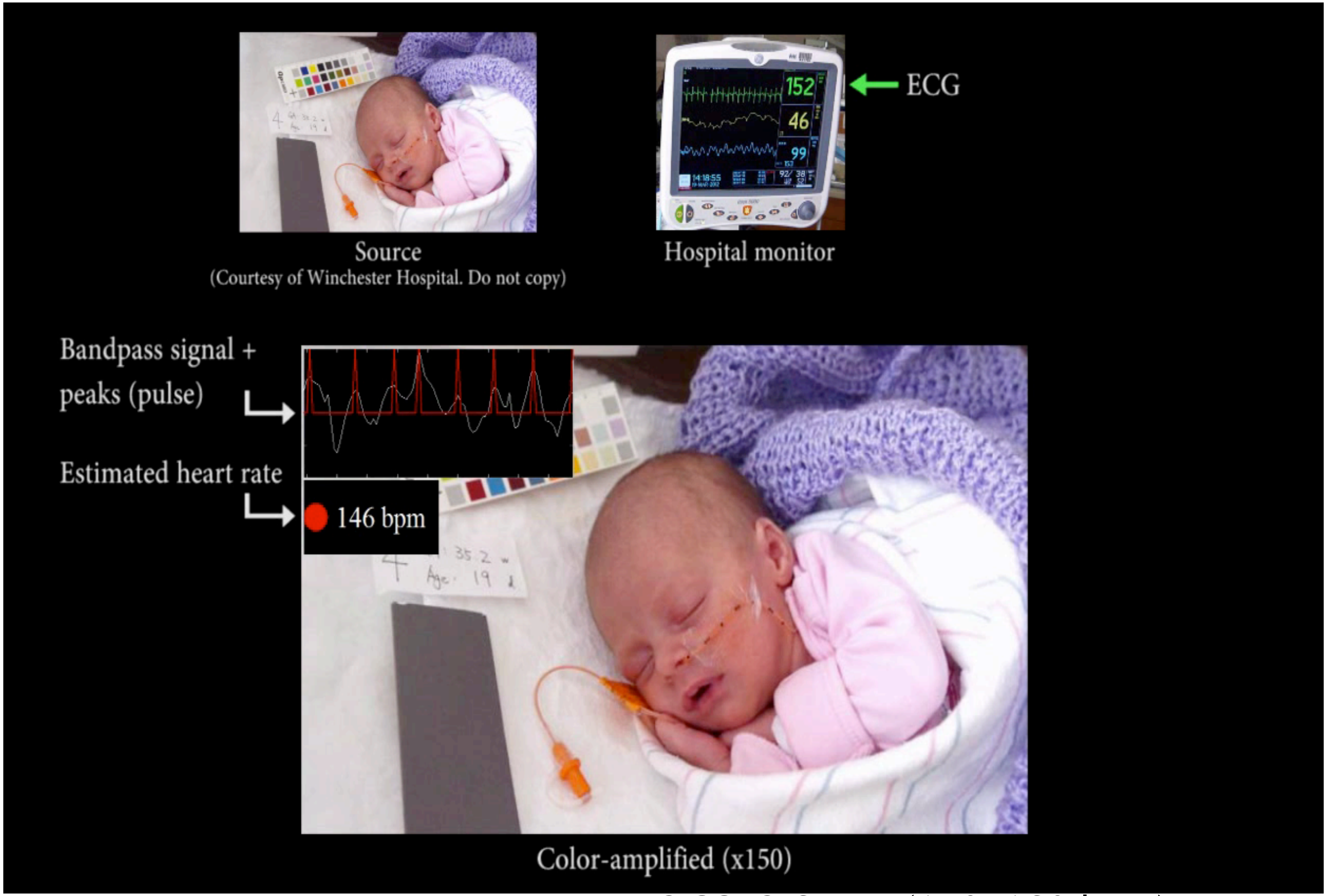


Temporally bandpassed trace (one patch)



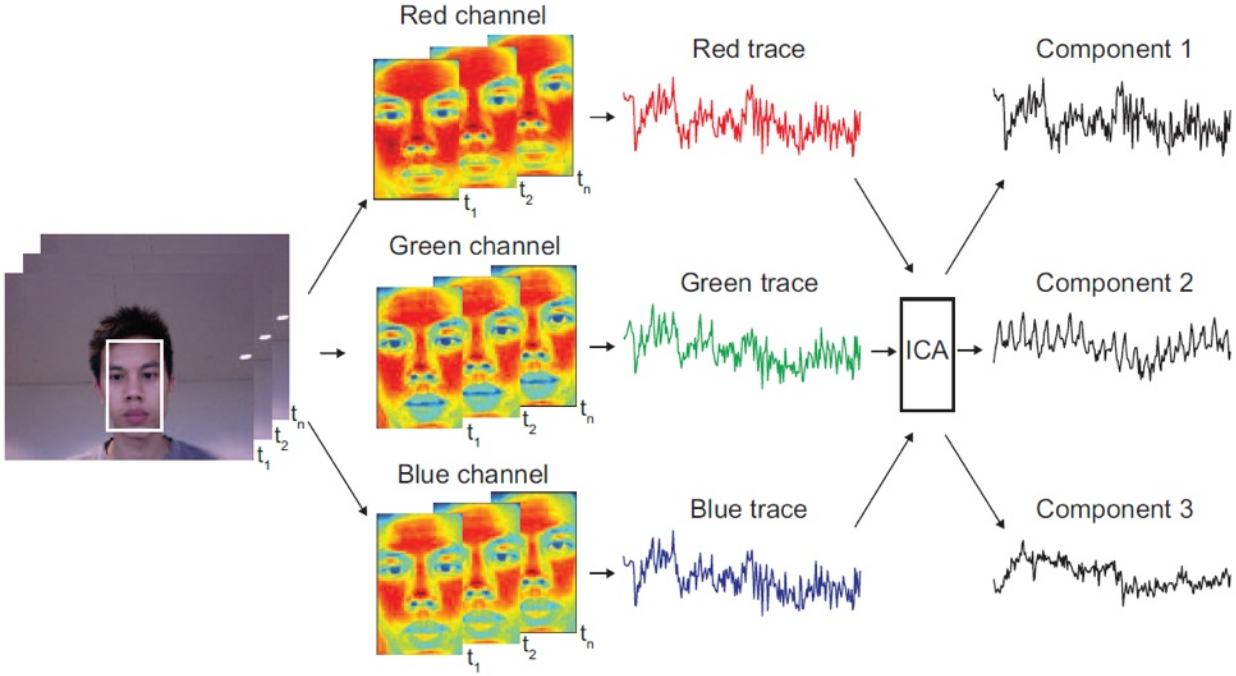
Pulse times

Heart Rate Extraction



Thanks to Dr. Donna Brezinski and the Winchester Hospital staff 2.33-2.67 Hz (140-160 bpm)

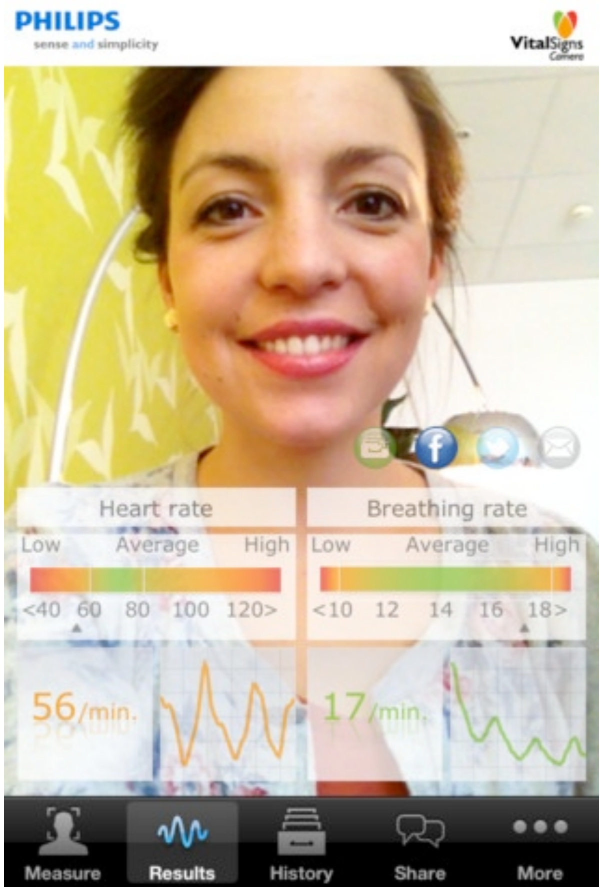
Related Work on Pulse Detection from Videos



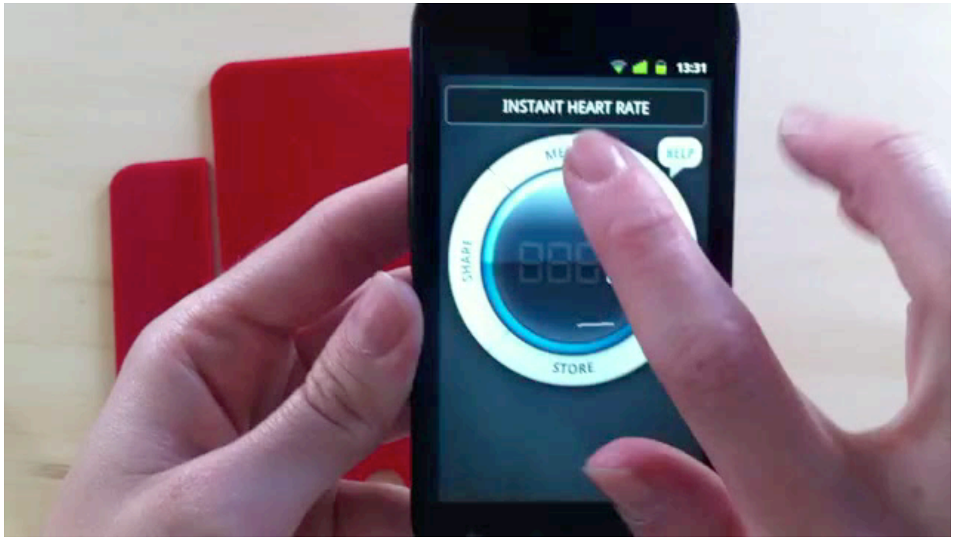
Poh, McDuff and Picard, MIT Media Lab, Non-contact, automated cardiac pulse measurements using video imaging and blind source separation, 2010

Face recognition and tracking → RGB averaged over ROI (single number) → Use ICA to find signal sources

Cell Phone Apps Developed Previously by Others…



“Vital Signs Camera” – Philips (proprietary)



“Instant Heart Rate” for Android
Photoplethysmography (PPG)

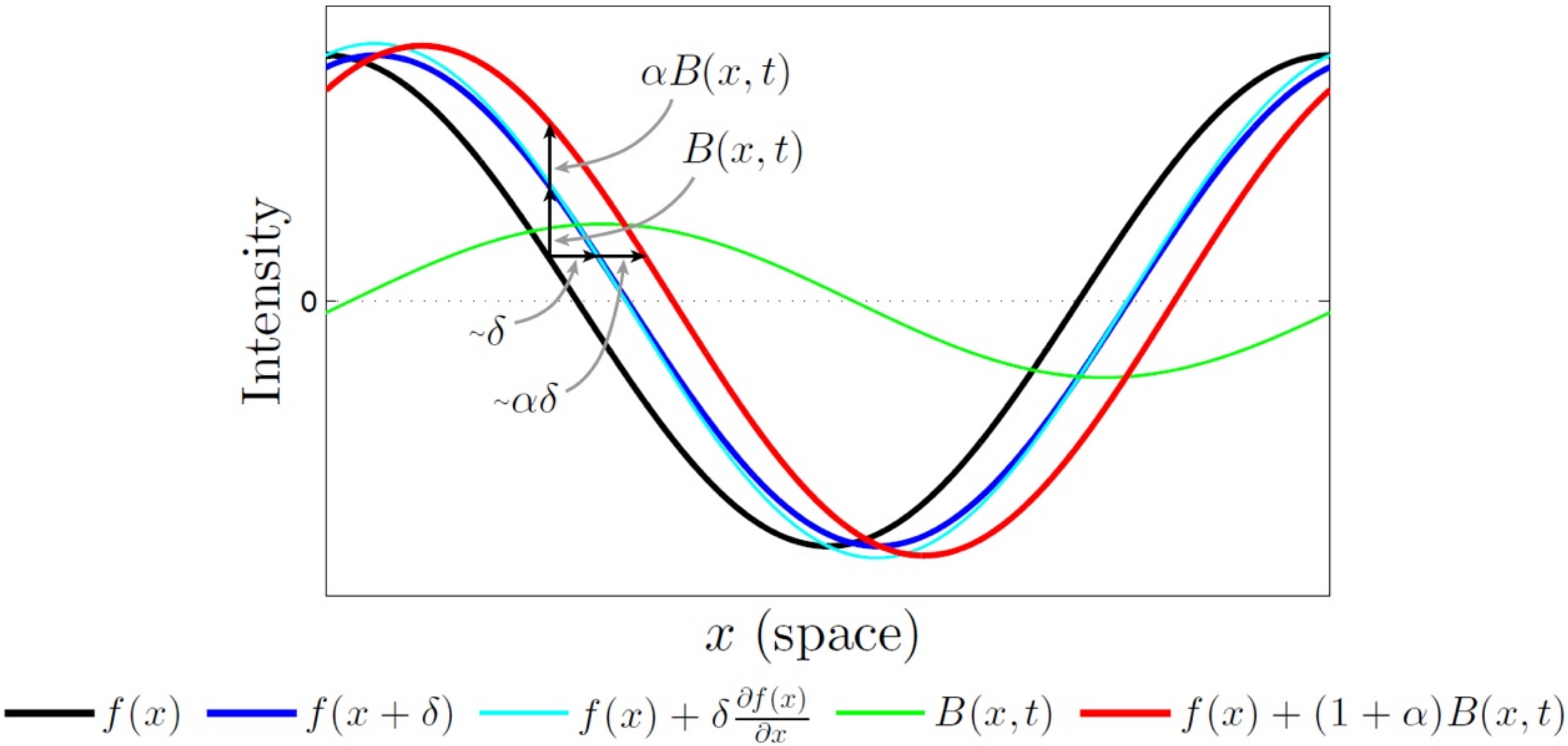
Spatio-temporal variations of pulse onset may reveal medically useful information

Source

Color-amplified (x100)
0.83-1 Hz (50-60 bpm)

The extra motion with the color amplification puzzled us...

Motion Magnification via Temporal Filtering



Linearized motion magnification

Rigid translation

$$I(x,t) = f(x+\delta(t))$$

Assume small translation relative to image structures and expand in 1st order Taylor series

$$I(x,t) \approx f(x) + \delta(t)\frac{\partial f(x)}{\partial x}$$

Amplify temporally bandpassed signal

$$\hat{I}(x,t) = I(x,t) + (\alpha - 1)B_t[I(x,t)]$$

Modified signal

$$\approx f(x) + \alpha\delta(t)\frac{\partial f(x)}{\partial x}$$

Assume the amplified translation is still small relative to image structures to use Taylor series approximation again

$$\approx f(x+\alpha\delta(t))$$

Where we expect this to break down

Let's look at it for f(x) being a sinusoid, $\cos(\omega x)$

Exact translation by $\alpha\omega\delta$

$$\cos(\omega(x+\alpha\delta)) = \cos(\omega x)\cos(\alpha\omega\delta) - \sin(\omega x)\sin(\alpha\omega\delta)$$

$$\approx \cos(\omega x) - \alpha\omega\delta \sin(\omega x)$$

Linearized

For the motion magnification approximation to hold:

$$\cos(\alpha\omega\delta) \approx 1$$

$$\sin(\alpha\omega\delta) \approx \alpha\omega\delta$$

Condition required for those conditions to be approximately true:

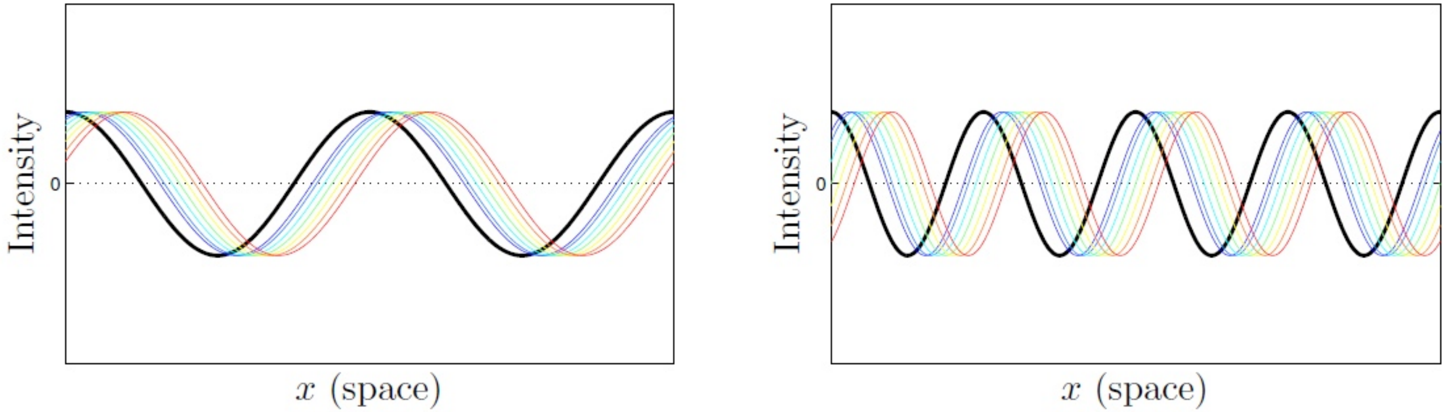
$$\sin(\frac{\pi}{4}) = 0.9\frac{\pi}{4} \qquad \alpha\omega\delta < \frac{\pi}{4}$$

true displacement

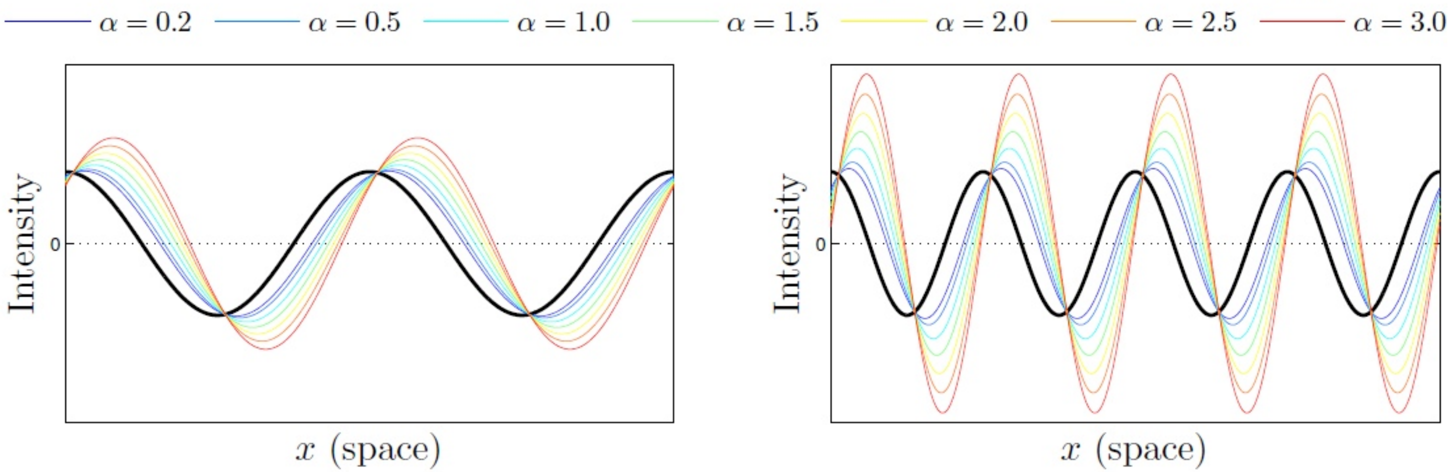
spatial frequency

magnification amount

Synthetic 1D Examples

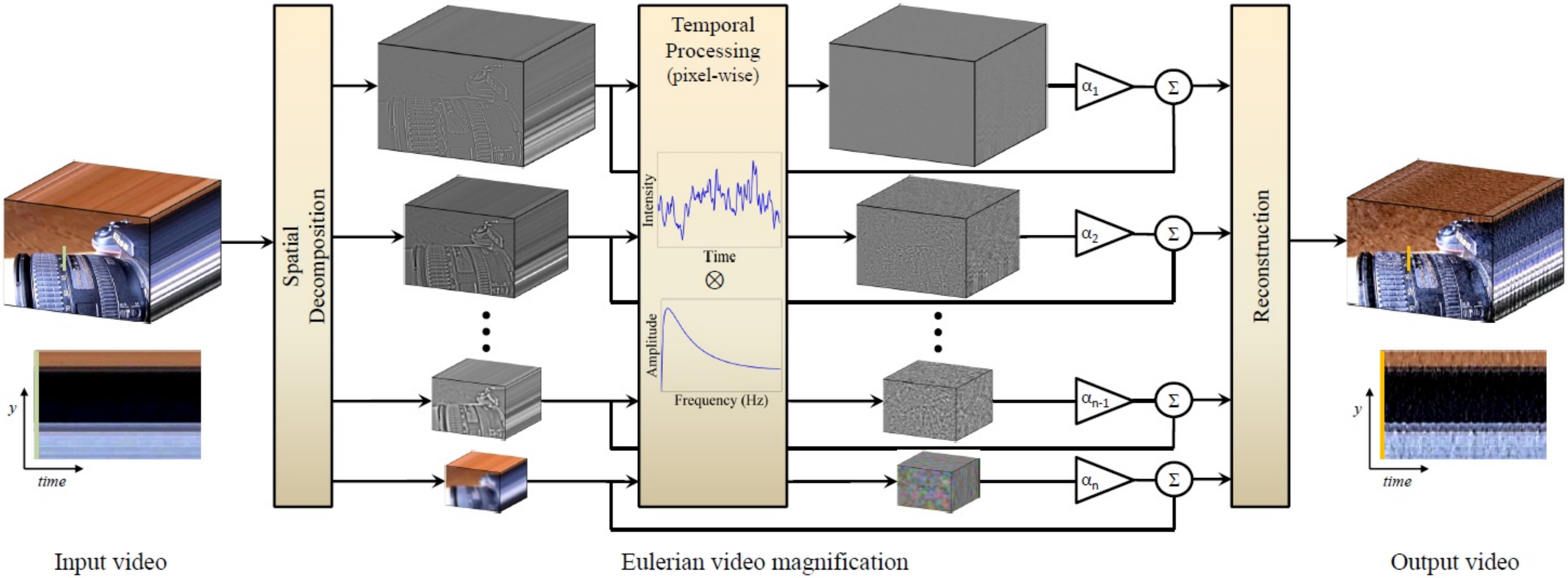


(a) True motion amplification: $\hat{I}(x,t) = f(x + (1+\alpha)\delta(t))$.

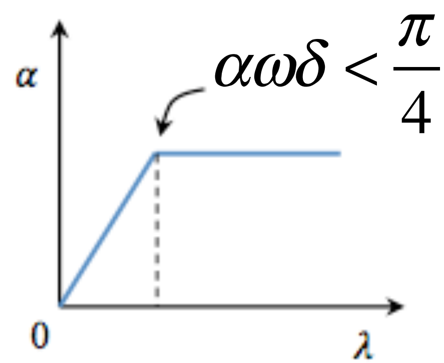


(b) Motion amplification via temporal filtering:
$\tilde{I}(x,t) = I(x,t) + \alpha B(x,t)$.

System Overview



Amplify spatial frequencies where approximation holds, otherwise fail toward zero



$$\lambda = \frac{2\pi}{\omega}$$

Figure 6: *Amplification factor, α, as function of spatial wavelength λ, for amplifying motion. The amplification factor is fixed to α for spatial bands that are within our derived bound (Eq. 14), and is attenuated linearly for higher spatial frequencies.*

Source

Motion-magnified (3.6-6.2 Hz, x60)

Motion Magnification Results

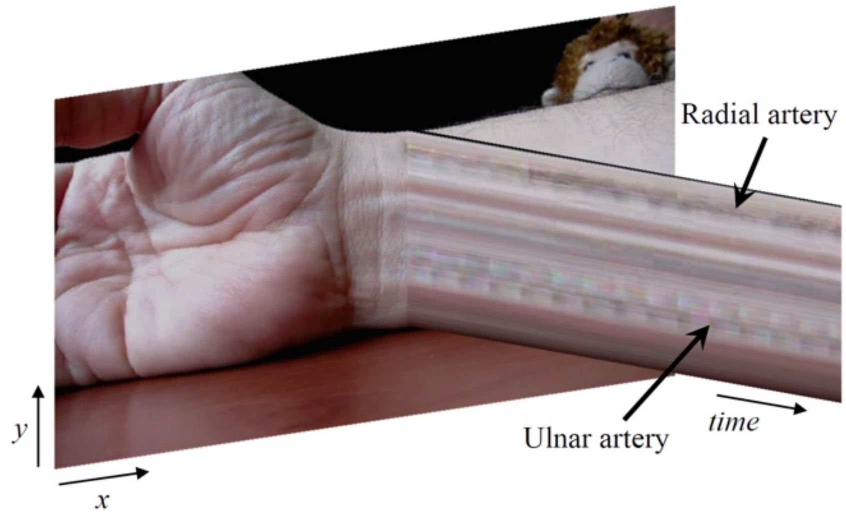
Source

Motion-magnified (0.4-3 Hz, x10)

Motion Magnification

Source

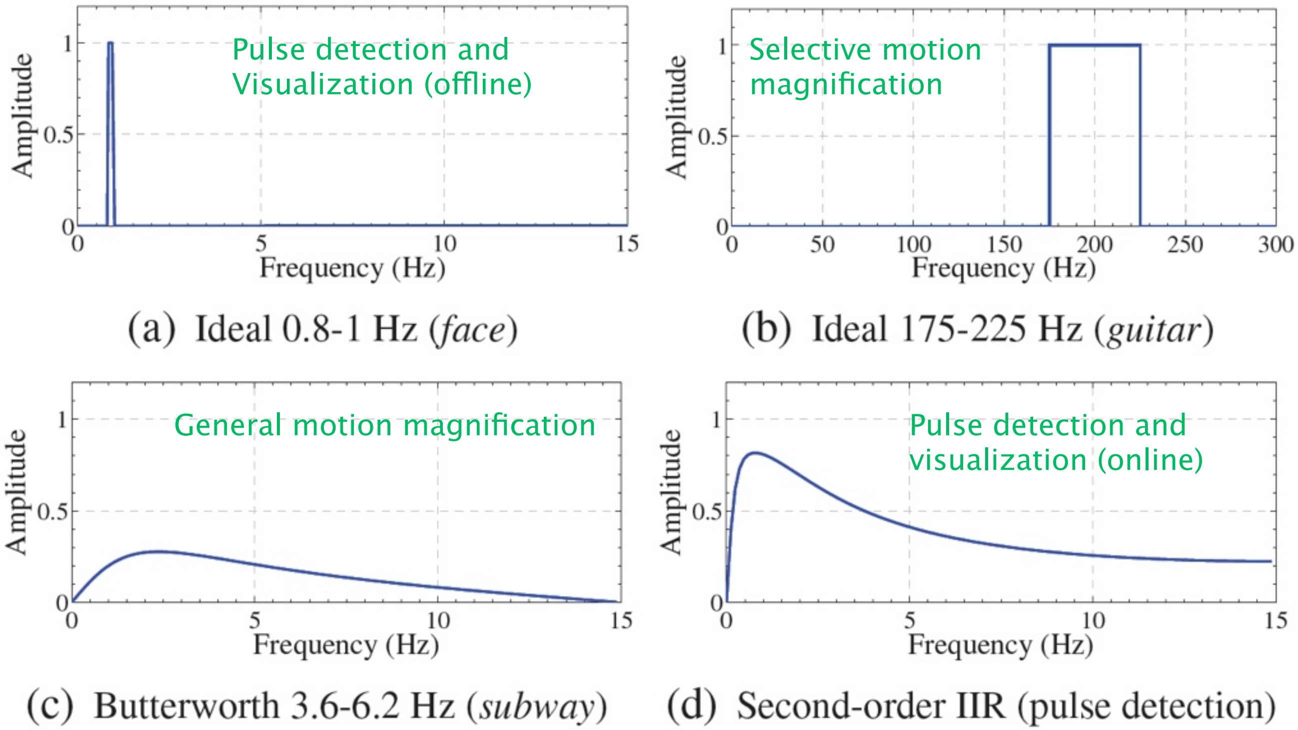
Motion-magnified (0.4-3 Hz, x10)



Temporal Filters

- Mostly application dependent
  - Configurable by the user

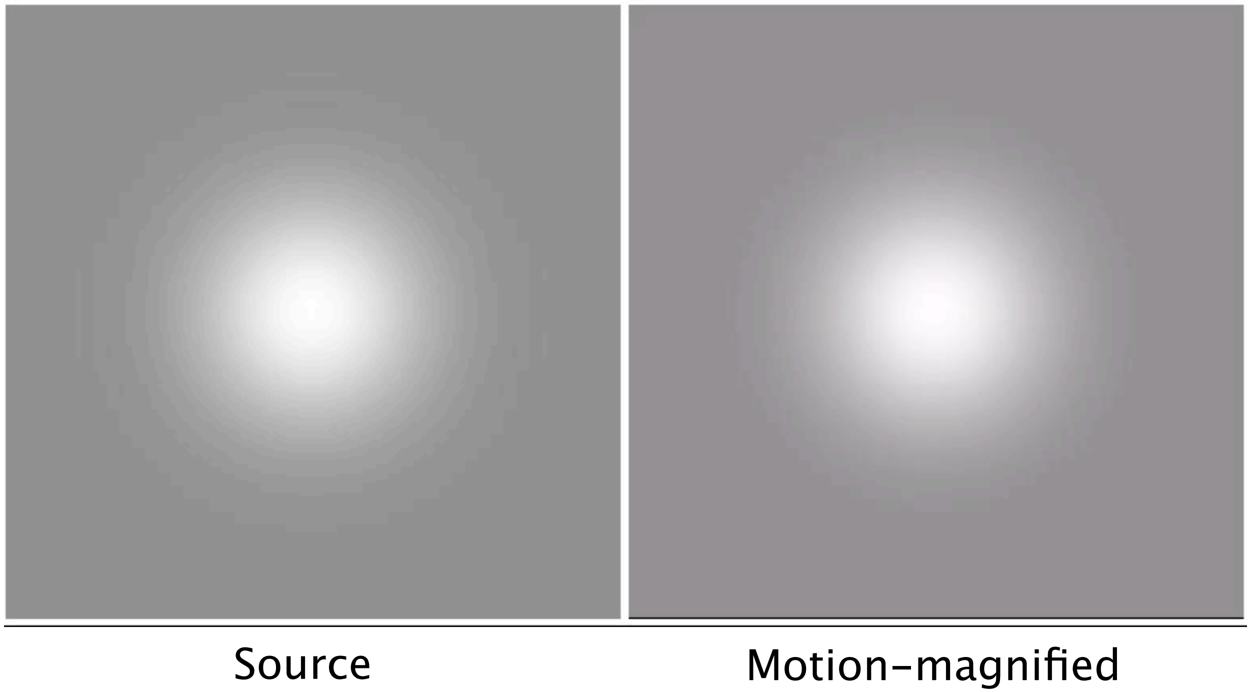
- Some of the filters we used (and their applications):



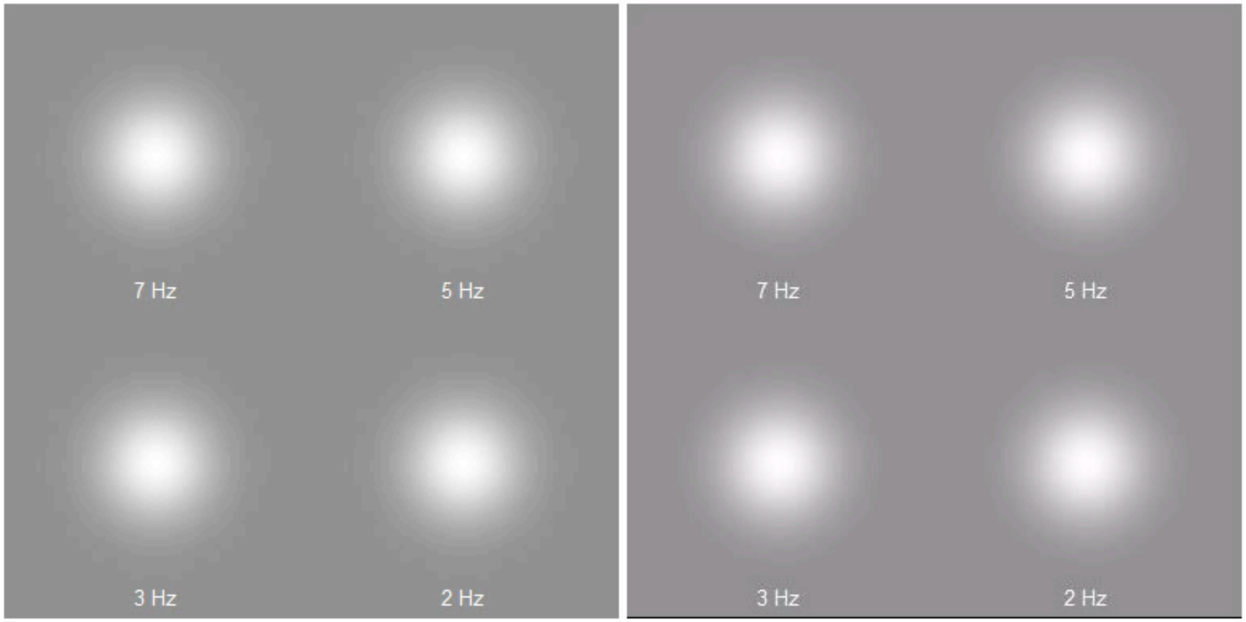
(a) Ideal 0.8-1 Hz (*face*) (b) Ideal 175-225 Hz (*guitar*)

(c) Butterworth 3.6-6.2 Hz (*subway*) (d) Second-order IIR (pulse detection)

Synthetic 2D Example



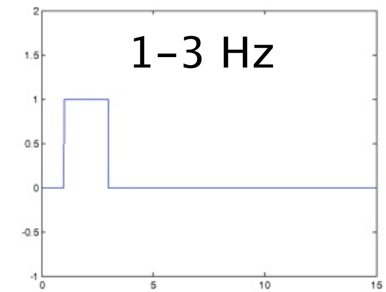
Selective Motion Magnification



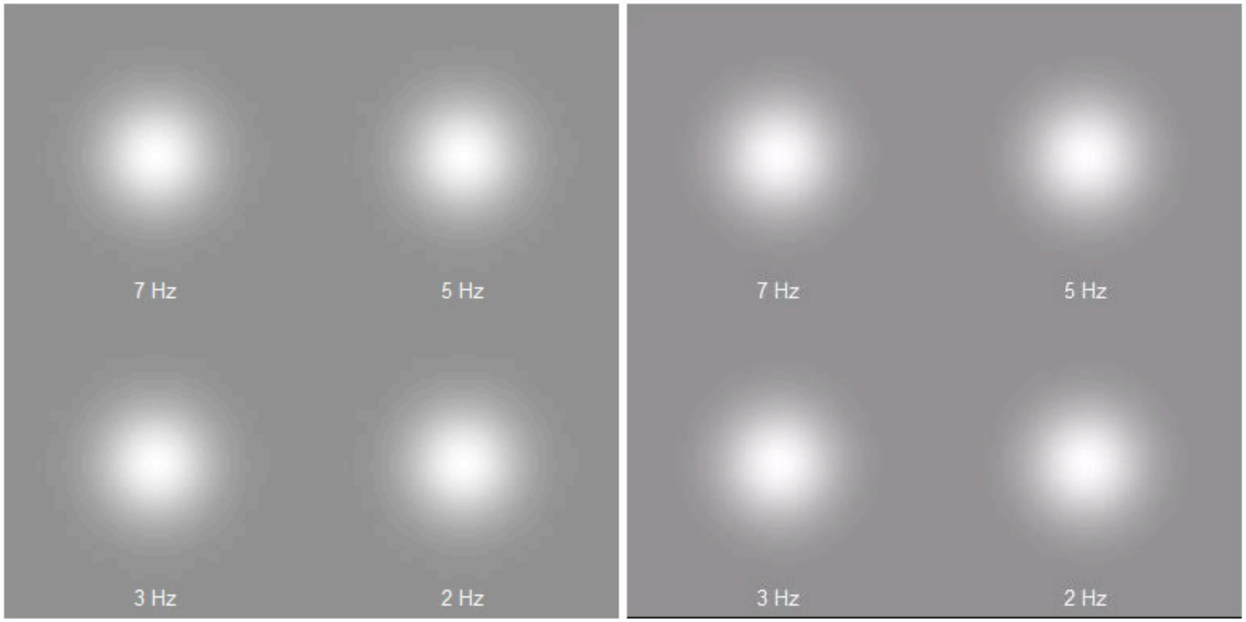
Source

Motion-magnified (2 Hz)

Temporal filter:



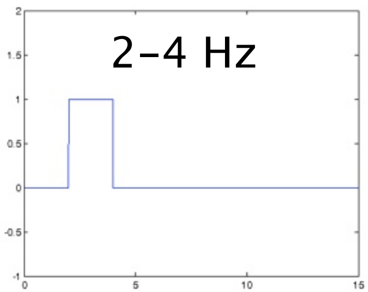
Selective Motion Magnification



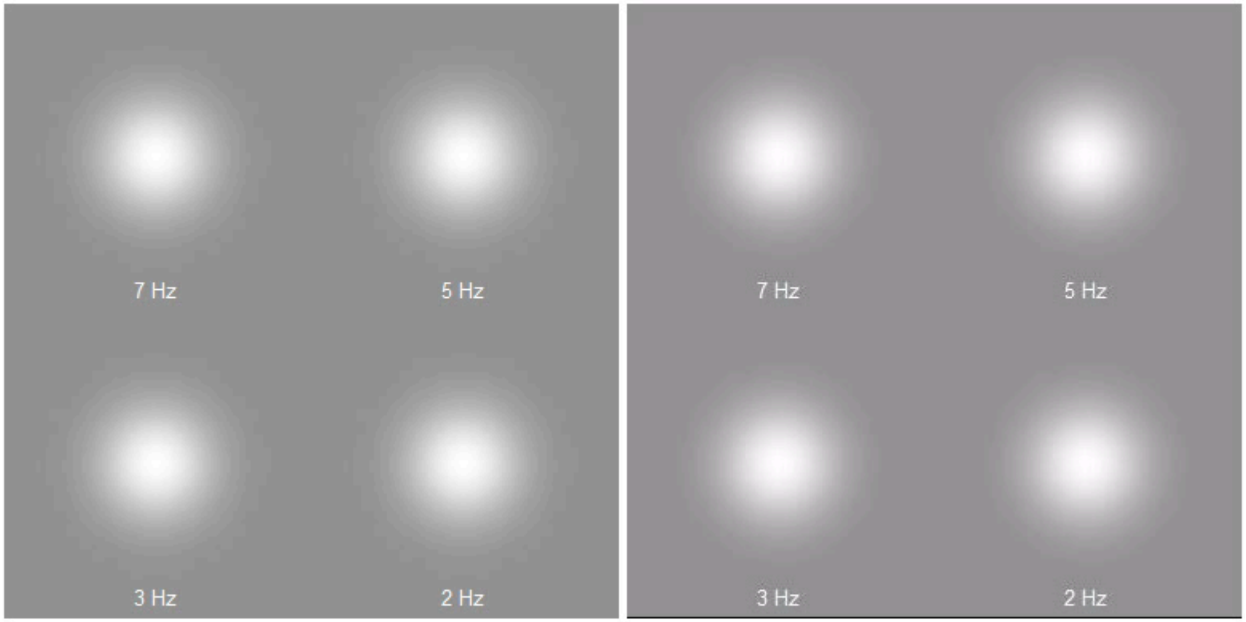
Source

Motion-magnified (3 Hz)

Temporal filter:



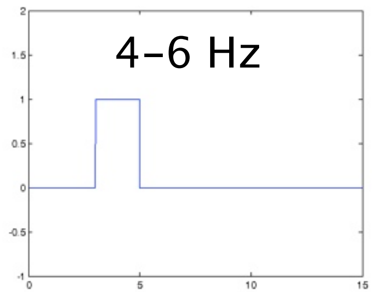
Selective Motion Magnification



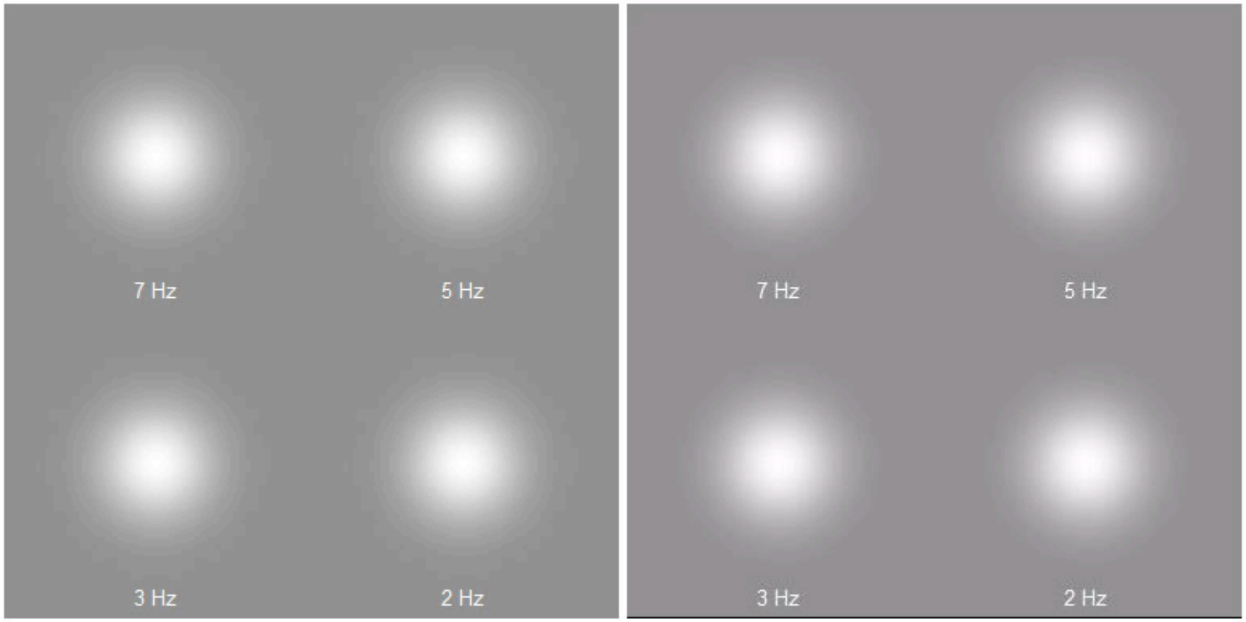
Source

Motion-magnified (5 Hz)

Temporal filter:



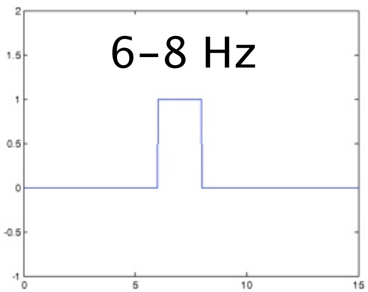
Selective Motion Magnification



Source

Motion-magnified (7 Hz)

Temporal filter:



Selective Motion Magnification in Natural Videos

Source
(600 fps)

72-92 Hz
Amplified

100-120 Hz
Amplified

Low E (82.4 Hz)

A (110 Hz)

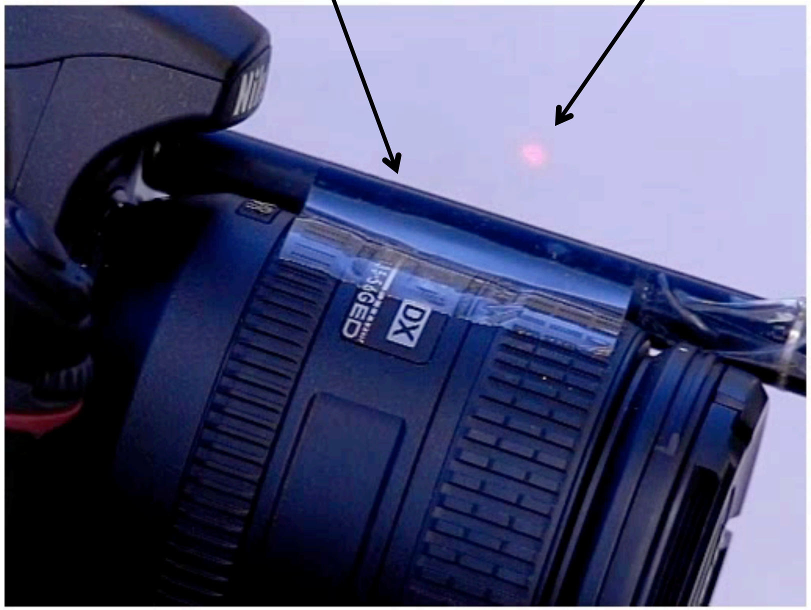
Motion Magnification Results

Source (300 fps)

Motion-magnified (45-100 Hz, x100)

DSLR Controlled Setup

Laser pointer Reflected laser point



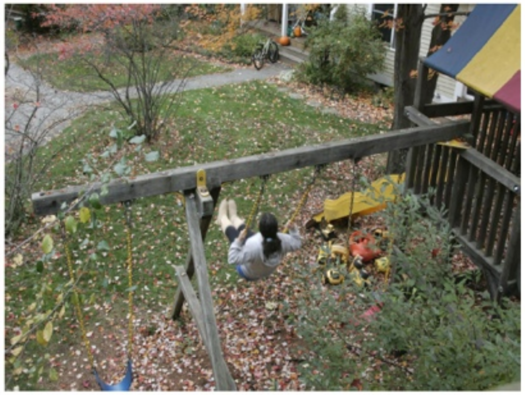
Source (300 fps)

Motion-magnified (45-100 Hz, x100)

Same DSLR camera
(Nikon D400)

Standard A4 sheet

Previous Work on Motion Magnification

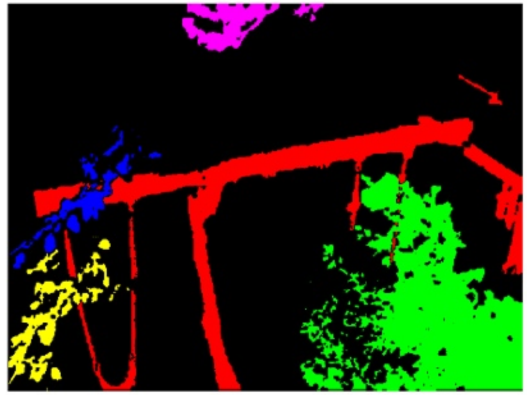


(a) Registered input frame

+

(b) Clustered trajectories of tracked features

+

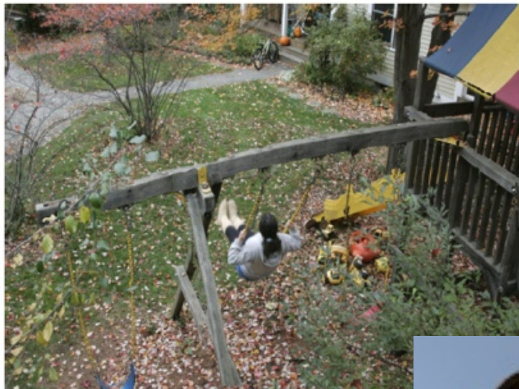


(c) Layers of related motion and appearance

+

(d) Motion magnified, showing holes

+

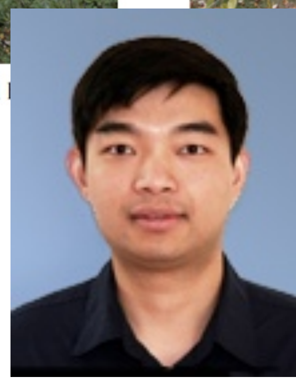


(e) After texture in-painting to fill

+

's modifications to segmentation map in (c)

+



Ce Liu

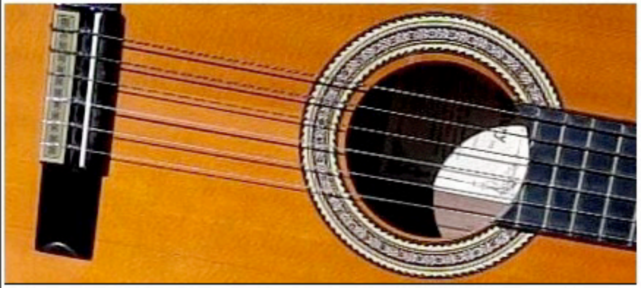
Motion Magnification

Source

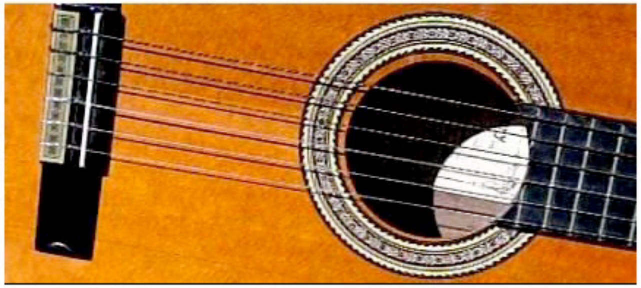
Motion-magnified

Liu, Torralba, Freeman, Durand, Adelson, Motion Magnification, SIGGRAPH 2005

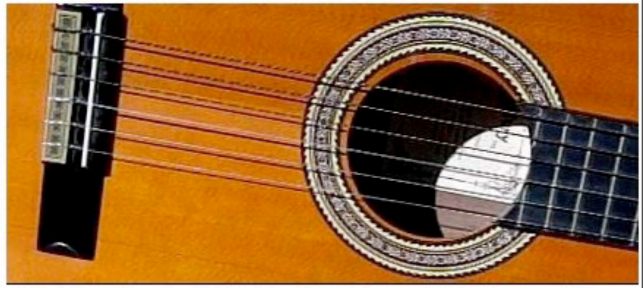
Eulerian vs. Lagrangian Motion Magnification



Source



Eulerian



Lagrangian

Source

Eulerian

Lagrangian

play NSF video

http://www.youtube.com/watch?v=Rm1UMbxnq-8

SIGGRAPH video

http://www.youtube.com/watch?feature=player_embedded&v=ONZcjs1Pjmk

and YouTube video

http://www.youtube.com/watch?v=J1wvFmWv7zY

Main Contributions

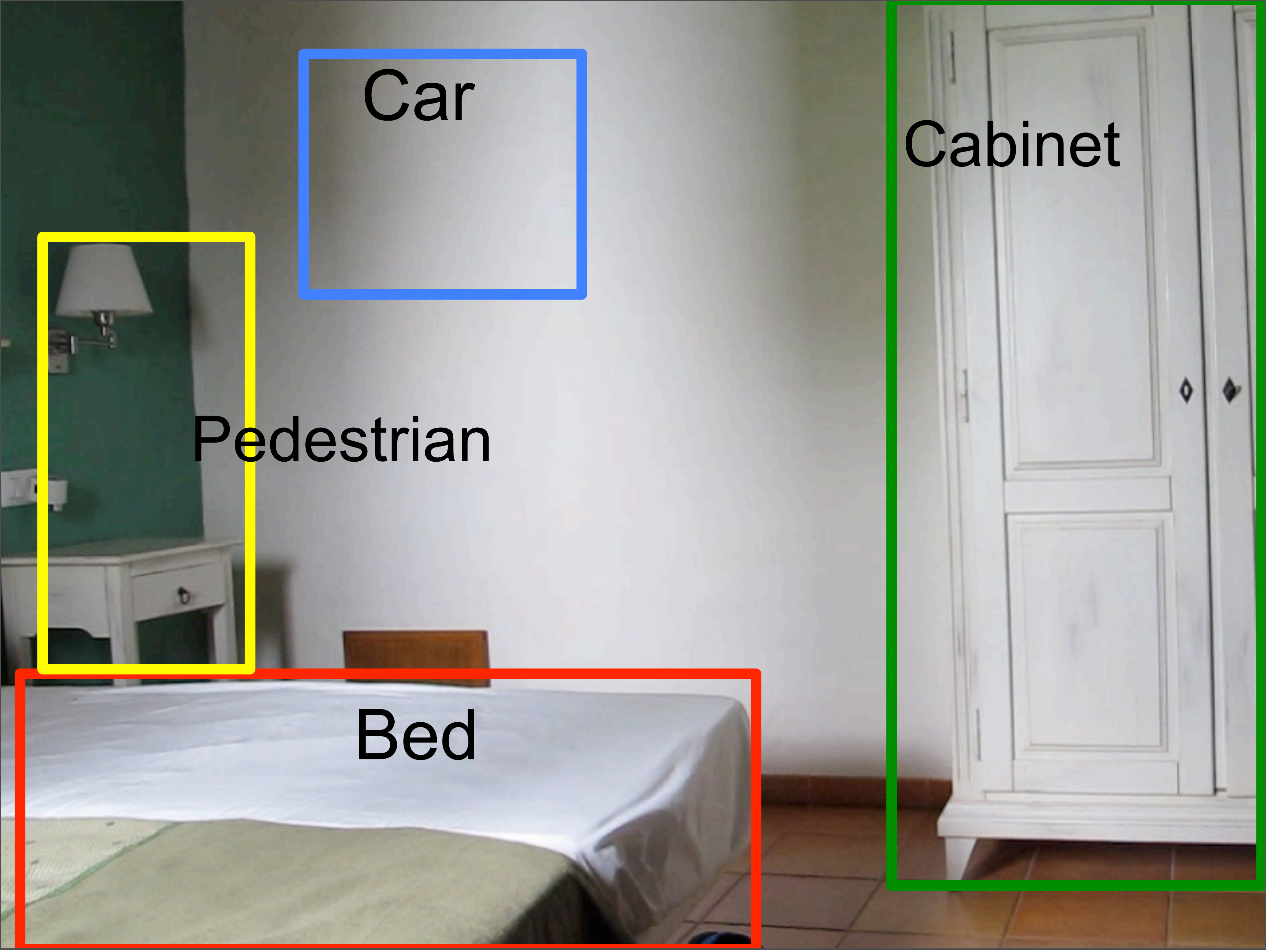
- Applying spatial decomposition and temporal filtering to standard videos can enhance details that are difficult or impossible to see with the naked eye
  - Eulerian perspective - properties of pixels/voxels over time
- Temporal filtering can be used to amplify spatial motion
  - No explicit motion estimation or tracking required
- There’s a big world of tiny motions out there.
- **“MIT's New Video Technology Could Give You Superhuman Sight”**

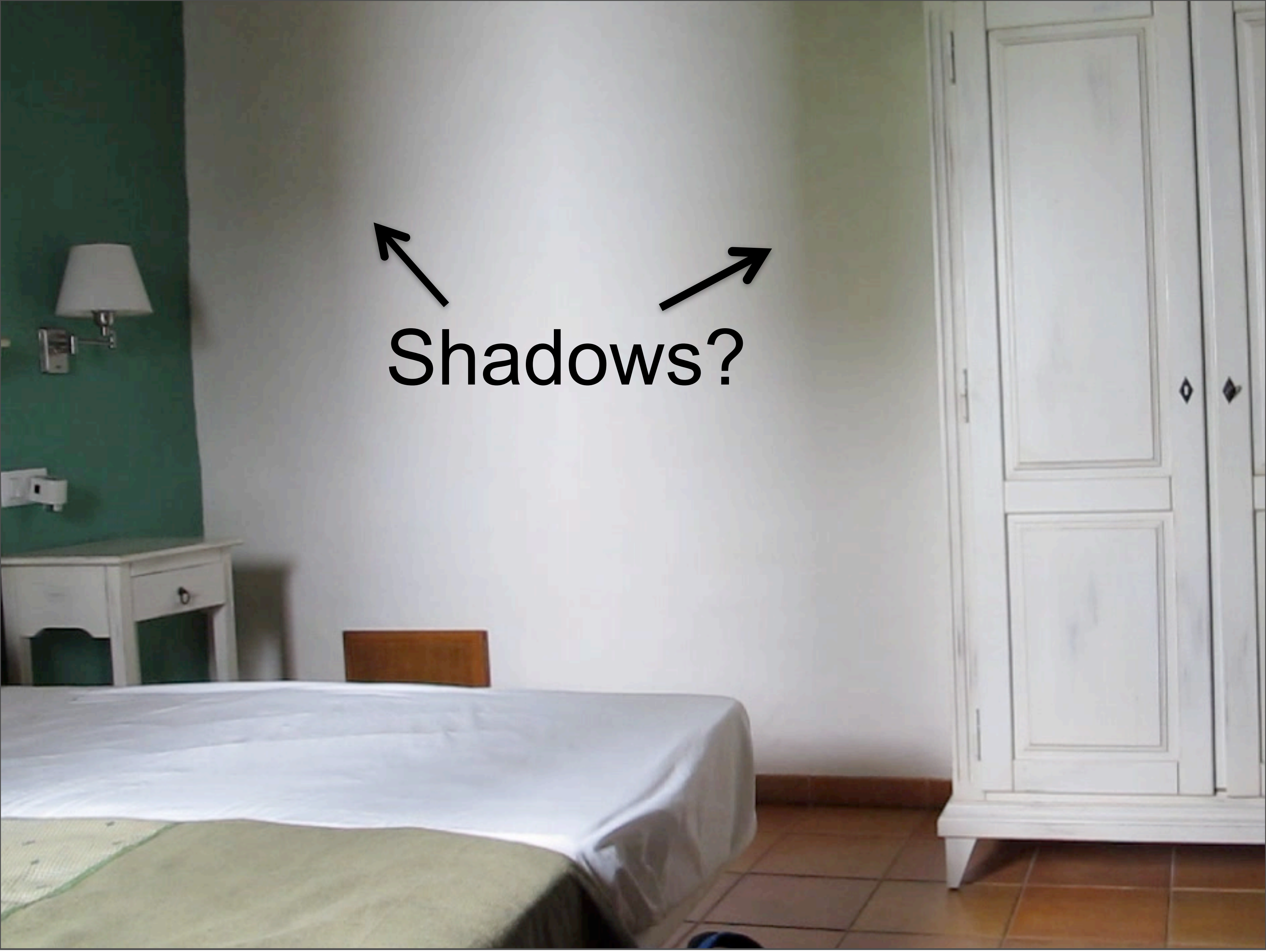
Accidental pinhole and pinspeck cameras: Revealing the scene outside the picture

Antonio Torralba
William T. Freeman

See project page for videos:
http://people.csail.mit.edu/torralba/research/accidentalcameras/

CVPR 2012


Car
Cabinet
Pedestrian
Bed

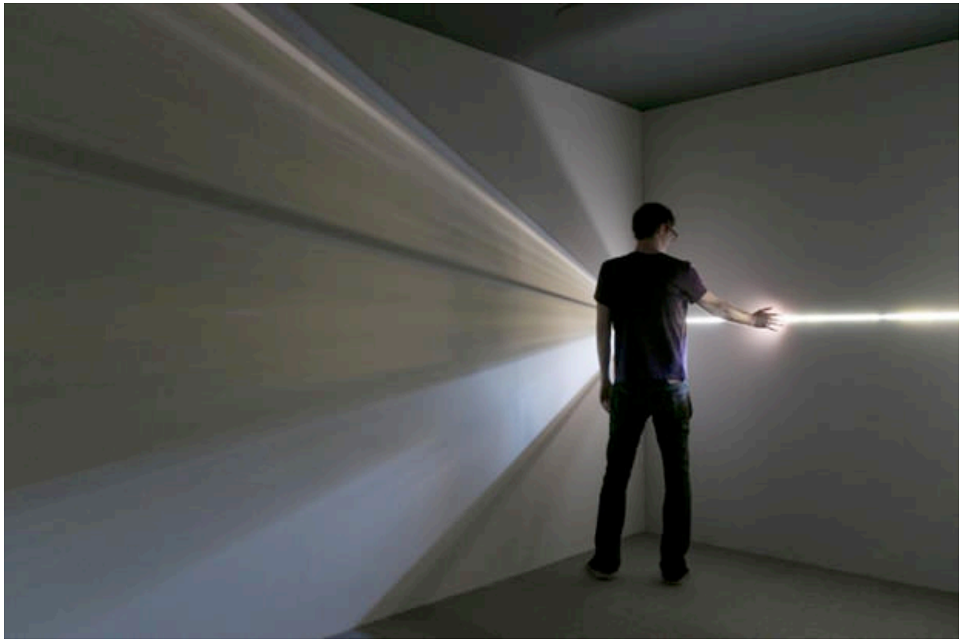

Shadows?

Accidental pinhole camera

Window turned into a pinhole

View outside

Source: wikipedia



Chris Fraser

”a camera obscura has been used ... to bring images from the outside into a darkened room”

Aberlado Morell

Window open

Window turned into a pinhole

Hotel room,
contrast enhanced.

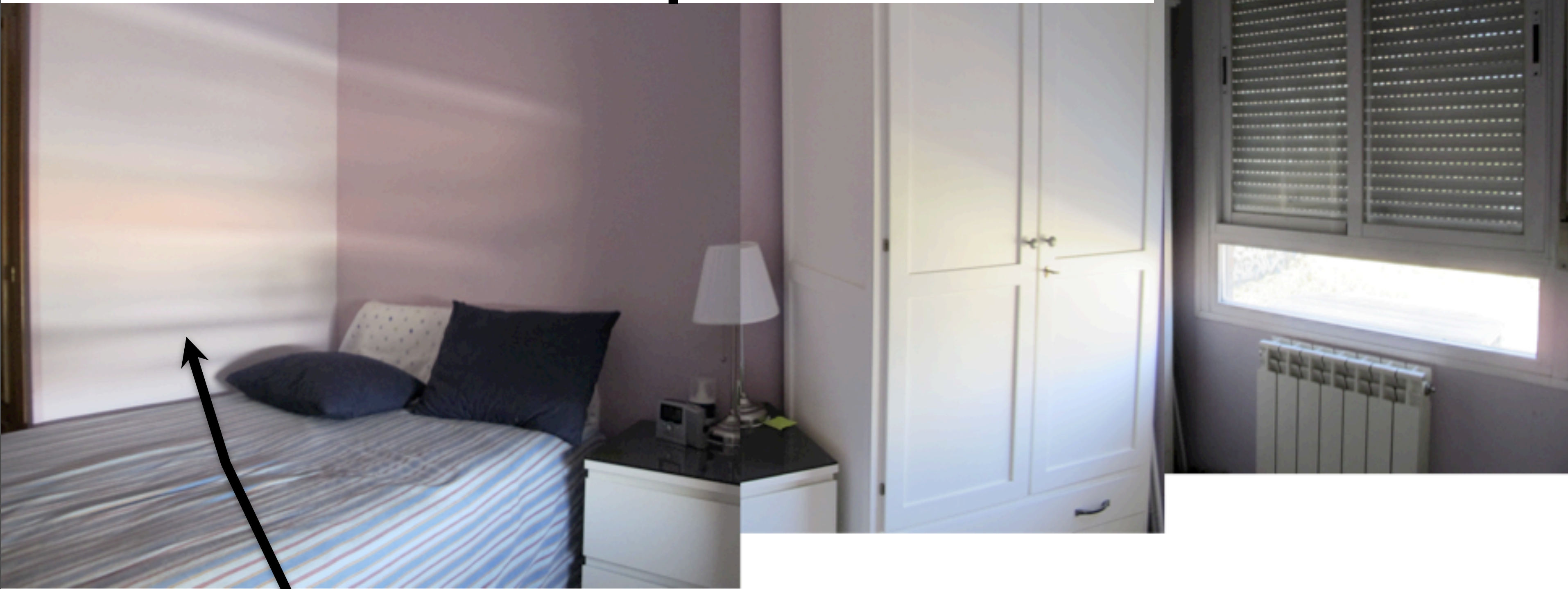
The view from the window

Accidental pinholes produce images that are unnoticed or misinterpreted as shadows

Accidental pinhole

If the window is big then the images recovered are too blurry

Accidental pinhole



Outside scene

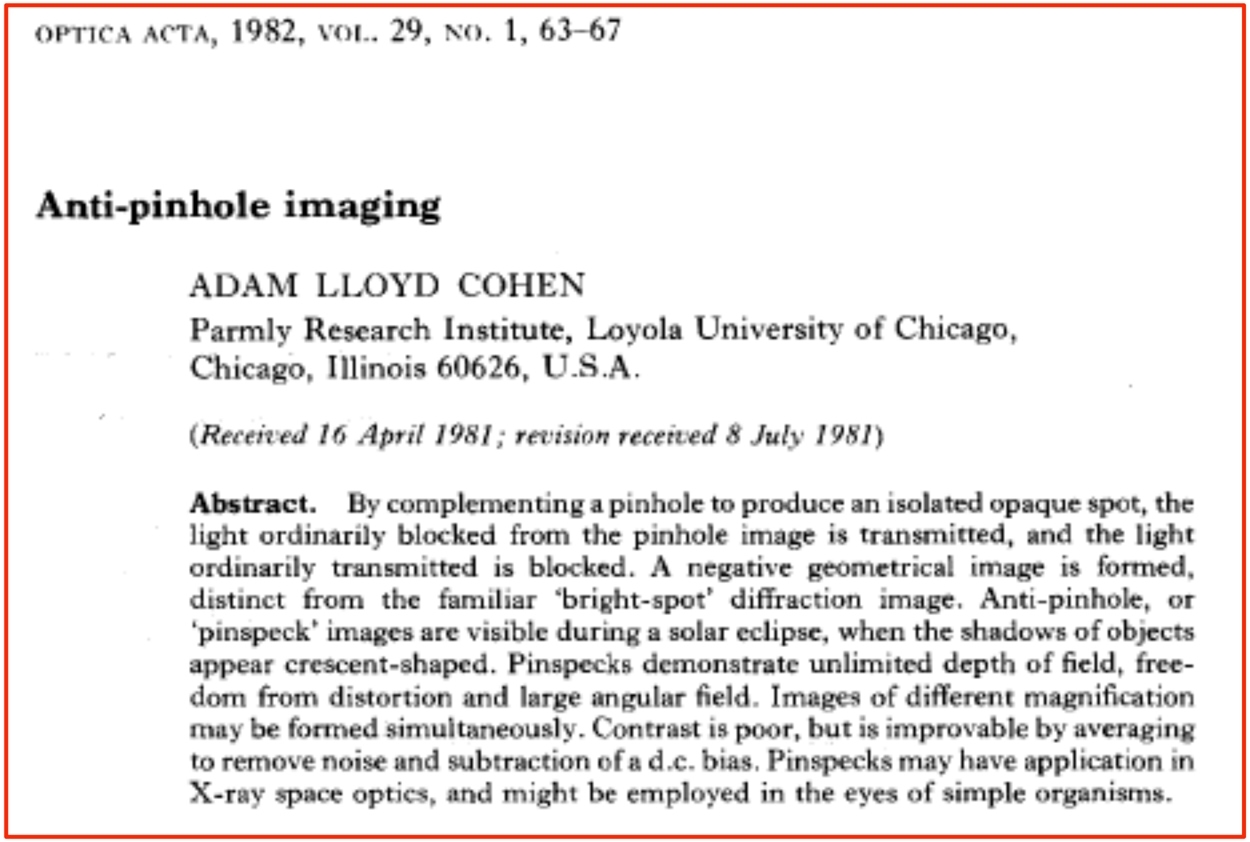
*

Aperture

See Zomet, A.; Nayar, S.K. CVPR 2006 for a detailed analysis.

Anti-pinhole or Pinspeck

Adam L. Cohen, 1982



OPTICA ACTA, 1982, VOL. 29, NO. 1, 63–67

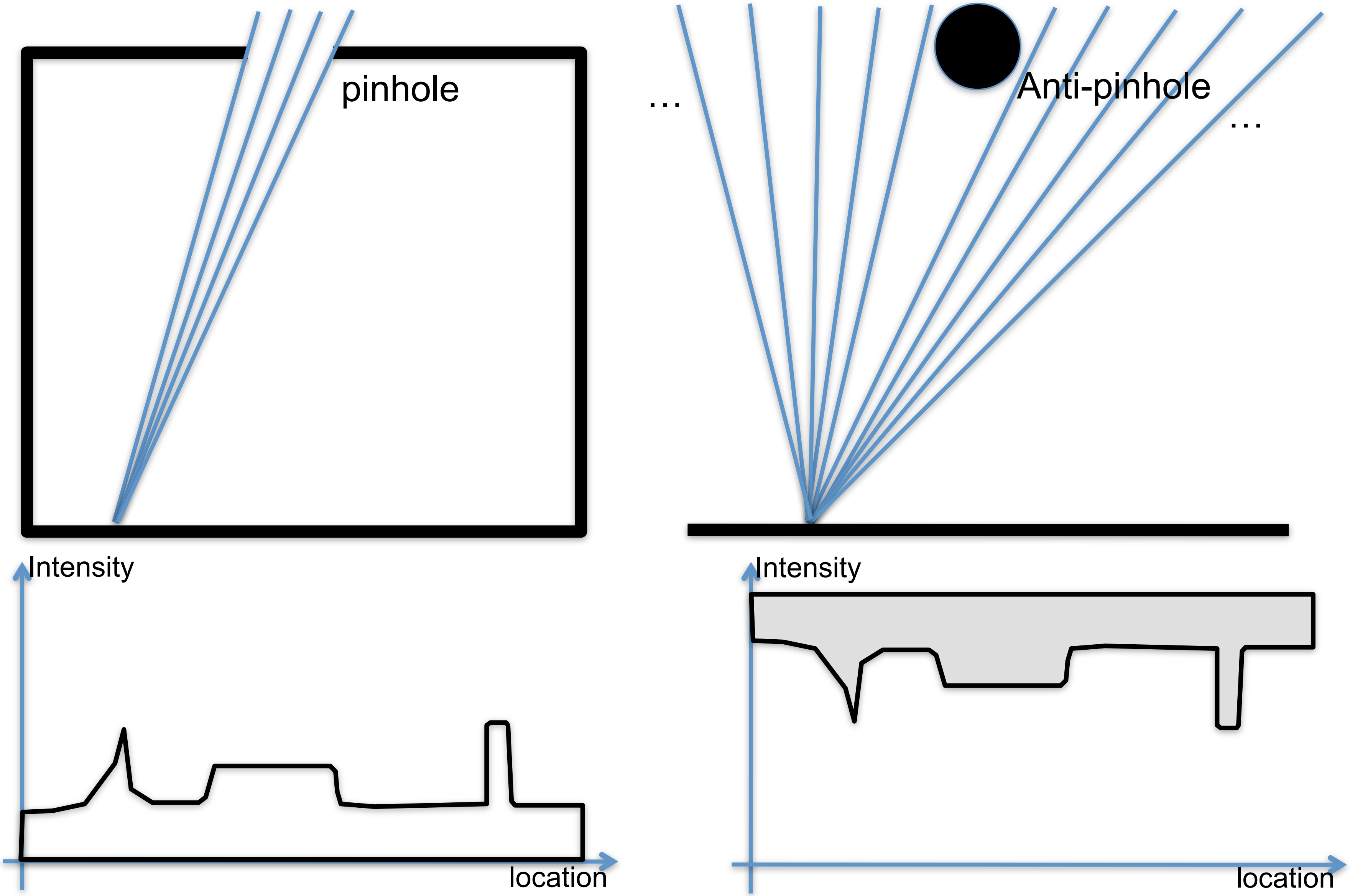
Anti-pinhole imaging

ADAM LLOYD COHEN
Parmly Research Institute, Loyola University of Chicago,
Chicago, Illinois 60626, U.S.A.

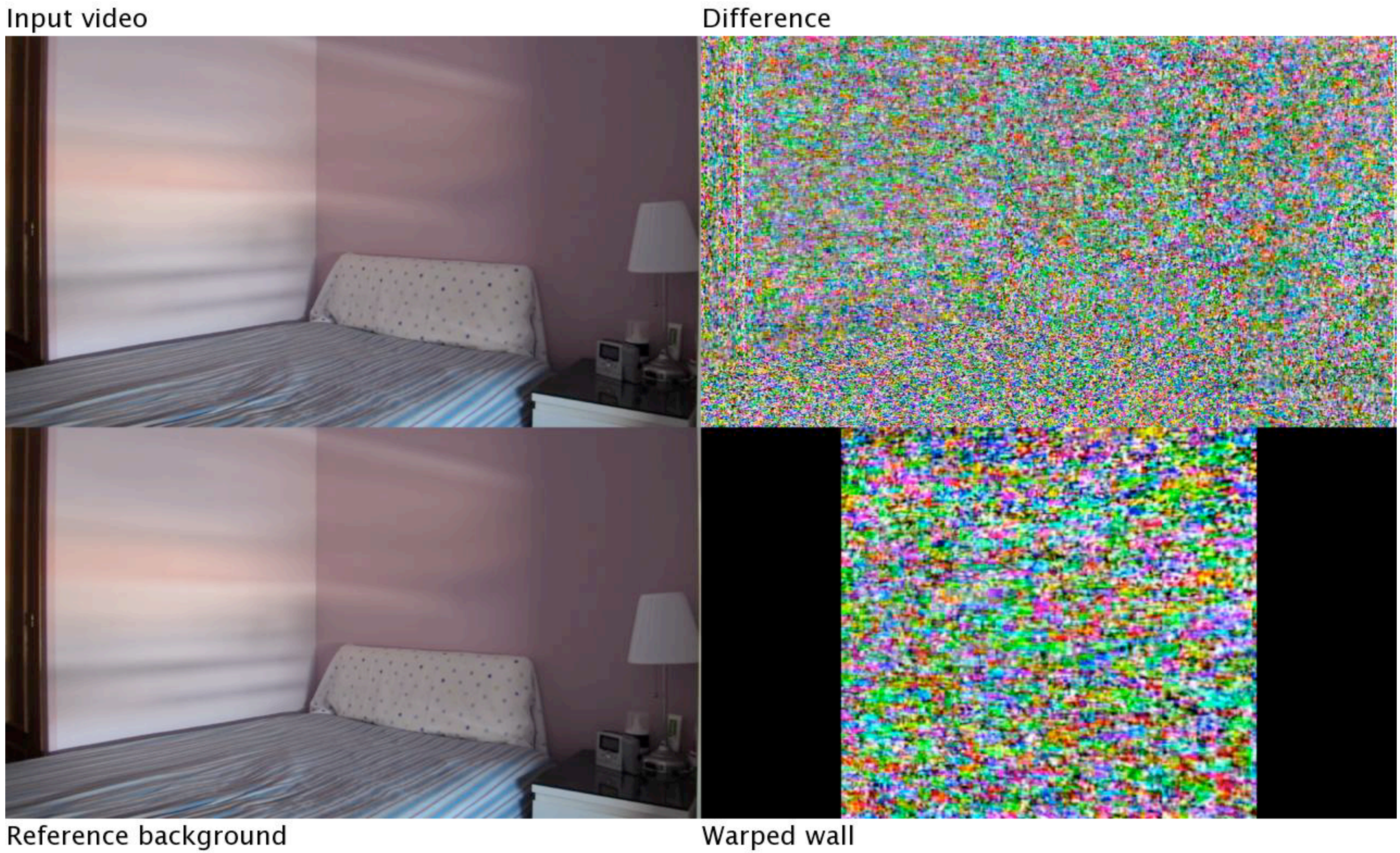
(*Received 16 April 1981; revision received 8 July 1981*)

Abstract. By complementing a pinhole to produce an isolated opaque spot, the light ordinarily blocked from the pinhole image is transmitted, and the light ordinarily transmitted is blocked. A negative geometrical image is formed, distinct from the familiar 'bright-spot' diffraction image. Anti-pinhole, or 'pinspeck' images are visible during a solar eclipse, when the shadows of objects appear crescent-shaped. Pinspecks demonstrate unlimited depth of field, freedom from distortion and large angular field. Images of different magnification may be formed simultaneously. Contrast is poor, but is improvable by averaging to remove noise and subtraction of a d.c. bias. Pinspecks may have application in X-ray space optics, and might be employed in the eyes of simple organisms.

Pinhole and Anti-pinhole



Adam L. Cohen, 1982


Input video
Difference
Reference background
Warped wall

Natural eyes

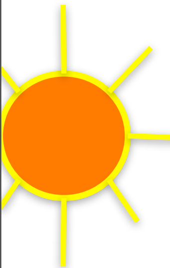
Lenses

Pinholes

Anti-pinholes

Like other Euglenoids, *Euglena* possess a red eyespot, an organelle composed of carotenoid pigment granules. The red spot itself is not thought to be photosensitive. Rather, it filters the sunlight that falls on a light-detecting structure at the base of the flagellum (a swelling, known as the paraflagellar body), allowing only certain wavelengths of light to reach it. As the cell rotates with respect to the light source, the eyespot partially blocks the source, permitting the *Euglena* to find the light and move toward it (a process known as phototaxis).[11]

Shadows
Accidental anti-pinhole cameras



Background image

-

Input video

=

Negative of the shadow

Background image

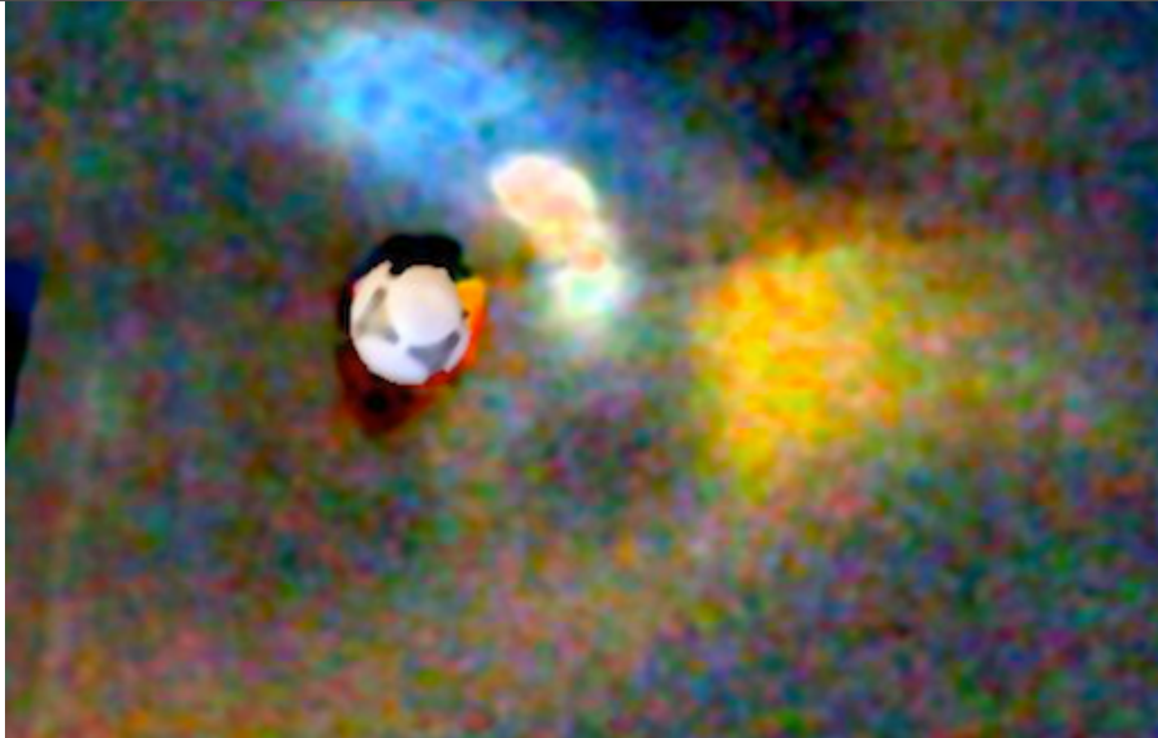
-

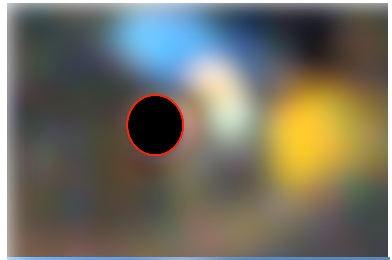
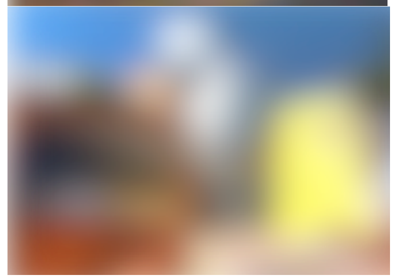
Input video

= Negative of the shadow

Input video

Negative of the shadow



The importance of the size of the occluder

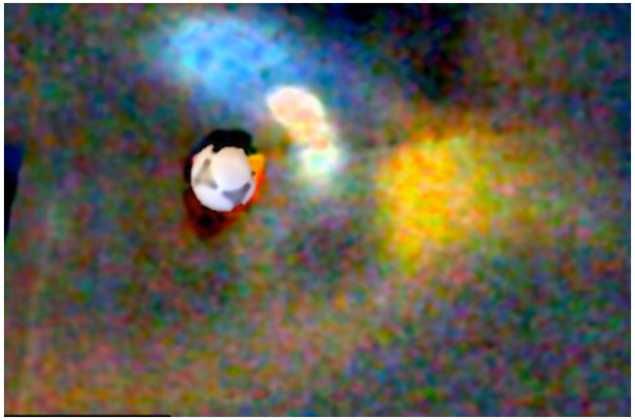
Input video

Negative of the shadow

Size of the occluder

Antonio

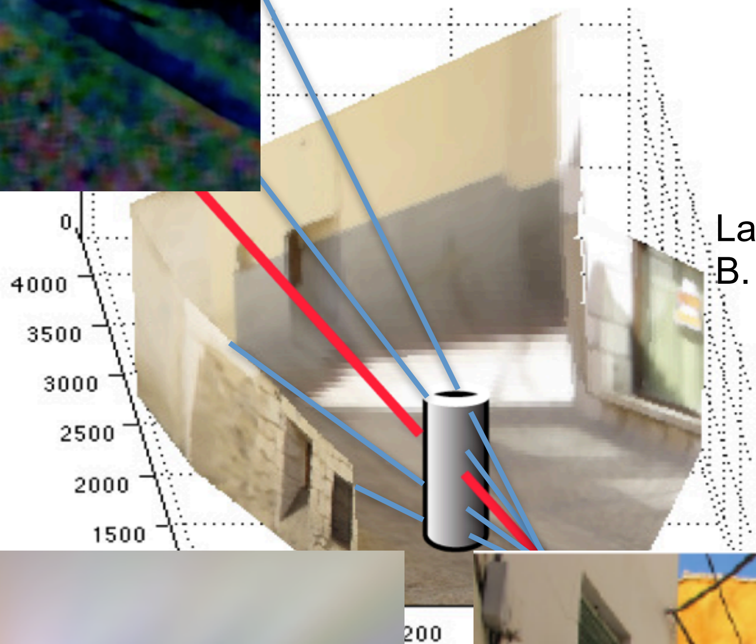
Ball



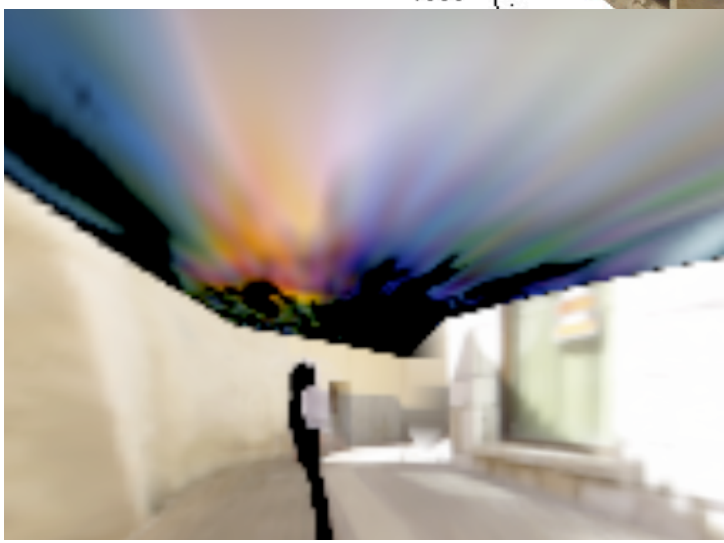
Input video

Negative of the shadow

Using some single view metrology. A. Criminisi, I. Reid, and A. Zisserman 1999



Labelme 3D toolbox.
B. Russell, A. Torralba. CVPR 2009



Mixed accidental pinhole and anti-pinhole cameras

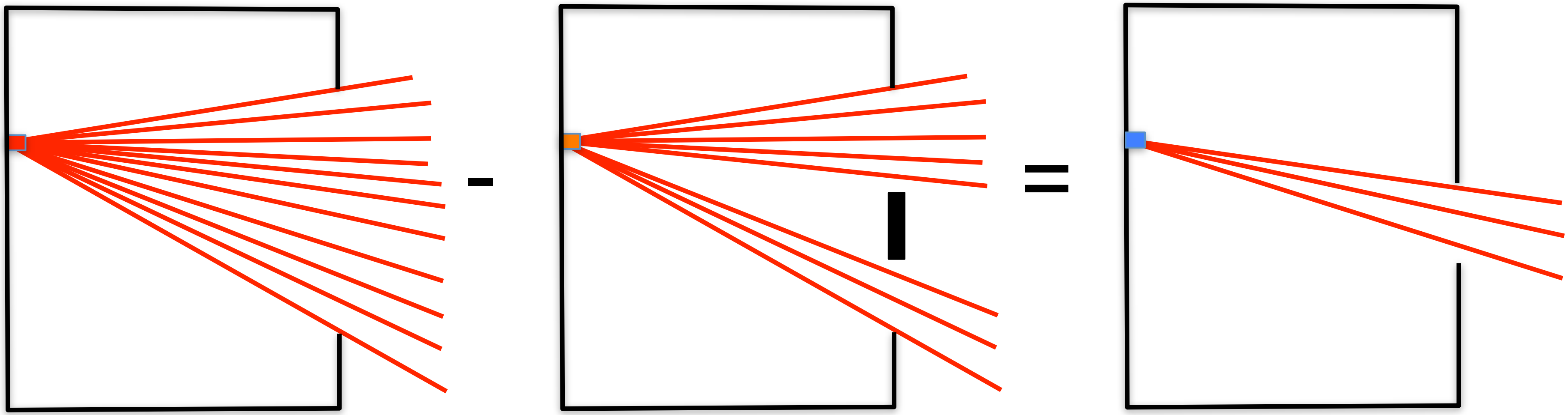
Mixed accidental pinhole and anti-pinhole cameras

Mixed accidental pinhole and anti-pinhole cameras

Room with a window

Person in front of the window

Difference image



-

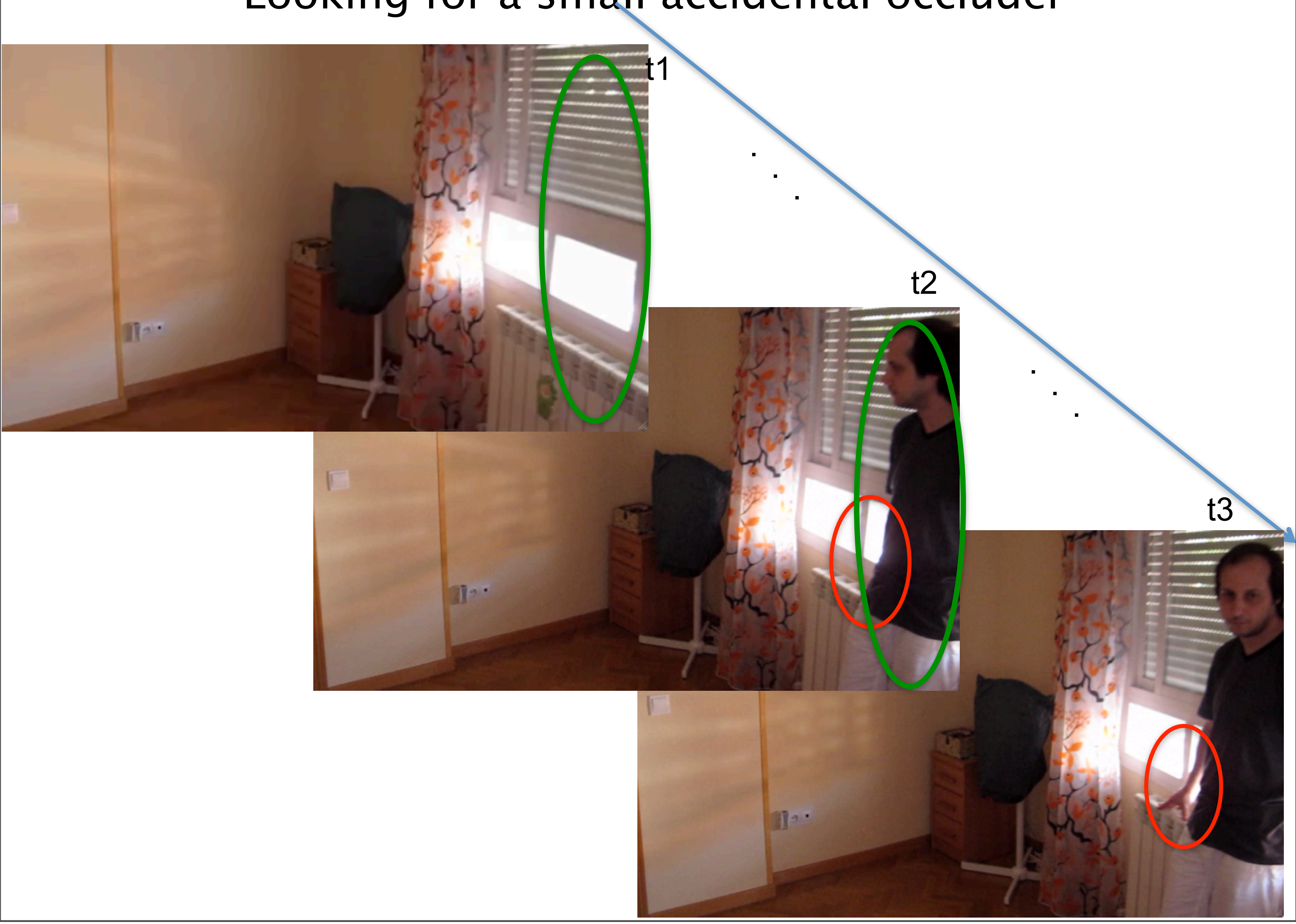
= ?

Mixed accidental pinhole and anti-pinhole cameras

Body as the occluder

View outside the window

Looking for a small accidental occluder



Reference

-

Video

=

Looking for a small accidental occluder

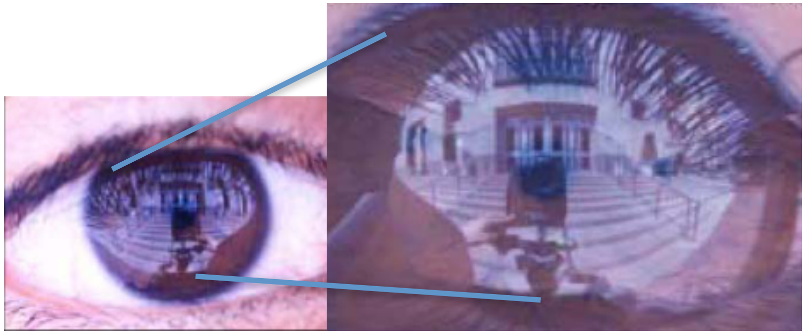
Body as the occluder

Hand as the occluder

View outside the window

Accidental cameras reveal

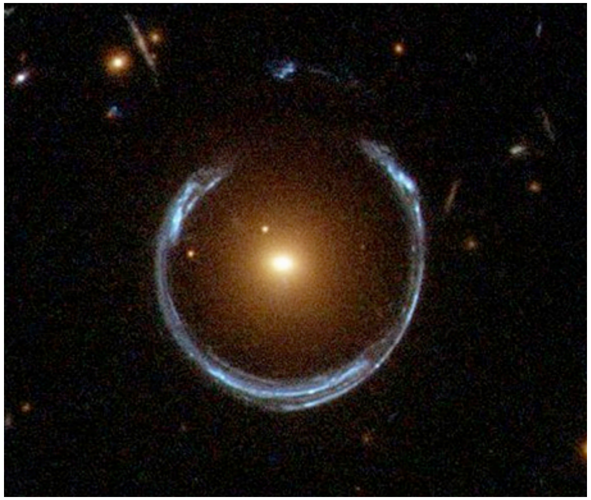
Mirrors



Nishino and Nayar, IJCV 2006

Pinholes

Lenses



Gravitational lensing

Anti-pinholes

Conclusion

Shadows and apertures produce accidental images that are unnoticed most of the time. Accidental cameras can reveal the scene outside the picture.

Applications:

- Image forensics (J. O’Brian & H. Farid, 2012)

- Computer graphics providing better light models

Funding for this work was provided by NSF Career award 0747120 to A.T, and NSF CGV 1111415 and NSF CGV 0964004 to W.T.F.

Small motions and small photometric changes reveal otherwise unseen features of the world both in and outside the picture frame

small motions

small intensity changes