



Content creation for light-field displays

Atanas Gotchev

3D Media Group

Department of Signal Processing

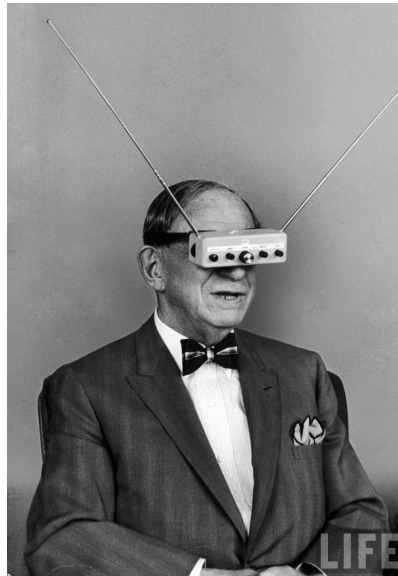
Tampere University of Technology

Towards Full Parallax Imaging

3D MEDIA GROUP OVERVIEW



Research Motivation



- Higher realism in representing visual scenes
- Better interactivity
- New forms of communication and display

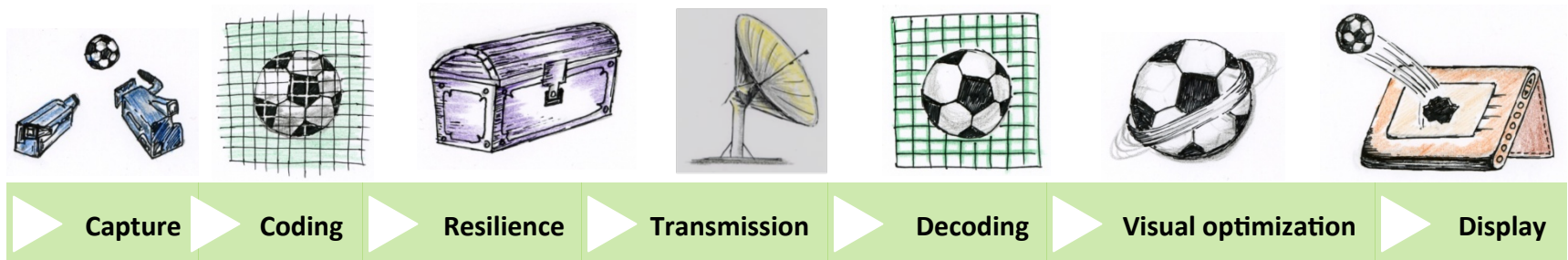
The group

- An associate professor, two doctors (one Marie Curie researcher), five researchers (doctoral students), two visiting researchers, four research assistants (MSc students)



Core Expertise

- Methods for capture, processing and display of real world visual 3D data
- Knowledge about perception of depth and visual cues
- Addresses specific problems related with
 - Massive amounts of data to be processed
 - Optimal visualization on state-of-the-art 3D displays



3D Media Lab (currently dismantled)



- 24 HD cameras, scalable multi-camera rigs
- Range sensors: PMD Tech ToF, Kinect
- State of the art 3D displays
- Currently being transformed into **Research Centre for Immersive Visual Technologies (CIViT)**



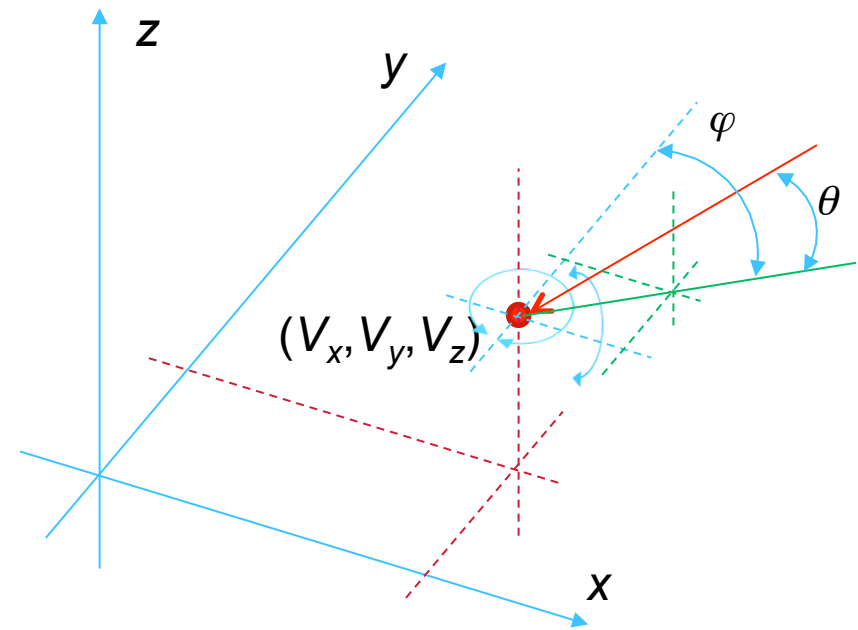
Ray-space aka Epipolar-plane images

LIGHT FIELD REPRESENTATION



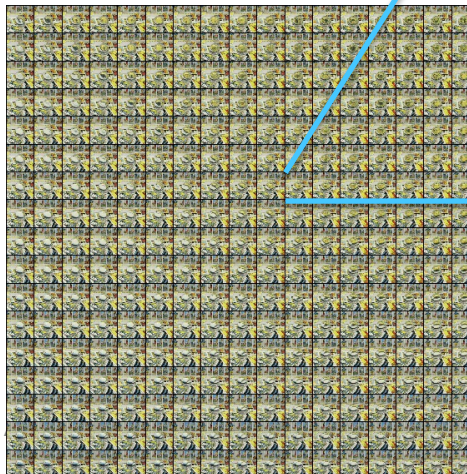
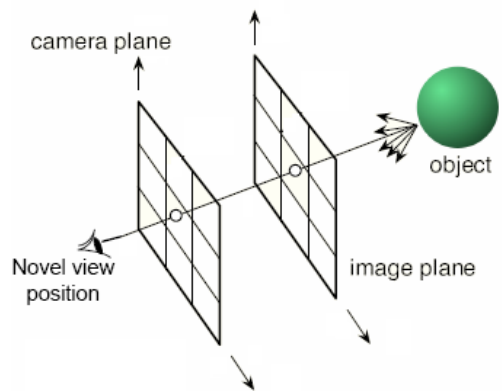
Plenoptic function (PF)

- Introduced by Adelson and Bergen (1991)
 - *Plenus* (complete) + *Optic* = Plenoptic
- Continuous function that describes the amount of light through every point in space in every direction
 - Describes the 'light field' in terms of ray optics
- 7-D function $P(\theta, \varphi, \lambda, t, V_x, V_y, V_z)$
 - (V_x, V_y, V_z) – location in 3D space
 - (θ, φ) – angles determining the direction
 - λ – wavelength
 - t – time



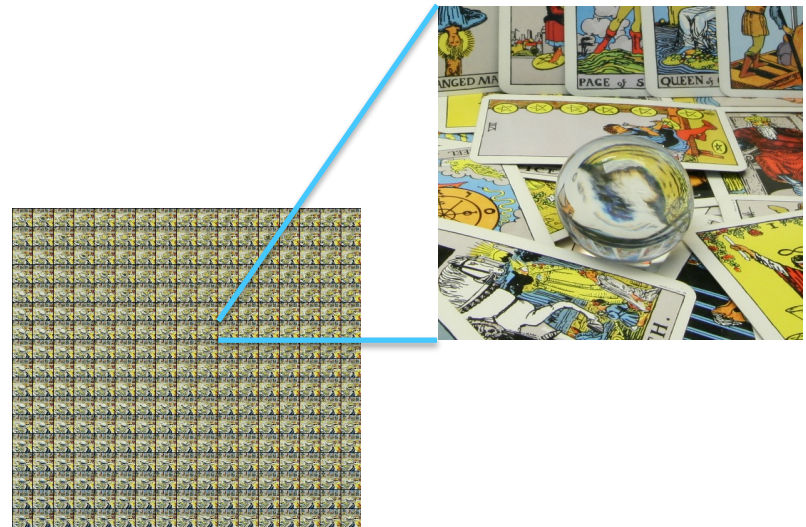
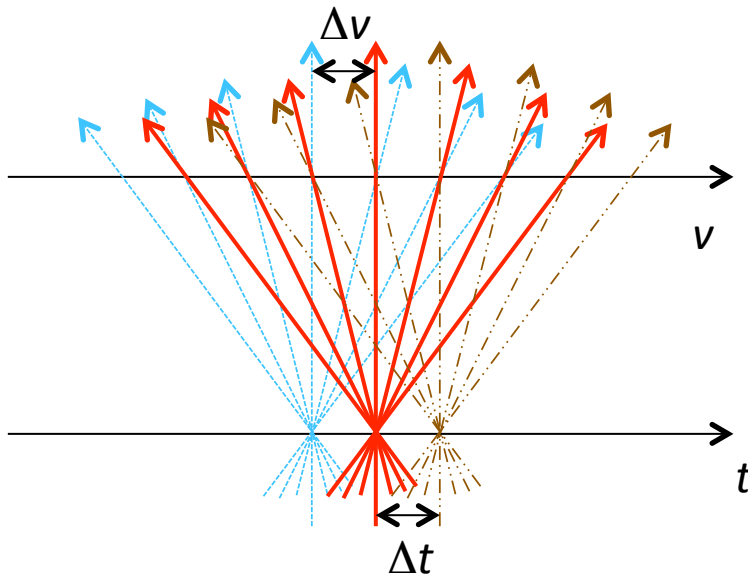
4D Light Field

- 4D Approximation of PF:
 - Replace wave with RGB;
 - Consider static scenes;
 - Constrain the viewing position
- Two-plane parameterization of the 4D LF
 - Parallel planes $L(u,v,s,t)$ (also called light slabs)
 - Levoy and Hanrahan (1996) – light field
 - Gortler et al. (1996) – Lumigraph

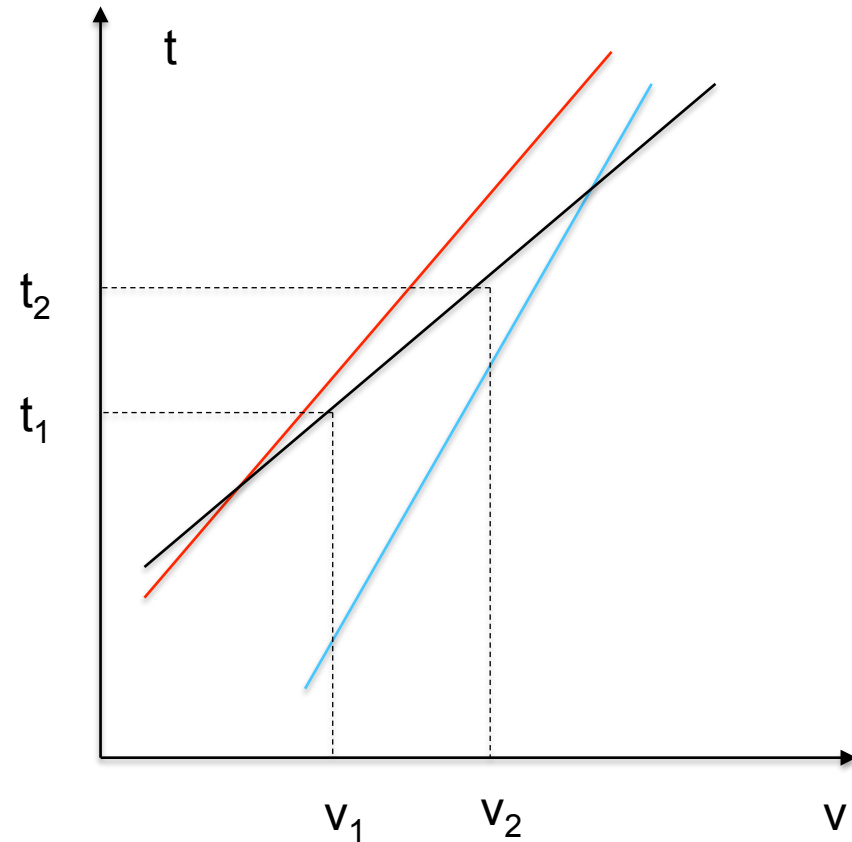
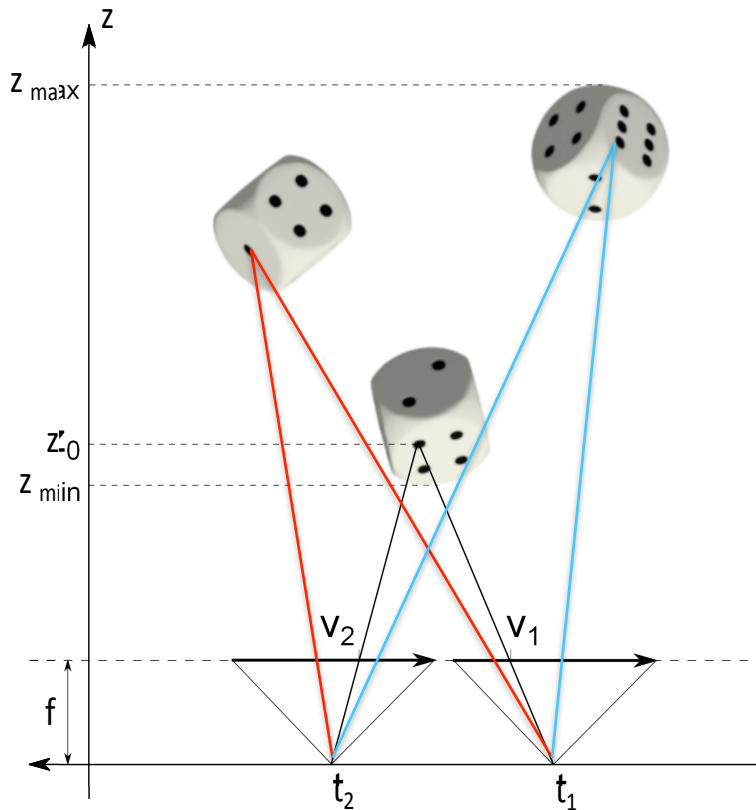


4D Light Field

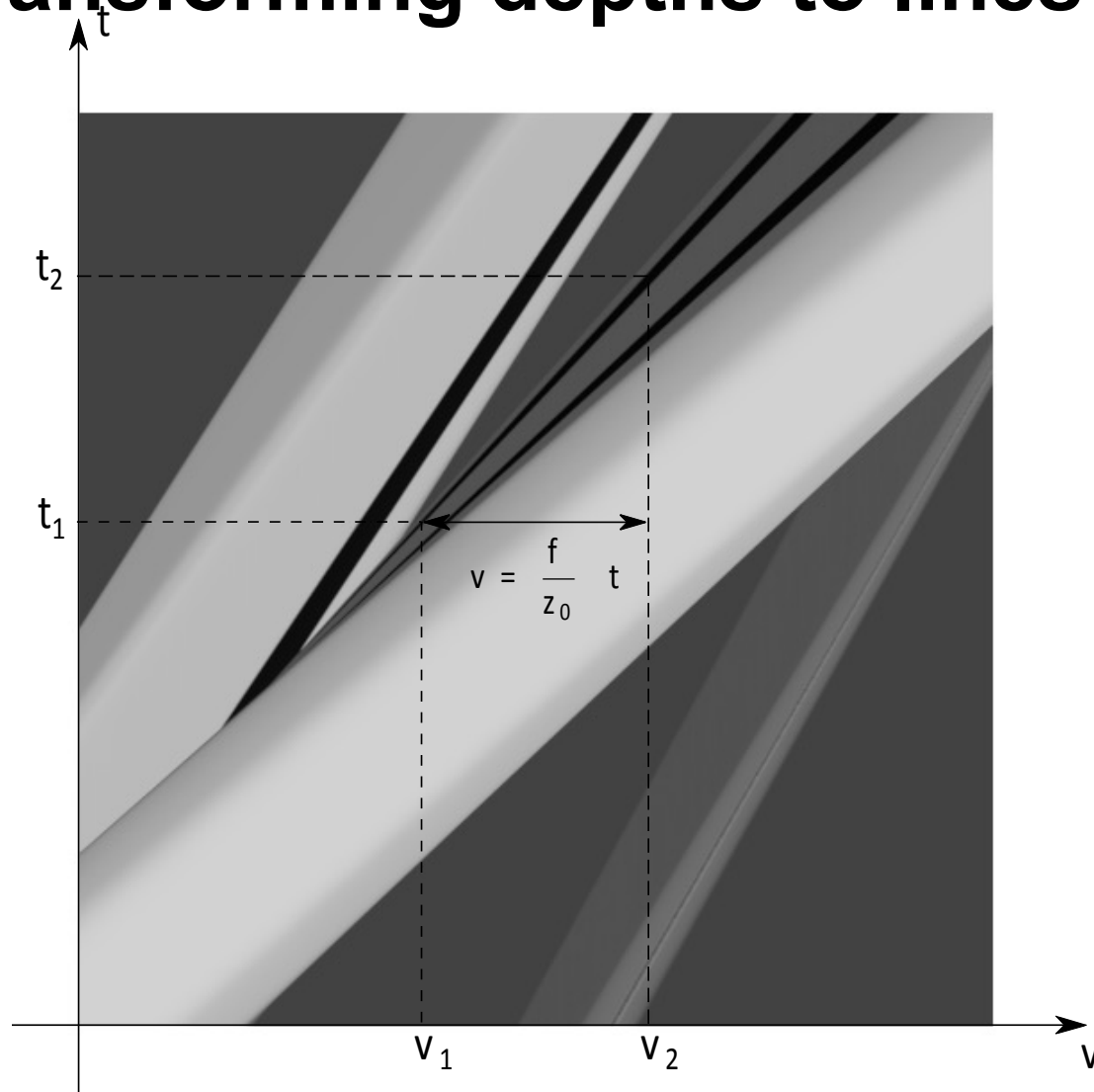
- The function $P(u, v, s, t)$ defines the intensity of the light ray at the pixel position (u, v) and camera position (s, t)



EPI: Transforming depths to lines

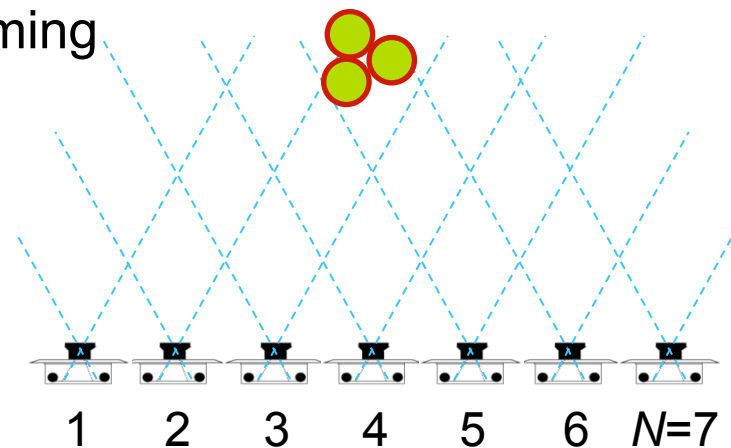
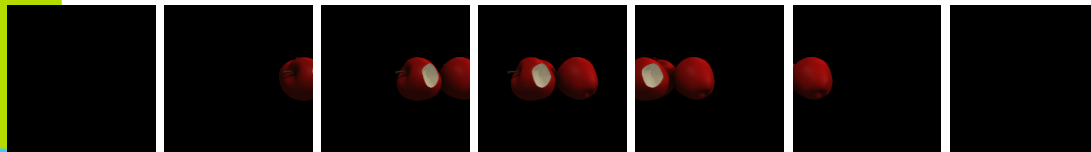


EPI: Transforming depths to lines

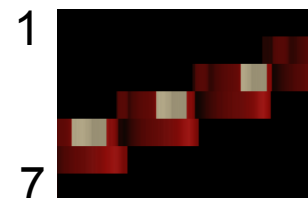
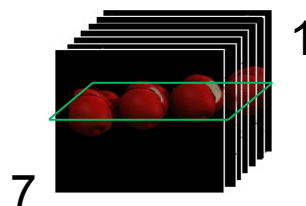


Epipolar plane image (EPI)

- Consider a horizontal camera array forming N images of resolution R_x by R_y



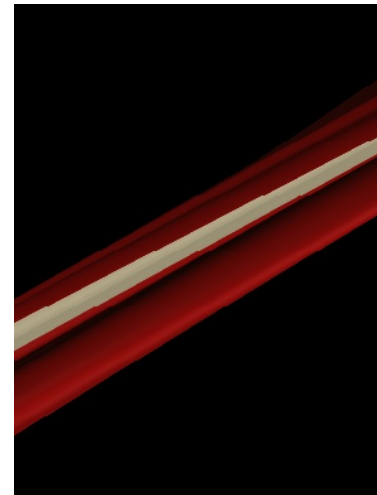
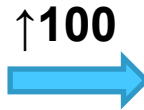
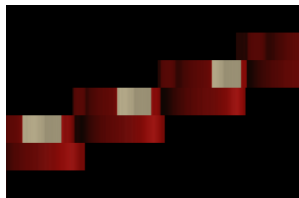
- Stack images to get a volume – a slice of that volume represents an epipolar image (EPI)



EPI and rendering

- 600 instead of 7 cameras

7 rows

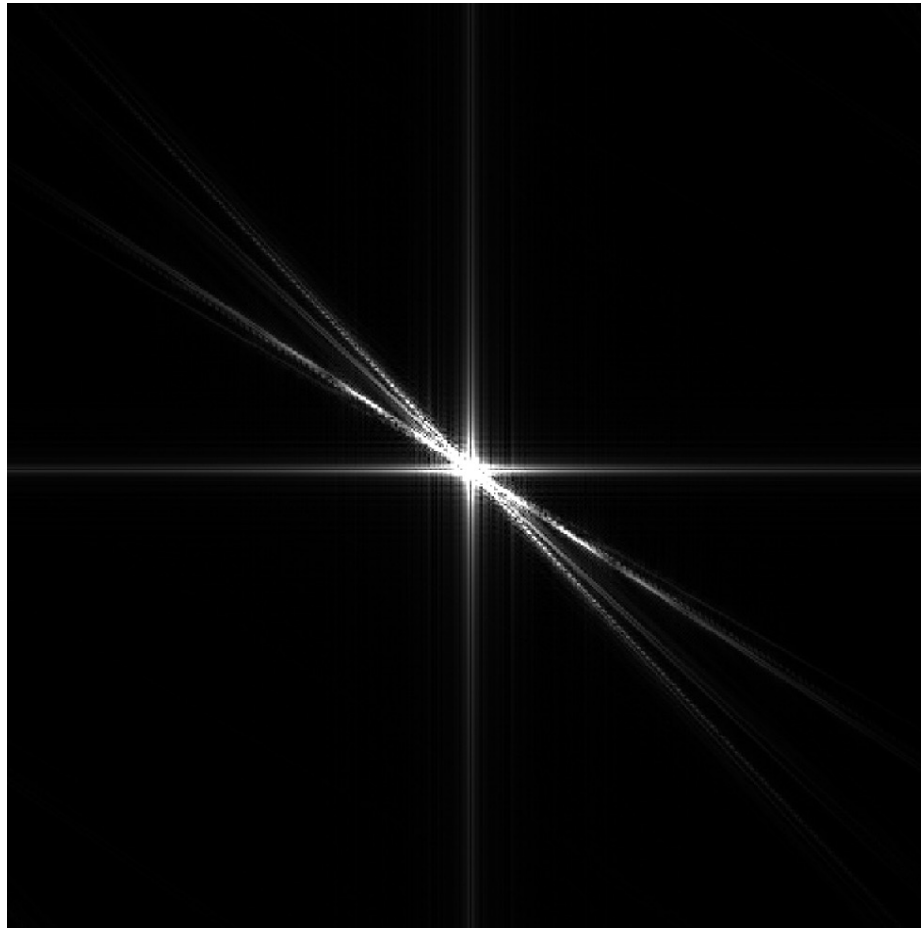


600 rows

- Rendering becomes interpolation of a desired set of cameras from a finite set of cameras

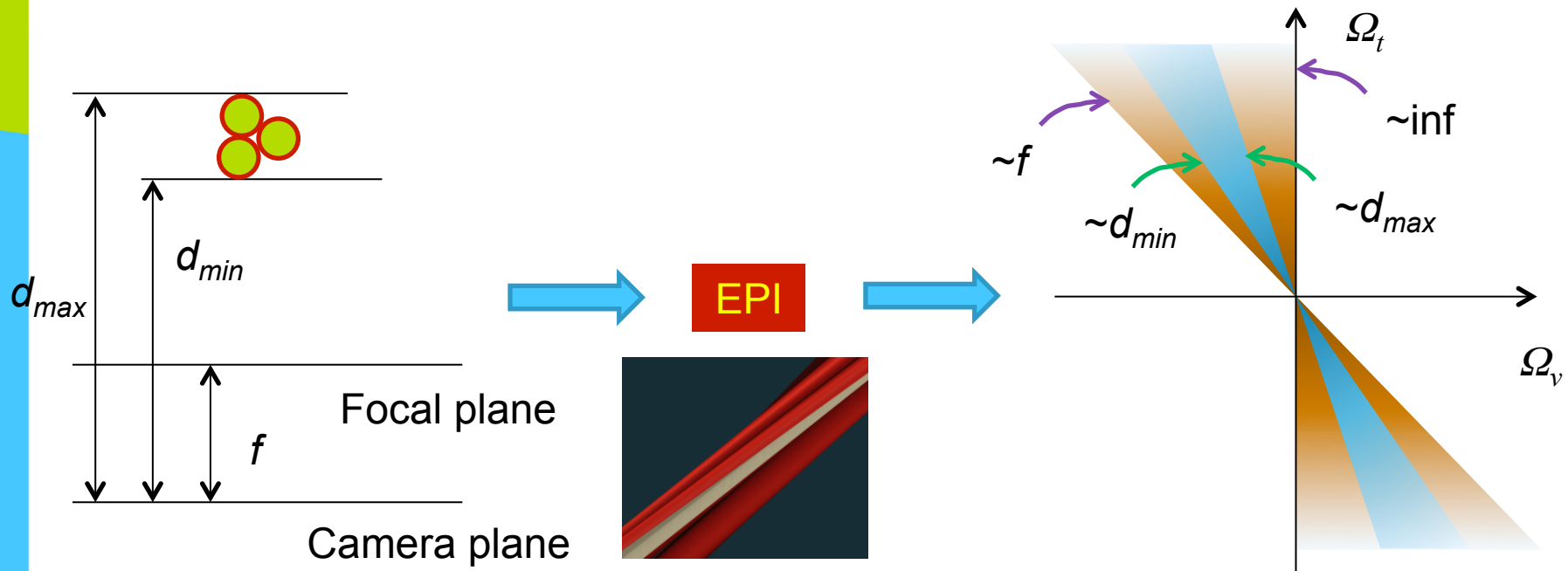
EPI in Fourier domain

- Lines in EPI transforms to lines in Fourier domain



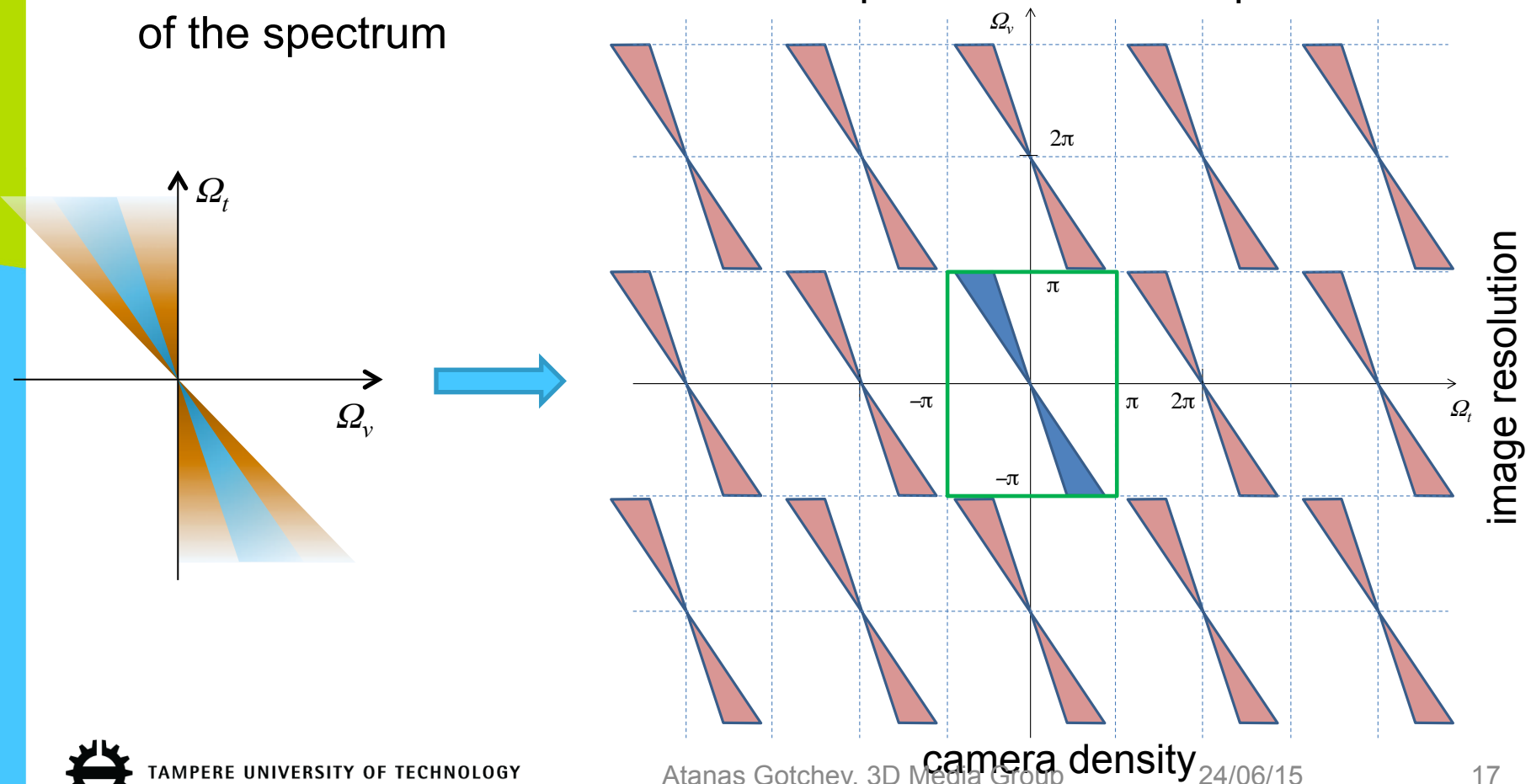
EPI – frequency support vs. scene depth

- Fourier transform of an EPI is limited by two lines determined by the min and max scene depth
 - Limited to non-occluded Lambertian scenes



EPI – frequency support vs. scene depth

- Finite number of cameras with finite pixel count lead to periodization of the spectrum



LF representation: recap

- We need a representation format efficiently representing EPIs
- We need a reconstruction method which reconstructs a densely sampled LF out of sparse camera sets and going beyond Nyquist



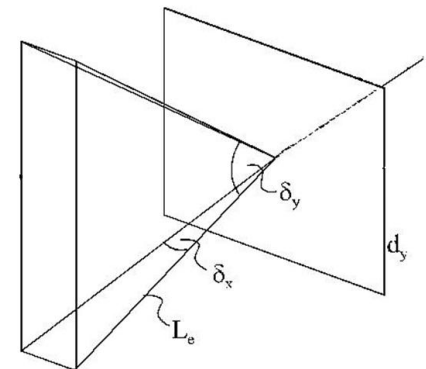
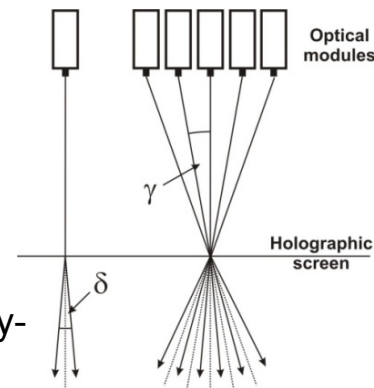
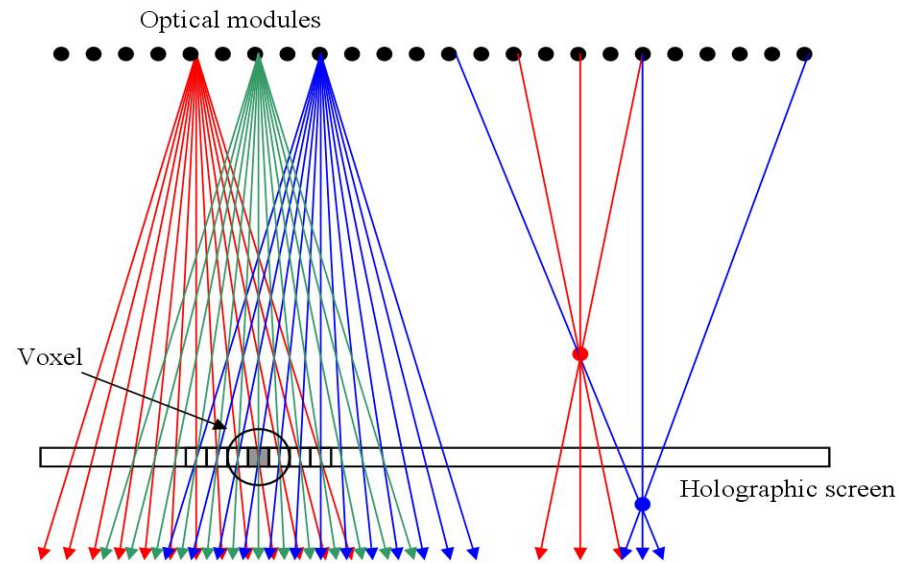
A match of sampling grids

DISPLAY SPECIFIC LIGHT FIELD ANALYSIS



Modeling of projector based LF displays

- Optical modules
 - light beams to hit the screen under with various angles of incidence
- Holographic screen
 - direction selective property
 - the screen diffusion angle δ is equal to the angle γ between the neighboring modules
- Horizontal-parallax only

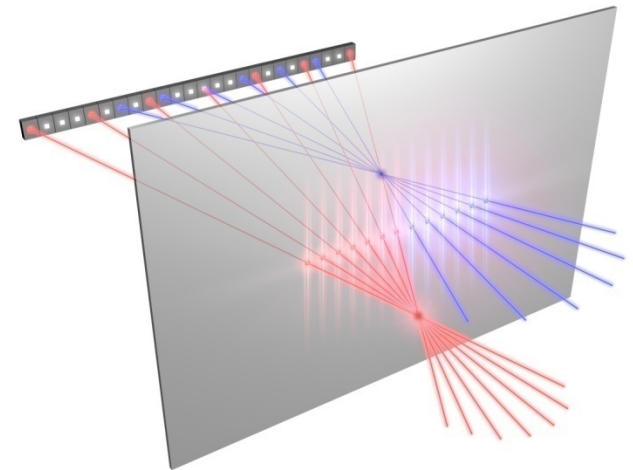
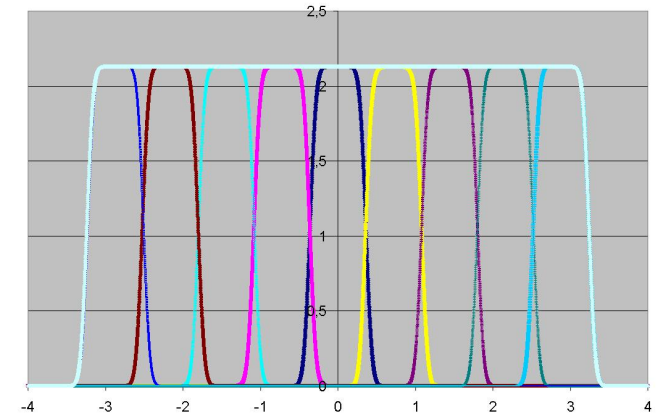


R. Bregović, P. Kovács, T. Balogh, A. Gotchev, 'Display-specific light-field analysis', in Proc. SPIE DSS 2014



Modeling of projector based LF displays

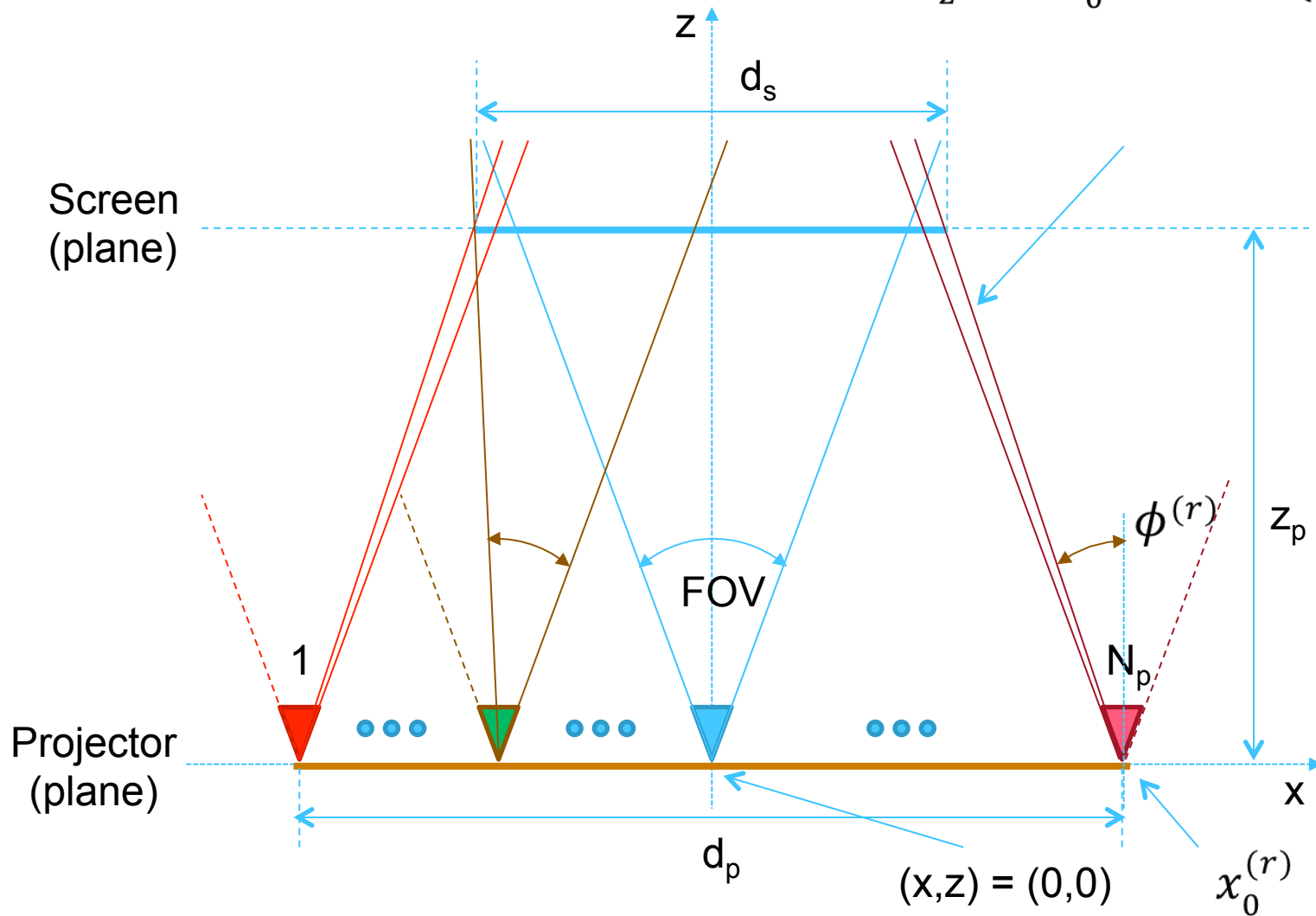
- Holographic screen
 - bell-shaped highly-selective diffuse profile eliminates crosstalk and provides proper intensity uniformity (continuous LF reconstruction)
 - LF reconstruction instead of views, specific distributed image organization



Display-camera setup

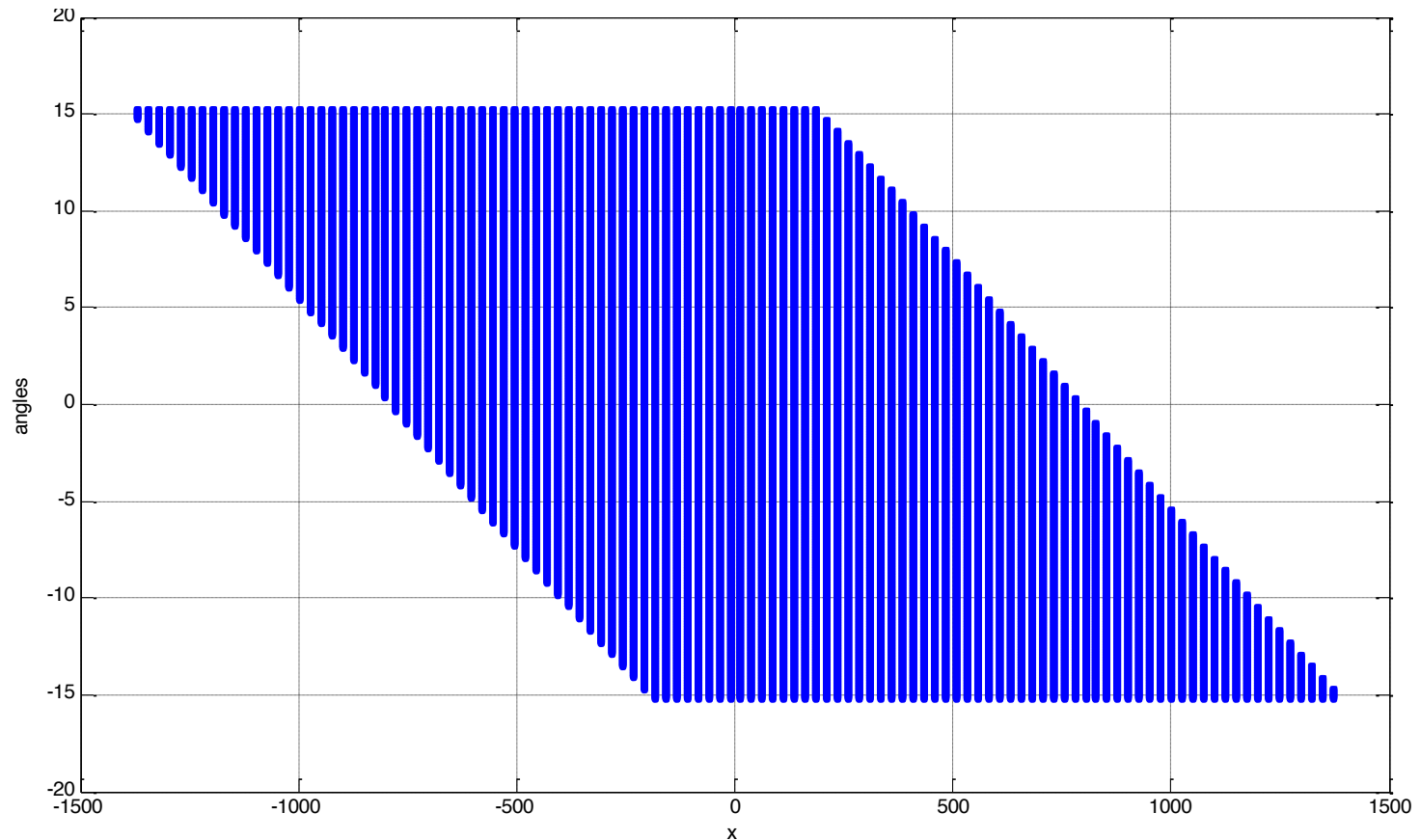
Ray r

$$x_z^{(r)} = x_0^{(r)} + z \tan(\phi^{(r)})$$



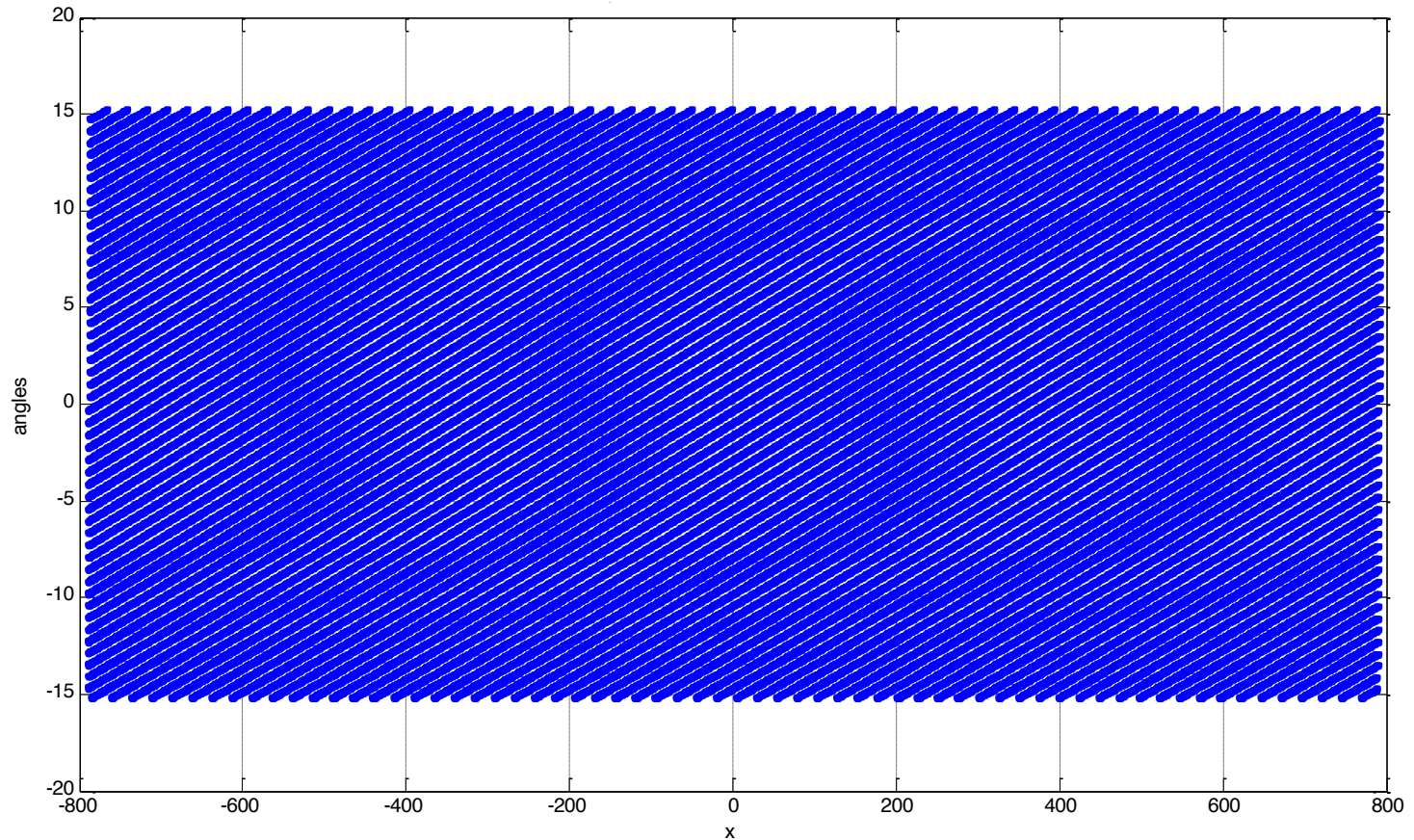
Projectors – ray space

Projector plane



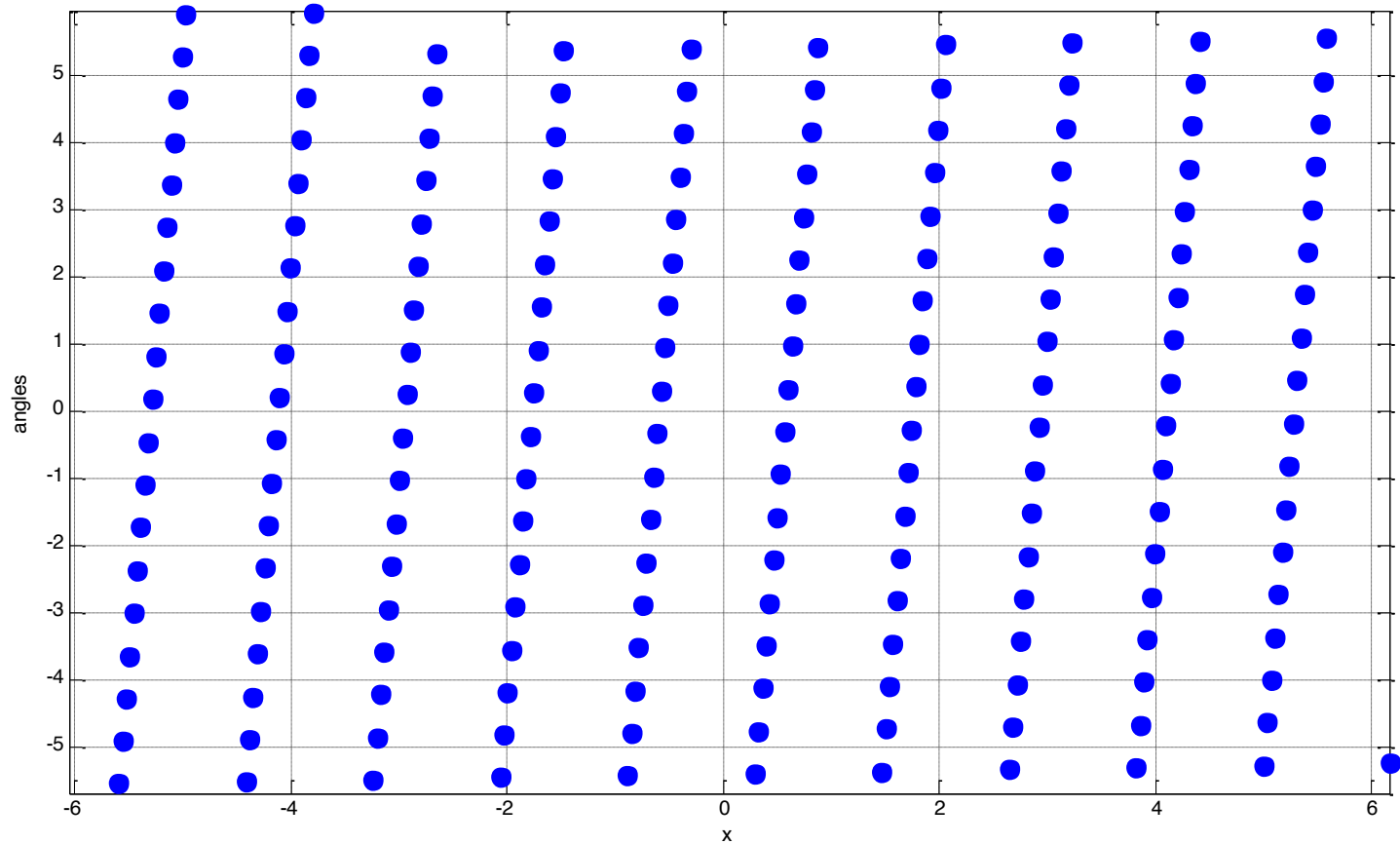
Projectors – ray space

Screen plane



Projectors – ray space

Screen plane

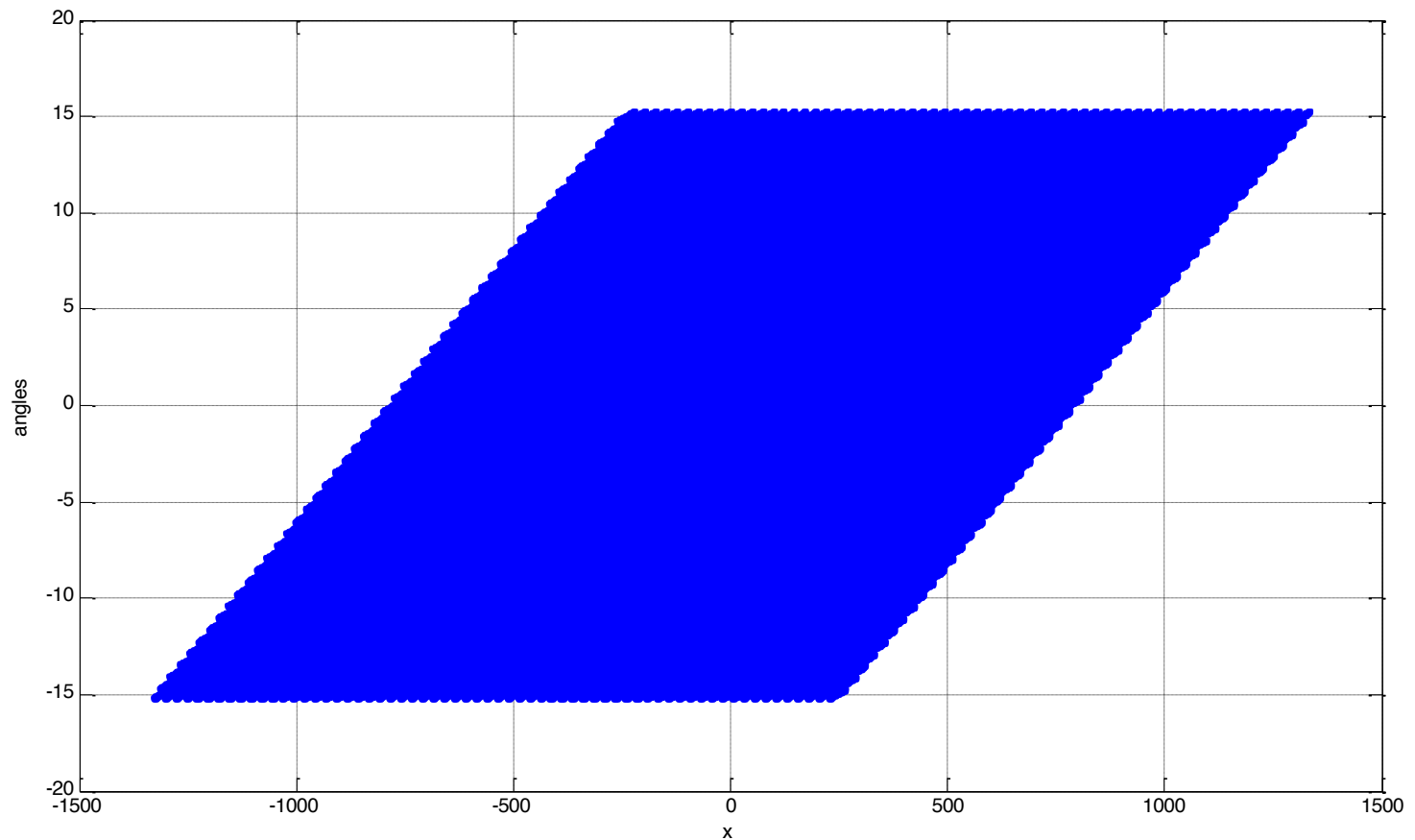


! Nonrectangular grid !



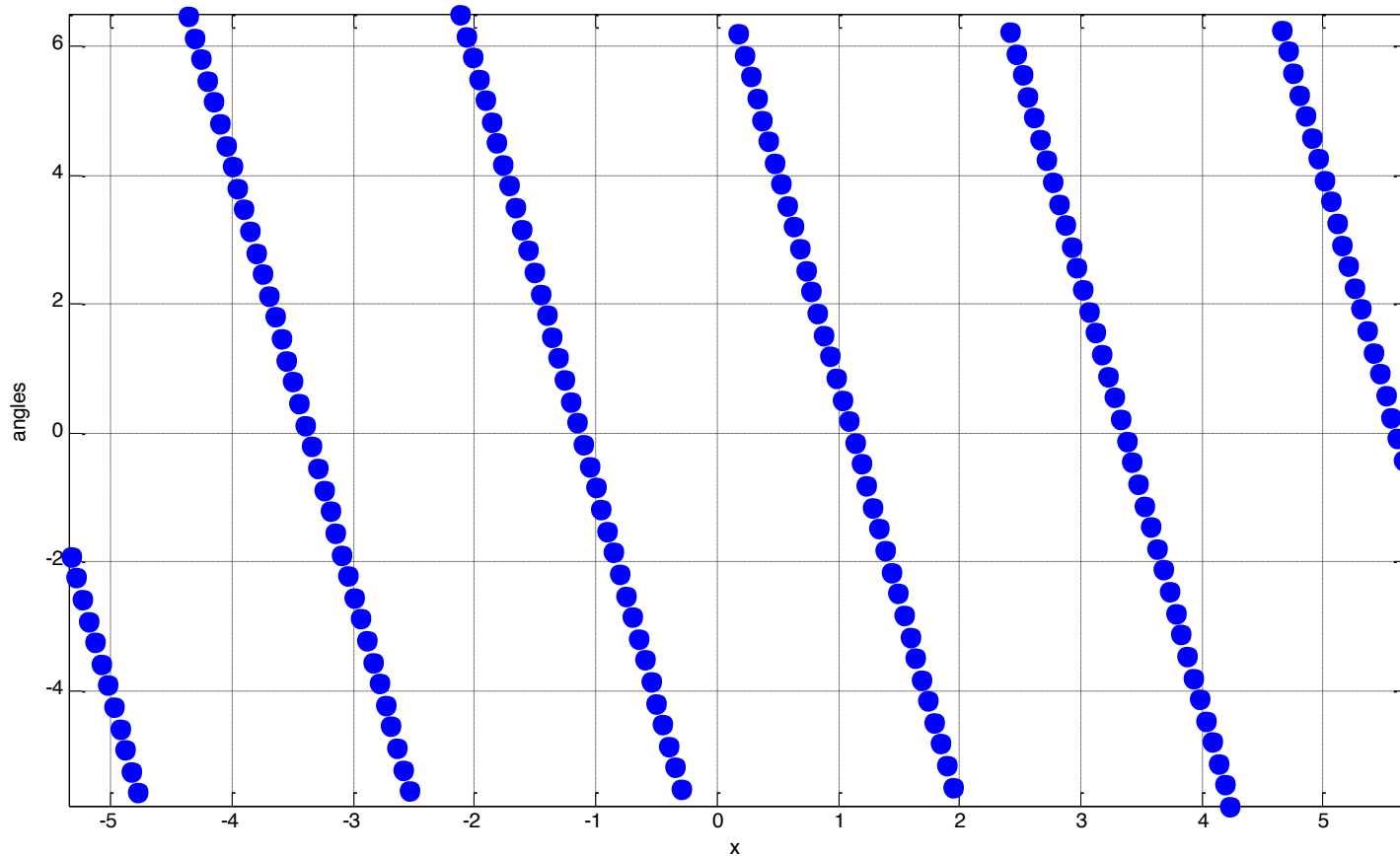
Projectors – ray space

Camera / viewer plane



Projectors – ray space

Camera / viewer plane



Content at some projectors (example)

image 32 / 112, shift = 0



image 86 / 112, shift = 0

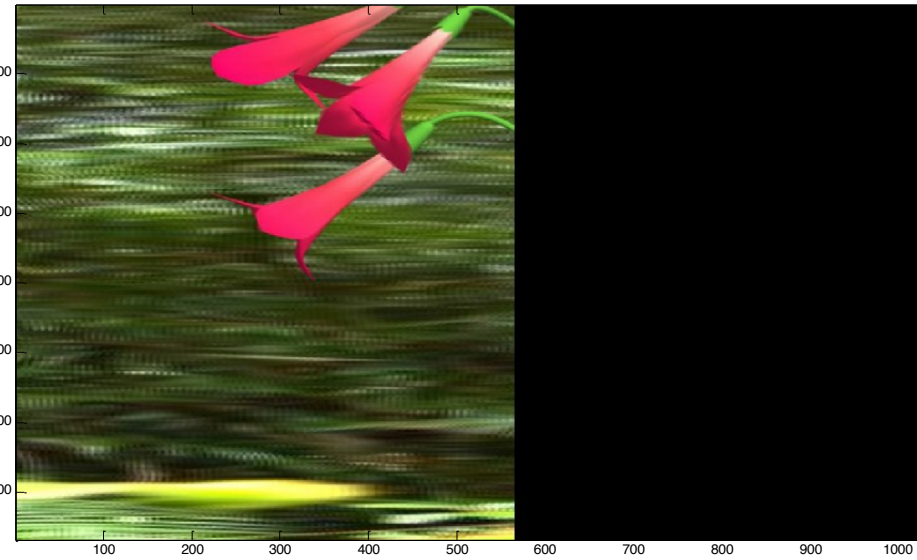
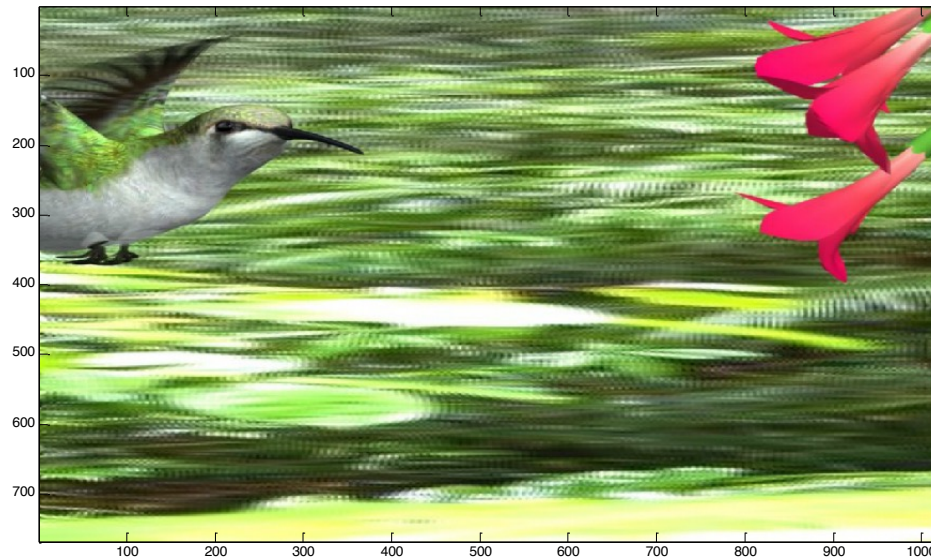


image 56 / 112, shift = 0

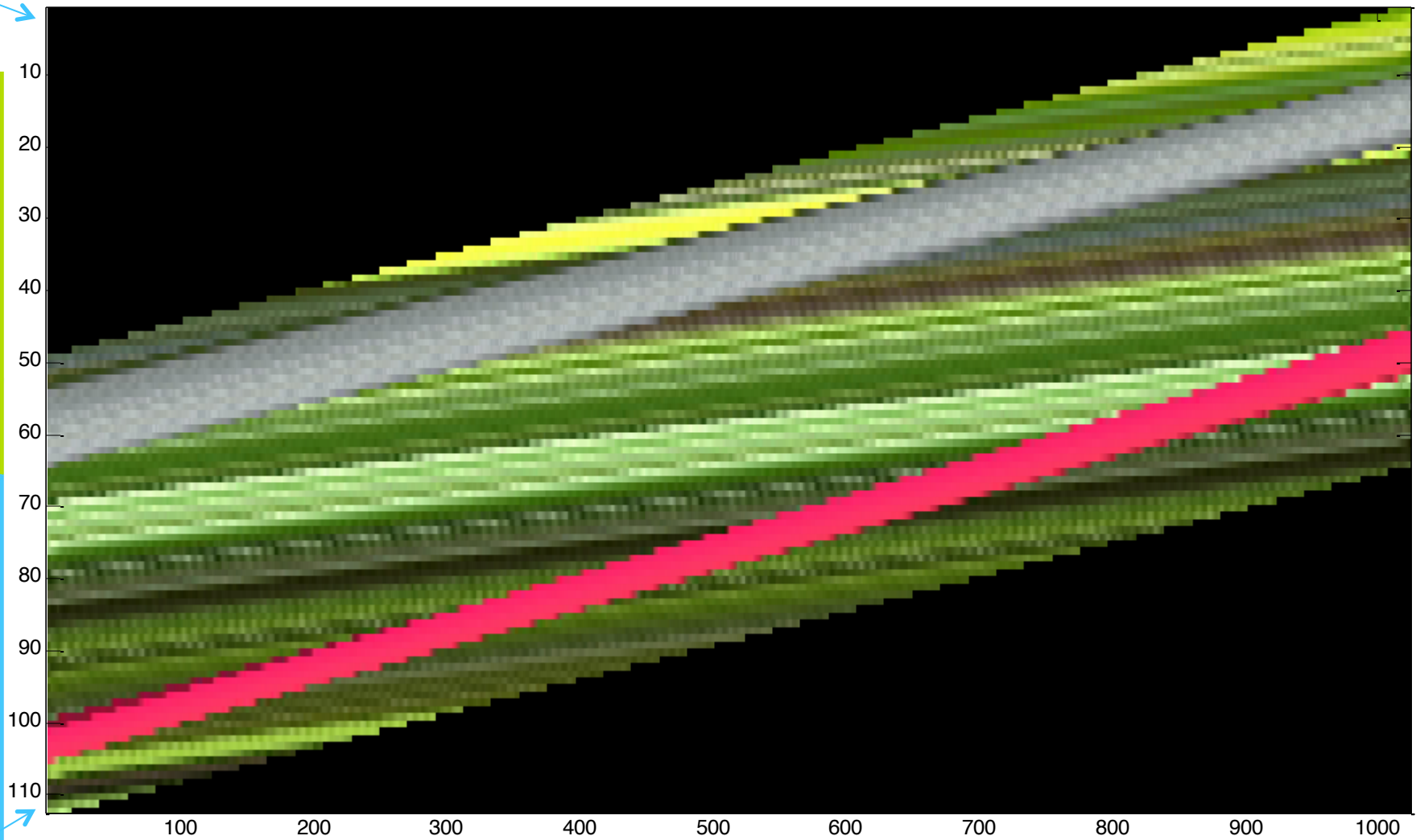


Projector
==
image



Left most projector

image 300 / 768, shift = 0



Right most projector



image 300 / 768, shift = 19

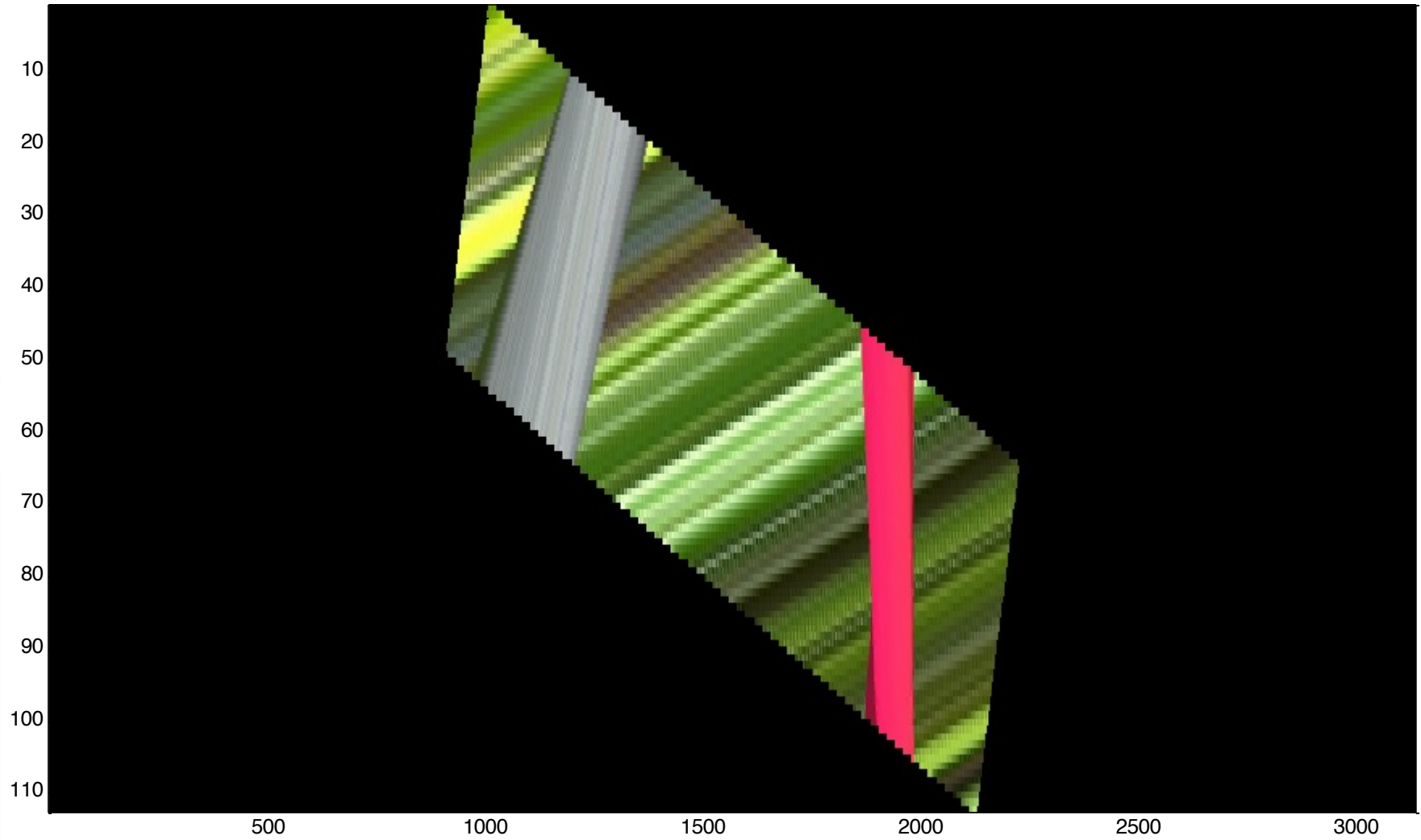
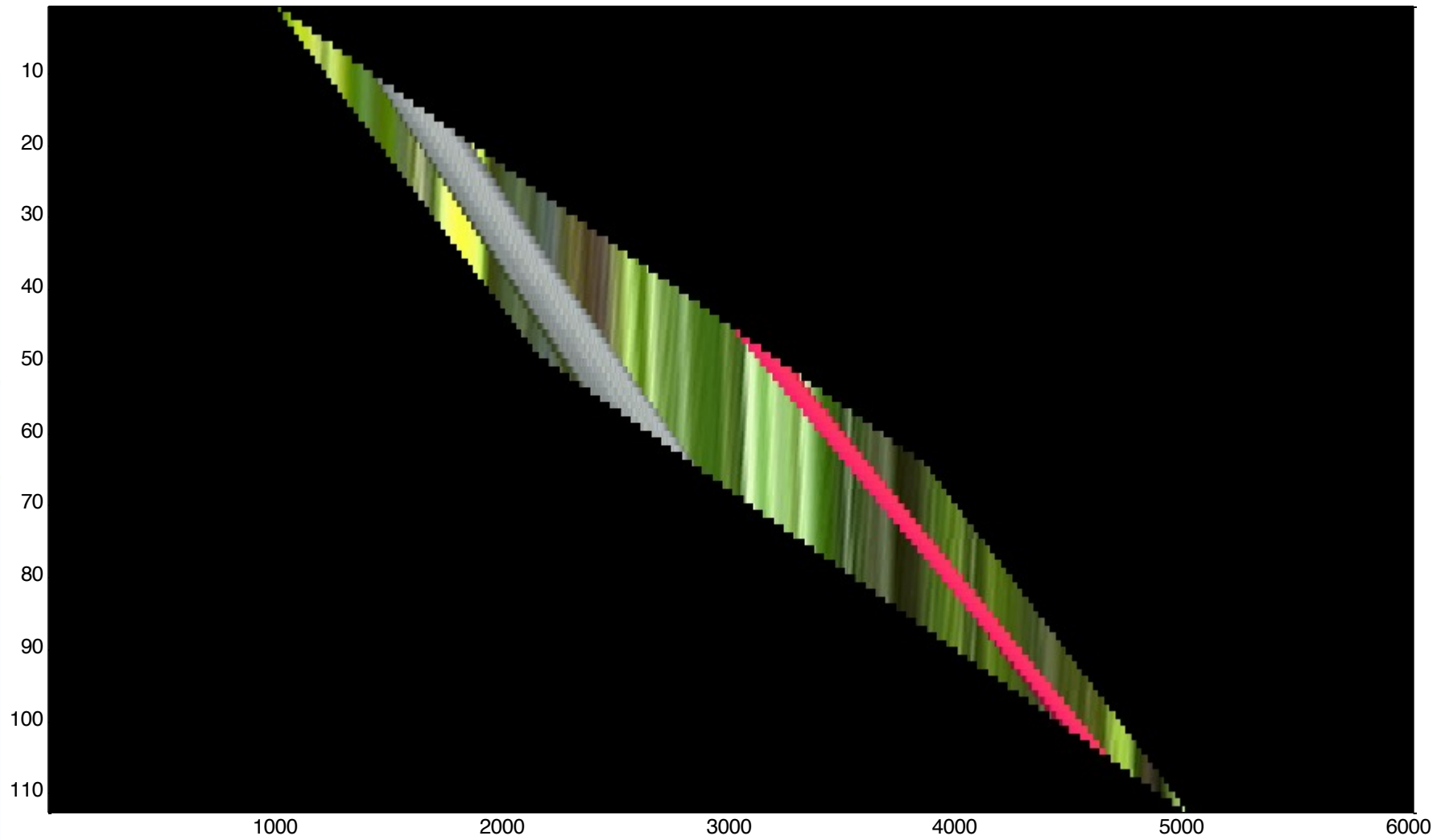
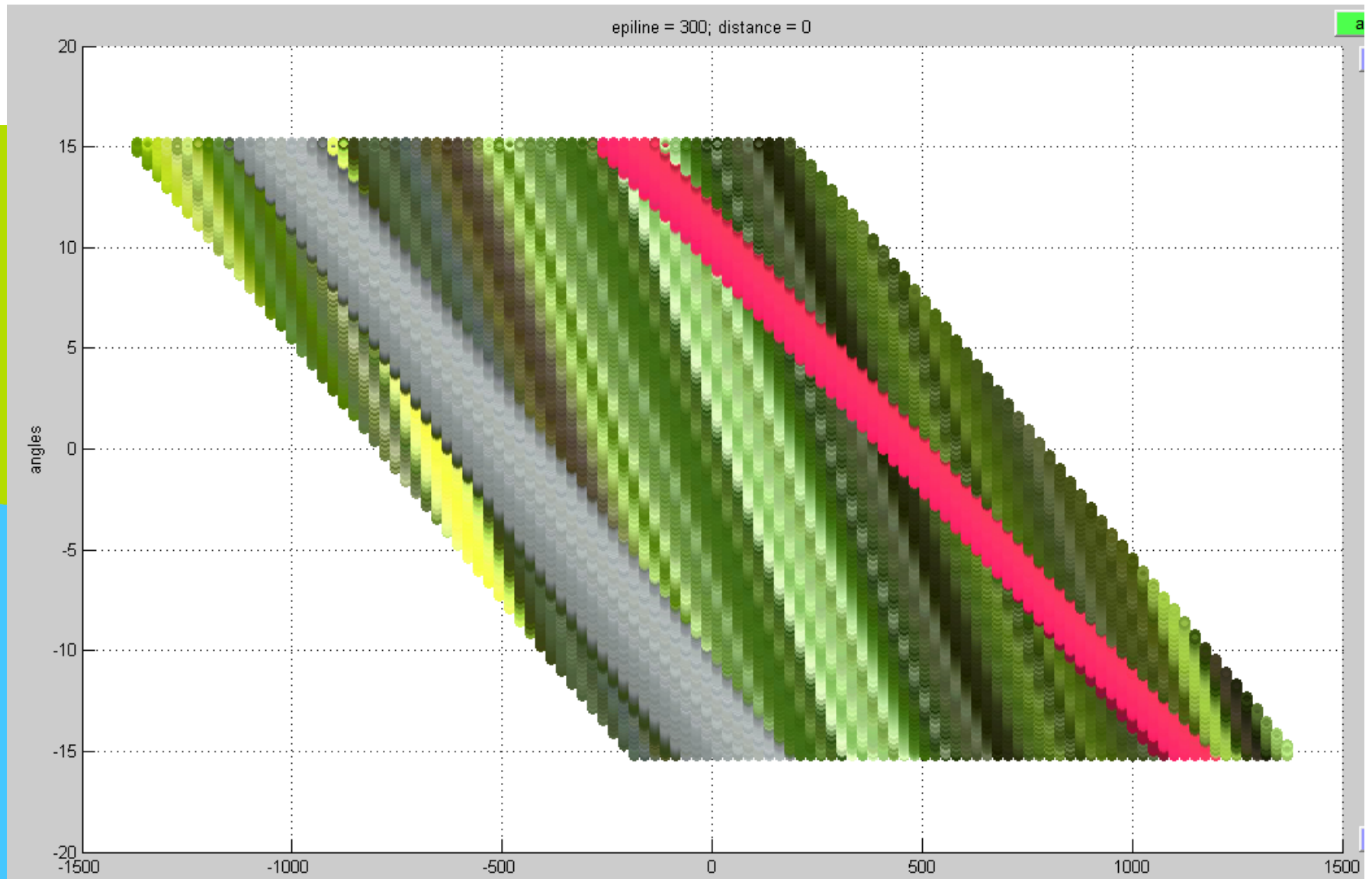
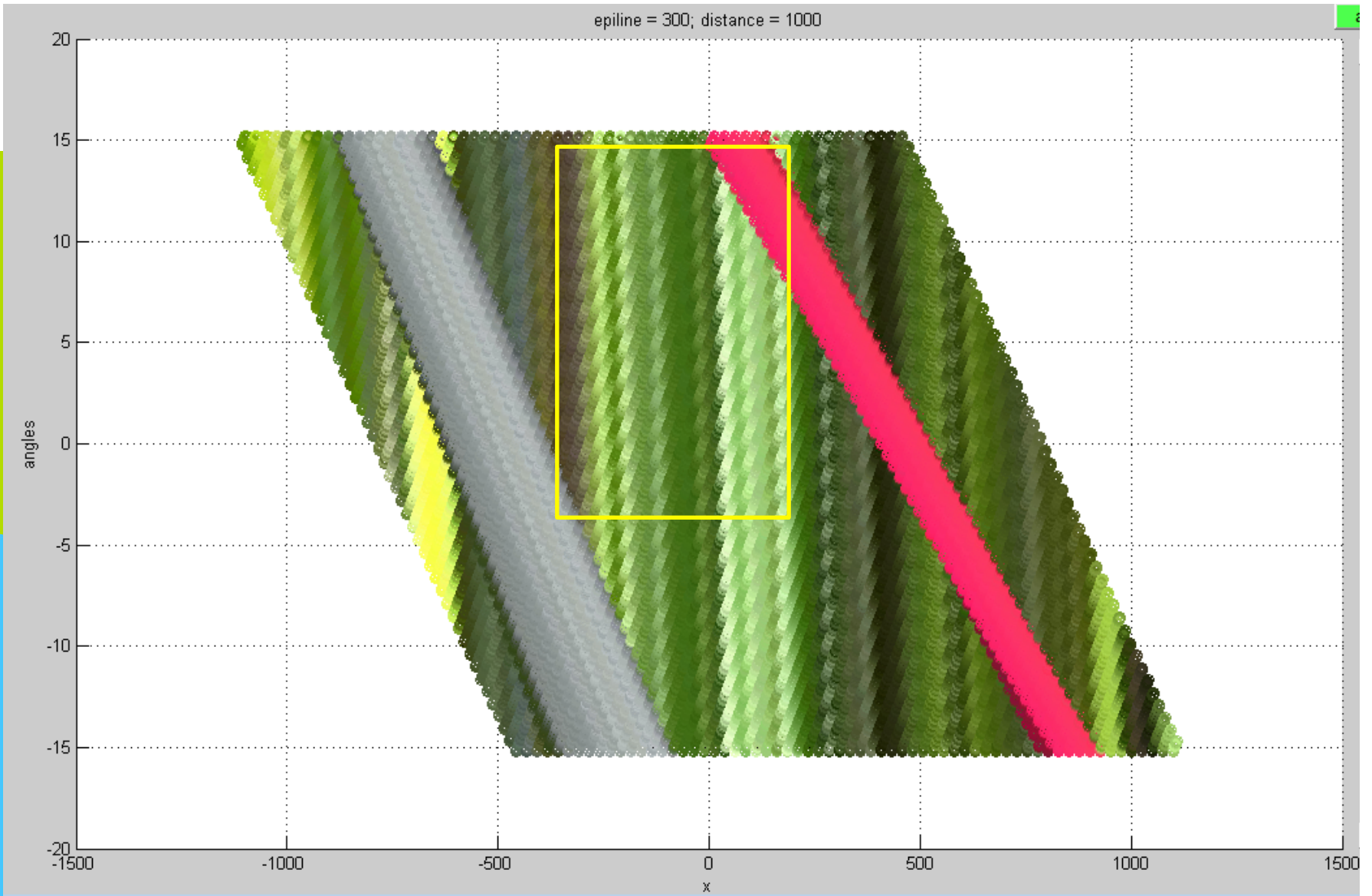


image 300 / 768, shift = 45







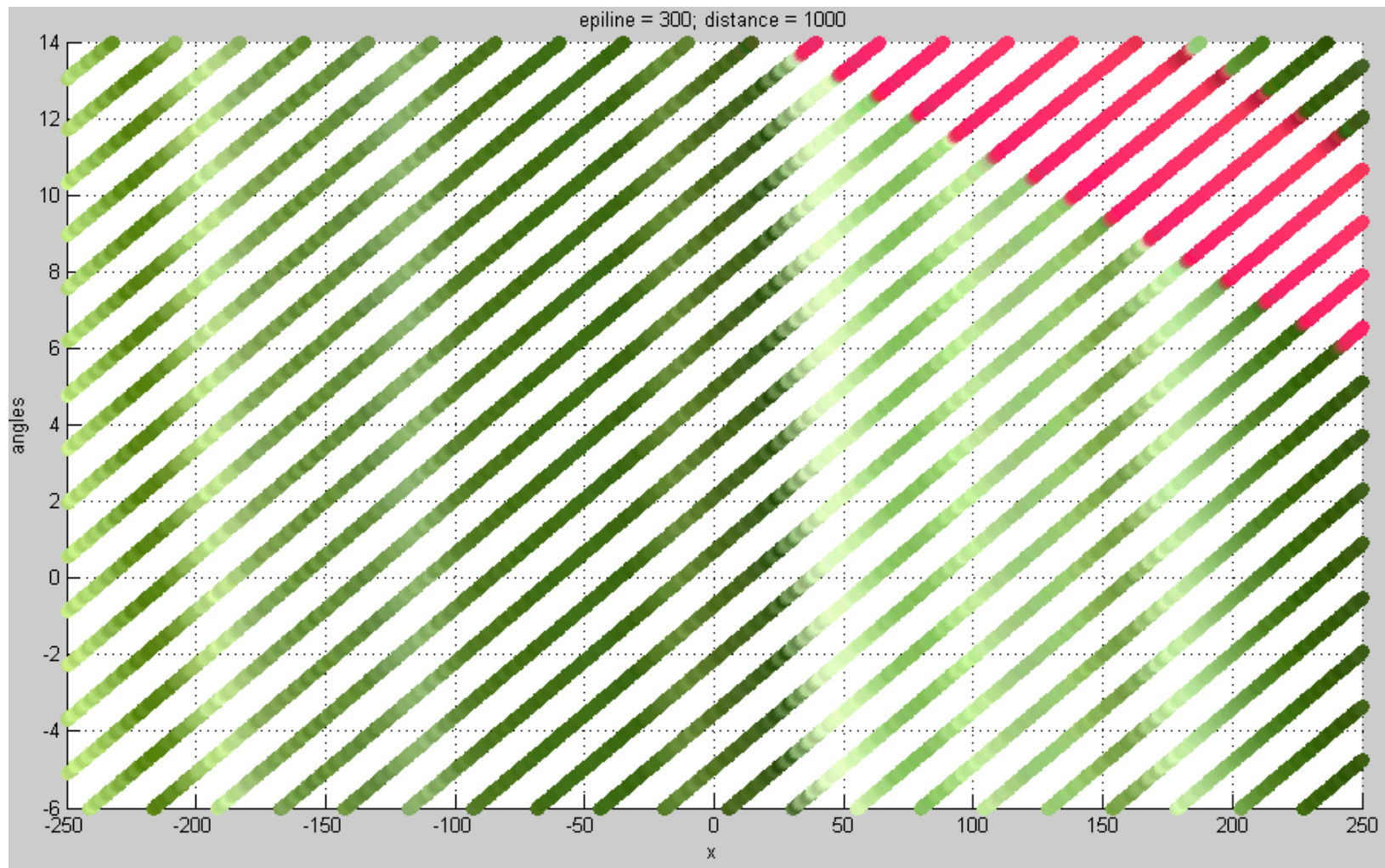
At depth 1000

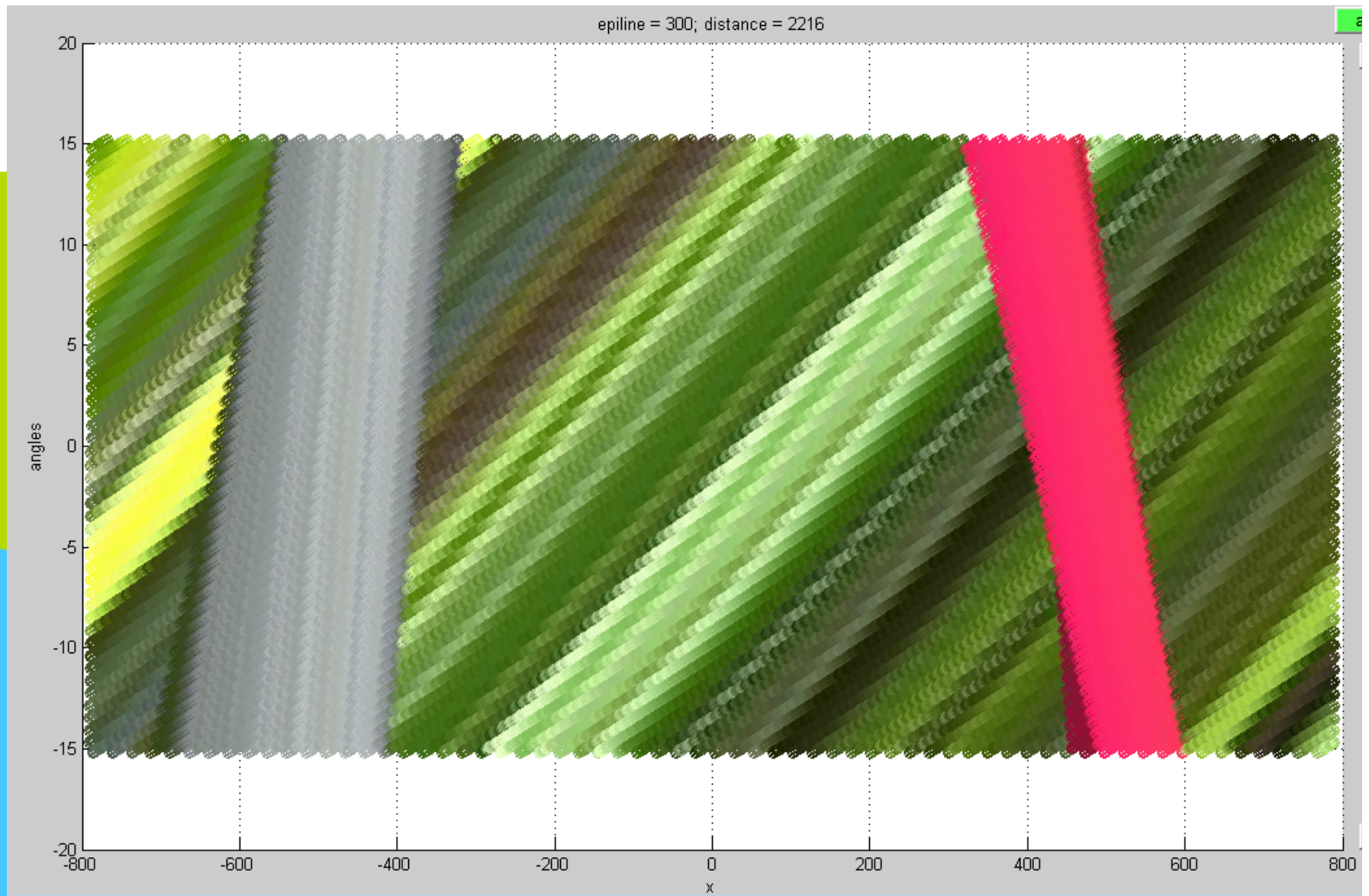


TAMPERE UNIVERSITY OF TECHNOLOGY

Atanas Gotchev, 3D Media Group

24/06/15

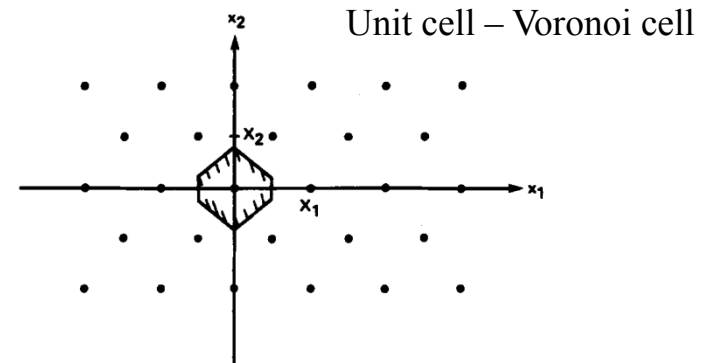
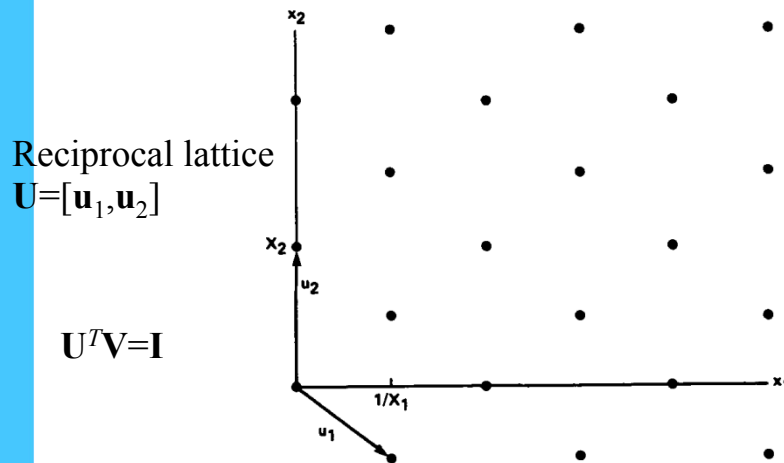
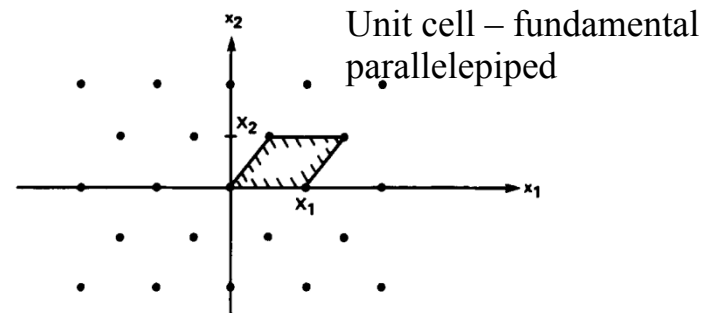
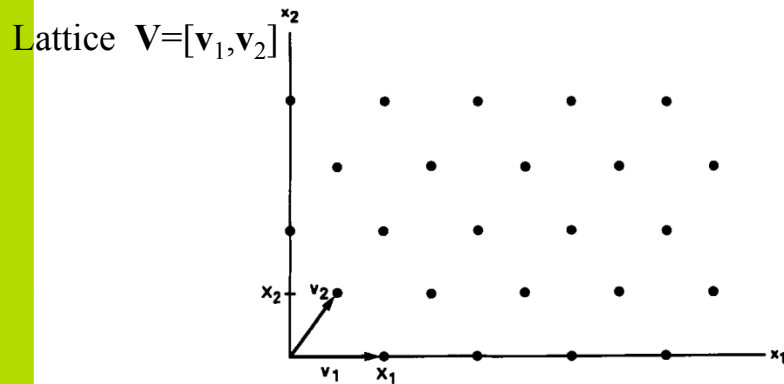




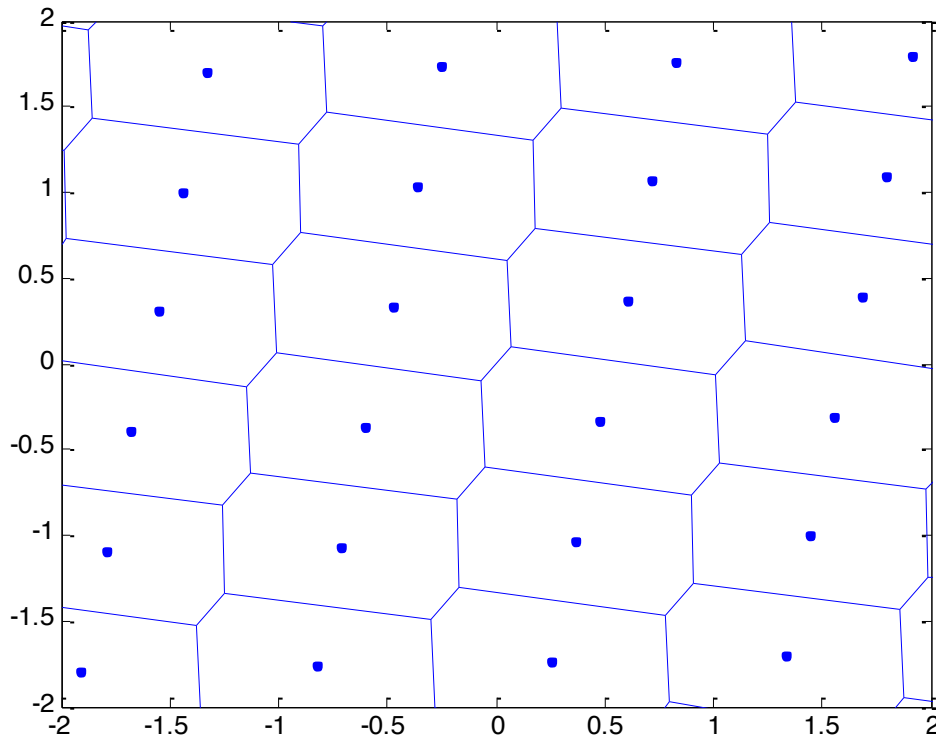
Content at Screen plane



Lattices and Voronoi cells



Cameras – ray space (screen plane)



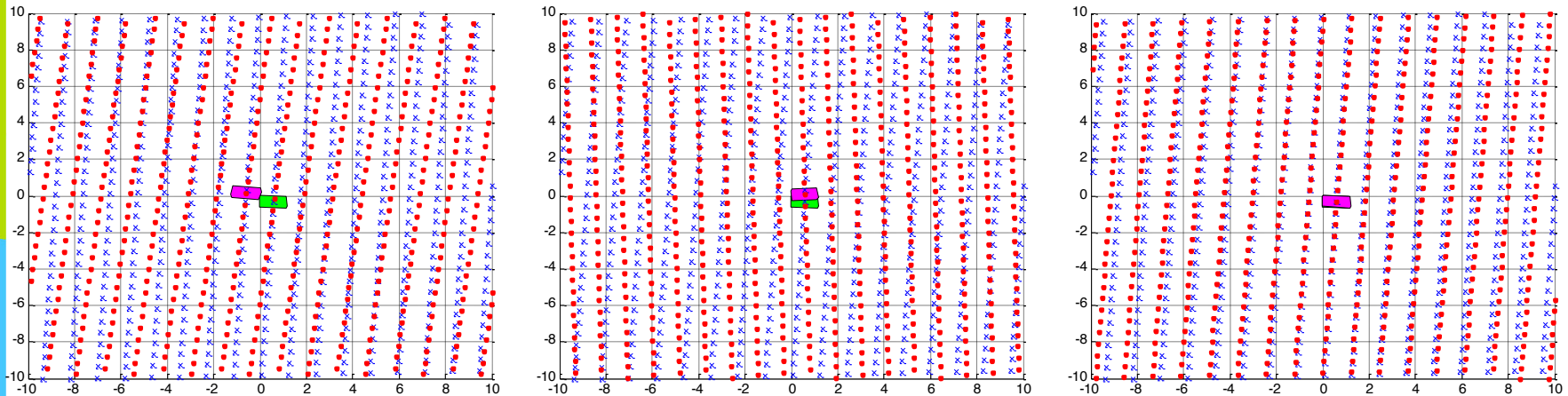
Fitting Voronoi cells to the sampling grid at the screen plane

Size of the cell defines angular-spatial bandwidth of the display at screen level



Cameras – ray space at screen plane

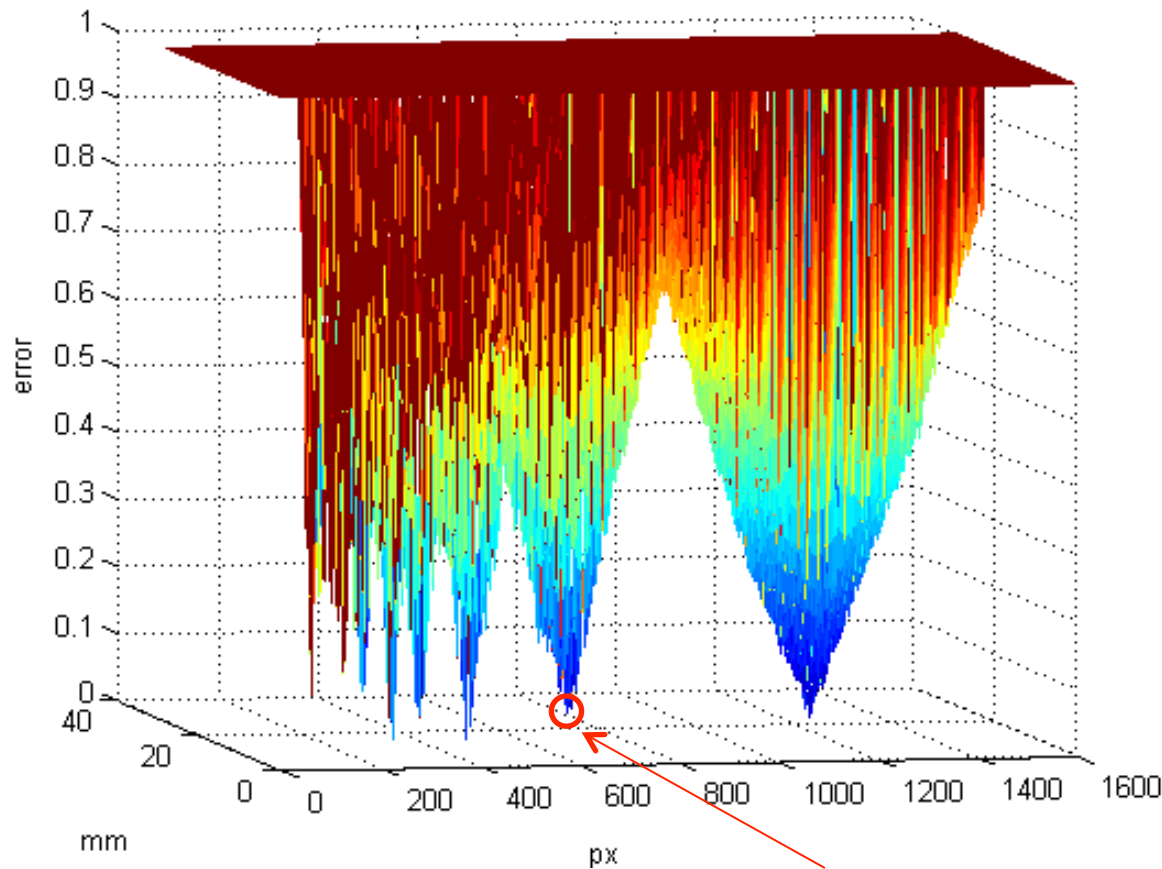
- Camera projected back to the screen plane
- Different camera setups leads to different sampling grids



- Adjust the camera-to-camera spacing and spatial resolution with the aim to minimize the difference between camera bandwidth and display bandwidth (match respective Voronoi cells)



Camera optimization - example



$x \sim$ resolution (px)

$y \sim$ camera distance (mm)

$z \sim$ error (thresholded to 1)

Multiple minimas

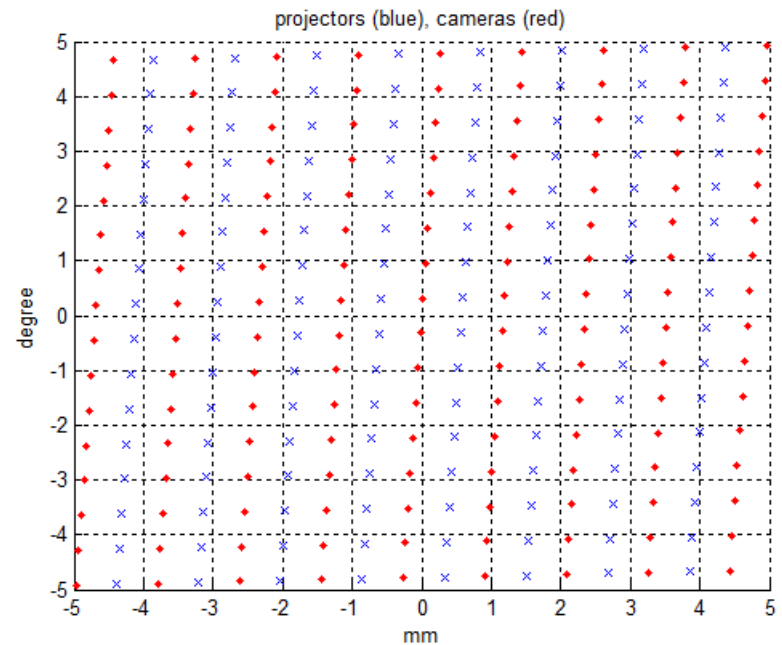
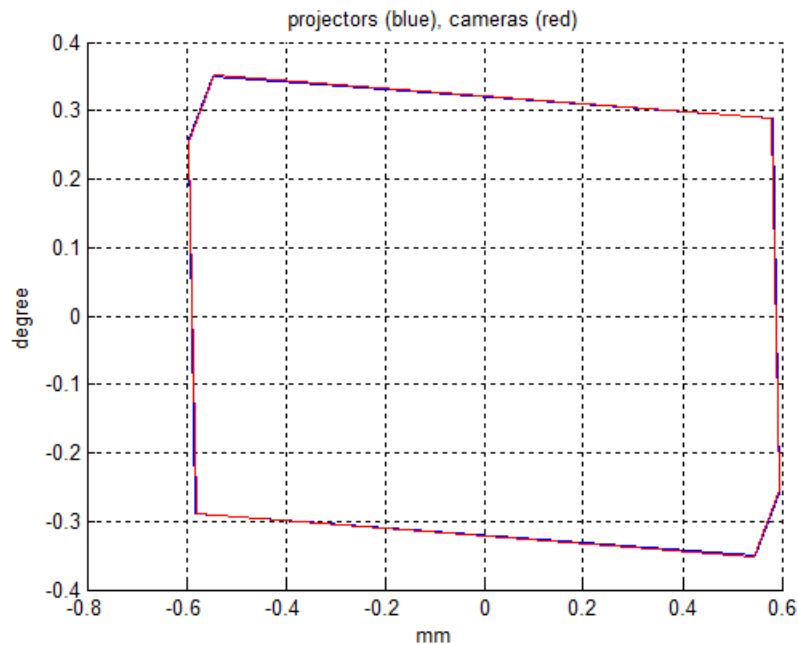
camera-to-camera
spacing depends on
resolution



Camera optimization - example

Optimized camera setup: 1240 px & 22.29 mm

Screen plane



Display specific LF analysis: recap

- An LF display is a band-limited device
 - Aliasing effects perfectly apply
- Sampling topology in ray-space allows to find (by optimization) the camera topology which delivers the same bandwidth
 - However, one still needs a denser capture (or a non-band-limited reconstruction)



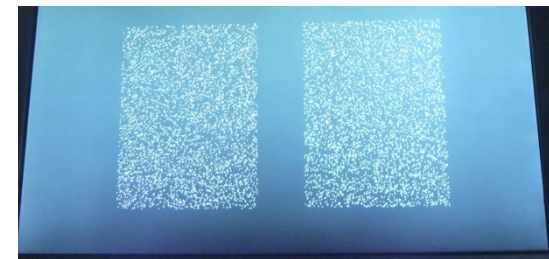
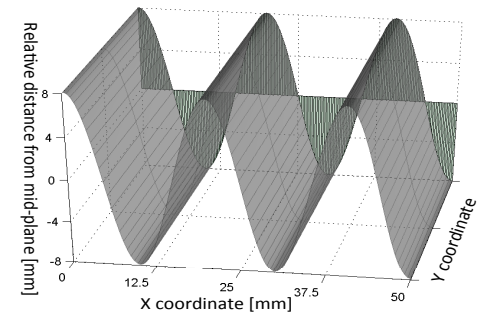
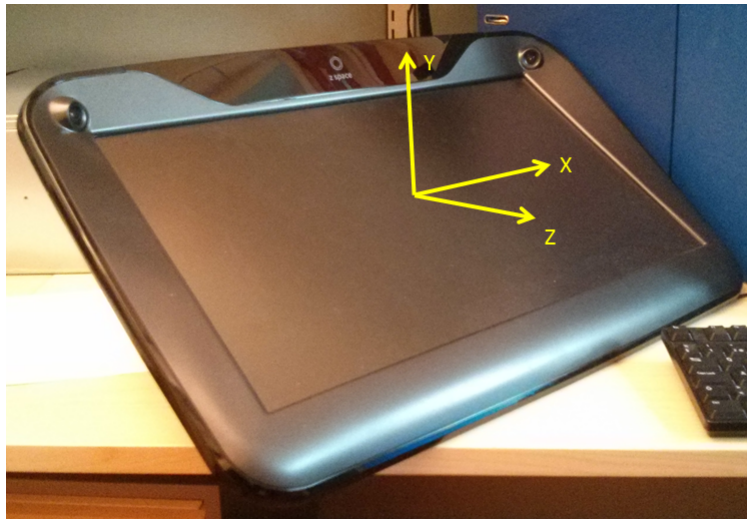
Quantification of perceptual thresholds of continuous parallax vs stereopsis

PERCEPTION OF PARALLAX



Head parallax for depth perception

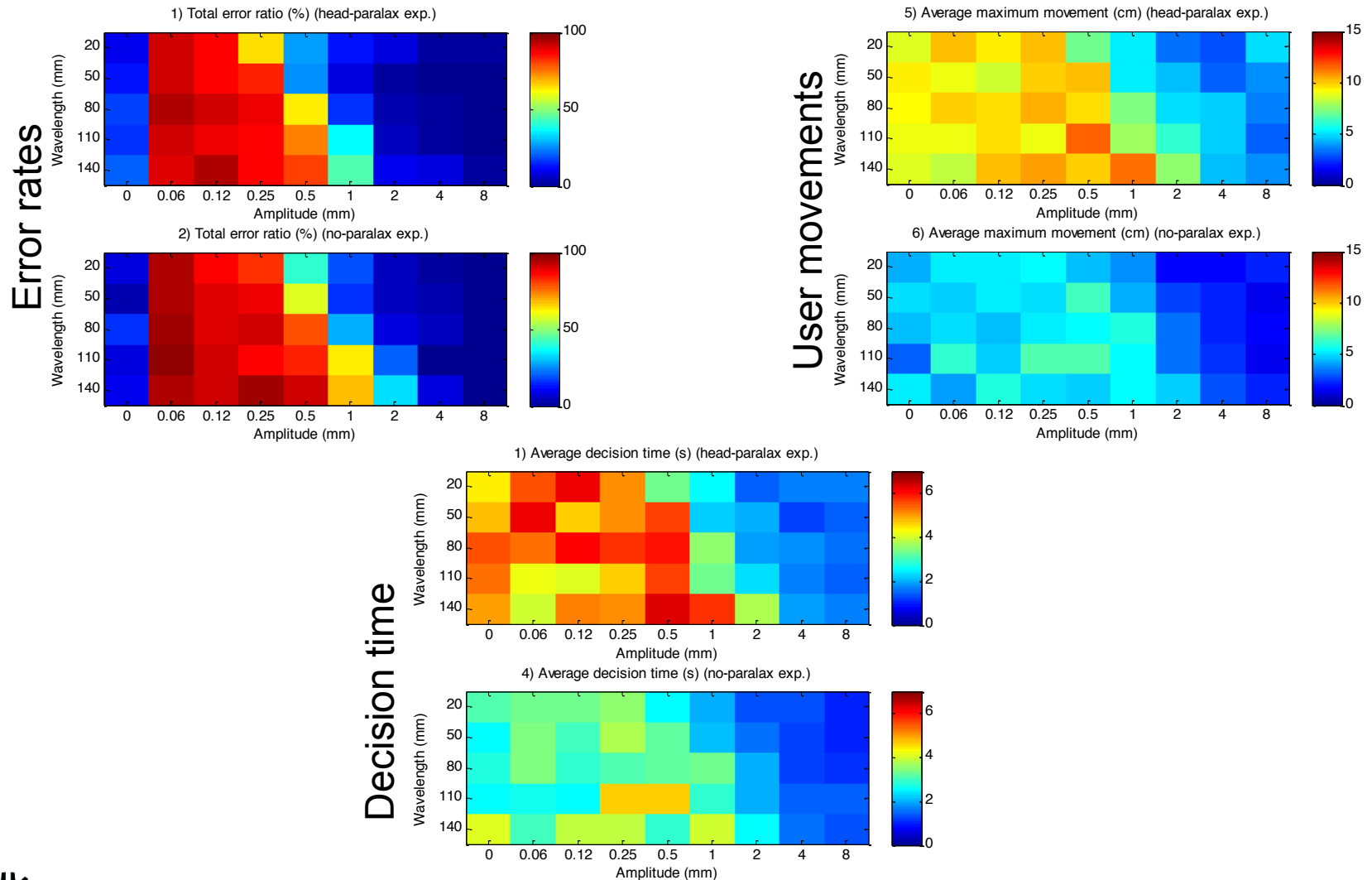
- Testing parallax vs stereo
 - Random dot stereograms for varying spatial frequencies



K. Lackner, A. Boev, A. Gotchev, 'Binocular depth perception: does head parallax help people see better in depth?', in Proc. 3DTV-CON 2014



Head parallax for depth perception



Content creation for LF displays

LF representation

- Multi-view multi-depth vs EPI representations
 - How to efficiently represent EPI?
 - How to reconstruct LF out of sparse camera views

LF displays

- Display bandwidth
 - Continuous LF reconstruction
- Optimal content representation in terms of number of cameras and their resolution is achievable
 - Avoids aliasing artifacts

Quantifying the full parallax perception might help in saving bits...



Looking forward to further discussions...

Contact

Assoc. Prof. **Atanas Gotchev**
Tampere University of Technology
P.O.Box 553, FI-33101 Tampere, Finland
Phone: +358 40 8490733,
Email: atanas.gotchev@tut.fi



Thank you for your attention!

