Moving Image Perception summary

Prof. Ioannis Pitas
Aristotle University of Thessaloniki
pitas@csd.auth.gr
www.aiia.csd.auth.gr

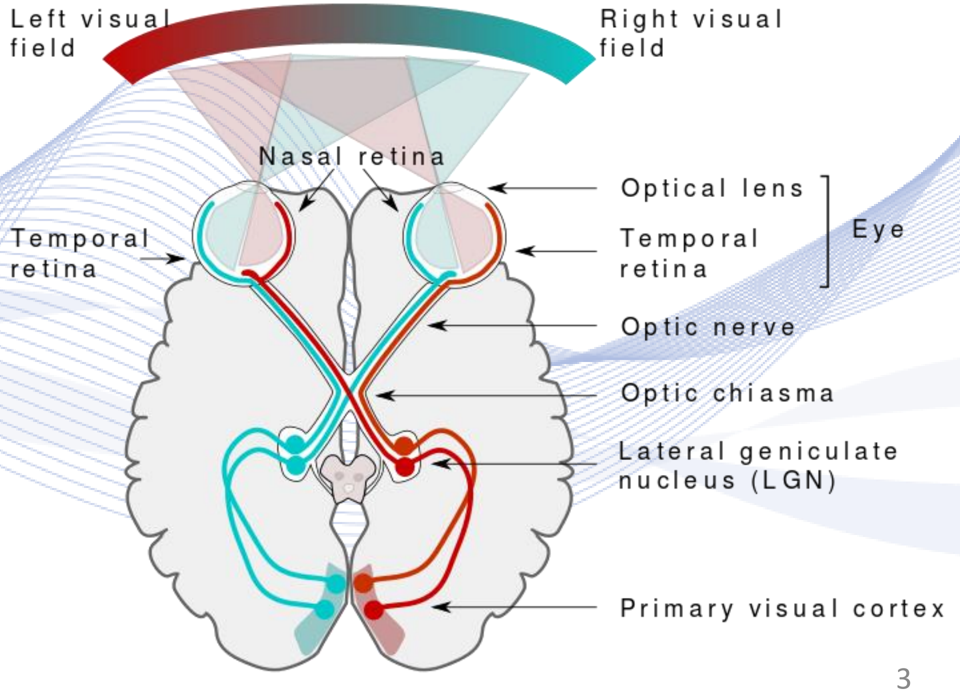
Version 2.6.1

Moving Image Perception

- **Human Vision Modeling**
- Video Frequency Content
- Spatiotemporal HVS Models
- Video Quality Assessment

Human Vision Modeling

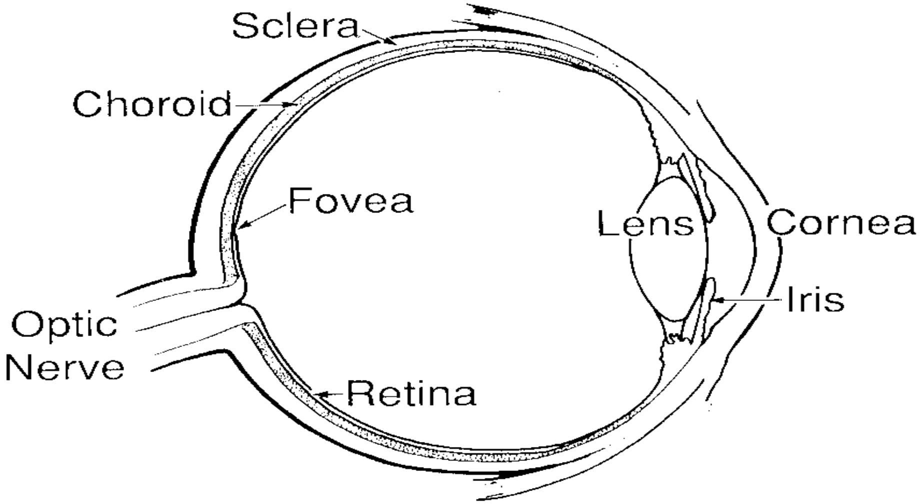
- One of digital image and video processing aims is image quality improvement.
- ***Human Visual System*** (***HVS***) modeling is difficult, because of its complex structure.



Human visual system [HVP].

Human Vision Modeling

- ***Human eye***: spherical shape with a diameter of 20 mm.
- Light enters the ***pupil*** of the ***iris*** (diameter 2 - 8 mm).
- It passes through ***lens***, ***vitreous humour*** and on the ***retina***.



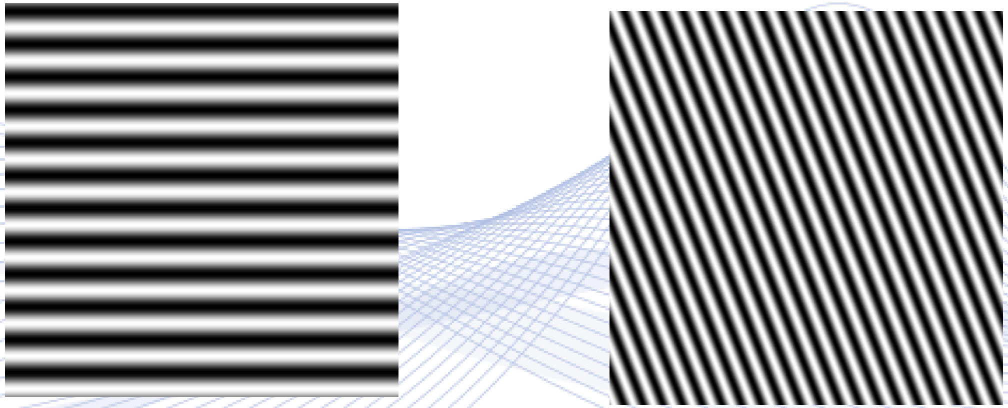
Human eye.

Moving Image Perception

- Human Vision Modeling
- **Video Frequency Content**
- Spatiotemporal HVS Models
- Video Quality Assessment

Video Frequency Content

- A frequency F is linked with angular frequency $\Omega = 2\pi F$.
- Spatial frequencies (video content changes along x, y axes):
  - $\Omega_x = 2\pi F_x$ and $\Omega_y = 2\pi F_y$.



2D sinusoidal signals: a) $(F_x, F_y) = (0,6)$; b) $(F_x, F_y) = (10,4)$.

Video Frequency Content

Temporal Frequency F_t:

- Video signal: moving image (2D video frames changing over time).
- The temporal frequency F_t depends on image content motion.
- The video content motion is due to:
  - camera motion and/or
  - object(s) motion.

Video Frequency Content

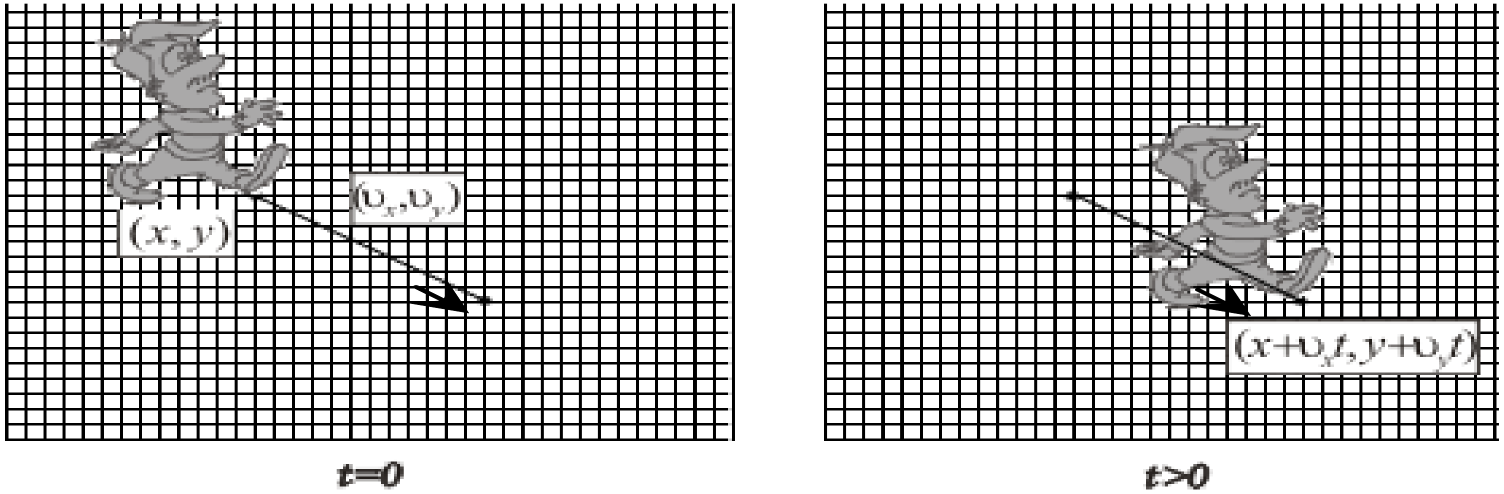
Constant velocity ***2D linear object motion***:

- $f_0(x, y) = f(x, y, 0)$: object image at time $t = 0$.
- $\mathbf{v} = [v_x, v_y]^T$: object motion vector.
- v_x, v_y: horizontal/vertical speed.

- Object image at time t (for homogeneous image background):

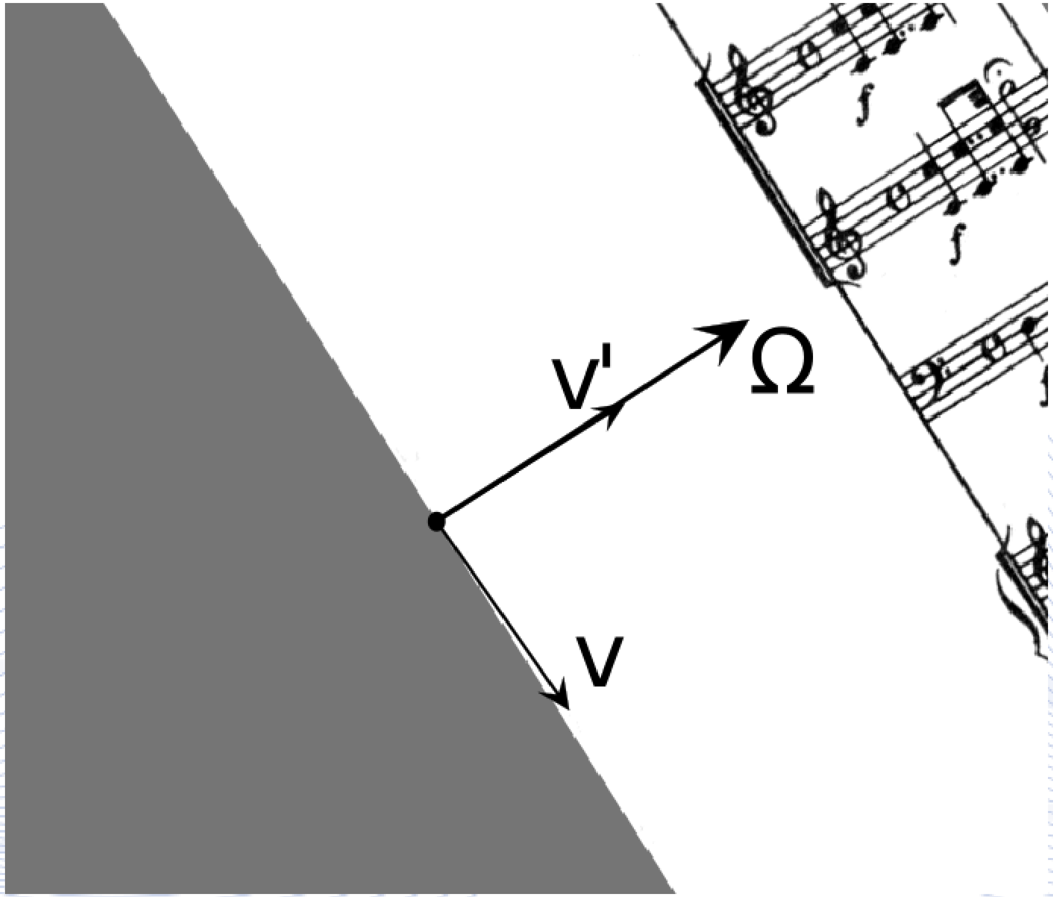
$$f(x, y, t) = f\left(x - v_x t, y - v_y t, 0\right) = f_0(x - v_x t, y - v_y t).$$

Video Frequency Content



Linear constant 2D object motion.

Video Frequency Content



Motion vector perpendicular to local image frequency vector.

Artificial Intelligence & Information Analysis Lab

Moving Image Perception

- Human Vision Modeling
- Video Frequency Content
- **Spatiotemporal HVS Models**
- Video Quality Assessment

Spatiotemporal HVS models

- For the ***spatiotemporal modelling*** of human vision, dynamic models of neurons must be used:
  - It is particularly difficult.

- The eye is a dynamic system:
  - pupil diameter changes with light intensity,
  - the human eye can rotate and perform smooth pursuit movements.
- Spatiotemporal image perception experiments.

Spatiotemporal HVS models

Temporal sensitivity of the human vision:

- HVS temporal frequency response refers to HVS sensitivity to temporal video content variations.

- ***Display flicker***.

Spatiotemporal HVS models

Kelly experiments: determination of the necessary video frame rate (fps).

- Observers were presented a flat screen whose luminance changed sinusoidally:

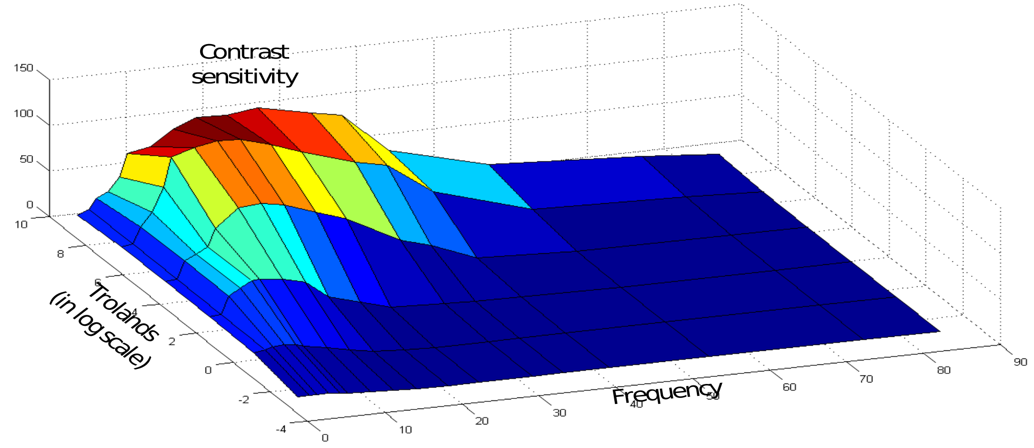
$$f(x,y,t) = C(1 + s\cos 2\pi F_t t).$$

- C: constant luminance,
- F_t: temporal frequency,
- s intensity modulation level.

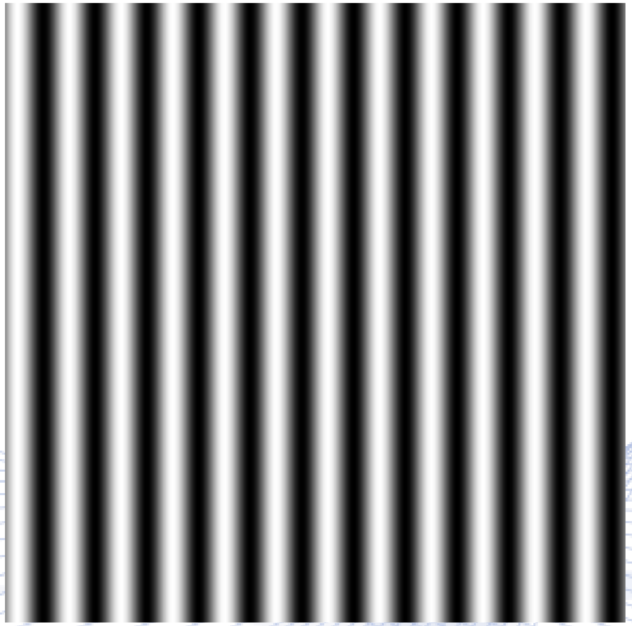
Spatiotemporal HVS models

Contrast Sensitivity Function **(*CSF*)** $s_e(C, F_t)$.

- ***Troland***: a unit to describe the light intensity entering the eye retina.
- CSF $s_e(C, F_t)$ is a band-pass function of both C, F_t.

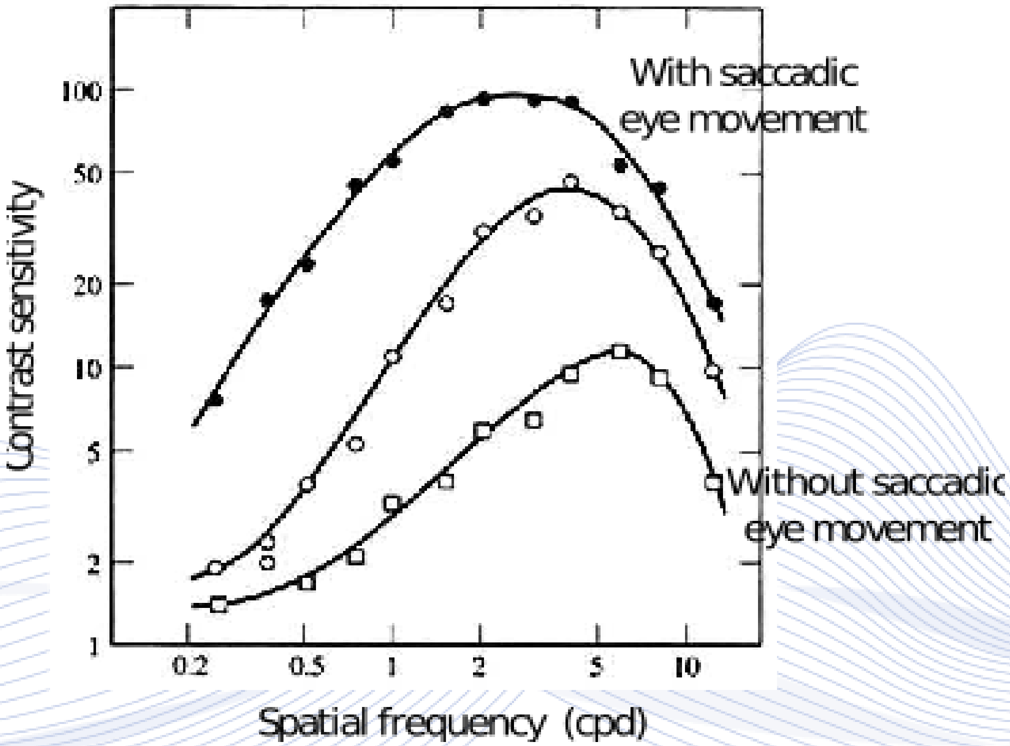


Spatiotemporal HVS models



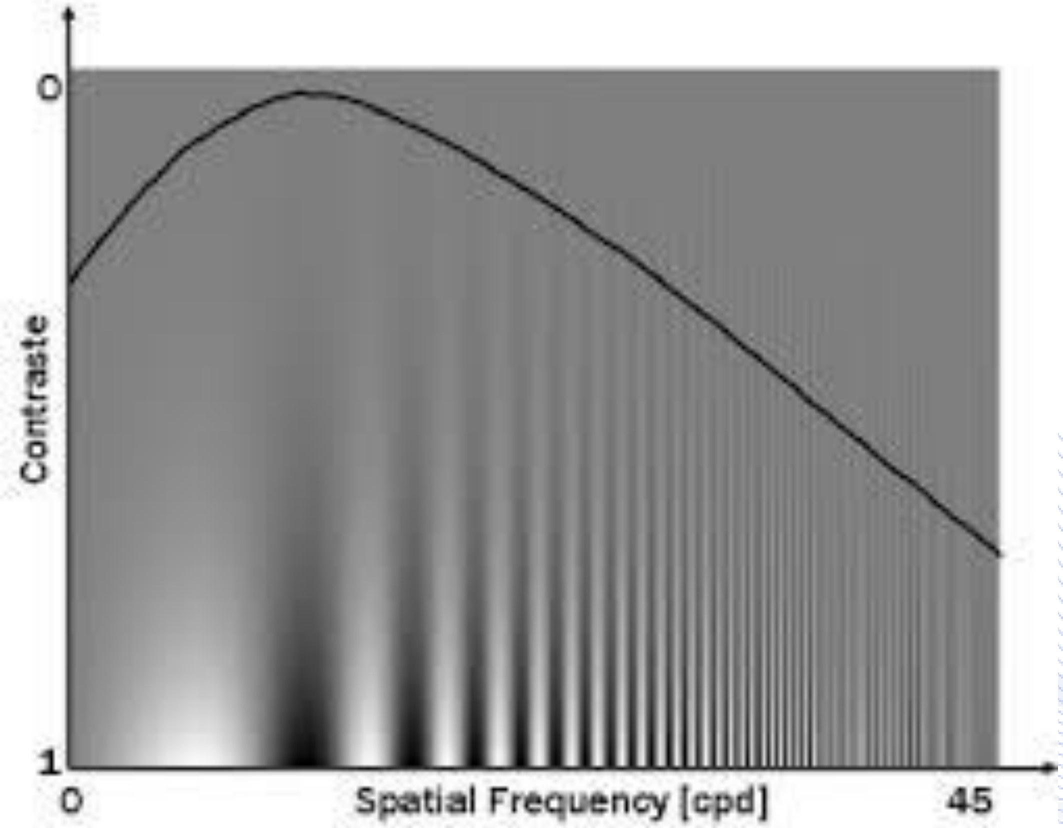
Horizontal 2D sinusoidal signals having $(F_x, F_y) = (6,0)$.

Spatiotemporal HVS models



Spatial HVS frequency response.

Spatiotemporal HVS models



Contrast sensitivity.

Spatiotemporal HVS models

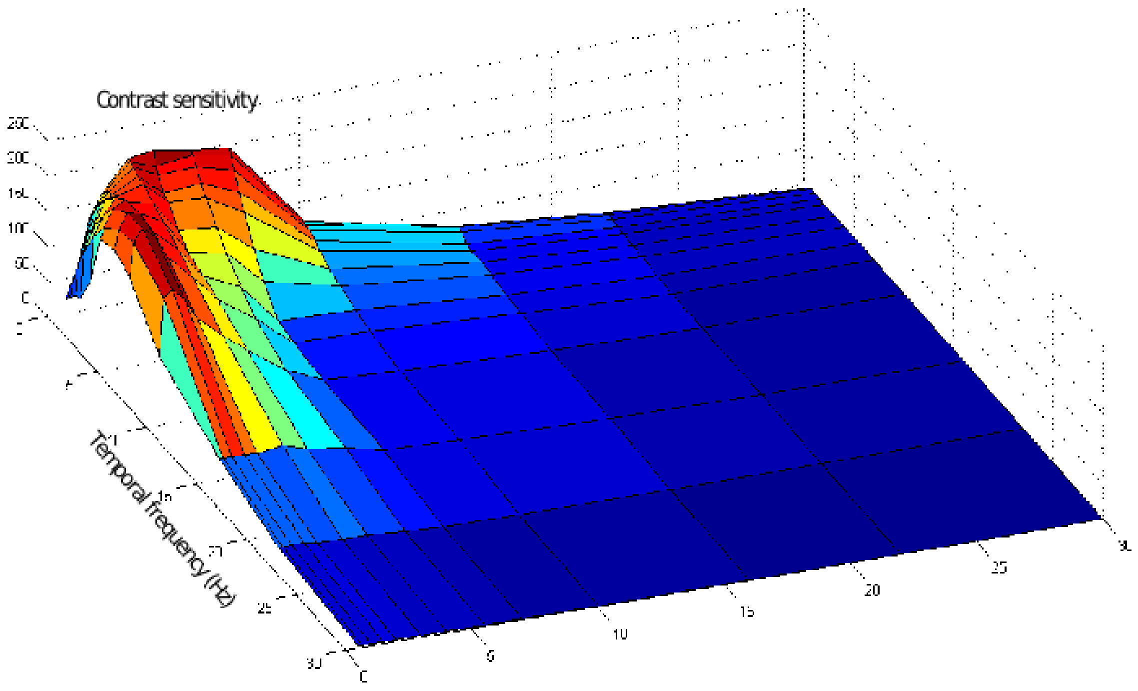
HVS spatiotemporal frequency response experiments:

- Test pattern:

$$f(x, y, t) = C(1 + s\cos(2\pi F_x x)\cos(2\pi F_t t)).$$

- For a fixed pair of F_x and F_y the modulation level s was changed.
- The observer was requested to determine the minimal observable modulation level s_{min}.

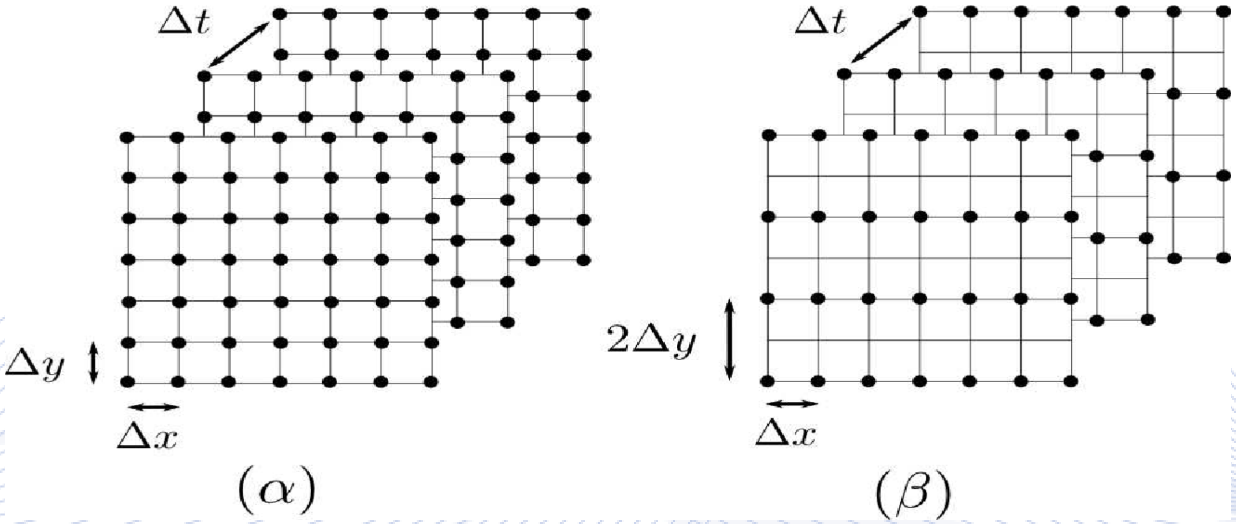
Spatiotemporal HVS models



Spatiotemporal HVS contrast sensitivity as a function of F_x , F_t for unconstrained eye motion.

Spatiotemporal HVS models

- 2:1 Interlaced video takes advantage of HVS properties.



Sampling grids for: a) progressive; b) 2:1 interlaced video.

Moving Image Perception

- Human Vision Modeling
- Video Frequency Content
- Spatiotemporal HVS Models
- **Video Quality Assessment**

Artificial Intelligence &
Information Analysis Lab

Video Quality Assessment

- In many cases, humans are the final video consumers.
- Perceived video quality must be quantified.
- ***Video Quality*** (***VQ***) is influenced by:
  - Acquisition noise;
  - Compression effects;
  - Transmission errors.

- VQ assessment can help meet video storage and transmission requirements.

Video Quality Assessment

Subjective video quality assessment:

- Ask humans to watch the video and assess its quality.
- ***Mean Opinion Score*** (***MoS***): scale $[1, \dots, 5]$.
- 1: worst, 5: best quality.
- Labor intensive and expensive.
- A large number of viewers is needed to lower score variability and provide statistical certainty.
- Impossible to assess all videos before broadcasting.
- Useful in providing a golden standard for automated VQA methods.

Video Quality Assessment

Objective video quality assessment:

- No human observers involved.
- ***Full reference VQA algorithms*** operate on distorted video, while employing the original video reference for comparison.
- VQA measures:
  - Mean Square Error,
  - Peak Signal to Noise Ratio,
  - SSIM.

Bibliography

[PIT2017] I. Pitas, “Digital video processing and analysis” , China Machine Press, 2017 (in Chinese).
[PIT2013] I. Pitas, “Digital Video and Television” , Createspace/Amazon, 2013.
[PIT2021] I. Pitas, “Computer vision”, Createspace/Amazon, in press.
[NIK2000] N. Nikolaidis and I. Pitas, “3D Image Processing Algorithms”, J. Wiley, 2000.
[PIT2000] I. Pitas, “Digital Image Processing Algorithms and Applications”, J. Wiley, 2000.

Q & A

Thank you very much for your attention!

More material in
http://icarus.csd.auth.gr/cvml-web-lecture-series/

Contact: Prof. I. Pitas
pitas@csd.auth.gr