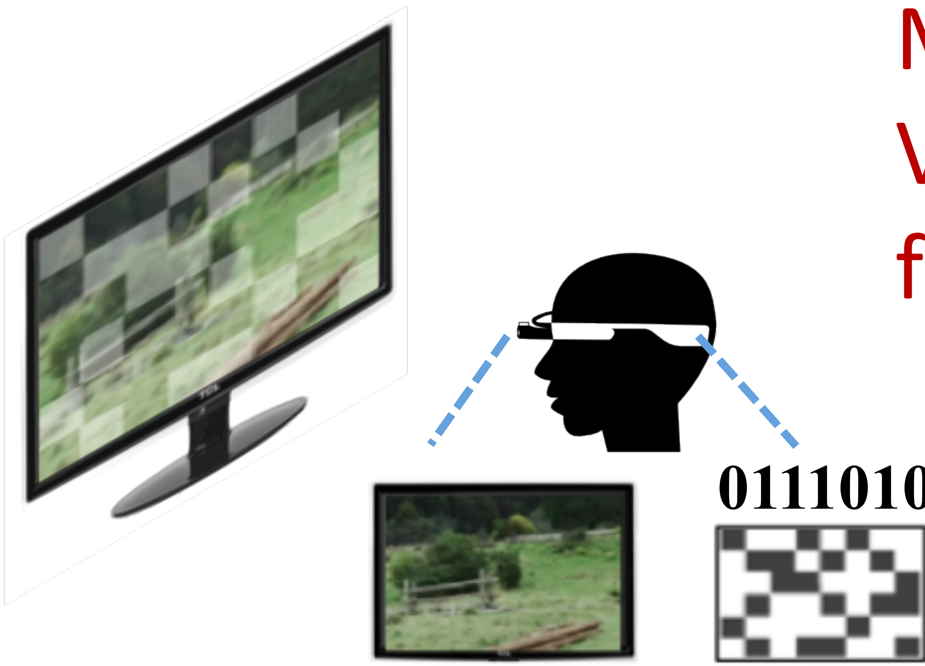
# InFrame:

## Multiflexing Full-Frame Visible Communication for Humans and Devices




Anran Wang **Chunyi Peng** Ouyang Zhang
Guobin Shen Bing Zeng

THE OHIO STATE UNIVERSITY
BEIHANG UNIVERSITY 1952
Microsoft Research
UESTC 1956

# Vision, the most important sense

# Display, not only for human eyes

# Emerging Screen-Camera Communication



# But, tension in display allocation arises


- Small area for device
  - Limit carried bytes
  - Add more pains
  - Still, distractive

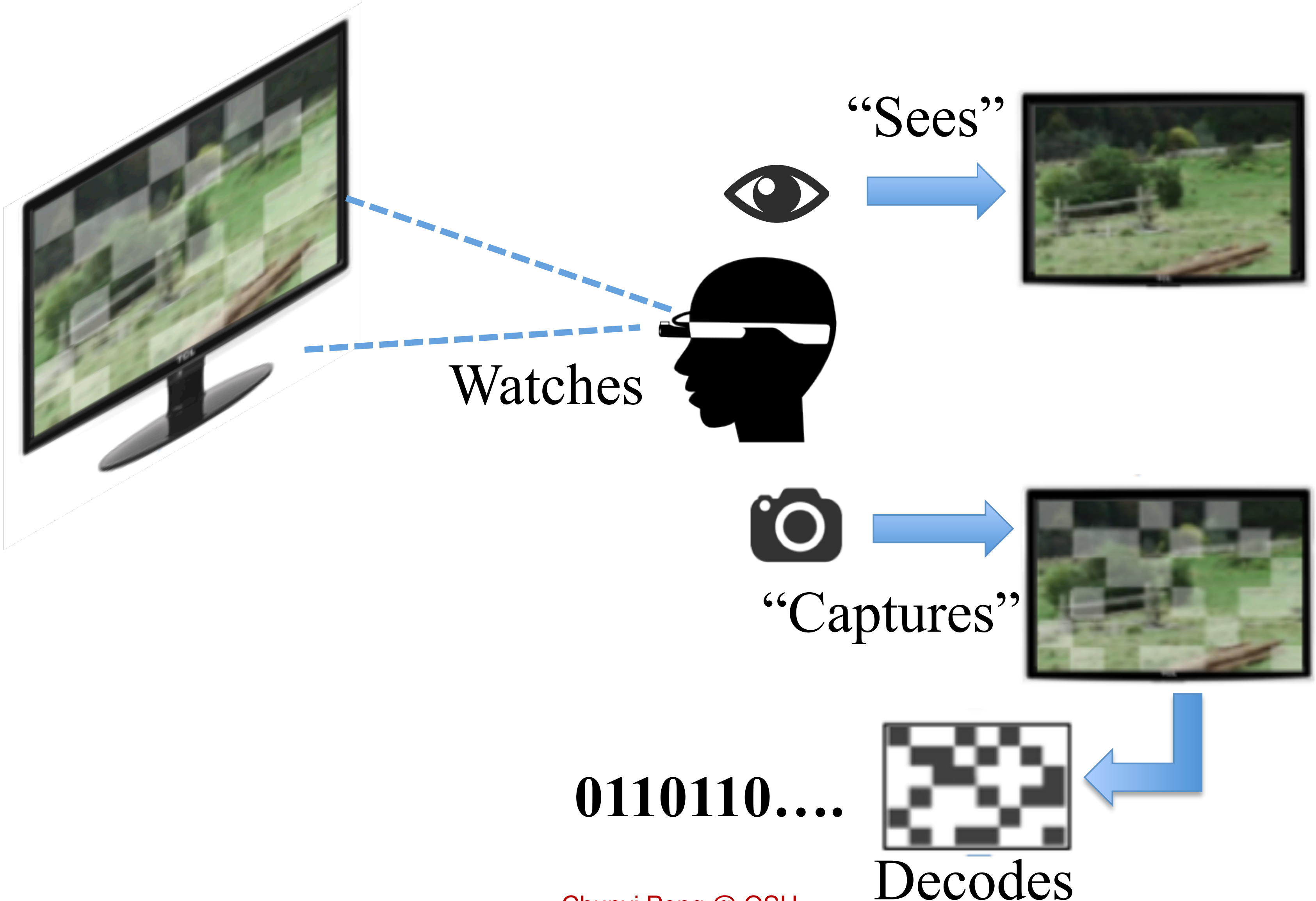
Device-favorable content (e.g., QR code)

Human-favorable content (e.g., video, image)

# Can we do both?

Enabling  full-frame viewing for both human and devices

# YES, WE CAN.

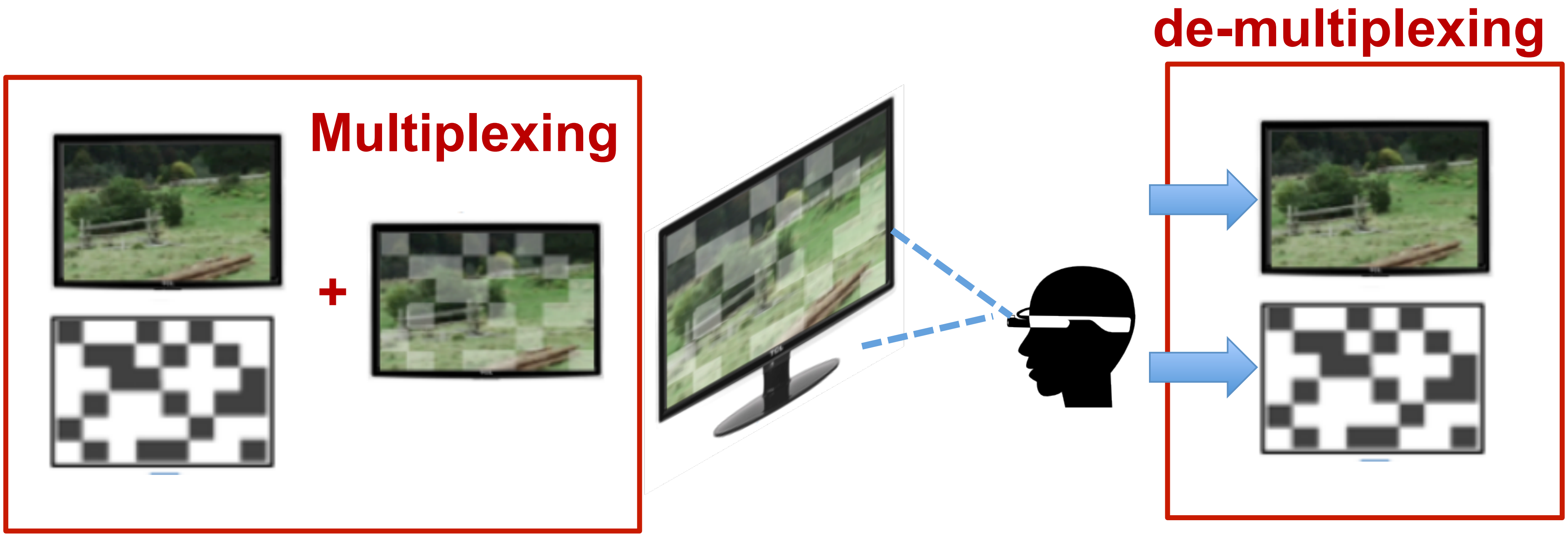


# Demo: video captured by camera

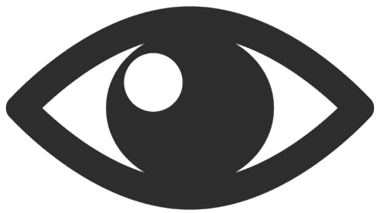
# In fact,  video frames are

# InFrame: Full-Frame, Dual-Mode

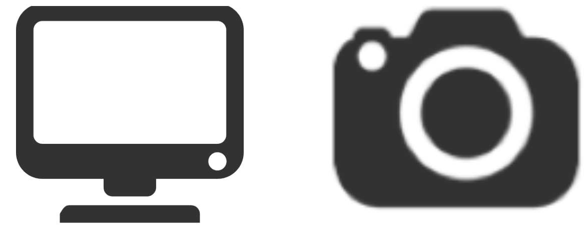
- Screen-Human video watching (primary)
  - Goal#1: no impairing video watching experience
- Screen-Camera data communication
  - Goal#2: reasonable performance (throughput)



# Opportunity: Perception Gap



- Physical limits of eyes
  - Temporal resolution (40-50Hz)
  - Imperceptible: too fast (>50Hz)
  - Flicker fusion: time-variant fluctuations of light intensity are not perceptible to human eyes if beyond a critical flicker frequency (CFF)



- Display: higher fresh rate
  - 120/240+ frame per second (FPS)
- Camera: high resolution, high capture rate, e.g.,
  - iPhone 6: 8M pixel, 240FPS
  - Samsung S5: 16M, 120FPS
- Moreover, device capability is advancing at a faster pace

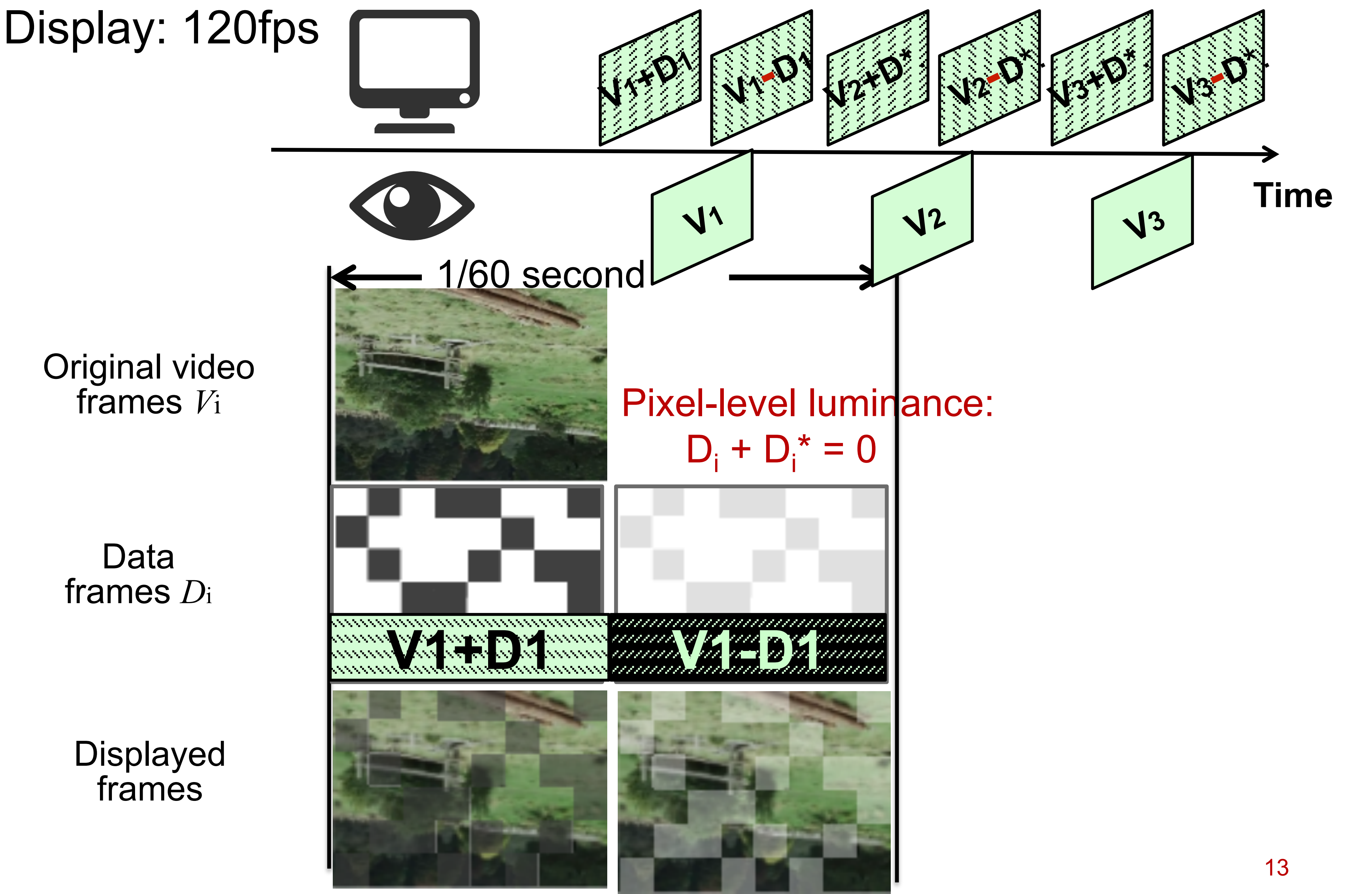
# I. Screen-Human video watching

Goal#1: without impairing video watching

Challenge: human perception is complex and sensitive

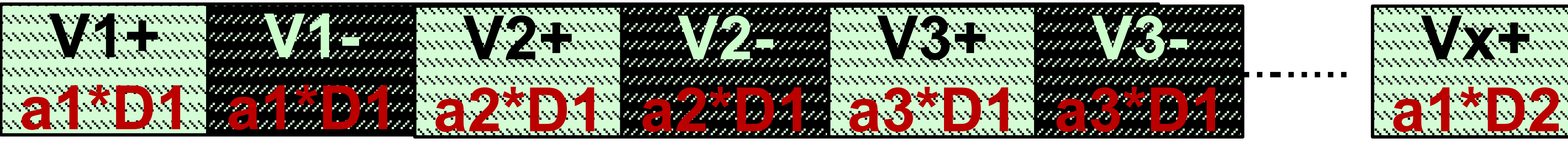
# Complementary Frames

Temporal Fusion

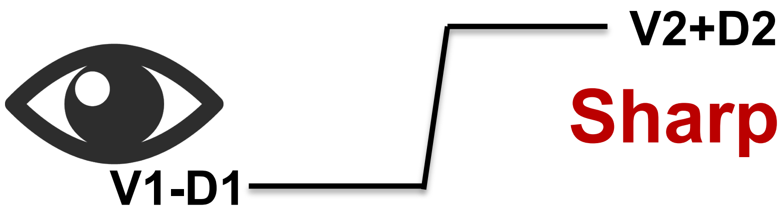


# Data Block Smoothing

Temporal Smoothing



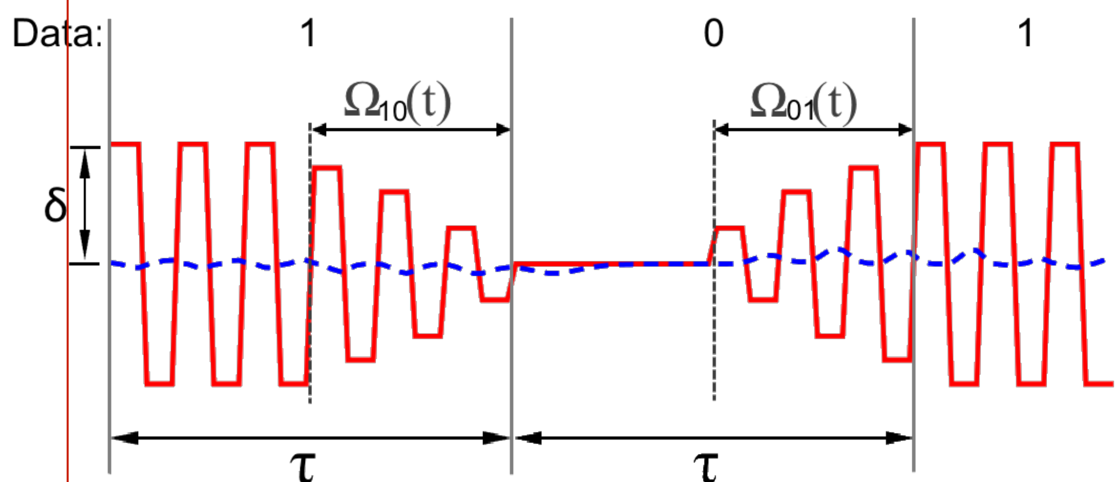

**D1 ≠ D2**




**Sharp transition -> flicker**

Transition duty cycles τ

- pixel-level
- 0 to 1 or 1 to 0
- Amplitude follows a waveform
  - half of the square-root raised Cosine



# II. Screen-Camera data communication

Goal#2: reasonable performance

Challenge: interference from the primary channel

\+ existing constraints in screen-camera comm.

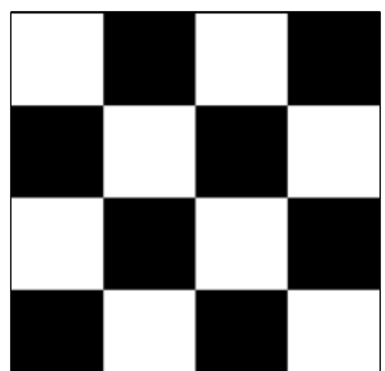
# Visual Pattern Design

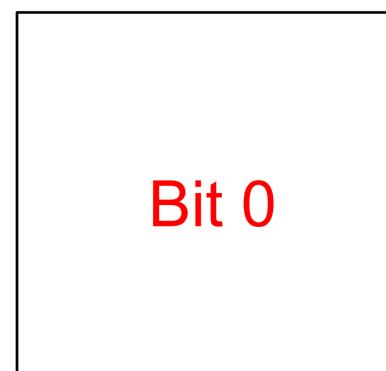
- What is bit 0 or 1?
  - Interference from original video

VS

  - Retrievable from any video

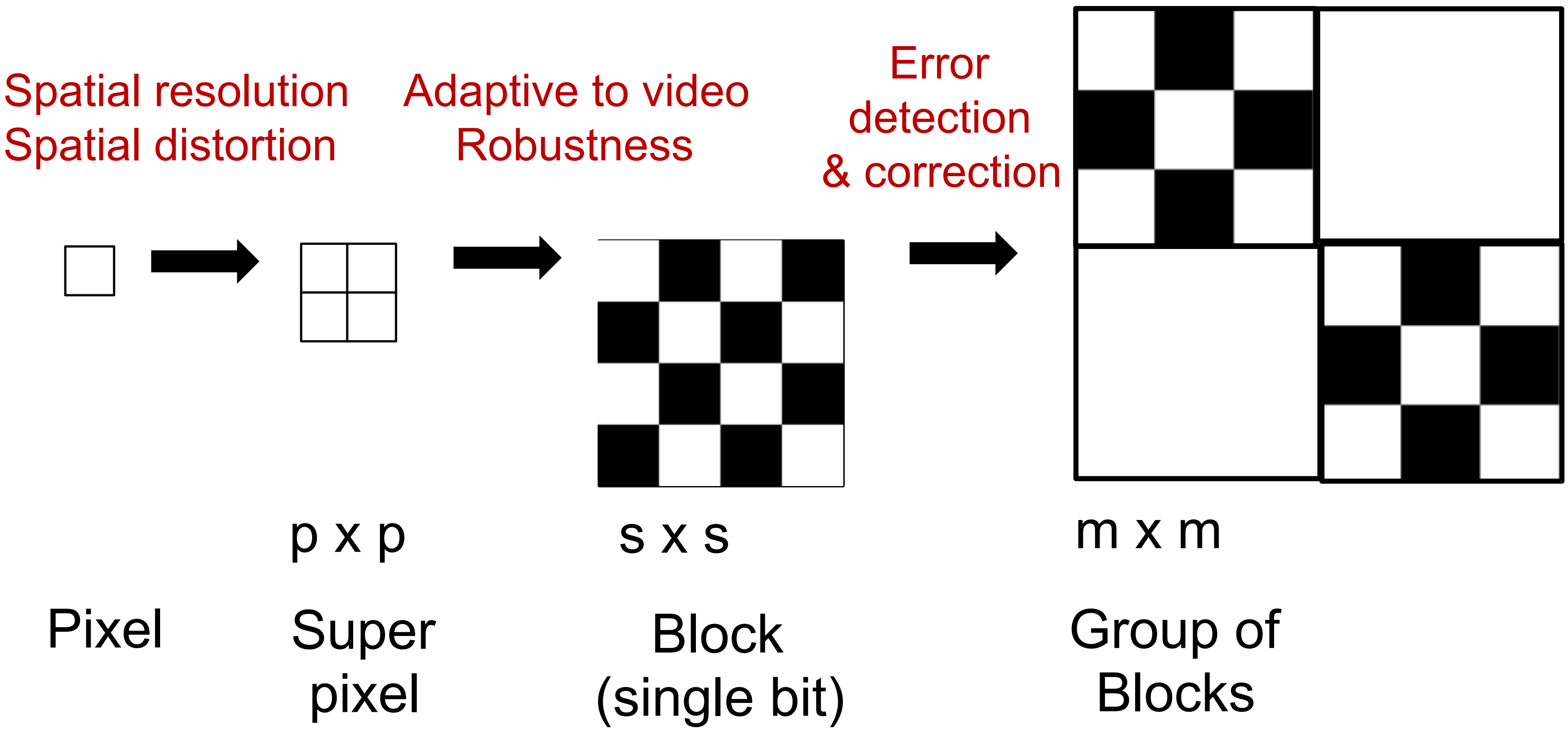
Bit 1





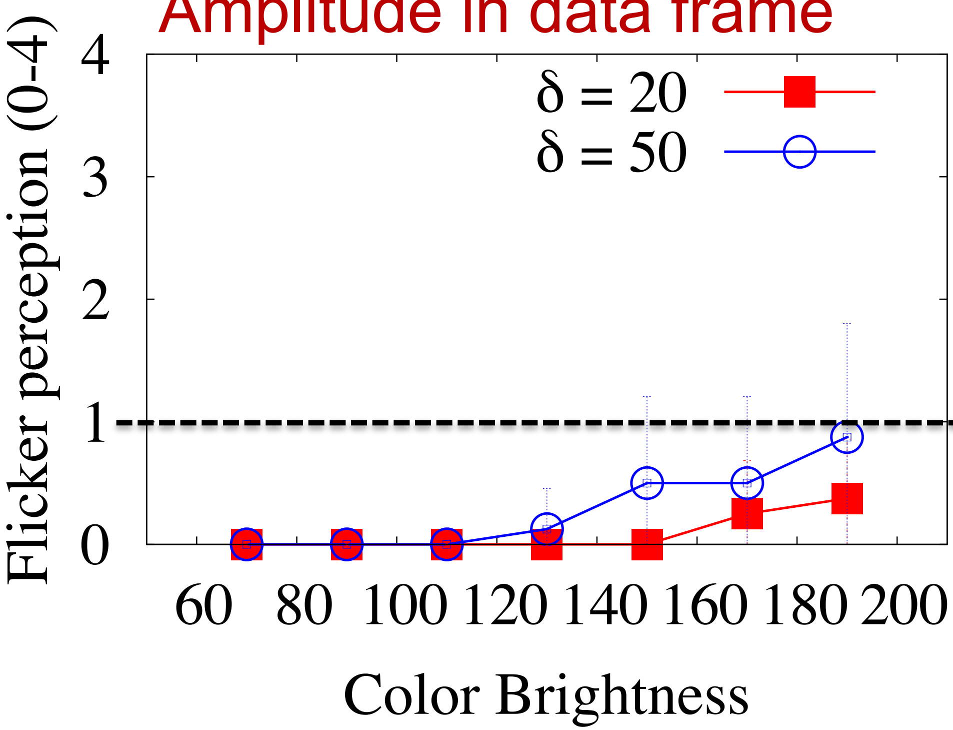
# Data Frame Structure

- Structure: super pixel, block, group of blocks

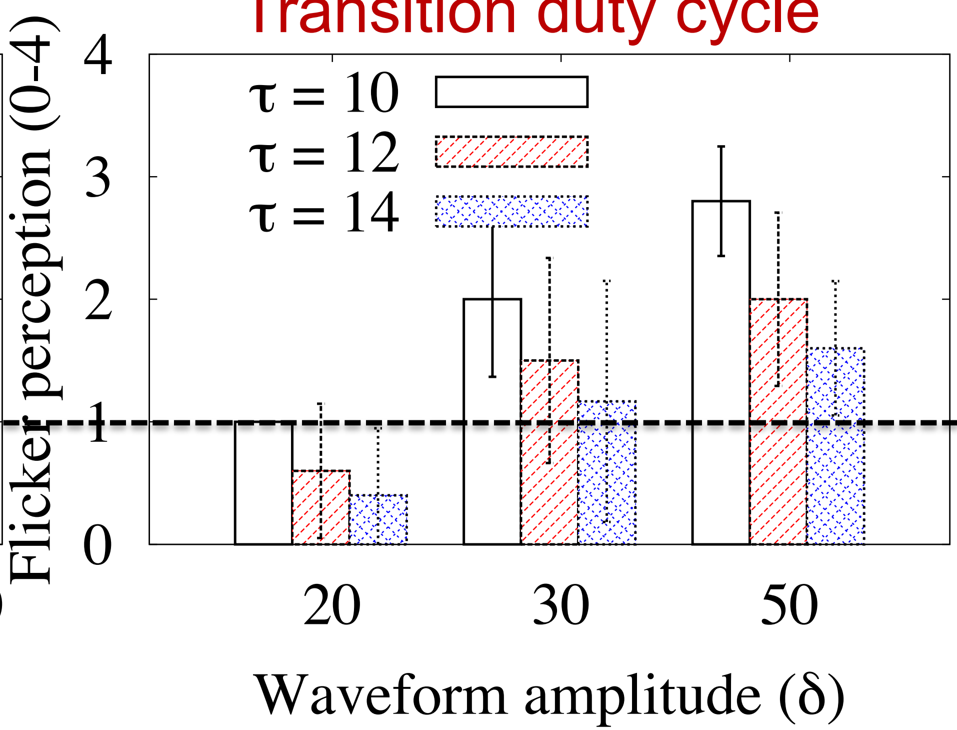


# Evaluation: Feasibility in User Study

- 8 users: 5 males and 3 females
  - Half wearing glasses
  - Two experts on design and video
- Flicker score (0-4): 0/1 = no flicker/almost unnoticeable




Color brightness = (R+G+B)/3



# Throughput

- Amplitude = 20, 30,   cycle = 10, 12, 14
- Three videos:

| Video | Throughput | Decoding Rate |
|---|---|---|
|  | 9.2~**12.6Kbps** | Available GOB: 95-98%<br>Correction rate: 98-99% |
|  | 9.2-**12.8Kbps** | Available GOB: 96-97%<br>Correction rate: 98-99% |
|  | 5.0~**7.0Kbps** | Available GOB: 60-70%<br>Correction rate: 80-90% |

# Still, Open issues and ongoing work

- Applications for InFrame
  - Screen-Human link:  context for human activities
  - Screen-Device link: eFormat (byte/text) -> automatic logging and processing
    - Video Ad + digital Coupons
- Throughput: a better multiplexing and frame design
- Practical issues
  - Viewing operations (video play, pause, forward…)
  - Camera capture quality; anchor design
  - Real-time rendering, computation and energy costs

# Summary

- Full-frame video viewing for both human and devices
  - One visible channel, two views
- Opportunity: perception gap between human and device
  - Even larger in the future
- Two-mode communication & sensing paradigm
  - Human: attention-free, or even effort-free
  - Device: opportunistic channel