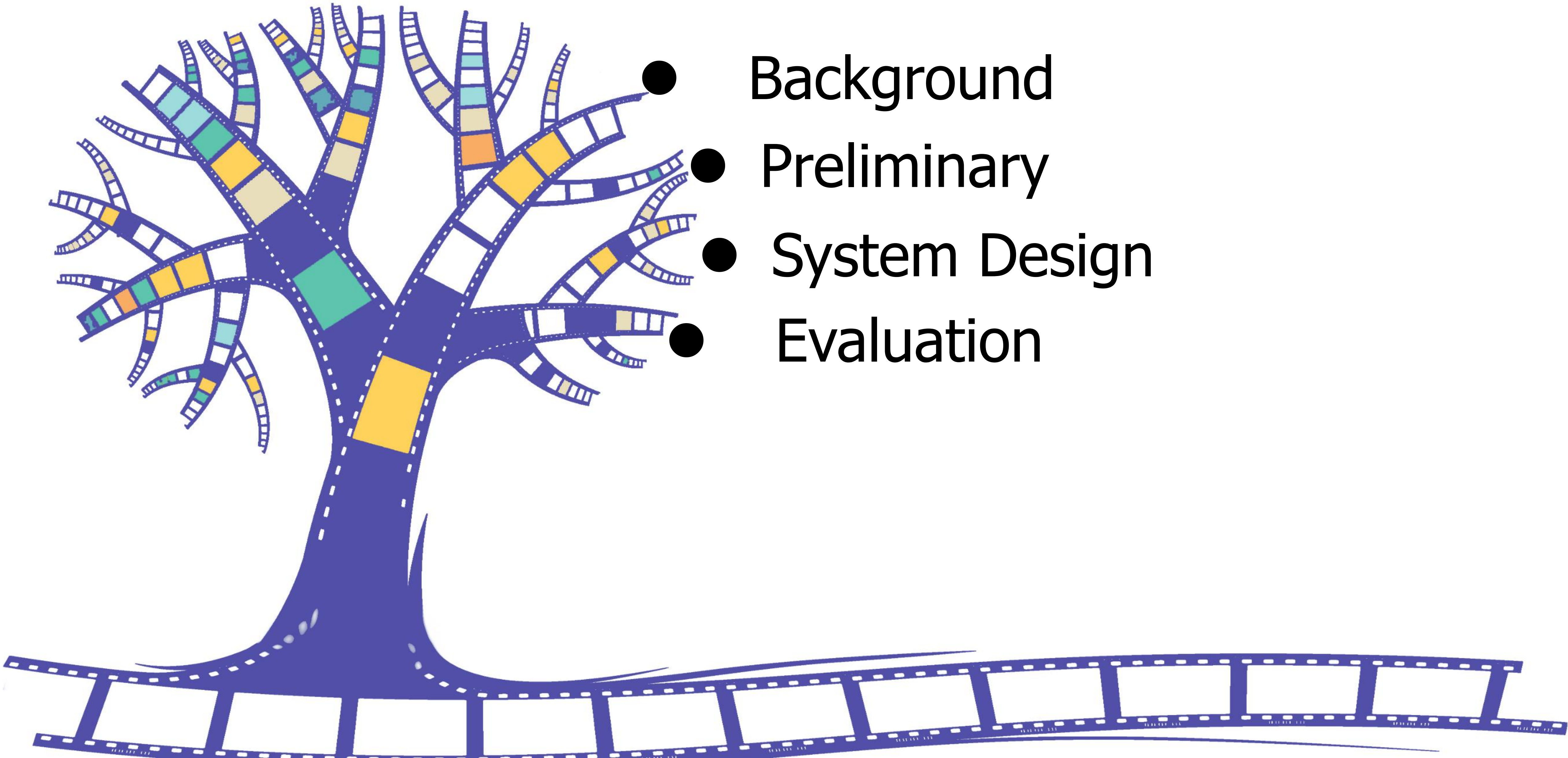
MobiCom'15

Kaleido: You Can Watch It But Cannot Record It

Lan Zhang, Cheng Bo, Jiahui Hou, Xiang-Yang Li, Yu Wang, Kebin Liu, Yunhao Liu

Tsinghua University, China
University of North Carolina at Charlotte, USA
Illinois Institute of Technology, USA

Outline



- Background
- Preliminary
- System Design
- Evaluation

Background

Blooming of Visual Displays

- Movie screen, phone, tablet, TV, electronic board…
- Entertainment, news, advertising, tour guide…

Video playback has contributed to about 80% of the Internet traffic [1].

[1] Cisco, I. Cisco visual networking index: Forecast and methodology, 2013–2018. CISCO White paper (2014), 2013–2018.

Piracy - A Growing Threat

Film piracy causes lost of revenue about $20.5 billion annually.

Over 90% of this illegal online content is delivered from these pirate movies.

Unauthorized videotaping during the exhibition, presentation or project demonstration could cause infringement of copyright and even plagiarism.

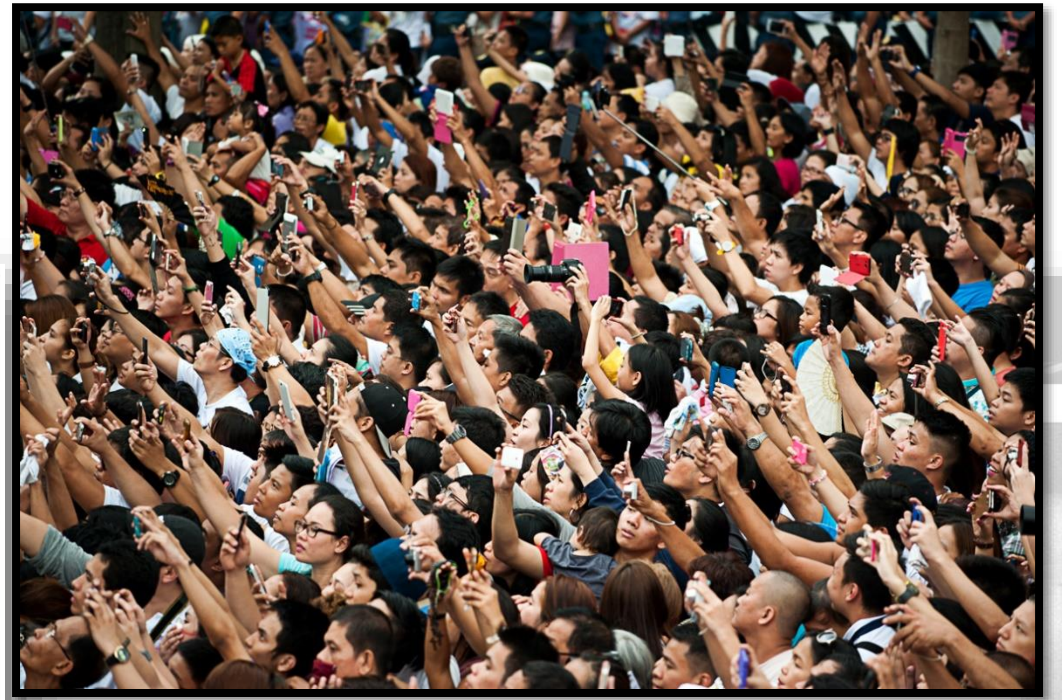
Copyright Protection

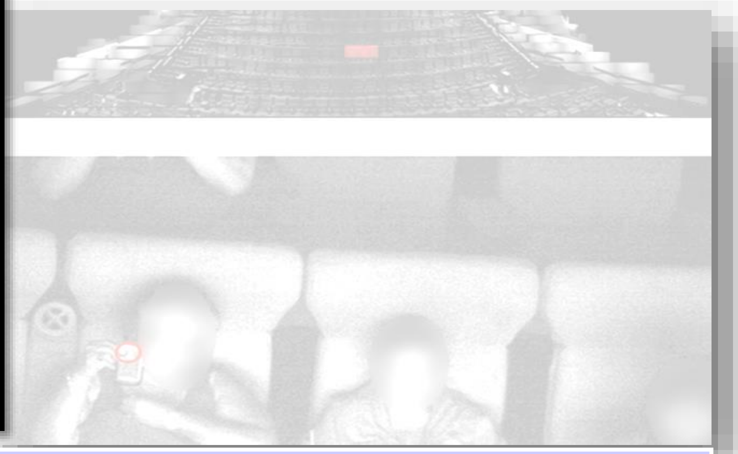
Watermark

No camera policy

Camera detect

Copyright Protection



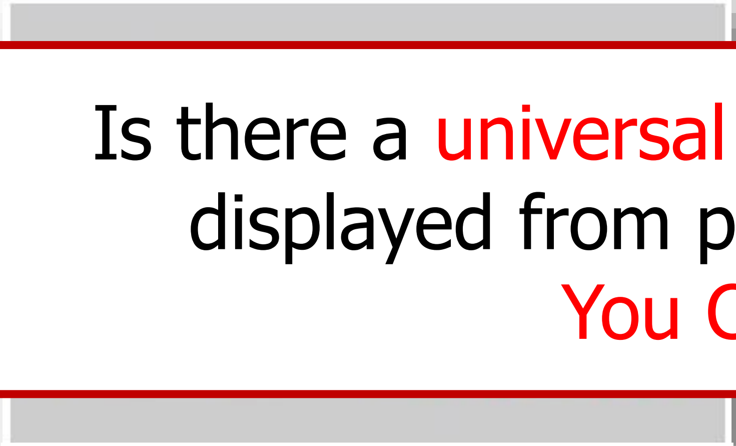


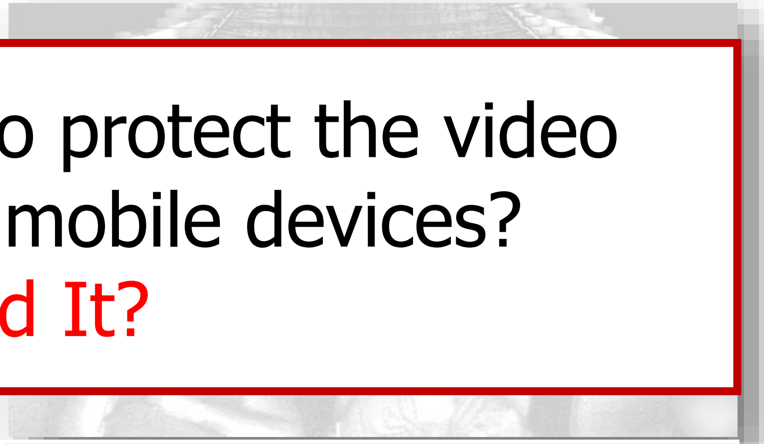
Wa

tect

Ineffective in preventing attendees from taking pirate video for later redisplay, especially when using mobile devices.

Copyright Protection





Is there a universal technology that can be used to protect the video displayed from pirate videotaping using typical mobile devices?

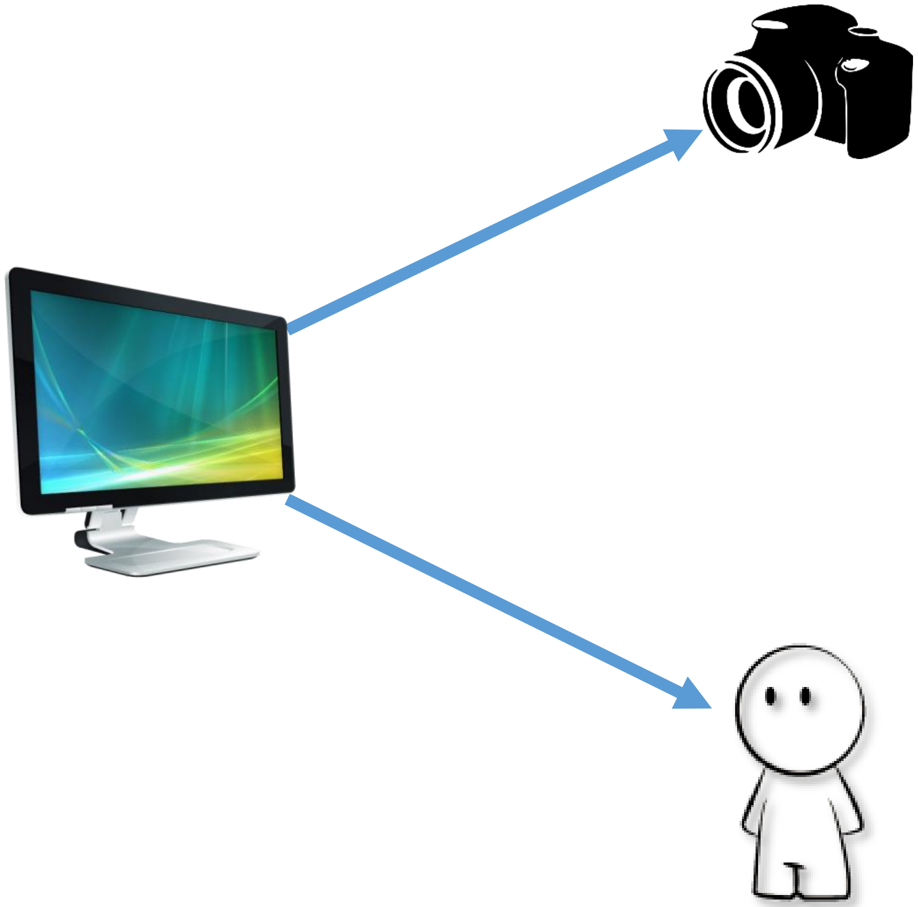
You Can Watch It But Cannot Record It?

Watermark

No camera policy

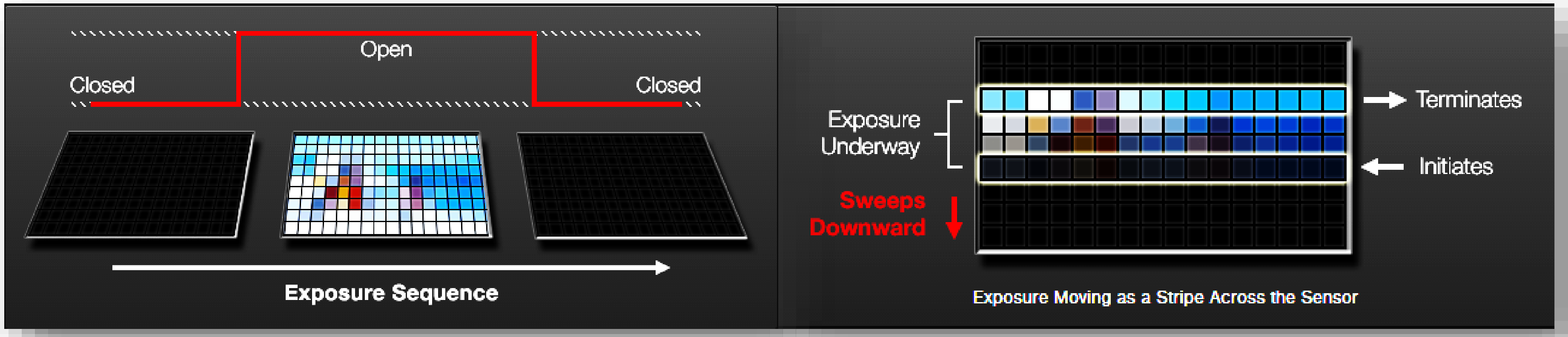
Camera Detect

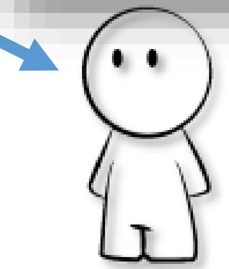
Screen-Camera v.s. Screen-Eye



Screen-Camera v.s. Screen-Eye

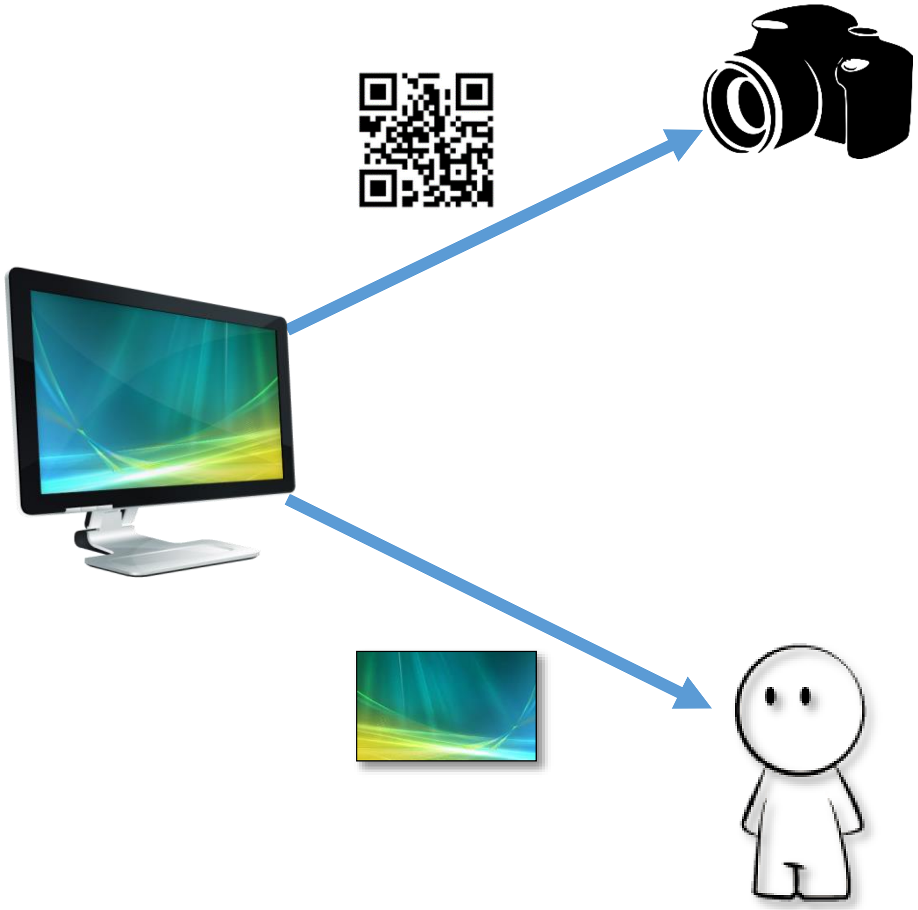
- Discrete global/rolling shutter , higher temporal resolution.





- Continuous “global shutter”, low-pass filter.

Screen-Camera Communication



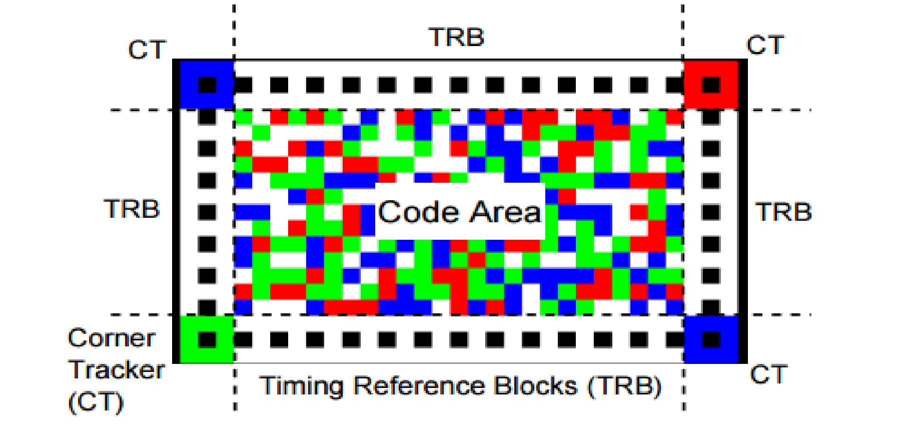
- Encode information into the screen-camera side-channel.
- The extra signal can be captured by camera but not the human eye.

Existing Methods

PixNet (SiGCOMM'11)

PixNet leverages 2D OFDM to modulate high-throughput 2D barcode frame.

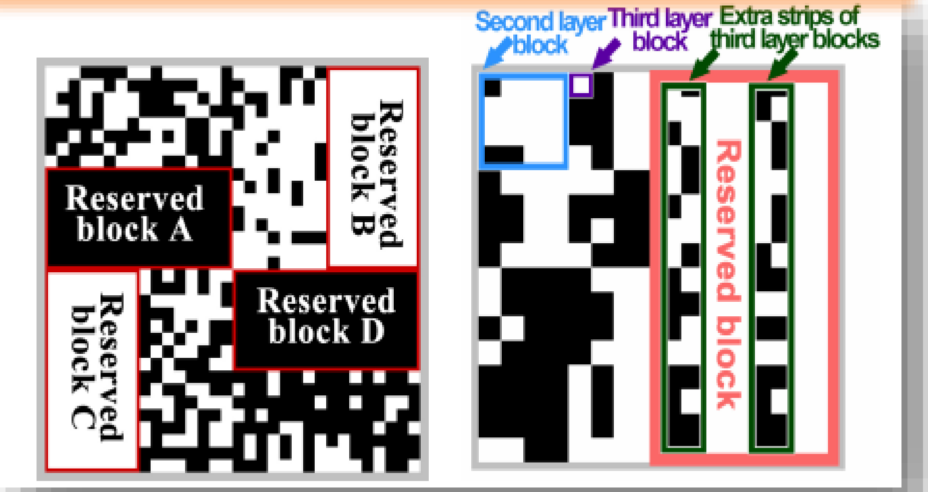
COBRA (MobiSys'12)



COBRA [20] achieves real-time phone-to-phone optical communication.

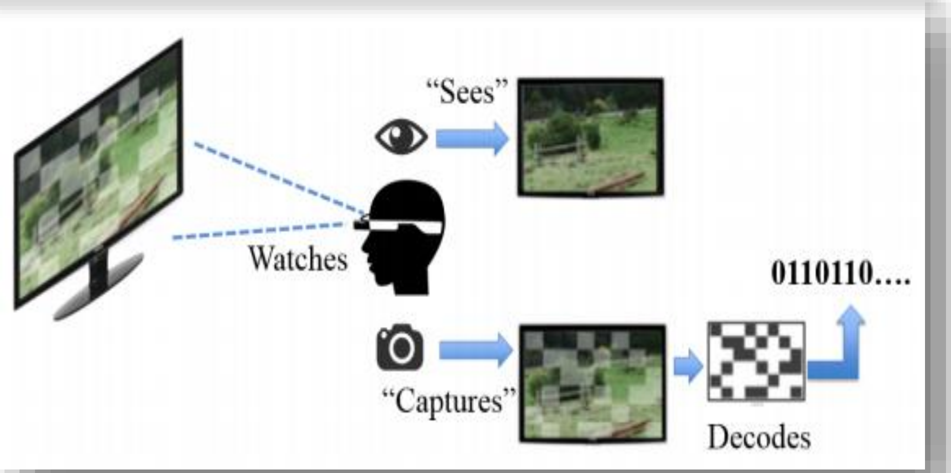
Existing Methods

Strata (MobiCom'14)



Strata supports wide range of frame capture resolutions and rates so as to deliver information rate correspondingly.

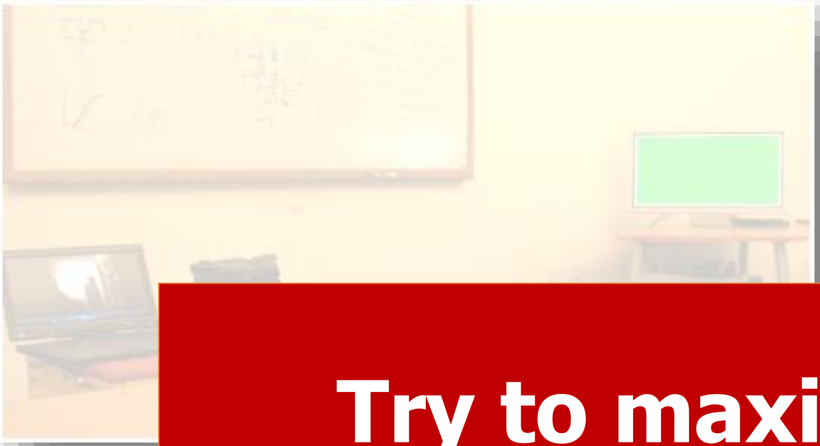
InFrame++ (MobiSys'15)



InFrame++ achieves dual-mode full frame communication between screen and both humans and devices simultaneously.

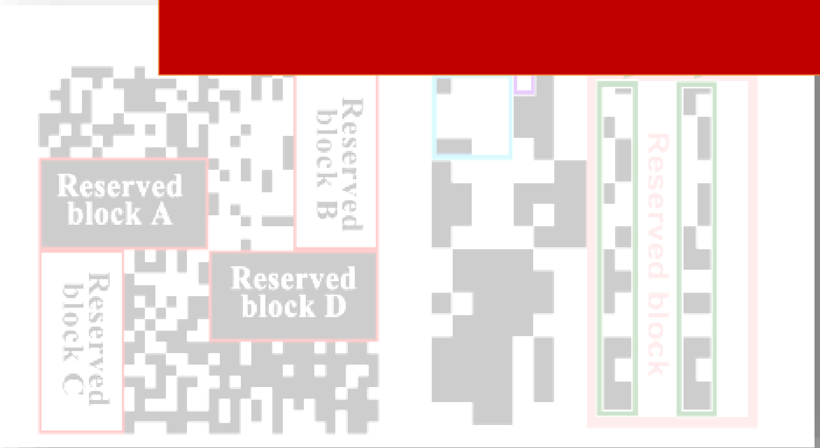
Existing Methods

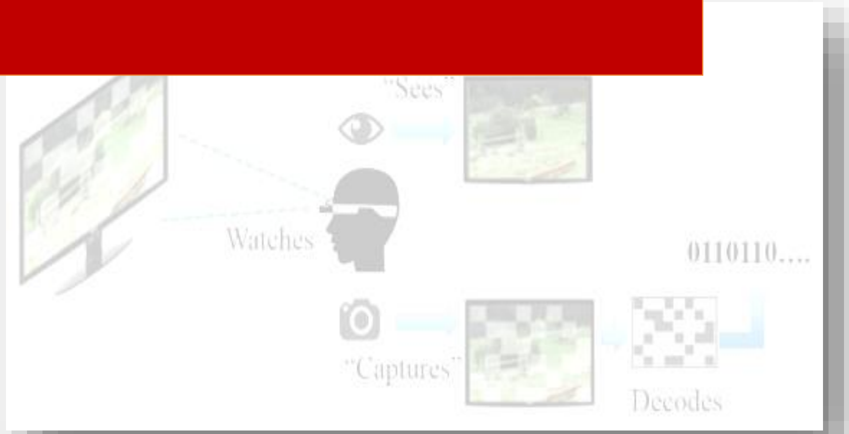
PixNet (SiGCOMM'11)



COBRA (MobiSys'12)

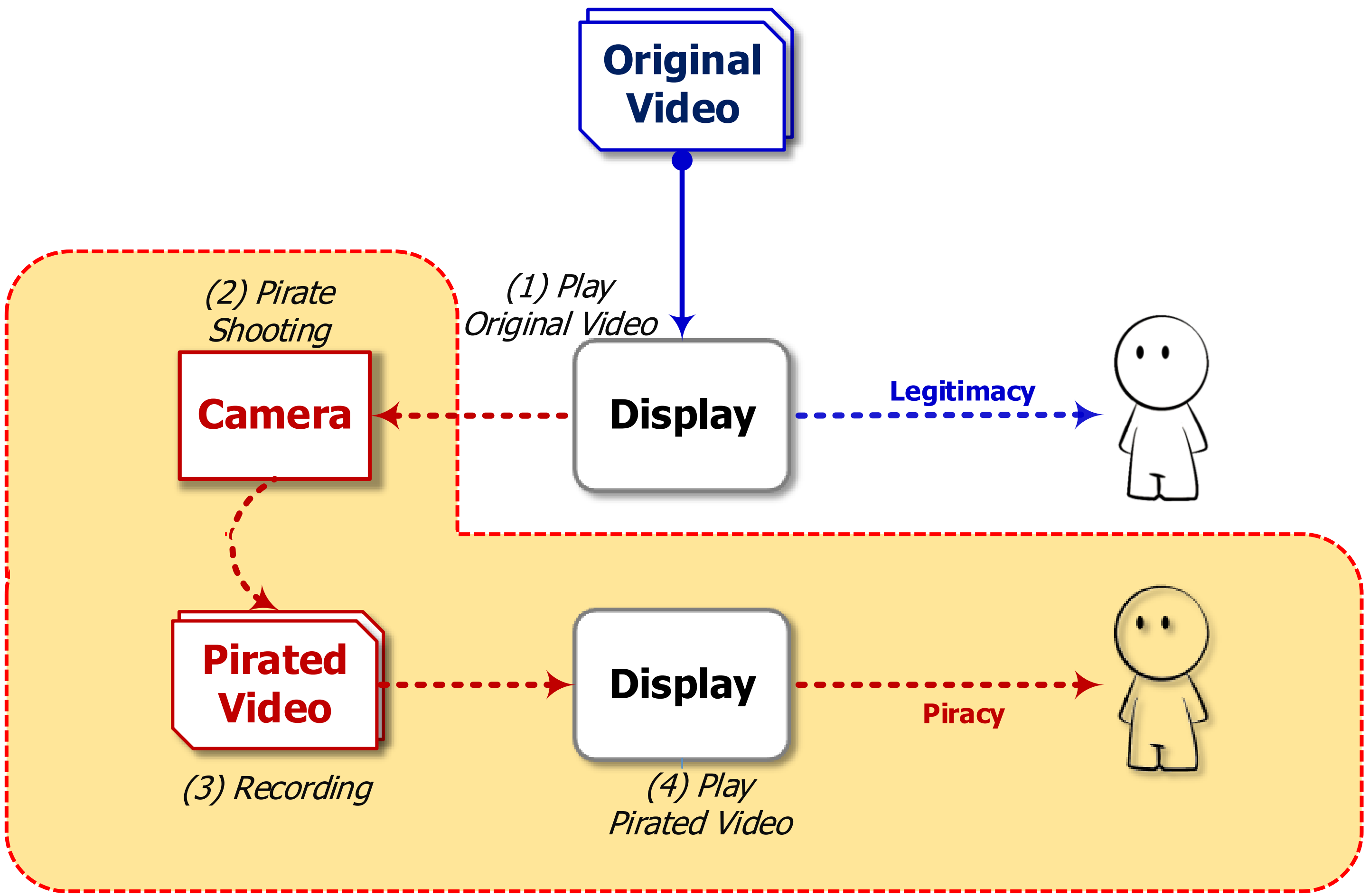
Str





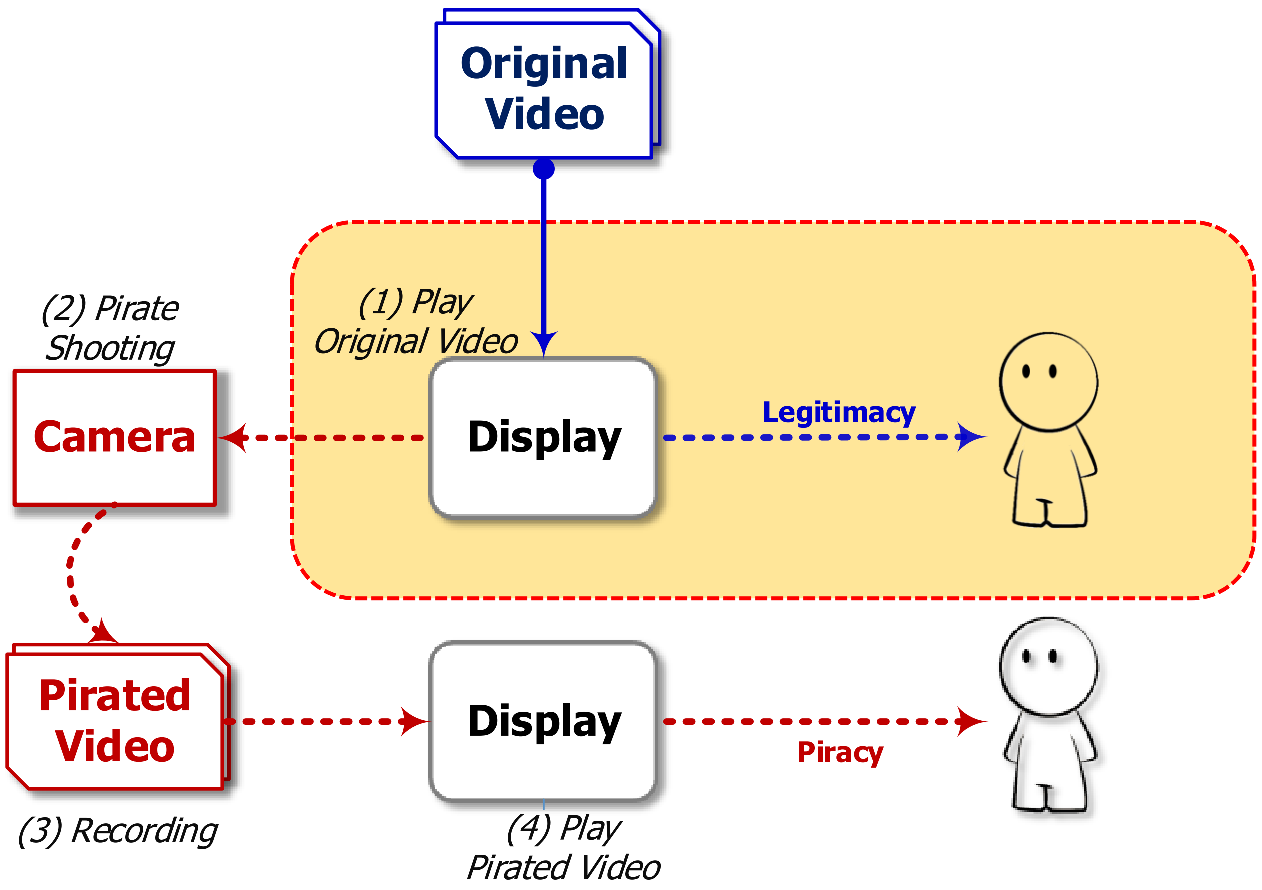
Try to maximize decodability of screen-camera channel.

Kaleido



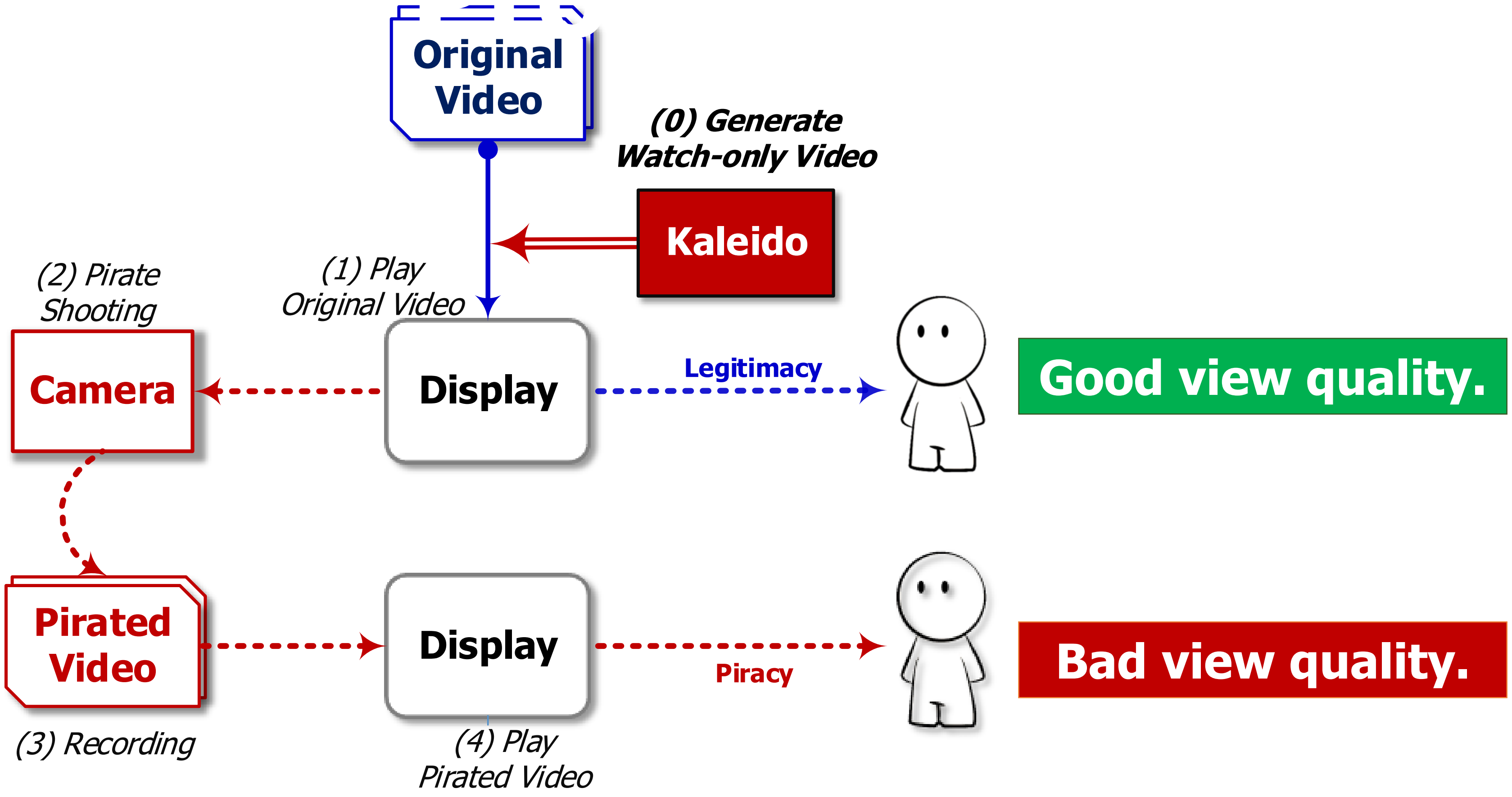
Prevent unauthorized users from videotaping a video played on a screen for high-quality redisplay.

Kaleido



Do not affect the viewing experience of live audiences.
Do not use any extra hardware device.

Kaleido



Kaleido: generate a **watch-only** version of the video .

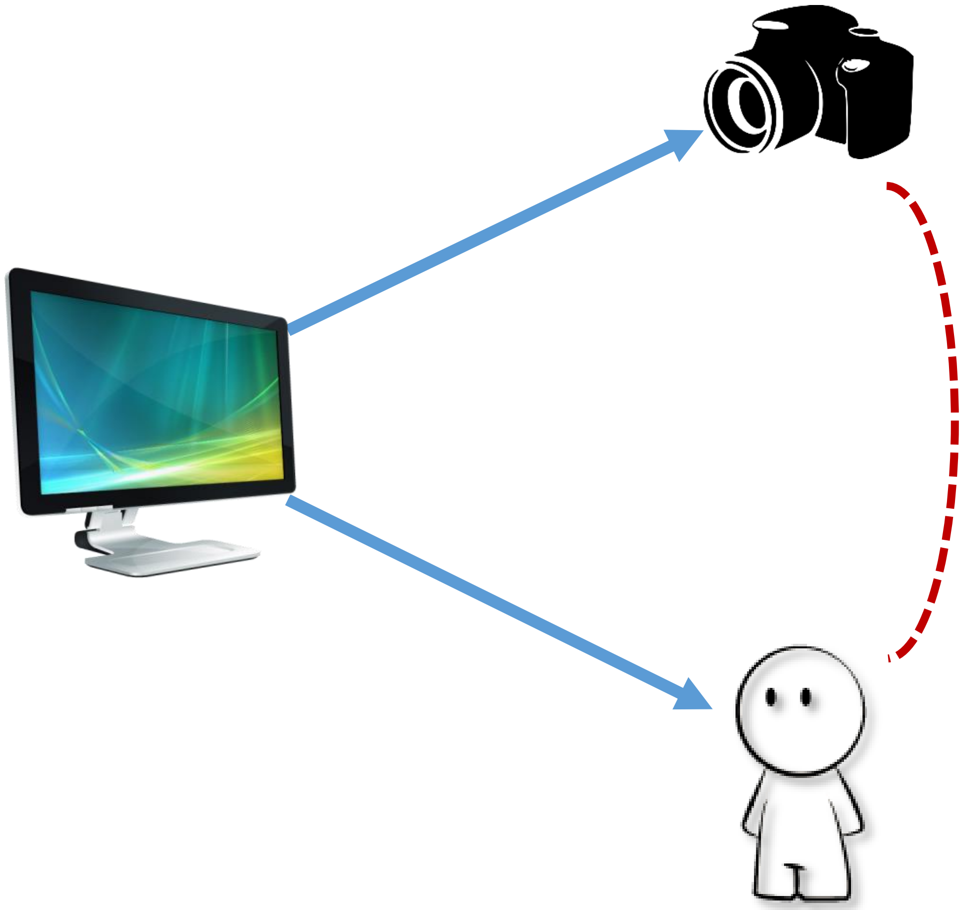
Existing

Try to maximize decodability of screen-camera channel.

Kaleido

We seek to maximize the quality degradation of the display-camera channel while retain the quality of the screen-eye channel.

Challenges

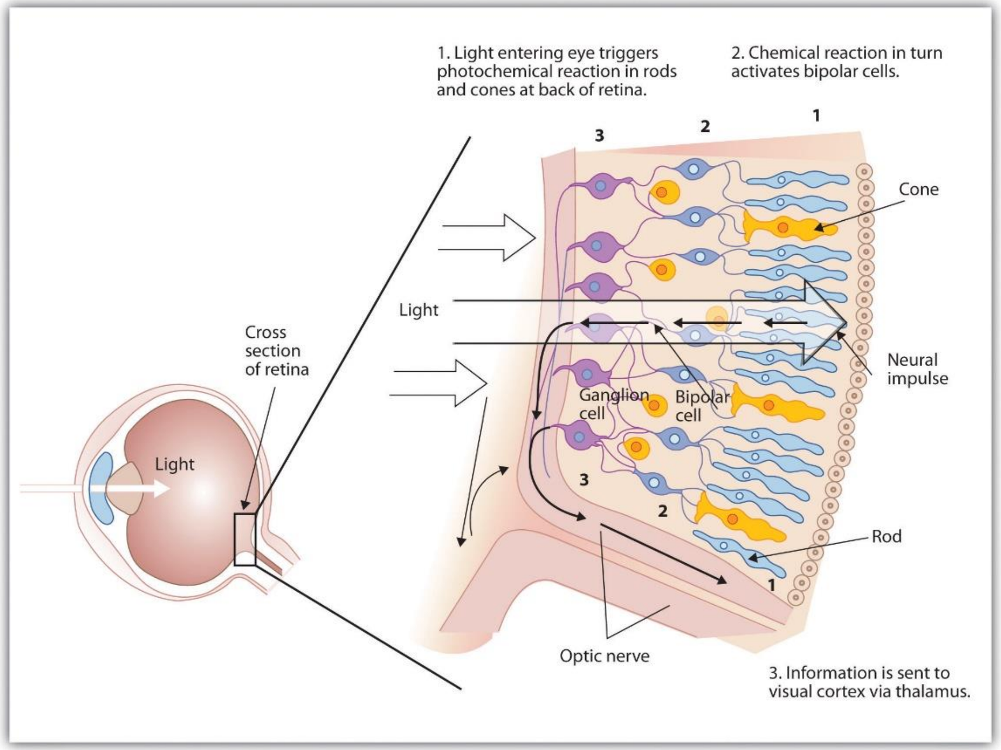


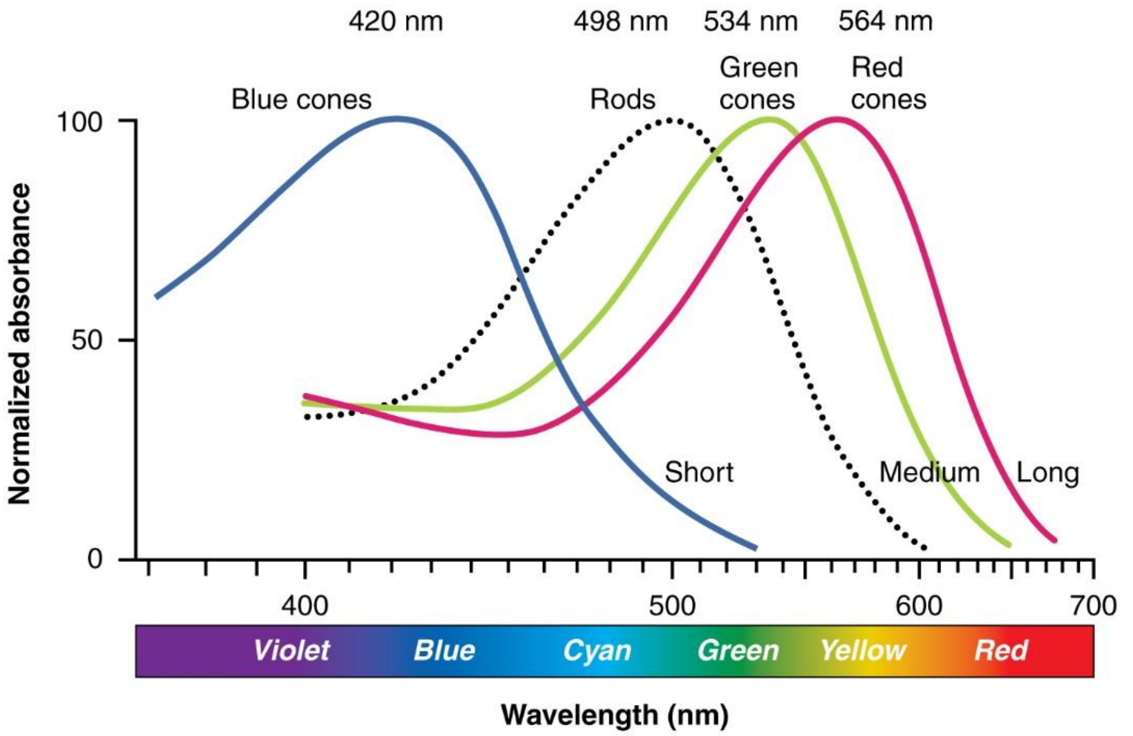
Cameras are designed to mimic the human eye.

Limited disparities, very small design opportunity.

Preliminary

Different light wavelengths stimulate the three kinds of cone cells of a viewer in different degrees, providing her perception of distinct colors.

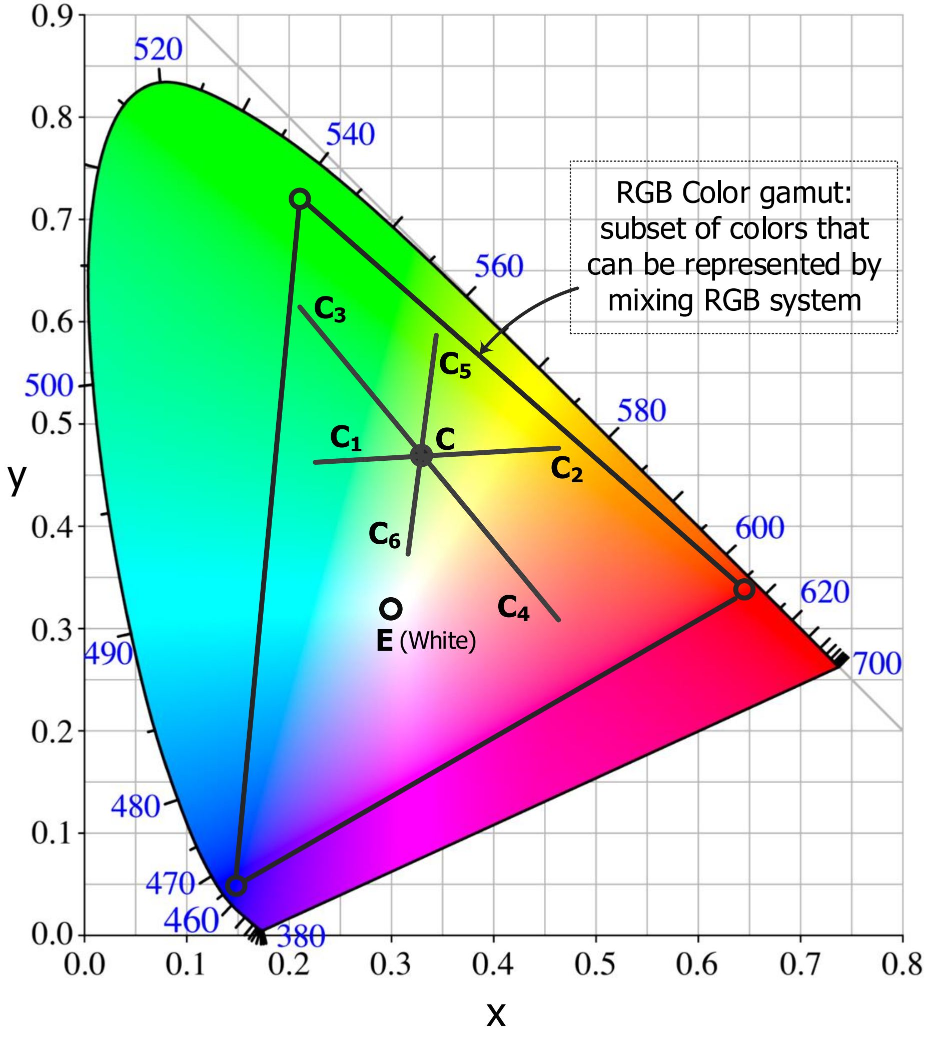




Human Vision

1931 CIE Color Space

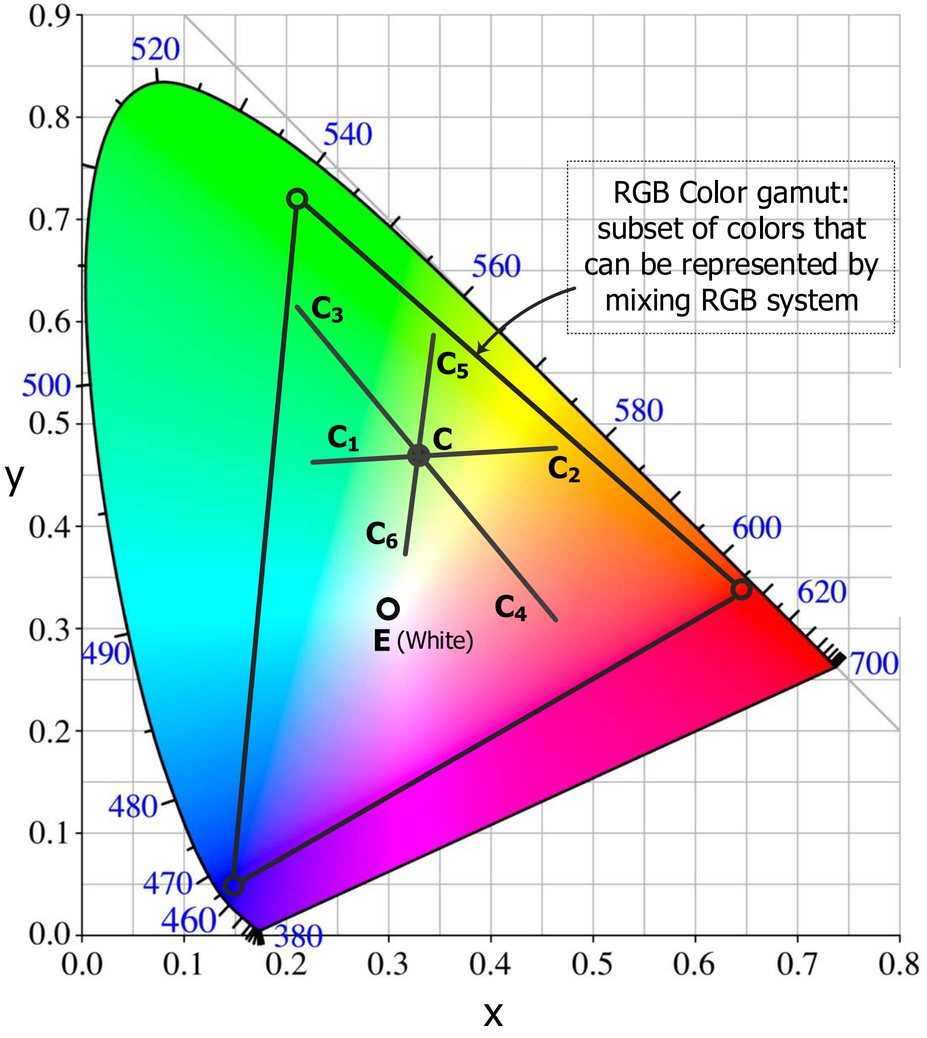
CIE chromatic diagram



- Two aspects:
  - illuminance
  - chromaticity (composition of the light spectra).
- (x, y, Y) to present a color
  - Y determines the illuminance
  - x and y give chromaticity at that luminance.

Spectral Color Mix

CIE chromatic diagram



Spectral Color Additive Rule

$$(x_1, y_1, Y_1) \quad (x_2, y_2, Y_2)$$

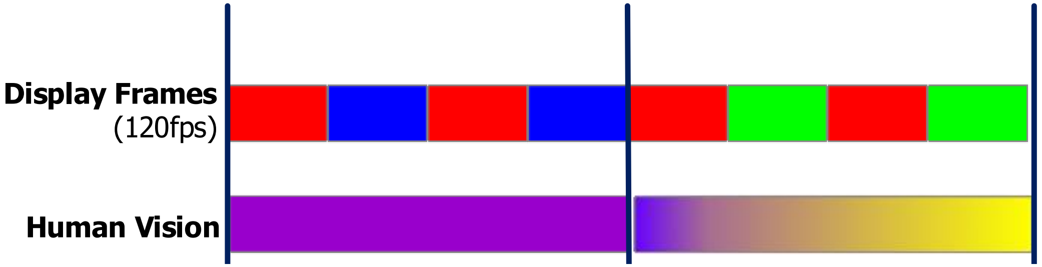
Shining them simultaneously

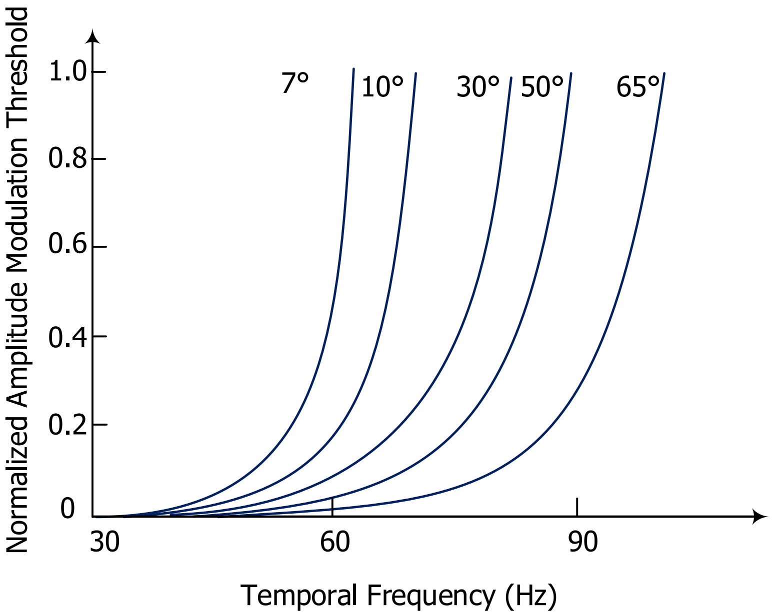
$$\begin{cases} (x, y) = \frac{Y_1}{Y_1+Y_2}(x_1, y_1) + \frac{Y_2}{Y_1+Y_2}(x_2, y_2) \\ Y = (Y_1 + Y_2)/2 \end{cases}$$

- Mixed color depending on the relative brightness.
- The **combination** of colors to produce a given perceived color **is not unique**.

Temporal Change Color

Temporal Color Additive Perception

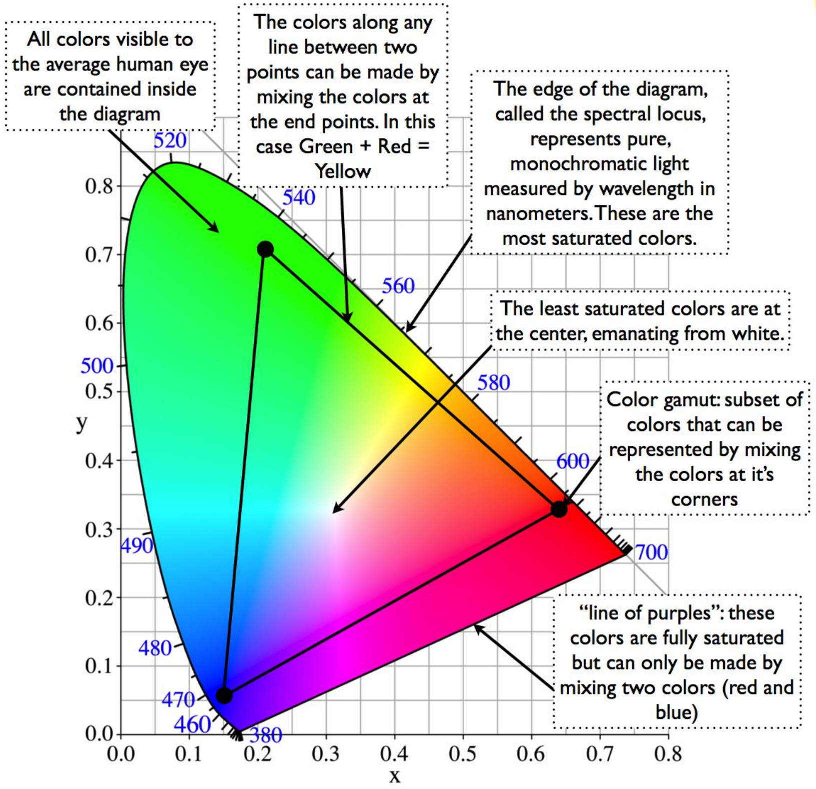




- When the flicker frequency is larger than Critical Flicker Frequency (CFF), the illuminance/chromatic flicker stimulus from a sequence of continuous frames are only perceived as time-averaged luminance/wavelength respectively.
- The averaged chromaticity is determined based on the spectral color additive rule.
- Typically, human eyes can only resolve up to 50Hz to luminance flicker and 25Hz to chromatic flicker .

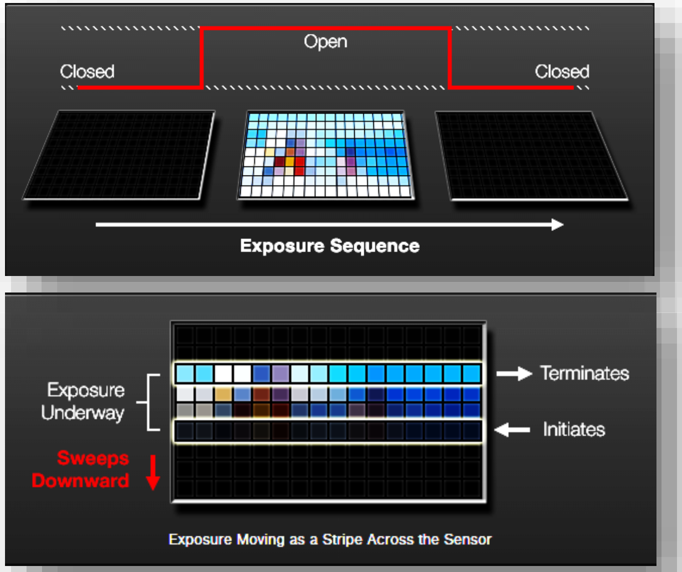
$$A(f) = a \cdot e^{bf} \rightarrow CFF = \frac{\ln[A(f)/a]}{b}$$

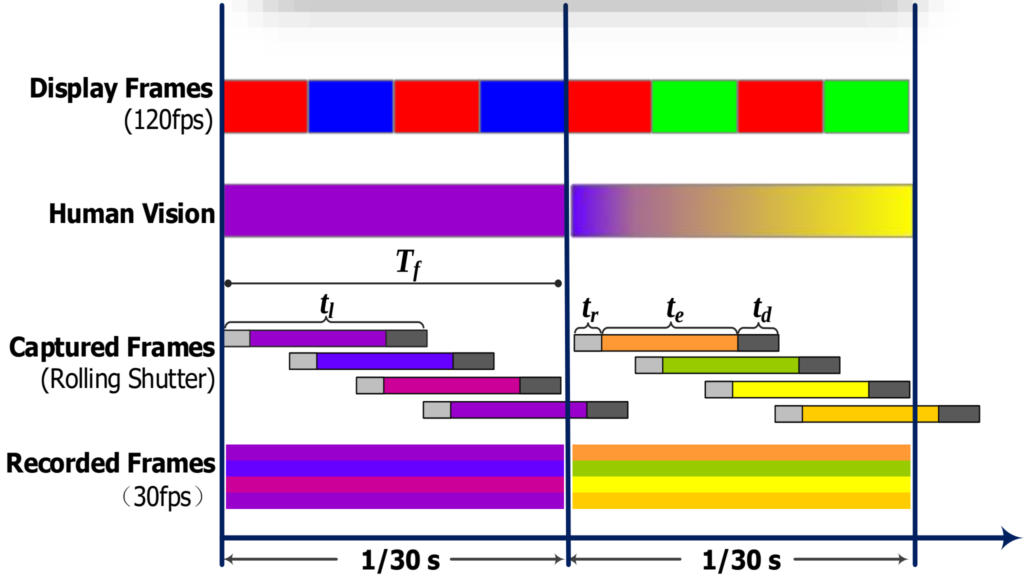
Gamut of an RGB display



- LCD monitor & Projector
  >=120 refresh rate
- Video: 24 or 30 frame rate

Video Encoding & Display





- **Rolling Shutter**
  - By 2013, CMOS (low cost) image sensors takes 97% market share.
  - Total acquisition duration

$$t_l = t_r + t_e + t_d$$

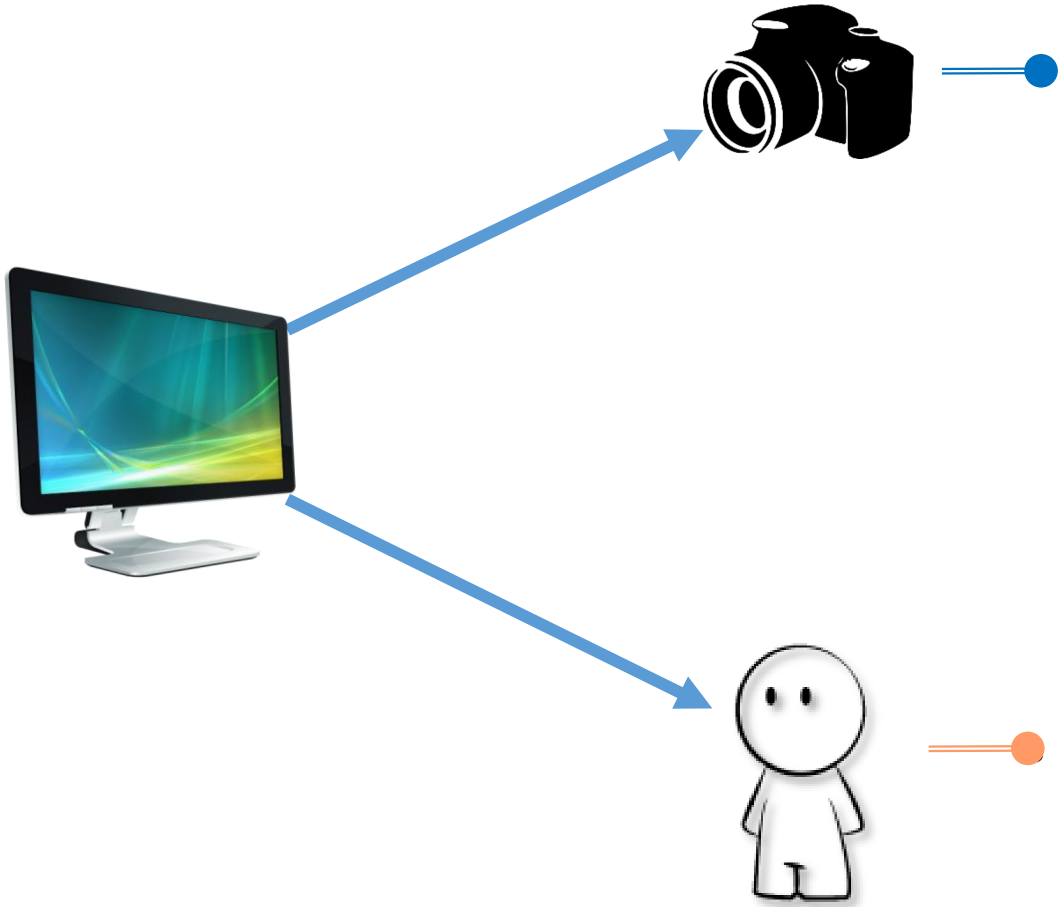
  - Effective light sampling frequency

$$f_s = f_c \times n$$

- **Varying Recording Rate**
  - 24, 30, 60, 120, 240 fps
- **Unstable Inter-frame Interval**

Video Recording

Screen-Camera v.s. Screen-Eye

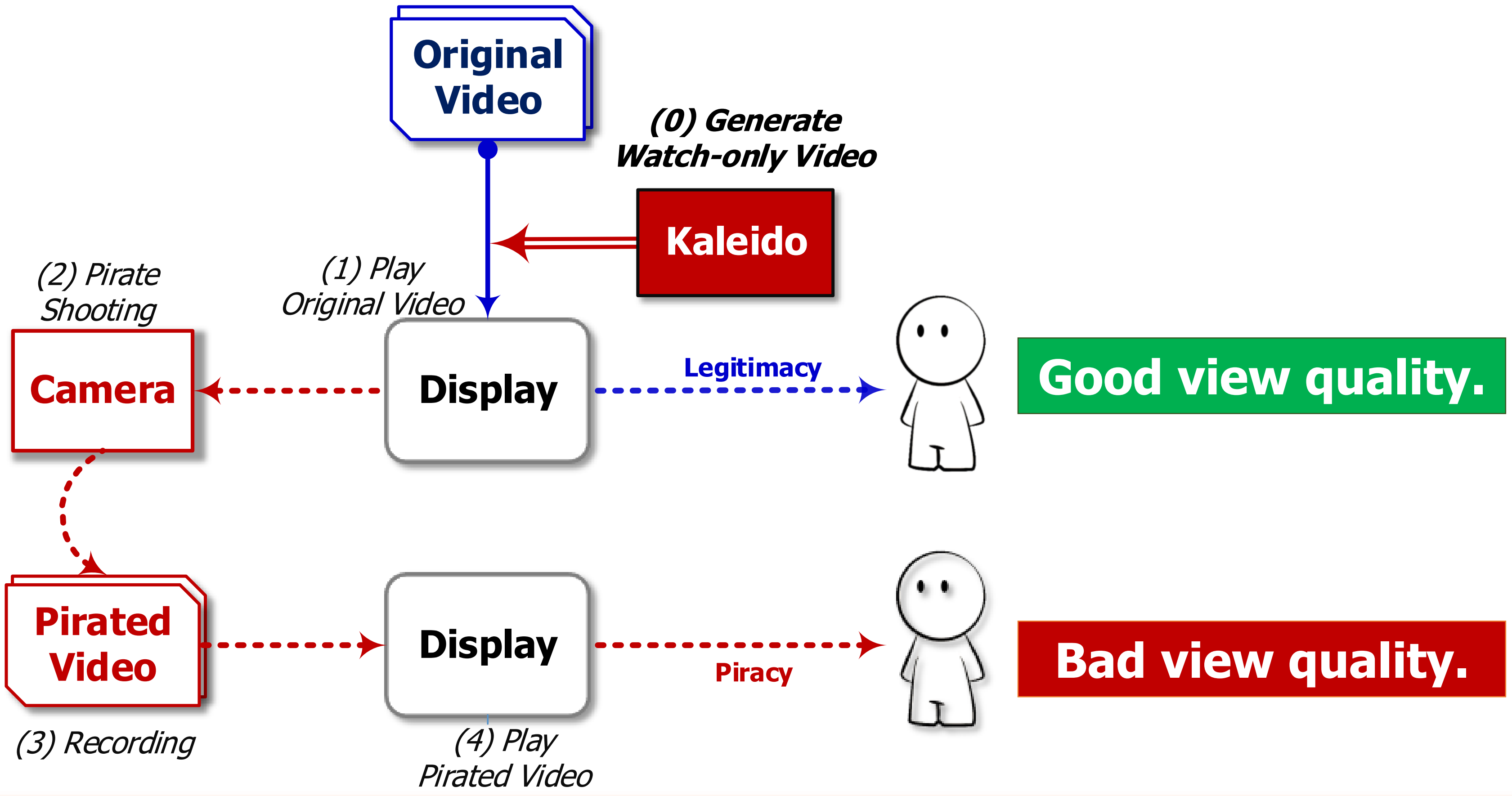


Discrete rolling shutter , higher temporal resolution.

Continuous “global shutter”, low-pass filter.

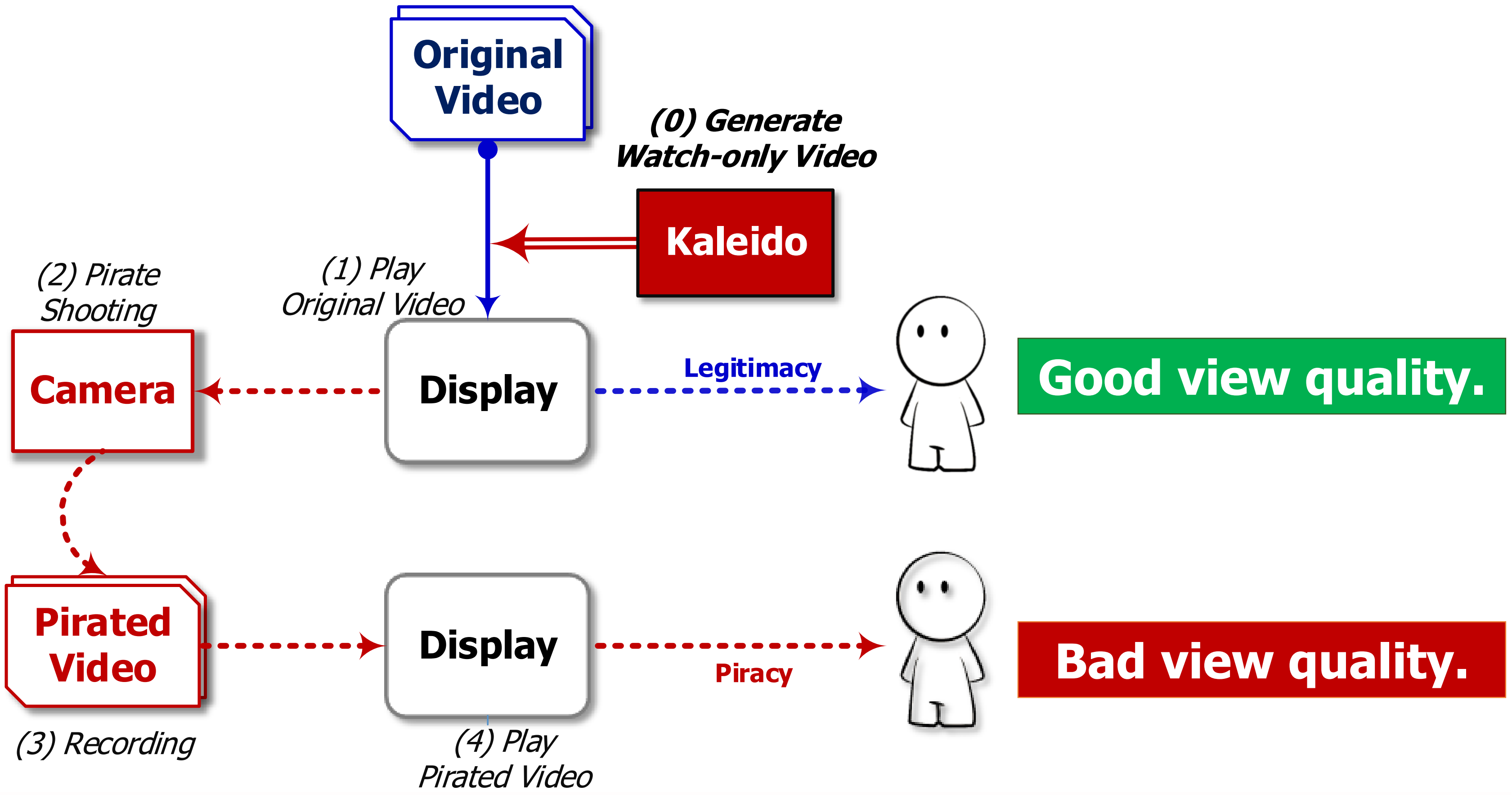
System Design

Challenge & Opportunity



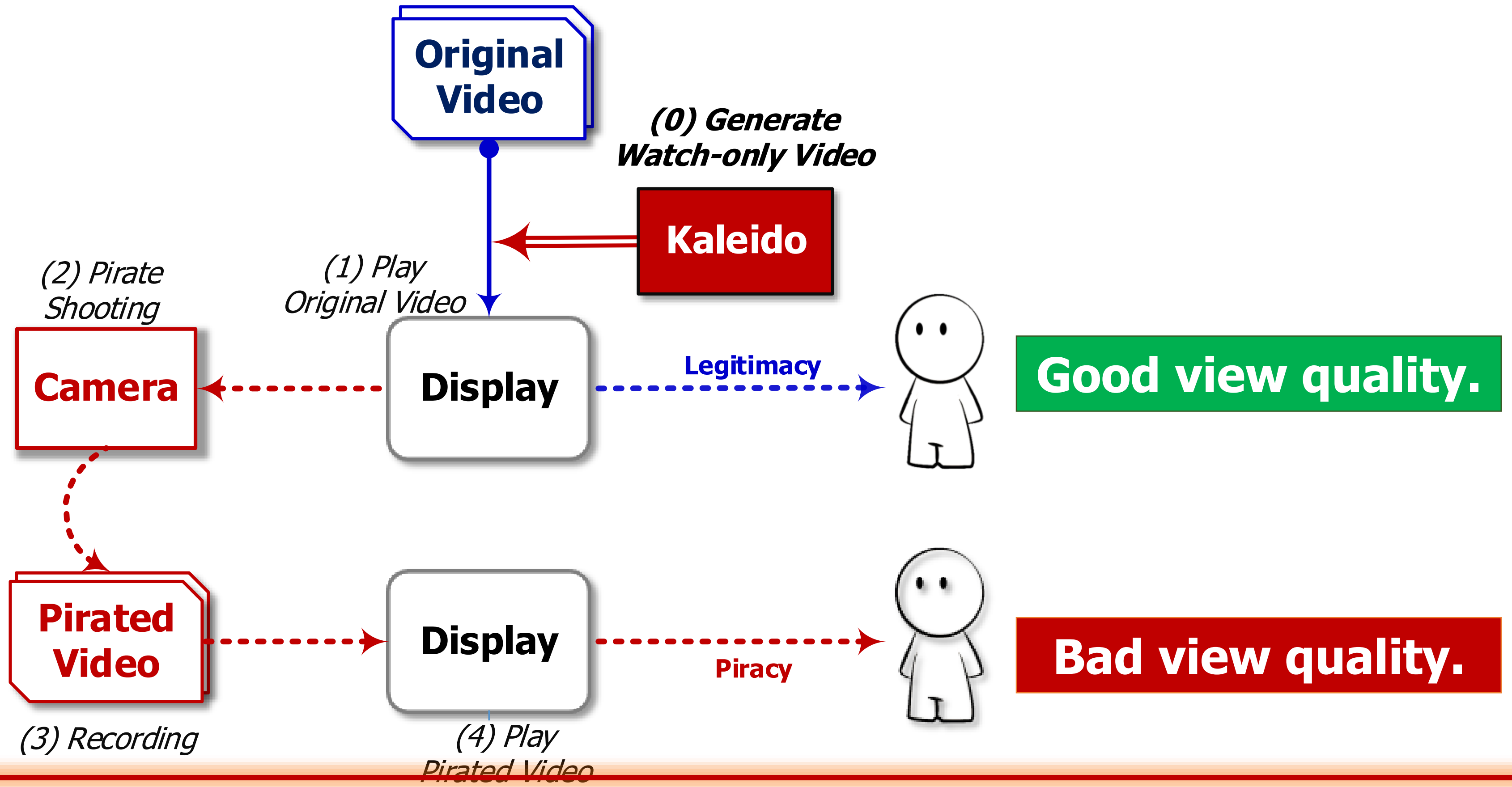
Kaleido: addon to generate a watch-only version of the video .

Challenge & Opportunity



Introduce illuminance flicker and chromatic distortion into the re-encoded frames.

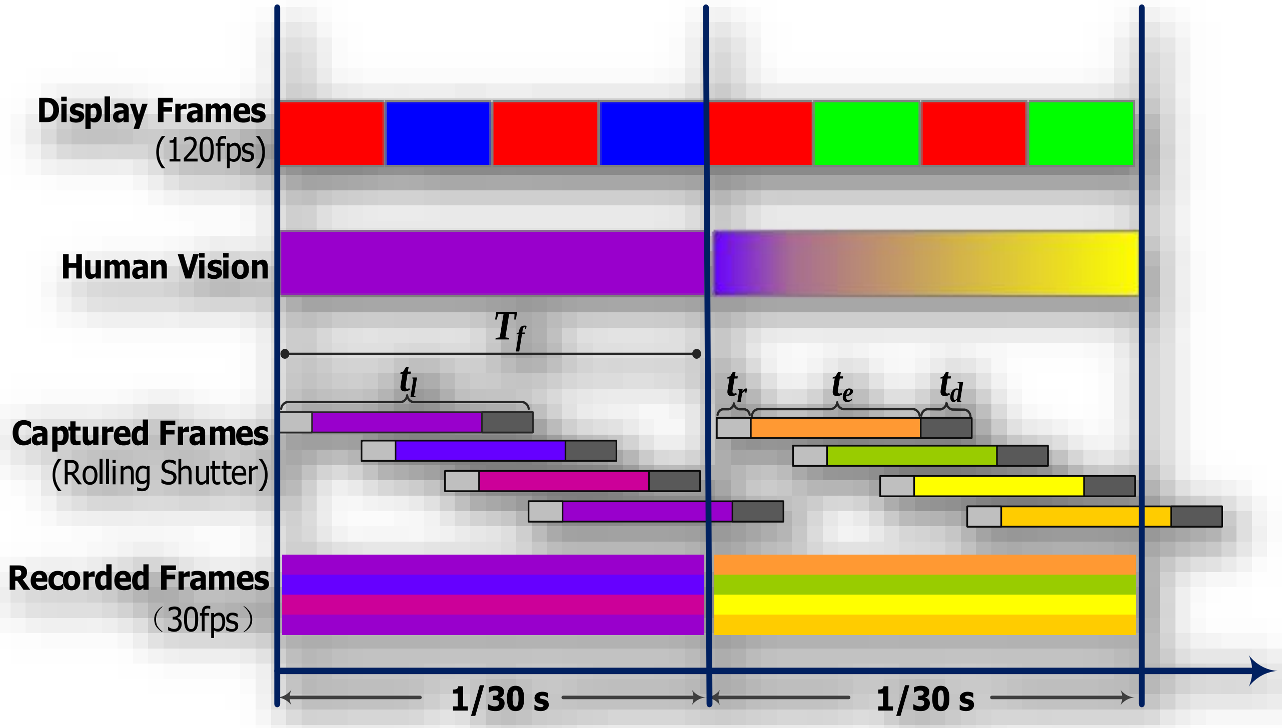
Challenge & Opportunity



How to ensure the encoded flicker and distortion are imperceivable to the legitimate viewers at first, and then become perceivable after a piracy procedure?

Case #1

Display rate is larger than record rate: $f_d > f_c$



Rolling shutter effect

$t_c \, (t_e < t_c/2)$

If signal is time-varying

Temporal variation loss

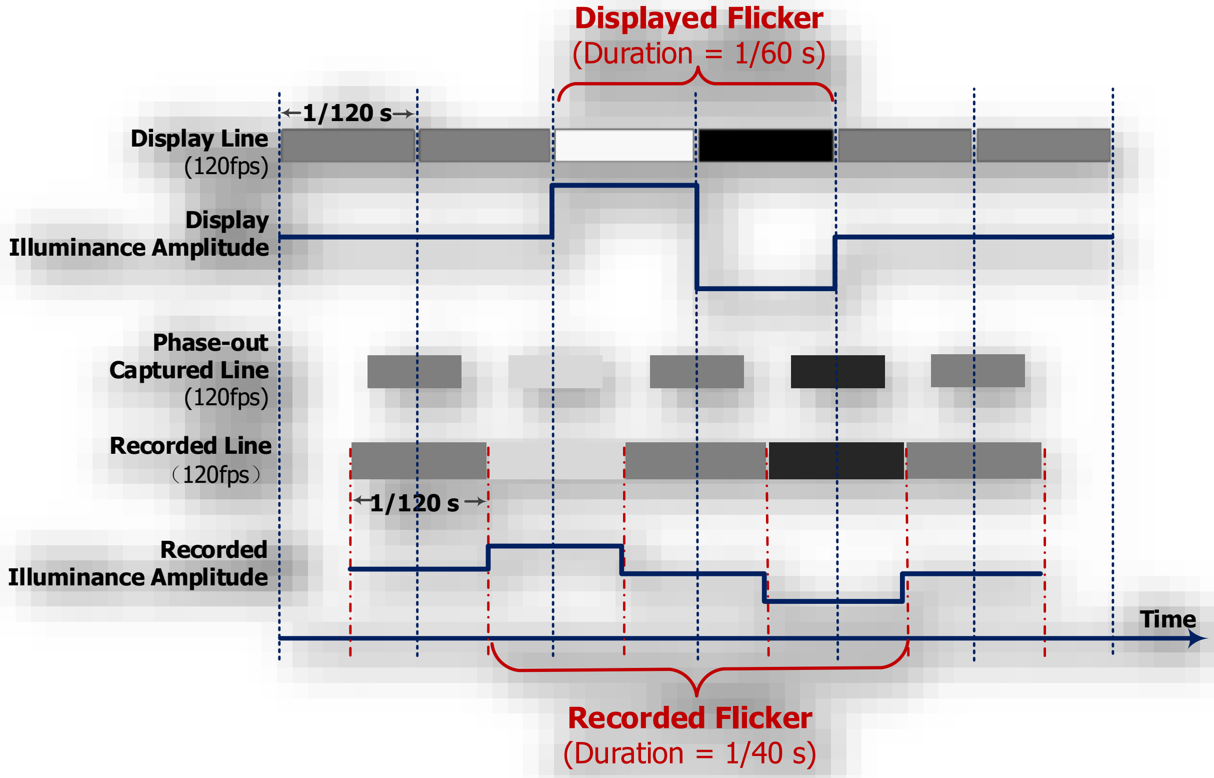
Eye perceives time-averaging signal

Perceivable distortion

Case #2

Display rate is less than record rate: $f_d \leq f_c$

Flicker frequency is down-converted from 60Hz to 40Hz.



Asynchronized frames

↓

Out-phase recorded lines

↓

Down-conversion flicker frequency

↓

f_f>CFF → f_f<CFF

↓

Perceivable flicker

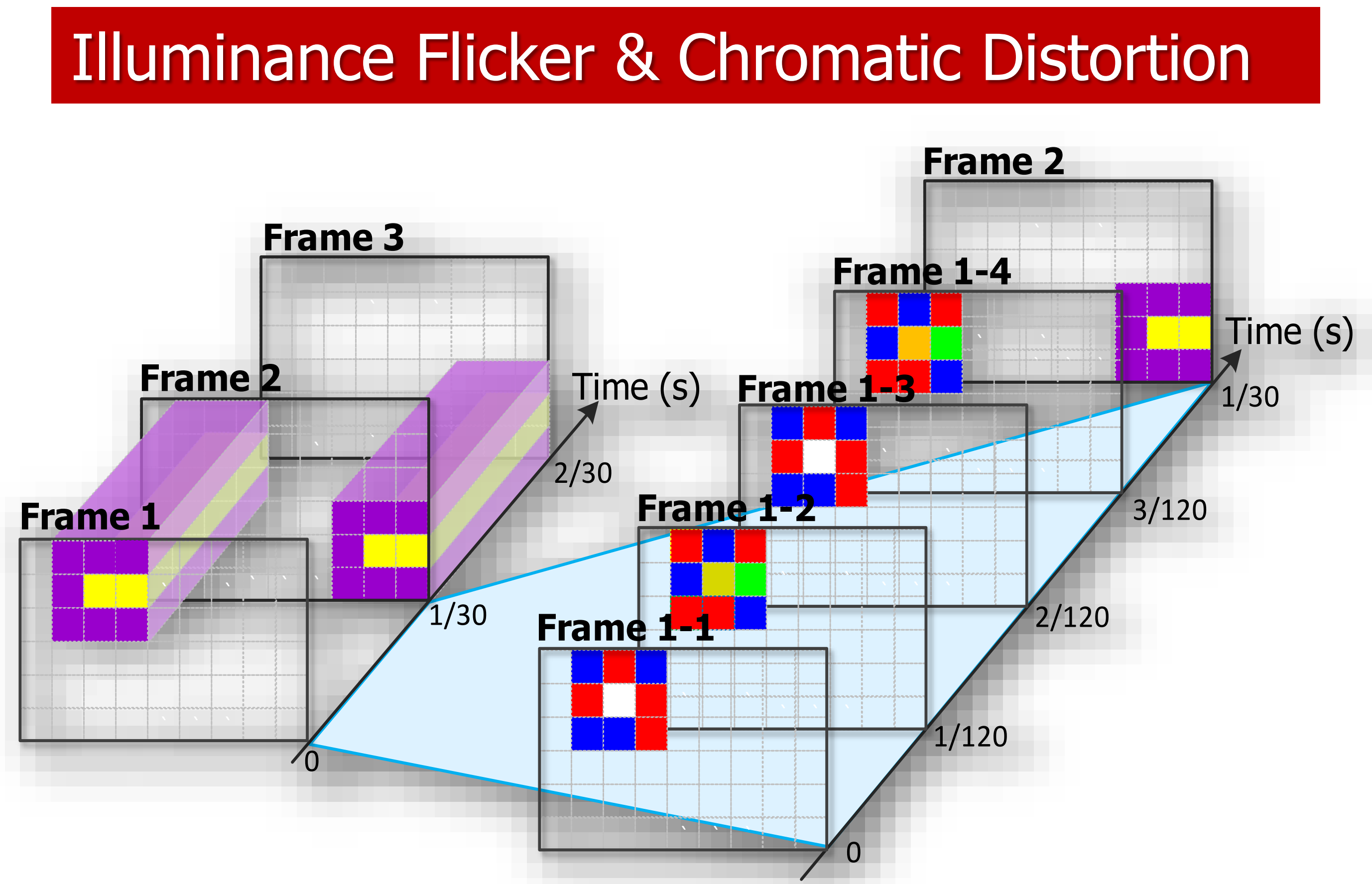
- A flicker at a frequency f_f will be captured by temporal successive out-phase lines.
- The flicker is recorded but its frequency $\frac{2}{2+f_d/f_c} f_f$
- Unstable inter-frame intervals of cameras aggravate the information loss and distortion.

Generate Watch-only Video

- Most current videos are 30fps
- High refresh rate display (e.g., 120Hz)
- One frame → 4 successive frames following the additive rule of human eyes.
- The flicker frequency is 60, which is larger than the CFF.
- Original Frame V^k

  Sub-frames $\{V^{k,1}, V^{k,2}, V^{k,3}, V^{k,4}\}$

  We need to determine $(x_{ij}^{k,l}, y_{ij}^{k,l}, Y_{ij}^{k,l})$ of each pixel $P_{ij}^{k,l}$



Tech #1: Illuminance Frame Pollution

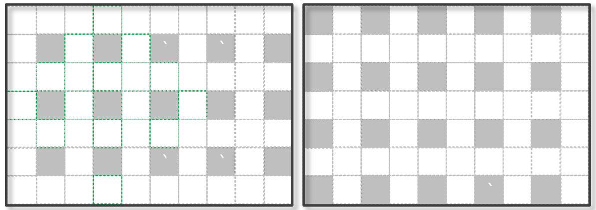
- A pair of pollution frames: add an illuminance complementary perturbation $(+\delta, -\delta)$ to each pixel pair.
- The time averaging illuminance of each pixel from two pollution frames equals 0.
- Flicker frequency is just above CFF and the amplitude and block size should be maximized to aggravate the pollution.

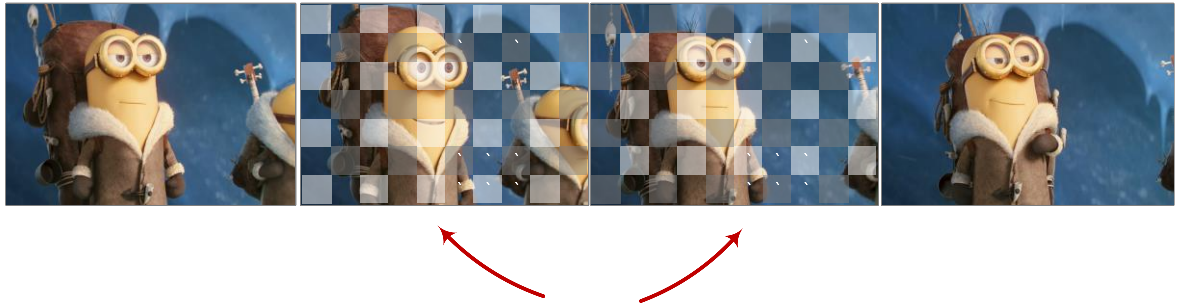
$$A(f) = a \cdot e^{bf} \rightarrow CFF = \frac{\ln[A(f)/a]}{b}$$

- If any down-conversion and instable interval happens, the flicker will become perceivable.

Successive Frames

Illuminance Flicker Pair

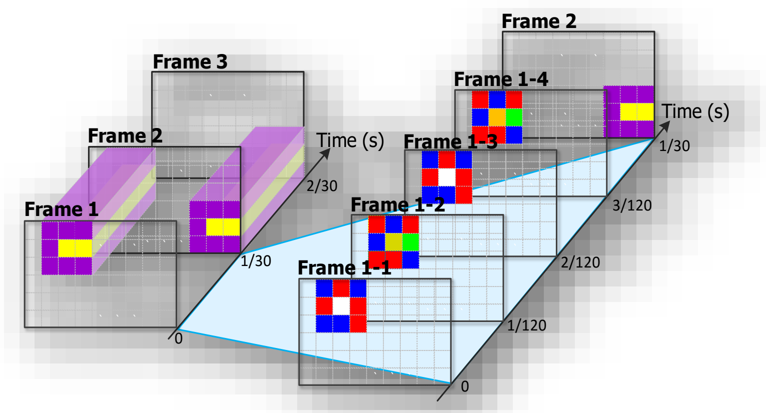




Polluted Frame Pair

Tech #2: Chromatic Frame Decomposition

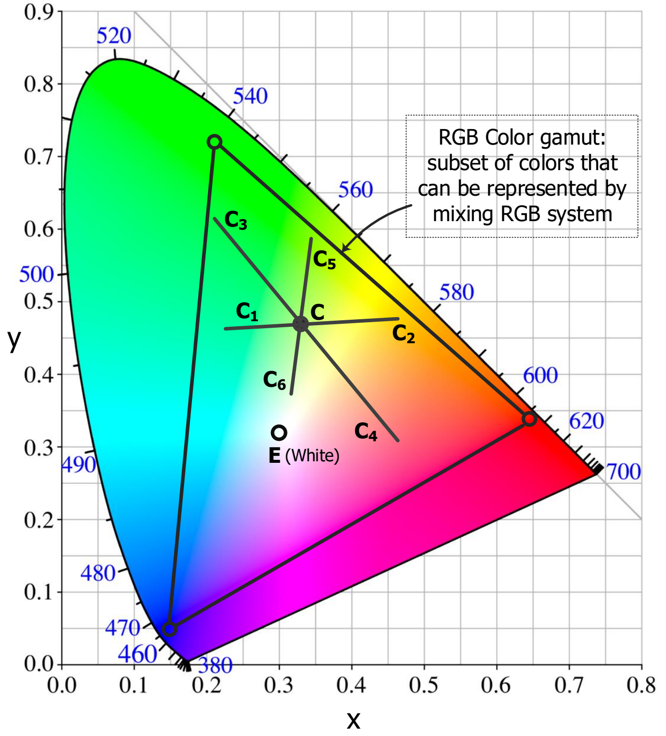
- Metamerism: a color can be decomposed to an infinite number of different color pairs.
- Choose a set of combinations that will (approximately) maximize the potential color distortion and spatial deformation.



$$D_c(C, C') = |\alpha - \beta| D_c(C_1, C_2)$$

Determined by the original video and the illuminance pollution.

Determined by camera parameter.



$$\max D_c(C_1, C_2) \quad \text{such that}$$
$$\begin{cases} \frac{D_c(C_1, C)}{D_c(C_2, C)} = \frac{Y_2}{Y_2} \\ \text{both } C_1 \text{ and } C_2 \text{ are within the RGB triangle.} \end{cases}$$

- We propose an algorithm to achieve the optimum with constant time complexity.

Tech #3: Embrace Spatial Deformation

- Make display colors appear as random as possible.
- Metamerism: a color can be decomposed to an infinite number of different color pairs.
- Randomizing different decomposition color pairs will make each display frame like a random noise.

Maximize Spatial Deformation

Tradeoff between maximization of color distortion and spatial deformation.

Radom

Pattern

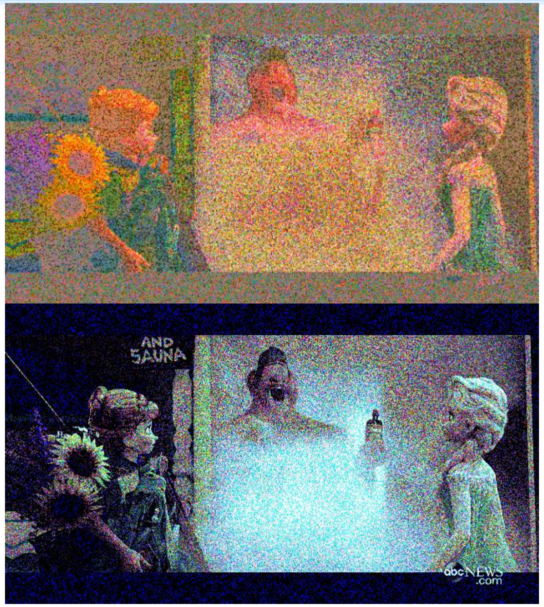
5	3	5	3	5	3
3	1	3	1	3	1
5	3	5	3	5	3
3	1	3	1	3	1

2	6	2	6	2	6
6	4	6	4	6	4
2	6	2	6	2	6
6	4	6	4	6	4

Mix

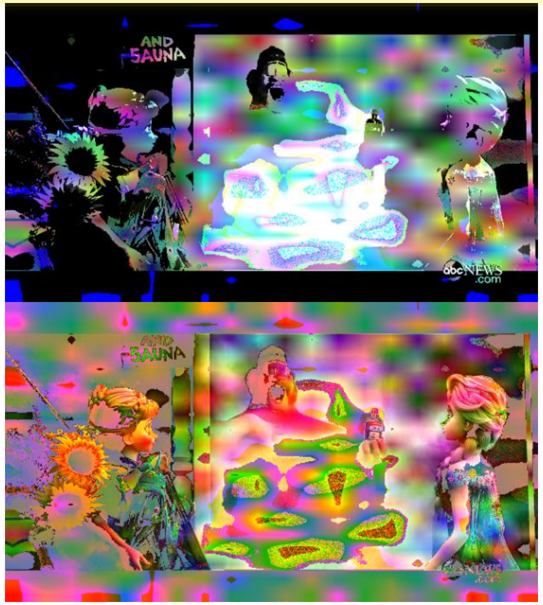
Maximize Spatial Deformation

Radom



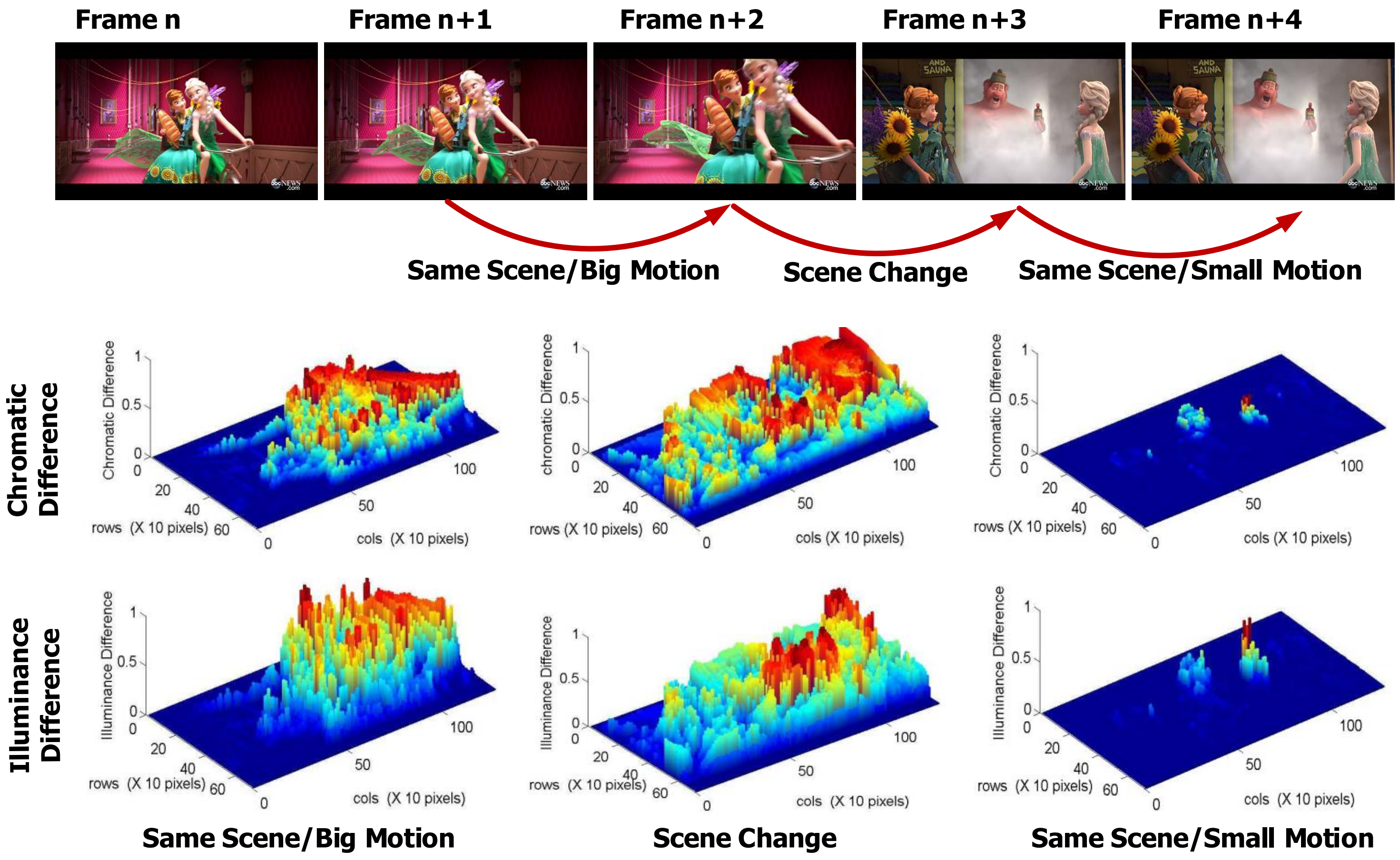
Pattern

Mix



Watch-Only Video Generation

Reducing the Encoding Cost



Watch-Only Video Generation

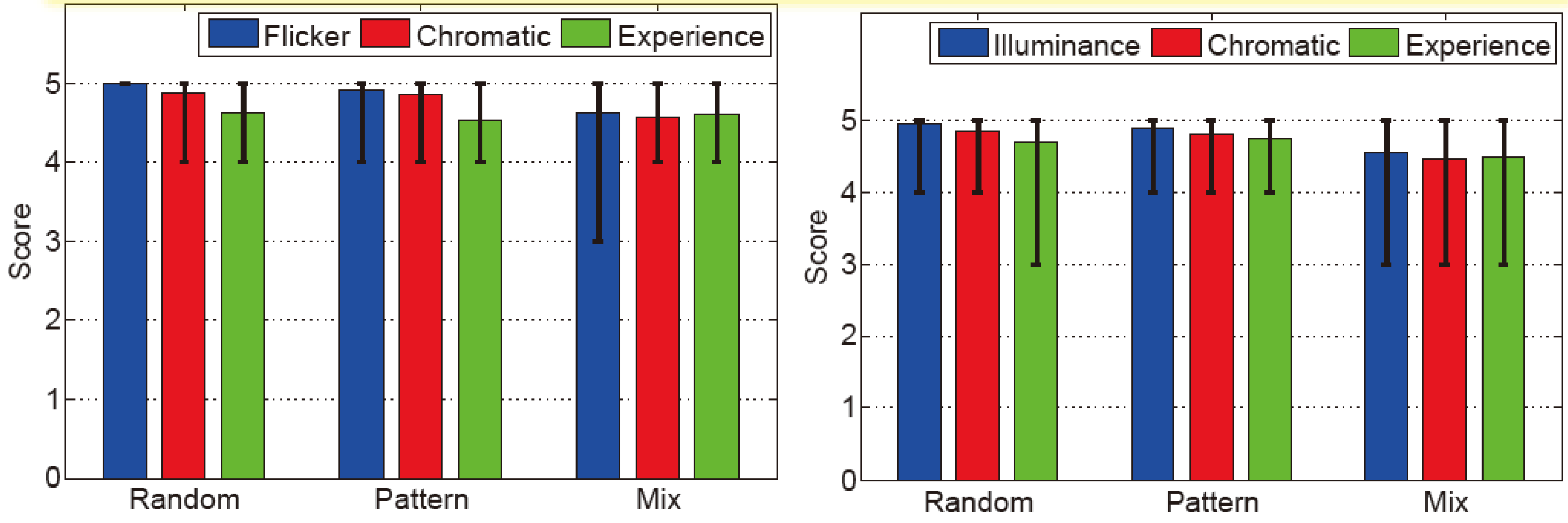
Evaluation

Experiment Setting

- LCD monitor and projector (120fps).
- 5 different smartphones (iPhone 5s, iPhone 6, Samsung Note3, Note4 Edge, HTC M8)
- Capturing rate is 1080p in 30fps or 60fps and 720p in 120fps
- 20 diverse high-definition (1280*720) video clips
- 50 volunteers

Watch-Only Video Quality Assessment

Average Subjective Score



(a) LCD Monitor (b) Projector

- **Random**: 96% volunteers did not even notice they are watch-only video clips.
- **Pattern**: 92% volunteers did not distinguish them.
- **Mix**: 38% volunteers noticed, but the degradation is acceptable (the average score is above 4).
- LCD monitors have a slightly better performance than projectors.
- Light condition and video type do not cause significant differences of watching experience.

Pirated Video Quality Assessment

- **Original**

- **Pirated**

Pirated Video Quality Assessment

Original Video

Pirate Video of Original Video
(camera fps=30)

Pirate Video of Watch-only Video
(LCD Screen, camera fps=30)

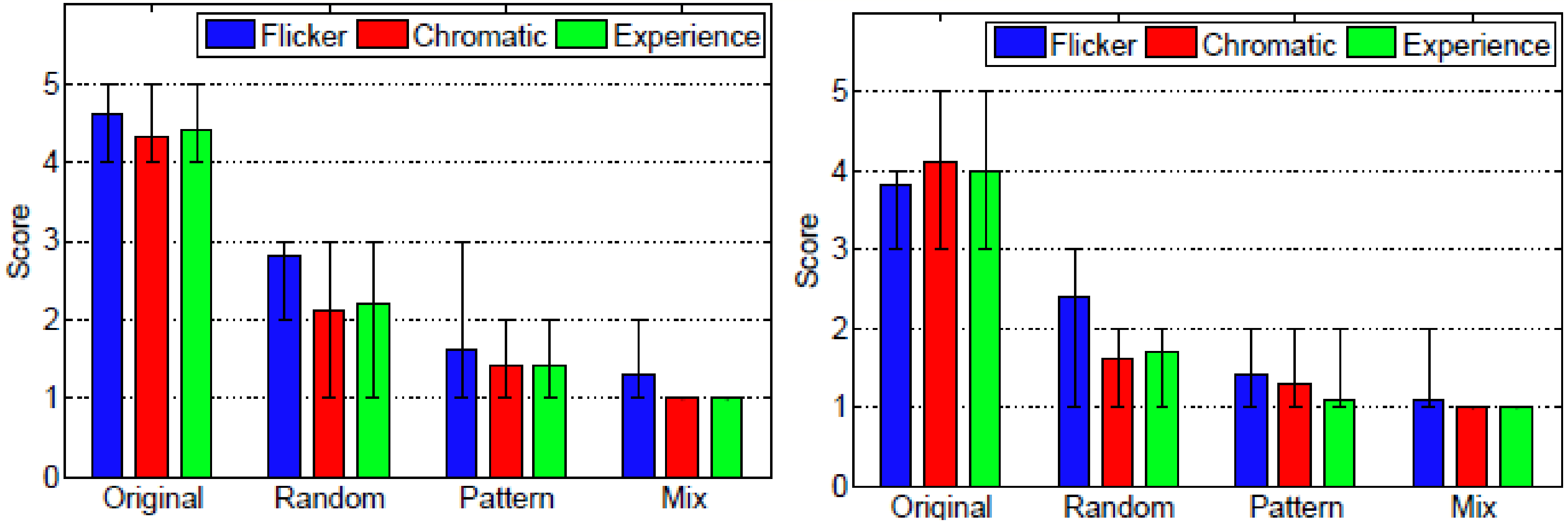
Pirate Video of Watch-only Video
(LCD Screen, camera fps=120)

Pirate Video of Watch-only Video
(Projector, camera fps=30)

Pirate Video of Watch-only Video
(Projector, camera fps=120)

Pirated Video Quality Assessment

Average Subjective Score

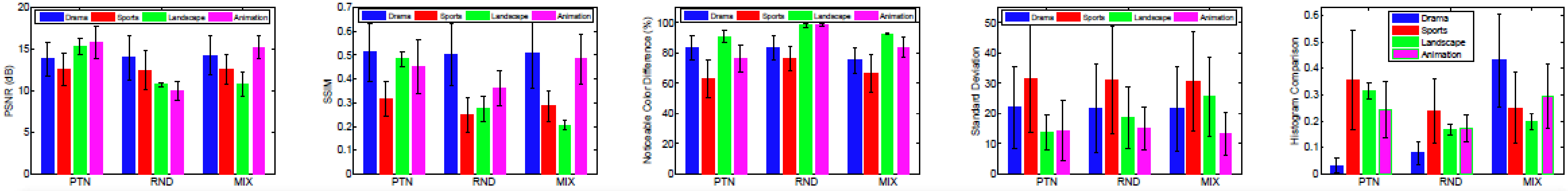


(a) LCD Monitor (b) Projector

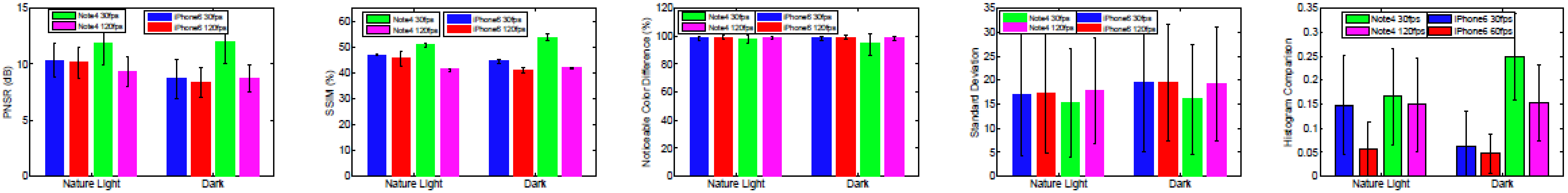
- The score for pirated original video is about **4**, which indicates acceptable quality.
- **96%** volunteers claim the quality degradation is intolerable, and the average rating score is below **2**.

Pirated Video Quality Assessment

Different decomposition methods



Different light conditions and capture rates

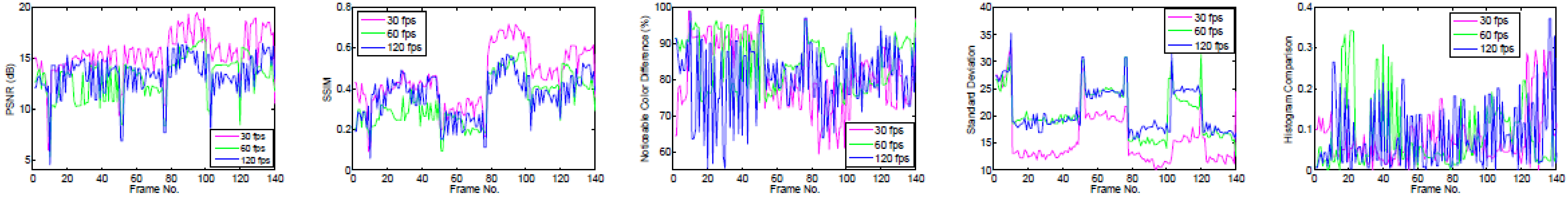


(a) PSNR (b) SSIM (c) CD Proportion (d) CD Std. Dev. (e) Histogram

- All our methods distort the color of original frames, which leads to significant quality degradation in all videos in all conditions.

Pirated Video Quality Assessment

Compare pirate video from original and watch-only



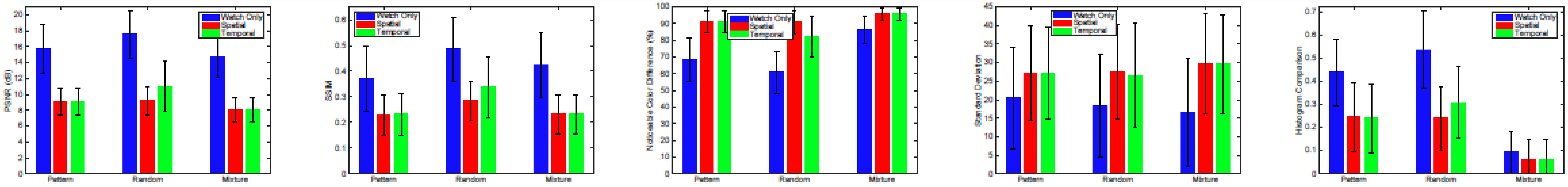
(a) PSNR (b) SSIM (c) CD Proportion (d) CD Std. Dev. (e) Histogram

- Pirate watch-only video also has severe quality degradation compared to the pirate original video of non-modified version.

Pirated Video Quality Assessment

We apply existing noise removal techniques to pirate watch-only video.

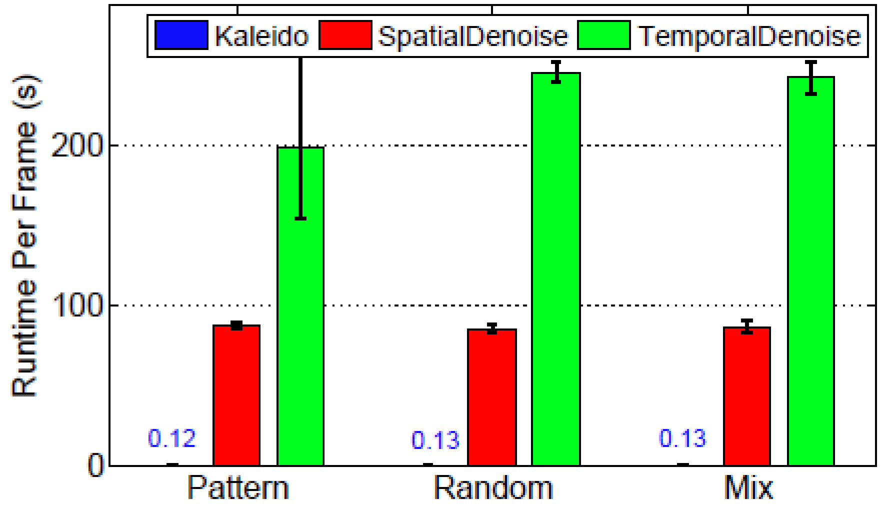
Video Quality after denoising



(a) PSNR (b) SSIM (c) CD Proportion (d) CD Std. Dev. (e) Histogram

- Common post denoising process not only cannot recover the original video, but also deteriorate the video quality compared to the watch-only version due to recognizing noise incorrectly.

Overhead



- It takes **0.12s** in average to process one 1280*720 video frame, i.e., the process speed is about **8.3fps**.

- Since our method increase the frame rate from 30fps to 120fps, the watch-only video quadruples the file size of the original one.

Thanks!

lan@greenorbs.com