Rainbow: Preventing Mobile-Camera-based Piracy in the Physical World

Lin Yang*, Wei Wang‡, Zeyu Wang*, Qian Zhang*
*Hong Kong University of Science and Technology,
‡School of Electronic Information and Communications, Huazhong University of Science and Technology.
Email: *{lyangab, zwangas, qianzh}@cse.ust.hk, ‡weiwangw@hust.edu.cn

***Abstract*— Since the mobile camera is often small in size and easy to conceal, existing anti-piracy solutions are inefficient in defeating mobile-camera-based piracy, thus it remains a serious threat to copyright of Intellectual property. This paper presents *Rainbow*, a low-cost lighting system which prevents mobile-camera-based piracy attacks on intellectual property in the physical world, *e.g.*, art paintings. Through embedding invisible illuminance flickers and chromatic changes into the light, our system can significantly degrade the imaging quality of a camera while maintaining a good visual experience for human eyes. Extensive objective evaluations under different scenarios demonstrate that *Rainbow* is robust to different confounding factors and can effectively defeat piracy attacks on various mobile devices. Subjective tests on volunteers further evidence that our system can not only significantly pollute pirated photos but also be able to provide good lighting conditions.**

I. Introduction

To protect the copyright on intellectual property, such as films and artwork, photo taking is often not allowed in many scenarios, *e.g.*, cinemas, museums, art galleries or exhibitions [1]. However, as modern mobile cameras are often small in size and easy to conceal, they are hard to detect, rendering mobile-camera-based piracy a serious threat to copyright protection. Existing no-photography policies are often implemented by security guards [2], which involve much human participation and cannot effectively defeat the mobile-camera-based piracy.

As a remedy, some researchers propose to defeat piracy by polluting the photos as much as possible. In this field, infrared light [3], [4] and watermarking [5], [6] are the most widely-adopted techniques in the film/photography community. However, infrared light is evidenced to be harmful to paintings and thus cannot be applied in many museums and galleries [7]. Also, watermarking is proved to be inefficient in preventing attackers from recording video clips for later redisplay [8]. Furthermore, some pioneer researchers use advanced display techniques [9] and video encoding schemes [8] to embed invisible noise into videos. Although these approaches have demonstrated to be effective, they require a modification to the video frames and thus can only work on digital content, but not the physical intellectual properties. In addition, several anti-piracy systems aim to localize the attacker by various tracking techniques, such as infrared scanning [10], distortion analysis [11], and audio watermarking tracking [12]. These solutions often rely on high-cost professional devices, which hinder their wide adoption.

In this paper, we aim to prevent mobile-camera-based piracy attacks on 2D physical intellectual property—such as paintings or photographs—in indoor scenarios, such as museums, art galleries, and exhibitions. To this end, we propose a low-cost anti-piracy system, *Rainbow*, which leverages existing light infrastructure to degrade the imaging quality of mobile cameras as much as possible while maintaining a good visual experience for human viewers. The key idea comes from a fact that modern mobile cameras mainly adopt a Complementary Metal-Oxide Semiconductor (CMOS) image sensors with a rolling shutter [13]. Due to hardware limitations, the rolling shutter mechanism introduces a small delay among the exposure of pixel rows. This implies that, if the light conditions vary temporally during the exposure, the variation will turn into spatial distortions due to the exposure delay in rows and eventually result in "band"-like distortions termed the *banding effect* on the image. In light of this idea, we modulate high-frequency illuminance flickers and chromatic changes into the light energy. As the light is reflected from the physical object and projected into the camera, these variations can cause a banding effect with obvious visual distortions. These distortions then serve as a "watermark" to significantly pollute the image, making it worthless to copy and thus the target's copyright can be protected. Meanwhile, as the human eye acts as a "global" shutter with low-bandpass characteristics, such variations cannot be perceived by the human viewers and a good visual experience can be maintained.

Fig. 1. Application of *Rainbow*: preventing mobile-camera-based piracy in museums. Our system can seriously pollute an pirated image while maintaining good visual quality for human viewers.

To realize this system, several challenges need to be addressed: First, it is not clear how to maximize the visual distortion caused by the banding effect. To find the answer, a theoretical model of the banding effect is defined and its confounding factors are well-investigated. Moreover, to defeat piracy attacks performed on diverse mobile cameras in various exposure settings, we need to ensure our system works under a wide range of exposure times. To this end, a collaborative exposure coverage algorithm is proposed to select a set of optimal light frequencies. By coordinating the selected light frequencies collaboratively, we can guarantee the pirated photos taken at various exposure times within a possible range can be obviously polluted. Extensive objective

evaluations under different scenarios indicate that our system is robust to various confounding factors and can effectively defeat piracy attacks performed on diverse mobile devices. Additionally, the subjective tests on 20 volunteers further evidence that our system is not only able to create severe quality degradation on a pirated photo but also provides an excellent visual experience for human viewers.

The contributions of this work lie in the following aspects:

- To the best of our knowledge, we are the first to explore the possibility of utilizing the banding effect to prevent mobile-camera-based piracy on physical targets. Our theoretical model and experimental tests have demonstrated the feasibility of creating significant illuminance fading and chromatic shift on the pirated photos with a banding effect.
- We have built Rainbow, which is an anti-piracy lighting system based on existing light infrastructure. To defeat piracy attacks performed on diverse mobile devices in various settings, we have designed a collaborative exposure coverage algorithm to cover a wide range of exposure times.
- Extensive evaluations show that our system can provide a good performance under different scenarios. Additionally, our subjective tests on 20 volunteers further evidence that our system is not only able to protect a target's copyright, but also provide a good lighting function.

The rest of the paper is organized as follows: Section II briefly reviews the preliminary knowledge and Section III presents the system design. The evaluation results are reported in Section V and practical issues are discussed in Section VI, followed by a literature review and conclusion in Sections VII and VIII, respectively.

II. Background

A. Understanding the Human Visual System

The generation of human vision involves two functioning units: the eye and the brain. While the complex cognition process is performed by the brain, it is the eye which functions as a biological equivalent of a camera to capture an image. When the light within our visible spectrum, *i.e.*, around 300 to 700 nm, passes through the pupil and projects into the retina, different types of photoreceptors in the retina are activated, generating the perception of colors [14].

While the human eye has the amazing ability to sense chromatic changes, it suffers severe limitations on its temporal resolution. Medical studies indicates that our eyes act as a low-frequency filter and can only perceive any change slower than a frequency threshold [15]. This phenomenon is related to the *persistence of vision* and the frequency threshold is termed the *Critical Flicker Frequency (CFF)*. Although many factors, *e.g.*, the illuminance level and stimulus size, can affect the CFF, a typical value is 60 Hz for the majority of people. This means that, if the flickering frequency of an intermittent light is larger than 60 Hz, it appears to be completely steady to the average human observer. Similarly, a quick chromatic change at a higher frequency than the CFF is perceived as the *color fusion* of all the individual colors. For example, a fast chromatic iteration over red, green, and blue leads to the perception of them being white color.

B. Characterizing the Mobile Camera

With the ability to precisely capture a scene, image sensors become the most commonly used sensors equipped on modern mobile devices. Two types of image sensors are used for consumer-level cameras: the *Charge Coupled Device (CCD)* and *Complementary Metal Oxide Semiconductor (CMOS)*. The major distinction is the way that the sensor reads the signal accumulated at a given pixel [13].

The CCD image sensor employs the *global shutter* mechanism, in which every pixel is exposed simultaneously and the signal of each pixel is serially transferred to a single Analog-to-Digital Converter (ADC). As a result, its frame rate is often limited by the ADC rate. To eliminate this bottleneck, the CMOS sensor, which is widely adopted on modern mobile cameras [16], utilizes an ADC for every column of pixels. Such a design can significantly reduce the number of pixels processed by a single ADC and enable a much shorter readout time. However, all the sensor pixels still need to be converted one row at a time. This results in a small time delay between each row's readout, making each row's exposure no longer simultaneous, which gives the name of this mechanism, *i.e.*, the *Rolling Shutter*.

Figure 2 gives an illustration of the rolling shutter mechanism. In this simplified example, the CMOS image sensor contains four rows. Each of them is exposed for the same amount of time, but due to the limitations of the single-line readout, a small delay, often in several nanoseconds, exists between two consecutive rows' exposures. Although this mechanism empowers the CMOS sensor with the ability to sense high-frequency temporal variation, it can also cause visual distortions on the resulting image.

In particular, if the light energy fluctuates during exposure, the temporal variation is reflected as a spatial variation on the image sensor due to the exposure delay among pixel rows, which leads to "band"-like spatial distortion termed the *banding effect* on the resulting image. A common cause of the banding effect is the lighting we use every day. Despite the differences in lighting technology, all commonly-used lights, including incandescent lights, compact fluorescent lights, as well as Light-Emitting Diodes (LEDs), exhibit different levels of illuminance flickers [17]. For instance, an incandescent lamp connected to AC power often creates an illuminance banding effect at 50 or 60 Hz.

III. System Design

According to the previous discussion, we know that the rolling shutter on a mobile camera introduces a small time delay between each pixel row's exposure, enabling it to sense high-frequency variations and causing the banding effect on the image. On the contrary, the human eye acts as a continuous "global" shutter with a low-frequency filter. It can only perceive changes slower than the CFF frequency, which is 60 Hz in the majority of humans.

Our system leverage the discrepancy between the mobile camera and the human eye to pollute pirated photos without affecting the human visual experience. In particular, we propose to embed a high-frequency illuminance flicker and chromatic change into the light. When the light is reflected by physical objects and projected into the camera, it can generate a banding effect on the image which includes obvious illuminance fading and chromatic shift. Such distortions can

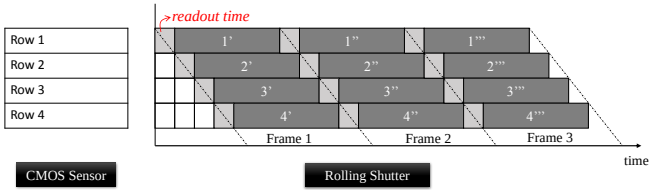


Fig. 2. Small delay exists between pixel rows due to the rolling shutter.

(a) Image w/o illuminance fading.

(b) Image w/i illuminance fading.

Fig. 3. An example of the illuminance banding effect.

then significantly degrade the quality of the resulting photo and serve as a "watermark" to protect the copyright of the targeted object. At the same time, as light modulation varies faster than the CFF frequency, human viewers cannot perceive any distortion and thus good visual quality can be maintained.

In this section, we first model the generation of the banding effect and explore the design space for embedding the illuminance fading and chromatic shift, then analyze the image pollution problem with the distortion hologram. To tackle the challenge of agnostic exposure time in real applications, we further propose a collaborative exposure coverage algorithm to cover a wide range of possible exposure times.

A. Embedding the Distortion with Banding Effect

1) Illuminance Fading: Consider a light with temporal illuminance variation as:

$$L(t) = A\sin^2(2\pi ft) \tag{1}$$

where A is the luminance intensity, $2f$ is the variation frequency, and $L(t)$ defines the illuminance variation function of light. In this case, the light energy E captured by each pixel row can be defined as:

$$\begin{aligned} E &= \int_{t_0}^{t_0+t_e} A\sin^2(2\pi ft)dt \\ &= \frac{A}{4\pi f}[\underbrace{2\pi f t_e}_{\text{DC component}} - \underbrace{\sin 2\pi f t_e}_{\text{Flicker ratio}}\underbrace{\cos 2\pi f(2t_0+t_e)}_{\text{Flicker component}}] \end{aligned} \tag{2}$$

where t_0 denotes the exposure starting time, and t_e is the exposure time of each row.

Several observations can be made from this equation:

1) The light energy captured by each pixel row comprises three parts: The *DC component* defines the base light energy received during the exposure. It is determined by the exposure time t_e and does not change among rows. Meanwhile, the illuminance fading is jointly produced by the *flicker ratio* and the *flicker component*.
2) Given an exposure time t_e, as the rolling shutter causes a small delay between the exposure starting times t_0 of different rows, the *flicker component* varies among rows and eventually leads to a "band-like" illuminance fading on the image.
3) The degree of illuminance fading is further controlled by the *flicker ratio*, which depends on the relationship between the light frequency f and the exposure time t_e. For example, if the exposure time is a multiple of the light period, *i.e.*, $t_e = n/2f$, the *flicker ratio* becomes zero and the illuminance fading vanishes, while its effect is maximized when the *flicker ratio* is equal to 1, *i.e.*, $t_e = (2n+1)/4f$.

In addition, we notice that, to address the illuminance banding effect caused by the light lamps, modern mobile cameras often enforce the exposure time t_e to be a multiple of either $1/50$ or $1/60$ seconds by time padding [18]. This can effectively alleviate the illuminance banding caused by AC power. However, such an anti-banding technique fails if the light frequency changes. Figure 3 shows photos taken under two identical scenarios, except one is lit by an LED light flickering at 60 Hz, while the other adopts a modified LED of 73 Hz. We can see the camera's anti-banding fails and an obvious illuminance fading occurs on the photo taken under the 73-Hz LED light.

2) Chromatic Distortion: To embed the chromatic distortion with the banding effect, we use an RGB LED light which can emit light of three primary colors—red, green, and blue—simultaneously. Consider the case in which the light switches among these three primary colors at a frequency f and the camera's exposure time is t_e, their relationship can be described as:

$$t_e = \frac{n}{f} + r + g + b, \text{ where } \begin{cases} n & = \lfloor t_e/\frac{1}{f} \rfloor \\ (r+g+b) & = t_e \ mod \ \frac{1}{f} \end{cases} \tag{3}$$

where n is the number of light periods contained in the camera's exposure t_e, while r, g, and b represent the residual durations of red, green, and blue colors in the remainder of $t_e/\frac{1}{f}$, respectively.

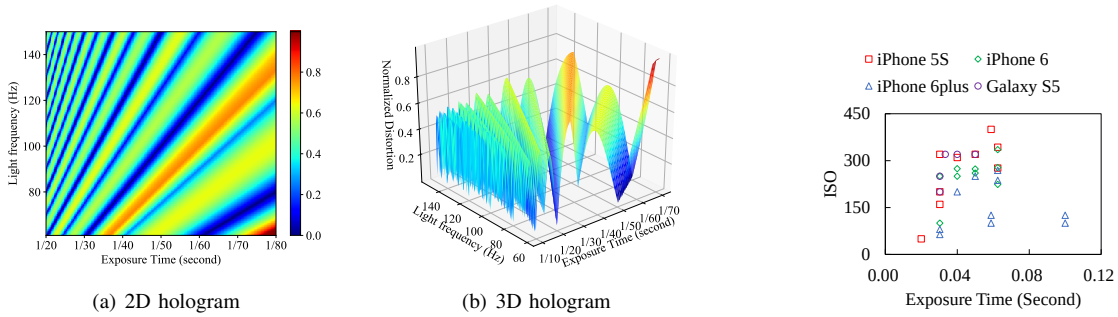
Recall that the low-frequency characteristics of human eyes make a chromatic change faster than the CFF frequency perceived as a *color fusion* of the individual colors. As a result, through carefully tunning the flickering frequency and proportion of the three colors, we can ensure that human viewer can not perceive any chromatic variation and the emitted light meets various illuminance requirements, *e.g.*, warm white around 2700-3000 kelvins used in many indoor scenarios [17].

However, unlike the human eye which acts as a continuous "global" shutter, the camera is exposed in a discrete way. Therefore, if the exposure time t_e is not a multiple of the light changing period $1/f$, some residual colors—r, g, and b—are left in the remainder of each row's exposure. Since the fusion result of these residual colors cannot be guaranteed to be white, they can introduce an obvious chromatic shift on each pixel row. Moreover, the rolling shutter mechanism further aggravates this problem by rendering the resulting color of each row distinct, which eventually causes a visual "color-band"-like chromatic distortion on the image.

Apparently, the degree of the chromatic distortion depends on the ratio of the residual color to the white color:

$$\begin{aligned} &\text{residual ratio} = \tfrac{\text{residual color}}{\text{white color}} = \tfrac{\max(r,g,b)-\min(r,g,b)}{n/f+3*\min(r,g,b)}, \\ &\text{where } \begin{cases} n = \lfloor t_e/\frac{1}{f} \rfloor \\ (r+g+b) = t_e \bmod (1/f) \end{cases} \end{aligned} \tag{4}$$

Note that all the variables in this function are jointly determined by the camera's exposure time t_e and the light frequency f. Similar to the case of illuminance fading, once



(a) 2D hologram (b) 3D hologram

Fig. 4. Hologram to exhibit the interaction between the light frequency and exposure time.

Fig. 5. Variations in the exposure settings.

the exposure time is a multiple of the light period, the residual color becomes zero and no chromatic distortion is induced. This implies that, to maximize the chromatic distortion, we need to carefully manipulate the light frequency according to the exposure time.

B. Polluting the Image

To magnify the image quality degradation, we would like to combine both the illuminance fading and chromatic shift. According to previous analysis, we know that the degree of illuminance fading is determined by the *flicker ratio*, while the chromatic distortion is controlled by the *residual ratio*. Both variables strongly depend on the interaction of camera's exposure time t_e and light frequency f. Therefore, we define the overall distortion function $Dist(\cdot)$ as follows:

$$Dist(f, t_e) = \alpha_1 \sin 2\pi f t_e + \alpha_2 \frac{max(r,g,b) - min(r,g,b)}{n/f + 3 * min(r,g,b)},$$
$$\text{where} \begin{cases} n = \lfloor t_e / \frac{1}{f} \rfloor \\ (r+g+b) = t_e \bmod \frac{1}{f} \end{cases}$$

where α_1 and α_2 are the weights of the illuminance fading and the chromatic shift, which are 0.5 by default in our system.

Obviously, this distortion function is not jointly convex. To study its characteristics, we first partition the parameter space into a finite grid $M \times N$. Then, we employ a *distortion hologram* to explore the interaction among the image distortion d, light frequency f and exposure time t_e. The distortion hologram is a distortion exhibition using an image to display the level of image pollution that be generated by the frequency-exposure combination in a partitioned grid. Given a $(f, t_e)_{M \times N}$ partition, a distortion hologram D is defined as:

$$D = \begin{pmatrix} d_{11} & d_{12} & \dots & d_{1N} \\ d_{21} & d_{22} & \dots & d_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ d_{M1} & d_{M2} & \dots & d_{MN} \end{pmatrix} \quad (5)$$

where d_{ij} represents the distortion generated at a given frequency-exposure combination, *i.e.*, $d_{ij} = Dist(f_i, t_{e_j})$, M and N denote the number of possible light frequencies and exposures, respectively.

Figure 4 gives an example of distortion hologram, in which the exposure time ranges from 1/80 to 1/20 seconds and the light frequency is from 60 to 140 Hz. We term that an exposure time is "covered" by a light frequency if the corresponding distortion value is larger than a predefined threshold ϵ ($\epsilon = 0.4$ by default). According to this figure, we find that:

- A single light frequency cannot cover all possible exposure times.
- However, a light frequency can cover multiple exposure times, and an exposure time can also be covered by several light frequencies with different distortion levels.

In theory, if the exposure time of the attacker's camera is known, we can easily find an optimal light frequency according to the hologram. In practice, however, this does not work as the exposure time of the attacker's camera cannot be known. In the next subsection, we explain the reason and discuss the solution for this issue.

C. Variation of Exposure Time

The design of a modern mobile camera generally follows the Additive System for the Photographic Exposure model (APEX) [18], which defines the relationship between the exposure time and its confounding factors:

$$\frac{F^2}{t_e} = \frac{B \cdot S}{k}, \quad (6)$$

where F is f-number of the camera lens, t_e represents the exposure time, B denotes the brightness, S and k are the gain and scaling factors of an image sensor, respectively. In this model, the exposure value EV can be defined on the logarithmic space of APEX:

$$EV = 2 \log F - \log t_e = \log B + \log S - \log k. \quad (7)$$

Given a requirement of the brightness level, the exposure time can be determined by an on-chip Auto-Exposure (AE) control algorithm. However, as the lighting conditions in the target scenes can be quite sophisticated, many advanced techniques are proposed in the AE to gain more accurate exposure control and most mobile device manufacturers run their own AE control algorithms on their cameras [18]. As a result, the exposure time determined on various devices can be distinct. Besides, in real applications, the attacker can perform piracy attacks from different distances and angles, in which the exposure time changes with the variation of illuminance level. Moreover, some camera applications even allow the users to set the exposure time manually, which further aggregates this problem.

To further understand this problem, we use the default camera applications on various mobile devices to determine the exposure time for a same scene. The results are reported in Figure 5, from which we find that the exposure settings vary with devices. Even on the same device, the exposure settings determined from various distances and angles can be significantly different. These results imply that an accurate estimation of the exposure time on an attacker's camera can be very hard, if not impossible.

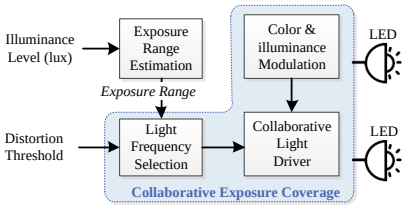


Fig. 6. Rainbow System Architecture.

D. Collaborative Exposure Coverage

While the heterogeneity of the camera's exposure control hinders the accurate estimation of the exposure time, another fact sheds light on the possible solution: due to the constraints in the image sensor, *e.g.*, size and cost, modern mobile cameras are often limited in their hardware variety, *i.e.*, lens aperture, gain and scaling factors [18]. This means that, given a scene of an illuminance level, it is possible to roughly estimate the possible range of the exposure times [19]. As a result, instead of targeting at an agnostic exposure time, we aim to cover all the exposure times within the possible range.

To this end, we propose to collaborate multiple light frequencies to cover different exposure times within the possible range. This approach is applicable as the indoor deployment of light lamps are generally dense and there is often more than one light inside a room. However, considering the deployment and maintenance cost, the number of lights used should be minimized.

In light of this idea, we now formulate the exposure coverage problem as follows: First, we define a step function $u(\cdot)$ on the distortion hologram D:

$$u(d_{ij}) = \begin{cases} 1, & d_{ij} \geq \epsilon \\ 0, & d_{ij} < \epsilon \end{cases} \tag{8}$$

in which if the distortion value $d_{ij} = Dist(f_i, t_{e_j})$ is larger than a threshold ϵ, the function outputs 1 and we say the corresponding exposure time t_{e_j} is covered by light frequency f_i. By applying such a step function to the distortion hologram, we can compute the covered exposure times of each light frequency. Let S_i be the set of all the exposure times covered by light frequency f_i.

Then, we define the cost function of set S_i to be:

$$C(S_i) = \sum_{t_{e_j} \in S_i} \big(1 - Dist(f_i, t_{e_j})\big) \tag{9}$$

where t_{e_j} is the exposure time covered by light frequency f_i.

Given the universe set U of all the exposure times within the possible range, and a collection $\Psi = \{S_1, S_2, \ldots, S_n\},\ S_i \subseteq U$, for each light frequency f_i and its corresponding set S_i, we associate a variable x_{S_i} that indicates whether S_i is chose.

In this way, the problem of polluting an image under a wide range of exposure times with limited lights becomes finding a sub-collection $S\prime \subseteq \Psi$ that covers all exposure times in U at minimum cost:

$$\begin{aligned} \min \quad & \mathit{Val}(x) = \sum_{S_i \in \Psi} C(S_i) x_{S_i} \\ s.t. \quad & \sum_{S_i : t_e \in S_i} X_{S_i} \geq 1 \qquad t_e \in U, \\ & x_{S_i} \in \{0,1\} \qquad S_i \in \Psi \end{aligned} \tag{10}$$

in which we can have solutions as a vector $\mathbf{x} \in \{0,1\}^n$.

Theoretically, this is *de facto* an NP-hard *SET COVER* problem [20]. To solve this problem, we propose a light frequency selection algorithm based on the primal-dual schema [21] as shown in Algorithm 1. This algorithm iteratively changes a primal and dual solution until the relaxed primal-dual complementary slackness conditions are satisfied. Define the *frequency* of an exposure time to be the number of sets it is contained in. Let k denote the frequency of the most frequent exposure time. It can be theoretically proved that this prime-dual-based algorithm can achieve a $k-$approximation for our problem [21].

Algorithm 1 Exposure Coverage Algorithm.

Input:
Exposure universe U with n possible values,
Collection $\Psi = \{S_1, S_2, \ldots, S_n\},\ S_i \subseteq U$,
Distortion hologram $D = (d_{ij}) \in \mathbb{R}^{M \times N}$.
Output:
Frequency selection vector $\mathbf{x} \in \{0,1\}^n$
Apply step function $u(\cdot)$ to the distortion hologram D.
Compute exposure coverage set for each light frequency.
Define the primal problem and its corresponding dual.
$x \leftarrow 0, y \leftarrow 0$, Declare all the exposure times uncovered.
while some exposure times are uncovered **do**
  Pick an uncovered exposure time, t_{e_j}, raise $y_{t_{e_j}}$ until some set goes tight.
  Pick all tight sets S_i in the cover, *i.e.*, set $\mathbf{x}_{S_i} = 1$.
  Declare all the exposure time in these sets as covered.
end while
return $\mathbf{x}$

In a real application, we can first measure the illuminance level of the target scene with a light meter and roughly estimate the possible range of the exposure time[1]. To ensure substantial image pollution under all possible exposure times, multiple light frequencies can be selected appropriately by the exposure coverage algorithm. For example, according to our experiment in Section V, two frequencies, *e.g.*, 73 Hz and 83 Hz, are sufficient to cover a wide range of exposure times in a room with an illuminance level of 400 lux.

IV. System Implementation

To realize our design, we build an anti-piracy lighting system,*Rainbow*, as shown in Figure 6. It comprises four components: 1) the *Exposure Range Estimation* calculates a coarse range of possible exposure times with the help of a light meter. 2) The *Light Frequency Selection* module finds a set of optimal light frequencies by solving the exposure coverage problem to ensure good performances under all possible exposure times. The selected frequencies are then used to configure the 3)*Collaborative Light Driver*, which synchronizes and collaborates multiple lights to embed noise with banding effect, while 4) the *Color & Illuminance Modulation* unit defines the illuminance and color modulation patterns.

Figure 7 shows a prototype of *Rainbow*, in which several 10-Watts RGB LED bulbs connected to a DC power are controlled by a light driver box, on which we implement our system in C. To ensure the light beams can conveniently concentrate on a specific target, the LEDs are designed in forms of spotlights.

V. Evaluation

To comprehensively evaluate the performance of our system, we set up an experiment environment as shown in Figure 8. In

[1]As a common practice in photography, the estimation of exposure range is ignored here due to page limitations. More details can be found in [18], [19].

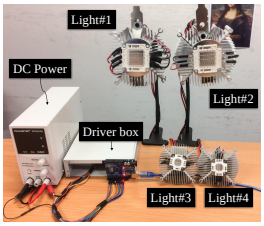


Fig. 7. Prototype system.

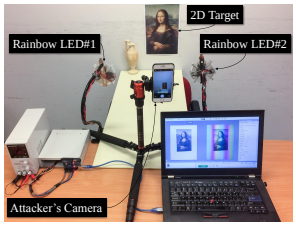


Fig. 8. Experiment setup.

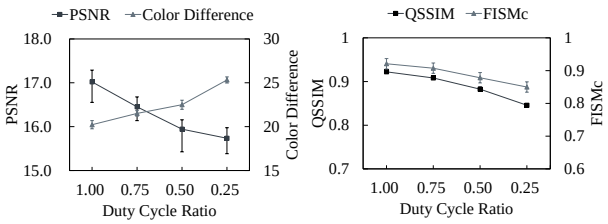


Fig. 9. System Performance under different duty cycles.

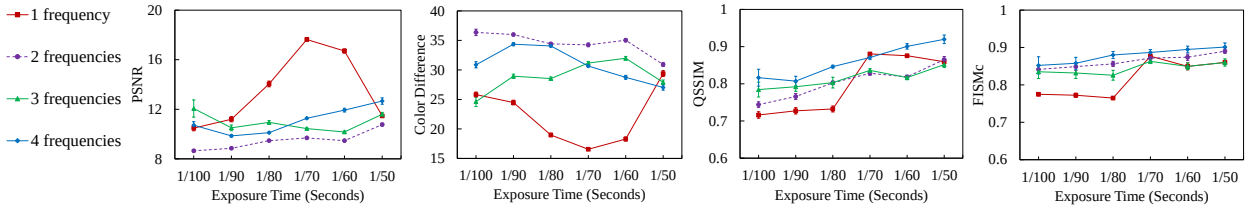


Fig. 10. System Performance with different numbers of lights. The dual-light setup outperforms the others.

a room of $4.3m \times 8.2m$, a *Rainbow* system with multiple lights is placed 0.5 meters away from the target and the light beams are carefully tuned to ensure a good coverage of the scene. Several mobile devices, including 4 Apple devices (iPhone 5S, iPhone 6, iPhone 6S, and iPhone 6S Plus) and 3 Android phones (Samsung Galaxy S5, Xiaomi Redmi 4, and Huawei Honor 4), are employed throughout the evaluation. Also, a tripod is used to avoid unnecessary image quality degradation brought about by hand shaking.

In each experiment, we first take a photo of the target scene under a unmodified light. The resulting image is used as the *reference image*. After that, several *pirated images* are taken at the same scene under our Rainbow system. By comparing the piracy images to the reference image, we can objectively measure the image quality degradation caused by our system. Apart from the objective evaluations, 20 volunteers, including 6 females and 14 males with good eye sights and chromatic viewing capabilities, are recruited for a subjective test. By querying volunteers' opinions about their visual experience and the quality difference between the reference and piracy images, we can subjectively quantify users' experience of our system.

Throughout the experiments, 5 quality metrics are adopted. 1) The **Peak Signal-to-Noise Ratio (PSNR)** evaluates the ratio of maximum signal power to the noise power at a pixel level. A PSNR value lower than 18 dB often implies significant quality degradation [22]. 2) The **Color Difference (CD)** computes the chromatic differences between the reference and piracy images according to the CIEDE2000 Delta-E formula [23]. A CD value larger than 6 indicates an obvious chromatic distortion occurs in the pirated image [8]. 3) The **Quaternion Structural Similarity Index (QSSIM)** leverages the quaternion image processing to quantify the structural similarity of two images in color space. Its value is normalized and decreases linearly with viewers' subjective experience [24]. 4) The **Feature-Similarity Index color (FISMc)** measures the local structure and contrast information to provide an excellent quantification of the visual experience [25]. An FISM lower than 0.85 means the viewers tend to give opinion scores less than 4 out of 10 to the polluted image, suggesting a significant visual distortion. 5) The **Mean Opinion Score (MOS)** reflects the viewers' subjective opinion upon their visual experience. Similar to previous work [8], we design a grading standard from 1 to 5, in which an MOS of 1 indicates the worst viewing perception with significant distortion/artifact, while a value of 5 represents an excellent visual experience.

A. *Effect of Parameters*

In this subsection, we evaluate some parameters which deeply affect the performance of our system, including the *light duty cycle* and the *multiple light frequencies* adopted.

1) Duty Cycle: The *duty cycle* determine the duration of lights-off state during the light flickering. To understand the effect of this parameter, we configure our system with different duty cycle ratios. Figure 9 shows the corresponding system performance in various duty cycle settings. We can observe the system performance increases with the decrease of duty cycle ratio. This is because a low duty cycle implies less light energy emitted within a light period, which results in a more obvious illuminance fading on the image. Nevertheless, a low duty cycle also reduces the overall luminance level and may cause an energy-efficiency problem. As a trade-off between system performance and energy efficiency, we set the duty cycle of Rainbow to 0.75.

2) Multiple Light Frequencies: To cover all the possible exposure times, multiple lights frequencies are selected according to the exposure coverage algorithm. This experiment examines the effectiveness of these selected frequencies.

Given the illuminance level in our evaluation setup (400 lux in this experiment), the possible range of exposure time is estimated to be from 1/100 to 1/50 seconds. The candidate light frequencies are chosen from 65 Hz to 155 Hz (with 1-Hz interval) and we empirically set the distortion threshold ϵ to 0.4. In this setting, the exposure coverage algorithm suggests a dual-frequency combination—73 Hz and 83 Hz—are sufficient to cover all the possible exposure times. For the comparison, we employ three other baselines. The *1-frequency* setup only uses a single light frequencies of 73 Hz, while the *3-frequencies* scheme adopts a combination of {67 Hz, 73 Hz, 83 Hz} and the *4-frequencies* setup employs a configuration of {67Hz, 73Hz, 83Hz, 89Hz}. By measuring the image pollution under all possible exposure times, we compare the quality degradation brought about by different frequency combinations in Figure 10.

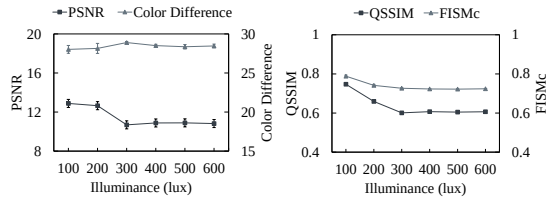


Fig. 11. System Performance with different illuminance levels. The results indicate our system perform well under various illuminance level requirements.

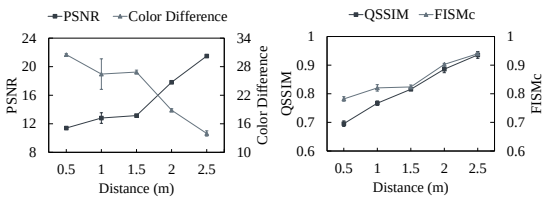


Fig. 12. System performance at different distances. Current effective distance is 2 meters, which can be further extended with higher-power light.

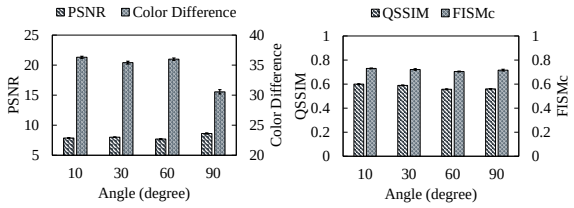


Fig. 13. System performance under various shooting angles. Note that the system performance are good and relatively stable, suggesting that our system can defeat piracy from different shooting angles.

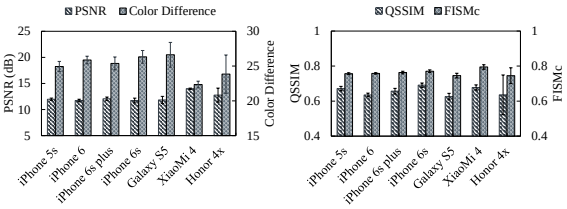


Fig. 14. System performances on different devices. Despite a slight performance variation due to the heterogeneity of cameras, the results show that the photos taken on all these devices are seriously polluted.

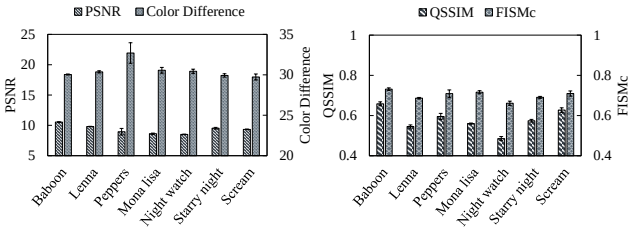


Fig. 15. System performance on various targets.

First, we can see that a single frequency is insufficient to cover all the exposure times. We can see the system performance experiences an obvious drop when the camera's exposure time approximates $1/73$ seconds. This is because both the flicker ratio and residual ratio are determined by $t_e \bmod 1/f$). Once the exposure time approximates a multiple of the light period, the banding effect declines dramatically, resulting in a significant performance degradation. In addition, we further find that the dual-frequency setup suggested by our system obviously outperforms others. Its average PSNR is 9.48 and color difference approximates 34.49, obviously better than other configurations. Even from the perspective of QSSIM and FISMc, its performance is relatively more stable at different exposure times. This may be explained by the fact that more frequencies imply higher interference among lights, which may lead to a variation in the overall performance.

B. Objective System Performance

Next, we evaluate our system under different confounding factors, including luminance level, photo-taking distance and angle, device type, and target object.

1) Illuminance Level: Different scenarios impose distinct requirements on illuminance level [17]. For example, many museums limit the illumination to 150 lux for most paintings, but the illuminance level of an exhibition room can be more than 600 lux according to our measurements.

Figure 11 shows the performance of our system under various illuminance levels. We can see that the degree of image pollution slightly increases with the increase in illuminance. This is because only a small proportion of light energy is captured by the camera in a low illuminance setting, making the banding effect relatively poor. With the growth in illuminance, more light energy is captured and the banding effect can be enhanced.

However, even in the worst cases with the illuminance less than 300 lux, the performance is sufficient for our purposes. The corresponding PSNR is less than 13 dB and the color difference is larger than 28, indicating significant noise on the pirated photos at the pixel level. Besides, the FISMc score is less than 0.8, which implies that the users' average opinion score should be less than 2.5, given the grading standard from 0 to 9. Nevertheless, the QSSIM scores are relatively poor, suggesting that only mild structural distortion occurs. This limitation derives from the fact that our system mainly induces illuminance fading and chromatic distortion on the image, but does not radically change its structural information. However, as we are only targeting copyright protection, rather than content protection, this is still acceptable.

2) Shooting Distance & Angle: In real applications, an attacker can take photos from various distances and angles. To examine the effective distance and angle of *Rainbow*, we place the attacker's camera at different distances and angles to the target and evaluate the corresponding performance.

Figure 12 shows that the system performance degrades with the growth of shooting distance, *e.g.*, as the shooting distance increases from 0.5 meters to 2.5 meters, the PSNR increases from 11.41 dB to 21.49 dB, while the color difference drops from 30.53 to 13.97. A similar trend can also be observed in the QSSIM and the FISMc metrics. This is because the light energy attenuates exponentially with its propagation distance. Therefore, as the shooting distance increases, less light energy is captured by the camera and the banding effect is reduced. According to the result, the working distance of our current implementation is around 2 meters. This distance can be further extended by using higher-powered lamps.

In the experiment of the shooting angle, the attacker's camera is placed 0.5 meters away at different shooting angles to the target. Since the setup is symmetric, only shooting angles from 10° to 90° are reported in Figure 13. We can see the system performance of different shooting angles is

good and relatively stable. This demonstrates that our system is robust to piracy attacks from various shooting angles.

3) Mobile Device: To further validate that our system can work on a variety of devices, we employ 7 mobile devices, including 4 iOS devices and 3 Android phones. The corresponding results are reported in Figure 14.

We can see a slight performance variation among the devices. The reason is that, given the same target scene, the exposure times determined on various devices can be distinct due to the differences in image sensor hardware. For example, a device with a sensitive CMOS image sensor, *e.g.*, iPhone 6 Plus, gives a relatively short exposure time, while a camera with a smaller aperture (such as the Huawei Honor 4x) needs longer exposure time.

However, our system works well on all of these devices. In general, the PSNR in the worst case is 13.99 dB and the average color difference is 25.04, which demonstrates that obvious distortion occurs on the pirated images at the pixel level. Meanwhile, although the QSSIM values around 0.66 suggest that only some moderate structural distortions are induced, an FISMc lower than 0.79 implies that, given a grading scale from 0 to 9, the viewers only give a mean opinion score of 2.3 to the pirated photos, suggesting significant visual quality degradation.

4) Various Targets: To examine the applicability of our system to different 2D physical objects, we employ two kinds of targets: 1) the **standard images** are selected from a standard test image database commonly-used in the computer vision community, *i.e.*, the USC-SIPI database [26]. Three images are printed in color and adopted in this test: the *baboon*, the *lenna*, and the *peppers*. Also, to examine the performance on real artwork, 2) several copies of **real paintings**, including Leonardo da Vinci's *Mona Lisa*, Rembrandt van Rijn's *Night Watch*, *Starry Night* from Vincent van Gogh and *Scream* by Edvard Munch, are adopted in this experiment. Figure 16 gives some examples of these test sets

The corresponding performance on each object is reported in Figure 15. We can observe *Rainbow* works well on all the targets. The average PSNR value is 9.33 dB while the color difference is larger than 29.73, revealing a significant discrepancy from the pirated images to the reference images at the pixel level. Apart from this, the low QSSIM and FISMc values further demonstrate that our system can induce serious visual quality degradation.

C. Subject Evaluation

Since human visual perception is subjective, the objective evaluation can not perfectly quantify the visual experience of viewers. As a complement, we recruited 20 volunteers, including 6 females and 14 males aging from 22 to 30. All of them have normal visual abilities and do not suffer color blindness. In this subjective test, the volunteers are required to provide an opinion score for their visual perceptions. Similar to [27], we use a grading scale from 1 to 5, which corresponds to five experience categories, *i.e.*, "bad", "poor", "fair", "good", and "excellent".

To examine whether a user can perceive any illuminance or chromatic flicker in our system, we present each viewer with the same scene lit by two lighting systems: one is lit by a normal LED and the other is by our system. Each lighting system is turned on alternately for 10 minutes and then the viewer is required to provide an opinion score on the *flicker perception* and the *overall experience* of our system compared to a normal LED. Table I presents the users' opinion scores about their visual experience.

(a) S1: ref. (b) S1: piracy (c) S2: ref. (d) S2: piracy

(e) S3: ref. (f) S3: piracy (g) S4: ref. (h) S4: piracy

Fig. 16. Some examples of the test sets.

TABLE I
USER'S EXPERIENCE TO OUR SYSTEM.

Performance	Grading	Avg.	Std.
Flicker Perception	5 - Excellent, no flicker. 4 - Good, bare flicker. 3 - Fair, noticeable flicker. 2 - Poor, obvious flicker. 1 - Bad, strong flicker.	4.91	0.30
Overall Experience	5 - Excellent, no discomfort. 4 - Good, bare discomfort. 3 - Fair, minor discomfort. 2 - Poor, mild discomfort. 1 - Bad, obvious discomfort.	4.55	0.69

According to the viewers' feedbacks, our system performs quite well regarding the flicker perception. The average score is 4.91, suggesting that flickering barely occurs. Also, a mean value of 4.55 on the overall experience indicates that users have a good viewing experience under our system.

VI. Discussion

As a first step towards preventing mobile-camera-based piracy on physical intellectual property, our system still has several limitations.

First, as our system relies on the banding effect caused by the rolling shutter to pollute the image, it does not work on the CCD cameras with global shutters. However, according to previous market reports [16], [28], the CMOS image sensor occupied over 83.8% of the mobile camera market in 2013 and its market is expected to grow at a CAGR of 11.8% between 2015 and 2020. This means our system already covers the majority of consumer-level cameras. Also, compared to the high-end professional camera, mobile-camera-based piracy is often harder to notice owing to their small size and ease of concealment, which renders them the main threat to copyright protection.

In addition, some medical studies point out that low-frequency light flicker could cause some discomfort [29]. As our pupils expand and shrink with the flickers, long-time exposure to a flickering light causes frequent pupillary constrictions and lead to the eye muscle relaxing, which is the main reason for eye strain and myopia. However, the minimal modulation frequency of our system is 73 Hz, which varies faster than the critical flicker frequency and thus can not be perceived by the human eye. Similarly, an incandescent lamp

which flicks at 50 or 60 Hz, is still widely used in many locations [7].

For now, our system only targets 2D physical intellectual properties, such as art paintings and photographs. We leave its extension to 3D targets, *e.g.*, sculptures or human performance, for future exploration.

VII. Related Work

Since the mobile camera is often small in size and easy to carry, photo/video-taking from a mobile device is one of the most perturbing issues. Aggregated with other context information, *e.g.*, temporal and spatial information, a malicious user can easily reveal much of a user's private information [30]. Apart from privacy violation, the copyright protection of intellectual property is another important reason why the camera is not allowed in many scenarios, *e.g.*, cinemas, museums, galleries or exhibitions [2], [8]. Existing no-photography policies are often imposed by security guards [10] which requires much human participation and is often inefficient.

As a remedy, various solutions have been proposed, one of which is to degrade the quality of pirated photo/video: intrusive methods, *e.g.*, infrared light [3], [4], are used to pollute the pirate photo/video in cinemas, while watermarking [5], [6] is also widely adopted in the film industry. Unfortunately, these approaches can be ineffective under some scenarios, *e.g.*, infrared has been evidenced to be harmful to historical paintings and cannot be deployed in many museums and galleries [7], while watermarking is not efficient enough to prevent audiences from taking videos for piracy purposes. To fill this gap, Zhang *et al.* proposed a novel video re-encoding scheme to maximize the distortion between video and camera while retaining good visual quality for the human eye [8]. However, this approach requires the re-encoding of the original digital content and can only work on digital content. Meanwhile, several anti-piracy systems aim to locate attackers in theaters by various techniques, such as infrared scanning [10], distortion analysis of the captured video [11], spread-spectrum audio watermarking [12]. These approaches either rely on a dedicated device or require modification of the content, which hinders their wide adoptions. Compared with these works, our system provides a low-cost and practical anti-piracy solution based on existing light infrastructures and extends the protection ability into the physical world.

VIII. Conclusion

In this work, we propose an anti-piracy lighting system to prevent mobile-camera-based piracy on 2D physical intellectual property. By modulating high-frequency illuminance flickers and chromatic change into existing light infrastructures, our system can create a serious visual distortion on the pirated images without affecting the human visual experience. Extensive experiments demonstrate that our system can defeat piracy attacks while providing a good lighting function in different scenarios.

Acknowledge

This work was supported in part by the RGC under Contract CERG 16212714, 16203215, Contract ITS/143/16FP-A, Contract R8015, and National Science Foundation of China under Grant 61502114, 91738202, and in part by the Huawei-HKUST Joint Laboratory.

References

[1] M. Yar, "The global epidemic of movie piracy: crime-wave or social construction?" *Media, Culture & Society*, vol. 27, pp. 677–696, 2005.
[2] C. A. Miranda, "Why can't we take pictures in art museums?"
[3] A. Ashok and *et al.*, "Do not share!: Invisible light beacons for signaling preferences to privacy-respecting cameras," in *VLCS*. ACM, 2014, pp. 39–44.
[4] T. Yamada and *et al.*, "Use of invisible noise signals to prevent privacy invasion through face recognition from camera images," in *MM*. ACM, 2012, pp. 1315–1316.
[5] I. J. Cox and *et al.*, "Secure spread spectrum watermarking for images, audio and video," in *ICIP*, vol. 3. IEEE, 1996, pp. 243–246.
[6] R. B. Wolfgang and E. J. Delp, "A watermark for digital images," in *ICIP*, vol. 3. IEEE, 1996, pp. 219–222.
[7] T. Perrin and *et al.*, "Ssl adoption by museums: survey results, analysis, and recommendations," PNNL, Tech. Rep., 2014.
[8] L. Zhang and et al., "Kaleido: You can watch it but cannot record it," in *MobiCom*. ACM, 2015, pp. 372–385.
[9] Z. Gao and *et al.*, "Dlp based anti-piracy display system," in *VCIP*. IEEE, 2014, pp. 145–148.
[10] P. E. Inc, "Pirateeye anti-piracy solution." [Online]. Available: http://www.pirateeye.com/pirateeye/technology/
[11] M.-J. Lee, K.-S. Kim, and H.-K. Lee, "Digital cinema watermarking for estimating the position of the pirate," *IEEE transactions on multimedia*, vol. 12, no. 7, pp. 605–621, 2010.
[12] Y. Nakashima, R. Tachibana, and N. Babaguchi, "Watermarked movie soundtrack finds the position of the camcorder in a theater," *IEEE Transactions on Multimedia*, vol. 11, no. 3, pp. 443–454, April 2009.
[13] QImage, "Rolling shutter vs. global shutter," 2014.
[14] T. Maintz, "Digital and medical image processing," *Universiteit Utrecht*, 2005.
[15] S. Hecht and S. Shlaer, "Intermittent stimulation by light," *The Journal of general physiology*, vol. 19, no. 6, pp. 965–977, 1936.
[16] M. Research, "Cmos image sensor market: Global trends and forecast to 2020," 2015.
[17] J. E. Kaufman and J. F. Christensen, *IES lighting handbook: The standard lighting guide*, 1972.
[18] S. Battiato and *et al.*, "Exposure correction for imaging devices: an overview," *Single-Sensor Imaging: Methods and Applications for Digital Cameras*, pp. 323–349, 2008.
[19] S. Kelby, *The digital photography book*. Peachpit Press, 2012.
[20] R. M. Karp, "Reducibility among combinatorial problems," in *Complexity of computer computations*. Springer, 1972, pp. 85–103.
[21] C. H. Papadimitriou and K. Steiglitz, *Combinatorial optimization: algorithms and complexity*. Courier Corporation, 1982.
[22] W. Lin and C.-C. J. Kuo, "Perceptual visual quality metrics: A survey," *Journal of Visual Communication and Image Representation*, vol. 22, no. 4, pp. 297–312, 2011.
[23] M. R. Luo, G. Cui, and B. Rigg, "The development of the cie 2000 colour-difference formula: Ciede2000," *Color Research & Application*, vol. 26, no. 5, pp. 340–350, 2001.
[24] A. Kolaman and O. Yadid-Pecht, "Quaternion structural similarity: a new quality index for color images," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 1526–1536, 2012.
[25] L. Zhang and *et al.*, "Fsim: A feature similarity index for image quality assessment," *IEEE transactions on Image Processing*, vol. 20, no. 8, pp. 2378–2386, 2011.
[26] A. G. Weber, "The usc-sipi image database version 5," *USC-SIPI Report*, vol. 315, pp. 1–24, 1997.
[27] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Transactions on image processing*, vol. 15, no. 11, pp. 3440–3451, 2006.
[28] T. gand view research, "Image sensor market analysis 2016," 2016.
[29] P. Drew and *et al.*, "Pupillary response to chromatic flicker," *Experimental brain research*, vol. 136, no. 2, pp. 256–262, 2001.
[30] W. Wang and Q. Zhang, "Privacy preservation for context sensing on smartphone," *IEEE/ACM Transactions on Networking*, vol. 24, no. 6, pp. 3235–3247, 2016.