# Movie Piracy Tracking using Temporal Psychovisual Modulation

Yuanchun Chen, Guangtao Zhai, Zhongpai Gao, Ke Gu, Wenjun Zhang, Menghan Hu, Jing Liu
Institute of Image Communication and Information Processing,
Shanghai Jiao Tong University, Shanghai, China
Email: {chenyuanchun, zhaiguangtao, gaozhongpai, gukesjtuee, zhangwenjun, humanghan, liujing} @sjtu.edu.cn

***Abstract*—Nowadays, camcorder piracy has great impact on the motion picture industry. Although some watermarking technologies can track the movie pirate, the video content viewed in the theater may be affected and they cannot obstruct the need of pirated movie because the watermarks in pirated moves are invisible. This paper presents a new method to defeat camcorder piracy and realize content protection in the theater using a new paradigm of information display technology, called Temporal Psychovisual Modulation (TPVM), which utilizes the differences between the human-eye perception and digital camera image-forming to stack an invisible pattern on digital screen and projector. The images formed in human vision are continuous integration of the light field, while discrete sampling is used in digital video acquisition which has "blackout" period in each sampling cycle. Based on this difference, we can decompose a movie into a set of display frames with specific patterns and broadcast them out at high speed so that the audience cannot notice any disturbance, while the video frames captured by camcorder will contain highly objectionable artifacts (i.e., the patterns). The pattern embedded in the movies can also serves as tracking information to reveal the one responsibility for the camcorder piracy.**

***Keywords*—camcorder piracy, watermarking technologies, content protection, Temporal Psychovisual Modulation (TPVM), display technology, invisible pattern, tracking information**

## I. Introduction

Movie piracy has a profound impact on the motion picture industry. The Motion Picture Association of America (MPAA) [1] conducted an investigation on the movie piracy in 2005. According to the statistics in the report, the major U.S. motion picture studios lost 6:1 billion or more annually. These losses in revenue will obviously cause serious financial problems for the studios and even contribute to their current downfall. In 2010, for example, over one million copies of James Cameron Avatar were downloaded illegally in just seven days [2]. In the view of the law, movie piracy is considered as crime all over the world. As an important source of movie piracy, the camcorder piracy accounts for about 23% of the piracy methods according to the BBC News [3]. As the source of infringing DVDs, camcorder movies spread rapidly on the internet.

The sources of camcorder pirated movies can be made prior to the theatrical release date and after the date, classified as pre-release piracy and post-release piracy respectively. Prior to the theatrical release, the film may be shown in private screenings to critics, sponsors and VIPs. There are two potential sources of piracy during this time. First, the audience, who have been

BMSB2017 – Cagliari, Italy; 978-1-5090-4937-0 Copyright ©by IEEE

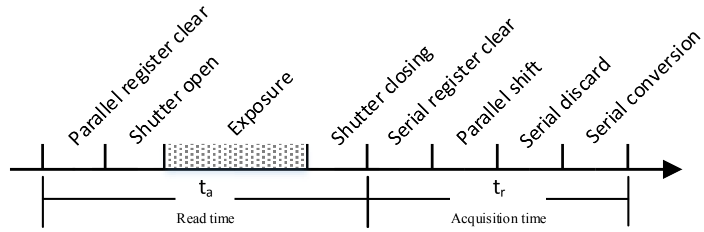



Fig. 1. Timeline of one frame recorded by CCD camera. The frame rate $f$ can be computed with the frame acquisition time taand frame read time $t_r$ as $f = 1/(t_a + t_r)$

invited as the guests of the studio may capture the movie on camcorder and make the bootleg available for unauthorized distribution. Such copies are typically poor quality with poor sound. This is known as cam. Second, the theater operator, working as an agent of the studio, can surreptitiously copy the movie at very high quality. This can be done simply, by showing the movie after hours with an empty house and a camcorder setup on a tripod in the back of the theater. The sound can be taken directly from the house sound system. This is known as telesync (TS). Once the film is released in the theaters, the potential for camcorder capture and theater operator piracy increases drastically, also classified as cam and telesync respectively [4].

As a deterrent against the camcorder piracy, several watermarking technologies have been proposed [5]. The main idea of these techniques is to embed a imperceptibly message (i.e., tracking information) into the movie. The message indicates the theater to which it was distributed, the equipment on which it was shown, the date and time of showing, and perhaps information identifying the projectionist. If movies are pirated and the illegal recordings are transmitted via the Internet or some other route, then the message can be extracted from the pirated movies to reveal the person or organization responsible for the unauthorized release. As a forensic tool, tracking information gives the content owner information to help manage the piracy problem and serves as a further surveillance and a deterrent to future piracy. However, there are two problems with watermarking techniques. First, they cannot obstruct or defeat camcorder piracy directly because the watermarks in the pirated movies are invisible and not objectionable enough. Second, although by strict definition, the alteration of watermarking must be imperceptible, both visually and audibly. Potential perceptual artifacts are still left by the watermark embedding process, the video content viewed in the theatre may be affected.

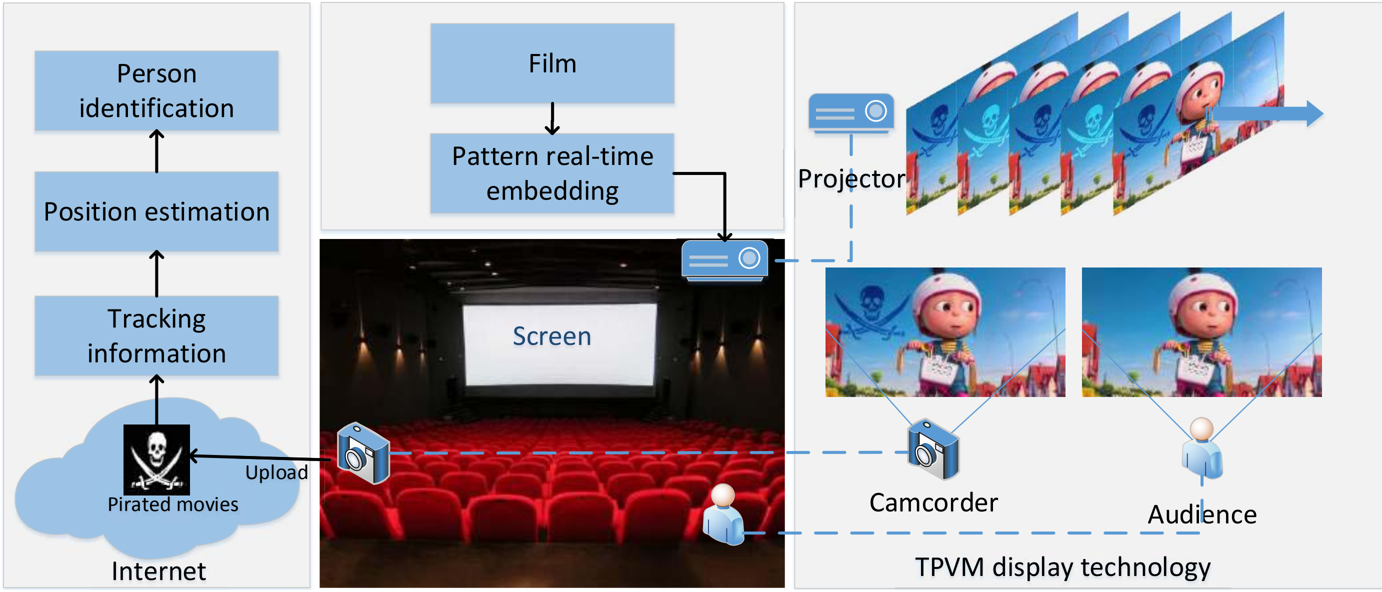



Fig. 2. Working mechanism of camcorder piracy tracking by temporal psychovisual modulation (TPVM) display technology.

This paper presents a new method to overcome these two problems by exploiting the difference of image formation mechanism between human eyes and imaging sensors. In human visual system, a steady image is formed by continuous integration of light field, whereas semiconductor imaging sensors take discrete samples of continuous light source in time. CCD (charge-coupled device) and CMOS (complimentary metal oxide semiconductor) [6] are two commonly used types of sensors. Fig. 1 shows the timeline of the imaging process for a single frame by full frame CCD in digital cameras[1]. We can find that the highest frame rate $f$ of the digital cameras is determined by the frame acquisition time $t_a$ and frame read time $t_r$ as $f = 1/(t_a + t_r)$. However, the integration period of the sensors, i.e. the exposure time is only a portion of the total imaging time $(t_a + t_r)$. Therefore, the digital cameras cannot capture all the information on the screen due to the "blackout" period [7]. They are limited by CCD working mechanism if the movie projector is fast enough.

As analysis above, digital camcorders cannot fully record all the optical signals emitted from the screen. Therefore, human eyes and semiconductor sensors apply different temporal convolution kernels to light sources when forming images. Based on this difference, we can design optical signals in the way that human and pirating camcorders perceive drastically different images. The intriguing bifurcation can be realized by the recently proposed paradigm of information display technology, called temporal psychovisual modulation (TPVM) [8][9][10]. The rationale behind TPVM is that, on one hand, the human visual system (HVS) cannot resolve temporally rapidly changing optical signals beyond flicker fusion frequency (about 60 Hz for most of people). On the other hand, nowadays, modern display can work at 120 Hz or even higher refresh rate [11]. For example, the pattern rate of spatial light modulator in digital light processing (DLP) technology is up to 32 kHz (beyond 1900 Hz for 8-bit gray) [12]. Meanwhile lots of similar study also was done as described in [13][14]. The TPVM based method can be applied to severely degrade visual quality of pirated movie contents without affecting the theatre audiences. Moreover, the pirated movies can be designed to contain some specific patterns that carry tracking information.

[1]CMOS works in a way much like CCD.

## II. PROPOSED ANTI-PIRACY SCHEME

### A. Working mechanism

Fig. 2 presents the working mechanism of the proposed anti-piracy scheme. The proposed anti-piracy scheme can be separated into three steps. 1) Pattern embedding, when playing movies in theater, the projector emits frames that are embedded some specific patterns in real time. Based on TPVM display technology, pirated movies recorded by camcorder will contain some artifacts which degrade the visual quality of the movie severely, where the visual quality of audiences will not be affected. 2) Pattern extraction, by analyzing the pirated movies uploaded to the Internet, the tracking information can be extracted to determine the theater and showtime at which the pirated movies were made, thus, revealing the organization responsibility for the illegal release and making the organization to improve the preventive measures to defeat anti-piracy. 3) Pirate identification, the position estimation system estimates the position in the theater where the pirate was, then combining with a person identification system (e.g., ticketing system, face recognition system), the person who are the pirate can be identified. In this paper, we just consider the first two steps and the third step will be studied in future works.

### B. Communication model

Fig. 3 shows the communication model of pattern based on TPVM technology. First, let $f_d$ be the flicker fusion frequency and $f_c = K f_d$ be the refresh rate of a digital display. The embedder, taking an image $\mathbf{i} \in \mathbb{R}^N$ and a pattern $\mathbf{c} \in \mathbb{R}^N$ as input ($N$ is the resolution of $\mathbf{i}$ and $\mathbf{c}$ is padded with zeros to have the same size with $\mathbf{i}$), generates a set of basis images $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_K$ which are subjected to $\sum_{j=1}^{K} \mathbf{x}_j = \mathbf{i}$. The

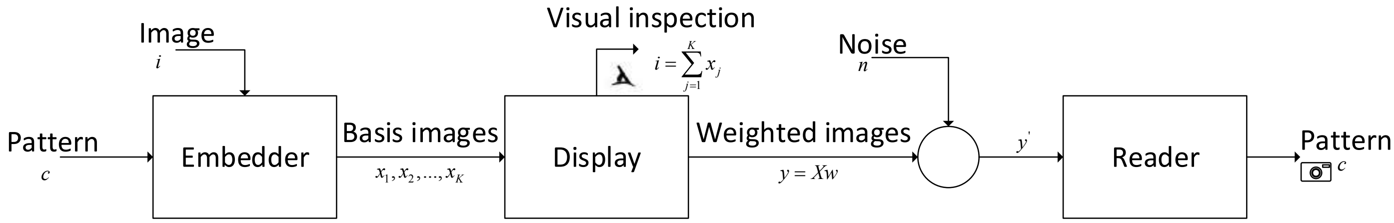
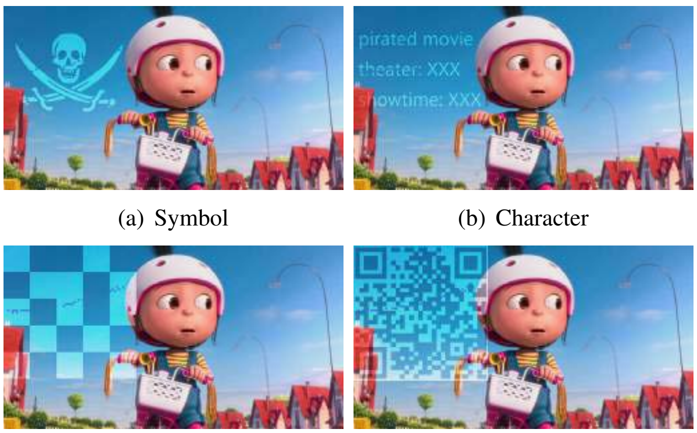



Fig. 3. Communication model of pattern based on temporal psychovisual modulation (TPVM).




(a) Symbol (b) Character

(c) Code (d) QR Code

Fig. 4. Four types of pattern embedded in movie.

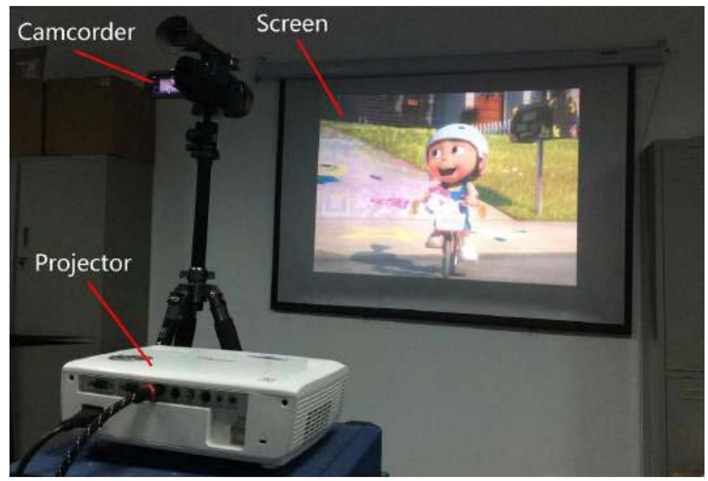



Fig. 5. Setup of the system. The QR Code in the top right corner of the display can be seen because the picture is taken by camera, and it can hardly be seen by naked eye.

high-speed display emits the basis images sequentially. For human, the HVS integrates all of light field of the basis images to form the complete image $\mathbf{i}$. But for imaging sensor, because of the "blackout" period, camera cannot receive all the optical signals emitted from screen, as shown in Fig. 1. That is, the basis images are weighted to form the image $\mathbf{y} = \mathbf{X}\mathbf{w}$, where $\mathbf{y} \in \mathbb{R}^N$, $\mathbf{w} \in \mathbb{R}^K$ and $\mathbf{X} \in \mathbb{R}^{N\times K}$ contains all $\mathbf{x}_j$'s as its columns. The channel from display device to the pattern reader is noisy, because that the relative position and angle between camera and screen are not fixed. Also, the luminance chance also impacts the image quality. All these changes can be represented by a noise term $\mathbf{n}$. Under the influence of $\mathbf{n}$, the weighted image $\mathbf{y}$ turns into $\mathbf{y}'$. For a specific pattern reader, we want to design an embedder that minimizes the visual distance between $\mathbf{c}$ and $\mathbf{y}'$, and yet $\mathbf{c}$ can be successfully extracted and decoded with high probability.

### C. Pattern selection

The pattern embedded in the movies serves two purposes: as artifacts that degrade the visual quality of the pirated movie and as semantides that carry tracking information. So, the pattern is expected to be visual unpleasing and can carry enough information. There are different types of pattern can be used in the proposed method. Fig. 4 presents four types of pattern as examples. First, some specific symbols are chosen as pattern, as shown in Fig. 4(a). Although this type pattern can be recognized easily, it is impracticable to design different symbols to distinguish different theaters and showtimes. Second, characters are chosen as pattern, as shown in Fig. 4(b). Although the tracking information is clear in the characters, it is difficult to detect the characters, especially when the images of movie are complicated. Third, some codes can be designed as pattern, as shown in Fig. 4(c). The code is similar to checkerboard and every grid carry 1-bit information. Although this type pattern can carry enough information, if one grid read error, the tracking information cannot be decoded correctly. Forth, Quick Response (QR) Codes are chosen as pattern, as shown in Fig 4(d). The QR Code can hold a large amount of data. In addition, an advanced error correction method and other unique characteristics allow the QR Code to be read more reliably and at higher speeds than other codes [8].

## III. SYSTEM DESIGN AND RESULTS

In this paper, we build a prototype system of anti-piracy based on TPVM using a projector with 120 Hz refresh rate, as shown in Fig. 5. The system is written in C++ language in Microsoft Visual Studio 2010. SDKs of NVIDIA 3D Vision, DirectX, CEGUI are used to fully realize of the application. Suppose the critical flicker frequency is 60 Hz, then the 120 Hz display is able to emit a pair of frames without causing flashing feeling of the HVS. The main problems in this system are pattern embedding and pattern extraction introduced in subsection II-A.

### A. Pattern embedding

The principle of the embedder is to minimize the visual distance between the pattern $\mathbf{c}$ and the image $\mathbf{y}'$, which is equivalent to minimize the visual distance between $\mathbf{c}$ and the weighted image $\mathbf{y}$, as shown in Fig. 3. Here we just need to consider the square area of the pattern. For an 120 Hz LCD display, $K = 2$. So we need to design two basis images $\mathbf{x}_1$ and $\mathbf{x}_2$ which are subjected to $\mathbf{x}_1 + \mathbf{x}_2 = \mathbf{i}$. Based on the TPVM technology as illustrated in subsection II-B, $\mathbf{x}_1$ and $\mathbf{x}_2$

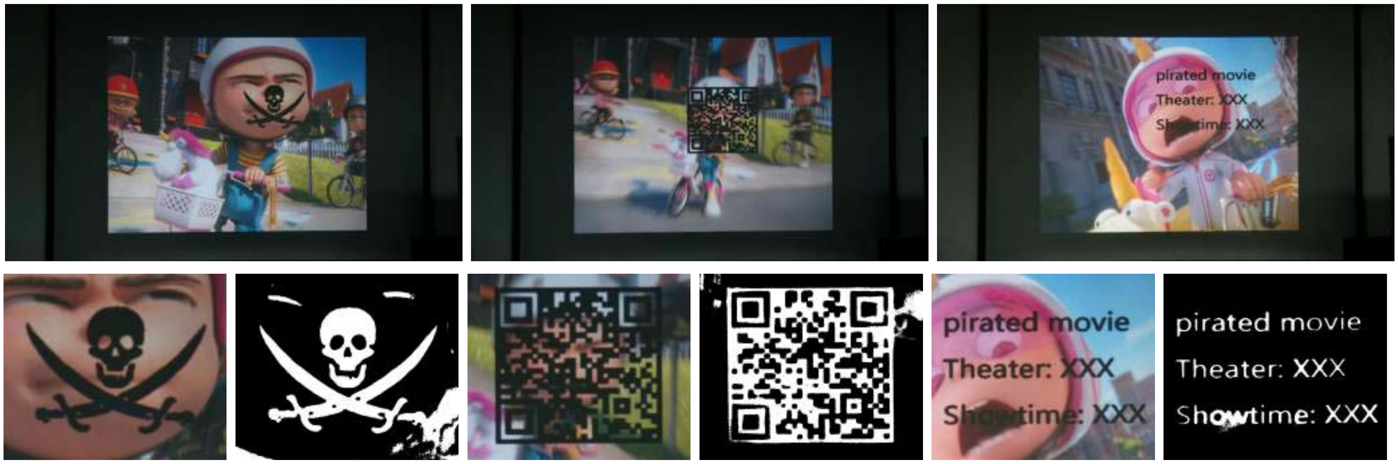



Fig. 6. Experimental Results. The first row are three images taken by camcorder with different types of patterns. The second row are the three pair results. The left ones are the subarea of the first row and the right ones are extracted from the left ones after binaryzation. The QR Code can be recognized by smartphone.

are emitted from the display alternately. Support $\mathbf{w} = [1, 0]^T$ , then $\mathbf{y} = \mathbf{x}_1 \times 1 + \mathbf{x}_2 \times 0 = \mathbf{x}_1$. Now, we just need to design $\mathbf{x}_1$ that has the minimum visual distance to the pattern $\mathbf{c}$. That is, we just need to design $\mathbf{x}_1$ such that it has the same structure as $\mathbf{c}$, as outlined in Algorithm 1.

**Algorithm 1** Algorithms for pattern embedding

**Require:** $\mathbf{i}, \mathbf{c} \in \mathbb{R}^N, K = 2$

1: $\mathbf{x}_1 = \mathbf{i}$;

2: $\mathbf{x}_1(find(\mathbf{c} == 0)) = 0$;

3: $\mathbf{x}_2 = \mathbf{i} - \mathbf{x}_1$;

### B. Pattern extraction

Patterns in the obtained images $\mathbf{y}'$ may change randomly, because the camera is not synchronized with the display. However, if we set the parameters of camera correctly, the obtained images are close to $\mathbf{x}_1$ or $\mathbf{x}_2$. In this paper, the patterns are extracted by the following two steps: 1) We combine the three RGB channels of $\mathbf{y}'$ to get a grayscale image $\mathbf{l}_y$. Second, we binarize every pixel of $\mathbf{l}_y$ according to the average of its neighbors. Fig. 6 shows the results of extracting patterns. In particular, the second column of Fig. 6 shows the results of extracted QR Code, which can be detected by standard module scanner. The method to extract pattern introduced in this paper is simple. In the future, the pattern recognition will be studied to make it be more reliable and at higher speeds.

## IV. Conclusion

The paper introduces a prototype system of defeating camcorder piracy in movie theatre based on Temporal psychovisual modulation (TPVM). By exploiting the differences between human eyes and semiconductor imaging sensors in temporal convolution of optical signals, we demonstrated how the T-PVM based technique can visually destroy the recorded movie contents while achieving visual transparency of embedding patterns to the theater audience. The pattern (e.g., QR Code) also serves as the tracking information to reveal the one responsibility for the camcorder piracy.

## V. Acknowledge

This work was supported by Science and Technology Commission of Shanghai Municipality (15DZ0500200) and NSFC (61422112,61371146,61331014).

## References

[1] S. E. Siwek, *The true cost of copyright industry piracy to the US economy*. IPI Center for Technology Freedom, 2007.

[2] J. Dorning, *Intellectual Property Theft: A Threat to U.S. Workers, Industries, and Our Economy*. DPE Research Department, 2014.

[3] B. NEWS, "The fact and fiction of camcorder piracy," [Online]. Available: http://news.bbc.co.uk/2/hi/technology/6334913.stm, 2015.

[4] J. Bloom and C. Polyzois, "Watermarking to track motion picture theft," in *Signals, Systems and Computers, 2004. Conference Record of the Thirty-Eighth Asilomar Conference on*, vol. 1, Nov 2004, pp. 363–367 Vol.1.

[5] J. Haitsma and T. Kalker, "A watermarking scheme for digital cinema," in *Image Processing, 2001. Proceedings. 2001 International Conference on*, vol. 2, Oct 2001, pp. 487–489 vol.2.

[6] S. A. Taylor *et al.*, "Ccd and cmos imaging array technologies: technology review," *UK: Xerox Research Centre Europe*, 1998.

[7] G. Zhai and X. Wu, "Defeating camcorder piracy by temporal psychovisual modulation," *J. Display Technol.*, vol. 10, no. 9, pp. 754–757, Sep 2014. [Online]. Available: http://jdt.osa.org/abstract.cfm?URI=jdt-10-9-754

[8] A. DENSO, "QR Code essentials," 2011, retrieved 12 March 2013.

[9] X. Wu and G. Zhai, "Temporal psychovisual modulation: A new paradigm of information display [exploratory dsp]," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 136–141, Jan 2013.

[10] Z. Gao, G. Zhai, and J. Zhou, "Factorization algorithms for temporal psychovisual modulation display," *IEEE Transactions on Multimedia*, vol. 18, no. 4, pp. 614–626, April 2016.

[11] C. Hu, G. Zhai, Z. Gao, and X. Min, "Information security display system based on spatial psychovisual modulation," in *2014 IEEE International Conference on Multimedia and Expo (ICME)*, July 2014, pp. 1–4.

[12] Z. Gao, G. Zhai, and X. Min, "Information security display system based on temporal psychovisual modulation," in *Circuits and Systems (ISCAS), 2014 IEEE International Symposium on*. IEEE, 2014, pp. 449–452.

[13] Y. Chen, N. Liu, G. Zhai, K. Gu, J. Wang, Z. Gao, and Y. Zhu, "Quality assessment for dual-view display system," in *2016 Visual Communications and Image Processing (VCIP)*, Nov 2016, pp. 1–4.

[14] Y. Chen, N. Liu, G. Zhai, Z. Gao, and K. Gu, "Information security display system on android device," in *2016 IEEE Region 10 Conference (TENCON)*, Nov 2016, pp. 1634–1637.