COMBINING INFORMATION DISPLAY AND VISIBLE LIGHT WIRELESS COMMUNICATION

Xiaolin Wu *Xiao Shu*

McMaster University
Shanghai Jiao Tong University

ABSTRACT

This paper opens up an unforeseen and intriguing application area of our recent pioneer work of temporal psychovisual modulation: wireless optical communication. It is demonstrated how a high-speed optoelectronic display functions as a 2D array of optical transmitters and at the same exhibits conventional images as usual. At the receiving end, digital cameras can download data via the optical MIMO link while user(s) can work and read the display as they are accustomed. The said unification of information display and wireless optical communication is made possible by psychovisually based image processing.

Index Terms**—** Wireless optical communication, temporal psychovisual modulation, information display, MIMO.

1. INTRODUCTION

Recently, we developed and published a new paradigm of information displays called temporal psychovisual modulation (TPVM) [1, 2]. When pioneering the TPVM technology, our main motivation is to create a new type of multiuser visual experiences with optoelectronic image/video displays. The landmark functionality of TPVM displays is to generate multiple interference-free images concurrently on the same physical display medium for different viewers, without sacrificing spatial resolution or confining viewing positions. This gives TPVM a wide range of image, video and graphics applications, including stereoscopic displays, virtual and augmented reality, electronic gaming, privacy-protection displays, antipiracy displays, etc.

In this position paper, we extend the TPVM applications beyond the realm of visual presentations and advocate an ingenious use of the TPVM display technology, which gives birth to a new modality of wireless data communication. The innovation is to make an optoelectronic display, provided that its refresh rate is higher than the flicker frequency of the human visual system, transmit spatially-modulated visible lights in such ways that the high-speed display simultaneously fulfils two roles: image/video presentation and wireless optical communication.

In a bid to combine conventional visual information display with wireless optical communication, we consider a high-speed optoelectronic display and a high-speed camera as paired 2D arrays of transmitters and receivers, resulting a wireless optical multiple-input and multiple-output (MIMO) link. By nature such MIMO optical links enjoy an unprecedented degree of spatial diversity and can thus sustain very high data transmission rates. This promise has brought a great deal of research efforts to wireless optical MIMO communications [3]. Many applications of wireless optical communication have been suggested and tested, including in-door communications [4], chip and board level optical interconnects [5], free-space long range optical communications [6], etc. In contrast to all published works in the field, the new technology to be detailed below offers a distinct and defining feature: the 2D array of optical transmitters can generate user-wanted normal images not just some semantically meaningless 2D barcode or noise-like patterns. The only known dual-purpose visible light communication apparatus is ceiling LED installations that provide both data down link and illumination [7].

In today's offices, homes and public spaces, optoelectronic displays and digital cameras are ubiquitous and inexpensive, either standalone or more commonly in form of accessories. It is not hard to find applications of the proposed dual-purpose technology in daily pedestrian settings. For example, the wire clutters on desks can be eliminated by enabling computer monitors or laptop screens to transmit data to smartphones, pads, peripherals, etc. and vice versa, meanwhile the users can still work and read their displays as usual. Furthermore, as the data link is confined by the line of sight, a strong measure of data security is gained conveniently compared against popular wifi links. Another application scenario is to double public bulletin boards, advertisement displays, television sets and the alike, which are widely deployed in airports, train stations, shopping malls, stadiums, restaurants, etc. as the transmitters of wireless optical MIMO links; users double their smartphone cameras as optical MIMO receivers to acquire desired data. In contrast to their predecessors

This research is supported by Natural Sciences and Engineering Research Council of Canada, and by National Natural Science Foundation of China (NSFC grant 61331014).

that need precisely aligned optical transceivers, the aforementioned TPVM-based wireless communication solutions are free of any installation and of any added costs, because the displays and cameras are deployed for other purposes anyways.

The remainder of the paper is structured as follows. Section 2 outlines the principle and basic elements of the TPVM display technology to lay the basis for TPVM-based wireless optical MIMO communication. Section 3 presents the main contribution of this research: a unified framework for information display and wireless optical communication. Section 4 discusses the design of camera-based receivers for the proposed wireless optical communication system. Section 5 includes experimental results as proof of concept, observations and remarks.

2. TPVM BACKGROUND

The premise of the TPVM information display paradigm is a well known psychophysical phenomenon: humans cannot resolve temporally rapidly changing optical signals beyond flicker fusion frequency, which is 60 Hz for most viewers and under most conditions [8]. On the other hand, nowadays many optoelectronic displays can crank up their frame rate well above 60 Hz. For examples, even main-stream liquid crystal displays (LCD) can operate at 120 Hz and above; the light-emitting diode (LED) and organic light-emitting diode (OLED) displays can easily reach frame rates that are orders of magnitude higher. In other words, a high-speed optoelectronic display can broadcast in visible spectrum a far greater amount of optical information than the human visual system (HVS) can possibly assimilate. This psychovisual redundancy of fast optoelectronic displays was exploited by Wu and Zhai in their design of the TPVM display system that can generate multiple visual percepts for different viewers [1].

In the first TPVM paper [1], concurrent exhibitions of different images on the same display surface were made possible by exploiting both psychovisual redundancy of a high-speed display and statistical redundancy of the concurrent visual signals. TPVM performs temporal modulation of 2D optical signals as explained in Fig. 1. A high-speed display sequentially emits a set of $M \geq 2$ so-called atom frames. These atom frames are not conventional images but rather constituent parts of $K \geq 2$ target images; the M atom frames are amplitude modulated by display-synchronized light-regulating viewing devices, such as liquid crystal (LC) glasses, to generate different target images to HVS. The LC glasses are used to control how much of incoming light to pass through and enter the viewer's retina up to a quantization precision (say, 256 levels); moreover, they can switch between transparency levels at sufficiently high speed [9]. By choosing suitable light attenuation coefficients (modulation weights) for his/her LC glass in synchronization with the atom frames, a viewer can perceive one of the K target images that is generated by the chosen amplitude modulation of the display-emitted 2D optical signals.

TPVM is a problem of signal decomposition $\mathbf{Y} = \mathbf{XW}$, where Y is an $N \times K$ matrix with column k being target image $\mathbf{y}_k$ of N pixels, $1 \leq k \leq K$; $\mathbf{X}$ is an $N \times M$ matrix whose columns are the M atom frames $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_M$ of N pixels; $\mathbf{W}$ is an $M \times K$ modulation matrix whose columns are the K weighting vectors $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_K$ to generate $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_K$ using $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_M$, respectively. Furthermore, $\mathbf{Y} = \mathbf{XW}$ has to be a non-negative matrix factorization (NMF) [10], because the light energy cannot be negative and active LC glasses cannot implement negative weights.

3. COMBINING VISUAL PRESENTATION AND WIRELESS OPTICAL COMMUNICATION

Now let us change perspective, and reexamine the TPVM display system as a 2D array of optical signal transmitters for MIMO wireless communication, but under the condition that any given semantically meaningful image can also be exhibited to viewers who watch the screen directly not through light-regulating glasses. This target image is the result of HVS fusing all M atom frames without any attenuation, which is the signal $\mathbf{y} = \mathbf{x}_1 + \mathbf{x}_2 + \ldots + \mathbf{x}_M$.

Our design objective is a hybrid optoelectronic signal processing system that can fulfil dual roles of conventional visual presentation and wireless optical communication, as depicted by Fig. 2. Here the atom frames $\mathbf{x}_m$, $1 \leq m \leq M$, are the data carrier frames that are transmitted by the display to both viewer(s) and receiver camera(s), at a frame rate above the critical flicker frequency. The underlying task is to embed the transmitted information into the M non-negative atom frames $\mathbf{x}_m$ under the constraint that $\mathbf{y} = \mathbf{x}_1 + \mathbf{x}_2 + \ldots + \mathbf{x}_M$ is a perceptually satisfactory approximation of a user-chosen natural image. The challenge is how to increase the wireless optical communication data rate and at the same time keep the embedded information in atom frames perceptually indiscernible to the viewers who are using the display functionalities of the dual display-communication system.

For the benefits of high data rate and high perceptual image quality, the frame rate f_d of the display should be made as high as possible; at least $f_d \geq 120$Hz because $M = f_d/60 \geq 2$ atom frames are required to make the individual atom frames perceptually unrecognizable to viewers. Fortunately, displays and cameras of frame rate 120Hz and above are now commonplace.

Next we discuss how to embed, at each pixel n (or superpixel as introduced in the next section), $M/2$ bits b_m, $1 \leq m \leq M/2$, into the corresponding pixel values $x_m(n)$ of M atom frames, $1 \leq m \leq M$, while satisfying $\mathbf{y} = \mathbf{x}_1 + \mathbf{x}_2 + \ldots + \mathbf{x}_M$ given a natural image $\mathbf{y}$. A simple scheme, which is more suited for normal dynamic range displays such as LCD

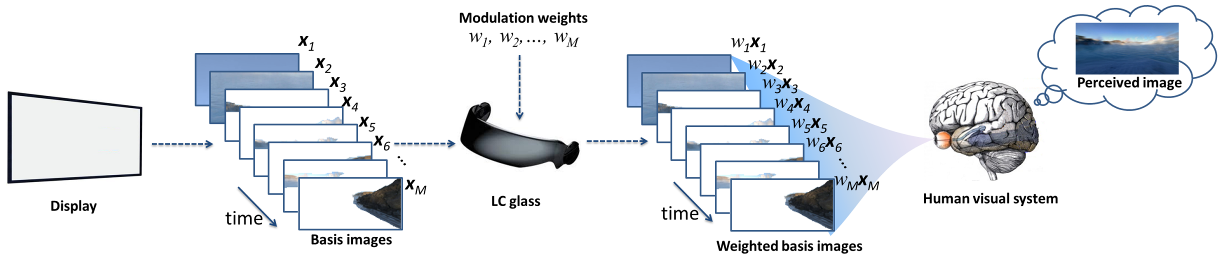

Fig. 1. Image formation via temporal psychovisual modulation in visible spectrum. A high-speed display of refresh rate f_d Hz emits $M = f_d/f_c$ atom frames, where f_c is the critical flicker frequency. The light fields of these M atom frames pass through and get amplitude-modulated by an active LC glass. The M modulated atom frames are temporally fused by HVS and perceived as an image.

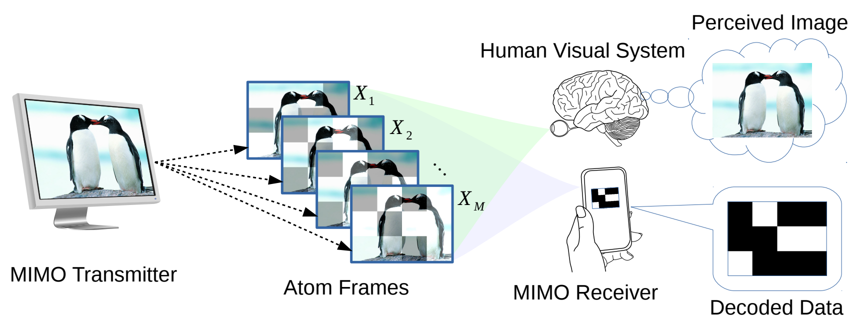

Fig. 2. TPVM-based joint information display and wireless optical communication system.

displays, is

$$\begin{aligned} x_{2m-1}(n) &= y(n) - (-1)^{b_m}\delta \\ x_{2m}(n) &= y(n) + (-1)^{b_m}\delta, \quad 1 \le m \le M/2 \end{aligned} \qquad (1)$$

where $0 < \delta < 1$ is the strength of the transmitted signal. In general, greater δ increases SNR of the received signal but decreases the perceptual quality of the displayed image, vice versa. Note that the matrix formula (1) is only a simplified expression of the information embedding; by no means the embedded signal strength δ should be constant everywhere in the image $\mathbf{y}$. Some local refinements are possible and necessary if the dynamic range of atom frames $\mathbf{x}_m$ is $[0, 1]$. At pixel positions where adding/subtracting δ to/from y exceeds the display dynamic range, clipping has to take place in (1). As the clipping causes $\mathbf{y} = \mathbf{x}_{2m-1} + \mathbf{x}_{2m}$, the embedded bits may become visible in vicinities where $\mathbf{y}$ is near 0 or 1 (0 or 255 for 8-bit intensity values). To mitigate this problem, one can adapt δ to the local waveform of image $\mathbf{y}$ by reducing δ when $\mathbf{y}$ is near either end of the display dynamic range.

For high dynamic range (HDR) displays, such as LED displays, the peak intensity of atom frames $\mathbf{x}_m$ in short durations can be significantly greater than maximum intensity of the target image $\mathbf{y}$. Keep in mind that in TPVM the perceived image $\mathbf{y} = \mathbf{x}_1 + \mathbf{x}_2 + \ldots + \mathbf{x}_M$ is the result of averaging the intensities of the atom frames. Thus, here we can take the advantage of high display dynamic range and use another scheme of embedding information bits into the atom frames:

$$\begin{aligned} x_{2m-1}(n) &= 2b_m y(n) \\ x_{2m}(n) &= 2y(n) - x_{2m-1}(n) \end{aligned} \qquad (2)$$

This scheme allows higher strength of embedded information bits and hence benefits the decoder wherever $\mathbf{y} > 0.5$.

4. SCREEN-CAMERA ALIGNMENT AND MIMO RECEIVER DESIGN

The design priority of this work is given to user friendliness. The most suited application scenario is that users aim their hand-held cameras at and shoot the TPVM screen to download data in an in-door environment. In such settings the screen is likely the strongest light source in the captured video, minimizing ambient light interferences. A main issue in the implementation of the proposed wireless optical MIMO communication method is how to align between the pixel array of the screen (transmitters) and the sensor array of the camera (receivers).

As the effective area of a TPVM screen is a planar rectangle in space, it is imaged by a camera as a quadrilateral. A pixel on a screen projects to a point inside the imaged quadrilateral. This mapping between corresponding pixel positions

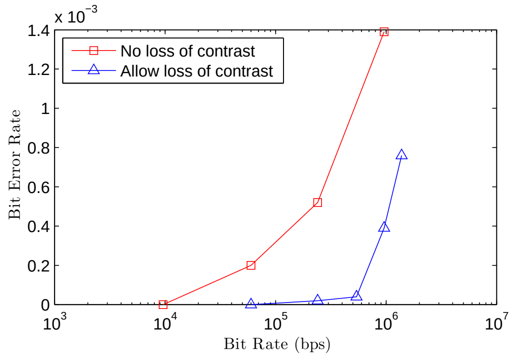

Fig. 3. The bit error rate as a function of the bit rate. Various bit rates can be achieved by changing the size of the superpixel blocks.

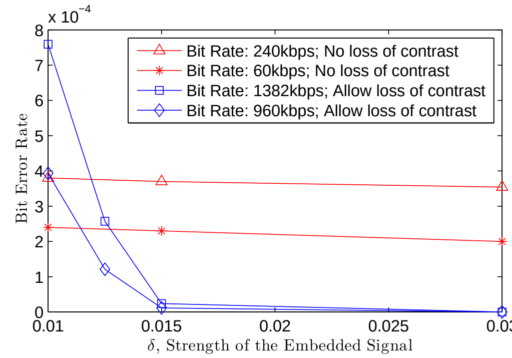

Fig. 4. The stronger the embedded signal, the lower the bit error rate is. Allowing slight compromise to the output contrast of the display could greatly reduce the bit error rate as the strength of the embedded signal increases.

can be modelled using homographic transformation:

$$\begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} \propto \underbrace{\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix}}_{\mathbf{H}} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \quad (3)$$

where (u, v) and (u', v') are pixel positions on the TPVM screen and in the captured screen image, respectively. Anchoring the four corners of the rectangular screen suffice to determine the eight coefficients in the homography matrix $\mathbf{H}$ [11]. With more point correspondences, such as pixels along the edges of the screen, the estimation of $\mathbf{H}$ can be more robust against the noise in the measurement of feature positions [12, 13].

Considering inevitable registration errors, sensor and ambient noises, we introduce redundancies into the system by partitioning the screen pixels into $k \times k$ blocks, each of which is treated as a superpixel. Within each subpixel the embedded information bit is repeated k^2 times to combat errors and noises.

The above pixel blocking strategy is chosen not only as an error resilience measure, it also happens to be a stipulation of digital camera hardware architecture. Many consumer grade cameras (e.g., iphone cameras, Panasonic LUMIX ZS30, CASIO E-100, Canon PowerShot G16, and many others) have to trade spatial resolution for high frame rate. This trade-off of spatial and temporal resolutions is an intrinsic camera design rule in the interest of price-performance ratio. Thus, the proposed TPVM-based wireless optical communication method extends the utilities of smartphone and cameras, and make them a convenient wireless data link by simply switching the camera mode.

5. EXPERIMENTAL RESULTS AND DISCUSSIONS

We built a simple prototype of the proposed information display system that offers visible light wireless communication as a free, side benefit. Empirical findings are reported in this section, but they are very preliminary and mainly serve as a proof of concept in this position paper, and also for discussions to identify important design issues and suggest the future research topics.

In the TPVM-based wireless MIMO communication system, each pixel could be programmed as an independent light transmitter. But the spatial resolution of modern information displays is too high for average cameras to resolve individual pixels, making the use of $k \times k$ superpixels necessary. The smaller the value k, the higher the nominal bit rate. However, as shown in Fig. 3, the bit error rate increases quickly as the nominal bit rate is beyond a threshold, making it difficult to reliably decode the embedded information. The reason is obvious: as the superpixel blocks become smaller, it becomes more difficult for the MIMO receiver (the imager) and image registration algorithm to accurately align and identify these blocks amidst sensor noises and geometric distortion of the imaged screen.

Another influence factor on bit error rate is the strength of the embedded signal δ. In Fig. 4, as expected, the stronger the embedded signal, the lower the bit error rate. But the embedded signal strength δ is constrained not to perceptually disturb the carrier natural image $\mathbf{y}$ being displayed. A design decision is how to balance the effective communication bandwidth and the perceptual quality of the carrier images. A strategy is to slightly reduce the dynamic range of the carrier image signal $\mathbf{y}$ without perceptually reducing its contrast too much. This allows a greater δ value to be added/subtracted to/from the pixel values of $\mathbf{y}$ without exceeding the display dynamic range, and thus reduces bit error rate without making embedded information bits visible. The better solution is to make the embedded signal strength δ spatially adaptive by setting δ_i for pixel location i to the maximum possible such that $y_i + \delta_i \leq 1$ and $y_i - \delta_i \geq 0$.

6. REFERENCES

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

[2] G. Zhai and X. Wu, "Multiuser collaborative viewport via temporal psychovisual modulation," *IEEE Signal Processing Magazine*, vol. 31, no. 5, pp. 144–149, 2014.

[3] D. K. Borah, A. C. Boucouvalas, C. C. Davis, S. Hranilovic, and K. Yiannopoulos, "A review of communication-oriented optical wireless systems," *EURASIP Journal on Wireless Communications and Networking*, vol. 1, pp. 1–28, 2012.

[4] S. Hranilovic and F. R. Kschischang, "A pixelated MIMO wireless optical communication system," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 12, no. 4, pp. 859–874, 2006.

[5] D. V. Plant and A. G. Kirk, "Optical interconnects at the chip and board level: challenges and solutions," *Proc. IEEE*, vol. 88, no. 6, pp. 806–818, 2000.

[6] N. Cvijetic, S. G. Wilson, and M. Brandt-Pearce, "Receiver optimization in turbulant free-space optical MIMO channels with APDs and Q-ary PPM," *IEEE Photonic Technology Letters*, vol. 19, no. 2, pp. 103–105, 2007.

[7] H. Elgala, R. Mesleh, and H. Haas, "Indoor broadcasting via white LEDs and OFDM," *IEEE Trans. Consumer Electron*, vol. 55, no. 3, pp. 1127–1134, 2009.

[8] M. Kalloniatis and C. Luu, "Temporal resolution," *Webvision: The Organization of the Retina and Visual System*, 2007.

[9] L. Komitov, G. Hegde, and D. Kolev, "Fast liquid crystal light shutter," *Journal of Physics D: Applied Physics*, vol. 44, pp. 1–5, 2011.

[10] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, pp. 788–791, 1999.

[11] R. Hartley and A. Zisserman, *Multiple View Geomerty in Computer Vision*, Cambridge University Press, second edition, 2003.

[12] J. J. Lee and G. Y. Kim, "Robust estimation of camera homographyusing fuzzy RANSAC," *ICCSA '07: International Conference on Computational Science and Its Applications*, pp. 992–1002, 2007.

[13] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.