Improved Temporal Psychovisual Modulation for Backward-Compatible Stereoscopic Display

Rui Ma, Oscar C. Au, Pengfei Wan, Lingfeng Xu, Wenxiu Sun, Wei Hu
The Hong Kong University of Science and Technology
{rmaaa, eeau, leoman, lingfengxu, eeshine, huwei}@ust.hk

***Abstract*—Recently a new information display technique named temporal psychovisual modulation (TPVM) was developed to extend the utilities of optoelectronic displays. A typical usage of TPVM is the backward-compatible stereoscopic display, where a stereoscopic view is perceived with 3-D glasses while a 2-D view of the same scene is concurrently available for naked-eye viewers. The most challenging part of this task is to display a clean 2-D image without ghosting artifacts. In order to reduce the ghosting artifacts, we present an improved TPVM system for backward-compatible stereoscopic display. Different from previous approaches, we rebuild the TPVM system by investigating the light perception process of human eyes, and reformulate the problem as a minimization of the light intensity difference between the target and perceived views. In order to derive the light intensities, we model it as a monotonic increasing function of the pixel intensity consisting of a scaling plus an offset, and incorporate this model into our system. By solving the optimization problem, ghosting artifacts are significantly reduced in the 2-D view. Moreover, a closed-form solution is provided for real-time applications. Experimental results demonstrate that the proposed method is superior to previous approaches in terms of the 2-D view quality, while the quality of the 3-D view remains unchanged.**

***Index Terms*—Information display, stereoscopy, light perception**

I. Introduction

With the rapid development of information display technology, the emerging 3-D stereoscopic displays have enriched and benefited our daily life. Stereoscopic displays in market generally fall into two categories: the ones requiring 3-D glasses [1] and the others free of glasses [2]. The naked-eye stereoscopic displays are also called autostereoscopic displays, in which spatial multiplexing technique is deployed. On the other hand, temporal multiplexing is used for time-sequential stereoscopic displays, where stereo frames are perceived by human eyes using a high refresh rate display and a pair of synchronized shutter glasses. Compared with autostereoscopic displays, a time-sequential stereoscopic display has much lower inter-channel crosstalk since the left-eye and right-eye signals are separated, and it is also much less restrictive of viewing angles and distances [3].

A backward-compatible stereoscopic display is capable of exhibiting a 3-D view to viewers wearing stereoscopic glasses and a 2-D view of the same scene to naked-eye viewers at the same time. It provides the compatibility with the traditional 2-D video signal as well as an optional 2-D choice for people who do not have stereoscopic glasses or feel uncomfortable with 3-D experiences. An early attempt to address this problem is [4], where a clean 2-D image is derived by flattening the disparity between stereo images at the cost of a degraded quality for the 3-D view. This method relies heavily on the accuracy of disparity estimation which is known to be computational expensive. Recently Wu et al. [5] present a new approach based on temporal psychovisual modulation (TPVM) technology [6]. TPVM provides a new paradigm of information display, where the modern high refreshing rate display as well as the traditional signal processing techniques are combined to enhance the current optoelectronic displays. The TPVM-based approach [5] is able to generate a naked-eye clean 2-D image as well as high quality 3-D images at the same time. Nevertheless it suffers from ghosting artifacts in 2-D views due to the simple blending of the time-sequential displayed frames. Jiao et al. [7] propose to reduce the ghosting artifacts in the original TPVM framework by controlling the display backlight, where a closed-form solution is provided for real-time applications. In previous TPVM-based approaches, the system is implicitly defined in the pixel intensity domain, where both target and perceived views are regarded as images composed of pixel values. However the brightness of an image is determined by the incoming light intensities that we perceived from the image rather than the pixel values. The same pixel value can appear bright or dark when the lighting conditions change. Therefore, it is inaccurate to use the pixel values to measure the actual signal perceived by human eyes.

In order to address the aforementioned issue, we propose an improved TPVM system for backward-compatible stereoscopic display, which considers the effect of light intensities during the image perception process of human eyes. In particular, the problem is reformulated as a minimization of the light intensity difference between target views and perceived views. In order to derive the light intensities, the relationship between the light intensity and the pixel intensity is exploited and modeled as a monotonic increasing function consisting of a scaling plus an offset. By incorporating this function into our improved TPVM system, the minimization framework is translated into a least-square problem, which is tractable using a generic solver. A closed-form solution is also derived for real-time applications. Experimental results demonstrate that the proposed method effectively reduces the ghosting artifacts and is superior to previous approaches in terms of the 2-D view quality, while the quality of the 3-D view remains unchanged.

The rest of the paper is organized as follows. The original TPVM system is reviewed in Section II. Our improved TPVM approach is elaborated on in Section III. Experimental results are presented and discussed in Section IV. Section V draws the conclusion.

II. Review of the original TPVM system

In a TPVM-based backward-compatible stereoscopic display system, four *atom frames* $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4 \in \mathbb{R}^{N\times 1}$ (N is the total number of pixels in a frame) are cyclically emitted on a 240Hz display. A pair of synchronized shutter glasses are used to attenuate the light transmittancy of individual atom frames with weights $\mathbf{w}^L = (w_1^L, w_2^L, w_3^L, w_4^L)^T$ and $\mathbf{w}^R = (w_1^R, w_2^R, w_3^R, w_4^R)^T$ for left and right eye respectively. Since the 240Hz refresh rate far exceeds the critical flicker frequency of 60Hz [8] for human visual system (HVS), the time-sequential atom frames would be fused by HVS and perceived as an image [6]. This process can be modeled as an accumulation of consecutive frames. Therefore with a pair of shutter glasses, the attenuated atom frames adjusted by $\mathbf{w}^L$ and $\mathbf{w}^R$ are fused to generate the left and right views, namely $\mathbf{y}^L$ and $\mathbf{y}^R$ respectively, which further form the 3-D view in HVS. While for a naked-eye viewer, all four atom frames would be fused by HVS

978-1-4799-7088-9/14/$31.00 ©2014 IEEE

without attenuation and a 2-D view of the same scene called $\mathbf{y}_0$ is perceived. The whole process can be described as follows:

$$[\mathbf{y}_0\ \mathbf{y}^L\ \mathbf{y}^R] = [\mathbf{x}_1\ \mathbf{x}_2\ \mathbf{x}_3\ \mathbf{x}_4] \begin{bmatrix} s & w_1^L & w_1^R \\ s & w_2^L & w_2^R \\ s & w_3^L & w_3^R \\ s & w_4^L & w_4^R \end{bmatrix}, \quad (1)$$

where $\mathbf{y}_0, \mathbf{y}^L, \mathbf{y}^R \in \mathbb{R}^{N\times 1}$, and the scaling factor s is used to adjust the intensity of the 2-D view. Therefore the problem of concurrently displaying 2-D and 3-D views can be formulated as a non-negative matrix factorization (NMF) [5]:

$$\begin{aligned} &\min_{\mathbf{X},\mathbf{W}} ||s\mathbf{Y} - \mathbf{X}\mathbf{W}||_F \\ &\text{s.t. } 0 \le \mathbf{X} \le 1,\ 0 \le \mathbf{W} \le 1, \end{aligned} \quad (2)$$

where the columns of matrix $\mathbf{Y}$ are target 2-D and 3-D views, and the columns of matrix $\mathbf{X}$ are the atom frames. $\mathbf{W} = [s\mathbf{1}\ \mathbf{w}^L\ \mathbf{w}^R]$ is the weighting matrix. $||\cdot||_F$ denotes the Frobenius norm. The two non-negative normalization constraints are due to the fact that 1) the light energy emitted by the display is non-negative and bounded, 2) the shutter glasses could not enhance the light energy and the weights are non-negative.

III. The Improved TPVM system

A. Introducing of Light Intensity

In the original TPVM system, both target views and atom frames are implicitly defined in the pixel intensity domain. Note that for the same displayed image, the perceived brightness may change according to different lighting conditions. A bright pixel under dim light can be perceptually darker than a dark pixel under sufficient light. These examples indicate that pixel intensity is not an appropriate measurement to describe the actually signal perceived by the viewer.

This problem motivates us to derive a better quantity to precisely describe the process of light perception. In HVS, a bright pixel is distinguished from a dark pixel because the amount of light perceived by the eyes in unit time of the bright pixel is greater than that of the dark pixel. Inspired by this, we reserve the term *light intensity* to describe the physical quantity of the total amount of light perceived by our eyes in unit time period.

With the concept of light intensity, the original TPVM system is rebuilt to better fit the visual perception of HVS. The atom frames $\mathbf{x}_k$ $(k = 1, ..., 4)$ are now redefined as the normalized light intensities of the display within the range $[0, 1]$. Using a shutter glasses of weights $\mathbf{w} = (w_1, w_2, w_3, w_4)^T$, the accumulated amount of light perceived in a $4t$ ($t = 1/240$sec) time period is expressed as $tw_1\mathbf{x}_1 + tw_2\mathbf{x}_2 + tw_3\mathbf{x}_3 + tw_4\mathbf{x}_4$. Then we redefine the notation $\mathbf{y}$ ($\mathbf{y}_0$, $\mathbf{y}^L$ or $\mathbf{y}^R$) as the light intensities of the target view. The accumulated amount of light desired to be achieved in a $4t$ time period is $4t\mathbf{y}$. In order to display the target view $\mathbf{y}$, the amount of light we perceived should be as close as possible to the amount of light of the target view, i.e.

$$4t\mathbf{y} = tw_1\mathbf{x}_1 + tw_2\mathbf{x}_2 + tw_3\mathbf{x}_3 + tw_4\mathbf{x}_4,$$

or in a simpler form by dropping the positive constant t

$$4\mathbf{y} = w_1\mathbf{x}_1 + w_2\mathbf{x}_2 + w_3\mathbf{x}_3 + w_4\mathbf{x}_4.$$

With the above equation, the original TPVM system (1) is reformulated as

$$4[\mathbf{y}_0\ \mathbf{y}^L\ \mathbf{y}^R] = [\mathbf{x}_1\ \mathbf{x}_2\ \mathbf{x}_3\ \mathbf{x}_4] \begin{bmatrix} 1 & w_1^L & w_1^R \\ 1 & w_2^L & w_2^R \\ 1 & w_3^L & w_3^R \\ 1 & w_4^L & w_4^R \end{bmatrix}. \quad (3)$$

Apparently the new formulation only adds a scaling factor, however it describes the light intensity aggregation process within a time period which could not be addressed in the pixel intensity domain. Therefore, the new formulation better presents the TPVM system by considering the perception of the image signal in HVS.

B. Modeling Light Intensity from Pixel Intensity

While in the improved TPVM system $\mathbf{y}_0$, $\mathbf{y}^L$ and $\mathbf{y}^R$ denote the light intensities of the target views (they denote the target images in the original TPVM system), to avoid confusion, *we use a new set of symbols, $\mathbf{t}_0, \mathbf{t}^L, \mathbf{t}^R$, to represent the target 2-D image, the left image and the right image of the 3-D view respectively.* All pixel intensities are normalized to the range of $[0, 1]$.

When investigating the original TPVM system (1) from the perspective of light intensity, the relationship between pixel intensities and light intensities of the target views is implicitly determined to be proportional as follows.

$$4\mathbf{y}_0 = s\mathbf{t}_0,\ 4\mathbf{y}^L = \mathbf{t}^L,\ 4\mathbf{y}^R = \mathbf{t}^R, \quad (4)$$

Note that (4) indicates that the two quantities are equal in amount, not in physics. In real cases the light intensities of the target view do not have to be necessarily proportional to the pixel intensities. Generally it can be designed as any monotonic increasing function. In our method we assume a simple function consisting of a scaling plus an offset, and (4) is modified as

$$\begin{aligned} 4\mathbf{y}_0 &= s\mathbf{t}_0 + b\mathbf{1}, \\ 4\mathbf{y}^L &= s^L\mathbf{t}^L + b^L\mathbf{1}, \\ 4\mathbf{y}^R &= s^R\mathbf{t}^R + b^R\mathbf{1}, \end{aligned} \quad (5)$$

where all the scaling factors should be positive. In each pair of the scaling factor and the offset, the scaling factor adjusts the light intensity range (LIR) between the maximal and minimal possible light intensity values, while the offset controls the global light intensity shift for the corresponding target view. The added offset introduces one more freedom to adjust the light intensities of the target views based on the target images, which plays a crucial role in our method. The previous TPVM system (4) can be regarded as a special case when the offsets are equal to zero. Substitute the above designed model into (3), the improved TPVM system becomes

$$\begin{aligned} &\underbrace{\begin{bmatrix}\mathbf{t}_0\ \mathbf{t}^L\ \mathbf{t}^R\ \mathbf{1}\end{bmatrix}}_{\mathbf{T}} \underbrace{\begin{bmatrix} s & 0 & 0 \\ 0 & s^L & 0 \\ 0 & 0 & s^R \\ b & b^L & b^R \end{bmatrix}}_{\mathbf{A}} \\ &\qquad = [\mathbf{x}_1\ \mathbf{x}_2\ \mathbf{x}_3\ \mathbf{x}_4] \begin{bmatrix} 1 & w_1^L & w_1^R \\ 1 & w_2^L & w_2^R \\ 1 & w_3^L & w_3^R \\ 1 & w_4^L & w_4^R \end{bmatrix}. \end{aligned} \quad (6)$$

Then the original NMF problem (2) of concurrently displaying 2-D and 3-D views is now reformulated as

$$\begin{aligned} &\min_{\mathbf{A},\mathbf{X},\mathbf{W}} ||\mathbf{T}\mathbf{A} - \mathbf{X}\mathbf{W}||_F \\ &\text{s.t. } 0 \le \mathbf{X} \le 1,\ 0 \le \mathbf{W} \le 1,\ s, s^L, s^R > 0, \end{aligned} \quad (7)$$

Since $\mathbf{A}$ acts as a modeling matrix mapping the target images into the target light intensities, it can be set as user-controlled parameters for flexible light intensity adjustment as the user like. Alternatively, $\mathbf{A}$ can also be derived optimally by the system in order to achieve the best objective quality. When $\mathbf{A}$ is fixed, the problem is still a NMF which is known to be computational complex.

C. Practical Solution

As described in [5], a practical solution of TPVM-based method is to fix the weighting matrix as $\mathbf{w}^L = [1\ 0\ 0\ 0]^T$, $\mathbf{w}^R = [0\ 0\ 1\ 0]^T$, and set $\mathbf{x}_1, \mathbf{x}_3$ equal to the left and right stereo images respectively. With this simplification, problem (7) reduces to the linear least-square problem

$$\begin{aligned} \min_{\mathbf{A},\mathbf{X}} \ & ||s^L\mathbf{t}^L + b^L\mathbf{1} - \mathbf{x}_1||_2^2 + ||s^R\mathbf{t}^R + b^R\mathbf{1} - \mathbf{x}_3||_2^2 + \\ & \lambda||s\mathbf{t}_0 + b\mathbf{1} - \mathbf{x}_1 - \mathbf{x}_2 - \mathbf{x}_3 - \mathbf{x}_4||_2^2 \\ \text{s.t. } & 0 \le \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4 \le 1;\ s, s^L, s^R \ge 0, \\ & s^L\mathbf{t}^L + b^L\mathbf{1} = \mathbf{x}_1,\ s^R\mathbf{t}^R + b^R\mathbf{1} = \mathbf{x}_3. \end{aligned} \tag{8}$$

where the Lagrangian multiplier λ is used to balance the distortion between 2-D and 3-D views. The two equal constraints have a feasible solution:

$$\begin{cases} \mathbf{x}_1 = \mathbf{t}^L,\ \mathbf{x}_3 = \mathbf{t}^R, \\ s^L = s^R = 1,\ b^L = b^R = 0. \end{cases} \tag{9}$$

This solution infers $\mathbf{y}^L = \mathbf{t}^L/4$ and $\mathbf{y}^R = \mathbf{t}^R/4$, indicating that $\mathbf{y}^L$ and $\mathbf{y}^R$ are not distorted but the LIRs of both views shrink to a value of 0.25. This partial solution for the 3-D view is exactly the same solution as in previous TPVM approaches, although we interpret in the newly defined light intensity domain. This solution leads to a high quality 3-D view without distortion and the reconstruction error only exists in the 2-D view.

Since the scaling factors and offsets are fixed for the 3-D stereoscopic views, only s and b is adjustable for controlling the relative brightness between 2-D and 3-D views. As mentioned in previous section, s and b can be treated as user-inputs in order to introduce more flexibility. Therefore problem (8) is reduced to the following least-square

$$\begin{aligned} \min_{\mathbf{x}_2,\mathbf{x}_4} \ & ||s\mathbf{t}_0 - \mathbf{t}^L - \mathbf{t}^R + b\mathbf{1} - \mathbf{x}_2 - \mathbf{x}_4||_2^2 \\ \text{s.t. } & 0 \le \mathbf{x}_2, \mathbf{x}_4 \le 1. \end{aligned} \tag{10}$$

Although this problem is tractable by a generic solver, a closed-form solution is preferred for real-time applications. Let $\mathbf{x} = \mathbf{x}_2 + \mathbf{x}_4$, (10) is relaxed into

$$\min_{\mathbf{x}} ||s\mathbf{t}_0 - \mathbf{t}^L - \mathbf{t}^R + b\mathbf{1} - \mathbf{x}||_2^2, \quad \text{s.t. } 0 \le \mathbf{x} \le 2. \tag{11}$$

Note that for any feasible solution $\mathbf{x}$ to problem (11), a feasible solution to problem (10) can be obtained as $\mathbf{x}_2 = \mathbf{x}_4 = \mathbf{x}/2$; and vice versa. Let $\mathbf{p} = s\mathbf{t}_0 - \mathbf{t}^L - \mathbf{t}^R$, and let p_i and x_i denotes the corresponding entries in $\mathbf{p}$ and $\mathbf{x}$ for each pixel i respectively. Problem (11) can be written as a summation of individual sub-problems,

$$\min_{x_i} \sum_i (p_i + b - x_i)^2, \quad \text{s.t. } 0 \le x_i \le 2, \forall i.$$

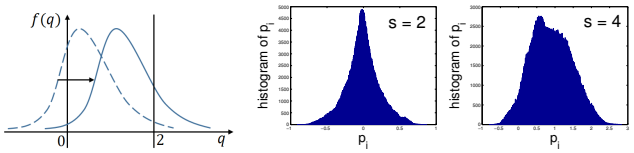
The optimum is achieved when each sub-problem is optimized. Obviously the optimal solution to each sub-problem is $x_i = Clip(p_i + b)$, where $Clip(\cdot)$ is the clipping function bounding the value to the range of $[0, 2]$. Therefore the optimal solution to problem (11) is given by

$$\mathbf{x} = Clip(\mathbf{p} + b\mathbf{1}). \tag{12}$$

With (12) we can derive the optimal solution of the improved TPVM system as follows given a distortion-free 3-D view

$$\begin{cases} \mathbf{x}_1 = \mathbf{t}^L,\ \mathbf{x}_3 = \mathbf{t}^R, \\ \mathbf{x}_2 = \mathbf{x}_4 = Clip(s\mathbf{t}_0 - \mathbf{t}^L - \mathbf{t}^R + b\mathbf{1})/2. \end{cases} \tag{13}$$

In this solution, the user-inputs of s^L, s^R, b^L and b^R are fixed in (9). Only s and b are adjustable in order to balance the distortion and brightness of the 2-D view.



(a) (b)

Fig. 1. (a) The distribution function of q; (b) Histogram of p_i for image `Cones` [9] under different s.

D. Derivation of Optimal Light Intensity Offset

The original TPVM system (4) is a special case when $b = 0$, therefore the solution (13) to our problem also includes the solution to the original TPVM-based approach. Given s, the original TPVM-based approach provides a unique solution when $b = 0$ while our approach provides multiple solutions corresponding to different b values. In order for comparison with the original TPVM-based approach and also for derivation of the best performance of our method, we next derive the optimal b that minimize problem (11).

Minimizing problem (11) is equivalent to minimize the clipping error in (12). Taking each p_i as an observation of a random variable q, the histogram of p_i indicates the probability distribution function $f(q)$ of q. Examples of histogram of p_i are shown in Fig. 1b. As illustrated in Fig. 1a, finding the minimum clipping distortion can be interpreted as shifting $f(q)$ along the horizontal axis and minimizing the energy lying outside the range of $[0, 2]$, or equivalently, maximizing the energy lying within $[0, 2]$. So the optimal b can be obtained by

$$b^\star = \arg\max_b \int_0^2 f^2(q - b)\, \mathrm{d}q. \tag{14}$$

Assuming $f(q)$ is continuous differentiable, $b^\star$ can be obtained by taking the derivative of the objective function and setting it to zero. Due to page limitation, we did not show the derivation process. It can be verified that $b^\star$ is obtained when the following equation is satisfied,

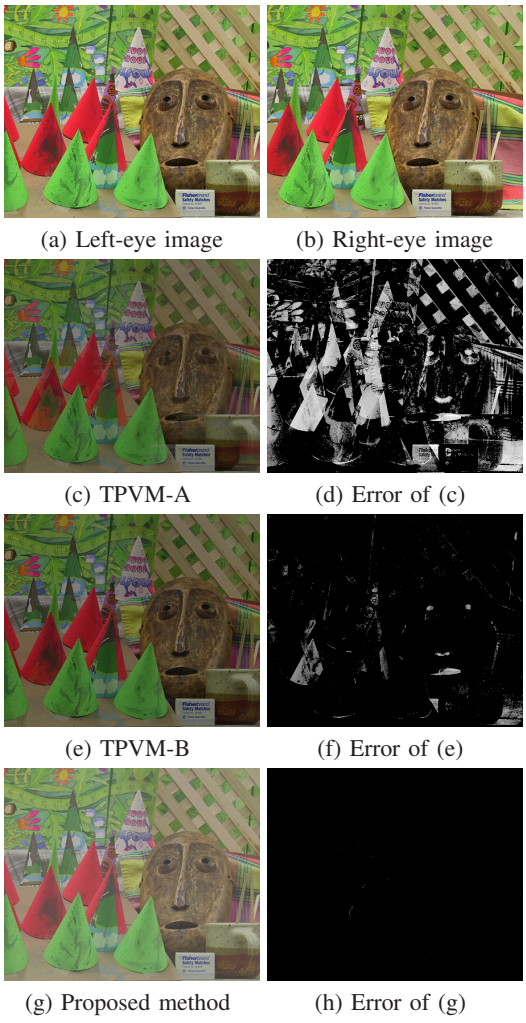
$$f(2 - b^\star) = f(-b^\star). \tag{15}$$

Once $f(q)$ is known, $b^\star$ can be directly computed from (15). $f(q)$ can be estimated from the histogram of p_i, however it is computational expensive. Note that Fig. 1b indicates that $f(q)$ approximately follows a symmetric unimodal distribution. In symmetric unimodal distribution, simply $b^\star = 1 - c$, where c is the location of the peak which is also the mean value. Therefore, assuming symmetric unimodal distribution, it is derived that

$$b^\star \approx 1 - E[q] = 1 - \frac{1}{N}\sum_i p_i = 1 - \frac{1}{N}\mathbf{1}^T\mathbf{p}. \tag{16}$$

IV. Experimental Results

The proposed method is evaluated using the stereo datasets [9] [10] [11]. A software simulated version of TPVM system is applied for verification. Without loss of generality, the naked eye 2-D view is set to be equal to the left stereo view, i.e. $\mathbf{t}_0 = \mathbf{t}^L$. For the sake of fair comparison with previous TPVM approaches, we first translate the TPVM system in [5] (TPVM-A) and [7] (TPVM-B) into the improved TPVM systems by considering the effect of light intensity. In [5] the light intensity is incorporated directly by adding a scaling factor as shown in (3). While in [7] the dynamic range of the display is first normalized when divided by the maximal backlight intensity, and the same scaling factor is added afterwards. When translated into the improved TPVM formulation, the distortion is no longer measured as the image difference between target and perceived views, but the



(a) Left-eye image (b) Right-eye image

(c) TPVM-A (d) Error of (c)

(e) TPVM-B (f) Error of (e)

(g) Proposed method (h) Error of (g)

Fig. 2. Visual results of the 2-D view for `Cones` [9], LIR=0.625. The log-scaled error maps are shown in the right column.

light intensity difference between them. Therefore PSNR is measured in the light intensity domain, and the ground-truth is set to be the light intensities of the target views, namely, $\mathbf{y}_0$, $\mathbf{y}^L$ and $\mathbf{y}^R$.

Since all methods can display the 3-D view perfectly without error, only the distortion of the 2-D view is compared. For fair comparison, the light intensity range (LIR) between the maximal and minimal possible light intensity values for the 2-D view is set to be equal for different methods. LIR is controlled by the scaling factor s. As mentioned before, our improved TPVM system introduces an additional offset b for the light intensity adjustment, while the original TPVM system is the special case when $b = 0$. Fig. 3 shows the PSNR of the 2-D view as a function of b for different methods when LIR=0.625. It verifies that the original TPVM method (TPVM-A) is the special case when $b = 0$. As b increases, the quality of the 2-D view first increases to a peak value then decreases. The offset corresponding to the peak value is $b^\star$ that achieves the minimum distortion. We also notice that TPVM-B provides a better PSNR when $b = 0$. This is at the expense of the brightness loss of the 3-D view, i.e. the LIR of the 3-D view is lower. However our method does not affect the 3-D view, and with a larger b our method achieves better quality than TPVM-B. The corresponding visual results are shown in Fig. 2 where the $b^\star$ is chosen for the proposed method. It is shown that our method significantly reduces the ghosting artifacts

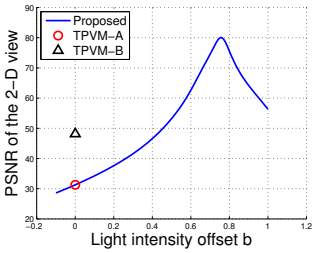
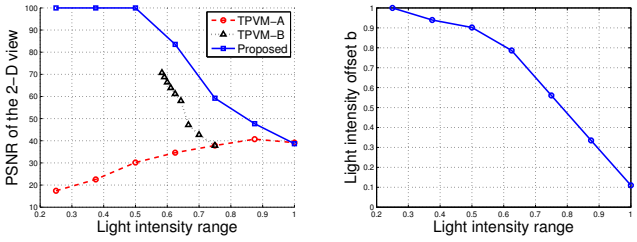


Fig. 3. PSNR of the 2-D view as a function of b for `Cones` [9], LIR=0.625.



(a) (b)

Fig. 4. (a) PSNR of the 2-D view versus LIR for various methods; (b) The relationship between light intensity offset b and LIR.

with almost no visible error compared with other methods. The ghosting artifacts occur in areas where the right view is brighter than the left view, making it impossible to find a compensated pixel to attenuate the brightness of the right view pixel in order to recover a darker pixel in the 2-D view (which is identical to the left view in our settings). The offset b has an effective resistance against the ghosting artifacts, which levels up the global light intensities of the 2-D view and increases the possibilities of finding a compensated pixel for the dark pixels.

The average PSNR curves of the 2-D view under different LIR are shown in Fig. 4a, where $b^\star$ is used in the proposed method. Note that PSNR values more than 99.99dB are truncated for numerical stability. The proposed method is superior to previous methods in all LIR values especially in low LIR areas, where it allows more freedom for choosing $b^\star$ and a smaller distortion can be achieved.

While $b^\star$ provides high quality 2-D view with minimum distortion, it increases the global light intensities, resulting in a brighter 2-D view (Fig. 2g) than previous methods. If $b^\star$ is too large, the viewing experience might be affected. Fig. 4b plots the relationship between $b^\star$ and LIR. For large LIR, $b^\star$ is small indicating a small global brightness increment which is good for viewing. When LIR decreases, $b^\star$ increases, resulting in a large global brightness increment which might affect the viewing experience. Since b can be viewed as a user-input parameter, considering the relationship between b and PSNR in Fig. 3, a user can actually lower down b from $b^\star$ for a better viewing at the cost of quality degradation. Since the quality at low LIR areas is super high, there is much space for lowing down b so that appropriate quality degradation still provides high quality 2-D view.

V. Conclusion

We present an improved TPVM system for backward-compatible stereoscopic display. We consider the light perception process of human eyes and build a new relationship between light intensities and pixel intensities. The proposed method significantly reduces the ghosting artifacts and achieves better performance than previous approaches in terms of the 2-D view quality. A closed-form solution is also derived for real-time applications.

REFERENCES

[1] J. Su, M. Johnson, D. Miller, G. Wernig, and D. Burckhard, "3-d stereo glasses," Sep. 22 2009, uS Patent D600,738. [Online]. Available: https://www.google.com/patents/USD600738

[2] J. Harrold and G. Woodgate, "Lcd spatial light modulator as electronic parallax barrier," Jun. 23 2004, eP Patent 0,833,183. [Online]. Available: https://www.google.com/patents/EP0833183B1?cl=en

[3] A. J. Woods, "Crosstalk in stereoscopic displays: a review," *Journal of Electronic Imaging*, vol. 21, no. 4, pp. 040 902–040 902, 2012. [Online]. Available: http://dx.doi.org/10.1117/1.JEI.21.4.040902

[4] P. Didyk, T. Ritschel, E. Eisemann, K. Myszkowski, and H.-P. Seidel, "A perceptual model for disparity," *ACM Trans. Graph.*, vol. 30, no. 4, pp. 96:1–96:10, Jul. 2011. [Online]. Available: http://doi.acm.org/10.1145/2010324.1964991

[5] X. Wu and G. Zhai, "Backward compatible stereoscopic displays via temporal psychovisual modulation," in *SIGGRAPH Asia 2012 Emerging Technologies*, ser. SA '12. New York, NY, USA: ACM, 2012, pp. 4:1–4:4. [Online]. Available: http://doi.acm.org/10.1145/2407707.2407711

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

[7] L. Jiao, X. Shu, and X. Wu, "Led backlight adjustment for backward-compatible stereoscopic display," *Signal Processing Letters, IEEE*, vol. 20, no. 12, pp. 1203–1206, Dec 2013.

[8] M. Kalloniatis and C. Luu, "Temporal resolution," Webvision: The Organization of the Retina and Visual System, Jun 2007. [Online]. Available: http://www.ncbi.nlm.nih.gov/books/NBK11559/

[9] D. Scharstein and R. Szeliski, "High-accuracy stereo depth maps using structured light," in *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, vol. 1, June 2003, pp. I–195–I–202 vol.1.

[10] D. Scharstein and C. Pal, "Learning conditional random fields for stereo," in *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on*, June 2007, pp. 1–8.

[11] H. Hirschmuller and D. Scharstein, "Evaluation of cost functions for stereo matching," in *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on*, June 2007, pp. 1–8.