Seamless Visual Sharing with Color Vision Deficiencies

Wuyao Shen   Xiangyu Mao   Xinghong Hu   Tien-Tsin Wong
The Chinese University of Hong Kong*
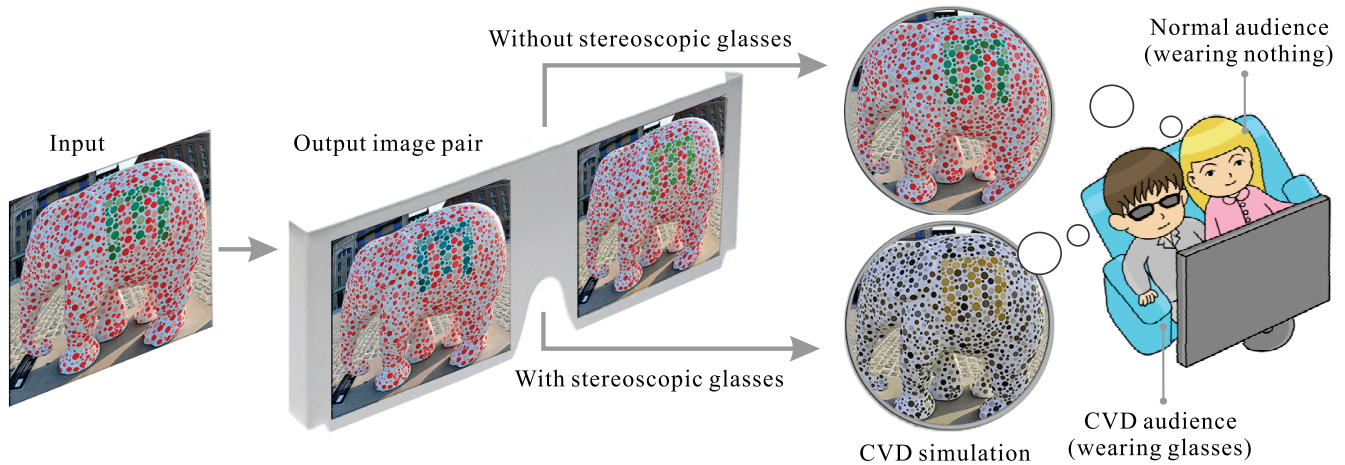Shenzhen Research Institute, The Chinese University of Hong Kong

Figure 1: *By allocating two visual experiences of ordinary stereoscopic display to different audiences, we firstly accomplish seamless visual sharing between color vision deficiencies (or colorblindness) and normal-vision audiences. Our system synthesizes an image pair from an input image. By wearing the stereoscopic glasses, CVD audiences can identify the original indistinguishable colors. Without wearing the stereoscopic glasses, normal-vision audiences are presented with original colors simultaneously without being aware of the change of visual content.*

Abstract

Approximately 250 million people suffer from color vision deficiency (CVD). They can hardly share the same visual content with normal-vision audiences. In this paper, we propose the first system that allows CVD and normal-vision audiences to share the same visual content simultaneously. The key that we can achieve this is because the ordinary stereoscopic display (non-autostereoscopic ones) offers users two visual experiences (with and without wearing stereoscopic glasses). By allocating one experience to CVD audiences and one to normal-vision audiences, we allow them to share. The core problem is to synthesize an image pair, that when they are presented binocularly, CVD audiences can distinguish the originally indistinguishable colors; and when it is in monocular presentation, normal-vision audiences cannot distinguish its difference from the original image. We solve the image-pair recoloring problem by optimizing an objective function that minimizes the color deviation for normal-vision audiences, and maximizes the color distinguishability and binocular fusibility for CVD audiences. Our method is extensively evaluated via multiple quantitative experiments and user studies. Convincing results are obtained in all our test cases.

*e-mail:{wyshen,xymao,huxh,ttwong}@cse.cuhk.edu.hk

Keywords: CVD, daltonization, binocular vision

Concepts: ●**Computing methodologies** → **3D imaging;** *Image processing;*

1 Introduction

People suffering from the color blindness or color vision deficiency (CVD) are usually not able to clearly discriminate certain colors, leading to misunderstanding, inconvenience, or even danger in daily life. There are approximately 250 million CVD people worldwide [Wong 2011]. Ninety-nine percent of the CVD people are inborn and there is no medical cure, unfortunately. It has been found that CVD is usually hereditary [Sharpe et al. 1999]. In other words, it is common for a family to have mostly normal-vision members and a very few CVD members. However, existing digital visual entertainment systems including computer games, movies, or TV usually do not care or are not aware of the CVD audiences. This leads to an inconvenient scenario that normal-vision and CVD family members may not be able to share the same visual content. Such sharing inconvenience is even more common in working environment.

One way to discriminate colors in daily life is to wear the tinted glasses. They amplify the red-green color difference while filter away the others [Sheedy and Stocker 1984]. However, the visual experience of wearing tinted glasses is not comfortable, and they may potentially impair the depth perception of users [Hartenbaum and Stack 1997]. For digital visual content presented via a display, recoloring techniques [Rasche et al. 2005a; Jefferson and Harvey 2007; Huang et al. 2007] can be applied. Several approaches have been proposed to modify the colors in image to maximize the color distinguishability for CVD individuals. Even though some [Huang

et al. 2007; Kuhn et al. 2008a; Huang et al. 2009] try to minimize the deviation from the originals, the change in color remains significant, and hence the resultant images are no longer shareable with normal-vision audiences. Recent works [Sajadi et al. 2013; Chua et al. 2015] begin to consider the visual sharing need. However, the texture pattern intentionally introduced by Sajadi et al. [2013] is noticeable to normal vision audiences. Colorbless method [2015] introduces obvious "false contour" artifact (to both CVD and normal-vision audiences) when applied to natural images (Fig. 12). Introduction of additional contour (pattern or false contour) should be avoided as audiences cannot differentiate such undesirable content from the original image content.

In this paper, we propose a method to synthesize the color-discriminable and undesirable-content free images that can be shared with CVD audiences, without the normal-vision audiences noticing the change. This sounds unachievable, but with an extra display channel on popular and low-cost stereoscopic display, this becomes feasible. Human binocular vision system can fuse the visual content from both eyes into a single percept, via a complex nonlinear neurophysiological process [MacMillan et al. 2007; Baker et al. 2007]. In other words, the visual perception of the binocular presentation of a dichoptic image pair (two views are different in color) is different from the visual perception of a monocular presentation of the blending of two views. These two different visual experiences can be allocated to the CVD and the normal-vision audiences, separately, so that two types of audiences can share the same visual content. Our configuration is illustrated in Fig. 1. While the normal-vision audiences can simply watch the 2D images via a stereoscopic display as usual (without wearing any glasses), CVD individual has to wear shutter glasses or polarized glasses to view the image pair that maximizes the color distinguishability.

To make such configuration feasible, the input image is fed to our system to synthesize an image pair that can maximize the color distinguishability for CVD audiences when displayed binocularly, and at the same time, the average of the left and right images is equivalent to the input. We formulate the image pair synthesis problem as an optimization problem that simultaneously maximizes the color distinguishability of the image pair for CVD audiences and minimizes the deviation of the left+right average from the input. However, there remains a challenge that the synthesized image pair may not be fusible (binocular rivalry), when two views are too different from each other. Yang et al. [2012] proposed a metric to predict the binocular fusibility. We make an attempt to refine and apply this metric to suit CVD audiences, during the optimization. With our approach, we are the first to achieve "seamless" visual sharing between CVD and normal-vision audiences, not just on toy-case images, but also challenging natural images that contain excessive amount of confusing color pairs. In addition, we propose a novel CVD calibration method that allows us to tailormade the image pair synthesis for different CVD individuals according to their type and severity of color blindness. Our results of various types of images are validated via multiple quantitative experiments and user studies. Our contributions can be summarized as followed.

- While there exist attempts to achieve visual sharing between the CVD and the normal audiences, we are the first to achieve "seamless" visual sharing without the normal vision audiences being aware of the change of visual content. We are the first to demonstrate the seamless visual sharing of challenging natural images, with many confusing colors simultaneously existing in the same image.

- We are the first to utilize the computer-controlled binocular display systems to tailor-made the color discrimination solution for each CVD individual.

- We also propose a novel calibration method to measure the severity of different CVD individuals.

2   Related Work

Humans perceive color via three kinds of cone cells in our retina, with their spectral sensitivities peak at different wavelengths. However for the CVD individuals, one type of cone cells mutates and its spectral sensitivity peak shifts towards that of another type of cone cells. Detailed description of this phenomenon can be found in [Machado et al. 2009].

To study CVD, several models of CVD have been proposed to simulate the visual experiences of the CVD individuals. Note that, there are different types of CVD, including protanopia, deuteranopia and tritanopia. Each type of CVD can have different levels of severity. Based on the experiments of unilateral dichromats [Graham and Hsia 1959; Judd 1948], Brettel et al. [1997] and Meyer and Greenberg [1988] calculated the projection of a given color for different types of dichromacy. According to the two-stage model [Judd 1966], Machado et al. [2009] proposed a unified model that can simulate different types of dicromats with different levels of severity. In our work, we adopt this simulation model.

Optical Approach   With these simulation models, several approaches have been proposed to assist the people with CVD in distinguishing colors. The existing approaches can be roughly classified into two main categories: optical approach and recoloring approach. Optical approach [Hovis 1997] requires the user to wear a physical tinted lens to filter certain colors in order to amplify the color difference. However, the tinted lens may unnecessarily filter away color (and hence information) that CVD individuals can well distinguish even without the lens.

Recoloring   On the other hand, recoloring approach is mainly for displaying digital visual content. The basic idea is to globally or locally recolor the visual content in order to avoid any color composition that CVD individuals cannot distinguish. Lau et al. [2011] redistributed colors in target color space by optimizing the distinguishability of CVD audiences. However, their solution is limited to colors existing in the projected original image. Laccarino et al. [2006] recolored web pages by increasing the contrast and lightness according to the customized parameters. Jefferson et al. [2006; 2007] developed a user interface to allow users to specify the type of CVD they belong to and adjusts the parameter to distinguish colors. However, since these systems require users to explicitly control the parameters, the quality of recoloring result highly depends on the skill of users. Based on the Kondo's model [1990], Ichikawa et al. [2004] proposed a method to automatically recolor the images to compensate the loss of color discrimination of CVD audiences. Rasche et al. [2005a; 2005b], Wakita et al. [2005] and Machado and Oliveira [2010] preserved the contrast between all pairs of colors and maintains luminance consistency via optimization. The above automatic methods only consider the CVD audiences during the recoloring, the colors in the resultant images may largely deviate from the original, and hence the results are hardly shareable with the normal-vision audiences.

Visual Shareability   Recent works attempt to achieve such shareability between CVD and normal audiences. Huang et al. [2007; 2009] and Kuhn [2008a] minimized the deviation of colors from the original. This relieves the problem but does not really solve it, because the color deviation remains large to avoid confusing the CVD audiences. Hung and Hiramatsu [2013], and Sajadi et al. [2013] proposed a visualization method to overlay patterns over the color
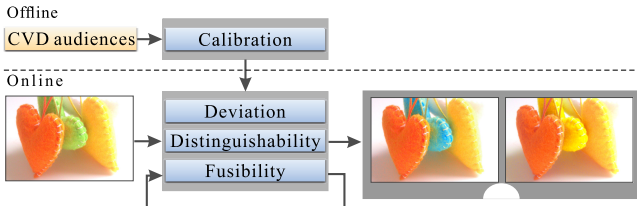
Figure 2: *System overview.*

(a) initial color pair   (b) just-distinguishable color pair

Figure 3: *Calibration interface. (a) The initial color pair. (b) Just-distinguishable color pair.*

regions to assist CVD audiences. However, the introduction of additional contours (texture in this case) confuses both CVD and normal audiences as audiences no longer can tell the introduced pattern belongs to the original image or serves for differentiating colors. Color discrimination for CVD audiences and preservation of the same visual content for normal vision audiences is always a dilemma, for a single-image configuration. With an extra display channel of stereoscopic devices, we propose to synthesize image pairs that allow CVD audiences to distinguish colors (by wearing stereoscopic glasses), without normal-vision audiences (without wearing any glasses) aware of the change in image. Independently, Chua et al. [2015] proposed a guideline to highlight confusion color regions for CVD audiences using binocular display. Their method is based on luster effect [Howard 2002] and modifies the luminance channel alone, to provide three comfortable and distinguisable luster levels only. Their application to natural images is questionable as there can be much more confusion colors simultaneously exist in the same image. False contours frequently appear when their method is applied on natural images (Fig. 12). As their method is based on the luster effect, they utilize the non-fusibility of binocular vision to discriminate colors. In contrast, our method does not rely on the luster effect and we want to avoid the discomfort caused by the non-fusibility.

So far, most previous methods are not tailormade solutions, as they do not measure the CVD severity of each individual. In contrast, our solution can be precisely calibrated for each CVD individual, regardless of his/her severity and his/her type of CVD.

3   Overview

Our system overview is illustrated in Fig. 2. It consists of an offline and an online phases. The offline calibration is a one-off step, and should be performed once for each CVD audience (Section 4). During the online recoloring phase (Section 5), our system simultaneously preserves contrasts in two domains, one for CVD audiences and one for normal-vision audiences. Given an input image, our system synthesizes an image pair that maximizes the color distinguishability for the characterized CVD audiences when displayed binocularly, and at the same time, minimizes the deviation of the average left and right images from the input, for normal-vision audiences. As CVD audiences are fed with image pair, we enclose the image pair with a gray boundary (as the one in Fig. 2) to indicate this, from now on.

We formulate our recoloring as an optimization problem. It adjusts the colors to optimize for an objective function that considers three factors: minimization of the color deviation between the input and the blending of left and right images (*deviation term*); maximization of the chrominance distinguishability for CVD audiences (*distinguishability term*), and the binocular fusibility for CVD audiences (*fusibility term*). We first present a global color mapping approach, and then we extend it to a local method, so as to enlarge the solution space and preserve fidelity.
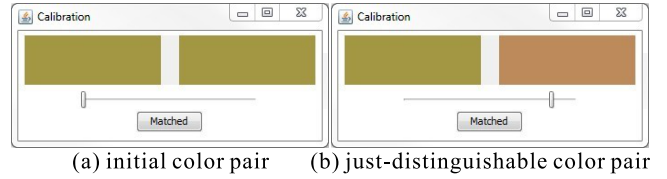
4   Calibration

In this paper, we adopt the CVD simulation model proposed by Machado and Oliveira [2009] due to its nice feature of luminance preservation that facilitates our subsequent computation. The color deficiency can be modeled as a color projection,

$$
\begin{bmatrix} X_{\text{CVD}} \\ Y_{\text{CVD}} \\ Z_{\text{CVD}} \end{bmatrix} = T \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix},
\tag{1}
$$

where T is a 3×3 color projection matrix; X, Y, Z are the color values; while X_{CVD}, Y_{CVD}, Z_{CVD} are the color values perceived by the CVD individual. In other words, the type and severity of a CVD individual can be sufficiently characterized by T. Hence, the goal of our calibration is to determine T for each CVD individual.

To do so, the target CVD individual takes part in a simple experiment modified from Farnsworth-Munsell 100 hue test [Farnsworth 1957]. The original Farnsworth-Munsell 100 hue test is time-consuming and tedious. Even worse, the user can learn from the past experience which harms the accuracy of the test. Fig. 3 shows the interface of our test. Each time it shows a pair of color patches. Initially, the two patches have the same color C_A. By moving the scroll bar, users can move the color of right patch away from its initial color, and towards another color C_B. The color pair (C_A, C_B) is randomly chosen from Farnsworth-Munsell 100 hue test. Users are asked to gradually move the scroll bar until he/she just notices the two patches are different. C_1 (the left color) and C_{JD} (the right color) form the just-distinguishable color pair. In other words, we can slightly move C_{JD} back towards C_1 to construct a confusing color pair (C_1, C_2), where $C_2 = \eta C_1 + (1 - \eta)C_{\text{JD}}$ and η=0.0001. The target CVD individual cannot distinguish C_2 from C_1, and our system records this confusing color pair for determining T. This process repeats, until the system collects sufficient number of confusing color pairs.

For each confusing color pair (C_1, C_2), CVD individual cannot distinguish them. Mathematically, it means,

$$
T \cdot \begin{bmatrix} R_1 \\ G_1 \\ B_1 \end{bmatrix} = T \cdot \begin{bmatrix} R_2 \\ G_2 \\ B_2 \end{bmatrix},
\tag{2}
$$

where (R_1, G_1, B_1), and (R_2, G_2, B_2) are RGB values of C_1 and C_2, respectively. This means each confusing color pair gives us a set of 3 equations. With 9 unknowns in T, 3 confusing pairs (9 equations) are sufficient for determining T. In practices, we collect 10 confusing color pairs for better accuracy. In addition, as stated in [Machado et al. 2009], each element in T has its own range on values (Section S5 of the supplement), and the sum of each row of T must be equal to 1. Together with these constraints and collected confusing pairs, we solve for a more accurate T in a least-square manner.

(a) input | (b) without deviation term | (c) with deviation term

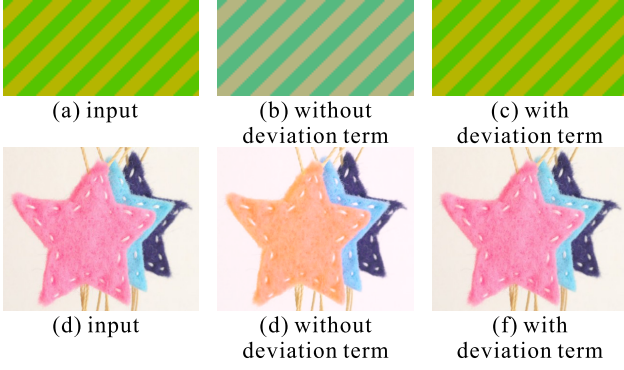(d) input | (d) without deviation term | (f) with deviation term

Figure 4: *Effectiveness of deviation term. (a)&(d): Input images. (b)&(e): The average of resultant left and right images without the deviation term. (c)&(f): With the deviation term, the difference between inputs and the average becomes very small. The left and right images before blending can be found in Fig. S3 & S5 of the supplement.*

(a) input gradient | (b) CVD simu. of (a) | (c) result gradient in CVD simulation

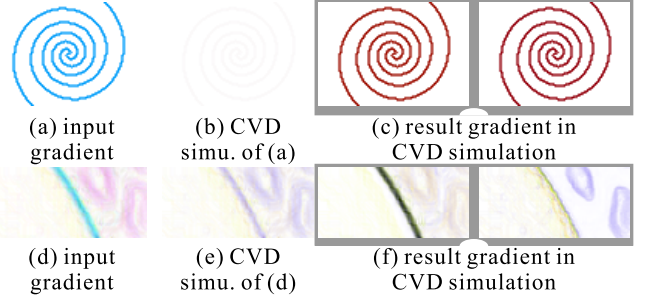(d) input gradient | (e) CVD simu. of (d) | (f) result gradient in CVD simulation

Figure 5: *Effectiveness of distinguishability term. (a)&(d): Gradient magnitudes of the input images (normal vision). (b)&(e): Gradient magnitudes of input as in CVD simulation. The contrast of edges becomes very weak (especially in the swirl example). (c)&(f): Gradient magnitudes of our resultant image pairs as in CVD simulation. Edge structures in both eyes are well preserved in both cases. The original images can be found in Fig. S3 & S8 of the supplement.*

5 Recoloring as Optimization

Synthesizing color-distinguishable images for CVD audiences is a problem to recolor images in a lower-dimensional gamut with the goal of preserving contrast. In this regard, it is a dimension reduction problem. Several methods [Gooch et al. 2005; Neumann et al. 2007; Grundland and Dodgson 2007; Kuhn et al. 2008b; Lu et al. 2012] have been proposed for the color-to-gray problem (a related, but not the same as our problem). In particular, Rasche et al. [2005b] synthesized images for CVD audiences using a constrained multidimensional scaling technique. However, existing methods only synthesize a single output image. They do not consider the binocular perception as in our case and therefore not applicable.

To synthesize the image pair for CVD audiences, we formulate this image synthesis as an optimization problem that optimizes an objective function. In our visual sharing application, we need to minimize the color deviation of the blending of left and right images from the original input, so that normal-vision audiences without wearing any stereo-glasses are not aware of the color differences. Simultaneously, we want to maximize the color distinguishability and binocular fusibility of the left and right images, when the CVD audiences is presented with the dichoptic image pair. These three requirements correspond to the three energy terms in the objective function, the deviation (E_d), the distinguishability (E_c), and the fusibility terms (E_f). The whole objective function is defined as,

$$\min \{\lambda_1 E_d + \lambda_2 E_c + E_f\}. \quad (3)$$

All terms E_d, E_c, and E_f are functions of the input image I, f_L and f_R, where f_L and f_R are the mapping functions of I to the left $f_L(I)$ and the right $f_R(I)$ images, respectively. λ_1 and λ_2 are weights.

Image Pair Synthesis Model We model the mapping f_L and f_R as a finite multivariate polynomial function [Lu et al. 2012]. That is, we Taylor-expand a continuous function f as

$$f(I) = \sum_{i=0}^{\infty} \omega_i \Phi_i(I), \quad (4)$$

where $\Phi_i(I)$ is the finite multivariate polynomial function. We want to obtain the optimum mapping f_L and f_R by seeking

their Taylor parameter sets $\Lambda_L = \{\omega_1^l, \omega_2^l, \ldots\}$ and $\Lambda_R = \{\omega_1^r, \omega_2^r, \ldots\}$. High-order parameters are close to zero as high-order Π_m's correspond to high-frequency information.

In our application, we keep only the first nine components by setting $\omega_i = 0$ for $i = m+1, \ldots, \infty$, where $m = 2$. Each pixel of the original image I contains three color channels, c_1, c_2 and c_3 in RGB color space. Each color channel of a pixel in the synthesized image is formulated as the linear combination $\Phi_i(I)$ of 9 components, $\{c_1, c_2, c_3, c_1c_2, c_1c_3, c_2c_3, c_1^2, c_2^2, c_3^2\}$. With 3 color channels and 2 views, these add up to 54 weights ω's to estimate. As the same set of ω's is applied to the whole image, changing these parameters only causes smooth color change across the images. Hence, false contour artifacts as in other methods never appear in ours.

Deviation Term The deviation term E_d is introduced for normal-vision audiences, as they look at the stereoscopic display (non-autostereoscopic ones) without wearing any stereoscopic glasses. Effectively, they observe the linear blending (average in our case) of the left and the right images. Hence, this deviation term is introduced to minimize the per-pixel color deviation of the blending and the input images. It is simply defined as,

$$E_d = \frac{1}{N} \sum_N \|G(p_i^L, p_i^R) - p_i^I\|, \quad (5)$$

where N is the total number of pixels in the input I and p_i^I is the pixel value at position i, and p_i^L and p_i^R are the corresponding pixel values in the left and right images synthesized with the parameter sets Λ_L and Λ_R, respectively. Function G is the linear blending function which depends on the stereoscopic display device. In our case, G is the average of p_i^L and p_i^R.

Fig. 4 demonstrates the significance of this deviation term. Without the deviation term, the blending of left and right images (Fig. 4(b)&(e)) may have an observable color difference when compared to the input (Fig. 4(a)&(d)). In contrast, by introducing the deviation term, we can effectively suppress the difference in Fig. 4(c)&(f) (the corresponding left and right images can be found in Fig. S3 & S5 of the supplement). This means that the normal-vision audiences may not be aware of the difference, and hence accomplishes half of our visual-sharing goal.

Distinguishability Term The color distinguishability term E_c and the binocular fusibility term E_f (explained shortly) are tailored

(a) without fusibility term   (b) with fusibility term   (c) without fusibility term   (d) with fusibility term
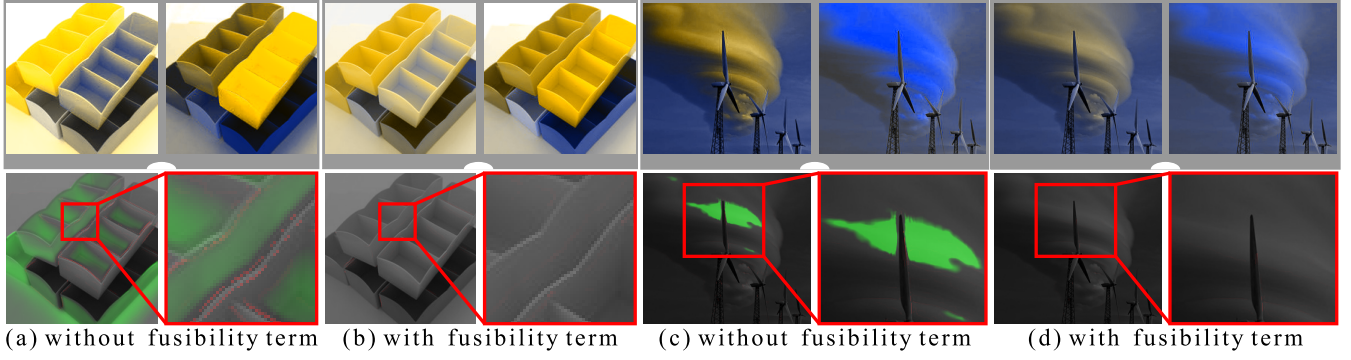
Figure 6: *Effectiveness of fusibility term. Top row of (a)&(c) show the resultant image pairs without the fusibility term. The contrast and contour differences are color-coded in green and red, respectively, in the bottom row of (a)&(c). After introducing the fusibility term, the discomfort zones are significantly reduced in (b)&(d).*

for CVD audiences as the visual content is presented to them in a dichoptic manner via the stereoscopic display device. The distinguishability term aims at preserving the structural information in the input during the synthesis of the image pair. Note that when the image pair is perceived by the CVD audiences, we need to preserve the contrast in the simulated color space of CVD audiences. This is where we utilize the personalized projection matrix T. The term is defined as,

$$E_c = -\frac{1}{N}\sum_N (S(T \cdot p_i^L, p_i^I) + S(T \cdot p_i^R, p_i^I)), \quad (6)$$

where the projection matrix T projects a pixel color from the color space of normal-vision audiences to the simulated color space of the target CVD audiences. Function $S(x, y)$ measures the structural similarity of x and y. Here, we adopt from the structure and contrast terms of SSIM [Wang et al. 2004] and defined it as,

$$S(x, y) = \frac{2\sigma_{xy} + \epsilon}{\sigma_x^2 + \sigma_y^2 + \epsilon}, \quad (7)$$

where

$$\sigma_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \mu_x)(y_i - \mu_y), \quad (8)$$

σ and μ are the standard deviation and the mean within the local neighborhood, respectively; n is the number of pixels in the local neighborhood, and the small constant $\epsilon = 0.0009$ avoids the division-by-zero. The above design minimizes the structural difference between images perceived by normal-vision audiences and the CVD audiences. Fig. 5 demonstrates the effectiveness of our distinguishability term. To visualize the structural information, we plot the gradients of the input (in RGB color space) and the synthesized image pair (in simulated CVD color space). The contour of the swirl (Fig. 5(a)) becomes hard to see without the distinguishability term (Fig. 5(b)). On the other hand, with the distinguishability term, we can effectively preserve the structural information in the simulated CVD color space and hence, allow the CVD audiences to distinguish colors (Fig. 5(c)&(d)).

Fusibility Term   As the image pair is presented in a dichoptic manner, this leads to the risk of binocular rivalry. This phenomenon happens when the left and right images deviate too much (e.g. strong stimulus and too many unmatched contours) and our brain cannot fuse them into a stable single percept [Lei and Schor 1994]. Yang et al. [2012] proposed a metric called binocular visual comfort predictor (BVCP) to measure the potential of binocular rivalry.

We adopt the BVCP in designing our binocular fusibility term and define it as,

$$E_f = \frac{1}{N}\sum_N (B_{cf}(T, p_i^L, p_i^R) \\ + B_{rc}(T, p_i^L, p_i^R)), \quad (9)$$

where functions B_{cf} and B_{rc} measure the contour fusion discomfort and the regional contrast discomfort in [Yang et al. 2012], respectively. B_{cf} measures the discomfort caused by the fusion of the contour difference between the left and right images, while B_{rc} measures the discomfort caused by the fusion of the contrast difference between the left and right images. Different from Yang's design which takes thresholds on the discomfort values, we minimize these discomfort values in our optimization formulation. The two functions are defined as,

$$B_{rc}(T, p_i^L, p_i^R) = \frac{1}{|\Omega|}\|\sum_{p \in \Omega(p_i^L)} (T \cdot p) - \sum_{p \in \Omega(p_i^R)} (T \cdot p)\|, \quad (10)$$

and

$$B_{cf}(T, p_i^L, p_i^R) = \frac{1}{|\Omega|}\|\sum_{p \in \Omega(p_i^L)} \zeta(T \cdot p) - \sum_{p \in \Omega(p_i^R)} \zeta(T \cdot p)\|, \quad (11)$$

where $\Omega(p_i)$ is the local neighborhood centered at p_i which approximates the projected fusional area. We assume users are viewing the laptop display (as in all our experiments) and set the size of this neighborhood as 11×11 pixels. Function ζ returns the luminance perceived by CVD audiences. As we adopt the simulation model proposed by Machado and Oliveira [2009], the same projection matrix that projects RGB to opponent-color space can be applied to both the CVD simulated and original images to obtain the luminance, because their model ensures the luminance perceived by CVD audiences from the original image is the same as that perceived by normal-vision audiences from the simulated image. The detailed description on computing ζ can be found in Section S5 of the supplement. Fig. 6 compares the results synthesized with and without the fusibility term. Here, we color-code those pixels that have high contrast differences and contour differences in green and red, respectively. We can see that the introduction of fusibility term cleans up those discomfort zones (Fig. 6(b)).

Note that we perform most computation in *Lab* color space, as it is more sensible for perceptual measurement. But the personalized CVD simulation ($T \cdot p_i$) has to be first computed in RGB space, as the calibration is conducted in RGB space.
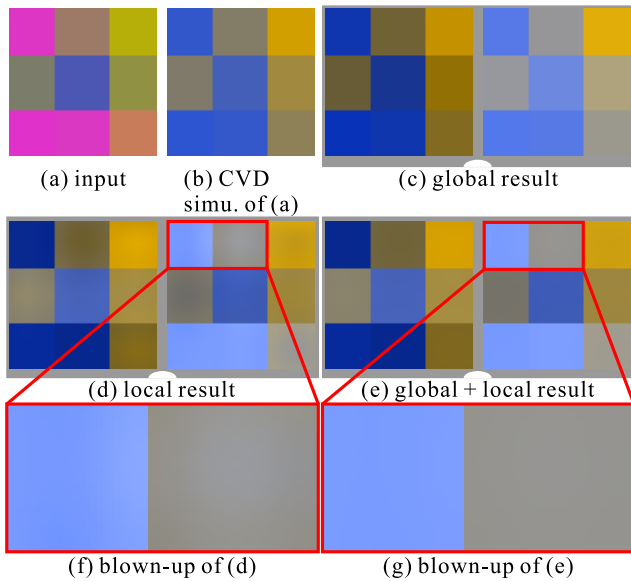
(a) input   (b) CVD simu. of (a)   (c) global result

(d) local result   (e) global + local result

(f) blown-up of (d)   (g) blown-up of (e)

Figure 7: *Local and global mappings. (a) The input image as in normal vision. (b) CVD simulation. (c) Globally mapped result as in CVD simulation. Fail to distinguish certain color region. (d) Locally mapped results as in CVD simulation. Halo artifact is observable. (e) Globally+locally mapped results as in CVD simulation. (f) local blown-up of (d). (g) local blown-up of (e). Now CVD audiences can distinguish all different color regions, with a significant reduction of halo artifact.*

(a) Input   (b) $\gamma = 0.8$

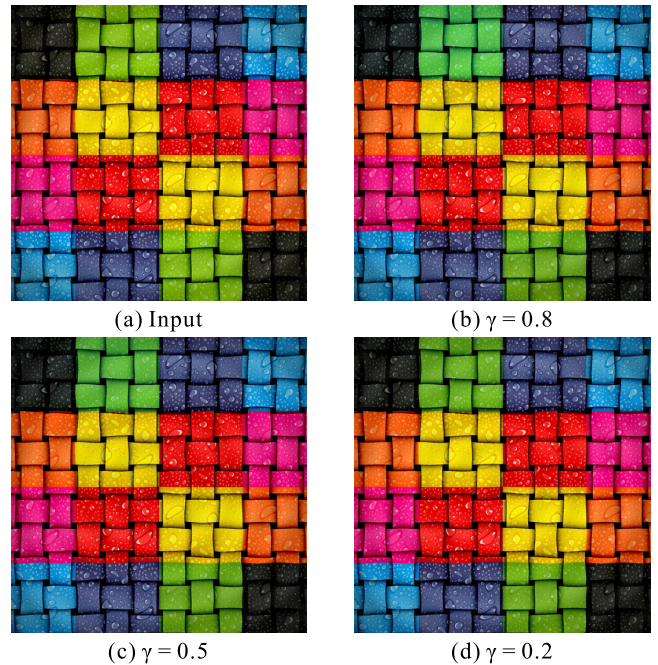(c) $\gamma = 0.5$ (d) $\gamma = 0.2$

Figure 8: *Tradeoff between global consistency and local variation. (a) The input image. (b) Result left image when $\gamma = 0.8$. (c) Result left image when $\gamma = 0.5$. (d) Result left image when $\gamma = 0.2$. The result right images can be found in Fig. S10 of the supplement.*

5.1 Local Mapping

So far, the mapping we discussed is a global one. It is not spatially varying. In very rare scenarios, the global constraint is too strong to produce a good solution. Fig. 7 shows the blown-up of one such example where the global mapping fails to synthesize an image pair with sufficient color distinguishability. It happens when the indistinguishable color patches scattered over the image and the color compositions are interconnected. The solution space can be significantly enlarged by relaxing this global constraint to a local one.

The proposed method can be easily extended to local mapping by applying the same method to local regions instead of the whole image. To do so, we randomly distribute t seeds over the image and apply our method to the region centered at each seed. The region size is set to $3\sqrt{hw/t}$, where h and w are the height and width of input image, respectively. This setting ensures each pixel is covered by 3 regions statistically. Obviously, we would like the local parameter set of the u-th local region, Λ_L^u and Λ_R^u, are similar to any regions overlapping with it. Hence, we can design a new smoothness term E_s as,

$$E_s = \frac{1}{K}\sum_{u,v}(\|\Lambda_L^u - \Lambda_L^v\| + \|\Lambda_R^u - \Lambda_R^v\|), \text{ s.t. } u, v \text{ overlap} \quad (12)$$

where u and v are regions that overlap and K is the total number of combination of region overlapping. Hence, our objective function is modified as follows,

$$\min\{\lambda_1 E_d + \lambda_2 E_c + \lambda_3 E_f + E_s\}. \quad (13)$$

where λ_i are weights. The overall mapping function f is modified to be the average of local mapping functions f_i. For each pixel location x, we overload the notation f and denote it as the overall mapping function at this location as,

$$f(x) = \frac{1}{k}\sum_i^k f_i(x), \quad (14)$$

Fig. 7(d) shows the recoloring result synthesized by this local approach. Halo artifact (the unnatural color change surrounding the strong boundary) appears as a result of color inconsistency from the global perspective. Besides, two isolated regions of the same color may result in color difference, using this local approach (e.g. the colors of the top and bottom green regions in Fig. 8(b) are drifted apart). To reduce the color inconsistency, we can combine the local and global mapping. The change is only on the construction of the overall mapping function f. For each pixel location, if it is overlapped by k regions, its overall mapping function is defined as

$$f(x) = (1 - \gamma)f'(x) + \gamma f''(x), \quad (15)$$

where f' and f'' are the global mapping and local mapping respectively. Parameter γ weights between the global and local mapping. Fig. 8 visualizes the effects of different γ values. Throughout all our experiments, we set γ=0.2, i.e. a higher weight on the global mapping for maintaining the color consistency. By combining the local and global mapping and optimizing them together, we enlarged the solution space while maintaining the color consistency. The result is further improved and no more observable haloing nor color drift exists (Fig. 7(e), Fig. 8(d)).

We solve the optimization with a typical gradient descent approach. We calculate the energy function in Eq.(13), take the derivative to the Taylor parameters ω_i, step forward along the tangent direction, and repeat this process until the energy value converges. It may take 10 to 100 iterations to converge, depending on image content. To
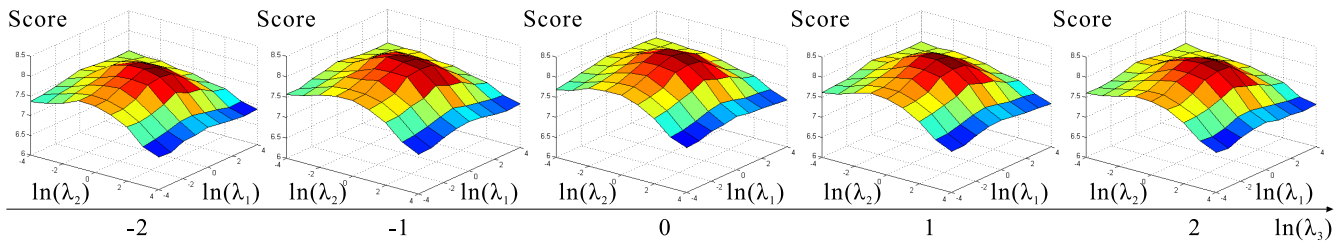
Figure 9: *Score against λ_1, λ_2 and λ_3. x axis represents λ_1, while y axis represents λ_2 and z axis represents score. Four instances of λ_3 are plotted. More details can be found in Fig. S17, S18 & S19 of the supplement.*

reduce the chance of being trapped by local optimum, we perform multiple times of gradient descent (50 times for all experiments), each with a different initial point. The one with the minimal energy is selected as our result. Since our current implementation is not optimized nor GPU-parallelized, it takes 1 to 10 minutes for optimizing a 1M-pixel image. It can be significantly sped up by applying speed-optimized solver together with GPU acceleration, as most computation can be straightforwardly parallelized.

5.2 Estimating Energy Weights

The energy weights (λ_1, λ_2, λ_3) of the objective function in Eq.(13) are estimated via a user study, so that the same set of weights are universally applied. We sample the 3D weight space of (λ_1, λ_2, λ_3) with λ_1, λ_2, $\lambda_3 \in [0.01, 100]$. For each sample point in the weight space, we optimized the corresponding objective function in order to synthesize the image pairs (30 images from a rich variety of visual content are selected). Then, twenty participants are invited in this weight estimation phase. Half of them are CVD audiences and the rest are normal vision audiences. For CVD audiences, they have to wear the stereo-glasses and are presented with the synthesized image pairs in dichoptic mode. They were asked to grade each synthesized image pair with a single score ([1,10], with 10 being the best) by collectively considering both the color distinguishability and the binocular fusion comfortability. For normal vision participants, they were monocularly presented with the synthesized images, as well as the original images, and are asked to grade the synthesized images in terms of deviation from the originals (also in the scale of [1,10]). The final score of each sample point in (λ_1, λ_2, λ_3) space is the average score of all participants on all correspondingly synthesized image pairs. Fig. 9 shows the result. Since the weight space is 3D, we visualize the result by fixing λ_3 at different instances. We found that the optimal weights can be found at (1,1,1). For this set of weights, the standard deviation (SD) of scores given by normal audiences is 0.85 and the SD by CVD audiences is 1.52, which are relatively small. Hence, we fixed our weights as (1,1,1) for all experiments. Note that our results are not very sensitive to the choice of weights, i.e. two results from two sets of slightly different λ_i values are sometimes visually quite similar. This can be observed in Fig. 9. Even though the graphs are plotted against the natural log of λ_i, the audiences scores of adjacent λ_i sets are still quite similar.

6 Results and Discussion

To evaluate the effectiveness of our method, we relied on visual comparison and multiple quantitative experiments as well as user studies with both CVD and normal-vision audiences.

Visual Comparison Firstly, we visually present our synthesized image pairs, by taking the most severe case of protanopia as our

target CVD audiences for generating results presented in this paper. Due to the space limit, results for other types of CVD audiences and other degrees of severity can be found in Fig. S11 to S16 of the supplement. Fig. 12 compares our results with our competitors. Both drawing (including test image from Ishihara test) and real photograph examples are evaluated. All test images exhibit at least one or multiple regions that CVD audiences has a difficulty in distinguishing colors. Two state-of-the-art methods proposed by [Huang et al. 2009] and [Chua et al. 2015] are compared. The former represents the traditional single-image recoloring method which attempts to minimize the deviation from original image, while the latter produces an image pair which is similar to our method. For a fair comparison, we have obtained the original implementations from [Huang et al. 2009] and [Chua et al. 2015] to generate their results for comparison. The left half of Fig. 12 shows the input and results as in normal vision, while the right half visualizes the results in CVD simulation. Note that Chua's and our results in Fig. 12(g)&(h) are presented binocularly to CVD audiences. For normal vision, the left and right images are blended (Fig. 12(c)&(d)) to simulate the blending effect of stereoscopic display without wearing the stereoscopic glasses. From the results, both Huang's and Chua's methods may fail to generate CVD-distinguishable results (e.g. the Ishihara image "42"). Moreover, their methods tend to introduce annoying false contours in natural images such as the results of "fruits" (as highlighted by red boxes). From the CVD simulation, the false contour of Chua's method is even more severe in almost all natural image results (Fig. 12(g)). In contrast, our results (blending of left and right) are almost identical to the input when presented to normal-vision audiences, and the distinguishability is maintained when displayed binocularly for CVD audiences. No false contour exists in any of our results thanks to our synthesis model. We achieve seamless visual sharing in which normal vision audiences are not aware of any discrepancy, while simultaneously CVD audiences' distinguishability are enhanced.

Readers are referred to Fig. S1 & S2 of the supplement, where we also compare our method to [Sajadi et al. 2013] and [Kuhn et al. 2008a]. We cannot fit in all results in the paper due to the page limit. As no original implementation of [Sajadi et al. 2013] can be found, we can only compare our results with the ones provided by their paper. The code of [Kuhn et al. 2008a] only produces recolored image in CVD simulated color space but not recolored image in normal vision, so we can only partially compare to their results. In summary, [Sajadi et al. 2013] introduces additional texture which limits its application to natural images, as audiences can no longer tell whether the texture is original or introduced by the synthesis. On the other hand, the method in [Kuhn et al. 2008a] does not consider normal vision audiences and may also produce undesirable false contours.

Quantitative Evaluation To quantitatively and objectively evaluate our results, we compare our images with those generated by

	SSIM (normal)	PSNR (normal)	CPR (CVD)
Chua	0.9573	33.1896 dB	0.8897
Huang	0.8844	22.5771 dB	0.9215
Ours	0.9719	40.0124 dB	0.9576

Table 1: *Statistics from quantitative evaluations in SSIM, PSNR and CPR.*

| (a) Ishihara image | (b) chart | (c) natural image |

Figure 10: *Example test images for CVD audiences user study. (a) Ishihara test image. (b) Color chart. (c) Natural image with color region indistinguishable for CVD audiences.*

[Huang et al. 2009] and [Chua et al. 2015] again. Thirty test cases of a wide variety of images are evaluated. They all exhibit color indistinguishability to CVD audiences. Note that our method and the method in [Chua et al. 2015] both generate an image pair, while the method in [Huang et al. 2009] generates a single image. We separate the quantitative evaluation into two parts, one for evaluating results presented to normal-vision audiences and the other for evaluating results presented to CVD audiences.

For results presented to normal-vision audiences, we evaluate how far our and competitors' results deviate from the input images. So we measure the SSIM [Wang et al. 2004] and PSNR. Since normal-vision audiences are presented with the blending of our left and right images, we compute the average of left and right images, and measure the SSIM and PSNR of the average image compared to the input. Table 1 shows the statistics. The higher the values are, the smaller deviation the images are. Both SSIM and PSNR scores of our results are the best. Huang's scores are the worst and significantly lower than ours, as their method considers no normal vision audiences and only utilizes a single display channel. Chua's SSIM score is slightly lower than ours, but their PSNR score is much lower than ours, due to the false contours. This is also evidenced in the above visual comparison.

To evaluate the results presented to the CVD audiences, we measure how well the contrast is preserved from the input image to the CVD simulated image pair. We extract the part corresponding to the contrast term in SSIM and regard it as the contrast preservation rate (CPR). It is defined as,

$$\frac{1}{n-1} \frac{\sum_{i=1}^{n} (x_i - \mu_x)(y_i - \mu_y) + \epsilon}{\sigma_x^2 + \sigma_y^2 + \epsilon}, \qquad (16)$$

where x_i and y_i are the two corresponding pixels from two compared images; μ and σ are the mean and standard deviation within a local neighborhood; n is the number of pixels in the local neighborhood; ϵ is a small constant to avoid divide-by-zero. In all our experiments, we take a local neighborhood of 11×11. This CPR falls in the range of [0,1]. The higher the value is, the better the contrast is preserved.

Next, we convert our left and right images to the CVD simulated color space, and compute the above CPR for each pixel in both images. The two resultant CPR maps are combined into one by taking the maximum of two corresponding pixels from the two maps. This per-pixel maximum operator is justified by the psychological finding [Scott et al. 2000] that, when two different contrast images are presented in a dichoptic manner, human vision system selectively perceives the sites with stronger contrast during the binocular single vision. Finally, the whole image CPR is simply the mean of all pixels in the combined map. The CPRs of our competitors are computed similarly, except that no maximum operation is needed for [Huang et al. 2009] as their method synthesizes only a single image. From the statistics, our method outperforms our competitors. Chua's score is much lower because of the false contours.

User Study on Functionality To directly evaluate our effectiveness, we invite 8 CVD individuals (2 deuteranopes and 6 deuteranomalous of age 18 to 27) with different severity to participate in a
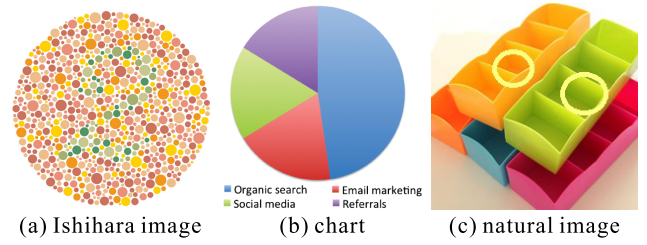
user study. Before the test, we first perform a short Ishihara test to classify the CVD type of each participant, then perform calibration to obtain his/her personal projection matrix, for generating results that tailored for each participant. Fifteen test images are chosen for the experiment. They contain Ishihara test images, color charts, and natural images (Fig. 10). All test cases exhibit at least one or more places where CVD individuals cannot distinguish colors. Test cases are presented to the participants in a random order.

During the experiment, we compare the visual experience of CVD audiences in using our solution to that of wearing tinted glasses (Enchroma Cx lenses), and that of two existing methods, [Huang et al. 2009] and [Chua et al. 2015]. We set up the experiment using a stereoscopic display on the laptop ASUS G750JX. Its displaying luminance is around 250 cd/m^2 and its size is 17.3 inches. User studies are conducted indoor with an ambient illumination of around 200 lux and the laptop display is calibrated with the colorimeter Spyder 3 in the same lighting condition. The display screen is positioned at 0.5 meter away from the participant. During testing our solution and [Chua et al. 2015], they have to wear the shutter glasses in which the images are presented to them in a dichoptic way. For tinted-glasses solution, participants wear the tinted glasses and are presented with the input images. For testing [Huang et al. 2009], participants are monocularly presented with the recolored images without wearing any glasses. Monocular presentation of the original input images to CVD participants without wearing any glasses is also provided as the control.

During the survey, we ask the CVD participants with questions on the functionality and the comfortability of the compared solutions. For functionality, we ask whether CVD participants can distinguish the colors and acquire the visual information. Since the nature of the types of test images are quite different, the way we query are adjusted accordingly for each type of images. For Ishihara test images, participants are asked to tell the number or object embedded in the image. For color chart images, participants are asked to link the color legend with the color regions in the chart. Only when they can correctly link all colors will the answer be considered as correct. For natural images, participants are asked to point out whether two circled regions are in difference colors (Fig. 10(c)). For each test case, we ask the participants twice, each with a different circled region pair. One of them is fake, i.e. the circled region pair is of the same color. Only when both questions are answered correctly, we regard the answer as correct.

Fig. 11(a) plots the average correctness of the above four solutions together with the control. The breakdown statistics for each type of test images can be found in Fig. S38 of the supplement. The vertical interval on each bar corresponds to the 95% confidence interval of user correctness. In general, all four solutions improve the correctness when compared to the control. Except for the type of natural images, the tinted glasses solution is inferior than the

Method	Mean	Standard Deviation	95% confidence interval	
			Upper bound	lower bound
Chua	88.3%	32.1%	94.1%	82.5%
Our	90.8%	28.9%	96.0%	85.6%

Table 2: *Stable vision. The statistics above shows the percentage that stable vision is formed for Chua's and our methods.*

Method	Mean	Standard Deviation	95% confidence interval	
			Upper bound	Lower bound
Huang	6.0%	23.7%	10.3%	1.7%
Chua	38.0%	48.5%	46.8%	29.2%
Ours	84.7%	36.0%	91.2%	78.2%

Table 3: *Deviation from the originals for normal-vision audiences*

control. To ensure the statistics is meaningful, we further apply the one-way analysis of variance (ANOVA) with a significant level 0.05 and a commonly used post hoc analysis method, least significant difference (LSD), to evaluate the result. Statistics shows that the difference between the control and tinted-glasses method is not significant, and that between the control and each of the three recoloring methods (Chua's, Huang's, and our methods) are statistically significant (p-values are much smaller than 0.05). That means the improvement of wearing tinted glasses is not significant, while the other three solutions improve the color distinguishability. In terms of functionality, our method is statistically better than Chua's method, and comparable to Huang's method. Note that Huang's method is designed without the consideration of normal vision audiences. Detailed statistics can be found in table S1 to S5 of the supplement.

Even if the participants can correctly answer, they may not be certain. So we also record their certainty of the correct answers. The certainty is in the scale of 1 to 5, with 5 being very certain. Fig. 11(b) plots the average certainty of each solution. From this overall statistics, there is no statistically significant differences in certainty between the control and each of the four solutions. Looking into the breakdown statistics of color chart images, the certainties of tinted glasses and Chua's method are significantly lower than that of the control. For the natural images, our method significantly improves the certainty and outperforms all competitors.

User Study on Comfortability Next, we evaluate the visual comfortability. As three out of four solutions require CVD participants to wear extra glasses, these three (the tinted glasses, Chua's and ours) certainly have a lower comfortability. We ask the user to grade their overall comfortability of these two solutions, again in the scale of 1 to 5, with 5 being the most comfortable. Fig. 11(c) plots the statistics. From the overall statistics, our solution is significantly better than tinted-glasses and Chua's method. CVD participants complain that tinted-glasses color the whole view, and leads to unnatural viewing experience, while Chua's solution introduces too much false contours. Similar observation is found in the breakdown statistics, except for the Ishihara image type, in which Chua's method and ours are comparable.

The second test is on binocular rivalry and is only for Chua's and our method, both utilize binocular display and dichoptic presentation. For each test case, we ask the participants whether the image pair can form a stable percept. Table. 2 shows the stability of perceived images (we assume our data follows normal distribution and calculate confidence intervals according to t-distribution table). Stable vision are generally formed in around 88.3% in Chua's results. This is as expected since all their results are produced under a certain amount of luminance deviation, which is below the rivalry threshold shown in their paper. The percentage of stable vision for our method is higher than Chua's, around 90.8%. Note that our optimization can only minimize, not eliminate, the chance of binocular rivalry. If the test case is very tough, it is possible the system has to pay less attention on the fusibility to trade for distinguishability and deviation. This happens in the cases of Ishihara images. These test cases are so tough that our competitors even cannot generate solutions (the top row of Fig. 12(g)&(h)). Our method has to relax the fusibility to generate solutions, leading to the higher

chance of binocular rivalry.

The third test is on the visual preference. This test is always the last among all tests for CVD participants. We tell the participants the right answer for each test case and present them the solutions of all competitors again. Ask them to grade their preference on each solution in the scale of 1 to 5, with 5 being the most preferred. The tinted-glasses and Chua's solution are the most unpopular, while Huang and our methods are comparable. Nevertheless, our solution offers the extra visual shareability with normal-vision audiences.

User Study on Visual Shareability We evaluate the visual shareability by studying the visual experience of normal-vision audiences when they are presented with the recolored images. We invited 10 participants with normal vision (age 18 to 30). Fifteen test images previously used for the above CVD user study are reused here. In each test session, the participant is presented with a 2×2 matrix of panels of images on the same stereoscopic display with the same illumination condition set up for the CVD user study. The normal-vision audiences do not wear any glasses. The top-left panel shows the original input images. The other three panels show the recolored images by [Huang et al. 2009], [Chua et al. 2015], and ours in a random order. Our and Chua's image pairs are displayed via the stereoscopic display, and hence effectively, the blending results of left and right images are presented to the participants.

During the experiment, participants are asked to select one or more results, that are closest to the original input, out of the three panels. We allow participants to choose more than one results when there is a tie. Table 3 shows the percentage of participant selections for all three solutions. Our results are selected over 84% of cases. This confirms to the statistics in quantitative evaluation, and evidences that our method introduces the minimal change. Note that for the remaining 16%, participants prefer our competitors' results because they make little changes on the original images, which also makes them fail to produce color-distinguishable solution (for CVD audiences).

Video Extension Video extension is an obvious next step to pursue. Machado and Oliveira [2010] and Huang et al. [2011] also consider temporal consistency in their methods, by sacrificing the deviation term a bit. Although our method is originally designed for static images, we have evaluated its feasibility in extension to video, via a naive strategy to maintain the temporal coherence. As each frame is synthesized with the corresponding Taylor parameter values, we can apply a temporal Gaussian filter (of a window size of 11 frames) on the parameter values. A result can be found in the supplementary materials.

To evaluate the result of video extension, we invite 8 deuteranopes and 10 normal people to perform a preliminary user study. One video is tested in this study. CVD participants are asked to watch the original video first, followed by our recolored video (wearing the stereo-glass). They are asked whether the color contrast in the recolored video is enhanced, and whether there exist any abrupt color changes in the recolored video. All CVD participants agree that our method improves the color contrast when compared to the original video, and 7 out of 8 CVD participants cannot aware of any tem-
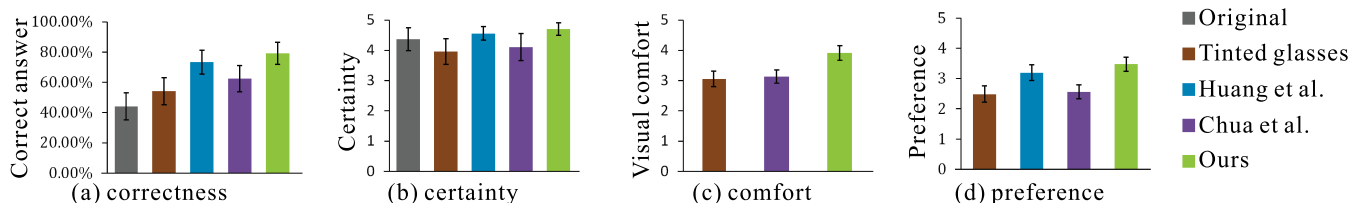
Figure 11: *Results of user study on functionality and comfortability.*

poral incoherence of the recolored video. For normal-vision participants, they are presented with the original and recolored videos in a side-by-side manner, and are asked to rate the color similarity between the two videos. The rating options include "not similar at all," "slightly similar," "moderately similar," "almost the same" and "totally the same." All of them rate the video as "almost the same" or "totally the same." This pilot test looks promising. Nevertheless, we believe a more sophisticated algorithm may be needed for more challenging video cases, and a more thorough evaluation may be required. This is out of the scope of this paper and should be our future direction.

Limitations Although our system can be calibrated for the target CVD audience, it can serve for only one type of CVD audiences at a time (either protanopia, deuteranopia or tritanopia). When the CVD audiences are in the same CVD type but different severity, we may calibrate our system according to the most severe individual so as to ensure the distinguishability for all audiences. Our current implementation also assumes that both eyes own the same type and the same severity of CVD. If this does not hold, we need to separately model the projection matrices for left and right eyes. We believe that, by modifying the implementation and utilizing binocular suppression, we can still generate a good solution. Another assumption is that the perceived image of normal audiences is the linear blending of left and right images on the screen, and here we ignore the influence of the nonlinear gamma correction of the display. BVCP is originally designed for normal-vision audiences. In this paper, we basically assume the BVCP model can be applied to CVD individual. Another limitation is that our method relies on the validity of the physiologically-based simulation of CVD vision. If the simulation cannot accurately model the visual perception of CVD individual, we may fail to produce good result for CVD individual. Fortunately, the CVD simulation in our recoloring optimization framework is replaceable. If a more accurate CVD simulation model is proposed in the future, we can simply replace our current model with the more advanced one. Recoloring-based techniques can be applied on digital visual content only, this limits its applications in comparison with the optical approach. Our current implementation is not real-time, further optimization is required.

7 Conclusions

By utilizing the extra display channel of stereoscopic display, we present the first system that allows CVD and normal-vision audiences to share the same visual content seamlessly and simultaneously, without sacrificing the original image color for normal-vision audiences or sacrificing the color distinguishability for CVD audiences. By wearing the stereoscopic glasses, CVD audiences can identify the indistinguishable colors. Without wearing the stereoscopic glasses, normal-vision audiences are presented with the blending of the left and right images, which is very close to the original image. We solve the image pair recoloring problem as optimization of an objective function that minimizes the color deviation for normal-vision audiences, and maximizes the color distinguisha-

bility and binocular fusibility for CVD audiences. Via extensive quantitative experiments and user studies, we demonstrate the effectiveness of the proposed method.

So far, we have tackled the still images and performed a pilot test on its extension to video. More sophisticated algorithm may be needed for maintaining the temporal coherence of more challenging cases (that may reduce the solution space) and further in-depth evaluation is necessary. This will be our future direction. Currently we do not consider the visual attention. We believe if the visual attention is taken into account, some of the constraints can be further relaxed and lead to an even larger solution space.

References

BAKER, D. H., MEESE, T. S., MANSOURI, B., AND HESS, R. F. 2007. Binocular summation of contrast remains intact in strabismic amblyopia. *Investigative ophthalmology & visual science 48*, 11, 5332–5338.

BRETTEL, H., VIÉNOT, F., AND MOLLON, J. D. 1997. Computerized simulation of color appearance for dichromats. *JOSA A 14*, 10, 2647–2655.

CHUA, S. H., ZHANG, H., HAMMAD, M., ZHAO, S., GOYAL, S., AND SINGH, K. 2015. Colorless: Augmenting visual information for colorblind people with binocular luster effect. *ACM Transactions on Computer-Human Interaction (TOCHI) 21*, 6, 32.

FARNSWORTH, D. 1957. The farnsworth-munsell 100-hue test for the examination of color discrimination.

GOOCH, A. A., OLSEN, S. C., TUMBLIN, J., AND GOOCH, B. 2005. Color2gray: salience-preserving color removal. *ACM Trans. Graph. 24*, 3, 634–639.

GRAHAM, C., AND HSIA, Y. 1959. Studies of color blindness: a unilaterally dichromatic subject. *PNAS 45*, 1, 96.

GRUNDLAND, M., AND DODGSON, N. A. 2007. Decolorize: Fast, contrast enhancing, color to grayscale conversion. *Pattern Recognition 40*, 11, 2891–2896.

HARTENBAUM, N. P., AND STACK, C. M. 1997. Color vision deficiency and the x-chrom lens. *Occupational health & safety (Waco, Tex.) 66*, 9, 36.

HOVIS, J. K. 1997. Long wavelength pass filters designed for the management of color vision deficiencies. *Optometry & Vision Science 74*, 4, 222–230.

HOWARD, I. P. 2002. *Seeing in depth, Vol. 1: Basic mechanisms.* University of Toronto Press.

HUANG, J.-B., TSENG, Y.-C., WU, S.-I., AND WANG, S.-J. 2007. Information preserving color transformation for protanopia and deuteranopia. *Signal Processing Letters, IEEE 14*, 10, 711–714.

HUANG, J.-B., CHEN, C.-S., JEN, T.-C., AND WANG, S.-J. 2009. Image recolorization for the colorblind. In *ICASSP*, 1161–1164.

HUANG, C.-R., CHIU, K.-C., AND CHEN, C.-S. 2011. Temporal color consistency-based video reproduction for dichromats. *Multimedia, IEEE Transactions on 13*, 5, 950–960.

HUNG, P., AND HIRAMATSU, N. 2013. A colour conversion method which allows colourblind and normal-vision people share documents with colour content. Tech. rep., tech. rep., Konica Minolta Tech. Report.

ICHIKAWA, M., TANAKA, K., KONDO, S., HIROSHIMA, K., ICHIKAWA, K., TANABE, S., AND FUKAMI, K. 2004. Preliminary study on color modification for still images to realize barrier-free color vision. In *Systems, Man and Cybernetics, 2004 IEEE International Conference on*, vol. 1, 36–41.

JEFFERSON, L., AND HARVEY, R. 2006. Accommodating color blind computer users. In *SIGACCESS*, 40–47.

JEFFERSON, L., AND HARVEY, R. 2007. An interface to support color blind computer users. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 1535–1538.

JUDD, D. B. 1948. Color perceptions of deuteranopic and protanopic observers. *J. Res. Natl. Bur. Stand 41*, 247–271.

JUDD, D. B. 1966. Fundamental studies of color vision from 1860 to 1960. *PNAS 55*, 6, 1313.

KONDO, S. 1990. A computer simulation of anomalous color vision. In *Color Vision Deficiencies, Symp. Int. Res. G. on CVD*, 145–159.

KUHN, G. R., OLIVEIRA, M. M., AND FERNANDES, L. A. 2008. An efficient naturalness-preserving image-recoloring method for dichromats. *Visualization and Computer Graphics, IEEE Transactions on 14*, 6, 1747–1754.

KUHN, G. R., OLIVEIRA, M. M., AND FERNANDES, L. A. 2008. An improved contrast enhancing approach for color-to-grayscale mappings. *The Visual Computer 24*, 7-9, 505–514.

LACCARINO, G., MALANDRINO, D., DEL PERCIO, M., AND SCARANO, V. 2006. Efficient edge-services for colorblind users. In *WWW*, 919–920.

LAU, C., HEIDRICH, W., AND MANTIUK, R. 2011. Cluster-based color space optimizations. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, IEEE, 1172–1179.

LEI, L., AND SCHOR, C. M. 1994. The spatial properties of binocular suppression zone. *Vision research 34*, 7, 937–947.

LU, C., XU, L., AND JIA, J. 2012. Contrast preserving decolorization. In *ICCP*, 1–7.

MACHADO, G. M., AND OLIVEIRA, M. M. 2010. Real-time temporal-coherent color contrast enhancement for dichromats. In *Comput. Graph. Forum*, vol. 29, 933–942.

MACHADO, G. M., OLIVEIRA, M. M., AND FERNANDES, L. A. 2009. A physiologically-based model for simulation of color vision deficiency. *IEEE Trans. Vis. Comput. Graph. 15*, 6, 1291–1298.

MACMILLAN, E. S., GRAY, L. S., AND HERON, G. 2007. Visual adaptation to interocular brightness differences induced by neutral-density filters. *Investigative ophthalmology & visual science 48*, 2, 935–942.

MEYER, G. W., AND GREENBERG, D. P. 1988. Color-defective vision and computer graphics displays. *IEEE CGA 8*, 5, 28–40.

NEUMANN, L., ČADÍK, M., AND NEMCSICS, A. 2007. An efficient perception-based adaptive color to gray transformation. In *Proceedings of the Third Eurographics conference on Computational Aesthetics in Graphics, Visualization and Imaging*, 73–80.

RASCHE, K., GEIST, R., AND WESTALL, J. 2005. Detail preserving reproduction of color images for monochromats and dichromats. *IEEE CGA 25*, 3, 22–30.

RASCHE, K., GEIST, R., AND WESTALL, J. 2005. Re-coloring images for gamuts of lower dimension. In *Comput. Graph. Forum*, vol. 24, 423–432.

SAJADI, B., MAJUMDER, A., OLIVEIRA, M. M., SCHNEIDER, R. G., AND RASKAR, R. 2013. Using patterns to encode color information for dichromats. *IEEE Trans. Vis. Comput. Graph. 19*, 1, 118–129.

SCOTT, S., BARBARA, S., AND GARZIA, R. 2000. Foundations of binocular vision: A clinical perspective. *McGraw-Hill Medical*.

SHARPE, L. T., STOCKMAN, A., JÄGLE, H., AND NATHANS, J. 1999. Opsin genes, cone photopigments, color vision, and color blindness. *Color vision: From genes to perception*, 3–51.

SHEEDY, J., AND STOCKER, E. 1984. Surrogate color vision by luster discrimination. *American journal of optometry and physiological optics 61*, 8, 499–505.

WAKITA, K., AND SHIMAMURA, K. 2005. Smartcolor: disambiguation framework for the colorblind. In *Proceedings of the 7th international ACM SIGACCESS conference on Computers and accessibility*, 158–165.

WANG, Z., BOVIK, A. C., SHEIKH, H. R., AND SIMONCELLI, E. P. 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing 13*, 4, 600–612.

WONG, B. 2011. Points of view: Color blindness. *nature methods 8*, 6, 441–441.

YANG, X., ZHANG, L., WONG, T.-T., AND HENG, P.-A. 2012. Binocular tone mapping. *ACM Trans. Graph. 31*, 4, 93:1–93:10.
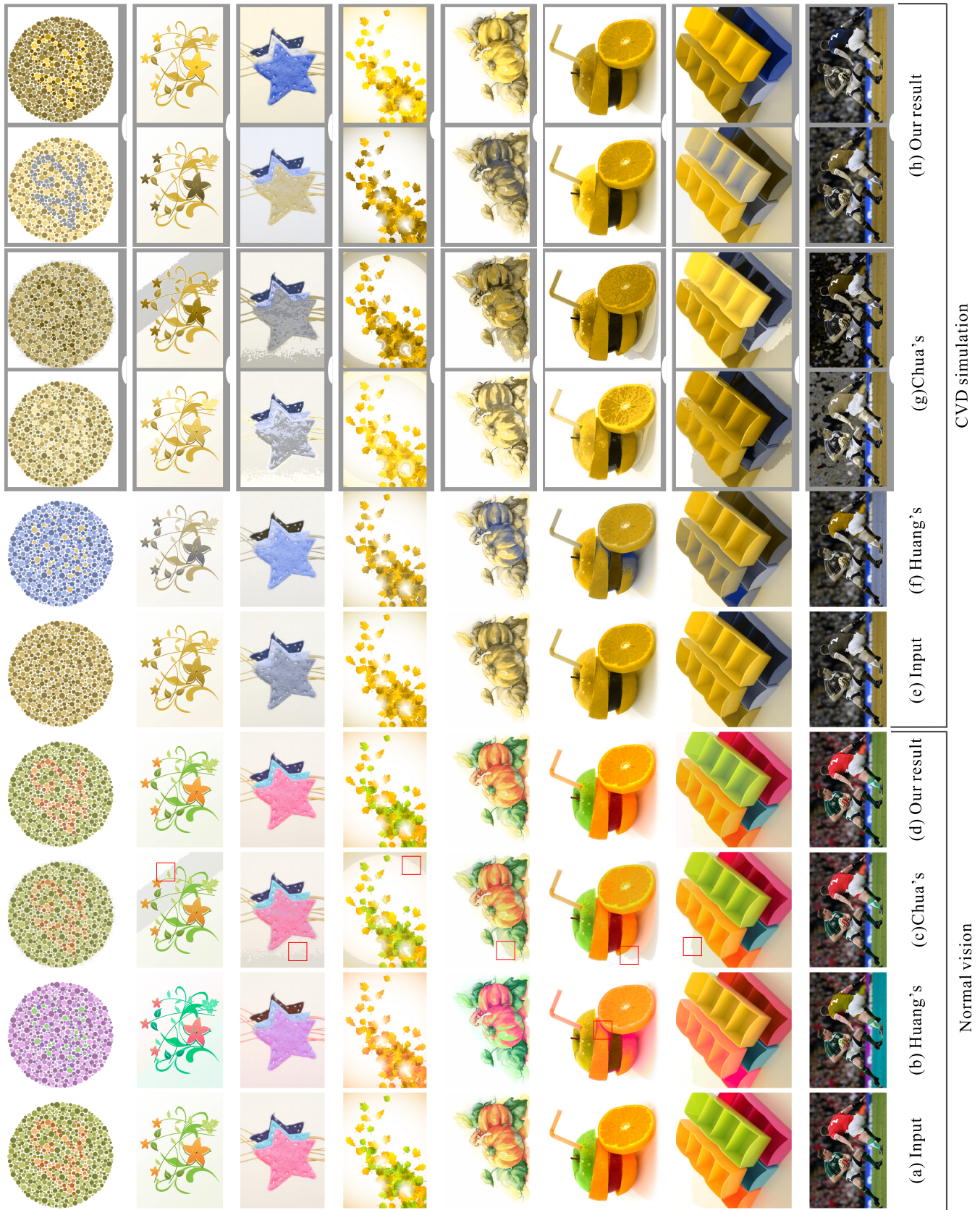
Figure 12: *Visual comparison of results. Column (a) shows the input. (b),(c)&(d): Results from Huang's, Chua's and our method as in normal vision. (e) Input in CVD simulation. (f),(g)&(h): Results from Huang's, Chua's and our method as in CVD simulation. Note that Chua's and our results are image pairs for CVD audiences. False contours can be found in red boxes.*