

Anti-Glare: Tightly Constrained Optimization for Eyeglass Reflection Removal

Tushar Sandhan and Jin Young Choi

Department of ECE, ASRI, Seoul National University, South Korea

tushar@snu.ac.kr, jychoi@snu.ac.kr

Abstract

Absence of a clear eye visibility not only degrades the aesthetic value of an entire face image but also creates difficulties in many computer vision tasks. Even mild reflections produce the undesired superpositions of visual information, whose decomposition into the background and reflection layers using a single image is a highly ill-posed problem. In this work, we enforce the tight constraints derived by thoroughly analysing the properties of an eyeglass reflection. In addition, our strategy regularizes gradients of the reflection layer to be highly sparse and proposes the facial symmetry prior via formulating a non-convex optimization scheme, which removes the reflections within a few iterations. Experiments on frontal face image inputs demonstrate the high quality reflection removal results and improvement of the iris detection rate.

1. Introduction

The French poet Du Bartas once praised eyes as, “These lovely lamps, these windows of the soul” [7]. If reflections are obstructing eye visibility, much can be lost!

Absence of a clear eye visibility degrades the aesthetic value of an entire face image as well as causes difficulties in many computer vision tasks, such as face recognition [37], eye detection [13], iris detection and segmentation [12], eye tracking and blink detection [8], eye center localization [55], iris and face liveness detection [36, 48, 49]. Even mild eyeglass reflections create more challenges by generating superposition of visual information. Since the presence of reflections is inevitable in real world, the natural need is to separate them and restore the underlying scene clearly.

Apart from illumination conditions, eyeglasses are the main culprits behind these annoying reflections, because photographing objects (eyes) are situated behind a semi-reflective medium (eyeglasses). As a result the captured image I , having N pixels, is a mixture between the desired background scene I_B and the reflection layer I_R , i.e. $I = I_B + I_R$. So the goal here is to separate I_B and I_R from a single input face image I , as shown in Figure 1.

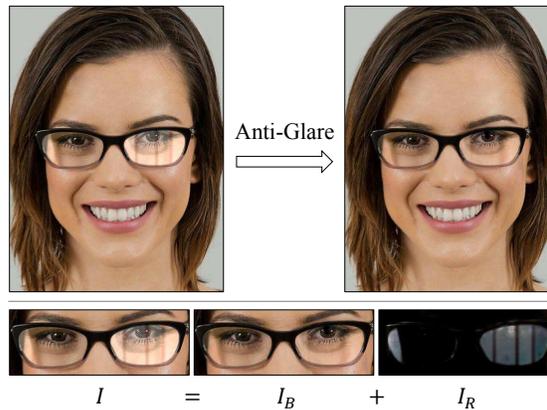


Figure 1: An input [39] having synthetic reflections and our result are shown at top. Decomposed background I_B and reflection I_R layers nearby eyeglasses are shown at bottom. For better visualization, the contrast of I_R is boosted.

This ‘layer separation’ problem is massively ill-posed, as it requires extraction of two unknown layers each of N pixels from a single image *i.e.* number of unknowns is twice the number of measurements. Prior information or additional measurements are required to make the problem tractable. So the previous work on reflection removal can be grouped into three main categories: *(i)* use of additional apparatus or controlled photographing conditions; *(ii)* use of multiple images of the same scene; and *(iii)* use of a single image along with good image priors.

Additional apparatus. The methods [22, 42] use a polarized filter to capture two images from the same position by rotating polarizing lens with different angles and then decompose the background and reflection using different priors on each layer [14, 52]. Reflections can be suppressed just by exploiting physical properties of the polarization [23, 46, 47]; or by using a pair of images captured with and without a flash [2, 3, 6]; or by using differential focusing technique [44, 45] with a pair of images taken, each focused only on one layer. These kinds of approaches require professional photographing skills and special tools, which puts limit on their applicability to daily use.

Multiple images. Generally reflections are formed closer to the glass as compared to the background scene, so the motion of a camera causes the reflection and background layers to move differently. This relative motion can be exploited for layer separation using either a sequence of images [19, 21, 30, 42, 54] or a video [24, 51, 60]. Similarly, the temporal misalignment between the layers helps in their decomposition, which can be detected effectively in gradient domain via gradient sparsity [15, 16, 59]. A more general approach [17] uses image statistics for separating the superimposed images. If one of the layers has repetitive dynamic motion then the layers can be decomposed easily by using temporal visual information [43]. Two-layer stereo information, which is obtained by limiting the camera motion to unidirectional parallel motion, can be leveraged for recovering the reflections [53, 56]. However, the absence of multiple images from different view points severely degrades the performance of these approaches.

Single image. A single image reflection removal task is extremely challenging, so previous work relies on sophisticated image priors or additional input from a user. If a user marks up on an image to denote gradient locations, color or texture that belong to either the background or the reflection, then the layers can be separated by imposing gradient sparsity or spatial smoothness prior [27, 28, 62]. Such methods become impractical when one has to deal with a lot of images, or reflection removal is used as the pre-processing block in some real time system (*e.g.* section 3.4); and their performance is governed by how precise the assistance of a user is. If the picture is taken by focusing only on the distant background scene then the reflection layer has significant blur. Thus reflection smoothness prior can be used for layer decomposition [31, 61]. Thicker glass causes the multiple reflections of the same object, which is known as the ‘ghosting effect’ and serves a good image prior for reflection removal [50]. Eyeglass reflections obey different illumination constraints and exhibit different properties (analysed in section 2.1) than general window glass reflections. Thus the existing methods are not adequate for efficient eyeglass reflection removal.

Contributions. In this work, we propose a tightly constrained optimization scheme to remove eyeglass reflections automatically using a single image. We first inspect the salient properties of eyeglass materials and ambient illuminations that characterize the reflections. Thereby we derive the priors as gradient sparsity, asymmetry, distinct color tint and piecewise constancy of the reflection layer. Then we embed these priors in a unified optimization framework, the solution of which gives the reflection-removed clean image. To the best of our knowledge, the proposed approach is the first attempt on eyeglass reflection removal, so we construct a new database to extensively verify its effectiveness for reflection removal and benefits for an iris detection.

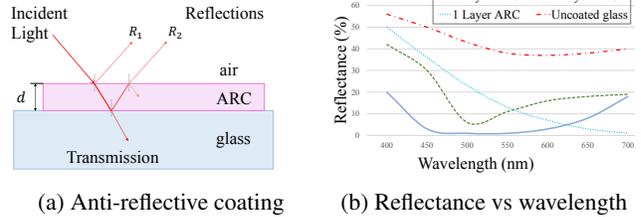


Figure 2: Interaction of the light with an eyeglass. For better comprehension, the figure (b) is recreated from [9, 40].

2. Our Approach

2.1. Eyeglass Reflection Properties

Unlike a window glass, which is thick and flat, an eyeglass is thin, curved and fabricated with variety of materials [1, 34]. So it is essential to inspect the structure of an eyeglass to discover properties of the eyeglass reflection.

2.1.1 Residual reflection

The relentless quest for understanding light and material science has led to the development of anti-reflective coatings (ARC). It refers to a transparent thin film that is applied to the glass for increasing the transmission of light and reducing the surface reflectance [40]. The basic working principal of ARC is summarized in Figure 2a. Thickness d of the single layer ARC is an odd multiple of $\lambda/4$, where λ is the wavelength of the incident beam. As a result, the reflected waves R_1 and R_2 become out of phase and cancel each other via destructive interference. These single layer ARC, although widely used due to their low cost and durability, work only at a single wavelength and a normal incidence. In order to achieve broadband optical performance a multi-layer ARC is necessary. Figure 2b summarizes the reflection reduction performance of different types of ARC [9, 40], which shows that even multi-layer ARC cannot reduce the reflections completely. Moreover, due to high cost, non durability and debonding issues, multi-layer ARC are not widely used [40]. In case of an eyeglass, as different wavelengths are attenuated differently (Figure 2b), the reflection layer shows a specific color or a hue like green, blue, light green, violet etc. [32]. These are known as the ‘residual reflections’ of an eyeglass [9, 38].

2.1.2 Sharp and sparse gradients

A transparent surface both reflects and refracts light, so its reflective properties depend upon what the ambient illumination is on its both sides. Figure 3a depicts typical glass window illuminations during the day and night. During the day, indoor glass reflections are overwhelmed by the profusion of light refracted from the outdoor. It renders re-

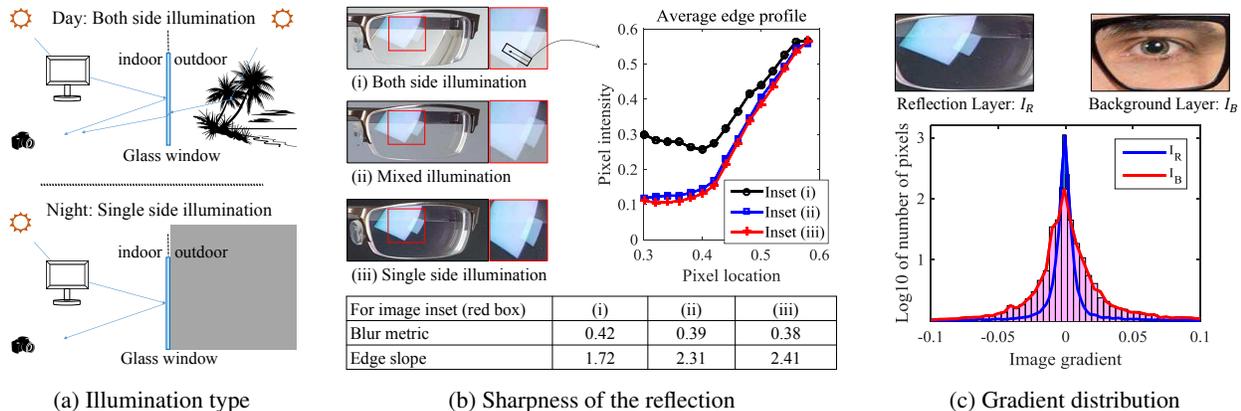


Figure 3: Properties of the eyeglass reflection: (a) The type of ambient illumination defines the reflection properties; (b) The blur metric [10] and edge slope of the reflection changes with ambient illumination. Edge profile is simply the average pixel intensity variation along a path (arrow in the inset of (i)); (c) Gradient distributions of the reflection and background layers.

reflections as smooth and diluted by the background scene. Contrary, the night has dimmed outdoor light and an eyeglass has a blocked side by an eye region, so being a single side illuminated, the reflections appear bright and compact. That is why a glass window of a train turns into a mirror like reflector at night or in a tunnel. Figure 3b shows the results of the experiment mimicking different illuminations. When one of the sides of an eyeglass has low or no light, the reflection shows higher edge slope and lower blur metric [10] than the case of full light on both sides. Figure 3c shows the typical eyeglass reflection, background layers and their corresponding gradient histograms. It reveals that the reflection layer has high and narrow peak distribution, whereas the background layer has low and broad peak distribution. Thus eyeglass reflections have sharp and sparse gradients.

2.2. Layer Separation Model

The layer separation problem can be addressed by adapting a probabilistic model to seek the most likely explanation of an image using prior knowledge about gradient distributions of the two layers (sec. 2.1.2) [29]. We model the high and narrow peak distribution of reflection layer as a hyper-Laplacian ($P(x) \propto \exp(-|x|^\alpha/b)$, for $\alpha < 1$) and the low and broad peak distribution of background layer as a Gaussian ($P(x) \propto \exp(-x^2/\sigma^2)$). The layers can be separated by maximizing joint probability $P(I_B, I_R)$, which is equivalent to minimizing $-\log P(I_B, I_R)$. Similar to [28, 59], assuming the two layers are independent and employing the gradient independence (i.e. $P(I_B, I_R) = P(I_B)P(I_R)$), the objective function becomes,

$$\min_{I_B, I_R} \sum_i \left\{ \sum_{\partial_j \in \mathbb{J}_R} |(\partial_j * I_R)_i|^\alpha + \sum_{\partial_j \in \mathbb{J}_B} \frac{\lambda}{2} \|(\partial_j * I_B)_i\|^2 \right\}, \quad (1)$$

where i is the pixel index $\in \{1, \dots, N\}$, ∂_j denotes the derivative filters belonging to the set \mathbb{J}_B or \mathbb{J}_R , λ is the scalar weight, $\alpha < 1$ and $*$ is the convolution operator. We used $\mathbb{J}_B = \{[1, -2, 1], [1, -2, 1]^T\}$ and $\mathbb{J}_R = \{[1, -1], [1, -1]^T\}$ in our implementation. Let denote the i^{th} element of a vector X as $X[i]$. Reflection should be non-negative ($0 \leq I_R[i]$) and cannot be more than the actual image ($I_R[i] \leq I[i]$). For simplicity, abusing the notations as $(\partial_j * X)_i = D_i^j X$ and $\partial_j \in \mathbb{J}_X \equiv j \in \mathbb{J}_X$; and after substituting $I_B = I - I_R$ in (1), the optimization problem for reflection I_R recovery is formulated as,

$$\min_{I_R} \sum_i \left\{ \sum_{j \in \mathbb{J}_R} |D_i^j I_R|^\alpha + \sum_{j \in \mathbb{J}_B} \frac{\lambda}{2} \|D_i^j I_R - D_i^j I\|^2 \right\} \\ \text{subject to } 0 \leq I_R[i] \leq I[i]. \quad (2)$$

2.3. Facial Symmetry Prior

As humans are bilaterally symmetric, the background layer I_B will have the symmetrical axis S_A . Whereas for the reflection layer I_R to possess the same axis of symmetry as S_A , the net illumination has to be symmetrical along S_A . In reality, the presence of multiple primary and secondary light sources makes the I_R less likely to be symmetrical. The recent methods [5, 26, 58] detect S_A in images by developing new symmetry descriptors, using oriented (e.g. Radon [41]) transforms or using dense feature correspondence, but it requires a long processing time [26] and high accuracy is not guaranteed. So instead, we use a simple trick of auto-flip-warping the image (summarized below and in Figure 5) for incorporating symmetry prior in the objective function.

Let an image be denoted in a vector form as I and in a matrix form as \mathbf{I} , and \mathbf{W}_f be the $N \times N$ flipping matrix such that, if $I_f = \mathbf{W}_f I$ then i^{th} column of the image matrix

				
PSNR [20]: 21.92 dB	21.93 dB	28.92 dB	35.80 dB	-
SSIM [20]: 0.8913	0.8915	0.9689	0.9880	-
(a) Synthetic input	(b) Reflection sparsity ($\gamma=0, \kappa=0$ in (7))	(c) Symmetry prior ($\gamma \neq 0, \kappa=0$ in (7))	(d) Tight constraints ($\gamma \neq 0, \kappa \neq 0$ in (7))	(e) Ground truth

Figure 4: A synthetic input showing importance of the different prior terms in (7) for reflection removal. (b) shows that relying only on the high sparsity of reflection layer without using any other prior leads to a $I_R \approx 0$ solution, which indeed is very sparse. After adding the symmetric prior term (section 2.3), the solution improves to (c) where still the symmetric parts of reflection prevail. Finally a good result is obtained after adding the tight constraints (section 2.4) to (7) as shown in (d). We show the PSNR and SSIM for the one to one comparison of images from (a) to (d) with the ground truth in (e).

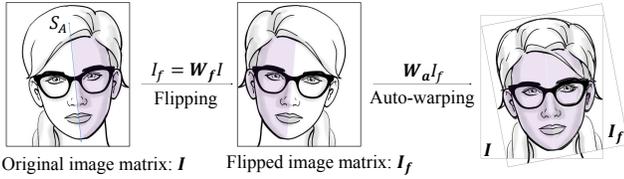


Figure 5: Illustrating the auto-flip-warping. Note how the in-plane rotation of a face can easily be handled by \mathbf{W}_a .

I_f will be identical to $(N-i+1)^{\text{th}}$ column of \mathbf{I} . Then the flipped image I_f can be warped or registered back to the original image I using $N \times N$ (auto) warping matrix \mathbf{W}_a as shown in Figure 5. So $\mathbf{W}_{af}I$ is the auto-flip warped image of I , where $\mathbf{W}_{af} = \mathbf{W}_a\mathbf{W}_f$. According to the symmetry prior, the background image will be invariant to this transformation *i.e.* $\mathbf{W}_{af}I_B \approx I_B$. Considering the $N \times N$ identity matrix \mathbf{I}_i , a scalar γ , defining $\widetilde{\mathbf{W}}_{af} = \mathbf{W}_{af} - \mathbf{I}_i$ and using $I = I_B + I_R$ lead us to the new prior term,

$$\frac{\gamma}{2} \left\| \widetilde{\mathbf{W}}_{af}I_R - \widetilde{\mathbf{W}}_{af}I \right\|^2. \quad (3)$$

2.4. Constraint Tightening by Residual Map

The prior information about the piecewise constancy and the distinct hue or color tint displayed by the residual reflections on an eyeglass (section 2.1.1) can be incorporated in the objective function via tightening the lower bound in the slack constraints of (2). So we define the hue map M_{H_t} as,

$$M_{H_t} = \exp(-\eta_1 \|H_I - \bar{\mu}(H_{R_t})\|^2), \quad (4)$$

where H_{R_t} is hue of the recovered reflection layer I_{R_t} at t^{th} iteration of the optimization scheme (described in section 2.5 and Figure 4); $\bar{\mu}(X)$ is the function calculating mean of the non-zero elements of a vector X ; η_1 is a positive scalar; H_I is hue of the input image I . That is (4)

searches the representative hue $\bar{\mu}(H_{R_t})$ of residual reflections over the input hue H_I . Thus the hue map M_{H_t} crudely locates the reflections using their hue distinctness.

Considering the piecewise constancy (*i.e.* compactness) of the reflections (section 2.1.2), we refine M_{H_t} at each iteration to construct the residual map M_{R_t} at a pixel i as,

$$M_{R_t}[i] = \frac{1}{w} \sum_{k \in \mathbb{N}_i} W[k] \exp(-\eta_2 \|i - k\|^2) \frac{M_{H_t}[k]}{\max(M_{H_t})},$$

$$W[k] = \exp\left(-\eta_3 \|\Delta_{i,k}H_I\|^2 - \eta_4 \|\Delta_{i,k}S_I\|^2\right), \quad (5)$$

where η_2, η_3, η_4 are positive scalars used for controlling width of the respective Gaussian kernels; S_I is the saturation of I ; w is the normalizing weight $w = \sum_{k \in \mathbb{N}_i} W[k] \exp(-\eta_2 \|i - k\|^2)$; \mathbb{N}_i is the neighbourhood of i in the image plane; $\Delta_{i,k}X = X[i] - X[k]$ and $0 \leq M_{R_t}[i] \leq 1$. The weight vector W avoids smoothing of M_H across the color edges by averaging hue and saturation of I . If we replace W in (5) as $\exp(-\eta_5 \|\Delta_{i,k}M_{H_t}\|^2)$ with η_5 being a positive scalar, then the residual map M_{R_t} simply becomes bilateral filtered version of the hue map M_{H_t} . However as the reflections are very sharp (Figure 3b), guiding the filtering by hue and saturation of the input image produces a residual map which is piecewise constant and consistent with the reflection regions on the eyeglasses. Typical examples of M_{H_t} and M_{R_t} are shown in Figure 6a.

Considering the priors (non-symmetry and specific color tint) exhibited by eyeglass reflections (sec. 2.1.1 and 2.3), the initial reflection color is estimated by using the non-symmetrical regions of an image that sprawl over the residual map *i.e.* $\bar{\mu}(I_{R_0}) \leftarrow \bar{\mu}(|(\widetilde{\mathbf{W}}_{af}I) \circ M_{R_t}|)$, where \circ denotes element-wise multiplication. Slack constraints of (2) are imposed upon the reflections so we tight the lower bound by disseminating the representative color of the reflection $\bar{\mu}(I_{R_t})$ over the residual map as,

$$\Gamma_t = \bar{\mu}(I_{R_t})M_{R_t}. \quad (6)$$

2.5. Optimization

After combining the priors (3) and (6) with the layer separation model (2), our final optimization problem becomes,

$$\min_{I_R} \sum_i \left\{ \sum_{j \in \mathbb{J}_R} |D_i^j I_R|^\alpha + \sum_{j \in \mathbb{J}_B} \frac{\lambda}{2} \|D_i^j I_R - D_i^j I\|^2 \right\} + \frac{\gamma}{2} \left\| \widetilde{\mathbf{W}}_{af} I_R - \widetilde{\mathbf{W}}_{af} I \right\|^2, \quad \text{s.t. } \kappa \Gamma_t[i] \leq I_R[i] \leq I[i], \quad (7)$$

where the scalar $\kappa \in [0, 1]$ helps to tight the lower bound gradually as the Γ_t becomes more reliable with each iteration for solving (7). Figure 4 shows the importance of different terms in (7) for the reflection removal on a synthetic input. Since $\alpha < 1$, (7) is non-convex, so we use half-quadratic splitting procedure [18, 25, 35] for solving it. Introducing auxiliary variables z_i^j at each pixel, we can split the cost function in (7) as,

$$\min_{I_R, z^j} \sum_i \sum_{j \in \mathbb{J}_R} \left\{ |z_i^j|^\alpha + \frac{\beta}{2} \|D_i^j I_R - z_i^j\|^2 \right\} + \sum_i \sum_{j \in \mathbb{J}_B} \frac{\lambda}{2} \|D_i^j I_R - D_i^j I\|^2 + \frac{\gamma}{2} \left\| \widetilde{\mathbf{W}}_{af} I_R - \widetilde{\mathbf{W}}_{af} I \right\|^2, \quad (8)$$

where as the scalar $\beta \rightarrow \infty$, (8) gets closer to the cost function in (7). So the optimization problem (7) is solved iteratively like: for a fixed β , we solve (8) by alternating between subproblems I_R and z^j with enforcing the constraints given in (7) at the end of each iteration t .

z-subproblem. For a fixed I_{R_t} , z^j can be updated for each pixel (ignoring the pixel index i) as,

$$z_{t+1}^j = \underset{z^j}{\operatorname{argmin}} \left\{ |z^j|^\alpha + \frac{\beta}{2} (z^j - D_i^j I_{R_t})^2 \right\}. \quad (9)$$

Since this is a single variable optimization problem, it can be rapidly solved by using a lookup table (LUT) based implementation [25], which maps the values from $D_i^j I_{R_t}$ to z^j . We fill the LUT by generating 10^4 different gradient values between -0.4 to 0.4 (this range is influenced by Figure 3c), for $\alpha = 0.6$ and various β . If the required value is missing from LUT then it is found via interpolation.

I_R -subproblem. With fixed z_i^j , (8) becomes quadratic in I_R , which can be solved by differentiating with respect to I_R and setting it to 0. We further apply 2D FFT \mathcal{F} to find the optimal update I_{R_t} quickly as given below in (11), where $\epsilon (=10^{-16})$ avoids division by 0, $\overline{\mathcal{F}(\cdot)}$ is a complex conjugate of $\mathcal{F}(\cdot)$, $\bar{1}$ is a vector of same size as I and having

$$I_{R_t} = \mathcal{F}^{-1} \left(\frac{\lambda \sum_{j \in \mathbb{J}_B} \overline{\mathcal{F}(D^j)} \circ \mathcal{F}(D^j) \circ \mathcal{F}(I) + \beta \sum_{j \in \mathbb{J}_R} \overline{\mathcal{F}(D^j)} \circ \mathcal{F}(z_t^j) + \gamma \mathcal{F}(\widetilde{\mathbf{W}}_{af}^T \widetilde{\mathbf{W}}_{af} I)}{\lambda \sum_{j \in \mathbb{J}_B} \overline{\mathcal{F}(D^j)} \circ \mathcal{F}(D^j) + \beta \sum_{j \in \mathbb{J}_R} \overline{\mathcal{F}(D^j)} \circ \mathcal{F}(D^j) + \gamma \mathcal{F}(\widetilde{\mathbf{W}}_{af}^T \widetilde{\mathbf{W}}_{af} \bar{1}) + \epsilon} \right) \quad (11)$$

all elements as 1, \circ denotes element-wise multiplication and the division in (11) is also performed element-wise. Only $\mathcal{F}(z_t^j)$ need to be computed at each iteration, whereas rest of the terms are precomputed once.

Satisfying constraints. Constraints of (7) can be satisfied by adding the adaptive normalizing constant ξ to I_{R_t} at each iteration such that $(\xi + I_{R_t}[i])$ falls within $[\kappa \Gamma_t[i], I[i]]$. The ξ can be obtained by solving following,

$$\min_{\xi} \sum_i \mathcal{U}(\xi + I_{R_t}[i] - I[i])^2 + \mathcal{U}(\kappa \Gamma_t[i] - \xi - I_{R_t}[i])^2, \quad (10)$$

where $\mathcal{U}(x)$ is the general unit step function (*i.e.* $\mathcal{U}(x) = 1 \forall x > 0$ otherwise $\mathcal{U}(x) = 0$), which penalizes only those elements that violet the constraints of (7). After solving (10) by a simple gradient descent, we update I_{R_t} for all i as $I_{R_t}[i] = \xi + I_{R_t}[i]$. Algorithm 1 summarizes the whole optimization scheme for an eyeglass reflection removal.

Algorithm 1 Anti-Glare: An Optimization Framework

- 1: **Input:** input image I ; optimization weights λ, γ ; total number of iterations T
 - 2: **Initialize:** $I_{R_0} \leftarrow I$, $\beta = \beta_0$
 - 3: **for** iteration t from 1 to T **do**
 - 4: update z_t^j using (9)
 - 5: update I_{R_t} using (11)
 - 6: $\beta = 2\beta$, $\kappa = \frac{t}{T}$
 - 7: update $I_{R_t} = I_{R_t} + \xi \bar{1}$ using (10)
 - 8: **output:** clean image $(I - I_{R_T})$
-

3. Experiments

In this section, first we present the experimental set up and implementation details; and then verify efficacy of the proposed method via the sets of experiments mentioned in the sections from 3.1 to 3.4. Eyeglass reflection removal results *i.e.* clean images obtained by using Algorithm 1 are shown in the Figures 1, 4, 8, 9, 10a and 11.

Dataset. We constructed a synthetic dataset with frontal face images of the people wearing variety of eyeglasses (*e.g.* equipped with or without ARC, narrow or wide glass aperture) under various illuminations (*e.g.* indoor, outdoor). Artificial reflection patches were added to those images which did not have much prior eyeglass reflections. Each image was annotated with the ground truth about the contours of eyeglasses and the outer boundaries of irises. This Eyes with Eyeglasses (EwE) dataset is available for download (<http://pil.snu.ac.kr>).

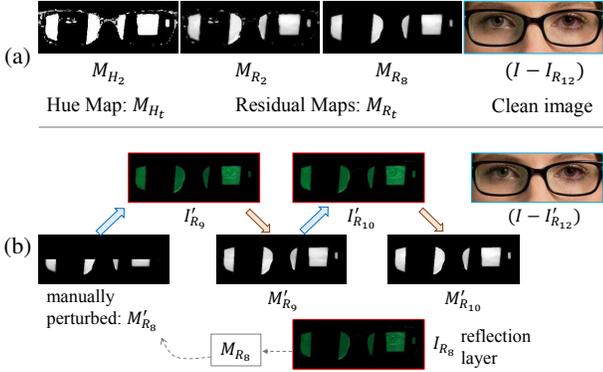


Figure 6: Evolution of the residual map for the synthetic input image given in Figure 4a. (a) shows, how using M_{H_t} from (4), the M_{R_t} gets refined iteration by iteration (as in (5)) to finally produce the clean image using Algorithm 1. (b) shows, how the reflection layer I_{R_t} indirectly influences the M_{R_t} via (4). So even if we manually perturb M_{R_8} to create the M'_{R_8} , it quickly recovers back to M'_{R_9} and eventually produces the $(I - I'_{R_{12}})$ which is visually similar to the $(I - I_{R_{12}})$ shown in (a).

Implementation details. The experiments were conducted on a PC equipped with Intel Core i7@ 3.4GHz CPU and 4GB RAM. The algorithm 1 was implemented in Matlab without any GPU acceleration. For simplicity, instead of eyeglass boundary detection, we first detect the face bounding box using [57] and crudely consider its top one-third part as our region of interest or treat as an input image I (resolution from 120×300 to 200×480) in our experiments. For constructing the auto-flip warping matrix \mathbf{W}_{af} any image registration technique is sufficient [33, 63], so we simply used a Matlab function `imregtform` with employing affine image transformation. Unless otherwise stated, the parameters of Algorithm 1 were fixed empirically as $\alpha=0.6$, $\beta_0=5$, $\lambda=3$, $\gamma=1.5$, $T=8$ (section 3.2) and \mathcal{N}_i as 8-pixel neighbourhood. For color images, each red, green and blue channel was processed independently.

3.1. Symbiotic Relationship between the Priors

Individual importance of the symmetry prior and the residual map prior has already been mentioned in the sections 2.3 and 2.4 respectively. Here, we try to analyse their inter-relationship. Figure 6a shows natural evolution of the residual maps (M_{R_2}, M_{R_8}) with iterations of the Algorithm 1. Even though we artificially degrade M_{R_8} to create the M'_{R_8} as shown in Figure 6b, it quickly recovers back to a good shape M'_{R_9} with the help of reflection layer I'_{R_9} in the subsequent iteration. And when we trace this process back, we know that initial reflection layers are discovered with the help of symmetry prior. Similarly when the sym-

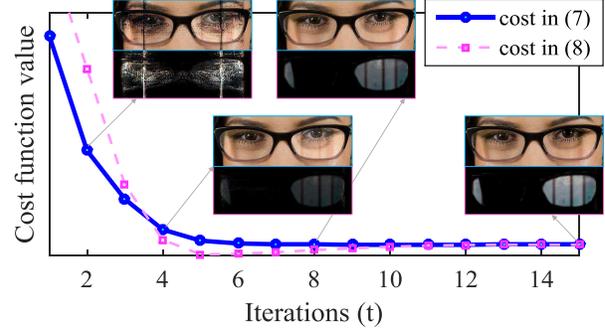


Figure 7: Variation of the costs in (7), (8); and some layer decomposition results (I_B is on top of I_R) with iterations.

metry prior got stuck and could not function further as in Figure 4c, the residual map had rescued it (Figure 4d) by growing over the remaining reflection regions and making them asymmetric iteration by iteration. Thus it is interesting to observe the beneficial (or obligate) symbiotic relationship between them *i.e.* in the optimization environment (7) both priors function together by helping one another for reducing hostility (cost function) of the environment. In the Figures 8 and 10a, even though some parts of the reflections are seemingly symmetric, the slight pixel level asymmetry is enough to work priors symbiotically and produce good quality reflection-removed images.

3.2. Convergence Analysis

Figure 7 shows the values of the cost function in (7) and (8) at each iteration. Even though (8) achieves the lowest values at around $t=5$, it does not resemble the cost in (7) at such early iterations as the β is small. Gradually as β increases, (8) approximates the cost in (7) closely. Four pairs of separated layers (I_B, I_R) are also shown in Figure 7 after 2, 4, 8 and 15 iterations. The results at $t=8$ are qualitatively similar to those at $t=15$. So by trading between speed and accuracy we set $T=8$ in our experiments. Thus the convergence speed of the Algorithm 1 is very fast.

3.3. Comparison with Previous Methods

We compare our method with the recent reflection removal approaches that use a single input image [28, 31, 50]. We used the implementations provided by the authors.

Qualitative comparison. Figure 8 compares the layer separation results (I_B, I_R) obtained by [28, 31, 50] on the images with different illuminations. Figure 8b shows the results of the method by Li and Brown [31], which relies on smoothness or blurriness of the reflection layer and therefore failed to remove the relatively sharp eyeglass reflections in the first and third (rows of Figure 8) input images; however the outdoor illuminated second (row) image has mild and smooth reflections where the method [31]

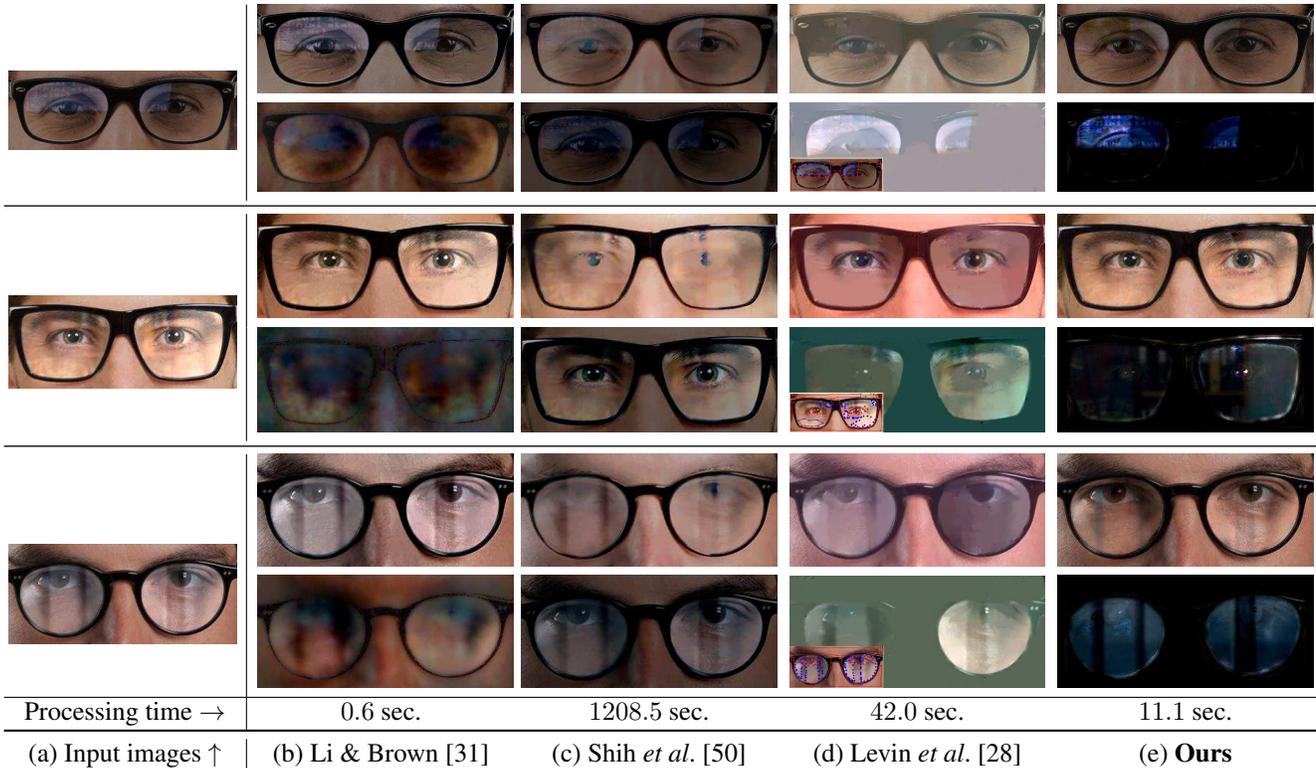


Figure 8: Qualitative comparison with previous methods. (a) shows the input images with different illuminations (from 1st to 3rd row as indoor dark, outdoor bright and outdoor dark). For these inputs, the columns (b) to (e) show the layer separation results (I_B on top of I_R) by various methods. We have created the annotations necessary for [28], which are shown in the inset of I_R images in (d). Average processing time for decomposing all inputs from (a) is compared in the bottom row.

produced good quality layer separation. Whereas Shih *et al.* [50] exploits the ghosting effect for reflection removal from the thick window glass, which is clearly absent on the thin eyeglasses. Thus [50] cannot handle the eyeglass reflections very well as shown in Figure 8c.

Levin *et al.* [28] requires user annotations for explicitly indicating reflection and background regions in an image. Following the instructions in [28], we have created the annotations as shown in the inset of the reflection layer images in Figure 8d. This method does relatively better job in removing sharp reflections, however it cannot recover the correct color tone of the I_B . Our method shows arguably the best results as shown in Figure 8e, considering the removal of sharp edges, high contrast of the I_R and preservation of the correct color configuration for I_B .

Quantitative comparison. In Figure 9, we quantitatively compare the reflection removal methods using the synthetic inputs and report the average SSIM (using gray and color images) and PSNR [20]. Levin *et al.* [28] shows visually good reflection removal results however achieves very less SSIM (0.5864) and PSNR (19.29 dB) on color

images. It may be due to the wrong recovery of the color tone and contrast in the result images (Figure 9d); because gray scale comparison of their results with the gray scale ground truth images shows improved SSIM (0.9428). The method [31] could not remove reflections (Figure 9b) because their prior (smooth reflections) is completely opposite of what we have here (sharp reflections). Reduction of the color SSIM values for the images in Figures 9b, 9c and 9d as compared to the values for original inputs 9a may be due to the mismatch in their color configuration, contrast *etc.* with the ground truth images 9f. As our priors (sections 2.3, 2.4) perfectly align with the properties of the reflection in synthetic images (section 2.1) and tight constraints (6) preserve the color configuration of original image, the proposed method shows (Figure 9e) significantly higher SSIM (0.9865) and PSNR (34.40 dB).

Computation time. Figure 8 reports the average computational time on the images of resolution about 180×400 . The proposed method consumes approximately 11 seconds, whereas Levin *et al.* [28] requires inconsistent time for getting user-annotations after which it spends 42 seconds to

24.43 dB	21.76 dB	20.03 dB	19.29 dB	34.40 dB	← avg. PSNR
0.9151	0.8687	0.8805	0.5864	0.9865	← color SSIM
0.9312	0.8904	0.7709	0.9428	0.9732	← gray SSIM
(a) Synthetic input	(b) Li & Brown [31]	(c) Shih <i>et al.</i> [50]	(d) Levin <i>et al.</i> [28]	(e) Ours	(f) Ground truth

Figure 9: Quantitative comparison using synthetic inputs. (b) to (e) show only the reflection-removed results I_B by different methods. We show average PSNR and SSIM of the color and gray images in (a) to (e) compared with the ground truth in (f).

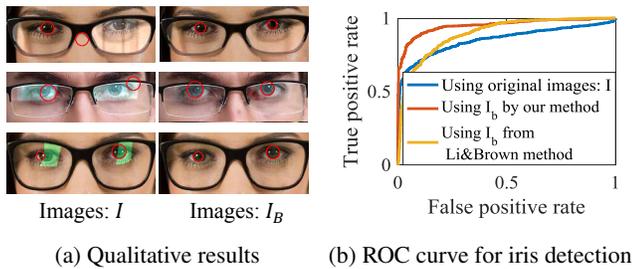


Figure 10: Experiments for iris detection. The clean images I_B from Algorithm 1 show better iris detection quality than the original images I with eyeglass reflections and the clean images from [31] as shown in (a) and validated in (b).

process the input data. The faster convergence (section 3.2) and thereby smaller computational time of the proposed method can be attributed to the LUT and FFT based iterative optimization framework.

3.4. Application: Iris Detection Improvement

Even mild or transparent eyeglass reflections degrade the performance of ocular biometric systems. Iris detection and segmentation are the precursors for iris recognition systems [11]. So here, we analyse how the eyeglass reflections affect the iris detection or localization performance. We have used EwE dataset for iris detection and have implemented [4] for detecting the outer boundary (circle) of an iris. Figure 10a visually compares the result of iris detection on original and reflection-removed images. Let C_g be the ground truth annotation and C_d be the detected bounding-circle of an iris, then the correct detection is obtained if the overlap area metric $\frac{C_g \cap C_d}{C_g \cup C_d}$ is greater than threshold τ . Receiver operating characteristic (ROC) curve is plotted by varying τ and shown in Figure 10b. It shows that the images cleaned using Algorithm 1 show greater iris detection accuracy than the original images or the clean images from [31].

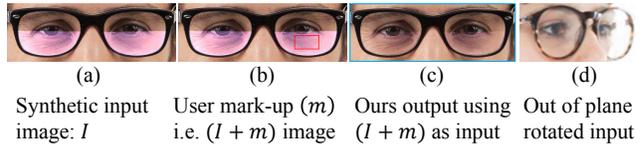


Figure 11: Challenging inputs for our method. (a) shows a synthetic input where the proposed method failed by producing $I_R \approx 0$, but after adding user annotations as in (b), it reduced the reflections as shown in (c). Our method cannot handle the largely out of plane rotated face images like (d).

4. Discussion

We have proposed the algorithm for eyeglass reflection removal from a single image. Our method exploits the sparsity, asymmetry, piece-wise constancy and specific color tint attributes of the reflections on an eyeglass, by incorporating all of them (priors) in a unified optimization scheme.

The challenging issue is that if the reflection layer turns out to be perfectly symmetric (*e.g.* Figure 11a), our priors break down and the method will fail to produce desired results. However if a user marks up a region containing reflections like Figure 11b, then the method starts working by building the residual map and gradually removes the reflections as in Figure 11c.

Even though the proposed method can handle slight out of plane rotated face images using dense (pixel-level) correspondence for building \mathbf{W}_{af} , it becomes less effective as the out of image-plane rotation increases like in Figure 11d.

In the future, utilization of spatio-temporal cues for mitigating above challenges as well as reducing saturated specular reflections could be an interesting research direction.

Acknowledgement

This work was partly supported by the ICT R&D program of MSIP/IITP (No.B0101-15-0552, Development of Predictive Visual Intelligence Technology), Brain Korea 21 Plus Project and Samsung Electronics Co. Ltd, MCB, Multimedia R&D group.

References

- [1] H. M. Abrams. Lamination of composite eyeglass lenses, 1995. United States Patent 5433810.
- [2] A. Agrawal, R. Raskar, and R. Chellappa. Edge suppression by gradient field transformation using cross-projection tensors. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, 2006.
- [3] A. Agrawal, R. Raskar, S. K. Nayar, and Y. Li. Removing photography artifacts using gradient projection and flash-exposure sampling. In *ACM SIGGRAPH*, 2005.
- [4] N. Q. AliAjdari Rad, Karim Faez. Fast circle detection using gradient pair vectors. In *Digital Image Computing: Techniques and Applications (DICTA)*, 2003.
- [5] I. Atadjanov and S. Lee. Reflection symmetry detection via appearance of structure descriptor. In *European Conference on Computer Vision (ECCV)*, 2016.
- [6] J. M. Baron. Method and apparatus for the removal of flash artifacts, 2005. United States Patent 6859565.
- [7] D. Bartas. The divine weeks and works of guillaume de saluste. (trans. by Josuah Sylvester, ed. by Susan Snyder), 2 vols. Oxford, 1979.
- [8] M. Chau and M. Betke. Real time eye tracking and blink detection with usb cameras, 2005. Boston University Computer Science Technical Report.
- [9] D. Chen. Anti-reflection (ar) coatings made by sol-gel processes: A review. *Solar Energy Materials and Solar Cells*, pages 313–336, 2001.
- [10] F. Crete, T. Dolmiere, P. Ladret, and M. Nicolas. The blur effect: perception and estimation with a new no-reference perceptual blur metric. *Human Vision and Electronic Imaging XII, Proc. SPIE*, 2007.
- [11] J. Daugman. How iris recognition works. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1):21–30, 2004.
- [12] J. Daugman. New methods in iris recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 37(5):1167–1175, 2007.
- [13] T. D’Orazio, M. Leo, G. Cicirelli, and A. Distanto. An algorithm for real time eye detection in face images. In *International Conference on Pattern Recognition (ICPR)*, 2004.
- [14] H. Farid and E. H. Adelson. Separating reflections and lighting using independent components analysis. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1999.
- [15] K. Gai, Z. Shi, and C. Zhang. Blindly separating mixtures of multiple layers with spatial shifts. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2008.
- [16] K. Gai, Z. Shi, and C. Zhang. Blind separation of superimposed images with unknown motions. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009.
- [17] K. Gai, Z. Shi, and C. Zhang. Blind separation of superimposed moving images using image statistics. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 34, 2012.
- [18] D. Geman and C. Yang. Nonlinear image recovery with half-quadratic regularization. *IEEE Transactions on Image Processing*, 4, 1995.
- [19] X. Guo, X. Cao, and Y. Ma. Robust separation of reflection from multiple images. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
- [20] A. Hore and D. Ziou. Image quality metrics: Psnr vs. ssim. In *International Conference on Pattern Recognition (ICPR)*, 2010.
- [21] M. Irani, B. Rousso, and S. Peleg. Computing occluding and transparent motions. *International Journal of Computer Vision (IJCV)*, 12, 1994.
- [22] N. Kong, Y. W. Tai, and J. S. Shin. A physically-based approach to reflection separation: From physical modeling to constrained optimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 36(2):209–221, 2014.
- [23] N. Kong, Y. W. Tai, and S. Y. Shin. A physically-based approach to reflection separation. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.
- [24] J. Kopf, F. Langguth, D. Scharstein, R. Szeliski, and M. Goesele. Image-based rendering in the gradient domain. *ACM Transactions on Graphics (TOG)*, 32, 2013.
- [25] D. Krishnan and R. Fergus. Fast image deconvolution using hyper-laplacian priors. In *Advances in Neural Information Processing Systems*. 2009.
- [26] S. Lee and Y. Liu. Curved glide-reflection symmetry detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34, 2012.
- [27] A. Levin and Y. Weiss. User assisted separation of reflections from a single image using a sparsity prior. In *European Conference on Computer Vision (ECCV)*, 2004.
- [28] A. Levin and Y. Weiss. User assisted separation of reflections from a single image using a sparsity prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29, 2007.
- [29] A. Levin, A. Zomet, and Y. Weiss. Separating reflections from a single image using local features. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2004.
- [30] Y. Li and M. S. Brown. Exploiting reflection change for automatic reflection removal. In *IEEE International Conference on Computer Vision (ICCV)*, 2013.
- [31] Y. Li and M. S. Brown. Single image layer separation using relative smoothness. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
- [32] N. G. Marechal and R. S. Blacker. Anti-static, anti-reflection coating, 2005. United States Patent 6852406.
- [33] S. Marenka and A. S. Ch. Watkins. *Modern Image Processing: Warping, Morphing and Classical Techniques*. Academic Press, UK, 1993.

- [34] C. W. Neeffe. Method of making high quality plastic lenses, 1979. United States Patent 4166088.
- [35] M. Nikolova and M. K. Ng. Analysis of half-quadratic minimization methods for signal and image recovery. *SIAM Journal on Scientific Computing*, 27, 2005.
- [36] G. Pan, L. Sun, Z. Wu, and S. Lao. Eyeblick-based anti-spoofing in face recognition from a generic webcam. In *IEEE International Conference on Computer Vision (ICCV)*, pages 1–8, 2007.
- [37] J.-S. Park, Y. H. Oh, S. C. Ahn, and S.-W. Lee. Glasses removal from facial image using recursive error compensation. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 27(5):805–811, 2005.
- [38] O. Pophillat and S. Gueu. Support, apparatus and method for performing a reflection measurement on an eyeglass, 2014. United States Patent 8891072.
- [39] Premium eyewear. Frames for america, Inc. accessed on, 2016. <https://www.framesdirect.com>.
- [40] H. K. Raut, V. A. Ganesh, A. S. Nair, and S. Ramakrishna. Anti-reflective coatings: A critical, in-depth review. *Energy Environ. Sci.*, 2011.
- [41] T. Sandhan, H. J. Chang, and J. Y. Choi. Abstracted radon profiles for fingerprint recognition. In *IEEE International Conference on Image Processing (ICIP)*, 2013.
- [42] B. Sarel and M. Irani. Separating transparent layers through layer information exchange. In *European Conference on Computer Vision (ECCV)*, 2004.
- [43] B. Sarel and M. Irani. Separating transparent layers of repetitive dynamic behaviors. In *IEEE International Conference on Computer Vision (ICCV)*, 2005.
- [44] Y. Y. Schechner, N. Kiryati, and R. Basri. Separation of transparent layers using focus. *International Journal of Computer Vision (IJCV)*, 39:2539, 2000.
- [45] Y. Y. Schechner, N. Kiryati, and J. Shamir. Blind recovery of transparent and semireflected scenes. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2000.
- [46] Y. Y. Schechner, J. Shamir, and N. Kiryati. Polarization-based decorrelation of transparent layers: The inclination angle of an invisible surface. In *IEEE International Conference on Computer Vision (ICCV)*, 1999.
- [47] Y. Y. Schechner, J. Shamir, and N. Kiryati. Polarization and statistical analysis of scenes containing a semireflector. *Journal of the Optical Society of America (JOSA)*, 17:276–284, 2000.
- [48] A. F. Sequeira, J. Murari, and J. S. Cardoso. Iris liveness detection methods in mobile applications. In *International Conference on Computer Vision Theory and Applications (VISAPP)*, 2014.
- [49] A. F. Sequeira, H. P. Oliveira, J. C. Monteiro, J. P. Monteiro, and J. S. Cardoso. Mobilive 2014 - mobile iris liveness detection competition. In *IEEE International Joint Conference on Biometrics (IJCB)*, pages 1–6, 2014.
- [50] Y. Shih, D. Krishnan, F. Durand, and W. T. Freeman. Reflection removal using ghosting cues. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
- [51] C. Simon and I. K. Park. Reflection removal for in-vehicle black box videos. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
- [52] M. Singh and X. Huang. Computing layered surface representations: an algorithm for detecting and separating transparent overlays. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2003.
- [53] S. Sinha, J. Kopf, M. Goesele, D. Scharstein, and R. Szeliski. Image-based rendering for scenes with reflections. *ACM Transactions on Graphics*, 31, 2012.
- [54] R. Szeliski, S. Avidan, and P. Anandan. Layer extraction from multiple images containing reflections and transparency. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2000.
- [55] F. Timm and E. Barth. Accurate eye centre localisation by means of gradients. In *Proceedings of the Int. Conference on Computer Theory and Applications (VISAPP)*, 2011.
- [56] Y. Tsin, S. B. Kang, and R. Szeliski. Stereo matching with linear superposition of layers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28, 2006.
- [57] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2001.
- [58] Z. Wang, Z. Tang, and X. Zhang. Reflection symmetry detection using locally affine invariant edge correspondence. *IEEE Transactions on Image Processing*, 24, 2015.
- [59] Y. Weiss. Deriving intrinsic images from image sequences. In *IEEE International Conference on Computer Vision (ICCV)*, 2001.
- [60] T. Xue, M. Rubinstein, C. Liu, and W. T. Freeman. A computational approach for obstruction-free photography. *ACM Transactions on Graphics (TOG)*, 34, 2015.
- [61] Q. Yan, E. E. Kuruoglu, X. Yang, Y. Xu, and K. Kayabol. Separating reflections from a single image using spatial smoothness and structure information. In *International Conference on Latent Variable Analysis and Signal Separation*, 2010.
- [62] S.-K. Yeung, T.-P. Wu, and C.-K. Tang. Extracting smooth and transparent layers from a single image. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2008.
- [63] B. Zitova and J. Flusser. Image registration methods: a survey. *International journal of Image and Vision Computing (IVC)*, 21:977–1000, 2003.