

# Reflection Removal for In-Vehicle Black Box Videos

Christian Simon and In Kyu Park

Department of Information and Communication Engineering, Inha University  
Incheon 402-751, Korea

{sen.christiansimon@gmail.com, pik@inha.ac.kr}

## Abstract

The in-vehicle black box camera (dashboard camera) has become a popular device in many countries for security monitoring and event capturing. The readability of video content is the most critical matter, however, the content is often degraded due to the windscreen reflection of objects inside. In this paper, we propose a novel method to remove the reflection on the windscreen from in-vehicle black box videos. The method exploits the spatio-temporal coherence of reflection, which states that a vehicle is moving forward while the reflection of the internal objects remains static. The average image prior is proposed by imposing a heavy-tail distribution with a higher peak to remove the reflection. The two-layered scene composed of reflection and background layers is the basis of the separation model. A non-convex cost function is developed based on this property and optimized in a fast way in a half quadratic form. Experimental results demonstrate that the proposed approach successfully separates the reflection layer in several real black box videos.

## 1. Introduction

The evolving paradigm of the automobile industry has changed the concept of a car from being not only a conventional driving machine but also a convergence of advanced IT technologies. Computer vision technologies, in particular, function to automatically provide the car with the capability of viewing and understanding the outside scene. The ultimate objectives of these technologies include vision-based lane detection, pedestrian detection, and collision avoidance.

In this regard, in-vehicle black box camera, as shown in Figure 1(a), has become a popular device for security monitoring and possible car accident recording. These devices are now being widely distributed in many countries. For example, in South Korea, more than 30% of personal vehicles and almost 100% of cabs and buses are equipped with

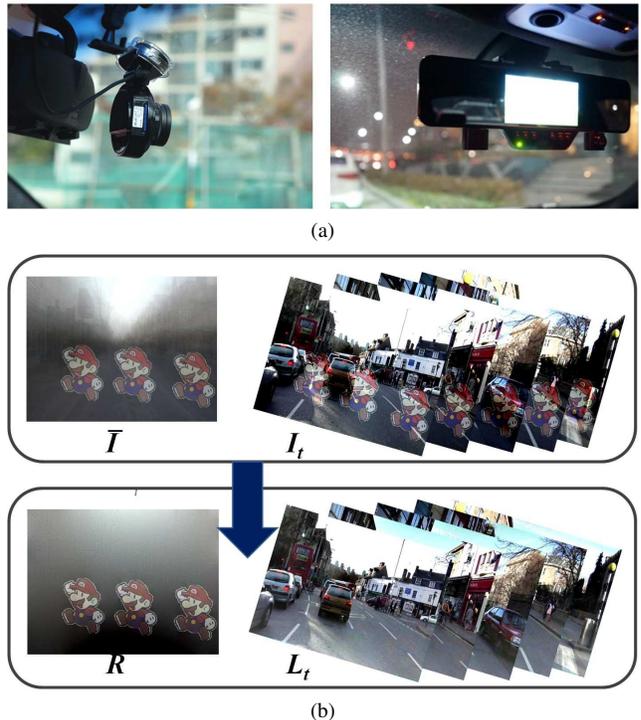


Figure 1: Reflection layer separation from the in-vehicle black box video. (a) Black box examples attached inside a car. (b) Example results of the proposed method in which the background scene ( $L$ ) and reflection ( $R$ ) are separated using the average image ( $\bar{I}$ ) of the frame sequence ( $I_t$ ) and the proposed optimization technique.

in-vehicle black box cameras.

However, most research on the in-vehicle black box camera has focused on enhancing its video resolution and compression. Despite the fact that the readability of video content is the most important aspect, few works have targeted this function. To address the readability issue, there are several problems that should be tackled. In this paper, we focus on video quality degradation due to the reflection on the

windscreen. In the video captured by an in-vehicle black box in a moving automobile, visibility is often diminished by the windscreen reflection of bright objects inside the vehicle. The unwanted reflection causes poor visibility in the captured videos. Therefore, the challenge is to remove the interference of the reflection, *i.e.* to separate the reflection layer from the background (outside scene) layer.

In the conventional layer separation problem, the observed image is comprised of the combination of the reflection layer  $R$  and the background layer  $L$ . The two layers are combined in a single image to create the mixture image  $I$  as follows.

$$I = L + R \quad (1)$$

This problem is highly ill-posed, even for a number of different mixture images. Attempts to solve the problem have been made to reconstruct the scene behind the glass surface [9, 13, 26]. In a black box video, the reflection on the windscreen seems easy to be separated from the background scene using existing methods. However, the black box video is very complicated to be handled because it typically includes non-uniform and perspective optical flow of the complex outdoor environment. For instance, it is not a trivial task to find the most salient gradient by taking the median in [24] on account of many moving objects outside. Moreover, in this condition, it is not straightforward to use conventional methodologies, such as optical flow [5], alignment, and warping.

To address the problem properly, we investigate the reflection in the captured videos to find the important difference from the imaging conditions of previous works. That is, the reflection on a windscreen tends to be stationary, whereas the background scene keeps moving along the direction of the vanishing line. This is natural because the camera and object inside the car are relatively static or slightly moving compared to rapidly moving objects in the outside scene. This unique circumstance of the in-vehicle black box video has inspired the concept of the effective approach proposed in this paper.

The main idea of the proposed method is to exploit the use of average image prior to remove the reflection with region-based optimization technique. The mathematical framework based on the sparsity prior like in [13] has shown significant contribution in reflection separation as well as in image deblurring [25, 28] and denoising [4, 23]. A non-convex cost function is developed based on the sparsity prior and subsequently optimized. Figure 1(b) shows an example of reflection removal by applying the proposed method. The contribution of this paper is summarized as follows.

- An efficient technique is presented to remove the reflection on the windscreen using the proposed average image prior and the region-based optimization technique.

- To our best knowledge, it is the first approach to separate the layer under fast forward motion of vehicle in outdoor environment.

## 2. Related Works

### 2.1. Single Image Layer Separation

Levin and Weiss [13] proposed a method using a mixture of Laplacian distributions to separate the layers on a single mixture image with human assistance. However, this method with human assistance is a tedious job, particularly to separate a reflection in a large number of images. In addition, Yu and Brown [27] proposed a method to separate a single latent image from reflection by assuming that reflection layer has significant blur while background layer remains sharp. However, this assumption is not applicable to black box videos in which both layers have less blur or sharp profile.

### 2.2. Layer Separation Using Multiple Images and Videos

In the earliest work, Weiss proposed a method of deriving intrinsic images from an image sequence captured in a video [24]. It minimizes the sparse representation by calculating the median of gradient in the temporal domain. Gradient in the image sequence indicates the possibility to separate the two layers. However, the median can incorrectly identify the gradient belonging to a specific layer especially when the gradient structure in image sequence is similar.

A few methods have utilized the camera's setting to capture multiple images. For instance, the reflection is removed from multiple images by employing focus adjustment [20] and polarizer rotation [9, 10, 20]. Moreover, Agrawal *et al.* [1] utilized gradient projection and the flash on a camera. Nevertheless, the additional setting and equipment are not relevant for application in portable mobile device such as black box. Note that the in-vehicle black box camera has fully automatic camera settings.

In [6, 7, 26], the most salient gradient in a set of images is utilized to separate layers. They assumed that the scene behind the glass surface is the layer to be reconstructed when capturing multiple images. Thus, it would have the most salient gradient. Nevertheless, this approach requires registration and warping processes which are unstable in a fast moving black box video.

In a video sequence, the layer separation problem becomes more tractable. Sarel and Irani [18] proposed a unique approach to separate non-rigid transparency using layer information exchange. However, they assumed that a set of images must have a different opacity for each layer. Moreover, the initial layer image mixture is obtained by intrinsic decomposition from the image sequence method [24]. The same authors also proposed a layer separa-

ration method using repetitive behavior detection [19]. This method requires an alignment method and motion estimation, which is difficult to be employed in a fast moving black box video with large displacement.

Although the above algorithms have shown notable performance, the black box video of a fast moving vehicle in an outdoor environment is quite complex for these methods to be applied in a straightforward manner. Therefore, we propose a novel algorithm.

### 3. Proposed Method

#### 3.1. Layer Separation Model

The reflection on a glass surface can be deduced as a layer separation model. The proposed model is inspired by the layer separation model in [13], in which the joint probability  $\Pr(L, R) (= \Pr(L)\Pr(R))$  of two layers  $L$  and  $R$  should be maximized. The probability can be derived as a minimization problem when the negative logarithm is applied. One approach minimizing the probability function is to utilize gradient independency between two layers. In this case, relevant prior knowledge should be applied to solve the minimization of the ill-posed problem. Therefore, it is equivalent to minimize  $E(L, R)$  as follows.

$$E(L, R) = \sum_{i \in N} \sum_{j \in J} \left( F_1((L \oplus g^j)_i) + F_2((R \oplus g^j)_i) \right) \quad (2)$$

where  $\oplus$  is the convolution operator and  $F_1(x)$  and  $F_2(x)$  denote negative log functions applied for distribution model of each layer. Moreover,  $N$  and  $J$  denote a set of pixels in the image space and derivative filters. In the remainder of the paper, the derivative filters are denoted as  $D_i^j X \equiv X \oplus g^j$  for  $j \in J$ .

In a video, this equation is extended to reconstruct the reflection in time domain ( $T$ ) with  $L = I - R$ , which thus becomes

$$\min_{R_t} \sum_{t \in T} \sum_{i \in N} \sum_{j \in J} (F_1(D_i^j R - D_i^j I)_t + F_2(D_i^j R)_t) \quad (3)$$

In this case, we observe the average of consecutive frames  $\bar{I}$  in a video, where

$$\bar{I} = \bar{L} + \bar{R}. \quad (4)$$

Because the reflection on the windscreen is stationary,  $R_t \approx R_{t+1}$  yields  $\bar{R} = R_t$ . Therefore, it is possible to omit the time domain subscript and regard  $R$  as a constant as follows.

$$E(R) = \min_R \sum_{i \in N} \sum_{j \in J} \left( F_1(D_i^j R - D_i^j \bar{I}) + F_2(D_i^j R) \right) \quad (5)$$

$\bar{I}$  is constant over several frames that are taken to obtain the average image. Stepping from Eq. (5), proper distribution

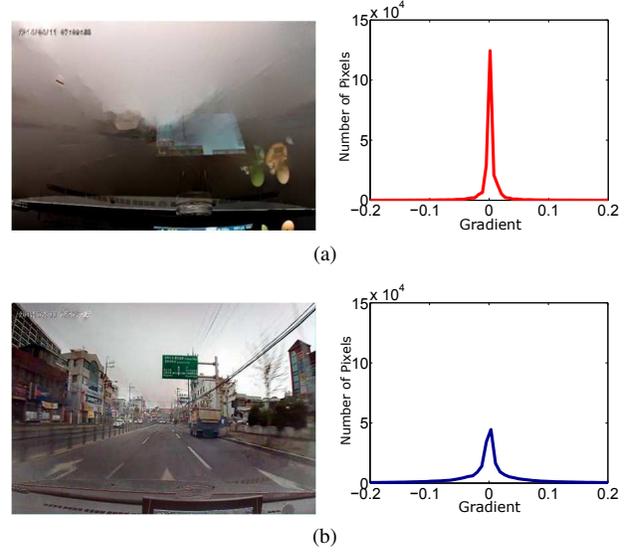


Figure 2: The distribution of image gradient from the background and reflection layers. (a) The reflection layer has high-peak distribution. (b) The background layer has low-peak distribution.

models for  $F_1$  and  $F_2$  should be imposed to perform reliable layer separation. In the proposed approach,  $F_1$  and  $F_2$  denote the narrow Gaussian and hyper-Laplacian models, respectively (*i.e.*  $F_1(x) \propto \frac{x^2}{\sigma^2}$  and  $F_2(x) \propto |x|^\alpha / s$  with  $\alpha < 1$ ). In this sense, unlike the previous methods, the proposed approach first reconstructs the reflection layer  $R$  rather than the scene behind the glass surface  $L$ . Therefore, the edges of the reflection layer should be separated robustly using the average image to solve the layer separation problem. The proposed average image prior and reflection removal method are described in the following subsections.

#### 3.2. Average Image Prior

An interesting observation from in-vehicle black box video is that the reflection is almost stationary over time. Based on the observation that the reflection layer is static, we can take average from the frame sequence to produce an extremely low-passed image in which the reflection still remains sharp. It is named *average image prior* in this paper.

To achieve reliable separation of two layers, it is necessary to have the beneficial prior information. As shown in Figure 2(a), the gradient of the reflection layer is quite sparse and the gradient in most area is very small or nearly zero. On the other hand, as shown in Figure 2(b), the distribution of background image has a lower peak. Thus, in modeling the distribution using the function  $p(x) = e^{-|x|^\alpha} / s$ , it can be deduced that the function to model the gradient distribution of reflection layer has a much smaller  $\alpha$  than  $\alpha$  to model that of the background layer.

**Hyper-Laplacian Distribution** There are several distributions that can model the non-Gaussian distribution in low-level image processing to approximate sparse distribution. For instance, student-t [17], Gaussian scale mixtures [22], Laplacian mixtures [13], and hyper-Laplacian [8] distributions are commonly used. As described in [12], high-peak with heavy-tail distribution is more effectively approximated by the hyper-Laplacian model (with  $\alpha = 0.8$ ) to produce sharper edges and less noise. Accordingly, the prior distribution model for reflection utilizes the robustness of the hyper-Laplacian model ( $(p(x) = e^{-k|x|^\alpha}/s)$ ) with  $\alpha < 1$ . However, in the average image, the gradient representing moving scene outside cannot be well-penalized if only a single  $\alpha$  is used because some areas have large gradient values while the others are not. Therefore, different  $\alpha$  values are applied using the proposed region-based model which is described in the following paragraph.

**Relative Sparsity Distribution using Region Division**

In the proposed approach, the average image is divided into several different angular regions to utilize different shapes of hyper-Laplacian distributions with different  $\alpha$  values. The number of different alpha values is denoted as  $H$ . Considering the forward motion of a vehicle, there is a single vanishing point in the center area of the image. The background layer consists of regions with different properties of gradient sparsity including the ground, building/wall/green, and sky areas. Therefore, it is better to assign different  $\alpha$  values to different angular regions. The vanishing point is easily found using a group of line segments with a similar direction as shown in Figure 3(b). The line segments are detected using the recent detector [21]. Then, the vanishing point is estimated using the dominant line direction from detected line segments with high confidence.

Figure 3(c) shows an example of uniformly divided angular bins centered at the detected vanishing point. For each region, the sum of length is computed for the line segments which converge to the vanishing point. Using the histogram of the sum of length as shown in Figure 3(d) (sorted in ascending order), classification is performed to generate  $H$  groups of angular regions with similar gradient density. For example, in Figure 3(c), 36 angular bins are grouped into 4 groups of similar gradient density. In Figure 3(c) and Figure 3(d), histogram bins and angular regions with similar gradient density are painted with the same color. Angular regions with higher density are assigned bigger  $\alpha$  values, with typical range of  $0.5 \leq \alpha \leq 0.8$ .

**3.3. Reflection Removal**

Reflection removal can be performed by applying the proposed prior and optimizing each term. Because the image is divided into angular subregions with different gradient distribution, the minimization in Eq. (5) becomes as

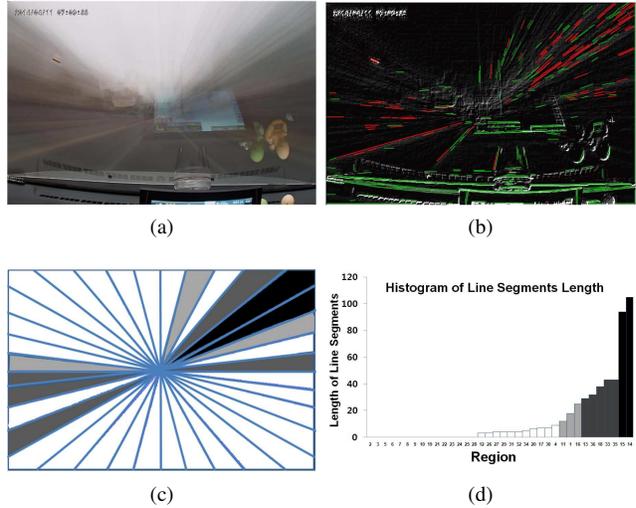


Figure 3: Line segments and angular region division. (a) The average image. (b) Detected line segments from the average image. Red line segments converge to the vanishing point while green ones do not. (c) Example of angularly divided regions (with different  $\alpha$  values in different shades). (d) The histogram of line segments length for each region.

follows.

$$\min_R \sum_{i \in N} F_1 \left( \sum_{j \in J_1} \{D_i^j R - D_i^j \bar{I}\} \right) + \sum_{h \in H} \sum_{i \in N_h} \sum_{k \in J_2} F_2(D_i^k R)_h \tag{6}$$

where narrow Gaussian  $F_1$  can be modeled as  $F_1(x) \propto \frac{x^2}{\sigma^2}$  with  $1/\sigma^2$  associated with  $\lambda$ . Moreover,  $F_2(x)_h$  is the function of hyper-Laplacian model with different  $\alpha$  value in each region and  $N_h$  denotes a set of pixels in regions assigned with  $\alpha_h$ . The second derivative filters are utilized in  $J_1$  to control the smoothness in the result, while the first derivative filters are used in  $J_2$  to recover the edges. In our implementation, the derivative filters in  $J_1$  are  $D^1 = [ 1 \ -2 \ 1 ]$ ,  $D^2 = [ 1 \ -2 \ 1 ]^T$  while the derivative filters for  $J_2$  are  $D^1 = [ 1 \ -1 ]$ ,  $D^2 = [ 1 \ -1 ]^T$ .

**Optimization** It is not a trivial problem to optimize Eq. (6) since  $\alpha < 1$ . This condition leads to a concave function. In this case, the half-quadratic method [11, 15] should be applied to simplify the optimization problem. The auxiliary variables  $y^1$  and  $y^2$  are introduced for non-convex

function  $F_2$  as follows.

$$\begin{aligned} & \min_{R, y^k} \sum_{i \in N} \left\{ \frac{\lambda}{2} \left( \sum_{j \in J_1} \{D_i^j R - D_i^j \bar{I}\} \right)^2 \right\} + \\ & \sum_{h \in H} \sum_{i \in N_h} \sum_{k \in J_2} \left\{ \frac{\beta}{2} (\|D_i^k R - y_i^k\|_2^2 + |y_i^k|^{\alpha_h}) \right\} \quad (7) \\ & \text{s.t. } 0 \leq (R)_i \leq \bar{I}_i \end{aligned}$$

Eq. (7) can be solved by optimizing two subproblems  $R$  and  $y$ . The optimal solution for subproblem  $y$  can be obtained by solving the following auxiliary minimization.

$$y^* = \arg \min_y |y|^{\alpha_h} + \frac{\beta}{2} (y - D_i^k R)^2 \quad (8)$$

In our case, same  $\alpha$  is repetitively accessed for a specific region  $h$ . Therefore, it is necessary to compute the optimized  $y$  and store the values to map the result after optimization. This problem can be rapidly tackled by employing the method proposed in [11]. A lookup table (LUT) is utilized to map the values from  $D_i^k R$  to  $y$ . To map the values, we generate 10,000 different gradient values between -0.6 to 0.6 for specified  $\alpha$  and  $\beta$  values. The gradient value which is not covered in the LUT can be found using extrapolation. Consequently, the optimization of  $y$  can be performed fast.

Afterwards, as described in [11], the  $R$  sub-problem is solved in a fast manner using FFT ( $\mathcal{F}$ ) and IFFT ( $\mathcal{F}^{-1}$ ) as follows.

$$R = \mathcal{F}^{-1} \left( \frac{\sum_{k \in J_2} \mathcal{F}(D^k)^* \circ \mathcal{F}(y^k) + (\frac{\lambda}{\beta}) \mathcal{F}(K_{J_1})^* \circ \mathcal{F}(K_{J_1})^* \circ \mathcal{F}(\bar{I})}{\sum_{k \in J_2} \mathcal{F}(D^k)^* \circ \mathcal{F}(D^k) + (\frac{\lambda}{\beta}) \mathcal{F}(K_{J_1})^* \circ \mathcal{F}(K_{J_1}) + \epsilon} \right) \quad (9)$$

where  $\epsilon$  and  $K_{J_1}$  are a small number to avoid division by zero and 2D matrix of two second derivative filters in  $J_1$ , respectively. In addition,  $*$  and  $\circ$  denote complex conjugate and component-wise multiplication, respectively. After  $R$  is obtained, it should be normalized such that it falls within a specific range  $[0, \bar{I}_i]$ . To achieve it, we need to obtain the constant  $g$  by minimizing as follows.

$$\min_g \sum_{i \in N} m_i((R)_i + g)^2 + n_i((R)_i + g - \bar{I}_i)^2 \quad (10)$$

where,

$$\begin{aligned} m_i(\cdot) &= \begin{cases} 1 & (R)_i + g < 0 \\ 0 & \text{otherwise} \end{cases} \\ n_i(\cdot) &= \begin{cases} 1 & (R)_i + g > \bar{I}_i \\ 0 & \text{otherwise} \end{cases} \end{aligned} \quad (11)$$

In Eq. (10),  $m_i$  and  $n_i$  are indicative functions. Traditional gradient descent is sufficient to optimize this normalization.

**Postprocessing** In the final step, the background scene can be computed by  $L_t = I_t - R$ . However, both  $R$  and  $L_t$  may have varying tone. To adjust the tone of  $R$  for each frame, we use luminance transfer method in [16] with the input frame as a reference image. Moreover, background scene  $L$  is adjusted to have similar brightness by calculating the mean of luminance in the newly reconstructed background ( $\bar{l}_L$ ) and the input frame ( $\bar{l}_f$ ) for angularly divided region. Lastly, luminance value  $l$  for each pixel  $i$  in  $L$  can be adjusted by  $l_i = l_i + G(\bar{l}_L - \bar{l}_f)$ , where  $G$  is the average filter.

## 4. Experimental Results

In this section, we describe the performance of the proposed method and the comparison with previous works. In our experiment, the proposed method is evaluated using both real-world videos with reflection on the windscreen and synthetic videos. We collect the data from various sources such as video sharing sites and private black box video collections. To demonstrate the robustness of the proposed method, we compare it with the state-of-the-art algorithms by Yu and Brown [26] and Sarel and Irani [18]. Furthermore, single image layer separation techniques in [27] and [13] are additionally considered for comparison. The proposed method is implemented using Matlab on a PC with Windows 7 32-bit operating system. The PC equips with an Intel® Core™ i7 3.5 GHz with 8 GB RAM. In angular region division, the average image is divided into 8 initial regions which are empirically set  $\alpha_1 = 0.5$ ,  $\alpha_2 = 0.6$ ,  $\alpha_3 = 0.7$ , and  $\alpha_4 = 0.8$  ( $H = 4$ ).

**Qualitative Comparison** First, we compare the proposed method with the layer separation algorithms using multiple images [26] and video [18]. These methods require many frames to find the correct gradient of layers and huge amount of memory which can hardly be handled properly with current PC memory. Therefore, we consider using five different frames to be applied for their methods. Moreover, in our test data, it is not feasible to apply the layer separation technique by Yu and Brown [27] directly from the multiple images because their method requires image registration and warping process to align gradient from the background scene. Since the reflection scene is static, SIFT-Flow [14] algorithm is not performed correctly to warp the multiple images. The global threshold values and parameters of the proposed method are tuned to obtain best results in the frame sequences of test videos. We cannot show the comparison with another layer separation from multiple images proposed by Gai *et al.*'s [6] because their method



Figure 4: Frame by frame results with comparison. The first row shows the average image from the input frames and the result of separated reflection layers from the initial sample frame. (a) Input frames. (b) Result of [26]. (c) Results of [18]. (d) Results of the proposed method. Video results including this example are provided in the supplementary material.

cannot handle memory well for the image resolution used in our data. Furthermore, their optimization takes a long processing time even with the small size image.

Figure 4 and Figure 5 show that the proposed method performs better in removing the reflection of the real-world videos. Note that the proposed method even works properly for videos in rural areas as long as the background is not homogeneous. In addition, the method also works

well for slightly changing lighting system. While the proposed method shows comparable performance with Sarel and Irani’s work [18], it is observed that [18] cannot produce the background scene correctly as shown in Figure 4(c) and Figure 5(c). Much of the objects’ gradients are degraded and the frames of the video have inconsistent tone in temporal domain. Not surprisingly, degraded details in the result occur because their method heavily de-



Figure 5: Additional comparison with single layer separation techniques. (a) Selected input frames. (b) Results of [26]. (c) Results of [18]. (d) Results of the proposed method.

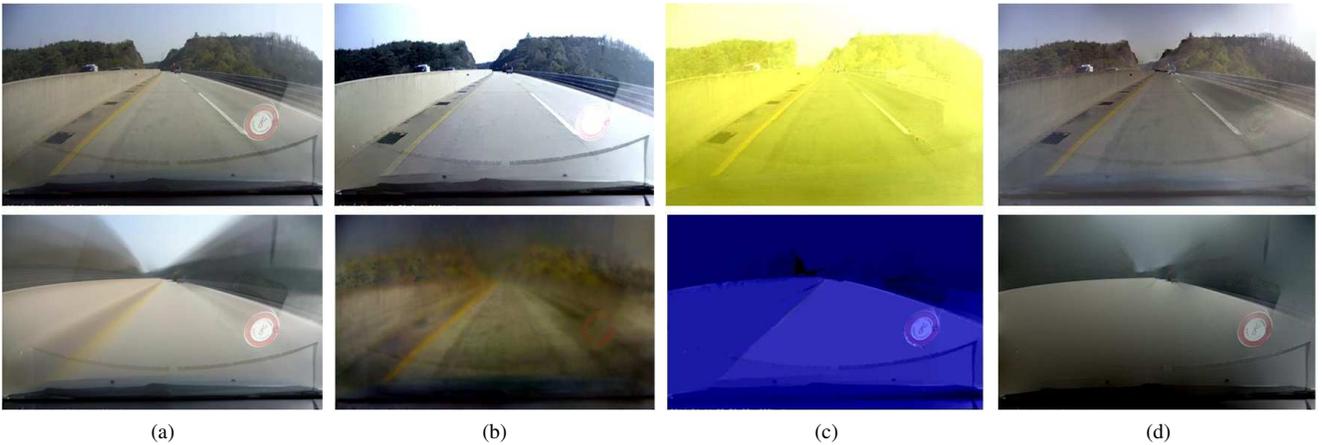


Figure 6: Comparison results with single layer separation techniques. (a) Selected frame with the average image. (b) Separation layer result from [27]. (c) Separation layer result from [13]. (d) Result from the proposed method.

depends on the median of gradients in the temporal domain when creating the initial mixture of images. Furthermore, initial layer mixture should be distinct in terms of gradient structure to obtain satisfactory result. In addition, as shown in Figure 4(b) and Figure 5(b), the method of Yu and Brown [26] produces unpleasing result with inconsistent color and ghost artifact because the complex outdoor scene distracts the gradient saliency detection.

We additionally compare the performance of the proposed method with single image layer separation techniques, as shown in Figure 6. Yu and Brown’s method [27] is not appropriate for our problem setting, because it as-

sumes that the reflection layer is smoother than the background scene. Edges separation of the method in [13] can be provided by manual user markup or the method in [26]. In our comparison, we use the method in [26] for edge selection. However, the rest is unchanged from [13] including Laplacian mixture distribution model and the optimization method. Nevertheless, the method in [13] produces unsatisfactory results with inconsistent color reconstruction.

**Quantitative Comparison** In order to validate the robustness of the proposed algorithm quantitatively, we test a few synthetic frames using *CamVid* dataset [2, 3]. Figure 7



Figure 7: Comparison on the synthetic frames. (a) Selected input frames. (b) Result of [26]. (c) Result of [18]. (d) Result of the proposed method. (e) Ground truth. Video results including this example are provided in the supplementary material.

Table 1: RMSE comparison using *CamVid* dataset [2, 3].

| Algorithm            | RMSE           |                |                | Mean  |
|----------------------|----------------|----------------|----------------|-------|
|                      | <i>seq06R0</i> | <i>seq16E5</i> | <i>seq05VD</i> |       |
| Proposed             | 29.14          | 28.63          | 28.67          | 28.81 |
| Yu and Brown [26]    | 51.48          | 52.43          | 49.85          | 51.26 |
| Sarel and Irani [18] | 49.52          | 52.47          | 56.52          | 52.83 |

shows the visual comparison, which shows clearly that the proposed method outperforms the others. In Table 1, the root-mean-square error (RMSE) is calculated for each selected frame to prove that the proposed method has low error rate than those of the previous methods in [18, 26]. Although [18] may seem to produce comparable visual results visually, Table 1 obviously shows that the RMSE error is significantly higher than the proposed method.

**Computation time** To evaluate the computation time, an additional experiment is performed using video frames with  $512 \times 288$  resolution. The proposed algorithm consumes approximately 2 seconds, while the methods in [26] and [18] spend about 2 minutes and 7 seconds on the same data, respectively. The faster computation time of the proposed method is achieved because LUT is applied to speed up the performance. Moreover, LUT is particularly useful when there are a lot of images to be processed because it does not need to recompute the same calculation redundantly.

**Limitation** The proposed method has a limitation because the use of average image prior and the assumption on static reflection restrain the application in general cases. The proposed method is mostly applicable to separate the

reflection when it is relatively static or slightly moving, which is the case for the forward motion videos captured by car black boxes. Therefore, the limitation is not a major concern in using the proposed algorithm in vehicular application we focus on.

## 5. Conclusion

Reflection on a glassy surface is a typical problem in reality. The complexity of removing the reflection increases particularly when dealing with outdoor environment. In this paper, the average image prior was specifically proposed to remove the reflection on the windscreen of a vehicle. A real world black box video was used in our experiment to compare the performance of the proposed method with the previous methods. Experimental results validated the robustness of the proposed method in separating the windscreen reflection and producing clear background images. In future, we will extend the current algorithm to complete the background region occluded by saturated reflection or by car dashboard. We believe it becomes possible by developing accurate flow estimation and image warping algorithms in spatio-temporal domain under fast forward motion.

## Acknowledgement

This work was supported by the IT R&D program of MSIP/KEIT. [10047078, 3D reconstruction technology development for scene of car accident using multi view black box image]. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. NRF-2013R1A2A2A01069181).

## References

- [1] A. Agrawal, R. Raskar, S. K. Nayar, and Y. Li. Removing photography artifacts using gradient projection and flash-exposure sampling. *ACM Trans. on Graphics*, 24(3):828–835, July 2005.
- [2] G. J. Brostow, J. Fauqueur, and R. Cipolla. Semantic object classes in video: A high-definition ground truth database. *Pattern Recognition Letters*, 30(2):88–97, January 2009.
- [3] G. J. Brostow, J. Shotton, J. Fauqueur, and R. Cipolla. Segmentation and recognition using structure from motion point clouds. In *Proc. of European Conference on Computer Vision*, pages 44–57, 2008.
- [4] W. Dong, X. Li, L. Zhang, and G. Shi. Sparsity-based image denoising via dictionary learning and structural clustering. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pages 457–464, 2011.
- [5] H. Farid and E. H. Adelson. Separating reflections and lighting using independent components analysis. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pages 261–267, 1999.
- [6] K. Gai, Z. Shi, and C. Zhang. Blind separation of superimposed moving images using image statistics. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 34(1):19–32, January 2012.
- [7] X. Guo, X. Cao, and Y. Ma. Robust separation of reflection from multiple images. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pages 2195–2202, 2014.
- [8] N. Joshi, C. L. Zitnick, R. Szeliski, and D. Kriegman. Image deblurring, and denoising using color priors. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pages 1550–1557, 2009.
- [9] N. Kong, Y.-W. Tai, and J. S. Shin. A physically-based approach to reflection separation: from physical modeling to constrained optimization. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 36(2):209–221, February 2014.
- [10] N. Kong, Y.-W. Tai, and S. Y. Shin. High-quality reflection separation using polarized images. *IEEE Trans. on Image Processing*, 20(12):3393–3405, December 2011.
- [11] D. Krishnan and R. Fergus. Fast image deconvolution using hyper-laplacian priors. In *Proc. of Advances in Neural Information Processing Systems*, pages 1033–1041, 2009.
- [12] A. Levin, R. Fergus, F. Durand, and W. T. Freeman. Image and depth from a conventional camera with a coded aperture. *ACM Trans. on Graphics*, 26(3):70:1–70:9, July 2007.
- [13] A. Levin and Y. Weiss. User assisted separation of reflections from a single image using a sparsity prior. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 29(9):1647–1654, September 2007.
- [14] C. Liu, J. Yuen, A. Torralba, J. Sivic, and W. T. Freeman. SIFT flow: Dense correspondence across different scenes. In *Proc. of European Conference on Computer Vision*, pages 28–42, 2008.
- [15] M. Nikolova and M. K. Ng. Analysis of half-quadratic minimization methods for signal and image recovery. *SIAM Journal on Scientific Computing*, 27(3):937–966, October 2005.
- [16] E. Reinhard, M. Adhikhmin, B. Gooch, and P. Shirley. Color transfer between images. *Computer Graphics and Applications*, 21(5):34–41, September 2001.
- [17] S. Roth and M. J. Black. Fields of experts: a framework for learning image priors. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pages 860–867, 2005.
- [18] B. Sarel and M. Irani. Separating transparent layers through layer information exchange. In *Proc. of European Conference on Computer Vision*, pages 328–341, 2004.
- [19] B. Sarel and M. Irani. Separating transparent layers of repetitive dynamic behaviors. In *Proc. of IEEE International Conference on Computer Vision*, pages 26–32, 2005.
- [20] Y. Schechner, N. Kiryati, and R. Basri. Separation of transparent layers using focus. In *Proc. of IEEE International Conference on Computer Vision*, pages 1061–1066, 1998.
- [21] R. von Gioi, J. Jakubowicz, J.-M. Morel, and G. Randall. LSD: A fast line segment detector with a false detection control. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 32(4):722–732, April 2010.
- [22] M. J. Wainwright and E. P. Simoncelli. Scale mixtures of Gaussians and the statistics of natural images. In *Proc. of Advances Neural Information Processing Systems*, pages 855–861, 2000.
- [23] B. Wang and Z. Tu. Sparse subspace denoising for image manifolds. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pages 468–475, 2013.
- [24] Y. Weiss. Deriving intrinsic images from image sequences. In *Proc. of IEEE International Conference on Computer Vision*, pages 68–75, 2001.
- [25] L. Xu, S. Zheng, and J. Jia. Unnatural L0 sparse representation for natural image deblurring. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pages 1107–1114, 2013.
- [26] L. Yu and M. S. Brown. Exploiting reflection change for automatic reflection removal. In *Proc. of IEEE International Conference on Computer Vision*, pages 2432–2439, 2013.
- [27] L. Yu and M. S. Brown. Single image layer separation using relative smoothness. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pages 2752–2759, 2014.
- [28] H. Zhang, D. Wipf, and Y. Zhang. Multi-image blind deblurring using a coupled adaptive sparse prior. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pages 1051–1058, 2013.