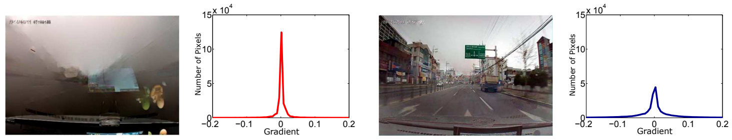
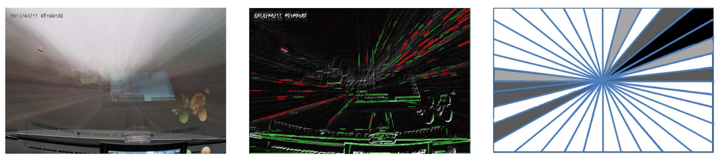
Reflection Removal for In-Vehicle Black Box Videos

Christian Simon[1], In Kyu Park[1]
[1]Department of Information and Communication Engineering, Inha University



(a) (b)

Figure 1: 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.



(a) (b) (c)

Figure 2: Line segments and circular region division. (a) The average image. (b) Detected line segments to find the vanishing point from the average image. Red line segments are converging to the vanishing point while green ones are not. (c) An example of circularly divided regions (with different α values in different shades).

Figure 3: Frame results from three different videos. First, second, and third columns are input images, background layers, and reflection layers, respectively. The first and third rows are from real world black box videos, while the second row is from a synthetic video.

The in-vehicle black box camera (dashboard camera) has become a popular device in many countries for security monitoring and event capturing but the content is often degraded due to the windscreen reflection of objects inside. In the literature, there are several existing methods for reflection removal that have been proposed using several images such as the methods proposed by Sarel and Irani [3] and Yu and Brown's [4]. However, there is no reflection removal method focusing on black box videos, thus, a novel method is proposed to remove the reflection on the windscreen in the outdoor environment from in-vehicle black box videos.

We use the basic equation inspired from [2] that the joint probability ($\Pr(L,R) = \Pr(L)\Pr(R)$) of two layers: background layer L and reflection layer R should be maximized. Therefore, it is equivalent to minimize $E(L,R)$ as follows.

$$E(L,R) = \sum_{i \in N} \sum_{j \in J} \left(F_1\left((L \oplus g^j)_i\right) + F_2\left((R \oplus g^j)_i\right) \right) \tag{1}$$

where $\oplus$ is the convolution operator to the derivative filters g^j and $F_1(x)$ and $F_2(x)$ denote negative log functions applied for distribution model of each layer.

The method exploits the spatio-temporal coherence of the reflection stating that a vehicle moves forward while the reflection layer of the internal object remains static. Based on this observation, we can take average from the frame sequence to produce an extremely low-passed image in which the reflection still remains sharp, named as *average image prior* in this paper. The average image is denoted as $\bar{I}$.

The objective function is inspired from the reflection layer and background layer distribution as shown in Figure 1. Figure 1(a) shows that the reflection layer has a higher peak than background layer, thus, hyper-Laplacian ($p(x) = e^{-k|x|^{\alpha}}/s$ with $\alpha < 1$) is utilized to model gradient distribution in the reflection layer. Additionally, we modified the hyper-Laplacian to adapt with the condition of the average image. The background layer consists of regions with different properties of gradient sparsity, which includes the ground, building/wall/green, and sky areas. These regions with different properties create different gradient structure in the average image. Consequently, the average image is divided into several regions (36 parts in the example) as shown in Figure 2 (c). Then, these regions are grouped into H groups as indicated with different shades. Region-based hyper-Laplacian is performed to give different sparse penalty with the adjustment to α value in a set of pixels (N_h) from regions in group h. Hence, the objective function is stated as follows.

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

where J_1 and J_2 are first order derivative filters and second order derivative filters in horizontal and vertical direction. The derivative filters are denoted as $D_i^j X \equiv X \oplus g^j$ for $j \in J$. The objective function is simplified with a half-quadratic method. The optimization for this problem is adopted from [1] that the optimization technique is fast and reliable with the use of FFT and lookup table(LUT).

The experiment is conducted to both real world and synthetic videos. Figure 3 shows the visual performance of the proposed method in several videos; more results can be found in the paper and the supplementary material. The limitation of the proposed approach is that it cannot handle rapidly moving reflection because of the stationary assumption. We conclude that it is possible to achieve reliable reflection removal from black box videos with the simplified assumptions and the proposed objective function.

[1] 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.

[2] 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.

[3] B. Sarel and M. Irani. Separating transparent layers through layer information exchange. In *Proc. of European Conference on Computer Vision*, pages 328–341, 2004.

[4] 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.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.