

Eliminating artifacts when inverting visual reverberations

Yaron Diamant and Yoav Y. Schechner

Dept. of Electrical Engineering

Technion - Israel Institute of Technology

Haifa 32000, Israel

yarondi@rafael.co.il , yoav@ee.technion.ac.il

Abstract

Reference [1] approximates the imaging model as orthographic. This report describes a consequence of perspective, when compared to the orthographic approximation. Slight misalignments are created between the true visual reverberations, and the shifts of the δ functions used in the recovery filters of Ref. [1]. The result is that reconstruction of L_r as presented in Sec. 6.2 of Ref. [1] contains weak residual edge artifacts. We present here a method that overcomes these artifacts. It is based on a variation of the method of [2] for eliminating inconsistent edges.

1. Perspective

Reference [1] approximates the imaging model as orthographic. We now describe a consequence of perspective, when compared to the orthographic approximation. Fig. 1 illustrates two possible paths of light rays originating from object L_r . The first path is termed *path 1*: rays taking this path hit the *front* interface of the window, and reflect to the camera, where they create an image. The second path is termed *path 2* in Fig. 1. Rays taking *path 2* hit the *back* interface of the window en route to the camera. These light rays create a shifted replica of the same image.

The total length of *path 2* is *greater* than the total length of *path 1*. The difference in the lengths of the paths is depicted in green color in Fig. 1. Due to the longer distance, light rays following *path 2* create a *smaller* image than light rays that following *path 1*. This effect repeats itself in the higher orders of the reverberations. The same phenomenon of different image sizes is valid for L_t as well. This is a result of perspective: if the projection was orthographic, then the changed path length would not have mattered. In our experimental setup, the difference of the paths was about 1% of the total path length. This path difference is sufficient to create a noticeable difference in the size of the replicas, up to several pixels.

Let us consider how different pixels in the original image are reverberated. Consider Fig. 2(a). A pixel in the upper

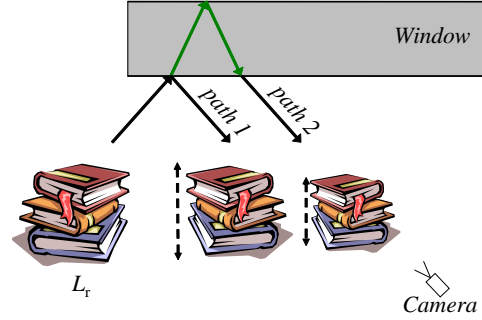


Figure 1. A light ray that creates the first image of the reverberation (denoted by *path 1*) travels a shorter way from the object L_r to the camera, than the ray that creates the second image of the reverberation (denoted by *path 2*). Thus, the second ray creates a smaller image. The difference in the paths is marked in green.

part of L_r reverberates to slightly lower pixels in the acquired frame. On the other hand, in Fig. 2(b), a pixel at the bottom of L_r reverberates to higher pixels in the acquired frame. This is the result of image size reduction, caused by perspective. Thus, different pixels generally have a direction of reverberation that is slightly different than others.

These observations yield two differences between this image formation model and the orthographic model presented in Sec. 2 of Ref. [1]. First, the perspective model has two dimensional (2D) shifts, since pixels reverberate to heights slightly different than their original one. Second, the model is *not* spatially invariant, since different pixels reverberate to different heights, in a shift that depends on the original location. Thus, a model of simple space-invariant convolution is not strictly correct, but an approximation.

2. Residual Edge Artifacts

This section shows residual edge artifacts that are created when recovery is not in full consistency with a simple orthographic approximation. In reality, the image formation

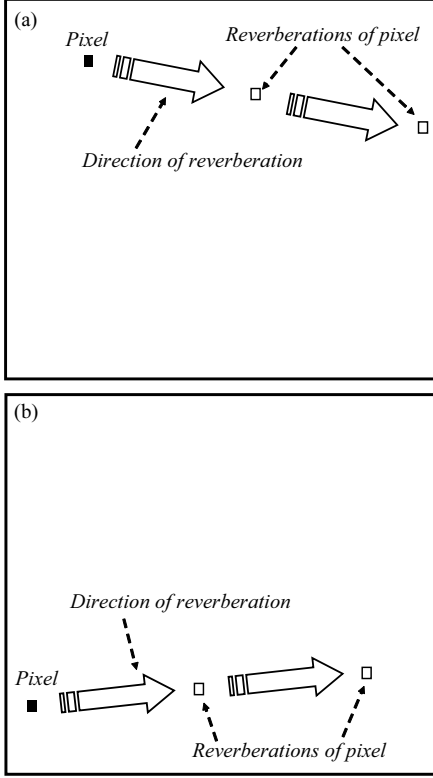


Figure 2. Reverberation in a perspective model. (a) The pixel depicted in this figure is at the upper half of the frame. This pixel is reverberated to lower pixels. (b) The pixel depicted here is at the lower part of the frame. This pixel is reverberated to higher pixels.

is perspective, but the recovery model assumes a space invariant model. This creates slight misalignments between the true visual reverberations, and the shifts of the δ functions used in the recovery filters.

Fig. 3 presents the reconstructed \hat{L}_r obtained directly from the method described in Sec 4.4 of Ref. [1], i.e., a method that assumes a spatially invariant (orthographic) model. Indeed, the small misalignment described, causes false edges in the reconstructed \hat{L}_r . This can be seen, for example, in the marked rectangular regions in Fig. 3. These residual false edges are *weak*, but they can be noticed.

We note that this effect is practically insignificant in \hat{L}_t . The reason is that, as described in Ref. [1], the reverberation intensity of L_t is typically significantly weaker than of L_r . Hence, the recovery of L_t is effectively a mild, stable operation, that does not emphasize noise effects.

3. Perturbed Solution

As described in Sec. 1, different pixels are reverberated to different heights. However, for the moment, let all pixels reverberate in a uniform direction and extent, for example

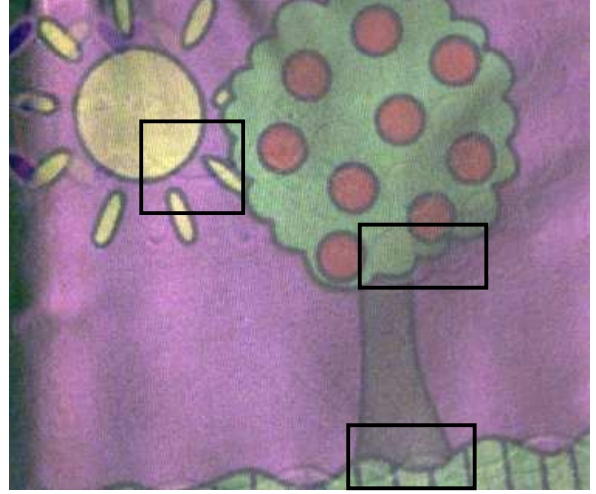


Figure 3. The reconstructed \hat{L}_r under the assumption of orthographic projection. It can be seen that the result contains residual edge artifacts, seen for example in the marked rectangular regions.

30 pixels to the right and 2 pixels up, while the model assumed only a horizontal $d = 30$. Hence, the horizontal shift is *perturbed* by a slight vertical shift of $\Delta y = 2$ pixels. Still, the mathematical solution presented in Sec. 4.4 of Ref. [1] applies: now p is a 2D filter¹

$$p^{\Delta x, \Delta y}(x, y) = a\delta(x, y) + b\delta(x - d + \Delta x, y + \Delta y). \quad (1)$$

where Δx is a horizontal perturbation, which in this example is null. The convolution filter $p^{\Delta x, \Delta y}$ in this special case has the form

$$p^{\Delta x, \Delta y}(x, y) = \begin{pmatrix} 0 & 0 & \cdots & b \\ 0 & \cdots & \cdots & 0 \\ a & \cdots & 0 & 0 \end{pmatrix}. \quad (2)$$

Then, we use Eq. (29) in Ref. [1],

$$\hat{L}_r = \arg \min_{L_r} \left(\|U - L_r \star p^{\Delta x, \Delta y}\|^2 + \lambda \|\nabla^2 L_r\|^2 \right). \quad (3)$$

In reality, the perturbation is not uniform, since the misalignment of the model is not spatially invariant. Therefore Applying Eq. (3) yields good reconstruction only at locations where the specific perturbation is valid. In other locations, we get a degradation of the reconstruction, in the form of new or stronger residual false edges (See Fig. 4).

4. Derivative-based Edge Removal

This section reviews a derivative-based algorithm for edge-removal, based on a series of images [2]. In Sec. 3

¹Alternatively, we can rotate the local coordinate system and the acquired frame accordingly, such that the model becomes one dimensional locally.



Figure 4. Reconstruction using $p^{\Delta x, \Delta y}$, when $\Delta x = 0$ and $\Delta y = 2$. Many new false edge artifacts are created in this perturbed result. This is seen by comparing to Fig. 3, in which $\Delta x = \Delta y = 0$.

we modify the algorithm of [2] to remove residual artifacts from the reconstructed images. Let $\mathbf{L}^{(k)}$, $k = 1 \dots K$ be a series of K images. The images are similar. They differ by the presence of unwanted edges in different locations. On the other hand, the desired edges appear in the same locations in most of the frames of the series. In [2], the series of images differed only by the presence of different shadows in different locations. The goal of [2] was to remove these shadows.

The spatial derivatives of the images are

$$g_x^{(k)}(x, y) = \frac{\partial L^{(k)}(x, y)}{\partial x}, \quad g_y^{(k)}(x, y) = \frac{\partial L^{(k)}(x, y)}{\partial y}. \quad (4)$$

In [2], new spatial derivatives were computed by

$$g_x^{(\text{new})}(x, y) = \text{median}_k \left[g_x^{(k)}(x, y) \right], \quad (5)$$

$$g_y^{(\text{new})}(x, y) = \text{median}_k \left[g_y^{(k)}(x, y) \right]. \quad (6)$$

These new derivatives preserve the consistent edges in the frames, but eliminate edges that appear only in a minority of the frames. Thus, the algorithm sets a gradient for a desired image $L^{(\text{new})}$, by

$$\frac{\partial L^{(\text{new})}(x, y)}{\partial x} \approx g_x^{(\text{new})}(x, y), \quad (7)$$

$$\frac{\partial L^{(\text{new})}(x, y)}{\partial y} \approx g_y^{(\text{new})}(x, y). \quad (8)$$

Eqs. (7,8) yield a set of equations for reconstructing $L^{(\text{new})}$ based on its spatial derivatives. This set of equations

is overconstrained. To solve it, define

$$\mathbf{f}_x = \begin{bmatrix} 1 & -1 \end{bmatrix}, \quad \mathbf{f}_x^{(\text{flip})} = \begin{bmatrix} -1 & 1 \end{bmatrix}, \quad (9)$$

$$\mathbf{f}_y = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \mathbf{f}_y^{(\text{flip})} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}. \quad (10)$$

Now, suppose there is a filter \mathbf{q} for which

$$\mathbf{q} \star \left[\sum_{m=\{x,y\}} \left(\mathbf{f}_m^{(\text{flip})} \star \mathbf{f}_m \right) \right] = \delta(x, y), \quad (11)$$

where \star denotes a convolution. An approximate filter \mathbf{q} is described in [2]. The solution for Eq. (7) is given [2] by

$$\mathbf{L}^{(\text{new})} = \mathbf{q} \star \sum_{m=\{x,y\}} \mathbf{f}_m^{(\text{flip})} \star \mathbf{g}_m^{(\text{new})}. \quad (12)$$

5. Overcoming Residual Artifacts

We now describe our method for overcoming the residual edge artifacts. We make a series of perturbed reconstructions, each using different $\{\Delta x, \Delta y\}$ values, i.e., using differently perturbed filters $p^{\Delta x, \Delta y}$. This results in a series of perturbed solutions. All these solutions are fused in a manner similar to the derivative-based edge-removal method described in Sec. 4. Frames having the smallest derivative (locally) are defined as

$$k_x(x, y) = \arg \min_k \left| g_x^{(k)}(x, y) \right| \quad (13)$$

and

$$k_y(x, y) = \arg \min_k \left| g_y^{(k)}(x, y) \right|. \quad (14)$$

We modify the algorithm described in Sec. 4 by replacing Eqs. (5) and (6) by

$$g_x^{(\text{new})}(x, y) = g_x^{[k_x(x,y)]}(x, y) \quad (15)$$

and

$$g_y^{(\text{new})}(x, y) = g_y^{[k_y(x,y)]}(x, y). \quad (16)$$

Then, we use the result of Eqs. (15,16) in Eq. (12), to obtain the final result.

To illustrate this, we demonstrate the edge-removal algorithm on the area marked by the upper rectangle in Fig. 3. Define the filters

$$\begin{aligned} p^{(1)} &= \begin{pmatrix} a & 0 & \cdots & 0 & b \end{pmatrix} \\ p^{(2)} &= \begin{pmatrix} 0 & 0 & \cdots & 0 & \frac{b}{2} \\ a & 0 & \cdots & 0 & \frac{b}{2} \end{pmatrix} \\ p^{(3)} &= \begin{pmatrix} a & 0 & \cdots & 0 & \frac{b}{2} \\ 0 & 0 & \cdots & 0 & \frac{b}{2} \end{pmatrix}. \end{aligned} \quad (17)$$

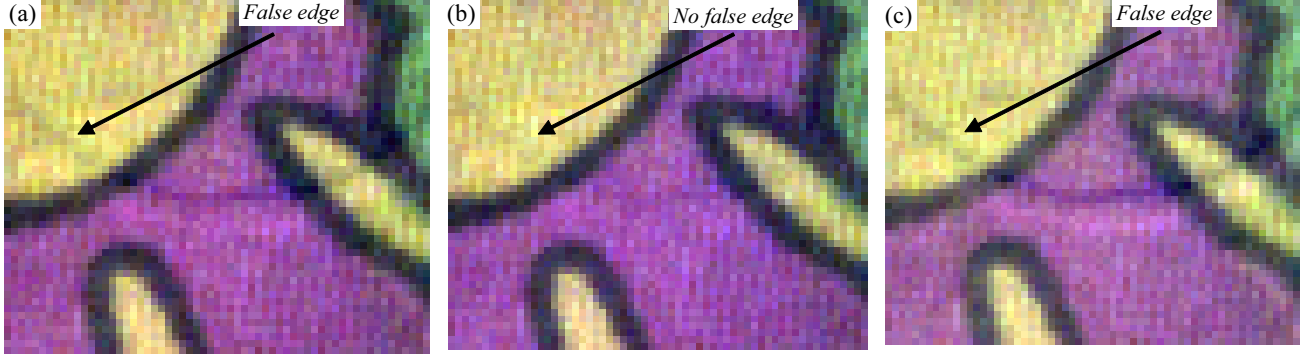


Figure 5. (a) A part of $\hat{L}_r^{(1)}$ reconstructed using $p^{(1)}$. A weak residual false edge is marked. (b) A part of $\hat{L}_r^{(2)}$ reconstructed using $p^{(2)}$. Here this residual false edge does *not* appear in the marked location. (c) A part of $\hat{L}_r^{(3)}$ reconstructed using $p^{(3)}$. Here, as in (a), a weak residual false edge is marked.

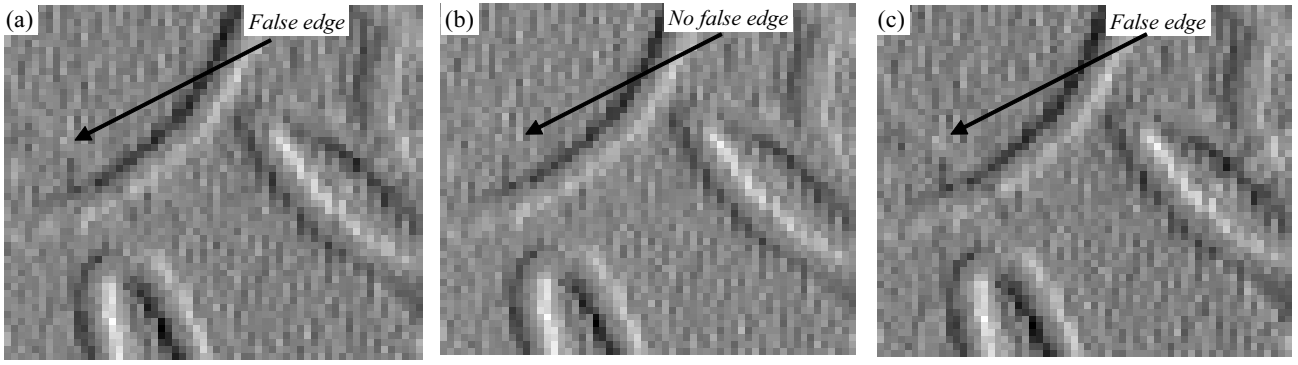


Figure 6. (a) The image $\mathbf{g}_x^{(1)}$. It is the x -derivative of $\hat{L}_r^{(1)}$. A weak residual false edge is marked. (b) The image $\mathbf{g}_x^{(2)}$. It is the x -derivative of $\hat{L}_r^{(2)}$. Notice there is *no* false edge in this marked location. (c) The image $\mathbf{g}_x^{(3)}$. It is the x -derivative of $\hat{L}_r^{(3)}$. Here, as in (a), a weak residual false edge is marked.

Use of $p^{(1)}$ yields a result denoted by $\hat{L}_r^{(1)}$. The filter $p^{(2)}$ corresponds to reverberation of the pixels 30 pixels to the right, and $\sim \frac{1}{2}$ a pixel up. Using it yields $\hat{L}_r^{(2)}$. Similarly, the filter $p^{(3)}$ corresponds to a slight downward shift of $\sim \frac{1}{2}$ pixel, and its result is denoted by $\hat{L}_r^{(3)}$. Fig. 5 shows close-ups of $\hat{L}_r^{(1)}$, $\hat{L}_r^{(2)}$ and $\hat{L}_r^{(3)}$ in one of our experiments.

The corresponding derivative fields of $\hat{L}_r^{(1)}$, $\hat{L}_r^{(2)}$ and $\hat{L}_r^{(3)}$ were derived using Eq. (4). Fig. 6 shows close-ups of $\mathbf{g}_x^{(1)}$, $\mathbf{g}_x^{(2)}$ and $\mathbf{g}_x^{(3)}$. A weak residual false edge artifact is marked in $\hat{L}_r^{(1)}$, $\hat{L}_r^{(3)}$ and consequently in $\mathbf{g}_x^{(1)}$, $\mathbf{g}_x^{(3)}$. This false edge is not apparent in $\hat{L}_r^{(2)}$ (and $\mathbf{g}_x^{(2)}$).

The field $\mathbf{g}_x^{(\text{new})}$ was derived using Eq. (15), and is shown in Fig. 7. Indeed, the false edge artifact does not appear there. At the same time, the details and edges that consistently appeared in all the perturbed solutions are preserved in $\mathbf{g}_x^{(\text{new})}$. Consequently, the final recovery \hat{L}_r obtained using Eq. (12), and similarly \hat{L}_t are largely clean of edge artifacts. They are shown in Figs. 8 and 9 respectively.

This residual edge-removal process is automatically performed in our other experiments as well.

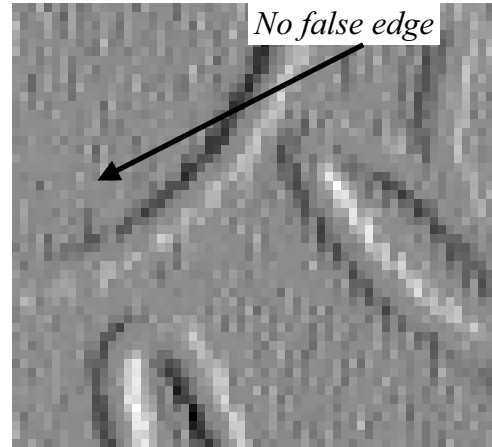


Figure 7. The image $\mathbf{g}_x^{(\text{new})}$, defined in Eq. (15). Notice the *elimination* of the false edge of Figs. 6(a) and 6(c). On the other hand, the consistent features are maintained.

Acknowledgments

Yoav Schechner is a Landau Fellow - supported by the Taub Foundation. The work was supported by the Israeli Ministry



Figure 8. Reconstruction of L_r using the fusion of perturbed solutions. Reverberations, residual artifacts and a mixture caused by L_t are eliminated. The true structure is maintained.



Figure 9. L_t reconstructed in the experiment.

of Science, Culture and Sport (Grant 3-3426). It was conducted in the Ollendorff Minerva Center. Minerva is funded through the BMBF.

References

- [1] Y. Diamant and Y. Y. Schechner. Overcoming visual reverberations. In *Proc. IEEE CVPR* 2008. 1, 2
- [2] Y. Weiss. Deriving intrinsic images from image sequences. In *Proc. IEEE ICCV*, vol. 2, 68–75, 2001. 1, 2, 3