Technical Report for Obstruction-free Photography

Tianfan Xue, MIT CSAIL
Sept 2015

1 Formulation

In this report, we use a lower-case letter a to denote a scalar, a bold lower-case letter $\mathbf{a}$ to denote a 2×1 vector, a normal capital letter A to denote a vector, and a bold capital letter $\mathbf{A}$ to denote a matrix. We denote the matrix product as $\mathbf{A}B$, where $\mathbf{A} \in \mathbb{R}^{n \times m}$ and $B \in \mathbb{R}^m$, and the element-wise product of two vectors as $A \circ B$, where $A, B \in \mathbb{R}^n$.

The task of obstruction removal is to decompose an input sequence $\{I^t\}$ into the background and obstruction components, I_B and I_O. We will first derive an algorithm for the more general case of an unknown, spatially varying alpha map, A, and then show a small simplification that can be used for reflection removal where we assume the alpha map is constant. Let $\{V_O^t\}$ and $\{V_B^t\}$ are the sets of motion vectors for the obstruction and background components, respectively. According to the Section 4.1 of [2].

$$\min_{I_O, I_B, A, \{V_O^t\}, \{V_B^t\}} \sum_t \|I^t - \mathbf{W}(V_O^t) I_O - \mathbf{W}(V_O^t) A \circ \mathbf{W}(V_B^t) I_B\|_1$$
$$+ \lambda_1 \|\nabla A\|_2^2 + \lambda_2 (\|\nabla I_O\|_1 + \|\nabla I_B\|_1) + \lambda_3 L(I_O, I_B) + \lambda_4 \sum_t \|\nabla V_O^t\|_1 + \|\nabla V_B^t\|_1$$

Subject to:

$$0 \leqslant I_O, I_B, A \leqslant 1, \tag{1}$$

where $I^t, I_B, I_R, A \in \mathbb{R}^n$, temporal index $t \in \{1, 2, ..., T\}$, and n is number of pixels per frame. $V_O, V_B \in \mathbb{R}^{2N}$, and $\mathbf{W}(V_B^t) \in R^{n \times n}$ is a warping matrix such that $\mathbf{W}(V_B^t) I_B$ is the warped background component I_B according to the motion field V_B^t. We define $L(I_O, I_B) = \sum_x \|\nabla I_O(x)\|^2 \|\nabla I_B(x)\|^2$, where x is the spatial index and $\nabla I_B(x)$ is the gradient of image I_B at position x. Also, recall that we use $\circ$ to denote element-wise product.

2 Optimization

We use an alternating gradient descent method to solve Eq. 1. We first fix the motion fields $\{V_O^t\}$ and $\{V_B^t\}$ and solve for I_O, I_B and A, and then fix I_O, I_B and A, and solve for $\{V_O^t\}$ and $\{V_B^t\}$. Similar alternating gradient descent approach for joint estimation has been used in video super resolution [1].

2.1 Decomposition step: fix motion fields $\{V_O^t\}$ and $\{V_B^t\}$, and solve for I_O, I_B, and A.

In this step, we ignore all the terms in Eq. 1 that only consist of V_O^t and V_B^t: Ignoring all the terms that only contain $\mathbf{v}_B^t$ and $\mathbf{v}_R^t$ in Eq. 1, we get:

$$\min_{\{I_O, I_B, A\}} \sum_t \|I^t - \mathbf{W}_O^t I_O - \mathbf{W}_O^t A \circ \mathbf{W}_B^t I_B\|_1 + \lambda_1 \|\nabla A\|^2$$
$$+ \lambda_2 (\|\nabla I_O\|_1 + \|\nabla I_B\|_1) + \lambda_3 L(I_O, I_B), \tag{2}$$

Subject to

$$0 \leqslant I_O, I_B, A \leqslant 1, .$$

where we denote $\mathbf{W}(V_B^t)$ and $\mathbf{W}(V_O^t)$ as $\mathbf{W}_B^t$ and $\mathbf{W}_O^t$ respectively for simplicity.

We solve this problem using a modified version of iterative reweighted least squares (IRLS). The original IRLS algorithm is designed for a non-constrained optimization with only l_1-and l_2-norms. To get this form, we linearize the higher-order terms in the objective function in Eq. 2.

First, we approximate the non-smooth L1-norm by the smooth function $\phi(x) = \sqrt{x + \epsilon^2}$, where ϵ is a very small number:

$$\min_{I_O, I_B, A} \sum_t \phi(\|I^t - \mathbf{W}_O^t I_O - \mathbf{W}_O^t A \circ \mathbf{W}_B^t I_B\|^2) + \lambda_1 \|\nabla A\|^2$$
$$+ \lambda_2 \left(\phi(\|D_x I_O\|^2 + \|D_y I_O\|^2) + \phi(\|D_x I_B\|^2 + \|D_y I_B\|^2)\right) + \lambda_3 L(I_O, I_B), \tag{3}$$

Subject to

$$0 \leqslant I_O, I_B, A \leqslant 1, .$$

where $\mathbf{D}_x, \mathbf{D}_y \in \mathbb{R}^{n \times n}$ are the derivative matrices defined as: $(\mathbf{D}_x I_B^t)(\mathbf{x}) = \frac{\partial I_B^t(\mathbf{x})}{\partial x}$ and $(\mathbf{D}_y I_B^t)(\mathbf{x}) = \frac{\partial I_B^t(\mathbf{x})}{\partial y}$ (note that $\mathbf{D}_x I_B^t \in \mathbb{R}^n$).

Let $\hat{I}_O$, $\hat{I}_B$ and $\hat{A}$ be the obstruction component, the background component, and the alpha map of the last iteration, respectively. Here we use another approximation used in IRLS:

$$\phi(x^2) = \sqrt{x^2 + \epsilon^2} \approx \frac{x^2}{\sqrt{x^2 + \epsilon^2}} = \frac{x^2}{\phi(x^2 + \epsilon^2)} \approx= \frac{x^2}{\phi(\hat{x}^2 + \epsilon^2)} \tag{4}$$

where $\hat{x}$ is very close to x. Eq. 4 can be further simplified as:

$$\min_{I_O, I_B, A} \sum_t \frac{\|I^t - \mathbf{W}_O^t I_O - \mathbf{W}_O^t A \circ \mathbf{W}_B^t I_B\|^2}{\phi(\|I^t - \mathbf{W}_O^t \hat{I}_O - \mathbf{W}_O^t \hat{A} \circ \mathbf{W}_B^t \hat{I}_B\|^2)} + \lambda_1 \|\nabla A\|^2$$
$$+ \lambda_2 \left(\frac{\|D_x I_O\|^2 + \|D_y I_O\|^2}{\phi(\|D_x \hat{I}_O\|^2 + \|D_y \hat{I}_O\|^2)} + \frac{\|D_x I_B\|^2 + \|D_y I_B\|^2}{\phi(\|D_x \hat{I}_B\|^2 + \|D_y \hat{I}_B\|^2)} \right) + \lambda_3 L(I_O, I_B), \tag{5}$$

Subject to

$$0 \leqslant I_O, I_B, A \leqslant 1, .$$

Note all the denominators are constant in Eq. 6. For simplicity, we denote:

$$w_1 = \phi(\|I^t - \mathbf{W}_O^t \hat{I}_O - \mathbf{W}_O^t \hat{A} \circ \mathbf{W}_B^t \hat{I}_B\|^2)^{-1}, \tag{6}$$
$$w_2 = \phi(\|D_x \hat{I}_B\|^2 + \|D_y \hat{I}_B\|^2)^{-1}, \tag{7}$$
$$w_3 = \phi(\|D_x \hat{I}_O\|^2 + \|D_y \hat{I}_O\|^2)^{-1}. \tag{8}$$
$$\tag{9}$$

The weights $\mathbf{w}_1$, $\mathbf{w}_2$, and $\mathbf{w}_3$ for each pixel are different, so that their vectors, not scalar.

Second, we linearize all the higher-order terms using first-order Taylor expansion. The data term is linearized as[1]

$$I^t - \mathbf{W}_O^t I_O - \mathbf{W}_O^t A \circ \mathbf{W}_B^t I_B = I^t - \mathbf{W}_O^t I_O - \mathbf{W}_O^t A \circ \mathbf{W}_B^t \hat{I}_B - \mathbf{W}_O^t \hat{A} \circ \mathbf{W}_B^t I_B + \mathbf{W}_O^t \hat{A} \circ \mathbf{W}_B^t \hat{I}_B. \tag{10}$$

Here we use an approximation commonly used in optimization:

$$xy \approx x\hat{y} + \hat{x}y - \hat{x}\hat{y}, \tag{11}$$

where $\hat{x}$ is very close to x and $\hat{y}$ is very close to y.

We also linearize the edge ownership term as:

$$\lambda_3 L(I_O, I_B) = \lambda_3 \sum_x \|\nabla I_O\|^2 \|\nabla I_B\|^2 \approx \sum_x \|\nabla \hat{I}_O\|^2 \|\nabla I_B\|^2 + \|\nabla I_O\|^2 \|\nabla \hat{I}_B\|^2 - \|\nabla \hat{I}_O\|^2 \|\nabla \hat{I}_B\|^2$$
$$= \lambda_3 L(\hat{I}_O, I_B) + L(I_O, \hat{I}_B) - L(\hat{I}_O, \hat{I}_B). \tag{12}$$

At last, we incorporate the two inequality constraints into our objective function using the penalty method [?]. For example, for the non-negativity constrain $I_B \geqslant 0$, we include the following penalty function into the objective function:

$$\lambda_p \min(0, I_B)^2 \tag{13}$$

[1] we make an approximation commonly used in optimization: $xy \approx x\hat{y} + \hat{x}y - \hat{x}\hat{y}$, where $\hat{x}$ is very close to x and $\hat{y}$ is very close to y.

where I_B^2 denotes element-wise square and λ_p is the weight for the penalty (we fix $\lambda_p = 10^5$). This function will apply a penalty proportional to the negativity of I_B (and will be zero if I_B is nonnegative). Then we add the following terms to enforce all the inequility constrains.

$$\lambda_p\big(\min(0,I_B)^2+\min(0,1-I_B)^2+\min(0,I_O)^2+\min(0,1-I_O)^2+\min(0,A)^2+\min(0,1-A)^2\big) \quad (14)$$

Plugging Eq. 10, Eq. 12, and Eq. 14 into Eq. 2, we have:

$$\begin{aligned}\min_{I_O,I_B,A}\sum_t & w_1(\|I^t-\mathbf{W}_O^t I_O-\mathbf{W}_O^t A\circ\mathbf{W}_B^t\hat{I}_B-\mathbf{W}_O^t\hat{A}\circ\mathbf{W}_B^t I_B+\mathbf{W}_O^t\hat{A}\circ\mathbf{W}_B^t\hat{I}_B\|+\|\nabla A\|^2\\ &+\lambda_2\left(w_2(\|D_x I_O\|^2+\|D_y I_O\|^2)+w_3(\|D_x I_B\|^2+\|D_y I_B\|^2)\right)+\lambda_3 L(I_O,I_B) \qquad (15)\\ &+\lambda_p\Big(\min(0,I_B)^2+\min(0,1-I_B)^2+\min(0,I_O)^2+\min(0,1-I_O)^2+\min(0,A)^2+\min(0,1-A)^2\Big)\end{aligned}$$

Eq. 15 is a quadratic objective function with respect to I_B, I_O, and A, and its minimum is found by solving the linear system which we get by setting the derivative of Eq. 15 to zero. We use Successive over-relaxation (SOR) to solve it [3].

2.2 Motion estimation step: fix I_O, I_B, A, and solve for the motion fields V_O^t and V_B^t

When fixing I_B and I_R, solving $\mathbf{v}_B$ and $\mathbf{v}_R$ become a joint optical flow problem. Ignoring all the terms that only contain I_B and I_R in Eq. 1, we get:

$$\min_{\mathbf{v}_B^t,\mathbf{v}_O^t}\|I^t-\mathbf{W}(V_O^t)I_O-\mathbf{W}(V_O^t)A\circ\mathbf{W}(V_B^t)I_B\|_1+\|\nabla V_O^t\|_1+\|\nabla V_B^t\|_1 \quad (16)$$

Note that when we solve the motion field, there is no dependency between frames, so that we can solve each frame independently.

First, we explicitly represent the warping matrices $\mathbf{W}(V_B^t)$ and $\mathbf{W}(V_O^t)$.

$$\min_{\mathbf{v}_B^t,\mathbf{v}_O^t}\sum_{\mathbf{x}}\|I^t(\mathbf{x})-I_O(\mathbf{x}-V_O^t(\mathbf{x}))-A(\mathbf{x}-V_O^t(\mathbf{x}))I_B(\mathbf{x}-V_B^t(\mathbf{x}))\|_1+\lambda_4\|\nabla V_O^t\|_1+\lambda_4\|\nabla V_B^t\|_1 \quad (17)$$

Similar to Eq. 2 and Eq. 4, we approximate the non-smooth $l1$-norm by the smooth function $\phi(x)=\sqrt{x+\epsilon^2}$:

$$\min_{\mathbf{v}_B^t,\mathbf{v}_O^t}\sum_{\mathbf{x}}w_1(\mathbf{x})\|I^t(\mathbf{x})-I_O(\mathbf{x}-V_O^t(\mathbf{x}))-A(\mathbf{x}-V_O^t(\mathbf{x}))I_B(\mathbf{x}-V_B^t(\mathbf{x}))\|^2 \quad (18)$$

$$+\lambda_4 w_2(\mathbf{x})\|\nabla V_O^t\|^2+\lambda_4 w_3(\mathbf{x})\|\nabla V_B^t\|^2 \quad (19)$$

where

$$w_1(\mathbf{x})=\frac{1}{\|I^t(\mathbf{x})-I_O(\mathbf{x}-\hat{V}_O^t(\mathbf{x}))-A(\mathbf{x}-\hat{V}_O^t(\mathbf{x}))I_B(\mathbf{x}-\hat{V}_B^t(\mathbf{x}))\|^2}, \quad (20)$$

$$w_2(\mathbf{x})=\frac{1}{\|\nabla\hat{V}_O^t\|^2}, \quad (21)$$

$$w_3(\mathbf{x})=\frac{1}{\|\nabla\hat{V}_B^t\|^2}. \quad (22)$$

where $\hat{V}_O^t$ and $\hat{V}_B^t$ be the motion fields of obstruction and background layers of the last iteration, respectively.

Second, we linearize all the higher-order terms using first-order Taylor expansion. The data term is linearized as:

$$\begin{aligned}&I^t(\mathbf{x})-I_O(\mathbf{x}-V_O^t(\mathbf{x}))-A(\mathbf{x}-V_O^t(\mathbf{x}))I_B(\mathbf{x}-V_B^t(\mathbf{x}))\\ =&I^t(\mathbf{x})-I_O(\mathbf{x}-V_O^t(\mathbf{x}))-A(\mathbf{x}-V_O^t(\mathbf{x}))I_B(\mathbf{x}-V_B^t(\mathbf{x}))\\ =&I^t(\mathbf{x})-I_O(\mathbf{x}-V_O^t(\mathbf{x}))-A(\mathbf{x}-\hat{V}_O^t(\mathbf{x}))I_B(\mathbf{x}-V_B^t(\mathbf{x}))-A(\mathbf{x}-V_O^t(\mathbf{x}))I_B(\mathbf{x}-\hat{V}_B^t(\mathbf{x}))+A(\mathbf{x}-\hat{V}_O^t(\mathbf{x}))I_B(\mathbf{x}-\hat{V}_B^t(\mathbf{x}))\\ \stackrel{def}{=}&I^t(\mathbf{x})-I_O(\mathbf{x}-V_O^t(\mathbf{x}))-\overline{A}(\mathbf{x})I_B(\mathbf{x}-V_B^t(\mathbf{x}))-A(\mathbf{x}-V_O^t(\mathbf{x}))\overline{I}_B(\mathbf{x})+\overline{A}(\mathbf{x})\overline{I}_B(\mathbf{x})\\ =&I^t(\mathbf{x})-I_O(\mathbf{x}-V_O^t(\mathbf{x}))-\nabla I_O(\mathbf{x}-\hat{V}_O^t(\mathbf{x}))^\intercal(V_O^t(\mathbf{x})-\hat{V}_O^t(\mathbf{x}))\\ &-\overline{A}(\mathbf{x})I_B(\mathbf{x}-\hat{V}_B^t(\mathbf{x}))-\overline{A}(\mathbf{x})\nabla I_B(\mathbf{x}-\hat{V}_B^t(\mathbf{x}))^\intercal(V_B^t(\mathbf{x})-\hat{V}_B^t(\mathbf{x}))\\ &-A(\mathbf{x}-\hat{V}_O^t(\mathbf{x}))\overline{I}_B(\mathbf{x})-\nabla A(\mathbf{x}-\hat{V}_O^t(\mathbf{x}))^\intercal(V_O^t(\mathbf{x})-\hat{V}_O^t(\mathbf{x}))\overline{I}_B(\mathbf{x})+\overline{A}(\mathbf{x})\overline{I}_B(\mathbf{x}) \qquad (23)\\ \stackrel{def}{=}&I^t(\mathbf{x})-\overline{I}_O(\mathbf{x})-\nabla\overline{I}_O(\mathbf{x})^\intercal\Delta V_O^t(\mathbf{x})-\overline{A}(\mathbf{x})\overline{I}_B(\mathbf{x})-\overline{A}(\mathbf{x})\nabla\overline{I}_B(\mathbf{x})^\intercal\Delta V_B^t(\mathbf{x})-\overline{I}_B(\mathbf{x})\nabla\overline{A}(\mathbf{x})^\intercal\Delta V_O^t(\mathbf{x}),\end{aligned}$$

where

$$
\begin{aligned}
&\overline{A}(\mathbf{x}) \stackrel{def}{=} A(\mathbf{x} - V_O^t(\mathbf{x})),\ \ \overline{I}_O(\mathbf{x}) \stackrel{def}{=} I_O(\mathbf{x} - \hat{V}_O^t(\mathbf{x})),\ \ \overline{A}_B(\mathbf{x}) \stackrel{def}{=} I_B(\mathbf{x} - \hat{V}_B^t(\mathbf{x})),\\
&\nabla\overline{I}_O^t(\mathbf{x}) \stackrel{def}{=} \nabla I_O(\mathbf{x} - \hat{V}_O^t(\mathbf{x})),\ \ \nabla\overline{I}_B^t(\mathbf{x}) \stackrel{def}{=} \nabla I_B(\mathbf{x} - \hat{V}_B^t(\mathbf{x})),\\
&\Delta V_O^t = V_O^t - \hat{V}_O^t,\ \ \Delta V_B^t = V_B^t - \hat{V}_B^t
\end{aligned} \tag{24}
$$

Plugging Eq. 23 into the Eq. 19, we a quadratic objective function with respect to V_O^t and V_B^t :

$$
\begin{aligned}
\min_{\mathbf{v}_B^t,\mathbf{v}_O^t} \sum_{\mathbf{x}} & w_1(\mathbf{x}) \| I^t(\mathbf{x}) - \overline{I}_O(\mathbf{x}) - \nabla\overline{I}_O(\mathbf{x})^\intercal \Delta V_O^t(\mathbf{x}) - \overline{A}(\mathbf{x})\overline{I}_B(\mathbf{x}) - \overline{A}(\mathbf{x})\nabla\overline{I}_B(\mathbf{x})^\intercal \Delta V_B^t(\mathbf{x}) - \overline{I}_B(\mathbf{x})\nabla\overline{A}(\mathbf{x})^\intercal \Delta V_O^t(\mathbf{x})\|^2 \\
&+ \lambda_4 w_2(\mathbf{x}) \|\nabla V_O^t\|^2 + \lambda_4 w_3(\mathbf{x}) \|\nabla V_B^t\|^2.
\end{aligned} \tag{25}
$$

Similarly, we find the minimum of Eq. 25 using Successive over-relaxation.

In the case of a reflective pane, the alpha map A is essentially absorbed into the background and reflection images and there is no need to solve for it separately. We thus remove the prior term $\|\nabla A\|^2$ (Eq. 1) from the objective function, and only solve for I_O, I_B, $\{V_O^t\}$, and $\{V_B^t\}$.

References

[1] Liu, C., Sun, D.: On bayesian adaptive video super resolution. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 36(2), 346–360 (2014)

[2] Xue, T., Rubinstein, M., Liu, C., Freeman, W.T.: A computational approach for obstruction-free photography. ACM Transactions on Graphics (Proc. SIGGRAPH) 34(4) (2015)

[3] Young, D.: Iterative methods for solving partial difference equations of elliptic type. Transactions of the American Mathematical Society pp. 92–111 (1954)