

Supplementary Material

Single Image Optical Flow Estimation with an Event Camera

Liyuan Pan^{1,2}, MiaoMiao Liu^{1,2}, and Richard Hartley^{1,2}

¹ Australian National University, Canberra, Australia, ² Australian Centre for Robotic Vision

liyuan.pan@anu.edu.au

Abstract

In this supplementary material, we first provide our new ‘look’ at the Primal-Dual optimisation and a new perspective of the primal dual method regarding solving an objective function involving a norm term defined on Hilbert space in Section 1 and Section 2. Then, we provide more details about the derivation of the primal-dual optimisation to our model in Section 3. Finally, we provide additional experimental evaluations of our method in Section 4.

1. A new look at primal dual optimization

We consider the problem

$$\min_{\mathbf{u} \in X} G(\mathbf{u}) + F(K\mathbf{u}), \quad (1)$$

where by $K\mathbf{u}$ is meant a linear mapping K acting on \mathbf{u} in some Hilbert space X , with inner product $\langle \cdot, \cdot \rangle_X$, mapping into a space Y with inner product $\langle \cdot, \cdot \rangle_Y$. More simply, K is a matrix and \mathbf{u} is a vector in $X = \mathbb{R}^N$ for some N . The mapping $F : Y \rightarrow \mathbb{R}$ is defined.

The basis of the primal-dual formulation is to replace F in (1) by its double Fenchel dual F^{**} , so it becomes $\min_{\mathbf{u} \in X} (G(\mathbf{u}) + F^{**}(K\mathbf{u}))$, which is

$$\min_{\mathbf{u} \in X} \left(G(\mathbf{u}) + \max_{\mathbf{p} \in Y} (\langle K\mathbf{u}, \mathbf{p} \rangle_Y - F^*(\mathbf{p})) \right). \quad (2)$$

Recall that the Fenchel dual (convex conjugate) F^* of function F is defined as

$$F^*(\mathbf{q}) = \sup_{\mathbf{p} \in Y} (\langle \mathbf{p}, \mathbf{q} \rangle - F(\mathbf{p})), \quad (3)$$

and that $F = F^{**}$ if F is a convex function (a norm is convex).

Now, consider the case where F is a norm $\|\cdot\|_F$, not necessarily the standard norm on X (induced by the inner product). In this case, it is well known that the Fenchel dual of a norm is the indicator function of the unit ball in the dual

norm (zero inside the ball, ∞ outside). We denote this as B_{F^*} . In this case,

$$\max_{\mathbf{p} \in Y} (\langle K\mathbf{u}, \mathbf{p} \rangle_Y - F^*(\mathbf{p}))$$

cannot be achieved outside of B_{F^*} , so the min-max problem becomes

$$\min_{\mathbf{u} \in X} \left(G(\mathbf{u}) + \max_{\mathbf{p} \in B_{F^*}} \langle K\mathbf{u}, \mathbf{p} \rangle_Y \right) \quad (4)$$

This, then is the min-max problem to be solved. Observe that $\max_{\mathbf{p} \in B_{F^*}} \langle K\mathbf{u}, \mathbf{p} \rangle_Y$ must be achieved on the boundary \bar{B}_{F^*} of the unit ball. In other words, this minimization can be written as

$$\min_{\mathbf{u} \in X} \max_{\mathbf{p} \in \bar{B}_{F^*}} (G(\mathbf{u}) + \langle K\mathbf{u}, \mathbf{p} \rangle_Y) \quad (5)$$

Let

$$\begin{aligned} C(\mathbf{u}, \mathbf{p}) &= G(\mathbf{u}) + \langle K\mathbf{u}, \mathbf{p} \rangle_Y \\ &= G(\mathbf{u}) + \langle \mathbf{u}, K^*\mathbf{p} \rangle_X \end{aligned} \quad (6)$$

The primal-dual algorithm of [1] effectively consists of a series of gradient ascents in \mathbf{p} and gradient descents in \mathbf{u} starting from initial estimates \mathbf{u}^0 and \mathbf{p}^0 :

$$\begin{aligned} \mathbf{p}^{n+1} &= \mathbf{p}^n + \sigma C^{(\mathbf{p})}(\mathbf{u}^n, \mathbf{p}^n) \\ &= \mathbf{p}^n + \sigma K\mathbf{u}^n \\ \mathbf{u}^{n+1} &= \mathbf{u}^n - \tau C^{(\mathbf{u})}(\mathbf{u}^n, \mathbf{p}^{n+1}) \\ &= \mathbf{u}^n - \tau (\nabla G(\mathbf{u}^n) + K^*\mathbf{p}^{n+1}). \end{aligned} \quad (7)$$

Here, the bracketed superscripts (\mathbf{u}) and (\mathbf{p}) represent partial derivatives of C . This method is simply gradient ascent in \mathbf{p} and gradient descent in \mathbf{u} , with step size controlled by σ and τ .

This general scheme is modified in two ways. First, because the maximization should take place over B_{F^*} , or \bar{B}_{F^*} , we add a modification after \mathbf{p} is updated, so as to project back into the unit ball (or onto the boundary – it

probably does not matter which). Thus, we define $\pi(\mathbf{p}) = \mathbf{p}/\|\mathbf{p}\|$ (or $\pi(\mathbf{p}) = \mathbf{p}/\max(1, \|\mathbf{p}\|)$). Second, we add extrapolation. The updates then are

$$\begin{aligned}\mathbf{p}^{n+1} &= \pi(\mathbf{p}^n + \sigma K \bar{\mathbf{u}}^n) \\ \mathbf{u}^{n+1} &= \mathbf{u}^n - \tau (\nabla G(\mathbf{u}^n) + K^* \mathbf{p}^{n+1}) \\ \bar{\mathbf{u}}^{n+1} &= \mathbf{u}^{n+1} + \theta(\mathbf{u}^{n+1} - \mathbf{u}^n)\end{aligned}\quad (8)$$

This gives an algorithm as in Algorithm 1.

Algorithm 1: Primal-Dual Minimization - Flow

Initialization: Choose $\tau, \sigma > 0$, $n = 0$, and set $\bar{\mathbf{u}}^0 = \mathbf{u}^0$.

Iterations : Update $\mathbf{u}^n, \mathbf{p}^n, \bar{\mathbf{u}}^n$ as follows

```

1 while n < 20 do
2   Dual ascent in p
3   p^{n+1} = pi(p^n + sigma K bar{u}^n)
4   Primal descent in u
5   u^{n+1} = u^n - tau(nabla G(u^n) + K^* p^{n+1})
6   Extrapolation step
7   bar{u}^{n+1} = u^{n+1} + (u^{n+1} - u^n)
8   n = n + 1
9 end

```

2. More details.

We fill in here some of the mathematical gaps. The main points will be stated as lemmas.

2.1. Fenchel double dual

We show that if a function F is convex, then its double dual is the function itself. Let $F(x)$ be a function, $F : X \rightarrow \mathbb{R}$, where X is a Hilbert space with inner product $\langle \cdot, \cdot \rangle$. We will not need to write $\langle \cdot, \cdot \rangle_X$. The Fenchel dual is

$$F^*(y) = \max_{z \in X} \langle z, y \rangle - F(z).$$

The double-dual is therefore

$$\begin{aligned}F^{**}(x) &= \max_{y \in X} (\langle y, x \rangle - F^*(y)) \\ &= \max_{y \in X} \left(\langle y, x \rangle - \max_{z \in X} (\langle z, y \rangle - F(z)) \right) \\ &= \max_{y \in X} \langle y, x \rangle + \min_{z \in X} F(z) - \langle z, y \rangle \\ &= \max_{y \in X} \min_{z \in X} \langle y, x \rangle + F(z) - \langle z, y \rangle \\ &= \max_{y \in X} \min_{z \in X} F(z) - \langle z - x, y \rangle\end{aligned}\quad (9)$$

Now, let x be fixed and consider $F(z) - \langle z - x, y \rangle$. When $z = x$, this is equal to $F(x)$ for all y . Therefore

$$\min_{z \in X} F(z) - \langle z - x, y \rangle \leq F(x)$$

for all y . Hence $F^{**}(x) \leq F(x)$.

To show the opposite inequality, note the definition of subgradient. The subgradient of a function $F(x)$ at a point x is a vector y such that

$$F(z) \geq \langle z - x, y \rangle + F(x)$$

for all z . Essentially, this says that $\langle z - x, y \rangle + F(x)$, as a function of z , defines a hyperplane, with gradient y sitting underneath the graph of $F(z)$, and touching it at point x . A convex function has a subgradient at any point x .

Now, fix x and choose $y'(x)$ to be a subgradient of F at point x . Then, for all z ,

$$F(z) - \langle z - x, y'(x) \rangle \geq F(x).$$

and so

$$\min_{z \in X} F(z) - \langle z - x, y'(x) \rangle \geq F(x).$$

This inequality holds for the particular $y'(x)$ chosen, namely the subgradient. It follows, therefore, that

$$\begin{aligned}F^{**}(x) &= \max_{y \in X} \min_{z \in X} F(z) - \langle z - x, y \rangle \\ &\geq \min_{z \in X} F(z) - \langle z - x, y'(x) \rangle \\ &\geq F(x).\end{aligned}\quad (10)$$

This shows the following.

Lemma 1. If F is a convex function, then $F(x) = F^{**}(x)$ for all x .

2.2. Norms and dual norms

Suppose that $\|\cdot\|$ is a norm defined on a Hilbert space X with inner product $\langle \cdot, \cdot \rangle$. The idea is that $\|\cdot\|$ is not the standard norm induced by the inner product. The dual norm $\|\cdot\|_*$ is defined by

$$\begin{aligned}\|y\|_* &= \max_{x \in X} \frac{\langle y, x \rangle}{\|x\|} \\ &= \max_{\|x\|=1} \langle y, x \rangle.\end{aligned}\quad (11)$$

This is well known to define a norm.

Now, suppose a function $F(x)$ is defined by $F(x) = \|x\|$. We look at the Fenchel dual of F , namely

$$F^*(y) = \max_{x \in X} (\langle y, x \rangle - \|x\|). \quad (12)$$

Now, suppose that $\|y\|_* \leq 1$. Then from (11),

$$\langle y, x \rangle - \|x\| \leq 0 \text{ for all } x \neq 0.$$

On the other hand, $\langle y, x \rangle - \|x\| = 0$, when $x = 0$. Therefore

$$F^*(y) = \max_{x \in X} (\langle y, x \rangle - \|x\|) = 0.$$

On the other hand, if $\|y\|_* > 1$, then there exists x such that $\langle y, x \rangle - \|x\| > 1$. Then $\langle y, \lambda x \rangle - \|\lambda x\| > \lambda$ for all $\lambda > 0$. So

$$F^*(y) = \max_{x \in X} (\langle y, x \rangle - \|x\|) = \infty.$$

This shows the following fact.

Lemma 2. *If $F(x) = \|x\|$ for some norm $\|\cdot\|$, then the Fenchel dual $F^*(y)$ is the indicator function for the unit ball under the dual norm $\|\cdot\|_*$.*

2.3. Fenchel duals and dual norms

Suppose that $F(x) = \|x\|$. It is not true that $F^*(y) = \|y\|_*$. They are quite different objects. One, $F^*(y)$, is an indicator function for the dual ball, whereas the other, $\|y\|_*$ is a norm. Nevertheless, the double duals are the same.

The double dual norm is given (by repeating the use of (11)) by

$$\|x\|_{**} = \max_{y \in X} \frac{\langle y, x \rangle}{\|y\|_*} \quad (13)$$

Then

$$\begin{aligned} F^{**}(x) &= \max_y \langle x, y \rangle - F^*(y) \\ &= \max_{\|y\|_*=1} \langle x, y \rangle \\ &= \|x\|_{**}. \end{aligned}$$

The second line comes from the form of $F^*(y)$ and the last equality comes from the definition (13). However, since $F(x) = \|x\|$ is convex, $F(x) = F^{**}(x)$. This shows the following.

Lemma 3. *If $\|\cdot\|$ is a norm on a Hilbert space X and $F(x) = \|x\|$, then*

$$\|x\| = F(x) = F^{**}(x) = \|x\|_{**}. \quad (14)$$

This allows us to write

$$\|x\| = \max_{y \in X} \frac{\langle y, x \rangle}{\|y\|_*} = \max_{\|y\|_*=1} \langle y, x \rangle. \quad (15)$$

This shows the following results.

Lemma 4. *If $\|\cdot\|$ is a norm defined on a Hilbert space X , then*

$$\|x\| = \max_{\|y\|_*=1} \langle y, x \rangle.$$

2.4. Simple interpretation of primal dual method

If we accept as a standard result that

$$\|x\| = \|x\|_{**} = \max_{\|y\|_*=1} \langle y, x \rangle \quad (16)$$

then we can explain the primal-dual method of the previous section without any reference to Fenchel duals. Starting with the problem (1) of finding $\min_{\mathbf{u} \in X} G(\mathbf{u}) + F(K\mathbf{u})$, where F is a norm, $F(K\mathbf{u}) = \|K\mathbf{u}\|$, we use (16) to write it as

$$\min_{\mathbf{u} \in X} \max_{\|\mathbf{p}\|=1} G(\mathbf{u}) + \langle K\mathbf{u}, \mathbf{p} \rangle_Y.$$

which is the same as the starting point (2) for the method of alternating gradient ascent and descent to find the saddle. By doing this, we avoid all the machinery of Fenchel duals, and interpret it as a simple application of the standard result that $\|x\| = \|x\|_{**}$.

3. Derivation details in our case

We will derive the convex conjugate function for the regularisation term used in our paper. First, we give a brief summary of our model. In this paper, our energy minimization model is formulated as:

$$\min_{\mathbf{L}, \mathbf{u}} \mu_1 \phi_{\text{eve}}(\mathbf{L}, \mathbf{u}) + \mu_2 \phi_{\text{colour}}(\mathbf{L}, \mathbf{u}) + \phi_{\text{flow}}(\nabla \mathbf{u}) + \phi_{\text{im}}(\nabla \mathbf{L}). \quad (17)$$

Here, \mathbf{u} denotes optical flow, \mathbf{L} denotes the latent image, and μ_1, μ_2 are weight parameters. Each term is denote as

$$\begin{aligned} \phi_{\text{eve}}(\mathbf{L}, \mathbf{u}) &= \sum_{\mathbf{x} \in \mathbf{x}} \|\mathbf{L}(\mathbf{x}, f)(\exp(c \mathbf{E}(\mathbf{x}, t)) - 1) \\ &\quad + [\mathbf{u}_{\mathbf{x}}, v_{\mathbf{x}}]^T \nabla \mathbf{L}(\mathbf{x}, f) \|, \\ \phi_{\text{colour}}(\mathbf{L}, \mathbf{u}) &= \sum_{\mathbf{x}, \mathbf{y} \in \mathbf{x}} \|\mathbf{k}_{\mathbf{u}'(\mathbf{x})}(\mathbf{y}) \mathbf{L}(\mathbf{x} - \mathbf{y}) - \mathbf{B}(\mathbf{x})\|^2, \\ \phi_{\text{flow}}(\nabla \mathbf{u}) &= \|w \nabla \mathbf{u}\|_{1,2} = \sum_{\mathbf{x} \in \mathbf{x}} \|w(\mathbf{x}) \nabla \mathbf{u}(\mathbf{x})\|, \\ \phi_{\text{im}}(\nabla \mathbf{L}) &= \sum_{\mathbf{x} \in \mathbf{x}} \|\nabla \mathbf{L}(\mathbf{x})\|_1, \end{aligned} \quad (18)$$

where ϕ_{eve} enforces the brightness constancy constraint by event data, ϕ_{colour} enforces the blurred image formation process, ϕ_{flow} and ϕ_{im} enforces the smoothness. Our model involves two sets of variables, namely optical flow \mathbf{u} , and latent images \mathbf{L} . We perform the optimization iteratively.

We fix the latent image, namely $\mathbf{L} = \hat{\mathbf{L}}$, and (17) reduces to

$$\min_{\mathbf{u}} \mu_1 \phi_{\text{colour}}(\mathbf{u}) + \mu_2 \phi_{\text{eve}}(\mathbf{u}) + \phi_{\text{flow}}(\nabla \mathbf{u}). \quad (19)$$

The primal-dual formulation of [1] in our case is

$$\min_{\mathbf{u}} G(\mathbf{u}) + F(K\mathbf{u}). \quad (20)$$

Here, we denote $\nabla \mathbf{u}(\mathbf{x})$ as

$$\nabla \mathbf{u}(\mathbf{x}) = \left(\frac{\partial u(\mathbf{x})}{\partial x}, \frac{\partial u(\mathbf{x})}{\partial y}, \frac{\partial v(\mathbf{x})}{\partial x}, \frac{\partial v(\mathbf{x})}{\partial y} \right)^T, \quad (21)$$

where we denote $\nabla \mathbf{u}(\mathbf{x}) = (u_{\mathbf{x}}^{(x)}, u_{\mathbf{x}}^{(y)}, v_{\mathbf{x}}^{(x)}, v_{\mathbf{x}}^{(y)})^T$ for short. Note that (here and elsewhere) superscripts in brackets represent differentiation with respect to x or y .

Casting (19) in the form of (20), we see that

$$\begin{aligned} G(\mathbf{u}) &= \mu_1 \phi_{\text{color}}(\mathbf{u}) + \mu_2 \phi_{\text{eve}}(\mathbf{u}) \\ F(K\mathbf{u}) &= F(K\mathbf{u}) = \|K\mathbf{u}\|_{1,2} = \phi_{\text{flow}}(\nabla \mathbf{u}). \end{aligned} \quad (22)$$

where $K\mathbf{u} = w\nabla \mathbf{u}$ is a linear function. Let $\mathbf{u} \in X = \mathbb{R}^{2N}$, and $\nabla \mathbf{u} \in Y = \mathbb{R}^{4N}$, so $G : X \rightarrow \mathbb{R}$, and $F : Y \rightarrow \mathbb{R}$, where $N = HW$ is the number of pixels.

The basis of the primal-dual formulation is to replace F in (20) by its double Fenchel dual F^{**} , so it becomes

$$\min_{\mathbf{u} \in X} (G(\mathbf{u}) + F^{**}(K\mathbf{u})),$$

which is

$$\min_{\mathbf{u} \in X} \left(G(\mathbf{u}) + \max_{\mathbf{p} \in Y} \langle K\mathbf{u}, \mathbf{p} \rangle_X - F^*(\mathbf{p}) \right). \quad (23)$$

We focus on the regularization term $F(K\mathbf{u})$ first. For any pixel \mathbf{x} , vector $w(\mathbf{x}) = (w_{\mathbf{x}}^x, w_{\mathbf{x}}^y) \in \mathbb{R}^2$ in (18), and $\nabla \mathbf{u}(\mathbf{x}) \in \mathbb{R}^4$, define

$$w(\mathbf{x})\nabla \mathbf{u}(\mathbf{x}) = \left(w_{\mathbf{x}}^x u_{\mathbf{x}}^{(x)}, w_{\mathbf{x}}^x u_{\mathbf{x}}^{(y)}, w_{\mathbf{x}}^y v_{\mathbf{x}}^{(x)}, w_{\mathbf{x}}^y v_{\mathbf{x}}^{(y)} \right)^T. \quad (24)$$

Putting all the pixels together, we define $w\nabla \mathbf{u}$, where $w \in \mathbb{R}^{2N}$.

$$\begin{aligned} F(K\mathbf{u}) &= \|w\nabla \mathbf{u}\|_{1,2} \\ &= \sum_{\mathbf{x} \in \Omega} \sqrt{(\mathbf{p}_{\mathbf{x}}^1)^2 + (\mathbf{p}_{\mathbf{x}}^2)^2 + (\mathbf{p}_{\mathbf{x}}^3)^2 + (\mathbf{p}_{\mathbf{x}}^4)^2} \\ &= \sum_{\mathbf{x} \in \Omega} \sqrt{\langle \mathbf{p}_{\mathbf{x}}, \mathbf{p}_{\mathbf{x}} \rangle}, \end{aligned} \quad (25)$$

where $\mathbf{p}_{\mathbf{x}} = w\nabla \mathbf{u}_{\mathbf{x}} = K\mathbf{u}$, and $\mathbf{p}_{\mathbf{x}} \in \mathbb{R}^4$. We define $\mathbf{p}_{\mathbf{x}}^i$ as the i th component of $\mathbf{p}_{\mathbf{x}}$.

Recall that the Fenchel dual (convex conjugate) F^* of function F is defined as

$$F^*(\mathbf{q}) = \sup_{\mathbf{p} \in Y} (\langle \mathbf{p}, \mathbf{q} \rangle - F(\mathbf{p})), \quad (26)$$

and that $F = F^{**}$ if F is a convex function (a norm is convex). It is well known that the Fenchel dual of a norm is the indicator function of the unit ball in the dual norm. In this case, $F^*(\cdot)$ is a mixed norm $\|\cdot\|_{1,2}$, and its dual is a norm $\|\cdot\|_{\infty,2}$. The indicator function is therefore a product \mathbb{B}^N of N Euclidean 2-balls (each in \mathbb{R}^4).

We compute the convex conjugate function F^* as follows

$$F^*(\mathbf{q}) = \sup_{\mathbf{p} \in Y} \sum_{\mathbf{x} \in \Omega} \langle \mathbf{p}_{\mathbf{x}}, \mathbf{q}_{\mathbf{x}} \rangle - \sqrt{\langle \mathbf{p}_{\mathbf{x}}, \mathbf{p}_{\mathbf{x}} \rangle} \quad (27)$$

$$= \sum_{\mathbf{x} \in \Omega} \sup_{\mathbf{p}_{\mathbf{x}} \in Y} \langle \mathbf{p}_{\mathbf{x}}, \mathbf{q}_{\mathbf{x}} \rangle - \sqrt{\langle \mathbf{p}_{\mathbf{x}}, \mathbf{p}_{\mathbf{x}} \rangle} \quad (28)$$

$$= \sum_{\mathbf{x} \in \Omega} f^*(\mathbf{p}_{\mathbf{x}}), \quad (29)$$

where $f^* : \mathbb{R}^4 \rightarrow \mathbb{R}$,

$$f^*(\mathbf{p}_{\mathbf{x}}) = \begin{cases} 0 & \text{if } \mathbf{p}_{\mathbf{x}} \in \mathbb{B} \\ \infty & \text{otherwise,} \end{cases} \quad (30)$$

and \mathbb{B} is the L^2 ball.

Proof. We distinguish $\mathbf{p}_{\mathbf{x}}$ into two cases,

- if $\mathbf{p}_{\mathbf{x}} \in \mathbb{B}$, then

$$\langle \mathbf{q}_{\mathbf{x}} \cdot \mathbf{p}_{\mathbf{x}} \rangle \leq |\mathbf{q}_{\mathbf{x}}|. \quad (31)$$

The equality holds if $\mathbf{q}_{\mathbf{x}} = 0$.

- if $\mathbf{p}_{\mathbf{x}} \notin \mathbb{B}$, there exists an \mathbf{q} with $|\mathbf{q}_{\mathbf{x}}| \leq 1$, $\mathbf{q}_{\mathbf{x}}^T \mathbf{p}_{\mathbf{x}} > 1$. Then

$$F^*(\mathbf{p}_{\mathbf{x}}) \geq \langle s\mathbf{q}_{\mathbf{x}}, \mathbf{p}_{\mathbf{x}} \rangle - \|s\mathbf{q}_{\mathbf{x}}\| = s(\mathbf{q}_{\mathbf{x}}^T \mathbf{p}_{\mathbf{x}} - |\mathbf{q}_{\mathbf{x}}|), \quad (32)$$

and the right hand side goes to infinity when $s \rightarrow +\infty$. \square

In this case, $F^*(\cdot)$ is a mixed norm $\|\cdot\|_{1,2}$, and its dual is a norm $\|\cdot\|_{\infty,2}$. The indicator function is therefore a product \mathbb{B}^N of N Euclidean 2-balls (each in \mathbb{R}^4). More precisely

$$F^*(\mathbf{p}) = \begin{cases} 0, & \text{if } \|\mathbf{p}_{\mathbf{x}}\| \leq 1 \text{ for all } \mathbf{x} \\ +\infty, & \text{otherwise.} \end{cases} \quad (33)$$

The proximal operator \mathcal{P}_{F^*} is therefore given by

$$\begin{aligned} F^*(\bar{\mathbf{p}}) &= \arg \min_{\mathbf{p} \in Y} (2F^*(\mathbf{p}) + \|\bar{\mathbf{p}} - \mathbf{p}\|^2) \\ &= \arg \min_{\mathbf{p} \in \mathbb{B}^N} \|\bar{\mathbf{p}} - \mathbf{p}\|^2. \end{aligned} \quad (34)$$

In other words, each $\bar{\mathbf{p}}_{\mathbf{x}}$ is projected to the nearest point in the unit ball, given by $\bar{\mathbf{p}}_{\mathbf{x}} / (\max(1, \|\bar{\mathbf{p}}_{\mathbf{x}}\|))$.

4. Experiments

4.1. Experimental setup details

Evaluations. Since our method jointly estimates optical flow and the deblurred images, we evaluate these two tasks separately. For optical flow estimation results, we evaluate optical flow by Mean Square Error (MSE), Average Endpoint Error (AEE) and Flow Error metric (FE), which is by counting the number of pixels having errors more than 3 pixels and 5% of its ground-truth. We adopt the PSNR to evaluate the deblurred image. Please watch the video included in the main file.

$$\text{AEE} = \frac{\sum_{\Omega} \|\mathbf{u}_{\text{est}} - \mathbf{u}_{\text{gt}}\|_1}{2HW}, \quad \text{MSE} = \frac{\sum_{\Omega} \|\mathbf{u}_{\text{est}} - \mathbf{u}_{\text{gt}}\|^2}{2HW}.$$

We adopt the PSNR to evaluate deblurred images. The error map shows the distribution of the endpoint error of measurements compared with the ground-truth flow and the success rate is defined as the percentage of results with errors

below a threshold. In other words, the curve shows the percentage of the measured flow for pixels that lower than a given threshold.

Implementation details. Regarding the model parameters, we perform grid search on 30 reserved images. In our experiments, we fix the model parameters as $\sigma = 5$, $\tau = 0.025$, $\gamma = 10$, $\eta = 6.25e - 3$, $\mu_1 = 2$, $\mu_2 = 5$, $\mu_3 = 0.08$, $\mu_4 = (30/255)$. In the coarse-to-fine framework, the pyramid level = 7. In our approach, we assume that the duty cycle $\theta = T/\Delta t$ is known for each sequence. The duty cycle is used to approximate the optical flow during the exposure time, where $\mathbf{u}' = \theta\mathbf{u}$ (seeing Section 4.2 in our submitted version). For the dataset that does not provide their exposure information, we can follow [2] to estimate the duty cycle.

4.2. More results

We compare our results with baselines on optical flow estimation and image deblurring on 5 (including real and synthetic) datasets. We report qualitative comparisons in the following figures. We also provide a video ‘demo.mp4’ to show an example result on BED dataset [8].

References

- [1] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40(1):120–145, 2011.
- [2] S. Cho, J. Wang, and S. Lee. Video deblurring for hand-held cameras using patch-based synthesis. *ACM Transactions on Graphics (TOG)*, 31(4):64, 2012.
- [3] D. Gong, J. Yang, L. Liu, Y. Zhang, I. Reid, C. Shen, A. van den Hengel, and Q. Shi. From motion blur to motion flow: A deep learning solution for removing heterogeneous motion blur. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2319–2328, 2017.
- [4] M. Jin, G. Meishvili, and P. Favaro. Learning to extract a video sequence from a single motion-blurred image. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, June 2018.
- [5] P. Liu, M. Lyu, I. King, and J. Xu. Selfflow: Self-supervised learning of optical flow. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, June 2019.
- [6] E. Mueggler, H. Rebecq, G. Gallego, T. Delbrück, and D. Scaramuzza. The event-camera dataset and simulator: Event-based data for pose estimation, visual odometry, and slam. *Int. J. Robot. Research*, 36(2):142–149, 2017.
- [7] S. Nah, T. H. Kim, and K. M. Lee. Deep multi-scale convolutional neural network for dynamic scene deblurring. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, July 2017.
- [8] L. Pan, C. Scheerlinck, X. Yu, R. Hartley, M. Liu, and Y. Dai. Bringing a blurry frame alive at high frame-rate with an event camera. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, June 2019.
- [9] D. Sun, X. Yang, M.-Y. Liu, and J. Kautz. Pwc-net: Cnns for optical flow using pyramid, warping, and cost volume. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, June 2018.
- [10] X. Tao, H. Gao, X. Shen, J. Wang, and J. Jia. Scale-recurrent network for deep image deblurring. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, June 2018.
- [11] H. Zhang, Y. Dai, H. Li, and P. Koniusz. Deep stacked hierarchical multi-patch network for image deblurring. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, June 2019.
- [12] A. Zhu, L. Yuan, K. Chaney, and K. Daniilidis. Ev-flownet: Self-supervised optical flow estimation for event-based cameras. In *Robotics: Science and Systems*, Pittsburgh, Pennsylvania, June 2018.

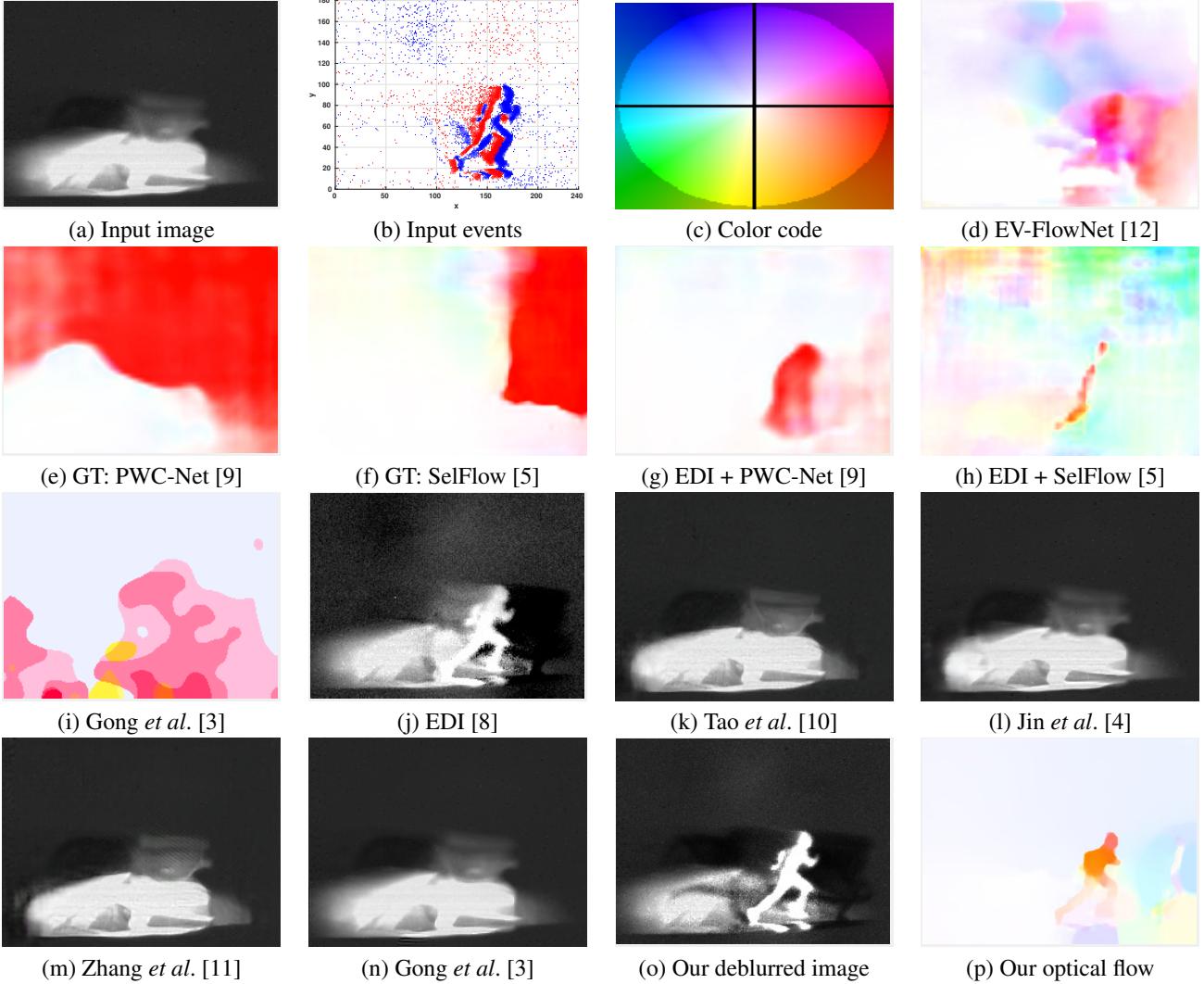


Figure 1. (a) and (b) are the input to our method, where (a) shows the intensity image from the event camera, and (b) visualizes the integrated events over a temporal window (blue: positive event; red: negative event). (c) Color-coded optical flow in all our experiments. (d) Flow result by [12], using the images and events as input. (e) Flow result by [9] based on two ground-truth images in the video, while the images are blurred. (f) Flow result by [5] based on two ground-truth images in the video. (g) Flow result by [9] based on images estimated by EDI model [8]. (h) Flow result by [5] based on images estimated by the EDI model. (i) Optical flow result of [3] by using a single blurred image. (j) Deblurred result by [8]. (k) Deblurred result by [10]. (l) Deblurred result by [4]. (m) Deblurred result by [11]. (n) Deblurred result by [3]. (o) Our deblurred image. (p) Our estimated optical flow. (Best viewed on screen).

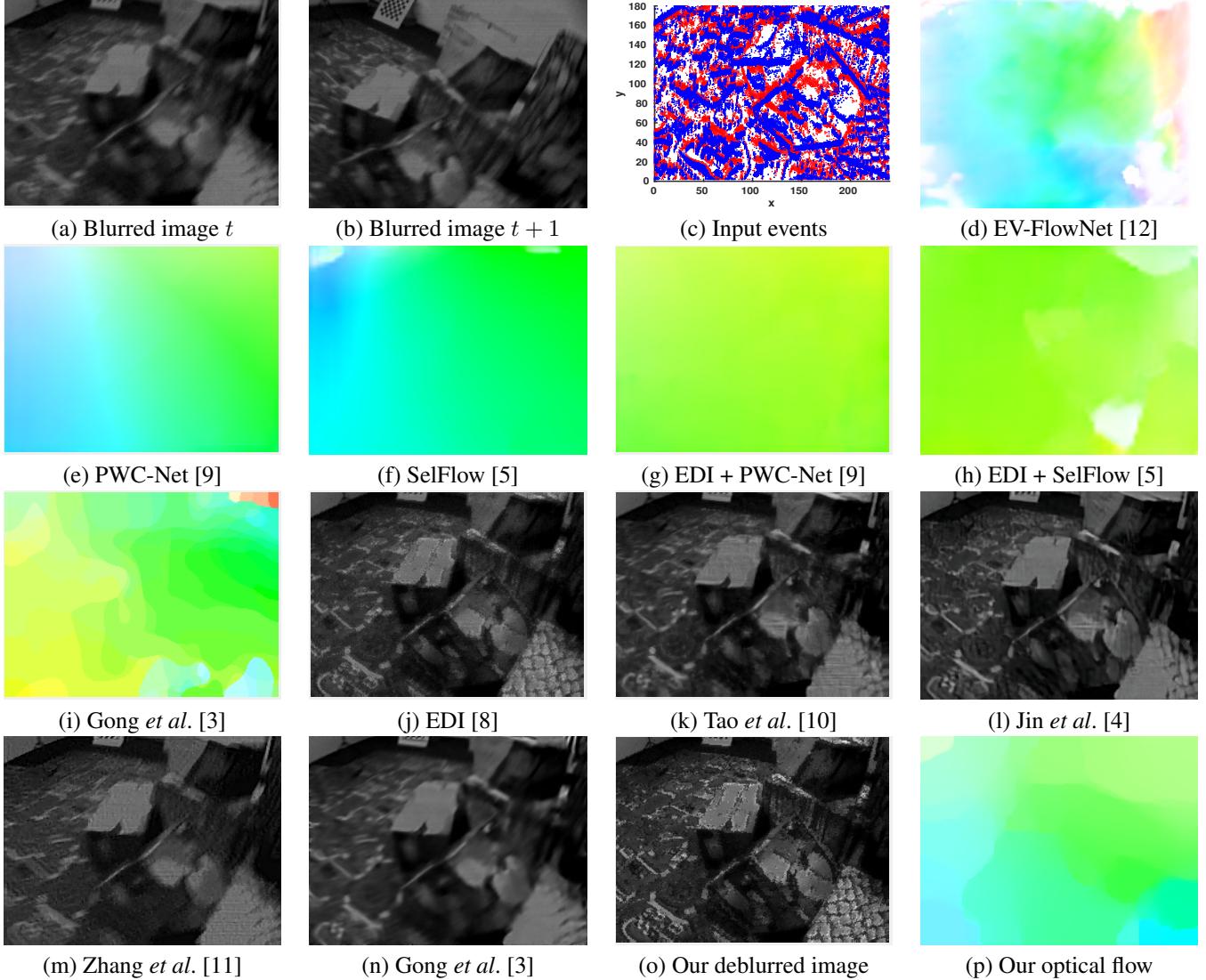


Figure 2. An example of our deblurring result on the real dataset EBC [6], and EBC is the first event camera datasets for high-speed robotics, the image size is 346×260 . This dataset contains a bunch of blurred images due to moving camera or moving objects. (a) and (b) are the input to our method, where (a) shows the intensity image from the event camera, and (b) visualises the integrated events over a temporal window (blue: positive event; red: negative event). (c) Color coded optical flow in all our experiments. (d) Flow result by [12], using the images and events as input. (e) Flow result by [9] based on (a) and (b), which are two blurred ground-truth images in the video. (f) Flow result by [5] based on two ground-truth images in the video. (g) Flow result by [9] based on images estimated by EDI model [8]. (h) Flow result by [5] based on images estimated by EDI model. (i) Optical flow result of [3] by using a single blurred image. (j) Deblurred result by [8]. (k) Deblurred result by [10]. (l) Deblurred result by [4]. (m) Deblurred result by [11]. (n) Deblurred result by [3]. (o) Our deblurred image. (p) Our estimated optical flow. (Best viewed on screen).

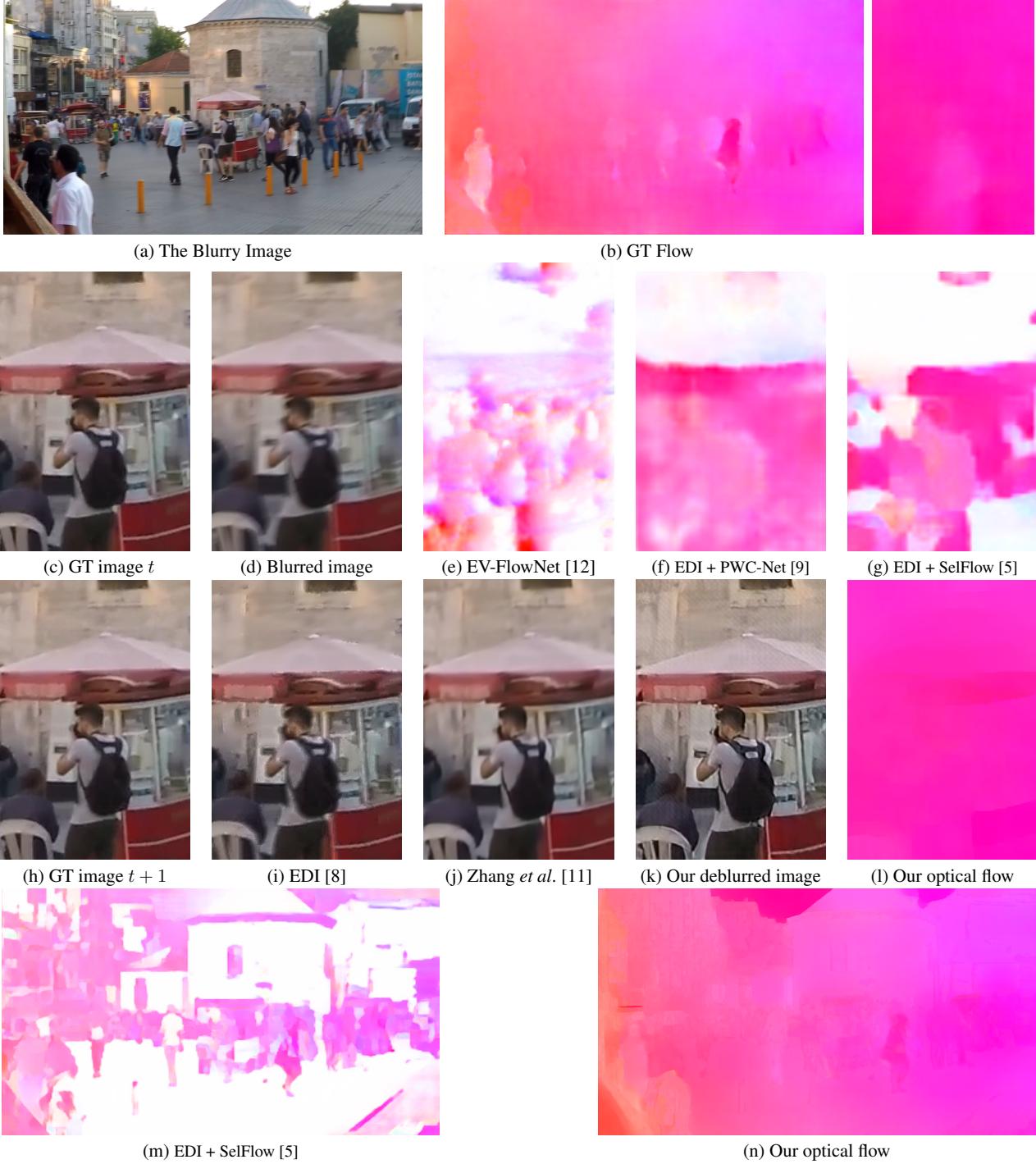


Figure 3. To provide a quantitative deblurring comparison, we generate another synthetic dataset with events and motion blur, based on the real GoPro video dataset [7], where the image size is 1280×720 . It has ground-truth latent images and associated motion blurred images. We additionally use PWC-Net to estimate flow from sharp images as the ground-truth for flow evaluation. (a) The blurred image. (b) The ground-truth flow and the cropped ground-truth flow. (c) and (f) are the cropped ground-truth latent images at time t and $t + 1$. (d) The cropped blurred image. (e) The cropped flow result by [12], using the images and events as input. (f) The cropped flow result by [9] based on images estimated by the EDI model [8]. (g) The cropped flow result by [5] based on images estimated by the EDI model. (i) The cropped deblurred result by [8]. (g) The cropped deblurred result by [11]. (k) Our deblurred image. (l) Our estimated optical flow. (m) and (n) are the non-cropped flow results. (Best viewed on screen).

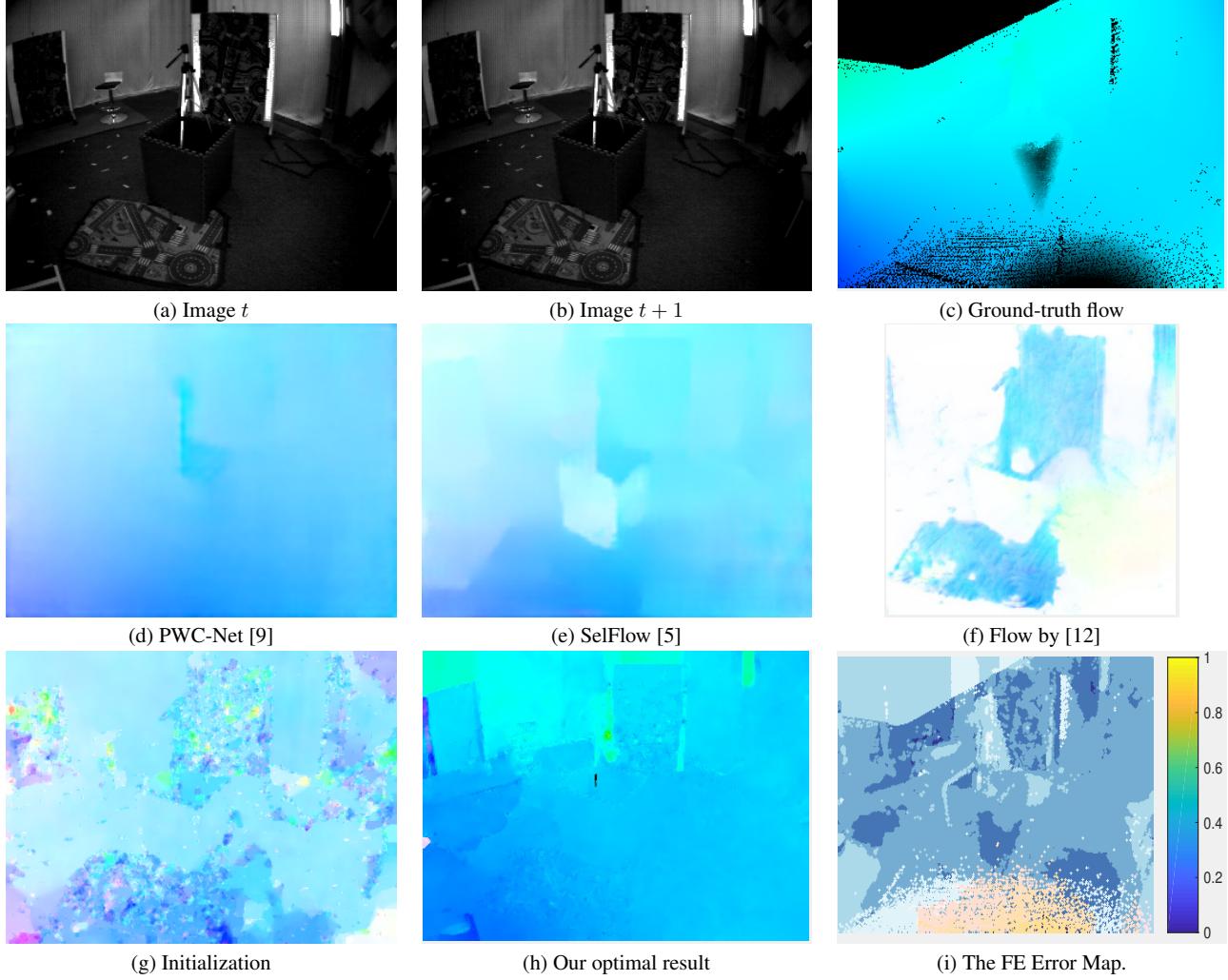


Figure 4. An example result of our method on a sharp image in the dataset [12]. Our method achieves competitive performance on both the blurred and the sharp image. (a) The input image to our model. (b) The reference image in time $t + 1$. (c) The ground-truth optical flow. (d) Flow result by [9] based on two images (a) and (b). (e) Flow result by [5] based on two images (a) and (b). (f) Flow result by [12], using events as input. Their provided results are cropped to size 256×256 . (g) Note, our initial flow is simply based on event frames by Eq (2) (in the submitted version). (h) Our optical flow result. (i) We use the end-point-error to evaluate the flow accuracy, which is by counting the number of pixels having errors more than 3 pixels and 5% of its ground-truth. (Best viewed on screen).