Homework 3: Optical Flow

Prof. Davi Geiger

Due April $13^{th}, 2017$

Introduction

We will develop an optical flow method. First, let us create an image of a square moving by 3 pixels in x and y, i.e., $\vec{v}_0 = (3, 3)$. Say the image is 100 x 100 pixels, black background I(x,y)=100, and the square I(x,y)=200, is in the center with size 31 x 31.

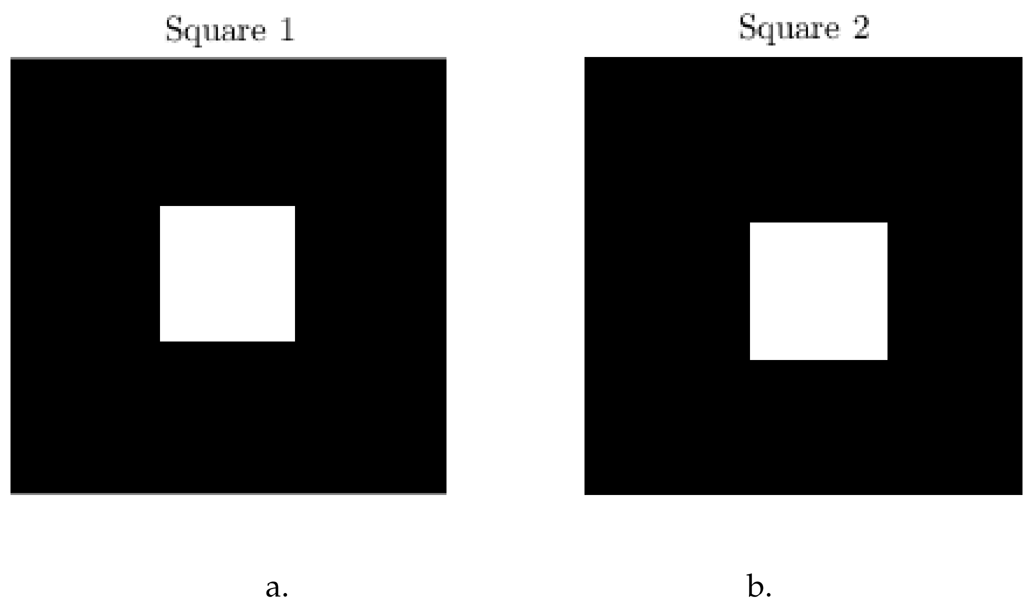


Figure 1: *A sequence of two Images of size 100 × 100 pixels. The square square of size 31 × 31 pixels in the middle moves by $\vec{v}_0 = (3, 3)$ pixels.*

After we develop the algorithm we will use the "mini cooper pictures" from the data set in http://vision.middlebury.edu/flow/data/

We want to explore the method to compute the optical flow through different scales. The main reason is that all possible motions is a very large space to exploit. So, the method starts at a large scale consideration, say $\sigma = 6$ pixels (or more), so that even

Figure 2: *A sequence of two Images of the "the mini cooper" series (frame07 and frame 08), see http://vision.middlebury.edu/flow/data/.*

large motions become small in the scale considered. Then it uses such coarse solutions to shift the search at the new solution at smaller scale. The use of the previous scale result allow to focus on the data at these shifted locations.

Of course, even at large scale, we need a method to compute optical flow.

Let us elaborate it further as an algorithm (by doing it).

Optical flow via $\frac{\partial I(\vec{u},t)}{\partial t} + \vec{\nabla} I(\vec{u},t) \cdot \vec{v}(\vec{u},t) = 0$

The equation of optical flow is $\frac{\partial I(\vec{u},t)}{\partial t} + \vec{\nabla} I(\vec{u},t) \cdot \vec{v}(\vec{u},t) = 0$. We assume that the best approximation to $\vec{\nabla} I(\vec{u},t)$, at scale σ, is the best angle response of the imaginary component of the wavelet

$$
\begin{aligned}
\vec{\nabla} I(\vec{u},t) \quad &\approx \quad \mathcal{W}^{im}_{\sigma,\theta^{max}}[I](\vec{u},t)\, \hat{e}_{\theta^{max}_{\vec{u},t}} \qquad (1)\\
&\text{where}\\
\theta^{max}(\vec{u},t) \quad &= \quad \arg\max_{\theta} \left| \mathcal{W}^{imaginary}_{\sigma,\theta}[I](\vec{u},t) \right|
\end{aligned}
$$

where $\hat{e}_{\theta^{max}_{\vec{u},t}} = (\cos\theta^{max}_{\vec{u},t}, \sin\theta^{max}_{\vec{u},t})$ and the optical flow equation, at scale σ, becomes

$$
\frac{\partial [G_\sigma * I(\vec{u},t)]}{\partial t} + \mathcal{W}^{im}_{\sigma,\theta^{max}}[I](\vec{u},t)\, \hat{e}_{\theta^{max}_{\vec{u},t}} \cdot \vec{v}(\vec{u},t) \quad = \quad 0 \qquad (2)
$$

Problem 1: Compute approximate Image Gradient and Image Time Derivative at scale $\sigma = 6$ pixels, and Downsample.

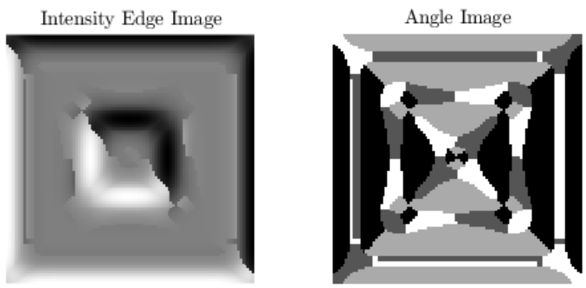
Consider $\theta \in \{0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}, \}$ and $\sigma = 6$ pixels.

Compute the x and y gradient components and the time derivative of the image, at the scale $\sigma = 6$ pixels, as follows (display of results in figure 3):

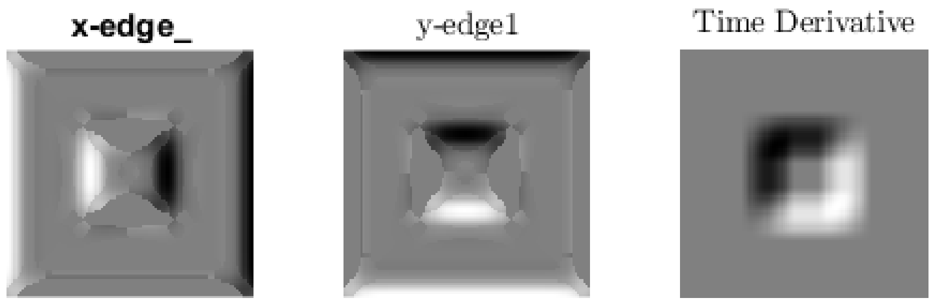
1. Compute $\mathcal{W}^{im}_{\sigma,\theta^{max}}[I](\vec{u}, t = 1)$ and $\theta^{max}(\vec{u}, t = 1)$ for the first image of the square (figure 1). Results shown in figure 3 a. and b.

2. Multiply point wise to get the x and y components, i.e., $\partial_x(G_\sigma * I) \approx \mathcal{W}^{im}_{\sigma,\theta^{max}}[I](\vec{u}, t)\ \cos(\theta^{max}(\vec{u}, t))$ and $\partial_y(G_\sigma * I) \approx \mathcal{W}^{im}_{\sigma,\theta^{max}}[I](\vec{u}, t)\ \sin(\theta^{max}(\vec{u}, t))$ (figure 3c. and d.).

3. Finally, compute $\frac{\partial[G_\sigma * I(\vec{u},t)]}{\partial t} \approx G_\sigma * I(\vec{u}, t = 2) - G_\sigma * I(\vec{u}, t = 1)$ (figure 3e.).



(a) $\mathcal{W}^{im}_{\sigma,\theta^{max}}[I](\vec{u},t)$ (b) $\theta^{max}_{\vec{u},t}$



(c) $\approx \partial_x(G_\sigma * I)$ (d) $\approx \partial_y(G_\sigma * I)$ (e) $\approx \partial_t(G_\sigma * I)$

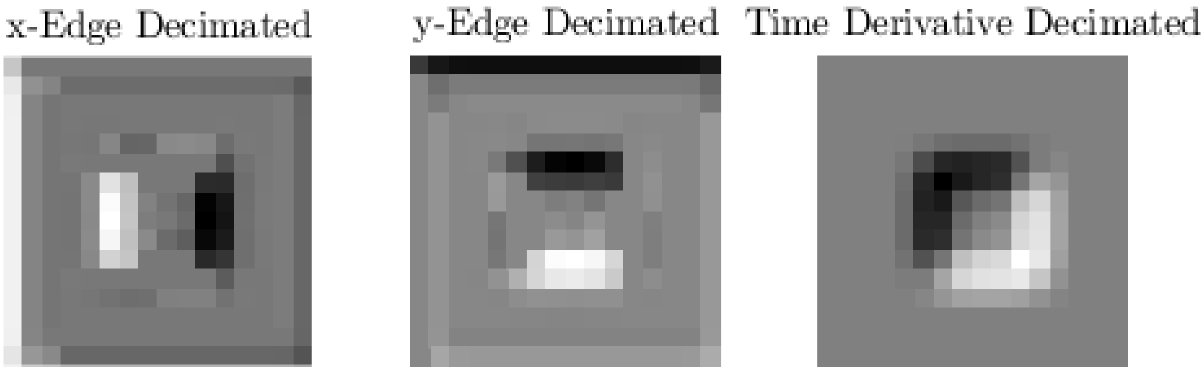
Figure 3: *All images computed at scale $\sigma = 6$ pixels. The image gradient, $\nabla I(\vec{u})$, is approximately computed at every pixel by the imaginary component of the wavelet response, selected per pixel as the maximum absolute value across all angles, i.e., $\mathcal{W}^{im}_{\sigma,\theta^{max}}[I](\vec{u},t)\ (\cos\theta^{max}_{\vec{u},t}, \sin\theta^{max}_{\vec{u},t})$. a. The best imaginary wavelet component per pixel, across the four angle responses, $\mathcal{W}^{im}_{\sigma,\theta^{max}}[I](\vec{u},t)$. b. $\theta^{max}(\vec{u})$, the best angle response of the imaginary Wavelet for each pixel. c. and d. are the x and y components of ∇I, i.e., c. $\mathcal{W}^{im}_{\sigma,\theta^{max}}[I](\vec{u},t)\ \cos\theta^{max}_{\vec{u},t}$ and d. $\mathcal{W}^{im}_{\sigma,\theta^{max}}[I](\vec{u},t)\ \sin\theta^{max}_{\vec{u},t}$. e. Temporal difference of the images, after Gaussian blur.*

Downsample (Decimate) it by a factor $\sigma = 6$ pixels, i.e., make new images smaller by a factor 6 by only keeping the image pixels for every other 6 pixels.

Display the quantities (which can be negative, so scale them appropriate by subtracting from the min values), you results should resemble figure 4.

Problem 2: Solve Optical Flow Equation

Say we focus on a pixel $\vec{u}_0 = (x_0, y_0)$ centered in a window 5×5, i.e., pixels $\{(x_0, y_0); (x_0 \pm 2, y_0 \pm 2)\}$ and assume all these pixels to be moving by the same opti-



(a) (b) (c)

Figure 4: *Images from c., d., e., of above figure, decimated for every* $\sigma = 6$ *pixels. c. and d. are the x and y components of* ∇I*, i.e., a.* $\mathcal{W}^{im}_{\sigma,\theta^{max}}[I](\vec{u},t)\,\cos\theta^{max}_{\vec{u},t}$ *and b.* $\mathcal{W}^{im}_{\sigma,\theta^{max}}[I](\vec{u},t)\,\sin\theta^{max}_{\vec{u},t}$*, decimated and c. Temporal difference of the images, after Gaussian blur, decimated.*

cal flow $\vec{v}(\vec{u}_0,t) = v\,\hat{e}_{\theta^v}$. We can enumerate these twenty five (25) points (in any order, with the center point $\vec{u}_0$ as the first one) giving the coordinate indices $\vec{u}_0, \vec{u}_1, ..., \vec{u}_{24}$.

Step 1: For each pixel $\vec{u}_0 = (x_0, y_0)$ consider a window $W(\vec{u}_0) = \{(x_0,y_0); (x_0 \pm 1,\, y_0 \pm 1)\}$ and perform the following test:

test 1: Check that at least for one pixel pixel $\vec{u}_i \in W(\vec{u}_0)$, $|\,\mathcal{W}^{im}_{\sigma,\theta^{max}}[I](\vec{u}_i,t)\,| > \delta_1$ and $|\,\partial_t[G_\sigma * I(\vec{u}_i,t)]\,| > \epsilon_1$, where δ_1 and ϵ_1 are very small values near 0 (noise tolerance). Say $\delta_1 = \epsilon_1 = 0.03$

If passed test1, move to test 2.

test 2:

Construct $\mathbf{A}_{\vec{u}_0} = \begin{bmatrix} \mathcal{W}^{im}_{\sigma,\theta^{max}}[I](\vec{u}_0,t)\,\cos(\theta^{max}_0) & \mathcal{W}^{im}_{\sigma,\theta^{max}}[I](\vec{u}_0,t)\,\sin(\theta^{max}_0) \\ \vdots & \vdots \\ \mathcal{W}^{im}_{\sigma,\theta^{max}}[I](\vec{u}_{24},t)\,\cos(\theta^{max}_{24}) & \mathcal{W}^{im}_{\sigma,\theta^{max}}[I](\vec{u}_{24},t)\,\sin(\theta^{max}_{24}) \end{bmatrix}$

create $\mathbf{A^T A}_{\vec{u}_0} = \mathbf{A}^T_{\vec{u}_0}\,\mathbf{A}_{\vec{u}_0}$ and find its two eigenvalues, $\lambda_1 \geq \lambda_2$.

Example shown in figure 5

If $\lambda_2 \geq \eta = 0.1$, we can invert $\mathbf{A^T A}_{\vec{u}_0}$ and solve

$$\mathbf{A}_{\vec{u}_0} \begin{bmatrix} v_x \\ v_y \end{bmatrix} = - \begin{bmatrix} \partial_t[G_\sigma * I(\vec{u}_0, t)] \\ \vdots \\ \partial_t[G_\sigma * I(\vec{u}_8, t)] \end{bmatrix} \quad \Rightarrow \quad \vec{v}_\sigma(\vec{u}) = \begin{bmatrix} v_x \\ v_y \end{bmatrix} = - \left(\mathbf{A}^T\,\mathbf{A}\right)^{-1}\,\mathbf{A}^T \begin{bmatrix} \partial_t[G_\sigma * I(\vec{u}_0, t)] \\ \vdots \\ \partial_t[G_\sigma * I(\vec{u}_{84}, t)] \end{bmatrix} \tag{3}$$

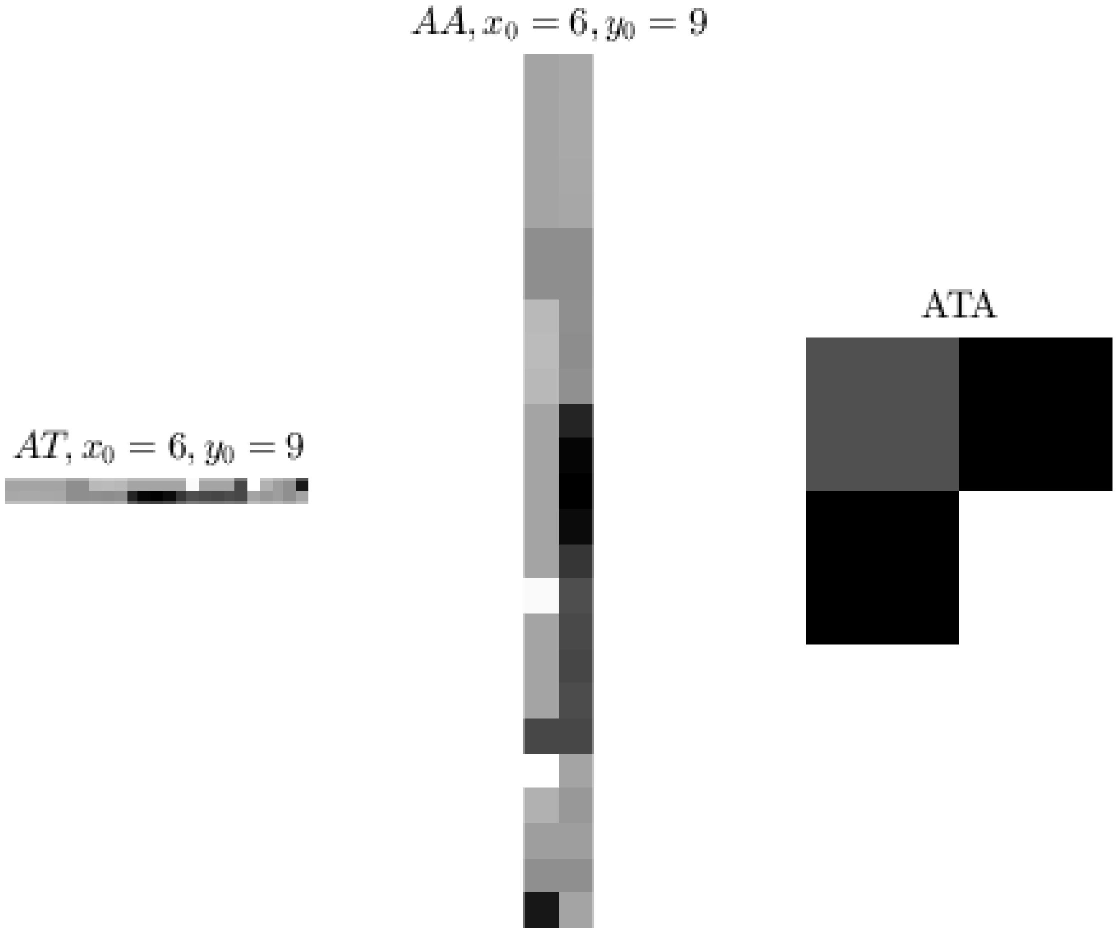
else if $\lambda_2 < \eta$ (test 2 is not passed), we at least know one component of the solution, namely

$$\vec{v}_{\sigma,\theta^{max}}(\vec{u}) \quad = \quad \frac{\partial_t[G_\sigma * I(\vec{u}, t)]}{\mathcal{W}^{im}_{\sigma,\theta^{max}}[I](\vec{u}, t)}\,\hat{e}_{\theta^{max}_{\vec{u},t}} \tag{4}$$

This is a partial solution, we only know the component of $\vec{v}_\sigma(\vec{u})$ at direction $\hat{e}_{\theta^{max}_{\vec{u},t}}$.

Step 2: Display the solution $\vec{v}^\sigma(\vec{u})$ in the 100 x 100 array, so the vector $\vec{v}^\sigma(\vec{u})$ is shown only at every 6 pixels, as arrows with different lengths and directions. Shown an array with the confidence of the solution, it is 0 if the test 1 is not passed. It is 1 if test 1 passes but not test 2 (only one non-zero eigenvalue), and it is 2 if both tests passes (two non-zero eigenvalues).

Solutions to the Mini Cooper stereo pair: we used $\sigma = 12$



(a) $\mathbf{A}^T$ (b) $\mathbf{A}$ (c) $\mathbf{A}^T\,\mathbf{A}$

Figure 5: *A window of size* 5×5 *pixels around each pixel is used to compute the following matrices. We show at pixel* $(x_0 = 6, y_0 = 9)$ *a. Matrix* $\mathbf{A}^T$*, a* 2×25 *matrix, the transpose of b.* $\mathbf{A}$*, which is a* 25×2 *matrix with x and y components of the image gradient at each of the 25 pixels. c. Finally, the* 2×2 *matrix* $\mathbf{A}^T\,\mathbf{A}$.

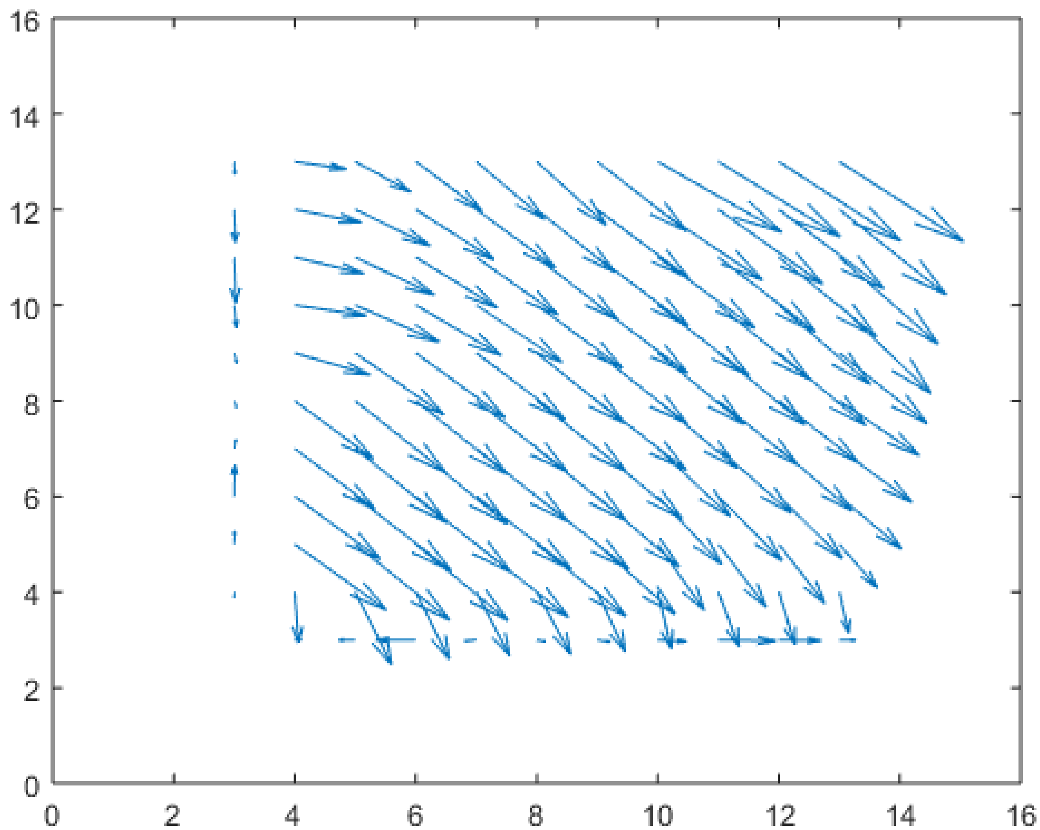
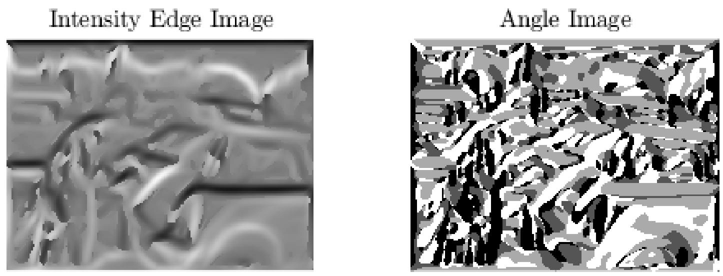
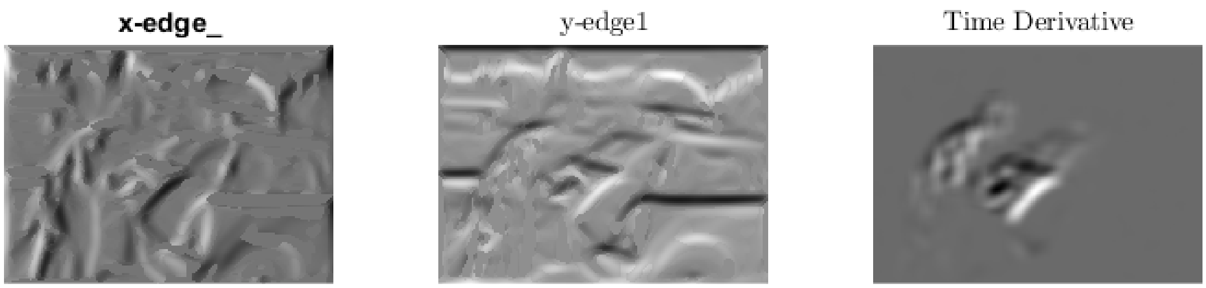


Figure 6: *An optical flow solution to the square problem at scale $\sigma = 6$ pixels.*

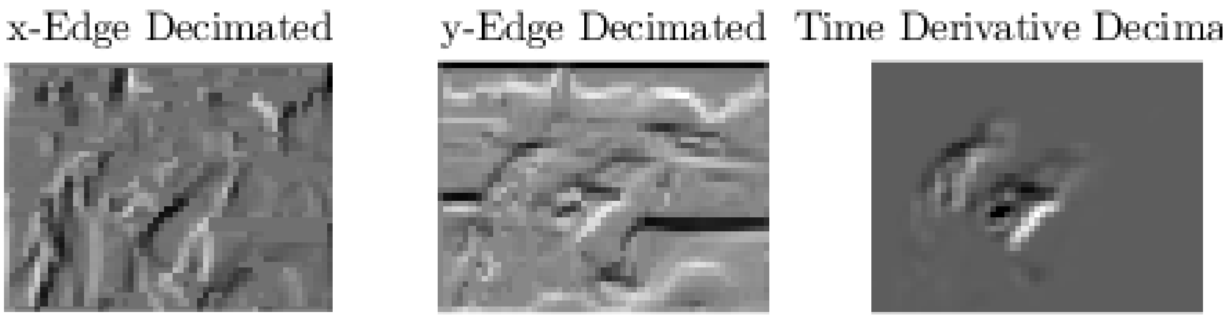


(a) $\mathcal{W}^{im}_{\sigma,\theta^{max}}[I](\vec{u},t)$ (b) $\theta^{max}_{\vec{u},t}$



(c) $\approx \partial_x(G_\sigma * I)$ (d) $\approx \partial_y(G_\sigma * I)$ (e) $\approx \partial_t(G_\sigma * I)$

Figure 7: *All images computed at scale $\sigma = 12$ pixels. The image gradient, $\nabla I(\vec{u})$, is approximately computed at every pixel by the imaginary component of the wavelet response, selected per pixel as the maximum absolute value across all angles, i.e., $\mathcal{W}^{im}_{\sigma,\theta^{max}}[I](\vec{u},t)\,(\cos\theta^{max}_{\vec{u},t}, \sin\theta^{max}_{\vec{u},t})$. a. The best imaginary wavelet component per pixel, across the four angle responses, $\mathcal{W}^{im}_{\sigma,\theta^{max}}[I](\vec{u},t)$. b. $\theta^{max}(\vec{u})$, the best angle response of the imaginary Wavelet for each pixel. c. and d. are the x and y components of ∇I, i.e., c. $\mathcal{W}^{im}_{\sigma,\theta^{max}}[I](\vec{u},t)\,\cos\theta^{max}_{\vec{u},t}$ and d. $\mathcal{W}^{im}_{\sigma,\theta^{max}}[I](\vec{u},t)\,\sin\theta^{max}_{\vec{u},t}$. e. Temporal difference of the images, after Gaussian blur.*



(a) (b) (c)

Figure 8: *Images from c., d., e., of above figure, decimated for every $\sigma = 12$ pixels. c. and d. are the x and y components of ∇I, i.e., a. $\mathcal{W}^{im}_{\sigma,\theta^{max}}[I](\vec{u},t)\ \cos\theta^{max}_{\vec{u},t}$ and b. $\mathcal{W}^{im}_{\sigma,\theta^{max}}[I](\vec{u},t)\ \sin\theta^{max}_{\vec{u},t}$, decimated and c. Temporal difference of the images, after Gaussian blur, decimated.*

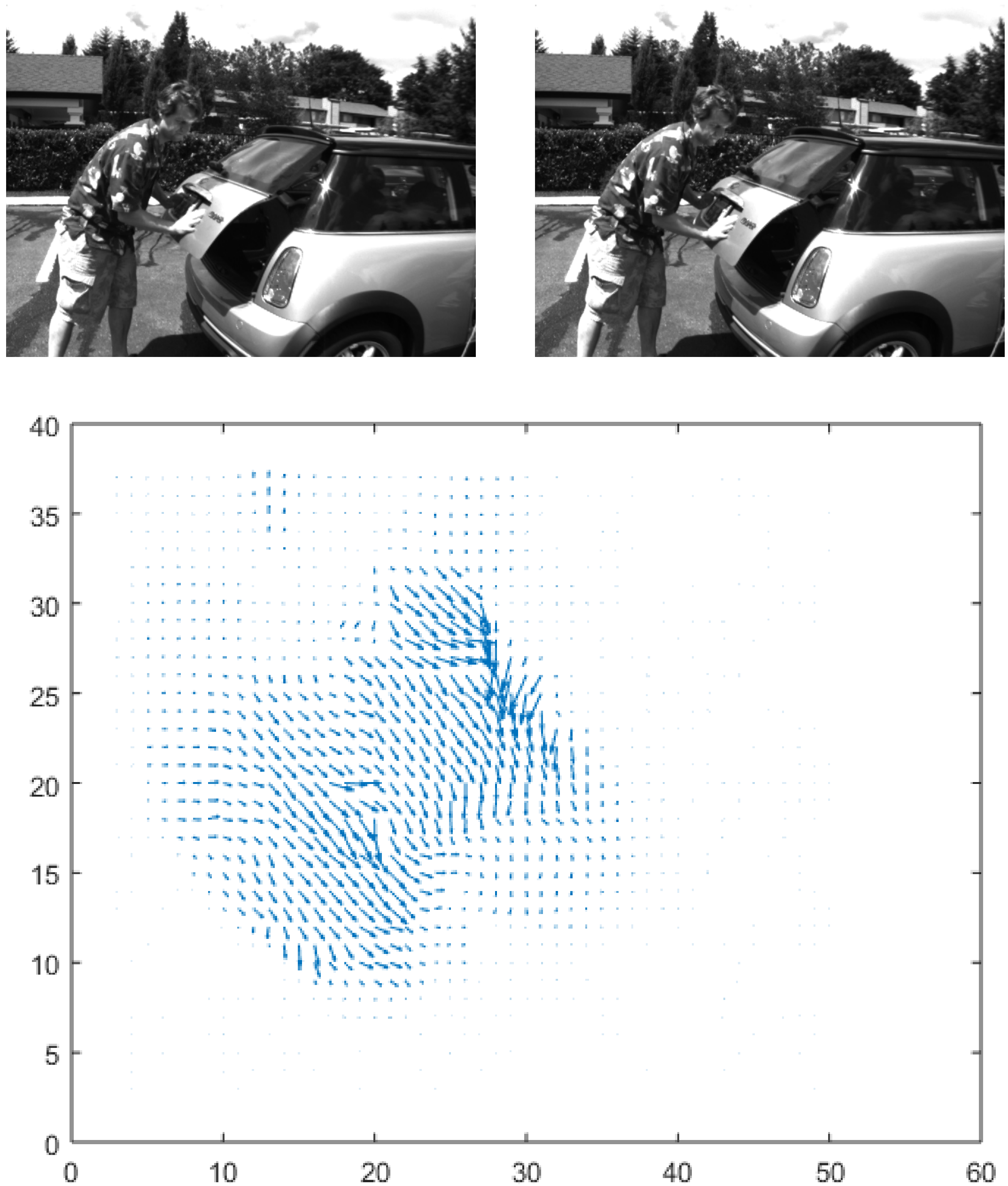


Figure 9: *An optical flow solution to the Mini Cooper problem at scale* $\sigma = 12$ *pixels.*