Discriminative Filters for Depth from Defocus

McGill Centre for Intelligent Machines

Fahim Mannan and Michael S. Langer
School of Computer Science, McGill University

Discriminative Filters Model

Contributions We propose:

1. A discriminative model for depth from defocus that encapsulates a large set of existing models.
2. An optimization objective for the proposed model that accepts solutions to existing models and more.

Summary Depth from defocus can be considered as applying a set of depth discriminative filters (bottom-left) to a set of defocused image patches (top). These filters can either be analytically derived, or estimated from calibrated PSFs, or learned from defocused images. The filters can be represented using a set of matrices $W = \{W_1, \ldots, W_d, \ldots, W_N\}$ of size $\mathbb{R}^{M \times K^2 L}$, where M is the number of filters of size $K \times K$ for L patches with depth $d \in [1, N]$. Depth from defocus solves the problem:

$$\operatorname*{argmin}_d \|W_d x\| \tag{1}$$

The estimated depth is d^*, if for any other depth d, $\|W_{d^*} x\| < \|W_d x\|$.

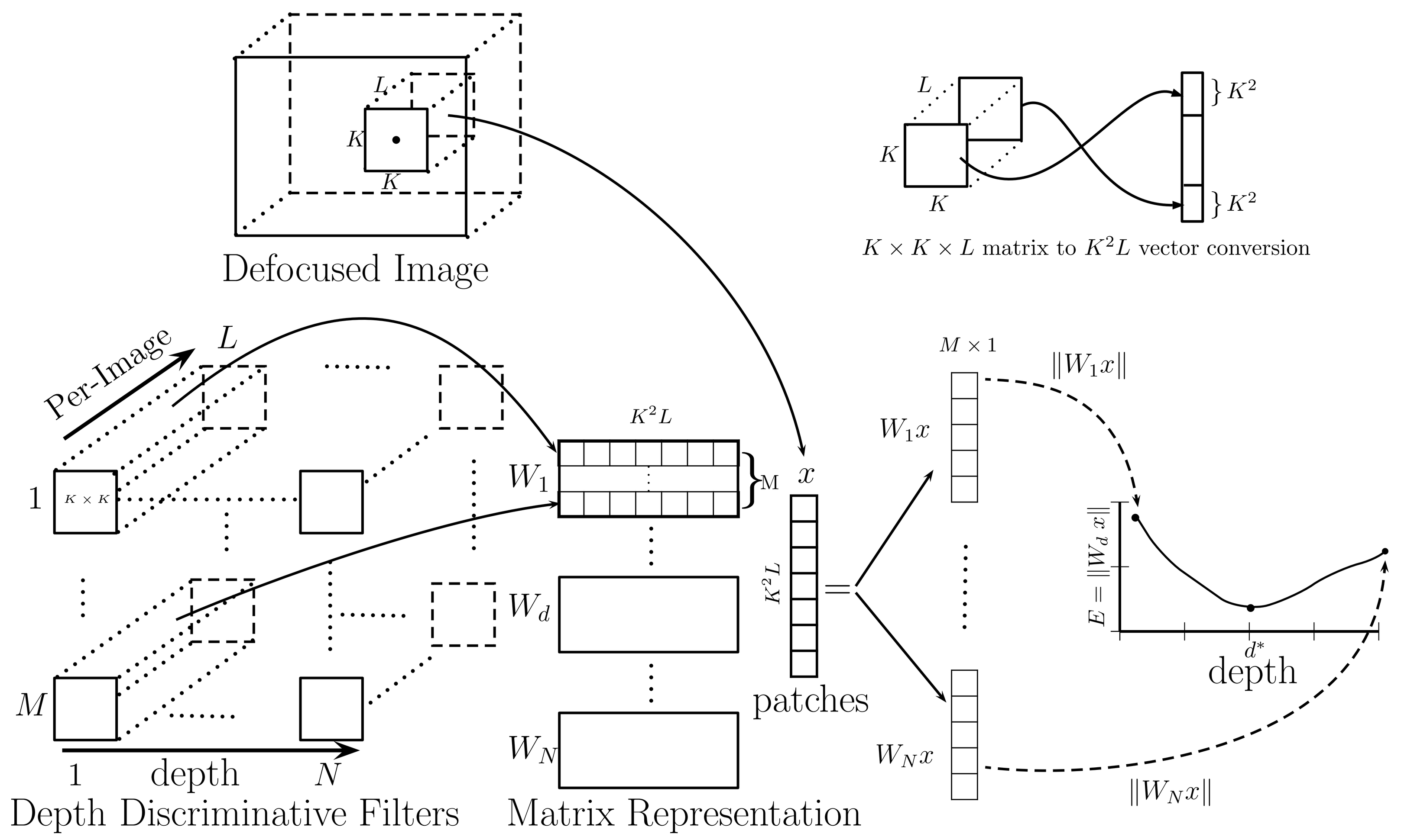

Figure 1: Depth estimation using blur discriminative filters. The filter bank with the lowest energy indicates the depth of the observed patches.

Representation of Different Models

Relative Blur $x_B = h_R * x_S + \mathcal{N}(0, \sigma_N^2)$.

$$\operatorname*{argmin}_d \left\| \begin{bmatrix} I, & -H_R(d) \end{bmatrix} \begin{bmatrix} x_B \\ x_S \end{bmatrix} \right\|^2. \tag{2}$$

Blur Equalization (BET) $x_1 * h_2 = x_2 * h_1 + \mathcal{N}(0, \sigma_N^2)$.

$$\operatorname*{argmin}_d \left\| \begin{bmatrix} H_2(d) & -H_1(d) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right\|^2. \tag{3}$$

Deconvolution $\operatorname{argmin}_{x_0, d} \|h_d * x_0 - x\|^2 + \lambda \|C * x_0\|^2$.

$$\operatorname*{argmin}_d \|(I - H_d H_d^+) x\|^2. \tag{4}$$

Subspace Projection

$$\text{Null space}: \operatorname*{argmin}_d \|U_N^T(d) x\|^2. \tag{5}$$

$$\text{Rank space}: \operatorname*{argmin}_d -\|U_R^T(d) x\|^2. \tag{6}$$

Optimal Filters

Let x^t be the t-th training image patch and y_t be the associated depth, then the filter set $W = \{W_1, \ldots, W_d, \ldots, W_N\}$ is found by solving

$$\operatorname*{argmin}_W \rho(W) \tag{7}$$
$$\text{subject to } \|W_{y_t} x^t\| < \|W_j x^t\| \; \forall t, j \; j \neq y_t$$

Here $\rho(W)$ is a regularization function on the filters which in our case is the squared Frobenius norm. The unconstrained form is,

$$\operatorname*{argmin}_W \lambda_1 \rho(W) + \frac{\lambda_2}{N} \sum_{t,j} \rho_l(y_t, j) \max(0, \|W_{y_t} x^t\| - \|W_j x^t\| + m). \tag{8}$$

In our experiments, margin $m = 1$, and $\rho_l(y_t, j) = (a|y_t - j|)^k$, with $k \in [0, 2]$ and $a \in (0, 1]$.

Evaluation

Average Cost Comparison Each column is an average of a few thousand image patches. Blue indicates 0 and bright yellow 1. Our method forces the minimum to be on the diagonal.

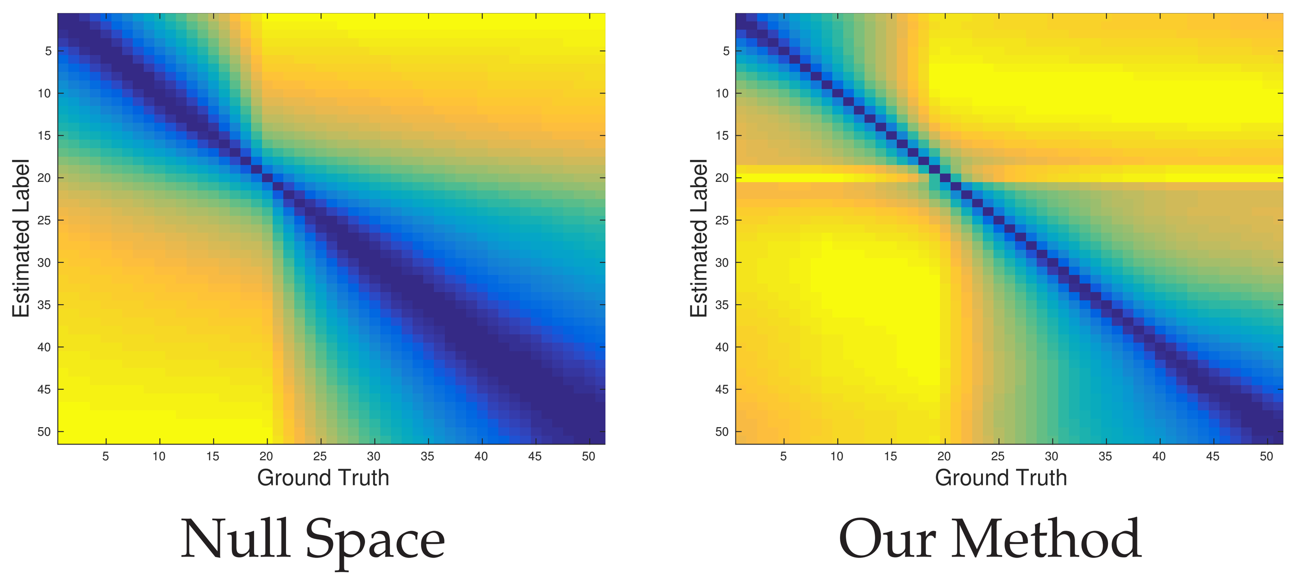

Figure 2: Results for a pair of defocused images with Gaussian PSFs. The scene depth range is 52.9 cm to 86.9 cm. The camera focal length is 25 mm and aperture $f/8.3$. The largest blur radius is ≈ 2.3 pixels. In this experiment, the blur scale is divided into 51 discrete values.

Mean and Variance Comparison