Just Noticeable Defocus Blur Detection and Estimation

Jianping Shi†, Li Xu‡, Jiaya Jia†
‡ Image & Visual Computing Lab, Lenovo R&T. ‡ Image & Visual Computing Lab, Lenovo R&T.

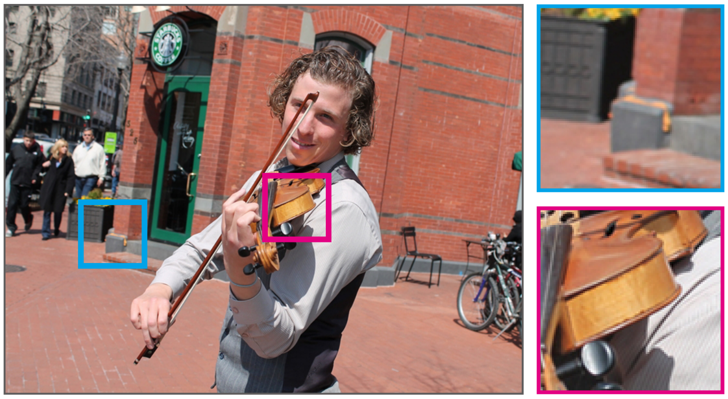


Figure 1: This image found over internet was captured with aperture size f/5.6 and exposure time 1/500s. When this supposedly clear image is viewed in its original resolution, slight blurriness can still be noticed. It is a general phenomenon.

Photos awaken our pictorial memories. A good photo generally contains clear and sharp objects that are important. With the prevalence of high-resolution imaging sensors, blurriness and its spatial change become perceivable. We name the commonly occurred small defocus blur *just noticeable blur* (JNB), which is formally defined as blur spanning about 3-9 pixels and losing a quantitatively insignificant level of structures. It commonly exists in images. It actually gives us useful information to understand the scene. A typical example is shown in Fig. 1, where sight blurriness implies foreground and the salient object we should notice.

We show a new direction to understand small image blur via sparse representation based on external data. Specifically, we found that when decomposing local image patches into dictionary atoms in an additive manner, clear and JNB dictionaries show quantitatively and visually different results. The diverged effect manifests that dictionary atoms can characterize structure in just noticeable blur images, thus amplifying the inherent difference between slight blur and clear regions. Based on it, we propose our simple but expressive JNB feature. It is verified on image data in accordance with our finding.

Sparsity JNB Feature Our new blur metric learns a blur dictionary D following Eq. (1)

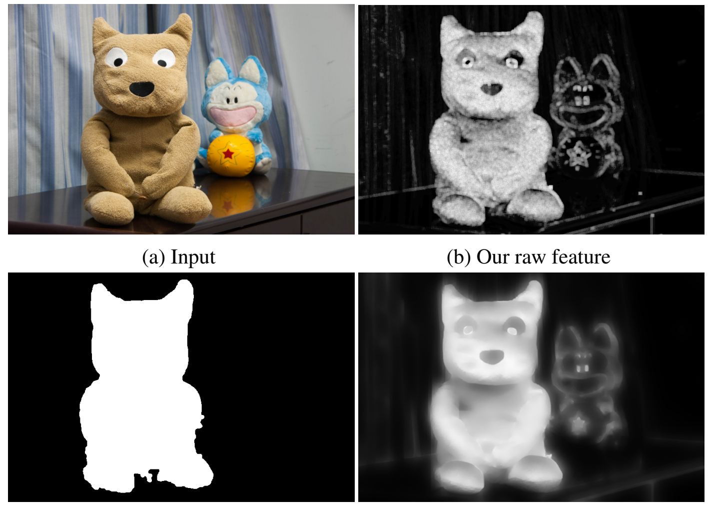
$$\min_{x_i} \|y_i - Dx_i\|_2^2 \qquad s.t.\ \|x_i\|_0 \leq k. \tag{1}$$

The input is a set of n signals $Y = \{y_1, \ldots, y_n\} \in \mathbb{R}^{d\times n}$, which consists 100,000 patches randomly cropped from 1,000 natural images blurred by the Gaussian kernel of $\sigma = 2$. $D \in \mathbb{R}^{d\times m}$ is an over-complete dictionary capturing all atomic information lying in Y. x_i is the coefficient to reconstruct y_i. The maximum number k corresponding to the used dictionary atoms is set to 5 in patch decomposition and the total dictionary size is 128. Basically, the sparse representation is to use dictionary atoms to capture elementary information.

After D is learned, it is applied to all image patches, both JNB and clear, for blur identification. For each new patch input y_i, we use another spare representation to decompose it into basic atoms. It is expressed as

$$\min_{x_i} \|x_i\|_1 \qquad s.t.\ \|y_i - Dx_i\|_2 \leq \varepsilon, \tag{2}$$

where ε is a constant (0.07 in our experiment). Different from the traditional form that selects a relatively large ε to resist noise and outliers, we set this value small to make the resulting averaged PSNR between the original and reconstructed patches over 50. This special setting is based on the consideration that detail-level structural information is central to image blur assessment in human perception.



(a) Input (b) Our raw feature

(c) Binary map (d) Our final result

Figure 2: Visual Result

The output atoms and corresponding coefficients reflect whether the input is blurred or not and how strong it is. We build our sparsity feature f_a for input y_i as the number of non-zero elements in x_i, expressed as

$$f_a = \|x_i\|_0. \tag{3}$$

Note these patches should not be flat in color in order to avoid classification ambiguity. Actually it does not matter that much if we label one flat patch as blur or clear for many applications such as deblurring and blur magnification.

We verify the generality of the phenomenon that less used dictionary atoms correspond to stronger blurriness. For blur standard deviation σ and sparsity feature values f, their relationships obey a logistic regression function as

$$f = \frac{a}{1+\exp(b\sigma + c)} + d, \tag{4}$$

where a, b, c and d are the fitted variables with corresponding values 39.49, 4.535, −3.538, and 18.53 respectively. Eq. (4) allows our system to even estimate the degree of blurriness for each patch even if it is small, and empowers spatial-varying blur strength estimation.

Experiments and Comparisons Our method does not handle flat regions due to their inherent ambiguity. As aforementioned, it does no matter to determine them as sharp or blur. We simply mask them out to indicate uncertain pixels. We fill in these holes using closed form matting [1]. The final blur map is bilateral filtered to remove noise and preserve sharp boundaries.

We provide an example in Fig. 2. (a) is the input image. Our raw feature in (b) is already powerful enough to classify the background toy as blurry. The final map in (d) is perceptually more reasonable. Given the color input and the blur map, we apply a graph-cut algorithm to label the blur region in (c). It is close to the ground-truth.

[1] Anat Levin, Dani Lischinski, and Yair Weiss. A closed-form solution to natural image matting. pages 228–242, 2008.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.