PSF RECOVERY FROM EXAMPLES FOR BLIND SUPER-RESOLUTION

Isabelle Bégin and Frank P. Ferrie

McGill University
Centre for Intelligent Machines
Montréal (Québec) Canada H3A 2A7
{ibegin, ferrie}@cim.mcgill.ca

ABSTRACT

This paper addresses the problem of super-resolving a single image and recovering the characteristics of the sensor using a learning-based approach. In particular, the Point Spread Function (PSF) of the camera is sought by minimizing the mean Euclidean distance function between patches from the input frame and from degraded versions of high-resolution training images. Once an estimate of the PSF is obtained, a supervised learning algorithm can then be used as is. Results are compared with another method for blind super-resolution by using a series of quality measures.

Index Terms— Super-Resolution, Image Quality, Point Spread Function, Learning, Markov Random Fields

1. INTRODUCTION

The subject of super-resolution has elicited considerable interest in the past few years. While standard approaches such as regularization [1], Fourier methods [2] or Bayesian techniques [3] have been widely used in the past, a lot of the recent research in super-resolution focusses on learning-based algorithms (*e.g.* [4, 5]). Initially, most methods using supervised learning techniques assumed knowledge of the camera Point Spread Function (PSF) in order to construct the training set. Generalizations were then developed to account for the case where the PSF is unknown [6, 7, 8]. However, most of these methods for blind super-resolution do not aim at obtaining the PSF, but are simply able to super-resolve an image without its knowledge. Therefore, there are very few analyses on the capacities of a learning-based method to actually recover the PSF, and on a quantification of its success.

In this paper, the problem of obtaining the PSF *as well as* a super-resolved image is addressed, from a single input frame and a set of high-resolution images from a database. The particular framework chosen was developed originally in [4]. Markov Random Fields (MRFs) are used to model a pair of low/high resolution training images and a belief propagation algorithm is used to obtain the Maximum a Posteriori estimate of the super-resolved image. While in the original method the PSF has to be known in advance, in this paper, the sensor PSF is also recovered from examples. The training dataset is artificially degraded with a parametric model of a PSF, and the model best representing the input image is chosen as the estimated PSF.

A framework based on a similar idea was developed in [9]. However, in that study, the training dataset is degraded with three different models and the PSF recovery amounts to determine which model was preferred by the belief propagation algorithm. Here, the PSF recovery is performed prior to the super-resolution process, to avoid potential tradeoffs between a good PSF and a good super-resolved image. Furthermore, here the super-resolution results are analyzed through the use of various similarity measures, representing different aspects of image quality: the quality of the signal, the perceptual quality and the sharpness of edges. The use of these measures is important not only to show an improvement in quality with respect to the input image, but also to detect the impact of various parameters on the results.

2. RECOVERING THE PSF FROM EXAMPLES

One of the main hypotheses of the method of [4] is that the belief propagation process can be applied on a set of candidates. Given an input image patch, a set of n closest patches (in terms of the Euclidean distance) are found in the low-resolution training database, previously obtained by blurring, subsampling and reinterpolating the high-resolution training images. Each input patch is thus associated with n pairs of low/high training patches. These n pairs are modeled in a MRF and the belief propagation algorithm determines which of the pairs best represents the input patch[1]. The method presented here is inspired from this idea. Since Euclidean distance is used to determine a set of candidates, it is also used to determine the best possible PSF.

Let an image I_{low} be an input image to be super-resolved and the set $T = \{T_1, T_2, ..., T_{N_t}\}$ be the N_T high-resolution training images. Breaking all images into patches, the set of patches for I_{low} is denoted by $I_p = \{p_1, p_2, ..., p_K\}$, and the set of patches for all images in T is: $T_p = \{t_{p1}, t_{p2}, ..., t_{pM}\}$.

[1] The reader is referred to [4] for more details about the algorithm

The training images in T are blurred with a Gaussian PSF of variance σ_b^2, and I_{low} is assumed to have been blurred with a Gaussian PSF of unknown variance σ_t^2.

What is proposed here is to obtain the PSF parameter only using the Euclidean distance between input and training local patches. The Mean Euclidean Distance (MED) between the training dataset and the input patches is calculated for a series of blurring parameters σ_b^2. The parameter giving the lowest MED will be the recovered parameter σ_r^2. In other words, the MED for a parameter σ_b^2 is:

$$D(\sigma_b^2) = \frac{1}{N} \sum_i d(p_i(\sigma_t^2), t_{c_i}(\sigma_b^2)), \tag{1}$$

where σ_t^2 is the true variance, $d(p_i, t_{c_i})$ is the Euclidean distance between input patch p at location i and its corresponding closest candidate t_{c_i} in the training database. Therefore, for a series of σ_b^2, the recovered parameter σ_r^2 is:

$$\sigma_r^2 = \underset{\sigma_b^2}{\text{argmin}}\ \mathbf{D}(\sigma_b^2). \tag{2}$$

To obtain the minimum of the function, a Golden Section Search (GSS) [10] was chosen. Given a range within which the true solution is assumed to lie, the function is sampled at values equidistant from both ends of the range. Given the output of the function, the range is narrowed and the process continues until a predefined tolerance is reached.

However, to avoid local minima and to limit the number of computations, the initial range must be fairly narrow. The work presented in [9] provides a way to limit the range of possibilities by making use of already existing algorithms for blind deconvolution in order to get a first estimate, and obtain an uncertainty around this estimate so that the search can be constrained. A short description is provided below.

First, a blind Lucy-Richardson (LR) algorithm [11] provides a first estimate of the PSF, and the resulting deconvolved image is not used by the system. The output of the algorithm is the convolution kernel, and the closest Gaussian is found by minimizing the Root-Mean-Square (RMS) distance between the Gaussian function and the values of the kernel. The variance of this Gaussian, σ_e^2, becomes the PSF estimate.

Second, an uncertainty is computed based on a synthetic texture, a horizontal line (1-pixel wide) of maximum intensity (here 255) while the rest of the image has minimum intensity. This texture provides a single sharp feature to illustrate the effect of the PSF. A Gaussian and rotationally symmetric PSF is chosen so that its convolution with the texture reduces the line's intensity value by half (at its original location), the rest of the intensity being dispersed to its neighbours.

This criterion allows the standard deviation of the PSF to be easily calculated. This can be expressed as (σ_f is the standard deviation and K_f is a constant):

$$\sum_{x=-1}^{1} \sum_{y=-1}^{1} G_f(x,y) = \sum_{x=-1}^{1} \sum_{y=-1}^{1} K_f\, e^{-\frac{(x^2+y^2)}{2\sigma_f^2}} = 1, \tag{3}$$

$$255\,(G_f(-1,0) + G_f(0,0) + G_f(1,0)) = \frac{255}{2}. \tag{4}$$

From Eqs. 3 and 4 it can be shown that $\sigma_f^2 = 0.72$. For a kernel of size 5×5, the uncertainty can be shown to be $\sigma_f^2 = 0.64$.

In this paper, the uncertainty obtained above determines the range within which the GSS will be performed: $r = [\sigma_e^2 - \sigma_f^2,\ \sigma_e^2 + \sigma_f^2]$. If $\sigma_f^2 > \sigma_e^2$, then a minimum value of 0.1 is imposed. Once σ_r^2 is obtained, the standard belief propagation algorithm of [4] can be applied as is. This procedure is denoted as the PSF-GSS/BP-MRF framework, and can be easily adapted to other PSF models, such as a pillbox.

3. SIMILARITY MEASURES

To assess the quality of the super-resolved images, three different quality measures are used, representing different aspects of image quality. In all cases, quality is assumed to be the level of similarity with the ground truth high-resolution image. First, the Peak Signal to Noise Ratio (PSNR) is used to measure the quality of the signal: $\text{PSNR} = 20 \log_{10} \frac{\text{MAX}_I}{\text{RMSE}}$, where MAX_I is the maximum possible intensity (here 255) and RMSE is the Root Mean Square Error.

Second, the Edge Stability Mean Square Error (ESMSE) [12] is used to quantify the localization and scale of edges. An edge detector at five different scales is applied to both the ground truth and the super-resolved images. The edge stability map is obtained by finding, for each pixel, the longest uninterrupted sequence of edge presence along scales. The EMSE is the MSE between the maps of the ground truth and super-resolved images. In this paper, the measure is transposed into a PSNR for easier comparison with the other measures, and is referred to as the Edge Measure (EM).

Finally, a measure that has attracted considerable interest recently is the Structural Similarity Measure (SSIM) [13], consisting of comparisons of the luminance, the contrast and the structure of both the distorted ($\mathbf{Y}$) and the true ($\mathbf{X}$) signals. The comparisons are performed in local windows leading to a SSIM map. The Mean SSIM (MSSIM) can then be calculated for the entire image. The reader is invited to consult [12, 13] for the formal definitions of the ESMSE and the SSIM measures.

All results are expressed in terms of the improvement of similarity (in percentage): $\text{SI} = 100 \cdot \frac{M_{rh} - M_{lh}}{M_{lh}}$, where M_{rh} is the value of the similarity measure between the super-resolved and the high-resolution images, and M_{lh} is the value of the measure between the low and the high-resolution images. A result above zero indicates an improvement with respect to the low-resolution image.

4. EXPERIMENTS AND RESULTS

The proposed framework was tested on images of various types, and the quality of the super-resolved images is assessed

a **b** **c** **d** **e** **f** **g** **h** **i**

Fig. 1. Details of images for the experiments. **a**, **d**, **g** are the original high-resolution images, **b**, **e**, **h** are the input low-resolution images (reinterpolated using bilinear interpolation), and **c**, **f i** are the super-resolved images using the proposed method. Three databases are used for the examples: a Fingerprint, a Car and a MRI database.

using the similarity measures described above. All ground truth images were blurred using a 5×5 Gaussian PSF of variance σ_t^2 and subsampled by a factor of 2 to obtain the input low-resolution images. For each input image, a training database of high-resolution images of the same nature is available. Examples of input and ground truth images, as well as super-resolved images are shown in Fig. 1[2].

Results of the similarity measures are shown in Fig. 2**a,c,e** with respect to a series of true parameters σ_t^2. What can be seen is that the similarity improvements vary greatly with the image used and the type of measure. In general, however, the MSSIM improvements are always higher than those of the PSNR. The EM improvements show more variations, with fairly low values for the Fingerprint database but high values for the Car and MRI image. Also, the values for all measures seem to have a peak around $\sigma_t^2 = 3$ and then reach a plateau for larger variances.

The error on the refined variance with respect to the true variance is shown in Fig. 2**b,d,f** for a blind deconvolution process and for the algorithm presented here. In general, the results are as good or better than the blind LR algorithm. It can be seen that the variance is fairly well estimated for values around $\sigma_t^2 = 3$, however the error raises sharply for higher values. This is due to the fact that the size of the blurring kernel was kept small (5×5), thus the effect of increasing the variance will be nil at some threshold.

For low values of σ_t^2 and for the Fingerprint and MRI examples, the proposed method allowed in general a good refinement of the PSF estimated by the blind deconvolution algorithm. For the Car example, however, results compare to the blind deconvolution estimates, thus indicating that the type of image and the training database available will greatly influence the results.

The same series of experiments were also performed using the method proposed in [9], referred to as the PSF-BP method. For space reasons, only the most significant results are reported. Comparisons for the PSF recovery procedure are shown in Figs. 3 for the Fingerprint and Car examples, and Fig. 4 show comparisons of the obtained similarity improvements for the Fingerprint and the MRI examples. In Fig. 3**a** (for the Fingerprint example), both the PSF-BP and the PSF-GSS methods in general perform better than standard blind deconvolution for low values of σ_t^2. For the Car example, shown in Fig. 3**b**, the PSF-GSS method gives in general higher errors. For the image quality measures, results vary strongly with the type of image used. For the MRI example the PSF-GSS method gives much better results, while for the Fingerprint image the PSF-GSS method gives a less pronounced improvement.

[2]The images were provided from the Fingerprint Verification Competition (http://bias.csr.unibo.it/fvc2000/), the McGill's Artificial Perception Laboratory and the Montreal Neurological Institute (McGill University).

5. CONCLUSIONS AND FUTURE WORK

A new PSF refinement procedure for blind super-resolution, was presented, based on the Euclidean distance between image patches from the input image and patches from degraded versions of high-resolution training images. Modeling the PSF as a rotationally symmetric Gaussian, the PSF-GSS procedure is able to refine an estimate of the variance, previously obtained from a standard blind deconvolution algorithm. Furthermore, a series of image similarity measures representing different aspects of quality was used to quantify the effect of this PSF refinement algorithm on a learning-based super-resolution process. It was found that the PSF-GSS method in general leads to higher image quality than another blind super-resolution algorithm. Another important contribution of this paper is comparison with a standard blind deconvolution algorithm for PSF recovery, which has generally not been done in blind super-resolution studies (e.g. [8]).

Future work will include the use of other distance metrics for a PSF refinement process from examples, as well as comparisons with other blind super-resolution methods.

6. REFERENCES

[1] N. Nguyen, P. Milanfar, and G. Golub, "Efficient generalized cross-validation with applications to parametric image restoration and resolution enhancement," *IEEE Trans. on Image Processing*, vol. 10, no. 9, pp. 1299–1308, 2001.

[2] R.Y. Tsai and T.S. Huang, "Multiframe image restora-

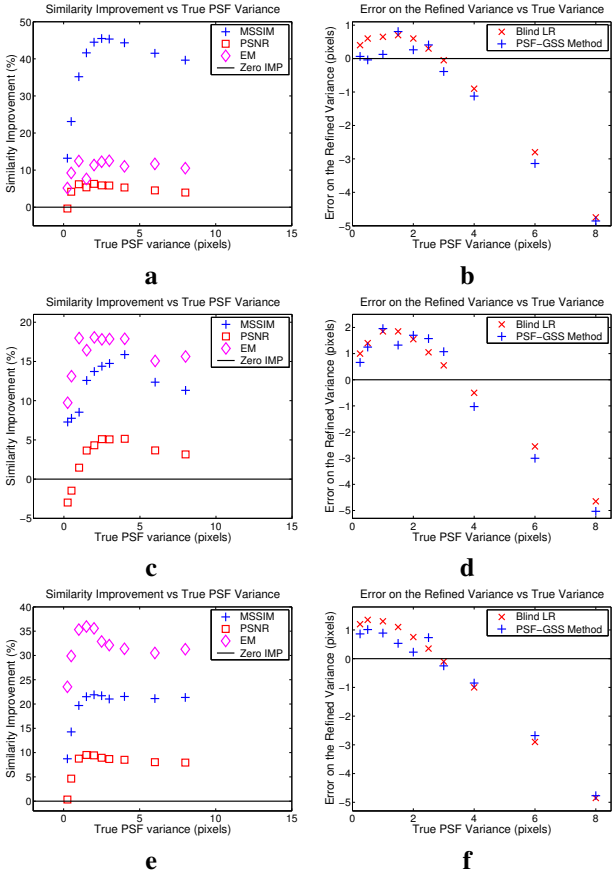

Fig. 2. Results using the PSF-GSS/BP-MRF method. **a-c:** Similarity improvement vs true PSF variance, **b-d:** Error on refined variance vs true PSF variance. The error on the refined variance is here defined as $e = \sigma_r^2 - \sigma_t^2$. The results in **a,b** are for the Fingerprint image, those in **c,d** are for the Car image and the images in **e,f** are for the MRI image.

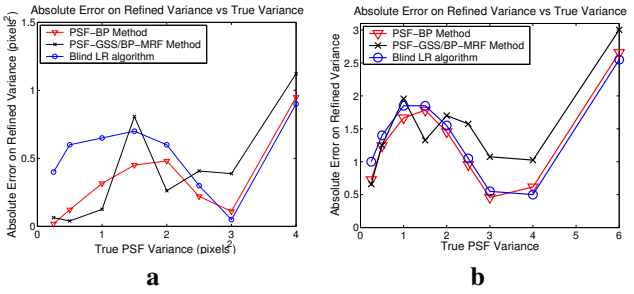

Fig. 3. Comparison of the PSF-GSS/BP-MRF and the PSF-BP methods: Absolute error on the refined variance vs true PSF variance for the Fingerprint (**a**) and the Car (**b**) databases.

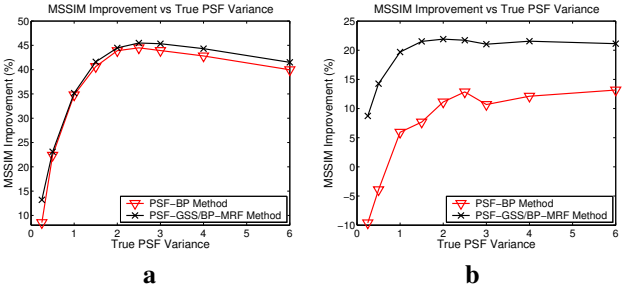

Fig. 4. Comparison of the MSSIM improvement for the PSF-GSS/BP-MRF and PSF-BP methods. **a** show results for the Fingerprint image and **b** show results for the MRI image.

tion and registration," in *Adv. in Computer Vision and Image Processing*. 1984, vol. 1, pp. 318–339, JAI Press.

[3] R.R. Schultz and R.L. Stevenson, "A Bayesian approach to image expansion for improved definition," *IEEE Trans. Image Processing*, vol. 3, no. 3, pp. 233–242, 1994.

[4] W.T. Freeman, E.C. Pasztor, and O.T. Carmichael, "Learning low-level vision," *Int. Journal of Computer Vision*, vol. 40(1), pp. 25–47, October 2000.

[5] S. Baker and T. Kanade, "Limits on super-resolution and how to break them," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2000, vol. 2, pp. 372–379.

[6] R. Rosales, K. Achan, and B. Frey, "Unsupervised image translation," in *Proc. Int. Conf. on Computer Vision*, 2003, pp. 472–478.

[7] M. D. Gupta, S. Rajaram, N. Petrovic, and T. S. Huang, "Restoration and recognition in a loop," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2005, vol. 1, pp. 638–644.

[8] Q. Wang, X. Tang, and H. Shum, "Patch based blind image super resolution," in *Proc. IEEE Int. Conf. on Computer Vision*, 2005, vol. 1, pp. 709–716.

[9] I. Bégin and F. P. Ferrie, "Blind super-resolution using a learning-based approach," in *Proc. Int. Conf. on Pattern Recognition*, 2004, vol. 2, pp. 85–89.

[10] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes in C*, Cambridge University Press, 2nd edition, 1999.

[11] D.S.C. Biggs and M. Andrews, "Acceleration of iterative image restoration algorithms," *Applied Optics*, vol. 36, no. 8, pp. 1766–1775, 1997.

[12] İ Avcibaş, B. Sankur, and K. Sayood, "Statistical evaluation of image quality measures," *Journal of Electronic Imaging*, vol. 11, no. 2, pp. 206–223, 2002.

[13] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. on Image Processing*, vol. 13, no. 4, 2004.