Defocus Inpainting

Paolo Favaro[1] and Enrico Grisan[2]

[1] Siemens Corporate Research,
Princeton, NJ 08540, USA
paolo.favaro@siemens.com
[2] Department of Information Engineering,
Universitá di Padova, Italy
enrico.grisan@dei.unipd.it
http://www.dei.unipd.it/enrigri

Abstract. In this paper, we propose a method to restore a single image affected by space-varying blur. The main novelty of our method is the use of recurring patterns as regularization during the restoration process. We postulate that restored patterns in the deblurred image should resemble other sharp details in the input image. To this purpose, we establish the correspondence of regions that are similar up to Gaussian blur. When two regions are in correspondence, one can perform deblurring by using the sharpest of the two as a proposal. Our solution consists of two steps: First, estimate correspondence of similar patches and their relative amount of blurring; second, restore the input image by imposing the similarity of such recurring patterns as a prior. Our approach has been successfully tested on both real and synthetic data.

1 Introduction

In many instances, images contain recurring patterns that are similar up to some transformation group. For example, the image of a tree may contain multiple instances of the same leaf at different locations, scales and orientations. In more specific applications, such as corneal imaging, one may find repeated patterns of cells or clusters of cells (see Figure 1). Due to the large aperture of the microscope, cells are not only similar up to an affine transformation, but also up to defocus. In other words, there may be cells in some locations that are blurred version of other cells. Then, one could think of restoring those cells by using the corresponding ones that are less blurred. More in general, if we are interested in restoring images belonging to a specific domain, such as corneal imaging, then one can exploit more than the lone input image. One could also use a database of corneal images to find more recurring patterns. This kind of approach is very similar in spirit to "hallucination" methods [1] which have been applied to faces with success. Our approach can be seen as an extension to these methods, which are limited to a single recurring pattern (e.g. a face) and whose position is known. In this paper, however, to keep the method focused, we consider the restoration problem in the simple case when the database is made of a single image. The extension to multiple images is straightforward.

Exemplar-based methods for inpainting [2] are also similar to our approach. As in [2], we look for *exemplars* that can be used to restore the input image. However, while

A. Leonardis, H. Bischof, and A. Pinz (Eds.): ECCV 2006, Part II, LNCS 3952, pp. 349–359, 2006.

Fig. 1. Oftentimes, images exhibit recurring patterns. For example, in natural images such patterns may be the leaves of a plant or the petals of a flower (left image). In more specific domains, such as in corneal imaging, the recurring pattern is made of cells of the cornea (right image). Notice how both the examples contain regions that are not only similar up to translation, rotation and scale, but also up to the amount of defocus.

we share the general idea of these popular methods, we do not use the same procedural methods to recover the missing information. Rather, we perform restoration by simultaneously considering all the corresponding patterns and their relative amount of defocus. This simultaneous integration allows us to automatically take into account the overlap of patterns and find the best tradeoff between them and the original input image.

In this paper, we propose a solution to the problem of deblurring a single image affected by defocus blur. Our main contribution is the formulation of a novel regularization method, which is based on the input data. As such, we relate to *image restoration* in the field of image processing [3] and *blind deconvolution* in the field of signal processing [4], which belong to the larger class of *inverse problems* [5]. Most of these problems are formulated as a linear model (either in discrete or continuous form), where the task is to infer the unknown object by *inverting* the model. The main challenge is that such inversion is *ill-posed*, as it may lead to multiple solution, or have no solution, or be such that small variations in the input data may cause large variations in the recovered unknown. The general recipe to solve ill-posed problems is to introduce *regularization* [6]. Regularization has been applied to the inverting operator [6, 7] and/or directly to the restored image. The latter approach is also known as Tikhonov regularization [6]. Our approach falls within this type of regularization methods, as we directly operate on the unknown unblurred image. Furthermore, our main strength is that we regularize the restoration of the image by using only the image itself, thus introducing texture that is familiar to the scene.

When the input image is made only of recurring patterns, our algorithm can be used to infer a depth map of the scene. This is reminiscent of shape from texture methods [8, 9], where one recovers the local orientation of an image patch. In our case, rather than using orientation, we consider the local amount of blur as a cue for shape as it has been done in *shape from defocus* [10, 11, 12, 13, 14]. Indeed, a byproduct of our algorithm is the estimation of the relative amount of blur between two similar regions, which can be directly related to depth.

In the next section, we will introduce our model of blurred images with recurring patterns. Then, we will formalize the problem of image restoration, so that it has a non-trivial and unambiguous solution (section 3). We show that despite the complexity of the model and the unknowns, the restoration problem can be solved into two steps: First, we determine the correspondences between recurring patterns and their relative amount of blur (section 4), and second, we integrate this information to restore the input image into a global optimization scheme (section 5).

2 Modeling Defocus and Recurring Patterns

In this section, we will introduce the image formation model for scenes with recurring patterns and captured by a real aperture camera. Let us start by defining the image formation model of a blurred image $I : \Omega \subset \mathbb{R}^2 \mapsto [0, \infty)$

$$I(\mathbf{y}) = \int K_\sigma(\mathbf{y}, \mathbf{x}) f(\mathbf{x}) d\mathbf{x} + n(\mathbf{y}) \quad \forall \mathbf{y} \in \Omega \tag{1}$$

where $K_\sigma : \Omega \times \mathbb{R} \times [0, \infty)$ is called the *point spread function* (PSF) [15], $f : \mathbb{R} \mapsto [0, \infty)$ is the unblurred texture of the scene and $n : \Omega \mapsto \mathbb{R}$ collects noise and distortions that are not captured by the linear term in eq. (1). The PSF K_σ depends on the amount of defocus encoded by the variable $\sigma : \Omega \mapsto [0, \infty)$, which is related to the depth of the scene $s : \Omega \mapsto [0, \infty)$ via [14]

$$\sigma(\mathbf{x}) = \frac{Dv}{2} \left| \frac{1}{v} + \frac{1}{s(\mathbf{x})} - \frac{1}{F} \right| \tag{2}$$

where D is the diameter of the lens, v the distance between the lens and the image plane and F is the focal length of the lens.

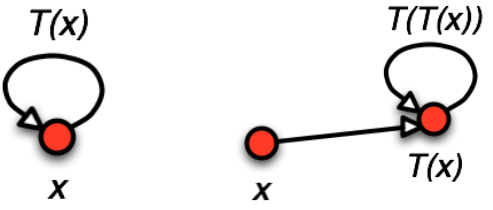
We now formalize the notion of recurrence of a pattern within a blurred image and propose a suitable representation for it. Suppose that two regions of the unblurred image $f, O \subset \Omega$ and $O' \subset \Omega$ with $O \bigcap O' = \emptyset$, are identical to each other. Define $T : \Omega \mapsto \Omega$ the mapping of points $\mathbf{x} \in O$ to points $\mathbf{x}' \in O'$, so that $T(\mathbf{x}) = \mathbf{x}'$. Then, we model a recurrence as

$$f(\mathbf{x}) = f(T(\mathbf{x})) \quad \forall \mathbf{x} \in O. \tag{3}$$

More in general, let us define the mapping T for all recurrences and let us call T the *correspondence map*. In other words, $T(\mathbf{x})$ tells us where to find the location of a region similar to the one around $\mathbf{x}$. When patterns are unique, then we simply have $T(\mathbf{x}) = \mathbf{x}$ $\forall \mathbf{x} \in O$, i.e. a *self-reference*. Notice that a generic correspondence map may generate *loops*. For example, there may be points $\mathbf{x} \neq \mathbf{y}$ such that $T(\mathbf{x}) = \mathbf{y}$ and $T(\mathbf{y}) = \mathbf{x}$. In our context such loops are unnecessary and undesirable. Hence, we will enforce that the correspondence map has only two types of mappings:

1. a self-reference, i.e. $T(\mathbf{x}) = \mathbf{x}$
2. a link to a self-reference, i.e. $T(T(\mathbf{x})) = T(\mathbf{x})$.

For clarity, see Figure 2.



(a) (b)

Fig. 2. The two types of links admitted by the correspondence map T. On the left we have a self-reference, while on the right we show a link to a self-reference.

The correspondence map T just defined is very general and captures any type of deformation of one region to another. For instance, local affine deformations are captured by using

$$T(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b} \quad \forall \mathbf{x} \in O \tag{4}$$

where $\mathbf{A}$ is a 2×2 matrix and $\mathbf{b}$ a 2-dimensional vector. Later on, we will restrict the class of parametric deformations modeled by T to simple translations, i.e. such that $\mathbf{A}$ is the identity matrix in eq. (4), and we will show how to recover the translation $\mathbf{b}$ from a blurred image.

So far, the model that we have introduced can be summarized as:

$$\begin{aligned} f(\mathbf{x}) &= f(T(\mathbf{x})) \\ I(\mathbf{y}) &= \int K_\sigma(\mathbf{y}, \mathbf{x}) f(\mathbf{x}) d\mathbf{x} \end{aligned} \tag{5}$$

where we have neglected the term n for simplicity. We assume that the blurring is locally constant, and therefore it can be modeled by a shift-invariant PSF K_σ, and that the PSF is Gaussian, i.e. such that

$$K_\sigma(\mathbf{y}, \mathbf{x}) = \frac{1}{\sqrt{2\pi\sigma^2(\mathbf{y})}} \exp^{-\frac{\|\mathbf{y}-\mathbf{x}\|^2}{2\sigma^2(\mathbf{y})}}. \tag{6}$$

We will now show in the next section how to pose the problem of deblurring with recurring regions.

3 Maximizing Deblurring

Suppose that $O \subset \Omega$ is a recurring region. Then, according to eq. (5) we have that

$$I(\mathbf{y}) = \int K_\sigma(\mathbf{y}, \mathbf{x}) f(\mathbf{x}) d\mathbf{x} = \int K_\sigma(\mathbf{y}, \mathbf{x}) f(T(\mathbf{x})) d\mathbf{x}. \tag{7}$$

If now we allow the correspondence map to capture only translations, then we have $T(\mathbf{x}) = \mathbf{x} + \mathbf{b}$. By substituting the explicit expression of T in the equation above, and

using the assumption that blurring is locally constant and Gaussian, it is easy to derive that

$$I(\mathbf{y}) = \int K_{\sigma}(\mathbf{y}+\mathbf{b},\mathbf{x}) f(\mathbf{x}) d\mathbf{x} = \int K_{\Delta\sigma}(\mathbf{y},\mathbf{x}) I(\mathbf{x}+\mathbf{b}) d\mathbf{x} \tag{8}$$

where $\Delta\sigma$ is called *relative blur* and it satisfies $\Delta\sigma^2(\mathbf{y}) = \sigma^2(\mathbf{y}) - \sigma^2(\mathbf{y}+\mathbf{b})$, $\forall \mathbf{y} \in O$. Since relative blur is meaningful if and only if $\Delta\sigma(\mathbf{y}) \geq 0$, we impose that the correspondence map T maps regions to regions that are less blurred, i.e. such that $\sigma^2(\mathbf{x}) \geq \sigma^2(T(\mathbf{x}))$. Hence, by definition, regions that are self-referencing will be subject to no blurring ($\Delta\sigma = 0$).

The main advantage of eq. (8) is that it does not depend on the unblurred image f, as eq. (5), but only on the the relative blur $\Delta\sigma$ and the correspondence map T. Hence, by using eq. (8) one can decouple the problem of simultaneously estimating all the unknowns into two problems where one first recovers the relative blur and the correspondence map and then restores the unblurred image f. In this section, we will introduce the problem of recovering the first two unknowns, while we will devote section 5 to the restoration of the unblurred image f.

Now, recall eq. (8). It is easy to see that this equation is satisfied by $\Delta\sigma = 0$ and $T(\mathbf{x}) = \mathbf{x}$. This means that given a blurred image I, a correspondence map T that is always admissible is the one such that all regions are unique and hence their mapping is a self-reference. As a consequence, the relative blur will be null everywhere. Such T and $\Delta\sigma$ do not give any advantage with respect to previous methods for deblurring. To avoid this situation, we pose the problem of finding the solution with largest relative blur. Hence, to recover $\Delta\sigma$ and T we pose the following maximization problem

$$\begin{aligned} \Delta\sigma, T = \arg\max_{\Delta\sigma} & \int \Delta\sigma^2(\mathbf{x}) d\mathbf{x} \\ \text{subject to: } & \begin{cases} I(\mathbf{y}) = \displaystyle\int K_{\Delta\sigma}(\mathbf{y},\mathbf{x}) I(\mathbf{x}+\mathbf{b}) d\mathbf{x} & \\ T(\mathbf{x}) = \mathbf{x} & \forall \mathbf{x} | \Delta\sigma(\mathbf{x}) = 0 \\ T(\mathbf{x}) = \mathbf{x}+\mathbf{b}, \mathbf{b} \neq 0 & \forall \mathbf{x} | \Delta\sigma(\mathbf{x}) > 0 \\ T(T(\mathbf{x})) = T(\mathbf{x}) & \end{cases} \end{aligned} \tag{9}$$

where the first constraint corresponds to eq. (8); the second one corresponds to having no relative blur between self-references; the third one corresponds to imposing the translational model whenever there is relative blur between two regions; finally, the fourth constraint imposes that T satisfies only the two types of mappings shown in Figure 2. This equation can also be interpreted as the maximization of the *amount of deblurring* that we will be able to perform in the second part of the algorithm (section 5).

4 Localization of Recurring Regions

In order to solve the problem in eq. (9), we need to recall the representation of the correspondence map T. The map T is defined jointly with the relative blur $\Delta\sigma$ as being either

$$T(\mathbf{x}) = \mathbf{x} \quad \forall \mathbf{x} | \Delta\sigma(\mathbf{x}) = 0 \tag{10}$$

or

$$T(\mathbf{x}) = \mathbf{x} + \mathbf{b}, \mathbf{b} \neq 0 \quad \forall \mathbf{x} | \Delta\sigma(\mathbf{x}) > 0. \tag{11}$$

In other words, the relative blur defines a partition of Ω into regions where T is equal to a constant vector $\mathbf{b}$ (Figure 3). Given this representation, we propose the following approximate algorithm to solve eq. (9):

- Initialize the map T such that $T(\mathbf{x}) = \mathbf{x} \quad \forall \mathbf{x} \in \Omega$
- Quantize the relative depth into L levels
- **for each level** l **from** L **to** 0
  - **for each translation** $\mathbf{b}$
    - Compute the region where all the constraints in eq. (9) are simultaneously satisfied; in particular, where eq. (8) is satisfied and where $\Delta\sigma(\mathbf{x}+\mathbf{b}) = 0$
  - Merge the new correspondences to the current map T so that the resulting map is *admissible*, i.e. such that $T(T(\mathbf{x})) = \mathbf{x} \quad \forall \mathbf{x} \in \Omega$. Indeed, although the new correspondences and the current map T are admissible on their own, when merging them there may be links with two hops. Since we start from the highest depth level and proceed to the lowest, such multiple hops are not possible and we set them to be self-references.

Once both $\Delta\sigma$ and T have been computed, we can proceed with the restoration of the unblurred image f. In the next section, we call such restoration *defocus inpainting* as we are filling in blurred regions with the corresponding sharp ones.

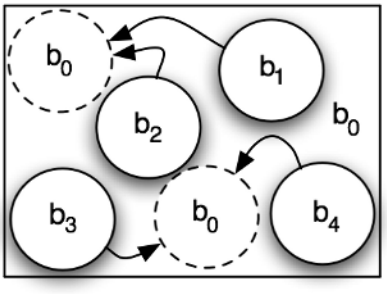


Fig. 3. A partition of the image domain Ω into regions where the relative blur is 0 ($\mathbf{b_0}$) and where it is strictly positive ($\mathbf{b_1}$, $\mathbf{b_2}$, $\mathbf{b_3}$, $\mathbf{b_4}$). Notice that multiple partitions may be in correspondence with the same region.

5 Defocus Inpainting

Image restoration is well-known to be an *ill-posed* problem [5]. To eliminate the ill-posedness, one can introduce *regularization* during the restoration process. Since we pose image restoration as an energy minimization problem, regularization can be added in the form of an additional energy term E_2, so that our solution can be found by minimizing

$$\hat{f} = \arg\min_f \int_\Omega \left(I(\mathbf{y}) - \int K_\sigma(\mathbf{y}, \mathbf{x}) f(\mathbf{x}) d\mathbf{x} \right)^2 d\mathbf{y} + \mu E_2 \tag{12}$$

where μ is a scalar that regulates the amount of regularization. Typically, the term E_2 is a prior that is introduced independently of the input image. For example, one could use a *measure of sharpness* of local patches such as the structure tensor [16].

In our approach instead, we exploit the recurrence of regions as a regularization term. We define E_2 to be

$$E_2 = \int_\Omega \left(I(\mathbf{y}) - \int K_{\Delta\sigma}(\mathbf{y}, \mathbf{z}) \int K_\sigma(\mathbf{z} + \mathbf{b}, \mathbf{x}) f(\mathbf{x}) d\mathbf{x} \right)^2 d\mathbf{y}. \tag{13}$$

Notice that in eq. (12) one has to recover both the depth map s (encoded by the amount of blur σ) and the unblurred image f. Furthermore, such reconstruction is possible only if one knows the camera parameters. In many instances, however, such parameters are not available. In this case, we propose a method to improve the restoration of the input image I, by introducing the following constraints:

$$\begin{array}{ll} f(\mathbf{x}) = I(\mathbf{x}) & \forall \mathbf{x} | T(\mathbf{x}) = \mathbf{x} \\ f(\mathbf{x}) = f(T(\mathbf{x})) & \forall \mathbf{x} | T(\mathbf{x}) \neq \mathbf{x}. \end{array} \tag{14}$$

The two equations above formalize the following procedure: if a region is self-referencing, then no restoration is performed; if a region maps to another region, since such region is sharper by construction of the correspondence map T (see section 3), then the latter one is used in place of the first one. Hence, the regularization term in this case becomes simply:

$$E_2 = \int_\Omega \left(f(\mathbf{x}) - I(T(\mathbf{x})) \right)^2 d\mathbf{x} \tag{15}$$

and the data term in eq. (12) has to be computed on the known relative blur $\Delta\sigma$ rather than σ resulting in

$$\hat{f} = \arg\min_f \int_\Omega \left(I(\mathbf{y}) - \int K_{\Delta\sigma}(\mathbf{y}, \mathbf{x}) f(\mathbf{x}) d\mathbf{x} \right)^2 d\mathbf{y} + \mu \int_\Omega \left(f(\mathbf{x}) - I(T(\mathbf{x})) \right)^2 d\mathbf{x}. \tag{16}$$

The computation of the unknown unblurred image f can then be done by performing a gradient descent on eq. (16) with the following energy gradient:

$$\nabla_f E(\mathbf{x}) = -2 \int_\Omega \left(I(\mathbf{y}) - \int K_{\Delta\sigma}(\mathbf{y}, \mathbf{x}') f(\mathbf{x}') d\mathbf{x}' \right) K_{\Delta\sigma}(\mathbf{y}, \mathbf{x}) d\mathbf{y} + 2\mu \left(f(\mathbf{x}) - I(T(\mathbf{x})) \right). \tag{17}$$

6 Experiments

We test our algorithm on both synthetic and real data. In the case of synthetic data, we show the average performance of the method on 50 experiments with fixed shape and variable texture. In Figure 4 we show one example of the texture that has been employed in the generation of the synthetic images (a), together with the corresponding blurred image (b). In (c) we show the true depth map which can be compared to the recovered

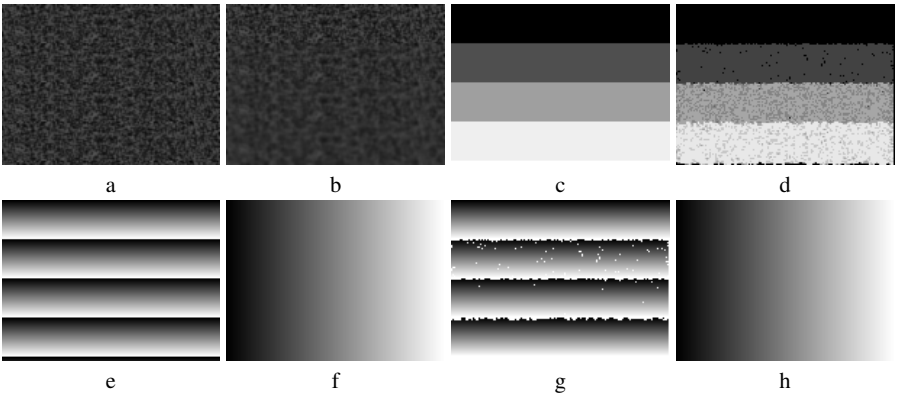


Fig. 4. One example of synthetic defocus inpainting. (a) The true unblurred image. (b) The input image. (c) The true depth map. (d) The recovered depth map. (e) and (f) the true correspondence map T where (e) corresponds to the x coordinates and (f) to the y coordinates. (g) and (h) the recovered correspondence map.

depth map in (d). In (e) and (f) we show the true correspondence map T where (e) corresponds to the x coordinates and (f) to the y coordinates; in (g) and (h) we show the recovered correspondence map. In Figure 5 we show a few snapshots of the restoration of one example (shown in Figure 4). On the leftmost image we show the given blurred image, while on the rightmost image we show the true unblurred texture.

We find that the mean restoration error is of 0.1441 with standard deviation of 0.0116, which, once compared to the error between the input image and the true unblurred image 0.3183 with standard deviation of 0.0177, shows an improvement of more that 2 times.

In the case of real experiments, we run our algorithm on images of the endothelium cell layer, that were acquired from several corneas at the Cornea Bank Berlin using an inverse phase-contrast microscope (CK 40, Olympus Co. Japan) at 100x and 200x magnification, and thus are subject to a substantial amount of defocus. The corneas were kept in hypotonic BSS for a better microscopy visualization of the endothelial

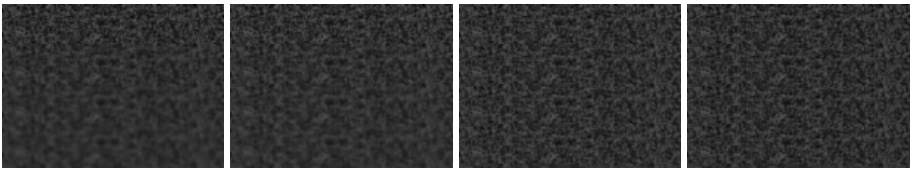


Fig. 5. Snapshots of the restoration process. For comparison, on the leftmost image we show the input image, while on the rightmost image we show the true unblurred image. In the middle we show the evolution of the gradient descent presented in section 5. Iteration time increases going from left to right.

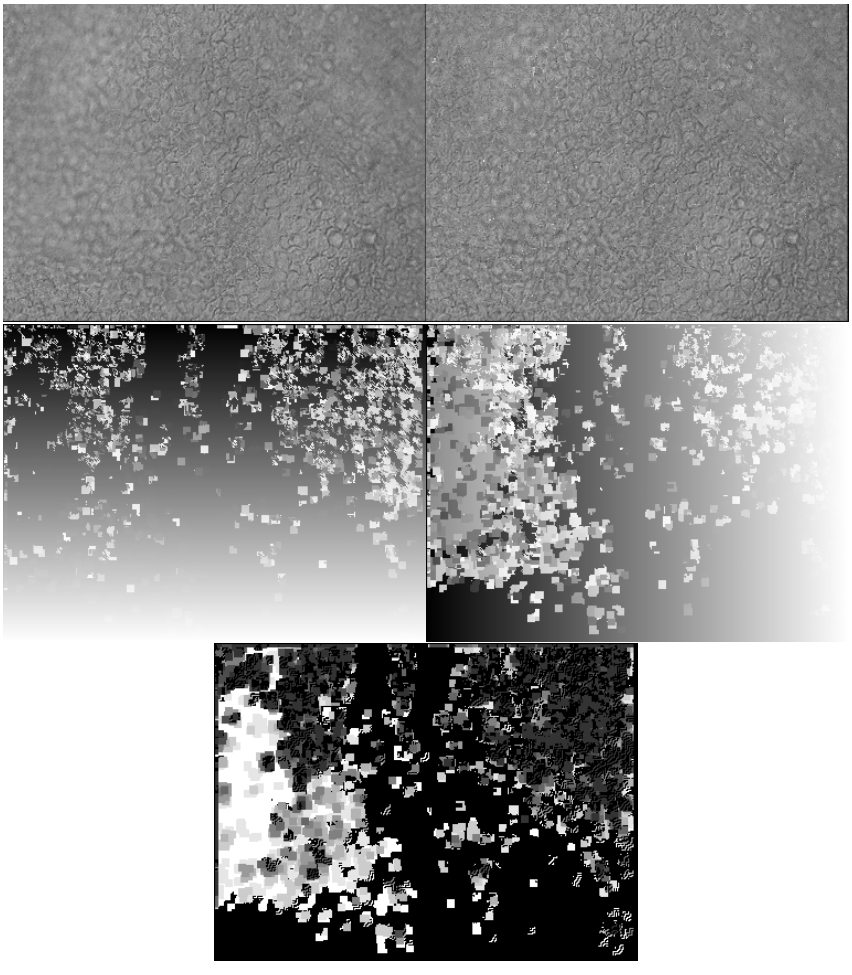


Fig. 6. (top row - left) Image of a cornea. Notice that in some portions of the image the cells are more defocused than in other locations due to the change in depth of the surface. (right) Restored image of the cornea. Notice that cells that were blurred in the original image are now restored and resemble other cells in the same image (data kindly provided by *Fondazione Banca degli Occhi di Venezia. Italy*). (second row) Visualization of the estimated correspondence map. (left) image showing the x coordinates of T. (right) image showing the y coordinates of T. Dark intensities correspond to lower values of the coordinates and vice versa for light intensities. Recall that the map $T(\mathbf{x})$ assigns a sharp patch at $T(\mathbf{x})$ to the blurred patch at $\mathbf{x}$. (bottom row) Visualization of the reconstructed blur map. Light intensities correspond to large amounts of blur, while dark intensities to low amounts of blur.

cells by osmotic stimulation of their cell membranes. In Figure 6 on the top row we show the input image (left) and the restored image (right). Notice that in the input image some location are more blurred than others due to changes in the depth of the cornea. Furthermore, notice that in the restored image most of the blurred cells are

Fig. 7. Examples of defocus inpainting on various images. (left) Input image. (right) Restored image.

restored and resemble the appearance of similar cells that are sharper. In Figure 6, second row, we show the estimated correspondence map T. For ease of visualization, the coordinates of this map are shown as two grayscale images. Notice that the algorithm detects that most sharp patches are located on the right of the input image, and that most of the blurred patches are located on the left. In Figure 6, third row, we show the estimated blur map of the scene. Notice that light intensities correspond to regions that are subject to a large amount of blur, while dark intensities correspond to regions that are subject to small amounts of blur. By visual inspection it is possible to verify that the recovered blur map correctly assigns high values to regions that are blurred in the input image.

In Figure 7 we show a number of examples where the left images are the input images, and the right images are the restored ones.

7 Conclusions

We introduced a novel paradigm for image restoration, where regularization is extracted directly from data. We exploit the assumption that the image contains recurrences of patterns that are similar up to translation and amount of defocus, and show how to model them in the context of defocused images. Then, we propose a novel solution to identify the recurring patterns, to estimate their difference in amount of blur and finally to restore the unblurred image. Our method can also be readily extended to work with multiple images, and we are currently working on handling similarity up to scale and rotations in addition to translations.

References

1. Baker, S., Kanade, T.: Limits on super-resolution and how to break them. In: IEEE Transactions on Pattern Analysis and Machine Intelligence. Volume 24. (September, 2002) 1167–83
2. Criminisi, A., Perez, P., Toyama, K.: Object removal by exemplar-based inpainting. In: CVPR03. (2003) II: 721–728
3. Katsaggelos, A.: Digital Image Restoration (Book). Springer (1991)
4. Yitzhaky, Y., Milberg, R., Yohaev, S., Kopeika, N.S.: Comparison of direct blind deconvolution methods for motion-blurred images. In: Applied Optics-IP. Volume 38. (July 1999) 4325–32
5. Bertero, M., Boccacci, P.: Introduction to inverse problems in imaging. Institute of Physics Publishing, Bristol and Philadelphia (1998)
6. Engl, H., Hanke, M., Neubauer, A.: Regularization of Inverse Problems. Kluwer Academic Publishers, Dordrecht (1996)
7. You, Y., Kaveh, M.: Blind image restoration by anisotropic diffusion. IEEE Trans. on Image Processing **8** (1999) 396–407
8. Aloimonos, Y., Swain, M.: Shape from texture. BioCyber **58** (1988) 345–360
9. Blostein, D., Ahuja, N.: Shape from texture: Integrating texture-element extraction and surface estimation. PAMI **11** (1989) 1233–1251
10. Ens, J., Lawrence, P.: An investigation of methods for determining depth from focus. IEEE Trans. Pattern Anal. Mach. Intell. **15** (1993) 97–108
11. Pentland, A.: A new sense for depth of field. IEEE Trans. Pattern Anal. Mach. Intell. **9** (1987) 523–531
12. Subbarao, M., Surya, G.: Depth from defocus: a spatial domain approach. Intl. J. of Computer Vision **13** (1994) 271–294
13. Watanabe, M., Nayar, S.: Rational filters for passive depth from defocus. Intl. J. of Comp. Vision **27** (1998) 203–225
14. Chaudhuri, S., Rajagopalan, A.: Depth from defocus: a real aperture imaging approach. Springer Verlag (1999)
15. Born, M., Wolf, E.: Principle of optics. Pergamon Press (1980)
16. Weickert, J.: Anisotropic Diffusion in Image Processing. B.G.Teubner Stuttgart (1998)