

# Robust Shape from Focus via Markov Random Fields

Victor Gaganov\*, Alexey Ignatenko\*\*

\*Keldysh Institute of Applied Mathematics RAS, Moscow, Russia

\*\*Lomonosov Moscow State University, Moscow, Russia

{vgaganov, ignatenko}@graphics.cs.msu.ru

## Abstract

In this paper we study a problem of 3D scene reconstruction from a set of differently focused images, also known as the *shape from focus* (SFF) problem. Existing shape from focus methods are known to produce unstable depth estimates in areas with poor texture and in presence of strong highlights. So in this work we focus on the robustness of 3D scene structure recovery. We formulate a shape from focus problem in a Bayesian framework using Markov Random Fields and present an SFF method that yields a globally optimal surface with enforced smoothness priors. Although shape from focus has been studied for quite a long time there is no widely accepted test set for evaluation of SFF algorithms. Therefore we present a test set composed of 27 image sets with hand-labeled ground truth. We quantitatively evaluate our method on this test set and present the comparison results. These results demonstrate that our method is robust to highlights and untextured regions and that it outperforms the state-of-the-art.

**Keywords:** *computer vision, 3D reconstruction, shape from focus, Markov Random Fields, MRF, energy minimization.*

## 1. INTRODUCTION

Recovering the 3D structure of the scene from images is one of key challenges in computer vision. There are many approaches to this problem, each exploiting different image cues. Shape from focus and shape from defocus (SFD) are two such approaches that exploit image focus to estimate 3D geometry of the scene. SFD approach [3,4] attempts to extract scene depth information measuring the relative blurriness of scene images, taken with different focus settings. SFF approach [1,2] scans the scene, taking a sequence of images with different lens focus settings. Instead of varying the lens settings one can gradually move the camera along the viewing direction, taking images. After that for each scene point a best-focused image is found. For a sharp feature on an image the position of this feature in 3D space can be determined, and the uncertainty of an estimate is, by definition, the depth of field.

Shape from defocus can obtain scene depth estimates given as few as two images of the scene. The reconstruction can be done fast and an SFD approach can even be applied to dynamic scenes. In the same time SFF approach needs a fairly large number of images to obtain reasonable reconstruction of the scene geometry and hence SFF is limited to static scenes. However there are still many applications that may benefit from a reliable SFF method. Robust and accurate SFF algorithm is especially desired in such fields as mineralogy [9] and industrial inspection of small-scale objects [16]. In these applications a perfectly static object is observed using an optical microscope and the goal is to obtain a 3D model of this object. Also, for a general static scene shape from focus has greater applicability than shape from defocus since

it makes mild assumptions about the scene and the image formation process. While SFD methods have to explicitly model the camera blurring process, the only assumption made by the SFF algorithms is that some quantitative measure of blur is minimized at the position of best focus. For these reasons we have chosen to investigate and improve shape from focus algorithms.

Although there are many different SFF methods [1,2,5-9] almost all of them are essentially local. For each pixel of each image these methods calculate a sharpness measure that is used to determine the quality of image focus. To reduce negative effects introduced by camera noise the values of sharpness measure for each pixel are averaged over a local window centered in this pixel. This local window is called the evaluation window and the averaging process itself is called aggregation. After the aggregation for each pixel SFF methods search for an image that has the largest value of sharpness measure. Clearly, for objects that have textureless regions on their surface such a procedure will produce unstable results. Local SFF algorithms also fail to produce reliable depth estimates when imaging conditions are imperfect. In the event that some parts of the image are underexposed or, to the contrary, contain a strong highlight local SFF methods would fail to produce reasonable depth estimates. One way to handle this is to use aggregation with larger evaluation windows. However using large evaluation windows results in coarse reconstruction and fine details of the scene surface may be lost. Also, in general, size of textureless regions can be beyond the reasonable size of an evaluation window.

In this paper we present a new SFF method that is robust and can produce stable depth estimates even in textureless regions. Our method is based on a Markov Random Fields (MRFs) theory. MRFs have a long history of use in computer vision and have proven to be a useful tool for many problems, such as image segmentation [10] and stereo correspondence [11]. We cast a shape from focus problem in a Bayesian framework using MRFs. In contrast to existing approaches, this yields a global SFF method. In our method the 3D model of the scene is found as a minimum of MRF energy function. This energy function combines sharpness cues taken from images and smoothness priors on reconstructed surface shape. Enforcing smoothness priors allows our algorithm to produce stable results in poorly exposed, textureless and highlighted regions. To verify the stability and robustness of the proposed method we have conducted experiments on the real world data. We have collected a diverse test set for SFF methods evaluation. Our test set is composed of images of different mineralogical samples. For each set of images in our test set we have hand-labeled the ground truth. We have quantitatively evaluated our method on this test set and the results show, that our method outperforms existing SFF methods in terms of robustness.

The rest of the paper is organized as follows. In Section 2 we introduce some notation and overview the related work. We

present our new robust shape from focus method in Section 3. In Section 4 we describe our test set for SFF methods evaluation and present experimental results. Finally, we provide concluding remarks in Section 5.

## 2. RELATED WORK

### 2.1 Image formation geometry

We begin with the essential concepts of optics that the SFF algorithms are based upon. Figure 1(a) shows a simple model of the image formation process – the thin lens model. All the light rays that are radiated by the scene point P and are refracted by the lens converge in the point Q. The relationship between object distance  $p$ , image distance  $q$  and focal length  $f$  of the lens is given by the thin lens law:

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{f} \quad (1)$$

If the point Q coincides with the camera sensor, then an image of scene point P would be perfectly focused on a photograph. However if for some scene point P' corresponding convergence point Q' is not on a sensor plane the energy received by the sensor from P' would be distributed over a patch on a sensor plane, resulting in image blur (see Figure 1(b)). The blur radius increases with the distance between Q' and a sensor plane.

The depth of field of an optical system is a part of 3D space that can be seen in-focus on a photograph. Size of the depth of field is a limiting factor for all shape from focus algorithms. However in practice the size of the depth of field can be decreased by using long focal lengths and large aperture sizes. For an optical microscope the depth of field size can be as small as several micrometers.

### 2.2 Shape From Focus

Shape from focus algorithms extract depth information from a set of differently focused images. We will denote sets of images of the same size by capital letters with hat  $\hat{I} = \{I_1, \dots, I_N\}$ . An

images set  $\hat{I}$  can be obtained either by gradually varying the focal settings of the camera, or by moving the camera along the viewing direction. SFF algorithms assume that for each image the internal parameters of an optical system and the camera position are known. Given this information for a sharp feature of an image its position in 3D space can be estimated. Only the depth of field of an optical system limits precision of such an estimate. Since all the images in the set  $\hat{I}$  are taken with the same camera orientation the 3D model of the scene can be represented as a depth map. Size of the depth map  $D = \{d(x, y)\}, d(x, y) \in [0, N-1]$  coincides with size of images in set  $\hat{I}$ .

Existing SFF algorithms do not operate directly on images from set  $\hat{I}$ . First, a sharpness measure function is computed for each image in the set, resulting in N sharpness images  $\hat{S} = \{S_1, \dots, S_N\}$ . Sharpness measure function should produce high response to image regions that contain high frequencies since focused images usually have more high frequencies than defocused images. The most popular sharpness measures for SFF algorithms

are the Tenengrad operator [6] and the Modified Laplacian (ML) that was proposed by Nayar and Nakagawa [2]. In this work we have used ML operator that is a discrete version of the following function

$$ML(x, y) = \left| \frac{\partial^2 I(x, y)}{\partial x^2} \right| + \left| \frac{\partial^2 I(x, y)}{\partial y^2} \right| \quad (2)$$

Raw values of sharpness measure are extremely sensitive to camera noise and for this reason existing SFF algorithms perform sharpness measure aggregation. In an SFF method proposed by Nayar [2] sharpness measure values are averaged over a square evaluation window centered in a target pixel

$$SML(x_0, y_0) = \sum_{(x, y) \in \Omega(x_0, y_0)} ML(x, y) \quad (3)$$

After averaging the resulting depth map is obtained by a winner-takes-all (WTA) strategy. For each pixel an image with a maximum value of averaged sharpness measure is selected

$$D(x, y) = \arg \max_i (SML_i(x, y)) \quad (4)$$

There are many different modifications of this basic scheme [1,7,8]. In [1] authors propose to refine depth estimates given by (4) finding a local plane with maximum average sharpness. This improves accuracy in areas where scene surface is not parallel to image plane.

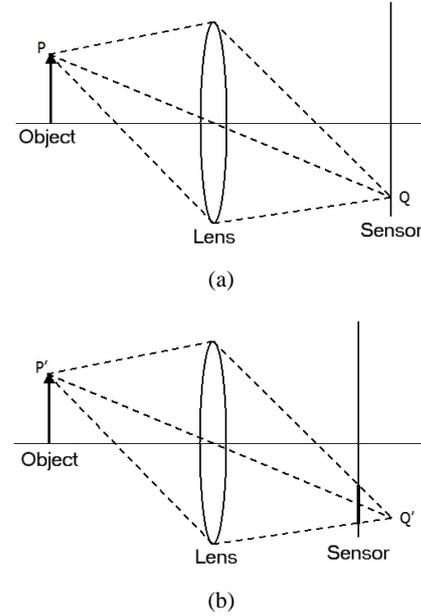
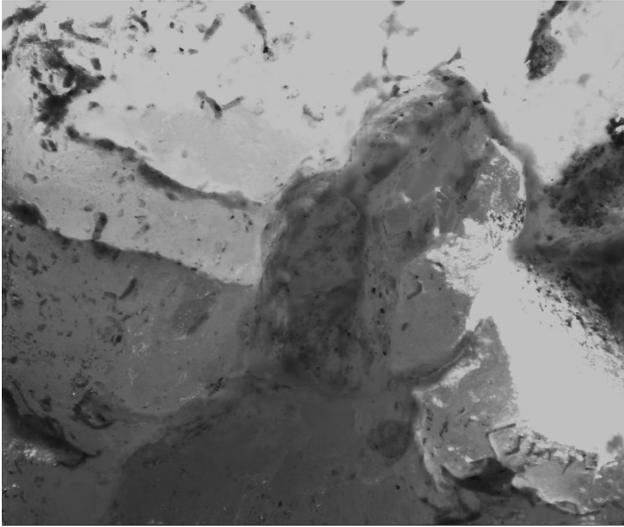


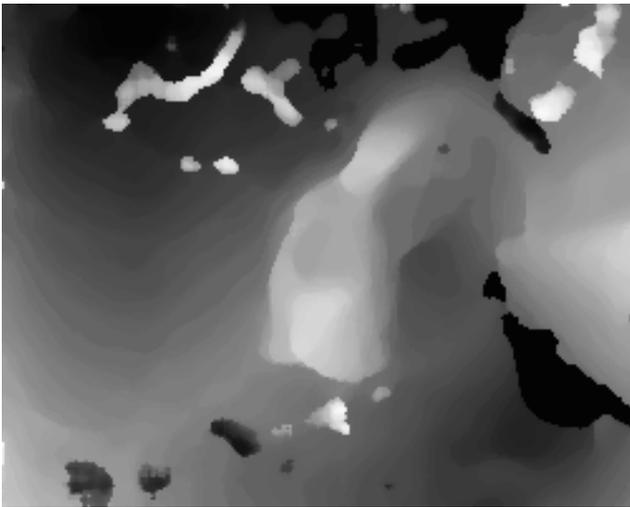
Figure 1. Formation of focused (a) and defocused (b) images

In [7] Ahmad and Choi propose to use dynamic programming to refine results, obtained by (4). This method is less computationally expensive than a brute force approach of [1]. Authors of [8] propose an adaptive averaging scheme that resembles bilateral filtering. This method is computationally efficient and allows obtaining more accurate estimates near sharp depth discontinuities. However this method utilizes an image of the scene taken with large depth of field. Large depth of field image cannot be obtained for strong magnifications of an optical microscope due to physical limitations.

The major drawback of all mentioned SFF algorithms is that they are essentially local. The WTA strategy that searches for a maximum value of sharpness measure fails to produce reasonable results in textureless and underexposed areas and in highlighted regions. In case a textureless region is not very large this can be handled by increasing the evaluation window size. However in general it is impossible to make an evaluation window bigger than the largest textureless region in the scene. Figure 2 shows a typical result of an SFF algorithm on a complex real-world example. As it can be seen the results are unstable in regions without texture. Surprisingly few works on SFF address this issue. Nair and Stewart [6] propose to reject depth estimates in areas where the sharpness measure is below a certain threshold. This results in a sparse but more reliable depth map. Authors of work [9] propose to apply median filter to the resulting depth map to suppress gross errors in reconstruction.



(a)



(b)

Figure 2. Results of a local SFF algorithm of a real world example. Depth map (b) is quite unstable in textureless regions. Multifocus image for this example can be seen in Figure 2(a)

However these simple countermeasures often fail in real-world applications, as it will be shown in our experiments section. To produce stable results in untextured, highlighted and underexposed regions an SFF algorithm should be based on global reasoning. Reliable SFF algorithm should take into account global constraints, along with local sharpness information.

### 3. PROPOSED METHOD

Our shape from focus method is based on Markov Random Fields theory. MRFs have proven to be an extremely useful tool for many computer vision problems such as stereo correspondence [11] segmentation [10] and image restoration [17]. In this section we show how an SFF problem can be formulated in MRF framework. In this framework the resulting depth map can be found as a maximum a posteriori (MAP) estimate of hidden state of an MRF. This yields an SFF algorithm based on global energy minimization that combines surface shape priors with observed sharpness information from images.

#### 3.1 MRF Formulation

In SFF problem the goal is to estimate the hidden state  $D = \{d(x, y)\}$  given the observed data  $\hat{I}$ . Like other SFF algorithms we would operate on sharpness measure, rather than on original images. Hence the Bayes rule gives us  $P(D | \hat{S}) = P(\hat{S} | D)P(D) / P(\hat{S})$ . To reduce clutter in notation we will write a matrix  $D = \{d(x, y)\}$  as a one-dimensional vector  $d$  and  $S_i$  as vector  $s_i$ . Thus a MAP estimate of  $d$  can be obtained as follows

$$d_{MAP} = \arg \min_d (-\ln(P(\hat{S} | d)) - \ln(P(d))) \quad (5)$$

First term in (5) is a likelihood of observing the data given a certain hidden state  $d$  and the second term is a prior probability of  $d$ . We factorize the likelihood  $P(\hat{S} | d) = \prod_i P(\bar{s}_i | d_i)$

where  $\bar{s}_i = \{s_1(i), \dots, s_N(i)\}$ . This corresponds to an assumption that values of sharpness measure in certain pixel are independent of depth values in all other pixels. We also assume that  $d$  is a Markov Random Field and we assume that only neighboring pixels are statistically dependent. In this case prior in (5) can be factorized using the Hammersley-Clifford theorem and minimum of (5) can be found as (see [12] for detailed derivation)

$$E(d) = \sum_i (V_i(d_i) + I \sum_{j \in N(i)} U_{\{i,j\}}(d_i, d_j)) \quad (6)$$

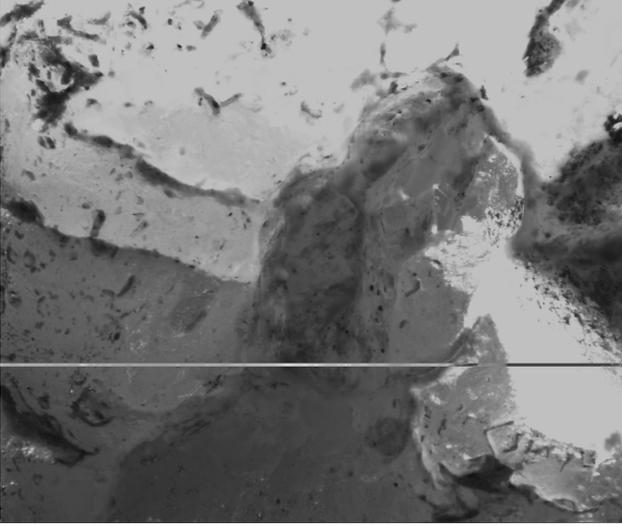
where  $N(i)$  defines the 4-connected neighborhood of element  $i$ . Function  $E(d)$  is called the MRF energy function. This function consists of sums of terms of two types. Terms of first type depend on single element of depth map and are called data terms. Terms of second type depend on two neighboring elements of depth map and are called the smoothness terms. Defining data and smoothness terms appropriately and optimizing energy function in (6) we will obtain  $d_{MAP}$  that is a MAP estimate of the scene depth map.

### 3.2 Data term

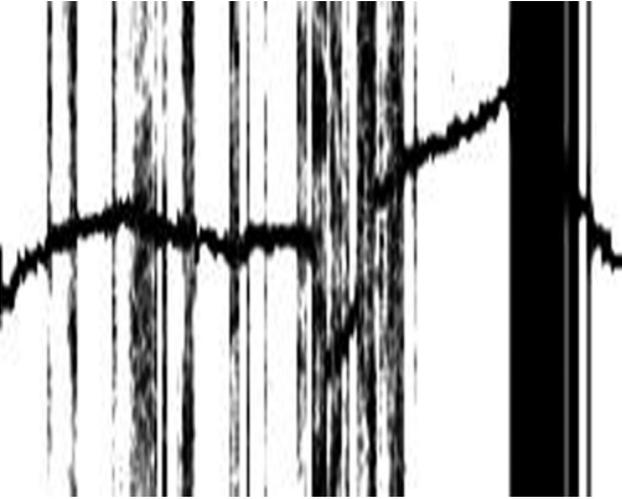
Data term in energy (6) encodes likelihood of observing particular  $\bar{s}_i$  given different values of  $d_i$ . We propose to define the data term as follows

$$V_i(d_i) = \min((\max(\bar{s}_i) - s_i(d_i))^2, T_d) \quad (7)$$

The data term (6) penalizes significant deviations from a maximum value of sharpness measure.



(a)



(b)

Figure 3. Illustration of our data term behavior. Figure 3(a) shows a multifocus image of the set and a position of data term slice.

Figure 3(b) shows a slice of our data term. The X axis corresponds to the X axis of images in set, Y axis corresponds to a number of image in set

It can be seen that our data term is minimized by the depth value that corresponds to an image with maximum sharpness. Hence our data term gives preference to depth maps that pass through sharp areas of  $\hat{S}$ . In case scene area that corresponds to  $d_i$  contains no significant texture all the variations in sharpness measure profile

$\bar{s}_i$  will be due to camera noise. Hence our data term will give no significant preference to any particular value of  $d_i$ . In such areas the surface will be primarily shaped by the smoothness term. However if there is strong texture sharpness profile maximum will be significant and the data term will give clear preference to depth values that correspond to sharp photographs. The behavior of our data term is illustrated in Figure 3. The data term is bounded by  $T_d$  to make the method less sensitive to outlier maximums in sharpness profile.

We would like to notice that unlike the existing SFF algorithms we omit the sharpness measure aggregation. This is motivated by the fact that (6) that is minimized by our method contains a sum of data terms for all elements of depth map  $d$ . Hence our algorithm implicitly averages the sharpness measure over the entire surface of the reconstructed scene. This choice is verified by our experiments.

### 3.3 Smoothness term

Smoothness term in (6) is responsible for regularization of the depth map estimate. In scene parts where sharpness measure doesn't provide reliable depth information smoothness prior shapes the depth map to make the results stable. We derive our smoothness term from truncated L2 norm

$$U_{\{i,j\}}(d_i, d_j) = \min((d_i - d_j)^2, T_s) \quad (8)$$

We prefer L2 norm rather than L1 norm or Generalized Potts model (GPM) [12] since it is a better prior for slanted surfaces. Truncation with  $T_s$  is necessary to preserve discontinuities. We propose the following smoothness term based on (8)

$$U_{\{i,j\}}(d_i, d_j) = \begin{cases} \min((d_i - d_j)^2, T_s), w_{ij} = 1 \\ (d_i - d_j)^2, w_{ij} = 0 \end{cases} \quad (9)$$

It can be seen that our smoothness term is spatially varying. Binary variable  $w_{ij}$  is an indicator of significant texture presence. In case significant texture is present around  $d_i$  we truncate the L2 norm with  $T_s$  thus allowing sharp jumps in depth map. Otherwise we use non-truncated L2 norm. As a result sharp jumps are not allowed in areas where significant texture isn't present.

We propose to calculate  $w_{ij}$  as follows. First we compose a maximum sharpness image

$$H(x, y) = \max(S_1(x, y), \dots, S_N(x, y)) \quad (10)$$

After that we threshold  $H(x, y)$  and obtain a binary mask of pixels with maximum sharpness greater than  $T_H$ . This mask is then processed with mathematical morphology. We apply *opening* with a circular structuring element of radius R to obtain the resulting significant texture presence image. After that we define  $w_{ij}$  to be equal to one in case there is significant texture in both pixels that correspond to  $d_i$  and  $d_j$ .

### 3.4 Energy optimization

Optimization of MRF energies is an area of active research and there are many different methods that can minimize energy (6). These methods include belief propagation, tree-reweighted message passing (TRW),  $\mathbf{a}$ -expansion and  $\mathbf{a}, \mathbf{b}$ -swap [13]. In this work we have chosen to use  $\mathbf{a}$ -expansion. This choice is motivated by the comparison of different MRF energy optimization methods presented in [13]. According to this comparison  $\mathbf{a}$ -expansion is the fastest energy optimization algorithm and it produces solutions with very low energy. Although the smoothness term that we have proposed in section 3.3 is non-submodular [18]  $\mathbf{a}$ -expansion still can be applied to optimize energy (6) as it was shown in [14].

## 4. EXPERIMENTS

To validate the proposed method and compare it with the existing methods we have implemented it in C++ and tested on a set of real-world example. Unfortunately there is no widely accepted test set for comparison of SFF algorithms. For this reason we have created a test set with data from our application. Inspired by the work of Scharstein and Szeliski on stereo correspondence algorithms evaluation [15] we have hand-labeled the ground truth for our data to evaluate algorithms quantitatively.

### 4.1 Test images for evaluation

For evaluation of our SFF algorithm we have composed a real-world test set with 27 different image sets. All images in our test set were taken with Leica Z16 APOA optical microscope equipped with a gray-scale digital camera. For our experiments we used 1280x1024 resolution. All images were taken at maximum magnification of the microscope with the depth of field  $\sim 30\mu\text{m}$ . Each of our 27 test sets contains from 50 up to 200 images. Leica Z16 is motorized and sets of differently focused images have been obtained moving the optical system of the microscope along the viewing direction automatically. Step between two consecutive images for each test set was one-half of the depth of field. Our test set is composed of images of different minerals. Our major application was reconstruction of cavities on the surface of the minerals. Our data set is divers and includes image sets with a lot of texture along with image sets with many untextured regions (see Figure 4).

To evaluate our algorithm quantitatively we have hand-labeled the ground truth for all our 27 test sets. We have implemented a tool that allows browsing a set of differently focused images and placing control points in these images. Control points have to be placed in areas where the sharp features are present. After a number of control points have been placed a ground truth depth map can be obtained via interpolation. For textureless regions near the image border interpolations sometimes fails to produce reasonable depth estimates. Such regions have been hand-labeled and excluded from evaluation. To obtain a high quality depth map lots of control points a required and described procedure becomes very time-consuming. For this reason in the future we are planning to obtain ground truth using an X-ray scanner. Figure 4 demonstrates a number of ground truth depth maps labeled with our tool along with multifocus images of corresponding image sets.

### 4.2 Comparison results

To evaluate our algorithm we have implemented it in C++. We have also implemented an algorithm presented by Nayar with 10x10 evaluation window [2]. We have implemented this algorithm rather than its modifications presented in [1,7] since all these modifications use results of Nayar algorithm as an initial estimate. These algorithms refine the initial estimate and if there is a gross error in the initial estimate the refinement result will also be erroneous. We could not implement methods presented in [6,8] for our data, since these algorithms rely on an image of the scene taken with large depth of field. For strong magnifications of a microscope it is impossible to obtain high quality image with a large depth of field due to physical limitations. However for our comparison we have implemented a modified algorithm of Nayar with rejection of unreliable depth estimates described in work on Nair and Stewart [6] followed by linear interpolation. The results of this algorithm have been post-processed with a median filter to suppress gross errors, as proposed by Niederoest et al in [9]. We denote this algorithm as 3N (Nayar, Nair, Niederoest).

To compare the algorithms quantitatively we have used error metric presented in [15]

$$Q(\hat{d}, \tilde{d}) = \frac{\sum_{i=1}^m q(\hat{d}_i - \tilde{d}_i, T_e)}{m}, \quad q(x, t) = \begin{cases} 1, & |x| > t \\ 0, & |x| \leq t \end{cases} \quad (11)$$

where  $m$  is the number of depth map elements. This metric measures the share of depth map elements with estimation error greater than a specified threshold. In our experiments  $T_e$  was twice as large as the depth field. We have selected such a large  $T_e$  to detect only gross errors since our major concern was robustness of the algorithms. Parameters of 3N and our algorithm have been fine tuned to produce optimal results on our test set. We used the same values of parameters for all image sets in our test base. Results of quantitative evaluation are presented in Table 1.

Method name	Significant errors share ( $Q$ )
Nayar	15.7%
3N	6.4%
Proposed method	3.6%

Table 1. Quantitative comparison of proposed method with existing methods

It is clear that in term of error metric (11) our algorithm produces the best results. For visual comparison in Figure 4 we present the results for 3 different image sets from our test base. It can be seen that basic algorithm of Nayar is extremely unstable in textureless regions. This perception is confirmed by our quantitative evaluation. According to it as much as 15% of depth estimate given by method of Nayar are unreliable. Results of 3N algorithm are much more stable than the results of Nayar algorithm. For the second image set (Figure 4(k)) 3N algorithm produced near perfect result. It's worse noticing that the best results for 3N algorithm were obtained with the median filter radius of approximately 50 pixels. This suggests that an algorithm tends to incorporate as much context information as possible to produce stable results. This verifies our idea that a stable SFF algorithm

should be based on global reasoning. However for large untextured regions the 3N algorithm fails to produce reliable depth estimates (Figure 4(j,l)). In contrast, the proposed method handles textureless regions of arbitrary size correctly. It can be seen that the results of our algorithm are smooth and consistent although we omit the sharpness measure aggregation. Optimal results for our algorithm have been obtained with  $T_s = 400$  and the resulting depth maps tend to be smooth almost everywhere. Sometimes this results in oversmoothing, as it can be seen from Figure 4(m). We are planning to address this issue in our future work. However on the average our algorithm produces the most reliable result among the compared methods.

## 5. CONCLUSION AND FUTURE WORK

In this work we have presented a new shape from focus algorithm that is based on Markov Random Fields theory. In contrast to existing SFF methods our algorithm yields globally optimal depth maps with enforced smoothness constraints. Presented experiments verify that our algorithm is able to produce stable depth estimates in untextured, underexposed and highlighted regions, where existing SFF methods fail to produce reliable results.

To evaluate our algorithm we have used the same values of all parameters for all image sets in our test base. However we have noticed that for some image sets this results in suboptimal depth estimates. So in the future we are planning to estimate algorithm parameters, along with the depth map in the unified framework. As it was noticed in the experiments section our algorithm sometimes produces results that are oversmoothed. To handle this realistic priors on sharp discontinuities location should be incorporated in the algorithm. This will allow our method to produce globally optimal depth maps without oversmoothing.

## 6. REFERENCES

- [1] Subbarao M., Choi T. *Accurate recovery of three-dimensional shape from image focus*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 1995, v. 17, n. 3, pp. 266–274.
- [2] Nayar S. K., Y. Nakagawa. *Shape from focus*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994, v.16, n.8, pp. 824-831.
- [3] Jin H., Favaro P. *A variational approach to shape from defocus*, Proc. IEEE European Conference on Computer Vision, 2002, pp. 18-30.
- [4] Subbarao M., Surya G. *Depth from defocus by changing camera aperture: A spatial domain approach*, Proc. Computer Vision and Pattern Recognition, 1993, pp. 61–67.
- [5] Grossmann P. *Depth from focus*, Pattern Recognition Letters, 1987, v.5 n.1, p.63-69
- [6] Nair H.N., and Stewart C.V. *Robust focus ranging*, Proc. Computer Vision and Pattern Recognition, 1992, pp. 309–314.
- [7] Ahmad, M.B., Tae-Sun Choi. *A heuristic approach for finding best focused shape*, IEEE Transactions on Circuits and Systems for Video Technology, 2005, v.15, n.4, pp. 566–574.
- [8] Aydin T., Akgul Y.S., *A New Adaptive Focus Measure for Shape From Focus* Proc. BMVC 2008
- [9] Niederoest M., Niederoest J., and Scucky J. *Automatic 3D Reconstruction and Visualization of Microscopic Objects from a Monoscopic Multifocus Image Sequence*. International Workshop on Visualization and Animation of Reality based 3D Models, 2002
- [10] Boykov Y., Jolly M.P. *Interactive Graph Cuts for Optimal Boundary and Region Segmentation of Objects in N-D Images*. Proc. International Conference On Computer Vision, 2001, pp. 105-112
- [11] Boykov Y., Veksler O., Zabih R. *Fast Approximate Energy Minimization via Graph Cuts*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, v.23, n.11, pp. 1222-1239
- [12] Boykov Y., Veksler O., Zabih R. *Markov Random Fields with Efficient Approximations*. Proc. Computer Vision and Pattern Recognition, 1998, pp. 648-655
- [13] Szeliski R., et al. *A comparative study of energy minimization methods for markov random fields*, In Proc. European Conference on Computer Vision, 2006, v.20, n.6, pp. 1068-1080
- [14] Rother, C., Kumar, S., Kolmogorov, V., Blake, A. *Digital tapestry*, Proc. Computer Vision and Pattern Recognition, 2005, pp. 589-596
- [15] Scharstein D., Szeliski R. *A taxonomy and evaluation of dense two-frame stereo correspondence algorithms*, International Journal of Computer Vision, 2002, v. 47, pp. 7-42.
- [16] Schaper D. *Automated quality control for micro-technology components using a depth from focus approach*, Image Analysis and Interpretation, 2002, pp. 50-54
- [17] Geman S., Geman D. *Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1984, v.6 pp. 721-741.
- [18] Kolmogorov V., Zabih R. *What Energy Functions Can Be Minimized via Graph Cuts?* IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, v. 26, pp. 147-159

## About the authors

Victor Gaganov is a PhD student at Keldysh Institute of Applied Mathematics of Russian Academy of Science. His research interests include computer vision, 3D reconstruction, machine learning and adjacent fields. His contact e-mail is [vgaganov@graphics.cs.msu.ru](mailto:vgaganov@graphics.cs.msu.ru).

Alexey Ignatenko is a researcher at Computational Mathematics and Cybernetics department of Moscow State University. His research interests include photorealistic 3D rendering, 3D modeling and reconstruction, image-based rendering and adjacent fields. His contact e-mail is [ignatenko@graphics.cs.msu.ru](mailto:ignatenko@graphics.cs.msu.ru).

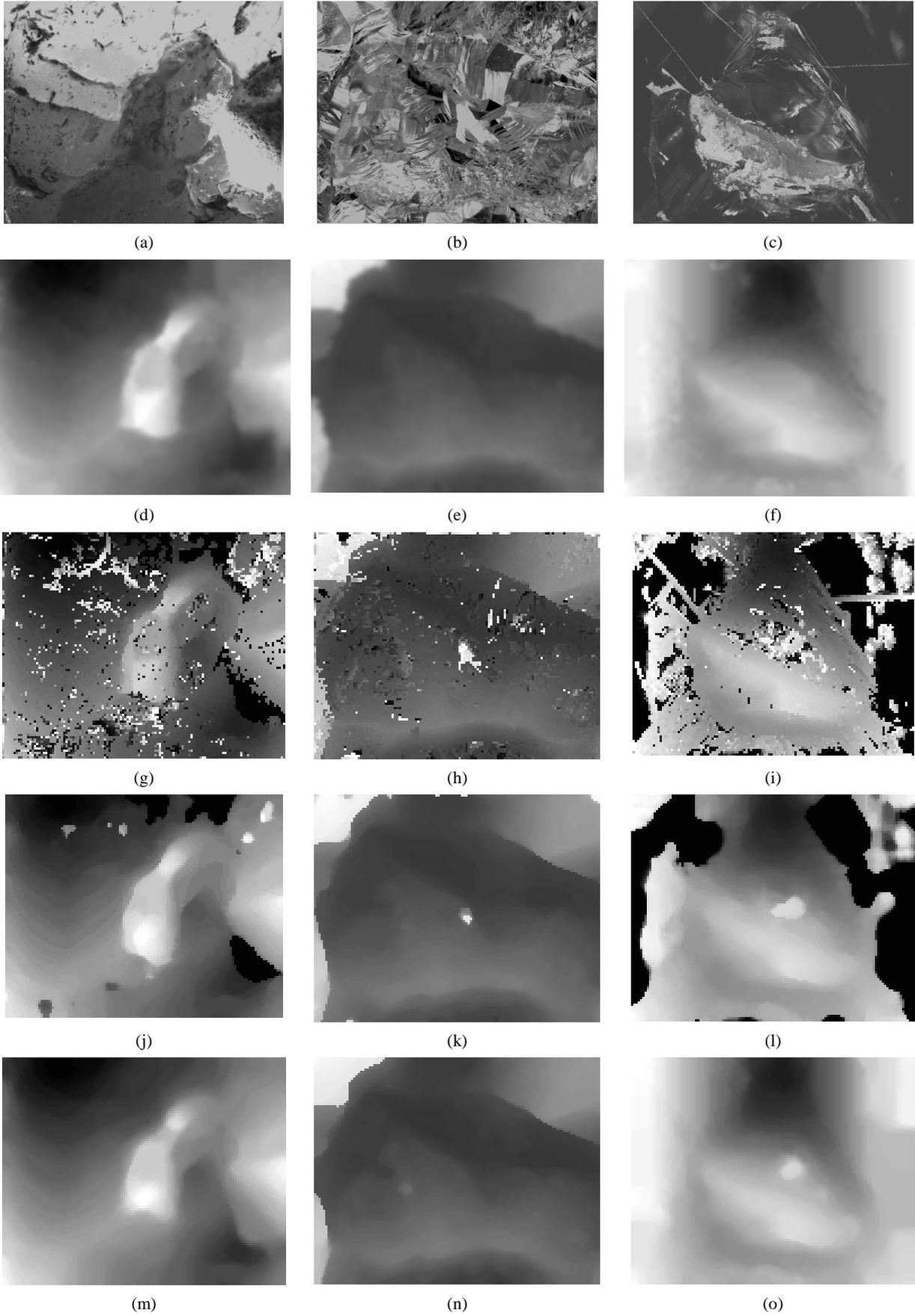


Figure 4. Comparison results for three examples from our test set. First row contains multifocus images for these three sets. Second row is ground truth. The rest three rows are the results of Nayar algorithm, 3N algorithm and the proposed method respectively