A BAYESIAN APPROACH TO SHAPE FROM CODED APERTURE

Manuel Martinello, Tom E. Bishop, Paolo Favaro

School of EPS, Heriot-Watt University, Edinburgh, EH14 4AS, United Kingdom

ABSTRACT

In this paper we present analysis and a novel algorithm to estimate depth from a single image captured by a coded aperture camera. This is a challenging problem which requires new tools and investigations, compared with multi-view reconstruction. Unlike previous approaches, which need to recover *both* sharp image and depth, we consider directly estimating *only* depth, whilst still accounting for the statistics of the sharp image. The problem is formulated in a Bayesian framework, which enables us to reduce the estimation of the original sharp image to the local space-varying statistics of the texture. This yields an algorithm that can be solved via graph cuts (without user interaction). Performance and results on both synthetic and real data are reported and compared with previous methods.

Index Terms— Coded aperture, depth estimation, single image, Bayesian methods.

1. INTRODUCTION

Shape estimation is the enabling step to performing tasks such as autonomous navigation, object recognition, and human-machine interaction. Typically, due to the difficulty of solving this problem, one considers multiple images as input. In this paper however, we are interested in investigating depth estimation when only one image is provided. The main advantage being that less data is transmitted, and the synchronization and calibration of multiple imaging devices is not needed.

The single image we use is captured by using a coded aperture camera. This technique has a long history starting from the pioneering work of Ables [1] and Dicke [4] in X-ray telescopes, and has been investigated in several research fields. The key idea is that each element in the binary mask either blocks or allows light to go through the main lens and generate a *view* of the scene on the sensor. Due to the additive nature of light, such views are then averaged together in a *coded* image (see Fig. 1). In this paper we address the task of recovering the depth of the scene from a single coded image, which we call *shape from coded aperture*. In its most general form, solving this task involves both recovering the depth map of a scene as well as restoring its sharp image. However, we show how this task can be formulated in an equivalent problem that depends only on the depth of the scene, whilst still taking into account uncertainties in the values of the sharp image.

This work has been supported by EPSRC grant EP/F023073/1(P).

Fig. 1. Left: Image captured using a coded aperture camera. Right: Depth map recovered by our algorithm given the single image at the left as input (white indicates closer objects).

1.1. Prior Work

Shape from coded aperture is, in general, an *inverse problem* and as such it can be formulated as blind deconvolution [12, 10], thus exhibiting all the symptoms of this class of problems. There are several passive methods geared specifically for estimating depth from a single lens single exposure image. For instance, [9] introduces a system for single-image passive ranging based on a wavefront-coded aperture and analyses its performance via information-theoretic principles, similarly to [5]. More recent work [8] investigates 3D point spread functions (PSF) whose transverse cross sections rotate as a result of diffraction, and shows that such PSFs yield an order of magnitude increase in the sensitivity with respect to depth variations. The main drawback however, is that the depth range and resolution is limited due to the angular resolution of the reconstructed PSF.

Coded aperture photography has been recently introduced in computer vision and computer graphics by Raskar *et al.*, who formalized the idea of coding images both in time [14] and space [15], and by Levin *et al.* who first proposed Bayesian priors to address the problem [12]. In these methods both the depth map and the sharp image are reconstructed from a single coded aperture image. However, the main difference between these methods and our approach is that while they estimate the depth and the sharp image in two separate steps and require post-processing, we explicitly avoid the estimation of the sharp image by marginalizing it. This strategy allows one to limit the computational cost. Depth estimation can also be obtained by employing alternative single camera systems, such as the multiplexed coded aperture [13], differ-

ential masking [6], and by changing camera settings [3, 7]. However all these methods are based on multiple images.

1.2. Contributions

We apply a sound Bayesian analysis to the problem. By marginalising the unknown texture we concentrate on estimating depth; however the statistical variation of the texture is still taken into account. Also we show how filtering the input can simplify the computation of the algorithm. This produces a novel algorithm incorporating the required prior information that avoids ambiguities in the solution. We obtain results on real data without any post-processing or user intervention.

2. SHAPE FROM CODED APERTURE

A generic object can be represented by a texture f and a depth d. We consider a mask composed of N small square apertures each offset by $\mathbf{\Delta_i}$, $i = 1 \dots N$. Then, the image g captured by a coded aperture camera with such a mask can be written as the linear combination of N views:

$$g(\mathbf{p}) = \frac{1}{N} \int \underbrace{\left(\sum_{i=1}^{N} \delta_d(\mathbf{p} + \boldsymbol{d}(\mathbf{p})\mathbf{\Delta_i}, \mathbf{q}) \right)}_{h_d(\mathbf{p},\mathbf{q})} \boldsymbol{f}(\mathbf{p})\, \mathrm{d}\mathbf{q} + w(\mathbf{p}), \tag{1}$$

where $\mathbf{p}$ is a pixel of the image g and $\mathbf{q}$ is a point of the object. The operator $h_d(\mathbf{p}, \mathbf{q})$ is called *point spread function* (PSF) and it depends upon the parameters of the camera as well as the 3D shape of the scene, and w is a zero-mean uncorrelated additive Gaussian noise $w(\mathbf{p}) \sim \mathcal{N}(0, \sigma^2)$. Eq. (1) can be written in a linear matrix-vector form:

$$\boldsymbol{g} = \boldsymbol{H_d f} + \boldsymbol{w} \tag{2}$$

where $\boldsymbol{H_d}$ represents the space-varying convolution with the aperture mask; this is conditional on the depth map $\boldsymbol{d}$, which is defined as a vector of depth values at each pixel of $\boldsymbol{f}$.

2.1. Image Prior Model

Similarly to [2], we define an image prior based on a set of P filtered versions of the original image $\boldsymbol{f}$:

$$\hat{\boldsymbol{f}}_k = \boldsymbol{C}_k \boldsymbol{f}, \qquad k = 1, \dots, P. \tag{3}$$

The operators $\boldsymbol{C}_k$ are zero mean conditional high-pass filters and each one of them is used to impose a particular contraint on the restored image $\boldsymbol{f}$.

Since $\boldsymbol{g}$ is Gaussian distributed and $\boldsymbol{C}_k$ is a linear operator, we can utilise the *commutative property*[1] and obtain that $\hat{\boldsymbol{g}}_k = \boldsymbol{C}_k \boldsymbol{g} = \boldsymbol{H_d} \hat{\boldsymbol{f}}_k + \boldsymbol{C}_k \boldsymbol{w}$ is also a Gaussian distributed, and its conditional distribution is given by

$$\mathbf{p}\left(\hat{\boldsymbol{g}}_k \,\middle|\, \hat{\boldsymbol{f}}_k, \boldsymbol{d} \right) = \mathcal{N}\left(\hat{\boldsymbol{g}}_k \,\middle|\, \boldsymbol{H_d} \hat{\boldsymbol{f}}_k, \boldsymbol{C}_k \sigma^2 \boldsymbol{I} \right). \tag{4}$$

The likelihood of our prior assumes that the k^{th} filtered versions of the sharp image $\boldsymbol{f}$ follows a Gaussian distribution with zero mean

$$\mathbf{p}\left(\hat{\boldsymbol{f}}_k \,|\, \boldsymbol{A}_k \right) = \mathcal{N}\left(\hat{\boldsymbol{f}}_k \,|\, 0, \boldsymbol{A}_k^{-1} \right). \tag{5}$$

where $\boldsymbol{A}_k$ is a diagonal matrix of variances $a_k(p)$ at each pixel p. Chantas *et al.*[2] model the distribution of $a_k(p)$ as a Gamma distribution, which leads to a heavy-tailed marginal distribution for $\hat{\boldsymbol{f}}_k$. A similar approach has been used in [12], but they impose $\boldsymbol{A}_k = \alpha \boldsymbol{I}$. Our assumption is that $\boldsymbol{A}_k$ is a diagonal matrix of unknown values which makes our marginalisation tractable. We write $\boldsymbol{A} = \{\boldsymbol{A}_1 \cdots \boldsymbol{A}_P\}$.

In general, the complete inference problem may be seen as estimating $\boldsymbol{d}$, $\hat{\boldsymbol{f}}$, and $\boldsymbol{A}$ from the observations $\hat{\boldsymbol{g}} = [\hat{\boldsymbol{g}}_1^T, \cdots, \hat{\boldsymbol{g}}_P^T]^T$. Since we are interested in depth estimation alone, we consider instead how to solve the problem

$$\boldsymbol{d}^* = \underset{\boldsymbol{d}}{\operatorname{argmax}}\, \mathbf{p}\left(\boldsymbol{d} \,|\, \hat{\boldsymbol{g}}, \boldsymbol{A} \right) \tag{6}$$

$$= \underset{\boldsymbol{d}}{\operatorname{argmax}}\, \mathbf{p}\left(\hat{\boldsymbol{g}} \,|\, \boldsymbol{d}, \boldsymbol{A} \right) \mathbf{p}\left(\boldsymbol{d} \right). \tag{7}$$

We call *shape from coded aperture* the problem of reconstructing the projected depth map $\boldsymbol{d}$ given the set of observed filtered images $\hat{\boldsymbol{g}}$, described in eq. (6). In the next section the marginal likelihood in eq. (7) is obtained.

3. BAYESIAN DEPTH INFERENCE

We now describe how to estimate the depth map directly from the observations without explicit estimation of the texture.

3.1. Marginalisation

To begin the analysis, we marginalise $\hat{\boldsymbol{f}}_k$ as follows:

$$\mathbf{p}\left(\hat{\boldsymbol{g}}_k \,|\, \boldsymbol{d}, \boldsymbol{A}_k \right) = \int \mathbf{p}\left(\hat{\boldsymbol{g}}_k, \hat{\boldsymbol{f}}_k \,|\, \boldsymbol{d}, \boldsymbol{A}_k \right) \mathrm{d}\hat{\boldsymbol{f}}_k \tag{8}$$

$$= \int \mathbf{p}\left(\hat{\boldsymbol{g}}_k \,\middle|\, \hat{\boldsymbol{f}}_k, \boldsymbol{d} \right) \mathbf{p}\left(\hat{\boldsymbol{f}}_k \,|\, \boldsymbol{A}_k \right) \mathrm{d}\hat{\boldsymbol{f}}_k \tag{9}$$

$$= \mathcal{N}\left(\hat{\boldsymbol{g}}_k \,|\, \boldsymbol{\mu}_k(\boldsymbol{A}_k), \boldsymbol{\Sigma}_k(\boldsymbol{A}_k) \right) \tag{10}$$

where[2] $\boldsymbol{\mu}_k(\boldsymbol{A}_k) = 0$ and $\boldsymbol{\Sigma}_k(\boldsymbol{A}_k) = \boldsymbol{H_d} \boldsymbol{A}_k^{-1} \boldsymbol{H_d}^T + \boldsymbol{C}_k \sigma^2 \boldsymbol{I}$.

This integration is achieved by applying the Gaussian integral.[3] One could estimate $\boldsymbol{A}_k$ and use the definition of $\boldsymbol{\Sigma}_k$ to evaluate the likelihood in eq. (10). For simplicity, we propose to estimate $\boldsymbol{\Sigma}_k$ directly from the data. This becomes tractable due to (i) the fact that eq. (10) is Gaussian, which allows us to work with local conditional distributions (Section 3.2) and (ii) the structure of $\boldsymbol{\Sigma}_k$ (Section 3.3).

3.2. Local Factorisation of $\boldsymbol{\Sigma}_k$

To work locally, the Markov Random Field (MRF) principle of conditional independence may be applied, if we can show that the pixel p only depends on certain neighbours in a given small region N_p:

[1] Strictly this only holds for planar scenes; however we find this is a reasonable approximation if we work with locally frontal-planar patches.

[2] The mean is given by $\boldsymbol{\mu}_k(\boldsymbol{A}_k) = \boldsymbol{\Sigma}_k^{-1} \left(\sigma^{-2} \boldsymbol{I} + \boldsymbol{H_d}^{-T} \boldsymbol{A}_k \boldsymbol{H_d}^{-1} \right) \sigma^{-2} \boldsymbol{H_d}^{-1} \boldsymbol{A}_k \boldsymbol{\mu}_{\hat{f}}$ where $\boldsymbol{\mu}_{\hat{f}} = 0$ in our image prior model.

[3] Due to normalisation of the Gaussian distribution, we have in general that $\int \cdots \int_{\mathbb{R}^{P \times 1}} \exp\left[-\frac{1}{2} \left(\boldsymbol{x}^T \boldsymbol{\Gamma} \boldsymbol{x} - 2 \boldsymbol{\beta}^T \boldsymbol{x} + \alpha \right) \right] \mathrm{d}\boldsymbol{x} = \frac{(2\pi)^{P/2}}{\sqrt{\det |\boldsymbol{\Gamma}|}} \exp\left[-\frac{1}{2} \left(\alpha - \boldsymbol{\beta}^T \boldsymbol{\Gamma}^{-1} \boldsymbol{\beta} \right) \right]$.

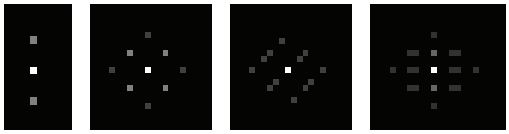


Fig. 2. **Structure of N_p for different masks**. Left to right: 2-hole, 4-hole symmetric, 4-hole asymmetric, and 5-hole mask. Notice that in the symmetric case some of the neighbours in N_p are counted more than once (brighter color).

$$\mathbf{p}\left(\hat{\boldsymbol{g}}_k[p] \,|\, \hat{\boldsymbol{g}}_k[\backslash p], \boldsymbol{d}\right) = \mathbf{p}\left(\hat{\boldsymbol{g}}_k[p] \,|\, \hat{\boldsymbol{g}}_k[N_p], \boldsymbol{d}\right). \tag{11}$$

In other words, rather than considering all other pixels $\hat{\boldsymbol{g}}_k[\backslash p]$ in the above expressions, we can just work with $\hat{\boldsymbol{g}}_k[N_p]$. This will be shown in Section 3.3.

Since $\hat{\boldsymbol{g}}_k$ is Gaussian, the conditional distribution of one pixel $\hat{\boldsymbol{g}}_k[p]$ in the image given the rest $\hat{\boldsymbol{g}}_k[\backslash p]$ is also Gaussian, with PDF

$$\mathbf{p}\left(\hat{\boldsymbol{g}}[p] \,|\, \hat{\boldsymbol{g}}[\backslash p], \boldsymbol{d}\right) = \mathcal{N}\left(\hat{\boldsymbol{g}}[p] \middle| \nu_{p|\backslash p}, \Gamma_{p|\backslash p}\right) \tag{12}$$

$$= \mathcal{N}\left(\hat{\boldsymbol{g}}[p] \middle| \nu_{p|N_p}, \Gamma_{p|N_p}\right), \tag{13}$$

with

$$\nu_{p|N_p} = \boldsymbol{\mu}[p] + \boldsymbol{\Sigma}[p, N_p]\boldsymbol{\Sigma}[N_p, N_p]^{-1}\left(\hat{\boldsymbol{g}}[N_p] - \boldsymbol{\mu}[N_p]\right) \tag{14}$$

$$\Gamma_{p|N_p} = \boldsymbol{\Sigma}[p, p] - \boldsymbol{\Sigma}[p, N_p]\boldsymbol{\Sigma}[N_p, N_p]^{-1}\boldsymbol{\Sigma}[N_p, p], \tag{15}$$

and $\boldsymbol{\mu}[p]$ and $\boldsymbol{\mu}[N_p]$ become zero from the assumption described in Section 2.1. The subscripts $(_k)$ are assumed but omitted for clarity and indices inside brackets address rows and columns of $\boldsymbol{\Sigma}_k$, such that the following structure contains all non-zero elements pertaining to the pixel p

$$\left[\begin{array}{c|c} \boldsymbol{\Sigma}[p, p] & \boldsymbol{\Sigma}[p, N_p] \\ \hline \boldsymbol{\Sigma}[N_p, p] & \boldsymbol{\Sigma}[N_p, N_p] \end{array}\right] \tag{16}$$

where $\boldsymbol{\Sigma}[p, N_p]$ is of size $1 \times |N_p|$ and $\boldsymbol{\Sigma}[N_p, p] = \Sigma[p, N_p]^T$.

3.3. Structure of the Local Neighbourhood in $\boldsymbol{\Sigma}_k$

Since $\boldsymbol{A}_k$ is diagonal, the neighbourhood structure N_p only depends on the offsets in $\boldsymbol{H_d}$. In Fig. 2 we show some examples of neighbourhood structure generated by $\boldsymbol{H_d}\boldsymbol{A}_k^{-1}\boldsymbol{H_d}^T$. The bright point at the center of each image indicates the pixel p and the surrounding points represent the neighbourhood N_p.

Thus the neighbourhood N_p of a pixel p can be defined as $N_p = \{\boldsymbol{p} + \boldsymbol{\delta}_{ij} d \,|\, i \neq j \wedge i, j \in \mathcal{M}\}$ where $\boldsymbol{\delta}_{ij} = (\boldsymbol{\Delta}_i - \boldsymbol{\Delta}_j)$ is a vector that represents the distance between the aperture i and the aperture j in the mask $\mathcal{M}$. The number of elements in N_p is given by

$$|N_p| = \frac{N!}{(N-2)!} = N(N-1), \tag{17}$$

which indicates that the amount of computations of our algorithm increases with the number of apertures N in the mask.

Since we have verified that the pixel p only depends on a small finite number of neighbours N_p, we can apply the MRF principle of conditional independence:

$$\mathbf{p}\left(\hat{\boldsymbol{g}} \,|\, \boldsymbol{A}, \boldsymbol{d}\right) = \prod_{p=1\ldots M} \mathbf{p}\left(\hat{\boldsymbol{g}}_k(p) \,|\, \hat{\boldsymbol{g}}_k(N_p), \boldsymbol{d}\right). \tag{18}$$

Due to just one observation of the image being available, we employ the ergodicity assumption of local stationarity, that is a local window can be used to estimate the required statistics at each point in the image.

4. MAP ESTIMATION OF DEPTH MAP

Given the local estimates of the image mean and variance conditional on each possible depth (we assume a discrete set of depth values corresponding to integer disparities), we can consider maximising the posterior for $\boldsymbol{d}$ in eq. (7). Due to the independence of the filtered observations [2], $\mathbf{p}\left(\hat{\boldsymbol{g}} \,|\, \boldsymbol{A}, \boldsymbol{d}\right) = \prod_{k=1\ldots P} \mathbf{p}\left(\hat{\boldsymbol{g}}_k \,|\, \boldsymbol{A}_k, \boldsymbol{d}\right)$. We define the prior $\mathbf{p}\left(\boldsymbol{d}\right)$ as the penalty term on the gradients of the depth map in the L_1 norm (Gibbs distribution). We can now take the negative logarithm of the likelihood in eq. (7), apply the MRF principle in eq. (18), and successively eq. (13); this yields

$$\boldsymbol{d}^* = \underset{\boldsymbol{d}}{\operatorname{argmin}}\left(E_{data}(\boldsymbol{d}) + E_{sm}(\boldsymbol{d})\right) \tag{19}$$

with

$$E_{data}(\boldsymbol{d}) = \frac{1}{2}\sum_k \sum_p \Big[(\hat{\boldsymbol{g}}_k(p) - \nu_{p|N_p})^T \Gamma_{p|N_p}^{-1}(\hat{\boldsymbol{g}}_k(p) - \nu_{p|N_p}) + \log(2\pi \det|\Gamma_{p|N_p}|)\Big] \tag{20}$$

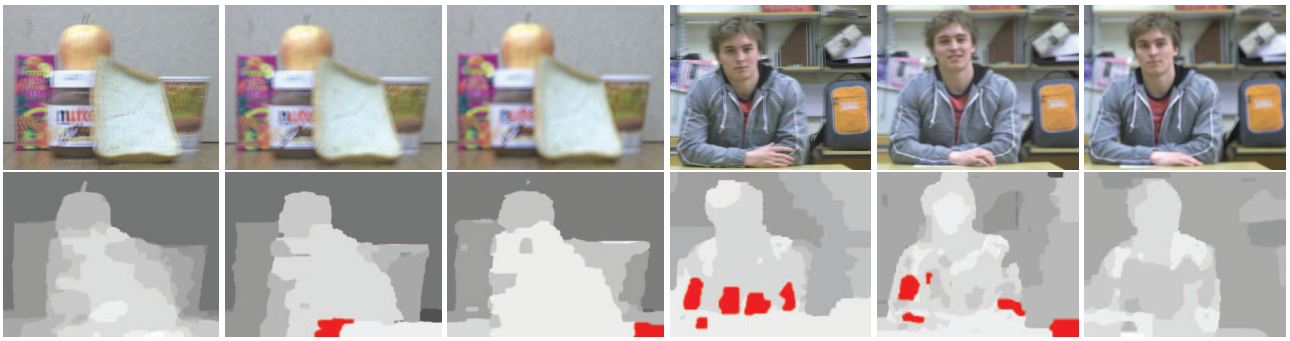
$$E_{sm}(\boldsymbol{d}) = -\log \mathbf{p}\left(\boldsymbol{d}\right) = \sum_{p,\{q \in V_p\}} \min(|d_p - d_q|, T), \tag{21}$$

where V_p is the neighborhood of a pixel p and T is a constant. Thus E_{sm} penalizes differences in the depths of neighboring pixels. Our inference procedure consist of minimising the energy given by eq. (19) via Graph-Cuts [11]. In our implementation the number of operators $\boldsymbol{C}_k$ is $P = 2$, and they correspond to discrete horizontal and vertical derivatives.

5. EXPERIMENTS

5.1. Performance

We have compared our algorithm with 5 methods previously proposed for coded aperture images on different types of aperture. Since the computational cost of our algorithm is rapidly increasing with the number of apertures in the mask (as described in eq. (17), consider only 3 simple ones: 2-hole, 3-hole, and 4-hole masks. We synthetically simulated some coded images by placing a plane of random texture at 33 different depths. Then we gave these images as input to the methods and reported the mean error in Table 1 (occlusions are not considered). SNR is taken into account by considering the amount of light that goes through each aperture. Since our algorithm does not restore, its computational time is very low for the types of masks we have analysed here: it takes less then 1 minute (in a Pentium Core2Duo 3.00GHz) to compute the depth map of a coded image of size 640×480 taken with a 2-hole mask, like the one shown in Fig. 1.



(a) (b) (c) (d) (e) (f)

Fig. 3. **Real data**. Images given as input to our algorithm (top) and their relative depth map (bottom). The two scenes has been both captured with 3 different masks: 2-hole (a, d), 3-hole (b, e), and 4-hole (c, f). Red color represents areas where depth has not been estimated. All the datasets shown here will be available online from the authors' website.

Methods	**Masks** - (image noise level $\sigma = 0.0001$)		
	2-hole	3-hole	4-hole
Lucy-Richardson	14.2	15.4	13.9
Regular filters	11.9	14.9	13.9
Wiener filters	17.0	17.0	15.1
Gaussian priors[12]	12.1	15.2	13.7
Levin *et al.*[12]	13.9	15.2	13.7
Our method	**8.7**	**9.3**	**8.7**

Table 1. Performance comparison (mean error).

5.2. Real Data

Coded aperture images were obtained by inserting a mask into a 50mm $f/1.4$ lens mounted on a Canon EOS-5D DSLR. The shutter speed of each exposure was set to 40ms (ISO 500) for images captured with the 2-hole mask, 33ms (ISO 400) with the 3-hole mask, and 20ms (ISO 400) for the 4-hole mask. Each aperture in the mask is a 4×4 mm square, and the distance between the centers of the holes is about 13 mm. We apply our algorithm to two different kinds of scenario to show how it performs with different ranges of depths and changes of mask. In order to maximise the disparities, we set the focal plane of the camera lens just after the object of interest (Fig. 1 and Fig. 3(a, c)) or just before them (Fig. 3(d, f)). Fig. 3(a-c) displays a scene with several objects placed at distances between 80cm and 120cm from the camera lens, while Fig. 3(d, f) represents a scene with a wider range of depths (from 200cm to 350cm). One can notice from the estimated depth maps that when we increase the number of the apertures in the mask we loose details but we solve some ambiguities due to occlusion or repeating texture which are present in images captured with masks with 2 or 3 apertures. Fig. 1 in front page shows a depth map obtained from a coded aperture image capture with a 2-holes mask. The focal plane is at 120cm and the objects are placed in a range of 50cm.

6. CONCLUSIONS

We have presented analysis and a simple algorithm to solve shape from coded aperture. We propose a novel depth inference procedure in a Bayesian framework that has higher performance then previous method based on a single image as input. We introduce both priors on the scene texture and depth map, and show how to bypass recovery of the sharp texture, without restrictive assumptions.

7. REFERENCES

[1] J. Ables. Fourier transform photography: a new method for x-ray astronomy. *Proceedings of the Astronomical Society of Australia*, 1:172–177, Dec 1968. 1

[2] G. Chantas, N. Galatsanos, A. Likas, and M. Saunders. Variational bayesian image restoration based on a product of t-distributions image prior. *IEEE Transactions on Image Processing*, 17(10):1795–1805, Oct 2008. 2, 3

[3] S. Chaudhuri and A. Rajagopalan. *Depth from Defocus: a Real Aperture Imaging Approach*. Springer-Verlag, 1999. 2

[4] R. Dicke. Scatter-hole cameras for x-rays and gamma rays. *Astrophys Journal*, 153:101–112, Aug 1968. 1

[5] E. R. Dowski and T. W. Cathey. Single-lens single-image incoherent passive-ranging systems. *Applied Optics*, 33(29):6762–6773, 1994. 1

[6] H. Farid. *Range Estimation by Optical Differentiation*. PhD thesis, University of Pennsylvania, 1997. 2

[7] P. Favaro and S. Soatto. *3-D Shape Reconstruction and Image Restoration: Exploiting Defocus and Motion-Blur*. Springer-Verlag, 2006. 2

[8] A. Greengard, Y. Y. Schechner, and R. Piestun. Depth from diffracted rotation. *Optics Letters*, 31(2):181–183, 2006. 1

[9] G. E. Johnson, E. R. Dowski, and W. T. Cathey. Passive ranging through wave-front coding: Information and application. *Applied Optics*, 39(11):1700–1710, 2000. 1

[10] D. Jones and D. Lamb. Analyzing the visual echo: Passive 3-d imaging with a multiple aperture camera. Technical report, McGill University, Feb 1993. 1

[11] V. Kolmogorov and R. Zabih. Computing visual correspondence with occlusions via graph cuts. *IEEE ICCV*, 2:508–515, 2001. 3

[12] A. Levin, R. Fergus, F. Durand, and W. T. Freeman. Image and depth from a conventional camera with a coded aperture. *SIGGRAPH*, 26(3):70, 2007. 1, 2, 4

[13] C.-K. Liang, T.-H. Lin, B.-Y. Wong, C. Liu, and H. Chen. Programmable aperture photography: Multiplexed light field acquisition. *ACM Transactions on Graphics*, 27(3):55:1–55:10, 2008. 1

[14] R. Raskar, A. K. Agrawal, and J. Tumblin. Coded exposure photography: Motion deblurring using fluttered shutter. *ACM Trans. Graph.*, 25(3):795–804, Jul 2006. 1

[15] A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin. Dappled photography: mask enhanced cameras for heterodyned light fields and coded aperture refocusing. *SIGGRAPH*, 26(3):69–80, Jul 2007. 1