

Coded Apertures for Defocus Deblurring

Belen Masia, Adrian Corrales, Lara Presa and Diego Gutierrez

Universidad de Zaragoza

Abstract

The field of computational photography, and in particular the design and implementation of coded apertures, has yielded impressive results in the last years. Among their applications lies defocus deblurring, in which we focus in this paper. Following the approach of previous works, we obtain near-optimal coded apertures using a genetic algorithm and an existing quality metric. We perform both synthetic and real experiments, testing the performance of the apertures along the dimensions of depth, size and shape. We additionally explore non-binary apertures, usually overlooked in the literature, and perform a comparative analysis with their binary counterparts.

Categories and Subject Descriptors (according to ACM CCS): I.4.3 [Image Processing and Computer Vision]: Enhancement—Sharpening and deblurring

1. Introduction

In the past few years, the field of computational photography has yielded spectacular advances in the imaging process. The main idea is to code the light information in novel ways before it reaches the sensor, in order to decode it later and obtain an improved, enhanced or extended representation of the scene being captured. Several different strategies exist, from structured lighting, to new optical devices, to modulated apertures or shutters. In this work we focus on *coded apertures*. These are masks obtained by means of computational algorithms which, placed at the camera lens, encode the defocus blur in order to better preserve high frequencies in the original image. They can be seen as an array of multiple ideal pinhole apertures (with infinite depth and no chromatic aberration), whose location on the 2D mask is determined computationally. Decoding the overlap of all pinhole images yields the final image.

Some existing works interpret the resulting coded blur attempting to recover *depth from defocus*. Given the nature of the blur as explained by simple geometrical optics, this approach imposes a multi-layered representation of the scene being depicted. While there is plenty of interesting on-going research in that direction, in this paper we limit ourselves to the problem of *defocus deblurring*: we aim to obtain good coded apertures that allow us to recover a sharp image from its blurred original version. We follow standard approaches and pose the imaging process as a convolution between the original scene being captured and the blur kernel (plus a

noise function). In principle, this would lead to a blind deconvolution problem, given that the such blur kernel is usually not known. Assuming no motion blur nor camera shake, this kernel is reduced to the point spread function of the optical system. Traditional circular apertures, however, have a very poor response in the frequency domain: not only do they lose energy at high frequencies, but they exhibit multiple zero-crossings as well; it is thus impossible to recover information at such frequencies during deconvolution.

In this paper, we present several coded apertures with better frequency response, which allow us to recover information apparently lost to blur during the capture process. We follow the approach of previous works, and rely on the average power spectra of natural images to guide our optimization process, which is in turn performed by means of genetic algorithms. Once the coded apertures have been obtained, we show the feasibility of our results by printing them out on a photomask sheet and inserting them in an off-the-shelf camera. The captured blurred images are then deconvolved using Wiener deconvolution. We analyze the performance of our apertures as a function of shape, depth and size. We additionally modify our genetic algorithm to allow for non-binary masks, and perform a comparative analysis with their binary counterparts.

2. Previous Work

Coded apertures have been traditionally used in astronomy, coding the direction of incoming rays as an alternative to fo-

cusing imaging techniques [Itz92]. Possibly the most popular patterns were the MURA patterns (Modified Uniformly Redundant Array) [GF89]. Veeraraghavan et al. [VRA*07] showed how a 4D light field can be reconstructed from 2D sensor information by means of a coded mask. Placed at the lens, the authors achieve refocusing of images at full resolution, provided the scene being captured contains only Lambertian objects. Nayar and Mitsunaga [NM00], extended the dynamic range capabilities of the imaging system by placing a mask of spatially varying transmittance next to the sensor, and then mapping the captured information to high dynamic range.

Other works have proposed different coded apertures for defocus deblurring or depth approximation. To restore a blurred image, the apertures are designed to have a broad-band frequency response, along with none (or distinguishable) zero-crossings in the Fourier domain. Hiura and Matsuyama [HM98] proposed a four-pinhole coded aperture to approximate the depth of the scene, along with a deblurred version of it, although their system required multiple images. Liang et al. [LLW*08] use a similar approach, combining tens of images captured with Hadamard-based coded patterns. Levin et al. [LDF07] attempted to achieve all-focus and depth recovery simultaneously, relying on image statistics to design an optimal aperture. Depth recovery is limited to a multi-layered representation of the scene. Last, the idea of spatial coding of the mask was transferred to the temporal domain by applying a coded exposure aimed at motion deblurring [RAT06].

In [ZLN09], the authors obtained paired apertures to recover both depth and focus from two images, using both genetic algorithms and gradient descent search. Last, a framework for evaluating coded apertures was recently presented [ZN09], based on the quality of the resulting deblurring and taking into account natural image statistics. Near-optimal apertures are obtained by means of a genetic algorithm. In this paper we follow the same approach, and analyze the obtained apertures along the size, depth and shape dimensions. Additionally, we extend our study by analyzing non-binary masks.

3. Optimal Aperture Design

Image blur due to defocus is caused by the loss of high frequency content when capturing the image. The capture process can be modeled as a convolution between the scene being captured and the point spread function (PSF) of the camera, which is defined as the response of the optical system of the camera to an impulse input in the spatial domain. Thus:

$$f = k_d * f_0 + \eta \quad (1)$$

where f_0 is the real scene being photographed, f is the captured image, k_d is the PSF and η accounts for the noise introduced in the imaging process. Subscript d accounts for depth, since the PSF varies with depth or, more specifically,

with the degree of defocus (strictly speaking, it also varies spatially with the position within the image). We will assume that the noise follows a Gaussian distribution of zero mean and standard deviation denoted by σ , $N(0, \sigma^2)$. By means of deconvolution, an approximation \hat{f}_0 to the original sharp image can be obtained. Note that in the frequency domain the convolution becomes a multiplication, and Equation 1 can be written as:

$$F = K_d \cdot F_0 + \zeta \quad (2)$$

As Figure 1 shows, the PSF, and thus the response of the camera, is characterized by the pattern of the aperture. The response to a coded aperture can also be seen in Figure 2, which depicts the calibration array used in our physical experiments. Since, as mentioned, blur is caused by the loss of information at certain frequencies, the response of an aperture is better analyzed in the frequency domain. Figure 3 depicts a 1D slice of the power spectrum of different aperture patterns, computed by Fourier transforming the aperture (note that the y-axis is log-scale). This shows the magnitude of the response for different frequencies. Circular apertures exhibit zero crossings at several frequencies, and thus information at those frequencies is lost during the imaging process. Optimal apertures for deblurring therefore seek a smooth power spectrum, while keeping the transmitted energy as high as possible.



Figure 1: Left: Images of the response to a point light of different apertures (from top to bottom: focused aperture, defocused circular aperture -defocus depth = 90 cm- and one of our coded apertures -defocus depth = 90 cm-, shown in the right). A LED and black cardboard were used to create the point light. Right: Canon EF 50mm f/1.8 lens with one of our coded apertures.

3.1. Aperture Quality Metric

Devising an aperture pattern whose frequency response is optimal can be done in different manners. In this paper we follow the approach of Zhou and Nayar [ZN09], which states the quality of an aperture pattern based on the quality of the deconvolution and on a prior model of natural images. In the following we briefly describe the metric and its foundation, and we refer the reader to the original paper for additional details.

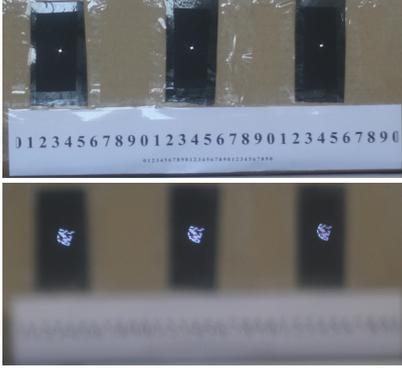


Figure 2: Our poor man's LED array used to calibrate the PSFs of the apertures. Top: Focused image. Bottom: Image taken with one of our coded apertures at a defocus depth of 70 cm.

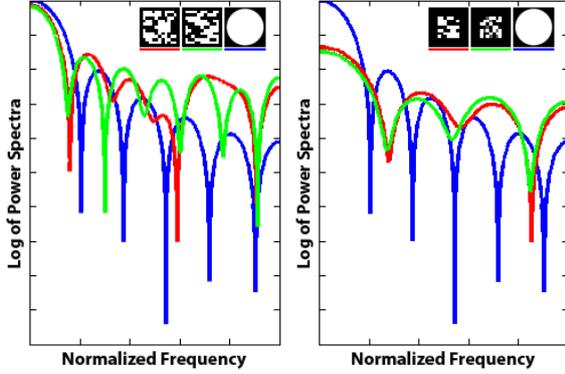


Figure 3: Power spectra comparison of different apertures with respect to a circular aperture (blue). Left: Our apertures for resolution 11×11 and noise levels $\sigma = 0.001$ (red) and $\sigma = 0.005$ (green). Right: Our apertures for resolution 7×7 , binary (red) and non-binary (green).

The quality metric chosen is the expectation of the L_2 distance between the deconvolved image \hat{F}_0 and the ground truth image F_0 with respect to ζ , which we want to be minimal (note that we have removed the subscript d for the sake of simplicity):

$$R(K, F_0, C) = \mathbb{E}_{\zeta} [\|\hat{F}_0 - F_0\|^2] \quad (3)$$

The recovered image \hat{F}_0 can be obtained using Wiener deconvolution as follows:

$$\hat{F}_0 = \frac{F \cdot \bar{K}}{|K|^2 + |C|^2} \quad (4)$$

where \bar{K} is the complex conjugate of K , and $|K|^2 = K \cdot \bar{K}$. $|C|^2 = C \cdot \bar{C}$ is the matrix of noise-to-signal power ratios (NSR) of the additive noise. Substituting this formulation in

Equation 3 we have:

$$R(K, F_0, C) = \mathbb{E}_{\zeta} \left[\left\| \frac{\zeta \cdot \bar{K} - F_0 \cdot |C|^2}{|K|^2 + |C|^2} \right\|^2 \right] \quad (5)$$

and assuming that ζ follows a Gaussian distribution with zero mean, $\zeta \sim N(0, \sigma^2)$:

$$R(K, F_0, C) = \left\| \frac{\sigma \cdot \bar{K}}{|K|^2 + |C|^2} \right\|^2 + \left\| \frac{F_0 \cdot |C|^2}{|K|^2 + |C|^2} \right\|^2 \quad (6)$$

Using a model of natural images as a prior, the expectation of $|F_0|^2$ is

$$A(\xi) = \int_{F_0} |F_0(\xi)|^2 d\mu(F_0), \quad (7)$$

where ξ represents frequency and A can be approximated by averaging the power spectra of a number of natural images. This way the dependence on F_0 , which is unknown, is circumscribed, obtaining:

$$R(K, C) = \left\| \frac{\sigma \cdot \bar{K}}{|K|^2 + |C|^2} \right\|^2 + \left\| \frac{A^{1/2} \cdot |C|^2}{|K|^2 + |C|^2} \right\|^2 \quad (8)$$

The value of $|C|^2$ which, for a given K , minimizes the value of R is $|C|^2 = \sigma^2/A$. Substituting this value in Equation 8 yields the sought quality metric, which depends only on the Fourier transform of the aperture pattern K , the estimated image noise σ and the average power spectra of natural images A :

$$R(K) = \frac{\sigma^2}{|K|^2 + \sigma^2/A} \quad (9)$$

3.2. Aperture Pattern Optimization

Once we have a way of evaluating a certain aperture with Equation 9, an optimization method can be used to obtain the minimum value of $R(K)$ over the range of possible apertures. The space of possible apertures is infinite, since the aperture can be of different resolutions, and each pixel can in principle take infinite values. A priori the solution is limited only by physical restrictions, i.e. apertures with negative values are not realizable in practice and resolution is limited by the printing process. Resolution is additionally limited by diffraction effects, which appear as the size of the pixels in the aperture gets smaller, and hinder the performance of the aperture. Transmissivity is an additional issue to be taken into account when designing an aperture. Coded apertures typically have lower transmission rates than their circular counterparts, and the use of a longer exposure time to obtain an equivalent brightness to that of the circular aperture can cause other problems such as motion blur. This metric does not consider transmissivity when evaluating an aperture, but still it yields satisfactory results for the majority of cases.

4. Experimental Setup and Results

In order to search for the best aperture pattern we have implemented a genetic algorithm which uses the quality metric described in Section 3 as evaluation function, resembling Zhou and Nayar’s work. The algorithm has the following scheme:

- *Initialization.* The initial population of N possible apertures is randomly generated. An aperture is defined by a vector of L elements, each element corresponding to a pixel.
- *Selection.* The quality metric of Equation 9 is used to evaluate the N possible apertures. They are then sorted according to this value and the best M apertures are selected.
- *Reproduction.* The selected M apertures, by means of crossover and mutation, populate the next generation. Crossover implies randomly selecting two apertures, duplicating them, and exchanging corresponding bits between them with probability c_1 , obtaining two new apertures. Mutation ensures diversity by modifying each bit of the aperture with probability c_2 .
- *Termination.* The two previous steps of reproduction and selection are repeated sequentially until the termination condition is met. We use a maximum number of generations G as stopping condition.

We have tested apertures of two different resolutions, 11×11 and 7×7 pixels (that is, $L = 121$ and $L = 49$, respectively), while the rest of the parameters we used for the algorithm are $N = 4000$, $M = 400$, $G = 80$, $c_1 = 0.2$ and $c_2 = 0.05$. Since the optimal aperture depends on the noise of the image we have run the algorithm for different noise levels and tested the resulting apertures. Apertures designed for σ values of 0.001 and 0.005 proved to work best for a wide variety of images. Regarding the possible values the pixels in the aperture can take, we have experimented both with binary and non-binary apertures, but at this first stage we show results just for binary apertures. Results for non-binary apertures are discussed in Section 5.

From all the obtained apertures we have chosen three, and a conventional circular aperture, to perform our experiments. We chose the ones which we saw performed best over a wide variety of images. Two of them are 11×11 apertures designed for noise levels of $\sigma = 0.001$ and $\sigma = 0.005$, and the third one is a 7×7 aperture designed for $\sigma = 0.005$. The three of them are depicted in Figure 4. For these apertures, we have performed both a synthetic validation and a validation with physical printed-out apertures.

The synthetic validation is done by simulating the capture process convolving a sharp image f_0 with the aperture (plus noise) as in Equation 1, and subsequently using Wiener deconvolution to recover a deblurred image \hat{f}_0 . The quality of the recovered image is measured using the L_2 norm. We did this for 10 images and computed the average L_2 value. Results are shown in Table 1 for the tested apertures. The minimum error is obtained by the 7×7 aperture and the

two 11×11 apertures perform very similarly, there is no significant difference, while, as expected, the circular aperture yields worse results. Another measure of the quality of

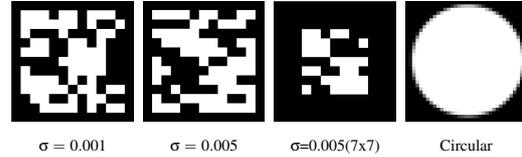


Figure 4: Apertures used in our experiments.

| | 0.001 | 0.005 | 0.005 (7x7) | circular |
|------------|-------|-------|-------------|----------|
| L_2 norm | 1.28 | 1.27 | 0.88 | 1.62 |

Table 1: Results of the L_2 norm for different apertures. The table shows percentages with respect to the maximum error.

the apertures is given by their power spectrum, depicted in Figure 3. The 11×11 apertures eliminate less frequencies than the circular aperture, and the 7×7 aperture has an even smoother spectrum, which correlates with the L_2 values previously obtained.

Experiments in real scenarios have been performed using a Canon EOS 550D with a Canon EF 50mm f/1.8 II lens shown (unmounted) in Figure 5. Our apertures were printed in a high resolution photomask sheet (see Figure 6 left) and inserted into the lens. The first step is the calibration of the



Figure 5: Camera and lens used in our experiments.

response of the camera (PSF) at different depths. We also calibrated the PSF for different image positions, since the response is spatially varying across the image plane. To do this we used an array of LEDs which we made as close as possible to point light sources with the aid of black cardboard. Figure 2 shows a close-up of the LED array. We locked the focus at 1 m and took an initial focused image, followed by images of the LEDs every 10 cm and until a distance of 2 m, thus having PSFs for defocus depths from 10 to 100 cm. For each position within the image and each depth, the actual cropped image of the LED served us as PSF, after appropriate thresholding of surrounding values which contain residual light. The resulting PSFs for three depths and the four tested apertures are shown in Figure 6 (right).

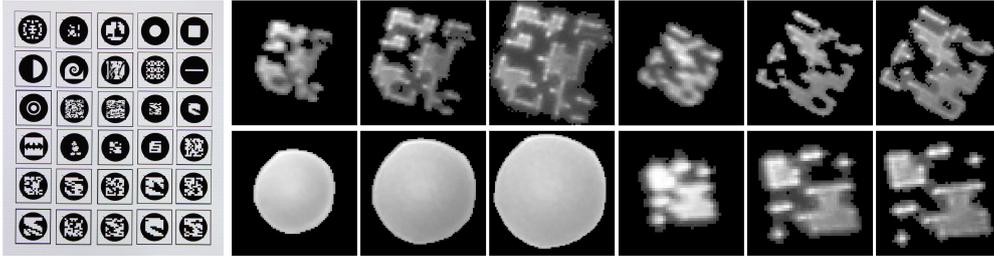


Figure 6: Left: Photomask sheet showing some of the apertures used. Right: PSFs at three different defocus depths (40, 70 and 90 cm) for the four apertures depicted in Figure 4.



Figure 7: Focused ground truth scenes.

Once calibration had been performed, images of three scenes at different depths were taken with each of the selected apertures. These images are then deblurred using the corresponding calibrated PSF by means of Wiener deconvolution. We used a NSR of 0.001 when deconvolving, since it gave the best results. The same exposure time and aperture was used for all the apertures, which results in some images being darker than others. Figure 7 shows the ground truth focused images of the three scenes, whereas Figure 8 depicts the defocused image captured with each aperture and the recovered image for the three different depths. Insets show the corresponding PSF. For all cases our apertures clearly outperform the circular one. The results of the other three apertures are fairly similar, with the 7×7 aperture revealing more detail than the others in some regions. However, we believe this may be due to the fact that because of its smaller size, it offers a wider depth of field, thus causing less defocus blur for the same settings as the others. The ringing artifacts which can be observed are probably partially caused by inaccuracies of the calibrated PSFs. Additionally, and although very minor, some of the apertures exhibit slight diffraction effects which can also be the cause of artifacts due to misalignments of the color channels [VRA*07].

5. Study of Non-Binary Apertures

Binary codes have the initial advantage of reducing the search space, and are usually preferred in the existing literature. However, there is no principled motivation to restrict the aperture pixel values to either black or white, other than apparent simplicity. A notable exception in this regard is the work by Veeraraghavan and colleagues [VRA*07], where the authors report the advantages of continuous-valued aper-

tures, found by gradient descent optimization, over their binary counterparts. In this section we perform an analysis of non-binary apertures focused on our specific context and optimization method; in order to limit the search space of the genetic algorithm, we restrict the set of possible values to $\{0, 0.5, 1\}$.

We have studied the quality of the resulting aperture and the computation time for different executions of the genetic algorithm for the cases of binary and non-binary apertures. We have varied both the initial population N and the number of generations G , yielding seven different combinations of these two parameters. For each combination of parameters we have performed three executions of the algorithm, plotting the average values. For all the figures in this section, the x-axis shows the initial population N and the number of generations G of each set of executions. The number of selected apertures, M , is always a 10% of the initial population, the crossover probability c_1 is set to 0.2 and the probability of mutation c_2 is 0.05. All the calculated apertures have a resolution of 7×7 . The reason of this is two-fold; first, computational cost of the algorithm is significantly reduced, and second, our previous experiments have shown that 7×7 apertures yield results on par with (or better than) 11×11 apertures. The value of σ (noise level) is set to 0.005 for all executions.

Figure 9 shows the average value of the quality metric to which the algorithm converged. Non-binary apertures tend to converge to slightly lower values of $R(K)$, potentially indicating a better performance. However, as expected, it also takes longer for non-binary apertures to converge to a stable value of $R(K)$. The execution times consumed until conver-

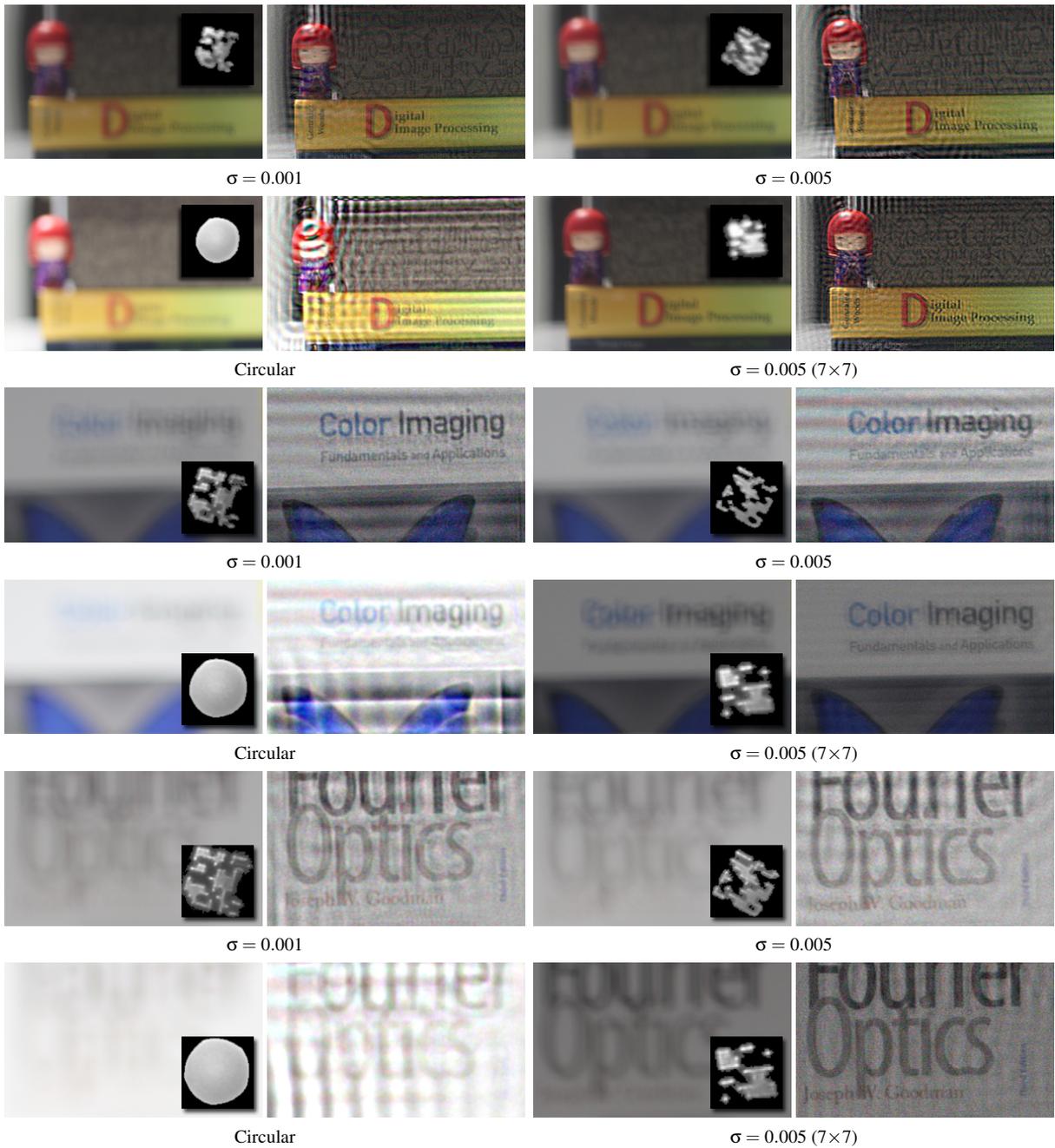


Figure 8: From top to bottom, each of the three scenes have been captured at a defocus depth of 40, 70 and 90 cm, respectively. For each pair of images, the left image shows the captured defocused image and the right image the recovered one. Insets depict the PSF of the aperture used in each case.

gence when running the algorithm on an Intel Core i7 930 @ 2.80GHz are shown in Figure 10.

For all the optimal apertures obtained in the different ex-

ecutions we have performed a synthetic evaluation similar to the one described in Section 4. We applied Equation 1 to an image f_0 of the ISO 12233 resolution chart, to simulate

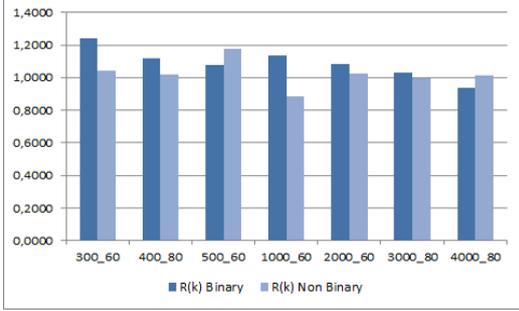


Figure 9: Average value of the quality metric $R(K)$ for binary and non-binary apertures and for different initial parameters of the genetic algorithm.

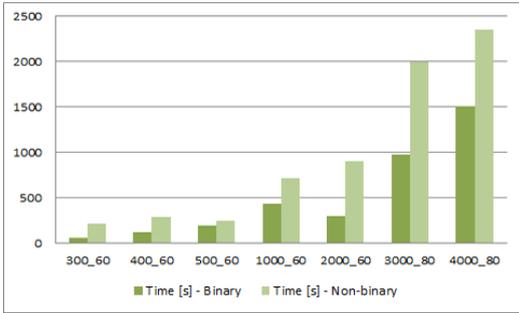


Figure 10: Average value of the time until convergence (in seconds) for binary and non-binary apertures and for different initial parameters of the genetic algorithm.

the capture process with the different apertures; we then performed Wiener deconvolution to recover the estimated sharp image \hat{f}_0 . We have computed the L_2 norm between \hat{f}_0 and f_0 and plotted the results in Figure 11. The non-binary apertures tend to behave better, the global tendency thus correlating with that of the quality metric $R(K)$. Nevertheless, this graphs shows how lower values of $R(K)$ not necessarily yield lower values of the L_2 norm. This can be explained by the fact that $R(K)$ is devised to give optimal performance over the entire space of natural images and thus may not be optimal for an image in particular [ZN09]. Figure 12 shows the image of the chart after convolution, and the recovered image for the best binary and non-binary apertures we obtained. For these two apertures we also plotted the power spectrum, shown in Figure 3 (right). Although both spectra are similar, overall the non-binary aperture has a more favorable response.

To further test the performance of binary vs. non-binary apertures, we printed out these two best apertures (shown in the insets of Figure 12) and captured real images with them. We have calibrated their PSF at different depths as explained in Section 4, and captured a set of images which we then have recovered using Wiener deconvolution. Figure

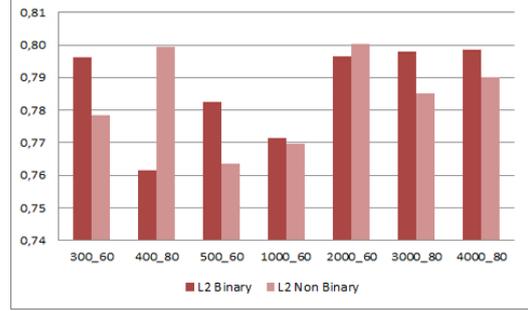


Figure 11: Average value of the L_2 norm for binary and non-binary apertures and for different initial parameters of the genetic algorithm. Values show percentage with respect to the maximum error.

13 shows the results for a defocus depth of 70 cm and 90 cm. The corresponding ground truth focused scenes are shown in Figure 7. We can see how even though both recover detail to a great extent, the non-binary aperture performs better.

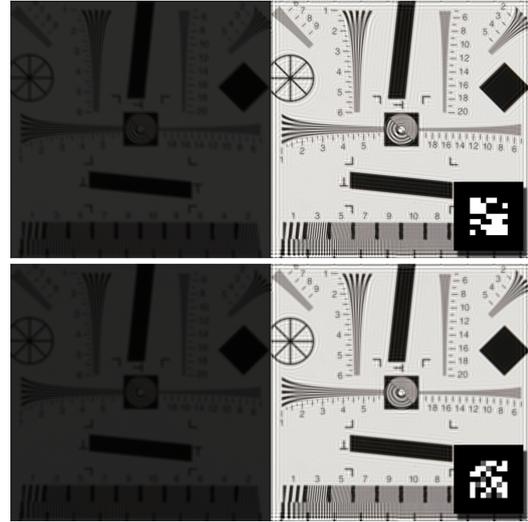


Figure 12: Top left: Defocused image of the ISO 12233 chart obtained using Equation 1 with the aperture shown in the inset in the right as convolution kernel. The aperture is the optimal binary aperture we obtained. Top right: Image recovered using Wiener deconvolution. Bottom row: Same for the optimal non-binary aperture, shown in the inset.

6. Conclusions and Future Work

In this paper we have introduced a comprehensive study of coded apertures for defocus deblurring, and implemented the full pipeline: from the genetic algorithms to obtain the codes, to their physical realization, and finally to the actual deblurring of out-of-focus images. We have analyzed the perfor-

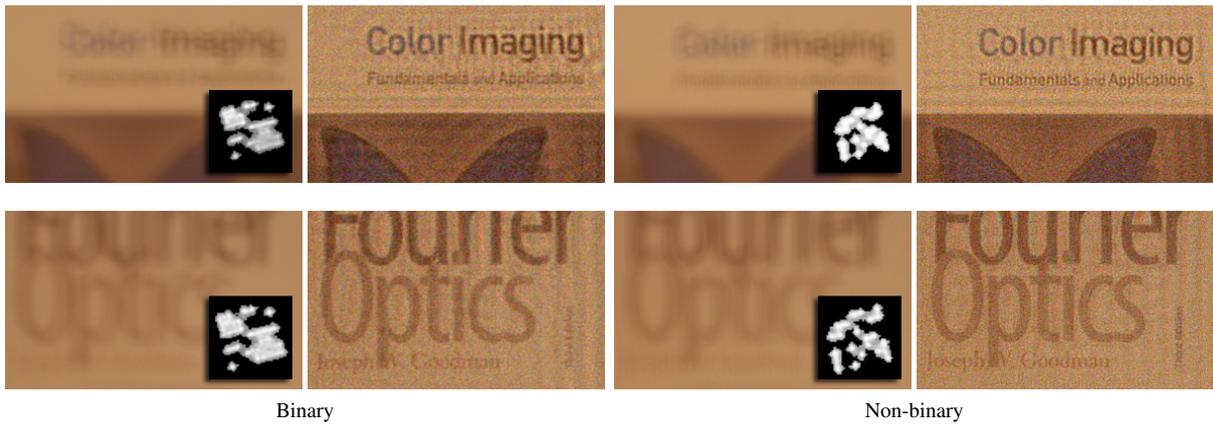


Figure 13: For each pair of images, the left image shows the captured defocused image and the right image the deblurred image. Insets depict the PSF used in each case. Defocus depths are 70 cm (top scene) and 90 cm (bottom scene). The color of the images differs from those in Figure 8; please note that this is due to the fact that illumination conditions during the capture process were different, and not to the coded apertures themselves.

mance of the different patterns along several dimensions, namely shape, depth and size. For instance, we found that 7×7 apertures are on par with, or outperform, higher resolution ones, which tend to be more computationally expensive to obtain. Additionally, we have extended previous works in the literature by lifting the binary restriction in our patterns, and allowing the genetic algorithms to add mid-gray to the binary (black or white) set of possible values. Although our results are not conclusive and more research needs to be carried out, initial findings suggest that there may be value in exploring continuous apertures, where several gray levels are allowed.

The inherent reduced light transmission when placing a modulating mask at the lens is also a factor that we would like to investigate further. By adding a term that maximizes transmission, we may come up with more efficient apertures. Similarly, finding coded apertures that optimize both defocus deblurring and depth is still an open problem where the community has barely scratched the surface. Last, we believe that the results shown in this paper show the viability and potential of this line of research, and we hope to raise awareness of this exciting field, fostering the creation of more research groups and potential collaborations.

7. Acknowledgements

We would like to thank the reviewers for their valuable comments. We also thank C. Zhou for his insights. This research has been partially funded by a Marie Curie grant from the 7th Framework Prog. (251415), the Spanish Ministry of Science and Technology (TIN2010-21543) and the Gobierno de Aragón (OTRI 2009/0411 and CTPP05/09). Belen Masia is

supported by a FPU grant from the Spanish Ministry of Education.

References

- [GF89] GOTTESMAN S., FENIMORE E.: New family of binary arrays new family of binary arrays for coded aperture imaging. *Applied Optics*, 20 (1989), 4344–4352. 2
- [HM98] HIURA S., MATSUYAMA T.: Depth measurement by the multi-focus camera. In *IEEE Conference on Computer Vision and Pattern Recognition* (1998). 2
- [ItZ92] IN 'T ZAND J.: *Coded Aperture Imaging in High-Energy Astronomy*. PhD thesis, University of Utrecht, 1992. 2
- [LFD07] LEVIN A., FERGUS R., DURAND F., FREEMAN W.: Image and depth from a conventional camera with a coded aperture. *ACM Transactions on Graphics* 26, 3 (2007). 2
- [LLW*08] LIANG C., LIN T., WONG B., LIU C., CHEN H.: Programmable aperture photography: multiplexed light field acquisition. *ACM Transactions on Graphics* 27, 3 (2008). 2
- [NM00] NAYAR S., MITSUNAGA T.: High dynamic range high dynamic range imaging: spatially varying pixel exposures. In *Computer Vision and Pattern Recognition* (2000), vol. 1, pp. 472–479. 2
- [RAT06] RASKAR R., AGRAWAL A., TUMBLIN J.: Coded exposure photography: Motion deblurring using uttered shutter. *ACM Transactions on Graphics* 25, 3 (2006), 795–804. 2
- [VRA*07] VEERARAGHAVAN A., RASKAR R., AGRAWAL A., MOHAN A., TUMBLIN J.: Dappled photography: mask enhanced cameras for heterodyned light fields and coded aperture refocusing. *ACM Trans. Graph.* 26 (July 2007). 2, 5
- [ZLN09] ZHOU C., LIN S., NAYAR S.: Coded aperture pairs for depth from defocus. In *ICCV (oral)* (2009). 2
- [ZN09] ZHOU C., NAYAR S. K.: What are Good Apertures for Defocus Deblurring? In *IEEE International Conference on Computational Photography* (2009). 2, 7