A Gentle Introduction to Coded Computational Photography

Horacio E. Fortunato   Manuel M. Oliveira
Instituto de Informática – UFRGS
Porto Alegre, RS, Brazil
E-mail: {hefortunato,oliveira}@inf.ufrgs.br

I. Abstract

Computational photography tries to expand the concept of traditional photography (a static two dimensional projection of a scene) using state-of-the-art technology. While this can be achieved by combining information from multiple conventional pictures, a more interesting challenge consists in encoding and recovering additional information from one (or more) image(s). Since a photograph results from the convolution of scene radiance with the camera's aperture (integrated over the exposure time), researchers have designed apertures with certain desirable spectral properties to facilitate the deconvolution process and, consequently, the recovery of scene information. Images captured using these so-called *coded apertures* can be deconvolved to create all-in-focus images, and to estimate scene depth, among other things. Images of moving objects acquired using a *coded exposure* (obtained by switching between a fully-closed and a fully-opened aperture, according to a predefined pattern) can be deconvolved to reduce motion blur. The notion of encoding information during image acquisition opens up new and exciting possibilities, which researchers have just begun to explore. This article provides a gentle introduction to coded photography, focusing on the fundamental concepts and essential mathematical tools.

Keywords- **computational photography, coded photography, image processing.**

II. Introduction

Digital photography replaced the photographic film by an electronic sensor. Despite of its practical implications, this phenomenon limited itself to replace the chemical process of film development by an electronic gathering of photons. Current technology, however, allows us to expand the capabilities of photography with new and exciting possibilities. The most straightforward way of achieving this is by combining information from multiple images acquired with a single camera. In this case, each image is obtained using slightly different parameter values, such as position/orientation, exposure time, and aperture size. Such a strategy can be used to construct panoramas [1], and high-dynamic-range images [2]. Multiple images can also be acquired using camera arrays [3], which can be used to improve resolution, signal-to-noise ratio, dynamic range, depth of field, frame rate, spectral sensitivity, or sample 4D light fields [4]. More sophisticated approaches can obtain additional information from a single picture by modifying one or more of the components of a traditional camera (*i.e.*, optics, sensor, and illumination). For instance, directional information of the scene radiance entering a camera can be recorded by placing an array of microlenses in front of the sensor [5], [6] , modifying the camera aperture [7] or the the camera optical system [8]. The resulting cameras are called *plenoptic cameras*. By trading spatial for angular resolution, plenoptic cameras can record a sampled version of the 4D light field [4]. Images acquired with plenoptic cameras can be refocused after acquisition time [5]. Other camera designs allow the recovery of a high-dynamic-range image from a single photograph by placing an optical mask corresponding to a mosaic of cells with different transmittances over the camera's sensor [9], [10]. A depth map can also be obtained from single pictures by placing a conical mirror in front of the camera's objective lens [11].

An important class of computational photography techniques use some coding strategies during image acquisition. Collectively known as **coded photography**, these approaches can be classified as *coded aperture*, *coded exposure*, *coded illumination* and *coded sensing*. Since a photograph results from the convolution of scene radiance with the camera's aperture (integrated over the exposure time), coded-aperture techniques use especially-designed apertures with certain spectral properties that facilitate the deconvolution process and the recovery of scene information. Although the images captured using coded apertures are usually not suitable for immediate visualization, with proper deconvolution they can be used, for instance, to recover all-in-focus pictures [12], [13], and to estimate scene depth [12], [14].

Coded-exposure techniques switch the camera's aperture between a fully-closed and a fully-opened situation during image acquisition, according to some predefined pattern. Given an image of a moving object acquired using coded aperture, the image can be deconvolved to reduce the occurrence of motion blur [15]. Coded illumination, on the other hand, consists in projecting some controlled light patterns into the scene to allow the extraction of scene properties.

This article provides a gentle introduction to coded photography, focusing on its fundamental concepts and essential mathematical tools. Several good surveys [16], [17], [18], [19] complement this material providing a comprehensive description of the computational photography field as a whole. Due to space constraints, we restrict the presentation to coded-

aperture and coded-exposure techniques. The article begins with a description of the concepts and tools required to master coded aperture and coded exposure. This is followed by sections relating the abstract concepts to actual techniques. It closes with a brief summary and a discussion of some research opportunities.

III. CONVOLUTION AND CORRELATION

This section briefly reviews the concepts of convolution and correlation, which are central to the image formation process. For this, matrix notation is used to represent images. Lower-case and uppercase symbols represent entities in the spatial and frequency domains, respectively. Thus, given an image f, $f_{r,c}$ refers to its pixel at row r and column c. F and $F_{u,v}$ denote f's Fourier transform, and its coefficients associated to frequency (u, v), respectively. We use circular convolution and modulo arithmetic for matrix index calculations. Assuming f has R rows and C columns, a row index $r < 0$ has its value replaced by $R + r$. If $r \geq R$, its value is then replaced by $r - R$. Likewise, a column index $c < 0$ has its value replaced by $C + c$, and if $c \geq C$, its value becomes $c - C$.

The circular convolution of two one dimensional vectors f and h with R elements is defined as

$$[f \otimes h]_i = \sum_{r=0}^{R-1} f_r \, h_{i-r}, \quad i \in [0, R-1], \tag{1}$$

where $[\square]_i$ is the i-th element of vector $\square$. The convolution of two matrices f and h with R rows and C columns can be expressed as

$$[f \otimes h]_{i,j} = \sum_{r=0}^{R-1} \sum_{c=0}^{C-1} f_{r,c} \, h_{i-r,j-c}, \tag{2}$$

where $i \in [0, R-1], j \in [0, C-1]$, and $[\square]_{i,j}$ is the element at row i and column j of matrix $\square$. The corresponding correlation operations are defined as

$$[f \circ h]_i = \sum_{r=0}^{R-1} f_r \, h_{i+r}, \tag{3}$$

and

$$[f \circ h]_{i,j} = \sum_{r=0}^{R-1} \sum_{c=0}^{C-1} f_{r,c} \, h_{i+r,j+c}. \tag{4}$$

Since these operations are fundamental to the image acquisition process, it is instructive to have some intuition on how they affect the processed data. For this, we will analyze their impact on some elementary matrices $d^{(r,c)}$, whose elements are all zeros, except for one, located at row r and column c, whose value is one. The elements of these matrices may be expressed as a product of two kronecker deltas. The *Kronecker delta* δ_{ij}, discrete analog of the Dirac delta, is a function of two variables i and j defined as

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j; \\ 0, & \text{otherwise.} \end{cases}$$
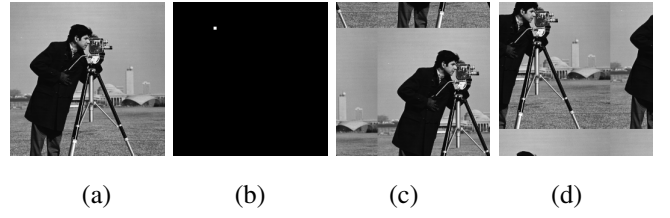

(a) (b) (c) (d)

Fig. 1. Convolution and correlation with a delta image. (a) Input image f. (b) Delta image $d^{(r,c)}$. (c) The convolution $f \otimes d^{(r,c)}$ circularly shifts f by r rows down and c columns to the right ($f_{i-r,j-c}$). (d) The correlation $f \circ d^{(r,c)}$ circularly shifts f in the opposite direction (r rows up and c columns to the left, producing $f_{i+r,j+c}$.

Thus, the elements of a matrix $d^{(r,c)}$ can be expressed as

$$d^{(r,c)}_{i,j} = \delta_{ri} \, \delta_{cj}.$$

We refer to the images corresponding to these elementary matrices simply as **delta images**. They define a basis for a vector space of images, as any 2D image f can be written as a linear combination of delta images:

$$f = \sum_{r=0}^{R-1} \sum_{c=0}^{C-1} f_{r,c} \, d^{(r,c)}. \tag{5}$$

The convolution of an image f with a delta image $d^{(r,c)}$ circularly shifts f by r rows down and c columns to the right (Fig. 1Convolution and correlation with a delta image. (a) Input image f. (b) Delta image $d^{(r,c)}$. (c) The convolution $f \otimes d^{(r,c)}$ circularly shifts f by r rows down and c columns to the right ($f_{i-r,j-c}$). (d) The correlation $f \circ d^{(r,c)}$ circularly shifts f in the opposite direction (r rows up and c columns to the left, producing $f_{i+r,j+c}$)figure.1, c):

$$[f \otimes d^{(r,c)}]_{i,j} = \sum_{k=0}^{R-1} \sum_{l=0}^{C-1} f_{k,l} \, d^{(r,c)}_{i-k,j-l},$$

$$[f \otimes d^{(r,c)}]_{i,j} = \sum_{k=0}^{R-1} \sum_{l=0}^{C-1} f_{k,l} \, \delta_{r,i-k} \, \delta_{c,j-l},$$

$$[f \otimes d^{(c,r)}]_{ij} = f_{i-r,j-c}. \tag{6}$$

Likewise, the correlation of f with a delta image $d^{(r,c)}$ circularly shifts f in the opposite direction, by r rows up and c columns to the left (Fig. 1Convolution and correlation with a delta image. (a) Input image f. (b) Delta image $d^{(r,c)}$. (c) The convolution $f \otimes d^{(r,c)}$ circularly shifts f by r rows down and c columns to the right ($f_{i-r,j-c}$). (d) The correlation $f \circ d^{(r,c)}$ circularly shifts f in the opposite direction (r rows up and c columns to the left, producing $f_{i+r,j+c}$)figure.1, d):

$$[f \circ d^{(c,r)}]_{ij} = f_{i+r,j+c}. \tag{7}$$

The relation between convolution and image formation will be detailed in section VIIImage Formation and Convolutionsection.6. Simply put, a captured image can be described as the convolution of an ideal all-in-focus image with the camera's *point spread function* (PSF), which describes
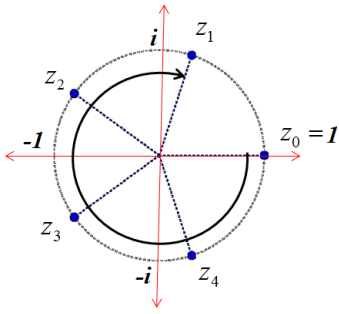
Fig. 2. N-th roots of unity on the complex plane. This example shows the roots of unit for $N = 5$.
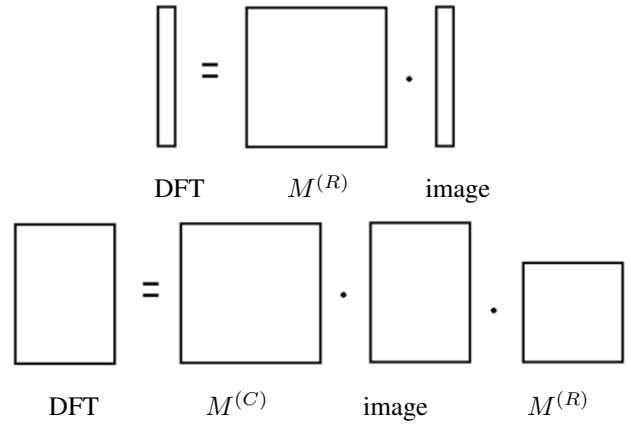

Fig. 3. Fourier transform computed as matrix products. (top) One-dimensional case: the DFT of (an image represented as) a 1D vector f can be obtained as the matrix-vector product $F = M^{(R)}f$ (Equation 8Convolution and Fourier transformsequation.4.8). (bottom) Two-dimensional case: the DFT of (an image represented as) a 2D matrix is another 2D matrix obtained as $F = M^{(R)}fM^{(C)}$ (Equation 10Convolution and Fourier transformsequation.4.10). $M^{(N)}$ is a square matrix constructed from the N-th roots of unity (Equation 9Convolution and Fourier transformsequation.4.9).

the camera's response to a point source. In turn, the PSF is largely shaped by the camera's aperture.

Convolution and correlation are linear operators. As we show later in section VMatrix Operatorssection.5, there is a one-to-one correspondence between linear operators in a vectorial space and matrices. This provides another way of writing the convolution and correlation operations as a conventional matrix-vector multiplication, which will be useful to simplify the algebraic formulations of deconvolution techniques.

IV. CONVOLUTION AND FOURIER TRANSFORMS

Given a column vector f with R rows, we define its discrete Fourier transform (DFT) F as the product of f by a square matrix $M^{(R)}$, with dimensions $R \times R$ and complex elements:

$$F = M^{(R)}f, \tag{8}$$

where

$$M^{(R)}_{r,c} = \exp\left(-2\pi i\left(\frac{rc}{R}\right)\right).$$

The elements of $M^{(R)}$ are all members of the group U_R of R-th roots of unity (i.e., $U_R = \{z \in \mathbb{C}|R \in \mathbb{Z}, z^R = 1\}$). For example, for $R = 5$, $M^{(5)}$ is expressed as:

$$M^{(5)} = \begin{bmatrix} z_0^0 = 1 & z_4^0 = 1 & z_3^0 = 1 & z_2^0 = 1 & z_1^0 = 1 \\ z_0^1 = 1 & z_4^1 = z_4 & z_3^1 = z_3 & z_2^1 = z_2 & z_1^1 = z_1 \\ z_0^2 = 1 & z_4^2 = z_3 & z_3^2 = z_1 & z_2^2 = z_4 & z_1^2 = z_2 \\ z_0^3 = 1 & z_4^3 = z_2 & z_3^3 = z_4 & z_2^3 = z_1 & z_1^3 = z_3 \\ z_0^4 = 1 & z_4^4 = z_1 & z_3^4 = z_2 & z_2^4 = z_3 & z_1^4 = z_4 \end{bmatrix}, \tag{9}$$

where z_0 to z_4 are the 5-th roots of unity. The columns of $M^{(R)}$ are built from these roots taken in a clockwise order on the complex plane (see Fig. 2N-th roots of unity on the complex plane. This example shows the roots of unit for $N = 5$figure.2).

For two dimensional matrices (or images) f, with R rows and C columns, we define their DFT as the product:

$$F = M^{(R)}fM^{(C)}. \tag{10}$$

The inverse of $M^{(R)}$ is simply its complex conjugate times a normalization factor $1/R$:

$$I^{(R)} = M^{(R)}\frac{1}{R}M^{(R)^*},$$

where $I^{(R)}$ is the identity matrix of order R. From this, it follows that the inverse discrete Fourier transform (IDFT) may be expressed as:

$$f = \frac{1}{R} M^{(R)^*} F,$$

for the one-dimensional case, and

$$f = \frac{1}{RC} M^{(R)^*} F M^{(C)^*},$$

for the two-dimensional case. The important connection between convolution and the DFT is given by the *convolution theorem*, which establishes that the DFT of the convolution of two vectors or matrices is the element-wise product of their DFT's:

$$DFT(f \otimes h) = F. * H.$$

Let $g = f \otimes h$ be the convolution of two images f and h in the spatial domain. The value of each pixel of g depends on the values of all pixels from both f and h. In the frequency domain, the coefficients associated with a pair of frequencies (u, v) result from the product of coefficients located in the same row and column of the corresponding DFT matrices: $G_{u,v} = F_{u,v}H_{u,v}$. The convolution theorem also shows that the convolution is commutative:

$$f \otimes h = h \otimes f.$$

V. MATRIX OPERATORS

Convolution, correlation, and the Fourier transform are all linear operators. However, it is not immediately obvious from their definitions that each of these operators can be expressed as the product of a matrix by a vector. One interesting consequence of this representation and of the convolution theorem is that the DFT can be seen as a change of basis that

Fig. 4. Representing a 2D image with R rows and C columns as a 1D vector containing $R \times C$ elements (right), which can then be transformed using a matrix-vector multiplication (center). The transformed vector is then decoded back into an image with R rows and C columns.
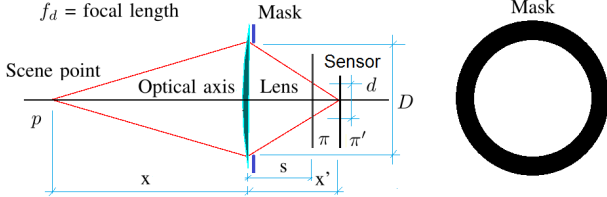
Fig. 5. Thin lens approximation. The image of a point p located at a distance x in front of the lens is formed at a distance x' behind the lens.

diagonalizes the convolution operator matrix. It also allows one to make use of a variety of tools available in linear algebra. This section briefly discusses how to obtain such matrix-vector representations.

Let V and U be C- and R-dimensional vectorial spaces, respectively. Also, let $T : V \to U$ be a linear operator that maps every vector v in V to a vector u in U. Then, there exists a matrix M, with R rows and C columns, such that

$$Mv = u = T(v); \quad \forall \, v \in V.$$

The matrix elements M_{ij} of M can be obtained applying the linear operator T to each vector v_j of the canonical basis of V, and computing the dot product of the resulting $T(v_j)$ with the vectors u_i of the canonical basis of U:

$$M_{ij} = u_i \cdot T(v_j) \tag{11}$$

There is a one-to-one correspondence between matrices and linear operators in vector spaces.

In order to apply these matrix operators to images, one must first rearrange its pixels to form a column vector v, as illustrated in Fig. 4Representing a 2D image with R rows and C columns as a 1D vector containing $R \times C$ elements (right), which can then be transformed using a matrix-vector multiplication (center). The transformed vector is then decoded back into an image with R rows and C columnsfigure.4, and then obtain the matrix M from Equation 11Matrix Operatorsequation.5.11. It is important to note that the column vectors so constructed have $R \times C$ rows and that M will have $R \times C$ rows and $R \times C$ columns. The usual definition of two-dimensional convolution, correlation, and Fourier transforms provide more compact expressions, as M tends to be huge.

Fig. 6. Motion blurr. (a) Ideal image f. (b) Trajectory mask h_{motion}, expressing the relative motion of the camera with respect to the scene objects. (c) Image containing motion blurr: $g = f \otimes h_{motion}$.

VI. IMAGE FORMATION AND CONVOLUTION

In this section we assume a scene with all the objects located at the same distance x from the camera lens. We also assume that their surfaces are Lambertian, *i.e.*, the radiance leaving a scene surface point does not vary with direction.

Fig. 5Thin lens approximation. The image of a point p located at a distance x in front of the lens is formed at a distance x' behind the lensfigure.5 shows how a scene point p located at a distance x from the camera lens is projected on the sensor. Rays leaving p go through the lens and converge to a sharp point behind the lens at a distance x' that depends both on x and on the lens focal distance f_d. The relation between x, x' and f_d is given by Gauss formula:

$$\frac{1}{x} + \frac{1}{x'} = \frac{1}{f_d} \tag{12}$$

If the sensor is located exactly at x', a sharp image of p is formed. If, however, the sensor is located at another distance $s < x'$, the rays leaving p will be spread over a region with the same shape as the aperture. The relation between the aperture diameter D and its projection d can be calculated using trigonometric relations as:

$$d = D\left(\frac{x' - s}{x'}\right) = D\left(1 - s\frac{1}{x'}\right). \tag{13}$$

From Equations 12Image Formation and Convolutionequation.6.12 and 13Image Formation and Convolutionequation.6.13, one can compute d as :

$$d = D\left(1 - s\left(\frac{1}{f_d} - \frac{1}{x}\right)\right). \tag{14}$$

Scene points located at a distance y from the optical axis will be projected on the sensor at a distance y' from the optical axis:

$$y' = y\frac{s}{x}. \tag{15}$$

The difference between an ideal all-in-focus image f and the captured one g is that in g, every point of f will appear scaled by d. Thus, for each point on the scene plane, its image will appear as a scaled version of the aperture shifted by some distance y'. As we saw in Section IIIConvolution and Correlationsection.3, this is exactly what one would obtain by convolving a scaled version of the aperture with a delta image. Using the fact that f itself may be expressed as a linear combination of deltas (Equation 5Convolution

and Correlationequation.3.5) and that convolution is a linear operator, one concludes that the image formed on the sensor can be described as the convolution of the ideal image f with a scaled version of the aperture h:

$$g = f \otimes h. \tag{16}$$

Using a similar analysis for a moving object, g will be formed by a linear combination of several *instantaneous* projections of its scaled and blurred aperture image. Thus, the blurring mask combines the aperture mask and a mask which is a scaled version of the object's trajectory projected on a plane perpendicular to the optical axis. If the shutter is opened and closed several times during the capture of g, only the portions of the object trajectory corresponding to times when the shutter is open must be considered. In short, for a moving object, Equation 16Image Formation and Convolutionequation.6.16 still holds but one should replace h by:

$$h = h_{aperture} \otimes h_{motion}. \tag{17}$$

For a real scene, the convolution kernel will vary with object distance to the camera's focal plane, as well as with its relative movement with respect to the camera. This means that, in general, one cannot use a single kernel to deconvolve an entire image.

A. Image formation and noise

The captured image g may be contaminated by noise n from multiple sources, some of which will be independent of the image content:

$$g = f \otimes h + n. \tag{18}$$

Although some sources of noise do depend on image content, following a common practice in computational photography works, we will only consider additive noise. One should be aware, however, that current technology allows the construction of cameras for which the principal source of noise is the random arrival of photons. This can be modeled by a Poisson process that depends on the image content.

VII. THE DECONVOLUTION PROBLEM

As discussed in the previous section, the image acquisition process can be modeled as the convolution of an ideal image with a kernel, which depends on the camera's aperture pattern. The captured image may be blurred if the scene objects are out of focus, or due to camera-object relative motion. A central question is whether this process is invertible. In other words, can one recover a sharp image f given a blurred picture g and knowledge about the camera's aperture used to acquire g? Another key question to many computational photography techniques is how can one design apertures (masks) that facilitate or improve the recovery of f from g. In fact, *this is the most fundamental question in coded photography*. This section addresses these important questions and introduces the notion of **deconvolution**. It also provides an intuitive introduction to various techniques commonly used in computational photography to perform deconvolution.
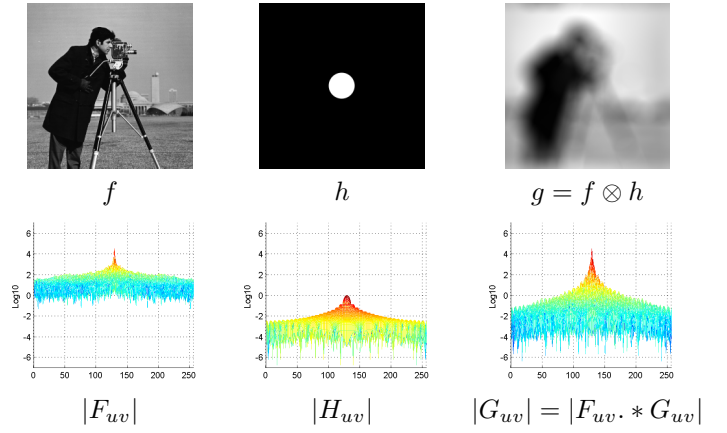

Fig. 7. The result of convolving an image with a circular aperture, represented both in the spatial and frequency domains. (left) Input image f (cameraman) and a side view of its amplitude spectrum ($|F_{u,v}|$) plotted as a surface in 3D, using a $\log_{10}$ scale. (center) Circular aperture h and a side view of its amplitude spectrum ($|H_{u,v}|$). Note the occurrence of very small values, of the order of 10^{-6}. (right) Blurred image $g = f \otimes h$, and a side view of its amplitude spectrum $|G_{u,v}|$.

The Noiseless Case: We will begin by the simple case of a convolution kernel that is constant all over the image. In this case, the captured image g will be the convolution of the ideal image f with a kernel h that models the aperture pattern plus motion blur. For now, we ignore the existence of noise. Thus,

$$g = f \otimes h.$$

In the frequency domain, this relationship can be expressed as

$$G_{u,v} = F_{u,v} H_{u,v}, \ \forall \ u, v, \tag{19}$$

where u and v are frequencies. This is illustrated for both domains in Fig. 7The result of convolving an image with a circular aperture, represented both in the spatial and frequency domains. (left) Input image f (cameraman) and a side view of its amplitude spectrum ($|F_{u,v}|$) plotted as a surface in 3D, using a $\log_{10}$ scale. (center) Circular aperture h and a side view of its amplitude spectrum ($|H_{u,v}|$). Note the occurrence of very small values, of the order of 10^{-6}. (right) Blurred image $g = f \otimes h$, and a side view of its amplitude spectrum $|G_{u,v}|$figure.7, which shows the result of convolving an image with a circular aperture. The ideal image f is shown on the left, followed by the circular aperture h, at the center, and by the resulting blurred image g, on the right. Side views of the amplitude spectra of f, h, and g (i.e., $|F_{u,v}|$, $|H_{u,v}|$, and $|G_{u,v}|$, respectively) are shown under the corresponding images, plotted as a surface in 3D, using a $\log_{10}$ scale. For $|H_{u,v}|$, one should note the existence of some very small values, in the order of 10^{-6}.

Assuming that H has no zero values, one can recover the ideal image representation F from the blurred one G and the

Fig. 8. Debluring by inverse filtering in the abscence of noise. (left) Blurred image g obtained in Fig. 7The result of convolving an image with a circular aperture, represented both in the spatial and frequency domains. (left) Input image f (cameraman) and a side view of its amplitude spectrum ($|F_{u,v}|$) plotted as a surface in 3D, using a $\log_{10}$ scale. (center) Circular aperture h and a side view of its amplitude spectrum ($|H_{u,v}|$). Note the occurrence of very small values, of the order of 10^{-6}. (right) Blurred image $g = f \otimes h$, and a side view of its amplitude spectrum $|G_{u,v}|$figure.7 as $g = f \otimes h$, and a side view of its amplitude spectrum. (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained as $F_{u,v} = G_{u,v}/H_{u,v}$ (Equation 20The Deconvolution Problemequation.7.20). $F_{u,v}$, and therefore f, can be exactly recovered.

Fig. 9. Debluring by inverse filtering in the presence of noise. (left) Blurred image g obtained as $g = f \otimes h + n$, and a side view of its amplitude spectrum. In this example, Gaussian noise with $\sigma = 0.001$ was added to the image shown in Fig. 8Debluring by inverse filtering in the absence of noise. (left) Blurred image g obtained in Fig. 7The result of convolving an image with a circular aperture, represented both in the spatial and frequency domains. (left) Input image f (cameraman) and a side view of its amplitude spectrum ($|F_{u,v}|$) plotted as a surface in 3D, using a $\log_{10}$ scale. (center) Circular aperture h and a side view of its amplitude spectrum ($|H_{u,v}|$). Note the occurrence of very small values, of the order of 10^{-6}. (right) Blurred image $g = f \otimes h$, and a side view of its amplitude spectrum $|G_{u,v}|$figure.7 as $g = f \otimes h$, and a side view of its amplitude spectrum. (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained as $F_{u,v} = G_{u,v}/H_{u,v}$ (Equation 20The Deconvolution Problemequation.7.20). $F_{u,v}$, and therefore f, can be exactly recoveredfigure.8 (left). The added noise has dominated the frequency components with low magnitude values (compare both spectra). (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained using Equation 22The Deconvolution Problemequation.7.22. As the noise dominates several frequency components, inverse filtering becomes ill posed, and the recovered image $\hat{f}$ is essentially noise.

blurring kernel H as

$$F_{u,v} = G_{u,v} / H_{u,v}. \qquad (20)$$

This method of obtaining F is known as **inverse filtering**, as one searches, in the frequency domain, for a sharp image f that convolved with h will give the blurred image g. A similar problem involves finding H from G and F. Fig. 8Debluring by inverse filtering in the abscence of noise. (left) Blurred image g obtained in Fig. 7The result of convolving an image with a circular aperture, represented both in the spatial and frequency domains. (left) Input image f (cameraman) and a side view of its amplitude spectrum ($|F_{u,v}|$) plotted as a surface in 3D, using a $\log_{10}$ scale. (center) Circular aperture h and a side view of its amplitude spectrum ($|H_{u,v}|$). Note the occurrence of very small values, of the order of 10^{-6}. (right) Blurred image $g = f \otimes h$, and a side view of its amplitude spectrum $|G_{u,v}|$figure.7 as $g = f \otimes h$, and a side view of its amplitude spectrum. (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained as $F_{u,v} = G_{u,v}/H_{u,v}$ (Equation 20The Deconvolution Problemequation.7.20). $F_{u,v}$, and therefore f, can be exactly recoveredfigure.8 illustrates the use of inverse filtering for the example shown in Fig. 7The result of convolving an image with a circular aperture, represented both in the spatial and frequency domains. (left) Input image f (cameraman) and a side view of its amplitude spectrum ($|F_{u,v}|$) plotted as a surface in 3D, using a $\log_{10}$ scale. (center) Circular aperture h and a side view of its amplitude spectrum ($|H_{u,v}|$). Note the occurrence of very small values, of the order of 10^{-6}. (right) Blurred image $g = f \otimes h$, and a side view of its amplitude spectrum $|G_{u,v}|$figure.7. Since we are assuming a noiseless acquisition process, f can be exactly recovered (Fig. 8Debluring by inverse filtering in the absence of noise. (left) Blurred image g obtained in Fig. 7The result of convolving an image with a circular aperture, represented both in the spatial and frequency domains. (left) Input image f (cameraman) and a side view of its amplitude spectrum ($|F_{u,v}|$) plotted as a surface in 3D, using a $\log_{10}$ scale. (center) Circular aperture h and a side view of its amplitude spectrum ($|H_{u,v}|$). Note the occurrence of very small values, of the order of 10^{-6}. (right) Blurred image $g = f \otimes h$, and a side view of its amplitude spectrum $|G_{u,v}|$figure.7 as $g = f \otimes h$, and a side view of its amplitude spectrum. (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained as $F_{u,v} = G_{u,v}/H_{u,v}$ (Equation 20The Deconvolution Problemequation.7.20). $F_{u,v}$, and therefore f, can be exactly recoveredfigure.8, right).

In case H contains zero values, it is not possible to recover F in general, since some frequencies from f might have been lost. This can be easily understood by considering the element-wise multiplication that takes place in the frequency domain (Equation 19The Deconvolution Problemequation.7.19). If some frequencies are lost, the problem is said to be **ill posed**, meaning that it has either none or multiple solutions.

For instance, given one solution, any other $\hat{F}$ differing from the first one only in the frequencies for which $H_{u,v}$ is zero will also be a plausible solution.

If H has no zeros, but instead it contains very small values associated to some frequencies (u, v), the exact solution exists and is unique, but it may be **ill conditioned**. In this case, the solution becomes very sensitive to small changes in the values of the corresponding frequencies in G. This situation is of serious practical concern and will be discussed next, when we analyze the noisy case.

The Noisy Case: In the presence of (additive) noise, the captured image g may be modeled as the convolution of the ideal image f with a blurring kernel h, plus some noise n:

$$g = f \otimes h + n.$$

In the frequency domain this can be expressed as

$$G_{u,v} = F_{u,v} H_{u,v} + N_{u,v}, \ \forall \ u, v, \tag{21}$$

where $N_{u,v}$ are the (u, v) DFT components of the noise. Assuming that H has no zeros, applying inverse filtering to Equation 21The Deconvolution Problemequation.7.21 produces

$$F_{u,v} = \frac{G_{u,v}}{H_{u,v}} - \frac{N_{u,v}}{H_{u,v}}. \tag{22}$$

The second term on the right hand side of Equation 22The Deconvolution Problemequation.7.22 is the *deconvolved noise*. Whenever $|H_{u,v}|$ is small and $|N_{u,v}| >> |G_{u,v}|$, the deconvolved noise will dominate the reconstruction process at frequencies (u, v). As a result, the recovered image $\hat{f}$ might not even resemble f. This is illustrated in Fig. 9Debluring by inverse filtering in the presence of noise. (left) Blurred image g obtained as $g = f \otimes h + n$, and a side view of its amplitude spectrum. In this example, Gaussian noise with $\sigma = 0.001$ was added to the image shown in Fig. 8Debluring by inverse filtering in the abscence of noise. (left) Blurred image g obtained in Fig. 7The result of convolving an image with a circular aperture, represented both in the spatial and frequency domains. (left) Input image f (cameraman) and a side view of its amplitude spectrum ($|F_{u,v}|$) plotted as a surface in 3D, using a $\log_{10}$ scale. (center) Circular aperture h and a side view of its amplitude spectrum ($|H_{u,v}|$). Note the occurrence of very small values, of the order of 10^{-6}. (right) Blurred image $g = f \otimes h$, and a side view of its amplitude spectrum $|G_{u,v}|$figure.7 as $g = f \otimes h$, and a side view of its amplitude spectrum. (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained as $F_{u,v} = G_{u,v}/H_{u,v}$ (Equation 20The Deconvolution Problemequation.7.20). $F_{u,v}$, and therefore f, can be exactly recoveredfigure.8 (left). The added noise has dominated the frequency components with low magnitude values (compare both spectra). (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained using Equation 22The Deconvolution Problemequation.7.22.

As the noise dominates several frequency components, inverse filtering becomes ill posed, and the recovered image $\hat{f}$ is essentially noisefigure.9, where Gaussian noise with $\sigma = 0.001$ has been added to the image shown in Fig. 8Debluring by inverse filtering in the abscence of noise. (left) Blurred image g obtained in Fig. 7The result of convolving an image with a circular aperture, represented both in the spatial and frequency domains. (left) Input image f (cameraman) and a side view of its amplitude spectrum ($|F_{u,v}|$) plotted as a surface in 3D, using a $\log_{10}$ scale. (center) Circular aperture h and a side view of its amplitude spectrum ($|H_{u,v}|$). Note the occurrence of very small values, of the order of 10^{-6}. (right) Blurred image $g = f \otimes h$, and a side view of its amplitude spectrum $|G_{u,v}|$figure.7 as $g = f \otimes h$, and a side view of its amplitude spectrum. (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained as $F_{u,v} = G_{u,v}/H_{u,v}$ (Equation 20The Deconvolution Problemequation.7.20). $F_{u,v}$, and therefore f, can be exactly recoveredfigure.8 (left). Although Figs. 8Debluring by inverse filtering in the abscence of noise. (left) Blurred image g obtained in Fig. 7The result of convolving an image with a circular aperture, represented both in the spatial and frequency domains. (left) Input image f (cameraman) and a side view of its amplitude spectrum ($|F_{u,v}|$) plotted as a surface in 3D, using a $\log_{10}$ scale. (center) Circular aperture h and a side view of its amplitude spectrum ($|H_{u,v}|$). Note the occurrence of very small values, of the order of 10^{-6}. (right) Blurred image $g = f \otimes h$, and a side view of its amplitude spectrum $|G_{u,v}|$figure.7 as $g = f \otimes h$, and a side view of its amplitude spectrum. (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained as $F_{u,v} = G_{u,v}/H_{u,v}$ (Equation 20The Deconvolution Problemequation.7.20). $F_{u,v}$, and therefore f, can be exactly recoveredfigure.8 (top left) and 9Debluring by inverse filtering in the presence of noise. (left) Blurred image g obtained as $g = f \otimes h + n$, and a side view of its amplitude spectrum. In this example, Gaussian noise with $\sigma = 0.001$ was added to the image shown in Fig. 8Debluring by inverse filtering in the abscence of noise. (left) Blurred image g obtained in Fig. 7The result of convolving an image with a circular aperture, represented both in the spatial and frequency domains. (left) Input image f (cameraman) and a side view of its amplitude spectrum ($|F_{u,v}|$) plotted as a surface in 3D, using a $\log_{10}$ scale. (center) Circular aperture h and a side view of its amplitude spectrum ($|H_{u,v}|$). Note the occurrence of very small values, of the order of 10^{-6}. (right) Blurred image $g = f \otimes h$, and a side view of its amplitude spectrum $|G_{u,v}|$figure.7 as $g = f \otimes h$, and a side view of its amplitude spectrum. (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained as $F_{u,v} = G_{u,v}/H_{u,v}$ (Equation 20The Deconvolution Problemequation.7.20). $F_{u,v}$, and therefore f, can be exactly recoveredfigure.8 (left). The added noise has dominated the frequency components with

low magnitude values (compare both spectra). (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained using Equation 22The Deconvolution Problemequation.7.22. As the noise dominates several frequency components, inverse filtering becomes ill posed, and the recovered image $\hat{f}$ is essentially noisefigure.9 (top left) are visually indistinguishable, the added noise has dominated the frequency components with low magnitude values in f. This can be verified by simply comparing both spectra. As a result, the image $\hat{f}$ obtained using Equation 22The Deconvolution Problemequation.7.22 is dominated by noise (Fig. 9Deblurring by inverse filtering in the presence of noise. (left) Blurred image g obtained as $g = f \otimes h + n$, and a side view of its amplitude spectrum. In this example, Gaussian noise with $\sigma = 0.001$ was added to the image shown in Fig. 8Deblurring by inverse filtering in the abscence of noise. (left) Blurred image g obtained in Fig. 7The result of convolving an image with a circular aperture, represented both in the spatial and frequency domains. (left) Input image f (cameraman) and a side view of its amplitude spectrum ($|F_{u,v}|$) plotted as a surface in 3D, using a $\log_{10}$ scale. (center) Circular aperture h and a side view of its amplitude spectrum ($|H_{u,v}|$). Note the occurrence of very small values, of the order of 10^{-6}. (right) Blurred image $g = f \otimes h$, and a side view of its amplitude spectrum $|G_{u,v}|$figure.7 as $g = f \otimes h$, and a side view of its amplitude spectrum. (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained as $F_{u,v} = G_{u,v}/H_{u,v}$ (Equation 20The Deconvolution Problemequation.7.20). $F_{u,v}$, and therefore f, can be exactly recoveredfigure.8 (left). The added noise has dominated the frequency components with low magnitude values (compare both spectra). (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained using Equation 22The Deconvolution Problemequation.7.22. As the noise dominates several frequency components, inverse filtering becomes ill posed, and the recovered image $\hat{f}$ is essentially noisefigure.9, right), characterizing a situation of ill conditioning.

Since noise is inherent to any image-capture process, it should be clear that more robust techniques than just inverse filtering are needed to perform deconvolution. This is the subject of following section.

VIII. Deconvolution techniques

This section discusses four popular approaches to perform deconvolution while trying to avoid noise amplification. First, we will describe the *autocorrelation method* used in X-ray and γ-ray astronomy. Then, we will review the *Wiener filter*, the *Richardson-Lucy algorithm*, and conclude the section describing techniques based on *image priors*.

Autocorrelation Methods: Coded-aperture techniques were introduced by Dicke [20] and by Ables [21] to image X- and γ-ray radiation in astronomy, overcoming the limitation that such rays cannot be refracted using regular lenses. While a pinhole camera can be used to capture images from high-energy sources, it is not light efficient. The light-efficiency problem is resolved with the use of a (coded-aperture) mask containing a large number of pinholes. However, such a solution results in multiple copies of the target image being projected on the sensor, which will require deconvolution. This situation is similar to superimposing the convolutions of a desired image f with a series of delta images, each one representing a pinhole, as discussed in Section IIIConvolution and Correlationsection.3.

The central idea behind autocorrelation methods is to find a mask whose autocorrelation (*i.e.*, its correlation with itself) results in an approximation of a delta image. Such masks consist of equal-sized holes on an opaque plate. They can be modeled as a binary matrix with R rows and C columns, having k cells with value 1. Assuming that the value 1 represents a hole, the mask transparency is defined as $T = \frac{k}{R \times C}$. In 1971, Golay [22] proposed a set of patterns called *Non-Redundant Arrays* (NRA) whose autocorrelations consist of a delta at the origin with a peak of height k (*i.e.*, $k\, d^{(0,0)}$), plus some few non zero elements of value $1 << k$. Other patterns known as *Uniform Redundant Arrays* (URA) [23], [24] have been found with autocorrelations producing a delta at the origin plus a matrix with constant element values equal to $\lambda \mathcal{U}$:

$$h \circ h = k\, d^{(0,0)} + \lambda \mathcal{U},$$

where λ is a scalar value, $\mathcal{U}$ is a matrix with all ones, and $k\, d^{(0,0)}$ is a delta image with peak magnitude k at the origin. A random pinhole mask approximately fulfills this condition. For example, a mask h of size $R \times C$ with half of its elements zeros and half ones, which are randomly spaced will have an autocorrelation:

$$(h \circ h)_{r,c} = \begin{cases} (R \times C)/2, & \text{if } r = 0 \text{ and } c = 0; \\ \approx (R \times C)/4, & \text{otherwise.} \end{cases}$$

Random pinhole patterns are described in [20] and [21]. Deconvolution in the presence of noise can then be expressed as

$$\begin{aligned} g \circ h &= (f \otimes h + n) \circ h \\ &= f \otimes h \circ h + n \circ h \\ &= f \otimes (k\, d^{(0,0)} + \lambda \mathcal{U}) + n \circ h \\ &= k\, f + \lambda (f \otimes \mathcal{U}) + n \circ h \\ &= k\, f + k\, \lambda\, \overline{f} + n \circ h, \end{aligned}$$

where $\overline{f}$ is f's average value. Since h is composed of zeros and ones, this approach does not suffer from the noise amplification problem that plagues inverse filtering. The reconstructed image $\hat{f}$ is defined as:

$$\widehat{f} = \frac{g \circ h}{k}. \tag{23}$$

It will differ from the ideal f by an approximately constant offset.

A variation of this restoration method known as **balanced autocorrelation** uses a modified mask $\widehat{h}$ to perform correlation:

$$\widehat{h}_{ij} = \begin{cases} 1, & \text{if } h_{ij} = 1; \\ T/(T\text{-}1), & \text{if } h_{ij} = 0, \end{cases}$$

where T is the mask transparency defined above. in this way the sum of all the elements of $\widehat{h}_{ij}$ is zero and any DC background is removed from the reconstructed image [23].

Modified Uniform Redundant Arrays (MURA) were introduced in [25] and are similar to URAs. To deconvolve a MURA, the following mask $\widehat{h}$ is used:

$$\widehat{h}_{ij} = \begin{cases} 1, & \text{if } i + j = 0; \\ 1, & \text{if } h_{ij} = 1 \text{ and } (i + j) \neq 0; \\ -1, & \text{if } h_{ij} = 0 \text{ and } (i + j) \neq 0. \end{cases}$$

Fig. 10MURA mask (a) Input image. (b) MURA mask. (c) Convolution of input image with the MURA mask plus Gaussian noise ($\sigma = 0.001$). (d) Restored imagefigure.10 shows an image convolved with a MURA mask, and restored after adding Gaussian noise.
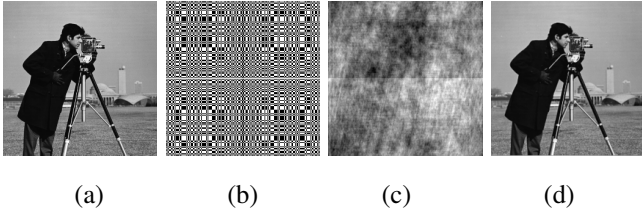
| (a) | (b) | (c) | (d) |

Fig. 10. MURA mask (a) Input image. (b) MURA mask. (c) Convolution of input image with the MURA mask plus Gaussian noise ($\sigma = 0.001$). (d) Restored image.

Wiener Filter: Comparing the amplitude spectra shown in Figs. 8Debluring by inverse filtering in the abscence of noise. (left) Blurred image g obtained in Fig. 7The result of convolving an image with a circular aperture, represented both in the spatial and frequency domains. (left) Input image f (cameraman) and a side view of its amplitude spectrum ($|F_{u,v}|$) plotted as a surface in 3D, using a $\log_{10}$ scale. (center) Circular aperture h and a side view of its amplitude spectrum ($|H_{u,v}|$). Note the occurrence of very small values, of the order of 10^{-6}. (right) Blurred image $g = f \otimes h$, and a side view of its amplitude spectrum $|G_{u,v}|$figure.7 as $g = f \otimes h$, and a side view of its amplitude spectrum. (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained as $F_{u,v} = G_{u,v}/H_{u,v}$ (Equation 20The Deconvolution Problemequation.7.20). $F_{u,v}$, and therefore f, can be exactly recoveredfigure.8 (left) (noiseless case) and 9Debluring by inverse filtering in the presence of noise. (left) Blurred image g obtained as $g = f \otimes h + n$, and a side view of its amplitude spectrum. In this example, Gaussian noise with $\sigma = 0.001$ was added to the image shown in Fig. 8Debluring by inverse filtering in the abscence of noise. (left) Blurred image g obtained in Fig. 7The result of convolving an image

with a circular aperture, represented both in the spatial and frequency domains. (left) Input image f (cameraman) and a side view of its amplitude spectrum ($|F_{u,v}|$) plotted as a surface in 3D, using a $\log_{10}$ scale. (center) Circular aperture h and a side view of its amplitude spectrum ($|H_{u,v}|$). Note the occurrence of very small values, of the order of 10^{-6}. (right) Blurred image $g = f \otimes h$, and a side view of its amplitude spectrum $|G_{u,v}|$figure.7 as $g = f \otimes h$, and a side view of its amplitude spectrum. (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained as $F_{u,v} = G_{u,v}/H_{u,v}$ (Equation 20The Deconvolution Problemequation.7.20). $F_{u,v}$, and therefore f, can be exactly recoveredfigure.8 (left). The added noise has dominated the frequency components with low magnitude values (compare both spectra). (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained using Equation 22The Deconvolution Problemequation.7.22. As the noise dominates several frequency components, inverse filtering becomes ill posed, and the recovered image $\hat{f}$ is essentially noisefigure.9 (left) (noisy case), one observes that lots of frequencies in the spectrum of Fig. 9Debluring by inverse filtering in the presence of noise. (left) Blurred image g obtained as $g = f \otimes h + n$, and a side view of its amplitude spectrum. In this example, Gaussian noise with $\sigma = 0.001$ was added to the image shown in Fig. 8Debluring by inverse filtering in the abscence of noise. (left) Blurred image g obtained in Fig. 7The result of convolving an image with a circular aperture, represented both in the spatial and frequency domains. (left) Input image f (cameraman) and a side view of its amplitude spectrum ($|F_{u,v}|$) plotted as a surface in 3D, using a $\log_{10}$ scale. (center) Circular aperture h and a side view of its amplitude spectrum ($|H_{u,v}|$). Note the occurrence of very small values, of the order of 10^{-6}. (right) Blurred image $g = f \otimes h$, and a side view of its amplitude spectrum $|G_{u,v}|$figure.7 as $g = f \otimes h$, and a side view of its amplitude spectrum. (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained as $F_{u,v} = G_{u,v}/H_{u,v}$ (Equation 20The Deconvolution Problemequation.7.20). $F_{u,v}$, and therefore f, can be exactly recoveredfigure.8 (left). The added noise has dominated the frequency components with low magnitude values (compare both spectra). (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained using Equation 22The Deconvolution Problemequation.7.22. As the noise dominates several frequency components, inverse filtering becomes ill posed, and the recovered image $\hat{f}$ is essentially noisefigure.9 (left) become dominated by noise, which has *erased* the original signal information. Thus, to avoid noise amplification, during deconvolution one can selectively ignore the information associated with such frequencies. One such approach is called Wiener filtering. It attenuates the noisy

frequencies before applying a conventional inverse filter:

$$\widehat{F}_{u,v} = \frac{1}{H_{u,v}} \left(B_{u,v} \, G_{u,v} \right), \qquad (24)$$

where $\widehat{F}$, G, and H are, respectively, the spectra of the restored image $\hat{f}$, of the observed image g, and of the blurring filter h. B is given by:

$$B_{u,v} = \frac{|H_{u,v}|^2}{|H_{u,v}|^2 + \frac{|N_{u,v}|^2}{|F_{u,v}|^2}}. \qquad (25)$$

Note that in the absence of noise, $B = 1$, and the Wiener filter reduces to an inverse filter.

Fig. 9Debluring by inverse filtering in the presence of noise. (left) Blurred image g obtained as $g = f \otimes h + n$, and a side view of its amplitude spectrum. In this example, Gaussian noise with $\sigma = 0.001$ was added to the image shown in Fig. 8Debluring by inverse filtering in the abscence of noise. (left) Blurred image g obtained in Fig. 7The result of convolving an image with a circular aperture, represented both in the spatial and frequency domains. (left) Input image f (cameraman) and a side view of its amplitude spectrum ($|F_{u,v}|$) plotted as a surface in 3D, using a $\log_{10}$ scale. (center) Circular aperture h and a side view of its amplitude spectrum ($|H_{u,v}|$). Note the occurrence of very small values, of the order of 10^{-6}. (right) Blurred image $g = f \otimes h$, and a side view of its amplitude spectrum $|G_{u,v}|$figure.7 as $g = f \otimes h$, and a side view of its amplitude spectrum. (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained as $F_{u,v} = G_{u,v}/H_{u,v}$ (Equation 20The Deconvolution Problemequation.7.20). $F_{u,v}$, and therefore f, can be exactly recoveredfigure.8 (left). The added noise has dominated the frequency components with low magnitude values (compare both spectra). (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained using Equation 22The Deconvolution Problemequation.7.22. As the noise dominates several frequency components, inverse filtering becomes ill posed, and the recovered image $\hat{f}$ is essentially noisefigure.9 (right) and Fig. 11Comparison of several deconvolution techniques for image deblurring in the presence of additive noise. (a) Ideal image f. (b) Blurred image with noise $g = f \otimes h + n$ (same as shown in Fig 9Debluring by inverse filtering in the presence of noise. (left) Blurred image g obtained as $g = f \otimes h + n$, and a side view of its amplitude spectrum. In this example, Gaussian noise with $\sigma = 0.001$ was added to the image shown in Fig. 8Debluring by inverse filtering in the abscence of noise. (left) Blurred image g obtained in Fig. 7The result of convolving an image with a circular aperture, represented both in the spatial and frequency domains. (left) Input image f (cameraman) and a side view of its amplitude spectrum ($|F_{u,v}|$) plotted as a surface in 3D, using a $\log_{10}$ scale. (center) Circular aperture h and a side view of its amplitude spectrum ($|H_{u,v}|$). Note the occurrence of very small values, of the order of 10^{-6}.

(right) Blurred image $g = f \otimes h$, and a side view of its amplitude spectrum $|G_{u,v}|$figure.7 as $g = f \otimes h$, and a side view of its amplitude spectrum. (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained as $F_{u,v} = G_{u,v}/H_{u,v}$ (Equation 20The Deconvolution Problemequation.7.20). $F_{u,v}$, and therefore f, can be exactly recoveredfigure.8 (left). The added noise has dominated the frequency components with low magnitude values (compare both spectra). (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained using Equation 22The Deconvolution Problemequation.7.22. As the noise dominates several frequency components, inverse filtering becomes ill posed, and the recovered image $\hat{f}$ is essentially noisefigure.9, left). Images (c) to (f) show the results obtained by deconvolving (b) using: (c) Wiener filtering; (d) the Richardson-Lucy algorithm; (e) image priors with Gaussian prior ($\beta = 2$); (f) image priors with a sparse prior ($\beta = 0.8$)figure.11 (c) show the results produced by inverse and Wiener filtering, respectively. Note that as several of the original frequencies in f got dominated by noise, Wiener filtering cannot reconstruct f exactly. However, by selectively modulating the contribution of each frequency according to its signal-to-noise ratio (SNR), the Wiener filtering can obtain a significantly better result. Fig. 12Plot of Wiener's $B_{u,v}$ function (Equation 25Deconvolution techniquesequation.8.25) for the circular aperture example shown in Fig. 7The result of convolving an image with a circular aperture, represented both in the spatial and frequency domains. (left) Input image f (cameraman) and a side view of its amplitude spectrum ($|F_{u,v}|$) plotted as a surface in 3D, using a $\log_{10}$ scale. (center) Circular aperture h and a side view of its amplitude spectrum ($|H_{u,v}|$). Note the occurrence of very small values, of the order of 10^{-6}. (right) Blurred image $g = f \otimes h$, and a side view of its amplitude spectrum $|G_{u,v}|$figure.7 in the presence of Gaussian additive noise ($\sigma = 0.001$). It resembles a selective low-pass filterfigure.12 plots the values of $B_{u,v}$ for the circular aperture example shown in Fig. 7The result of convolving an image with a circular aperture, represented both in the spatial and frequency domains. (left) Input image f (cameraman) and a side view of its amplitude spectrum ($|F_{u,v}|$) plotted as a surface in 3D, using a $\log_{10}$ scale. (center) Circular aperture h and a side view of its amplitude spectrum ($|H_{u,v}|$). Note the occurrence of very small values, of the order of 10^{-6}. (right) Blurred image $g = f \otimes h$, and a side view of its amplitude spectrum $|G_{u,v}|$figure.7 in the presence of Gaussian additive noise ($\sigma = 0.001$). Note that for this example B resembles a *selective* low-pass filter.

The solution obtained with the Wiener filter minimizes the mean-square reconstruction error among those $\hat{f}$ that can be obtained from g using a linear operator. For this to be true, the noise and the image must be uncorrelated and the noise must have zero mean.

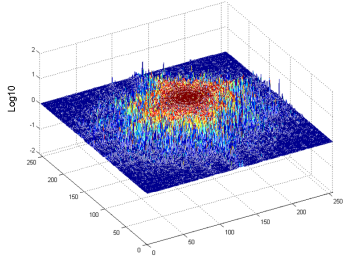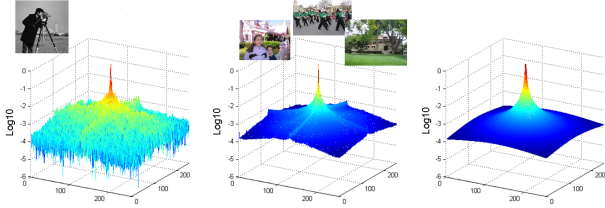One should note that the filter formulation (Equations 24Deconvolution techniquesequation.8.24 and 25Deconvolu-

Fig. 12. Plot of Wiener's $B_{u,v}$ function (Equation 25Deconvolution techniquesequation.8.25) for the circular aperture example shown in Fig. 7The result of convolving an image with a circular aperture, represented both in the spatial and frequency domains. (left) Input image f (cameraman) and a side view of its amplitude spectrum ($|F_{u,v}|$) plotted as a surface in 3D, using a $\log_{10}$ scale. (center) Circular aperture h and a side view of its amplitude spectrum ($|H_{u,v}|$). Note the occurrence of very small values, of the order of 10^{-6}. (right) Blurred image $g = f \otimes h$, and a side view of its amplitude spectrum $|G_{u,v}|$figure.7 in the presence of Gaussian additive noise ($\sigma = 0.001$). It resembles a selective low-pass filter.

Fig. 13. The $1/f$ law for the frequency spectrum of natural images. (left) Log of the amplitude spectrum of the cameraman image: $log|F_{u,v}|$. (center) Log of the average amplitude spectrum over over several natural images. (right) Plot of Log of $1/f^{1.7}$ function. All these plots have similar shapes.

tion techniquesequation.8.25) requires the power spectra of the noise, $|N(u,v)|^2$, and of the ideal image, $|F(u,v)|^2$. Since these quantities are not known, in practice one needs to provide approximations to both. A commonly used approximation for Gaussian noise is to assume a noise power distribution constant and equal to the noise variance σ^2. For the case of $|F(u,v)|^2$, for instance, Zhou et. al. [13] suggest using an approximation computed as the average over a set of *natural images*. Fig. 13The $1/f$ law for the frequency spectrum of natural images. (left) Log of the amplitude spectrum of the cameraman image: $log|F_{u,v}|$. (center) Log of the average amplitude spectrum over over several natural images. (right) Plot of Log of $1/f^{1.7}$ function. All these plots have similar shapesfigure.13 illustrates the idea: on average, the magnitude of the frequency spectra of natural images seem to follow a simple $1/f$ relation. A discussion on natural image models can be found in [26].

Richardson-Lucy: The Richardson-Lucy algorithm [27], [28],[29],[30],[31] is an iterative deconvolution approach derived by applying the Bayes' theorem to a probabilistic model of images. Using the matrix representation for linear operators presented in Section VMatrix Operatorssection.5, the image formation equation (Equation 18Image formation and

noiseequation.6.18) can be rewritten as

$$g_i = \sum_{j=0}^{m-1} h_{ij}\, f_j + n_i,$$

where g, f, and n are the vector representations of images g, and f, and the noise n, respectively. h is the matrix representation of the kernel h. Thus, for any given row i of h, its elements h_{ij} are the weights defining the contributions of pixels in the ideal image f to g_i, the i-th pixel (element) in g. The algorithm begins by initializing an approximation to f:

$$\widehat{f}_j^{(1)} = 1, \forall j.$$

Then, it iteratively computes successive approximations to f as:

$$\widehat{f}_j^{(t+1)} = \widehat{f}_j^{(t)} \sum_i \frac{g_i\, h_{ij}}{\sum_k h_{ik}\, \widehat{f}_k^{(t)}} = \widehat{f}_j^{(t)} \sum_i g_i \frac{h_{ij}}{\sum_k h_{ik}\, \widehat{f}_k^{(t)}}.$$

The superscript $(t+1)$ indicates the result of the t-th iteration. As the algorithm converges slowly, acceleration techniques have been developed [32]. Fig.11Comparison of several deconvolution techniques for image deblurring in the presence of additive noise. (a) Ideal image f. (b) Blurred image with noise $g = f \otimes h + n$ (same as shown in Fig 9Debluring by inverse filtering in the presence of noise. (left) Blurred image g obtained as $g = f \otimes h + n$, and a side view of its amplitude spectrum. In this example, Gaussian noise with $\sigma = 0.001$ was added to the image shown in Fig. 8Debluring by inverse filtering in the abscence of noise. (left) Blurred image g obtained in Fig. 7The result of convolving an image with a circular aperture, represented both in the spatial and frequency domains. (left) Input image f (cameraman) and a side view of its amplitude spectrum ($|F_{u,v}|$) plotted as a surface in 3D, using a $\log_{10}$ scale. (center) Circular aperture h and a side view of its amplitude spectrum ($|H_{u,v}|$). Note the occurrence of very small values, of the order of 10^{-6}. (right) Blurred image $g = f \otimes h$, and a side view of its amplitude spectrum $|G_{u,v}|$figure.7 as $g = f \otimes h$, and a side view of its amplitude spectrum. (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained as $F_{u,v} = G_{u,v}/H_{u,v}$ (Equation 20The Deconvolution Problemequation.7.20). $F_{u,v}$, and therefore f, can be exactly recoveredfigure.8 (left). The added noise has dominated the frequency components with low magnitude values (compare both spectra). (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained using Equation 22The Deconvolution Problemequation.7.22. As the noise dominates several frequency components, inverse filtering becomes ill posed, and the recovered image $\hat{f}$ is essentially noisefigure.9, left). Images (c) to (f) show the results obtained by deconvolving (b) using: (c) Wiener filtering; (d) the Richardson-Lucy algorithm; (e) image priors with Gaussian prior ($\beta = 2$); (f) image priors with a sparse prior
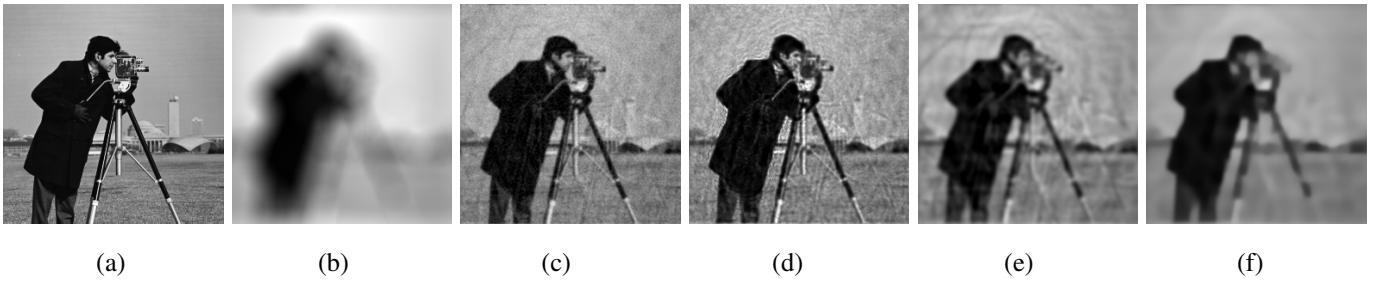
(a)	(b)	(c)	(d)	(e)	(f)

Fig. 11. Comparison of several deconvolution techniques for image deblurring in the presence of additive noise. (a) Ideal image f. (b) Blurred image with noise $g = f \otimes h + n$ (same as shown in Fig 9Debluring by inverse filtering in the presence of noise. (left) Blurred image g obtained as $g = f \otimes h + n$, and a side view of its amplitude spectrum. In this example, Gaussian noise with $\sigma = 0.001$ was added to the image shown in Fig. 8Debluring by inverse filtering in the abscence of noise. (left) Blurred image g obtained in Fig. 7The result of convolving an image with a circular aperture, represented both in the spatial and frequency domains. (left) Input image f (cameraman) and a side view of its amplitude spectrum ($|F_{u,v}|$) plotted as a surface in 3D, using a $\log_{10}$ scale. (center) Circular aperture h and a side view of its amplitude spectrum ($|H_{u,v}|$). Note the occurrence of very small values, of the order of 10^{-6}. (right) Blurred image $g = f \otimes h$, and a side view of its amplitude spectrum $|G_{u,v}|$figure.7 as $g = f \otimes h$, and a side view of its amplitude spectrum. (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained as $F_{u,v} = G_{u,v}/H_{u,v}$ (Equation 20The Deconvolution Problemequation.7.20). $F_{u,v}$, and therefore f, can be exactly recoveredfigure.8 (left). The added noise has dominated the frequency components with low magnitude values (compare both spectra). (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained using Equation 22The Deconvolution Problemequation.7.22. As the noise dominates several frequency components, inverse filtering becomes ill posed, and the recovered image $\hat{f}$ is essentially noisefigure.9, left). Images (c) to (f) show the results obtained by deconvolving (b) using: (c) Wiener filtering; (d) the Richardson-Lucy algorithm; (e) image priors with Gaussian prior ($\beta = 2$); (f) image priors with a sparse prior ($\beta = 0.8$)

($\beta = 0.8$)figure.11 (d) illustrates the result of the Richardson-Lucy algorithm used to deblur the noisy image shown in Fig.11Comparison of several deconvolution techniques for image deblurring in the presence of additive noise. (a) Ideal image f. (b) Blurred image with noise $g = f \otimes h + n$ (same as shown in Fig 9Debluring by inverse filtering in the presence of noise. (left) Blurred image g obtained as $g = f \otimes h + n$, and a side view of its amplitude spectrum. In this example, Gaussian noise with $\sigma = 0.001$ was added to the image shown in Fig. 8Debluring by inverse filtering in the abscence of noise. (left) Blurred image g obtained in Fig. 7The result of convolving an image with a circular aperture, represented both in the spatial and frequency domains. (left) Input image f (cameraman) and a side view of its amplitude spectrum ($|F_{u,v}|$) plotted as a surface in 3D, using a $\log_{10}$ scale. (center) Circular aperture h and a side view of its amplitude spectrum ($|H_{u,v}|$). Note the occurrence of very small values, of the order of 10^{-6}. (right) Blurred image $g = f \otimes h$, and a side view of its amplitude spectrum $|G_{u,v}|$figure.7 as $g = f \otimes h$, and a side view of its amplitude spectrum. (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained as $F_{u,v} = G_{u,v}/H_{u,v}$ (Equation 20The Deconvolution Problemequation.7.20). $F_{u,v}$, and therefore f, can be exactly recoveredfigure.8 (left). The added noise has dominated the frequency components with low magnitude values (compare both spectra). (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained using Equation 22The Deconvolution Problemequation.7.22. As the noise dominates several frequency components, inverse filtering becomes ill posed, and the recovered image $\hat{f}$ is essentially noisefigure.9, left). Images (c) to (f) show the results obtained by deconvolving (b) using: (c) Wiener filtering; (d) the Richardson-Lucy algorithm; (e) image priors with Gaussian prior ($\beta = 2$); (f) image priors with a sparse prior ($\beta = 0.8$)figure.11 (b).

Image-Prior-based Techniques: Other deconvolution methods involve the use of regularization techniques. The idea behind these approaches is to obtain a sharp image $\hat{f}$ that better *explains* the acquired image g under convolution with the kernel h and, at the same time, qualifies as a *natural* image. A camera's finite aperture may cause out-of-focus blurring. In this case, the kernel h that models such an aperture can be seen as a low-pass filter. As a result, many noisy variations of $\hat{f}$ will produce good approximations to g when convolved with h. To avoid the recovery of a noisy $\hat{f}$, a second term is added to the objective function to penalize such a selection. Since image derivatives can be used to detect noise, let $D^{[0]}$ to $D^{[3]}$ be matrix operators that compute the first and second image derivatives in the horizontal and vertical directions, respectively. Using the same matrix and vector representations used in the description of the Richardson-Lucy algorithm, one can express the following objective function to be minimized:

$$\delta = \sum_i \sum_k (h_{ik}\,\widehat{f}_k - g_i)^2 + \sum_i \sum_{q=0}^{3} \lambda_q (\sum_k D_{ik}^{[q]}\,\widehat{f}_k)^\beta, \quad (26)$$

where λ_q weights the contribution of $D^{[q]}$, and β is a parameter that adjusts the model to the distribution of derivatives in natural images. The second term penalizes the choice of noisy images. Differentiating Equation 26Deconvolution techniquesequation.8.26 with respect to each desired pixel $\widehat{f}_m$, and requiring the resulting expressions to be zero, one gets:

$$[h^T g]_m = [h^T h\,\widehat{f}]_m + \frac{\beta}{2} \sum_{q=0}^{3} \lambda_q (\sum_i (D_{mi}^{[q]})^T [(D^{[q]}\,\widehat{f})_i]^{\beta-1}).$$

Here, the superscript ^{T} indicates matrix transpose, and the op-

erator $[\square]_m$ returns the m-th element of the vector represented by the symbol $\square$. When the parameter β takes the value 2, it is said that a Gaussian prior is being used. In this case, the solution that minimizes δ is a linear system:

$$[\boldsymbol{h}^T \boldsymbol{g}]_m = [\boldsymbol{h}^T \boldsymbol{h}\, \widehat{\boldsymbol{f}}]_m + \sum_{q=0}^{3} \lambda_q (\sum_i (D_{mi}^{[q]})^T [(D^{[q]}\, \widehat{\boldsymbol{f}})_i]).$$

For $\beta \neq 2$, for example $\beta = 0.8$, it is said that a sparse prior is being used, and the resulting system is not linear, but produces better results (*i.e.*, less reconstruction artifacts):

$$[\boldsymbol{h}^T \boldsymbol{g}]_m = [\boldsymbol{h}^T \boldsymbol{h}\, \widehat{\boldsymbol{f}}]_m + 0.4 \sum_{q=0}^{3} \lambda_q (\sum_i (D_{mi}^{[q]})^T [(D^{[q]}\, \widehat{\boldsymbol{f}})_i]^{-0.2}).$$

Figs.11Comparison of several deconvolution techniques for image deblurring in the presence of additive noise. (a) Ideal image f. (b) Blurred image with noise $g = f \otimes h + n$ (same as shown in Fig 9Debluring by inverse filtering in the presence of noise. (left) Blurred image g obtained as $g = f \otimes h + n$, and a side view of its amplitude spectrum. In this example, Gaussian noise with $\sigma = 0.001$ was added to the image shown in Fig. 8Debluring by inverse filtering in the abscence of noise. (left) Blurred image g obtained in Fig. 7The result of convolving an image with a circular aperture, represented both in the spatial and frequency domains. (left) Input image f (cameraman) and a side view of its amplitude spectrum ($|F_{u,v}|$) plotted as a surface in 3D, using a $\log_{10}$ scale. (center) Circular aperture h and a side view of its amplitude spectrum ($|H_{u,v}|$). Note the occurrence of very small values, of the order of 10^{-6}. (right) Blurred image $g = f \otimes h$, and a side view of its amplitude spectrum $|G_{u,v}|$figure.7 as $g = f \otimes h$, and a side view of its amplitude spectrum. (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained as $F_{u,v} = G_{u,v}/H_{u,v}$ (Equation 20The Deconvolution Problemequation.7.20). $F_{u,v}$, and therefore f, can be exactly recoveredfigure.8 (left). The added noise has dominated the frequency components with low magnitude values (compare both spectra). (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained using Equation 22The Deconvolution Problemequation.7.22. As the noise dominates several frequency components, inverse filtering becomes ill posed, and the recovered image $\hat{f}$ is essentially noisefigure.9, left). Images (c) to (f) show the results obtained by deconvolving (b) using: (c) Wiener filtering; (d) the Richardson-Lucy algorithm; (e) image priors with Gaussian prior ($\beta = 2$); (f) image priors with a sparse prior ($\beta = 0.8$)figure.11 (e) and (f) show the results obtained with the above equations to deblur the noisy image shown in Fig.11Comparison of several deconvolution techniques for image deblurring in the presence of additive noise. (a) Ideal image f. (b) Blurred image with noise $g = f \otimes h + n$ (same as shown in Fig 9Debluring by inverse filtering in the presence of noise. (left) Blurred image g obtained as $g = f \otimes h + n$,

and a side view of its amplitude spectrum. In this example, Gaussian noise with $\sigma = 0.001$ was added to the image shown in Fig. 8Debluring by inverse filtering in the abscence of noise. (left) Blurred image g obtained in Fig. 7The result of convolving an image with a circular aperture, represented both in the spatial and frequency domains. (left) Input image f (cameraman) and a side view of its amplitude spectrum ($|F_{u,v}|$) plotted as a surface in 3D, using a $\log_{10}$ scale. (center) Circular aperture h and a side view of its amplitude spectrum ($|H_{u,v}|$). Note the occurrence of very small values, of the order of 10^{-6}. (right) Blurred image $g = f \otimes h$, and a side view of its amplitude spectrum $|G_{u,v}|$figure.7 as $g = f \otimes h$, and a side view of its amplitude spectrum. (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained as $F_{u,v} = G_{u,v}/H_{u,v}$ (Equation 20The Deconvolution Problemequation.7.20). $F_{u,v}$, and therefore f, can be exactly recoveredfigure.8 (left). The added noise has dominated the frequency components with low magnitude values (compare both spectra). (center) Image representation of the frequency domain inverse filter $1/H_{u,v}$ and a side view of its corresponding spectrum. (right) Deblurred image obtained using Equation 22The Deconvolution Problemequation.7.22. As the noise dominates several frequency components, inverse filtering becomes ill posed, and the recovered image $\hat{f}$ is essentially noisefigure.9, left). Images (c) to (f) show the results obtained by deconvolving (b) using: (c) Wiener filtering; (d) the Richardson-Lucy algorithm; (e) image priors with Gaussian prior ($\beta = 2$); (f) image priors with a sparse prior ($\beta = 0.8$)figure.11 (b) using $\beta = 2$ and $\beta = 0.8$, respectively. A detailed derivation of these expressions, based on statistical image models, and MATLAB scripts for implementing deconvolution using image priors, can be found in the accompanying materials of [12].

IX. Coded Aperture and Coded Exposure

This section enumerates and provides brief notes about a set of relevant papers and other sources of information on coded aperture and coded exposure. Its goal is to provide references to guide readers interested in learning more about these subjects.

A. Coded Aperture in Astronomy

First proposed by Dicke [20] and Ables [21] for X-ray and γ-ray astronomy, a lot of work has been done in coded aperture. Jean in 't Zand [33] maintains a site at NASA with introductory tutorials and references. A review of coded-aperture imaging with descriptions of practical implementations in astronomy can be found in [34]. Random Pinhole Pattern is described in [20] and [21]. A description of Uniformly Redundant Arrays (URA) can be found in [23], [24]. Modified Uniformly Redundant Arrays (MURA) were introduced in [25]. Pseudo Noise Product Arrays (PNP) are described in [35], and a discussion about Geometric Coded Aperture masks can be found in [36].
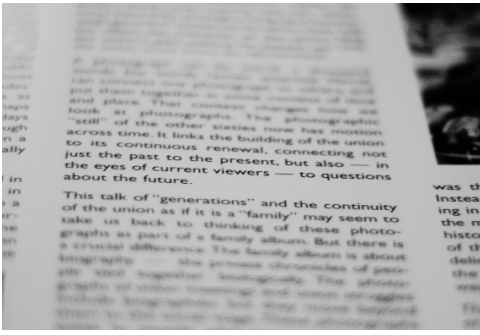
Fig. 14. Depth of field: The amount of blurring increases with object distance to the in focus plane

B. Coded Aperture in Computational Photography

In computational photography, coded apertures replace the camera's conventional circular aperture with patterns designed to preserve high-frequency content that would otherwise be lost by the circular aperture. Fig. 15Image capture using coded aperture. (a) Reference scene (image) f. (b) Circular aperture h. (c) Image acquired using the circular aperture in (b): $g = f \otimes h$. (d) Example of a simple coded aperture consisting of an array of pinholes: h'. (e) Image acquired using the coded aperture in (d) = $g' = f \otimes h'$. Note some high-frequency information not present in (c)figure.15 compares the blurring caused by a circular aperture with the one produced by a simple patterned aperture. Depending on its spatial location with respect to the camera's sensor and optical system, an optical mask can be used for different purposes. For instance, placing a mask over the sensor can be used to capture a light field [37] or to increase dynamic range [9]. If instead the mask replaces the optics diaphragm, it can be use to improve image deblurring [37], [13], encode information for distance computation from a single [12] or multiple images [38],[39], [40], or to encode light-field information [7]. Coded aperture has also been used to produce super-resolution [41]. A particular type of light-field modulation called *wavefront coding* is used in [42] to obtain depth-independent blurring, allowing the recovery of an all-in-focus image from deconvolution with a single kernel.

As explained in sections VIImage Formation and Convolutionsection.6 and VIIThe Deconvolution Problemsection.7 and can be observed in Fig. 14Depth of field: The amount of blurring increases with object distance to the in focus planefigure.14, the camera aperture introduces blurring that depends on object distance to the focal plane, which can be used to obtain distance information. Many computer vision techniques have been developed to estimate camera-object distance, and a discussion of this subject is outside the scope of this article. However, two of these approaches include *depth from focus* and *depth from defocus*. In the first, several images are taken while the aperture varies dynamically. In the second, a static set of images is used. The relation between depth from defocus and stereo is analyzed

in [43].

The central idea behind depth from defocus is that object distances to the camera may be determined if two or more images of the same object are taken with different apertures [39]. The use of coded apertures may improve distance discrimination. Hiura and Matsuyama [40] describe a multi-exposure camera for depth estimation that uses simple pinhole delta patterns as coded masks. Zhou et al. [14] discuss the problem of finding optimal coded aperture pairs to estimate depth from defocus, and Levin [44] analyses the problem of depth discrimination from a set of images captured with different coded apertures.

In principle, one cannot distinguish a blurred image of a sharp object from a sharp image of an object with blurred appearance. The use of coded aperture can be used to reduce this kind of ambiguity. Former work in this area can be found in [45]. Levin et al. [12] presented a technique to recover an all-in-focus image and estimate a scene depth map from a single photograph. They propose the use of the Kullback-Leibler (KL) divergence among the aperture masks at different scales as a measure of the depth-discrimination power of an aperture mask. KL is a concept of probability theory that measures the difference between two probability distributions. The optical mask is searched in a space of randomly sampled binary 13×13 patterns. The founded optimal mask is shown in Fig. 16Comparison of the Fourier amplitude spectra of different coded-aperture masks used for deconvolution. (top left) A conventional circular aperture with radius of 6 pixels. Note the existence of very small values, of the order of 10^{-5}, which causes the deconvolution process to become ill conditioned in the presence of noisy. (top right) The 13×13 binary mask proposed by Levin et al. [12]. (bottom left) A 7×7 grayscale mask proposed by Veeraraghavan et al. [37]. (bottom right) A 13×13 binary mask proposed by Zhou and Nayar [13]figure.16 (top right). The existence of zeros in this mask's spectrum complicates the deconvolution process. Thus, deblurring relies on the use of natural image priors (see Section VIIIDeconvolution techniquessection.8). As featureless regions of the image do not provide enough information for depth recovery, the estimated scene-distance maps are incomplete. They regularize such maps using graph-cuts and user assistance.

Veeraraghavan et al. [37] designed coded apertures for capturing light fields, as well as for defocus deblurring. For the later, they search for a broadband filter to improve the deconvolution process. The search criteria tries to maximize the mask's Fourier transform minimum magnitude, while minimizing its variance. They considered both binary and gray scale patterns, and concluded that the second are easier to find and, in addition, give superior performance. For gray scale patterns, they use a gradient-descent optimization to search for patterns formed by 7×7 pixels. An optimal mask is shown in Fig. 16Comparison of the Fourier amplitude spectra of different coded-aperture masks used for deconvolution. (top left) A conventional circular aperture with radius of 6 pixels.
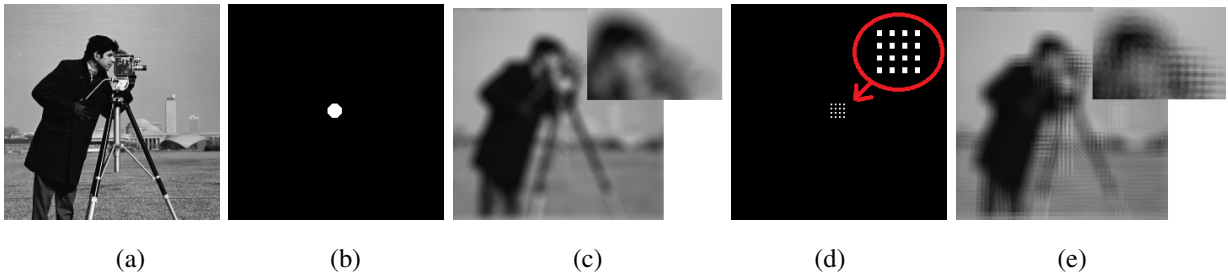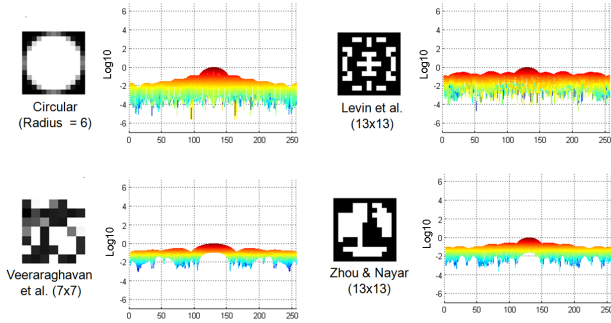
Fig. 15. Image capture using coded aperture. (a) Reference scene (image) f. (b) Circular aperture h. (c) Image acquired using the circular aperture in (b): $g = f \otimes h$. (d) Example of a simple coded aperture consisting of an array of pinholes: h'. (e) Image acquired using the coded aperture in (d) = $g' = f \otimes h'$. Note some high-frequency information not present in (c).

Fig. 16. Comparison of the Fourier amplitude spectra of different coded-aperture masks used for deconvolution. (top left) A conventional circular aperture with radius of 6 pixels. Note the existence of very small values, of the order of 10^{-5}, which causes the deconvolution process to become ill conditioned in the presence of noisy. (top right) The 13×13 binary mask proposed by Levin et al. [12]. (bottom left) A 7×7 grayscale mask proposed by Veeraraghavan et al. [37]. (bottom right) A 13×13 binary mask proposed by Zhou and Nayar [13].

Note the existence of very small values, of the order of 10^{-5}, which causes the deconvolution process to become ill conditioned in the presence of noisy. (top right) The 13×13 binary mask proposed by Levin et al. [12]. (bottom left) A 7×7 grayscale mask proposed by Veeraraghavan et al. [37]. (bottom right) A 13×13 binary mask proposed by Zhou and Nayar [13]figure.16 (bottom left).

Zhou et al. analyze the problem of finding optimal masks for deblurring [13], and the use of pairs of masks for depth-from-defocus estimation [38]. They use a modified version of the Wiener filter discussed in Section VIIIDeconvolution techniquesfigure.10. To find optimal masks for deblurring, they use a genetic algorithm to search the space of 13×13 binary masks [13]. One of their masks is shown in Fig. 16Comparison of the Fourier amplitude spectra of different coded-aperture masks used for deconvolution. (top left) A conventional circular aperture with radius of 6 pixels. Note the existence of very small values, of the order of 10^{-5}, which causes the deconvolution process to become ill conditioned in the presence of noisy. (top right) The 13×13 binary mask proposed by Levin et al. [12]. (bottom left) A 7×7 grayscale mask proposed by Veeraraghavan et al. [37]. (bottom right) A 13×13 binary mask proposed by Zhou and Nayar [13]figure.16 (bottom right). In their approach, the optimal mask varies with

image-noise level. To find optimal coded-aperture pairs for depth from defocus, Zhou et al. initially use the same genetic algorithm to find an 11×11 binary pattern. Starting from such a pattern, they progressively enlarge the mask (for use at different scales) up to 33×33 pixels. At each scale, they refine the initial estimate using a gradient descent optimization.

C. Coded Exposure

Coded exposure consists in changing the way the camera's sensor is exposed to the scene radiance, by opening and closing the shutter during image capture. The usual way is to keep the shutter opened during the entire acquisition period, which causes moving objects to appear blurred, over attenuating high frequencies. If the movement is uniform and rectilinear, the blurring can be approximated by convolution with a box filter.

Raskar et al. [15] proposed to open and close the shutter several times using a temporal coded pattern. This way, they can reduce high-frequency attenuation and improve image restoration. Figs. 17Comparison of image acquisition using conventional and coded exposure. (a) Reference scene. (b) A continuous exposure pattern. The yellow segment represents the time period in which the shutter is kept opened. (c) Image acquired using a continuous exposure during a horizontal camera movement. Note the continuous blurring. (d) A coded-exposure pattern. The exposure consists of several discrete time intervals. (e) Image acquired using the coded-exposure pattern in (d) for the same camera movement as in (c). Some high-frequency information has been capturedfigure.17 (c) and (e) show examples of images acquired using conventional and coded exposure, respectively. In the conventional case, the blurring is continuous (Figs. 17Comparison of image acquisition using conventional and coded exposure. (a) Reference scene. (b) A continuous exposure pattern. The yellow segment represents the time period in which the shutter is kept opened. (c) Image acquired using a continuous exposure during a horizontal camera movement. Note the continuous blurring. (d) A coded-exposure pattern. The exposure consists of several discrete time intervals. (e) Image acquired using the coded-exposure pattern in (d) for the same camera movement as in (c). Some high-frequency information has been capturedfigure.17 c). When using coded exposure, the camera captures a series of discrete blurred images (one for each exposure sub-interval) of the moving objects. Such images
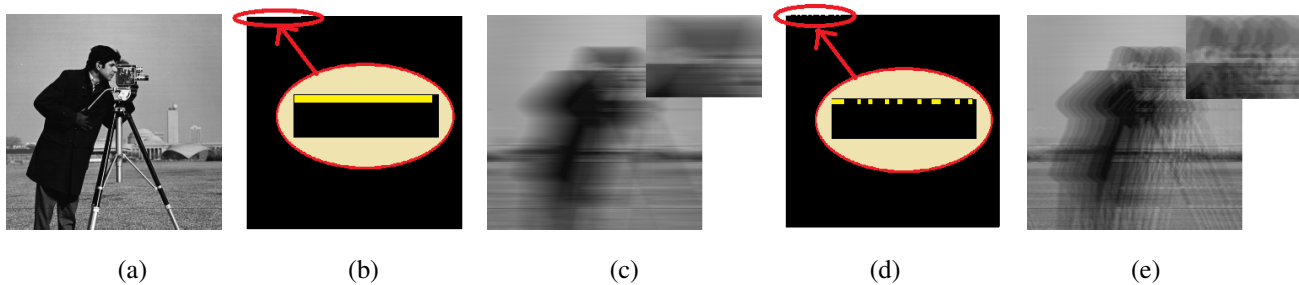
Fig. 17. Comparison of image acquisition using conventional and coded exposure. (a) Reference scene. (b) A continuous exposure pattern. The yellow segment represents the time period in which the shutter is kept opened. (c) Image acquired using a continuous exposure during a horizontal camera movement. Note the continuous blurring. (d) A coded-exposure pattern. The exposure consists of several discrete time intervals. (e) Image acquired using the coded-exposure pattern in (d) for the same camera movement as in (c). Some high-frequency information has been captured.

contain high-frequency information (Figs. 17Comparison of image acquisition using conventional and coded exposure. (a) Reference scene. (b) A continuous exposure pattern. The yellow segment represents the time period in which the shutter is kept opened. (c) Image acquired using a continuous exposure during a horizontal camera movement. Note the continuous blurring. (d) A coded-exposure pattern. The exposure consists of several discrete time intervals. (e) Image acquired using the coded-exposure pattern in (d) for the same camera movement as in (c). Some high-frequency information has been capturedfigure.17 e), which can be recovered through deconvolution.

Raskar et al. select the exposure pattern by performing a random brute-force linear search over the space of 1D binary patterns composed of 52 bits, half of which are 1's. Among all possibilities, they choose the pattern that maximizes the minimum amplitude value of the its Fourier transform, while minimizing DFT amplitude variance.

X. Conclusion

Computational photography is a new research area that tries to expand the concept of traditional photography using state-of-the-art technology. It houses many exciting challenges that will require new creative solutions. This article presented a gentle introduction to coded computational photography, focusing on coded aperture and coded exposure. The goal was to provide a rigorous, although intuitive, introduction to the fundamental concepts and tools required to master the subject. It also tried to relate these abstract concepts to actual algorithms and techniques. It is our hope that this document will be a valuable source of information for anyone interested in learning about computational photography.

A. Research Opportunities

Coded photography is a very young discipline, with several problems still waiting to be understood and solved. Many simplifying assumptions made by current techniques are not satisfied by most real scenes. For instance, objects tend to be distributed over relatively large depth ranges, are not composed by Lambertian surfaces, may be moving in arbitrary directions and with different speeds. Any of these situations brake the simple *one-mask deconvolution model* assumption,

and demand a point-by-point deconvolution strategy. So far, these problems have been treated with approximate solutions based on image segmentation and fusion. Current solutions still require certain degree of manual intervention to produce reasonable results.

It has been shown that image priors can be applied to the deconvolution problem yielding good results. Although such a strategy can alleviate some artifacts common to other approaches, a significant amount of work still need to be done with respect to the development of new priors to model natural images and techniques to explore this knowledge.

Light fields are a fundamental concept in computational photography. To properly sample it, sensor cells have to be allocated to capture directional information. This imposes a trade-off between spatial and angular resolution. Surely, a lot of work needs to be done in this area, especially related to how to explore light-field redundancy to reduce the sampling requirements using natural image priors.

Even though significant progress have been made regarding image capture and rendering quality, a photograph continues to be static and two dimensional projection of a 4-D light field. What other dimensions can be explored? Do photographs need to be two dimensional? instantaneous? How can we expand the sense of reality of a captured scene in all dimensions, including a cognitive sense?

References

[1] A. Agarwala, M. Agrawala, M. F. Cohen, D. Salesin, and R. Szeliski, "Photographing long scenes with multi-viewpoint panoramas," *ACM Transactions on Graphics*, vol. 25, pp. 853–861, 2006.

[2] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," *SIGGRAPH 97*, pp. 369–378, 1997.

[3] B. Wilburn, N. Joshi, V. Vaish, E.-V. Talvala, E. Antunez, A. Barth, A. Adams, M. Horowitz, and M. Levoy, "High performance imaging using large camera arrays," *ACM Trans. Graph.*, vol. 24, pp. 765–776, July 2005.

[4] M. Levoy and P. Hanrahan, "Light field rendering," in *Proc. SIGGRAPH '96*, 1996, pp. 31–42.

[5] R. Ng, "Fourier slice photography," *ACM Trans. Graph.*, vol. 24, pp. 735–744, July 2005.

[6] T. Georgeiv and C. Intwala, "Light field camera design for integral view photography," 2008.

[7] C.-K. Liang, T.-H. Lin, B.-Y. Wong, C. Liu, and H. H. Chen, "Programmable aperture photography: Multiplexed light field acquisition," *ACM Trans. Graph.*, vol. 27, pp. 55:1–55:10, 2008.

[8] T. Georgiev and A. Lumsdaine, "Focused plenoptic camera and rendering," *Journal of Electronic Imaging*, vol. 19, p. 021106, 2010.

[9] S. Nayar and S. Narasimhan, "Assorted Pixels: Multi-Sampled Imaging With Structural Models," in *ECCV*, vol. IV, 2002, pp. 636–652.

[10] S. Nayar and T. Mitsunaga, "High Dynamic Range Imaging: Spatially Varying Pixel Exposures," in *IEEE CVPR*, vol. 1, 2000, pp. 472–479.

[11] S. Kuthirummal and S. K. Nayar, "Multiview Radial Catadioptric Imaging for Scene Capture," *ACM Trans. Graph.*, pp. 916–923, Jul 2006.

[12] A. Levin, R. Fergus, F. Durand, and W. T. Freeman, "Image and depth from a conventional camera with a coded aperture," *ACM Trans. Graph.*, vol. 26, July 2007, Article 70.

[13] C. Zhou and S. K. Nayar, "What are good apertures for defocus deblurring?" in *IEEE Intern. Conf. Computational Photography*, 2009, pp. 1–8.

[14] C. Zhou, S. Lin, and S. Nayar, "Coded aperture pairs for depth from defocus," in *ICCV*, Kyoto, Japan, Oct 2009, pp. 325–332.

[15] R. Raskar, A. Agrawal, and J. Tumblin, "Coded exposure photography: motion deblurring using fluttered shutter," *ACM Trans. Graph.*, vol. 25, no. 3, pp. 795–804, 2006.

[16] S. Nayar, "Computational cameras: Approaches, benefits and limits," Tech. Rep., Jan 2011.

[17] G. Wetzstein, I. Ihrke, D. Lanman, and W. Heidrich, "State of the Art in Computational Plenoptic Imaging," in *Proc. Eurographics (STAR)*, 2011, pp. 25–48.

[18] R. Raskar, "Computational photography: Epsilon to coded photography," pp. 238–253, 2009.

[19] B. Hayes, "Computational photography," *American Scientist*, vol. 96, no. 2, pp. 94–99, March-April 2008.

[20] R. H. Dicke, "Scatter-hole cameras for x-rays and gamma rays," *Astrophysical Journal*, vol. 153, p. L101, 1968.

[21] J. G. Ables, "Fourier transform photography: A new method for x-ray astronomy," vol. 1, 1968, p. 172.

[22] M. J. E. Golay, "Point arrays having compact, nonredundant autocorrelations," *J. Opt. Soc. Am.*, vol. 61, no. 2, pp. 272–273, Feb 1971.

[23] E. E. Fenimore and T. M. Cannon, "Coded aperture imaging with uniformly redundant arrays," *Appl. Opt.*, vol. 17, no. 3, pp. 337–347, Feb 1978.

[24] E. E. Fenimore, "Coded aperture imaging: Predicted performance of uniformly redundant arrays," *Applied Optics*, vol. 17, pp. 3562–3570, 1978.

[25] S. R. Gottesman and E. E. Fenimore, "New family of binary arrays for coded aperture imaging," *Applied Optics*, vol. 28, pp. 4344–4352, 1989.

[26] Y. Weiss and W. Freeman, "What makes a good model of natural images," in *Proc. CVPR*, 2007, pp. 1–8.

[27] W. Richardson, "Bayesian-based iterative method of image restoration," *J. Opt. Soc. Am.*, vol. 62, no. 1, pp. 55–59, 1972.

[28] L. B. Lucy, "An iterative technique for the rectification of observed distributions," *The Astronomical Journal*, vol. 79, pp. 745+, 1974.

[29] E. S. Meinel, "Origins of linear and nonlinear recursive restoration algorithms," *J. Opt. Soc. Am. A*, vol. 3, no. 6, pp. 787–799, Jun 1986.

[30] R. Molina, J. Nunez, F. Cortijo, and J. Mateos, "Image restoration in astronomy: a bayesian perspective," *Signal Processing Magazine, IEEE*, vol. 18, no. 2, pp. 11–29, mar 2001.

[31] H. Bi and G. Boerner, "When does the Richardson-Lucy deconvolution converge?" *AAPS*, vol. 108, pp. 409–415, Dec. 1994.

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

[33] J. in 't Zand. (2010) Coded aperture imaging in high-energy astronomy. NASA. [Online]. Available: http://astrophysics.gsfc.nasa.gov/cai/

[34] E. Caroli, J. B. Stephen, G. Cocco, L. Natalucci, and A. Spizzichino, "Coded aperture imaging in X- and gamma-ray astronomy," *Space Science Reviews*, vol. 45, no. 3, pp. 349–403, 1987.

[35] S. R. Gottesman and E. J. Schneid, "PNP - a new class of coded aperture arrays," *IEEE Trans. on Nuclear Science*, vol. 33, pp. 745–749, 1986.

[36] A. R. Gourlay and J. B. Stephen, "Geometric coded aperture masks," *Applied Optics*, vol. 22, pp. 4042–4047, 1983.

[37] A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin, "Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing," *ACM Trans. Graph*, vol. 26, pp. 3562–3570, 2007, Article 69.

[38] C. Zhou, S. Lin, and S. Nayar, "Coded Aperture Pairs for Depth from Defocus and Defocus Deblurring," *International Journal on Computer Vision*, vol. 93, no. 1, p. 53, May 2011.

[39] A. P. Pentland, "A new sense for depth of field," *IEEE Trans. PAMI*, vol. 9, pp. 523–531, July 1987.

[40] S. Hiura and T. Matsuyama, "Depth measurement by the multi-focus camera," in *Proc. CVPR*. Washington, DC, USA: IEEE Computer Society, 1998, pp. 953–959.

[41] A. Mohan, X. Huang, R. Raskar, and J. Tumblin, "Sensing increased image resolution using aperture masks," in *Proc. CVPR*, 2008, pp. 1–8.

[42] W. T. Cathey and E. R. Dowski, "New paradigm for imaging systems," *Appl. Opt.*, vol. 41, no. 29, pp. 6080–6092, Oct 2002. [Online]. Available: http://ao.osa.org/abstract.cfm?URI=ao-41-29-6080

[43] Y. Y. Schechner and N. Kiryati, "Depth from defocus vs. stereo: How different really are they?" *Int. J. Comput. Vision*, vol. 39, pp. 141–162, September 2000.

[44] A. Levin, "Analyzing depth from coded aperture sets," in *Proc. ECCV: Part I*. Berlin, Heidelberg: Springer-Verlag, 2010, pp. 214–227.

[45] E. R. Dowski and W. T. Cathey, "Single-lens single-image incoherent passive-ranging systems," *Appl. Opt.*, vol. 33, no. 29, pp. 6762–6773, Oct 1994.