Master thesis and internship[BR]- Master's thesis : Improvement of the optical setup of a compressive sensing imager[BR]- Internship

Auteur : Ruwet, Emile
Promoteur(s) : Georges, Marc
Faculté : Faculté des Sciences appliquées
Diplôme : Master en ingénieur civil en aérospatiale, à finalité spécialisée en "aerospace engineering"
Année académique : 2023-2024
URI/URL : http://hdl.handle.net/2268.2/20509

Improvement of the optical setup of a compressive sensing imager

Author:

Emile RUWET

Academic supervisor:

Marc GEORGES

Committee members:

Lionel CLERMONT

Serge HABRAKEN

Laurent JACQUES

Master's thesis completed in order to obtain the degree of Master of Science in Aerospace Engineering by Emile RUWET

University of Liège - Faculty of Applied Sciences

Academic year 2023-2024

Abstract

Compressed sensing is an emerging field that allows for the recovery of a sparse signal from fewer measurements than permitted by the Nyquist theorem. This new paradigm can be used to create new imager architectures that are simpler, more compact, and cheaper than traditional imagers, acquiring images in a compressed manner and thereby reducing the amount of data to handle. These characteristics are appealing for potential implementation in Earth observation satellites, where size and weight are critical factors, and where the amount of collected data is substantial, with limited storage capacity and transfer rates to the ground.

This master's thesis focuses on the implementation of the optical part of a compressive sensing imager in the laboratory, with the objective of performing a particular compressive sensing reconstruction method, known as inpainting. A comprehensive review of the state of the art in compressive sensing and various imager architectures is first provided. Afterwards, the design of the instrument and the selection of all its components are extensively detailed. A digital micromirror device is used for producing incomplete, damaged images of a scene, and a camera detector records the resulting image. Subsequently, a calibration procedure for the damaged images was established to prepare them for reconstruction through inpainting. This calibration includes dark and flat frame corrections, as well as post-reconstruction image perspective correction. The instrument's point spread function is also measured, and a dithering method is implemented to improve its resolution. Furthermore, the pattern mask used for the reconstruction is studied and calibrated using morphological erosion.

Keywords— compressive sensing, inpainting, optical imaging device, digital micromirror device, Scheimpflug principle, pattern mask, point spread function, dithering

Acknowledgements

First, I would like to thank my academic supervisor Marc Georges for giving me the opportunity to carry out my master's thesis at the Centre Spatial de Liège (CSL). I would also like to thank him for his presence, as well as his suggestions and advice, which helped me make progress with this thesis. I would also like to thank Lionel Clermont (CSL) for his suggestions and feedback on my work, and Laurent Jacques (UCLouvain) for the many progress meetings and for sharing his experience in the field of compressive sensing.

Next, I would like to thank Clément Thomas (CSL) for all his help and advice during my internship, as well as his proofreading of my manuscript. I would also like to thank all the other PhD students as well as the other students present at the CSL for their company and discussions during lunchtimes. Additionally, I would like to thank all the CSL staff, and in particular those present in my office, for all the interesting discussions and the good working atmosphere during this internship.

Finally, I would like to thank my family for their unconditional support, and in particular my parents, who enabled me to complete these five years of study at university. I would also like to express my warmest thanks to my partner Mathilde, for her presence, her full support and help in proofreading this work.

Contents

List of Figures

List of Tables

Acronyms

BP	Basis Pursuit
CAD	Computer-aided Design
CASSI	Coded Aperture Snapshot Spectral Imager
CCD	Charge-coupled Device
CMOS	Complementary Metal Oxide Semiconductor
CS	Compressive Sensing
CSL	Centre Spatial De Liège
DCT	Discrete Cosine Transform
DFT	Discrete Fourier Transform
DL	Deep Learning
DMD	Digital Micromirror Device
DOF	Depth Of Field
FOV	Field Of View
FPA	Focal-plane Array
FPGA	Field-programmable Gate Array
IHT	Iterative Hard Thresholding
ISTA	Iterative Shrinkage Thresholding Algorithm
LC	Liquid Crystal
LCOS	Liquid Crystal On Silicon
MDF	Master Dark Frame
MFF	Master Flat Frame
MP	Matching Pursuit
MURA	Modified Uniformly Redundant Array
OMP	Orthogonal Matching Pursuit
ONB	Orthonormal Basis
PSF	Point Spread Function
RIP	Restricted Isometry Property
SLM	Spatial Light Modulator
SNR	Signal-to-noise Ratio
SPI	Single-pixel Imaging
SR	Super-resolution
URA	Uniformly Redundant Array
WT	Wavelet Transform

1 Introduction

In most modern cameras and imaging devices, images are captured using detector arrays made of millions of small pixels. Each of them collects a small portion of the overall image, by converting the incident photons into a measurable current. This type of technology, facilitated by the development of Complementary Metal Oxide Semiconductor (CMOS) and Charge-Coupled Device (CCD) sensors, mainly made of silicon, has enabled high-definition, cheap and high-performance imaging instruments, with a constant increase in the number of pixels on those detectors [1].

However, this multi-pixel camera design is at the expense of a continuously larger amount of data per image, as each pixel collects a finite quantity of information stored in a certain number of bits in a memory. Additionally, imagers generally do not acquire data in a single wavelength, but in multiple spectral bands. Most consumer cameras capture images in three spectral bands: red, blue and green. This increases the amount of image information by a factor of three. Furthermore, many domains in modern science demand ever-increasing amounts of data both spatially and temporally, to the point where the limitations of this traditional data sensing method begin to appear [2]. This problem is particularly true in the field of Earth observation, which requires the acquisition of a large quantity of images almost continuously. Additionally, it is now more and more common to acquire several hundred spectral bands at once, as it is the case of hyperspectral imaging. This poses challenges in data storage, transmission to Earth, and ground-based processing, among other issues.

Combined with a constant increase in computing power, recent advances in Compressive Sensing (CS) have led to the creation of numerous new imaging device designs that could be credible candidates for addressing these challenges. CS is a new branch in the field of data acquisition, which involves acquiring a signal, such as an image, in a compressed manner and subsequently reconstructing this compressed signal to retrieve the original signal [3]. CS is made possible by the fact that most natural signals are considered sparse or can be well approximated by sparse representations. A signal is considered sparse when it can be represented by only a few non-zero coefficients in a certain domain. In a simplified way, the compressed signal used for the image reconstruction is generally obtained by applying sequentially different masks to the image to acquire, blocking or not its pixels, which then forms a series of modulated measurements. Knowing which masks were used to collect these measurements, it is then possible to retrieve the original image using reconstruction algorithms. The image is effectively compressed when the number of these measurements is lower than the number of pixels in the original image. This new paradigm therefore offers the possibility of reducing the amount of data to handle because the imager only measures a compressed version of the image. Additionally, CS allows for optical instruments with much simpler, smaller, and cheaper architectures than traditional imagers. For example, the instrument might not contain any lenses, or it might be equipped with only one pixel on its detector.

While a major difficulty of CS lies in the reconstruction algorithms and their optimization, this work focuses on the implementation and calibration of a compressive sensing imager in the laboratory. As a starting point for future work, this imager focuses first on performing inpainting. Inpainting is an image reconstruction technique that can be carried out through CS. This allows for an easier implementation of a compressive sensing imager. This manuscript is structured into three main parts. First, an overview of the state of the art in compressive sensing are presented, covering its mathematical concept, various reconstruction techniques, the different architectures and components used in CS imagers. Secondly, the design and fabrication of the imager in the laboratory are detailed and explained. Finally, the calibration phase and post-processing of the images taken with the instrument are presented. The calibration process prepares the images so that they can be properly injected into a reconstruction algorithm.

2 State of the art

In this chapter, the state of the art of compressive sensing will be described. Its basics will first be introduced to have a better intuition of how it works. Next, compressive sensing and the concept of inpainting will be mathematically explained. Finally, a state of the art concerning the hardware and imagers using CS will be introduced.

2.1 Basics of compressive sensing

The traditional approach of acquiring a signal, such as an image, is to take at least as many measurements as the number of elements in the signal. In image processing, the hypothesis of the sparsity of the image is often taken. Sparsity principle corresponds to saying that a signal can be represented with much fewer data than the number of elements in the signal. With the knowledge of the signal sparsity, it seems that the traditional signal acquisition approach is not optimized and wastes resources. Indeed, all the efforts of measuring all the elements of the signal are lost because the signal is finally compressed and most of the null elements are discarded. Instead, it would be more interesting to directly measure the compressed signal to maximize the use of the observational and hardware resources. This is the basic idea behind compressive sensing [3]. Although pixelated images cannot directly be seen as sparse signals, it turns out that they are generally sparse after an appropriate change of basis such as Wavelet Transform (WT) and Discrete Cosine Transform (DCT). In these bases, images can therefore be compressively sensed.

Before going into the mathematical formulation of compressive sensing, it is important to have a good intuition of how CS is possible and how fewer samples than the number of pixels of an image can be used to reconstruct the image. One possible interpretation of why CS works is presented below and follows the work of [4].

Suppose that an imager is used to acquire a picture of a scene. Let us also consider that the observed scene is known ahead of time and consists of only one single illuminated pixel located in the picture, and all the other pixels are black. This image is hence considered very sparse, because most of the data it contains are zeros, corresponding to the dark background. Only one component of the image vector will be non-zero and will have a value equal to the intensity of this point source. In conventional imaging with a Focal-Plane Array (FPA) comprising N detector elements, every pixel of the image will be individually measured by each element of the detector. As a consequence, most of the measurements will have a null value, and only one will contain the source. This observational strategy is depicted in Fig. 2.1a, the measurement x_j is the inner product of the scene image f^* and the indicator image I_j of the j-th active pixel. However, a better strategy can also be applied to localize the position of the bright pixel. One can use a binary sensing strategy shown in Fig. 2.1b. It is obvious that each pattern will provide an increasing precision on the position of the non-zero pixel of the image. Indeed, each additional measure divides by 2 the search area where the bright source is located. Hence, it is possible to accurately localize the pixel with a number of measurements equal

to $M = \log_2 N$, whereas the first traditional method uses N samples. As explained in [1], this observational method is comparable to playing the "twenty questions" game where one player has to guess a specific person by asking questions to the second player who has to answer only by "yes" or "no". It is far more interesting to ask question that concerns as many individuals as possible such as the gender, hairstyle and eye color, rather than guessing names randomly.

It is clear that this intuitive method only works for images with only one single non-zero pixel. When the image is composed of some non-zero pixels, or when the image is not arbitrary but contains well-defined structures such as edges and uniform surfaces that can be expressed with few non-zero elements in other bases, compressive sensing provides mechanisms to still be able to reconstruct the image with $M \ll N$ measurements. These two cases of images are shown in Fig. 2.1c and Fig. 2.1d respectively, where the sensing matrices are random binary patterns. It corresponds to the same underlying intuition than the case shown in Fig. 2.1b but applied for more complex scenes.

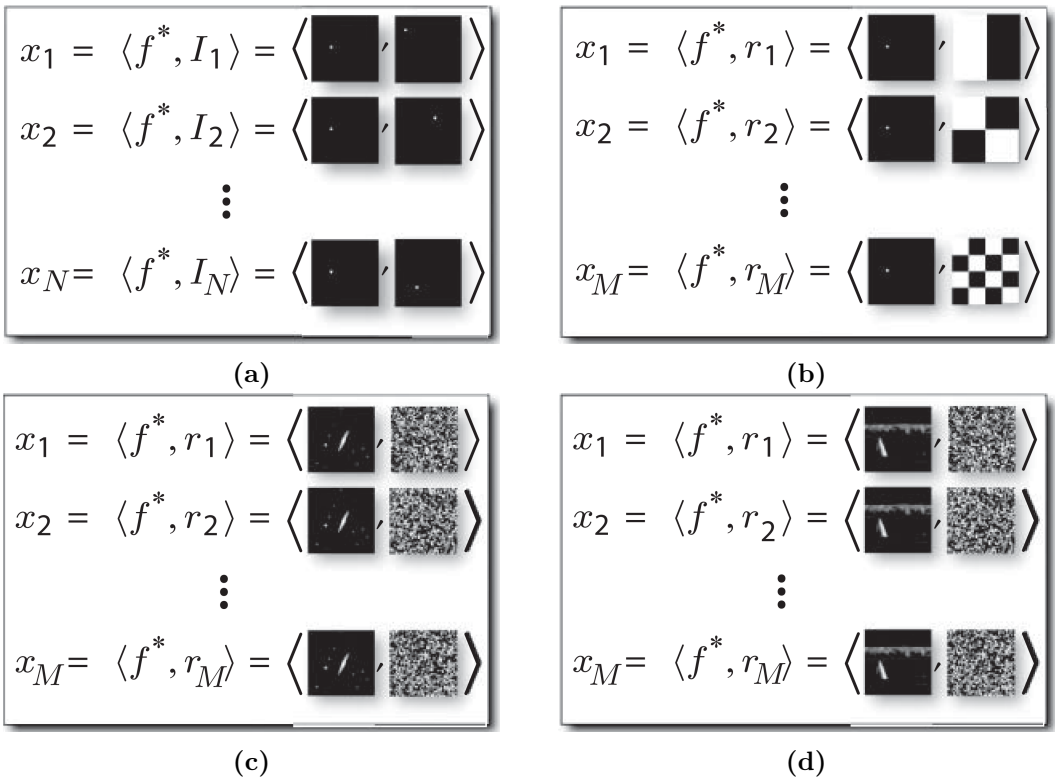


Figure 2.1: Different observational strategies to sense an image. Each measurement corresponds to the inner product of the image and an indicator image. (a) It corresponds to a classical FPA sensor where each pixel detector samples a small portion of the scene. (b) This intuitive method uses increasing smaller binary checkerboard patterns to locate the position of the bright pixel in $M = \log_2 N$ measurements. (c) CS strategy to acquire an image with some unknown nonzero pixels with, for this example, pseudo-random binary patterns. (d) This strategy is identical to the previous method but for the case where the image contains structures and that are sparse in other bases. From [4].

2.2 Mathematical description of compressive sensing

This section describes the mathematical background of CS. The developments presented below follow the work of [3, 5] and the lecture notes of [6].

Let us suppose a signal $\mathbf{x} \in \mathbb{R}^N$. This signal can be, for example, a gray levels image containing N pixels in it. The compressive sensing problem consists in reconstructing $\mathbf{x}$ based on its indirect measurement $\mathbf{y} \in \mathbb{R}^M$ with $M < N$. The relation between these two vectors is

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}, \tag{2.1}$$

where the matrix $\mathbf{A} \in \mathbb{R}^{M \times N}$ is called the measurement matrix and represents the linear measurement system that will sample the signal $\mathbf{x}$. Therefore, this matrix $\mathbf{A}$ models the whole imager for the case of image acquisition. The vector $\mathbf{n}$ is the additional noise during the measurements and is associated with the physics of the sensing device. In the following of this section, one will consider a noiseless system, therefore neglecting $\mathbf{n}$.

The system in Eq. 2.1 has N unknowns and only M equations. As $M < N$, this system is underdetermined and there is an infinite number of solutions, provided that there exists at least one. This result is closely related to the Nyquist-Shannon sampling theorem which describes how to sample a continuous-time signal without loss of information. It states that the signal can be perfectly reconstructed if the sampling frequency is at least twice the highest frequency of the original signal. Therefore, one needs at least N different samples of x in order to retrieve the signal with classical linear algebra.

However, as it has previously been said, compressive sensing introduces the hypothesis that the signal is sparse. A signal of size N is called K-sparse if it contains K non-zero elements with $K \ll N$. Images are generally not sparse in the canonical ("pixel") basis, but it is possible to change the basis of the image and express it into a different Orthonormal Basis (ONB), where the images are sparse. Discrete Cosine Transform (DCT), Discrete Fourier Transform (DFT) or Wavelet Transform (WT) are examples of ONBs in which images can be considered as sparse vectors [2, 7, 4]. Under a certain orthonormal transform domain, the signal $\mathbf{x}$ can be expressed as a weighted sum of simple functions and is written as

$$\mathbf{x} = \mathbf{\Psi}\mathbf{s} = \sum_{i=1}^{N} s_i \psi_i, \tag{2.2}$$

where $\mathbf{\Psi} \in \mathbb{R}^{N \times N}$ is the orthonormal basis matrix, also called the sparsity matrix, and $\mathbf{s}$ is a sparse vector containing the coefficients corresponding to each basis vector of $\mathbf{\Psi}$. Eq. 2.1 without noise can now be written as

$$\mathbf{y} = \mathbf{A}\Psi\mathbf{s} = \mathbf{A}'\mathbf{s}, \tag{2.3}$$

where $\mathbf{A}' \triangleq \mathbf{A}\Psi$ is the sensing matrix. If a sparse solution of $\mathbf{y} = \mathbf{A}'\mathbf{s}$ is found, the original signal can then be recovered using Eq. 2.2. A graphical representation of the equation above can be seen in Fig. 2.2 with a 3-sparse $\mathbf{s}$ vector as an example. The difficulty of CS lies in the fact that the non-zero, or at least, the most significant coefficients of $\mathbf{s}$ are not known *a priori*. Indeed, knowing the K significant coefficients in advance would be easier, as this would make it possible to measure them directly and reconstruct the signal. Hence,

CS aims to take M well-chosen measurements with $K < M \ll N$ to retrieve accurately the relevant K coefficients of $\mathbf{s}$.

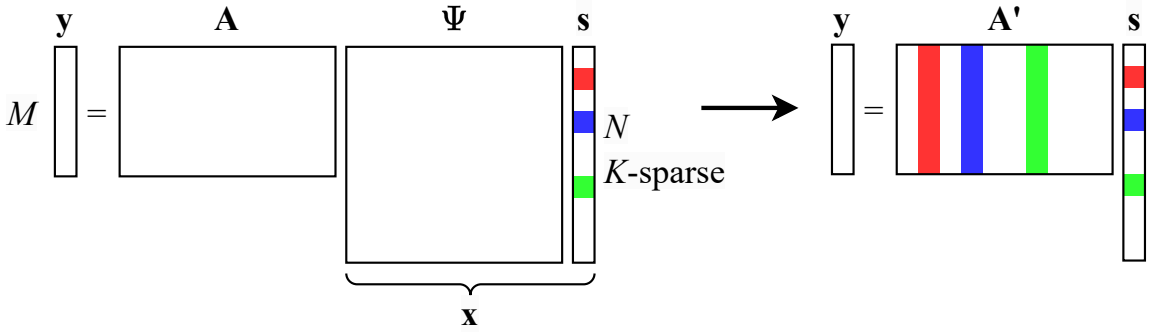


Figure 2.2: Graphical representation of the compressive sensing measurement process of Eq. 2.3. In this example, the $\mathbf{s}$ vector is K-sparse with $K = 3$, and the significant coefficients are highlighted as well as their corresponding columns in the sensing matrix $\mathbf{A}'$. Adapted from [4, 2].

For practical applications, one needs fast and efficient reconstruction algorithms to solve the underdetermined system $\mathbf{y} = \mathbf{A}'\mathbf{s}$ with the hypothesis of a sparse solution and retrieve the desired signal $\mathbf{x}$. Some of the main reconstruction algorithms used in compressive sensing are listed and explained in Sec. 2.3.

In addition to an efficient recovery method, the sensing matrix $\mathbf{A}'$ is an essential parameter in CS. It must allow an accurate signal reconstruction using the appropriate algorithm. Hence, the choice of the measurement matrix $\mathbf{A}$ and the sparsity matrix $\mathbf{\Psi}$ is primordial in the quality of the compressive sensing device. In this discussion, a commonly used criteria is the Restricted Isometry Property (RIP). The sensing matrix $\mathbf{A}'$ satisfies RIP of order $\delta_S \in (0, 1)$ if

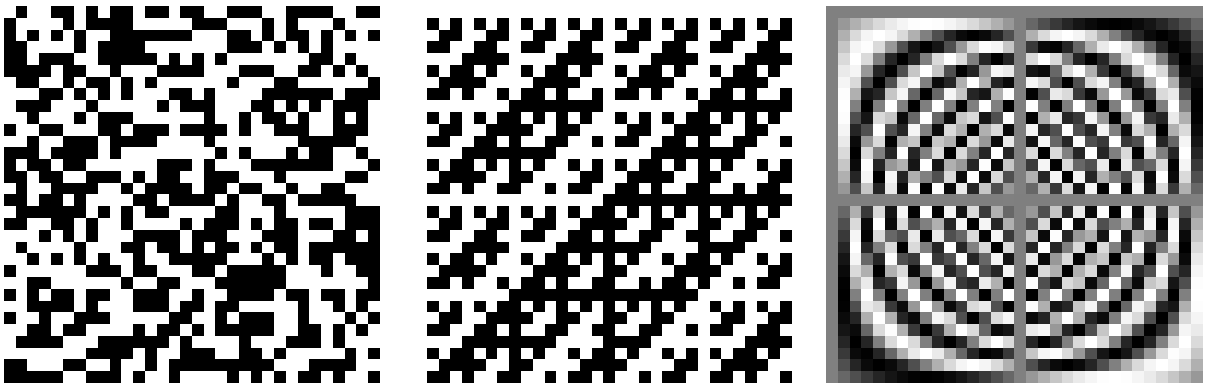
$$1 - \delta_S \leq \frac{\|\mathbf{A}'\mathbf{s}\|_2^2}{\|\mathbf{s}\|_2^2} \leq 1 + \delta_S \tag{2.4}$$

where $\mathbf{s}$ is an arbitrary sparse vector, and δ_S is called the restricted isometry constant and corresponds to the smallest constant to which Eq. 2.4 holds. The above equation uses the p-norm notation and is defined as $\|\mathbf{s}\|_p = \left(\sum_{i=1}^{N} |s_i|^p\right)^{1/p}$. Physically, the RIP ensures that the signal is modulated such that all the compressed samples contain all the needed information to reconstruct it, and that one can find a unique solution to the system of Eq. 2.3 [4]. Furthermore, the sensing matrix should satisfy the incoherence condition: it requires that the rows $\mathbf{A}_j$ of $\mathbf{A}$ cannot sparsely represent the columns Ψ_i of the basis matrix $\mathbf{\Psi}$, and conversely [8]. A low incoherence will prevent the measurement matrix to be aligned with the basis matrix, and will allow each measurement to be sufficiently varied to capture all the information of $\mathbf{s}$.

It can be demonstrated that pseudo-random matrices satisfy with a high probability the RIP property and the incoherence condition, and can therefore be employed as a sensing matrix for CS applications. For example, the matrix elements of $\mathbf{A}'$ can be i.i.d.

random variables from a Normal distribution $\mathcal{N}(0, 1/N)$. A sensing matrix of this form allows the retrieval of the original signal with $M \geq cK \log(N/K) \ll N$ with c a small constant [8]. However, the use of pseudo-random matrices can introduce several problems, according to [9] and a personal communication with L. Jacques of UCLouvain. First, it requires to store each element of the sensing matrix, as there is very low coherence between the elements, giving rise to a memory complexity $O(MN)$. Secondly, the computational complexity of the reconstruction is also $O(MN)$ which can rapidly become too demanding for large signals, especially for high resolution images.

These issues of pseudo-random matrices can be solved by using deterministic families of measurement matrices, such as Hadamard or Fourier matrices [10]. Hadamard patterns are binary and have a mosaic shape, while Fourier matrices are composed of grayscale fringes. Such deterministic measurement matrices offer some advantages. First, they theoretically allow perfect image reconstruction, as the patterns are orthogonal. The image can be losslessly retrieved when it is fully sampled. Secondly, because natural images can often be sparsely represented in the Hadamard and Fourier domain, these matrices enable reducing the number of measurements needed to achieve a clear image reconstruction. Lastly, efficient algorithms exist to deal with signals in the Hadamard/Fourier domain. Therefore, the computational complexity can be reduced to $O(N \log N)$. Examples of pseudo-random, Hadamard, and Fourier measurement matrices can be seen in Fig. 2.3.



(a) **(b)** **(c)**

Figure 2.3: Examples of 32×32 measurement matrices. (a) Pseudo-random binary matrix with a filling ratio of 50%. (b) Example of a binary Hadamard basis matrix. (c) Sine basis matrix in the Fourier domain, the matrix is no longer binary but in grayscale.

2.3 Reconstruction algorithms

As it has been said, the goal of CS is to solve the underdetermined system of Eq. 2.3 knowing a sparse solution. This additional information on the solution is a form of regularization, *i.e.*, a process that supposes the solution to be simple. The problem becomes better posed and can be solved with the right algorithm. The key of CS imaging lies in the efficiency and the quality of the reconstruction algorithms. Different recovery methods have been found since the beginning of compressive sensing. In the following of

this section, the primary algorithms will be explained: optimization, greedy, thresholding-based methods and deep-learning methods. These methods are presented in [3, 2, 11].

2.3.1 Optimization methods

In order to find a sparse vector $\widehat{\mathbf{s}}$, solution of Eq. 2.3, it could be adequate to use the ℓ_q-minimization problem, written as

$$\widehat{\mathbf{s}} = \arg\min_{\mathbf{s}} \|\mathbf{s}\|_q \text{ subject to } \mathbf{y} = \mathbf{A}'\mathbf{s}, \tag{2.5}$$

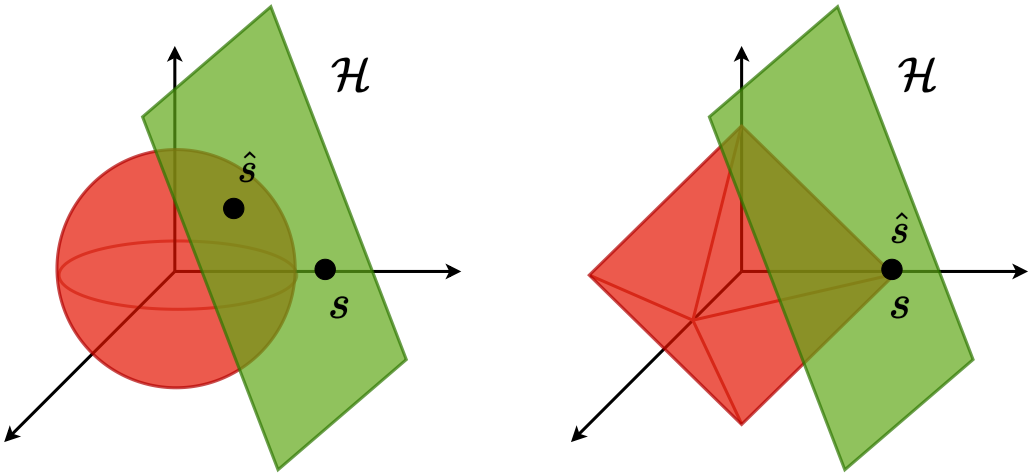
where $q \in [0, \infty[$. A classical method for solving inverse problem is to use the ℓ_2-minimization problem, where it minimizes the norm of the vector. Unfortunately, this approach is not suited for the system, because it will never find a sparse solution and will return a $\widehat{\mathbf{s}}$ vector containing many low amplitude elements.

Alternately, one can define an optimization problem that tries to minimize the number of non-zero elements in $\widehat{\mathbf{s}}$, or in other words, find the sparsest vector consistent with the measurement $\mathbf{y}$. This is equivalent to Eq. 2.5 with $q = 0$, and where $\|\cdot\|_0$, the 0-norm, simply corresponds to the number of non-zero elements in the vector. This ℓ_0-minimization problem is also a poor choice because it is a non-convex problem, and NP-complete. It means that this approach requires to enumerate all of the $\binom{N}{K}$ possible locations of the non-zero entries in $\mathbf{s}$ [8]. Therefore, it is impossible to use the ℓ_0-optimization for practical applications. Actually, the method stays non-convex for $0 < q < 1$, and is in general impossible to solve. Fortunately, for the limit case $q = 1$, the ℓ_1-minimization problem written as

$$\widehat{\mathbf{s}} = \arg\min_{\mathbf{s}} \|\mathbf{s}\|_1 \text{ subject to } \mathbf{y} = \mathbf{A}'\mathbf{s} \tag{2.6}$$

becomes a convex problem, and it is now possible to solve the system. This approach, also known as Basis Pursuit (BP), can exactly recover K-sparse signal, as long as the sensing matrix satisfies RIP of order δ_{2K} [2]. Many extensively studied algorithms exist for solving efficiently the ℓ_1-minimization problem, such as the homotopy method [12] and NESTA algorithm [13].

In order to understand why the ℓ_2-optimization fails to find a sparse solution, Fig. 2.4 depicts the ℓ_1- and ℓ_2-minimization problem in a three-dimensional space. Here, the sparse solution of interest $\mathbf{s}$ is a point located on one of the axes, its coordinate has thus only one non-zero value. In both Fig. 2.4a and Fig. 2.4b, the red structure represents the ℓ_p ball, and is a generalization of the hypersphere in the p-norm. The green plane corresponds to the translated sensing matrix null space $\mathcal{H} = \mathcal{N}(\mathbf{A}') + \mathbf{s}$, with $\mathcal{N}(\mathbf{A}') = \{\mathbf{z} \mid \mathbf{A}'\mathbf{z} = 0\}$. This plane thus contains all the vectors $\mathbf{s}^*$ such that $\mathbf{y} = \mathbf{A}'\mathbf{s}^*$. The ℓ_p-optimization method consists in intersecting the ℓ_p ball with the null space plane $\mathcal{H}$, giving an approximate solution of the problem. It can be seen that the roundness and the isotropy of the ℓ_2 ball will not give a sparse solution because the intersection will most often be far away from the main axes, where the K-sparse subspace vectors are [14]. On the other hand, the pointiness of the ℓ_1 ball represented by an octahedron in $\mathbb{R}^3$ has points aligned with the coordinate axes. It will therefore first contact $\mathcal{H}$ at a point near the coordinate axes, precisely where the sparse vector $\mathbf{s}$ is positioned [8].



(a) **(b)**

Figure 2.4: Representation of the optimization problem in $\mathbb{R}^3$, the sparse solution is located at the point $\mathbf{s}$. (a) Visualization of the l_2-minimization problem, with the ℓ_2 ball in red, and the sensing matrix null space $\mathcal{H}$ in green. The solution of the optimization problem is at the intersection point $\widehat{\mathbf{s}}$ of the two structures. (b) Visualization of the ℓ_1-minimization problem, where the pointiness of the ℓ_1 ball allows finding a sparse solution. Adapted from [8].

So far, the presence of noise was disregarded in the measurements and the sensing system. However, Eq. 2.6 can be generalized taking into account noise such that $\mathbf{y} = \mathbf{A}'\mathbf{s} + \mathbf{n}$, leading to the following algorithm known as the quadratically constrained basis pursuit:

$$\widehat{\mathbf{s}} = \arg\min_{\mathbf{s}} \|\mathbf{s}\|_1 \text{ subject to } \|\mathbf{A}'\mathbf{s} - \mathbf{y}\|_2 \leq \epsilon, \tag{2.7}$$

where ϵ is the size of the noise term $\mathbf{n}$. If the sensing matrix $\mathbf{A}'$ satisfies RIP of order $\delta_{2K} < \sqrt{2} - 1$, it is proved that the error of $\mathbf{s}_K$, the best K-sparse approximation of $\mathbf{s}$, obeys

$$\|\widehat{\mathbf{s}} - \mathbf{s}\|_2 \leq C_{1,K}\epsilon + \frac{C_{2,K}}{\sqrt{K}} \|\mathbf{s} - \mathbf{s}_K\|_1 , \tag{2.8}$$

where $\widehat{\mathbf{s}}$ is the solution to Eq. 2.7, $\mathbf{s}$ is the signal to recover, and $\mathbf{s}_K$ is the best K-sparse approximation of $\mathbf{s}$ [15]. The two constants $C_{1,K}$ and $C_{2,K}$ are function of the restricted isometry constant δ_{2K}. This relation shows that, if the signal to recover is K-sparse, the accuracy of the solution obtained with Eq. 2.7 is bounded by the noise level ϵ only.

2.3.2 Greedy methods

Greedy algorithms are methods that make the locally optimal choice at each iteration, regardless if this choice leads to the global optimal solution [16]. In many computational problems, greedy methods do not produce an optimal solution, but generally can lead to a local solution that approximates well the global solution. The basic principle of greedy algorithms for the case of CS is to iteratively find the support set of the sparse vector $\mathbf{s}$, and use the constrained least-squares estimation method to reconstruct the signal [5].

A common greedy algorithm used for sparse signal recovery is the Orthogonal Matching Pursuit (OMP). This method is based on another reconstruction algorithm called Matching Pursuit (MP). MP tries to find a solution that best fits the signal, while strictly imposing a K-sparse solution:

$$\widehat{\mathbf{s}} = \arg\min_{\mathbf{s}} \frac{1}{2} \|\mathbf{x} - \mathbf{\Psi}\mathbf{s}\|_2^2 \text{ subject to } \|\mathbf{s}\|_0 \leq K. \tag{2.9}$$

Orthogonal matching pursuit performs this algorithm but in a slightly different manner, and is capable to reliably recover a K-sparse signal $\mathbf{s}$ of size N with a number of random linear measurements equal to $M = O(K \ln N)$. It shows comparable performance to basis pursuit, but can recover the signal faster and with more ease in some settings [17]. The idea behind OMP is to determine and select in a greedy fashion the columns of the sensing matrix $\mathbf{A}'$ that take part in the measurement vector $\mathbf{y}$. At each iteration, the most correlated column of $\mathbf{A}'$ to $\mathbf{y}$ is chosen. Subsequently, its influence on $\mathbf{y}$ is subtracted, and the iteration continues on the residual. At the end of the iterations, the algorithm should have found the index set $\mathcal{S}$ comprising the right columns of $\mathbf{A}'$, and as a consequence, the reconstructed signal $\mathbf{s}$. More precisely, the OMP algorithm is structured as follows and is adapted from [3, 17]:

1) Start with the residual $\mathbf{r}^0 = \mathbf{y}$ and the index set $S^0 = \emptyset$. The iteration starts at $t = 1$.

2) Iterate with the following equations until $t = M$:

$$S^{(t)} = S^{(t-1)} \cup \{j_t\}, \text{ with } j_t = \arg\max_{j=1,\ldots,N} \left| \left\langle \mathbf{r}^{(t-1)}, \mathbf{A}'_j \right\rangle \right|, \tag{2.10}$$

$$\mathbf{s}^{(t)} = \arg\min_{\mathbf{z}} \left\{ \|\mathbf{y} - \mathbf{A}'\mathbf{z}\|_2, \text{ supp}(\mathbf{z}) \subset S^t \right\}, \tag{2.11}$$

where $\mathbf{A}'_j$ is the j-th column of the sensing matrix. The residual for the next iteration is $\mathbf{r}^{(t)} = \mathbf{y} - \mathbf{A}'\mathbf{s}^{(t)}$.

2.3.3 Thresholding-based methods

The idea behind thresholding-based algorithms is to iteratively find the solution of the under-determined system and perform a threshold on the solution at the end of each iteration, *i.e.*, only the largest elements of the result are kept for the next iteration. The intuition of this method is hence similar to Eq. 2.9 where the sparsity of the solution is strictly imposed.

One of the simplest algorithms of this category is called the Iterative Hard Thresholding (IHT) [18]. This iterative method performs a gradient descent to tend to the solution, and then proceed to a hard threshold on the result to ensure sparsity of the solution. One starts the iterative process with $\mathbf{s}^{(0)} = \mathbf{0}$ and $k = 0$:

$$\beta = \mathbf{s}^{(k)} + \mu \mathbf{A}'^{\intercal}(\mathbf{y} - \mathbf{A}'\mathbf{s}^{(k)}), \tag{2.12}$$

$$\mathbf{s}^{(k+1)} = \mathcal{H}_S^H(\beta). \tag{2.13}$$

Eq. 2.12 corresponds to a gradient descent controlled by $\mu > 0$, and Eq. 2.13 is the hard thresholding of the estimate β. The thresholding operator $\mathcal{H}_S^H$ is graphically represented in Fig. 2.5a and is expressed as follows:

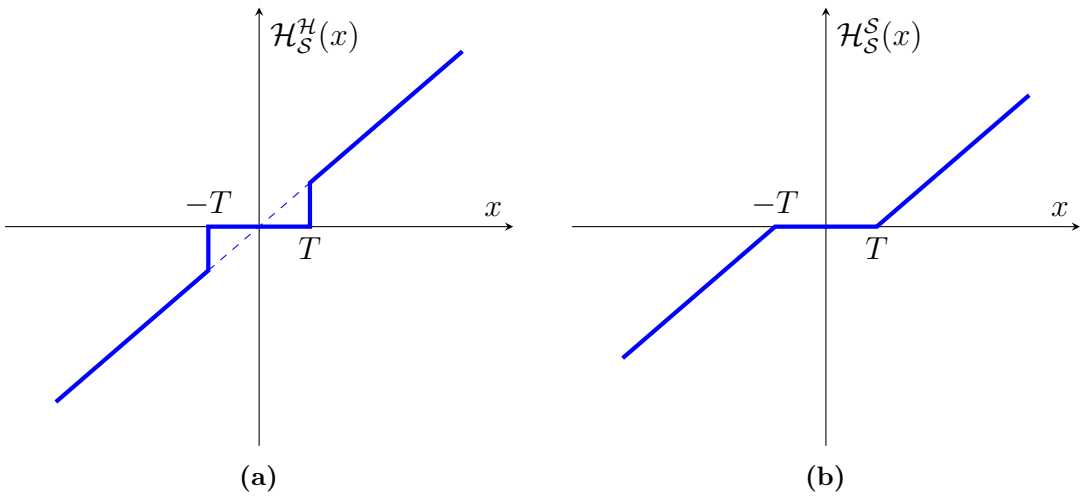
$$\mathcal{H}_S^H(x) = \begin{cases} x, & \text{if } |x| \geq T, \\ 0, & \text{otherwise.} \end{cases} \tag{2.14}$$

$\mathcal{H}_S^H$ allows controlling the sparsity of the solution, by means of tuning T. This parameter can either be a fixed value, or it can change during the iterations.

A variation of IHT is called the Iterative Shrinkage Thresholding Algorithm (ISTA) [19]. This algorithm improves the previous method in the thresholding of the solution estimate at each iteration, by introducing a soft thresholding operator $\mathcal{H}_S^S$. The latter, shown in Fig. 2.5b, scales all the values by T, and then set the elements smaller than T in absolute value to zero:

$$\mathcal{H}_S^S(x) = \begin{cases} x - \text{sgn}(x)T, & \text{if } |x| \geq T, \\ 0, & \text{otherwise,} \end{cases} \tag{2.15}$$

where $\text{sgn}(x)$ is the sign function of x.



(a) **(b)**

Figure 2.5: Thresholding operators. (a) Hard thresholding operator used in the iterative hard thresholding algorithm. (b) Soft thresholding operator used in the iterative shrinkage thresholding algorithm. Adapted from [2].

2.3.4 Deep-learning methods

Even though the previous reconstruction methods have been extensively studied in the literature and their mathematical properties are well known, these techniques are limited for practical applications. Indeed, they suffer from relatively long computational time, which is not suitable for real-time image acquisition. In this context, the introduction of Deep Learning (DL) algorithms in compressive sensing helps to improve the time complexity of the reconstruction often by two orders of magnitude with respect to traditional

recovery methods, because the majority of the algorithm's complexity is shifted to the training phase [11]. In addition, the use of DL has shown improvements in terms of quality of the signal reconstruction. Therefore, DL techniques are good candidates for the implementation of compressive sensing algorithms for real-life application, even though their properties have not yet been thoroughly studied.

There exist different possible ways to implement deep learning in the CS chain, as shown in Fig. 2.6. The first application, depicted in Fig. 2.6a, is to handle the reconstruc-

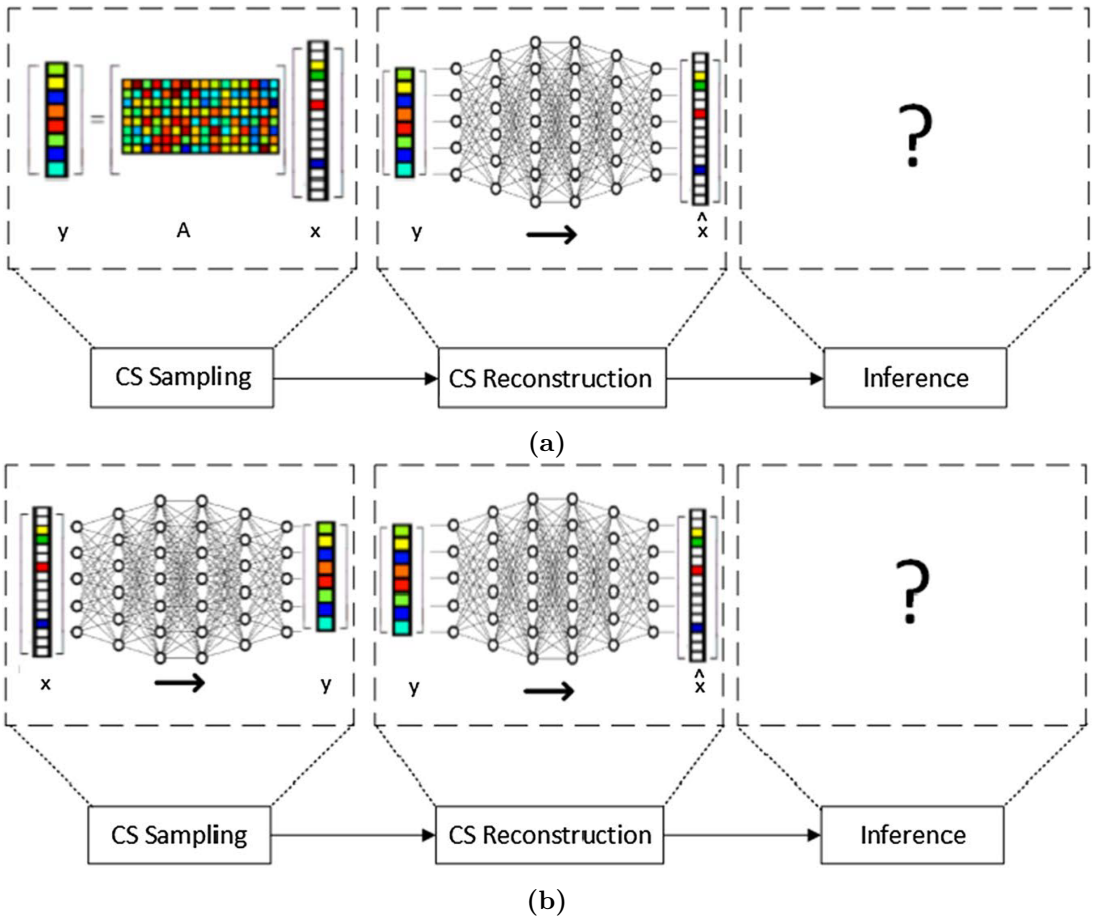


Figure 2.6: Possible ways of DL implementation in a compressive sensing sampling and reconstruction chain. (a) DL is used for the reconstruction of the signal. (b) DL is used for the sampling and the reconstruction process. From [11].

tion of the signal with a deep learning algorithm, based on the collected samples. It can be divided into two subtypes: iterative reconstruction or direct reconstruction. The first tries to copy the iterative methods mentioned in the previous sections. This technique is based on an algorithm architecture called algorithm unrolling: it maps each iteration into a network layer, and each layer is eventually stacked. The weights of the networks are to be learned, and they are optimized during the iterations. ISTA-Net$^+$ [20] is an example of this type of algorithm. On the other hand, the DL algorithm for direct reconstruction does not imitate the iterative schemes, but reconstructs the sparse signal with a complete neural network. This method offers a great flexibility in its implementation. However, it

has more of a black box nature, which makes it less interpretable, and it requires more learning parameters than DL iterative reconstruction for comparable recovery quality [21].

Secondly, the sampling matrix can also be controlled by a DL network, as it is shown in Fig. 2.6b. In this manner, the sampling matrix is learned with the goal of optimizing signal acquisition. Applied to CS imaging, this technique determines the most effective way to code the image for efficient sensing [22].

2.4 The case of inpainting

While compressive sensing focuses on capturing a signal, the reconstruction algorithms mentioned above and the sparsity principle also finds utility in the reconstruction of damaged images, where part of the information is missing. This technique, called inpainting, consists indeed in solving an underdetermined system, where the measured signal $\mathbf{y}$ is the damaged image, and the objective is to retrieve the undamaged image, *i.e.*, $\mathbf{x}$.

The damaged image of size $n \times n = N$ can be linked to the original image by the element-wise product of the undamaged image with a binary mask:

$$\boldsymbol{f} = \boldsymbol{\Phi} \odot \boldsymbol{f_0} \tag{2.16}$$

where $\boldsymbol{f} \in \mathbb{R}^{n \times n}$ is the damaged image, $\boldsymbol{f_0} \in \mathbb{R}^{n \times n}$ is the original image to recover, and $\boldsymbol{\Phi} \in \mathbb{R}^{n \times n}$ is the binary pattern mask defining which pixel has to be reconstructed. Eq. 2.16 can easily be rearranged into the compressive sensing basic vector equation $\mathbf{y} = \mathbf{A}\mathbf{x}$, with $\mathbf{A} \in \mathbb{R}^{M \times N}$. The value of M is equal to the number of original pixels of the image, and therefore the number of one-valued pixels in $\boldsymbol{\Phi}$. Knowing the binary mask, *i.e.*, the location of the holes in the image, the latter can be reconstructed with one of the recovery methods shown in Sec. 2.3.

2.5 Optical multiplexing technologies

Now that the mathematical and algorithmic formalisms are well established, this section aims to introduce the different optical multiplexing devices that are used to code the image in the CS acquisition process. Borrowed from the field of telecommunications, the term multiplexing means the combination of signals to share a scarce resource, such as a communication cable. In the field of CS, signal information shares the scarce resource of limited measurements [2]. These devices can be categorized into passive and active multiplexing. Passive technologies are composed of static optical elements that are not controllable. On the other hand, active multiplexing allows changing the optical multiplexing of the scene during its recording, via a Spatial Light Modulator (SLM). In the following of this section, two passive and two active multiplexers will be presented.

2.5.1 Coded aperture

A coded aperture is a type of passive multiplexer, made of a perforated opaque material [2]. The pattern formed by the holes on the aperture constitutes a binary mask, and allows the incident light to be multiplexed. Coded apertures first found utility in X-rays and γ-rays imaging, where the use of lenses is made impossible for these types of radiations, and where massive grazing incidence telescope mirrors are not feasible. This type of aperture

can drastically reduce the complexity of the imager, simplifying the system down to a coded aperture in front of a FPA, as shown in Fig. 2.7. The coded aperture design is an improved version of the pinhole camera, which is the simplest form of image modulation, but that suffers from high light losses through the very limited aperture. Coded apertures hence increase the amount of light coming to the detector by adding multiple holes. As a consequence, this introduces a more complex point spread function than that from a single pinhole, but it is used advantageously to reconstruct the images with the knowledge of the pattern on the aperture [23].

Several families of patterns have been studied for coded aperture imaging. It has been proven that randomly distributed pinholes with 50% opening can efficiently image stars in the X-rays domain [24]. Structured patterns have also been explored, such as Uniformly Redundant Array (URA) [25] and Modified Uniformly Redundant Array (MURA) [26].

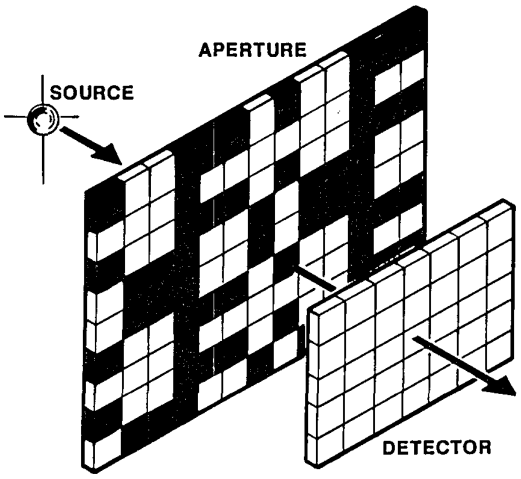


Figure 2.7: Schematic of a coded aperture detector. From [26].

2.5.2 Diffuser

One way to obtain a passive pseudo-random multiplexing of the image is to use diffusers. A diffuser is a transparent material that refracts and scatters the incident light in a diffusive manner, thanks to its surface roughness [2]. As shown in Fig. 2.8, a diffuser has a spatially random refractive index, which results in a pseudo-random coding of the scene. This spatially varying refraction induces a spatial phase modulation of the signal, in contrast to the coded aperture, which modulates the amplitude. This phase modulation creates caustic patterns on the detector. Because this pattern is not predictable in advance, the diffuser has to be calibrated by imaging a point source far from the detector. Any other points located on the scene will shift and/or change size of the caustic pattern on the camera sensor. A difference with coded aperture is that a diffuser does not block light, but refracts it. Therefore, more light can hit the detector and the system can provide a better Signal-to-Noise Ratio (SNR) than coded apertures [2].

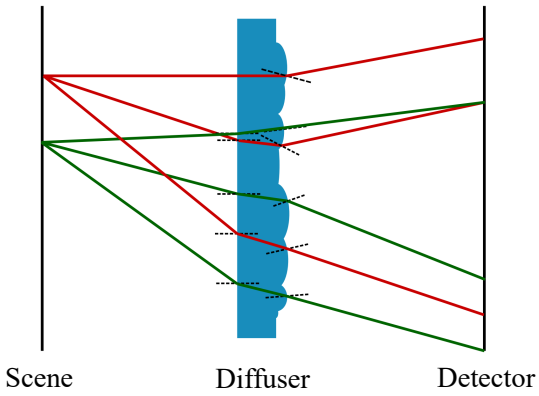


Figure 2.8: Schematic of diffuser used to randomly multiplex information from the scene onto the camera detector. Adapted from [2].

2.5.3 Digital micromirror device

A Digital Micromirror Device (DMD) is a microoptoelectromechanical system constituted of a rectangular array of tiltable micromirrors [27]. Consequently, it constitutes an active multiplexing system. Each mirror has a typical size of about 10 µm and can be rotated around their diagonal axis with a torsional hinge, as it can be seen in Fig. 2.9. The mirrors have two stable orientations noted as an ON/OFF state, and usually corresponding to a ± 10 ° or ± 12 ° tilt respectively. The orientation of each mirror is electrostatically controlled with a memory cell formed from two CMOS memory elements, the memory state stored in the memory cell is transferred to the mirror by a clocking pulse. A DMD allows coding binary patterns at a frequency of several tens of thousands of Hertz. When displaying a pattern on the device, a normal incident light is either reflected to one or another direction, and the reflected light is thus coded according to the pattern. In addition, one can play with the fast pattern rate and the duration of the tilt of each mirror to tune precisely the reflected light of the mirrors. This allows creating shades of gray in the pattern mask, rather than a simple binary pattern, at the cost of a slower pattern rate. Due to the small size of the micromirrors and their periodicity on the 2D array, a DMD can be seen as a grating for wavelengths of the same order of magnitude as the dimensions of the mirrors [28]. Therefore, the reflected light will be diffracted, and create unwanted rays, leading to reduced performance. The usable wavelength range of the DMD is by consequence limited by diffraction effects.

DMDs have originally been designed by Texas Instruments and are part of the Digital Light Processing technology to display images with video projectors. DMDs have enabled the miniaturization of such devices.

2.5.4 Liquid crystal spatial light modulators

A Liquid Crystal (LC) modulator is a type of active SLM, which uses the optical anisotropy of liquid crystals to modulate the incident light [30]. This property is due to the long molecules which compose the crystal and that have a tendency to get aligned to each other. This anisotropy, such as birefringence, depends on the orientation of the crystal molecules, which can be modified by the presence of an electric field. The incident light

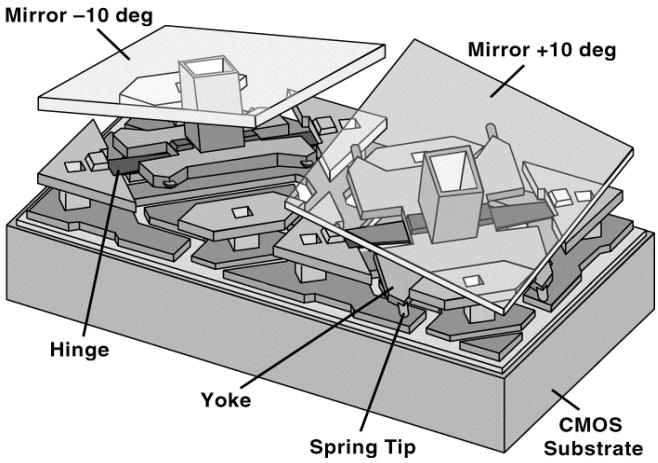


Figure 2.9: Schematic of the mechanism inside the pixel mirrors of a DMD with the different components highlighted. Figure from [29].

can therefore be modulated by controlling the applied electric field on the LC. In the context of compressive sensing, the LC equivalent of the DMD is the Liquid Crystal On Silicon (LCOS). It is a small two-dimensional array of individual LC modulators, allowing the spatial patterning of light. As shown in Fig. 2.10, each liquid crystal pixel is built on a silicon backplane, which contains CMOS electronics for the control of the individual modulator. Above this backplane and below the liquid crystal layer lays an aluminum layer, which reflects the incident light. This modulation device is therefore used in reflection, in the same fashion as DMDs.

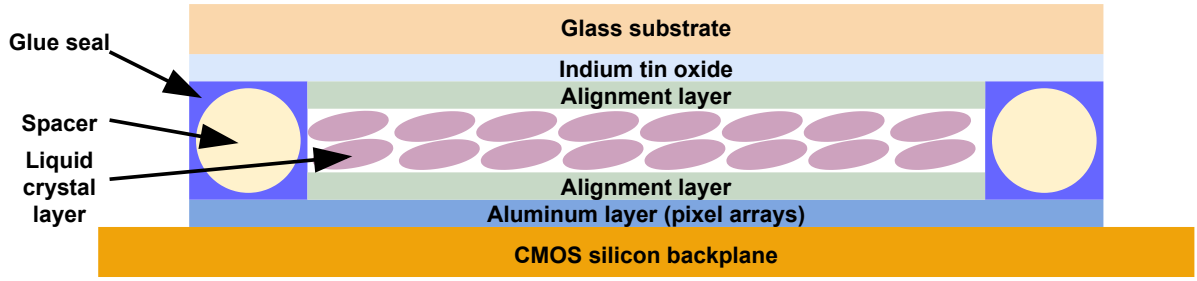


Figure 2.10: Structure of one pixel of a LCOS device. On the top, each pixel has a glass substrate, and a transparent indium tin oxide electrode. The LC layer is between two alignment layers that define the preferred direction of the crystals. Below that are placed the reflective aluminum coating and the CMOS silicon backplane. Each pixel is separated from each other by a glue seal and a spacer. Adapted from [31].

The modulation of light can be done in amplitude, in phase, and in grayscale. On the other hand, liquid crystal SLM are typically the optical modulators with the lowest refresh rate, being limited to around 100 Hz. This slow frame rate can be compromising for real-time CS applications.

2.6 Imaging devices using compressive sensing

This next section introduces several imaging device architectures that use compressive sensing to acquire images.

2.6.1 Single-pixel imaging

The single-pixel camera is certainly one of the most well-known imaging applications of compressive sensing. As its name implies, it is a camera composed of only one single photon detector [32]. Traditionally, cameras are equipped with a Focal-Plane Array (FPA) made of millions of pixels. This architecture of detector, boosted by the democratization of CCD and CMOS technology, allows acquiring high-quality images mainly in the visible domain. However, this comes at the cost of a large quantity of data to handle, and more importantly, imaging with FPA in wavelengths that are invisible to silicon such as infrared is substantially more challenging and expensive than in the visible, for comparable image resolution. It is in this context that Single-Pixel Imaging (SPI) was proposed, with the aim to build a cheaper and simpler camera, capable to image in a broader range of wavelengths than silicon-based FPA. Indeed, using only one detector pixel enables the use of exotic detectors that can operate in specific wavelengths, and that would be impossible to use in classical focal-plane arrays. The use of a single-pixel detector can offer improved performance with respect to FPA, in terms of quantum efficiency, dark counts, and thus can have a better imaging in low light conditions [1].

The general architecture of a single-pixel camera is shown in Fig. 2.11. The incident light of the scene is focused by a collecting lens on a DMD, which modulates the image. A condensing lens focuses the reflected light on a photodetector that measures the signal. The DMD displays random or structured patterns, for each of which an individual light

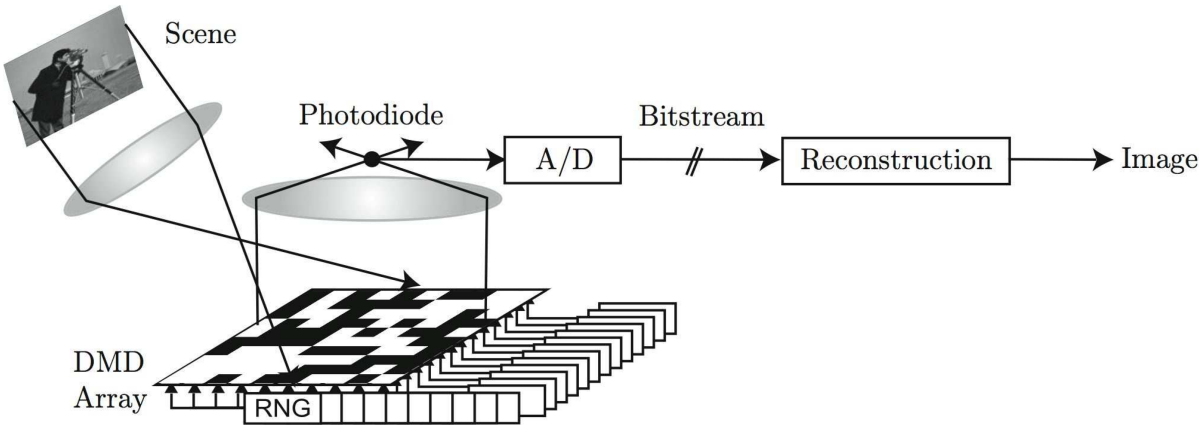


Figure 2.11: General architecture of a single-pixel imager. Figure taken from [33].

intensity is measured with the photodetector. The resulting set of measurements is fed to the proper compressive sensing algorithm to retrieve the image. Because of its interesting characteristics such as its fast modulation rate, the DMD has been the preferred choice of SLM since the first developments of SPI [34]. For one binary pattern $\mathbf{A_m}$ applied on the DMD, the photodiode measures $y_m = \langle \mathbf{A_m}, \mathbf{x} \rangle$, which corresponds to the inner product of the mask and the scene. By taking M measurements with different masks, it allows

returning to Eq. 2.1. The spatial resolution of SPI is hence limited by the resolution of the DMD.

The SPI architecture presented in Fig. 2.11 is not the only possibility to acquire spatial information of a scene. This DMD can also be used to illuminate the scene in a structured manner. As illustrated in Fig. 2.12, the light that illuminates the scene is coded by a DMD. The backscatter of the exposed scene is then returned to a photodetector, where its intensity is measured. By consequence, this reciprocal configuration is a passive sensing method. Equipped with a pulsed laser as illumination source, the structured illumina-

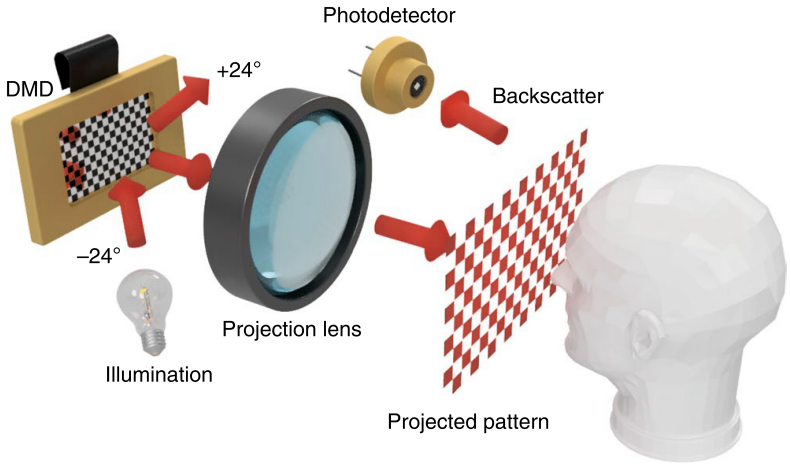


Figure 2.12: General architecture of the structured illumination single-pixel imaging. Figure from [1].

tion single-pixel architecture can be used to perform time-of-flight imaging, allowing 3D representation of a scene [34]. Equipped with a time-resolving detector, one mask of the SLM now corresponds to a series of measured intensities at different depths. A 3D image cube can then be obtained by associating the different masks with their corresponding intensities at different depths. By processing this data cube, the reflectivity and depth information of the scene can be retrieved [35].

The simplification of the architecture of the imager and the broadening of the spectral range with the use of a single-pixel detector comes with a price. Indeed, with only one measuring element, the different measurements needed to reconstruct the image have to be taken sequentially, which lengthens the acquisition time of a scene. On the other side, a classical imager equipped with a FPA can acquire a picture in one shot. Therefore, there is a space-time trade-off to make between the number of measurements and the number of pixels in the camera. It is thus essential to have fast-modulating SLMs to shorten the duration of the acquisition. Additionally, the reconstruction of the image can also take a substantial amount of time, which handicaps SPI for real-time applications. Nevertheless, it is possible to make a compromise between single-pixel and classical FPA imaging by using more than one pixel but less than the resolution of the imager. The number of measurements is then divided by the number of pixels present on the detector.

Single-pixel imaging and its trade-off version with multiple pixels introduce the concept of Super-Resolution (SR) imaging. SR is an imaging configuration where the final image

has a greater resolution than the detector array. For the case of CS imaging, the resolution of the image is determined by the resolution of the SLM that multiplexes the image. The number of pixels on the detector can therefore be fewer than the number of pixels on the SLM. The super-resolution factor is used to measure the degree of SR of an imager. This factor corresponds to the size of the group of SLM micro-pixels that are integrated into one single detector pixel [36]. The working principle of super-resolution is illustrated in Fig. 2.13. It is clear that the single-pixel imager is the extreme case of super-resolution, with a SR factor equal to the whole size of the SLM.

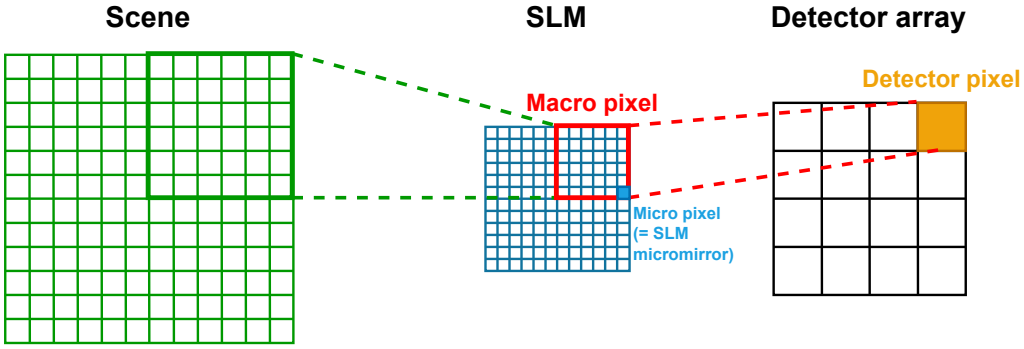


Figure 2.13: Working principle of the super-resolution. Adapted from [36].

2.6.2 Coded aperture spectral imaging

As its name implies, compressive sensing has the ability to acquire a signal compressively. By consequence, the technique is of interest in domains where large amount of data are involved, such as hyperspectral imaging. Therefore, several CS spectral imager architectures have been proposed in the literature, including the Coded Aperture Snapshot Spectral Imager (CASSI). This imager capable of sensing an entire data cube with just a few FPA measurements, and in most cases, in a single FPA shot [37].

There exist several optical architectures of CASSIs. The simplest design is certainly the single disperser CASSI, which is presented in Fig. 2.14. It has three main components: a coded aperture used to code the image, a dispersive element used to separate the different wavelengths of the data cube, and finally, a FPA detector to take the measurements. The

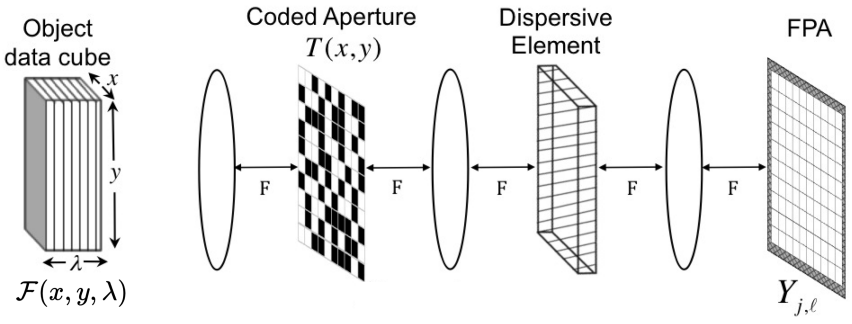


Figure 2.14: Optical arrangement of single disperser CASSI. Adapted from [38].

sensing mechanism of the imager is illustrated in Fig. 2.15. The data cube $\mathcal{F}(x, y, \lambda)$ of

the scene has $N \times N$ spatial pixels and L different spectral bands. One vertical slice of this data cube hence represents the scene viewed at one specific wavelength. This data cube is first modulated by the coded aperture $T(x, y)$ of size $N \times N$, which has the effect of allowing the spectral information of only some selected spatial pixels to pass through. Next, the modulated data cube traverses a dispersive element such as a prism. The spectral dispersion induces a spatial shear of the data cube along one direction, because each vertical slice will be shifted proportionally to the wavelength of that slice, *i.e.*, its spectral position in the data cube. Lastly, the sheared data is integrated and measured by means of the FPA detector. The dispersion of light implies the detector to be larger in the direction of dispersion, along x in this case, in order to completely measure all the data matrix. To have completely resolved spectral bands, the detector should have a size of $N \times (N + L - 1)$.

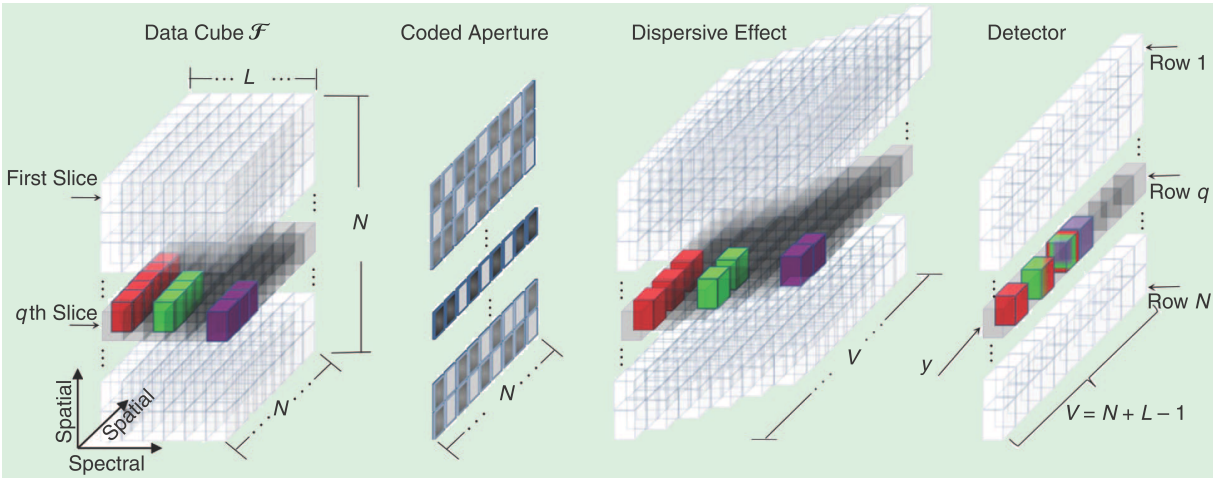


Figure 2.15: Schematics of the spectral optical flow of CASSI. The q-th horizontal slice of the data matrix $\mathcal{F}$ containing the spectral information of a horizontal row of the scene is coded by its corresponding row of the coded aperture and then dispersed by the prism. The detector then measures the intensity of the coded and dispersed light. From [37].

The data cube $\mathcal{F} = [\mathcal{F}_0, \ldots, \mathcal{F}_{L-1}]$ with L spectral bands can be denoted by $\mathcal{F}_{j,\ell,k}$, where j, ℓ are the two spatial indices and k is the index of the wavelength band. The modulated and dispersed cube is integrated on the detector and the measured intensity of the pixel (j, ℓ) on the $N \times (N + L - 1)$ FPA array is [39]

$$Y_{j,\ell} = \sum_{k=0}^{L-1} \mathcal{F}_{j,(\ell+k),k} T_{j,(\ell+k)} + \omega_{j,\ell}, \tag{2.17}$$

with $\omega_{j,\ell}$ the additional noise of the system. Note that this equation is only valid for the general case near the center of the detector, but not necessarily at its borders. For simple scenes, a single shot might be sufficient to compressively capture their spatial and spectral information. However, for more complex scenes, multiple shots with different coded apertures are required. For the i-th measurement, Eq. 2.17 can be written in a vectorized form:

$$\mathbf{y}^i = \mathbf{H}^i \mathbf{f} + \boldsymbol{\omega}^i, \tag{2.18}$$

where $\mathbf{H}^i \in \mathbb{R}^{N(N+L-1)\times(N^2L)}$ combines the effects of the i-th coded aperture and the dispersive element, and $\mathbf{f}$ is the data cube of the scene in vector form. The different measurements can be regrouped as $\mathbf{y} = [(\mathbf{y}^0)^\top, \ldots, (\mathbf{y}^{K-1})^\top]^\top$, with K the total number of acquisitions. The expression of Eq. 2.18 of each shot can be regrouped into one system of equations

$$\mathbf{y} = \mathbf{H}\mathbf{f} + \boldsymbol{\omega}. \tag{2.19}$$

The problem is now stated in the standard form of an underdetermined system of linear equations, which can be solved with the adequate algorithms of Sec. 2.3.

2.6.3 Lensless imaging

A lensless camera is an optical instrument that is not equipped with any complex optical elements such as lenses or mirrors. It is only composed of a diffuser in front of a focal-plane array detector. This type of camera offers a certain number of advantages compared to conventional lens-based devices [40]. The use of lenses requires to have a certain focal length between them and the detector. By contrast, lensless cameras can have the diffuser located very close to the sensor, reducing incomparably the total size of the imager. In addition, the camera can be much lighter, by removing the weight of the lenses, which makes up most of the mass of traditional cameras. The cost can also be drastically reduced, because high-quality optics is no longer required. Furthermore, lensless cameras can provide a better Field Of View (FOV) than lens-based cameras for comparable resolution, which can break the FOV versus resolution trade-off of traditional optical devices.

On the opposite, there are some limitations in the use of lensless imaging [40]. Most importantly, the quality of the images is not as good as lens-based imagers. Hence, for 2D-imaging applications that do not require the benefits of lensless cameras such as cost, weight, and size, traditional imaging is preferred. Another limiting factor, which is inherited from compressive sensing, is the reconstruction time. As for the previous CS imagers, the recovery of the scene with computational algorithms adds data processing difficulties, additional power consumption, and can make it difficult to use lensless cameras for real-time applications. Moreover, the light collected by the camera is limited by the size of the detector, while camera lenses are often designed to be larger than the detector to collect more light.

One of the most well known lensless camera is the DiffuserCam, developped by [41]. This camera is capable of taking one-shot 3D images of a scene. The DiffuserCam has a diffuser made of a thin transparent phase plate with smoothly varying thickness, placed at a small distance in front of the sensor. As mentioned in Sec. 2.5.2, the diffuser modulates the phase of the incident light and, for a point source placed in the scene, forms pseudo-random caustic patterns on the detector, which is the Point Spread Function (PSF) of the system. As shown in Fig. 2.16, a lateral shift of a point source causes a lateral translation of the caustic pattern, and a change in depth causes a scaling of the PSF. Therefore, different point sources located in the 3D scene will create a unique caustic pattern on the detector. A scene to image can be assumed to be a combination of different incoherent point sources. The measured intensity on the detector of the DiffuserCam can then be modelled as a linear combination of all the different caustic patterns of these points. This

problem can be written in the form of [41]

$$\mathbf{b} = \mathbf{H}\mathbf{v}, \tag{2.20}$$

where $\mathbf{b}$ is the vector of the 2D sensor measurements, and $\mathbf{v}$ is the vector containing the intensity of all the 3D points of the scene to image. $\mathbf{H}$ is called the forward model matrix, and contains in its columns the different caustic patterns created by all the 3D points of the scene. The shape of the PSF can be obtained with a calibration procedure by imaging a point source at different positions. Such DiffuserCam and its associated calibration procedure and processing were also developed in the master's thesis of S. Gramegna, as first exploration of CS imaging at the CSL [42]. Eq. 2.20 is an underdetermined inverse problem, and can be solved with CS algorithms, relying on the sparsity of the 3D scene in some basis. However, the addition of a third dimension increases dramatically the amount of data to handle, and the computation time. These issues can be solved by implementing specific algorithms for 3D problems.

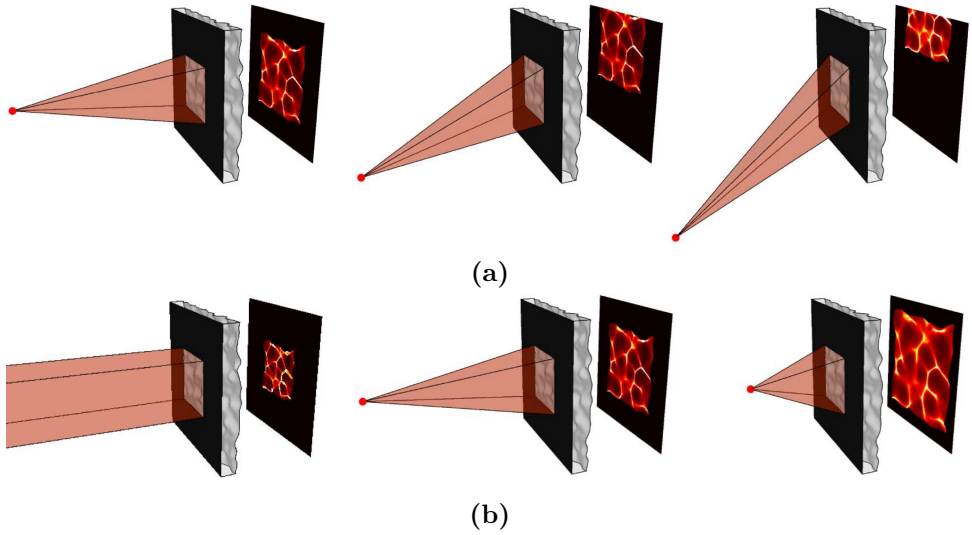


Figure 2.16: Effect of the lateral and axial displacement of a point source on the pattern created on the camera detector of the DiffuserCam. (a) Effect of a lateral shift. (b) Effect of a change in depth. From [41].

3 Optical setup of a compressive sensing imager

The main objective of this work is to build and improve a complete compressive sensing imaging device. This chapter will present and explain the design and realization of a CS imager in the visible light in laboratory. First, its general layout will be discussed. Afterwards, the choice of the optical elements and their final disposition will be analyzed. The computer control of the built imager will eventually be detailed. This work falls within the continuity of the master's thesis of C. Thomas [43] at the CSL.

The realization of a true CS imager such as a single-pixel camera is particularly challenging, both at the level of the optical system and at the level of image reconstruction. The largest difficulty lies in the acquisition of the large number of image samples. For example, compressively sensing an image with a resolution of $256 \times 256 = 65536$ pixels with CS ratio of only 10 % requires taking 6554 measurements of the image with the same number of different masks. As a consequence, the pattern rate of the SLM used in the imager must be sufficiently high to be able to capture such samples in an acceptable amount of time. Such pattern rate is currently not reachable in the CSL laboratory, and would require to optimize the SLM. The reconstruction process must also be optimized to handle these data. As a beginning, it has thus been decided to start with a simpler reconstruction algorithm: image inpainting.

Starting with inpainting has several advantages. First, the reconstruction process and the algorithm are simpler to put in place. Secondly, the optical layout is very similar to a CS imager such as a single-pixel camera. The image is intentionally damaged by the masking of the SLM, and then captured with a FPA requiring only one acquisition to perform the reconstruction. Therefore, it constitutes a good means to familiarize with the optical constraints of CS imagers, and it lays the groundwork for future work in the realization of CS cameras.

3.1 General layout

In the following of this section, the general optical layout of the setup will be described. As it has been explained in Sec. 2.5 and Sec. 2.6, there exist a multitude of different designs of compressive sensing imagers, with a large choice of modulating devices. The goal of this inpainting imager is to intentionally damage the image via a modulating device and then acquire the image, with the aim to reconstruct it afterwards with CS algorithms. In fact, the damaging operation is identical to the coding of a scene done by a CS imager to perform one measurement. Therefore, the general layout of the imager will have similarities with the architecture of a CS imaging system.

After reviewing the existing configurations that have already been done in the literature, the layout presented in Fig. 3.1 has been chosen. This simple configuration is similar to the one used for single-pixel imaging [32]. The imager is first composed of a collecting

lens that gathers the light coming from the scene. This lens focuses the light on the surface of a DMD that will serve to damage the image. As it has already been seen, a DMD is an array of micro-mirrors that can tilt in ON and OFF directions via an electronic control. The light focused on the ON-state mirrors is reflected toward a condensing lens, while the OFF-state mirrors direct the light out of the imager. The condensing lens collects the light coded by the DMD and focuses it on a camera detector that will capture the modulated image. Contrary to SPI, the detector here is a FPA sensor, because the goal is to obtain a good-quality image, but not to perform any super-resolution.

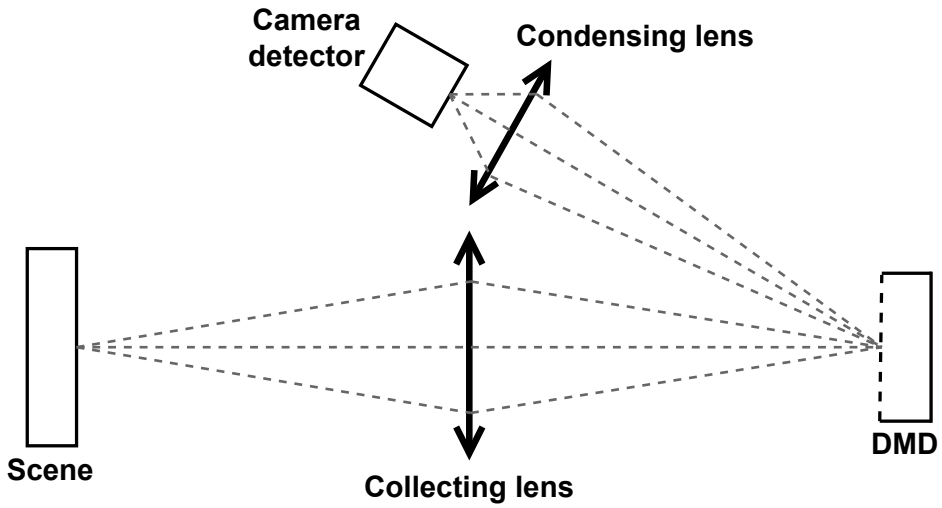


Figure 3.1: General layout of the imager. The gray dashed lines indicate some rays of light coming from the scene and projected on the detector via the DMD.

It has been seen that the square mirrors of a DMD rotate around their diagonal line. Therefore, if the DMD is placed vertically on the laboratory bench with its longest sides parallel to the surface of the table, a ray parallel to the table hitting a tilted mirror will be reflected with a 45° angle with respect to the plane of the table. This effect is not desirable for practical reasons. Indeed, in this configuration, the secondary arm composed of the condensing lens and the detector would have to be tilted and lifted above the table, which is not functional for precise tuning of the position of these elements. Instead, it is easier to have the rays of light that stay parallel to the bench, so that the optics stays in the plane of the table. This is achieved by rotating the DMD by 45° around the optical axis. As a consequence, the scene element as well as the camera detector will also have to be rotated by 45° to have the image well contained within the sensor array.

3.2 Description of the DMD

Since the DMD is a central component in the optical setup of the imager, it is important to first describe its characteristics and specifications. The DMD model available in the optical laboratory and used in the assembly is the DLP7000 manufactured by Texas Instruments [44]. A picture of this DMD is shown in Fig. 3.2 and its main characteristics are summarized in Tab. 3.1.

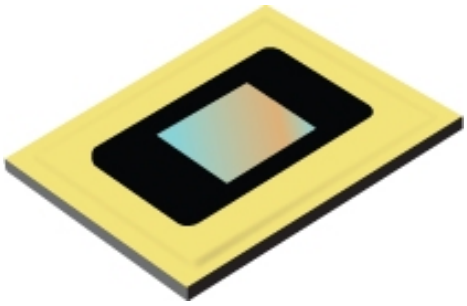


Figure 3.2: DLP7000 [44].

Table 3.1: Main characteristics of the DLP7000. From [44].

Parameter	Value
Number of columns (micromirrors)	1024
Number of rows (micromirrors)	768
Micromirror pitch (µm)	13.68
Micromirror array width (mm)	14.008
Micromirror array height (mm)	10.506
Micromirror tilt angle	±12°
Maximum binary pattern rate (Hz)	32552

The DLP7000 is a high-resolution DMD, with high-performance for spatial light modulation. It is optimized for use in visible light (between 400 and 700 nm), with a reflectivity of its micromirrors reaching 88%.

A DMD is a pertinent choice of SLM for the imager, because inpainting requires a binary coding in amplitude of the image while having a full control of the damaging of the scene. This is precisely what the DMD is capable of. The computer control of the DLP7000 will be explained in more detail in Sec. 3.5.

3.3 Choice of the optical elements

In this section, the choice of the different optical elements constituting the laboratory setup will be described in detail, as well as their characteristics. Each element will be discussed starting by the scene to image, and following the light path to finish with the detector.

3.3.1 Scene object

The setup has to be designed to acquire 2D images of a scene, in the visible light. Because it is a laboratory demonstration, one has a total control of the characteristics of the scene, such as its luminosity and the elements present in it. The scene must meet a number of criteria. Firstly, it must be sufficiently bright to provide enough photons to the imager. If its luminosity is too low, the Signal-to-Noise Ratio (SNR) of the instrument will not be high enough, but the luminosity should not be too high either to avoid saturating the detector. Secondly, this scene must be easy to set up in the laboratory and readily interchangeable. This is in order to be able to easily image different scenes in a reduced amount of time, and without too many manipulations.

The solution chosen for the display of the scene is an image printed on a sheet of paper, and held vertically by the mechanism shown in Fig. 3.3. This method allows for obtaining a scene with good-quality images, and easily interchangeable since it simply requires replacing the paper image on the support. This solution also ensures that the entire scene is at the same distance from the imager, thereby enabling a sharp focus across the entire image. The illumination of the scene is managed by a desk lamp positioned in front of the image and out of the imager's Field Of View (FOV). In this way, the scene is uniformly illuminated, and in a controlled and repeatable manner. In addition, standard printer paper is not a shiny material, its surface being relatively rough, giving it a diffuse

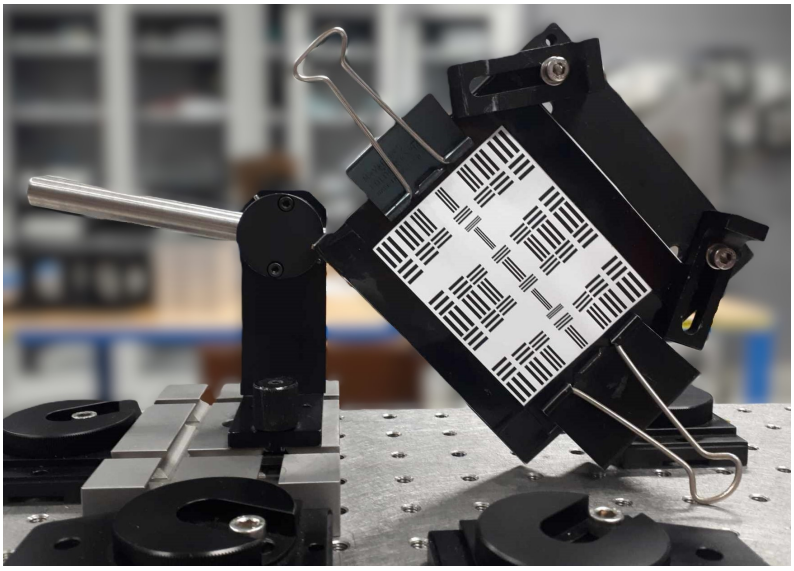


Figure 3.3: Mechanism holding the scene in the right position and orientation on the laboratory bench. The scene is printed on a sheet of paper and held flat at 45° with respect to the table. Here, the scene is a calibration target made to test the instrument.

reflectivity property. Therefore, the light of the lamp will be reflected uniformly towards the imager, thus avoiding the presence of unwanted specular reflection.

Regarding the dimensions of the scene and its position with respect to the instrument, several parameters have to be taken into account. First, for practical reasons, the instrument and the scene should both be positioned on the laboratory table. By consequence, the two objects cannot be spaced more than the usable length of the table, which is the distance between the DMD and the end of the table. Secondly, the dimensions of the scene and its distance from the imager can be summarized by one parameter: the field of view. The FOV is the solid angle through which the imager sees the scene, and is equal in the case of a rectangular scene to [45]

$$\mathrm{FOV} = 4 \arctan \frac{wh}{2d\sqrt{4d^2 + w^2 + h^2}}, \tag{3.1}$$

with w and h, the width and height of the scene, and d, its distance to the imager. It means that, in terms of FOV, a small scene close to the aperture of the instrument is equivalent to a large scene placed further away. The field of view is an important parameter in the design of optical instruments, as it is linked to the image quality. Indeed, a larger FOV increases optical aberration that can degrade the quality of the pictures [46]. In order to achieve a good image quality, a small FOV is thus preferred. On the other hand, because the scene is an image printed on paper, it should be sufficiently large to have good printed details on it. Based on that, the scene must be positioned as far away as possible from the imager. The precise position of the scene will be determined in the next section, when the collecting lens will be chosen.

Concerning its dimensions, the height of the image is set to be 65 mm, which allows obtaining a good printing quality of detailed images, and also provides a relatively small

FOV, as it will be seen during its calculation in the following of this work. This height is also based on the size of a phone screen, which was used as a scene by [43] in a similar optical configuration than the one shown here. The width of the image is determined by the aspect ratio of the screen of the DMD, because all the image has to be displayed on it. The DMD used in this laboratory setup has an aspect ratio of 4 : 3 (see Sec. 3.2). This gives a scene width of 87 mm.

3.3.2 Collecting lens

The collecting lens is placed between the scene and the DMD, as shown in Fig. 3.1. The role of this lens is to collect the light coming from the objects of the scene and focus this light on the screen of the DMD to produce a sharp image on it. In this section, the characteristics of the collecting lens will be determined.

The main physical properties of a lens are its focal length f, and its diameter ϕ. These two parameters influence the qualities of optics. For a given focal length, a larger diameter allows for more light to be collected, resulting in a brighter image, but it also increases aberrations as it deviates from the paraxial approximation. On the other hand, for a given diameter, a longer focal length increases the magnification and depth of field of the image but reduces its field of view, brightness, and increases the effects of diffraction. It is clear that these two parameters are interconnected, and this is mediated through the f-number

$$F\# = \frac{f}{\phi}. \tag{3.2}$$

This dimensionless parameter, often used to characterize the performance of telescopes, expresses the angular aperture of a lens. The angular aperture θ of a lens focused at infinity is the half-angle made by a marginal ray with the optical axis, as illustrated in Fig. 3.4. An optical system with a very low f-number can be a source of severe aberrations in the image, necessitating complex optics to correct them. In the case of a simple lens, a too low f-number would induce too many aberrations and greatly impair image quality. Conversely, a lens with a large f-number would take up more space on the laboratory bench and also produce a dimmer image. Determining the f-number of the collecting lens, *i.e.*, its focal length and diameter, is therefore a compromise, among other factors, between optical quality, image brightness, and device size.

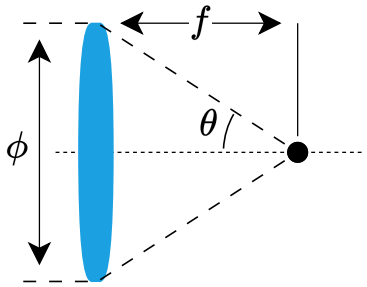


Figure 3.4: Schematic illustrating the angular aperture defined by the f-number.

The choice of the collecting lens is subject to several geometric constraints within the

optical system. The first obvious constraint is that the lens must be available in the laboratory's optical store. Since this store is not infinite, the selection of the lens must be limited to the available materials. Secondly, the size of the optical table is of finite length, and the entire primary arm of the imager consisting of the scene, the collecting lens, and the DMD must be accommodated on this table. The geometry of the primary arm is illustrated in Fig. 3.5. The object distance s_o and the image distance s_i are varying as a function of the focal length of the lens, accordingly to the thin lens equation

$$\frac{1}{s_o} + \frac{1}{s_i} = \frac{1}{f}. \tag{3.3}$$

In the case of simple and relatively thin lenses, and within the scope of a simple setup on an optical bench, the thin lens approximation is considered acceptable. Additionally, the transverse magnification M_T of a lens is a dimensionless parameter that represents the ratio of the height of the image and the height of the object, and can also be expressed in terms of image and object distances:

$$M_T = \frac{h_i}{h_o} = \frac{s_i}{s_o}. \tag{3.4}$$

The transverse magnification of the collecting lens is thus equal to the ratio between the width (or height) of the scene and that of the DMD. The dimensions of the DMD are 14×10.5 mm (see Sec. 3.2), which gives $M_T = 0.16$, which means that the image is about 6.25 times smaller than the object.

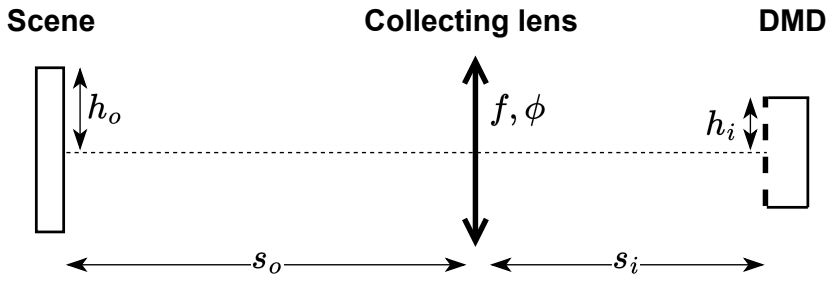


Figure 3.5: Schematic of the primary arm, including the scene, the collecting lens, and the DMD.

In the simple geometry of Fig. 3.5, the total distance between the scene and the DMD should not be larger than the usable length of the laboratory bench

$$s_o + s_i \leq l_t, \tag{3.5}$$

where $l_t = 1.35$ m is the usable length of the table. From this inequality, Eq. 3.3 and Eq. 3.4, a condition on the focal length of the lens can be found

$$f \leq l_t \frac{M_T}{(1 + M_T)^2}. \tag{3.6}$$

Replacing the variables with the numerical values, the upper limit of the focal length is found to be $f < 162$ mm. The object and image distances, as well as the total distance between the scene and the DMD as a function of the focal length of the collecting lens is shown in Fig. 3.6, with the maximum focal length allowed highlighted. These distances increase linearly with the focal length. This graph indicates the relative position of each element of the primary arm, depending on the focal length of the lens that will be chosen.

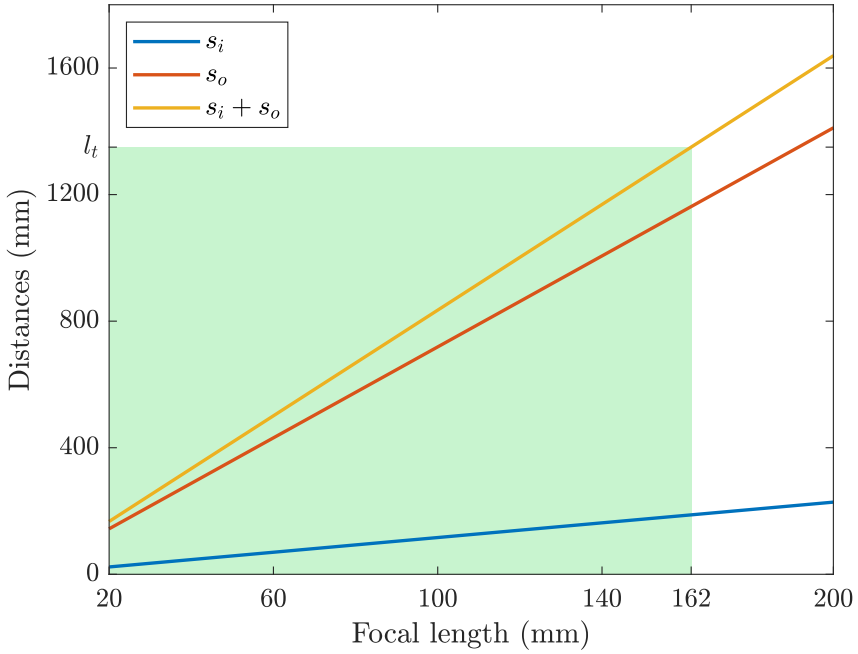


Figure 3.6: Evolution of the object distance s_o, the image distance s_i and the total distance $s_i + s_o$ as a function of the focal length of the collecting lens from 20 mm to 200 mm. The range of permissible focal lengths is highlighted in green.

The third constraint on the collecting lens is that the latter must not obstruct the light reflected by the DMD, and must give enough place for the condensing lens, as shown in the schematic of the setup in Fig. 3.1. This constraint is directly related to the angle of the light cone from the lens, which is expressed by its f-number . Indeed, the light cone reflected by the DMD and directed towards the condensing lens has the same aperture as the incident cone. If this aperture is too large (small f-number), the two cones will overlap, and the collecting lens will obstruct the passage of light in the secondary arm, hindering the view of the condensing lens. This non-obstruction constraint will therefore define the minimum f-number of the lens.

Geometrically, the critical light ray that will be obstructed first by the collecting lens is the ray starting from the upper end of the lens and reflecting at the lower end of the DMD. It is also necessary to consider the additional diameter taken by the lens mount in the non-obstruction condition, as illustrated in Fig. 3.7, where the path of the critical ray is depicted along with the geometry of the problem. Instead of performing expensive computations with ray-tracing, a simple 2D modelling of the problem can be set. This

model introduces the characteristic diameter of the DMD defined as

$$D_{DMD} = \sqrt{w_{DMD}^2 + h_{DMD}^2}, \tag{3.7}$$

where w_{DMD} and h_{DMD} are respectively the width and height of the DMD. This characteristic diameter is used to take into account the rectangular shape of the DMD in the 2D model, and corresponds to the diagonal of the SLM. The critical ray will not hit the lens mount if

$$\theta_r > \theta^*, \tag{3.8}$$

where θ_r is the angle of the reflected critical ray with respect to the optical axis, and θ^* is the angle made by the lens mount when viewed from the point of reflection of the critical ray on the DMD. The critical angle θ^* is determined by the geometry of the problem

$$\theta^* = \arctan\left(\frac{\phi_m + D_{DMD}}{2s_i}\right), \tag{3.9}$$

with ϕ_m, the total diameter of the lens with its mount. The incident ray with an angle θ_i is reflected on the micromirror surface, whose normal is inclined by the DMD mirror tilt angle $\psi_{DMD} = 12°$ with respect to the optical axis. By the law of reflection, the incident and reflected angles are linked by

$$\theta_r = 2\psi_{DMD} - \theta_i, \tag{3.10}$$

all the angles being considered positive when defined clock-wise. The incident ray angle θ_i can also be expressed as a function of the lens diameter, in the same fashion as θ^* in Eq. 3.9. The non-obstruction condition in Eq. 3.8 can now be rewritten as a function of the parameters of the lens and the DMD, therefore giving a condition on the diameter of the lens as well as its focal length by the presence of the image distance $s_i = f(1 + M_T)$. The non-obstruction condition is fulfilled if

$$\arctan\left(\frac{\phi + D_{DMD}}{2f(1 + M_T)}\right) + \arctan\left(\frac{p_m\phi + D_{DMD}}{2f(1 + M_T)}\right) < 2\psi_{DMD}. \tag{3.11}$$

The parameter p_m is the ratio between ϕ_m and ϕ. After measuring the different lenses and the corresponding mounts available in the laboratory, a value of $p_m = 1.5$ is found to cover the majority of the lenses.

To summarize, the constraint defined by Eq. 3.6 limits the focal length of the lens to 162 mm, and the non-obstruction constraint in Eq. 3.11 allows for defining a set of diameter and focal length that avoids obstructing the light reflected by the DMD. As mentioned above and in the previous section, the largest focal length possible is desired to provide the best image quality and reduce the aberrations. Therefore, the chosen collecting lens should have a focal length as close as, but below $f_{max} = 162$ mm. By fixing $f = f_{max}$ in the non-obstruction condition expressed by Eq. 3.11, the maximum lens diameter allowed is $\phi_{max} = 50$ mm. A lens with diameter and focal length equal to ϕ_{max} and f_{max} has an f-number equal to 3.2. This f-number is considered to be not too low, thus preventing significant optical degradation, but it is also sufficiently small to provide enough photons to the imager and achieve a bright image. The lens available in the optical store with

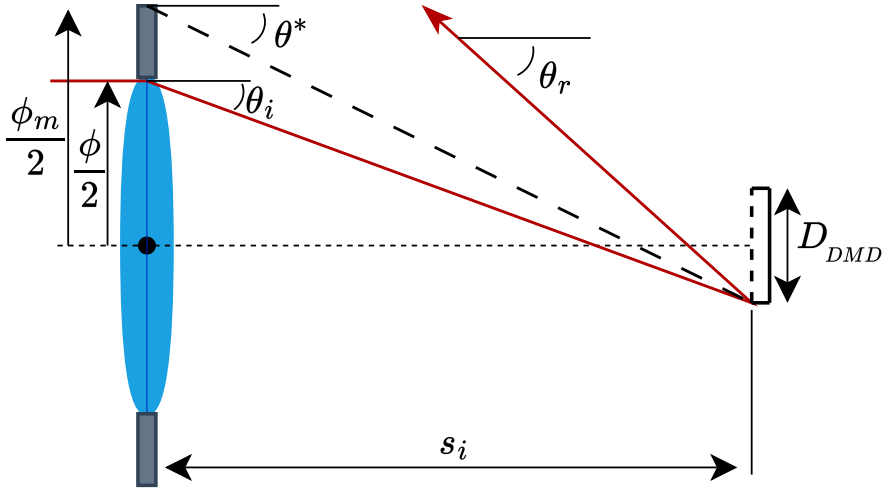


Figure 3.7: Geometry of the non-obstruction constraint, with the critical ray represented in red, and the lens mount drawn as gray boxes. The drawing is not to scale.

characteristics closest to ϕ_{max} and f_{max} is a lens with a diameter of 40 mm and a focal length of 160 mm. Its main characteristics are summarized in Tab. 3.2. This lens has the advantage to be a doublet lens. Composed of two pieces of glass with different optical properties glued together, this type of lens corrects chromatic aberrations, which cause the different wavelengths of the image to be focused differently and degrade the overall quality. The f-number of the lens is higher than the goal lens due to a smaller diameter, but the diameter is considered sufficiently large to give bright images for the purpose of a laboratory setup.

Table 3.2: Characteristics of the chosen collecting lens.

Parameter	Value
Type of lens	doublet
Diameter (mm)	40
Outer diameter (with mount) (mm)	60
Focal length (mm)	160
f-number (-)	4

Now that the characteristics of the lens are known, its position can be calculated using Eq. 3.3. The object distance s_o is 1150 mm and the image distance s_i is 186 mm. The total distance between the scene and the DMD is thus 1336 mm, which is indeed less than the total length of the laboratory table. The non-obstruction condition is also met, as shown by the ray-tracing simulation of the collecting lens-DMD assembly in Fig. 3.8, calculated with a personal code in Matlab. Thanks to Eq. 3.1, the field of view of the primary arm can be computed, and is 0.0064 sr or 20.94 deg^2.

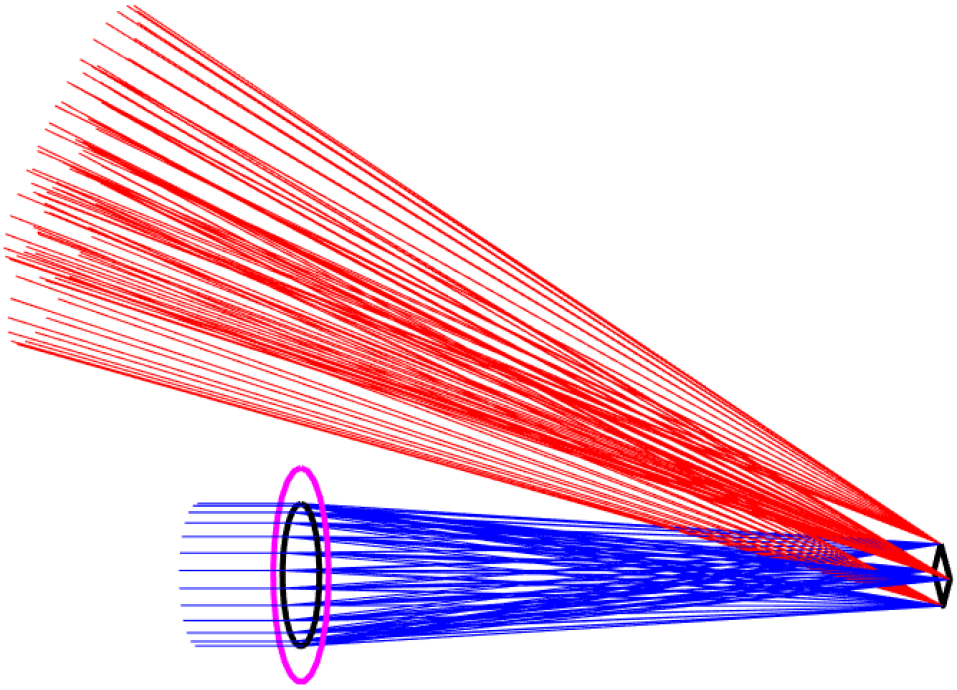
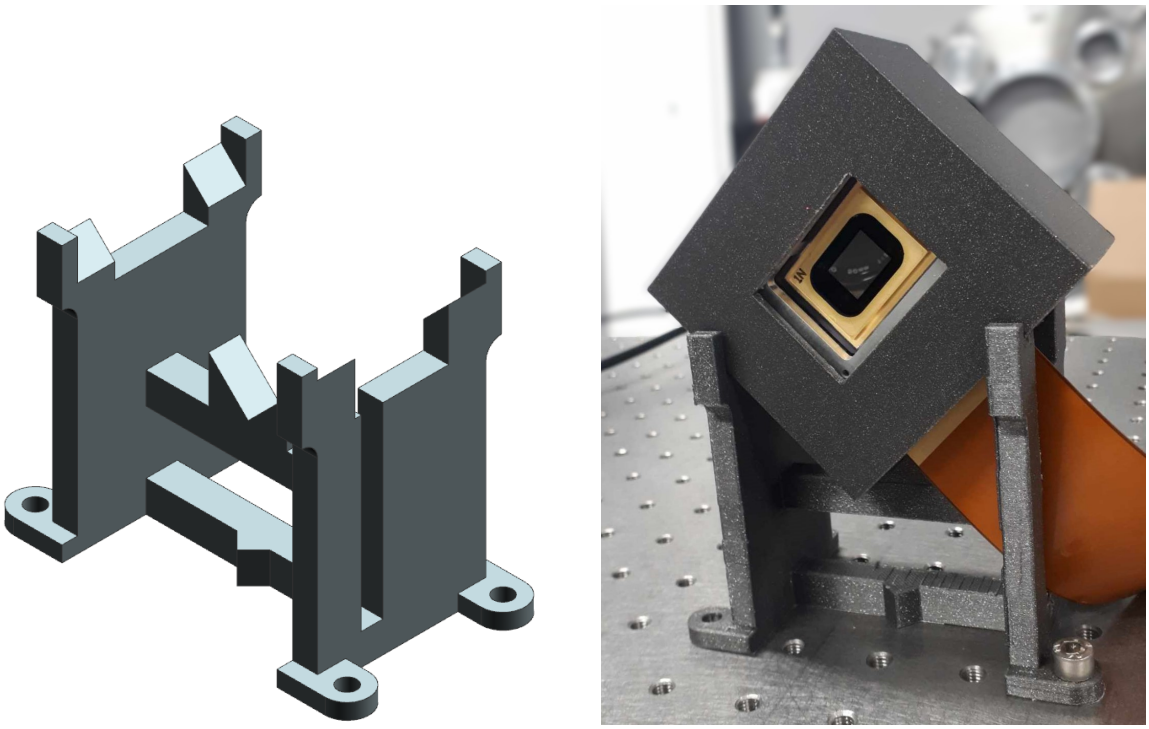


Figure 3.8: Ray-tracing simulation of the chosen collecting lens (left) and the DMD (right) done with a personal code in Matlab. The lens mount is represented in pink, the rays incident to the DMD are drawn in blue, and the reflected rays are in red. Only the rays focused on the four corners of the DMD are shown.

3.3.3 Spatial light modulator

As previously mentioned, the DMD used to modulate the images is a DLP7000. It is placed at the focal plane of the collecting lens and is tilted at 45° along the optical axis. Due to the rotation of the micromirrors around their diagonal, this 45° tilt keeps the optical path in a plane parallel to the table, thus simplifying the setup construction and facilitating its alignment. To properly position the DMD, a protective box was originally 3D printed and placed horizontally on the table. To rotate this box at the right angle, a new support had to be created. This support was designed to hold the protective box that was originally made but with a tilt of 45° with respect to the horizontal while facing the scene. This new support was designed with the Computer-Aided Design (CAD) software NX and was further 3D printed. The 3D CAD view of this support is shown in Fig. 3.9a, and the complete device mounted on the optical bench is visible in Fig. 3.9b.



(a) **(b)**

Figure 3.9: Assembly of the DMD. (a) CAD view of the support holding the DMD at the correction orientation. (b) Picture of the DMD and its 3D-printed support fixed on the laboratory bench.

3.3.4 Condensing lens

At this stage, the collecting lens has focused an image of the scene on the DMD. Placed after the DMD, the condensing lens collects the light rays reflected by it and focuses them onto the FPA of the detector located at the focal plane of the lens. From the point of view of the condensing lens, the scene to image now corresponds to the DMD screen and the image must be located within the camera detector.

The condensing lens must be capable of collecting as much light as possible reflected

by the DMD in order to achieve the best possible radiometric resolution at the detector. As shown in Fig. 3.8, this reflected light (in red) diverges, forming a cone. Therefore, to collect the same amount of light, a condensing lens placed further away from the DMD must have a larger diameter than a lens placed closer to the DMD. The focal length of this lens defines the distance from the DMD. The larger the focal length, the further the lens is placed from the object, and thus, the larger the diameter must be to capture enough light. The motivation to keep the imager compact and given the limited diameters of the lenses available in the optical store implies choosing a lens with a relatively short focal length. Consequently, decreasing the focal length increases the field of view and also increases optical aberrations. Therefore, it was decided to choose a camera objective as the condensing lens rather than a simple lens or a doublet. The presence of multiple complex lenses inside the objective allows for better aberration correction and much better performance than a simple lens. Moreover, the small size of the object (the DMD screen) and the small size of the FPA where the image will be focused require high-quality optics to ensure that the obtained image is sharp. All the camera objectives available in the laboratory are listed in Tab. 3.3.

Table 3.3: List of the available camera objectives in the laboratory.

Camera objective	Focal length (mm)	Diameter (mm)	f-number (-)
Pentax 6 mm	6	5	1.2
HR 16 mm	16	11.4	1.4
HR 25 mm	25	17.9	1.4
Rodenstock Rodagon 28 mm	28	7	4
FA 2/3”	35	21.9	1.6
Ernitec TV lens	50	38.5	1.3
Zuiko auto-macro	50	14.3	3.5
Tamron	75	19.2	3.9
Rodenstock Rodagon 210 mm	210	37.5	5.6

By handling and testing different camera lenses, it was noticed that lenses with an f-number lower than approximately 2 produce too many aberrations in the corners of the image. In terms of focal length, lenses with focal below 35 mm are considered too close to the DMD, and the distance between the detector and the lens is too short to achieve proper focus on the FPA. Conversely, above 75 mm focal length, the condensing lens is placed too far from the DMD, resulting in a dim image. This reduces the possibilities to the Zuiko 50 mm and the Tamron 75 mm.

It can be noticed that the f-number of the Zuiko objective is slightly lower than the Tamron lens, and most importantly, the Zuiko is a macro camera objective. It means that it has been designed to be used at close distances, for photographing small objects [47]. Considering the dimensions of the sensor (see Sec. 3.3.5) and the DMD, the transverse magnification of this lens is $M_T = 0.39$. The magnification being relatively large, the use of a macro lens is hence more appropriate to obtain good quality pictures of the image formed on the DMD. Therefore, using the Zuiko lens allows a more compact imager, with slightly brighter images. Based on all this, the Zuiko lens is chosen to be the condensing

lens for the secondary arm of the imager. A picture of this camera objective is shown in Fig. 3.10.

Figure 3.10: Picture of the Zuiko auto-macro 50 mm $f/3.5$ objective used as the condensing lens of the imager. From [48].

One particularity to notice is that, unlike the collecting lens, the object to be imaged here is not perpendicular to the optical axis. Indeed, due to the off-axis reflection of the DMD, its screen is seen by the condensing lens at an angle of $2\psi_{DMD} = 24°$. This angle is calculated thanks to Eq. 3.10 with $\theta_i = 0°$. As a result, the image plane is also no longer perpendicular to the optical axis. This phenomenon, illustrated in Fig. 3.11, is called the Scheimpflug principle. It states that the object, lens, and image planes intersect in a common point. This principle will be further explained and detailed in Sec. 3.3.5, as it has a great impact of how the image forms on the detector.

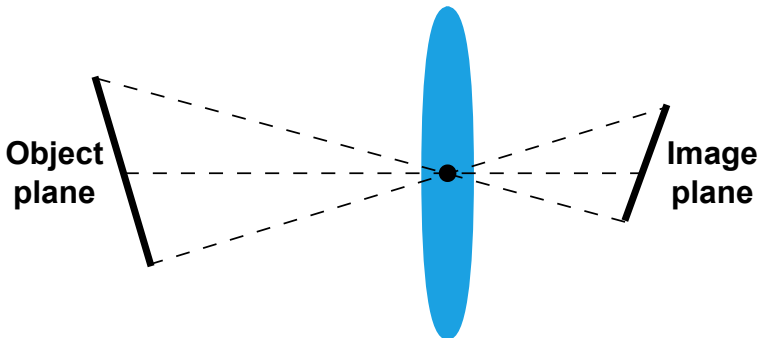


Figure 3.11: Scheimpflug principle. Because the lens and the object plane are not parallel, the image plane is not parallel to the lens.

The chosen lens is composed of several complex lenses, and it has a non-negligible thickness compared to its diameter. Hence, the thin lens approximation does not stand anymore. Nevertheless, the thin lens equation can still be used in the context of multiple lens systems or thick lenses by introducing the concept of principal planes. The principal planes of a system of lenses are two hypothetical planes at which all the refraction can be considered to happen [49]. As shown in Fig. 3.12, they are composed of the principal

object plane H_1 and the principal image plane H_2. These planes are the location from which the effective focal plane of the optical system is measured. They depend on the characteristics of the lens system but do not change with the position of the object to image. Their positions, which can be located outside the instrument, are measured in practice with a collimated laser, by measuring the distance between the focus point and a fixed reference on the lens. The position of the principal planes with respect to this reference is then found thanks to Fig. 3.12. For the Zuiko 50 mm $f/3.5$, the reference has been set to the end plate of the lens mount, and the principal object and image planes are located respectively 0 and 7 mm toward the front of the objective. The thin lens equation can be used with multiple lens systems, but it leaves out the distance between the two principal planes. It is thus necessary to add this relative distance to obtain the correct position of the optics, and the position of the image.

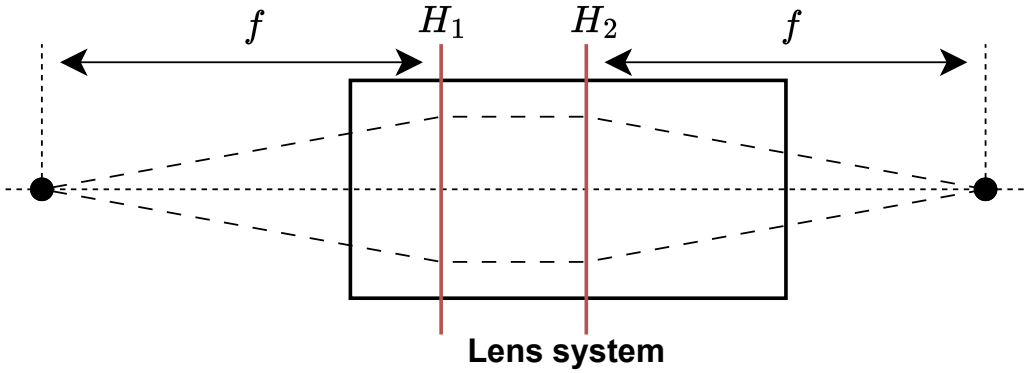


Figure 3.12: Illustration of the principal planes H_1 and H_2 of a multiple lens system.

The image and object distances s_i and s_o can also be calculated to determine the position of the condensing lens with respect to the DMD, and with the detector. To do this, the angle made by the DMD and the object plane due to the Scheimpflug principle must be taken into account. Thus, for a given object distance, the images of the ends of the DMD's medians are calculated using the simple lens equation. Then, the lengths of these median images in the image plane are calculated. The object distance is incremented until the lengths of these medians match the dimensions of the detector. When this condition is met, the image of the DMD is contained across the entire FPA, and we obtain the distances s_o and s_i for the condensing lens. For a thin lens, these distances are measured from the center of the lens, but for a lens system like in the present case, they are measured from their respective principal planes H_1 and H_2. For the chosen camera objective, the object and image distances are respectively equal to 148 mm and 76 mm. It should be noted that, especially for the condensing lens, the optical elements are manually positioned according to the distances calculated above, but slight adjustments are made to take into account more complex behaviors in terms of object and image distance that could not be predicted with the simple thin lens equation, with the goal to achieve the best focus and image quality.

3.3.5 Detector

The camera used to acquire the images is the UI 3240CP-M-GL-TL from IDS [50], and is shown in Fig. 3.13. It is a monochromatic CMOS camera working in the visible spectrum, and whose main characteristics are summarized in Tab. 3.4. Because it has a bit-depth of 8, the pixel intensities of the output image can take values between 0 and 255.

Figure 3.13: Picture of the camera [50].

Table 3.4: Characteristics of the camera [50].

Characteristic	Value
Type	CMOS Mono
Resolution (pixel)	1280×1024
Aspect ratio	5:4
Pixel size (µm)	5.3
Sensor dimensions (mm)	6.784×5.427
Bit-depth (bits)	8

In a conventional imager such as a digital camera, the detector is fixed to its objective lens, and the sensor's surface is perpendicular to the optical axis. However, as previously mentioned, the image plane is not perpendicular to the optical axis due to the Scheimpflug principle. By maintaining a classic positioning of the camera relative to the condensing lens, the image plane where the rays are focused will not be exactly aligned with the detector. This misalignment can result in a lack of sharpness and blurriness in most of the image. The Depth Of Field (DOF) of the condensing lens allows for measuring the impact of the Scheimpflug effect on sharpness and determining whether it is necessary to correct it. The depth of field is the range of distances within which objects appear acceptably sharp in an image captured by a camera [51], and for close object photography, it can be expressed as [52]

$$\mathrm{DOF} = 2CF\#\frac{(1+M_T)}{M_T^2}, \tag{3.12}$$

where M_T is the transverse magnification of the condensing lens, $F\#$ is its f-number, and C is the circle of confusion. The circle of confusion is the diameter of the circular light disk that can form on the detector and still be perceived as a point in the final image. This value of the circle of confusion is typically equal to the size of 1-3 sensor pixels [53]. Because the DMD is seen by the condensing lens with an angle, a side of the SLM is closer to the lens than the other side. If the difference in depth between the two sides is larger than the DOF of the lens, the effects of the Scheimpflug principle will begin to be visible and degrade the image quality. For the condensing lens, the DOF is 1.03 mm, which is more than 5 times smaller than the depth of the DMD viewed from the lens. Therefore, most of the image will appear blurry, making it crucial to correct this to achieve good image quality.

The blur caused by the Scheimpflug principle can be corrected in two ways. First, it can be compensated by tilting the condensing lens so that the lens plane is once again parallel to the object plane, and thus the image plane is also perpendicular to the optical axis. This first solution implies that the incident rays enter the lens at a much greater

angle of incidence, which induces many non-desired optical aberrations. The second solution is to keep the lens plane perpendicular to the optical axis and tilt the DMD so that it is perfectly aligned with the image plane. This second solution is chosen for the imager because it provides the best image quality while eliminating the Scheimpflug blur. Therefore, the condensing lens and the camera detector have to be decoupled to allow this correction.

The rotation angle ψ_s of the camera relative to the lens plane can be calculated geometrically using the Scheimpflug principle. As previously mentioned, this principle states that the object plane, lens plane, and image plane intersect at a single point when the object plane is not parallel to the lens plane. Therefore, two right triangles sharing a common side can be constructed, as illustrated in Fig. 3.14. From this geometric construction, the following relationship is derived:

$$\psi_s = \arctan\left(\frac{s_i}{s_o}\tan(24°)\right), \tag{3.13}$$

where s_i and s_o are the image and object distances of the condensing lens computed in the previous section. The computed horizontal rotation angle of the camera is $\psi_s = 12.6°$. Complementary to this rotation, the camera must also be rotated 45° relative to the horizontal so that the detector can capture the image of the DMD with the correct orientation, because the latter is tilted 45° relative to the table, as explained in Sec. 3.3.3.

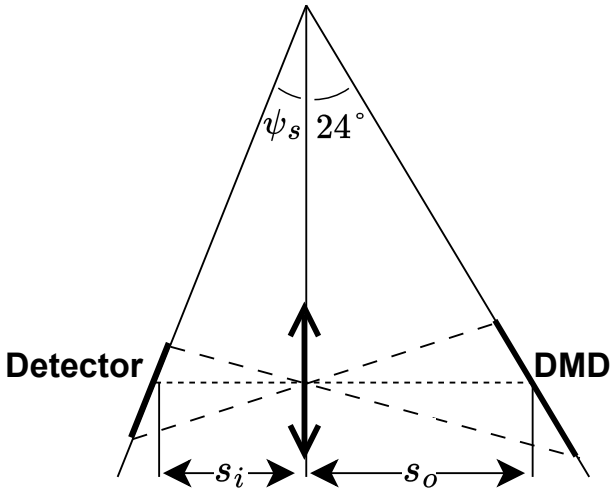
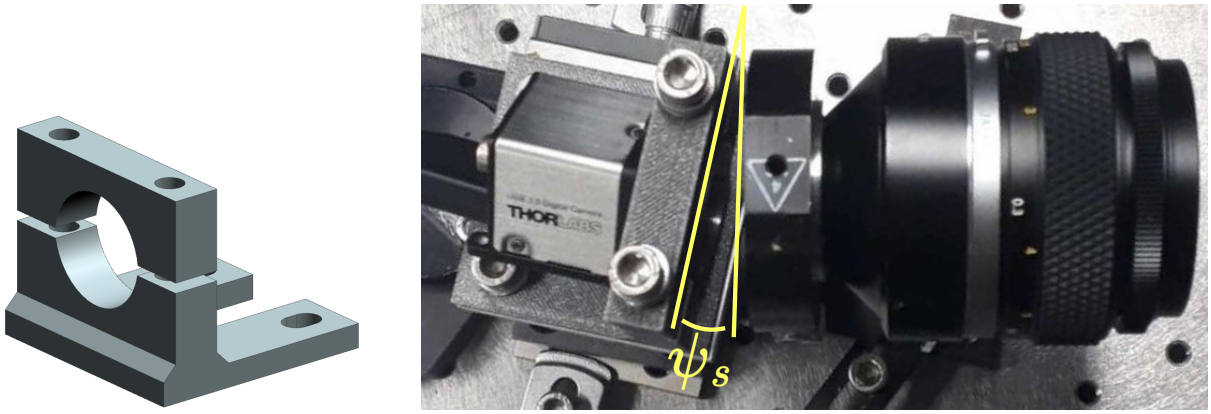


Figure 3.14: Schematic of the secondary arm with the Scheimpflug principle. The DMD is inclined of 24° with respect to the condensing lens. The lens is represented as a thin lens for the sake of clarity.

In order to correct the Scheimpflug blur, it is necessary to precisely position the detector at the correct distance from the camera lens and with the right orientation. To achieve this, a camera mount was designed in NX and 3D printed. This mount, whose CAD model is shown in Fig. 3.15a, consists of two parts, the base and the cover, each with a U-shaped portion, and joined together by two bolts. The cylindrical mount of the camera fits between the base and the cover, and tightening these two pieces with the bolts holds

the camera firmly in place like a hose clamp. This setup allows the camera to be freely tilted relative to the horizontal axis. This camera mount is, in turn, bolted onto a 1-axis precision vertical stage, allowing fine adjustment of the camera's height. The entire camera assembly is positioned at the correct tilt angle ψ_s relative to the condensing lens plane using a stop fixed on the optical table. A top view of the final camera assembly is shown in Fig. 3.15b.



(a) **(b)**

Figure 3.15: Camera detector assembly. (a) CAD view of the camera holder. Four bolts are inserted into the premade holes to fix the structure and the camera. (b) Top view of the complete assembly of the camera and the condensing lens. The tilt angle ψ_s of the camera with respect to the condensing lens' plane is highlighted.

3.4 Final setup

The imager is now assembled, and is capable to image a scene. Fig. 3.16 shows the full imager with the different components highlighted. As a summary of the results obtained in the previous sections, a top view scheme of the imager with the main measurements relative to the position, orientation and dimensions of the optical elements is shown in Fig. 3.17.

Fig. 3.18 shows the original image used as a scene and the raw image of the scene. All the mirrors of the DMD have been set to the ON-state to obtain an undamaged image of the scene. The original image in Fig. 3.18a is an aerial view of the CSL and the Liège Science Park. This image has been printed on a sheet of paper and placed in front of the imager. It can be seen in Fig. 3.18b that the obtained image is in gray levels because the sensor is monochromatic. The image also appears relatively sharp with a good image quality, demonstrating the effective assembly and selection of the imager's optical elements. It can be noted that print marks of the image on the paper are slightly visible as closely spaced lines, but their presence is not considered disruptive for the continuation of the work. Furthermore, the image of the scene does not take all the available space of the detector, which leaves dark margins all around the image. This is due to the placement of the camera sensor with respect to the condensing lens that is not perfect. However, even if the camera placement was perfect, the image would not be fully framed on the detector, because as can be seen, the image is not perfectly rectangular but rather trapezoidal. This

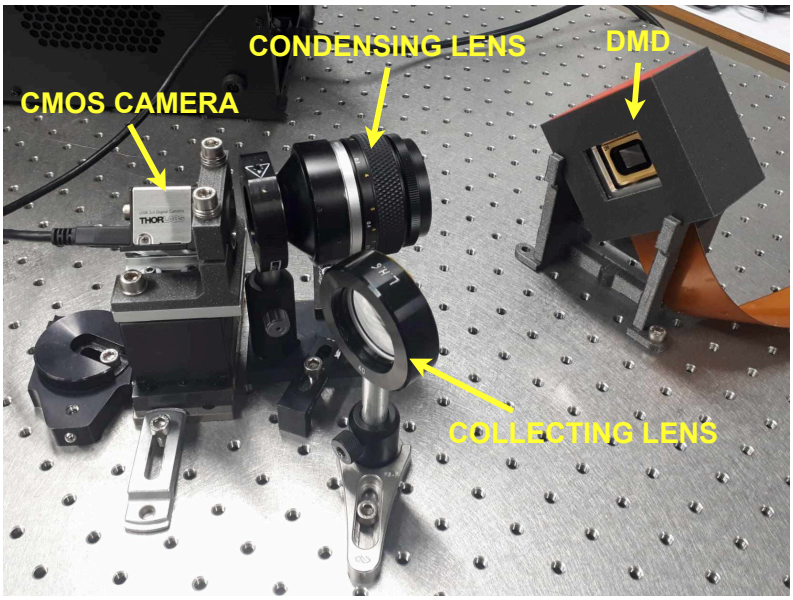


Figure 3.16: Picture of the assembled imager in the laboratory. The scene is located outside the picture toward the left of the optical setup. The different optical components are highlighted.

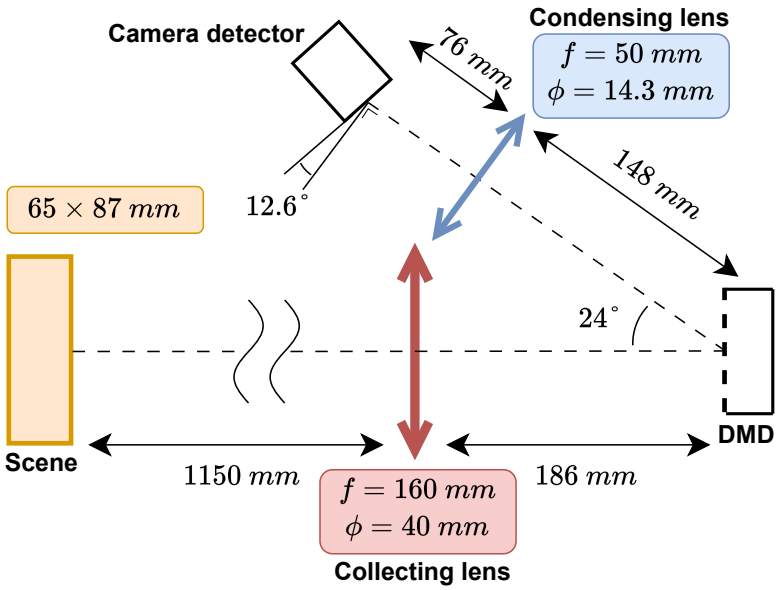
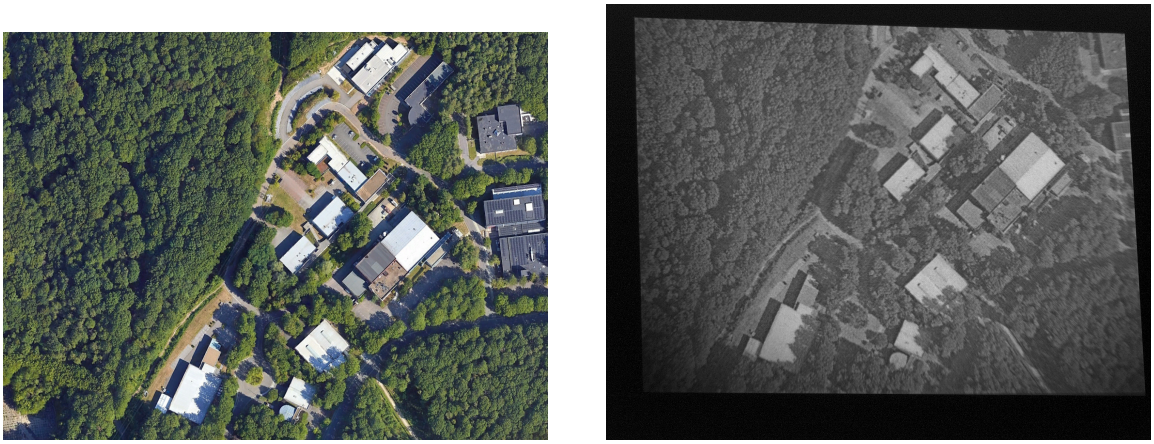


Figure 3.17: Top view scheme of the imager with its main dimensions.

distortion is due to the tilt of the DMD relative to the condensing lens and the proximity of these two elements. This creates a perspective effect that enlarges the parts of the DMD closer to the lens and shrinks the parts that are further away. Additionally, the raw image does not have a uniform illumination, with the presence of a darker area near the bottom left corner of the image. This effect, known as vignetting, is certainly due to a slight misalignment between the DMD, the condensing lens, and the camera sensor. The calibration and post-processing of the raw image is necessary to correct all these defects, and is presented in Chap. 4. Eventually, it is important to note that the framing of the image taken by the instrument is deliberately different from that of the original image. This is because, for a scene where orientation matters little, such as the aerial view of the CSL, it is easier to simply print the image in dimensions larger than the size of the scene and place this paper at the correct distance from the imager. However, for scenes requiring precise orientation, the device shown in Fig. 3.3 must be used.

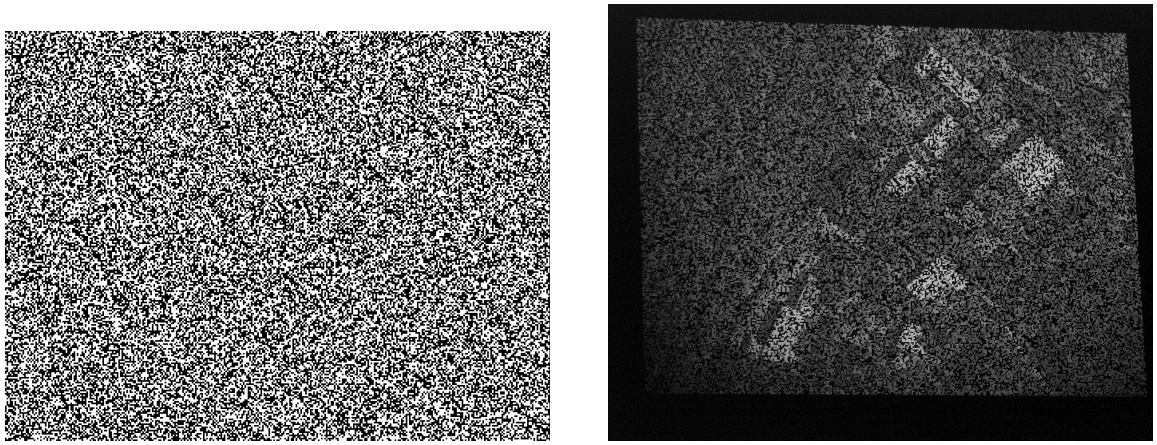


(a) **(b)**

Figure 3.18: Raw images of a scene taken by the imager. All the mirrors of the DMD are set to the ON-state. (a) Satellite view of the CSL used as the original image for the scene. (b) Raw image of the scene taken by the imager. Note that the framing of the two images is deliberately different.

The previously obtained image was captured with all the mirrors of the DMD activated. However, the imager needs to perform inpainting. By applying a binary pattern on the DMD, the image of the scene can be damaged in a controlled manner. Fig. 3.19a shows an example of binary pattern given by the DMD to damage the image. It is randomly generated with a filling ratio of 50%. The filling ratio is defined as the ratio of the number of active (white) pixels, and the total number of pixels on the DMD. This parameter allows controlling the amount of data of the original image that is kept or destroyed during the damage process. The random pattern has a binning of 3×3, which means that it is made of pixels of 3×3 micromirrors. The raw image damaged by the DMD is shown in Fig. 3.19b. It can be seen that the mask is accurately printed on the scene.

The instrument is located on an optical bench in a laboratory. Surrounding it are light sources such as a desk lamp used to illuminate the scene, the light for the computer screen used to control the imager, and, when turned on, the room lighting. All these non desired



(a) **(b)**

Figure 3.19: Raw images of a scene damaged by applying a binary pattern on the DMD. (a) Binary pattern applied on the DMD. (b) Raw image of the damaged scene taken by the instrument.

lights can contaminate the final image and greatly reduce its contrast. To prevent this, a housing has been installed around the imager, as shown in Fig. 3.20. This housing is made using two metal optical covers and black cardboard to fill all the openings. Only one opening at the collecting lens allows light from the scene to enter the imager. A trapdoor located on the top of the housing allows accessing and manipulating the optical elements of the imager. To measure the effects of adding the housing on the final image, two images of the same scene were taken: one before the housing was placed around the imager, and one after its installation. To compare the results, these two photos were taken under identical external illumination conditions, with the room lighting left on to simulate the worst external illumination conditions. The two pictures without and with the housing are shown in Fig. 3.21. The addition of the housing to the imager has an undeniable effect on the final result, significantly improving the contrast of the image. The principal cause of the non-desired illumination detector is by the gap between the camera and the condensing lens. This gap leaves indeed the CMOS sensor at open air, unprotected from the external light. A particular care has thus been taken to fully cover this gap during the housing process. As it has been said, the images of Fig. 3.21 are taken with the room lights turned on to simulate the worst case. However, the room lights are in practice turned off when imaging, to not pollute the results with unwanted light.

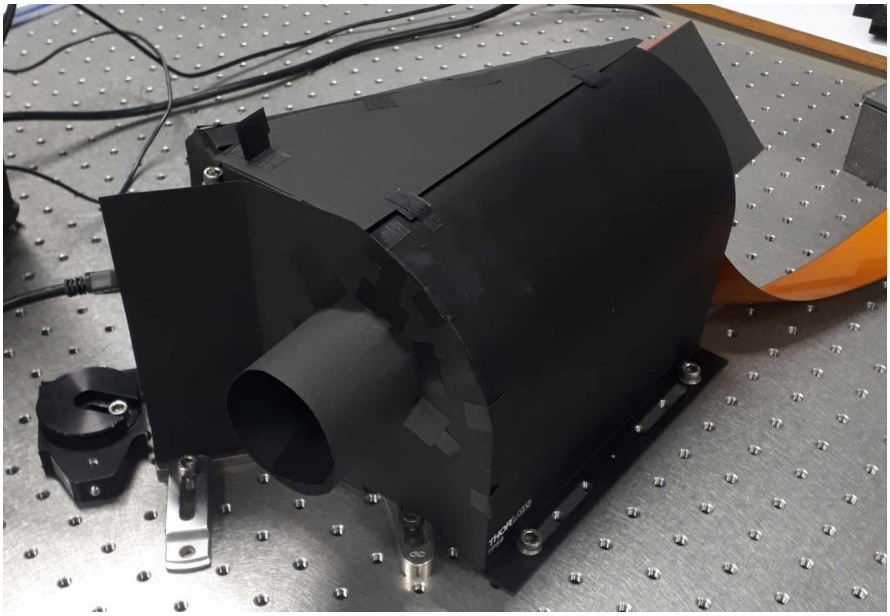
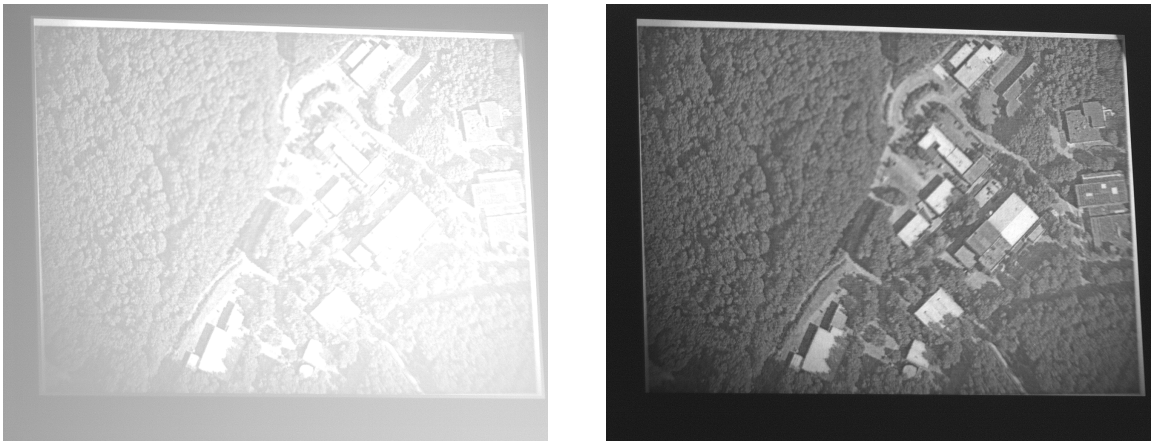


Figure 3.20: Picture of the complete imager with the housing built around it.



(a) without housing **(b)** with housing

Figure 3.21: Raw images of a scene taken before and after installing a housing around the optical setup. All the mirrors of the DMD are set to the ON-state. The lights of the laboratory room are turned on for both pictures.

3.5 Control of the imager with LabVIEW

This section focuses on the computer control of the imager during image acquisition. Automatic control of the instrument is an important feature to avoid having to perform all acquisitions manually, which could be very time-consuming and potentially cause errors. The goal is that the user only needs to provide a list of patterns to be displayed on the DMD, and the program will handle initializing the DMD, displaying the pattern, and finally saving on the computer the obtained image. To automate this acquisition, the LabVIEW software is used. LabVIEW is a system development platform commonly used for data acquisition, instrument control, and industrial automation. This software uses a visual programming language and is developed by National Instruments. The operation of the visual code that has been designed and programmed in LabVIEW to control the instrument is represented by the flowchart in Fig. 3.22. First, the list of binary patterns provided by the users is loaded in the software and transformed to be readable by the DMD. Then, the program will loop through all the patterns, apply the current pattern to the DMD, take a picture with the camera detector, and then save this image in a folder on the computer.

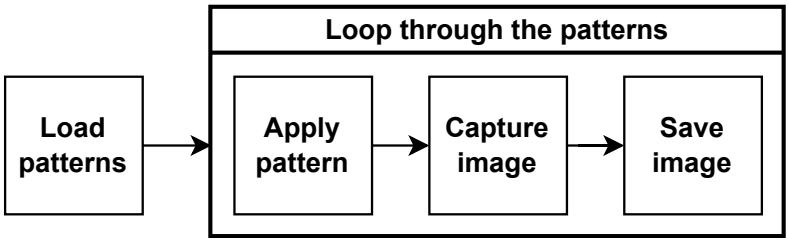


Figure 3.22: Flowchart of the LabVIEW program that controls the imager.

To control the DMD, it is connected to a controller board using a flexible PCB cable. The controller board is the DLP Discovery 4100 Development Platform. This controller, which is capable to precisely apply patterns at speed rates of up to 32 kHz [54], is connected to a computer via a USB 2.0 cable. The development platform is directly controllable in LabVIEW by built-in functions present in the DLP Discovery 4100 Development Platform ActiveX API. To understand how the DMD is controlled, it is interesting to know how the device loads and displays a pattern. First, each micromirror has a memory state and a mirror state. The memory state corresponds to the state of the mirror loaded in the CMOS memory underneath the pixel. On the other hand, the mirror state is the actual binary state of the pixel displayed on the DMD, *i.e.*, the tilt of the micromirror. These two states are linked together but can be different. Indeed, when a pattern is loaded in the DMD, the memory state is changed but not the mirror state. To transfer the memory state to the mirror state, the micromirror must undergo a reset operation. This reset releases the mirror and then re-lands the mirror based on the state of the CMOS memory below [55].

In the LabVIEW program, the patterns are first taken from a folder where they are placed by the user. These binary patterns are under the format of binary bitmap images, and are first converted into binary files that are readable by the DLP Discovery 4100 Development Platform. These binary files are then stored into another folder. During the

acquisition loop illustrated in Fig. 3.22, the current binary file is taken from the binary folder, and then loaded in the DMD via a buffer, and finally proceeds to a reset operation in order to display the pattern. In the current setup, these two operations that happen in the loop take about 0.52 s to complete, giving a pattern rate of 1.92 Hz. This speed is well below the achievable 32 kHz rate of the DMD. This slowness is primarily due to the way we interact with the DLP Discovery 4100 Development Platform. Indeed, the USB 2.0 connection between the computer and the control module greatly limits the data transfer rate of the patterns, limiting the module from operating at its full potential speed [56]. In the context of image acquisition using compressive sensing, for example, in single-pixel imaging, this speed is much too low. As mentioned at the beginning of this chapter, imaging with a single-pixel camera requires several thousand individual measurements with the same number of different patterns. At the current pattern rate, this would take several hours to be able to compressively sense just a single image. The control of the DMD is therefore not perfect, and the improvement of its pattern rate is primordial to acquire images with CS camera equipped with low resolution detectors in a short time. Fast pattern rates can be achieved using Field-Programmable Gate Array (FPGA) circuits, and using more sophisticated electronics to communicate with the controller board [56]. However, this low speed is sufficient to perform inpainting, because it only requires a few image acquisitions with a FPA. The slow pattern change speed is one of the main reasons why this work focuses on inpainting rather than compressive sensing acquisition.

After applying the pattern on the DMD, the camera detector can take a picture. The control of the camera is also performed in LabVIEW with the IMAQdx interface, allowing to control all the settings of the camera by code. These settings and their respective range of value are shown in Tab. 3.5. Brightness allows controlling the overall luminosity of the image by increasing or decreasing the intensity values of all the pixels. Exposure is the time interval during which the detector's pixels accumulate photons from the scene to form an electrical signal. Contrast allows controlling the differences in brightness between the dark and bright areas of the image. Gamma is a parameter that controls the nonlinear relationship between the incident light intensity at the pixel level and the numerical value of that intensity. Finally, sharpness controls the level of sharpening applied to the image during its processing by the camera's internal digital processing software. After adjusting the correct settings on the detector, an image of the scene is captured. This image is then saved in a folder and is now available to the user.

Table 3.5: Controllable parameters of the camera and their range of value.

	Brightness	Exposure (s)	Contrast	Gamma	Sharpness
Min value	0	$1.6 \cdot 10^{-5}$	0	0.01	0
Max value	255	0.125	1	10	2

The LabVIEW software has the advantage of being able to create user interfaces for programs. Fig. 3.23 shows the interface of the imager control program. In the top left, there is the IMAQdx session of the camera, allowing its control. Additionally, it will be seen later in this work that it is necessary to take two types of photos with different

camera settings. Therefore, it is beneficial to be able to easily switch from one camera configuration to another. The "Perform flat-frame" button at the top right of the interface allows switching from one configuration to another. In the middle, the camera settings for the two types of shots can be found. These settings can be modified as desired. Finally, the addresses of the file where the patterns are located and the file where the images taken by the instrument are to be placed must be specified in the two boxes at the bottom of the interface.

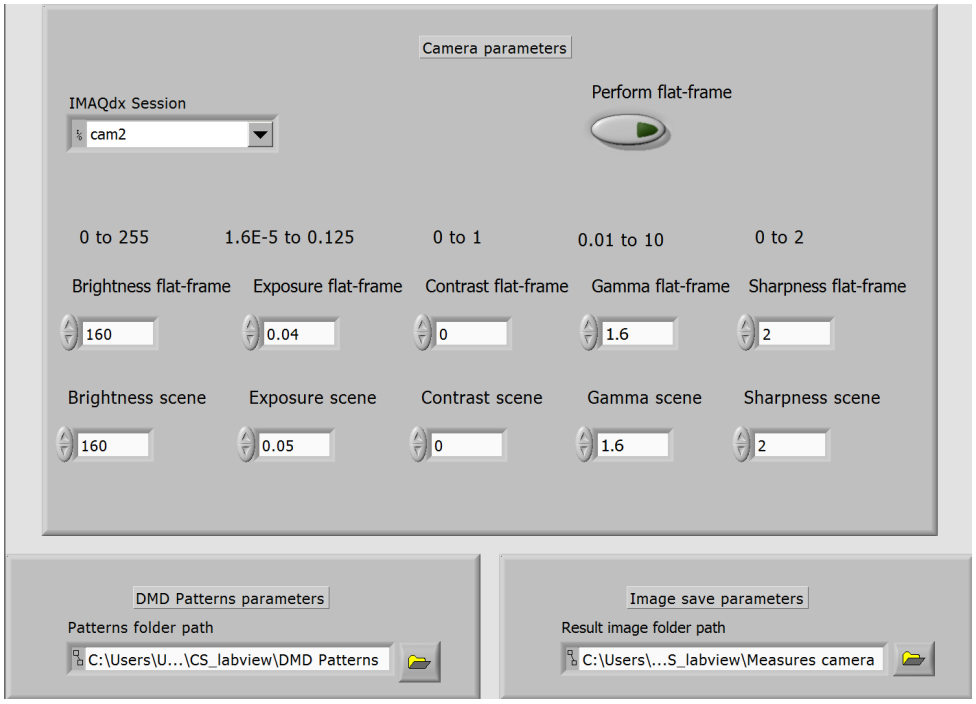


Figure 3.23: Picture of the control panel interface of the imager in LabVIEW.

4 Calibration and post-processing

In the previous chapter, an imager capable of photographing a scene and encoding the image with a DMD was constructed. Although the optical system was built with the intention of obtaining the best possible image, the image obtained on the detector contains defects. These defects are the instrument's mark on the light passing through it. As discussed in Sec. 3.4, the main defects of the raw images coming from the imager are, firstly, that the image projected by the DMD does not cover the entire surface of the detector, leaving dark edges around the image. Secondly, the DMD image is distorted by a perspective effect. And thirdly, the image on the detector shows vignetting, resulting in darker area, independently of the content of the scene.

Apart from these three optical effects, the scene is also degraded in other ways. Due to the imperfect nature of the optics, and the possibly imperfect adjustments made previously, some aberrations are inevitable. A point source present in the scene will not be acquired by the detector as a point after passing through the instrument, but rather as a spread-out spot. This results in blurring of the image and an overall decrease in sharpness. Additionally, the detector that records the image is not made of perfect pixels, and introduces noise into the measurements.

The objective of this chapter is to calibrate the image and perform processing to achieve the best possible quality before the reconstruction by inpainting. First, the calibration procedure will be explained in detail, along with all the images necessary for calibration and later for reconstruction. Then, each element of the calibration will be explained in detail. Finally, the results after calibration, as well as a reconstruction by inpainting of an image, will be presented.

4.1 Calibration procedure

The goal of calibration is to take the raw images from the instrument as input and produce usable images ready for reconstruction. The entire process of image reconstruction is not explored in this work. Reconstructing an image through inpainting requires three different images:

- A damaged image: This is the image of the scene that has been damaged by the DMD. It should be calibrated such that the remaining pixels are as close as possible as the original picture of the scene.
- A pattern mask: This is a binary mask with the same dimensions as the damaged image, indicating which pixel of this image is part of the scene or which is damaged and needs to be reconstructed by inpainting. In this mask, a white pixel corresponds to a pixel of the scene, and a black pixel corresponds to a damaged pixel.
- An undamaged image: This is the image of the scene taken with all the mirrors of the DMD in the ON-state. This image is not strictly necessary for the reconstruction, but it is useful as a ground truth to evaluate the quality of the reconstruction.

To obtain these three images, a calibration procedure was developed and is diagrammed in Fig. 4.1. Two images are present (in orange) at the input of this procedure. The first image is the uncalibrated damaged image coming from the imager. The second image is an image of a completely white scene that has been damaged by the same pattern used for the scene image. It is based on this illuminated pattern image that the binary pattern mask defined above is obtained. These two images first undergo a subtraction of the dark

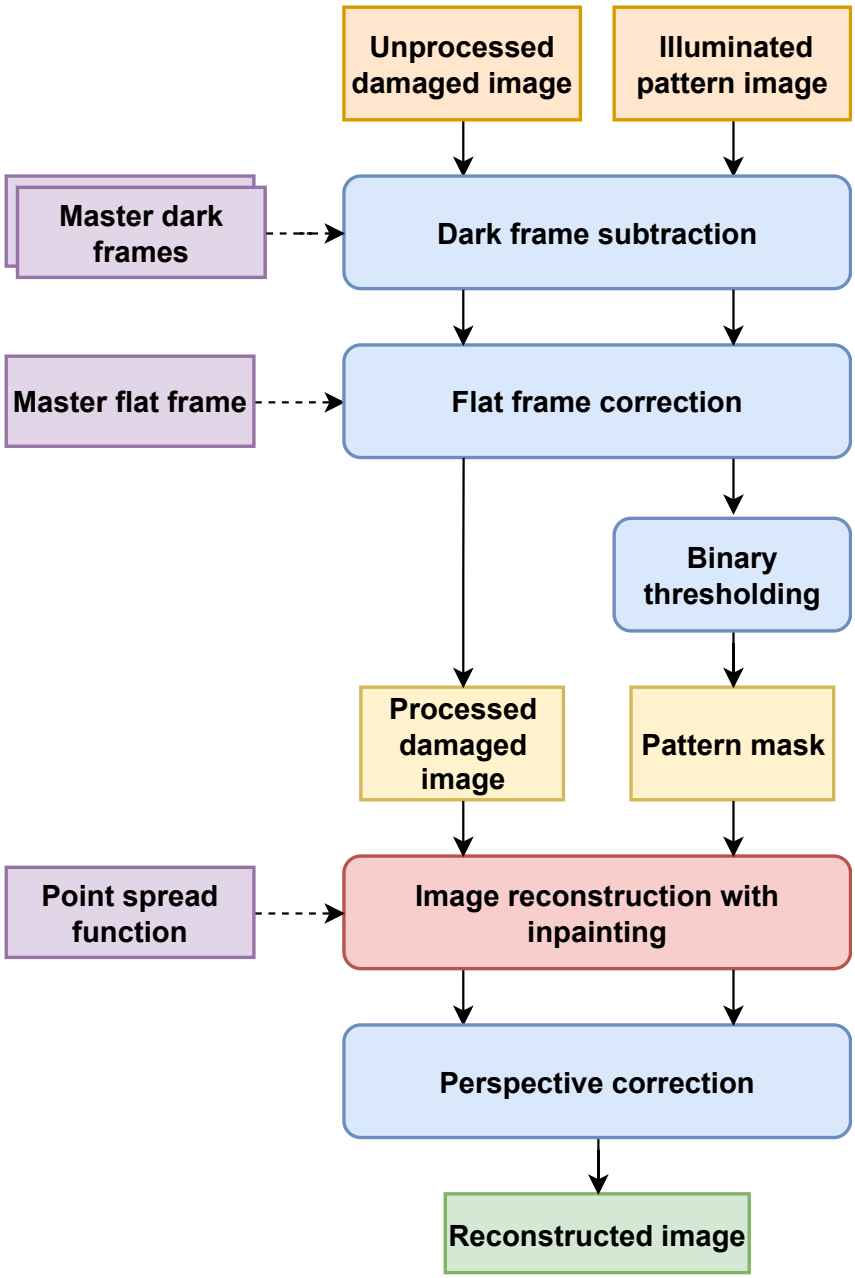


Figure 4.1: General flowchart of the calibration and post-processing of an image.

frame. This initial correction removes the noise caused by the detector and the electronics of the imager. For this, the correction uses Master Dark Frames (MDFs), which will be explained in more detail in Sec. 4.2. Next, both images are corrected for vignetting and

spatial illumination variations through flat frame correction, which is explained in Sec. 4.3. As with the dark frame, this correction uses a Master Flat Frame (MFF). At this point, the damaged image is ready to be reconstructed by an algorithm. The illuminated pattern image, on the other hand, must undergo binary thresholding to become the binary pattern mask usable during reconstruction. The process of pattern mask creation is covered in Sec. 4.4.

Then, the actual reconstruction of the image takes place. As it will be seen in Sec. 4.5, this reconstruction is coupled with image deblurring. The blur in the image caused by the optical instrument itself can be corrected using the Point Spread Function (PSF), which represents the impulse response of the optical system. By injecting a deblurring operator using the PSF into the forward model of the reconstruction algorithm, the image is deblurred and reconstructed simultaneously. However, it is still not rectangular because the perspective effect has not yet been corrected. This final correction is developed in Sec. 4.6. After this step, the reconstruction and calibration is over.

This calibration procedure is made to calibrate and reconstruct a damaged image, but it can also be used to calibrate the undamaged image of the scene. This is done by injecting a mask of a completely white pattern. Thus, no pixel will be reconstructed by the algorithm, and only the deblurring of the image will take place. It is important that the undamaged image undergoes the same calibration process as the damaged image so that the reconstruction phase is the only difference between these two images. Thanks to this, the reconstruction quality can be objectively evaluated. This statement needs to be nuanced because the deblurring correction with the PSF is included in the reconstruction algorithm, and the results of this deblurring could depend on the input image. In any case, it is important that both images undergo calibration as identically as possible.

The camera uses two different settings to image the damaged scene and the illuminated pattern (or flat frame), depending on the luminosity of the imaged scene. These are summarized in Tab. 4.1. These settings were obtained by adjusting each individual parameter to achieve a good output image without saturation or underexposure. It can be seen that only the exposure time varies between the two types of images, and it is shorter for the illuminated pattern. This is because the pattern is photographed with a completely white scene, which is brighter than a structured scene like the aerial view of the CSL in Fig. 3.18a. The exposure must therefore be reduced to avoid saturation.

Table 4.1: Camera settings for the two different types of images.

	Brightness	Exposure (s)	Contrast	Gamma	Sharpness
Scene	160	0.05	0	1.6	2
Illuminated pattern (flat frame)	160	0.04	0	1.6	2

4.2 Dark frame subtraction

The camera detector is an electronic device that converts photons into electrical signals. In any electrical system, unwanted disturbances creep into the signal, degrading it. This is

no exception for the camera detector, and noise from the electronics and photon conversion thus finds its way into the raw images. A correction that can be made to mitigate the noise on the image is called the dark frame subtraction.

A dark frame is an image taken by the imager with the instrument's aperture blocked, allowing no light from the scene to enter. The resulting image therefore contains only the noise generated by the detector and electronics. This noise can then be subtracted from the images taken with the instrument. The dark frames subtraction process only removes the noise components that remain the same from one shot to another. In particular, the dark frame contains noise caused by the dark current and fixed-pattern noise. Dark current noise is the constant response of the detector when it is not illuminated [57], and is due to small currents that flows through the CMOS pixels even when no photons hit them [58]. Fixed-pattern noise is a type of systematic noise where particular pixels of the detector show brighter intensities than the others and forming a fixed-pattern shape. Since systematic noise does not vary from one exposure to another, it can be subtracted from the image using a dark frame. The noise generally varies with the detector's exposure time and temperature. Therefore, it is important to perform dark frame correction using a dark frame taken under the same camera settings and conditions as the image to be processed. Here, two different camera settings are used: one for photographing the scene, and another for capturing the illuminated pattern image and the flat frames (see next section). Consequently, two types of dark frames need to be created. Additionally, the dark frame must be taken under the same external lighting conditions as during the scene captures. This is done to ensure that any parasitic light potentially entering the instrument and reaching the detector is captured in the dark frame.

Image noise also has a variable component that is fluctuating from image to image. This is due to the randomness of noise. To counter this random noise, one solution involves taking multiple photos of the dark frame and averaging all the images pixel by pixel. This stacking of images averages the random fluctuations of the noise and tend to diminish it, while the dark-current and fixed-pattern noise remains the same. This image stacking effectively increases the SNR of the image. As illustrated in Fig. 4.2, this process is used for the dark frame. Typically, 50 dark photos are taken for each of the two camera settings. These 50 frames are then stacked to form a single image, called the Master Dark Frame (MDF). This MDF is used for subtracting the noise of the scene image. Similarly, the image (whether damaged or not) of the scene can also be stacked to reduce random image noise. Since the initial SNR of this image is much higher than that of a dark frame, stacking only a few photos is sufficient to reduce this noise.

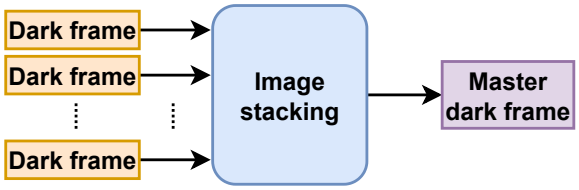
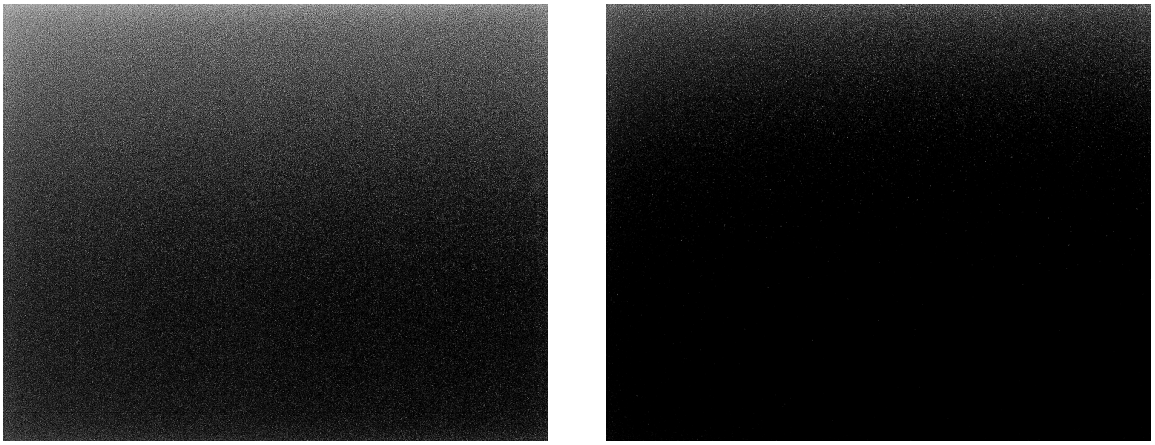


Figure 4.2: Flowchart of the creation of the master dark frame.

The MDFs of the two different camera settings are shown in Fig. 4.3 with their contrast increased by 8 to emphasize the noise. The image noise has a salt-and-pepper appearance that is not uniform across the image. It is observed that for both MDFs, the noise is more prevalent in the upper part of the image. This could be due to a higher level of parasitic external light affecting one side of the detector, or it could be that this part of the detector has electronics behind it that heats up its surface, giving it a higher temperature than the rest, which increases the noise. Additionally, it is noticeable that the noise is more present in the MDF taken with the scene camera settings. This is because the scene's exposure time is longer than that used for imaging the flat frame, and the image noise increases with exposure time.



(a) Scene image settings **(b)** Flat frame settings

Figure 4.3: Master scene and flat dark frames with the contrast of the image increased by 8 to emphasize the noise.

It is interesting to analyze the noise content in the MDFs and compare the noise present in the images for the two types of camera settings. To do this, the intensity distribution of the pixels in the two MDFs is shown in Fig. 4.4. On the x-axis is the pixel intensity, and on the y-axis is the number of pixels in the image having that intensity. It can be seen that both curves follow the same trend, but the curve with the scene settings is above that of the flat frame, indicating the presence of brighter noise. This is indeed clearly visible in the MDFs in Fig. 4.3. It can also be noted that for both types of dark frames, the intensity distribution can be divided into three parts. From left to right on the graph, we find:

- 0 in intensity. It corresponds to completely dark pixels that do not contain any measurable noise. The noise level in these pixels is below their sensitivity. The majority of the image contains this type of pixels.
- Between 1 and 60 in intensity. It corresponds to the dark current noise. The number of pixels decreases exponentially with the intensity. In the master dark frames, it corresponds to the salt-and-pepper regions.
- Above 60 in intensity. These are pixels with a relatively high intensity, but in very few numbers, generally 1 or 2 pixels at some intensities. These pixels, called hot

pixels, are defective pixels that accumulate a lot of signal even when no light is hitting them [59]. Therefore, they have higher dark currents than the other pixels. These hot pixels are in very few numbers and are generally the same from one shot to the other.

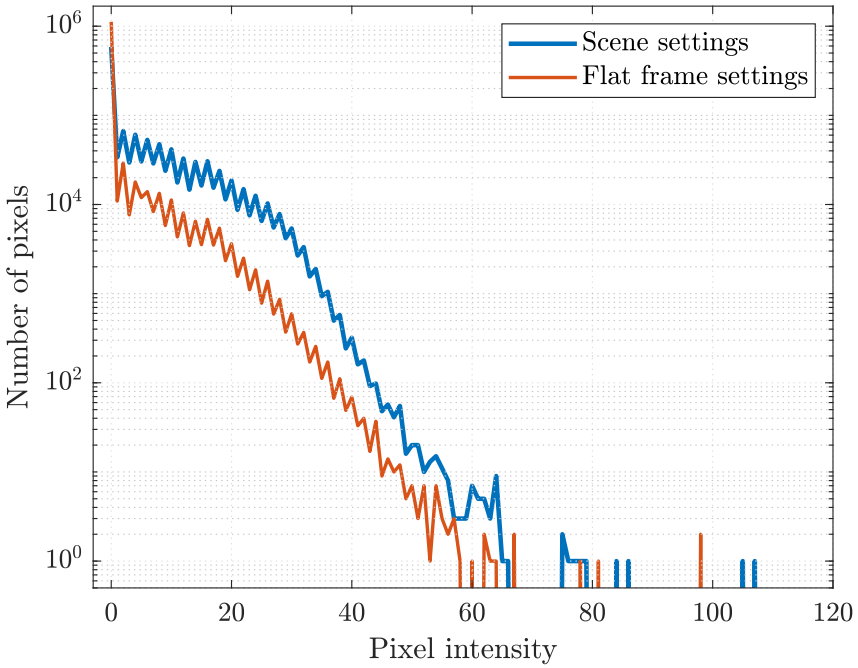


Figure 4.4: Pixel intensity distribution of the two master dark frames.

The MDFs created earlier were made with all the DMD mirrors turned on. However, in practice, the mirrors are either turned towards the detector or away from it depending on the pattern loaded on the DMD. This can slightly vary the parasitic light hitting the detector, and therefore change the dark frames. To measure the impact of the mirrors' state on the dark frame, a MDF was taken with all the DMD mirrors turned on or off, and their pixel intensity distributions are compared in Fig. 4.5. It can be observed that the two curves overlap almost perfectly up to an intensity of about 65 with the curve of the mirrors-OFF being slightly above the mirrors-ON curve. Then, the distribution is similar for hot pixels with an intensity above 65. In order to quantify this difference, the average of the pixel-wise intensity difference between the two images can be computed

$$\Delta D = \frac{\sum_{i=0}^{w} \sum_{j=0}^{h} |D_{\mathrm{ON}}(i,j) - D_{\mathrm{OFF}}(i,j)|}{wh}, \tag{4.1}$$

where D_{ON} and D_{OFF} are the pixel intensities of the MDF with all mirrors ON and OFF respectively. The number of pixels on the width and height of the image is noted w and h. This average intensity difference is equal to $\Delta D = 0.903$. Therefore, on average, the pixels of the two dark frames have less than 1 unit of intensity in difference. By consequence, the state of the mirrors on the DMD are considered to have negligible effects on the MDFs. These computations were made for the MDF with the scene camera settings, but the conclusion is identical for the MDF with the illuminated pattern settings.

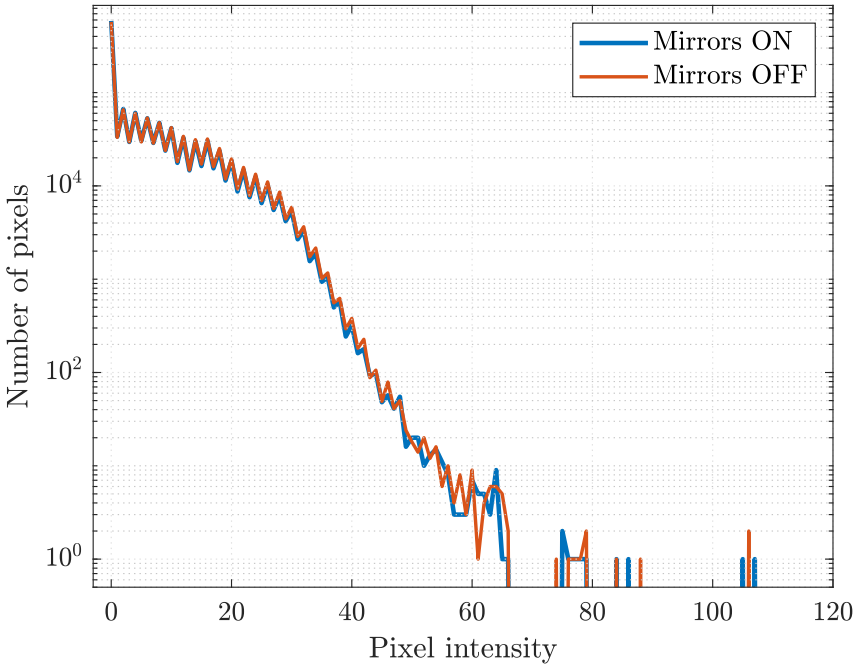


Figure 4.5: Pixel intensity distribution of the master dark frame for the scene camera settings, with the DMD's mirrors fully ON or OFF.

4.3 Flat frame correction

With an ideal imager, photographing a completely white and perfectly homogeneous scene should result in an entirely white image uniformly distributed across the entire detector surface. However, the design of the optics, distortions in the light path, imperfections in the lenses, and many other factors cause the image formed on the detector to have non-uniform exposure. Additionally, none of the detector's pixels are identical, and each one has slight differences in terms of light sensitivity. By consequence, the images obtained by the imager have darker regions which need to be corrected. A solution to this non-uniform exposure is to perform flat frame correction. A flat frame is the image obtained by taking a picture of a uniform white field. This gives a map summarizing all the non-uniformities of the instrument which can then be corrected.

Similarly to master dark frames, a Master Flat Frame (MFF) can be created. As illustrated in Fig. 4.6, this is done by taking several flat frames and proceed for each of them to a dark frame subtraction using the master dark frame corresponding to the camera settings used for capturing the flat frames. These flat frames are then stacked to form the master flat frame. Since the signal from the scene is relatively high compared to the camera noise, it is not necessary to stack many flat frames to obtain a high-quality master flat frame.

In practice, the MFF is created by imaging a white sheet of paper illuminated by a desk lamp and placed at an out-of-focus distance from the camera. This prevents the texture of the paper from appearing on the flat frame and results in a more uniform scene. Additionally, the paper is slightly offset relative to the instrument for each flat frame to

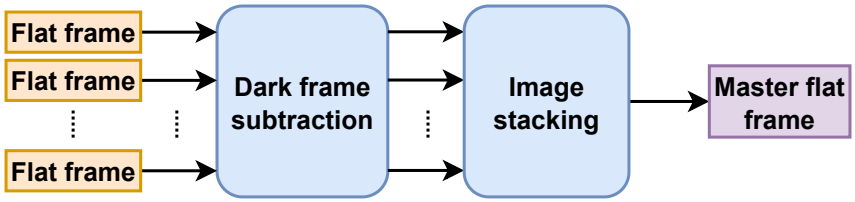


Figure 4.6: Flowchart of the creation of the master flat frame.

further avoid capturing the paper's texture in the MFF. The MFF of the imager is shown in Fig. 4.7. It can be noted that the obtained image is far from being uniformly white. Apart from the upper right corner of the DMD image, all other corners are darkened, particularly the lower left corner. This can be explained by a vignetting effect, which can be caused by the design of the optical setup or by a misalignment of the secondary arm of the imager. In addition to this vignetting, the brightness can naturally be reduced in the periphery of images due to the natural light falloff [60]. This is due to the cos^4 law of illumination, which states that the light falloff is proportional to the fourth power of the cosine of the angle at which the incident angle hit the detector. Because rays at the periphery of the image arrive at greater angle than in the center, the intensity of the incident light as seen by the pixel is lower. In addition to these two effects, dark circular spots can be seen in some areas of the image, indicating the presence of some dust on the surface of the optical elements. All these imperfections are present in all images taken with the instrument, and the flat frame correction aims to effectively correct them.

Figure 4.7: Master flat frame.

The camera settings used for the flat frame are summarized in Tab. 4.1. They were chosen such that the image brightness is in the upper part of the detector's dynamic range,

but without saturating it. The intensity distribution of the MFF, shown in Fig. 4.8, illustrates the composition of the MFF. The high peak corresponds to the majority of white pixels in the flat frame. If the latter was perfect, there would be only this narrow peak in the distribution. However, it extends to the left of its base over an intensity range of around 50 to 130. This spread corresponds to the darker areas and vignetting present in the MFF. It must be noted that this intensity distribution was computed only on the detector pixels comprising the image of the DMD. Therefore, the pixels of the dark margin around the image as shown in Fig. 4.7 are not accounted for in the intensity distribution.

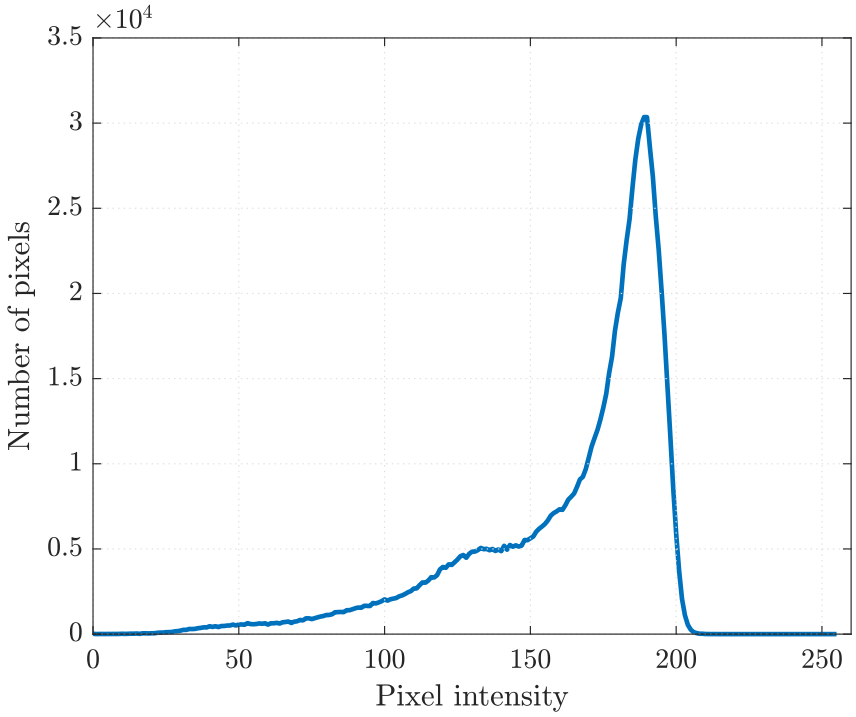


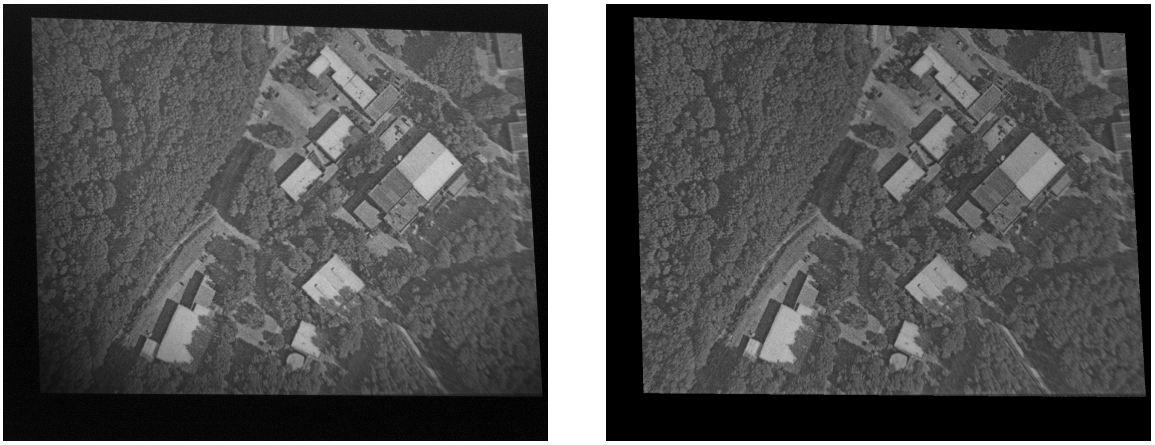
Figure 4.8: Pixel intensity distribution of the master flat frame. This distribution only accounts for pixels being part of the image of the DMD.

Once the master flat frame is obtained, the non-uniformities of the image can be corrected. The corrected image C after the flat frame correction can be expressed as

$$C = \frac{R - D_R}{F} m, \tag{4.2}$$

where R is the raw image to calibrate, F is the master flat frame image, and D_R is the MDF taken with the scene camera settings. The scalar m is the mean intensity of the pixels composing the MFF. In this equation, the terms $R - D_R$ corresponds to the dark frame correction of the raw scene image.

A comparison between a raw image of the aerial view of the CSL and this image corrected with the MDF and the MFF is shown in Fig. 4.9. The corrected image now has a uniform exposure, and the previously dark corners are now at the same level of brightness as the rest of the image. Apart from this flat frame correction, the intensity of the pixels composing the dark borders of the image have all been set to 0, because they are not part of the image of the scene.



(a) Raw image **(b)** Corrected image

Figure 4.9: Comparison between the raw image, and the image corrected by the dark and flat frames.

4.4 Pattern mask

At this stage, the damaged image of the scene is ready to be reconstructed. For this reconstruction to take place, the algorithm needs information about the positions of the pixels that belong to the image as well as the positions of the pixels that were destroyed by the DMD pattern and therefore need to be reconstructed. For this purpose, a pattern mask must be created. This is a binary image where each black pixel corresponds to a pixel to be reconstructed by the algorithm, and each white pixel is an undamaged pixel. Using the binary pattern employed to code the DMD is not an adequate way to obtain an accurate pattern mask. Indeed, it does not pass through the entire instrument and therefore does not undergo the optical distortions that the damaged image of the scene has undergone. As a result, there will not be an exact match between the pixels of the mask and those of the image. A solution to have a pixel-perfect pattern mask is to take a picture of a uniformly illuminated white scene with the imager while the pattern is projected on the DMD. Thus, the resulting image, visible in Fig. 4.10, is similar to a flat frame, except that the DMD pattern is imprinted on it. To obtain the binary mask, this uniform illumination of the pattern is used, and a binary thresholding is applied to it. Image thresholding is a method of binarizing a grayscale image, based on the intensity of each pixel to define the binary state of the binary image.

4.4.1 Binary thresholding

As explained just above, thresholding the image allows converting the grayscale image of the pattern into a binary mask of the pattern. The simplest solution for achieving this is the simple binary thresholding operation. Each pixel in the image is replaced by a black pixel if its intensity is below a predefined threshold, otherwise it is replaced by a white pixel. Mathematically, the simple binary threshold operator can be written as:

$$I'(x,y) = \begin{cases} 1, & \text{if } I(x,y) \geq T, \\ 0, & \text{otherwise,} \end{cases} \tag{4.3}$$

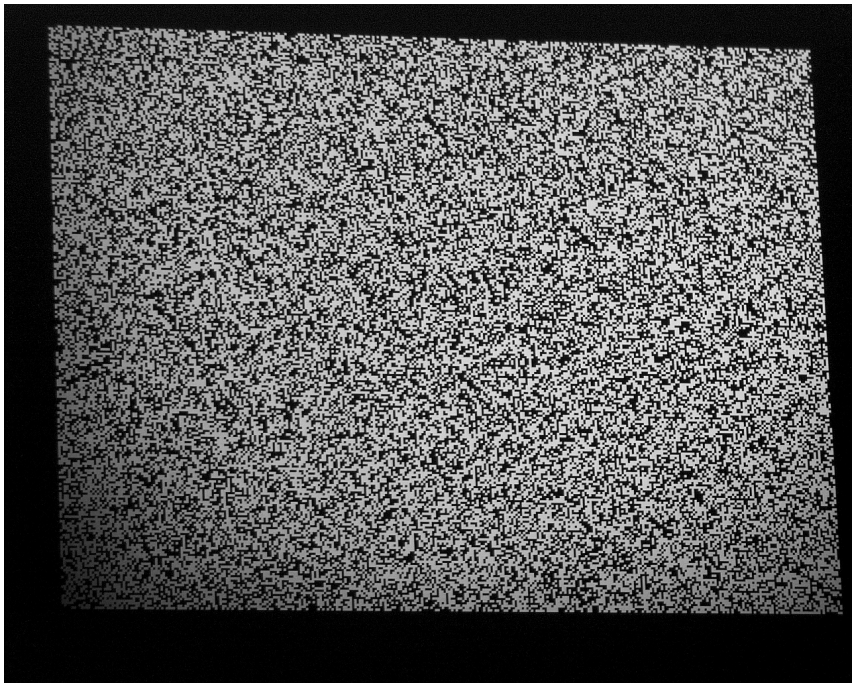
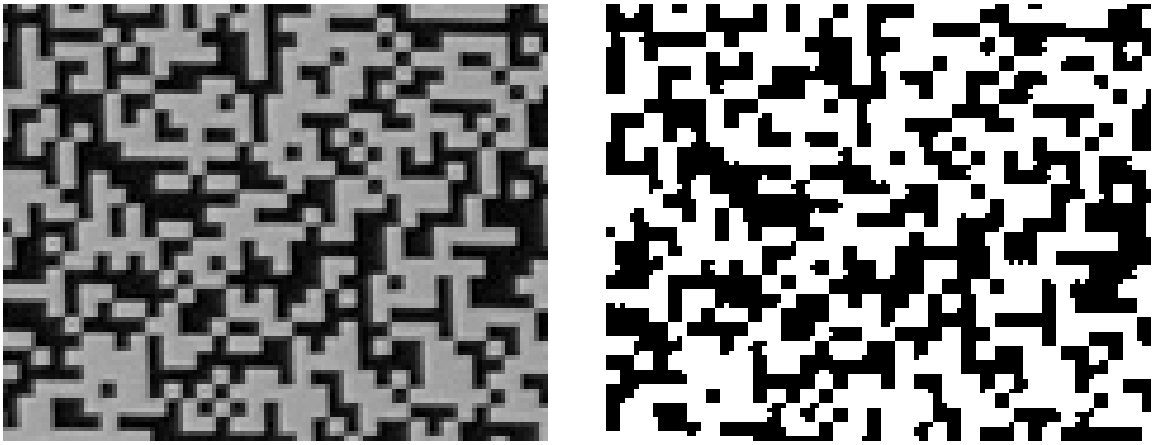


Figure 4.10: Raw image of the uniformly illuminated pattern. The applied mask is a random pattern, and binning of 3×3 and has a filling ratio of 50%.

where $I(x, y)$ is the grayscale image, T is the threshold, and $I'(x, y)$ is the binarized image. Although this thresholding method is simple, it yields good results for images with constant illumination across the entire field of view, which is the case here since, after flat frame correction, the image has a generally homogeneous illumination. However, the main drawback of this method is that the threshold parameter T must be chosen manually. This can be tedious, especially during development phases and fine-tuning of the imager, because it is necessary to repeat the operation every time the camera settings are changed, as this affects the pixel intensities of the image.

An improved version of the simple binary thresholding is the Otsu's binarization method [61]. This method avoids the need to manually choose a threshold, as it selects one automatically. To do this, it uses the image intensity distribution to find the optimal threshold. This method works particularly well when the image consists of two distinct categories of pixels, *i.e.*, when the intensity distribution has two well-separated peaks. It is thus well adapted for the uniformly illuminated pattern image, as it has well-defined dark and light areas. A zoomed-in view of the uniformly illuminated calibrated pattern image and the zoomed-in view of its binary mask obtained by the Otsu's method are shown in Fig. 4.11. It can be observed that the binary mask faithfully represents the general shape of the pattern, but a few pixels are still missing in certain areas.

To determine if the created mask accurately represents the mask applied to the image, the pixels of the damaged image marked as black on the pattern mask can be replaced with the pixels from the undamaged image. If the pattern mask accurately reflects the damaged image, this reconstruction using the original image should result in an image without any defects and without black pixels from the DMD mask. The result of this



(a) (b)

Figure 4.11: Pattern binary thresholding with the Otsu's method. (a) Zoom at the center of the processed illuminated pattern image. (b) Binary thresholding of the same portion obtained with the Otsu's binarization method.

artificial image reconstruction is shown in Fig. 4.12. It can be seen that this reconstruction using the undamaged image is poor. The reconstructed image shows small dark patterns scattered across the image. These small patterns correspond to the edges around the pattern present in the damaged image that were not perceived as damaged pixels during the thresholding operation. These edges, visible in the processed image in Fig. 4.11a, are gray transition pixels between the dark areas of the pattern and the white areas. The gray edges are caused by several factors. Firstly, the damaged image reflected by the DMD passes through the condensing lens and lands on the detector. Like any optical system, achieving infinite sharpness is impossible, because the optics blurs the image. This blurring causes a spreading of the pattern edges. Additionally, the focus on the detector

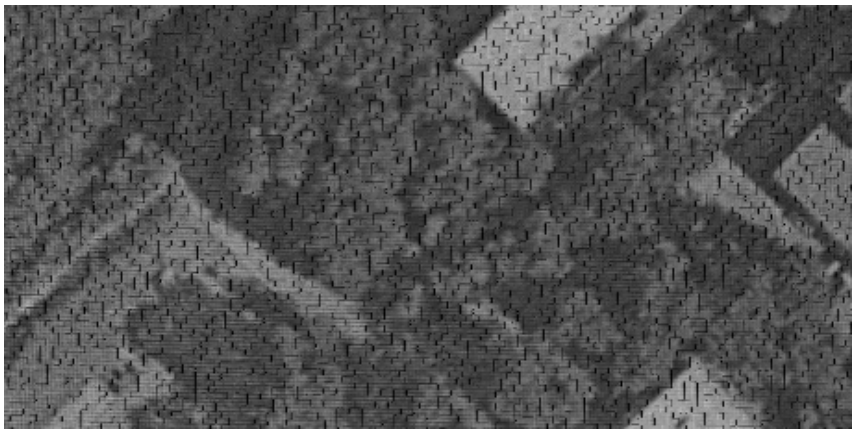


Figure 4.12: Zoomed-in view of the damaged image where the damaged pixels are replaced by the pixels of the undamaged image of the scene. The damaged pixels are determined based on the pattern mask binarized with the Otsu's method.

is certainly not perfect either, which also leads to the spreading of the edges. Moreover, the pixel grid of the DMD is not aligned with that of the detector. This misalignment

causes the edges of the pattern contours to sometimes fall in the middle of the detector's pixels. These pixels will thus capture a portion of the white image and a portion of the black. The light integration of this edge by the pixel will therefore produce a gray pixel.

The result in Fig. 4.12 therefore indicates that the current pattern mask has too many white pixels and that the mask considers some pixels of the image as undamaged when they are actually partially or completely damaged by the DMD. This is confirmed by the filling ratio of the created pattern mask, which is 55.9%, while the filling ratio of the binary pattern displayed on the DMD is 50%. If this problem is not addressed, all these edge pixels of the mask will be included in the reconstruction algorithm and will be considered as part of the image, thus distorting the measurements. This will ultimately reduce the quality of the image reconstruction.

4.4.2 Morphological erosion

Binary thresholding alone does not allow for obtaining a satisfactory pattern mask. One solution to eliminate the gray edges of the pattern is to erode the white areas in the pattern mask by a sufficient number of pixels. By proceeding in this manner, the gray edges will no longer be considered among the undamaged pixels of the scene image. The technique used to perform this operation is called morphological erosion. Morphological erosion uses a structuring element of a specified size and shape. A structuring element is a 2D matrix that identifies the pixel to be processed as well as the adjacent pixels used in this processing. The algorithm loops through all the different pixels of the image. For each pixel, the center of the structuring element is placed on the current pixel. The new intensity of this pixel is the minimum intensity of all the pixels contained within this superimposed structuring element. In other words, the erosion of a pixel is the minimum intensity of the pixels in its neighborhood defined by the structuring element. In the case of the pattern mask, it is a binary image composed of 0s and 1s. Therefore, as soon as a black pixel is included in a structuring element, the pixel associated with this structuring element will become a black pixel. Thus, this is indeed an erosion at the expense of the white pixels in the pattern mask.

As it can be seen in Fig. 4.11a, the thickness of the gray edges is generally one pixel. Therefore, the structuring element must have the smallest possible size, which is 3 pixels on each side. After several tests of the structuring element's shape, it turns out that the rectangular shape is the best for eliminating the gray edges while preserving the maximum number of undamaged white pixels. Hence, the structuring element is simply a 3×3 square. The result of the eroded pattern mask is shown in Fig. 4.13. It can be seen that the erosion has thickened the pattern on average by one pixel on all the edges.

To determine if this erosion has effectively removed the gray edges, a reconstruction using the undamaged image, as in Fig. 4.12, can be performed with the eroded pattern mask. A zoom of this new reconstructed image is shown in Fig. 4.14. It can be observed that the image is perfectly recovered, and that the gray edges are no longer present. This ensures that the pattern mask only selects the pixels undamaged by the DMD, thus allowing for optimal image reconstruction. Another more precise way to verify that the pattern mask no longer includes the gray edges is to use the pixel intensity distribution of the uniformly illuminated pattern image. More specifically, the intensity distribution of

(a) (b)

Figure 4.13: Comparison between the pattern mask before and after morphological erosion. (a) Zoomed-in view at the center of the pattern mask before erosion. (b) Same portion of the pattern mask after the morphological erosion.

the pixels in the pattern image before binarization is calculated only at the positions in the image that have a white pixel on the binary pattern mask. Thus, the distribution should contain only the pixels of the uniformly illuminated scene and should therefore present a single, well-defined sharp peak. Fig. 4.15 shows the intensity distributions of the uniformly illuminated pattern masked by the pattern mask before and after the erosion. First, it can

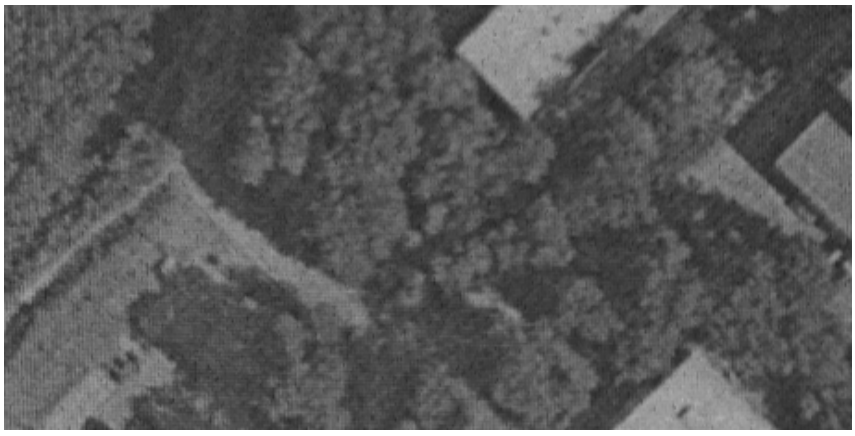


Figure 4.14: Zoomed-in view of the damaged image reconstructed by the undamaged image of the scene. The damaged pixels are determined based on the pattern mask binarized with the Otsu's method, followed by a morphological erosion.

be observed that before erosion, the remaining pixels in the pattern image form a peak distribution, but there is also a plateau between the pixels with intensities between 85 and 150. These are mainly the gray pixels composing the edges of the image. The effect caused by the thresholding can also clearly be seen, as no pixels with intensities below 85 are present. The second curve, in orange, represents the remaining pixels in the pattern image after the erosion of the pattern mask. This second distribution now contains only a sharp peak, representing the undamaged pixels of the uniformly illuminated scene. Thus,

it can be concluded that the erosion operation has effectively eliminated the gray edges around the pattern. It can also be observed that the peak in the distribution after erosion

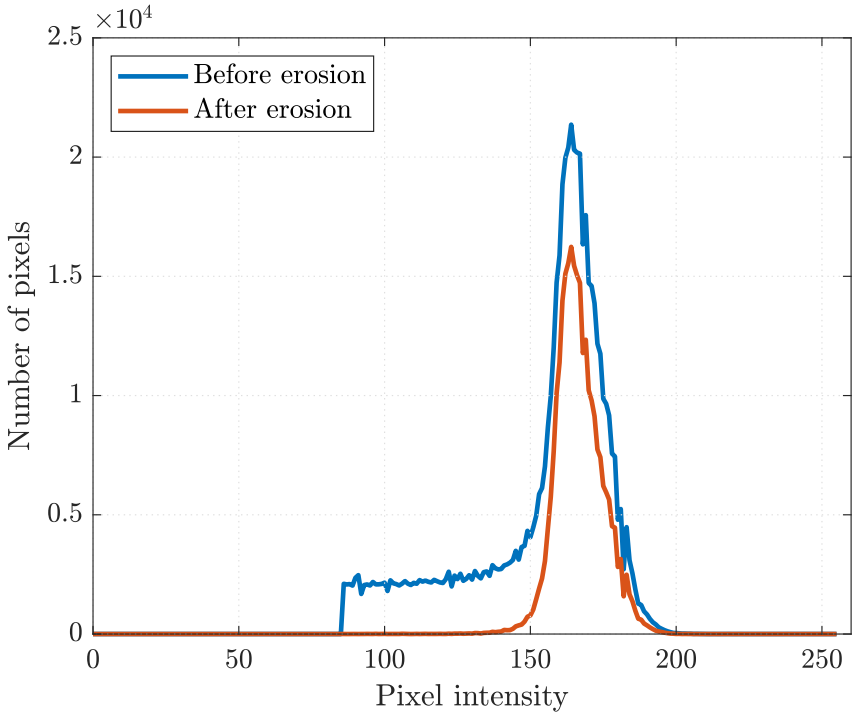


Figure 4.15: Pixel intensity distribution of the uniformly illuminated pattern image covered by the binary pattern mask before and after the morphological erosion.

is lower than before erosion. This means that the erosion also impacts white pixels that are not supposed to be damaged. This is an undesirable effect of the morphological erosion. The consequence of this effect is that some undamaged information from the image is lost and will not be used in the reconstruction process. This is particularly evident in the filling ratio of the pattern mask after erosion, which is now only 26.9%. This decrease in the filling ratio is due both to the elimination of the gray borders and to the removal of undamaged pixels by the erosion. It is clear that the impact of erosion on the filling ratio is related to the initial shape of the pattern applied to the DMD. More specifically, the filling ratio is influenced by the number of pixels forming the boundary between the white and black areas of the binary pattern mask. The larger the contact surface between the two areas, the more gray borders there will be in the image, and the more morphological erosion will remove pixels. Random patterns, like the one used in this section, have a very large contact surface between the white and black regions, so the erosion operation has a significant effect on the filling ratio. Conversely, for the same filling ratio, a pattern composed of a simple rectangular shape offers much less contact surface between the two binary areas of the pattern. Therefore, morphological erosion will have less effect on the filling ratio of this simple pattern mask.

It must be noted that this simpler type of mask is not advantageous for image reconstruction because it damages a large area in the photo. The more one moves inside this area, the less undamaged information is available in the neighborhood. As a result,

the reconstruction in this large area is poor. On the contrary, using masks with random patterns allows for a better distribution of damaged pixels across the entire image. Thus, each damaged area is smaller and always has information about the image nearby, enabling higher quality reconstruction.

Looking at Fig. 4.15, one might think that a better-positioned binary threshold (around an intensity of 150) could remove the gray borders without eliminating pixels in the peak of the distribution. However, experimentation with the threshold shows that this method alone does not entirely remove the gray borders as erosion does. This can be explained by the fact that binary thresholding applies to the entire image based solely on each pixel's intensity. In contrast, morphological erosion specifically targets contact areas between light and dark pixels, effectively eliminating the gray borders. Additionally, this method requires manually setting the threshold, which was not desirable initially.

4.5 Point spread function

When light from the scene enters the instrument, the optics through which it passes deforms it. Due to aberrations generated by the lenses, diffraction effects, or lack of focus on the detector, the image becomes blurred in one way or another. The quality of an optical instrument can be characterized by the Point Spread Function (PSF). A PSF is the image of a point source placed at infinity and captured by the instrument. The PSF hence represents the instrument's impulse response and includes all the deformations the wavefront undergoes as it passes through the imager. These deformations cause a spreading of the image of the point source, reducing the overall image quality. When the optical system is considered invariant, meaning that a shift in the scene in the object plane simply results in a corresponding shift of the image in the image plane, then the image of the scene through the instrument is expressed as the convolution of the original image with the PSF

$$\boldsymbol{f} = \mathbf{h} * \boldsymbol{f}_{\text{pure}}, \tag{4.4}$$

where $\boldsymbol{f}$ is the observed image on the detector, $\mathbf{h}$ is the PSF of the instrument, and $\boldsymbol{f}_{\text{pure}}$ is the pure image before its acquisition with the imager. The knowledge of the instrument's PSF therefore allows deconvolving the image, which is the inverse operation of Eq. 4.4, to recover the original image. This deconvolution operation is widely used in astrophotography to improve the quality and sharpness of images.

In the case of our imager, the presence of the DMD in the optical chain alters the method for acquiring the PSF. Instead of imaging a point source with all the mirrors of the DMD activated, it is obtained by activating only a single micromirror at the center of the DMD and imaging a uniformly illuminated scene. The image obtained on the detector is therefore the impulse response of a single micromirror. When imaging a scene with a pattern applied to the DMD, the image on the detector is the combination of the individual contributions of each micromirror. By knowing the impulse response of a single micromirror, it is thus possible to return to the expression of Eq. 4.4. An image of the PSF of the imager is shown in Fig. 4.16. It can be seen that the PSF has a non-symmetrical shape and spreads over several pixels. The lack of symmetry in the PSF is caused by the fact that the instrument itself is not symmetrical, as shown in the general schematics of the imager in Fig. 3.1. This spreading of the response from a single micromirror over

several pixels on the detector will induce blurring of the image. With this PSF, it is now possible to deconvolve the image obtained on the detector with this PSF, thereby unblurring it and improving image quality. In fact, the deconvolution operation can be performed with compressive sensing and sparsity principles, and can be implemented during the reconstruction process. Indeed, the convolution of Eq. 4.4 can be expressed as

$$\boldsymbol{f} = \mathbf{H}\boldsymbol{f}_{\text{pure}}, \tag{4.5}$$

where $\mathbf{H}$ is a block circulant matrix representing the convolution of the pure image by the PSF. This system cannot be directly inverted to retrieve the original image because the matrix $\mathbf{H}$ is generally badly conditioned or non-invertible [6]. Instead, it can be solved by adding the prior information that $\boldsymbol{f}_{\text{pure}}$ is sparse in some basis. With this hypothesis, the algorithms seen in Sec. 2.3 can be used to deconvolve the image. In the context of image inpainting, the initial reconstruction problem can be expressed as $\mathbf{y} = \mathbf{A}\mathbf{x}$ (see Sec. 2.4), where $\mathbf{A}$ expresses the effect of the patterning of the image by the DMD. The addition of the deconvolution in this equation yields

$$\mathbf{y} = \mathbf{A}\mathbf{x} = \mathbf{A}\mathbf{H}\boldsymbol{f}_{\text{pure}}, \tag{4.6}$$

where matrix $\mathbf{AH}$ now becomes the new measurement matrix. The deconvolution of the image now takes place during the reconstruction process of the image.

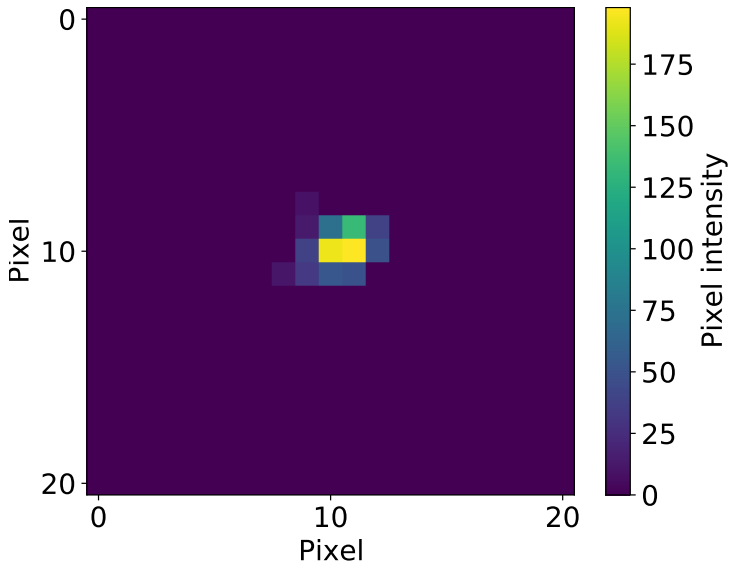


Figure 4.16: Zoomed-in view of the central PSF of the imager. The image is obtained by activating only the central micromirror of the DMD, and imaging a uniformly illuminated scene.

4.5.1 Improvement of the PSF quality

Since the region of interest of a PSF is generally very small, it only spreads across a few pixels of the detector, as seen in Fig. 4.16. In its operation, a detector pixel samples a portion of the scene by integrating all the incident light of this portion on its surface and outputs a single intensity. Therefore, small objects like the PSF are significantly affected by the sampling of the detector. The integration by the detector thus results in

a substantial loss of information about the initial shape of the PSF. Moreover, due to this sampling, the pixelated image of the PSF obtained on the detector strongly depends on the relative position of the PSF image on this detector. In order to be less impacted by the detector's sampling effect on the PSF, it is beneficial to enhance the PSF image quality by increasing its resolution. This will firstly allow for better characterization of the instrument, and secondly, improve the quality of the deconvolution.

To increase the resolution of the PSF, a straightforward solution would be to replace the current camera with a detector that has a higher resolution. However, this is not ideal because there are no other higher resolution cameras available in the optical store, and even if there were, the PSF resolution would rapidly become limited. A solution that does not require different or more advanced equipment is to implement dithering. Dithering is a process that introduces small fluctuations in the instrument's orientation or the detector's position between exposures to improve the final image quality. It can be employed in various ways. In astrophotography, dithering involves slightly shifting the telescope's pointing between exposures of the same object [62]. This causes the scene to move slightly on the detector in each image. During post-processing, the images are stacked and realigned. This method eliminates fixed-pattern noise and defective pixels, thereby increasing the SNR of the final image. Dithering can also be achieved by slightly shifting the detector on the image plane between each acquisition. This way, each image corresponds to the same scene but sampled differently by the camera's pixels. This increases the number of different measurements for the same object. Each of these acquisitions are then stacked and realigned in a precise manner to form a single image of the scene. Thanks to this realignment of images, the sampling rate of the new image is artificially increased, thereby enhancing its resolution [63]. This dithering method thus allows for achieving a higher image resolution than the resolution of the instrument's detector. It is this second use of dithering that will be employed here to increase the sampling rate of the PSF.

Here, it is not possible to move the camera relative to the image plane to perform dithering. However, the DMD can be used to artificially shift the PSF image on the detector. Instead of illuminating the central micromirror of the DMD, an adjacent mirror can be illuminated, creating a slightly shifted PSF image on the detector. The detector will sample the PSF differently and extract additional information from this image. By repeating this operation for several micromirrors adjacent to the center of the DMD, all the different images of the shifted PSFs can then be realigned and stacked, and a complete PSF image with a higher sampling rate than initially can be obtained.

To realign the different PSF images, it is important to quantify the shift of the PSFs image on the detector based on which micromirror is activated on the DMD. To do this, two coordinate systems are introduced and represented in Fig. 4.17. The first coordinate system defines the coordinates of points in the detector plane and is shown in blue in the figure. It has a domain corresponding to $(x, y) \in [0, 1280] \times [0, 1024]$, which corresponds to the number of pixels on the detector. This coordinate system represents the positions of the camera pixels in the image. The second coordinate system, illustrated in red in the figure, represents the coordinates of the micromirrors in the plane of the DMD surface. This coordinate system has its origin located in the lower right corner of the

DMD image since the collecting lens flips the image of the DMD. It can be observed that this micromirror axis system is not orthonormal in the detector domain. Indeed, due to the perspective effect, the image of the DMD and thus the position of the micromirrors on it are distorted. The domain of this coordinate system corresponds to $(x, y) \in [0, 1024] \times [0, 768]$. By knowing the coordinates of the four corners of the DMD image in the detector domain (see Sec. 4.6.1), it is now possible to establish the relation between a coordinate on the DMD and its coordinates on the detector. In other words, this allows us to obtain the position (x, y) on the image of a micromirror for which we only know the coordinates (x', y') on the DMD.

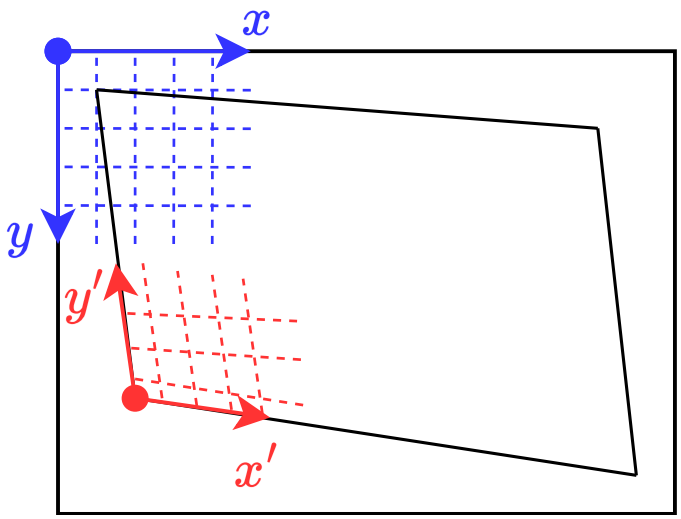


Figure 4.17: Definition of the different coordinate systems. The rectangular frame represents the border of the image taken by the camera, and the trapezoidal frame represents the image of the DMD. The coordinate system of the detector pixels is shown in blue while the coordinate system of the DMD's micromirrors is shown in red.

First, let's see how to realize dithering using only two PSF acquisitions. In the detector domain, the central micromirror m_c is located at the coordinate (x_c, y_c). Let introduce a micromirror m_i neighboring this central micromirror. In the detector domain, its coordinate is

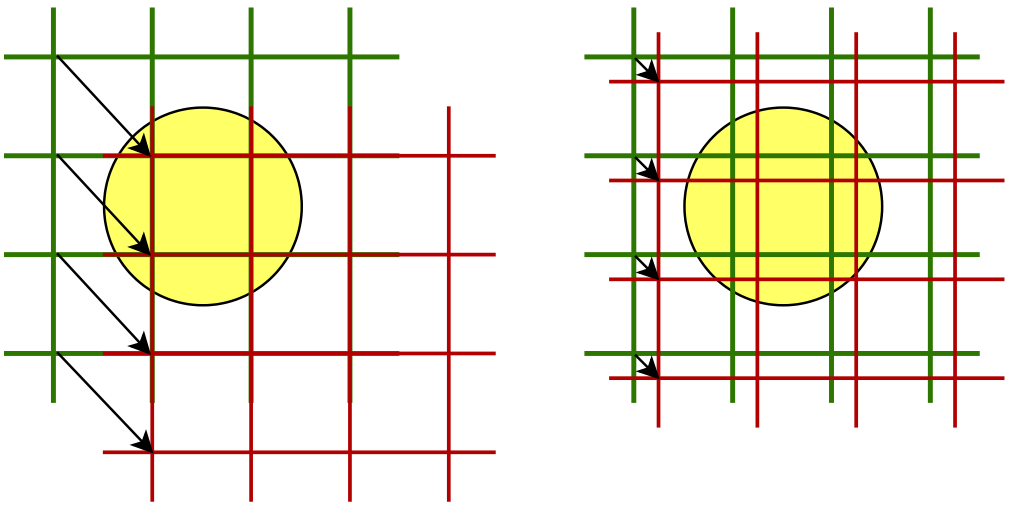
$$(x_i, y_i) = (x_c + dx, y_c + dy), \tag{4.7}$$

where dx and dy are the relative shifts in position in the detector domain between the two mirrors. They can be decomposed into two components

$$\begin{cases} dx &= dX + dp_x \\ dy &= dY + dp_y \end{cases} \tag{4.8}$$

where dX and dY are the integer parts of the displacement, and dp_x and dp_y are the decimal parts of the displacement. As previously mentioned, a detector pixel takes a measurement by integrating the intensity of the incident light on its surface. Knowing that, if $dp_x = dp_y = 0$, the two images are then identical, except for an integer pixel translation. This is not interesting because, in both images, the pixels sample the PSF

in the same way, just with different pixels. This case, where the image undergoes only an integer pixel shift, is illustrated in Fig. 4.18a. In this figure, the PSF is represented by a yellow disc. The pixel grid of the first image is shown in green, while the pixel grid of the shifted image is shown in red. It can be seen that the two images contain the same information about the scene. However, if the decimal shifts dp_x and dp_y are non-zero, then the two images will be different. As shown in Fig. 4.18b, different pixels sample the PSF differently. It is possible to recombine the two images taking this shift into account. Therefore, the image of a PSF generated by the mirror m_i that is shifted by (dx, dy) relative to m_c, is equivalent to obtaining the image of this central PSF but sampled with a pixel grid shifted by (dp_x, dp_y). In other words, photographing the two PSFs individually generated by the mirrors m_c and m_i is equivalent to photographing the same object from a different viewpoint, with each pixel integrating the object in a slightly different way. Since the micromirror grid of the DMD is not aligned with the pixel grid of the detector, and the mirrors do not have the same apparent size as the pixels in the image, it is unlikely to obtain $dp_x = dp_y = 0$.



(a) **(b)**

Figure 4.18: Illustration of the effect of the shift between two dithered images. The object to image in the scene is depicted as yellow disc. The pixels grid of the first image is shown in green while the pixels grid of the dithered image is presented in red. (a) The shift between the two images is an integer, the two pixels grids superimposes themselves. (b) the shift between the two images is not an integer, the two grids do not overlap.

In order to recombine these two images into one single image, let $S_{i,j,dx,dy}$ be the intensity measured by the pixel (i, j) on the PSF image obtained by the activating the micromirror m_i, which is shifted by (dx, dy) relative to the central micromirror m_c in the detector domain. Fig. 4.19 illustrates the stacking and alignment of the two images. In the black grid are shown the intensities of each pixel of the PSF image corresponding to the mirror m_c. The intensities of this black grid is therefore $S_{i,j,0,0}$. In purple are displayed the intensities of the pixels from the PSF image generated by m_i. These intensities form

a grid that is shifted by (dp_x, dp_y) relative to the first grid. By placing these two grids in this manner, we now obtain a PSF image with twice as much information as initially. The dithered image has filled in the information between the pixels of the original PSF image generated by m_c. By adding more PSF images from other micromirrors, the sample rate of the image can be significantly increased. Unlike dithering achieved by moving the detector on the image plane, where the offsets between two images can be finely adjusted, it is not possible to have complete control over the offsets dp_x and dp_y here. These offsets are indeed imposed by the array of micromirrors on the DMD, which determines the displacement between each PSF image on the detector.

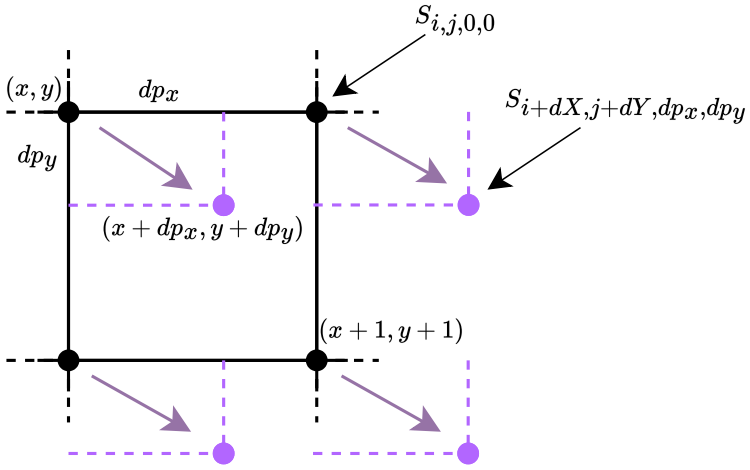


Figure 4.19: Stacking of the PSF images obtained by m_c and m_i. The two measurements are shifted from each other by dp_x and dp_y in the respective directions.

It is important to note that for the dithering to work, it is assumed that the PSFs created by all the individual micromirrors are identical. Considering this hypothesis to be true, all these PSFs can be superimposed. To ensure they are as similar as possible and not altered by a field effect due to the instrument, the micromirrors closest to the central mirror m_c are used in the dithering operation. Specifically, we define a square of micromirrors centered on m_c with a half-side length of N_{dit}. The parameter N_{dit} is an integer value that controls the number of images to include in the dithering. Thus, this square composed of $(2N_{dit} + 1) \times (2N_{dit} + 1)$ mirrors includes all the micromirrors that will be used to generate the different PSFs in the dithering operation. Increasing N_{dit} allows for more data, but it also quadratically increases the number of images to be acquired. Therefore, it is important to find the right value for N_{dit}, to obtain the best superposition of the data to cover the detector plane while minimizing the number of pictures to take. After several manipulations, it turns out that $N_{dit} = 3$ is a good compromise between the quantity of images to make and the sample rate of the final image. For N_{dit} below 3, there are not enough measurements, which leaves large area in the detector plane empty. On the other hand, for N_{dit} larger than 3, the data from the different images begin to overlap without bringing any essential new information about the PSF. Consequently, with $N_{dit} = 3$, 49 PSF images are needed to form the dithered

image. Each image is taken with the flat frame settings of the camera and is composed of 5 identical images stacked to increase its SNR. These images are then calibrated with a dark frame subtraction and a flat frame correction. A zoomed-in view of the raw result of the PSF obtained by dithering with $N_{dit} = 3$ is presented in Fig. 4.20. It can be seen that

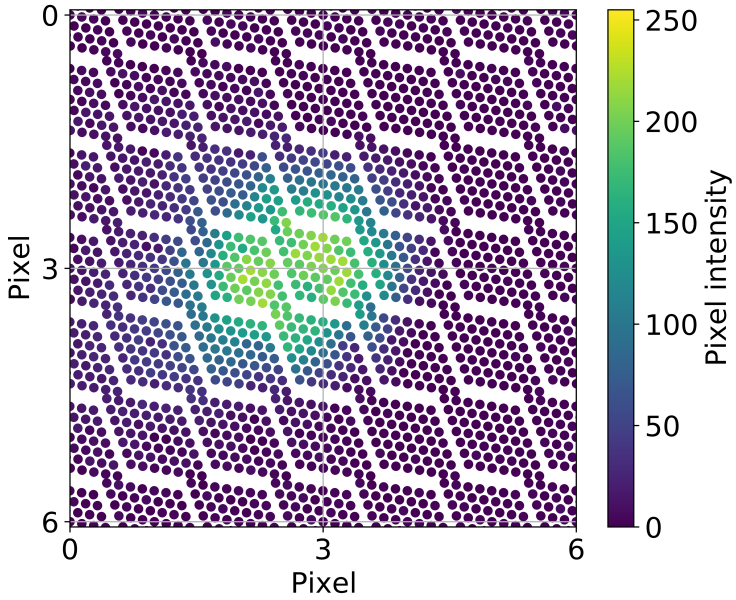


Figure 4.20: Raw data of the PSF obtained with dithering and $N_{dit} = 3$.

this is not an image composed of square pixels. Instead, it is an arrangement of points, each with its own intensity, as shown in Fig. 4.19. Each point corresponds to the intensity of a pixel from one of the 49 PSFs images. The points are relatively evenly spaced, but there are some areas not covered by data. Additionally, the points do not form a perfectly straight grid. These two observations are directly caused by the shape of the micromirror grid as seen from the camera and the fact that the PSFs can only be shifted to discrete, predetermined positions by the DMD. Aside from this, it can be observed that all the images have been successfully combined, and the PSF is visible as a well-defined spot with high intensity. This PSF is approximately 2 pixels in diameter, which is consistent with the PSF image obtained without dithering in Fig. 4.16.

As it stands, the measurements in Fig. 4.20 are not usable. To obtain a manipulable image, the values of the points must be mapped onto a straight and rectilinear grid. To achieve this, all the data from this figure are linearly interpolated onto a regular grid. The size of this grid, *i.e.*, the sampling rate, should not be too small to avoid over-interpolation, which could create artifacts in the image. Likewise, the grid size should not be too large to ensure that the dithering operation's efforts are not wasted by excessively reducing the sampling rate. The grid size is determined based on the average spacing between the points. This yields to a mesh size of 0.14. This mesh size indicates that the dithering operation has increased the sampling rate by approximately 7. The interpolated image of the PSF on a regular grid is shown in Fig. 4.21. One can see in this image that the PSF consists of a bright central part and a tail spreading to the left of it. It can be noted that the shape of the obtained structure is generally the same as the low-resolution PSF

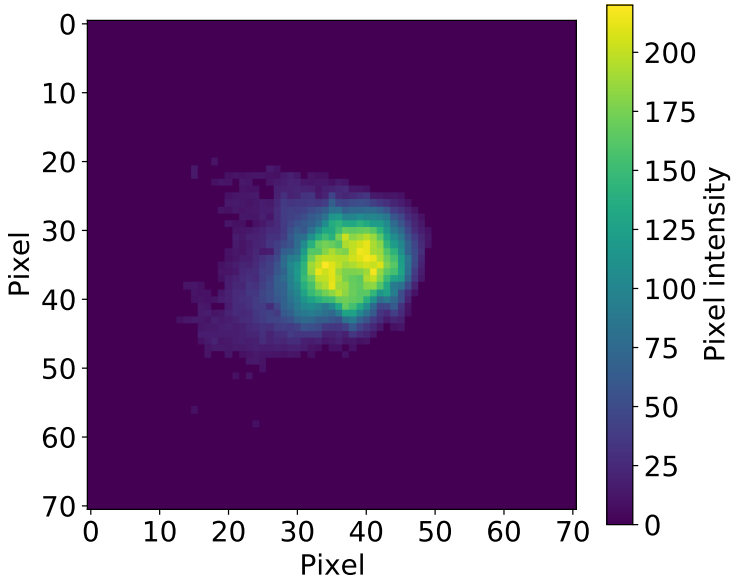


Figure 4.21: Zoomed-in view of the PSF obtained by linearly interpolating the data of the PSF retrieved by dithering.

obtained in Fig. 4.16. Both images have a bright central part roughly the size of two detector pixels. They also both have a tail extending to the left and downward from the center of the spot. Furthermore, it can be inferred that the central part of the PSF does not have a circular shape but rather a square shape. This shape corresponds to that of the micromirror tilted towards the camera, forming the instrument's PSF. This increase in the sampling rate of the image naturally comes at a cost. To capture a single photo of the PSF, it is necessary to obtain $(2N_{dit} + 1) \times (2N_{dit} + 1)$ PSF images, which amounts to 49 images for $N_{dit} = 3$. Each of these 49 images are also taken multiple times to be stacked in order to improve the SNR of the image. Ultimately, several hundred images are needed to achieve the final result. The number of images required can rapidly become prohibitive if a higher sampling rate is desired by increasing N_{dit}, as the total number of images required grows quadratically with N_{dit}.

The PSF obtained through dithering can be used to improve the deconvolution process of the image. To do this, the resolution of the image to be deconvolved must also be increased so that the PSF has the appropriate size relative to this scene image. This can be achieved by sampling each pixel of the image in such a way that these new pixels are the same size as the pixels of the PSF shown in Fig. 4.21. In this case, the resolution increase should approximately be by a factor 7, corresponding to the increase in the PSF sampling rate by the dithering operation. A single pixel of the image will thus be converted into a 7x7 pixel square with the same intensity as the original pixel. Consequently, the dimensions of the image to be reconstructed are also enlarged, which is detrimental for the reconstruction step as it increases the problem size and computation time. To limit this effect while retaining the benefits of dithering, the mesh size of the PSF during the linear interpolation operation on the structured grid should be increased. This way, the image size will be less enlarged, at the expense of the PSF resolution.

To test this dithering method with the instrument, it is interesting to verify if the result obtained in Fig. 4.21 is repeatable across multiple observations. Indeed, due to the constant presence of noise in measurements, the result may be heavily influenced by noise, rendering it unreliable and unusable. Therefore, obtaining the same result twice indicates that the process is robust to noise and that the results can be trusted. Thus, the same series of PSF images was captured one day after the one presented in Fig. 4.21, under the same conditions. The comparison of the two images obtained after dithering is shown in Fig. 4.22. The image on the left corresponds to the previously obtained PSF, and the

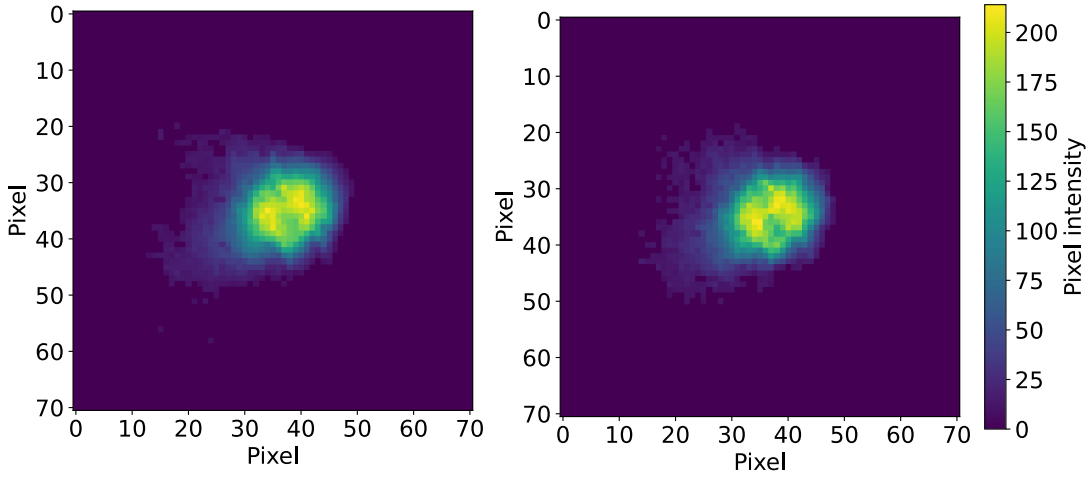


Figure 4.22: Repeatability test of the image of the PSF obtained with dithering. The PSF on the left is the image of Fig. 4.21. The PSF on the right is the image obtained one day after, with the same dithering parameters as in Fig. 4.21.

image on the right corresponds to the new PSF captured one day after the first. We notice that the two images are almost identical and exhibit the same characteristics. The center of the PSF has the same shape in both, and the overall shape of the tail is the same. It is in this tail area that slight differences appear, likely caused by noise, as the signal level is much lower there. These results demonstrate that the dithering method produces reliable results that are not significantly affected by noise. We can see in both images that there is a dimmer area in the central part of the PSF. One possible explanation for this spot is the characteristics of the DMD micromirrors' surface. As shown in Fig. 2.9, each mirror is attached to its mechanical support by a square stem. At the point where this stem connects to the mirror, it leaves a small non-reflective square area. Thus, the center of the mirror does not reflect incident light and remains dark. However, it is not possible to definitively assert that this non-reflective part is the cause of the observed decrease in brightness near the center of the PSF.

4.5.2 Variability of the PSF

To perform the deconvolution operation, it was assumed that the PSF was identical across the entire field of view of the instrument, and thus only one PSF was needed for the deconvolution. However, in reality, the PSF varies depending on the field of view

and is never identical to the one at the center of the image. There will generally be more aberrations on the edges of the image, resulting in a more spread-out and less concentrated PSF. If these edges of the image are deconvolved with a smaller central PSF, this region will remain blurry or will be distorted by the deblurring operation. Therefore, it is important to study the behavior of the instrument's PSF as a function of its position on the image. To do this, multiple micromirrors can be activated across the entire DMD screen. These micromirrors must be sufficiently spaced apart to ensure that one micromirror does not influence the PSFs of the others. Thus, a 5×5 grid of micromirrors is displayed on the DMD, spaced out over its entire surface. The image obtained by the imager is shown in Fig. 4.23, with the different PSFs highlighted by a yellow circle. A zoomed-in view of each of the 25 PSFs present in this image is shown in

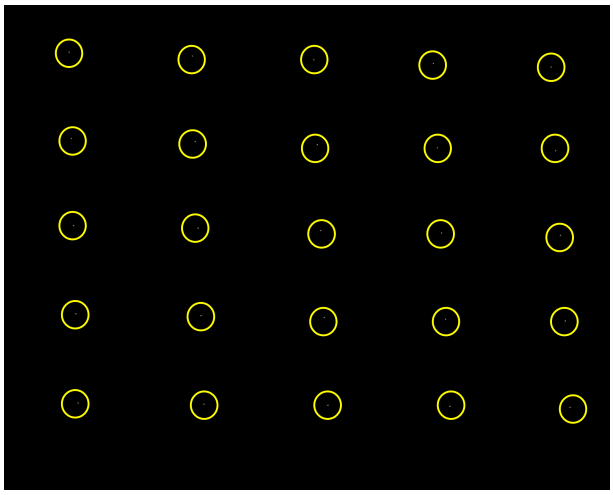


Figure 4.23: Image of the grid of PSFs taken by the instrument. The positions of the different PSFs are highlighted with yellow circles.

Fig. 4.24. Regardless of their location, the PSFs generally have the same size, with their bright central part being approximately two to three pixels wide everywhere. Also, They all have a lower intensity tail extending from this central part. This indicates that the PSFs exhibit a uniform behavior throughout the entire image, indicating the absence of substantial aberrations in any specific areas. However, these tails do not all point in the same direction. For the PSFs located in the upper left corner of the image, the spreading of the tail is directed towards the upper left corner of the image. Conversely, for the PSFs in the lower left corner of the image, the tail is directed towards the lower left corner. In fact, all the tails of the PSFs seem to align along lines that converge at the same point. This point appears to be located towards the center of the right edge of the image. This phenomenon, can be caused by the asymmetry of the optical instrument or imperfections in the lenses or alignment. However, these PSF images should not be over-interpreted since they are of low resolution and thus suffer from the sampling effect of the detector. Nevertheless, the main characteristics of the PSFs, such as size and general shape, are sufficiently visible to not consider using dithering to increase the sampling rate in order to study their behavior.

Knowing the shape of the PSF at different locations in the image allows for increased

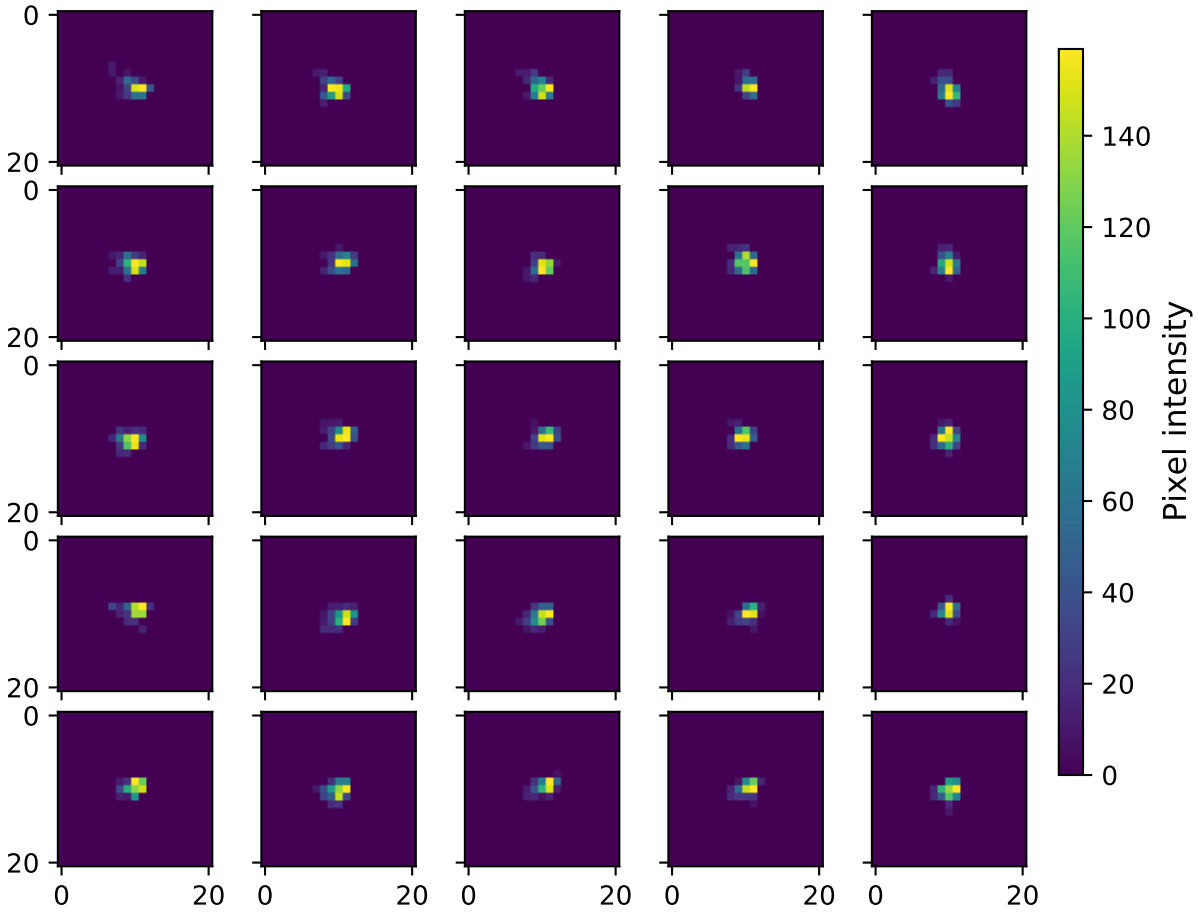


Figure 4.24: Zoomed-in view of the different PSFs shown in Fig. 4.23.

accuracy in the deconvolution process through block deconvolution. This method divides the image into regions, each associated with different PSFs. Each of these regions are deconvolved using their respective PSFs. The deconvolved regions are then recombined to reconstruct the image. The areas between each region can be obtained through interpolation to avoid sharp transitions between regions.

Although they are similar, the PSFs exhibit variations in shape depending on their position within the detector's field of view. Therefore, the assumption of PSF invariance made in the previous section for performing dithering is not valid across the entire detector. Consequently, during the dithering operation, it is crucial to only recombine PSF images from the same location, very locally on the detector. Failing to do so would result in recombining images with different PSFs, which would distort the outcome.

In conclusion, several deconvolution techniques have been highlighted. The first is to perform dithering on the PSF at the center of the image and deconvolve the entire image with this PSF. The second method is block deconvolution with a grid of PSFs without dithering, as explained earlier in this section. Each of these methods has its own characteristics and will certainly yield different results. Determining which of these methods is best suited for this problem requires the practical implementation of these various deconvolution methods in the reconstruction algorithm, which is beyond the scope of this work. A hybrid method can also be envisioned, combining the two approaches. This would involve block deconvolution where each PSF in the grid is dithered to increase its resolution.

4.6 Perspective correction

At this stage, the image taken with the instrument has been calibrated and reconstructed using an inpainting algorithm. The resulting image is no longer damaged, but it is still distorted by the perspective, and the scene occupies only a portion of the total image. This final step presented in this section, which serves as post-processing for the reconstructed image, should produce a straightened and correctly oriented image of the scene. This step is performed last in the calibration chain because it involves image deformation. Therefore, it is placed after the reconstruction step to avoid influencing it by degrading the scene image and the mask before reconstruction. The algorithm used to redress the image is called perspective warping. It takes as input the coordinates of the four corners of the image to be straightened. The output is a rectangular, straightened image with the desired dimensions. The first step is to locate the corners of the DMD in the image, which is done in the initial part of this section. Then, the algorithm will be explained, and the results of the perspective correction will be shown.

4.6.1 Corner detection algorithm

The image to be straightened is a quadrilateral with sides that are not parallel to each other. As mentioned earlier, the perspective correction algorithm requires as an input the coordinates of the four corners of this quadrilateral. The most straightforward solution would be to manually select the coordinates of these points based on the images obtained by the imager. However, this assumes that the DMD will never move in the detector's image. A slight change in a geometric parameter in the imager's secondary arm, such as

the focus of the condensing lens, would result in a different position of the DMD in the final image. This would necessitate redetermining the corner positions with each change, which would rapidly become tedious. A more suitable solution would be to develop an algorithm for detecting automatically the DMD corners. Thus, the corners of the image will be detected with each new image processed. Assuming the algorithm is effective, this ensures that the corners correspond to the image regardless of any changes made to the optical setup. It is this automatized solution that was chosen, and the principle of this algorithm is presented in the following of this section.

As seen in the MFF shown in Fig. 4.7, the sides of the scene image are straight. The corners of the image are located at the ends of these sides, so finding the Cartesian equations of the DMD sides is necessary to then find the intersections of these lines to obtain the coordinates of the four corners. An illustration of the corner detection method is shown in Fig. 4.25. The first step is to find these line equations. The starting point of this step is the MFF, as it is the image where the corners are the most clearly visible. Based on this MFF, a screen mask is created by performing a simple binary thresholding, which will make appear the DMD as fully white and the outer margins as black. Next, the transition pixels between the DMD and the margins, which constitute the sides, can be found by applying a gradient operator to the screen mask. Indeed, since the sides are characterized by a transition between a pixel of maximum intensity (white) and a pixel of minimum intensity (black), the gradient is high at these locations. Denoting $I(x, y)$ the pixel intensity of the screen mask for the pixel (x, y), the gradient operator of this image can be approximated by a finite central difference method

$$\vec{\nabla} I(x,y) = \begin{pmatrix} \dfrac{dI(x,y)}{dx} \\ \dfrac{dI(x,y)}{dy} \end{pmatrix} \approx \begin{pmatrix} I(x+1,y) - 2I(x,y) + I(x-1,y) \\ I(x,y+1) - 2I(x,y) + I(x,y-1) \end{pmatrix}. \tag{4.9}$$

By taking the norm of this gradient, and performing again a binary thresholding on the results, an image is obtained where only the edges of the DMD are marked as white pixels, forming an edge mask. Then, the image is divided into 9 identical regions as shown in Fig. 4.25, in order to isolate each side and obtain their Cartesian equations. Since the DMD occupies most of the space in the image, its corners must each be located in one of the image's corners. From this, we can identify four regions, each containing a segment of one side. For each segment, a linear interpolation is performed on the white pixels of the edge mask, which allows obtaining the coefficients a_i and b_i of the linear Cartesian equation $f_i(x) = a_i x + b_i$. Once the equations of the four sides are calculated, it is now possible to find the coordinates of the four corners. To do this, we calculate the intersection of the two lines corresponding to the two sides forming that corner. The coordinate of one corner is

$$(x, y) = \left(\frac{b_j - b_i}{a_i - a_j},\ a_i \frac{b_j - b_i}{a_i - a_j} + b_i \right), \tag{4.10}$$

where a_i, b_i, a_j and b_j are the respective coefficient of $f_i(x) = a_i x + b_i$ and $f_j(x) = a_j x + b_j$ the two intersecting edges. The coordinates of the four corners are now known. The

algorithm is able to provide results that are only a few pixels away from the visually identified corner in the image. The accuracy of the method is therefore more than sufficient for the perspective correction application.

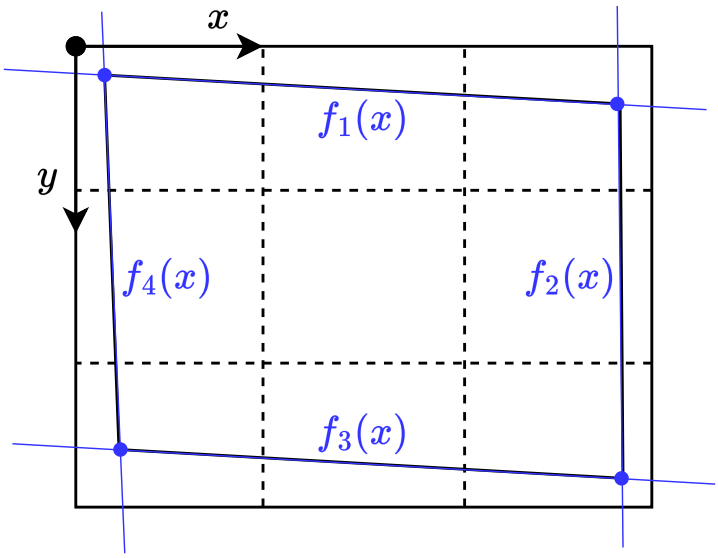


Figure 4.25: Illustration of the algorithm for the detection of the corners of the DMD. The image is divided into 9 equal regions. When the equations of the edges of the DMD are found, the coordinates of the corners are the intersections of these edges.

4.6.2 Perspective warping

The perspective warping algorithm, based on four points in an image, allows for the perspective transformation of the image subtended by these four points. It returns the corrected, rectangular image in the desired dimensions. The principle of this algorithm is illustrated in Fig. 4.26. In general, the perspective transformation can be expressed as [64]

$$\begin{pmatrix} t_i x' \\ t_i y' \\ t_i \end{pmatrix} = \begin{pmatrix} \alpha_1 & \alpha_2 & \beta_1 \\ \alpha_3 & \alpha_4 & \beta_2 \\ \gamma_1 & \gamma_2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \mathbf{M} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}, \tag{4.11}$$

where (x, y) is a point from the input image, (x', y') is the transformed coordinate of the input point and $\mathbf{M}$ is called the transformation matrix. Within this transformation matrix, the terms in α_i define transformations related to rotation and scaling, and other deformations. The terms in β_i define the translation vector, that does not deform the image but shift it in the image plane. Finally, the terms in γ_i define a projection vector. In the transformed domain, the term t_i is a scaling factor. It can be seen that t_i is directly related to the value of γ_1 and γ_2. In total, the transformation matrix $\mathbf{M}$ has 8 unknowns. In order to find these unknowns, the algorithm uses the coordinates of the four corners, which are injected into Eq. 4.11. Given that the coordinates in the transformed domain of the four corners are known, since we are defining the final dimensions of the image, this gives a linear system of 8 equations with 8 unknowns that is easy to solve.

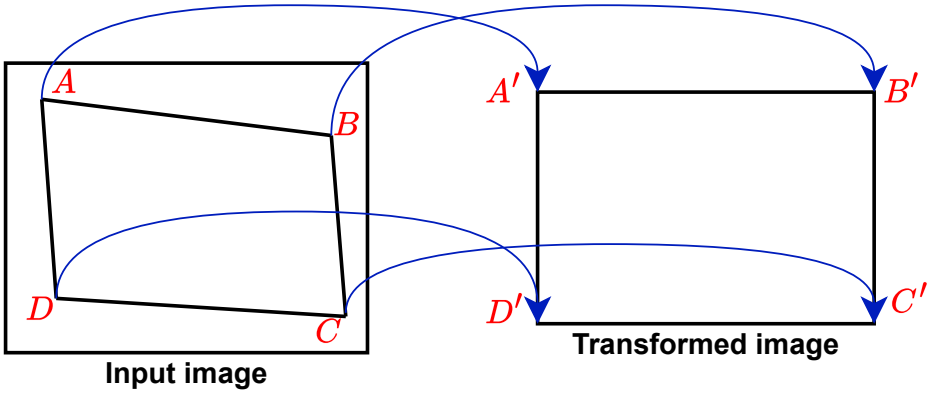
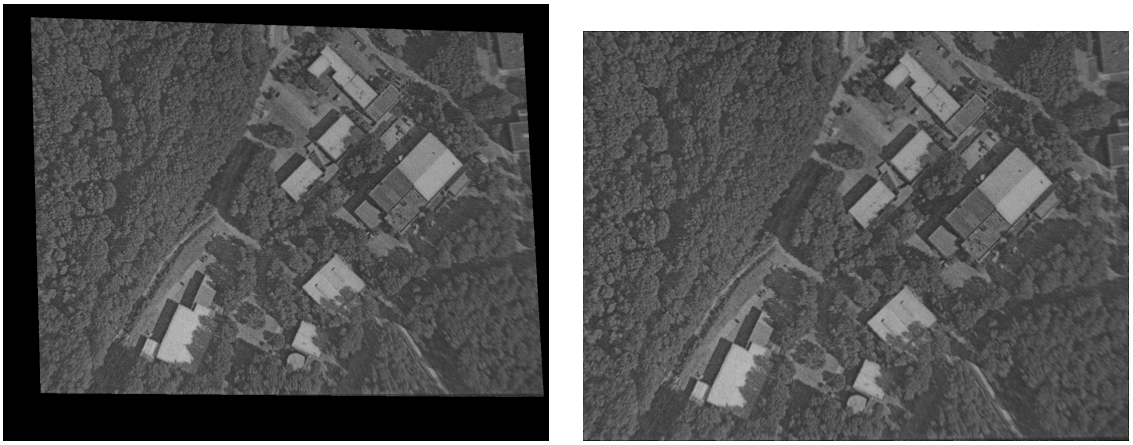


Figure 4.26: Principle of the perspective warping algorithm. On the left, the input image which contains the DMD image inside it. Based on the coordinates of the four corners of this image, the image is transformed into a rectangular image, as seen on the right of the figure. Adapted from [64].

Regarding the dimensions of the final image, they should be as close as possible to the dimensions of the trapezoidal image to minimize image scaling and avoid resolution loss. If the final dimensions are too small compared to the original image, compression will occur, leading to a loss of information. Conversely, if the dimensions are too large, the image will be enlarged. This means that the image will contain more pixels than the original. The pixels will therefore be artificially enlarged, resulting in a loss of resolution, even though no information is lost. Therefore, to choose these dimensions appropriately, we start with the dimensions of the initial distorted image. The largest width, either AB or CD in Fig. 4.26, is chosen as the width of the transformed image. Then, the height of the image is determined using the aspect ratio of the DMD, which is 4:3. Consequently, the transformed image has the right aspect ratio and has undergone minimal information loss with the scaling. The result of the perspective correction on an undamaged image of the scene taken by the imager is shown in Fig. 4.27. It can be observed that the image has been successfully rectified. It no longer exhibits perspective distortion, and the black margins around the scene are no longer present. Additionally, the detection of the corners and the framing of the DMD during the transformation were precise, as the scene was not cropped or cut by the perspective correction. For this image, the new dimensions are 1180×885.



(a) **(b)**

Figure 4.27: Comparison of the scene image (a) before, and (b) after perspective correction. The observed scene is the undamaged aerial view of the CSL, imaged with all the mirrors of the DMD activated. The image after this correction is straightened and rectangular.

4.7 Results

This last section summarizes the results obtained in this chapter on the imager calibration phase. The image shown in Fig. 4.28 represents the calibrated damaged image of the scene, ready for reconstruction. This image has undergone dark frame subtraction and flat frame correction. The scene represents the aerial view of the CSL and the image has been damaged with the random pattern shown in Fig. 3.19a and having a filling ratio of 50% on the DMD. It should be noted that this calibrated image has been masked by the pattern mask, which now has a filling ratio of only 26.9% due to the morphological erosion operation. This pattern mask is shown in full in Fig. 4.29. As can be seen, the pixels corresponding to the black margins in the damaged image are indicated as white pixels in the pattern mask. Indeed, since these margins are not part of the scene image, the algorithm should not attempt to reconstruct them. These two images in Fig. 4.28 and Fig. 4.29 thus correspond to the two images presented in yellow in the calibration flowchart in Fig. 4.1.

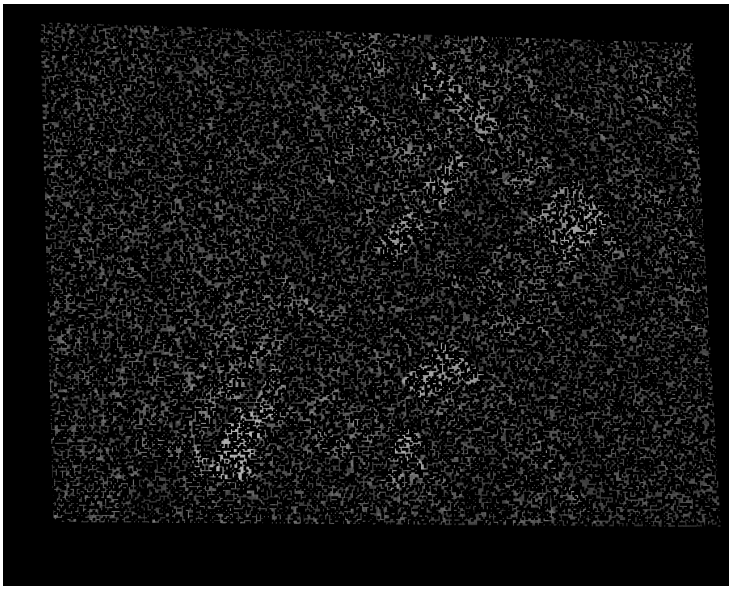


Figure 4.28: Processed damaged image of the scene.

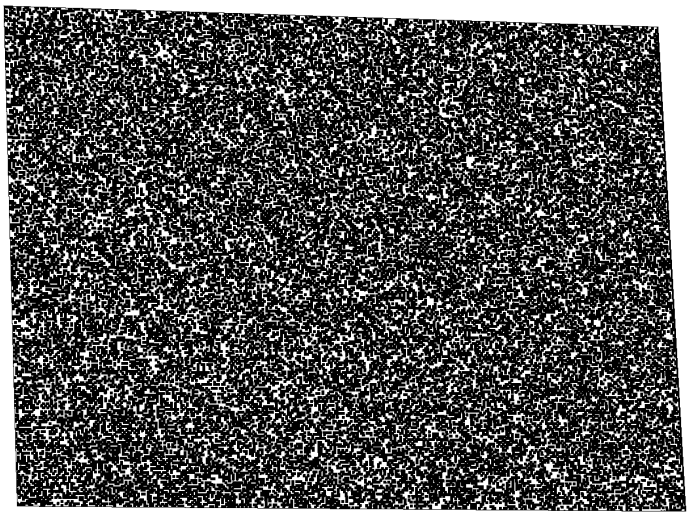


Figure 4.29: Processed pattern mask of the random pattern.

The present work focused only on the optical part of an imager performing inpainting, as well as the calibration of the images obtained to prepare them for reconstruction. As such, no discussion of image reconstruction is included. Nevertheless, it is worth showing the result of the reconstruction of the image of Fig. 4.28 to illustrate the potential of compressive sensing and its applications. This reconstruction is shown in Fig. 4.30 and was carried out by C. Thomas using an IHT algorithm. It can be seen that all the areas damaged by the pattern have been filled in and reconstructed by the algorithm. Most of the objects and details present in the scene are recognizable, despite a number of artifacts in the image. It is important to note that most of the difficulty in compressive sensing lies in the quality of the algorithms and their optimization, which is not the subject of this master's thesis. The reconstruction of the image in Fig. 4.30 has therefore not been pushed very far and has been carried out with a basic algorithm without much optimization. Furthermore, the deconvolution of the image by the PSF is not included in the reconstruction. This image reconstruction is presented in this manuscript as a glimpse of the CS's capabilities, but is not the subject of this work.

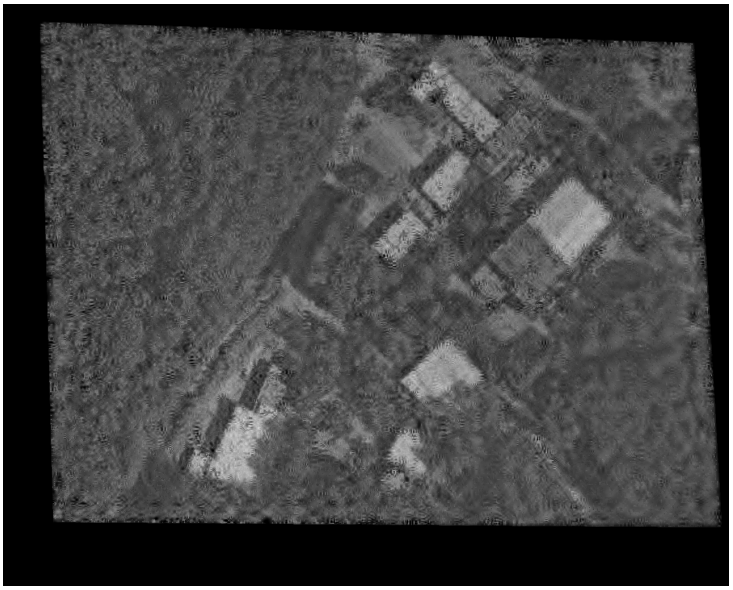
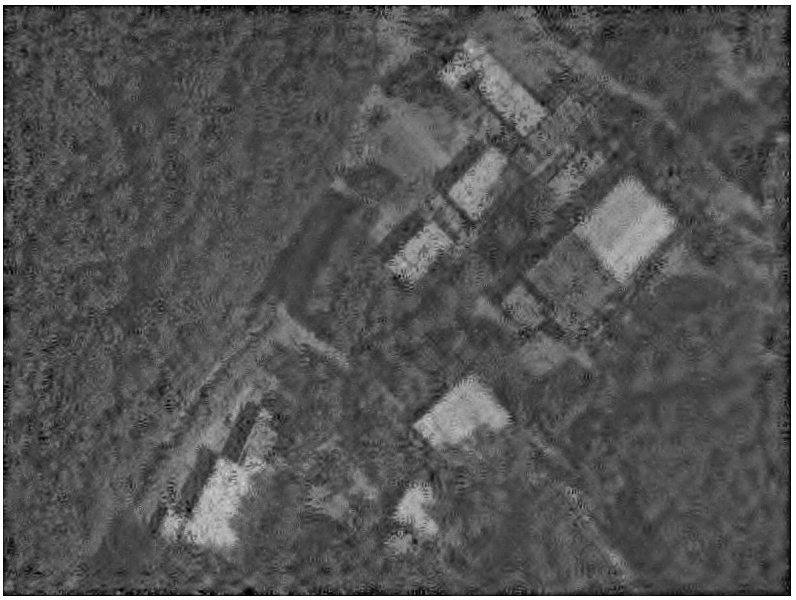


Figure 4.30: Example of the reconstruction of the damaged image shown in Fig. 4.28 performed by C. Thomas.

To follow the calibration flowchart of Fig. 4.1, the perspective correction can be applied to the reconstructed image. The image after this correction is shown in Fig. 4.31, placed next the undamaged image of the same scene. It can be seen that indeed the reconstruction allowed to retrieve the full image with most of the details of the scene being present. The fairly quality of the reconstruction can be partially due to the choice of the damaging pattern. Indeed, the pattern being random, the damaged pixels are spread out over the entire scene, which always leaves undamaged pixels close to these damaged pixels. Therefore, there are no large regions where no information concerning the scene is present. This recovered image concludes the calibration and reconstruction flowchart of the imager.



(a)

(b)

Figure 4.31: Comparison between the reconstructed image of the scene with perspective correction (top) and the undamaged image taken by the imager, also with its corrected perspective (bottom).

5 Conclusion

The objective of this master's thesis was to study the optical part of a compressive sensing imager and to build this optical system in the laboratory. This imager must be capable of masking images using a spatial light modulator so that these damaged images can be reconstructed afterwards through inpainting using compressive sensing algorithms. This work follows the work done by C. Thomas at CSL, who implemented a compressive sensing chain and investigated the applications of this technology for Earth observation. The first chapter of this work focused on the state of the art in compressive sensing, in terms of reconstruction algorithms, and the different architectures of CS imagers.

Afterwards, the architecture of the imager was decided. It uses a DMD to apply the desired masks to the image and has an FPA detector to record the image. The characteristics of the scene were studied, as well as all the components constituting the instrument. The instrument is divided into a primary and a secondary arm. The primary arm includes a collecting lens that focuses the light onto the DMD. The light is then reflected to the secondary arm, which consists of the condensing lens and the detector. The detector is tilted relative to the condensing lens to eliminate image blur caused by the Scheimpflug principle. The imager is controlled by a computer thanks to an interface programmed with the LabVIEW software.

The final chapter focuses on the calibration and post-processing of the images obtained by the instrument. Although the instrument was designed to avoid image defects, some remain and must be corrected before reconstruction. This correction includes a dark frame subtraction and a flat frame correction. The PSF of the instrument is also measured by activating only one micromirror of the DMD. This PSF is used for deconvolving the image during reconstruction. Additionally, a dithering method for improving the quality of this PSF has been developed and implemented. A cropping and perspective correction method has also been developed to obtain a rectified and rectangular image of the scene after reconstruction. Furthermore, a pattern mask was created based on the uniform illumination of the pattern on the DMD. This pattern mask indicates to the reconstruction algorithm which pixels of the image are to be reconstructed or not. A morphological erosion was applied to this mask to completely eliminate artifacts present around the pattern that could degrade the image reconstruction.

This work has explored the practical implementation of a compressive sensing imager in a laboratory. Concrete solutions were provided to achieve the best possible image quality with the available setup and resources in the laboratory. Additionally, the calibration phase allowed the improvement of the image quality and helped to better understand the effects of image masking by the DMD. Subsequently, a method for measuring the instrument's PSF was tested.

In light of this work, several improvements and future perspectives naturally emerge. First, the control of the DMD could be enhanced with suitable electronics to achieve higher pattern change speeds. This would enable compressed image acquisition. Sec-

ondly, another perspective would be to implement the acquisition of images at multiple spectral bands, for example, using a color camera. It would also be interesting to explore the architecture of multispectral imagers such as CASSI. Additionally, investigating the behavior of the DMD under the harsh temperature and pressure conditions of space would be valuable to assess the feasibility of its implementation for Earth observation. Alternatively, exploring other types of SLMs besides the DMD would be interesting. This non-exhaustive list showcases the vast range of possibilities and research directions in the implementation of compressive sensing for image acquisition.

References

[1] Matthew P. Edgar, Graham M. Gibson, and Miles J. Padgett. "Principles and prospects for single-pixel imaging". In: *Nature Photonics* 13.1 (Dec. 2018), pp. 13–20. ISSN: 1749-4893. DOI: 10.1038/s41566-018-0300-7. URL: http://dx.doi.org/10.1038/s41566-018-0300-7.

[2] Velat Kilic, Trac D. Tran, and Mark A. Foster. "Compressed sensing in photonics: tutorial". In: *J. Opt. Soc. Am. B* 40.1 (Jan. 2023), pp. 28–52. DOI: 10.1364/JOSAB.469865. URL: https://opg.optica.org/josab/abstract.cfm?URI=josab-40-1-28.

[3] Simon Foucart and Holger Rauhut. *A Mathematical Introduction to Compressive Sensing*. Springer New York, 2013. ISBN: 9780817649487. DOI: 10.1007/978-0-8176-4948-7. URL: http://dx.doi.org/10.1007/978-0-8176-4948-7.

[4] Rebecca M. Willett, Roummel F. Marcia, and Jonathan M. Nichols. "Compressed sensing for practical optical imaging systems: a tutorial". In: *Optical Engineering* 50.7 (2011), p. 072601. DOI: 10.1117/1.3596602. URL: https://doi.org/10.1117/1.3596602.

[5] Lixiang Li et al. "Overview of Compressed Sensing: Sensing Model, Reconstruction Algorithm, and Its Applications". In: *Applied Sciences* 10.17 (2020). ISSN: 2076-3417. DOI: 10.3390/app10175909. URL: https://www.mdpi.com/2076-3417/10/17/5909.

[6] Laurent Jacques and Christophe De Vleeschouwer. *Lecture notes in image processing and computer vision*. Ecole polytechnique de Louvain. 2024.

[7] D.L. Donoho. "Compressed sensing". In: *IEEE Transactions on Information Theory* 52.4 (2006), pp. 1289–1306. DOI: 10.1109/TIT.2006.871582.

[8] Richard G. Baraniuk. "Compressive Sensing [Lecture Notes]". In: *IEEE Signal Processing Magazine* 24.4 (2007), pp. 118–121. DOI: 10.1109/MSP.2007.4286571.

[9] Thu L. N. Nguyen and Yoan Shin. "Deterministic Sensing Matrices in Compressive Sensing: A Survey". In: *The Scientific World Journal* 2013 (Nov. 2013), p. 192795. ISSN: 2356-6140. DOI: 10.1155/2013/192795. URL: https://doi.org/10.1155/2013/192795.

[10] Zibang Zhang et al. "Hadamard single-pixel imaging versus Fourier single-pixel imaging". In: *Opt. Express* 25.16 (Aug. 2017), pp. 19619–19639. DOI: 10.1364/OE.25.019619. URL: https://opg.optica.org/oe/abstract.cfm?URI=oe-25-16-19619.

[11] Alina L. Machidon and Veljko Pejović. "Deep learning for compressive sensing: a ubiquitous systems perspective". In: *Artificial Intelligence Review* 56.4 (Apr. 2023), pp. 3619–3658. ISSN: 1573-7462. DOI: 10.1007/s10462-022-10259-5. URL: https://doi.org/10.1007/s10462-022-10259-5.

[12] David L. Donoho and Yaakov Tsaig. "Fast Solution of ℓ_1 -Norm Minimization Problems When the Solution May Be Sparse". In: *IEEE Transactions on Information Theory* 54 (2008), pp. 4789–4812. URL: https://api.semanticscholar.org/CorpusID:6961958.

[13] Stephen Becker, Jérôme Bobin, and Emmanuel J. Candès. "NESTA: A Fast and Accurate First-Order Method for Sparse Recovery". In: *SIAM Journal on Imaging Sciences* 4.1 (Jan. 2011), pp. 1–39. ISSN: 1936-4954. DOI: 10.1137/090756855. URL: http://dx.doi.org/10.1137/090756855.

[14] Justin Romberg. "Imaging via Compressive Sampling". In: *IEEE Signal Processing Magazine* 25.2 (Mar. 2008), pp. 14–20. ISSN: 1558-0792. DOI: 10.1109/MSP.2007.914729.

[15] Emmanuel Candes, Justin Romberg, and Terence Tao. *Stable Signal Recovery from Incomplete and Inaccurate Measurements*. 2005. arXiv: math/0503066 [math.NA].

[16] Wikipedia. *Greedy algorithm — Wikipedia, The Free Encyclopedia*. http://en.wikipedia.org/w/index.php?title=Greedy%20algorithm&oldid=1212155202. [Online; accessed 15-March-2024]. 2024.

[17] Joel A. Tropp and Anna C. Gilbert. "Signal Recovery From Random Measurements Via Orthogonal Matching Pursuit". In: *IEEE Transactions on Information Theory* 53.12 (Dec. 2007), pp. 4655–4666. ISSN: 1557-9654. DOI: 10.1109/TIT.2007.909108.

[18] Thomas Blumensath and Mike E. Davies. "Iterative Thresholding for Sparse Approximations". In: *Journal of Fourier Analysis and Applications* 14.5 (Dec. 2008), pp. 629–654. ISSN: 1531-5851. DOI: 10.1007/s00041-008-9035-z. URL: https://doi.org/10.1007/s00041-008-9035-z.

[19] Amir Beck and Marc Teboulle. "A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems". In: *SIAM Journal on Imaging Sciences* 2.1 (2009), pp. 183–202. DOI: 10.1137/080716542. eprint: https://doi.org/10.1137/080716542. URL: https://doi.org/10.1137/080716542.

[20] J. Zhang and B. Ghanem. "ISTA-Net: Interpretable Optimization-Inspired Deep Network for Image Compressive Sensing". In: *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Los Alamitos, CA, USA: IEEE Computer Society, June 2018, pp. 1828–1837. DOI: 10.1109/CVPR.2018.00196. URL: https://doi.ieeecomputersociety.org/10.1109/CVPR.2018.00196.

[21] V. Pavitra, V. B. S. Srilatha Indira Dutt, and G. V. S. Raj Kumar. "Deep Learning Based Compressive Sensing for Image Reconstruction and Inference". In: *2022 IEEE 7th International conference for Convergence in Technology (I2CT)*. 2022, pp. 1–7. DOI: 10.1109/I2CT54291.2022.9824890.

[22] P Hanumanth, P Bhavana, and Shreyanka Subbarayappa. "Application of deep learning and compressed sensing for reconstruction of images". In: *Journal of Physics: Conference Series* 1706.1 (Dec. 2020), p. 012068. DOI: 10.1088/1742-6596/1706/1/012068. URL: https://dx.doi.org/10.1088/1742-6596/1706/1/012068.

[23] Roummel F. Marcia and Rebecca M. Willett. "Compressive coded aperture superresolution image reconstruction". In: *2008 IEEE International Conference on Acoustics, Speech and Signal Processing* (2008), pp. 833–836. URL: https://api.semanticscholar.org/CorpusID:15207907.

[24] R. H. Dicke. "Scatter-Hole Cameras for X-Rays and Gamma Rays". In: *The Astrophysical Journal* 153 (Aug. 1968), p. L101. DOI: 10.1086/180230.

[25] E. E. Fenimore and T. M. Cannon. "Coded aperture imaging with uniformly redundant arrays". In: *Appl. Opt.* 17.3 (Feb. 1978), pp. 337–347. DOI: 10.1364/AO.17.000337. URL: https://opg.optica.org/ao/abstract.cfm?URI=ao-17-3-337.

[26] Stephen Gottesman and E Fenimore. "New family of binary arrays for coded aperture imaging". In: *Applied optics* 28 (Oct. 1989), pp. 4344–52. DOI: 10.1364/AO.28.004344.

[27] Texas Instruments. *Introduction to ±12 Degree Orthogonal Digital Micromirror Devices (DMDs)*. 2008. URL: https://www.ti.com/lit/pdf/dlpa008.

[28] Joseph P. Rice et al. "DMD diffraction measurements to support design of projectors for test and evaluation of multispectral and hyperspectral imaging sensors". In: *Emerging Digital Micromirror Device Based Systems and Applications*. Ed. by Larry J. Hornbeck and Michael R. Douglass. Vol. 7210. International Society for Optics and Photonics. SPIE, 2009, p. 72100D. DOI: 10.1117/12.808990. URL: https://doi.org/10.1117/12.808990.

[29] Michael Douglass. "DMD reliability: a MEMS success story". In: *Reliability, Testing, and Characterization of MEMS/MOEMS II*. Ed. by Rajeshuni Ramesham and Danelle M. Tanner. Vol. 4980. International Society for Optics and Photonics. SPIE, 2003, pp. 1–11. DOI: 10.1117/12.478212. URL: https://doi.org/10.1117/12.478212.

[30] R. Paschotta. *Liquid Crystal Modulators*. RP Photonics Encyclopedia. [Online; accessed 05-May-2024]. DOI: 10.61835/tve. URL: https://www.rp-photonics.com/liquid_crystal_modulators.html.

[31] Zichen Zhang, Zheng You, and Daping Chu. "Fundamentals of phase-only liquid crystal on silicon (LCOS) devices". In: *Light: Science & Applications* 3.10 (Oct. 2014), e213–e213. ISSN: 2047-7538. DOI: 10.1038/lsa.2014.94. URL: https://doi.org/10.1038/lsa.2014.94.

[32] Marco F. Duarte et al. "Single-pixel imaging via compressive sampling". In: *IEEE Signal Processing Magazine* 25.2 (Mar. 2008), pp. 83–91. ISSN: 1558-0792. DOI: 10.1109/MSP.2007.914730.

[33] Michael B. Wakin et al. "An Architecture for Compressive Imaging". In: *2006 International Conference on Image Processing*. Oct. 2006, pp. 1273–1276. DOI: 10.1109/ICIP.2006.312577.

[34] Graham M. Gibson, Steven D. Johnson, and Miles J. Padgett. "Single-pixel imaging 12 years on: a review". In: *Opt. Express* 28.19 (Sept. 2020), pp. 28190–28208. DOI: 10.1364/OE.403195. URL: https://opg.optica.org/oe/abstract.cfm?URI=oe-28-19-28190.

[35] Ming-Jie Sun and Jia-Min Zhang. "Single-Pixel Imaging and Its Application in Three-Dimensional Reconstruction: A Brief Review". In: *Sensors* 19.3 (2019). ISSN: 1424-8220. DOI: 10.3390/s19030732. URL: https://www.mdpi.com/1424-8220/19/3/732.

[36] Valentina Raimondi et al. "A feasability study for a compressive sensing imager in the medium infrared for hotspot detection". In: *International Conference on Space Optics — ICSO 2020*. Ed. by Bruno Cugny, Zoran Sodnik, and Nikos Karafolas. Vol. 11852. International Society for Optics and Photonics. SPIE, 2021, p. 1185259. DOI: 10.1117/12.2599938. URL: https://doi.org/10.1117/12.2599938.

[37] Gonzalo R. Arce et al. "Compressive Coded Aperture Spectral Imaging: An Introduction". In: *IEEE Signal Processing Magazine* 31.1 (Jan. 2014), pp. 105–115. ISSN: 1558-0792. DOI: 10.1109/MSP.2013.2278763.

[38] Henry Arguello, Claudia V. Correa, and Gonzalo R. Arce. "Fast lapped block reconstructions in compressive spectral imaging". In: *Appl. Opt.* 52.10 (Apr. 2013), pp. D32–D45. DOI: 10.1364/AO.52.000D32. URL: https://opg.optica.org/ao/abstract.cfm?URI=ao-52-10-D32.

[39] Ashwin Wagadarikar et al. "Single disperser design for coded aperture snapshot spectral imaging". In: *Appl. Opt.* 47.10 (Apr. 2008), B44–B51. DOI: 10.1364/AO.47.000B44. URL: https://opg.optica.org/ao/abstract.cfm?URI=ao-47-10-B44.

[40] Vivek Boominathan et al. "Recent advances in lensless imaging". In: *Optica* 9.1 (Jan. 2022), pp. 1–16. DOI: 10.1364/OPTICA.431361. URL: https://opg.optica.org/optica/abstract.cfm?URI=optica-9-1-1.

[41] Nick Antipa et al. "DiffuserCam: lensless single-exposure 3D imaging". In: *Optica* 5.1 (Jan. 2018), pp. 1–9. DOI: 10.1364/OPTICA.5.000001. URL: https://opg.optica.org/optica/abstract.cfm?URI=optica-5-1-1.

[42] Sabrina Gramegna. "Study of compressive sensing in view of space imaging applications". Available at http://hdl.handle.net/2268.2/14390. Unpublished master's thesis. University of Liège, June 2022.

[43] Clément Thomas. "Exploring Compressive Sensing for Earth Observation". Available at http://hdl.handle.net/2268.2/17460. Unpublished master's thesis. University of Liège, June 2023.

[44] Texas Instruments. *DLP7000 DLP® 0.7 XGA 2x LVDS Type A DMD*. Version G. 2012. URL: https://www.ti.com/product/DLP7000.

[45] Wikipedia. *Solid angle — Wikipedia, The Free Encyclopedia*. http://en.wikipedia.org/w/index.php?title=Solid%20angle&oldid=1222673862. [Online; accessed 17-May-2024]. 2024.

[46] R. Paschotta. *Field of View*. RP Photonics Encyclopedia. [Online; accessed 05-May-2024]. DOI: 10.61835/459. URL: https://www.rp-photonics.com/field_of_view.html.

[47] Wikipedia. *Objectif macro — Wikipedia, The Free Encyclopedia*. http://fr.wikipedia.org/w/index.php?title=Objectif%20macro&oldid=209997850. [Online; accessed 19-May-2024]. 2024.

[48] Phillip Reeve. *Olympus OM Macro 50mm 1:3.5: Small and affordable Macro with good performance - phillipreeve.net — phillipreeve.net*. https://phillipreeve.net/blog/review-olympus-om-zuiko-auto-macro-50mm-13-5/. [Online; Accessed 19-05-2024].

[49] *Principal planes*. http://hyperphysics.phy-astr.gsu.edu/hbase/hframe.html. [Online; Accessed 19-05-2024].

[50] IDS. *UI-3240CP Rev. 2*. URL: https://fr.ids-imaging.com/store/ui-3240cp-rev-2.html.

[51] Wikipedia. *Depth of field — Wikipedia, The Free Encyclopedia*. http://en.wikipedia.org/w/index.php?title=Depth%20of%20field&oldid=1207712571. [Online; accessed 21-May-2024]. 2024.

[52] Barbara London, Jim Stone, and John Upton. *Photography*. 12th ed. Upper Saddle River, NJ: Pearson, June 2016. ISBN: 978-0-13-448202-6.

[53] E. Nocerino et al. "Experiments on calibrating tilt-shift lenses for close-range photogrammetry". In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLI-B5 (2016), pp. 99–105. DOI: 10.5194/isprs-archives-XLI-B5-99-2016. URL: https://isprs-archives.copernicus.org/articles/XLI-B5/99/2016/.

[54] Texas Instruments. *DLP® Discovery™ 4100 Development Platform*. 2012. URL: https://www.ti.com/document-viewer/lit/html/dlpu040.

[55] R. M. Joany and E. Logashanmugam. "Evaluation of efficient method to control the functionality of DMD". In: *2015 International Conference on Communications and Signal Processing (ICCSP)*. 2015, pp. 0015–0019. DOI: 10.1109/ICCSP.2015.7322857.

[56] Xiaorui Yin. *High Frame-Rate and Low-Latency Control of Digital Micromirror Devices (DMD) using FPGA for Ultracold Atom-based Quantum Experiments*. en. Student Paper. Zurich, 2021. DOI: 10.3929/ethz-b-000510370.

[57] Jérome Loicq. *Lecture notes in Space Optics*. Faculté des sciences appliquées de l'université de Liège. 2023.

[58] Wikipedia. *Dark current (physics) — Wikipedia, The Free Encyclopedia*. http://en.wikipedia.org/w/index.php?title=Dark%20current%20(physics)&oldid=1193266484. [Online; accessed 24-May-2024]. 2024.

[59] *Bad Pixels! - Mastcam-Z — mastcamz.asu.edu*. https://mastcamz.asu.edu/bad-pixels/. [Accessed 25-05-2024].

[60] Nasim Mansurov. *What is Vignetting? — photographylife.com*. https://photographylife.com/what-is-vignetting. [Accessed 26-05-2024].

[61] Nobuyuki Otsu. "A Threshold Selection Method from Gray-Level Histograms". In: *IEEE Transactions on Systems, Man, and Cybernetics* 9.1 (1979), pp. 62–66. DOI: 10.1109/TSMC.1979.4310076.

[62] *What is Dithering in Astrophotography? — astrobackyard.com*. https://astrobackyard.com/dithering-astrophotography/. [Accessed 30-05-2024].

[63] L. Clermont et al. "Automatized alignment of the focal plane assemblies on the PLATO cameras". In: *Space Telescopes and Instrumentation 2018: Optical, Infrared, and Millimeter Wave*. Ed. by Makenzie Lystrup et al. Vol. 10698. International Society for Optics and Photonics. SPIE, 2018, p. 106986L. DOI: 10.1117/12.2314167. URL: https://doi.org/10.1117/12.2314167.

[64] *cv2.getPerspectiveTransform() – TheAILearner — theailearner.com*. https://theailearner.com/tag/cv2-getperspectivetransform/. [Accessed 03-06-2024].