



**HAL**  
open science

## VisAdapt: Catadioptric Adaptive Camera for scenes of variable density of visual information

Julien Jean Ducrocq, Guillaume Caron, Hervé Midavaine, Christian Duriez, Jérémie Dequidt, El Mustapha Mouaddib

► **To cite this version:**

Julien Jean Ducrocq, Guillaume Caron, Hervé Midavaine, Christian Duriez, Jérémie Dequidt, et al.. VisAdapt: Catadioptric Adaptive Camera for scenes of variable density of visual information. Applied optics, 2023, 62 (35), pp.9207-9214. 10.1364/AO.500663. hal-04276452

**HAL Id: hal-04276452**

**<https://hal.science/hal-04276452v1>**

Submitted on 14 Dec 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# VisAdapt: Catadioptric Adaptive Camera for scenes of variable density of visual information

JULIEN DUCROQC<sup>1,2,\*</sup>, GUILLAUME CARON<sup>1,3</sup>, HERVÉ MIDAVAIN<sup>1</sup>, CHRISTIAN DURIEZ<sup>4</sup>, JÉRÉMIE DEQUIDT<sup>4</sup>, AND EL MUSTAPHA MOUADDIB<sup>1</sup>

<sup>1</sup> Université de Picardie Jules Verne (UPJV), MIS laboratory, 33 rue saint Leu, 80039 Amiens, France

<sup>2</sup> Nara Institute of Sciences and Technology (NAIST), CARE (Cybernetics and Reality Engineering) laboratory, B315 8916-5 Takayama-cho, Ikoma, Nara 630-0192, Japan

<sup>3</sup> CNRS-AIST JRL (Joint Robotics Laboratory), IRL, 1-1-1 Umezono, Tsukuba, Ibaraki, 305-8560, Japan

<sup>4</sup> Université de Lille, Inria, CNRS, Centrale Lille, UMR 9189 CRISTAL, F-59000 Lille, France

\*julien.ducroqc.jz9@naist.ac.jp

Compiled December 14, 2023

This paper presents the design method of a multi-resolution camera, named *Visadapt*. It is made of a conventional compact camera with a sensor and a lens pointed to a new deformable mirror so that the mirror at flat state is parallel to the image plane. The main novelty of the latter mirror is the ability to control automatically strokes of several millimeters. This allows *Visadapt* to capture scenes with a spatially variable density of visual information. A grid of linear actuators, set underneath the mirror surface, deforms the mirror to reach the desired shape computed to capture several areas of different resolutions. Mechanical simulations allowed to iterate on *Visadapt*'s design, to reduce the geometrical distortions in the images. Evaluations made on an actual prototype of *Visadapt* show that, by adapting the mirror shape, this camera can magnify a scene object up to 20%, even off-centered in the field-of-view, while still perceiving the rest of the scene.

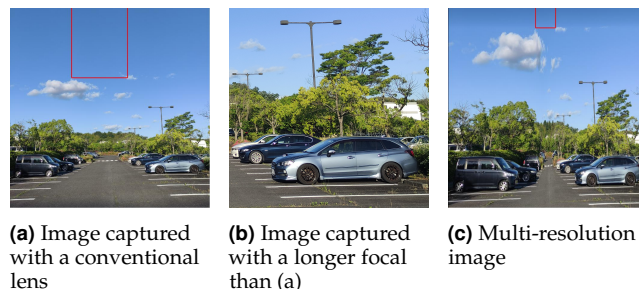
<http://dx.doi.org/10.1364/ao.500663>

## 1. INTRODUCTION

Everyday life environments are characterized by variety: they feature uniform regions, several objects and/or complex textures. Since they carry more visual information, objects and textures can be considered as key elements of the scene. Within such heterogeneous environments, the density of visual information is spatially variable. Therefore, when they are captured by a conventional camera, uniform scene regions may use the majority of the pixels and key elements of the scene only a few pixels (Fig. 1a). Thanks to the optical zoom, classical cameras increase the resolution of a specific region of the scene by increasing their focal length [1]. Nonetheless, the objects which are not targeted by the zoom are not visible anymore (Fig. 1b versus Fig. 1a).

Several applications are impacted by heterogeneous environments. In robotics, visual servoing methods [2, 3] allow mobile robots to navigate only with visual information. If the scene

is made of wide uniform regions, the robot trajectory may fall under local minima tens of centimeters away from the desired location due to the lack of features [4, 5]. Moreover, in video-surveillance, faces or silhouettes may not occupy enough pixels to allow identification. A pan-tilt-zoom (PTZ) camera may zoom-in a person to capture enough information for identification [6, 7], but it would not be able to monitor the rest of the scene anymore.



**Fig. 1.** Different images of an heterogeneous environment. The area framed in red occupies less pixels in (c) than in (a), while the rest is zoomed.

Thus, in these contexts, a camera should capture images with a locally variable angular resolution, to make key scene elements occupy more pixels than uniform regions that could shrink to a few pixels. For this purpose, such a camera has to capture  $P \in \mathbb{N}$  scene regions, each with a desired resolution, leading to a multi-resolution image (Fig. 1c). Furthermore, instead of tailored for a specific static scene, the local resolution of this camera needs to adapt to different scene geometries for versatile uses. Conceptually, such camera could be implemented with either a large pixel density followed by hardware local binning, spatially variable image sensor with photodiodes of adaptive sizes, or a deformable optical surface. We focus on the latter because the first is bounded by the noise of tiny photodiodes [8], no local hardware binning technology exists, and the second is not available yet. Deformable lenses *e.g.* electrowetting lenses [9], only change the magnification globally [10] or defocus blur,

locally [11]. Therefore, we propose to design and actuate an optical surface centimeters away the camera chip: a deformable mirror, facing a conventional camera.

The new catadioptric camera is named Visadapt. Its mirror is made of a thin reflective surface above a matrix of linear actuators, and its shape is adaptively optimized to capture a desired spatially variable resolution. The design of Visadapt has been iteratively improved thanks to mechanical simulations to reach the smooth curvatures of the ideal deformed optical surface, hence limiting the geometrical discontinuities of the captured images. Simulation results, completed by experiments on a proof-of-concept prototype, show that Visadapt is able to capture multi-resolution images with up to nine contiguous scene regions of different resolutions.

In the rest of this article, Section 2 relates the cameras who inspired Visadapt's design and existing methods to optimize a mirror shape. Then, Section 3 describes the design method of Visadapt involving mechanical simulations of deformable materials and optical geometry. After that, Section 4 reports and evaluates results with simulated scenes and Section 5 validates the practical design by images captured with a deformable mirror prototype, before Conclusion (Sec. 6).

## 2. RELATED WORKS

**Adaptive cameras** The adaptive cameras we consider are those able to change optical properties across time with hardware adaptation at the image capture time, contrary to cameras such as plenoptic cameras [12] that capture the light-field for software adaptation as post-processing. Among them, several are related to Visadapt's purpose. Adaptive cameras with a grid of deformable lenses can change their FOV [13] or magnification [14] globally. Thanks to several layers of pinhole grids, a lensless camera [15] is able to capture up to nine scene regions, each with a different magnification, but all the areas in between are lost. Moreover, a conventional camera facing a bendable mirror, can capture images with a non-uniform resolution [16]. The deformations of the latter consist in manual global bending of a few centimeters. Conversely, the deformable mirrors (DMs) used in the context of adaptive optics [17–19] are automatically controllable and capable of multiple local deformations. But since they modulate the optical phase of light beams in the entire FOV to increase locally the sharpness of images, they deform at high frequency of up to tens of micrometers. Thus, the image geometry is not changed by adaptive optics' DMs.

**Mirror shape optimization** Several methods exist to design a static mirror shape to capture scene regions with a desired resolution, *e.g.* with vector fields [20] or polynomial functions [21]. [22] computes b-splines-based mirror surfaces, which reflect light rays from scene points to desired pixel locations on the camera image sensor, matching a *scene-to-image* mapping. On the one hand, optimized mirrors are made once, *e.g.* to be put in cars for the driver gaze check while observing the surrounding road [20], or to be put at the center of crossroads for traffic monitoring [21]. On the other hand, existing DM-based adaptive cameras are aimed to capture astrophotographs [18] or retinal images [23], requiring a  $\mu\text{m}$ -scale stroke at high frequency.

## 3. THE DESIGN OF VISADAPT

Visadapt combines a DM of centimeters stroke (as the manually flexible mirror [16]) with an automatic actuation, to locally modify the image geometry. It is the first controllable camera able to

magnify several scene regions at once within a multi-resolution image, without omitting any object within the camera FOV (unlike *e.g.* lensless cameras with several pinhole grid layers [15]). When the scene changes, a new mirror surface can be computed with b-splines optimization [22], then the actuators will deform Visadapt's mirror to reach the optimal surface, *e.g.* to better emphasize the objects featuring more visual information.

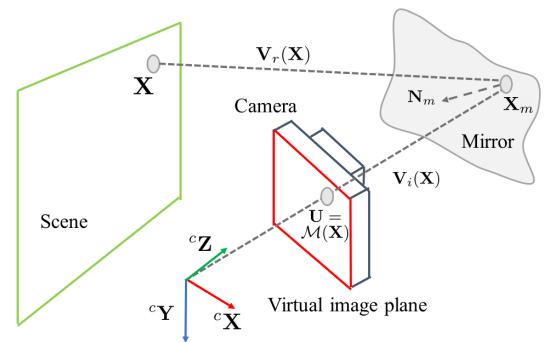
### A. The scene-to-image mapping model

Visadapt aims at capturing a scene with a desired spatially variable resolution, hence satisfying a *scene-to-image* mapping  $\mathcal{M}$  (which projects a scene point  $\mathbf{X} = (X, Y, Z)^T \in \mathbb{R}^3$  to an image pixel  $\mathbf{U} = (U, V) \in \mathbb{R}^2$ ). Fig. 2 illustrates the path of a light ray from  $\mathbf{X}$  to  $\mathbf{U}$ . Incident light rays  $\mathbf{V}_i \in \mathbb{R}^3$  from scene points  $\mathbf{X}$  intercept the mirror surface  $\mathcal{S}$  at  $\mathbf{X}_m = (X_m, Y_m, Z_m)^T \in \mathbb{R}^3$ , and then are reflected as  $\mathbf{V}_r \in \mathbb{R}^3$ . The coordinates of all these points are expressed within the camera orthonormal frame  $\mathcal{F}_c = \{^c\mathbf{X}, ^c\mathbf{Y}, ^c\mathbf{Z}\}$ .

$\mathbf{V}_r$  is expressed with the local normal  $\mathbf{N}_m = (N_x, N_y, N_z)^T \in \mathbb{R}^3$  to the mirror surface at  $\mathbf{X}_m$ , following the Snell-Descartes law  $\mathbf{V}_r = \mathbf{V}_i - (2\mathbf{N}_m^T \mathbf{V}_i) \mathbf{N}_m$ . If  $\mathcal{S}$  is optimized to satisfy  $\mathcal{M}$ , the real image plane coordinates  $\mathbf{x} = (x, y, z)^T \in \mathbb{R}^3$  of the desired pixel  $\mathbf{U} = (U, V)$  will belong to  $\mathbf{V}_r$ .

To simplify, consider the scene as  $P$  co-planar scene regions made of  $m_p \times n_p \in \mathbb{N}$  points of metrical known positions  $\mathbf{X}$ , delimited by borders  $\mathbf{X}_{p-}, \mathbf{X}_{p+} \in \mathbb{R}^3$ , with  $p \in [0, P] \subset \mathbb{N}$ . The image captured,  $\mathbf{I}$ , is made of  $P$  sub-images  $\mathbf{I}_p$ , each capturing a scene region  $p$  by  $m_p \times n_p$  pixels, of desired location  $\mathbf{U}$  in between borders  $\mathbf{U}_{p-}, \mathbf{U}_{p+} \in \mathbb{R}^2$ .  $\mathbf{I}$  is a multi-resolution image if at least one couple of sub-image borders  $\mathbf{U}_{p-}, \mathbf{U}_{p+}$  does not match the pinhole projection of the corresponding scene borders  $\mathbf{X}_{p-}, \mathbf{X}_{p+}$ . Indeed, it implies that the catadioptric camera can be modeled as  $P$  sub-cameras, with different intrinsic parameters  $\Gamma_p = (k_{u_p}, k_{v_p}, u_{0_p}, v_{0_p}) \in \mathbb{R}^4$  ( $k_{u_p}, k_{v_p}$  are the pixel size and  $u_{0_p}, v_{0_p}$  are the coordinates of the principal point for each subcamera  $p$ ). In order to capture an image of a non-uniform resolution, the mirror of a catadioptric camera must have curvatures [20, 22], so does Visadapt's mirror.

To satisfy  $\mathcal{M}$ ,  $\mathcal{S}$  is optimized iteratively, *e.g.* by minimizing the residual errors between a set of desired normals  $\mathbf{N}_m^* \in \mathbb{R}^3$ , each associated with a couple  $(\mathbf{X}, \mathbf{U})$ , and the local tangent of  $\mathcal{S}$ ,  $\mathbf{T}_m \in \mathbb{R}^3$  [22]. For the first iteration, the mirror is initially assumed planar. Then, all  $\mathbf{N}_m^*$  are recomputed for each iteration, considering the mirror surface obtained at the end of the



**Fig. 2.** The path of a light ray through the scene-to-image mapping (from the scene to the image sensor).

146 previous iteration, generally not planar.

## 147 B. The core idea of the design

148 In order to reach the optimal surface  $\mathcal{S}$ , Visadapt uses a de-  
149 formable mirror made of a reflective membrane with a matrix of  
150  $R \times C \in \mathbb{N}$  linear actuators underneath, regularly spaced, each  
151 one moving along the normal to the mirror in planar state. The  
152 number of actuators  $R \times C$  is chosen as a compromise between  
153 the resolution of actuation and the compactness of the design.  
154 The actuators are considered completely rigid, whereas the mirror  
155 surface is considered flexible enough to be locally deformed  
156 by an underneath push actuation. Visadapt relies on several  
157 parameters, which have been assessed through mechanical sim-  
158 ulations and formally described thereafter.

## 159 C. Design parameters evaluated through simulation

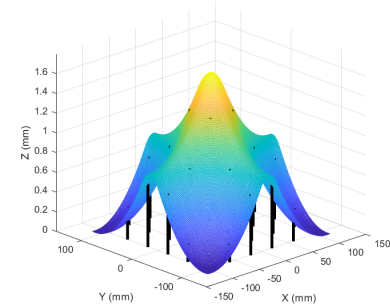
160 Thanks to a realistic mechanical simulation, several parameter-  
161 ized configurations of the mirror can be compared, to assess their  
162 impact on the mirror reflection and chose an optimal design for  
163 the prototype. A first parameter, the actuation surface ( $S_a \in \mathbb{R}$ ,  
164 expressed in  $mm^2$ ) changes the geometry of the contact area  
165 between the actuator and the mirror surface. This parameter has  
166 a strong impact on the local shape of the deformations. Then, in  
167 order to smooth the curvature of the mirror, a layer of a flexible  
168 material (referred here as a *smoothing layer*) is added between  
169 the mirror membrane and the actuators heads. The reflection  
170 will therefore depend on the thickness  $e \in \mathbb{R}_+$  of this layer, a  
171 second critical design parameter.

172 It is worth noting that the mechanical model includes sev-  
173 eral other parameters for the reflective membrane that are its  
174 thickness, its elasticity, the clamping model to a frame and the  
175 resulting tensile stress. However, they are set constant such as,  
176 according to the rectangular aspect of an image, the shape of the  
177 frame is set rectangular and the clamping of the membrane to the  
178 frame is modeled as dense fixed-point constraints along the four  
179 sides of the frame and the resulting tensile stress is neglected to  
180 not modify the mirror membrane elasticity parameter. Indeed,  
181 Section 5 reports that a single mirror membrane product was  
182 found to reach a satisfying visual quality, and describes as well  
183 the practical details that allow to make these above-mentioned  
184 assumptions realistic.

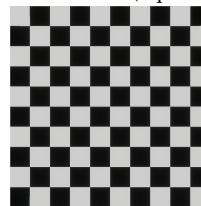
185 A poor choice of the actuation surface  $S_a$  and the smooth-  
186 ing layer thickness  $e$  would lead the images captured with the  
187 Visadapt camera to show many distortions. Hence,  $S_a$  and  $e$   
188 will be chosen minimizing the distortions in the image thanks  
189 to a simulation which scene considers a chessboard made of  
190  $j \times k \in \mathbb{N}^*$  squares, set fronto-parallel to the image, their axes  
191 aligned (Fig. 3b). In the simulation, a mapping  $\mathcal{M}$  is set to re-  
192 duce the resolution of the central scene region, hence increasing  
193 the resolution of every other region. To make the target area  
194 reach the desired resolution, an optimal mirror shape (Fig. 3a) is  
195 computed at first using b-splines surface optimization [22] with  
196 default parameters, that are  $nb_K = 36$  knots, splines of order  
197  $m_{ord} = 5$  and the number of iterations  $n_{iter} = 3$  (thanks to the  
198 *Matlab Reflex toolbox* of [22]). Then, the actuator heights  $h \in \mathbb{R}$   
199 are computed as the intersection of the actuator axes (see the  
200 black bars of Fig. 3a) and the optimal mirror shape (Fig. 3a).  $h$   
201 also considers the smoothing layer thickness and that only push  
202 actuations ( $h > 0$ ) are considered with the Visadapt prototype  
203 (Sec. 5).

204 Afterwards,  $S \in \mathbb{N}$  virtual models of the deformable mirror  
205 (considering different values of  $S_a$  and  $e$ ) are actuated and set in

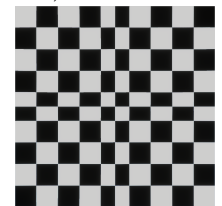
206 front of a virtual camera to the scene reflection thanks to ray trac-  
207 ing within the [Unity® rendering engine](#). Since the command is a  
208 desired image grid, we evaluate the mirror through its reflection  
209 of the scene. We thus quantify the undesired distortions within  
210 the  $S$  images obtained,  $\hat{\mathbf{I}}_s$ ,  $s \in [0, S - 1] \subset \mathbb{N}$ , by leveraging  
211 the TV distortion measure of the ISO 9039 [24]. It is easy to use  
212 here since the chessboard is fronto-parallel to the virtual camera  
213 image plane and the chessboard axes are parallel to those of the  
214 image. However, the TV distortion method measures a percent-  
215 age of global distortion that the curvy side of the chessboard  
216 shows in the image with respect to the straight line it should be.  
217 So, to deal with the local distortions that Visadapt images can  
218 show and to provide a single distortion measure for the whole  
219 chessboard instead of one per chessboard side, we extend the  
220 TV distortion method as follows. First, we do not only consider  
221 the chessboard borders but all  $L = (j - 1) + (k - 1) \in \mathbb{N}^*$  ver-  
222 tical and horizontal lines other than the chessboard borders to  
223 account for local distortions. Then, we measure the distortions  
224 with every pixel of each line  $l \in [0, L - 1]$ , each pixel being  
225 automatically detected thanks to Sobel's edge detection, as the  
226 horizontal standard deviation  $\sigma_{h_l}$  of all pixels corresponding to a  
227 vertical line  $l$ , respectively the vertical standard deviation  $\sigma_{v_l}$   
228 of all pixels corresponding to a horizontal line  $l$ . Finally, to output a  
229 single distortion measure for the whole chessboard, we compute  
230  $\Delta \in \mathbb{R}_+$  as the mean of all  $\sigma_{h_l}$  and  $\sigma_{v_l}$  for all lines  $l$ . With such  
231 extensions to the seminal TV distortion method, the percentage  
232 representation has no more meaning and is thus dropped, but  
233 anyway the lower  $\Delta$  the least the distortions. This quantitative  
234 evaluation is completed by the qualitative assessment of the  
235 image in general (*e.g.* locating visually where the image suffers  
236 more important local distortions).



(a) Mirror surface and actuators under-  
neath (represented as black bars).



(b) The chessboard scene.



(c) The scene reflection through  
the optimal mirror surface.

**Fig. 3.** The optimal mirror with the actuators underneath and the scene tested for the simulation experiments.

## D. Evaluation method

To evaluate the ability of Visadapt to reach mappings  $\mathcal{M}$ , we introduce  $\eta \in \mathbb{R}$  characterizing the ratio between the desired resolution of the area of interest and its reflected image resolution. Writing  $n^* \in \mathbb{N}^+$  the desired number of pixels in the target area (its desired definition), and  $\hat{n} \in \mathbb{N}^*$  the actual number of pixels in the corresponding area in the image reflected by the mirror,  $\eta$  is defined as follows:

$$\eta = \frac{n^*}{\hat{n}}. \quad (1)$$

The closer  $\eta$  is to 1.0, the more faithful is the mirror w.r.t. the mapping  $\mathcal{M}$ .  $\eta$  is considered to evaluate both simulation and experimental results with the final design.

## 4. MECHANICAL SIMULATION TO DETERMINE THE DEFORMABLE MIRROR PARAMETERS

This section presents, first, the simulation environment, then, the materials and the geometrical configuration of Visadapt, before results for various actuation surfaces and smoothing layers. Afterwards, another set of results on a scene other than a chessboard proves the ability of the virtual Visadapt to reach several *scene-to-image* mappings  $\mathcal{M}$ , thanks to the actuators changing the mirror shape.

For the first evaluation, the criterion  $\Delta$  (Sec. D) quantifies the distortions of  $L = 18$  inner lines (see Fig. 4), consisting in 9 horizontal lines (in green) and 9 vertical lines (in red). To detect those lines without outliers, the intensity threshold of the Sobel's edge detection is manually set to  $\epsilon = 0.2$ .

### A. Simulation environment

Actuators are modeled as rigid moving objects. Finite Element Method is used for the deformations of the mirror. Some bilateral constraints [25] allow to model the interactions between the motors and the mirror. Our mechanical simulation is using the Sofa framework, an open-source and realistic [mechanical simulator](#). For a single simulation step, the actuators change of heights, then an Euler implicit solver [26] computes local forces to deform realistically the mechanical model of the mirror [27], until they reach their final height.

At the end of the mechanical simulation, to compute realistic reflections thanks to ray tracing, the surface of the deformable mirror is transferred to a virtual environment, here rendered with the Unity® engine. Within this virtual environment, the camera is set at the origin of the world frame, the scene plane is set at  $Z_{scn} < 0$  away from the camera and the mirror (initially planar) at  $Z_{mir} > 0$ . The camera image plane, the scene plane and the mirror are all parallel such that the flat mirror leads to a perspective image. In the simulations, we set  $Z_{scn} = -150 \text{ mm}$  and  $Z_{mir} = 396 \text{ mm}$ , such that the whole experimental volume will be easy to reproduce for the prototype in the laboratory, with a camera which captures images of  $1000 \times 1000$  pixels, and of intrinsic parameters  $\Gamma_c = (k_u, k_v, u_0, v_0)$ , with  $k_u = k_v = 0.047 \text{ mm}$ , and  $u_0 = v_0 = 499.5 \text{ pixels}$  and of focal length  $f = 8 \text{ mm}$ .

### B. The materials of Visadapt

Available servomotors are made of a rigid material, which cannot be deformed. Within the mechanical simulation, the properties (especially Poisson's ratio  $\nu \in \mathbb{R}_+$  and Young's modulus  $E \in \mathbb{R}_+^*$ ) of all the different materials must be specified, except for the fully rigid actuators.

In order to determine the material of the mirror membrane, we tested several sheets of reflective materials, *i.e.*, a  $1.0 \text{ mm}$  thick sheet of Polyvinyl chloride (PVC), a  $0.20 \text{ mm}$  thick polyethylene terephthalate (PET) adhesive sheet and a  $0.12 \text{ mm}$  thin PET sheet. Each of them is covered by a thin layer of silver. For a local actuation, performed by a single actuator, underneath, only the thin PET sheet is deformed locally, the thicker PET and PVC sheets globally change of curvature, thus cannot be considered for the Visadapt prototype. Since we use PET, to ensure that the mirror membrane is initially planar and of the highest reflection quality, it must be tensed by a rigid frame (otherwise the scene objects are distorted by the reflection, see Fig. 7a versus Fig. 7b).

The PET is characterised by a Young's modulus of  $E = 3.5 \text{ GPa}$  and a Poisson's ratio of  $\nu = 0.38$ . For the smoothing layer, we consider a more deformable material, to enable the mirror of Visadapt to reach diverse shapes, here the [silicone rubber](#), of  $E = 25 \text{ MPa}$  and  $\nu = 0.48$ . In order to enlarge the actuation base, a square pad, made of [polypropylene](#), is added to the head of each actuator, a material defined by  $E = 1.325 \text{ GPa}$  and  $\nu = 0.43$ .

### C. The configuration of the deformable mirror

As a compromise between resolution of actuation, compactness, the amount of electronic interfaces and available products, we set  $R = C = 6$  to reach 36 actuators that have a  $5 \text{ mm} \times 5 \text{ mm}$  square base. They are regularly spaced on a grid of  $230 \text{ mm} \times 230 \text{ mm}$ , of  $46 \text{ mm}$  step size, which leads to  $f = 8 \text{ mm}$  to ensure that the camera (with a CMOS of about  $5.9 \text{ mm} \times 4.8 \text{ mm}$ ) at  $Z = 0$ , perceives all the actuators (set at  $Z_{mir} = 396 \text{ mm}$ ) at once. The rigid frame is set as far as possible from the actuated area to limit the constraints on the actuation while keeping a reasonable size. Within the virtual environment, this frame is modelled as fixed constraints on the four borders of the mirror.

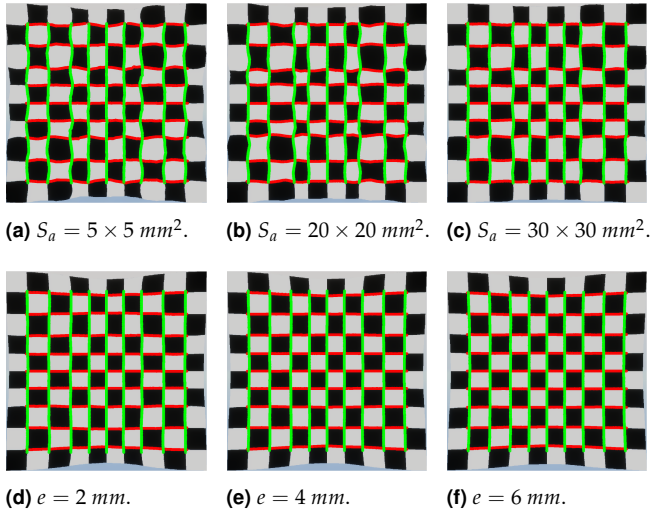
To create a virtual deformable mirror, a mesh made of hexahedra gathers the actuators, the thin mirror surface and, when necessary, the smoothing layer. The actuators are rigidly attached to the smoothing layer (or the mirror surface directly). This mesh is used as a mechanical model which can be deformed within the simulation. To prepare the simulations, each part of the mesh is associated to different mechanical properties, such as Young's modulus and Poisson's ratio, to make a realistic mechanical model, apart from the actuators which are completely rigid (see Sec. B).

### D. Impact of the actuation surface

Within the mechanical model, the impact of the actuation surface  $S_a$  is measured, for three different values  $S_a \in \{5 \times 5 \text{ mm}^2, 20 \times 20 \text{ mm}^2, 30 \times 30 \text{ mm}^2\}$ . When the actuation surface is small, the  $5 \times 5 \text{ mm}^2$  square head of the actuator is directly touching the mirror membrane, inducing undesired local discontinuities severely distorting the captured image (Fig. 4a). As the actuation surface grows, the lines get less distorted, as shown in Fig. 4b-4c. Indeed, the criterion  $\Delta$  is, respectively, of 2.81, 2.27 and 1.58 pixels:  $\Delta$  decreases linearly as  $S_a$  increases. Nonetheless, for any  $S_a$  tested, the borders of the square pads enlarging the actuation surface still induce distortions. These results emphasize the need to add a smoothing layer in between the actuators and the mirror membrane.

### E. Impact of the smoothing layer thickness

The goal of the smoothing layer is to attenuate the discontinuities of the mirror and reduce the geometrical distortions within the



**Fig. 4.** Reflection of the chessboard for 3 different actuation surfaces  $S_a$  and 3 different smoothing layer thicknesses  $e$ .

352 image captured. Thicknesses  $e \in \{2 \text{ mm}, 4 \text{ mm}, 6 \text{ mm}\}$  have  
 353 been successively tested, while the actuation surface remains  
 354  $S_a = 30 \times 30 \text{ mm}^2$  (the less distorting, Sec. D). When considering  
 355 all the 18 chessboard lines detected, the criterion  $\Delta$  changes for  
 356 each  $e$ , reaching successively 1.53, 1.26 and 1.40 pixels. Since the  
 357 borders of the mirror membrane are constrained by the virtual  
 358 frame, the smoothing layer has a different impact in the center  
 359 than in the border of the image. On the one hand, the image  
 360 borders display scene regions outside of the initial camera field  
 361 of view, of which areas increase as the thickness of the smoothing  
 362 layer (see the three reflection images on the second line of Fig. 4).  
 363 On the other hand, for the 10 lines of the  $4 \times 4$  central squares,  
 364 the distortions decrease as the smoothing layer gets thicker,  $\Delta$   
 365 being successively of 0.65, 0.49 and 0.42 pixels. Even though the  
 366 central region is less distorted for the thickest smoothing layer,  
 367 the reflected image suffers less distortions with  $e = 4 \text{ mm}$ .

368 For the simulations of the rest of this section and the design  
 369 of the actual prototype, we consider an actuation surface of  
 370  $S_a = 30 \times 30 \text{ mm}^2$  and a smoothing layer thickness of  $e = 4 \text{ mm}$ .

#### 371 F. Ability to reach diverse scene-to-image mappings

372 In this section, we consider two different mappings  $\mathcal{M}_1$  and  
 373  $\mathcal{M}_2$ . One simulation is done for each mapping, where the abil-  
 374 ity of Visadapt to capture a desired multi-resolution image is  
 375 assessed by the criterion  $\eta$  (Sec. D) computed on the reflection  
 376 of the deformed mirror. The scene used here is a photograph  
 377 of the cathedral of Amiens, an environment with heterogenous  
 378 densities of visual information, since it features many statues of  
 379 different sizes and uniform areas of flat stone.

380  $\mathcal{M}_1$  and  $\mathcal{M}_2$  change the resolution of a target area, res-  
 381 pectively starting at coordinates  $\mathbf{U}_1 = (300, 300)$  and  $\mathbf{U}_2 =$   
 382  $(480, 415)$ , from  $400 \times 400$  pixels to  $300 \times 300$  pixels and from  
 383  $200 \times 200$  pixels to  $150 \times 150$  pixels.  $\mathcal{M}_1$  and  $\mathcal{M}_2$  are illus-  
 384 trated by Fig. 5: Fig. 5a and Fig. 5c represent the conventional images,  
 385 acquired by the camera facing a planar mirror, and Fig. 5b and  
 386 Fig. 5d represent the multi-resolution desired images. Since  $\mathcal{M}_1$   
 387 and  $\mathcal{M}_2$  decrease the resolution of the target area (red areas in  
 388 Fig. 5) by the same horizontal and vertical factor, the resolution  
 389 of the areas at the four corners of the FOV increase too. About  
 390 the four remaining regions (central top and bottom, middle right  
 391 and left) they are transformed differently: on the one hand, the

392 left and right regions are shrunk vertically but enlarged horizon-  
 393 tally; on the other hand, the top and bottom regions are shrunk  
 394 horizontally but enlarged vertically.

395  $\mathcal{M}_1$  projects a centered scene region through a centered im-  
 396 age region, therefore, the corresponding optimal mirror shape  
 397 computed with [22] is symmetrical, with a high convexity at  
 398 its center and small convexities at its borders (Fig. 6a). The  
 399 mapping  $\mathcal{M}_2$  considers a scene area which is not centered, there-  
 400 fore, the mirror shape (Fig. 6c) is not symmetrical anymore, and  
 401 all the points on approximately one fourth of the mirror are  
 402 characterized by  $Z \approx 3.0 \text{ mm}$ .

403 For  $\mathcal{M}_1$ , the mirror reflection is still captured as a multi-  
 404 resolution image (Fig. 6b), but it suffers from distortions on its  
 405 borders (caused by the frame countering the larger actuation on  
 406 the borders). However, the target area is quite faithful to the  
 407 desired one, since  $\eta = 1.1$ . For  $\mathcal{M}_2$ , since several actuators on  
 408 the borders are set to  $3 \text{ mm}$  of height, but are constrained by the  
 409 frame nearby, the borders of the image suffer high distortions.  
 410 However, the target area, close to the center, is less affected, and  
 411 changes of resolution to reach  $174 \times 182$  pixels ( $\eta \approx 1.19$ ).

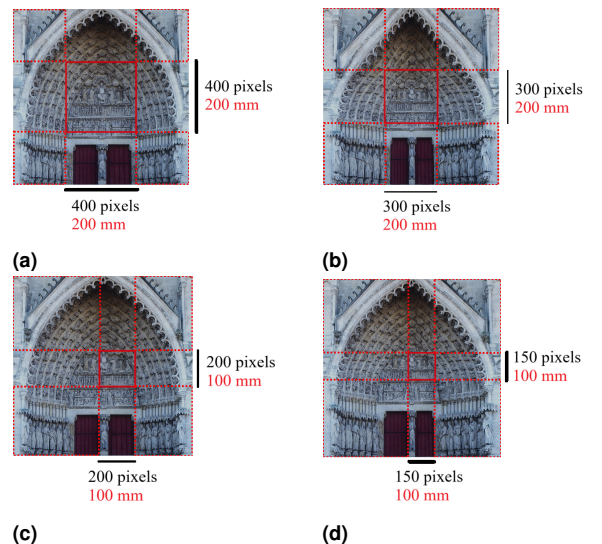
412 In summary, within the mechanical and visual simulation,  
 413 Visadapt allows capturing multi-resolution images close to the  
 414 desired ones,  $\eta$  being from 1.1 to 1.19 in the considered examples,  
 415 even though the images suffer from off-center distortions.

## 416 5. PROTOTYPE AND RESULTS

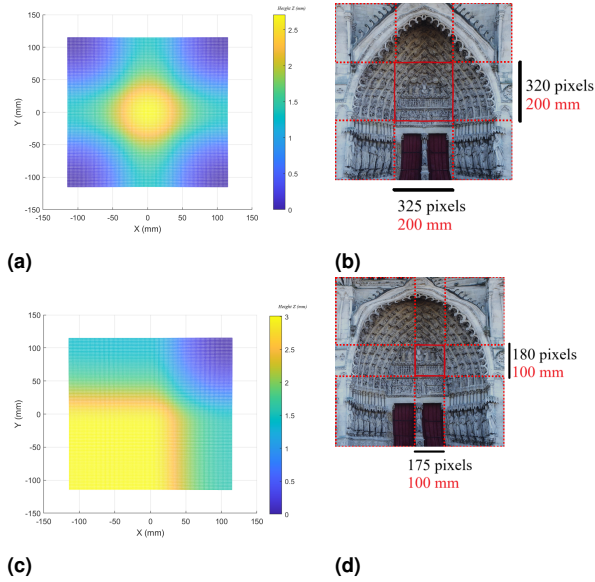
417 This section introduces the prototype of Visadapt and shows  
 418 preliminary results of image capture, as a proof-of-concept con-  
 419 firming the design decisions made following the simulation  
 420 study.

### 421 A. The actual prototype of Visadapt

422 The prototype of Visadapt (Fig. 7) is based on the results obtained  
 423 through the simulations reported in Section 4. The dimensions  
 424 match the description of Section C, and the respective locations  
 425 of the camera, the mirror and the scene remain as in Section A.  
 426 Moreover, Visadapt is made of the materials listed in Section B.



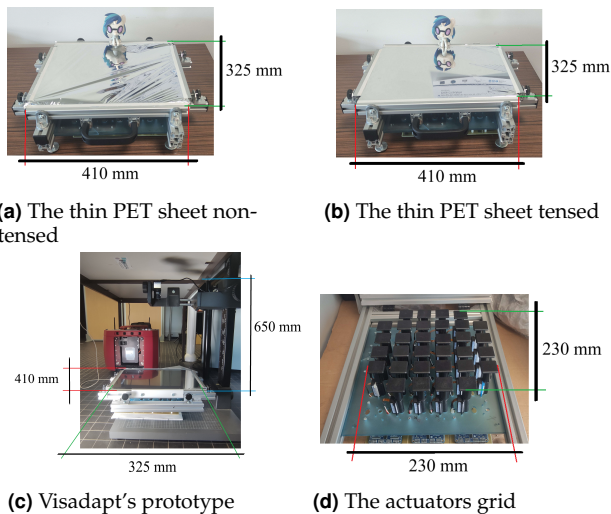
**Fig. 5.** Conventional (Fig. 5a and 5c) and desired (Fig. 5b and 5d) multi-resolution images (the target area is framed in red). Whereas the first line corresponds to the scene-to-image mapping  $\mathcal{M}_1$ , the second illustrates the mapping  $\mathcal{M}_2$ .



**Fig. 6.** For the two mappings  $\mathcal{M}_1$  (first line) and  $\mathcal{M}_2$  (second line), X-Y view of the optimal mirror shapes (a,c) and ray traced reflections, through the actuated mirror, of the scene, with the target area (defined in Fig. 5) framed in red (b,d).

427 For the prototype, we used servomotors [Actuonix® PQ12](#)  
 428 which move linearly along a main axis, on a 20 mm range. In  
 429 order to enlarge the actuation base, a 30 mm × 30 mm square  
 430 pad is added to the head of each actuator (see Sec. B). The pads  
 431 are not rigidly attached to the 4 mm thick silicone smoothing  
 432 layer. There is no issue when pushing nor when the actuator  
 433 moves back to its minimal height since the smoothing layer is  
 434 between the actuators grid and the tensed PET sheet. The PET  
 435 sheet is tensed by a rigid frame along the four borders of the  
 436 sheet (see Fig. 7a and Fig. 7b for a comparison between reflection  
 437 of relaxed and tensed PET sheets).

438 Practically, the prototype of Visadapt can be modified easily  
 439 with a different number of actuators, thanks to a latch that also  
 440 allows to easily change the silicone sheet used as a smoothing



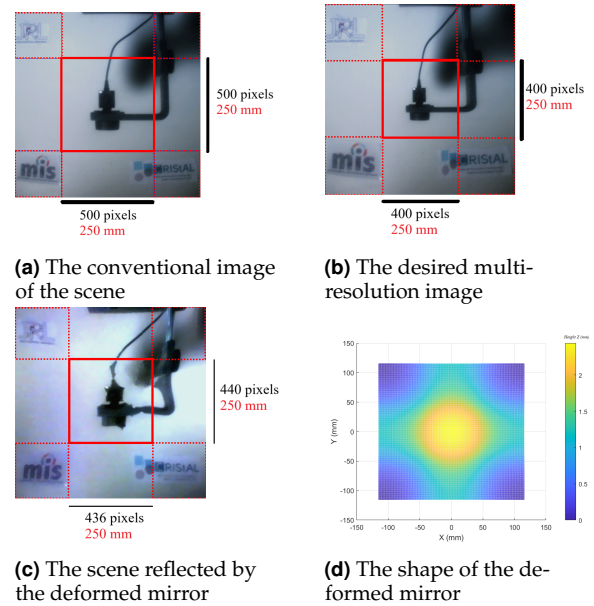
**Fig. 7.** The deformable mirror, the prototype of Visadapt and the actuators grid.

441 layer.

## 442 B. Results

443 The prototype is facing a planar scene made of the logos of the  
 444 three laboratories involved in this project, placed on three dif-  
 445 ferent corners of the field-of-view (Fig. 8a). For this experiment,  
 446 we consider the *scene-to-image* mapping  $\mathcal{M}_3$ , which reduces the  
 447 target area definition of 20%, from 500 × 500 pixels to 400 × 400  
 448 pixels (Fig 8b). Subsequently, the resolution of the four corners  
 449 of the captured field-of-view increases from 250 × 250 pixels to  
 450 300 × 300 pixels.

451 The mapping  $\mathcal{M}_3$  is considered to compute the optimal shape  
 452 that the deformable mirror should reach. The shape computed  
 453 with b-splines [22] is mainly convex, with an actuation up to  
 454 2.5 mm (Fig. 8d). The image captured by the prototype (Fig. 8c)  
 455 indeed reduces the resolution of the target area to 440 × 440  
 456 pixels. Therefore,  $\eta = 1.1$ , the resolution of the target area is  
 457 10% higher than the desired one. Moreover, the three logos are  
 458 respectively magnified by  $G_1 = 1.20$  (for the JRL logo),  $G_2 =$   
 459 1.15 (for the MIS logo) and  $G_3 = 1.08$  (for the CRISTAL logo),  
 460 highlighting the spatially variable resolution of the image. These  
 461 results confirm that Visadapt is able to capture images with an  
 462 adaptive local resolution.



**Fig. 8.** Experiments with the prototype, images of the scene captured, with the target area framed in red, and the computed mirror shape.

## 463 6. CONCLUSION AND FUTURE WORKS

464 To capture more visual information in environments of hetero-  
 465 geneous density, we designed an adaptive multi-resolution cata-  
 466 dioptic camera, Visadapt, by introducing a new deformable  
 467 and controllable mirror, a tensed PET sheet deformed by a grid  
 468 of servomotors, in front of a conventional camera. The new  
 469 camera was iteratively co-designed with realistic mechanical  
 470 simulations to select proper materials and refine the geometry  
 471 of elements interfacing them. Results of the first prototype show  
 472 that Visadapt is able to magnify several disjoint areas at once,  
 473 increasing by 20% the resolution of off-centered parts of the

474 field-of-view, without losing any intermediary area, of course  
475 shrinking the resolution of the latter.

476 In order to take full benefit of its dynamic adaptation poten-  
477 tial, future works aim at reducing the size of the Visadapt proto-  
478 type in order to mount it on mobile robots, *e.g.* to improve au-  
479 tonomous navigation in challenging environments where large  
480 areas are uniform and a few are visually rich.

481 **Funding.** The work presented in this paper has been partially funded  
482 by Amiens métropole, the French Ministère de l'Enseignement Supérieur  
483 et de la Recherche (MESR) and by the "Visadapt" project, CNRS INSIS  
484 PEPS 2021-2022.

485 **Acknowledgments.** We would like to thank Rayane Méchik, a Mas-  
486 ter student at Polytech Lille, for the technical support he gave to this  
487 project during his internship at CRIStAL.

488 **Disclosures.** The authors declare no conflicts of interest.

489 **Data Availability Statement.** No data were generated or analyzed  
490 in the presented research.

## 491 REFERENCES

- 492 1. H. Maître, *From photon to pixel: the digital camera handbook* (John  
493 Wiley & Sons, 2017).
- 494 2. G. Caron, E. Marchand, and E. M. Mouaddib, *Auton. Robots* **35**, 177  
495 (2013).
- 496 3. J. Gaspar and J. Santos-Victor, "Visual path following with a catadioptric  
497 panoramic camera," in *IEEE Int. Symp. on Intelligent Robotic Systems*,  
498 (1999), pp. 139–147.
- 499 4. M. Sheckells, G. Garimella, and M. Kobilarov, "Optimal visual servoing  
500 for differentially flat underactuated systems," in *IEEE/RSJ Int. Conf. on*  
501 *Intelligent Robots and Systems*, (2016), pp. 5541–5548.
- 502 5. X. Sun, X. Zhu, P. Wang, and H. Chen, "A review of robot control with  
503 visual servoing," in *IEEE Annual Int. Conf. on Cyber Technology in*  
504 *Automation, Control, and Intelligent Systems*, (2018), pp. 116–121.
- 505 6. P. Kumar, A. Dick, and T. S. Sheng, "Real time target tracking with pan  
506 tilt zoom camera," in *IEEE Digital Image Computing: Techniques and*  
507 *Applications*, (2009), pp. 492–497.
- 508 7. R. Cucchiara, A. Prati, R. Vezzani *et al.*, "Advanced video surveil-  
509 lance with pan tilt zoom cameras," in *IEEE Int. Workshop on Visual*  
510 *Surveillance*, (2006), pp. 334–352.
- 511 8. J. Goy, B. Courtois, J. Karam, and F. Presseccq, *Analog. Integr. Circuits*  
512 *Signal Process.* **29**, 95 (2001).
- 513 9. C. Quilliet and B. Berge, *Curr. Opin. Colloid & Interface Sci.* **6**, pp. 34  
514 (2001).
- 515 10. L. Li, D. Wang, C. Liu, and Q.-H. Wang, *Opt. express* **24**, pp. 2931  
516 (2016).
- 517 11. G. S. Jung and Y. H. Won, *Appl. Opt.* **59**, pp. 5527 (2020).
- 518 12. M. Levoy, *Computer* **39**, 46 (2006).
- 519 13. D. C. Sims, Y. Yue, and S. K. Nayar, "Towards flexible sheet cameras:  
520 Deformable lens arrays with intrinsic optical adaptation," in *IEEE Int.*  
521 *Conf. on Computational Photography*, (2016), pp. 1–11.
- 522 14. D. C. Sims, O. Cossairt, Y. Yu, and S. K. Nayar, arXiv preprint  
523 1804.07052 (2018).
- 524 15. A. Zomet and S. K. Nayar, "Lensless imaging with a controllable aper-  
525 ture," in *IEEE Conf. on Computer Vision and Pattern Recognition*, ,  
526 vol. 1 (2006), pp. 339–346.
- 527 16. S. Kuthirummal and S. K. Nayar, "Flexible mirror imaging," in *IEEE Int.*  
528 *Conf. on Computer Vision*, (2007), pp. 1–8.
- 529 17. R. Freeman and J. E. Pearson, *Appl. Opt.* **21**, 580 (1982).
- 530 18. R. Davies and M. Kasper, *Annu. Rev. Astron. Astrophys.* **50**, p305  
531 (2012).
- 532 19. P.-Y. Madec, "Overview of deformable mirror technologies for adaptive  
533 optics and astronomy," in *Int. Society for Optics and Photonics, Adaptive*  
534 *Optics Systems III*, , vol. 8447 (2012).
- 535 20. J.-F. LAYERLE, X. SAVATIER, E. MOUADDIB, and J.-Y. ERTAUD, "Catadioptric  
536 sensor for a simultaneous tracking of the driver's face and the road  
537 scene," in *8th Workshop on Omnidirectional Vision, Camera Networks*  
538 *and Non-classical Cameras*, (2008).
- 539 21. A. Ghorayeb, A. Potelle, L. Devendeville, and E. M. Mouaddib, "Optimal  
540 omnidirectional sensor for urban traffic diagnosis in crossroads," in  
541 *IEEE Intelligent Vehicles Symposium*, (2010), pp. 597–602.
- 542 22. R. Swaminathan, S. Nayar, and M. Grossberg, "Designing mirrors for  
543 catadioptric systems that minimize image errors," in *5th Workshop on*  
544 *Omnidirectional Vision*, (2004).
- 545 23. J. S. Gill, M. Moosajee, and A. M. Dubis, *Nat. Publ. Group, Eye* **33**, pp.  
546 1683 (2019).
- 547 24. ISO®, "ISO 9039:2008 Optics and photonics – Quality evaluation of  
548 optical systems – Determination of distortion," (2008).
- 549 25. F. Faure, C. Duriez, H. Delingette, *et al.*, "Soft tissue biomechanical  
550 modeling for computer assisted surgery," pp. 283–321 (2012).
- 551 26. B. Biswas, S. Chatterjee, S. Mukherjee, and S. Pal, *Electron. J. Math.*  
552 *Anal. Appl.* **1**, pp. 2090 (2013).
- 553 27. J. Allard, S. Cotin, F. Faure, *et al.*, "Sofa-an open source framework for  
554 medical simulation," in *IOP Press, Medicine Meets Virtual Reality*, , vol.  
555 125 (2007), pp. 13–18.