A Survey on Adaptive Cameras

Julien Ducrocq, Guillaume Caron

International Journal of Computer Vision manuscript No.
(will be inserted by the editor)

A survey on adaptive cameras

Julien Ducrocq · Guillaume Caron.

Received: date / Accepted: date

Abstract This paper surveys adaptive cameras, *i.e.* any camera device able to change its geometric settings. We consider their classification in four categories: lensless, dioptric, catadioptric and polydioptric cameras. In each category, we report and describe all the existing adaptive cameras. Then, the known applications of these devices are summarized. Finally, we discuss open research lines for new adaptations of cameras, and their promising uses.

Keywords state-of-the-art survey ; artificial vision ; adaptive cameras ; non-conventional cameras ; optics ; image sensor.

1 Introduction

1.1 Motivation

More than a thousand years ago, Ibn-Al-Haytham modeled the path of scene light rays through an infinitesimal *pinhole* on the front side of a chamber, later called *camera obscura* or *dark room*. On the back part of the chamber, the light rays form an *image*, *i.e.* a left-right and top-bottom inverse bi-dimensional representation of the three-dimensional scene. After hundreds of years, the 19^{th} century saw the birth of the first *view cameras* (Evens 2008). To get more light rays at once, the pinhole is replaced by lenses. Light rays enter a movable accordion box with lenses on its front part. A lot more compact, film cameras with objective made of lenses and a photographic film emerged one century later. Since two decades, digital cameras, with a matrix of photodiodes instead of the photographic film, have become standard. Nowadays, almost everyone have a miniature digital camera built into one's phone and use them for everyday life.

Beyond consumer-grade cameras, *adaptive cameras* are actively being researched. They are designed to adapt to changes in the scene interactively or automatically. Below, we draw a first general definition of adaptive cameras.

Definition 1 – **Adaptive camera:** A camera is considered adaptive if the geometrical properties of its optics (aperture and lenses) and image sensor are spatially and temporally variable. ◇

Adaptive cameras feature a large scope of adaptations, which modify particularly:

- the camera viewpoint and number of viewpoints;
- the field of view, possibly non-linear, discontinuous;
- the magnification of scene elements;
- the image definition and resolution;
- the depth of field of the camera;
- the local sharpness within the image.

Section 2 dives deeper into the definitions of camera parts and of the various subjects of adaptation.

Section 3 defines accurately four types of adaptive cameras: lensless, dioptric, catadioptric and polydioptric. Overall, this article surveys the state-of-the-art of adaptive cameras in which lensless cameras primarily target adapting the focus, catadioptric and polydioptric cameras mainly concentrate on adapting the field-of-view and the viewpoint whereas dioptric cameras aim at handling them three.

J. Ducrocq and G. Caron are with
Université de Picardie Jules Verne, MIS lab, Amiens, France
E-mail: julien.ducrocq07@gmail.com, guillaume.caron@u-picardie.fr

G. Caron is also with
CNRS-AIST JRL (Joint Robotics Laboratory), IRL, Tsukuba, Japan

1.2 Previous surveys

In the literature, several articles survey different types of cameras among which a subset is adaptive.

Imaging devices using coded apertures (Liang 2020), including lensless cameras, are among computational cameras (Zhou and Nayar 2011). They combine optical settings and processing to produce images that conventional cameras are unable to capture. This includes light field capture devices (Levoy 2006) of which some (Conti et al. 2020) feature dioptric designs and a few polycameras. Polycameras were themselves surveyed separately in (Olagoke et al. 2020).

Another type of camera where one finds adaptive cameras is the bioinspired type, *i.e.* cameras inspired by mammal or arthropod eyes (Kim et al. 2020). Among them, the Artificial Compound Eyes (ACE) are comprehensively reported in (Wu et al. 2017; Cheng et al. 2019).

Inversely, adaptive optics is a field where we find only adaptive cameras (Madec 2012), mostly cameras facing deformable mirrors (DM) to compensate atmosphere perturbations and improve the quality of astronomical observations.

All these surveys, despite very vast, focus on different kinds of cameras and none provides a global synthesis about adaptive cameras.

1.3 Contributions

We propose a comprehensive survey on adaptive cameras featuring:

- a general structure of the characteristics of adaptive cameras ;
- the classified state-of-the-art on adaptive cameras ;
- a list of the applications of adaptive camera ;
- guidelines for future research.

This survey is also the opportunity to update older ones (Evens 2008; Madec 2012; Levoy 2006) (almost 10 years old), without repeating all their content.

The scope of this survey is precisely defined: it does not feature post-processing algorithms and focuses on visible light. Finally, it does not include optical devices meant for non-imaging purposes, such as optical fibers.

1.4 Outline

Section 2 shortly recalls the properties and characteristics of conventional cameras in order to better settle the adaptations of adaptive cameras. Section 3 proposes the global structure of adaptive camera types. Then, Sections 4, 5, 6 report, respectively, all adaptive lensless, dioptric, catadioptric cameras, and Section 7 presents the polycameras. Section 8 reports all applications of adaptive cameras. Finally, Section 9 opens the article with a discussion on future research lines before conclusion (Sec. 10).

2 General properties of conventional cameras

In this section, we recall the basics of a conventional camera with respect to which adaptations of adaptive cameras are described in farther sections.

Definition 2 – **Light field:** The plenoptic function expresses the radiance $\mathcal{L} \in \mathbb{R}_+$ at every scene point $(x, y, z) \in \mathbb{R}^3$, for any incident direction $(\phi, \psi) \in \mathbb{R}^2$ and any wavelength $\lambda \in \mathbb{R}_+$ across time $t \in \mathbb{R}_+$ (Ihrke et al. 2016; Adelson et al. 1991). This is thus a seven parameters function. The *light field* is a simplification of the plenoptic function to four parameters, *i.e.* the time is constant, the wavelength is monochromatic and the radiance is assumed constant along the light ray direction (Ihrke et al. 2016). It is represented by a 2D position $(u, v) \in \mathbb{R}^2$ and the 2D direction $(s, t) \in \mathbb{R}^2$ of a light ray (Ng et al. 2005). ◇

Definition 3 – **Conventional camera:** A *conventional camera* is a device that captures images, *i.e.* 2D slices of the 4D light field. It uses an objective made of several lenses, a diaphragm and a planar image sensor, often with a rectangular shape. Since cameras available for consumers have a wide variety of fields of view, this article qualifies as conventional any camera with a field of view comparable with the human vision signal recognition area, about 60° (Maître 2017). ◇

Definition 4 – **Perspective camera:** A *perspective camera* images an object at different Z coordinates (depth) with different sizes, whereas an orthographic camera images it always at the same size, though with different blurriness (Poling 2015). ◇

The rest of this section formally recalls optical characteristics and the image sensor characteristics of a conventional camera.

2.1 Image formation through a lens

Even though referred as *lens* in consumer products, cameras have complex objectives cascading several lenses. However, they can be considered as *centered optical systems* with a single focal length $f \in \mathbb{R}_+^*$, and

approximated as a unique thin lens (Maître 2017) because they are made to share the same optical axis. Then, thanks to geometrical optics, image formation can be computed.

Approximated as a thin lens, a camera objective makes light rays from *object points*, in the world, focusing into *image points*, or pixels. The *thin lens equation* relates point coordinate $Z_o \in \mathbb{R}_+^*$ of the object and its counterpart $Z_i \in \mathbb{R}_+^*$ onto the image plane along the camera optical axis ${}^c\vec{\mathbf{Z}} \in \mathbb{R}^3$, inside the camera frame $\mathcal{F}_c$ of axes ${}^c\vec{\mathbf{X}} = [1,0,0]^\top$, ${}^c\vec{\mathbf{Y}} = [0,1,0]^\top$, ${}^c\vec{\mathbf{Z}} = [0,0,1]^\top$, centered at the lens:

$$\frac{1}{Z_o} + \frac{1}{Z_i} = \frac{1}{f}. \tag{1}$$

Furthermore, scene elements are observed smaller and inverted after passing through a lens (Maître 2017). This is due to the transverse magnification $G \in \mathbb{R}$, the ratio of the object observation size over the true object size, that can be expressed as:

$$G = -f/s = -s'/f, \tag{2}$$

with $s = Z_o - f$ and $s' = Z_i - f$.

In most cases, the object is far away the camera objective so f is negligible compared to Z_o. Therefore, the modulus of G increases as f does, so objects are imaged bigger with lenses of longer focal length.

2.2 Image in focus

The lenses inside the objective have a diameter obviously larger than a pinhole. Therefore, they can capture bright images without requiring a long exposure time.

The diaphragm controls the flow of light by modifying the diameter $D \in \mathbb{R}_+^*$ of the optics entrance pupil. The objective of a camera is also characterized by the aperture number, $N = f/D$ (Corke 2017). All the rays penetrating the camera objective during an exposure time $t_e \in \mathbb{R}_+$ pass through the camera *optical center*, *i.e.* the origin $\mathbf{C} = [X_c, Y_c, Z_c]^\top \in \mathbb{R}^3$ of the camera frame. Further away, they focus on the *focal plane* (also known as *focal surface*), to form the *circle of illumination* (Kingslake 1992).

Pixels are captured sharp only if the light from their scene point counterpart does not spread onto their neighbors. Such scene points are *in focus* and belong to the plane conjugated to the focal plane, *the focus plane* (Maître 2017).

2.3 Central projection model

When the points are in focus, their image counterpart can be computed by the *central projection model*. In computer vision, this model assimilates the camera to a pinhole chamber, with a no dimension hole. Within this model, the light rays focus at $Z = f$ into a virtual image plane, symmetric of the focal plane, *i.e.* the actual image plane in this model, with respect to the camera optical center. Thus, the image is a small version of the scene, but not inverted (which is not the case physically, with camera having lenses) (Corke 2017).

The three-dimensional (3D) points observed by the camera are called *object points* ${}^c\mathbf{P} = [X, Y, Z]^\top \in \mathbb{R}^3$, expressed in $\mathcal{F}_c$. They project into bi-dimensional (2D) image points $\mathbf{p} = [x, y]^\top \in \mathbb{R}^2$:

$$x = \frac{fX}{Z},\ y = \frac{fY}{Z}. \tag{3}$$

Image points are expressed in the homogeneous form $\tilde{\mathbf{p}} = [\tilde{x}, \tilde{y}, \tilde{z}]^T \in \mathbb{P}^2$ by (Corke 2017):

$$\tilde{x} = fX,\ \tilde{y} = fY,\ \tilde{z} = Z. \tag{4}$$

Introducing the homogeneous coordinates ${}^c\tilde{\mathbf{P}} = [X, Y, Z, 1]^T \in \mathbb{P}^3$ of object points in $\mathcal{F}_c$, the perspective projection described above can be expressed as a matrix product:

$$\tilde{\mathbf{p}} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} {}^c\tilde{\mathbf{P}}. \tag{5}$$

Perspective cameras (Def. 4) follow this model. As implied by (Eq.(4)), using a bigger focal length magnifies the image (which confirms the impact of f on G, the transverse magnification (2)).

2.4 The points out of focus

When the points are not in focus, this model is not valid anymore. Indeed, for each object point, several light rays enter the lens instead of one. They project onto the *circle of confusion*, a spot usually assumed circular of diameter $\epsilon \in \mathbb{R}_+^*$ (Kingslake 1992). In practice, a tolerance $\mathcal{E} \in \mathbb{R}_+^*$ is accepted on ϵ to consider the image points sharp. When the focus is made on a point ${}^c\mathbf{P}$, and the point is far from the camera ($Z >> f$), the depth of field $\Delta \in \mathbb{R}_+$ expresses the distance between the two extreme planes in focus of equation $Z = Z^+ \in \mathbb{R}$ and $Z = Z^- \in \mathbb{R}$ (Maître 2017; Potmesil and Chakravarty 1982), $Z^+ \geq Z^-$:

$$\Delta = Z^+ - Z^- = \frac{\mathcal{E} N Z^2}{f^2}. \tag{6}$$

Points with Z out of the $[Z^-, Z^+]$ interval are *out of focus* and the diameter of their circle of confusion increases as their distance to the planes of focus (Maître 2017). Δ increases as f decreases or N increases. Then, reducing the lens diameter makes more scene elements appearing sharp in the resulting image, at the price of letting enter less light into the camera. To compensate those losses, a longer exposure time may be set, if appropriate.

In addition, because of imperfections on camera lenses, image points appear sharper inside the *circle of definition* than in the rest of the larger circle of illumination (Sec. 2.2).

Unlike depth of field, the *depth of focus* $\mathcal{Z}_f$ is internal to the camera. It sets the tolerance on the image sensor placement relatively to $\mathbf{C}$, to ensure $\epsilon < \mathcal{E}$ (Allen and Triantaphillidou 2012). For far objects, it is expressed as:

$$\Delta_f = 2N\ \mathcal{E}. \tag{7}$$

The depth of focus, depth of field and thin lens equation principle, additionally to a 2D, simplified, representation of the projection model, with the image reversed since a lens is considered are all schematised in Figure 1.

2.5 The field of view

The scene elements appear sharp or blurry in the image depending on their position with respect to the camera. However, they can only be observed if they fall inside the camera field of view (FOV). It is an open rectangular pyramid, starting from the camera viewpoint, which subtends two angles $\theta_h \in [0, 2\pi]$ and $\theta_v \in [0, \pi]$, in the horizontal and vertical planes respectively (Corke 2017):

$$\theta_h = 2\ \arctan\left(\frac{H}{2f}\right),\ \theta_v = 2\ \arctan\left(\frac{V}{2f}\right), \tag{8}$$

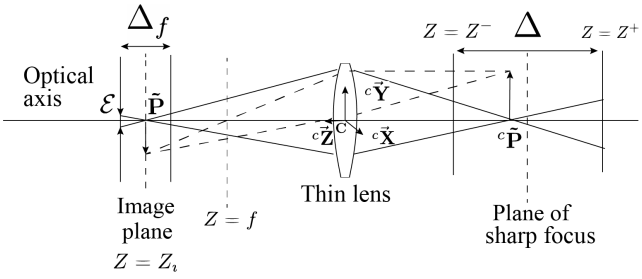


Fig. 1: Image projection of world points, and simplified representations of depth of field Δ and depth of focus Δ_f.

with $H \in \mathbb{R}_+^*$ the horizontal and $V \in \mathbb{R}_+^*$ the vertical dimensions of the image sensor. Depending on the case, it can be expressed as $\theta_h \times \theta_v$, usually in degrees unit (noted °), or only one of these two values. It is important to notice that the shorter the focal length, the wider the FOV (Maître 2017).

The *omnidirectional* FOV, 360° × 180°, is the whole sphere surrounding the camera. However, some works (Nayar 1997) qualify as omnidirectional a FOV of $180° \times 180°$, that this article names *panoramic* FOV, to distinguish with the former.

2.6 The image sensor

From (Eq. (8)), it comes that the FOV also depends on the image sensor dimensions. Image sensors, also called *photosensitive matrices*, are, in general, rectangular array of photodiodes. Two technologies, Charge Coupled Devices (CCDs) and Complementary Metal-Oxyde Semiconductors (CMOS) are used on digital cameras, but CMOS are the most used image sensor nowadays (Maître 2017). A large variety of formats are used in digital cameras, from less than 4 mm width on smart-phones (Maître 2017) to 200 mm on highly sensitive cameras for sky observation (Sako et al. 2016). The diagonal of image sensors is chosen to fit into the circle of definition in order to minimize image blurriness (Langford 2000). The photodiodes, also called *photosites*, or *photodetectors*, but mostly known as *pixels*, convert the incoming light photons on the image plane into pixel brightness, coded on several bits (Maître 2017).

Definition quantifies the number of pixels ($R \times C$, $R, C \in \mathbb{N}^*$) along the image sensor width and height, whereas *resolution* gives the size ($k_u \times k_v$, $k_u, k_v \in \mathbb{R}_+^*$) of its photodiodes. Photodiodes are generally square-shaped and measure between 0.9 and 10 μm of side (Corke 2017; Maître 2017). The smallest the photodiodes, the highest the level of perceptible scene details. However, photodiodes of a few micrometers side generate spatial noise in images (Goy et al. 2001). Thanks to the resolution of the camera, the image plane coordinates $\mathbf{p}$ (Eq. (3)) can be converted into digital, or pixel, coordinates $\mathbf{u} = [u, v]^T \in \mathbb{R}^2$ by (Corke 2017):

$$u = \frac{x}{k_u} + u_0,\ v = \frac{y}{k_v} + v_0, \tag{9}$$

with $(u_0, v_0) \in \mathbb{R}^2$ the *principal point*, where the optical axis intercepts the image plane.

Finally, one can directly derive the pixel coordinates of a 3D point of the scene in focus by:

$$\tilde{\mathbf{u}} = \begin{bmatrix} 1/k_u & 0 & u_0 \\ 0 & 1/k_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} {}^c\tilde{\mathbf{P}}, \tag{10}$$

with $\tilde{\mathbf{u}} = [\tilde{u}, \tilde{v}, \tilde{w}]^T \in \mathbb{P}^2$ the homogeneous pixel coordinates where $\tilde{u} = u\tilde{w}$, $\tilde{v} = v\tilde{w}$ and $\tilde{w} \in \mathbb{R}^*$.

Both optical settings and image sensor characterize the cameras. The previously recalled definitions are bases used for adaptive cameras, of which a general classification is proposed next.

3 State-of-the-Art structure

In this article, we classify the cameras in four main categories:

- lensless cameras
- dioptric cameras
- catadioptric cameras
- polycameras.

Each category is represented in Figure 2 and defined below.

Definition 5 – Lensless camera: A lensless camera is only made of apertures and an image sensor. ◇

The most basic lensless camera is a pinhole, but it is not energy efficient, since only little light pass through this small aperture. Over the years, designs with more apertures (Fig. 2a) have been proposed. Optical masks with transparent and opaque regions or array of pinholes are used by lensless cameras instead of the lenses and diaphragm of conventional cameras (Boominathan et al. 2016).

Definition 6 – Dioptric camera: Components of a dioptric camera are an objective made of lenses and a diaphragm, and an image sensor (Fig. 2b). ◇

Conventional cameras (Def. 3) are dioptric. However, properties differ between designs, for instance for a field of view wider than the recognition area of the human vision system (Sec. 2.5).

Definition 7 – Catadioptric camera: A catadioptric camera combines mirrors to objective components of a dioptric camera. ◇

Mirrors of a catadioptric camera can be in the objective, especially the ones with a very long focal length (Sec. 2.1), to reduce bulkiness, compared to much more common dioptric objectives (Maître 2017). Other catadioptric cameras are dioptric cameras pointed to a mirror (Fig. 2c), possibly curvy to change the camera FOV (Baker and Nayar 1999).

Definition 8 – Polycamera: A polycamera is a cluster of several lensless, dioptric or catadioptric cameras, possibly a combination of several types (Fig. 2d), therefore, relying on several image sensors. Images acquired by them are merged into the polycamera image. ◇

Then, we consider as adaptive any camera able to change its optical settings online, in contrast to static cameras with constant properties over time. Hence, we divide each of the four main categories of cameras into three subcategories, depending on camera adaptability: static, manually adaptive, and automatically adaptive.

Definition 9 – Camera with static properties: These cameras do not change their optical or image sensor properties over time. ◇

In this survey, very few cameras of static properties are presented, only for the purpose of positioning the adaptive designs.

Definition 10 – Manually adaptive cameras: When the user directly acts on a physical degree of freedom of the camera to change one of its properties, thus impacting captured images, the adaptation is considered manual. ◇

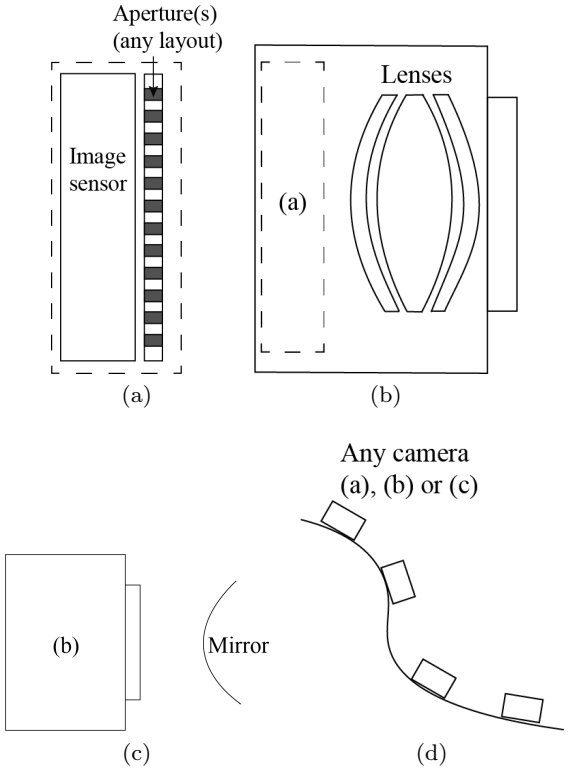


Fig. 2: Schematic design of the considered cameras: a lensless camera (a), a dioptric camera (b), a catadioptric camera (c), a polycamera (d).

A basic example of manual adaptation is acting on a ring of a camera objective to move the lenses allowing to directly change the plane of focus.

Definition 11 –Automatic adaptation: Any adaptation which is not directly controlled by the user is automatic. The user gives a target, then the camera modifies one or more of its optical or image sensor geometrical properties to impact the captured image accordingly. ◇

For instance, if the user selects an area in the live view of a dioptric camera, optical blocks therein are automatically controlled to change the focus in order to maximize the sharpness of the selected image area.

Figure 3 proposes a set representation of the hierarchical structure of camera types, interlinked with their adaptability. Every adaptability are comprehensively surveyed for each camera type in Sections 4 to 7.

Remark 1 (Automatic lensless cameras) So far, there is no automatic lensless cameras in the state-of-the-art. Therefore, Section 4 reports existing static and manually adaptive lensless cameras. The latter are separated in two different categories, the lensless cameras with an aperture changing of shape and the ones which can be curved as a whole.

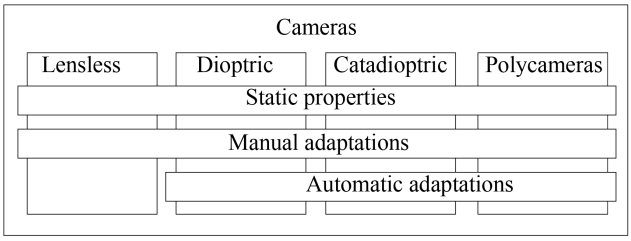


Fig. 3: The interlinked structure of the state-of-the-art.

4 Lensless cameras

With several to many apertures, lensless cameras capture much more light rays at once than a pinhole, even more than conventional dioptric designs (Boominathan et al. 2016). Marginal in the state-of-the-art, lensless cameras (Def. 5) are hundreds of times thinner (10 to 500 μm thick) than conventional ones (10 to 20 mm thick) (Asif et al. 2016). Instead of "bulky" optics, these designs are made of a *coded aperture* parallel to an image sensor (Boominathan et al. 2016). Hence, for each light source, a shadow of the mask is cast on the sensor, so reconstruction algorithms are needed to exploit the images. Among them, *compressive sensing* produces images of a higher resolution than the number of captured measurements (Qaisar et al. 2013).

4.1 Lensless cameras with static properties

Without any lens, an image sensor would simply record the average light intensity from the scene (Boominathan et al. 2016). To solve this issue, *Coded aperture imaging* uses arrays of pinholes or optical masks with optimized patterns to focus an image on the sensor. Mainly, two types of mask are suitable for lensless cameras (Boominathan et al. 2020):

- phase masks, which modulate the phase of the incoming light rays thanks to constructive and destructive interference ;
- amplitude masks are patterns of opaque and transparent rectangles. Opaque ones block light rays while transparent rectangles let them pass through.

This section focuses on lensless cameras based on amplitude masks or arrays of pinholes, because adaptive cameras are inspired by them. Among manufacturers, Hitachi®is developing a lensless camera with a phase mask, consisting of concentric opaque rings printed on a permeable film (Hitachi® 2017). Such mask, alternating transparent and opaque rings, is called a *zone plate*, and extends the camera depth of field (Sec. 2.4) (Indebetouw and Bai 1984).

Amplitude masks were used more than 50 years ago for X-ray and gamma rays imaging because of their high signal to noise ratio (Fenimore and Cannon 1978; Gottesman and Fenimore 1989). At first, such masks were implemented on dioptric camera lenses (see Sec. 5.3.1). Only $500\mu m$ thick, FlatCam (Asif et al. 2016) is a lot more compact than any dioptric camera. This lensless camera directly combines an image sensor of $6.7mm$ diagonal to capture 512×512 pixels color images. It uses an amplitude mask separable in two orthogonal patterns of large stripes, optimized for light efficiency. This camera is more compact than conventional ones, though is only able to image the scene within a narrow set of depths (approximated by a single depth plane) (Hua et al. 2020), whereas pinhole cameras are supposed to have a virtually infinite depth of field (Young 1971).

Lensless cameras based on pinhole arrays produce multiple images of the same scene (one for each pinhole) (Newman and Rible 1966). The optimal pinhole size, which depends on the distance from the sensor to the pinholes, is a few hundreds of μm for lensless

cameras. Several adaptive lensless cameras are based on that principle to build up 3D images or discontinuous FOV (Sec. 4.2.1).

In theory, pinhole cameras have a 180° FOV (Sec. 2.5) with no distortion (Young 1971). Therefore, the FOV of a lensless camera is limited to 180° in the forward direction only. To increase this maximum FOV, the *super FOV lensless camera* captures both its front and rear spaces (Fig. 4) (Nakamura et al. 2019). It consists in two CMOS sensors of 384×384 pixels, face to face, both of them with randomly $40\mu m \times 40\mu m$ square holes in blocks of 4×4 pixels. To obtain the best image quality, the ratio of air holes on the surface is 50%, determined by simulation. This system relies on *compressive sensing* algorithms (Qaisar et al. 2013) to reconstruct high resolution images with a small number of measurements. This camera should be able to acquire the omnidirectional FOV, though the proposed prototype acquires images of conventional FOV on both sides of the sensor (Nakamura et al. 2019).

4.2 Manual adaptations

To our knowledge, lensless cameras are all manually adapted, none of them is automated. Nonetheless, since their adaptations are very different in nature, the adaptive lensless cameras are subclassified below in two categories, depending if the aperture or the camera curvature can change.

4.2.1 Adaptive optics: aperture

Apart from one, all adaptive lensless cameras have a modifiable aperture. They either rely on array of pinholes, each being either opened or closed, or optimal masks with transparent or opaque rectangles.

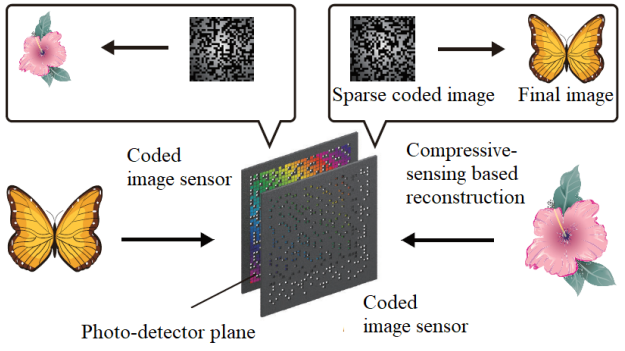


Fig. 4: Working principle of the super FOV lensless camera (reproduced from (Nakamura et al. 2019)).

Viewpoint adaptation: Opening a single pinhole sets the viewing direction of a lensless camera (Zomet and Nayar 2006). Changing the opened pinhole avoids to pan and tilt the camera, motion blur and mechanical constraints. This is implemented by a single attenuating liquid crystal (LC) sheet, with a locally controlled transmittance, just in front of an image sensor. On a prototype, a LC sheet placed at 55 mm of the (dismounted) lens mount of a camera, simulates arrays of 0.33 mm × 0.33 mm closed or opened pinholes (Zomet and Nayar 2006). Shifting the open pinhole changes the *scene-to-image mapping* W_x, *i.e.* the region of the scene represented in the image. Adding layers of LC sheets can produce several pinholes simultaneously. Then, the mappings W_x are multiplied, hence multiple sub-images i_x are captured simultaneously (Zomet and Nayar 2006). The image I_{merged} merges all the sub-images i_x, gathering disjoint scene regions (Fig. 5c).

If the viewpoints get closer, their field of views overlap. Therefore, multiple views of the same scene content can be captured, *e.g.* for 3D reconstruction. However, the proximity of the pinholes makes the reconstructed images blurred (Schwarz et al. 2015, 2016). To capture all these viewpoints sharp, (Schwarz et al. 2016) proposes to multiplex the pinhole array. During a single acquisition on an exposure time $t_{max} \in \mathbb{R}_+^*$, the pinholes viewing a scene region (3D) $\boldsymbol{X} = (X, Y, Z) \in \mathbb{R}^3$ are opened during time samples $t_k < t_{max}$, capturing a sub-image $i_k(\mathbf{X})$, and closed the rest of the time. Then, gathering all the sub-images $i_k(\boldsymbol{X})$ produces a bright and sharp image $I_{\text{capture}}(\boldsymbol{X}) = \sum_{k=1}^{K} i_k(x, y) \cdot t_k$ of the scene region $(\mathbf{X})$ (Schwarz et al. 2016).

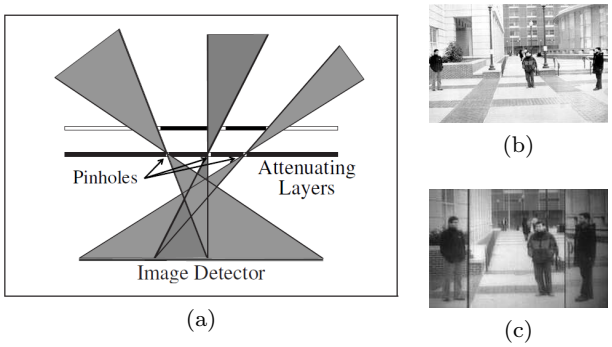


Fig. 5: A lensless camera with 3 pinholes (a), an image of a scene captured by a conventional camera (b) and the same image, taken at the same viewpoint, captured with the lensless camera configured as Fig.5a (c) (reproduced from (Zomet and Nayar 2006)).

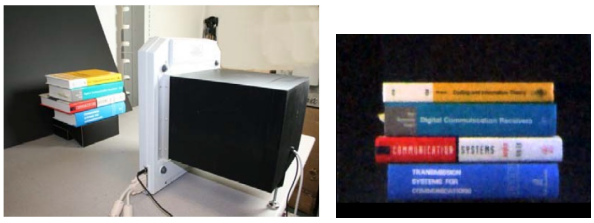
In order to capture all the sub-images $i_k(\mathbf{X})$, the camera of (Schwarz et al. 2016) is made of 3×3 individual units of controllable apertures. Each unit consists in three layers with 17 pinholes in total, each being $150\mu m$ in diameter. Practically, the multiplexing consists in switching the layers during a single exposure. The aperture layers are switched by settings.

Compressive sensing: Opening and closing pinholes during the exposure time can be used for compressive sensing (Sec. 4.1) (Huang et al. 2013).

In contrast with aforementioned lensless cameras, when scene elements are about ten centimeters close to a camera with an aperture array, closing the apertures $(u, v) \in \mathbb{R}^2$ makes the photodiode in the image sensor at the same position (u, v) not recording light anymore. Then, the corresponding pixel (u, v) is black in the image (Huang et al. 2013). This only applies in the specific case where scene elements are close to the camera. The adaptive aperture of this camera consists in a LC sheet of 302×217 square elements (like pixels). Each one is individually controllable to transmit or block the light rays. Therefore, the number of transmitting apertures directly changes the number of measurements, as a ratio of the 65534 pixels. Figure 6b displays an example image reconstructed from 25% of measurements. Nonetheless, the prototype (Fig. 6a) is clearly tens of centimeters thick, which is way thicker than any other lensless camera.

Focusing: The last adaptation reported in this section is focusing. Changing the shape of a coded aperture can also change the location of the plane of sharp focus (Hua et al. 2020). Inspired by FlatCam (Asif et al. 2016), *SweepCam* can image sharp scene elements lying in up to $j = 34$ settable planes at different depths $Z_{o_i} \in \mathbb{R}_+^*$, with $i \in [0, 33] \subset \mathbb{N}$. Translating the amplitude mask of this camera by 0.78 *cm* along the X axis (parallel to the u axis of the image sensor), moves the plane of sharp focus from 2.7 *cm* to 18 *cm*.

Instead of a mechanically translated mask, the prototype actually uses a programmable amplitude mask, made of a LC sheet between two cross polarizers. It is made of a 13×13 pattern of 2.27 *mm* variable optical and transparent regions of different areas. The pixel pitch of the LC sheet, *i.e.* the size of the smallest resolvable element, of $36\mu m$ constrains the image definition (up to 960×600 pixels). In this hardware implementation, the contrast ratio of the LC sheet (200 : 1) bounds the signal-to-noise ratio (SNR) of the captured images.



(a) (b)

Fig. 6: Camera prototype achieving adaptive compressive sensing (a) and image reconstructed from sparse measurements (25% of the maximum amount of pixels) (b). Reproduced from (Huang et al. 2013).

4.2.2 Adaptive optics and sensor: deforming aperture and sensor for focusing

This lensless camera is made of a $300\mu m$ thick *Söller collimator* and two luminescent concentrators as an image sensor (Koppelhuber and Bimber 2017). In practice, since it is very difficult to make a $300\mu m$ thick Söller collimator, the actual one is 6 *mm* thick.

A *Söller collimator* is a stack of several pinhole arrays. Unlike a single aperture (Fig. 7a) and thin aperture arrays (Fig. 7b), a Söller collimator (Fig. 7c) is designed to only let pass light rays perpendicular to its surface (Koppelhuber and Bimber 2017). Actually, each aperture lets enter light rays of incident angle in $[0, \alpha]$, with $\alpha \in \mathbb{R}_+$ the *collimation angle* (expressed in degrees here, for convenience).

The extent $\beta \in \mathbb{R}_+^*$ of captured light ray directions increases as the curvature $\kappa \in \mathbb{R}_+^*$ of the collimator

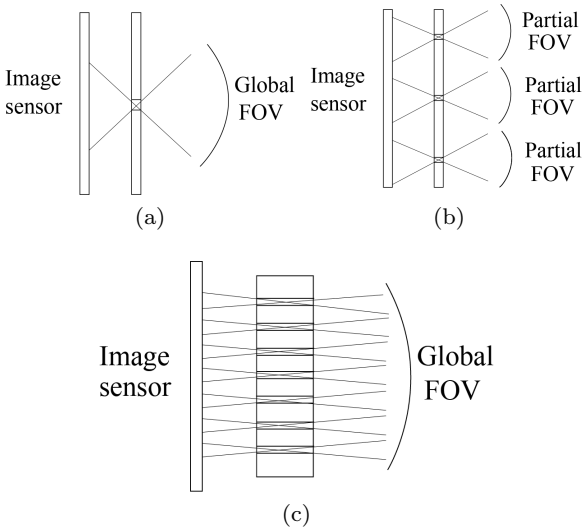


Fig. 7: The path of the incoming light rays through a single aperture (a), several apertures (b) and a Söller collimator (c).

(Fig. 8a):

$$\beta = \frac{W\kappa 180}{\pi}, \tag{11}$$

with W the width of the collimator.

Curving the Söller collimator can increase the FOV of this camera up to 195° (Koppelhuber and Bimber 2017). Indeed, the overall horizontal FOV of this lensless camera is $\theta_h = \alpha + \beta$. Hence, this lensless camera (with $W = 20cm$) views an horizontal FOV of $\theta_h = 15.19° + 180°$ when flexed with a radius of curvature $r = 1/\kappa = 62.5\ mm$. However, because of Söller collimators properties, this camera is mostly suitable for short-distance imaging, capturing sharp the objects within a few centimeters.

4.3 Wrap-up

Despite the bulkiness of some of the prototypes presented above, such as the compressive sensing camera (Sec. 4.2.1), lensless cameras are the thinnest cameras. Since their optics do not include lenses, they can be only a few mm thin (Asif et al. 2016). However, this absence of lenses reduces the light-efficiency of these cameras (Liang 2020; Boominathan et al. 2016, 2020). Hence, they use coded apertures and the raw images cannot be interpreted directly. They acquire multiple sub-images which need to be merged (Asif et al. 2016; Zomet and Nayar 2006; Schwarz et al. 2016), or images with a few measurements only (Huang et al. 2013; Nakamura et al. 2015, 2019). Therefore, the data captured need to be post-processed before obtaining an image (Qaisar et al. 2013), this procesing inducing delays (*e. g.* 100 ms in (Boominathan et al. 2016)). Recent approaches use neural networks to reconstruct images with higher quality, but they require even more time, up to 10 minutes (Liang 2020).

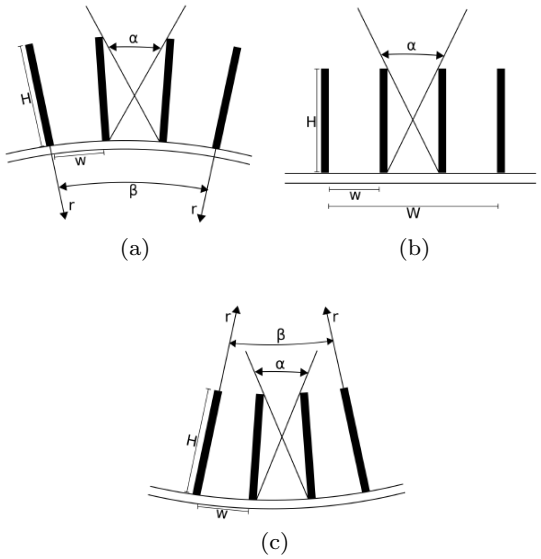


Fig. 8: The light rays captured for a concave (a), planar (b) and convex (c) curvature of the Söller Collimator. Reproduced from (Koppelhuber and Bimber 2017).

Because of these drawbacks, researches focus on designing static lensless cameras and developing the associated reconstruction algorithms. Indeed, lensless cameras are still emerging, and only a few of them can be considered adaptive. On the contrary, there are many adaptive dioptric cameras.

5 Dioptric cameras

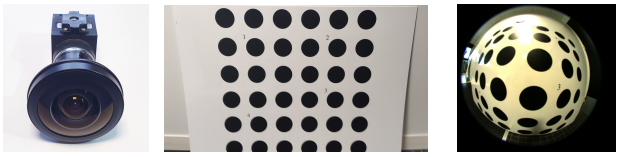
By using lenses, dioptric cameras (Def. 6), including conventional cameras (Def. 3), show a lot more diverse set of designs than lensless cameras (Sec. 4). Because of their huge number, they are classified below depending if their optics, image sensor or both of them are adaptive, manually (Sec. 5.2) or automatically (Sec. 5.3). But first, we review the non-adaptive dioptric cameras, focusing on the unconventional ones (Sec. 5.1) that inspire adaptive dioptric cameras.

5.1 Dioptric cameras with static properties

5.1.1 Unconventional optics

Recall that the FOV of a conventional camera is larger when its focal length is short (Sec. 2.5). For instance, a lens of $f = 10\ mm$, mounted on a $24\ mm \times 16\ mm$ image sensor (Zeiss® 2021), perceives a $107° \times 74°$ FOV.

In order to sense a wide FOV on the image plane, peripheral angles are compacted by *fisheye lenses* (Maître 2017). One distinguishes full-frame fisheyes, which expose the whole sensor but have a cropped FOV, with circular fisheye (Fig. 9) which project the whole panoramic FOV in a subpart of the sensor (Bettonvil 2005). Moreover, hyper fisheye lenses capture a FOV larger than 200°, *e.g.* 220° (Okumura 2017) and even 280° for the Entaniya280® (Entaniya® 2021). Whereas fisheye lenses available for consumers cannot capture a FOV wider than 280°, researches tend to increase this value *e.g.* the fisheye objective capturing a FOV of 310° (Martin 2004). Optically, fisheye lenses consist in an objective with a large diameter entry lens (Bettonvil 2005) (for instance, 77.5 *mm* for (Nikkor® 2023)) with an important curvature, and several lens groups. Nonetheless, such optics allows to



(a) (b) (c)

Fig. 9: A camera equipped with a fisheye lens (a), a calibration dot pattern (all dots are the same size) (b) and its image captured by the fisheye camera (c).

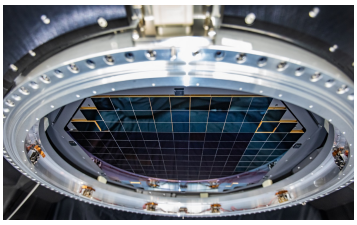


Fig. 10: The sensor surface of the future camera of the LSST (reproduced from (Standford 2020)).

sense a wide FOV at the price of important distortions. Therefore, the image resolution (10) is not uniform (Maître 2017). Surfaces of the scene with the same area, equally distant to the camera, occupy more pixels if they are imaged at the center than the borders of the panoramic image.

Another type of unconventional dioptric camera is the plenoptic camera[1] made with a microlens array to capture the incoming light field (Def. 2). Recall that pinhole-array based lensless cameras can capture several sub-images $i_k(\mathbf{X})$ of a scene region neighboring $\mathbf{X}$ (Sec. 4.2.1). The microlens array have a similar effect, capturing mutiple viewpoints as i_k of the same scene object (Levoy 2006; Ng et al. 2005). Then, the stack of subimages i_k can be merged to reconstruct depth (Jeon et al. 2015) or to perform *digital refocusing*, *i.e.* changing the plane of sharp focus after the light acquisition (Ng et al. 2005). Since each microlens in the array covers a fraction of the sensor, the resolution of the captured image is lower than for a conventional camera using the same number of photodetectors.

5.1.2 Unconventional image sensor

Apart from the optics, the image sensor of a dioptric camera can also differ, in dimension and shape, from a conventional camera.

To capture *gigapixel images*, *i.e.* with a definition (Sec. 2.1) of billions of pixels, several cameras use a wide sensor surface of tens of cm (Flaugher et al. 2015; McLeod et al. 2015; Kahn et al. 2010). The largest image sensor ever made consists in 189 individual CMOS sensors distributed within a circle of 634 mm diameter (Fig. 10) to capture images with a definition of 3.2 gigapixels (Kahn et al. 2010). The lenses of such large image sensor are obviously large, *e.g.* the camera of the Legacy Survey of Space and Time (LSST) will combine the 634 mm image sensor and three lenses, up to 1.57 meters in diameter (Lawrence Livermore National Laboratory 2019).

Very large or not, lenses imply a non-uniform illumination of the image plane because, physically, after passing through a lens, light rays are really focused on a curved surface, not a plane, because of the *Petzval field curvature*. Therefore, either the center or the borders of the image can appear sharp at once, on a planar sensor (Jung and Won 2020). Similar to the retina shape in the human eye, a curved image sensor matches the *Petzval surface* to capture sharp images. Combining specific lenses of opposite curvature may *flatten the field* but makes larger objectives which contribute to additional aberrations (Guenter et al. 2017).

One of the lens aberrations is chromatic optical aberrations, *i.e.* blurriness for several wavelengths in color images (Korneliussen and Hirakawa 2014). A curved image sensor with a concave surface reduces these aberrations to capture sharp images.

The *electronic camera eye* combines a focal surface of hemispherical shape with a transparent dome, a single lens placed at the top of the dome. This camera captures images with a wider and more uniform focus and intensities than planar image sensors (Ko et al. 2008). During the manufacturing process, an array of photodiodes is attached to a disk membrane made of silicone, which is deformed by ten independent paddle arms, to make an hemispherical focal surface (Sec. 2.2). Its radius of curvature is about 1 *cm* (corresponding approximately to 60° of curvature). The pixel density is severely reduced by the pixel pitch of about 1 *mm*, which constrains the image definition by 16×16 pixels. Practically, only square patterns, 75 mm away from the sensor, have been captured by the prototype, yet.

Curved image sensors are also capable of improving the brightness in the image corners. Conventional cameras suffer from the *vignetting effect*, the image brightness in the corner is severely reduced compared to the center (Zheng et al. 2008)[2]. Bending a CMOS sensor

[1] Two brands, Lytro® (Lightfield Forum® 2023) and Raytrix® (Raytrix® 2021) have provided light field cameras, until 2018 for Lytro®.

[2] the brightness loss can reach 90% in the case of Canon 50 mm f/1.2 lens mounted on a Canon 1DS Mark III body (Guenter et al. 2017).

dramatically reduces the vignetting effect, cancelling almost all losses, but the maximum curvature is limited (Guenter et al. 2017). Indeed, curving the sensor more than 36° would cause damages to its components. Therefore, the applicability of wide-angle lenses is limited, since the optimal curvature of the sensor depends on the lens. For instance, with a monocentric lens[3], the degree of curvature must equal the FOV to capture sharp images with a uniform illumination (Guenter et al. 2017). Some techniques, based on lithography and molding, allow to directly manufacture highly-curved image sensor surfaces, but they are still very expensive (Zhang et al. 2017).

5.1.3 Unconventional optics and image sensor together

Each of unconventional optics and unconventional image sensor impacts camera characteristics. But none of them is flawless. Combining them together has led to benefit of the advantages of both.

First, images captured by fisheye lenses are degraded by a high off-axis illumination (Martin 2004). Most light rays with a relatively large angle of incidence, *e.g.* 65° for a 310° FOV fisheye, wont pass through the aperture, hence the peripheral regions of the image will appear darker (Martin 2004). Curved image sensors compensate such illumination falloff. Indeed, a camera combining an hyper fisheye lens with a concave curved image sensor captures a 200° wide FOV almost exempt of chromatic aberrations and peripheral illumination losses (Lee et al. 2017). The latter combination improves the image quality but the panoramic images still feature distortions.

Whereas cameras with concave image sensors are inspired by mammal eyes, arthropod eyes have inspired the most recent convex artificial compound eyes (ACE) (Kim et al. 2020). ACE are cameras with multiple lenses and apertures which mimic the facets of arthropod eyes, the *ommatidiae* (Wu et al. 2017; Cheng et al. 2019).

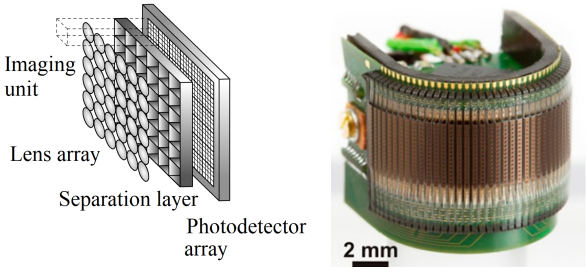
Remark 2 (Polycamera and Artificial Compound Eyes) Since ACEs use only one image sensor, not always CMOS or CCD but still relying on a single electronic board, they are considered as dioptric cameras in this survey. In (Cheng et al. 2019), polycameras (Def. 8) are classified as a subpart of ACEs. However, because they require one image sensor per individual camera, polycameras are reported in a different section (Sec. 7).

[3] A monocentric lens consists in spherical optical elements arranged concentrically around a central point. Its FOV can reach more than 120° (Ford et al. 2018).

The first ACEs, developed during the late 1990's and early 2000's, are compact planar cameras capturing multiple images of the scene simultaneously (Tan et al. 2004; Hamanaka and Koshi 1996). The most famous one, TOMBO (Fig. 11a), is made of a microlens array, a separation layer (a 120 μm thick array of square-shaped holes) and a planar image sensor (Tan et al. 2004). Closer to arthropod eyes, a curved ACE captures 180° × 60° FOV. The *curvACE* prototype consists in a convex image sensor, of 630 *ommatidiae* distributed in 42 arrays of 15 photodetectors, under a curved microlens array (Floreano et al. 2013) (Fig. 11b). This wide FOV is recorded as a 112×112 pixels image. The size of a photodetector prevents a higher definition for a compact ACE.

Curved ACEs may capture images with a *spatially variable field-of-view*, *i.e.* with different transverse magnifications (2) in the single captured image. Thus, the number of pixels that occupies an object in the image depends on its location in the scene. A trifocal ACE images scene objects with three different magnifications. It is made of a curved microlens array and a planar image sensor (Li et al. 2018). The microlens array is divided into three fan-shaped areas of 120°, each having its own focal length: 1.927 mm, 2.227 mm and 2.527 mm.

The latter spatially non-constant focal length leads to several different magnifications but always the same three. Hence, it is tailored for a scene where each considered object needs to be at an accurate distance of the camera. Otherwise, the camera should feature some degrees of freedom, *e.g.* variable focal length, either manually actuated (Sec. 5.2) or automatically controlled (Sec. 5.3).



(a) (b)

Fig. 11: (a) TOMBO architecture (reproduced from (Chan et al. 2006)), (b) CurvACE prototype (reproduced from (Floreano et al. 2013)).

5.2 Manual adaptations

In this section, dioptric cameras are classified depending if their optics, image sensor or both vary temporally from direct manual input.

In regards with adaptive optics (Sec. 5.2.1), adaptations of conventional cameras are described before less common optical adaptations. Then, deformable or moving image sensors behind static optics are discussed (Sec. 5.2.2). Finally, this section reports dioptric cameras with both adaptive optics and sensor, including manually deformable ACEs (Sec. 5.2.3).

5.2.1 Adaptive optics

5.2.1.1 Moving lenses

Adaptive FOV: As implied by (2), the transverse magnification G increases as the focal length f of a lens does, allowing to *zoom in* elements of a scene. The zoom can be approximated by a homothety for planar objects in frontal position inside the plane of sharp focus (Maître 2017). By moving the elements of an afocal system, *i.e.* a divergent lens between two convergent ones, the focal length of the camera changes while keeping (approximately) the same scene elements in focus (Fig. 12) (Maître 2017). Lens movements inside the camera objective are often non-linear, therefore (Lenk et al. 2019) proposes to replace one of the middle lens by a *tunable lens*, which can change focal length without moving, to simplify the movements of the other lenses while zooming. Tunable lenses, detailed in a survey on biosinspired cameras (Kim et al. 2020) and a survey on varifocal lenses (Kang et al. 2020), are mostly made of deformable membranes, actuated thanks to mechanical or electronic elements, or the variation of pressure of air or liquids.

Objectives which can change focal length in a *zoom range* $[f_{min}, f_{max}]$ are characterized by their *zoom ratio* f_{max}/f_{min} (Clark and Wright 1973), also called *zoom factor* (Demenikov et al. 2009). On available products, the zoom factor is up to several tens [4].

Zooming lenses of conventional cameras are bulky since they involve a mechanism which moves the lenses. Replacing the movements of lenses by tunable lenses, made of liquid crystal (LC), minimizes the size of the whole objective to about 10 *cm* (Lin et al. 2011). Such lenses change focal length thanks to intrinsic deformations, *i.e.* the orientation of crystals inside LC lenses. However, one only reaches a zoom ratio up to 7.9, when the object is only 10 *cm* away from the lenses, and decreases as the object goes farther.

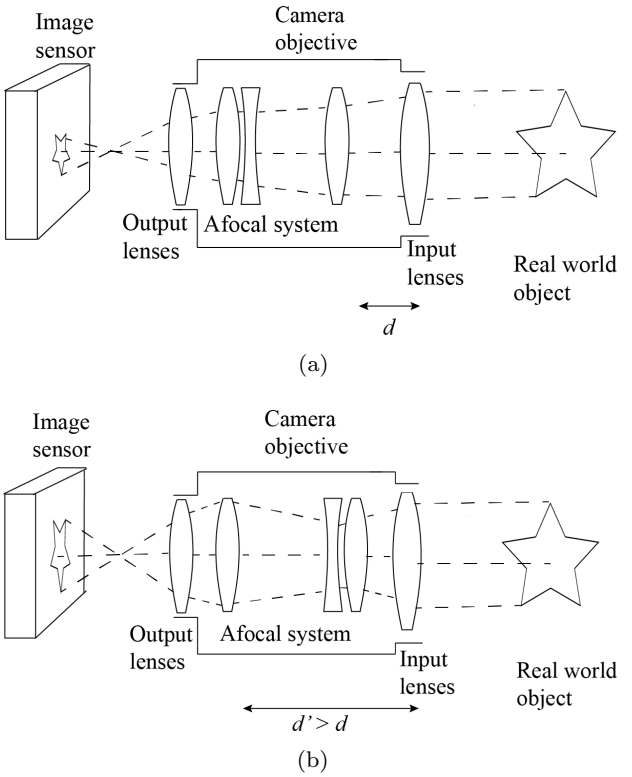


Fig. 12: Principle of the zoom: an afocal system, *i.e.* a divergent lens between two converging ones move to change the magnification of objects. Example of small (a) and large (b) magnifications.

Electronically deformable liquid lenses, between two solid lenses, are another intrinsic lens adaptation. With three of them, a microscope zooming objective has a continuous zoom change in the range $[7.8, 13.2]$ mm (Li et al. 2016), contrary to conventional microscopes, which switch between a few preset values of zoom.

It behaves similarly as the eye accommodation, where the eye lens modifies its own shape to *focus* on something, rounding to focus at near object and flattening for further ones (Allen and Triantaphillidou 2012).

Focusing: When focusing on an object point ${}^c\mathbf{P}$, all the objects within the two extreme plane of focus Z^+ and Z^- appear sharp (Sec. 2.4). Focusing satisfies the lens law (1), *i.e.* for a fixed focal length f, the depth $Z = Z_i$ of the image point perfectly in focus depends on the depth $Z = Z_o$ of ${}^c\mathbf{P}$ (Allen and Triantaphillidou 2012). Therefore, the camera is focused at infinity ($Z_0 = \infty$) when the distance between the camera optical center (main lens) and the image sensor is $Z_i = f$.

[4] 60 on the Samsung Digital Camera WB2200® (Samsung® 2023) or even 125 for Nikon COOLPIX P1000® (Nikon® 2022).

As the objects get nearer the camera, Z_o decreases so Z_i must be increased by a *focusing extension* Z_e to image them sharply. Practically, focusing a conventional camera objective is achieved by the movement of all its lenses as a single unit which directly increases Z_i (Allen and Triantaphillidou 2012) (Fig. 13).

Even though harder to achieve, alternative methods allow focusing (Allen and Triantaphillidou 2012). Moving the front lens changes the focal length instead of the lens-to-image-plane distance, which also changes the elements in focus. Adding a lens in front of an objective completes the previous methods to focus on scene objects of less than 1 meter distant to the camera. With *internal focusing*, an internal optical group moves between the front and the back lenses to modify f instead of Z_i without changing the external objective shape[5].

Furthermore, the depth of field Δ can also be modified by changing the diaphragm aperture D (Sec. 2.4). The smaller the aperture, the higher the value of the depth of field.

Adaptive viewpoint and focus: Whereas those methods accurately set the depth of the two extreme planes in focus, moving the objective (all optical blocks) changes the orientation of the plane of sharp focus and the image position on five different ways, w.r.t. a fixed image sensor (Herbert 1960). To avoid bulky motors, each movement is possible with miniaturized designs as Perspective-Control (PC) adapter (Phillips et al. 1984). This shifts the camera objective up and down or left and right, and also rotates around the camera optical axis, while keeping the objective and the image planes parallel. The circle of illumination (Sec. 2.2), which is bigger than the sensor, is thus moved relatively to the image sensor (Langford 2000). The image is then virtually translated, or rotated, by approximately the same distance as the lens actually is (Stroebel 1999). Therefore, shifting the camera lens acts exactly like moving the image sensor inside an image circle with a wide diameter. According to the projection model (3), a movement $\delta\boldsymbol{x} = \{\delta_x, \delta_y\} \in \mathbb{R}^2$ inside the image plane corresponds to a movement[6] $\delta\boldsymbol{X} = \frac{\delta\boldsymbol{x}\, Z}{f} \in \mathbb{R}^2$ of the camera as a whole.

[5] A lot of camera objectives, like Nikkor Z 58 mm for Nikon® cameras (Nikon® 2019) or FE C 16-35 mm T3.1 from Sony® (Sony® 2023), have rings to manually set the distance of the two extreme planes of sharp focus Z^+ and Z^- (in meters or in feet), but not independently.

[6] With a Nikon 35mm PC Nikkor® ($f = 35\ mm$) (Nikkor® 2022) for a 6 m distant scene object, a 2 mm lens translation moves the viewed object of 36 mm and a 10 mm translation up to 1.78 m (Phillips et al. 1984).

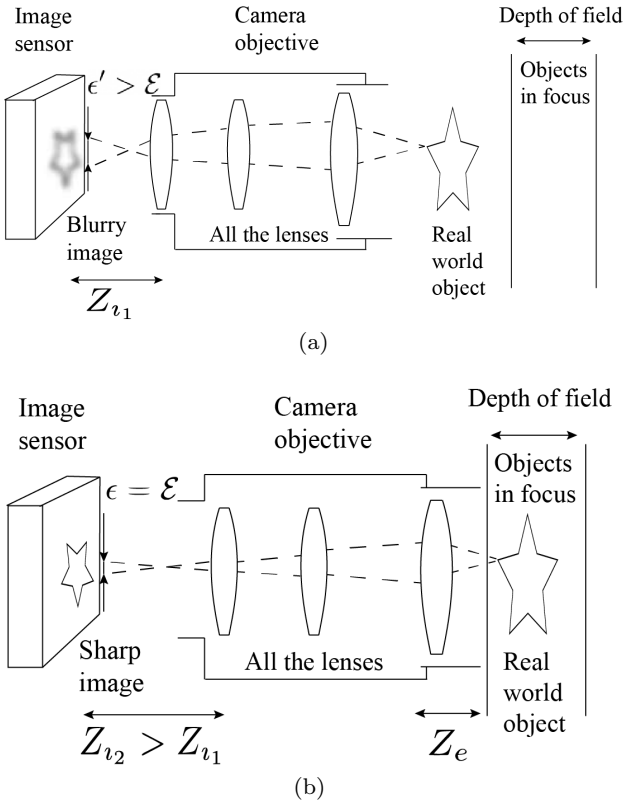


Fig. 13: The objects are within the depth of field δ far from the camera if Z_i is close to f (a). Extending Z_i by a focusing extension Z_e allows to image near objects (b).

Optical adaptations can also make disjoint blurry and sharp areas in an image, by the mean of rotating the plane of sharp focus. For instance, one could have sharp objects both in the foreground and the background, and all the surrounding objects blurry. Rotating the objective of $\boldsymbol{\theta} = \{\theta_x, \theta_y\} \in \mathbb{R}^2$ (in rads) around the vertical and horizontal axes is called *swinging* and *tilting* (Stroebel 1999). Hence, the orientation $\boldsymbol{\psi} = \{\psi_x, \psi_y\} \in \mathbb{R}^2$ (in rads) of the plane of focus (Sec. 2.4) is computed with the Scheimpflung principle: the image plane, the lens plane and the plane of focus must intersect along one line (Scheimpflug 1904). $\boldsymbol{\theta}$ and $\boldsymbol{\psi}$ are related by:

$$\tan\psi_x = \frac{Z_i}{Z_i \cos\theta_x - f} \sin\theta_x \tag{12}$$

$$\tan\psi_y = \frac{Z_i}{Z_i \cos\theta_y - f} \sin\theta_y \tag{13}$$

with Z_i the distance between the optical center of the camera and the image plane. These two movements make the image and lens planes not parallel anymore.

Therefore, the two extreme planes of focus are not parallel either. Indeed, the distance between them, *i.e.* the depth of field (Sec.2.4), increases as the depth of the scene elements does (Merklinger 1996). A 10 cm length objective combining a tilting part of three lenses and a fixed part indeed rotates the plane of focus (Nanjo and Sueyoshi 2011).

Tilt-shift lenses consist in adapters which can be mounted on a conventional camera objective (Scholz et al. 2014). These lenses enable to both change the orientation of the plane of sharp focus and move the camera viewpoint, thanks to the tilt (Fig. 14a) and the shift (Fig. 14b) described above. A magnetic position sensor can be added to locate the tilt-shift lens with respect to the camera image sensor (Poulsen 2011). On post-processing, the data provided by this sensor are used to correct the geometric optical aberrations on the recorded images[7].

(a) (b)

Fig. 14: Nikkor PC 19mm f/4E ED for Nikon®objectives (Nikon® 2023): in tilt (a) and shift (b) positions. Reproduced from (Nikon® 2020).

5.2.1.2 Deforming lenses

In summary, the commonly known adaptations of conventional cameras act globally on the camera properties. The zoom, indeed, modifies directly the focal length. Hence, it narrows the field of view as it magnifies objects. Reversely, zooming out increases the camera FOV at the price of reducing the size of scene elements of interest in the image.

Adaptive FOV: Increasing the FOV even more than zooming out, while reducing less the viewed objects size, is made possible by microlens arrays. Used in plenoptic cameras (Sec. 5.1.1) to multiply the viewpoints, microlens arrays are suitable for cameras of adaptive FOV.

A *sheet camera* consists in a microlens array combined with a flexible image sensor, which could be, in theory, attached to a curved surface of any shape to capture any FOV, assuming that the scene surface is curved with the same center of curvature than the array (Sims et al. 2016). Such camera does not exist yet, and only the optical part has been fully studied and optimized. A microlens array increases the FOV of a conventional camera while keeping the scene elements the same size. In that case, the conventional camera must be facing the microlens array. Its FOV increases as the curvature of the microlens array does, but if the individual lenses remain rigid the FOV becomes discontinuous, under-sampled. Indeed, the gaps between two consecutive lenses increase with the curvature, so one should use deformable lenses to counteract this effect and capture images with a continuous FOV (Sims et al. 2016). Therefore, optics of sheet cameras consist in a silicon-molded microlens array, flexible and fully transparent. Actually, as the curvature $\kappa \in \mathbb{R}_+$ of the sheet increases, the focal length of each individual lens changes to ensure a local FOV of $\omega\kappa$, with $\omega \in \mathbb{R}_+^*$ the pitch between two consecutive microlenses. On simulation, a 23 mm thick *sheet camera* wrapped around a cylinder of radius $r = 7.5$ mm, with a 7 mm lens pitch, can capture a panoramic FOV of 180°, either vertical or horizontal, depending on the direction of the deformation. On the prototype of (Sims et al. 2016), the camera is placed several tens of centimeters away from the microlens array (Fig. 16a). Hence, the camera FOV varies between 10° and 52°. Furthermore, the array is only 33×33 microlenses, so the acquired images are 33×33 dots and the pixels between them must be interpolated by post-processing to generate images with higher definition (see an example in Fig. 15). Yet, this prototype has not been used to image real scenes. A LCD screen displaying specific scenes is placed in front of the microlens array.

Instead of being flexed, the microlens array can be stretched to zoom-in. This array is even thinner, only 2 mm thick (Sims et al. 2018), than the microscope zoom (Li et al. 2016). Stretching it by 1 mm already increases the magnification of the whole scene, zooming-in globally.

As in (Sims et al. 2016), the microlens array is facing a conventional camera to change its properties. Stretching the microlens array by $\Delta\rho \in \mathbb{R}_+^*$ changes its focal length f:

$$f(\Delta\rho) = \frac{1}{\frac{\eta-1}{R(\Delta\rho)} - \frac{\eta}{T(\Delta\rho)}} \tag{14}$$

[7] For several thousands of US dollars, one can buy tilt-shift lenses such as the TS-E models by Canon® (Canon® 2022) or the Nikkor PC 19mm f/4E ED for Nikon® (Nikon® 2023).

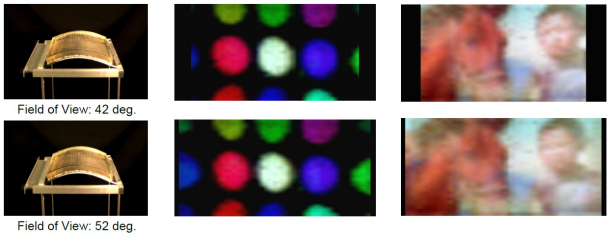


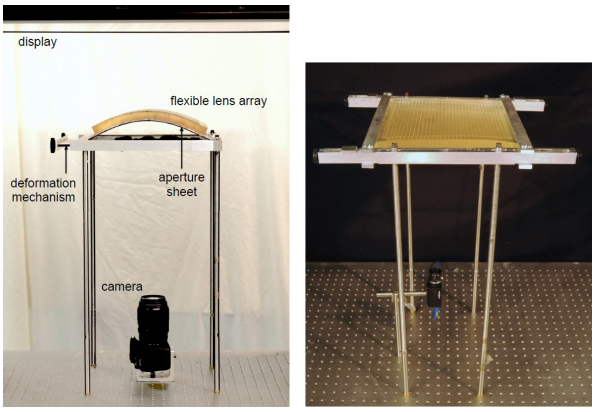
Fig. 15: Two different curvatures of the microlens array (on the left), dot images captured (middle), reconstructed images thanks to interpolation algorithms (on the right). Reproduced from (Sims et al. 2016).

with $\eta \in \mathbb{R}_+^*$ the index of refraction of the lenses material, $R(\Delta\rho) \in \mathbb{R}_+^*$ and $T(\Delta\rho) \in \mathbb{R}_+^*$ the lenses radius and thickness (see details in (Sims et al. 2018)). The prototype, named *Stretchcam* combines a conventional camera with a 80×80 mm microlens array (Fig. 16b). Stretched by 3% of its original length, it magnifies 300 mm distant scene elements by $G = 1.5$. This magnification is far below the zooming factor of several tens of the available digital cameras, but it requires much less space. Furthermore, *Stretchcam* acquires low-resolution images of 33×33 bright spots.

Adaptive FOV and focus: Also relying on a microlens array, a camera enables both adaptive FOV imaging and manual focusing (Cao et al. 2020). Its first adaptation increases dynamically up to 180°, a lot more than Flexcam and Stretchcam. This camera is made of a flexible microlens array, of about 1200 lenses, covering a microfluidic liquid chamber. Light rays pass through the lens array and the chamber before being refocused by an objective lens onto a CCD (Sec. 2.6). The lens array is the focal plane (Sec. 2.2) of the camera. As the chamber is filled by liquid, the focal plane gets curved to become an hemispherical shape of 2 mm in diameter. Then, the FOV increases as:

$$\text{FOV} = 2 \arcsin\left(\frac{2rh}{r^2 + h^2}\right) \tag{15}$$

with $r \in \mathbb{R}_+^*$ the radius of the hemispherical base and $h \in \mathbb{R}_+^*$ the height of the hemispherical base (when $r = h$, the FOV is panoramic). The curvature of the microlens array is concave with respect to the CCD. All the objects on the hemisphere in front of the camera (*i.e.* the whole optical system) are captured. Although this camera can dramatically change the FOV, no experiments to capture real scenes have been done, yet. Instead, to prove that this camera is able to perceive a wide FOV, the experiments of (Cao et al. 2020) place a laser, successively at 0°, 30° and 60° incident angles with the hemispherical microlens array. For these three



(a) (b)

Fig. 16: Prototypes of the flexible sheet camera (a) (reproduced from (Sims et al. 2016)) and Stretchcam (reproduced from (Sims et al. 2018)).

incident angles, the laser intercepts the focal plane of this dioptric camera. Since this camera can perceive incident rays at more than 60° of incident angle, the FOV of this camera is proven to be at least $2 \times 60° = 120°$.

This camera features a second adaptation, completely different from adaptive FOV imaging. For a fixed curved surface, this camera is unable to capture objects located at different depths sharply (Cao et al. 2020). Therefore, to distinguish target objects at different distances one by one, it also performs manual refocusing. As the liquid chamber changes volume, the focal length f of the camera changes from ∞ to 3.03 mm. Like in conventional focusing, modifying f changes the locations of the planes of sharp focus Z^+ and Z^- (Sec. 2.4). To demonstrate this focusing ability, patterns are projected onto the microlens array thanks to a red light and a shaped mask. Then, they lead to the CCD for imaging. The two masks are opaque with transparent "K" and "S" shapes, respectively about $1\ mm \times 1\ mm$ and $2\ mm \times 2\ mm$. The two masks are placed at different distances away from the microlens array, the one with "K" shapes at $1833\mu m$ and the one with "S" shapes at $1165\mu m$. As the microlens array increases its curvature, the "K" and "S" shapes form clear images successively.

Although promising a lot of interesting features, experiments made on cameras with microlens arrays are completely controlled. They capture projected dots or images displayed by LCD screens rather than real scenes. Also, all these cameras achieve global and uniform adaptations. For instance, they cannot perform spatially variable zoom. Furthermore, because they use mechanical systems to flex or stretch the lenses, or piping to fill a liquid chamber, the current

prototypes require several tens of centimeters of space.

5.2.1.3 Adaptive aperture: focusing

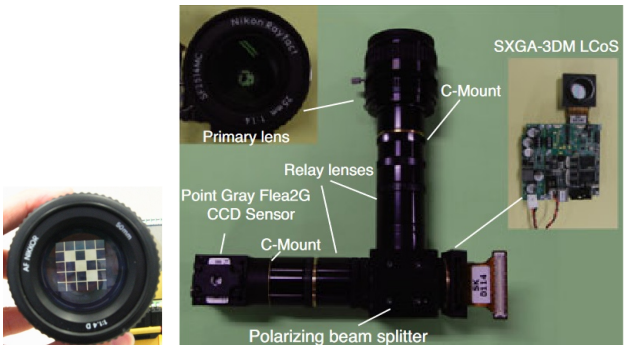
Recall that the optics of dioptric cameras are made of lenses and a diaphragm (Def. 3). After cameras with translating lenses, this section reports dioptric cameras with changeable apertures, similar than the ones in adaptive lensless cameras (Sec. 4.2).

In order to modify the shape of the aperture, one could set an array of opaque and transparent rectangles in front of the lenses of the camera. To do so, *pupil plane coding* puts an optical mask on the entry lens of a conventional camera (Zhou and Nayar 2011), thus changing the aperture shape exactly like the lensless cameras with coded aperture (Sec. 4.1). Apertures made of one transparent rectangle among a set of 5×5 (or 7×7) opaque ones capture a part of the incoming light rays on a small specified region (Liang et al. 2008). A sequential acquisition of $N \in \mathbb{N}^*$ images, during a single exposure time, captures the light field. N is then the number of viewpoints captured in the sequence, which defines the angular resolution of the light field, up to 4×4 in (Liang et al. 2008). As N increases, the number of refocused images with a different depth of field does. However, the image is noisy since most of the mask is closed and the acquisition is made during a fraction $1/N$ of the total exposure time. By opening several transparent rectangles inside the pattern (Fig. 17a), the acquisition of multiple views is *multiplexed*, *i.e.* light rays coming from several different angles are captured at once. To eliminate all the noise artifacts, N acquisitions of multiplexed patterns are required. Since the acquisition is sequential, it requires the scene to be static during the exposure time (160 ms), constraining the use to dynamic-less scenes. Then, a *demultiplexing* post-process reconstructs the light field image.

The sharpness of the refocused images depends on the number of viewpoints captured in a sequence. For $N = 3 \times 3$ viewpoints, aliasing makes all the objects in the image blurry, but for $N = 4 \times 4$ viewpoints, the objects in focus appear sharp. This system can be implemented as a scroll of paper patterns or a liquid crystal sheet or a liquid crystal array directly placed inside a camera objective. The light field images need to be down-sampled from 3008×2000 to 640×426 to save computational cost (Liang et al. 2008), but their capture still require 3 to 5 seconds.

By still capturing several viewpoints, each corresponding to a specific aperture shape, at once, cameras with changeable apertures also make *defocus deblurring* possible. *Defocus deblurring* consists in reducing the blurry areas, associated to object points out of focus, in the images (Zhou and Nayar 2009). The *programmable aperture camera* uses a LCoS (Sec. 4.2) device to implement a modifiable aperture (Nagahara et al. 2010). The LCoS can change the state (opaque or transparent) of any of its 1280×1024 squares at 5 kHz. On the one hand, using the same multiplexing method than (Liang et al. 2008), the programmable aperture camera of (Nagahara et al. 2010) captures 31 multiplexed images in 1.2 seconds. Hence, this camera captures light field images with a definition of 1280×960 pixels which combine 31 different views. On the other hand, for defocus deblurring, the shape of the aperture of the programmable aperture camera of (Nagahara et al. 2010) is optimized by a criterion which considers the noise level in the image (Zhou et al. 2009). Then, defocus deblurring successfully makes sharp previously defocused images.

Unlike (Liang et al. 2008), which inserts a mask inside the camera objective (Fig. 17a), the prototype of (Nagahara et al. 2010) is more complex (Fig. 17b). It consists in two objectives forming a 90° angle. Each of them, about 10 cm long, contains a relay lens. A primary lens drives light rays into the first objective, then, they are led to the second one thanks to a beam splitter. Finally, the last relay lens conducts the rays to the CCD sensor (Sec. 2.6). The lenses induce vignetting and field curvature, which impact the captured image. Especially, the light efficiency is low, hence the captured images are dark.



(a) (b)

Fig. 17: (a) Mask inserted into a camera objective (reproduced from (Liang et al. 2008)) and (b) the prototype of the programmable aperture camera (reproduced from (Nagahara et al. 2010)).

5.2.2 Adaptive image sensor

An ideal *Sheet camera*, *i.e.* compatible with any surface, requires a fully deformable image sensor (Sims et al. 2016). Such image sensor does not exist yet, but is targeted by several works.

5.2.2.1 Flexible image sensor: toward adaptive focus and FOV

As proven by the manufacturing process of the curved image sensors (Sec. 5.1.2), photodiode arrays can be integrated on flexible or curved substrates. The first organic photodiodes, made of polymers, have been released in 1998 (Yu et al. 1998). Aligned on a row, they could not capture 2D images directly. Instead, they acquired 1D images for different heights successively. Such organic photodiodes are the base of two image sensors: a 5 $cm \times 5$ cm one and a 5 $cm \times 3.8$ cm one, which can be both flexed up to 90° of curvature (Someya et al. 2005; Ng et al. 2008).

Circuits embedded on thin fabric or deformable silicone (PDMS), hundreds of μm thick, have been researched for more than 10 years (Kim et al. 2009). When they embed photodiodes, they are mostly used as stretchable photovoltaics panels (Lee et al. 2011; Zhao et al. 2019) instead of bendable image sensors. Only few cameras rely on such sensor, and since they also use deformable optics, they are discussed in Section 5.2.3.

5.2.2.2 Adaptive image sensor: moving the sensor

Focusing: Still emerging, flexible image sensors are not in adaptive dioptric cameras yet. Instead, they change the way they capture images by moving a conventional image sensor. A single axis translation of the image sensor allows to manipulate the depth of field accurately, on several ways (Kuthirummal et al. 2010). First, the plane of sharp focus location is directly related with the image sensor position along the camera optical axis ${}^c\vec{\mathbf{Z}}$. Indeed, the diameter of the circle of confusion ϵ (Sec 2.4) is expressed as:

$$\epsilon = \frac{D}{Z_i}|Z_i - \delta Z_s|, \tag{16}$$

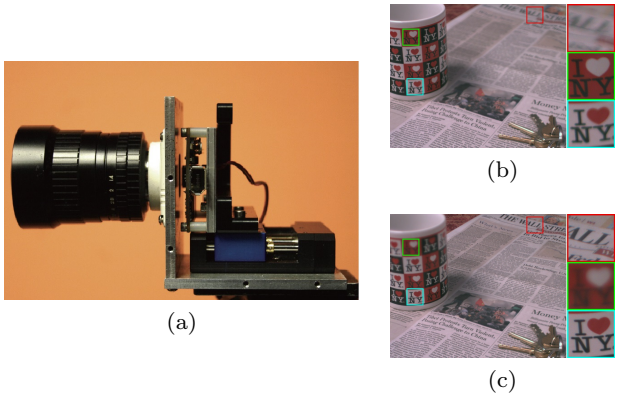
with $D \in \mathbb{R}_+^*$ the aperture diameter, $Z_i \in \mathbb{R}$ the sharp image points Z coordinate (Sec. 2.3) and $\delta Z_s \in \mathbb{R}_+^*$ the sensor shift. Results show that only a few hundreds of μm sensor displacements can change the scene depth ranges in sharp focus. Indeed, with focal length $f = 9.0$ mm, the focused depths range is $[1\ m, \infty]$ for a 81.5 μm translation and $[0.2\ m, 0.5\ m]$ for a 259.2 μm translation. Second, it is possible to capture a whole *focal stack* during the camera exposure time to effectively capture images with all scene elements in focus. This is made with a uniform translation, *i.e* with a constant speed, of the image sensor.

Moreover, this flexible depth of field camera records images with a *discontinuous depth of field* (Kuthirummal et al. 2010). A non-uniform sensor translation, where the sensor moves at speed v_1 in the foreground and the background and a speed $v_2 >> v_1$ in between, achieves two ranges of planes in focus in the image, one in the foreground and one in the background, and all the objects in between appearing blurry.

Finally, this camera also emulates the tilted plane of sharp focus of a tilt-shift lens thanks to a uniform sensor translation with an electronic rolling shutter (ERS). The Figures 18b and 18c compare an image taken with a planar, conventional, and tilted planes of sharp focus. The rows of photodetectors are enabled one by one so that a different scene height Y is captured in sharp focus for each depth Z. At last, the volume of the scene in focus can also have a curved shape, by combining a non-uniform translation and an ERS.

Practically, the one-axis translation is implemented by a micro-actuator with top speed of 2.7 mm/sec, which is very fast compared to the manual movements of tilt-shift lenses (Kuthirummal et al. 2010). The moving sensor is just placed behind a static camera objective (Fig. 18a), outside of a camera housing. Even though being fast, the translation motion constrains the exposure time to $1/3^{rd}$ of second.

Adaptive image definition and FOV: Translating the image sensor can also increase the image definition and both the vertical and horizontal FOV. Then, several individual sub-images, or *tiles*, are merged in a



(a) (b) (c)

Fig. 18: The camera with a flexible depth of field prototype (a), a scene captured with a conventional camera (b), the same scene captured with a plane of sharp focus tilted by 53° (c). Reproduced from (Kuthirummal and Nayar 2007).

gigapixel image (Sec. 5.1.1). Indeed, a camera captures high-quality images, without any aberration caused by conventional lenses, up to 65000 × 25000 pixels and 104° × 40° FOV (Cossairt et al. 2011). The prototype consists in a 5 MegaPixels image sensor sequentially moved by a pan-tilt motor around the frontal hemisphere of a static ball lens. Moving this sensor in 14×14 different positions creates a virtual image surface of 14×14 *tiles*. The sensor is moved onto a $75\ mm \times 50\ mm$ rectangle to assemble the different tiles thanks to small overlapping FOVs.

Fusing these tiles, like on a puzzle, generates a 60° × 40° FOV gigapixel image. The speed of the motor is not given in (Cossairt et al. 2011), but the sequential image acquisition requires at least one exposure time by individual tile, so it cannot be done instantaneously. Additionally, the bigger the desired image, the slower the acquisition.

Other adaptive designs have a moving image sensor, but since they are fully automated, they are discussed in Section 5.3.

5.2.3 Adaptive optics and image sensor: deforming shapes

In the literature, there is no camera with a translating or rotating image sensor combined with adaptive optics. Closer to the *sheet camera* concept, the designs reported here modify the shape of both optics and sensor parts of a dioptric camera.

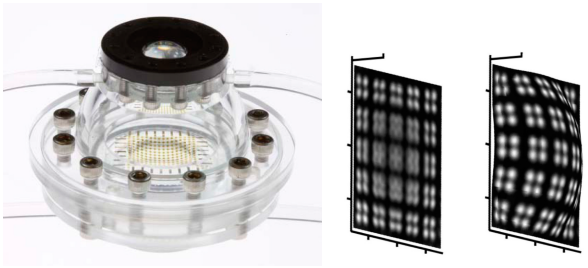
Focusing: Curved image sensors capture sharper images with a more uniform illumination than planar ones (Guenter et al. 2017; Ishihara 2015; Ko et al. 2008). However, the Petzval surface, where all the sharp image points lie (Sec. 5.1.2), changes with the magnification of a camera. The image sensor needs to change curvature accordingly to avoid significant image quality decrease (Jung et al. 2011). Then, the shape of the sensor matches the location of the image points in sharp focus, computed with the thin lens equation (1). To do so, (Jung et al. 2011) proposes a photodiode array attached to a flexible PDMS membrane (as in Sec. 5.2.2), above a water-filled cavity. As water is pumped in the cavity, the radius of curvature R_D of this adaptive image sensor changes:

$$R_D = \frac{d_{ph}^2 + 4\ h_p^2}{8\ h_p} \tag{17}$$

with $d_{ph} \in \mathbb{R}_+^*$ the diameter of the support of the photodiode array (49 mm on the prototype of (Jung et al. 2011)) and $h_p \in \mathbb{R}_+^*$ the peak deflection of the center diode (function of the pressure Pr). The camera of (Jung et al. 2011) associates this sensor with a tunable lens (Fig. 19a). The radius of curvature R_L of this fluid-filled lens increases as the fluid pressure does. When R_L increases, it simulates a zoom, as the focal length of the camera increases and the magnification as well. Then, to match the appropriate Petzval surface, R_D must increase. No general formula is given, but, in the extreme cases reported in the experiments, if R_L is set at 4.9 mm and 11.5 mm, R_D must be changed to 11.4 mm and 25.7 mm. These experiments have been made on dot patterns (Fig. 19b) placed 75 mm in front of the flexible lens. Due to the complexity of the fluid variation mechanism and the required electrodes connection, the prototype requires about ten centimeters of depth, width and height.

Instead of photodiodes, adaptive image sensors also rely on luminescent concentrators. They transport light rays of defined wavelength on their edges to capture images (Koppelhuber and Bimber 2013).

Recall that a lensless camera with an adaptive FOV (Sec. 4.2) works with a Söller collimator, a micro-aperture array, and a luminescent concentrator (Koppelhuber and Bimber 2017). To increase the amount of light getting to the concentrator, the collimator can be replaced by a Microlens Aperture Array (MLAA), *i.e.* an array of hexagonal apertures, each of them embedding a lens (Kurmi et al. 2018). Unlike (Koppelhuber and Bimber 2017), the dioptric camera with the MLAA highlights its focusing ability instead of an adaptive FOV (Kurmi et al. 2018). As the curvature of the MLAA gets more convex, the depth of field increases but less light rays are led to the sensor. Making it more concave does the opposite. As stated by (11), the sets of light rays entering the Söller collimator grow as the curvature does. Adding



(a) (b)

Fig. 19: Prototype of a camera with a flexible lens and adaptive image sensor (a), 3D representation of images captured with the sensor in a planar form, on the left, or concave form, on the right (b). Reproduced from (Jung et al. 2011).

lenses emphazises this property, the camera with MLAA capturing up to four times more light rays than the lensless one. Although capturing more light than (Koppelhuber and Bimber 2017), the camera of (Kurmi et al. 2018) needs to reconstruct the image by post-processing.

Adaptive FOV and focus: Alternately, deformable ACEs (Sec. 5.1.3) also are adaptive cameras with optics and sensor changing of curvature. An ACE with two configurations is able to capture images either with a wide FOV of more than 100° or a uniform sharpness (Zhang et al. 2017). It is made of 676 optoelectronic modules, *i.e.* a microlens and a photodetector, which are hexagons of 113 μm diameter distributed in a truncated icosahedron. Inspired by origamis, this structure is very flexible and can take a lot of other shapes, such as a ball, but only the truncated icosahedron had been tested for imaging yet. This assembly is attached to a silicon membrane which can change curvature.

Contrary to aforementioned designs, the prototype has not been tested on a whole range of dynamically varying curvatures, but only a single convex (like static ACEs) curvature of 2.27 mm and two concave curvatures (like curved image sensor in Sec. 5.1.2) of 2.27 mm and 7.20 mm, all of hemispherical shapes. During the experiments, a plano-convex lens, of 10 mm in diameter and in focal length, is put above the hemispherical photodiode array. The convex curvature allows to view a laser beam oriented 36° to the plane of the photodiodes support, therefore it has a FOV of $180° - 2 \times 36° = 108°$. The concave curvature of 2.27 mm captures sharp scene objects whereas the highest curvature, not optimized for the truncated icosahedron shape, captures blurry images. In experiments, the prototype captures a projected pattern of a single "W" shape. Moreover, there are two prerequisites to obtain sharp images, not spoiled by multiple dark spots caused by defective pixels. First, the plano-convex lens must be set at specific distances depending on the radius of curvature of the silicon membrane, 10.3 cm for the 2.27 mm curvature and 20 cm for the 7.20 mm curvature. Second, the camera must record 5 images of the pattern, with orientation from 0° to 60° with 12° increments. Finally, the 5 images can be merged into a sharp one.

To capture a set of multiple FOVs, ACEs should be able to take a lot more than 3 curvatures. Even though not fully functional yet, a bendable ACE with an adaptive FOV goes in that direction (Saito et al. 2005). Each of its 15 *ommatidiae* consists in a microlens, a transparent tube and an organic photodiode. Two configurations have been tested on the prototype, the ommatidiae successively placed on two cylinders of radius 40 mm and 80 mm. The ommatidiae, distributed on a single line, need to be moved on the cylinder height axis to capture 2D images. On the 40 mm-cylinder-radius setup, the camera has a 37° FOV, which is doubled in the 80 mm-cylinder-radius setup. This adaptation is global, hence the scene objects appear twice smaller in the second setup.

5.3 Automatic adaptations

Contrary to most works on manually deformable cameras that get close to a fully-flexible *sheet camera*, all the automatic adaptive dioptric cameras move the optics with respect to a fixed image sensor or the other way around, apart from one using a modifiable aperture.

5.3.1 Adaptive optics: moving lenses

This section starts with usual automatic adaptation, found in conventional cameras, toward the most original ones.

Focusing: First, *autofocus* automatically sets the regions captured sharp in an image. Just like manual focus (Sec. 5.1.1), lenses inside the camera objective moves to change the location of the plane of sharp focus. Contrary to the manual case, lenses are not directly moved by the user thanks to a ring or a button. First, the user selects an image area desired to be sharp, where points should be in focus, called here the *targeted area* (Maître 2017). Then, a rangefinder determines the distance between the objects in this area and the camera. (Maître 2017) reports several kinds of rangefinders:

- *stigmometers*, where prisms split the targeted area in two sub-images with a lateral shift. The smaller the shift, the sharper is the area ;
- active telemetry, which measure the travel time of emitted waves, *e.g* acoustic or infrared rays ;
- maximum contrast detectors, which compare the sharpness of the targeted area for two different focal lengths, are implicit rangefinders. The higher the contrast, the sharper are the scene objects ;
- phase detectors, using an internal mirror to lead light rays towards two telemetry sensors.

Finally, a motor moves the lenses relatively to the image sensor to shift the depth of field (6), in order to set the two extreme planes of focus Z^+ and Z^- (Sec. 2.4) so that all the objects to image sharp are located in-between. Then, this motor can be controlled in closed-loop (Fig. 20) by the distance measured by the rangefinder.

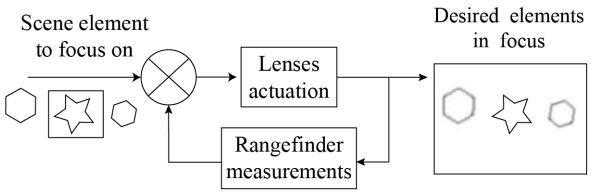


Fig. 20: Principle of autofocus.

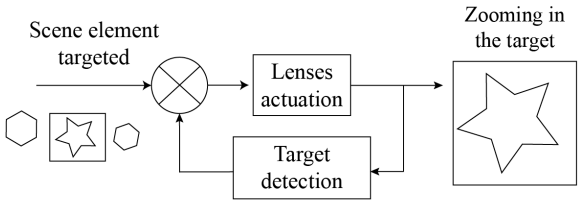


Fig. 21: Principle of automatic zooming in closed-loop.

Nowadays, most digital cameras feature an autofocus mode[8]. Whereas being very practical for the user, autofocus fails in several cases. For passive rangefinders based on contrasts, too dark or uniform areas make the lenses perpetually moving, looking for the optimal position. Also, objects moving fast (*e.g* insect wings), translated mirrors and waves in motion can disturb the optical measurements, and therefore, make the autofocus harder to achieve (Maître 2017).

Adaptive FOV: As the manual focus has its automated counterpart, zooming has *zoom tracking* or *autozoom*. This automated zoom tracks an object of the scene and ensures that it remains the same size in the image (Fig. 21) (Fayman et al. 1998). The focal length is closed-loop controlled, originally thanks to the data of a rangefinder or the estimated optical flow (Fayman et al. 1998). Both methods detect a moving object to directly control motors inside the camera objective. However, because of the perspective distortion, zoom tracking fails to make objects lying on different depths remaining the same size during their whole motion.

Since first researches in the 1990's, the control law has been improved to keep the element of interest, which is zoomed-in, in sharp focus (Zou et al. 2012). The most recent algorithms even rely on deep-learning methods (Wang et al. 2021) but the zooming requires about 5 seconds to automatically zoom-in an object, thus not suitable for real-time tracking yet.

Viewpoint stabilization: Another auto-adaptation compensates camera slight motions during the exposure time such as *hand shake* leading to blurry images. The Optical Image Stabilisation (OIS) (Kim et al. 2019) automatically shifts the image sensor or the optics to compensate these movements. The optical shifting is more popular, since it can be easily performed by a single lens movement. First, a gyroscope measures the angles θ_x , $\theta_y \in \mathbb{R}$ caused by the hand tremor. Then, the $\{t_x, t_y\} \in \mathbb{R}^2$ lens shifts to compensate this movement are computed by:

$$t_x = \frac{\alpha_{\theta x}}{\alpha_{dx}} \cdot \theta_y, \qquad t_y = \frac{\alpha_{\theta y}}{\alpha_{dy}} \cdot \theta_x \tag{18}$$

where $\alpha_{dy}, \alpha_{dx}, \alpha_{\theta y}, \alpha_{\theta x}$ are coefficients mapping movements and translation angles to compensate the linear deviation of the focusing point, estimated with the lens law (1). Measurements in the image (not precised in (Kim et al. 2019)) complete the gyroscope to make the movements actuation even more accurate. Finally, a voice coil motor (working like speakers) performs the t_x, t_y movements of the main lens. Unlike Digital Image Stabilization, which relies on image processing algorithms, so time consuming as the number of pixel increases, OIS can compensate the motion to avoid image blur in real time and is suitable for dark scenes. However, the optimal control law for OIS just came out very recently (Kim et al. 2019), therefore it is still not implemented in cameras available for consumers [9]. Furthermore, disturbances such as friction of the lens impacts the linear control and the experiments only implement a control law minimizing hand shakes blur simulated by an hand-shake controller (thus not generalized to every tremor) (Kim et al. 2019).

5.3.2 Adaptive image sensor: moving the sensor

The first two adaptive dioptric cameras reported below work reversely than tilt-shift lenses (Sec. 5.2.1). Their image sensor moves relatively to fixed optics, but contrary to designs in Sec. 5.2.2, the movements are automated. Both of these methods intent to automatically set accurate scene regions in sharp focus.

Focusing: Similarly to *autofocus* (Sec. 5.3.1), a camera with a translating image sensor sets precisely an area in sharp focus defined by the user (Mutze 2000).

[8] for instance, Canon RF24-105mm F4-7.1 IS STM® (Canon® 2023) and Panasonic Lumix GX9® (Panasonic® 2023).

[9] but simpler anti-shake modes do exist in some cameras, for instance Panasonic Lumix G Vario Lens®with a MegaOIS®system (Panasonic® 2022).

Thanks to electromechanical actuators, the image sensor can move with 5 degrees of freedom (DoF), implemented in 4 modes. First, *the automatic inclination mode* controls the three translations t_x, t_y, t_z (around the three axes of the camera frame $\mathcal{F}_c$, Sec. 2.1) to improve image sharpness, thanks to focus estimation based on image data. Second, *the relevant region mode* combines the three translations with two rotations R_x and R_y (around ${}^c\vec{\mathbf{X}}$ and ${}^c\vec{\mathbf{Y}}$) to set in sharp focus an area defined by the user, bounded by three points in the viewfinder. Third, *the gravitation mode* translates the sensor to view completely objects which are not fully-included in the image rectangle when the camera remains static. Like the shifting of tilt-shift lenses, it allows to capture the top of tall buildings without tilting the sensor and changing the perspective. Fourth, *the continuous focusing mode* translates the sensor along the optical axis ${}^c\vec{\mathbf{Y}}$ to generate a focal stack (Sec. 4.2).

In order to maximize the image sharpness in defined regions automatically, the two rotations set in (Mutze 2000) (the tilt) are implemented with two linear actuators at each side of the sensor (Sadlo and Dachsbacher 2011) (Fig. 22a). At first, the image sensor is orthogonal to the optical axis. Then, the sensor is tilted to maximize the contrast of recorded images. The images get sharper as the contrast increases (Sadlo and Dachsbacher 2011). A least-square-based control law, based on the contrast measurements, controls the actuators to tilt the image sensor. To make the user able to set an area in sharp focus, the space needs to be sampled. To do so, between 30 and 100 images of the scene, with the sensor tilted on different angles, are previously recorded. Hence, to use this camera in a different scene, this operation needs to be redone.

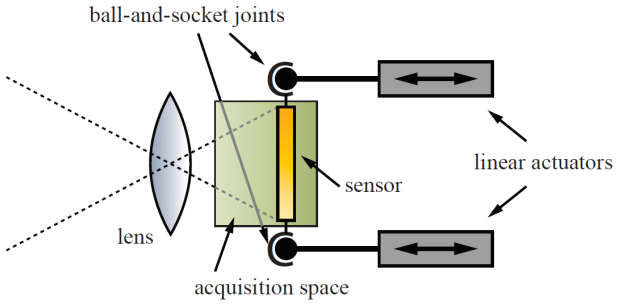
Viewpoints distance adaptation: The last dioptric camera is on the edge to be considered as a polycamera (Def. 8), since it uses two distinct moving image sensors. It is also made of two lenses, but they are fixed.

This camera captures images with a variable *parallax* (Lipton and Meyer 1992), the shift between the two images merged for binocular vision. The closer this parallax is to zero, the more accurate are the 3D geometry of scene objects. The disparity between the pixels of the two images is then minimized. For a specific *baseline*, *i.e.* the gap between the two image sensors[10], only objects at a specific depth are captured with a disparity close to zero. As these objects get farther from the camera, the baseline must decrease to maintain this zero disparity on them. To change the baseline accordingly, the two image sensors are translated horizontally and vertically w.r.t two fixed lenses. In closed-loop, the position of the left and right images are compared, and the image sensors are moved to minimize the parallax. Even though (Lipton and Meyer 1992) clearly sets the principle of a variable parallax camera, no tangible results are given and no prototype has been experimented.

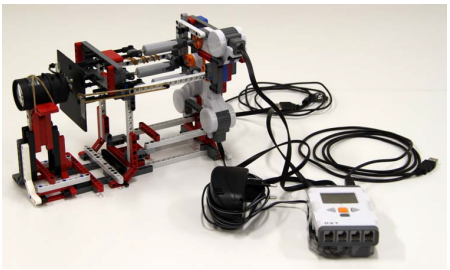
[10] in stereovision in general, the gap between two distinct cameras

5.4 Wrap-up

From all the cameras reported in this survey, the dioptric cameras are by far the most numerous. Most of their adaptations are still manual, being still in a preliminary state. Diverse conventional dioptric cameras are available for consumers, and their adaptations, such as *zooming* and *focusing* (Sec. 5.2.1) are well-known. To change their viewpoint, these conventional cameras need to be rotated or translated. Unconventional adaptations aim to avoid these external movements, by making only the sensor or optics embedded in the camera moving, *e.g.* shifting the objective changes the camera viewpoint (Sec. 5.2.1). However, most of these cameras perform global adaptations, they change uniformly the way they perceive the scene. Flexible *sheet cameras* (Sec. 5.2.1), stuck on any surface, could capture the scene with a non-uniform zoom or field of view, but



(a)



(b)

Fig. 22: The concept of the auto-tilt imaging sensor (a) and the actual prototype (b). Reproduced from (Sadlo and Dachsbacher 2011).

the cameras with both deformable optics and sensor (Sec. 5.2.3) are still emerging.

6 Catadioptric cameras

Catadioptric cameras (Def. 7) add mirrors to the optical parts of dioptric cameras. Some rather usual cameras use mirrors directly inside their objective to reduce their volume. Since those static objective are not related with any adaptive catadioptric camera, they are not discussed in this survey. Instead, this section describes dioptric cameras pointing to static, rotating or deformable mirrors.

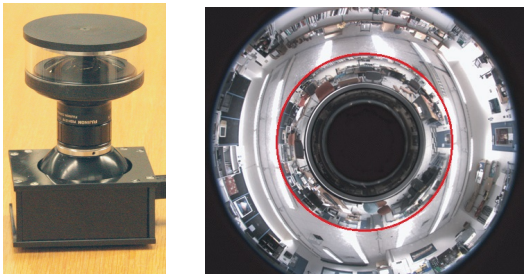
6.1 Catadioptric cameras with static properties

Pointing a conventional camera to a convex mirror increases its FOV (Sec. 2.5) (Nayar 1997). Therefore, static catadioptric cameras are mostly used for panoramic imaging. Since they use an external mirror as deported optics, they need more space than dioptric cameras with fisheye lenses, used for the same purpose (Sec. 5.1.1). They compensate this disadvantage by capturing panoramic images with a more regular resolution than fisheye cameras (Nayar 1997).

To capture panoramic images with a single effective viewpoint, catadioptric cameras need to meet specific requirements (Baker and Nayar 1999; Bakstein and Pajdla 2001). First, they can be made of a perspective camera (Def. 4) facing an hyperboloidal or ellipsoidal mirror[11]. Second, they can point an orthographic camera (Def. 4) to a paraboloidal mirror[12]. The catadioptric cameras with a single viewpoint follow a geometric model using a sphere of equivalence to project a 3D point of the world to the image plane (Geyer and Daniilidis 2001). For any other shape of mirror, this model is not valid anymore.

The shape of the curved mirror directly impacts the geometry of the image. Indeed, only a few mirror shapes allow to capture panoramic images with a single viewpoint. Nonetheless, other mirror shapes can be designed for specific applications. For instance, a pyramid-shaped mirror magnifies the four exits of a roundabout in order to monitor traffic diagnosis in crossroads (Ghorayeb et al. 2010). When this mirror is placed at the center of the roundabout, at a height of 9 meters from the ground, each face of the pyramid views a different road. The rest of the scene is ignored to maximize the usefulness of the image regarding the trafic monitoring application.

[11] *e.g.* manufactured by VStone® (Vstone® 2010)
[12] *e.g.* made by Asphericon ® (Asphericon® 2023).



(a) (b)

Fig. 23: The *cata-fisheye camera* prototype (a) and an image it acquires, made of two rings (the one with red border is captured by the mirror, the rest by a fisheye lens) (b). Reproduced from (Krishnan and Nayar 2008).

Distinct scene regions can also be captured simultaneously by combining a fisheye lens facing a spherical mirror (Fig. 23a). Such *cata-fisheye camera* captures concurrently two panoramic rings corresponding to one scene region facing the camera and another one behind the camera (Krishnan and Nayar 2008). In the captured image, the panoramic ring of the mirror is inside the fisheye one (Fig. 23b). Since this camera captures different areas along the optical axis Z (Sec.2.1), (Krishnan and Nayar 2008) expresses its FOV differently than Sec. 2.5. Instead of (θ_h, θ_v), it is expressed therein as $\theta_z \times \theta_x$, where $\theta_z = 360°$ is the FOV around the Z axis and $\theta_x = 55°$ is the FOV around the X axis (parallel to the u axis of the image plane, Sec. 2.6). However, the FOV along Z is bounded by the inter-occlusions and the combination of the different optics generates a gap between the two rings. Moreover, the resolution is not uniform, *i.e.* coarser for the inner ring than the outer.

A second example shows that a mirror made of two parts allows capturing twice in a single image the same scene region, each one with different viewpoint, resolution and FOV (Layerle et al. 2008). First, a panoramic ring, of $214° \times 180°$ FOV, is captured by a truncated, convex, paraboloidal part of the mirror. Set in a car, the image contains the inside of the car and the surroundings. Second, scene elements located at a specific depth (here, the driver's face) are magnified by a slightly concave mirror part. Designed by optimizing vector fields (Hicks and Perline 2001), this mirror is roughly a disc tilted at 22° w.r.t. the image plane. Like the aforementioned camera with a pyramid mirror, this catadioptric camera is useful for very specific applications.

In setups like (Layerle et al. 2008) where two viewpoints are captured at once, stereovision becomes possible. Recall that stereovision combines at least two

different viewpoints of an object to estimate its depth (Sec. 5.3.2). Indeed, the disparity between the pixels of this object in the two images directly gives the distance of scene elements to the camera. Therefore, estimating depth can be achieved by putting two planar mirrors in front of a conventional camera (Gluckman and Nayar 1999). However, since the mirrors are planar, this setup does not increase the FOV of the conventional camera facing them. To view a larger FOV, one should replace the planar mirrors by curved ones.

As explained above, an orthographic camera pointing to a paraboloidal mirror captures a panoramic FOV (Nayar 1997). Therefore, a camera with a telecentric objective (nearly orthographic) in front of a square containing four paraboloidal mirrors captures four panoramic viewpoints of the scene in parallel (Mouaddib et al. 2005; Caron and Eynard 2011). Combining all the pixel data brought by the four images gives the depth of all objects on the upper hemisphere facing the mirror.

Because of the occlusions due to the reflection of the mirrors on each other, an omnidirectional stereo catadioptric camera is optimal with four coplanar paraboloidal mirrors (Mouaddib et al. 2005). Capturing more viewpoints requires a different design principle so that self-occlusions do no increase. A class of cameras, named *radial imaging systems*, captures multiple views of a single 3D point of the scene, allowing to reconstruct textures, faces and even entire 3D objects (Kuthirummal and Nayar 2006). Its hardware implementation consists in a camera facing a hollow cone or a cylinder mirrored on the inside.

Compared with catadioptric cameras acquiring multiple viewpoints in parallel to perform stereovision, the following camera is very different in purpose. *Stellar Maris* is a *virtual periscope* able to capture sharp images of an airborne scene from under water. This catadioptric camera, fully immerged under water, compensates the distortions caused by the water-air interface (Fig. 24). *Stellar Maris* is made out of two components, a slope sensor and a conventional camera. They both capture data synchronously. First, the slope sensor is an horizontal array of pinholes, with the sun observation taken as a reference for measurements. Second, the conventional camera captures images of the scene above the water surface. Thanks to two planar mirrors, for every image captured by the camera, the slope sensor data are available. These data are then post-processed to reduce the distortions of the recorded images.

Apart from (Gluckman and Nayar 1999) and (Alterman et al. 2014), which rely on planar mirrors, all the systems described above use mirrors of specific shape. Such mirrors capture the scene differently than conventional cameras. But they make their specific mirror useless in scenes of different geometry, highlighting the need of adaptation.

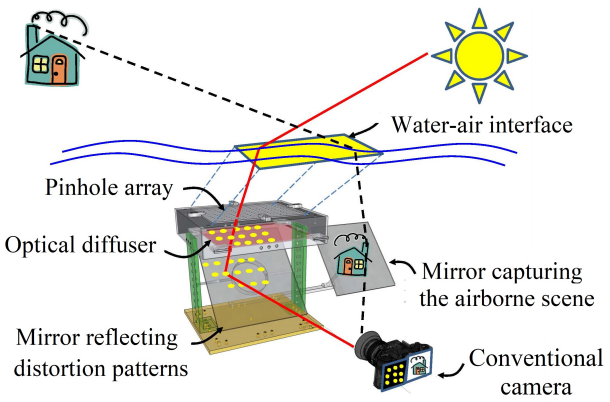


Fig. 24: *Stellar Maris* working principle (c) (reproduced from (Alterman et al. 2014)).

6.2 Manual adaptations: deforming the mirror

In a catadioptric camera, the FOV captured depends on the curvature of the mirror. Subsequently, a conventional camera in front of a flexible mirror, which can change curvature, would capture a variable FOV (Kuthirummal and Nayar 2007).

Adaptive FOV: Flexible mirror imaging modifies the FOV of a conventional camera directly by bending a mirror facing it (Kuthirummal and Nayar 2007). In experiments, the conventional camera has a $28.22° \times 20.89°$ FOV. Flexing the mirror by several centimeters on each side increases its FOV up to $120° \times 40°$. One must note that bending the mirror does not only increase the FOV value, but constructs FOVs of more complex shapes than the pyramid FOV of conventional cameras (Def. 3). Such FOV shapes are irregular, and they allow to magnify several elements of interest at once while minimizing the ones in between them. Therefore, the magnification of scene elements is not uniform. An example is given, when the top edge of the mirror is flexed. Then, the left and right sides of a street (see Fig. 25a for a conventional image of it) are magnified whereas the road between them is small (see Fig. 25b).

The flexible planar mirror is $465\ mm \times 355\ mm$ of size, and is made of an acrylic layer (for the reflection) above a plastic sheet (Kuthirummal and Nayar 2007) (Fig. 26). The dimensions of the mirror are chosen to minimize blurriness. Indeed, when flexed to reach a defined shape, a smaller mirror takes a higher curvature, inducing optical aberrations.

(a)

(b)

(c)

Fig. 25: The conventional image of a street (a), an image captured with the mirror flexed (b) and the corresponding equirectangular image (c). Reproduced from (Kuthirummal and Nayar 2007).

Since the mirror is flexed manually to magnify scene regions, the shape of the mirror is not directly controlled by changes in the scene. To track an object, the user must deform the mirror as fast as the object moves, and in the good direction, which is a very complex task for a human. Also, the captured images are severely distorted and require post-processing to construct rectangular images of uniform resolution (*equirectangular images*, see Fig. 25c), which needs to accurately estimate the mirror shape.

Adaptive FOV and HDR imaging: In summary, a flexible mirror with a continuous surface can be curved to capture a variable FOV. The manually adaptative catadioptric cameras can also rely on a segmented mirror surface to enable different adaptations.

A Deformable Mirror Device (DMD) is an array of micromirrors, *i.e.* square mirrors of less than one millimeter of side [13]. This segmented mirror modulates the incident light in a spatial pattern of black and white pixels. Therefore, they can reproduce an image given as an input, each micromirror representing a pixel in the image (Hornbeck 1983). To change the aspect of the pixel, the micromirrors switch between three different orientations, -10 °, 0 °, +10 °, around two axes (parallel to the (u, v) axes of the image plane), in a few microseconds.

[13] generally provided by Texas Instruments ® (Texas Instruments® 2018).

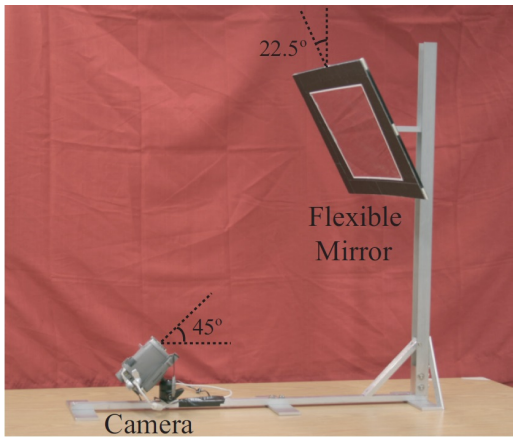


Fig. 26: The flexible mirror imaging prototype. Reproduced from (Kuthirummal and Nayar 2007).

The *programmable imaging* system consists in a camera pointing to a DMD, enabling two different adaptations (Nayar et al. 2006). Indeed, depending on the orientation of the micromirrors, either the camera viewpoint or the brightness of scene regions can be changed. On the one hand, changing the orientation of all the micromirrors from 0° to ±10° (around either the u or the v axes) virtually rotates the camera viewpoint of 20°. Nonetheless, since the individual mirrors can only have two different non-zeros orientations around two axes, the camera can only turn its viewpoint by 20° on the X or Y directions of the scene. On the other hand, since tilting the mirrors modulates the incident light, tilting the mirrors can reduce the brightness of scene captured. Then, the exposure time of the camera is *virtually* decreased. By setting a different orientation in patterns of 2×2 neighbouring mirrors, one can capture four different exposures of the scene at once, which can be merged into an HDR image. This principle can be extended to capture HDR videos without any artifact. For this purpose, the neighbouring mirrors can change of orientation on a cyclic way, at 30 Hertz.

Whereas allowing two different adaptations, the *programmable imaging* system has limits. Even though the DMD alone is only a few centimeters of side, the full prototype is not compact since it uses a dismounted video-projector (Fig. 27). Its lens focuses an image of the scene on the DMD, which is then re-focused by a lens on a monochrome camera. These two lenses reduce the resolution of the acquired images. Indeed, each pixel of the CCD corresponds to a group of 3×3 micromirrors of the DMD. Even though the DMD used here can display patterns of 800×600 black or white

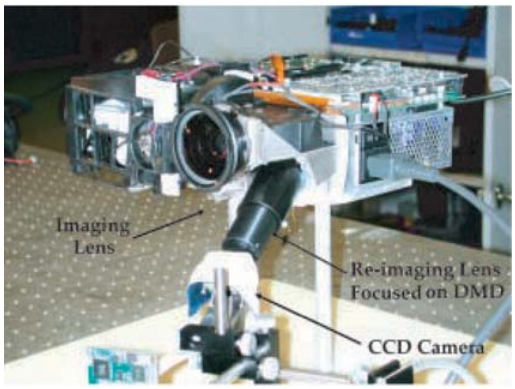


Fig. 27: The *Programmable imaging* prototype (reproduced from (Nayar et al. 2006))

pixels, the camera will only capture images of 200×150 pixels. Pointing directly the camera on the DMD, without these intermediary lenses, would be an interesting solution. However, since it generates blurriness in the image, it cannot be considered.

6.3 Automatic adaptations

The manually adaptive catadioptric cameras all rely on a camera pointing to a mirror which changes shape across time to modify the properties of the images captured. Whereas a majority of the designs reported in this section are based on that principle, the mirror can also be moved as a whole in front of the camera.

6.3.1 Adaptive optics: moving the mirror

When a planar mirror is in front of a camera, all the light rays hitting the mirror change of direction on the same way, because of reflection laws. Hence, rotating a planar mirror in front of a camera changes its viewpoint.

Rotating the viewpoint: A monopixel camera relies on this principle. It captures images made of a single pixel for any potential orientations ($\{\theta_x, \theta_y\} \in \mathbb{R}^2$) of the mirror (Krishnan and Nayar 2009). The planar mirror is held by a robotic arm. It moves all around a ball lens, with a 640×480 image sensor covered by a small aperture beneath. Then, all the pixels captured at different orientations can be gathered in an image. However, this one-by-one pixel acquistion needs 3 seconds to capture a $12.79° \times 9.28°$ FOV as an image of 240×175 pixels. Higher resolution images have been obtained by simulation only.

Adaptive viewpoint and FOV: Since each orientation $\{\theta_x, \theta_y\} \in \mathbb{R}^2$ of the mirror changes the viewpoint of the camera facing it, this principle can be extended to capture full images from different camera angles (Marchand and Chaumette 2017). Moreover, the translation $t_z \in \mathbb{R}$, along the axis perpendicular to the mirror, changes the magnification G of scene elements (Sec. 2.1). Then, an object of the scene appears bigger as the mirror gets closer to the camera. Practically, both the viewpoint and the magnification can be closed-loop controlled by visual servoing (Marchand and Chaumette 2017). The mirror is automatically rotated and translated by the end-effector of a robotic arm until the four corners of a target object are at the desired location.

Moving the viewpoint at high frequency: In the visual servoing case, the mirror movements are relatively slow since they depend on the camera framerate and the response of the robotic arm. In order to track several objects in different scene regions within a few milliseconds, the viewpoint adaptation needs to be faster. *FoveaCam* achieves this goal, it is able to track objects of interest and change the camera viewpoint at high frequency (Tilmon et al. 2020). It relies on a small circular mirror of only a few centimeters in diameter. Thanks to an electromagnetic actuation, this mirror is rotated in closed-loop around the X and Y axes at a frequency of tens of kilo-Hertz. Since the mirror is very small, the camera captures images of a 8.6° FOV.

Moving a planar mirror generates global adaptations in the images captured. Indeed, the camera viewpoint and magnification globally change as the mirror is rotated or translated. In order to perform local adaptations, the planar mirror is replaced by a deformable one, able to take a wide variety of shapes.

6.3.2 Adaptive optics: deforming the mirror

The surface of Deformable Mirrors (DMs) can be locally curved on many locations simultaneously. Therefore, they can be set on many different and complex shapes, described by 2D polynomials (Davies and Kasper 2012). Such mirrors are used in Adaptive Optics (AO), aiming at counterbalancing atmospheric perturbations of the light to improve the image sharpness locally.

Focusing: The locus of the incoming light rays from an object of interest, for instance a star, is a *wavefront*. When light passes through a *medium* with temporally and spatially irregular refraction indexes, such as in the atmosphere, its path is modified, thus disturbing the

wavefronts. Then, the image of a star appears blurry. Adaptive Optics rectifies the path of light, to make all the rays coming from the objects of interest reaching simultaneously the image sensor. The wavefront is thus corrected to become planar (Davies and Kasper 2012). AO implements cameras based on a DM, which surface changes dynamically from wavefronts measurements feedback at high frequency (see the working principle in Figure 28) (Davies and Kasper 2012).

Some stars are not bright enough to directly analyze their wavefront. Instead, the wavefronts coming from natural or artificial Guide Stars (GS) are measured by the Wave Front Sensor and corrected by the DM (Davies and Kasper 2012; Roddier 1999; Tyson 2015). Systems relying on a single GS to actuate only one DM, Single Conjugate AO (SCAO), capture sharp a small region around the GS. To extend this sharp region, Multi Conjugate AO (MCAO) and Multi Object AO (MOAO) multiply the number of GSs and DMs (Davies and Kasper 2012).

Remark 3 (The case of Stellar Maris:) Stellar Maris (Alterman et al. 2014) (Sec. 6.1) also intends to correct the distortions caused by a change of propagation medium, but uses the sun as a GS and rely on post-processing instead of a mirror changing its shape in real-time for optical adaptation.

In general, DMs are characterized by a large number of actuators, going from a few dozens to several thousands, and the high-frequency of deformations, about 1 kilo-Hertz (Madec 2012). The actuators are densely distributed above the mirror surface, separated by a pitch between a few millimeters and several centimeters. The DM mechanical stroke, the range of the local deformations caused by the actuators, are up to tens of micrometers, enough to undistort upcoming wavefronts.

Most kind of DMs are listed in (Madec 2012), the most comprehensive survey about DMs used in astronomy. In this survey, DMs are classified depending if the

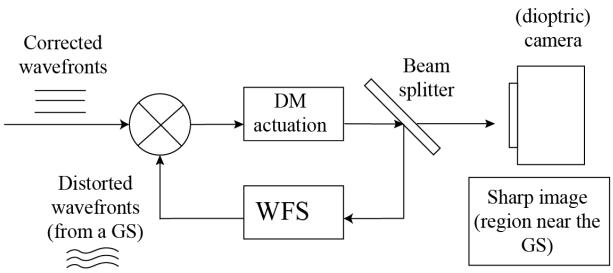
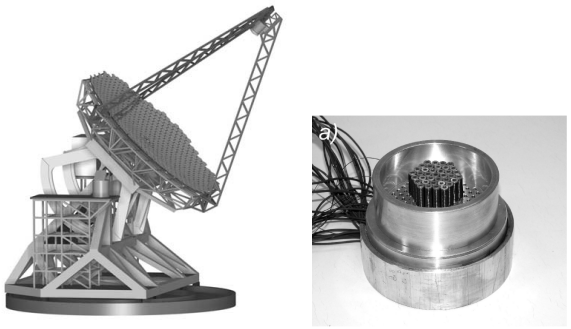


Fig. 28: Adaptive Optics with a Deformable Mirror (simplified) principle.



(a) (b)

Fig. 29: The GSMT design (reproduced from (Strom et al. 2003)) (a), an array of actuators for a liquid mirror (reproduced from (Brousseau et al. 2007)) (b).

surface of the mirror is continuous or segmented. A majority of DMs consists in a continuous mirror surface placed above actuators. They can be arranged in arrays, *e.g.* Stacked Arrays DMs (SAM)[14] or voice coil actuators[15]. Alternatively, they can consist in disks of deformable materials, which change of curvature like bimorph DMs (Huisman et al. 2020).

Whereas the aforementioned DMs all deform a continuous mirror membrane placed above the actuators, DMs can have a segmented surface. However, contrary to DMDs (Sec. 6.2), each segment can take many different angles and positions, thanks to tens of actuators controlling each, to make the mirror surface smooth (Strom et al. 2003). Several astronomical telescopes (detailed in Sec. 8) are based on the Giant Segmented Mirror Telescope (GMST) design, of 30 meters in diameter (Fig 29a).

Instead of a soft membrane or a set of rigid mirrors above actuators, DMs can also be made of liquid rotating in containers, therefore named *liquid mirrors* (Borra 1982; Brousseau et al. 2007; Ritcey and Borra 2010). Arrays of coil actuators (Fig. 29b) induce a local magnetic field to change the curvature of the surface of the liquid in several different places simultaneously. Recently, a liquid mirror can take an even more diverse set of shapes thanks to an electromagnet deforming a very smooth ferrogel (Falahati et al. 2020).

All the DMs reported above require a lot of space for the electronics to monitor and control their actuation. Micro-Opto-Electro-Mechanical Systems (MOEMS), the latest development in DMs, are much more compact. They use miniaturized actuators of less than 5 micro-meters of size, gathered in arrays of

[14] provided by CILAS® (Cilas® 2023) and Boston Micromachines® (Boston Micromachines® 2023)
[15] supplied by Microgate® (Microgate® 2011).

up to tens of thousands. MOEMS are fast actuated, at a frequency up to tens of kilo-Hertz. MOEMS are classified in four categories (Motamedi 2005): membrane mirrors, continuous mirrors, segmented mirrors and tip-tilt-piston segmented mirrors. Because of their reduced size, the MOEMS mechanical stroke (maximum deformation) is limited to 90 micro-meters [16].

Remark 4 (A specific kind of MOEMS:) Instead of enabling several local deformations of their surface at once, a specific kind of MOEMS is moved as a whole. They are circular MOEMS placed above magnets, which can be tilted by a few degrees on two directions by variations of the electromagnetic field. *FoveaCam* combines a camera with such circular MOEMS (Tilmon et al. 2020).

Adaptive FOV: A wide majority of DMs is used to locally improve the sharpness of images. Their shape can be accurately set and optimized to correct the shape of wavefronts automatically. Nonetheless, the adaptations of DMs can be downgraded to manual and global adaptations. Indeed, two cameras can switch between two different focal lengths f_1 and f_2 (Sec. 2) thanks to DM actuation. Instead of the lens movements performed by the zoom of conventional cameras (Sec. 5.3.1), they both rely on DM changing shape inside the objective of a camera (Zhao et al. 2013; Lin et al. 2012). Whereas they aim at reducing the size of conventional camera optics, they are far from enabling as much changes of magnification as conventional zooms (Sec. 5.2.1).

6.4 Wrap-up

Pointed to a mirror with a non-planar shape, a camera views its environment differently. Indeed, convex mirrors increase the FOV of the camera (Baker and Nayar 1999; Bakstein and Pajdla 2001; Nayar 1997) whereas concave ones can increase the resolution of specific parts of the scene (Layerle et al. 2008). The image resolution is not uniform anymore, hence scene objects of same size at same depths may appear with different sizes in the image. Therefore, the FOV of a camera facing a flexible mirror changes as the mirror is curved (Kuthirummal and Nayar 2007). Instead of curving it globally, deforming the mirror locally and at high-frequency sets sharp previously blurred image regions (Madec 2012; Motamedi 2005; Davies and Kasper 2012). Whereas most adaptive catadioptric cameras have a mirror which dynamically changes shape, some of them rely on a rotating mirror to move the camera viewpoint (Marchand and Chaumette 2017; Krishnan and Nayar 2009; Tilmon et al. 2020) or a translating one to change its magnification (Marchand and Chaumette 2017).

[16] continuous MOEMS provided by ALPAO (ALPAO® 2023) have a 90 micro-meters stroke.

7 Polycameras

Polycameras (Def 8) are diverse, since they consider any layout of cameras, once they are synchronized to capture several viewpoints of the scene simultaneously. After reviewing the major trends of static polycameras (Sec. 7.1), the few manual and automatically adaptive are surveyed (Sec. 7.2, resp. 7.3).

7.1 Polycameras with static properties

Polycameras are classified in (Wang et al. 2017), according to the layout of the camera cluster or array.

Parallel layouts allow to capture either high definition (Sec. 2.6) or wide FOV (Sec. 2.5) images. The individual cameras are patched together tightly, spaced by only a few millimeters, along one or two dimensions (Fig. 30a). The images captured simultaneously by all the cameras are merged to compute the final polycamera image, *the polymage*. Since the individual cameras

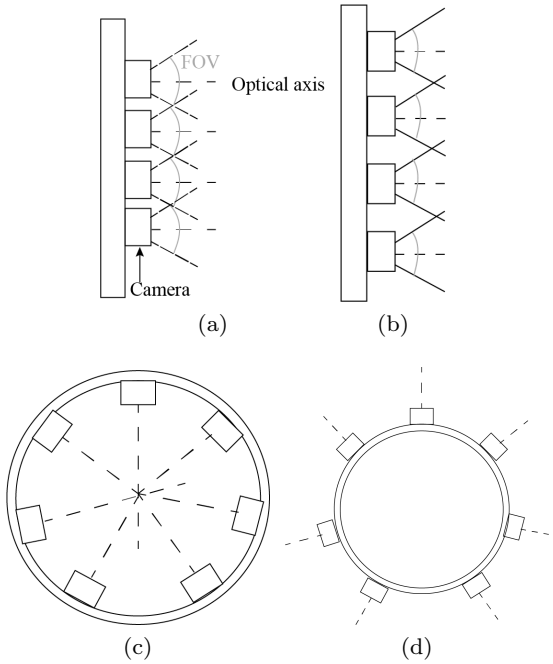


Fig. 30: The different polycamera layouts: first row, coplanar layouts where camera have a (a) large and (b) small overlapping FOV ; second row, radial layouts (a) toed-in or (d) toed-out.

are very close to each other, if all their optical axes are parallel, a large part of their FOVs overlap. Therefore, the definition of such polycamera is several times the one of a single camera constituing it [17].

When the cameras of such tightly packed parallel layouts are oriented so that their optical axes diverge, the FOV captured by the polymage is the union of the FOVs captured by all the individual cameras. Therefore, the less their FOVs overlap, the higher the polymage FOV. Prototypes are presented by the survey (Cheng et al. 2019), notably a 5×5 dioptric cameras array approximately spaced by one centimeter from each other. In this setup, if all the optical axes are parallel, a major part of the cameras FOVs overlap. Then, the definition of the polymage is significantly higher than the one of an individual camera, contrary to the FOV captured which remains almost the same. To increase more the FOV, one should increase the angle between the neighboring cameras. Since all the cameras are set on a planar support and located close to each other, the rotation of each individual camera is bounded to avoid collisions. Hence the FOV is increased, but still limited.

In this case, the baseline is small. Therefore, the multiple viewpoints captured at once are very close to each other. The corresponding images are then close enough to each other to be combined to reconstruct the light field coming from a small region of the scene (Xie et al. 2016). When the baseline between two consecutive cameras is increased up to several centimeters, their viewpoints are more distinct and their own FOVs overlap only slightly (Fig. 30b). Hence, the polymage becomes a mosaic of neighbouring, but distinct, regions of the scene. Then a panoramic FOV can be captured by only three aligned and ten-centimeters distant cameras (Baker et al. 2005).

The second type of polycameras, radial layouts, consist in arrays of cameras set up in curved supports They are sub-classified as *toed-in* (Fig. 30c) and *toed-out* (Fig. 30d) camera arrays, respectively with optical axes converging or not to a common point (Wang et al. 2017). The *Panoptic Studio*, a geodesic dome, 4.15 m height and 5.49 m diameter, is a toed-in layout of 480 cameras, which records movement of people inside the dome as 3D data (Joo et al. 2015).

Toed-out radial layouts consist in monocentric camera arrays where each camera views a different direction. They can be arranged in circles for omnidirectional imaging[18], or in hemispheres like *panoptic cameras*. If the viewpoints are close to each other, the panoptic cameras capture omnidirectional images with depth information (Afshari et al. 2013; Baker et al. 2004).

30 cameras of $66° \times 68°$ with focal length $f = 1.17\ mm$, mounted in an hemisphere, capture a 180° FOV (Afshari et al. 2013). Since neighbouring cameras are only spaced by a few degrees on the sphere surface, their viewpoints are redundant, *i.e.* they capture slightly different version of the same scene region, allowing depth reconstruction.

The majority of polycameras is classified into coplanar or radial dense layouts but some made of sparse sets of only a few cameras are at the bordeline of Def 8. For instance, equipped with fisheye lenses and set back-to-back, two cameras can capture the omnidirectional FOV (Forutanpour et al. 2019): they are dual-fisheye cameras[19].

7.2 Manual adaptations: deforming the layout

Three polycameras can be considered manually adaptive, in regards with definition 10. Whereas all of them capture an adaptive FOV, the third one can also capture a variable number of viewpoints (Sec. 2.5).

Adaptive FOV: The first polycamera reported in this section has both an horizontal and vertical adaptive FOV (Nomura et al. 2007). It consists in a 2D array of 5×4 camera modules, each being $41\ mm \times 24\ mm \times 34\ mm$ of size and capturing 752×480 pixels images. The cameras are attached on a flexible plastic sheet (Fig. 31a). The *scene collage* algorithm stitches together the images acquired by each camera, even if there are holes between the partial views of the scene (Fig. 31b) (Nomura et al. 2007). Nonetheless, this post-processing stage requires 700 seconds to merge the 20 input images into a final frame.

The second design, *FlexCam*, has an horizontal versatile FOV. Contrary to the aforementioned flexible polycamera, it does not require long post-processing steps to stitch several images (Dickie et al. 2012). The prototype of Flexcam consists in three cameras aligned and mounted on a bendable support. Each camera is $29\ mm \times 29\ mm \times 38\ mm$ and captures images of 640×480 pixels. Two horizontal flex sensors measure the shape of the bendable support accurately, which is used to stitch the images captured by the three cameras into a polymage. Then, this polymage is displayed by

[17] Sized like a smart-phone, Light L16® (Light® 2019) gathers 16 dioptric cameras of different focal lengths (28 mm, 70 mm and 150 mm) to capture 52-megapixel photographs (Sahin and Laroia 2017).

[18] *e.g.* GoPro 360® (GoPro® 2015) camera array.

[19] available as Ricoh Theta® (Ricoh® 2019) products, for instance.

a Flexible Organic Light Emitting Diode (FOLED) photographic viewfinder directly embedded in the bendable support. This polycamera acts similarly than many cameras available for consumers, which shows the images recorded by the camera on a small screen. In (Dickie et al. 2012)'s design, the FOLED is of 320×240 pixels of definition, spoiled by edge artifacts.

Adaptive FOV and number of viewpoints: The last manual design, *CrossbowCam* is not only able to capture images with an adaptive FOV. Indeed, by bending the support of its cameras, one can capture several distinct views of the same scene elements at once with this polycamera (Hsu et al. 2017). The *CrossbowCam* prototype consists in eight cameras aligned on a thin steel plate of length $l = 320\ mm$, spaced by $38\ mm$ (Fig. 32a). From operator inputs, a scroll and a stepping motor directly connected to the center of the plate changes the curvature of the thin plate, mimicking a crossbow mechanism. Thus, the camera array can change its layout between three modes.

In the first mode, the support of cameras is convex, therefore, the camera optical axes diverge such that *CrossbowCam* is a radial toed-out layout (Sec. 7.1). When this polycamera gets from planar to convex shape, its FOV increases from 60° to 90°, and is expressed as:

$$\theta = f(h) = 2 \cdot \left(90° - \tan^{-1}\left(\frac{h}{l/2}\right) + \tan^{-1}\left(\frac{l/2}{h}\right)\right) \tag{19}$$

with $h = 20\ cm$ the distance between the central camera and the center position, when the support of the cameras is set on a planar shape.

In the second mode, the optical axes of all the cameras are parallel. Therefore, *CrossbowCam* is a parallel layout of cameras, their optical axes being all parallel (Sec. 7.1). In the third mode, the camera support is concave, mimicking radial toed-in layouts. In both modes, all the cameras capture overlapping viewpoints which enable 3D reconstruction (see Sec 8 for more details).

7.3 Automatic adaptations: move each camera for adaptive viewpoints

To our knowledge, the self-reconfigurable camera array (Zhang and Chen 2004) is the single automatically actuated polycamera designs, according to Definition 11.

The *self-reconfigurable camera array*, like the second and third modes of *CrossbowCam*, aims at capturing overlapping viewpoints to improve the scene depth

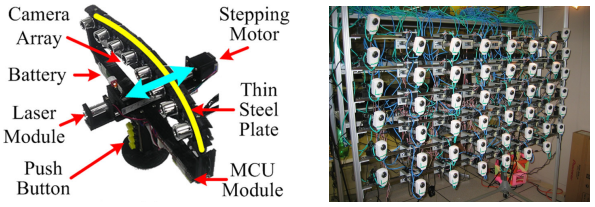
(a) (b)

Fig. 31: The Flexible Camera Array prototype, flexed, (a) and the image acquired, stitched by the *scene collage* algorithm (b) (reproduced from (Nomura et al. 2007)).

reconstruction (Zhang and Chen 2004). It consists in an array of 48 cameras (Fig. 32b) individually able to perform two movements (sidestep and pan). They are placed in 6 linear guides, with 8 cameras spaced by 20 *mm* each. Each camera captures images of 320×240 pixels. In total, 84 servo-motors move the cameras (48 for the pan and 36 for the sidestep). The whole polycamera is closed-loop controlled. At first, the images captured by two neighbouring cameras are compared. Then, the cameras pan and sidestep to optimize the difference between the viewpoints, in order to capture depth information. Controlling all the cameras for this purpose requires hundred of milliseconds. According to (Zhang and Chen 2004), the case of two neighbouring cameras colliding is actually possible.

7.4 Wrap-up

The layouts of polycameras can have either a coplanar or radial geometry (Sec. 7.1). Therefore, to change the way they capture the scene, adaptive polycameras change their geometry. Both the flexible camera array (Nomura et al. 2007) and *Crossbowcam* (Hsu et al. 2017) the shape of the support of the individual



(a) (b)

Fig. 32: Working principle of the *CrossbowCam* (reproduced from (Hsu et al. 2017)) (a) and prototype of the self-reconfigurable array (reproduced from (Zhang and Chen 2004)) (b).

camera to adapt (by increasing the FOV or the number of views). The *self-reconfigurable polycamera* rather changes the orientation of each individual camera to capture redundant viewpoints of a specific scene object (Zhang and Chen 2004). The biggest issue of versatile polycameras is the computational cost induced by their adaptation. As the two former rely on post-processing steps to construct a final image, the latter needs hundreds of milliseconds to be controlled.

8 Applications of adaptive cameras

In this section, the applications made possible by adaptive cameras are described, following the order of the designs reported in this survey (Sec. 3).

Lensless cameras (Sec. 4) are mainly employed in biomedical engineering (Boominathan et al. 2016). Indeed, their compactness allows to embed them in optical fibers to serve as endoscopes (Papadopoulos et al. 2013). They are also the base of microscopes (Moon et al. 2009) to image fluids inside of the body. Moreover, the adaptive aperture of some lensless cameras makes them able to capture several distinct regions of the scene during a single acquisition, *e.g.* to detect several parts of a face simultaneously and ease the person identification (Zomet and Nayar 2006). Differently, subregions of captured images can be redundant and combined to get depth information of the scene, then allowing 3D reconstruction (Schwarz et al. 2016). The ability of the bendable lensless camera (Koppelhuber and Bimber 2013) to change FOV as its structure is getting curved allows to use it for contactless sensing and optical see-through monitoring, but the prototype needs to be made thinner (300 μm instead of 6 mm currently).

Several adaptations of dioptric cameras (Sec. 5), like zooming and focusing, are well-known and, therefore, used in quantity of applications, including surveillance, animal observation, artistic photography and fashion (Maître 2017; Wang et al. 2021; Roberts 1995). In entertainment and arts, two camera adaptations are able to make real landscapes like models and the other way around. On the one hand, acting on the diaphragm to close partially the camera aperture increases the depth of field (Sec. 2.4). Therefore, for a specific aperture size, some parts of a scene object can look blurry whereas the others look sharp. For cinema, this makes miniature models looking like life-size landscapes (Held et al. 2010). On the other hand, tilting the objective rotates the plane of sharp focus (Sec. 5.2.1). Then, the disjoint blurry areas in the image turn real landscapes into models in the image, for art (Qian and Miura 2019). In order to take panoramic photographs of landscape and tall buildings in a single frame without any distortion (mostly for art and architecture), the objective may be translated thanks to tilt-shift lenses (Schulz 2015; Ewing 2016). The less conventional adaptation of all is to change the shape of a whole *sheet camera* to wrap it around any curved shape. Therefore, its FOV could take many different extents and shapes (Sims et al. 2016, 2018). This could be used in autonomous cars, where the sheet camera could monitor the surrounding road, or surveillance, where cameras could be wrapped around objects of public areas. However, several improvements of these cameras, discussed in Section 9, should be done beforehand to achieve these applications.

In regards with catadioptric cameras (Sec. 6), the applications depend of the amplitude of the deformations of the mirror shape. Indeed, small deformations of less than one millimeter allows to locally change the blurriness of small image regions. Adaptive Optics (AO, Sec. 6.3) uses this principle to image sky on astrophotographs[20], be used in ophtalmology for accurate retinal imaging (Gill et al. 2019) or even allows to capture microscopic images with nanometric accuracy (Lee et al. 2020). Changes of curvature of more than 5 mm cause two different effects: a concavity magnifies a scene region, which can be useful for surveillance (Kuthirummal and Nayar 2007) or autonomous vehicles by detecting fatigue on driver's face while monitoring the surrounding road (Layerle et al. 2008) ; a convexity, on the contrary, enlarges the FOV of the camera pointing to the mirror, allowing to capture a large group of people without stepping back or monitor a whole street even if the camera is in the middle of it (Kuthirummal and Nayar 2007).

Polycameras (Sec. 7) are getting more and more common for consumers, since even smarthphones are using them in order to capture landscapes with a large FOV Dxomark® (2021). Moreover, they are used for defense applications, since parallel layouts can acquire gigapixel resolution images with HDR ability (Wilburn et al. 2005; Brady et al. 2012). Flexible camera arrays (Nomura et al. 2007) change shape in order to track people and also capture artistic photographs, with some parts of the FOV left blank. Setting a polycamera like *CrossbowCam* on a concave curvature allows to capture multiple viewpoints at once for 3D reconstruction, *e.g.* autostereoscopic displays (*i.e.* showing images on a planar screen which appear embossed from any spectator

[20] 3 telescopes are based on the Giant Segmented Mirror Telescope concept (Strom et al. 2003): the European Extremely Large Telescope (EELT) European Southern Observatory (ESO) (2018), the Keck telescopes Observatory (2023) and the James Webb Space Telescope (JWST) NASA (2023).

viewpoint) for entertainment (Hsu et al. 2017). They may also produce visual effects for action movies, such as "bullet times" which gives a feeling of time freezing (Hsu et al. 2017).

9 Future works on adaptive cameras

This section proposes future research lines to improve the existing adaptive cameras, reported in the order followed throughout this survey (Sec. 3). For each camera category, the future works given by the authors of the papers surveyed are followed by ideas of our own.

Since they do not require any lens, lensless cameras (Sec. 4) are compact. However, because they use only apertures as optics, they have a lower SNR than other cameras (Schwarz et al. 2015, 2016). Moreover, the image reconstruction algorithms amplify the noise in the final image (Boominathan et al. 2016). Therefore, to improve the SNR, the aperture of lensless cameras has to be optimized (Boominathan et al. 2016). When the scene changes, the aperture shape needs to change accordingly to record each region of the scene with as much light as possible and avoiding blur (Schwarz et al. 2016). The programmable aperture mask of SweepCam (Hua et al. 2020) is a liquid crystal on silicon (LCoS) screen which limits the image definition and does not capture enough light at once. Instead, a translating mask with better light-efficiency, constituted of smaller apertures, would improve both the SNR and the image definition of the *Sweepcam* prototype. A different camera with an adaptive aperture, which relies on 2 layers of apertures stuck together, is currently able to capture at least 3 distinct scene regions at once (Zomet and Nayar 2006). The authors plan to add more aperture layers in order to capture even more distinct regions at once, and capture each of them with a different magnification, therefore making this camera able to perform a spatially variable zoom. Indeed, since transverse magnification increases as the focal length does (2), the magnification changes as the distance from aperture layers to the image sensor does. Also, those aperture grids would be closed-loop controlled in order to track a moving object in real-time.

In our opinion, lensless cameras must solve their main flaw: most of them do not capture images directly. Instead, they record data which is then processed into images by reconstruction algorithms which induce a delay (at least 100 ms (Asif et al. 2016; Boominathan et al. 2016)). In that direction, the lensless camera of (Zomet and Nayar 2006) already computes final images with the difference of images captured with two different aperture patterns. This principle could be applied by lensless cameras in general to reduce the time before actually seeing the picture captured. Indeed, the cost in time of the image processing prevent automatically adaptive cameras to be conceived. If lensless cameras took less than 50 ms to capture images, then their aperture or their image sensor could change automatically in real time, to adapt to changes in the scene.

On the other hand, all the lensless cameras only capture grayscale images. With a simple image processing just computing the difference of two images, adding R, G and B channels to the image sensor would add colors to the images recorded.

In regards with dioptric cameras (Sec. 5), apart from the translating image sensor which could be replaced by actuated tilt-shift lenses (Sadlo and Dachsbacher 2011), and other improvements on the control laws responsibble for focusing, OIS and automatic zoom (see Sec. 5.3), the most challenging and numerous future improvements would go to the *sheet cameras*. Indeed, their potential is high, but they are not yet efficient enough to be used for the applications described in Section 8. Several research lines are proposed in (Sims et al. 2016, 2018). At first, the current prototypes of (Sims et al. 2016, 2018) capture 33×33 colored dots and the final image, of 264×264 pixels, is interpolated. A microlens array with more lenses, and tinier, would directly record images of higher resolution. Second, sensing the deformations of the sheet camera, with stress sensors or accelerometers, would allow to record the geometry of the camera and make it easier to undistort the captured images. In order to reduce the space of these cameras, (Sims et al. 2018) proposes to replace the whole dioptric camera facing the flexible microlens array by a sparse array of pixels, with the same pitch than the one with two successive lenses, almost stuck to the microlens array (getting similar with a compact dioptric camera).

We believe that ideal *sheet cameras* should combine a flexible image sensor with deformable lenses, in order to significantly reduce their size. Therefore, the microlens arrays used in the prototypes of (Sims et al. 2016, 2018) should be combined with a fully-flexible image sensor, for instance an array of organic photodiodes (Sec. 5.2.2) printed on a plastic sheet.

Catadioptric cameras (Sec. 6) can rely on a segmented or continuous mirror. On the first case, DMDs (Sec. 6.2) should use mirrors of more intermediary orientations than just the two extreme ones, in order to change the FOV captured instead of tilting the camera viewpoint by a fixed number of degrees (Nayar et al. 2006). On the second case, the flexible continuous mirror of (Kuthirummal and Nayar 2007) must solve the trade-off between reducing the size of the mirror for compacity and minimizing the post-processing

time to make the images displayable. Indeed, small mirrors of a few centimeters of side induce optical aberrations when they are curved, which generates noise. Continuous mirrors with accurate and small deformations, *i.e.* DMs (Sec. 6.3), should improve their actuation to take an even more diverse set of shapes, according to many authors (Madec 2012; Huisman et al. 2020; Falahati et al. 2020). Also, the maximum deformations of MOEMS DMs (10 micrometers currently) could be increased with double-stage actuation, *i.e.* putting actuators of bigger DMs under the microscopic ones of the MOEMS (Zamkotsian and Dohlen 2002).

We think that two different principles of flexible mirrors could be combined to create a new camera which could automatically magnify several scene regions at once, while keeping the intermediary regions, though degraded. Indeed, deformations of amplitude higher than 5 millimeters allow to make important resolution changes in the image and the automation of DMs could be adapted to automatically change in closed-loop the mirror shape to capture temporally variable multi-resolution images.

Last, polycameras (Sec. 7) should be miniaturized to be used as the *new family camcorder* (Nomura et al. 2007). To do so, the camera modules must be small, of less than one centimeter of size, like the ones in smartphones (Dickie et al. 2012). The polycameras changing of shape, *e.g.* by a simple button pressed like CrossbowCam (Hsu et al. 2017), should embed all the electronics on the camera instead of being connected to a remote computer. Also, the torque of the actuators must be increased to allow more deformations with a higher amplitude.

The shape of the support of the camera modules could be automatically changed in order to automatically zoom-in different scene regions at once, but the space between two different cameras would limit the amplitude of the deformations. Indeed, unlike the continuous mirror case, the image would be discontinuous if the camera modules are spaced, and the higher the deformation, the more important the discontinuities.

10 Conclusion

This article surveys adaptive cameras, classified in four categories: lensless, dioptric, catadioptric and polydioptric cameras. For each design, the nature of the adaptation and its impact on the images is described. This article attempts to give the most comprehensive list of adaptive cameras, by reporting the works made in various research fields, related with specific kind of imaging devices, such as deformable mirrors or artificial compound eyes, alongside with many original designs.

Adaptive cameras meet the requirements of a wide range of applications, for which conventional cameras are not efficient enough. Nonetheless, most of these designs are emerging, further researches have to be done to over-take several limitations. First, lensless cameras should be able to directly acquire images of the scene, without image reconstruction steps. Then, the adaptations met in dioptric cameras have to be sized-down. Catadioptric cameras could combine the advantages of the self-actuated deformable mirrors and the wide and discontinuous field of view deformations allowed by flexible mirrors. Finally, adaptive polycameras are either able to capture a changing FOV or a various number of viewpoints of the same scene objects at once, but they need to be sized-down as well.

Data availability

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

References

Adelson EH, Bergen JR, et al. (1991) The plenoptic function and the elements of early vision, vol 2. Vision and Modeling Group, Media Laboratory, Massachusetts Institute of ...

Afshari H, Jacques L, Bagnato L, Schmid A, Vandergheynst P, Leblebici Y (2013) The panoptic camera: a plenoptic sensor with real-time omnidirectional capability. Journal of Signal Processing Systems 70(3):p305–328

Allen E, Triantaphillidou S (2012) The manual of photography. CRC Press

Alterman M, Swirski Y, Schechner YY (2014) STELLA MARIS: stellar marine refractive imaging sensor. In: IEEE Int. Conf. on Computational Photography (ICCP), IEEE, pp p1–10

Asif MS, Ayremlou A, Sankaranarayanan A, Veeraraghavan A, Baraniuk RG (2016) Flatcam: Thin, lensless cameras using coded aperture and computation. IEEE Trans on Computational Imaging 3(3):p384–397

Baker HH, Tanguay D, Papadas C (2005) Multi-viewpoint uncompressed capture and mosaicking with a high-bandwidth pc camera array. In: Proc. Workshop on Omnidirectional Vision (OMNIVIS 2005)

Baker P, Ogale AS, Fermuller C (2004) The argus eye: a new imaging system designed to facilitate robotic tasks of motion. IEEE Robotics & Automation Magazine 11(4):31–38

Baker S, Nayar SK (1999) A theory of single-viewpoint catadioptric image formation. Int journal of computer vision 35(2):p175–196

Bakstein H, Pajdla T (2001) An overview of non-central cameras. In: Computer vision winter workshop, vol 2

Bettonvil F (2005) Fisheye lenses. WGN, Journal of the International Meteor Organization 33:9–14

Boominathan V, Adams JK, Asif MS, Avants BW, Robinson JT, Baraniuk RG, Sankaranarayanan AC, Veeraraghavan

A (2016) Lensless imaging: A computational renaissance. IEEE Signal Processing Magazine 33(5):p23–35
Boominathan V, Adams JK, Robinson JT, Veeraraghavan A (2020) Phlatcam: Designed phase-mask based thin lensless camera. IEEE transactions on pattern analysis and machine intelligence 42(7):1618–1629
Borra EF (1982) The liquid-mirror telescope as a viable astronomical tool. Journal of the Royal Astronomical Society of Canada 76:p245–256
Brady DJ, Gehm ME, Stack RA, Marks DL, Kittle DS, Golish DR, Vera E, Feller SD (2012) Multiscale gigapixel photography. Nature 486(7403):p386–389
Brousseau D, Borra EF, Thibault S (2007) Wavefront correction with a 37-actuator ferrofluid deformable mirror. Optics express 15(26):p18190–18199
Cao JJ, Hou ZS, Tian ZN, Hua JG, Zhang YL, Chen QD (2020) Bioinspired zoom compound eyes enable variable-focus imaging. ACS Applied Materials & Interfaces 12(9):p10107–10117
Caron G, Eynard D (2011) Multiple camera types simultaneous stereo calibration. In: IEEE Int. Conf. on Robotics and Automation, pp p2933–2938
Chan WS, Lam EY, Ng MK (2006) Extending the depth of field in a compound-eye imaging system with super-resolution reconstruction. In: 18th Int. Conf. on Pattern Recognition (ICPR'06), IEEE, vol 3, pp 623–626
Cheng Y, Cao J, Zhang Y, Hao Q (2019) Review of state-of-the-art artificial compound eye imaging systems. Bioinspiration & biomimetics 14(3)
Clark AD, Wright W (1973) Zoom lenses. 7, Hilger
Conti C, Soares LD, Nunes P (2020) Dense light field coding: A survey. IEEE Access 8:p49244–49284
Corke P (2017) Robotics, vision and control: fundamental algorithms in MATLAB® second, completely revised, vol 118. Springer
Cossairt OS, Miau D, Nayar SK (2011) Gigapixel computational imaging. In: IEEE Int. Conf. on Computational Photography (ICCP), pp p1–8
Davies R, Kasper M (2012) Adaptive optics for astronomy. Annual Review of Astronomy and Astrophysics 50:p305–351
Demenikov M, Findlay E, Harvey AR (2009) Miniaturization of zoom lenses with a single moving element. Optics express 17(8):6118–6127
Dickie C, Fellion N, Vertegaal R (2012) Flexcam: Using thin-film flexible oled color prints as a camera array. In: CHI '12 Extended Abstracts on Human Factors in Computing Systems, Association for Computing Machinery, New York, NY, USA, CHI EA '12, p 1051–1054, DOI 10.1145/2212776.2212383, URL https://doi.org/10.1145/2212776.2212383
Evens L (2008) View camera geometry
Ewing J (2016) Follow the Sun: A Field Guide to Architectural Photography in the Digital Age. CRC Press
Falahati M, Zhou W, Yi A, Li L (2020) Development of an adjustable-focus ferrogel mirror. Optics & Laser Technology 125
Fayman JA, Sudarsky O, Rivlin E (1998) Zoom tracking. In: Proceedings. 1998 IEEE International Conference on Robotics and Automation (Cat. No. 98CH36146), IEEE, vol 4, pp 2783–2788
Fenimore EE, Cannon TM (1978) Coded aperture imaging with uniformly redundant arrays. Applied optics 17(3):p337–347
Flaugher B, Diehl H, Honscheid K, Abbott T, Alvarez O, Angstadt R, Annis J, Antonik M, Ballester O, Beaufore L, et al. (2015) The dark energy camera. The Astronomical Journal 150(5)
Floreano D, Pericet-Camara R, Viollet S, Ruffier F, Brückner A, Leitel R, Buss W, Menouni M, Expert F, Juston R, et al. (2013) Miniature curved artificial compound eyes. Proc of the National Academy of Sciences 110(23):p9267–9272
Ford JE, Agurok I, Stamenov I (2018) Monocentric lens designs and associated imaging systems having wide field of view and high resolution. US Patent 9,860,443
Forutanpour B, Le Nguyen PH, Bi N (2019) Dual fisheye image stitching for spherical image content. US Patent 10,275,928
Geyer C, Daniilidis K (2001) Catadioptric projective geometry. International journal of computer vision 45(3):223–243
Ghorayeb A, Potelle A, Devendeville L, Mouaddib EM (2010) Optimal omnidirectional sensor for urban traffic diagnosis in crossroads. In: IEEE Intelligent Vehicles Symposium, pp p597–602
Gill JS, Moosajee M, Dubis AM (2019) Cellular imaging of inherited retinal diseases using adaptive optics. Eye 33(11):p1683–1698
Gluckman J, Nayar SK (1999) Planar catadioptric stereo: Geometry and calibration. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR), vol 1, pp p22–28
Gottesman SR, Fenimore EE (1989) New family of binary arrays for coded aperture imaging. Applied optics 28(20):p4344–4352
Goy J, Courtois B, Karam J, Pressecq F (2001) Design of an APS CMOS image sensor for low light level applications using standard cmos technology. Analog Integrated Circuits and Signal Processing 29(1):95–104
Guenter B, Joshi N, Stoakley R, Keefe A, Geary K, Freeman R, Hundley J, Patterson P, Hammon D, Herrera G, et al. (2017) Highly curved image sensors: a practical approach for improved optical performance. Optics Express 25(12):p13010–13023
Hamanaka K, Koshi H (1996) An artificial compound eye using a microlens array and its application to scale-invariant processing. Optical Review 3(4):264–268
Held RT, Cooper EA, O'brien JF, Banks MS (2010) Using blur to affect perceived distance and size. ACM Trans on graphics 29(2)
Herbert G (1960) Apparatus for orienting photographic images. US Patent 2,931,268
Hicks RA, Perline RK (2001) Geometric distributions for catadioptric sensor design. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR), vol 1
Hornbeck LJ (1983) 128× 128 deformable mirror device. IEEE Trans on Electron Devices 30(5):539–545
Hsu CH, Cheng WH, Wu YL, Huang WS, Mei T, Hua KL (2017) Crossbowcam: a handheld adjustable multi-camera system. Multimedia Tools and Applications 76(23):p24961–24981
Hua Y, Nakamura S, Asif S, Sankaranarayanan A (2020) Sweepcam-depth-aware lensless imaing using programmable masks. IEEE Trans on Pattern Analysis and Machine Intelligence (TPAMI)
Huang G, Jiang H, Matthews K, Wilford P (2013) Lensless imaging by compressive sensing. In: 2013 IEEE International Conference on Image Processing, IEEE, pp 2101–2105

Huisman R, Bruijn M, Damerio S, Eggens M, Kazmi S, Schmerbauch A, Smit H, Vasquez-Beltran M, van der Veer E, Acuautla M, et al. (2020) High pixel number deformable mirror concept utilizing piezoelectric hysteresis for stable shape configurations. arXiv preprint arXiv:200809338

Ihrke I, Restrepo J, Mignard-Debise L (2016) Principles of light field imaging: Briefly revisiting 25 years of research. IEEE Signal Processing Magazine 33(5):59–69

Indebetouw G, Bai H (1984) Imaging with fresnel zone pupil masks: extended depth of field. Applied Optics 23(23):p4299–4302

Ishihara K (2015) Imaging apparatus having a curved image surface. US Patent 9,104,018

Jeon HG, Park J, Choe G, Park J, Bok Y, Tai YW, So Kweon I (2015) Accurate depth map estimation from a lenslet light field camera. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1547–1555

Joo H, Liu H, Tan L, Gui L, Nabbe B, Matthews I, Kanade T, Nobuhara S, Sheikh Y (2015) Panoptic studio: A massively multiview system for social motion capture. In: Proc. of the IEEE Int. Conf. on Computer Vision, pp p3334–3342

Jung GS, Won YH (2020) Simple and fast field curvature measurement by depth from defocus using electrowetting liquid lens. Applied Optics 59(18):5527–5531

Jung I, Xiao J, Malyarchuk V, Lu C, Li M, Liu Z, Yoon J, Huang Y, Rogers JA (2011) Dynamically tunable hemispherical electronic eye camera system with adjustable zoom capability. Proceedings of the National Academy of Sciences 108(5):1788–1793

Kahn S, Kurita N, Gilmore K, Nordby M, O'Connor P, Schindler R, Oliver J, Van Berg R, Olivier S, Riot V, et al. (2010) Design and development of the 3.2 gigapixel camera for the Large Synoptic Survey Telescope. In: Ground-based and airborne instrumentation for astronomy III, Int. Society for Optics and Photonics, vol 7735

Kang S, Duocastella M, Arnold CB (2020) Variable optical elements for fast focus control. Nature Photonics 14(9):p533–542

Kim DH, Kim YS, Wu J, Liu Z, Song J, Kim HS, Huang YY, Hwang KC, Rogers JA (2009) Ultrathin silicon circuits with strain-isolation layers and mesh layouts for high-performance electronics on fabric, vinyl, leather, and paper. Advanced Materials 21(36):3703–3707

Kim JJ, Liu H, Ashtiani AO, Jiang H (2020) Biologically inspired artificial eyes and photonics. Reports on Progress in Physics 83(4):047101

Kim WY, Seo HT, Kim S, Kim KS (2019) Practical approach for controlling optical image stabilization system. International Journal of Control, Automation and Systems pp 1–10

Kingslake R (1992) Optics in photography. SPIE press

Ko HC, Stoykovich MP, Song J, Malyarchuk V, Choi WM, Yu CJ, Geddes Iii JB, Xiao J, Wang S, Huang Y, et al. (2008) A hemispherical electronic eye camera based on compressible silicon optoelectronics. Nature 454(7205):p748–753

Koppelhuber A, Bimber O (2013) Towards a transparent, flexible, scalable and disposable image sensor using thin-film luminescent concentrators. Optics express 21(4):4796–4810

Koppelhuber A, Bimber O (2017) Thin-film camera using luminescent concentrators and an optical söller collimator. Optics Express 25(16):p18526–18536

Korneliussen JT, Hirakawa K (2014) Camera processing with chromatic aberration. IEEE Transactions on Image Processing 23(10):4539–4552

Krishnan G, Nayar SK (2008) Cata-fisheye camera for panoramic imaging. In: IEEE Workshop on Applications of Computer Vision, pp 1–8

Krishnan G, Nayar SK (2009) Towards a true spherical camera. In: Human Vision and Electronic Imaging XIV, Int. Society for Optics and Photonics, vol 7240

Kurmi I, Schedl DC, Bimber O (2018) Micro-lens aperture array for enhanced thin-film imaging using luminescent concentrators. Optics Express 26(22):29253–29261

Kuthirummal S, Nayar SK (2006) Multiview radial catadioptric imaging for scene capture. In: ACM SIGGRAPH 2006 Papers, Association for Computing Machinery, New York, NY, USA, SIGGRAPH '06, p 916–923, DOI 10.1145/1179352.1141975, URL https://doi.org/10.1145/1179352.1141975

Kuthirummal S, Nayar SK (2007) Flexible mirror imaging. In: IEEE 11th Int. Conf. on Computer Vision, pp p1–8

Kuthirummal S, Nagahara H, Zhou C, Nayar SK (2010) Flexible depth of field photography. IEEE Transactions on Pattern Analysis and Machine Intelligence 33(1):58–71

Langford M (2000) Basic photography. Taylor & Francis

Layerle JF, Savatier X, Mouaddib E, Ertaud JY (2008) Catadioptric sensor for a simultaneous tracking of the driver s face and the road scene. In: In OMNIVIS'2008, the Eighth Workshop on Omnidirectional Vision, Camera Networks and Non-classical Cameras, in conjunction with ECCV 2008

Lee GJ, Nam WI, Song YM (2017) Robustness of an artificially tailored fisheye imaging system with a curvilinear image surface. Optics & Laser Technology 96:p50–57

Lee J, Wu J, Shi M, Yoon J, Park SI, Li M, Liu Z, Huang Y, Rogers JA (2011) Stretchable gaas photovoltaics with designs that enable high areal coverage. Advanced Materials 23(8):986–991

Lee JH, You BG, Park SW, Kim H (2020) Motion-free TSOM using a deformable mirror. Optics Express 28(11):p16352–16362

Lenk L, Mitschunas B, Sinzinger S (2019) Zoom systems with tuneable lenses and linear lens movements. Journal of the European Optical Society-Rapid Publications 15(1):1–10

Levoy M (2006) Light fields and computational imaging. Computer 39(8):p46–55

Li L, Wang D, Liu C, Wang QH (2016) Zoom microscope objective using electrowetting lenses. Optics express 24(3):2931–2940

Li L, Hao Y, Xu J, Liu F, Lu J (2018) The design and positioning method of a flexible zoom artificial compound eye. Micromachines 9(7)

Liang CK, Lin TH, Wong BY, Liu C, Chen HH (2008) Programmable aperture photography: Multiplexed light field acquisition. In: ACM SIGGRAPH 2008 Papers, Association for Computing Machinery, New York, NY, USA, SIGGRAPH '08, DOI 10.1145/1399504.1360654, URL https://doi.org/10.1145/1399504.1360654

Liang J (2020) Punching holes in light: recent progress in single-shot coded-aperture optical imaging. Reports on Progress in Physics 83(11)

Lin YH, Chen MS, Lin HC (2011) An electrically tunable optical zoom system using two composite liquid crystal lenses with a large zoom ratio. Optics express 19(5):4714–4721

Lin YH, Liu YL, Su GDJ (2012) Optical zoom module based on two deformable mirrors for mobile device applications.

Applied Optics 51(11):1804–1810
Lipton L, Meyer LD (1992) Stereoscopic video camera with image sensors having variable effective position. US Patent 5,142,357
Madec PY (2012) Overview of deformable mirror technologies for adaptive optics and astronomy. In: Adaptive Optics Systems III, Int. Society for Optics and Photonics, vol 8447
Maître H (2017) From photon to pixel: the digital camera handbook. John Wiley & Sons
Marchand E, Chaumette F (2017) Visual servoing through mirror reflection. In: 2017 IEEE International Conference on Robotics and Automation (ICRA), IEEE, pp 3798–3804
Martin CB (2004) Design issues of a hyperfield fisheye lens. In: Novel Optical Systems Design and Optimization VII, Int. Society for Optics and Photonics, vol 5524, pp p84–92
McLeod B, Geary J, Conroy M, Fabricant D, Ordway M, Szentgyorgyi A, Amato S, Ashby M, Caldwell N, Curley D, et al. (2015) Megacam: A wide-field CCD imager for the MMT and Magellan. Publications of the Astronomical Society of the Pacific 127(950):366
Merklinger HM (1996) Focusing the view camera. Seaboard Printing Limited 5
Moon S, Keles HO, Ozcan A, Khademhosseini A, Hæggstrom E, Kuritzkes D, Demirci U (2009) Integrating microfluidics and lensless imaging for point-of-care testing. Biosensors and Bioelectronics 24(11):p3208–3214
Motamedi ME (2005) MOEMS: Micro-opto-electro-mechanical Systems, vol 126. SPIE press
Mouaddib EM, Sagawa R, Echigo T, Yagi Y (2005) Stereovision with a single camera and multiple mirrors. In: Proc. of the IEEE Int. Conf. on Robotics and Automation, pp p800–805
Mutze U (2000) Electronic camera for the realization of the imaging properties of a studio bellow camera. US Patent 6,072,529
Nagahara H, Zhou C, Watanabe T, Ishiguro H, Nayar SK (2010) Programmable aperture camera using lcos. In: European Conference on Computer Vision, Springer, pp 337–350
Nakamura T, Horisaki R, Tanida J (2015) Compact wide-field-of-view imager with a designed disordered medium. Optical Review 22(1):p19–24
Nakamura T, Kagawa K, Torashima S, Yamaguchi M (2019) Super field-of-view lensless camera by coded image sensors. Sensors 19(6)
Nanjo Y, Sueyoshi M (2011) Tilt lens system and image pickup apparatus. US Patent 7,880,797
Nayar SK (1997) Catadioptric omnidirectional camera. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR), pp p482–488
Nayar SK, Branzoi V, Boult TE (2006) Programmable imaging: Towards a flexible camera. Int Journal of Computer Vision 70(1):p7–22
Newman PA, Rible VE (1966) Pinhole array camera for integrated circuits. Applied optics 5(7):1225–1228
Ng R, Levoy M, Brédif M, Duval G, Horowitz M, Hanrahan P (2005) Light field photography with a hand-held plenoptic camera. PhD thesis, Stanford University
Ng TN, Wong WS, Chabinyc ML, Sambandan S, Street RA (2008) Flexible image sensor array with bulk heterojunction organic photodiode. Applied Physics Letters 92(21):191
Nomura Y, Zhang L, Nayar SK (2007) Scene collages and flexible camera arrays. In: Proc. of the 18th Eurographics Conf. on Rendering Techniques, pp p127–138
Okumura M (2017) Panoramic-imaging digital camera, and panoramic imaging system. US Patent 9,756,244
Olagoke AS, Ibrahim H, Teoh SS (2020) Literature survey on multi-camera system and its application. IEEE Access 8:p172892–172922
Papadopoulos IN, Farahi S, Moser C, Psaltis D (2013) High-resolution, lensless endoscope based on digital scanning through a multimode optical fiber. Biomedical optics express 4(2):p260–270
Phillips SJ, Kelley DL, Prassas SG (1984) Accuracy of a perspective control lens. Research Quarterly for Exercise and Sport 55(2):197–200
Poling B (2015) A tutorial on camera models. University of Minnesota pp 1–10
Potmesil M, Chakravarty I (1982) Synthetic image generation with a lens and aperture camera model. ACM Trans on Graphics (TOG) 1(2):p85–108
Poulsen A (2011) Tilt and shift adaptor, camera and image correction method. US Patent App. 13/000,591
Qaisar S, Bilal RM, Iqbal W, Naureen M, Lee S (2013) Compressive sensing: From theory to applications, a survey. Journal of Communications and networks 15(5):443–456
Qian K, Miura K (2019) The roles of color lightness and saturation in inducing the perception of miniature faking. In: 2019 11th Int. Conf. on Knowledge and Smart Technology (KST), IEEE, pp 194–198
Ritcey AM, Borra E (2010) Magnetically deformable liquid mirrors from surface films of silver nanoparticles. ChemPhysChem 11(5):p981–986
Roberts G (1995) A real-time response of vigilance behaviour to changes in group size. Animal Behaviour 50(5):1371–1374, DOI https://doi.org/10.1016/0003-3472(95)80052-2, URL `https://www.sciencedirect.com/science/article/pii/0003347295800522`
Roddier F (1999) Adaptive optics in astronomy. Cambridge university press
Sadlo F, Dachsbacher C (2011) Auto-Tilt Photography. In: Eisert P, Hornegger J, Polthier K (eds) Vision, Modeling, and Visualization (2011), The Eurographics Association, DOI 10.2312/PE/VMV/VMV11/239-246
Sahin FE, Laroia R (2017) Light L16 Computational Camera. In: Applied Industrial Optics: Spectroscopy, Imaging and Metrology, Optical Society of America, pp JTu5A–20
Saito H, Hoshino K, Matsumoto K, Shimoyama I (2005) Compound eye shaped flexible organic image sensor with a tunable visual field. In: 18th IEEE International Conference on Micro Electro Mechanical Systems, 2005. MEMS 2005., IEEE, pp 96–99
Sako S, Osawa R, Takahashi H, Kikuchi Y, Doi M, Kobayashi N, Aoki T, Arimatsu K, Ichiki M, Ikeda S, et al. (2016) Development of a prototype of the Tomo-e Gozen wide-field CMOS camera. In: Ground-based and Airborne Instrumentation for Astronomy VI, Int. Society for Optics and Photonics, vol 9908
Scheimpflug T (1904) Improved method and apparatus for the systematic alteration or distortion of plane pictures and images by means of lenses and mirrors for photography and for other purposes. GB patent 1196
Scholz E, Bittner WAA, Caldwell JB (2014) Tilt shift lens adapter. US Patent 8,678,676
Schulz A (2015) Architectural photography: composition, capture, and digital image processing. Rocky Nook, Inc.
Schwarz A, Wang J, Shemer A, Zalevsky Z, Javidi B (2015) Lensless three-dimensional integral imaging using variable and time multiplexed pinhole array. Optics letters

40(8):1814–1817
Schwarz A, Wang J, Shemer A, Zalevsky Z, Javidi B (2016) Time multiplexed pinhole array based lensless three-dimensional imager. In: Three-Dimensional Imaging, Visualization, and Display 2016, International Society for Optics and Photonics, vol 9867, p 98670R
Sims DC, Yue Y, Nayar SK (2016) Towards flexible sheet cameras: Deformable lens arrays with intrinsic optical adaptation. In: IEEE Int. Conf. on Computational Photography (ICCP), pp p1–11
Sims DC, Cossairt O, Yu Y, Nayar SK (2018) Stretchcam: Zooming using thin, elastic optics. arXiv preprint arXiv:180407052
Someya T, Kato Y, Iba S, Noguchi Y, Sekitani T, Kawaguchi H, Sakurai T (2005) Integration of organic fets with organic photodiodes for a large area, flexible, and lightweight sheet image scanners. IEEE transactions on electron devices 52(11):2502–2511
Stroebel L (1999) View camera technique. CRC Press
Strom SE, Stepp LM, Gregory B (2003) Giant segmented mirror telescope: a point design based on science drivers. In: Future Giant Telescopes, Int. Society for Optics and Photonics, vol 4840, pp 116–128
Tan KH, Hua H, Ahuja N (2004) Multiview panoramic cameras using mirror pyramids. IEEE Trans on Pattern Analysis and Machine Intelligence (TPAMI) 26(7):p941–946
Tilmon B, Jain E, Ferrari S, Koppal S (2020) FoveaCam: A MEMS Mirror-Enabled Foveating Camera. In: IEEE Int. Conf. on Computational Photography (ICCP), pp p1–11
Tyson RK (2015) Principles of adaptive optics. CRC press
Wang D, Pan Q, Zhao C, Hu J, Xu Z, Yang F, Zhou Y (2017) A study on camera array and its applications. IFAC-PapersOnLine 50(1):p10323–10328
Wang X, Ji J, Zhu Y (2021) A zoom tracking algorithm based on deep learning. Multimedia Tools and Applications pp 1–19
Wilburn B, Joshi N, Vaish V, Talvala EV, Antunez E, Barth A, Adams A, Horowitz M, Levoy M (2005) High performance imaging using large camera arrays. In: ACM SIGGRAPH 2005 Papers, Association for Computing Machinery, New York, NY, USA, SIGGRAPH '05, p 765–776, DOI 10.1145/1186822.1073259, URL https://doi.org/10.1145/1186822.1073259
Wu S, Jiang T, Zhang G, Schoenemann B, Neri F, Zhu M, Bu C, Han J, Kuhnert KD (2017) Artificial compound eye: a survey of the state-of-the-art. Artificial Intelligence Review 48(4):p573–603
Xie S, Wang P, Sang X, Li C (2016) Augmented reality three-dimensional display with light field fusion. Optics express 24(11):p11483–11494
Young M (1971) Pinhole optics. Applied Optics 10(12):2763–2767
Yu G, Wang J, McElvain J, Heeger AJ (1998) Large-area, full-color image sensors made with semiconducting polymers. Advanced Materials 10(17):1431–1434
Zamkotsian F, Dohlen K (2002) Prospects for moems-based adaptive optical systems on extremely large telescopes. In: European Southern Observatory Conference and Workshop Proceedings, vol 58, p 293
Zhang C, Chen T (2004) A self-reconfigurable camera array. In: Proceedings of the Fifteenth Eurographics conference on Rendering Techniques, pp 243–254
Zhang K, Jung YH, Mikael S, Seo JH, Kim M, Mi H, Zhou H, Xia Z, Zhou W, Gong S, et al. (2017) Origami silicon optoelectronics for hemispherical electronic eye systems. Nature communications 8(1):p1–8
Zhao H, Fan X, Zou G, Pang Z, Wang W, Ren G, Du Y, Su Y (2013) All-reflective optical bifocal zooming system without moving elements based on deformable mirror for space camera application. Applied Optics 52(6):p1192–1210
Zhao J, Zhang Y, Li X, Shi M (2019) An improved design of the substrate of stretchable gallium arsenide photovoltaics. Journal of Applied Mechanics 86(3):031009
Zheng Y, Lin S, Kambhamettu C, Yu J, Kang SB (2008) Single-image vignetting correction. IEEE transactions on pattern analysis and machine intelligence 31(12):2243–2256
Zhou C, Nayar S (2009) What are good apertures for defocus deblurring? In: IEEE Int. Conf. on Computational Photography (ICCP), pp p1–8
Zhou C, Nayar SK (2011) Computational cameras: convergence of optics and processing. IEEE Trans on Image Processing 20(12):p3322–3340
Zhou C, Lin S, Nayar S (2009) Coded aperture pairs for depth from defocus. In: IEEE 12th Int. Conf. on Computer Vision (ICCV), pp p325–332
Zomet A, Nayar SK (2006) Lensless imaging with a controllable aperture. In: IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR), IEEE, vol 1, pp p339–346
Zou T, Tang X, Song B, Wang J, Chen J (2012) Robust feedback zoom tracking for digital video surveillance. Sensors 12(6):8073–8099

Technical References

ALPAO®(2023) Deformable mirrors. URL https://www.alpao.com/adaptive-optics/deformable-mirrors.html
Asphericon®(2023) Parabolic mirror. URL https://www.asphericon.com/en/products/parabolic-mirror?\gclid=CjwKCAiA2O39BRBjEiwApB2Iks5rpILOEag23Q_Qd5UR2ENSYIrzqCO3xrHG-w_b7dQeAxyqg0s0RBoCrecQAvD_BwE
Boston Micromachines® (2023) Standard deformable mirrors. URL https://bostonmicromachines.com/standard-deformable-mirrors/
Canon®(2022) Tilt-shift lenses. URL https://www.canon-europe.com/lenses/tilt-and-shift-lenses/
Canon®(2023) RF24-105mm F4-7.1 IS STM. URL https://www.usa.canon.com/internet/portal/us/home/products/details/lenses/ef/standard-zoom/rf24-105mm-f4-7-1-is-stm
Cilas®(2023) Cilas, expert in laser and optronics. URL https://www.cilas.com/
Dxomark®(2021) All camera tests. URL https://www.dxomark.com/category/mobile-reviews/#tabs
Entaniya®(2021) Industrial fisheye lens m12 280. URL https://products.entaniya.co.jp/en/products/entaniya-fisheye-m12-280/
European Southern Observatory (ESO) (2018) The european extremely large telescope ("elt") project. URL http://www.eso.org/sci/facilities/eelt/
GoPro®(2015) Happy Sweet Sixteen: Introducing GoPro's 360° Camera Array. URL https://gopro.com/en/us/news/happy-sweet-sixteen-introducing-gopros-360-camera-array
Hitachi®(2017) Lensless camera – innovative camera supporting the evolution of iot. URL https://www.hitachi.com/rd/sc/story/lensless/index.html
Lawrence Livermore National Laboratory (2019) World's largest optical lens shipped to slac. URL https://www.

llnl.gov/news/world%E2%80%99s-largest-optical-lens-shipped-slac
Lightfield Forum® (2023) Lytro lightfield camera. URL http://lightfield-forum.com/lytro/lytro-lightfield-camera/
Light®(2019) The technology behind the l16 camera pt. 1. URL https://support.light.co/l16-photography/l16-tech-part-1
Microgate®(2011) Adaptive deformable mirrors. URL https://engineering.microgate.it/en/adaptive-deformable-mirrors/concept
NASA (2023) James webb space telescope. URL https://www.jwst.nasa.gov/
Nikkor®(2022) PC Nikkor 35mm f/2.8. URL http://cdn-10.nikon-cdn.com/pdf/manuals/archive/PC-Nikkor%2035%20mm%20f-2.8.pdf
Nikkor®(2023) AF-S Fisheye NIKKOR 8-15mm f/3.5-4.5E ED. URL https://www.nikon.co.in/af-s-fisheye-nikkor-8-15mm-f-3-5-4-5e-ed
Nikon®(2019) Nikon releases the NIKKOR Z 58mm f/0.95 S. URL https://www.nikon.com/news/2019/1010_lens_02.htm
Nikon®(2020) The PC Lens Advantage: What You See Is What You'll Get. URL https://www.nikonusa.com/en/learn-and-explore/a/tips-and-techniques/the-pc-lens-advantage-what-you-see-is-what-youll-get.html
Nikon®(2022) Coolpix P1000. URL https://www.nikonusa.com/en/nikon-products/product/compact-digital-cameras/coolpix-p1000.html
Nikon®(2023) PC Nikkor 19mm f/4E ED. URL https://www.nikonusa.com/en/nikon-products/product/camera-lenses/pc-nikkor-19mm-f%252f4e-ed.html
Observatory WMK (2023) Keck I And Keck II Telescopes. URL https://keckobservatory.org/about/telescopes-instrumentation/
Panasonic®(2022) G Series 14-42mm F3.5-5.6 ASPH X Vario Lens. URL https://shop.panasonic.com/products/g-series-14-42mm-f3-5-5-6-asph-x-vario-lens
Panasonic®(2023) LUMIX Compact System (Mirrorless) Camera DC-GX9 Body Only. URL https://www.panasonic.com/uk/consumer/cameras-camcorders/lumix-mirrorless-cameras/lumix-g-cameras/dc-gx9.specs.html
Raytrix®(2021) 3D Light-Field Cameras. URL https://raytrix.de/products/
Ricoh®(2019) Ricoh theta lineup comparison. URL https://theta360.com/en/about/theta/
Samsung®(2023) WB2200F 16MP Digital Camera with 60x optical zoom - Black. URL https://www.samsung.com/uk/support/model/EC-WB2200BPBGB/
Sony®(2023) FE C 16-35 MM T3.1: full specifications and features. URL https://pro.sony/ue_US/products/camera-lenses/selc1635g#ProductSpecificationsBlock-selc1635g
Standford (2020) Sensors of world's largest digital camera snap first 3,200-megapixel images at slac. URL https://www6.slac.stanford.edu/news/2020-09-08-sensors-world-largest-digital-camera-snap-first-3200-megapixel-images-slac.aspx
Texas Instruments® (2018) Introduction to ±12 Degree Orthogonal Digital Micromirror Devices (DMDs). URL https://www.ti.com/lit/an/dlpa008b/dlpa008b.pdf?ts
Vstone®(2010) Omni Directional Sensor. URL https://www.vstone.co.jp/english/products/sensor_camera/
Zeiss®(2021) Making sense of sensors – full frame vs. aps-c. URL https://lenspire.zeiss.com/photo/en/article/making-sense-of-sensors-full-frame-vs-aps-c