Preface

The advances in micro and nanotechnologies nowadays are nothing less than spectacular. Materials, structures, devices, and systems never seen before are appearing at an accelerated pace, while new properties and novel phenomena are being discovered almost every day. The microelectromechanical system (MEMS) and microoptoelectromechanical (MOEMS) technologies have brought a multitude of devices with ultracompact dimensions and with vastly enhanced or even completely novel properties and characteristics. Nanotechnologies are continuously expanding and improving the toolbox of available materials and systems, smudging the limits between technological and natural, even biological, and the potentials for further expansion appear literally limitless. The quest toward ever smaller dimensions and ever-increasing control over structures and properties is continuing and is already having a profound influence on practically all fields of applied science, steadily permeating more and more facets of ordinary life.

What is valid generally, is also valid in optoelectronics and photonics, actually maybe even more so than in some other fields. Completely novel areas are being created, ensuring manipulation with light at an unprecedented level. Micro and nanofabrication of structures with a periodicity of refractive index that may extend to all three dimensions have resulted in photonic bandgap materials, also known as photonic crystals, ensuring manipulation with photons in a manner similar to that with electrons in semiconductors. Subwavelength structuring of metal-dielectric materials brought metamaterials, structures possessing electromagnetic properties not readily met in nature and often unexpected and sometimes even counterintuitive. These include materials with values of refractive index below zero or equal to zero. It also created the field of plasmonics, dedicated to manipulation with evanescent, surface-bound electromagnetic waves that build a link between electronics and optics, ensuring devices with compact dimensions of micro and nanoelectronic circuits that operate at the speed of optical signals. Structures and devices are being created which until recently were deemed impossible.

The advent of micro and nanophotonics has brought astounding enhancements to different passive and active devices. In the field of infrared technologies, MEMS and NEMS, it ensured the appearance of a new generation of thermal detectors, with properties vastly surpassing traditional devices. Novel nanodevices are being produced where quantum structures are used to reach new levels of performance.

However, there is a glaring gap in the list of infrared devices enhanced by the micro and nanophotonic technologies. For decades, simple intrinsic photonic (semiconductor) detectors have been far superior to other devices for infrared detection. Even today they remain the optimum choice if one simultaneously requires sensitivity, selectivity, and speed at an acceptable cost. Yet relatively little has been done to improve their performance using the same technologies that brought the renaissance to so many other fields. There is no obvious reason why micro and nanophotonics should not bring about similar improvements to photonic devices that were superior to begin with. The results actually do exist, but are scattered throughout the literature and are sometimes quite underwhelming. The author is not aware of a systematic treatise dedicated to the application of novel approaches stemming from microphotonic and nanophotonic technologies to traditional semiconductor infrared detectors. The goal of this book is to fill this gap.

Nowadays, two large classes of optoelectronic devices are used for detection of mid- (MWIR) and long-wavelength (LWIR) infrared radiation in the atmospheric windows (3–5) and (8–14) μm: photonic (quantum) detector and thermal detectors. Photonic detectors have been ruling the market for several decades, mostly of intrinsic type, based on indium antimonide and mercury cadmium telluride. Besides having high values of specific detectivity and generally high performance, they are characterized by high response speeds. Even today they largely outperform the novel generation of thermal infrared detectors.

The largest drawback of photonic detectors of infrared radiation compared to thermal devices is their requirement for cooling. For this purpose one typically uses Dewar flasks with liquid nitrogen or various kinds of active coolers, including Joule–Thompson cryostats or multistage thermoelectric coolers. Such cooling systems obviously make the detector units overly complex, impractical for handling, and vastly increase their price.

It is no wonder that through the decades a large part of the research and development connected with semiconductor detector elements for mid- and long-wavelength infrared radiation was dedicated exactly to the question of how to fabricate devices with high specific detectivities and high speeds at elevated temperatures (the “HOT”—Higher Operating Temperature detectors). The seemingly simple task of fabricating uncooled infrared detectors with background-limited operation engaged armies of researchers and for some of them became the holy grail of photonic detector design.

This book is dedicated to the same task, observed however through the prism of modern microphotonic and nanophotonic technologies. We offer a systematic approach to different available methods for increase in the operating temperatures of photonic detectors up to room temperature, while simultaneously retaining the performance of cryogenically cooled devices. The avoidance of cooling requirements would reduce the detector price to a fraction of the previous one, but the main gain should be expended in improved performance, higher robustness and ease of use, and much more widespread utilization.

It is necessary to mention that this text is dedicated solely to intrinsic photonic infrared detectors. This means that we do not consider more complex structures of detector materials themselves, which include different nanocomposites like quantum wells, quantum wires, and quantum dots. The topic of this work is micro and nanophotonic technologies, and not advanced detector materials, although they also represent a huge field for potential improvements. However, practically all of the methods and techniques described here are also applicable to devices based on advanced detector materials.

An important motivation for this kind of book is obviously the wide and ever expanding applicability of infrared detectors, especially those for the long-wavelength range. Naturally, in today's society all kinds of sensors are in increasing demand, but infrared devices assume a special position.

The situation was quite different in the past. The majority of the MWIR and LWIR detectors were used for military applications, for implementations like night vision, heat seeking and guidance, etc. Materials used for these detectors were also specific and were mostly limited to narrow-bandgap semiconductors. These detectors were also used for lab work and generally scientific applications in physics, meteorology, geology, material science, astronomy, remote sensing, airspace research, etc. Mostly their users were highly demanding. This explains why cooled detectors became widespread—for such users the price was most often, albeit not always, second to performance.

Usually, a large field of use for MWIR and LWIR detectors is overlooked—the free-space optical (FSO) telecommunications. This is certainly unusual when one bears in mind that communications themselves are one of the pillars of civilization itself, actually one could maintain that communications were those that created civilization in the first place.

In the field of free-space communication a special position belongs to microwave and optical wavelength ranges. In highly urbanized environments the so-called last-mile solutions are becoming increasingly important—instead of having to build transmission lines into the already existing infrastructure, which often requires astronomical expenses, it is much easier to utilize various wireless approaches. Since the majority of the existing FSO systems belong to the near infrared range, the available bands are becoming quite cramped. The long-wavelength IR range is thus mentioned as an alternative. This range is less sensitive to atmospheric disturbances like fog, smoke, rain, etc.; LWIR wavelengths are much less likely to accidentally damage eyes since vastly higher radiation thresholds are tolerated in this range. Last but not least, the LWIR range is not yet covered by strict regulations: it is literally a new territory waiting to be conquered. Low-cost uncooled detectors with high performance would be a significant step toward a more widespread use of such systems.

The applicability of the most of the methods presented in this book does not end with infrared detectors and is in reality much wider. An obvious and important field of use of different micro and nanophotonic methods for detection improvement elaborated in this book are solar cells, especially thin-film photovoltaics. Solar energy harvesting represents a wide area of its own, but many of the approaches presented here may be applied to it with little or no modifications.

This book is organized into three chapters. The first chapter is introductory and considers various demands that may be posed to a general infrared detector, as well as which real types of detectors satisfy these demands best. General figures of merit of infrared detectors are analyzed and stress is given to the product of specific detectivity and bandwidth, which is taken to be one of the key factors of detector quality. The work then proceeds to present various particular mechanisms that define the mentioned figures of merit, especially the prevailing generation-recombination processes and the related noise mechanisms. The goal of this chapter is to estimate the most convenient choice among different infrared detectors and to consider which of its figures of merit could be optimized and how.

The second chapter of the book analyzes the possibility to enhance the detector performance utilizing photon management or light management. These methods may be denoted as equilibrium ones, since they do not generate excess charge carrier concentrations for their function and are related solely to handling of light flux outside the detector and within it. The equilibrium methods include various light trapping schemes both in the far field and in the near field with a goal to enhance the detector response. Since there are numerous approaches to this kind of enhancement, after a short literature review a possible classification of these approaches is proposed. Besides including all photon management methods known to the author, the use of the classification points out to some possible novel approaches, as well as to the use of certain methods that until now were utilized only in other fields of optical engineering.

Photon management methods are divided into two large groups. One group comprises structures that increase the incident optical flux. Among these, the text first considers optical concentrators, which include refractive and diffractive lenses, either separate or integrated with the detector, as well as some reflective nonimaging concentrators. The next subgroup is structures for decrease in reflection at the incident plane, which include interference antireflection films and diffractive subwavelength antireflection structures. An interesting new field in this domain is the use of plasmonics which brings optical concentration to the micro and nano levels—the enhancement occurs in the near field and utilizes evanescent waves.

The second equilibrium group encompasses structures for increase in the optical path of the beam which already entered the active area of the detector, the so-called light trapping structures. These structures simultaneously increase radiative lifetime through the mechanism of reabsorption—photon recycling. They include different surface relief structures for the increase in total internal reflection, from diffractive to macroscopic ones. Reflective detector surfaces also belong to this group, both the back-side ones and full resonant cavities (RCE—resonant cavity enhancement) with reflective surfaces both of the front and the back side of the detector. The most advanced structures for optical path and radiative time increase are radiative shields and photonic crystal enhancement structures, which represent a full cavity enhancement and may support the existence of multiple modes.

The third chapter of the book considers nonequilibrium methods for thermal noise management, which include manipulation with charge carrier concentration within the detectors utilizing external fields and which are implemented through the application of general micro and nanosystem techniques. These are very important, since they attack the very reason why photonic detectors are cooled. Since the operating wavelengths are long, i.e., the energies of incident photons are low, the bandgap of the utilized material must also be small. This means that thermal processes of carrier generation will start to compete with the detection process itself and thus cause excessive increase of generation-recombination noise. This is especially valid for noise caused by Auger processes that prevail in sufficiently pure semiconductors. The goal of nonequilibrium methods is to cause controlled perturbation of carrier distribution within the detector and thus locally decrease the carrier concentration below its equilibrium value, thereby directly decreasing the level of total generation-recombination noise. A systematic approach to these methods and a newly proposed classification stemming from it again show that in addition to the methods published in the literature one could utilize some new ones.

Special care is dedicated to the consideration of properties common to all types of nonequilibrium devices. Based on this a general model is defined to be utilized in simulation of any nonequilibrium detector, including the novel ones. The model is derived for isothermal processes, but can be extended to nonisothermal ones. When deriving the model, special care is dedicated to the limits of the model and the generality of the introduced approximations.

The third chapter considers the use of contact phenomena, i.e., structures utilizing a combination of a built-in dopant concentration gradient and the external electric field for suppression of Auger processes. Two main types are considered, exclusion-based photoconductors and extraction-based photodiodes. The text further considers the use of crossed electric and magnetic field and analyzes the use of galvanomagnetic phenomena, i.e., the magnetoconcentration effect for suppression of Auger processes. Finally, conclusions are drawn and assessment of possible future directions is given.

Contents

Chapter 2
Photon Management

2.1 Fundamentals of Photon Management in Photodetectors

2.1.1 Introduction

The optical methods of optimization have a goal to maximize the quantum efficiency of a photodetector for a given geometry (thickness and active area) and a given spatial and frequency dependence of the absorption coefficient. They are described as photon management or light management methods.

Various photon management methods considered in this chapter are mostly generally applicable, not only for infrared semiconductor devices. We consider a general case of a photodetector as a device that converts optical energy into another form of energy. Most often this energy is electrical signal, although other forms may be used like thermal [6], motion (like in some microcantilever-based detectors) [75], optical signal at another frequency (up- or down-converted) [76, 77], etc. A notable example of devices to which a majority of the methods described here can be applied are thin-film solar cells [78–81].

There are different methods of light management in a photodetector that can be divided into four groups, Fig. 2.1: optical concentration, use of antireflection structures, optical path increase, and light localization.

Optical concentration methods actually collect incident radiation from a larger area (denoted as “optical” area) and concentrate it (focus) to the smaller active area of photodetector (the “electrical” area). The concentration efficiency can be then defined as the ratio between the optical and the electrical area, minus absorption and scattering losses. This method of light management is typically done by utilizing structures which are not themselves a part of the detector, but can be integrated with it. The simplest case of a concentrator would be an immersion lens, but there are a number of different other structures to serve the same purpose. Roughly, one could

Z. Jakšić, *Micro and Nanophotonics for Semiconductor Infrared Detectors*,
DOI 10.1007/978-3-319-09674-2_2

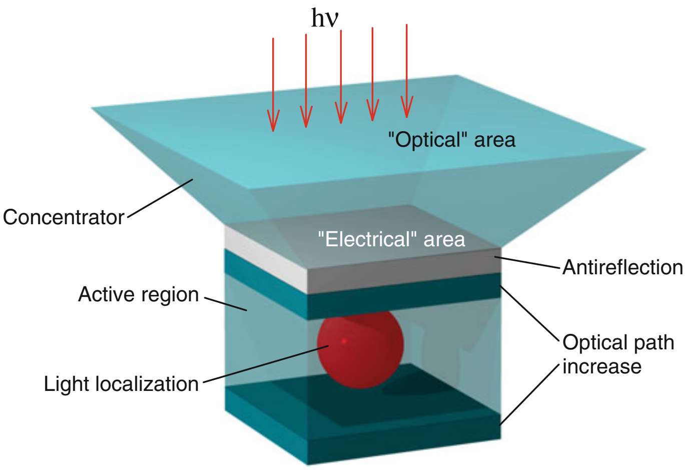


Fig. 2.1 Methods of photon management: use of optical concentrator, antireflection structure, structures for optical path increase (cavity enhancement) and light localization structures

divide them into refractive, reflective, and diffractive elements, although hybrid solutions are also possible.

After arriving to the detector surface, the incident beam should enter the photodetector area with losses as low as possible. Semiconductor materials used for photodetectors have large values of refractive index and thus large values of reflection coefficient at the device surface. This reflection is minimized using antireflection structures. Basically, these structures serve as impedance matching media between the detector environment (most often, but not always free space) and the active region.

Once useful radiation enters the detector-active region, the goal is to keep it there as long as possible and thus obtain maximum interaction between it and the active material. Probably the simplest way to do that is to place a mirror at the back side of the detector, thus doubling the optical path through the active region. However, the methods of optical path increase can be much more sophisticated and may include different cavities with reflective walls, as well as surface structuring. These methods belong to the optical trapping approaches.

There is another method of optical trapping, enabled by the advent of nanophotonics: the subwavelength optical localization utilizing plasmonic nanocomposites. Some metal-dielectric structures ensure the possibility of light localization on a level much smaller than the operating wavelength. To this purpose they utilize the propagating or localized surface plasmons polaritons.

This chapter investigates in some detail all of the mentioned methods of photodetector enhancement.

2.1.2 Fundamental Limits of Photon Management

Basically, all photon management techniques are intended to improve absorption of light in the detector, thus ensuring a higher degree of the conversion of incident optical radiation into a useful signal. Obviously, the efficiency of any conversion is limited by basic physical laws. A question is posed what are the fundamental limits of photodetector enhancement through light management.

A detector system is presented in Fig. 2.2. A background optical flux is incident to the active area of a photodetector with a thickness d. Both in the case of solar cells and night vision photodetectors, the optical flux is blackbody radiation, described by the Planck's law. In a general case, the detector material may incorporate nanostructuring that could localize optical field and create hotspots with high density of states. A perfect mirror is placed at the rear side of the device—i.e., it is assumed that the incident light is unidirectional, while the internal radiation is bidirectional.

The detector surface is corrugated in order to increase the optical path through the detector. The corrugation may be random or ordered, but in both cases its basic purpose is to change the direction of light incident upon the active surface and to make use of total internal reflection to ensure repeated passing of the beams through the active region. Light can escape if the direction of the internal beam falls within the escape cone, for which according to Snell's law $\sin \theta_{cr} = 1/n$ (θ_{cr} is the critical angle of total reflection, n is the refractive index of the active region).

We first consider the case limited by geometrical optics, which has been established by Yablonovitch [82–85]. In the literature it is variably denoted as the conventional limit, the ergodic light trapping limit, the ray-optics limit and the Lambertian limit. It is assumed that the detector-active material can be described by an effective absorption coefficient α isotropic throughout the device and that the detector thickness is much larger than the operating wavelength in free space ($d \gg \lambda/2n$), so that one considers a bulk process. The absorbance within the photodetector for a single pass across the structure (absorption without enhancement) is

$$A(\omega) = 1 - \exp(-\alpha(\omega)d) \approx \alpha(w)d, \tag{2.1}$$

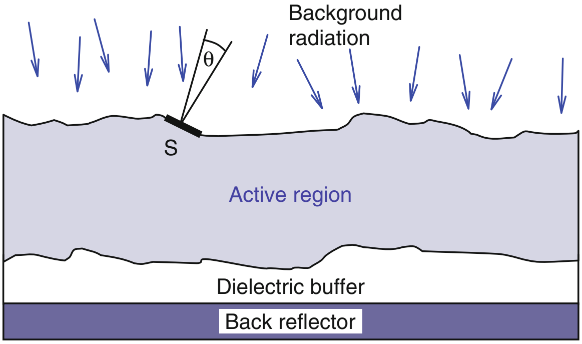


Fig. 2.2 The structure of the active region of a detector with corrugated surface and ideal backside reflector

i.e., the absorbance is equal to the optical thickness of a photodetector, which is defined as the $\alpha(\omega)d$ product.

Since a bulk case is considered, it is further assumed that interference/diffraction effects can be neglected and that the intensity of light within the detector medium is in equilibrium with external blackbody radiation. The density of states within the medium is proportional to n^2. The next assumptions are that the equipartition theorem is valid (the internal occupation of states is equal to the external one, the internal states are ergodic) and that the surface corrugation performs a full randomization of the incident signal over space. This is not always satisfied, but the assumption holds in a vast majority of cases [82]. A sufficient condition for randomization of light by multiple scattering corrugated surfaces is that these surfaces upon averaging behave as Lambertian. The internal distribution of the light within the medium is then isotropic.

According to the statistical ray optics approach [82], the relation between internal and external intensity of light is

$$I_{\mathrm{int}}(\omega, x) = 2n^2(\omega, x) I_{\mathrm{ext}}(\omega), \tag{2.2}$$

The same result is also obtained according to the principle of detailed balancing of the light [86] applied between the light incident to a small surface element of the detector-active area and escaping from that same element through the loss cone and by applying the brightness or radiance theorem (e.g., [87]) stating that the spectral radiance of light cannot be increased by passive optical devices (based on the principle of reversibility).

To determine the enhancement of absorption, one has to consider the loss of light due to various mechanisms. According to Yablonovitch [82, 83], there are three such mechanisms: the escape of light through the light cone, the losses due to imperfect reflection at the surfaces, and the absorption in bulk. The absorbance of a photon is the ratio of the rate at which absorption occurs and the sum of the absorption and the photon loss through the escape cone. For the volume absorption in the limiting case when $\alpha(\omega)d \ll 1$ and taking account the angle of the loss cone θ, this expression is

$$A(\omega) = \frac{\alpha(\omega)}{\alpha(\omega) + \frac{\sin^2\theta}{4n^2 d}}, \tag{2.3}$$

so that the absorption enhancement limit in the bulk case with internal randomization becomes $4n^2$. For $\theta = \pi/2$, this assumes the more often used simple form

$$A(\omega) = \frac{\alpha(\omega)}{\alpha(\omega) + \frac{1}{4n^2 d}}. \tag{2.4}$$

A more complex situation is encountered with the devices using plasmonic localization for absorption enhancement. In this case, many of the above assumptions introduced for ergodic limit are not valid. The crucial points are that the light distribution is not isotropic anymore (and actually the volumes with a strongly enhanced density of electromagnetic states may be deeply subwavelength) and the thickness of the detector is usually subwavelength.

A number of treatises are dedicated to the situations in which the ray optics limit is exceeded and optical modes are confined at subwavelength scale [88–90]. However, until now no generally valid solution has been given for the extension of the ray optics limit [91].

2.2 Nonimaging and Imaging Optical Concentrators

One of the obvious methods to increase the optical input to the detector is to collect optical signal from a larger area and to concentrate it to the smaller electrically active area. This can be done by some kind of optical concentrator, i.e., focusing optics integrated with the detector unit.

The usual way to concentrate light from a larger area to a smaller one is to utilize optical elements that can create an image of an object. However, besides such structures, there are also concentrators that belong to the wide class of nonimaging optics [92–97]. This is actually the branch of geometrical optics in which the main goal is not to form the best possible object image, but rather to gather as large portion of the incident radiation as possible and to concentrate it to the active area of the detector, regardless of the price to be paid from the point of image fidelity. A single photodetector element in infrared technique is not an imaging device and only the total intensity of the radiation impinging on its active surface is important, not its spatial distribution. A similar situation is met in the case of solar cells. Thus for such photodetectors one can apply both imaging and nonimaging concentrators.

A schematic presentation of an optical concentrator is given in Fig. 2.3. The shape of the system can be arbitrary, as well as its mechanism of focusing. The important factor of the system is its concentration ratio, defined as a ratio between the input (collector) area and the exit area. In an ideal concentration system, this ratio should be equal to the ratio of incident optical flux to the exit flux. In a real system, possible absorption losses (for instance in metal parts) will have to be deducted. Thus the concentration ratio of a light concentrator can be defined either as the geometric concentration ratio C, the ratio of the entry aperture area ($A_{\mathrm{collector}}$) to the exit aperture area (A_{exit}), or as the irradiance gain C_i, the ratio of the irradiance on the collector (Ir_c) to that on the entry aperture (Ir_m). The two concentration terms are related by the fraction of total incident power entering the module that reaches the concentrator exit aperture.

Different strategies can be used to concentrate the incident optical power. There are three main groups of optical concentrators. One of them is based on refractive optics (conventional lenses). The second one is reflective concentrators (mirrors), while the third group is diffractive optical elements. Obviously, a system may simultaneously incorporate two or even all three of the mentioned structures.

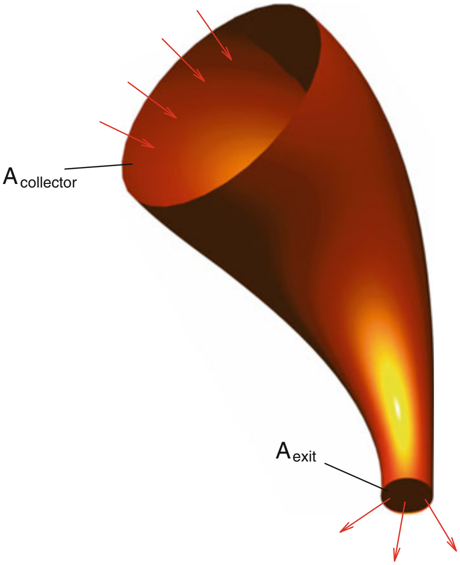


Fig. 2.3 Schematic presentation of an arbitrary light concentration system. The area at the entry of the system is $A_{\mathrm{collector}}$, and at the output A_{exit}

Monolithic or hybrid integration may be used to integrate concentrators with the photodetector (optical immersion of detectors and focusing optics). The concentrators may be partially integrated with the detector, i.e., positioned immediately to the detector surface, but without physical integration. In that case, there is a gap filled with air or some transparent buffer material with a real part of its refractive index comparable to that of the detector-active area or smaller than it and at a distance of an order of the operating wavelength (pseudoimmersion). Finally, concentrators may be integrated with the detector housing, i.e., encapsulated together with the active element, but separated from the detector surface by a distance much larger than the operating wavelength (discrete lenses).

From the point of view of coupling efficiency (minimization of reflection losses), the best solution is monolithic or hybrid integration; however, in practical situations one can encounter all of the mentioned approaches. Two main types of focusing lenses may be used—either refractive lenses fabricated in material with high real part of refractive index and low absorption coefficient at IR wavelengths, or diffractive lenses. Any of them may be either discrete or arrayed. Reflective optical concentrators may be used, also reflective holographic optical elements or any of their combinations.

Besides the maximum increase of incident flux, the main requirements posed to the concentrators in infrared photodetection are the simplicity of fabrication technology, minimal losses, and good reproductivity during fabrication. The dimension tolerances for optics are of secondary importance and allowed aberrations are higher than those posed, e.g., to the microlenses used with optical fibers. Even in imaging optics for the infrared range the tolerances (especially those for MWIR and LWIR) are much less stringent than those for visible light, since larger wavelengths are used. The nonimaging nature of the concentrators further relaxes the tolerances.

2.3 Refractive Concentrators

2.3.1 Immersion Microlenses

Probably the most obvious way to collect illumination from a larger area and to concentrate it to the active area of a photodetector is to use a microlens. Such a lens should have dimensions sufficiently small to be mounted on a detector. Since detector diameters are typically below 1 mm and very often below 0.1 mm, this means that the "micro" in the name of the microlens only means "small," since the dimensions of the lens itself will be of the order of millimeters.

In order to avoid reflective losses between the detector and the lens, typically some kind of refractive index-matching immersion will be used. The lenses for the photodetector improvement can be thus called immersion lenses. Obviously, the material of the lens has to be transparent in the wavelength range of interest and is typically used with some kind of antireflective coating.

There are a number of different geometries convenient for immersion microlenses in photodetection. Probably, the most well known and widely used form is hemisphere. The use of microsystem technologies allows the fabrication of various discrete or arrayed microlenses, with spherical surfaces (calottes, hemispheres and truncated spheres, full spheres), aspheric (ellipsoids, paraboloids, cylinders, cones), toroid, as well as various nonmonotonic surfaces consisting of two or more monotonous segments. Most of the microlenses convenient to increase the incident flux to the detector are plano-convex ones.

Figure 2.4 shows some types of refractive microlenses that can be fabricated utilizing the standard microfabrication procedures in materials convenient for the MWIR and LWIR ranges. Most of them are loosely based on the solutions for microlenses used in fiber optics to improve coupling between laser sources and fibers [98]. These immersion lenses were thus intended for the operation with coherent and monochromatic radiation, while most of the microlenses in the field of IR detector technology are intended for incoherent, mono- or polychromatic Lambertian sources and, of course, they operate in different atmospheric windows.

Among spherical microlenses utilized in optoelectronics literature quotes spherical calottes [99] (Fig. 2.4a), hemispheres [100, 101] (Fig. 2.4b) and truncated spheres [102] (Fig. 2.4e). These lenses may be defined as the refractive surfaces that can be represented as a part of a sphere with a radius r and with its center in the focus of the lens. The immersion lenses are usually formed by bringing into contact the planar surface of such a lens and the incident surface of the detector. Besides the above-mentioned lenses, ball-shaped microlenses are also used [103].

The following microlens shapes were also fabricated by microsystem technologies: elliptic calottes [104], hemiellipsoids [105], ellipsoids [103], hemicylinders [101, 105] (Fig. 2.4g), and cylinders [106] (Fig. 2.4h). These types of microlenses are especially convenient if the beam shape is elongated. Also among aspherics are conic microlenses with straight or curved generatrix [107], (Fig. 2.4i). In some optical systems, paraboloid or hyperboloid surfaces are sometimes used, usually

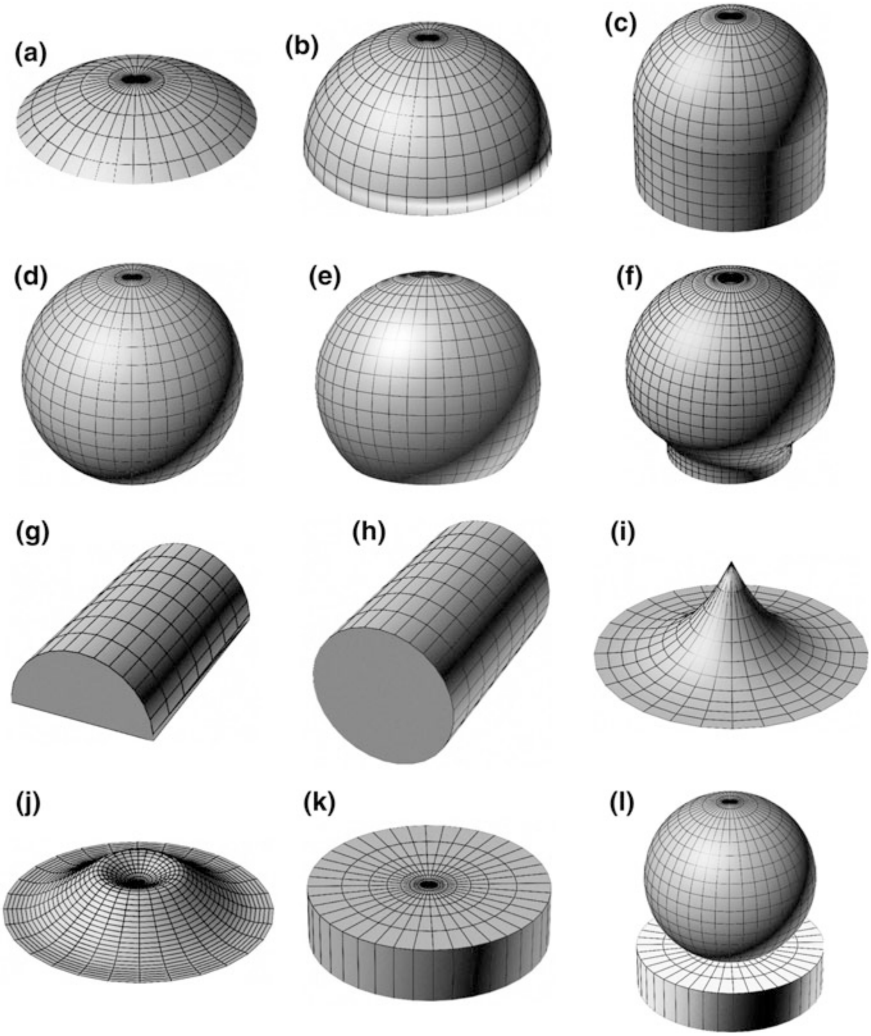


Fig. 2.4 Some spherical and aspheric discrete microlenses for optical concentrators in MWIR and LWIR range that may be fabricated by microsystem technologies. **a** calotte; **b** hemisphere; **c** hyperhemisphere; **d** ball lens; **e** truncated sphere; **f** "bulb"; **g** hemi-cylinder; **h** cylinder; **i** curvilinear cone; **j** concave concentrator; **k** gradient-index lens (GRIN); **l** complex two-element lens (sphere/GRIN). The grid corresponds to the homogeneity of refractive index, i.e., describes its gradient

fabricated by polishing. We should mention here optical concentrators with concave middle area [98] (Fig. 2.4j).

A very important type of optical concentrators, both in optical telecommunications and for MWIR and LWIR photodetectors are hyperhemispheric microlenses (combinations of a hemisphere and a cylinder with joined bases, i.e., the "rod lens")

[103] (Fig. 2.4c). Literature also mentions microlenses shaped as rounded pyramids, fabricated by micromachining [108].

An important class of microlenses are planar lenses with gradient refractive index (GRIN-microlenses, Gradient Index), Selfoc-lenses or Luneberg's lenses [103] (Fig. 2.4k; the grid shows the nonuniformity of refractive index, high in the middle, and decreasing toward the edges.) Within this context one could also mention the "electronic lenses" where the refractive index profile depends on impurity concentration and thus allows an external electrical control of the focus [109].

Confocal systems with two microlenses are often met in telecommunications [110], most often a combination ball lens-GRIN (Fig. 2.4l). Typical procedure is to focus one of the lenses and to align the other subsequently, so that the combination itself improves the allowed position tolerances several times. Also, for some applications (irregular beam shape) it is convenient to have more than one lens to allow a larger degree of freedom when adjusting beam collimation. Dependent on application, systems with three or more microlenses are also used [110].

2.3.2 MEMS Fabrication of Infrared Microlenses

As mentioned before, most of the structures for infrared detection can be fabricated by microsystem technologies (MST) and the microlenses for optical concentration are not an exception, regardless of the type of the optical element in question (refractive, reflective or DOE/HOE).

A point of special importance is the choice of proper material for MWIR and LWIR optical concentrators. A number of materials are customarily used to fabricate infrared optics. Besides these, optical concentrators may be fabricated directly in the photodetector material and indeed be monolithically integrated with it.

The usual requirements posed for infrared optics and regarding mechanical and water resistance, robustness and durability, insensitivity to temperature changes are significantly less stringent here, owing to the fact that microlenses are integrated within the detector housing. Some water-soluble materials are used for IR concentrator microlenses, because they are readily polished with water and cut by a wet string saw [1]. Among the materials used for infrared is germanium, since it ensures relatively simple machining. The most important requirements remain a maximized transmission in a given spectral range and the possibility to machine the microlens by some of the MST procedures.

A number of procedures are used to fabricate microlenses for telecommunications and some of them can be modified and adapted to fabricate optical concentrators for the MWIR and LWIR range. Some of these procedures already became standard.

Chronologically, the oldest approach to microlens fabrication is thermal melting. The action of heat is used to soften or melt convenient material (e.g., germanium in LWIR range) and form it into a desired shape. An alternative is to use material that polymerizes into a desired shape, or a convenient solver may be used that evaporates and leaves a solid object. Melting is the technology utilized for the least demanding applications, but at the same time it is probably the most often used one.

One of the oldest methods used with molding, but at the same time the least accurate, is cast molding (e.g., injection molding, casting, etc.) [111]. This method is used both for refractive and diffractive lenses. First, a high-accuracy "master" is fabricated. It is subsequently used to produce large series of microlenses. The method is especially convenient for work with plastic materials transparent in IR, like TPX (polymethylpentane, polymerized dimer of propylene.)

The procedures to fabricate masters are handled later in this section. Practically, all higher accuracy methods used to fabricate microlenses can be used to produce masters. In all of the quoted methods, it is important to choose polymers transparent in the IR range.

The methods for inexpensive replication and mass production of microlenses include

- Conventional injection molding (the most often used technique for low-end diffractive elements, for instance for the best-known diffractive optical elements, CD, DVD, and BluRay disks) and epoxy molding (the latter being more expensive because of longer curing cycles).
- Embossing: heat and pressure are applied to transfer surface relief from a mold into a plastic substrate, typically a thermoplastic foil.
- UV replication: probably the most convenient method for DOE mass production and also the most usual for higher quality and low-cost diffractive elements. The substrate is usually glass, and the mold is fabricated by electroforming. UV curable resin is potted between the mold and the glass and replicas are molded by photopolymerization. Alternatively, the mold may be used for e-beam evaporation of various inorganic materials to the substrate.

One of the molding methods often used to fabricate refractive lenses is to deposit a cube of lens material on the substrate using, e.g., some of the planar technology procedures, and then to treat it thermally until it melts. Surface tension turns it then into a hemisphere, calotte, or a flattened hemisphere [112]. An even better method is to use a similar procedure to deposit a "base" for the microlens with a desired diameter, and then to deposit lens material on it [113]. After melting, the lens material flows exactly to the edge, where surface tension prevents it to go farther (melt stop.) In this way, the lens surface is formed with an extreme accuracy.

One of the often used methods is to dip the lens support into the melt and remove it. In the course of this, a "drop" is formed at the end of the support. After cooling and hardening it is mechanically cut away. Finally, one of the used approaches is to drop the melt into cooling liquid. The shape and the size of the cooled drop define the lens.

The next important technology is photolithography. It is used both to fabricate lenses ready for use and masters for further replication, for any of refractive, reflective, or diffractive concentrators.

The most often used lithography procedures are electron-beam writing and optical lithography. Electron-beam writing is actually scanning of the photoresist-covered (spin-coated) surface by a focused electron beam, and thus exposing the resist. This technique enables the fabrication of continuous and curved profiles. It poses problems with positioning, requires computer control, and very high mechanical and vibration stability and must be performed in vacuum. Optical lithography uses light beam for writing and may be performed in the air. Some of its variants are direct laser writing in photoresist (laser beam lithography), holographic lithography (the use of interference pattern to write a DOE/HOE without a mask), and conventional photolithography (where mask can be also a hologram). The latter can use contact-printing (direct contact between the mask and the resist surface, furnishes the best results if performed in vacuum), proximity printing, and projection printing (conventional for volume production, but less accurate).

The next step in all the versions of lithography is removal of the photoresist, either the exposed part (positive photoresist) or unexposed (negative p.) and the final step is etching.

Two dry etching methods are often used: Reactive-Ion Etching (RIE) and ion milling. RIE is actually a simultaneous etching by chemical reaction and by particle bombing. It is anisotropic (vertical component is much more pronounced than the lateral). Ion milling uses a stream of inert gas to strike the surface and remove material by physical sputtering.

Wet etching includes chemical micromachining (etching in a chemically aggressive solution) and, depending on the properties on the substrate and the etching solution, it can be isotropic or anisotropic. Among the available techniques is laser-assisted chemical etching (convenient for very deep vertical holes). Blazing can be done by the melted resin technique and the mass transport method. Masters for simpler kinds of diffractive concentrators can be formed by holography or by mechanical ruling using, for example, a diamond cutter.

The fabrication of hemispheres and hemicylinders was reported in Cohen (1974), Bear (1980), Lee (1985). The microlenses here were fabricated in negative photoresist by point UV illumination to obtain a hemisphere, or by bar-like illumination for hemiellipsoides and hemicylinders. The excess resist is removed by solvent. A negative property of this method is that resist lenses are mechanically sensitive, stronger sources (e.g., industrial lasers) may easily damage them and they are applicable only for certain wavelength ranges.

This method is convenient for fabrication of GRIN lenses, where the required refractive index gradient is obtained by point diffusion [114]. Refractive microlenses with arbitrary geometry may be fabricated by photolithography so that first masks and etching are used to form a mesa or multiple-step structure directly in the substrate, and then additional etching is used (bromine-methanol for HgCdTe or InSb) to round it into a spherical or aspheric surface [115].

Another approach to form a lens directly in the substrate material is to form first a lens in photoresist deposited onto the substrate using some of the above procedures. After that reactive-ion etching (fast, but relatively small choice of materials) or ion milling (slower, more materials) is used to transfer lens into the substrate

[116]. The method is convenient for fabrication of optical elements in the wavelength ranges where resist is not transparent and generally offers a larger degree of freedom in choice of lens material. Finally, one of the options is to immerse the whole system into an electrolyte and to perform photochemical etching through a mask [117].

Most of the above photolithographic procedures can be utilized to fabricate diffractive lenses [118].

We mentioned holographic lithography as one of the types of photolithography used for IR concentrator fabrication [119]. Its main advantage is that the key steps of forming and exposing a mask are omitted. Resist (or some other medium, e.g., bichromated gelatin [120]) can be directly exposed by an interferometric pattern obtained by laser irradiation [107, 121]. A stable optical bench is necessary for that (or, alternatively, a pulse laser with high enough output power). The second alternative is to use computer-generated holography. In that case, a mask is necessary, but very complex systems can be fabricated [122]. Some authors mention the use of bichromated gelatin for the infrared optical elements [3].

Chemical micromachining is one of the methods of choice for fabrication of microlenses [123]. Silicon and germanium were used as lens materials. For spherical and aspheric geometries solutions for anisotropic etching are used, while nonmonotonous surfaces can be fabricated by combining masks and anisotropic etching. This method was used to fabricate various types of optical concentrators, among them cones and inflected surfaces [107].

Single microlenses with relatively larger dimensions can be efficiently produced mechanically, e.g., using abrasion or polishing. To fabricate both refractive and diffractive microlenses diamond turning is used [124]. One of the methods in this group is microgrinding [125].

One of the modifications of this approach is the use of a combination of mechanical and chemomechanical polishing, for instance, by simultaneous action of abrasive and bromine-methanol for lenses in Ge, CdTe, and InP [126].

2.4 Reflective Concentrators

Reflective nonimaging concentrators are not often used with infrared photodetectors, but are very well known and intensively applied in solar cell energy production [93, 123, 127–129]. Typically their structure is a tapering hollow with reflective walls, its wider aperture serving as the collecting surface, while multiple reflections from the walls collect beams toward the exit opening. The cavity may be empty (i.e., filled with air) or filled with some dielectric. The geometry may be two-dimensional (various kinds of trough) or three-dimensional, the latter being either a body of revolution or some other kind of tapering structure. The prototypical 3D structure is the so-called Winston cone, while various other tapering profiles may be used as well. They start from the simple ones like parabolic, hyperbolic, circular, conical shapes, extend to their combination, but may have different complex forms

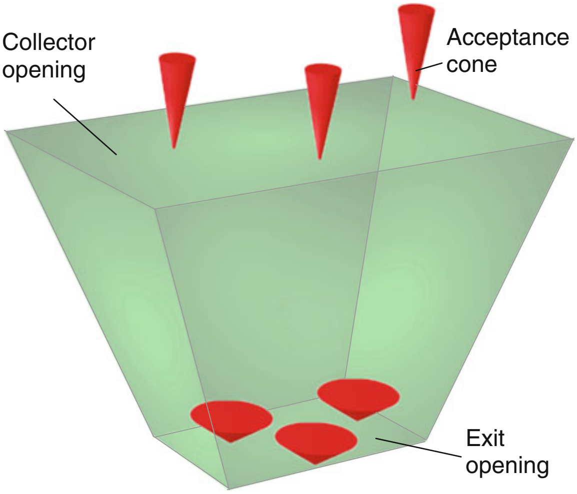


Fig. 2.5 Angular distribution of light showing larger acceptance angle cones at the exit opening and smaller at the collector opening

as determined by a particular application and the optimization procedure implemented to obtain maximum performance. The surface also may be multipart/segmented, with each segment having a different profile. Other solutions include Fresnel or echelon reflectors.

Figure 2.5 shows the connection between the spatial angle from which incident flux may arrive in order to be concentrated (acceptance cone) and the spatial angle under which radiation leaves the concentrator. Any radiation outside the acceptance cone will be rejected, i.e., the system will not bring it to the exit aperture. It can be seen that the acceptance cones at the input cover much smaller spatial angle than the light cones at the exit.

Regarding materials, the structures of microreflectors and their arrays are usually made of metal with good reflection coefficient. There are a wide choice of metals and alloys convenient for the fabrication of reflector-type concentrators.

The shape of a 3D reflector top and basis may be rounded, but also polygonal, the latter being more convenient if a high filling factor is desired when producing 2D arrays of reflectors. Figure 2.6 shows some possible shapes and the manners of their stacking into larger matrices. The rightmost example shows that the possible shapes may include combinations of polygonal entrance with round exit.

2.4.1 CPC (Winston Collector)

Compound parabolic concentrator (CPC, also called Winston collector or Winston cone) [95–97] is a reflective nonimaging light concentrator intended to collect incident light from a spatial angle larger than that of imaging concentrators and to

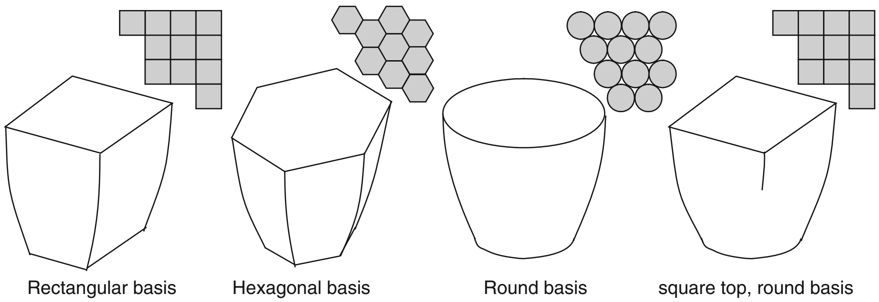


Fig. 2.6 Different entrance and exit geometries shapes for reflective light concentrators. *Gray matrices* in the right upper angle of each shape show the manner of concentrator packaging into arrays

direct it toward its exit opening. It has been the first discovered nonimaging reflective concentrator. Its 2D depiction is schematically shown in Fig. 2.7.

Two parabolas p_1, p_2 can be seen in the figure, with their respective focal points in F_1, F_2, and their axes being a_1, a_2. The parabolas are cut at the line of their focal points and then mutually rotated so that the focal point of one of them coincides with the curve of the other parabola and the very edge of the exit opening. A 3D version of this structure is obtained as a body of revolution whose sides are parabolic segments satisfying the same condition. Thus a Winston cone is an off-axis parabola of revolution.

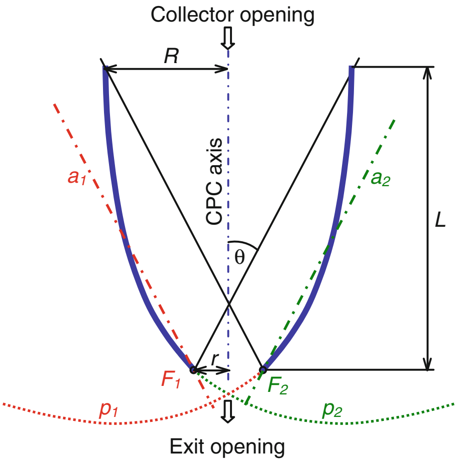


Fig. 2.7 Schematic presentation of CPC collector. R is the entry (collector opening) radius, r is the exit opening radius, a_1 is the axis of the parabola p_1, a_2 is the axis of the parabola p_2, F_1 and F_2 are focal points of the parabolas 1 and 2, respectively. Parabolas are rotated so that the focal point of one of the parabola coincides with the curve of the other parabola at the edge of the exit opening

A CPC concentrator increases the efficiency of concentration by collecting off-axis rays that would else be lost and directing them toward the exit opening. These rays reflect several times within the CPC and after such multiple bounces reach the exit. The rays arriving at an angle larger than the acceptance angle do enter the concentrator, but after bouncing several times off its walls they are lost, i.e., they do not reach the exit opening.

The total length of the CPC concentrator is

$$L = \frac{r(1 + \sin\theta)\cos\theta}{\sin^2\theta} \tag{2.5}$$

Obviously, a CPC can be truncated and in real situations often is. This will result in a lower concentration ratio, but the overall length might be more acceptable for implementations.

There is a relation between the radii of the entry and exit openings

$$R = \frac{r}{\sin\theta} \tag{2.6}$$

The equation defining the surface of a CPC in a radial system in which the CPC axis is in z-direction while the radial distance is described as the coordinate ρ is

$$\begin{aligned} 2r\rho\,(1 + \sin\theta)^2 + (\rho\cos\theta + z\sin\theta)^2 \\ = 2rz\cos\theta(2 + \sin\,\theta)^2 + r^2(1 + \sin\theta)(3 + \sin\theta) \end{aligned} \tag{2.7}$$

A special case of the rotational Winston collector is a CPC trough. Both 2D and 3D CPC structures are ideal concentrators. 2D CPC is convenient if the target travels along a line since then there is no need for its tracking.

One possible improvement of the CPC is to fill it with dielectric having a refractive index n. This shortens the structure and increases the acceptance angle. At the same time it uses total internal reflection which ensures a higher reflection than from metal surrounded by air. The maximum concentration ratio in this case is larger by a factor of n^2. Another useful modification is to utilize a two-stage CPC where the exit opening of the first stage CPC coincides with the entry opening of the second stage.

2.4.2 Parabolic Reflector

Geometrically, a 3D parabolic reflector represents a truncated paraboloid, as shown in Fig. 2.8. Parabolic reflectors are actually imaging optical elements, used in various optical systems. For instance, a parabolic reflector is often combined with a diverging mirror that directs light beams through the exit aperture. This is the well-known Cassegrain configuration, the standard design for reflector telescopes.

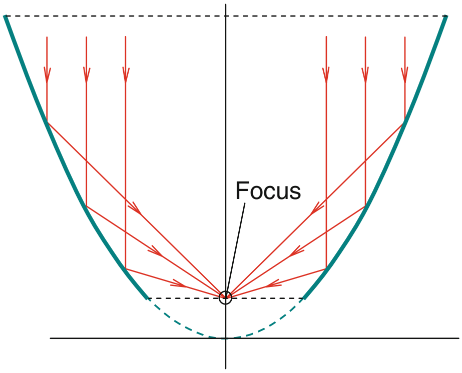


Fig. 2.8 Schematic presentation of a parabolic reflector. The parabola is cut off in the focal point

Fig. 2.9 Parabolic trough

The microfabricated version of parabolic reflector is used in solar energy concentration [130].

Contrary to the CPC, a parabolic concentrator is a high-concentrating device, but is more sensitive to scattering of incident radiation than the CPC. A problem with parabolic concentrators is that their total transmission is much smaller than that of CPC and reaches about 0.60.

A 2D implementation of the full paraboloid geometry is the parabolic trough [131], Fig. 2.9. Although a full three-dimensional paraboloid is more energy efficient than a trough, 2D implementation is less complex and can be useful in situations where there is a linear array of photodetectors.

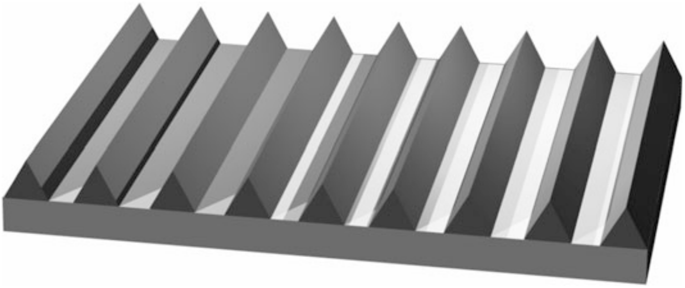


Fig. 2.10 A V-trough for reflective nonimaging concentrators

2.4.3 Conic Concentrator

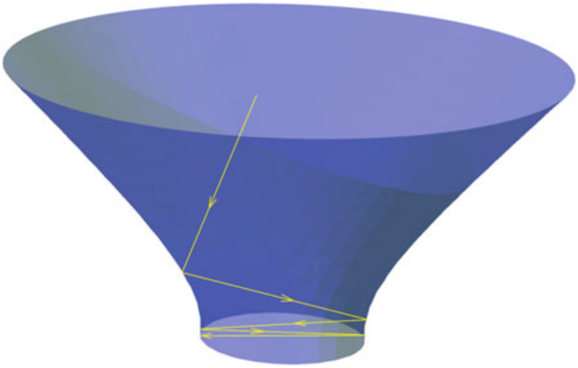
The conic concentrator is actually a truncated cone with reflective sides. It is not an ideal concentrator, and its transmission reaches about 0.8. Its 2D version (with a triangular profile) is called a V-trough, Fig. 2.10.

2.4.4 Hyperbolic (Trumpet) Geometry

A geometry for reflective concentrators is the trumpet concentrator [128], Fig. 2.11. This is actually a hyperboloid of revolution. Besides the CPC it is denoted as one of the "maximum collection efficiency" nonimaging concentrators. It is characterized by multitudinous reflections near the exit aperture.

Trumpet concentrators are ideal concentrators both in 2D and 3D case. They also offer the advantage of concentrating not only internal but also external rays (those latter are concentrated to an annular shape around the exit aperture). Trumpet concentrators are characterized by a small material expenditure. External rays are not usable for detector matrices.

Fig. 2.11 'Trumpet' concentrator



2.4.5 *Other Reflector Geometries*

There is a plethora of reflector geometries for nonimaging concentrator optics. Among them are elliptic and compound elliptic concentrators (CEC) as well as high-flux two-stage optical designs based on tailored edge-ray concentrators (ordinary and compound). They also include different lens-mirror combinations, utilizing reflection (X), refraction (R), and total internal reflection (I), for instance XR, RX, XX, XRI, etc. [128]. Complex forms for ultra-high concentrations of input radiation are obtained by using different algorithms, for instance global optimization procedures.

2.4.6 *Fresnel Reflector*

The idea of the Fresnel reflector is to divide a curved reflector into segments and to move these segments to a single joint plane [132, 133]. Thus one is able to effectively reduce a 3D object to a thin (quasi-2D) object while keeping the identical optical path to that of the 3D object. For instance, if one starts a 3D body of revolution like, for instance, a parabolic reflective concentrator or a Winston cone, this body can be divided into a number of annular segments which are subsequently placed concentrically on a flat surface. This process is illustrated in Fig. 2.12a, where a curved body is cut into segments and placed on a surface. An incident beam reflects under the same angles as it would from the full 3D body.

The surface of the original body of revolution can be turned into a Fresnel reflector by being cut into annular segments with the reflective area identical with the original, i.e., with curvilinear surfaces, Fig. 2.12b. It can be also approximated with blazed geometry, Fig. 2.12c, where the cross-section of the annular segments is fabricated in triangular form, with the internal side of the triangle following the original form as accurately as possible.

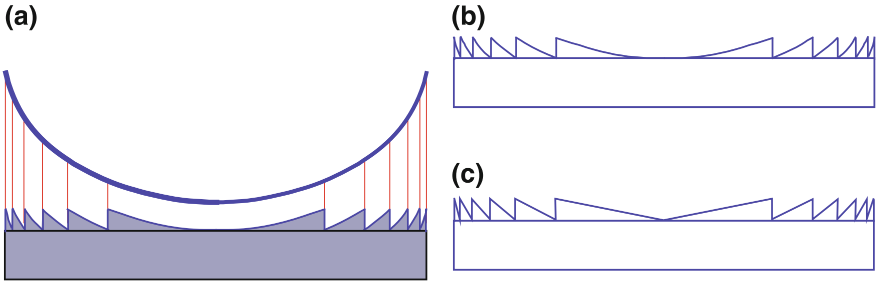


Fig. 2.12 Fresnel reflector. **a** Method of cutting reflective surface. **b** Reflective surface represented by in-plane curvilinear annular segments. **c** Reflective surface approximated by flat (blazed) segments with triangular profile