TECHNICAL UNIVERSITY OF LIBEREC
FACULTY OF MECHANICAL ENGINEERING
DEPARTMENT OF GLASS PRODUCING MACHINES AND ROBOTICS

INTRODUCTION TO MACHINE VISION

PART 1

FUNDAMENTAL PRINCIPLES AND HARDWARE

Vlastimil Hotař
Translation: Adam Babka

2017

Reviewer
Assoc. Prof. Maroš Tunák

Vlastimil Hotař

Translation: Adam Babka

ISBN 978-80-7494-391-1

INTRODUCTION TO MACHINE VISION, PART 1: FUNDAMENTAL PRINCIPLES AND HARDWARE

This educational text was created based on long-term research of machine vision and its applications, and uses the latest knowledge available in this field.

This textbook is intended for use during subject Scanning and Processing of Industrial Data, studied by students of the second year of N2301 Mechanical Engineering follow-up study program, specialisation in Glass Producing Machines and Robotics.

Vlastimil Hotař

Translation: Adam Babka

Reviewers	Assoc. Prof. Maroš Tunák
Published by	Technical University of Liberec
Approved by	The Rector's Office at TUL on 21st of December 2017 Ref. No.: RE 72/17
Published on	December 2017
Pages	126
Issue	1st
Publication No.	55-072-17
ISBN	978-80-7494-391-1
Electronic version	

Content

Acknowledgement

When I was preparing material for this publication, and during its processing, I had help of many people. Colleagues from the Department and the Faculty, academic staff from other universities and people from industrial practice. This textbook was published under significant merit of Assoc. Prof. Vladimir Klebsa, who contributed with valuable comments and proofreading. Also, I would like to thank the reviewer of this textbook, Assoc. Prof. Maroš Tunák (Faculty of Textile, Technical University of Liberec).

Additionally, I would like to thank my family, especially Lianka, my wife, who allowed me to work on this textbook in my free time.

Preface

This educational text is intended for students of Bachelor and follow-up Master's study programs. Its goal is to familiarise students with theoretical basis of machine vision and present basics of designing machine vision systems. The problems of optical sensors are rather extensive and description of current states of hardware and mathematical instruments related to image processing would form at least one large publication, therefore, this text should be perceived as a mere introduction to the problems. Goal of this textbook is to provide students of Bachelor and follow-up Master's study programs at the Faculty of Engineering with brief and comprehensible overview of problems as well as targeted links to other studies.

This publication has two parts. The first part presents fundamental principles of image capturing and hardware issues. The second part addresses image processing. Hardware is described in a way so that technically educated students understand the fundamental principles, requirements and restrictions. As for image processing, the text contains fundamental principles for obtaining useful data from obtained image.

The whole concept of this textbook is similar. Fundamental principles of image capturing are described first, while the following chapters explain hardware issues (first part). The following chapters are addressing image processing and basic methods of working with images (second part).

Introduction

Machine vision is used more and more frequently in industrial practice as well as applications for service equipment, military technology and in consumer products, toys and model making. It is becoming increasingly available, and in common practice, every university educated mechanical engineer encounters it at some point.

Machine vision is a field of science trying to mimic some aspects of human vision. Beside own image perceptions, intelligence of a person and his/her previous experience have a significant impact on evaluation of image information. Machine vision works similarly. A mere capture and transmission of image provides basically nothing in the field of automation. However, with image processing, it is possible to achieve certain, limited (in comparison to human) intelligence of the given machine. That is due to the fact that image processing allows extracting necessary information from the given image for the given type of task. This information is then used by the machine control system for decision-making.

In terms of image processing, it is necessary to use correct hardware and software, while both parts must “respect” one another and must be properly interconnected. Designing suitable equipment for image processing and related issues requires at least good knowledge about both parts. Additionally, with regard to profitability, proper solution must be selected as well and must be sufficient to handle the given problems and will not unnecessarily increase price and therefore increase the return on investment. For example, it is logically needless to use an expensive lighting system for shape detection of white parts, when it is simpler to place the part on black background. Similar solutions that are simple in their principle should be prioritised when designing the system. Also, hardware and software should be of similar level of complexity.

In order to implement proper and efficient machine vision, one must have sufficient knowledge regarding the problems and then choose a suitable strategy. Even the initial design of image processing system requires large amount of experience and invention. Not everything can be exactly calculated in advance, and any incorrect initial estimate may cause failure of the whole solution.

Knowledge of machine vision principles is not crucial only when designing a visualisation system, but when designing technological operation before and after the visual evaluation as well. When operating the system, knowledge of the principles are important as well, since the operators may prevent a malfunction or shorten the downtime.

This publication has two parts. The first part addresses fundamental principles of image capturing and required hardware, while the second part described image processing using various mathematical approaches [1].

1 Fundamental principles of image capturing

Knowledge in a very good textbook [2], as well as other literature, is used in this chapter. Many other relevant publications (e.g. [3, 4, 5]) can be found.

Despite the fact that this textbook focuses on capturing and processing a visible spectrum of radiation (eventually infra-red and ultraviolet radiation), it is wise describe the general perspectives regarding image display, which is performed using three types of energy:

1. Electromagnetic radiation (from gamma rays, through x-ray and ultraviolet radiation, to visible radiation, infra-red radiation, microwaves or radio waves Fig. 1.1; image example is shown in Fig. 1.2).
2. Particle radiation, for example electrons (used in electron microscopes) or neutrons, image example is shown in Fig. 1.2.
3. Acoustic waves (propagation velocity of acoustic waves is proportional to elastic properties of the medium), for example sonar (echo sounder) principle that uses ultrasound, or medical sonography, which is based on registering ultrasound bouncing off from tissue; image example is shown in Fig. 1.2.

In the first case, a portion of radiation bounces off the observed object, portion is absorbed, portion will cause secondary emission of radiation, portion of radiation may pass through the object and portion of it may be emitted by the object itself, Fig. 1.3. These phenomena are used based on object properties (e.g. surface or material properties, temperature) and properties of electromagnetic radiation (mostly wavelength) interacting with the object. Other types of energy are not regarded or described in this textbook.

The object capturing process itself is basically a radiometric measurement performed by means of appropriate probes. These probes are sensitive to the observed range of radiation, while individual probe “cells” (called pixels) measure the amount of incident radiation (intensity, which is brightness) in real time. The radiation information that can be used to identify the seen objects and measure some of their properties is:

- radiation wavelength (can also be expressed as frequency),

- amplitude, i.e. intensity,
- polarization mode for transverse waves,
- phase, which can however be measured in case of coherent display techniques such as interferometry and holography.

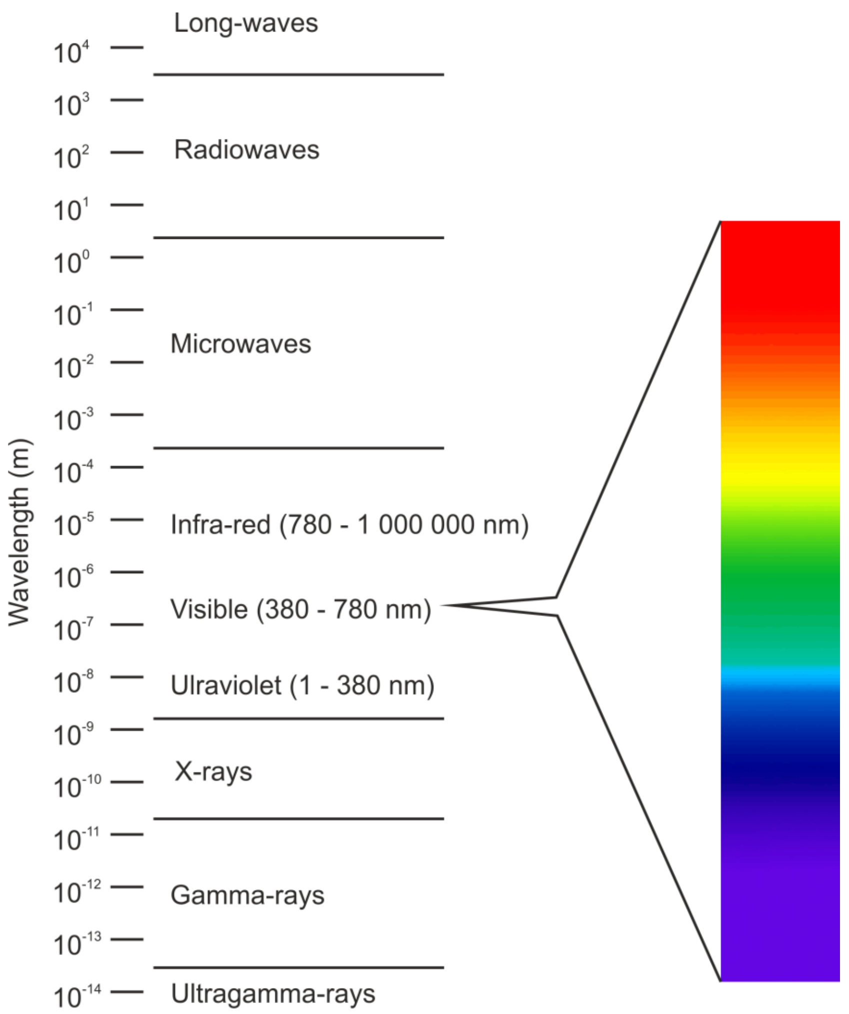

Fig. 1.1 Electromagnetic radiation and display of coloured light spectrum

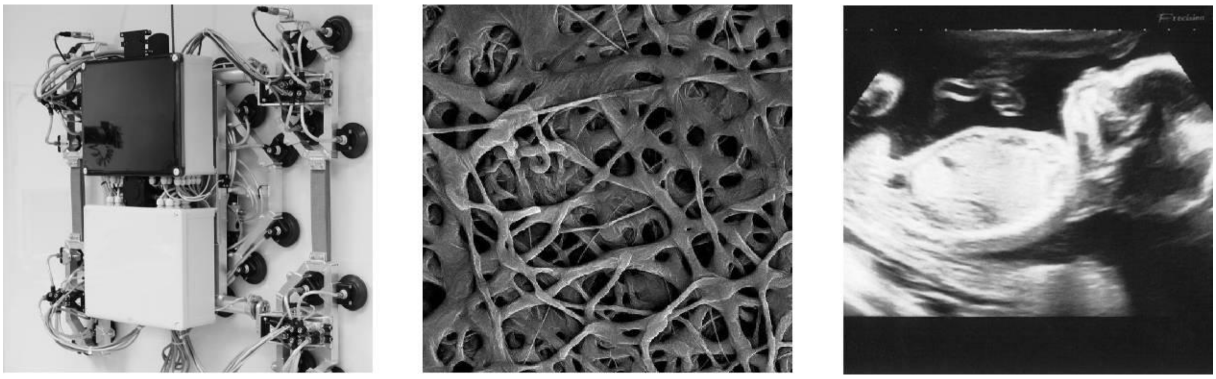

Fig. 1.2 Example of images captured using various types of energy

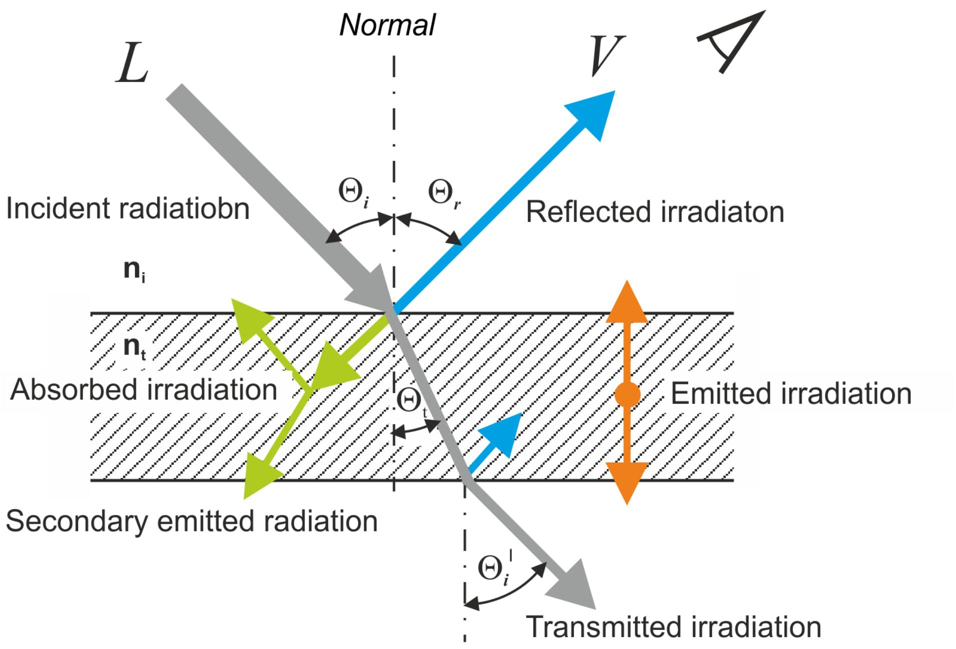

Fig. 1.3 Interaction of radiation and observed object

1.1 Visual signals processing

Before we start, it is necessary to determine Cartesian coordinates used in this textbook. The x, y, z coordinates will be used for real space, the subject of capturing. The u, v coordinates, or eventually w will be used to create an image and models created from the image.

In this chapter, the term 'image' will have a broader sense than 'frame' used in other chapters. Both terms can be defined as a signal obtained from probes and transformed to at least two dimensional expressions. Frame can be perceived as an individual signal captured within a given period of time. Image can also contain time information (another dimension of the signal) not contained in the individual Frame. Image may consist of individual frames.

Image can be formed mathematically by means of continuous scalar function $f_b$of two or three variables, called image function. When having static image captured in grey-scale, the image is described using a function of two coordinates $f_b(u, v)$ in one plane (generally speaking, on surface, e.g. approximate spherical surface of human eye retina). In case of surface image captured in a time period t, the image function has three variables $f_b(u, v, t)$. The image function has three variables even in case of 3D images, i.e. $f_b(u, v, w)$. Image function with four variables is a special case and is used for presenting 3D images in a time period $f_b(u, v, w, t)$. Values of the image function are then connected with the measured ones. In case of monochromatic cameras, the value is brightness (amount of visible spectrum photons striking upon a photosensitive cell, see chapter 1.3.1), electron microscope measures the amount of electrons striking upon a sensitive cell, in thermal imaging, temperature of observed object (intensity and frequency of incident radiation of higher wave-lengths), the observed value in case of x-ray tomography is the ability to absorb radiation at a single spot.

Image obtained by means of a camera represents the matrix-form image with pixel elements. The value of pixels is proportional to the amount of incident energy, or incident photons in case of monochromatic cameras scanning in visible spectrum. (So called line scan cameras described in chapters 2.1 and 3.4 scans only the lines in a scene, however, the resulting image is a matrix formed by combining individual lines.)

Mere capturing and displaying the image on a computer is not machine vision. In order to make the image usable in practice, it must be appropriately interpreted (properly evaluated using appropriate algorithms) with appropriate degree of robustness. The aforementioned algorithms represent so called image and are based on several levels of processing. The goal is to make the machine vision system understand the image and obtain information usable for further processing. There is an apparent effort to imitate the process of human perception based on information contained in the image. Image interpretation is affected by several fundamental problems that make it complicated:

1. ***Loss of information during perspective view.*** Image obtained using industrial cameras is a two-dimensional view of a 3D world (chapter 1.2). Logically, there is a loss of information required for complete description of real world. An inverse process to restore 3D properties of an object using an image of a single camera has infinite amount of solution. One of the solutions

is to use other types of knowledge (projective geometry[1]) and (or) information about the observed object (additional image obtained by means of another camera, definition of object position towards the camera, etc.). Respecting the loss of part of data is necessary even in case of relatively simple tasks such as shape detection.

2. ***Complicated and unclear relationship between brightness and shape of scanned 3D object surfaces***, chapter 1.3.1. Brightness of individual points (pixels) measured by the image probes (cameras) is affected by many factors. Those are for example reflectance of individual observed objects, background of the technological scene, orientation of objects and elements in the background, position and parameters of lighting. Detection of objects with various reflectances on various spots may be very complicated. Selecting proper lighting conditions and position of the observed object, eliminating intrusiveness of the background and other actions are required to achieve effective machine vision.

3. ***Large amounts of image data.*** Even with continuously increasing performance of camera control systems (e.g. classic PCs), this problem must be respected. With increasing performance of computers, the resolution of cameras increases as well, see chapter 3.1.2. If the processes would use full resolution of cameras, maximum frame rate and unmodified data format (RAW) offered by the current devices, the data stream would be so large that even the simple tasks would be practically impossible. When designing the system, it is necessary to consider the speed of observed technological process, required camera resolution, data pre-processing performed by the camera (including compression) and complexity of image analysis.

4. ***There is always a noise in the image from real scene*** (chapter 2.6). In case you need to eliminate the noise, it is necessary to apply certain measured performed by hardware and software tools [1, chapter 7]. However, these measures complicate and therefore slow down the image analysis. This must be respected when developing a real machine vision application and the effect

[1] Projective geometry is used to transform 3D into 2D. In this geometry, points and lines are defined instead of angles and distances.

of measures in complex application should be decreased, and various simplifications should be performed. Alternatively, robust analysis can be created.

5. ***Relationship between observed detail and detected whole.*** For example, due to sophistication of the analysis, only a small cut from the whole image is often evaluated in real implementation of machine vision. However, the cut prevents determination of global properties of an image, which can be important as well.

Image processing includes several levels of tasks allowing obtaining necessary information from the captured image about the observed object and interpreting these information correctly. According to [2], interpretation of the image means display: observed data → model. Model always represents the necessary simplification of real world, but due to display of 3D world in 2D, the representation has its restrictions. Individual levels of image function representation (data processing) can be divided to:

1. Iconic – these are digital images in a form of integer matrix after eventual pre-processing of images with the aim to improve the image (software filters, geometric correction).
2. Features – parts of the image are distributed to groups belonging to individual objects, or groups with texture or surface structure ready for further analysis.
3. Objects – these are the result of segmentation, i.e. total data interpretation. Obtained parameters quantifying given texture and surface structure belong to this level as well.
4. Relational models – describe quantitative and qualitative properties of an object. Machine vision uses recognition techniques and artificial intelligence.

Individual steps for processing two dimensional images can be expressed as tasks for achieving the goal of image analysis and allowing transition between the aforementioned representations. Processing and analysing 2D images can be expressed using a block diagram based on [2], Fig. 1.4.

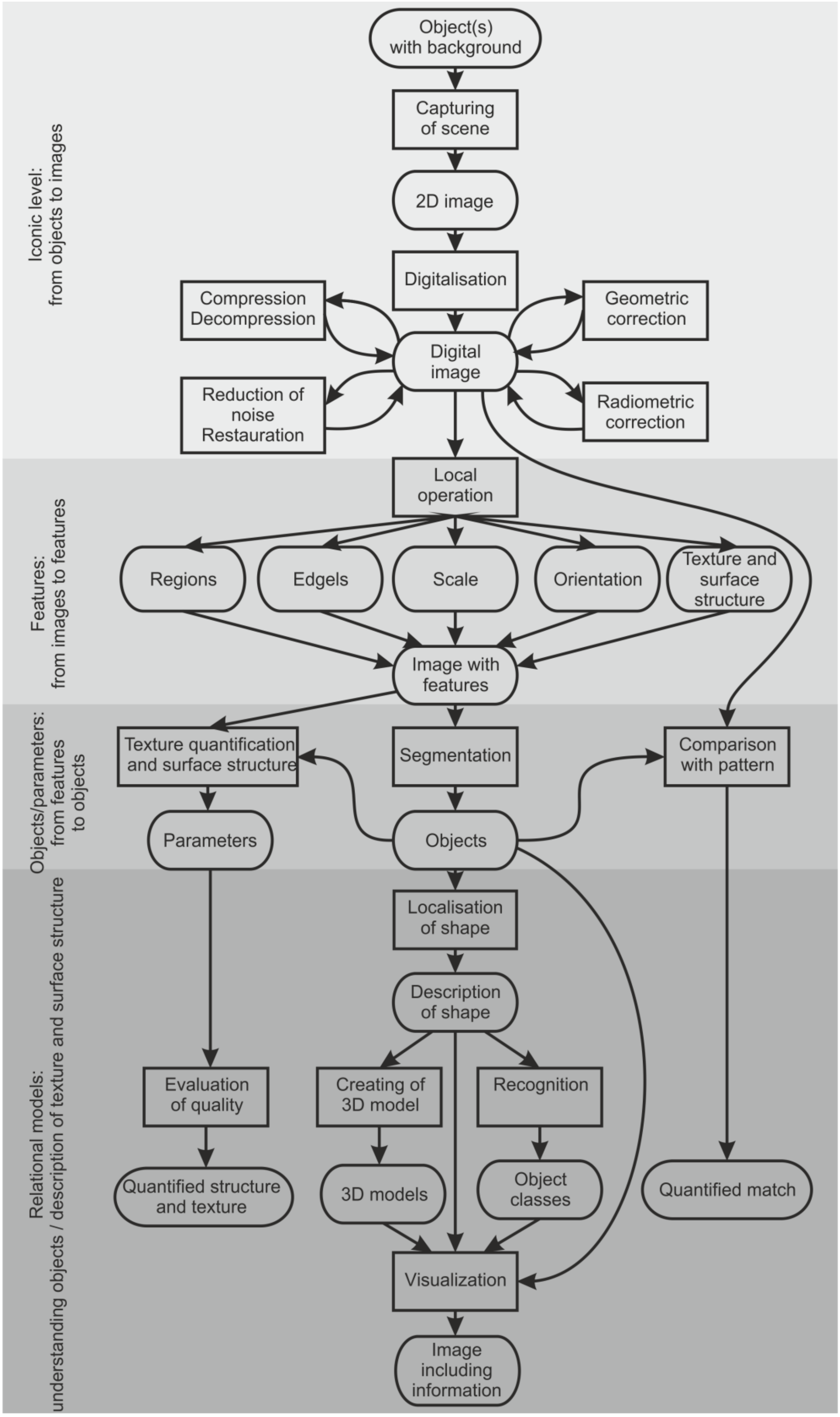

Fig. 1.4 Hierarchy of image processing tasks, is based on [2]

1.2 Image function and three-dimensional space

From the topology perspective, world can be considered three-dimensional (3D). In most cases, the obtained image is topologically two-dimensional (2D). This applies in case of area sensor and also in case of line sensors for line scan camera, where the resulting 2D image is obtained by means of moving the captured object or camera. (Use of 3D cameras is not considered in this text.) Image function (image obtained by probe/sensor or eye retina) is a result of perspective view, or so called central projection, modelled, in simplified description, by pinhole camera, Fig. 1.5.

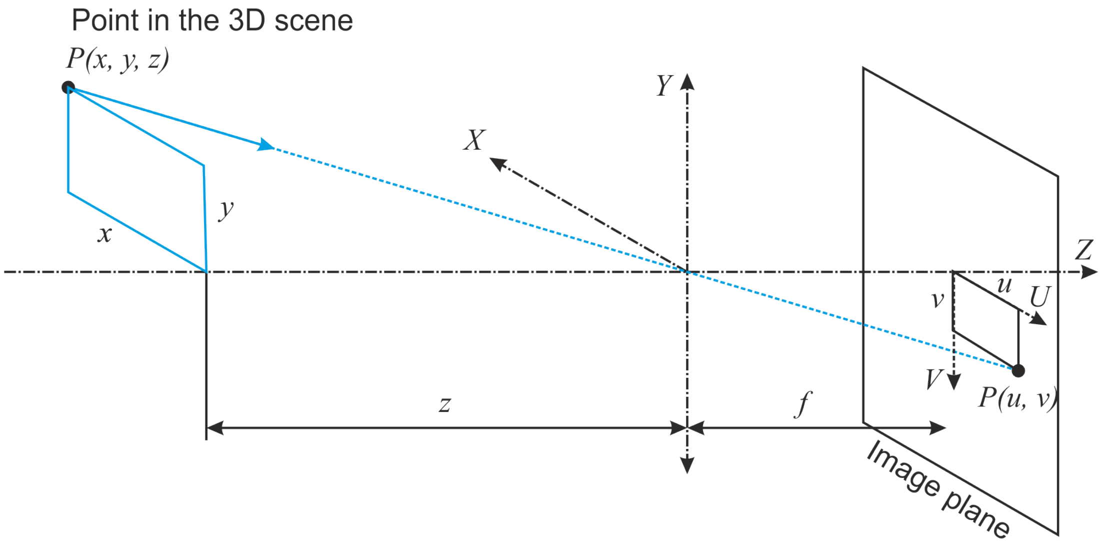

Fig. 1.5 Geometry of perspective projection, so called pinhole camera model

The P point has x, y, z coordinates within the 3D scene and f represents distance of image plane from centre of projection (in case of lenses, the image plane is in the distance corresponding to focal length). A point projected to 2D image has u, v coordinates. The following applies due to similarity of triangles:

$$u = \frac{xf}{z}, v = \frac{yf}{z}. \qquad (1.1)$$

Perspective view given by this formula is non-linear and is sometimes replaced by orthographical (or parallel) view formed under the condition $f \rightarrow \infty$.

Fig. 1.5 clearly shows that all points in 3D space on one half-line from centre of projection towards the scene are displayed as a single point. That results in a loss of information and perspective view is not invertible. The geometric task that is supposed to measure, reconstruct or identify 3D objects from a single 2D image is insufficiently conditioned. However, there are procedures that allow at least partial description of an object or scene in a 3D space [6, 7]. Such procedure does not usually use a single image, since the model is transformed to a 3D model using two or more images.

Projective geometry that addresses projection of 3D world into 2D world observes properties unchanged by projection transformation (collinearity) [8, 9]. Model for this geometry is usually a projective plane or projective space. In this geometry, points and lines are defined instead of angles and distances. In this geometry, the term *plane of projection* corresponds to the term image plane, which is a plane in space onto which all the projection beams are incident and create a projection (image plane). Due to the fact that image plane displays an inverted image, the plane of projection can be imagined as a plane in front of the centre of projection moved by the focal length $f' = f$, Fig. 1.6. Additionally, a term *projection rays* is implemented, which is a line directed through the projected (space) point, while its direction depends on the projection method.

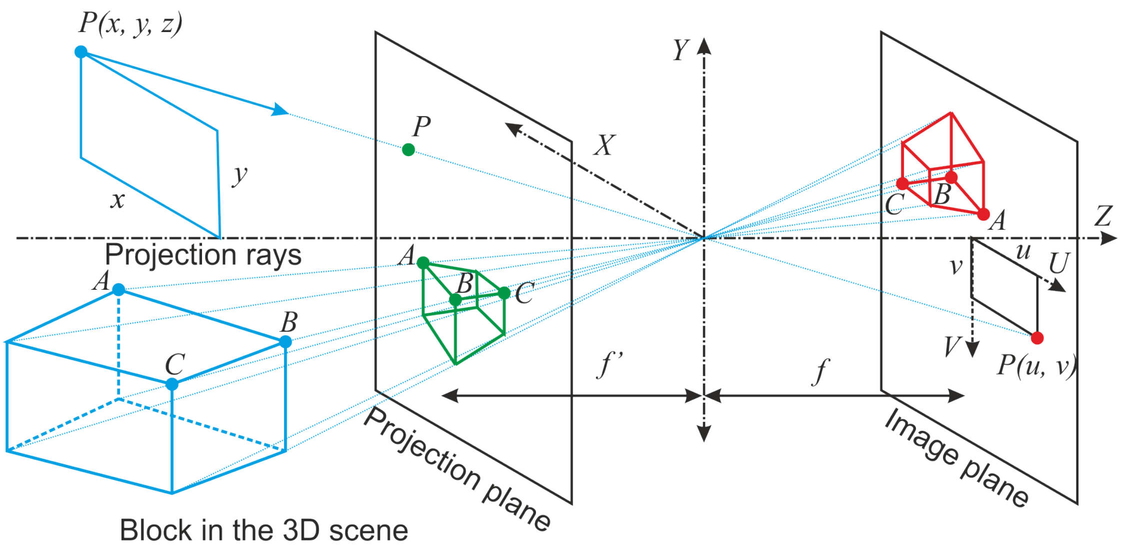

Fig. 1.6 Geometry of perspective projection of a point and a block, display of a plane of projection, so called pinhole camera model

Basically, there is a difference between one point, two point or three point projection.

When using telecentric lens, one may encounter a parallel projection, where the perspective will no manifest itself under certain conditions, chapter 4.3.

1.3 Image capture

CCD (chapter 2.1) and CMOS APS (chapter 2.2) sensors are usually used for capturing image of visible electromagnetic radiation overlapping to ultraviolet and infra-red radiation, Fig. 1.1. The sensors basically measure the amount of light energy incident on the pixels. The measured values are then relatively corresponding to the radiation intensity E [W m^{-2}], which describes energy incident on the surface point of the irradiated object. The radiation intensity is often informally referred to as brightness (which is a photometric value). One may also encounter terms specific luminosity or grey-scale. In physics, radiometry [10] addresses the measurement of electromagnetic radiation and its flux and transmission, and allows explanation of the image creation mechanism. Absolute values are used in radiometry. The elementary value in radiometry is radiant flux φ[*W*].

Photometry is a field related to radiometry and observes similar values describing responses of human eye. A photometric value depends on spectral characteristics on the radiation source and sensitivity of photosensitive cells on eye retina. The ability of the human eye to perceive intensity based on wave length of light is non-linear and differs from person to person. Therefore, a standard photometric observer was empirically determined and corresponds to statistically identified average of human abilities. Radiometry and photometry describes similar phenomena by using similar values. Radiant flux is similar to luminous flux in lumen, lm, 1W = 680 lm, radiation intensity corresponds to light intensity used in photometry E_f, lm.m^{-2} (lx, lux).

1.3.1 Radiation formula

The radiation formula defines a relation between radiation intensity E incident on the sensor in an image plane based on lighting L in the observed scene. A pinhole camera model is considered in the Fig. 1.7, while the optical axis is in the direction of Z axis. The lens is located at the start and has a focal length of f. $\delta O$is an elementary surface on the object surface and δI is its view in image. δO is in distance of z, n is its normal, while Θ is the angle between normal of the surface and connector with the zero-point of the coordinate system. β is the angle between optical axis and beam incident on the sensor.

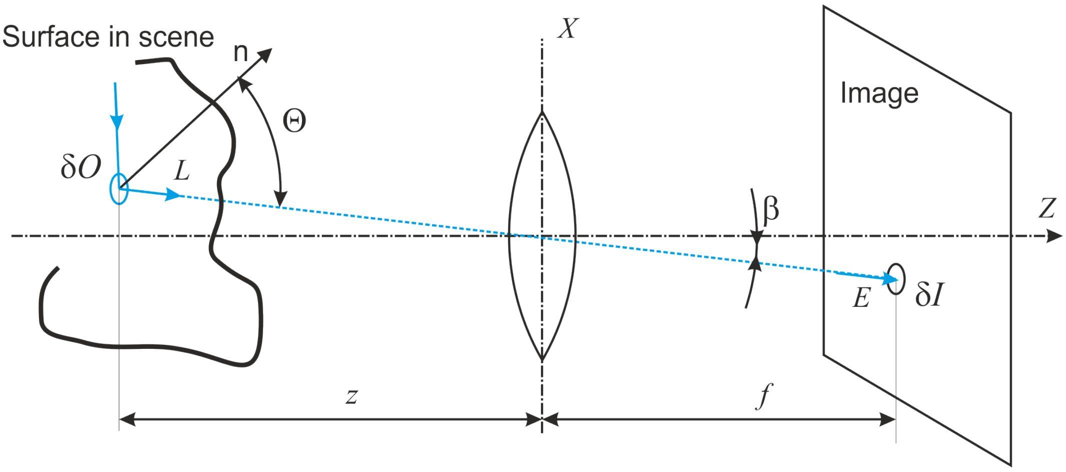

Fig. 1.7 Relation between radiation intensity E *and lighting* L

By performing derivation, e.g. in [2], a formula for defining lighting can be created describing the manifestation of lighting throughout the whole scene:

$$E = L\frac{\pi}{4}\left(\frac{d}{f}\right)^2 cos^4\beta \quad , \qquad (1.2)$$

Where d is the lens diameter. The aforementioned formula is applicable for the whole captured scene, while for individual points of the sensor, a *reflectance function* should be used (further information in the following chapter).

The $\cos^4\beta$ factor describes a systematic optical fault called natural vignetting. The result is dimming effect of the obtained image from the centre to edges. In case radiation intensity under the E_β angle and radiation intensity in E_0 optical axis (same lighting from the observed surface in both cases) and the radiation will be compared using their ratio, the L, f and d are constant values:

$$\frac{E_\beta}{E_0} = cos^4\beta. \qquad (1.3)$$

The natural vignetting is caused by the $\cos^4\beta$ agent. Beams refracting on the image side of the optical system are attenuated when under higher β angle. There are other vignetting types beside the natural one: optical and mechanical vignetting, see chapter 4.2.

1.3.2 Image as a radiometric measurement

Based on radiometric or photometric measurement, it is possible to connect a given radiation intensity (brightness) with a physical meaning. In machine vision and computer graphics, the value of image function f_b is perceived as estimated lighting L occurred by reflecting light energy from scene surface. Generally speaking, brightness (radiation intensity E) in a given pixel depends on the shape of an object, reflection properties of its surface, position of the observation element, position and type of lighting, lenses and optical filters. That can be used when describing captured technological scenes, where the obtained values were used for modelling the shape of observed object. However, this possibility is not used very often in practice due to complexity of the analysis, highly restricting requirements on the properties of observed objects, numerical instability of the available methods (radiometric accuracy of common cameras is much lower than the geometrical one) and, above all, the relation between measured brightness and shape is too complicated. Simply speaking, brightness of the given sensor pixel depends on shape of the observed object, its properties (primarily reflective, but also emission and other), sensor position and its properties (this may be a CCD sensor or human eye) and position and properties of electromagnetic radiation sources (lighting). Due to the aforementioned, using the calculation potential of radiometry and photometry in the field of machine vision is rather problematic [2]. The following text contains only some of the important and interesting terms.

The reflectance coefficient or ***albedo*** describes the ratio of incident energy reflected back to the space:

$$\rho(\lambda) = \frac{E_r(\lambda)}{E_i(\lambda)}, \quad (1.4)$$

Where $E_i(\lambda)$ is the intensity of radiation incident on the scene surface spot, $E_r(\lambda)$ is the intensity radiated back to the space due to reflection and λ is the wave length of electromagnetic radiation. The coefficient depends on wave length of the incident electromagnetic radiation, properties of surface spot (amount of absorbed radiation) and three angles describing the mutual relationship between the direction towards the light source L, towards the observing element V and local orientation given by the normal line n, Fig. 1.8. Cosines of these three angles can be considered as a scalar product of vectors, which is why the ***reflectance function*** (reflectivity) R is described by means of three scalar products:

$$R = R(n \cdot L, n \cdot V, V \cdot L) \, . \qquad (1.5)$$

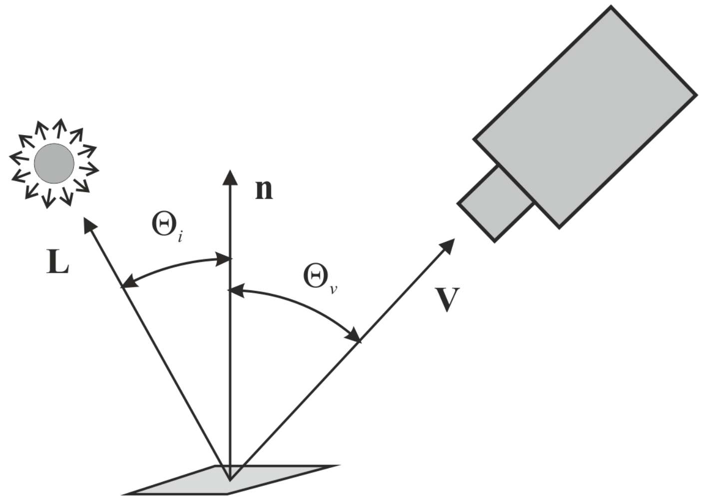

Fig. 1.8 Description of direction vectors towards the light source and observing element

The reflection occurs from a surface, whose properties are located between two extremes.

Lambert surface (ideally matt, ideal diffuse surface) reflects the light energy to all direction equally. Radiance (brightness) from all direction is constant (not depending on the view angle). Example of such surface for centre of the visible spectrum is a white blotting-paper with reflection coefficient of 0.8, white writing paper with 0.68, white ceiling or yellow paper with 0.6, dark brown paper with 0.14 or dark velvet with 0.004.

Ideal mirror surface reflects radiation based on the Law of Reflection (the angle of incidence is identical to the angle of reflection). The own surface is not visible and shows only a seeming mirror image of light sources.

Generally, the reflectance of objects is described using ***BRDF*** (Bidirectional Reflectance Distribution Function) f_r [sr^{-1}]:

$$f_r(\Theta_i, \phi_i; \Theta_v, \phi_v) = \frac{\delta L(\Theta_v, \phi_v)}{\delta E(\Theta_i, \phi_i)}, \text{ (1.6)}$$

where E is the sensor radiation intensity, L is the radiance of surface of the observed object (chapter 1.3.1). Orientation of spot is described in a spherical coordinate system using the ϕ azimuth and Θ polar angle, as clearly shown in Fig. 1.8 and Fig. 1.9.

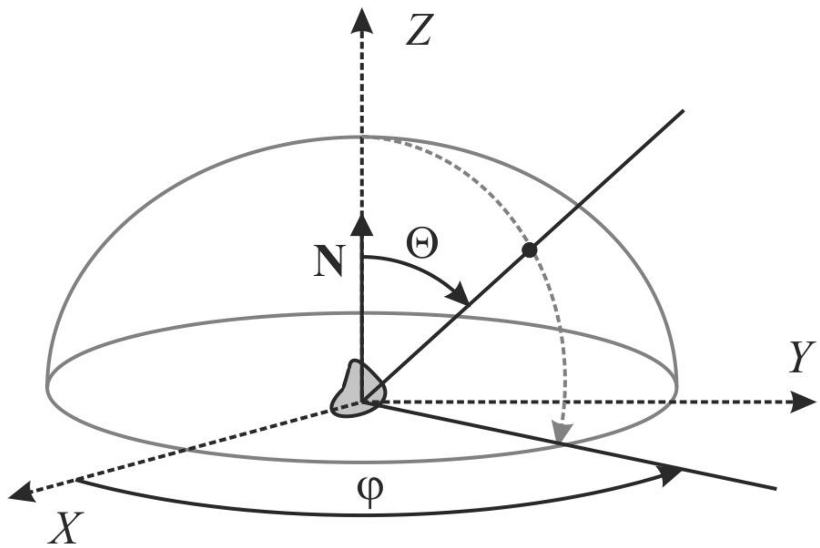

Fig. 1.9 Description of spot orientation in a spherical coordinate system

BRDF f_r describes the brightness of elementary surface spots for a given surface (material) of the observed object, light (and its properties) and the direction of observation. Modelling this function is important in computer graphics and is used for realistic shading. However, this function is very complicated and is used only when the surfaces change the reflection depending on their rotation around the normal line directed to the surface. An example of such surface with oriented micro-structure may be for example a tiger's eye (quartz permeated with fibres of crocidolite, tawny colour, chatoyant glitter), peacock feathers or for example rough aluminium cut.

For most surfaces significant for practice, the *BRDF* function depends only on difference between azimuths of directions towards the light source and observing element $\phi_i - \phi_v$, which is $f_r(\Theta_i, \Theta_v, (\phi_i - \phi_v))$ The simplification applies to Lambert surfaces, ideal mirror surfaces and their combinations.

Shape from shading is a task that allows determining the shape of surface based on change of reflectance manifested by change of brightness in the captured image. The determination should be based on the radiation formula $E(u,v)$ of an infinitely small light sensor located in u,v of an image plane. Radiation $E(u,v)$ is in relation with corresponding surface spot determined by parameters x,y. Assuming that the light beam is not significantly attenuated by the mean between the surface and sensor, this relation can be expressed as a less important radiation formula describing relation between orientation of surface and measured brightness in an image:

$$E(u,v) = \rho(x,y)R(N(x,y) \cdot L, N(x,y) \cdot V, V \cdot L). \quad (1.7)$$

In other words, the formula describes the relation between sensor surface radiation and the reflectance coefficient together with the reflectance function in a form of parameters. However, finding solution to the task is often very complex, numerically unstable and can be often achieved only by finding drastic simplifying assumptions.

The problems of radiance, determination of values for given shapes as well as gradient space is addressed in [2] and in [11] in much more detail.

1.4 Digital image

Digital image is used for further processing of an image that is a modified image of the real world (object, scene). Information in the digital image are final and limited by properties of imaging devices. Sensors for input of image function $f_b(u,v)$ are often a source of continuous signal amplified and transformed to a discrete signal by means of A/D converter (digitised, chapter 2.1.1) to *k* intervals. The obtained brightness function is transformed from continuous to discrete signal and is formed by integer values. This process is called quantization. Due to the principle of imaging, the digital image is represented by a matrix $U \times V$ of points (Fig.) called pixels and have final dimensions. This process is called sampling. The image is represented by a matrix of integers f, [1, chapter 6][2].

$$\mathbf{f} = \begin{bmatrix} f(0,0) & f(0,1) & \ldots & f(0,V-1) \\ \vdots & \vdots & & \vdots \\ f(U-1,0) & f(U-1,1) & \ldots & f(U-1,V-1) \end{bmatrix} \quad (1.8)$$

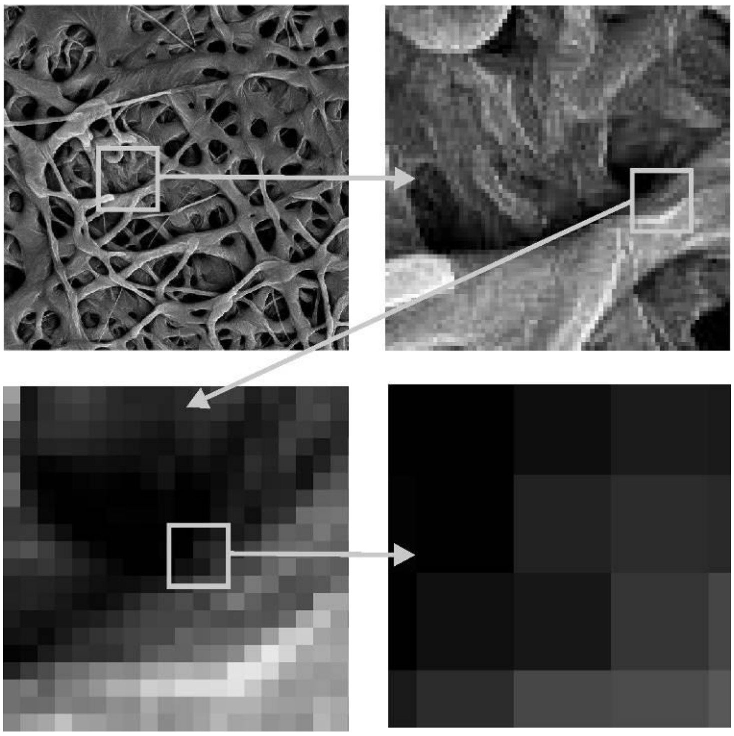

Fig. 1.10 Magnification of part of digital image up to an apparent matrix of points (pixels)

Digital image is a simplified representation of the real world. The higher the surface resolution of the sensor (higher U, V – softer sampling) and the amount of intervals, the better the approximation to the continuous image signal.

Generally speaking, the quality of a digital image is proportional to the surface, spectral, radiometric and time resolution. *Surface resolution* is determined by spacing of individual sampling points in an image – a sampling interval (distance between the closest sampling points in an image), chapter. 3.1.2. When Fourier spectrum is

[2] In some software (e.g. in Matlab), $f(0,0)$ corresponds to $f(1,1)$

created, a *spectral resolution* is used, which corresponds to the surface resolution. *Radiometric resolution* corresponds to the amount of quantization levels (amount of brightness values), chapter. 2.1.1. In case of images varying in time, another parameter is used – a *time resolution*, which is given by a time interval between capturing two consecutive images.

1.4.1 Colour image

The aforementioned relationships apply for an obtained brightness of individual pixels in grey scale. Images in grey scale (monochromatic) are sufficient for majority of industrial applications. In this case, the brightness function is $f_b(u, v)$ and is represented by a matrix. If scanning in colour, the radiometric measurement process is becoming even more complicated. Colour of objects is determined by selective absorption of components of the electromagnetic visible radiation. Parts of visible spectrum of white daylight incident on a colour surface is absorbed more, while other parts are reflected more (absorption is also discussed in chapters 1 a 6). Matters of radiation absorption and reflection also applies to wave lengths outside the spectrum visible by human eye (multi-spectral images). When performing capturing and computer processing, it is not possible to process all wave lengths at one, which is why the electromagnetic spectrum is divided into several spectral bands. These bands cover the required spectrum of wave lengths. The image is obtained by means of multiple sensors, which each of them is sensitive to relatively narrow spectrum of wave lengths. In case of colour and multi-spectral images of each surface resolution (u, v), there are brightness matrices for each colour component captured by individual sensors. Example: the visible spectrum can be divided to three individual spectral components – red (wave length of 700 nm), green (546.1 nm) and blue (435.8 nm), thus forming an RGB model. The colour image is created by combining these components. As mentioned in chapter 2.4, other colour models can be used. Spectrum observed by the camera may also capture ultraviolet and (or) infra-red part, chapter 3.1.4.

Digital processing of colour and multi-spectral images often use similar methods and algorithms as in processing monochromatic images. Procedures are applied to individual spectral components. The proportionally increasing amount of data and significant similarity of corresponding data in each component must be respected when working with colour and multi-spectral images [2].

1.5 Vision systems

The Fig. 1.4 shows the hierarchy of image processing. From the hardware perspective, it is a rather complicated system containing many elements. The basic configuration diagram of individual (and complementary) elements is shown in Fig. 1.11.

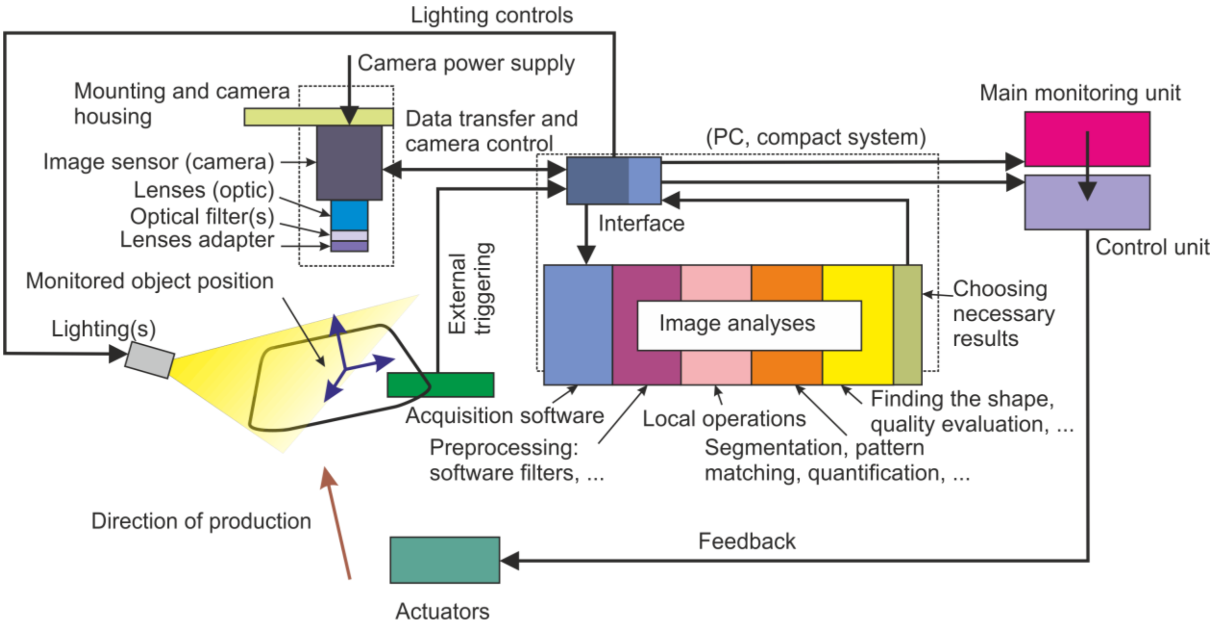

Fig. 1.11 Diagram of the vision system

In order to achieve the required quality of capture and evaluation that follows, many conditions that influence one another must be met. Mostly, the specific requirements for application in industrial practice are as follows:

a) lighting must be appropriate for scanning the given material; in some cases, it is necessary to correct the right combination of several types of lighting;

b) position of the captured object must be adapted to lighting and position of the camera or lighting, while the camera must be adapted to the position of the object;

c) it is possible to use lens adapters instead of (or together with) lighting in order to achieve better display of the object (polarized light adapter, aligned perpendicular lighting, diffuse lighting adapter, ...);

d) Filters for separating specific wave lengths from the captured scenes may be used as well;

e) Optics (lens) including aperture;

f) capturing element – sensor with the necessary devices (camera), selected by the type of observed object and its motion (line or area camera) as well as the required resolution; the capturing element can also be a digital camera;

g) mounting of the camera or its cover, including isolating from undesired vibrations, dust, low or high ambient temperatures;

h) required power supply for the camera (pursuant to manufacturer and type of camera);

i) transfer of data to an image processing device (cables, A/D converter) and camera control;

j) optional devices and cables for triggering the camera;

k) Interface pursuant to the type of selected camera (FireWire, USB, Camera Link, GigE Vision, …);

l) acquisition software (software for acquiring images from the camera);

m) very reliable image processing; the results must be obtained in time and in proper form:
   - preprocessing: software filters that are able to replace or complement the physical camera filters and are important for the consecutive image processing;
   - local operations: areas, edges, scale, texture, ... (Fig. 1.4);
   - segmentation,, comparison with the original model, quantification of surface, ...;
   - search of shape, quality assessment, ...;

n) selection of the required results;

o) data transfer to main monitoring unit and (or) control unit (extended by feedback to actuating elements).

Steps *k* through *n* are performed by means of an control vision unit, for example in a form of classic computer or compact and resistant system for industrial applications. A compact computer is basically a computer optimised for image processing using minimum mechanical elements that limits a more widespread usage of classic computers in industrial practice.

Naturally, it is possible to split individual elements of the system and complement them with other important information such as type of camera sensor. This is shown in the following chapter. The design of image analysis system must be based on an assumption that the brightness value of the given pixel depends on the shape of object, properties of surface reflectance, position of the observing element, position and type of lighting, lens and optical filters and then on processing the obtained signal from the sensor (amplification, preprocessing, compression, ...). All these aspects must be regarded in the design.

2 Types of camera sensors

Sensors of light energy were used from the 70‘s of 20^{th} century and can be divided to two groups [2]:

1. Sensors based on photo-emission principle use a photoelectric effect. Photon absorption provides enough energy to release a free electron. This principle was used in vacuum scanning electron tubes and photo-multipliers (this type of sensors is not described in this textbook due to current limited usage).
2. With development of semiconductor, sensors operating on photovoltaic or photo-current principles are used more and more and are based on internal photoelectric phenomena. These principles are also used by other two basic types of sensors often implemented for real industrial use. The older type is CCD, the more recent and more used is CMOS APS.

2.1 CCD sensor

CCD (Charge-Coupled Device) sensor is manufacturing using technology developed specifically for the camera industry in 1969 in Bell laboratories by Willard Boyle and George E. Smith. Basically, it is a shift register exposed to light. These sensors use the aforementioned photoelectric phenomenon (photo effect). Upon collision with an atom, a particle of light (photon) is able to transpose one of the atom’s electrons from normal to so called excited state (valence electrons in the electron shell of the atom are excited to energy levels higher than normal). The released electrons in the semiconductor can be transported using attached electrodes, similarly to common photo-diode. That is why electric current is created upon incidence of light. Photocells work on the same principle and are used as a source of electric energy. However, the CCD electrode is separated from the semiconductor via a very thin layer of Silicon dioxide (SiO_2) acting as a perfect electrical insulator so the electrons released by the photo effect cannot be transferred away. Each sensor consists of large amount of individual semiconductor miniature cells, while each of them captures light by itself. The image is exposed to isolated in a potential well – column-formed miniature cells. The images are then created using the points, named pixels.

Activity of the CCD sensor consists of three phases (Fig. 2.1):

1. *Preparation of CCD* - all free electrons are removed from the sensor as well as all residue from the previously captured image.
2. *Image exposure* – positive voltage is applied to electrodes 1 and a shutter is opened (the sensor is exposed to light). Photons of visible spectrum and nearby IR radiation has enough energy to release electrons from atoms forming a silicon structure, while the amount of released electrons is proportional to the amount of incident light. The excited electrons are drawn to positively charged electrodes Fig. 2.1 A. The created holes are drawn by the negative electrodes located on the bottom of CCD. The higher the amount of incident light, the higher the amount of electrons is released, creating higher charge in the given cell.
3. *Image capture* – the chips are arranged so that two more electrodes 2, 3 with positive voltage are placed next to electrodes 1 with positive voltage connected upon image exposition, while the electrode 1 is located between electrodes 2 and 3. A three-phase clock signal is then connected to the electrodes (there are also CCDs with four-phase or two-phase reading using two or four electrodes). Voltage of electrodes 2 slowly starts to increase, while decreasing on electrodes 1. As a result, electron clusters are drawn beneath electrodes 2 Fig. 2.1 B. The whole process is then repeated between electrodes 2 and 3 (Fig. 2.1 C), then between 3 and 1 (Fig. 2.1 D) and then repeated all over again. Electron clusters from individual pixels move over neighbouring pixels towards the output amplifier. This amplifier then amplifies the low voltage corresponding to the amount of drawn electrons in individual pixels to voltage levels suitable for further image processing. The next part is analog-to-digital converter (A/D converter) that performs so called quantization. The process is based on a principle, where the cell produces voltage in range of for example 0 to 5 mV in a form of continuous signal. Since the output consists of discrete values ranging from for example 0 to 255, the converter splits the analogue range of 0 to 5 mV to 256 parts. For example the voltage of 2.5 mV will have the value of 127. That is how A/D converter transforms (digitises) an analogue (continuous) signal to a digital (discrete) one [12].

After the image is captures, the sensor is being prepared and the whole process repeats.

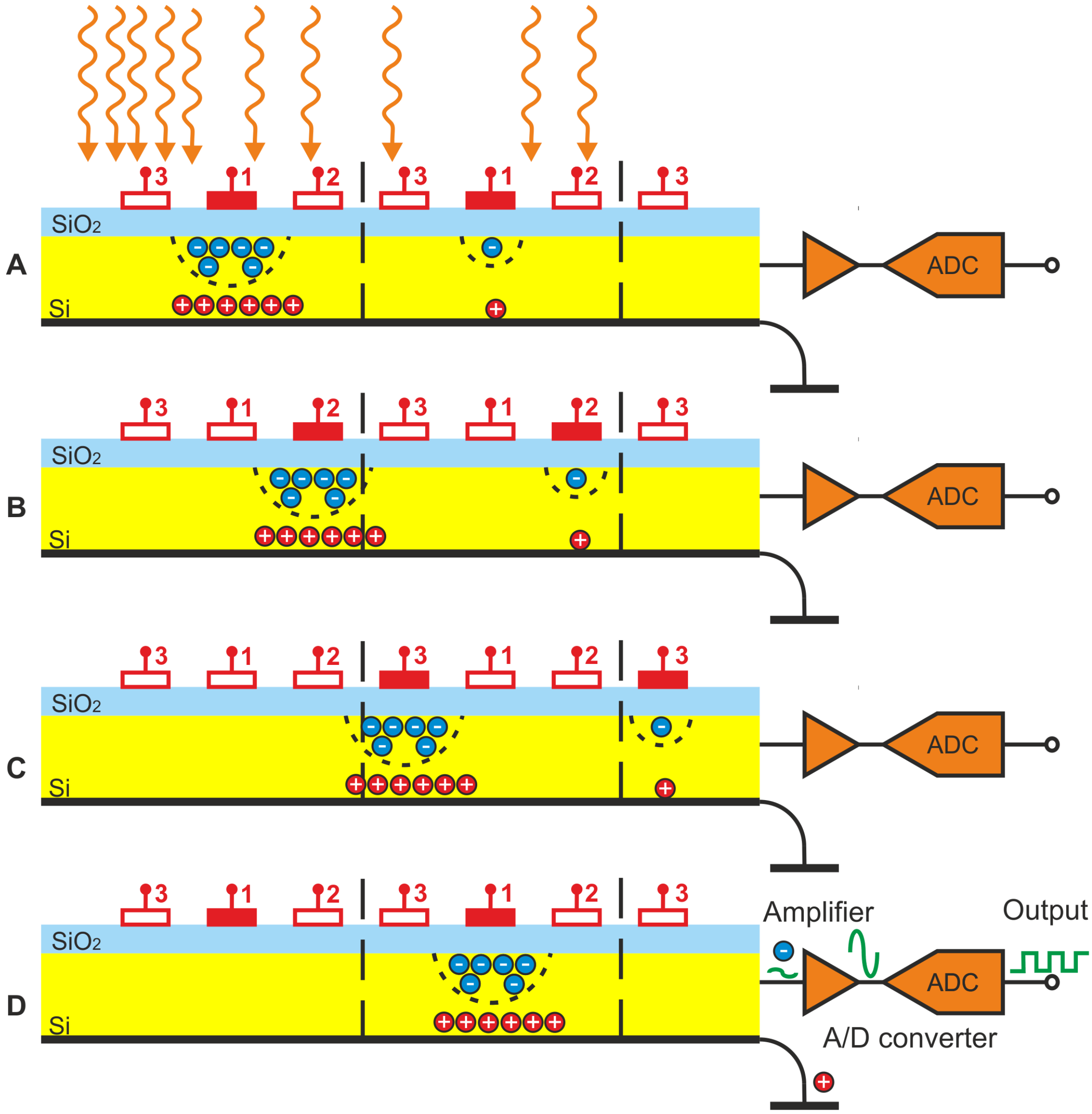

Fig. 2.1 The principle of CCD sensor operation

2.1.1 Image quantization – radiometric resolution

Quantization, a transformation to a discrete (digital) signal, should be performed so that the amount of quantization levels is large enough to express slight details of an image and prevent creating false outlines. Additionally, the sensitivity of the device should be approximating the sensitivity of human eye. Most systems for digital image processing use quantization to *k* identical intervals. If the information about an image element (pixel) is represented by *b* bits, the amount of brightness levels is $k = 2^b$. Currently, most cameras are able to process 12 bits per pixel. A single pixel of 12 bit monochromatic image can have 4096 values. However, in most applications using monochromatic image, 8 bits per pixel (256 values) is sufficient. That way, the image takes up less space in the memory and its transmission and processing is faster and sufficient for most applications. Convention says that the lowest value of a monochromatic image represents black, while the highest value is white. Values between these extremes are scales of grey. In case of 8 bit expression, 0 represents black and 255 represents white.

Human eye is able to differentiate approximately 50 levels of brightness (shades of grey) in monochromatic expression. False contours occur in images quantized for low amount of levels, therefore, quantization should be performed for 50 and more brightness levels [2].

On the contrary, quantization higher than 12 bits per pixel is meaningless with common cameras due to thermal noise. Cameras using A/D converter for 16 bit per pixel quantization of signal created by light-sensitive cells are often cooled and are used for example in microscopy.

2.1.2 Image capture with area sensors

The described CCD function creates linear information (line of pixels) that is suitable for line scan cameras, however, area sensors must capture two-dimensional image at one time and the alignment must be therefore extended. Area sensors consist of several linear ones positioned next to each other, thus creating a matrix of pixels, Fig. 2.2. The charge then enters another line CCD that will use the three-phase shift **v** to displace the first pixel of all vertical CCDs to bottom horizontal one. By repeatedly doing so, the three-phase shift **u** will push the whole row towards the image amplifier. Consequently, the other three-phase shift **v** pushes the second pixels of all columns to a horizontal CCD. This cycle continues until all rows are emptied.

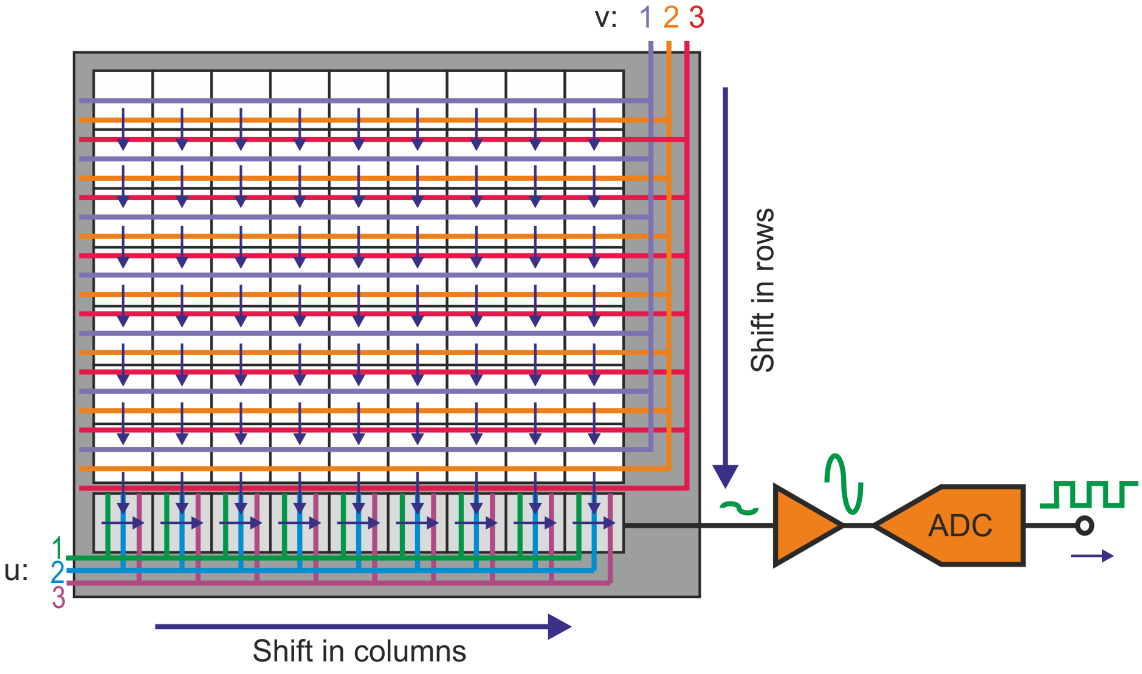

Fig. 2.2 Principle of operation of area (matrix) CCD sensor, FF reading for mechanical shutter

Information about charge accumulated in each pixel (amount of light incident on each pixel) creates a set of data representing an image function $f_b(u,v)$ (image). The image file consists of numerical matrix, while each number represents a pixel.

2.1.3 Mechanical and electronic shutter

A so called *shutter* is necessary for the operation of the whole systems. Shutter is a device that allows exposition of sensor for a certain period of time, thus enabling loading individual pixels. The sensor should therefore be covered to avoid affecting the charge. Mechanical shutters were used in the past, as in all analogue cameras and camcoders, these days however, we mostly use electronic shutters. That is related to the area sensor reading systems that basically have the following three approaches:

FF - Full Frame – the whole surface of the sensor (i.e. all available pixels) is exposed to light (Fig. 2.2 and Fig. 2.3 A). Transmission and measurement of charge from individual pixels can be done only when the system features a mechanical shutter that covers the sensor during the process.

FTr - Frame Transfer - the sensor is divided to an area exposed to incident light (Imaging Area – photosensitive part for capturing) and a permanently covered area (Storage Area – covered memory part). In practice, to have the same area of the chip exposed to light, must be twice as large as FF chips Fig. 2.3 B. During the given moment of the image capturing (generally tens of ms). The whole sensor matrix is transmitted to the covered matrix and its charge is “loaded” to an amplifier and then transmitted to an A/D converter. This principle is called electronic shutter [13].

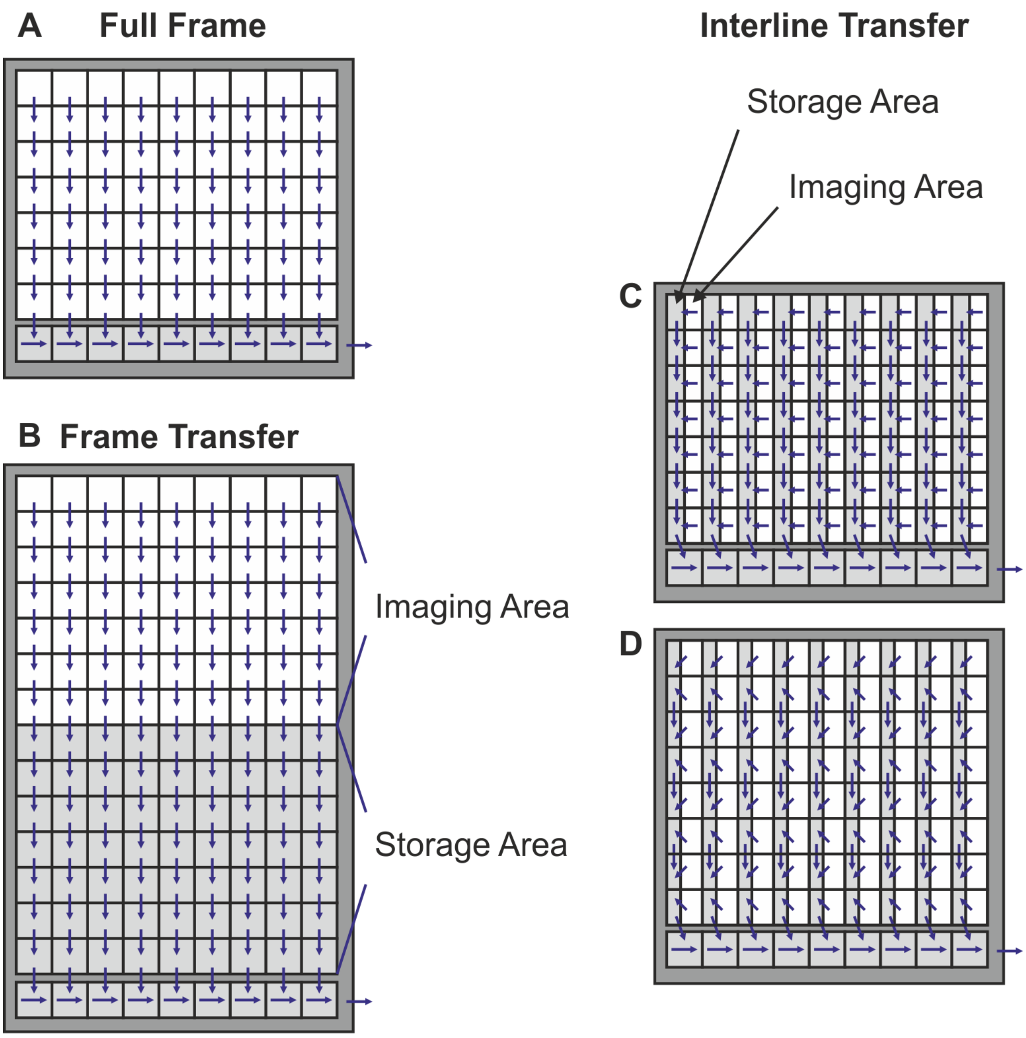

Fig. 2.3 Sensor diagram according to their reading method A - FF, B - FTR, C - ITR progressive reading, D - ITR interlaced reading (reading even images)

ITr - Interline Transfer - principle is similar to FTr. This principle uses electronic shutter as well, however, the covered area sensor is interlaced with active sensitive area (Fig. 2.3 C). Each odd column accumulates light, while even column are covered with an opaque layer. As soon as the exposition ends, the odd columns are quickly transmitted to even columns. Those are then sequentially placed into a horizontal register and digitised. The size of the sensor remains unchanged, however, interlacing the area with active and passive columns limits the area where the chip captures the light, thus decreasing the fill factor[3]. This effect can be partially compensated by advanced production procedures (e.g. micro-lens application) described later.

From historical reasons, a TV signal does not consist of a sequence of individual images, but alternating images with half the number of lines, so called field (two fields comprise one video frame). Therefore, the odd field consists of lines 1, 3, 5, etc., while the even field consists of 2, 4, 6, etc. The manufacturers of CCD chips adapted to these conventions and change the architecture of sensors used in TV so that these CCD chips are able to read individual fields. If only every second line is being read, the sensitivity would decrease to half. For this reason, the classic "TV" CCD chips features electronic summing of the exposed lines (see binning, chapter 3.4) so that the odd field contains the first line alone followed by a sum of 2nd and 3rd line, a sum of 4th and 5th, etc. The odd field then sums the 1st and 2nd line, 3th and 4th line, etc. (Fig. 2.3 D). This architecture is called an interlaced read architecture. On the contrary, the chip architecture allowing reading all pixels at once is called a progressive read.

2.1.4 Global and rolling shutter

Global shutter is a term used to describe an image sensor. It is characterised by sensor being exposed to light in a single period of time. Image data from the top, central and bottom part of the sensor are transmitted almost simultaneously.

Rolling shutter (line scan) is often used with commercial mobile devices with CMOS APS sensors. This sensor is often cheaper, and unlike a global shutter, the capturing process consists of reading individual cell rows (of the whole image)

[3] Fill factor is a ratio of sensitive part of the cell (photo-diodes) and a total cell area. The term is also used in other fields of technology.

sequentially from top to bottom. Since the image is displayed as a whole (unlike the loading phase, where the rows are loaded with varying time), the perpendicular edges of a moving object may be displayed in the image as oblique.

2.2 CMOS APS sensor

CMOS APS sensors (Active Pixel Sensor) are manufactured by means of technology often used when producing memory chips (CMOS – Complementary Metal-Oxide-Semiconductor). This technology was discovered in 1963 and was gradually being developed in the field of aviation and later in commercial applications. In the beginning of 90s, this technology became a fundamental part of most logic circuits and microprocessors. In the 70s, CMOS chips for optical sensors became a subject of experiments. However, there were specific problems, which is why in years 1993 through 1995, Jet Propulsion Laboratory manufactured prototypes of CMOS APS sensors that were a competition to CCD sensors [14, 15].

Cells of this sensor use the photo-effect as well and in this regard, their structure is similar to CCD sensors. But unlike the CCD sensors, the cell is directly read. Each photosensitive cell chip (each pixel) has its own amplifier and can be directly addressed and read by means of their X, Y coordinates, Fig. 2.4. This improves the image reading speed, decreases energy consumption and production of residual heat. The CMOS sensors are easier to manufactures and are used in large (close to a motion picture film) as well as extremely small sizes. Another positive aspect is that CMOS technologies are extended, relatively cheap and allow fitting the chip with other necessary elements for obtaining image such as control circuits and image digitization systems. On the contrary, having an amplifier on each cell results in lower total photo-sensitive area (the fill factor is decreased). A smaller photo-sensitive cell requires higher amplification, thus increasing an image noise. CMOS sensor manufacturers prevents this issue by using micro-lens above each photo-sensitive cell. Comparison of CCD and CMOS APS sensors is listed in Tab. 2.1, [15, 16].

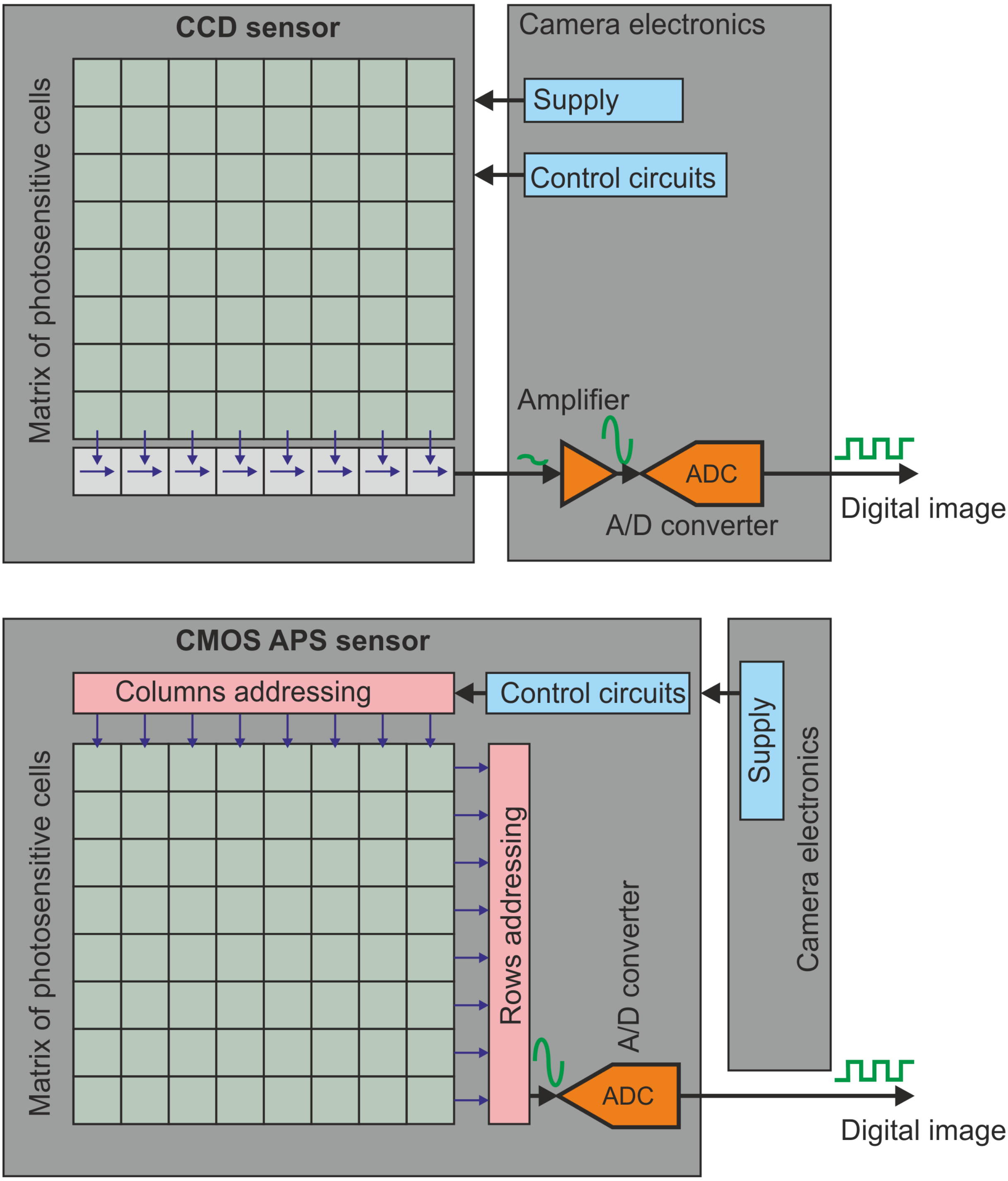

Fig. 2.4 Differences between CCD and CMOS APS sensor architecture

Tab.2.1 Comparison of CCD and CMOS APS sensors

Parameter	CCD	CMOS APS
Pixel output signal	Packet of electrons	Voltage
Chip output signal	Voltage (analogue signal)	Bits (digital signal)
Speed	Low, sequential read	High, cell addressed as matrix
Sensitivity	Very good, a high quality image can be reached, better colours	Hard to reach a high quality image in deteriorated light conditions, worse colours
Dynamic range	High	Average
Noise	Low, high image quality	Larger with regard to lower fill factor, resulting in the necessity to use micro-lens. Manifested by so called fixed pattern noise.
Energy consumption	Higher (stated up to 50 times more than CMOS APS)	Relatively small
Chip complexity and development costs	Lesser	High
System complexity	High, many circuits outside the chip	Lesser, most circuits are directly on the chip
Price	High, uses specialised manufacturing technology	Cheaper, uses standard digital logic circuits manufacturing technology.

2.3 Classification by sensor shape

Both types of sensor are used for two basic types of cameras:

- area scan cameras (chapter 3.1),
- line scan cameras (chapter 3.4).

Apart from industrial applications, area sensor cameras are used in consumer electronics such as digital cameras and camcorders of all kinds. The line scan cameras use a principle similar to scanners. Line sensors captures a single row of image points (pixels) at a time, while the resulting 2D image is obtained by moving the camera sensor of object.

2.4 Capturing in colour scale

The image capturing principle of both sensors described above is able to present frames in grey scale (often 8 bit with 256 shades of grey). Currently, grey scale cameras are the most widespread cameras in industrial practice. Naturally, image capturing in colour scale may be required in some industrial practice application. Two basic principles can be used in order to obtain a colour image: three-chip configuration and single-chip configuration fitted with a colour filter.

The three-chip configuration is based on distributing white daylight to primary colours: red (R), green (G) and blue (B). The distribution is performed by means of optical prisms featuring two semi-permeable (dichromatic) mirrors with applied colour filters (Fig. 2.5). This optical system splits the image to three CCD sensors capturing individual colour component of the image. The results is a 3-dimensional matrix of three levels.

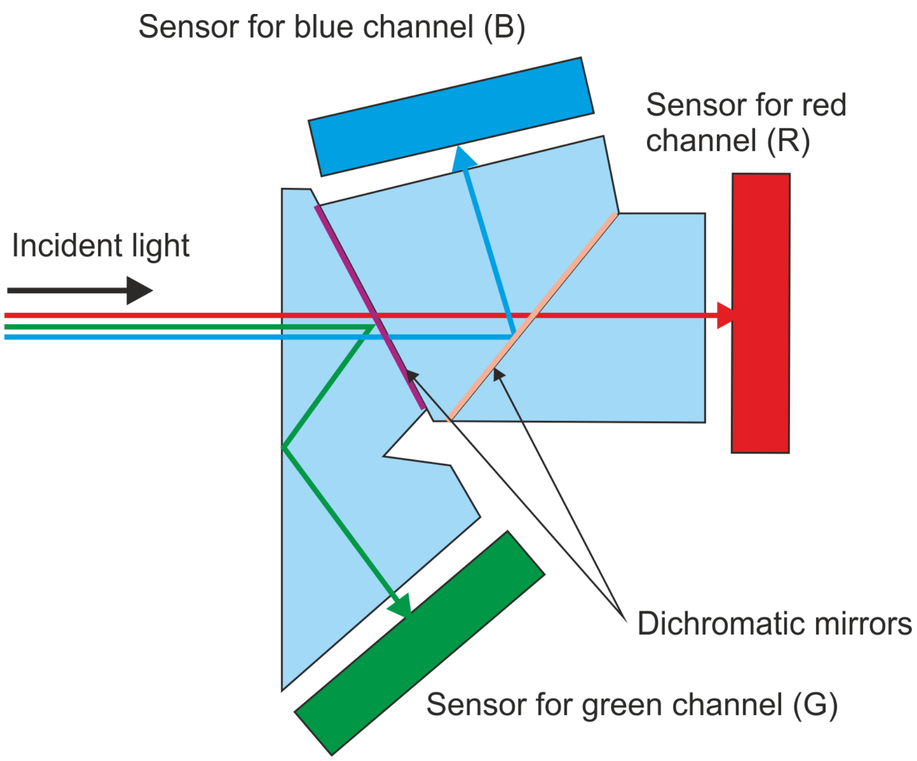

Fig. 2.5 Principle of three-chip configuration for image capturing in colour using area sensor

The advantage of this solution lies in having a full resolution RGB frames with high sensitivity, since the light is processed in its original form with its unambiguous characteristics. Due to minimal loss of light, this solution offers an image with higher brightness, smooth tonal gradations, superior depth, less colour noise, wider dynamic range. Overall, the image representation is much more natural than the one provided by single-chip systems. Clear disadvantages lie in high price, its size and weight and, last but not least, a requirement to have special (more expensive) optics.

Single-chip image capturing with colour filter is based on placing a chess patter in front of pixels of a sensor. This system often use RGB colour system, where the problem with configuration into a matrix is performed by means of colour filters, usually so called Bayer filter (Fig. 2.6, developed by Bryce Bayer in 1976, also called Bayer mask). The filter features twice as many green cells as the red and blue ones. The reason for this is that a human eye is most sensitive to yellow-green colour, therefore, information on this colour is the most important one. The image is loaded as usual, and further circuits are then used to interpolate the full scale colours of individual pixels from the closes pixels with RGB colours. The interpolation is often based on 4 neighbouring pixels, Fig. 2.7. It can cause artefacts such as colour bleeding in areas where there are abrupt changes in colour or brightness especially noticeable along sharp edges in the image. If the interpolation would not be performed, the resulting resolution would be three times smaller. That is related to lower sensitivity, since two thirds of the spectrum are filtered on each cells (e.g. Blue and green are filtered for a red filter). Reduction of this problem can be performed using so called complementary colours (cyan, magenta and yellow – CMY Fig. 2.6). The filter removes 1/3 of the spectrum of complementary colours (e.g. blue for yellow filter), therefore, this sensor is twice as sensitive. On the contrary, calculation of RGB values is somewhat more complex, thus creating an increased amount of noise [17]. Furthermore, Canon cameras and some other manufacturers use CYGM filters (cyan, yellow, green and magenta) instead of RGB. RGBE is another possible configuration, where the green pixel is replaced with colour “emerald”, like cyan (Fig. 2.6). Industrial practice applications also sometimes use filters with vertical configuration of cells (Fig. 2.6). When using this solution, it is necessary to take the 2/3 decrease of horizontal resolution instead of vertical into consideration.

In comparison with the three-chip solution, the advantage of single-chip solution lies in lower price, size and weight. The disadvantages are lower sensitivity and necessity of interpolation. However, the single-chip configuration is sufficient for most industrial applications requiring image capture in colour scale.

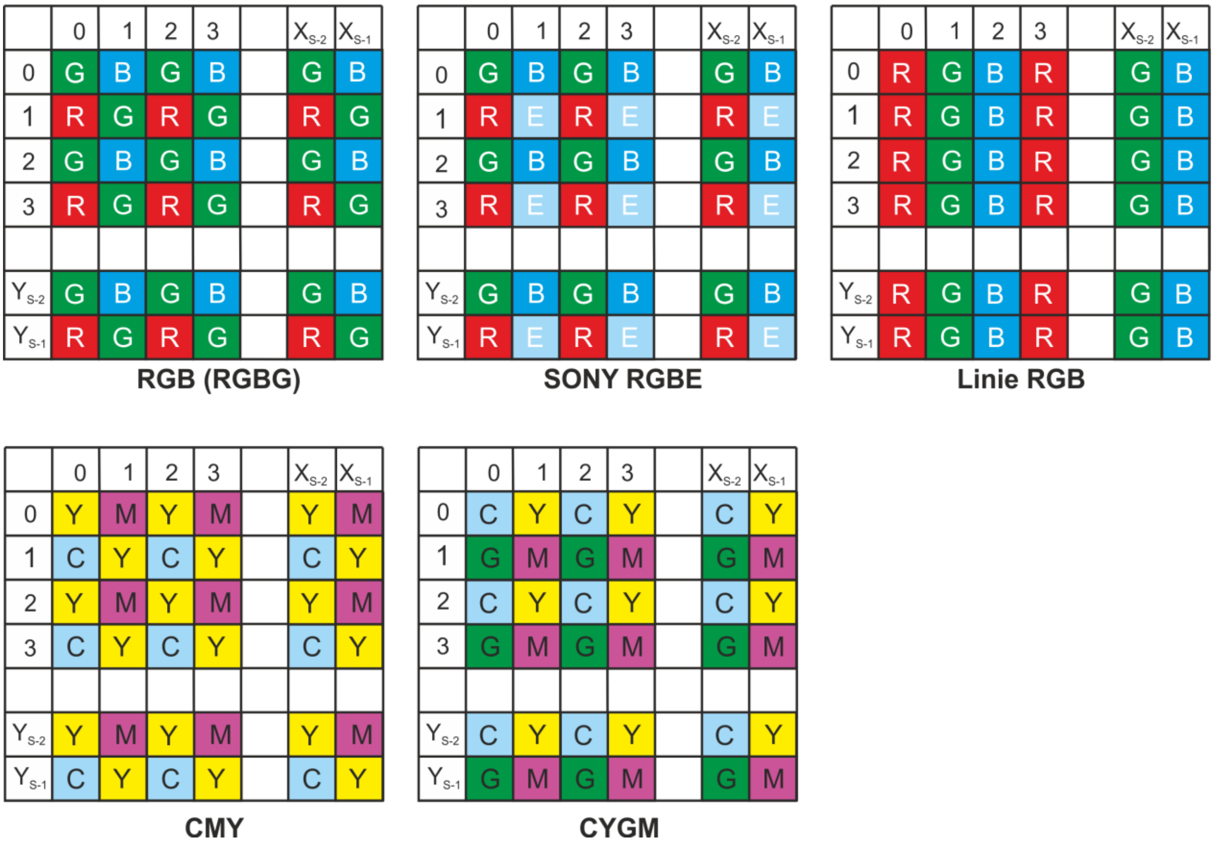

Fig. 2.6 Bayer RGB filter, Sony RGBE, CMY a C-Y-G-M filter for area sensor cameras

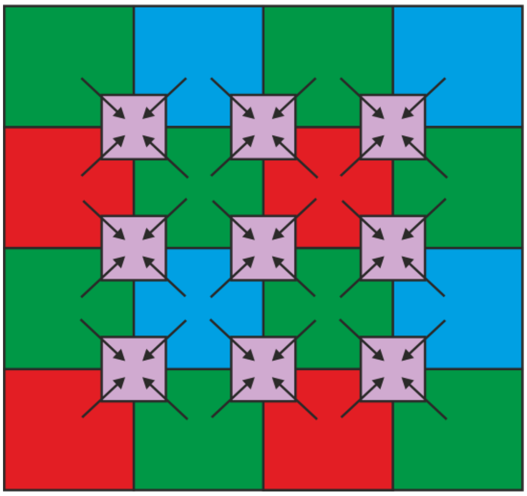

Fig. 2.7 Bayer filter interpolation

In case of **line scan cameras**, the colour image capture is performed using several principles. In the first principle, the **line sensor** is fitted with a filter with repeating basic RGB colours, Fig. 2.8 A. Advantage of this principle lies in simple synchronisation, however, as in previous case, it is necessary to take interpolation using neighbouring pixels into consideration. The interpolation again leads to decrease of sensitivity and accuracy of detail.

The second variation contains a **chip with three line sensors** below each other, while **each sensor is fitted with a filter of primary colour** Fig 2.8 B (so called three-line scan camera). The advantage is that we obtain image with true colours without losing any sensitivity. The disadvantage lies in the necessity of accurate synchronisation and following compensation, because the fact that the chip is fitted with three line sensors causes having three different parts of an object scanned at once. In case the image would not be synchronised and processed, there would be a so called halo (shift of colour components in an image). This effect must be compensated for specific camera and lens configuration. Additionally, the three-line camera cannot be used for capturing curved, bent or cylindrical surfaces.

The third option is a compromise between single-line and three-line configuration – **a dual-line chip**. Its advantage in comparison to a three-line chip is zero space between the lines, Fig. 2.8 C. This principle, as well as the single-line chip, uses Bayer coding. That leads to the disadvantage – the resolution is lowered as a result of colour interpolation. In comparison to a single-line, the resolution decrease is not as significant.

The fourth type of colour line camera is a three-chip line camera based on a principle similar to three-chip configuration of area sensor, Fig. 2.8 D. The light is separated to basic RGB colours by a set of prisms (two semi-permeable mirrors fitted with filters). By doing so, the image is split and directed to three line chips. This solution offers zero necessity of synchronisation and does not decrease resolution. On the contrary, this solution is the most expensive.

One of the terms in image capturing is colour depth. Colour depth is described in bits and represents how many colours can be recognised on the resulting frame. A 24 bit RGB colour depth is split to 8 bits for each colour (8 for red, 8 for green and 8 for blue). Therefore, it is possible to recognise 2^{24} (16 777 216) colours. Another term often related to colour depth is the A/D converter width. If the colour depth is 10 bits, then 10 bits are used for each colour channel. In total, the RGB format processed by the converter has a colour depth of 30 bits. However, the resulting colour depth can be lower – a cut of “unnecessary” bits is often performed after the A/D conversion.

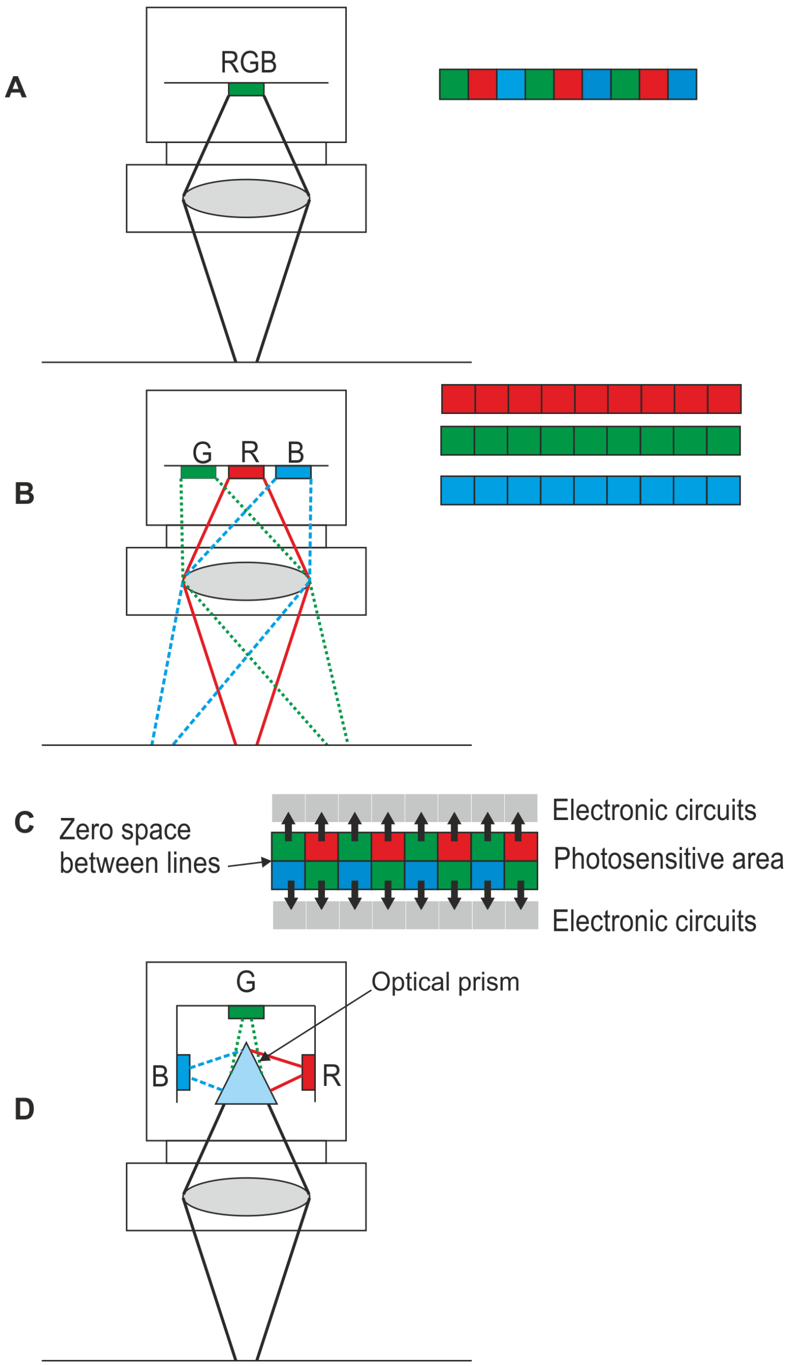

Fig. 2.8 Variants of line cameras fitted with filters for colour image capture

Aside from the aforementioned principle, one should realise that various sensors produced by various manufacturer may offer slightly different characteristics and properties, which also applies to expression of colours and range of captured spectrum. Therefore, when designing specific applications, this phenomenon should be regarded and appropriate camera and sensor manufacturer should be selected.

The aforementioned sensors are not the only one in the field of digital photography. There are many other types being developed, although these are usually modifications of the described sensors. For example Super CCD that differs from the classic CCD only by the shape of photo-sensitive cells those are octagonal and therefore provide better cover of the sensor surface [18]. The fourth generation of these chips is even fitter with secondary photoreceptor cells used to obtain complementary balance information (mostly for balancing white). Another example can be Super CCD EXR showing very interesting results, especially when capturing in difficult light conditions (such as very contrast scenes or insufficient lighting). One may encounter a variation of CMOS APS, where the Bayer filter is replaced by a filter using characteristics of silicone absorbing various ranges of light according to its layer thickness. When light passes through the filter layer, blue component is filter first, then green followed by red. That way, the photo-sensitive cell obtains information about the intensity of all three colours (CMOS Foveon X3 [18]).

2.5 Pixel binning

These days, many cameras feature pixel binning. It is a principle of merging neighbouring pixels and can be used for example in deteriorated lighting conditions or when it is necessary to increase frame rate. Pixel binning is based on merging neighbouring pixels (Fig. 2.9) acting as a single pixel with multiplied area. The camera sensitivity is multiplied as a result and/or offer increase of maximum frame rate while maintaining the field of view. On the contrary, the resolution is decreased proportionally. Pixel binning can be performed in area sensors as well as line sensors. Fig. 2.9 A shows an area sensor without the pixel merging feature, while Fig. 2.9 B shows horizontal pixel binning. Finally, Fig. 2.9 C shows vertical pixel binning. In both cases, the merging process results in doubling the pixel surface, decreasing the resolution to half. When performing horizontal AND horizontal merging, four neighbouring pixels are merged (Full Pixel Binning). The pixel surface (sensitivity) increases four times, while the resolution decreases to a quarter, Fig. 2.9 D. When using line sensors, two neighbouring pixels are merged, Fig. 2.9 F. In case of two lines, it is possible to merge two vertical (Fig. 2.9 G) or even four (vertical and horizontal pixel binning, Fig. 2.9 H) neighbouring pixels.

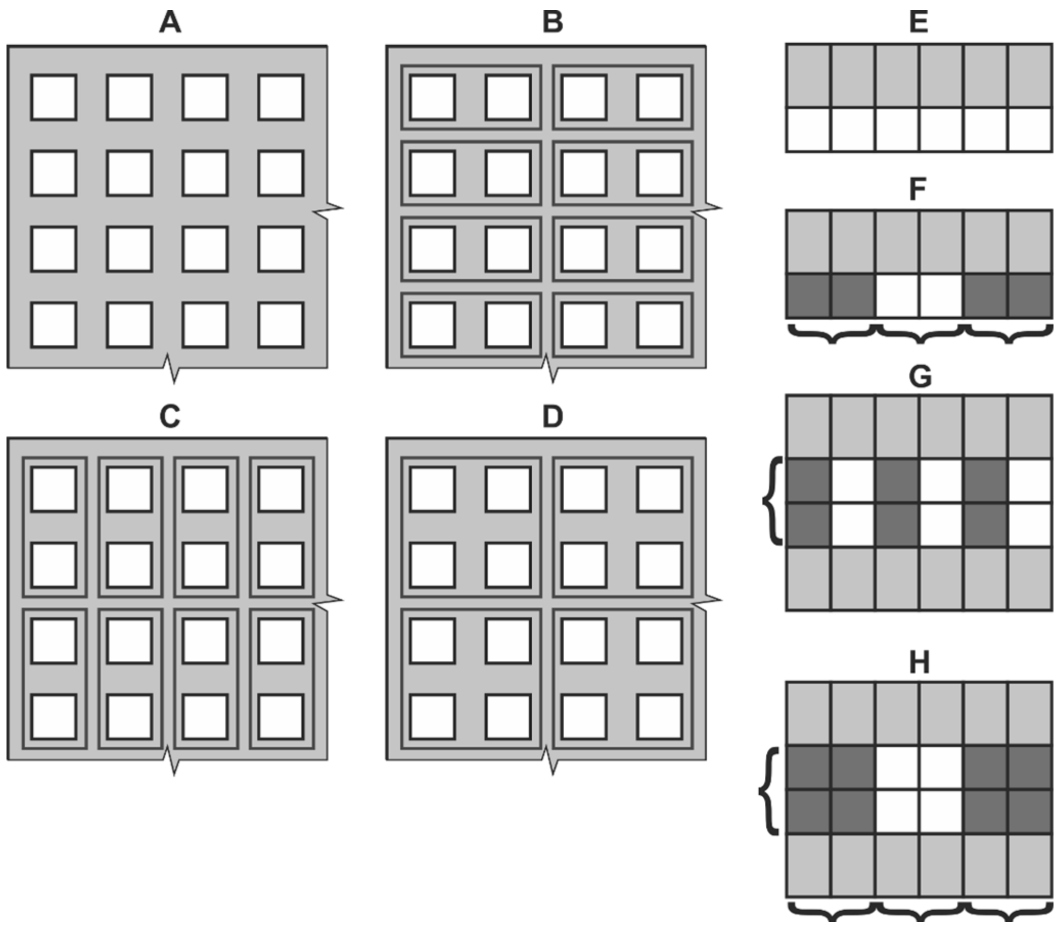

Fig. 2.9 Pixel binning – principle of merging neighbouring pixels: A - area sensor without merging neighbouring pixels, B – area sensor with horizontal merging, C – area sensor with vertical merging, D – area sensor with horizontal and vertical merging, E – line sensor without merging neighbouring pixels, F – line sensor with horizontal merging, G – dual-line sensor with vertical merging, H – vertical and horizontal pixel binning

2.6 Noise, vignetting, blooming

There are several characteristic problems related to obtaining a digital image. Some of these problems are directly or indirectly related to the sensor – e.g. noise,, vignetting and blooming.

Noise occurs due to many various causes. However, the primary one is the ***thermal motion of the semiconductor crystal lattice*** (referred to as "dark current"). During this phenomenon, an electron is released without being affected by a photon. Such electron is drawn to the exposition electrode and is then added to the light exposition value of such cell. Since the immediate noise is different from cell to cell and exposition to exposition, it is impossible to entirely eliminate the noise. The noise is primarily influence by the following factors: size of the photo-sensitive light (noise generated in a small photo-sensitive cell is basically the same as in three times larger cell), sensor heating (higher temperatures of a sensor lead to higher probability of random electron release), intensity (brightness) of the light incident on a sensor (lesser amount of incident light increases the relative noise ratio SNR[4]), exposition time (longer expositions cause higher noise) and electromagnetic field (the noise increases with increasing intensity of electromagnetic field).

The issue can be addressed in several ways. The first one is to increase the sensor cell size. The cell should be as large as possible. By having large capacity of a cell results in relative decrease of SNR. That is why small-format sensor(s) with high resolution always has significantly deteriorated noise properties than larger sensors with lower resolution. The other approach is implemented in scientific devices, where the noise is being decreased by cooling the sensor (usually by means of liquid nitrogen, or thermoelectric like Peltier modules). The third and the most suitable for technical practice is increasing intensity of light incident to a sensor, thus decreasing the SNR. That means to increase light intensity of the technological scene that is being captured (most usually performed by increasing light source luminous flux). Although opening the aperture and decreasing shutter speed (while keeping the EV value) increases the amount of incident light, the shutter speed is

[4] SNR – Signal Noise Ratio, ratio of undistorted signal and noise

prolonged as well, resulting in increased noise. It is possible to eliminate the noise by means of various software filters [1, chapter 7]. An interesting solution is to capture two and more frames shortly after each other, while these frames are basically identical, except the random noise. That way, it is possible to make a quite accurate determination regarding the noise intensity in individual photo-sensitive cells and then eliminate the noise by means of software.

Vignetting is based on optical laws of camera lens system and is manifested by gradual dimming of the image towards the edges, for more information about vignetting, see chapter 1.3.1 and chapter 4.2. The vignetting effect is also increased by using micro-lens fitted to sensors with the aim to increase the fill factor. Vignetting can be suppressed by using optical elements or eliminated by means of pre-processing software filters.

Blooming is a phenomenon that occurs when using electronic shutter with CCD sensors. If there is too much light incident on some of the pixels, their capacity overflows. The excess electrons are then dispersed to other pixels in a line, forming parallel lines of irregular lengths around the strong lighted areas. This phenomenon can be prevented by using anti-blooming gates, however, these take up more than 30 % of sensor surface, thus decreasing the fill factor. The second solution is to use a series of short expositions instead of a long one (used in special photography, e.g. in astronomy).

3 Industrial cameras for image capture

When performing image analysis, the data are used mostly by industrial or digital cameras. Cameras can be generally divided by shape of the capturing element to:

1. Area scan cameras with their capturing pixels in a matrix:

   a) Standard (common cameras):

   b) Intelligent (camera features advanced hardware and software for frame processing, while the output is the image and results of image analysis);

   c) Camera sensors (camera includes a simple hardware and software for image processing, however, the output is only in a form of logic output – yes/no, for example is the product present or not? Does it have satisfactory quality or not?).

2. Line scan cameras with capturing cells in one, two, three or four lines.

3. 3D cameras.

Logically, digital commercial cameras have only area sensors. On the contrary, scanner features linear configuration (those can be used in for example in laboratories for image capture of textile texture for analysis purposes).

3.1 Standard cameras with area sensor

When selecting standard camera, it is important to know what technological scene needs to be captured, how to capture it. It is also important to be aware of limiting conditions of the camera and to respect the purpose of capturing and purpose of subsequent evaluation by means of image analysis. Between the basic characteristics of area sensor cameras are: type of sensor (listed in chapters 2 and 2.2), sensor size, area resolution of the sensor, capturing (frequency of) speed (time resolution of the sensor), spectral range, colour scale, range of shutter speeds, interface, output data format, camera control features (inputs), lens mount, mounting method of camera, power supply, resistance to external influences (cover, resistance to vibrations, resistance to mechanical stress, range of ambient temperature and humidity during camera's operation, …) and more (manufacturer of camera and sensor, …).

3.1.1 Sensor size

Sensor size is given in inches and does not correspond directly with size of the sensor. It is rather a historic label of image capturing tubes used in TV cameras before CCD sensors were developed – those tubed were called vidicon tubes. The size was given by the outer diameter of glass cylinder of the tube. Diagonal of the usable sensitive layer for image capture was only approx. 2/3 of this diameter. With CCD and CMOS APS, this convention was maintained. The common manufactured sizes of optical sensors are listed in Tab. 3.1.

Tab. 3.1 Sizes of common optical sensors

Type	Ratio W:H [-]	Width [mm]	Height [mm]	Diagonal [mm]	Surface [mm^2]	Relative surface [-]
1/6"	4:3	2.300	1.730	2.878	3.979	1.000
1/4"	4:3	3.200	2.400	4.000	7.680	1.930
1/3.6"	4:3	4.000	3.000	5.000	12.000	3.016
1/3.2"	4:3	4.536	3.416	5.678	15.495	3.894
1/3"	4:3	4.800	3.600	6.000	17.280	4.343
1/2.7"	4:3	5.270	3.960	6.592	20.869	5.245
1/2"	4:3	6.400	4.800	8.000	30.720	7.721
1/1.8"	4:3	7.176	5.319	8.932	38.169	9.593
2/3"	4:3	8.800	6.600	11.000	58.080	14.597
1"	4:3	12.800	9.600	16.000	122.880	30.882
4/3"	4:3	18.000	13.500	22.500	243.000	61.070
APS-C	3:2	25.100	16.700	30.148	419.170	105.346
35mm	3:2	36.000	24.000	43.267	864.000	217.140
645	4:3	56.000	41.500	69.701	2324.00	584.066

Sensor size affects the size of photo-sensitive cell. Therefore, larger sensors with identical resolution have larger pixels, resulting in decreased noise and image with higher quality. On the contrary, it is necessary to take the lens focal length for a given angular field of view and a given field of view into consideration. To achieve identical angular field of view, large sensors require lens with higher focal length. Lens with higher focal length are larger, heavier and more expensive (but can be used with common lens). Increasing focal length results in decreasing depth of field, which can be undesirable in some technical applications.

As suggested by the aforementioned, the sensor size is important for selecting an appropriate lens for capturing the given technological scene. Size of individual pixels is important as well. Sensors with smaller pixels require more expensive lens with higher quality. Price of the sensor is important as well and is determined by its size. When deciding about size of the sensor, while the area resolution remains the same, the properties listed in Tab. 3.2 should be taken into consideration.

Tab. 3.2 Comparison of sensor properties by size, while having the same resolution

Small sensor	Cheap, low energy consumption, more expensive lens, but has blurred image, higher noise, low quality colours, low sensitivity.
Large sensor	More expensive, higher energy consumption, common lens, but for large formats, sharp and contrast image, lesser colour deficiency, better colours, higher sensitivity.

3.1.2 Area resolution of the sensor

When selecting area resolution of a sensor, one should address two issues. The first one is the resolution (sampling), which is the distance between the closest captured (sampled) image points (pixels), itself. Issue regarding the area sampling frequency (distance of sampling points) can be solved by means of Shannon theorem about sampling [2]. The theorem says that the sampling frequency must be at least twice as high as the highest “interesting” frequency in the sampled signal. When a lower sampling frequency is used (lower resolution) a so called aliasing [19, 20] may occur. During this phenomenon, the reconstructed signal may be slightly different from the original one. The sensor resolution is therefore primarily determined by the required accuracy of measurement. Due to the aforementioned conditions, the resolution must be at least twice the smallest measurable dimension.

However, this theoretical accuracy and the related measurement repeatability are decreased by the image noise, lossy compression of image data and other factors. When using a single-chip colour sensor, the accuracy is decreased due to the performed colour interpolation and the mutual shift of image in individual colour channels. The considered linear resolution should be approximately half. In both, black and white as well as colour image capturing, one should consider creating a reserve depending on the observed object and the available sensors with appropriate resolution. Experience shows that when processing an image, the sampling should be at least 5 times softer than the theoretical limit given by the sampling theorem. Therefore, resolution should be at least five times higher than the theoretical limit.

The other issue is selecting a sampling grid – areal configuration of points for sampling. There are only three polygonal shapes available entirely covering the network: triangles, squares, regular hexagons. In practice, we usually use squares to form the grid, even though squares cause problems with connecting individual areas, [1, chapter 1]. Individual cells of the matrix then represent pixels.

A hexagonal grid is used in rather exceptional cases (the aforementioned regular hexagons) and solves some of the paradoxes mentioned below. All six neighbouring cells of this grid has identical properties, especially then the distance. Further usage of this shape is prevented by massive spread of square raster used by most digitisation devices as well as by the fact that hexagonal raster is unsuitable for some operations such as Fourier frequency filtration [2]. These days, the matrix configuration is used in vast majority of cases.

Area resolution of a sensor is an important parameter stating the amount of pixels on a chip used to achieve the required image resolution and is stated in mega-pixels. These days, this parameter is often overemphasises, since resolution of up to 2 Mpx is sufficient for the most common technical applications. Additionally, it is important to realise the dependence of sensor price on resolution. Industrial cameras are several times more expensive than common commercial cameras, and the increase of resolution is directly linked to significant increase of price. When designing the hardware of the system, it is therefore necessary to consider what objects will be observed and determine real requirements on resolution and size of the sensor.

These days, the resolution of industrial camera sensors are ranging from 0.31 Mpx (labelled VGA) to 29 Mpx. The lower values are often labelled by the computer screen resolution, Tab. 3.3. There are many other labels.

Tab. 3.3 The most common area resolution of sensors

Label	Amount of pixels on edges *	Resolution
VGA	640 x 480	0.31 Mpx
WVGA	800 x 480	0.38 Mpx
CCIR	720 x 576	0.41 Mpx
SVGA	800 x 600	0.48 Mpx
XGA	1024 x 768	0.79 Mpx
SXGA (1.3 Megapixel)	1280 x 1024	1.3 Mpx
1.4 Megapixel (XGA-2)	1360 x 1024	1.4 Mpx
UXGA (2 Megapixel)	1600 x 1200	1.9 Mpx
HDTV 1080p	1920 x 1080	2.1 Mpx
QXGA	2048 x 1536	3.1 Mpx
4 Megapixel	2048 x 2048	4.2 Mpx
QSXGA (5 Megapixel)	2560 x 2048	5.2 Mpx
8 Megapixel**	3296 x 2472	8.1 Mpx
16 Megapixel**	4872 x 3248	15.8 Mpx
20 Megapixel**	5120 x 3840	19.7 Mpx
29 Megapixel***	6576 x 4384	29 Mpx

** amount of pixels on the edges can be slightly difference from manufacturer to manufacturer*

*** according to offer of JAI company [21]*

**** according to offer of PixeLINK company, camera PL-H9629 [22]*

3.1.3 Frame rate with full resolution – time resolution of the sensor

Frame rate (speed) with full resolution is an important parameter for obtaining images in relatively fast processes. When selecting a camera, it is necessary to know speed of the observed process and use a camera with appropriate reserve. Frame rate is stated in frames per second (fps). Currently, the fps values are commonly ranging from 3 fps (colour cameras fitted with 16 Mpx CCD sensors, e.g. JAI AB-

1600CL camera, but PixeLINK PL-H9629 with its resolution of 29 Mpx is able to capture with frequency of 6.2 fps of colour signal, CCD) up to 500 fps ("High Speed Cameras" with 1.3 Mpx resolution exclusively fitted with CMOS APS sensor). Today camera producers offer camera with region of interest (ROI) reducing and increasing of frame rate (e.g. Basler acA2500-60um, ROI: 640 x 480 @ 978 fps). Frame rates are not standardised, one may therefore see various values. However, the 1394 Trade Association document [23] states for example a range of ..., 1.875; 3.75; 7.5; 15; 30; 60; 120; 240; ... In general, the camera frame rate is proportional to sensor resolution and the applied communication interface that is able to transmit only a limited amount of data. Frame rate is also often affected by the output image format of the camera, specifically its data requirements. Using RAW format means the highest requirements on data transmission. Even though the manufacturer of cameras states frame rate for the given cameras, it is necessary to consider the limits of an image processing system – the performance of a control unit. Frame rate is therefore not just an issue of camera and given communication interface. The real amount of captured frames is affected by the interface, acquisition software, the selected image processing analysis, hardware of the evaluation unit (computer) and more. Special cameras are able to achieve the order of thousands frames per second. Cameras with extreme frame rates (trillion of frames per second) are no longer fitted with CCD or CMOS APS [24] sensors.

3.1.4 Spectral range

Cameras can be split into categories by the spectral range of electromagnetic waves they are operating in: LWIR (long-wave infra-red radiation), MWIR (medium-wave infra-red radiation), SWIR (short-wave infra-red radiation), NIR (near infra-red radiation), visible radiation, NUV (near ultraviolet radiation), UV (ultraviolet radiation) and X-Ray radiation.

Most cameras are manufactured to capture a range of visible spectrum – 380 to 780 nm (labelled VIS as visible). The sensors are able to capture other wave lengths, which is why we use an IR-cut filter that prevents the infra-red component to exist in the resulting image. IR-cut filter can be placed by the manufacturer either directly in front of the protective glass in front of the sensor (in case of monochromatic cameras) or to camera colour filters (e.g. Bayer filter) when capturing colours. Also, these filters can be places in the lens. Lens with IR correction are optimised for VIS as well as NIR. When using other lens in NIR, the plane of sharpness is moved beyond the chip rendering the image blurry.

Additionally, there are variants for visible spectrum with near infra-red radiation (VIS + NIR), where the range of scanned spectrum is moved to near range of infra-red radiation (IR). The stated spectral range is from 400 to 1000 nm [25]. Radiation of NIR wave length is not visible by human eye, but unexpected reflection may occur if the NIR radiation is of higher intensity and may easily cause overflow of capacity of individual pixels. The excess electrons are then dispersed to other pixels. The result of such occurrence is a devalued image. The sensitivity to near infra-red radiation is used mainly for transportation, e.g. for measurement of vehicle speed complemented by driver photography. In order to avoid a case where a driver is dazzled by the camera flash, systems use infra-red lighting that is invisible to human eye but is visible to cameras operating in the NIR range. Applications for detection of faults, surface adjustments and insufficiencies in industrial practice are usually using infra-red ranges, since the detection is easier to perform. The visible radiation can be filtered by means of IR filters so that the area can be captured only in NIR range. Additionally, one should consider using specially adjusted lens (with the required characteristics), which is also more expensive [26].

It might be also wise to use cameras with a range shifted to the second border of visible spectrum – into the ultraviolet radiation (VIS + UV). The spectrum is stated to be ranging from 200 to 800 nm. Sensitivity to ultraviolet radiation can be achieved in several ways. First method is to remove the lens as well as the cover in front of the sensor. In this case however, the guarantee of camera becomes void. Another option is to use sealing quartz glass above the sensor and to remove lens. The best and the most expensive is using a special sensor for UV radiation. This sensor is, again, without micro-lens and is fitted with highly sensitive quartz glass. Cameras operating in this spectrum allow better observation of surface dirt and adjustments as well as presence of water or other small particles [21].

3.1.5 Colour scale

The colour scale is given by the camera configuration and may be either monochromatic (image capturing in grey scale) or colour. The principles of colour cameras were described in chapter 2.4 and in general, requirement of higher resolution and sensitivity of a single-sensor camera makes using a grey scale camera necessary. Each photo-sensitive cell generates one pixel (colour is not then interpolated using neighbouring cells). Also, parts of spectrum are not filtered.

3.1.6 Range of shutter speeds and shutter

The ranges of shutter speed and shutter are usually from 0.04 ms (1/25000 s) to 10 000 ms. In older devices, the exposure (illumination time of photosensitive cells) was performed by means of a mechanical shutter. These days, most technical applications use an electronic one (chapter 2 and 2.6).

3.1.7 Communication interface

Communication interface allows communication between a camera and a control unit (computer). There are several types of interfaces, while each of them offers different properties and is usable for different types of applications. Selecting camera depends on the communication methods of the image processing system. Sometimes, camera manufacturers offer the same model with more variants, each with different type of connection. Selection of the interface is primarily based on:

- data transmission speed (amount of data that must be transmitted to a computer for a given period of time),
- latency (a time between image capturing and its display/delivery for processing);
- distance of data transmission (distance between the computer and camera),
- power supply method (communication interface may ensure camera power supply),
- delay of commands (computer-to-camera or vice versa, this parameter is important mainly with high-speed processes),
- possibility of connecting multiple cameras (if required),

- CPU or GPU load,
- price etc.

Basic types of interface for connecting cameras to a computer are: USB 2.0 and 3.0, FireWire (IEEE 1397 a/b), Gigabit Ethernet (GigE Vision) and Camera Link. Tab. 3.4 lists an overview of main properties of individual interfaces.

Tab. 3.4 Selected properties of the most widespread communication interface

Label	USB 2.0	USB 3.0 (USB3 Vision)	IEEE 1394a	IEEE 1394b	Gigabit Ethernet (GigE Vision)	Camera Link
Maximum transmission speed	480 Mb/s	4800 Mb/s	400 Mb/s	800 Mb/s	1000 Mb/s	>2000 Mb/s (up to 5 440 Mb/s)
Bandwidth	45 MB/s	600 MB/s	Video: 32 MB/s (80%) Total: 40 MB/s	Video: 64 MB/s (80%) Total: 80 MB/s	90 MB/s (120 MB/s for GigE bus)	255 MB/s (basic) 680 MB/s (full)
Maximum cable length	5 m	5 m	4.5 m	4.5 m	100 m	10 m
Maximum length with amplification	30 m	30 m	72 m	72 m	unlimited	30 m
Camera power supply	up to 0.5 A, 5V, via USB	up to 0.9 A, 5V, via USB	up to 1.5A, 36V, using a FireWire cable	up to 1.5A, 36V, using a FireWire cable	separately (except PoE)	separately (except PoCL)
Motherboard support	all in general	relatively new motherboards	many	some	some	none
OS support	Windows, Linux	Windows, Linux	Windows, Linux	Windows, Linux	Windows, Linux	depends on the supplier
Delay of command (computer-to-camera)				350 – 500 μs	600 – 900 μs	up to 100 μs
Delay of command (camera-to-computer)				13 – 14 ms	13 – 14 ms	
Computer processor load	average (5 to 10%)	very low (less than 2%)	very low (less than 1%)	very low (less than 1%)	average (5 to 10%)	very low (less than 1 %)
Connection of multiple cameras		yes, via HUB	theoretically up to 63 devices	4 PCI cards ports, HUB is usable	theoretically unlimited	difficult
Wireless support	no	no	no	no	yes	no
Relative price	low	low	medium	medium	low	high

USB 2.0

Using USB 2.0 interface is good for short distances between the computer and the camera and for low data flow (low resolution or frame rate). Interface allows connection and power supply of the camera by means of a single cable, allowing simpler solution and connection. Unfortunately, the power supply may not be sufficient in some cases. The general support and the possibility to connect the camera practically to any computer with installed driver and software for image display is important as well. Unlike the USB3 Vision, this interface does not have a standard and each camera manufacturer offers own drivers. Camera can be theoretically connected to a distance of up to 30 m using a signal amplifier integrated to each cable. However, electromagnetic interference may occur. The aforementioned USB 2.0 interface is not widely used in industry and one may usually find this solution in laboratories or non-industrial applications. When designing a system, it is currently much wiser to use USB 3.0 or GigE.

USB 3.0

The upcoming USB 3.0 should be mentioned as well (presented in 2010), since it offer better conditions for communicating with industrial cameras (Fig. 3.1). Transmission speed is up to 4 800 Mb/s, therefore a theoretical bandwidth of up to 600 MB/s. It also offers better camera power supply (higher power of the USB interface) of up to 0.9 A. Cable length remains at "up to 5m" if not fitted with an amplifier. The primary advantage lies in simple connection of these cameras. In comparison to USB 2.0, connection management was improved and offers higher connection reliability and significantly lesser CPU load. Most USB 3.0 ports allow using DMA (Direct Memory Access) in order to minimise the CPU load when transmitting high data volumes. Similarly to a Gigabit Ethernet, USB 3.0 offers a USB3 Vision protocol to allow reliable use of cameras in machine vision systems as a real-time device. Compatibility of USB3 vision and GenlCam[5] allows mutual

[5] The GenlCam standard defines a common programming interface for industrial cameras (or devices in general) in terms of their configuration and image capture. The system is independent on the transmission technology (GigE Vision, Camera Link, USB, Fire Wire).

interconnection of systems and cameras of various manufacturers without the necessity to change drivers or applications [26].

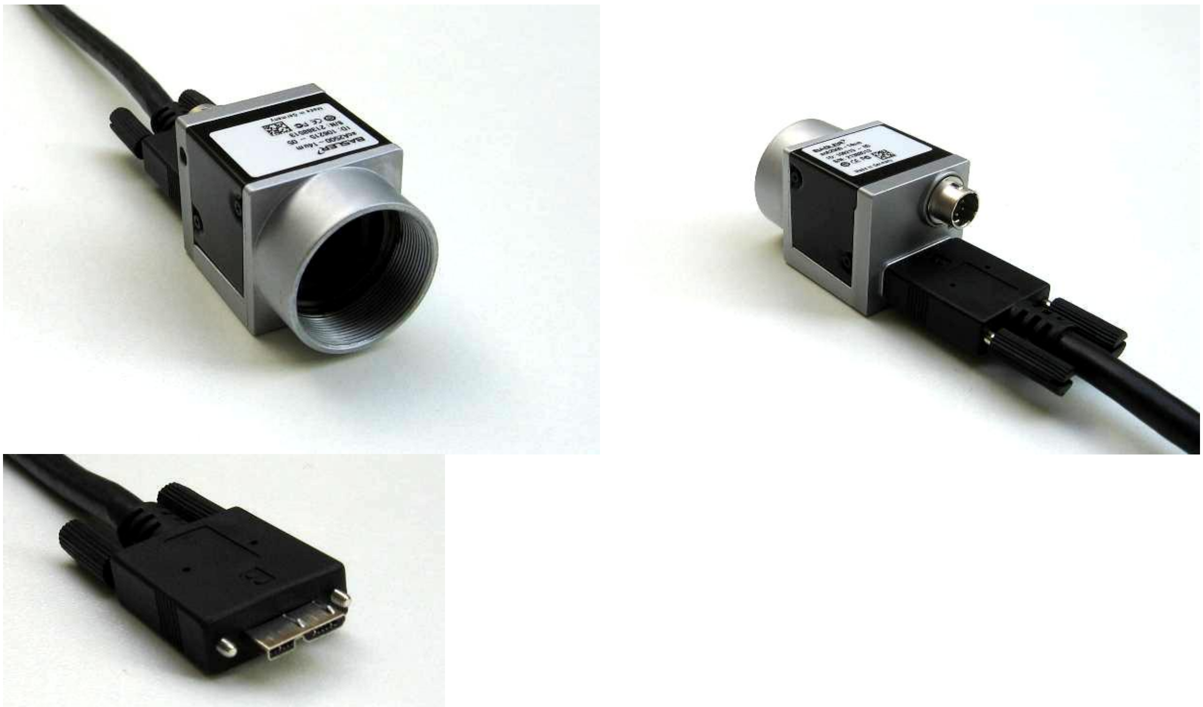

Fig. 3.1 Camera with an area sensor and USB 3.0 connector

IEEE1394a/b

This communication interface is called FireWire, and as in USB, this method is suitable for short ranges of up to 4.5m (using a repeater increases the range to up to 72m). Maximum transmission speed changes according to the variants, which are now formed by two standards: 393Mb/s for type “a” (FireWire 400) and 786Mb/s for type “b” (FireWire 800). Applicable specification IEEE 1394b from 2002 provides transmission speeds of 3 200Mb/s (which is 400MB/s) with distance of up to 100m (by means of a beta mode and optical connection). There are also specifications such as FireWire S1600 and S3200 with transmission speed of 1 600 or 3 200 Mb/s, while having the same connectors as FireWire 800. Furthermore, there is a definition of IEEE 1394c standard being developed that uses an Ethernet connector. However, the device does not exist yet and its implementation to industrial camera systems is not currently planned. That is even due to development of USB 3.0 that offers similar properties.

FireWire 400 uses a classic 6pin connector, Fig. 3.2. Laptops and other peripheries use 4pin micro-connector that however contains only data pins and it is therefore necessary to use auxiliary power supply via cable. These days, this interface is used with older cameras. Newer systems usually use FireWire 800.

Fig. 3.2 Camera with area sensor and IEEE 1394a connector

FireWire 800 features a 9pin connector. The bus is compatible with the previous standard (400), therefore, it is possible to connect a camera operating on IEEE 1394b to an IEEE1394a interface and vice versa, but the communication is performed with lower speed. For this purpose, there are cables with both connectors.

FireWire can use a tree topology using a HUB to connect up to 63 devices. The system allows peer-to-peer communication, similar to scanner-printer communication, in order to avoid usage of system memory and CPU. FireWire features Plug and play support and hot swap. The six- and nine-line variant may feature power supply of up to 45W with 30V, allowing operation of device with medium power consumption without any additional power supply.

Comparison with USB 2.0: PC does not usually offer transmission speed higher than 280 Mb/s, and is usually 240 Mb/s. Lower speed is probably caused by relying on computer CPU that manages the low-level protocol elements, while

FireWire delegates the same tasks to hardware interface (requiring lower CPU usage). Furthermore, FireWire is considered to be more stable for transmission of image data and provides better power supply to devices connected to the device [27].

Comparison with USB 3.0: When comparing FireWire and USB 3.0, USB wins on all fronts. New devices does not feature FireWire any more, therefore its usage in applications is decreasing. The main reason for this is the limited cable length and low communication speed. When planning new projects, we recommend using cameras with USB 3.0 interface with characteristics similar to FireWire [26].

GigE Vision[6]

GigE interface is used when multiple cameras are connected, and (or) when the distance between the camera and computer is long. The advantage is the easy connection, where the camera acts as a standard network device, enabling to connect tens of cameras connected to the Ethernet network to a computer with a single network card, Fig. 3.3. It is therefore theoretically possible to use a single computer for observing and processing image from unlimited amount of cameras. Another advantage is that a standard Ethernet cable of category 5e or 6 as well as usual network elements (NIC, switch, hub) can be used for communication. It is also possible to use remote connection for example by means of VPN or Wi-Fi.

The disadvantage is the higher load of CPU and longer delay when executing the exposition by means of a software trigger.

Although the communication works with any network card, Basler and JAI recommend to use proven adapters with Intel PRO/1000 chip set that is optimised for transmission and buffer usage. The CPU is not necessarily loaded. It is recommended to use devices supporting so called Jumbo Frames in order to achieve optimal performance with higher data flows. The cameras and devices should be set to the highest possible size of transmitted packet. When using GigE and IEEE1394b,

[6] Standard GigE Vision defines a set of communication protocols for industrial cameras that use Ethernet as the transmission medium.

it is wise to use cables with screw connectors, at least on the camera side. Doing so prevents an undesired communication failure due to vibrations or unwary intervention of the operator.

Fig. 3.3 Camera with area sensor and GigE connector

Comparison to IEEE1394b: Practical experience with cameras of both interfaces states that IEEE1394b seems to be more reliable and suffers less from communication failure than Ethernet with classic TPC/IP protocol used for Ethernet data communication. GigE Vision protocol was developed for camera applications using gigabit Ethernet, making the communication faster and more reliable. (Unlike TCP IP, GigE Vision is above the UDP protocol – lower latency, better control and data throughput.) Also, GigE Vision and GenlCam standards allow combining cameras, hardware and software of various manufacturers.

The advantage of IEEE1394b lies in its ability to power the camera directly via the communication cable, therefore the system does not require an auxiliary power source. On the other hand, there is a limitation of the amount of cameras and the distance between the evaluation unit and camera. GigE interface also allows wireless transmission.

There are also GigE cameras powered through the Ethernet – so called PoE (Power over Ethernet). In such case, camera is powered by Ethernet cable lines plugged into the camera. The camera input can be up to 15.4W or 25.5W for PoE+, while maintaining the same properties as GigE.

Camera Link

The capacity of gigabit Ethernet is insufficient for high resolution and frame rates, which is why we use the Camera Link bus with a transmission speed of 5,540 Mb/s (there are also cards allowing transmission speed of up to 6,800 Mb/s, for example NI PCIe-1433). The interface is used for high-performance cameras with a frame rate of hundreds of frames per second or high resolution line scan cameras. The disadvantage of Camera Link is the necessity to use a PC special card (frame grabber) along with the limited maximum cable length of 10 m (extendible to 30 m with amplifiers).

The connection with a camera is performed by means of a 26pin MDR connector consisting of a Channel Link and system for serial communication with the camera, Fig. 3.4. Due to small size of some cameras, many of them use an HDR connector, often called Mini-Camera Link. The basic configuration is called camera Link Base and its maximum capacity is 255 MB/s. Applications with higher capacities use an extended configuration called Medium or Full that uses an additional cable with two Channel Link channels. By doing so, the bandwidth increases to 510 MB/s (Medium) or 680 MB/s (Full) respectively. There is also an additional Extended-Full configuration that extends that bandwidth up to 850 MB/s [28]. External trigger is not connected to digital output of the camera, but directly to frame grabber fitted with a corresponding input.

Many small cameras can be powered directly from the Camera Link interface (must be supported by the frame grabber). This technology is called PoCL (Power over Camera Link) or PMCL (Power over Mini-Camera Link). The camera power consumption must be only 4 W (theoretically 8 W).

Comparison with GigE: The Camera Link interface currently offers the highest transmission speed of all the standard interfaces. Its disadvantage lies in short maximum distance between camera and computer as well as the necessity to use a special PCI card, which is not necessary with GigE communication interface.

Fig. 3.4 Line scan camera and Camera Link connector

Naturally, there are other, less usual, communication interfaces. One of them is ***RS-644*** interface called LVDS (Low Voltage Differential Signalling) on which Camera Link is based [29].

Another example is less usual interfaces is ***CoaXPress*** [30, 26], a system that uses data transmission with 780 MB/s bandwidth within a distance of 40 metres and 390 MB/s up to a distance of 100 m. This interface also allows 2.5 MB/s camera management. If the transmission speed is to be increased, all you need to do is connect more coaxial cables (in theory, a bandwidth of 3,125 MB/s can be reached when using 4 cables). CoaXPress is called a “new generation” interface. Its task is to reach high data throughput without limitations of for example Camera Link. Price of the cables as well as its maximum possible length is a great advantage (plain coaxial cable). Furthermore, this method uses a elaborate communication protocol that allows synchronisation with the camera with very low latency.

Analogue interfaces are no longer an up-to-date alternative. These interfaces are used to digitize the analogue camera signal. In comparison to other communication interfaces, the transmission speeds are rather low.

3.1.8 Format of output video signal

The format of output video signal is a parameter that is often ignored. The best image quality is offered by cameras providing raw image data (so called RAW). The cameras do not transform the RAW image in any way, including colour balancing, interpolation. Additionally, the image is not compressed in any way. That makes image with high accuracy available in the connected computer. On the other hand, the high data flow may be a problem, since the requirements on communication interface are high, the total image processing process is often difficult in terms of calculation, and in case of single-chip colour image capturing, e.g. by means of Bayer filter, the system allows only a one-third resolution.

High image quality is not guaranteed by mere use of digital connections. Individual camera designs are very similar to each other – most digital CCD cameras contain a similar integrated camera bus that digitises data from the CCD sensor, performs colour balancing and interpolates colours from the Bayer filter; in case of commercial or security cameras, the image is always compressed to MJPEG format (4 % RAW data flow) or MPEG4 format (1 % RAW data flow). CMOS APS sensors operate on similar base, however, the digitisation and certain pre-processing operations are performed directly by the sensor. The aforementioned compression can be also performed in industrial cameras, while the data flow is then high, does not load the communication tables and the data processing is quite fast. That is very beneficial in case of observing the processes remotely via cameras connected to the Internet or Ethernet, which are not unnecessarily loaded by the image transmission. However, the quality of this operation with regard to limited properties of the integrated image processor has its limits and the resulting image is burdened with inaccuracies and is not sharp when zoomed. That may affect the required accuracy of information obtained from the image analysis that follows. When designing the machine vision system, it is necessary to specify the image quality requirements for the given solution. Clean, stable and precise image is not necessary in all types of applications; however, more complex applications often require high image quality to offer stable, thus successful, operation of the inspection system.

The output video signal format determines the resolution of the image capture, colour space (RGB, YUV, MONO), the amount of bits used to represent a single pixel

(8, 12, 16 and 24 bits/pixel), and in what form are the obtained data (lossy compressed, RAW, ...). Tab. 3.5 lists examples of formats for industrial practice using IIDC 1394-based Digital Camera Specification standard for FireWire by 1394 Trade Association Specification, [23].

Tab. 3.5 Examples of formats for industrial use with IIDC 1394-based Digital Camera Specification standard for FireWire by 1394 Trade Association Specification

Resolution	Mono	RGB	YUV
160 X 120			(4:4:4) 24 bits/pixel
320 X 240			(4:2:2) 16 bits/pixel
640 X 480	(Mono) 8 bits/pixel	24 bits/pixel	(4:1:1) 12 bits/pixel
	(Mono16) 16 bits/pixel		(4:2:2) 16 bits/pixel
800 X 600	(Mono) 8 bits/pixel	24 bits/pixel	(4:2:2) 16 bits/pixel
	(Mono16) 16 bits/pixel		
1024 X 768	(Mono) 8 bits/pixel	24 bits/pixel	(4:2:2) 16 bits/pixel
	(Mono16) 16 bits/pixel		
1280 X 960	(Mono) 8 bits/pixel	24 bits/pixel	(4:2:2) 16 bits/pixel
	(Mono16) 16 bits/pixel		
1600 X 1200	(Mono) 8 bits/pixel	24 bits/pixel	(4:2:2) 16 bits/pixel
	(Mono16) 16 bits/pixel		

Digital cameras often feature other circuits such as microprocessor, image memory, software gate array, etc. These systems offer interesting ways to increase intelligent-related abilities that allows to pre-process, filter, binning, compressing or performing shading-correction etc. of the image data in real time. Microprocessor assumes the camera control and is able to control its parameters by means of commands sent from PC. The parameters are for example the exposure time (sensor irradiation), field of vision, gain, offset, etc. Flexibility of the serial bus allows expanding the flow or image data and stream the pre-processed image data along with the original one, as well as various statistics, histograms, etc.

3.1.9 Input signals for camera control

Nowadays, input signals for camera control are widely used to control some of the camera parameters remotely from the evaluation unit. Input signal can be transmitted via a separate cable, however, the usual solution is to merge the input cable with the cable for output (when using GigE and CameraLink interface). There are several standards for the separate cable, for example RS-485, RS-422, RS-322 and more.

3.1.10 Lens mount

Lens mount is an interface — mechanical and often also electrical — between a camera body and a lens. Lens mount (thread parameter) is important for purchase of the correct lens. Currently, there are several basic types of mounts for industrial cameras:

C-mount – originally designed for 16 mm cameras (C stands for “cine”) has a thread of 1 inch diameter with a pitch of 32 threads per inch. The C-mount is a universal lens thread used with many sensor sizes. This type of mount is used for cameras with area sensor and less expensive line scan cameras. Flange to Focal Distance (FFD) from the end of mounting thread to the image focus (from the active surface of the image sensor) is 17.52 mm. There is also a CS-mount that differs in FFD, which is 12.52 mm. The C-mount can be mounted to camera fitted with CS-mount by means of 5 mm thick spacer. However, CS-mount lens on camera for C-mount lens cannot be focused.

F-mount (Nikon F-mount) – was developed for 35 mm cameras (chips up to 35 mm). It is a bayonet featuring three-lug and a socket with diameter of 44 mm; its FFD is 46.5 mm. Industrial applications use this system for macro lens, line scan cameras and more. Since F-mount was presented in 1959, it is one of the longest used standards for interchangeable lens.

Leica M39 x 1/26” has a specific thread with diameter of 39 mm with 26 Whitworth threads per inch and was developed for 35 mm films; its FFD is 38 mm.

The ***V-mount*** was designed for Schneider Kreuznach lens (V-groove); its FFD is 38 mm. The unique characteristics of this system is that the lens is fitted in the adapter by means of three adjusting screws that are supposed to hold the lens in place in case of vibrations in industrial applications. This system is used with line scan cameras and macro systems.

P-mount (M42, Praktica mount or Pentax mount) uses M42 x 1 mm metric thread and is called a universal lens thread. The mount was used for cameras in 1938 by Carl Zeiss and its FFD is 45.5 mm. There is also a T-mount (T2) variation featured metric thread of M42 × 0.75 mm, its FFD is 55 mm. Both variations are used in cameras and camcorders with chips above 35 mm of size and find its uses in line scan cameras.

M72 x 0,75 mm is a metric thread used by chips with size above 35 mm in line scan cameras and as a connector between filter and lens. There are other types of threads, for example the metric M52 used in line scan cameras and more. The [32] contains good list of mounts.

Mounting of filters to lens is performed by means of many threads, usually metric, e.g. M37 x 0.75 mm, M40,5 x 0.5 mm, M43 x 0.75 mm, M46 x 0.75 mm, M52 x 0.75 mm, M58 x 0.75 mm, M62 x 0.75 mm, M72 x 0.75 mm, M86 x 1 mm, M105 x 2 mm and more.

3.1.11 Camera mounting methods

Method of mounting cameras to supports (tripods), directly to a device or to protective cases is not governed by any standards and differs from manufacturer to manufacturer. Camera cases are often fitted with at least four female threads (metric threads in most cameras on the European market) located on one of the wider sides of the case, often on the other sides. The camera manufacturers are primarily trying to ensure an optimal amount of threads for any mounting operation. Due to the absence of any standard, it is necessary to use mounts provided by the manufacturer or, which is the more often solution, design an own adapter or mounting system.

3.1.12 Power supply

As in previous case, camera power supply is not standardised. Camera connection is performed by means of various types of connectors, but one should always remember to purchase an appropriate power source for the camera as well. Overview of various types of power supply interface is listed in Tab. 3.4.

3.1.13 Resistance to environmental influences

There are standards for resistance to water and dust by means of so called protection of electronic equipment that describes the structural protection from liquid ingress, dangerous contact and foreign objects. Level of protection is expressed by abbr. IP (International Protection) and is standardised pursuant to EN 60529. The letters IP are followed by two numbers (Tab. 3.6 a Tab. 3.7) or even a complementary letter describing the test method. Industrial cameras usually offer IP30 protection, which may not be sufficient for some industrial applications. In such case, a proper cover should be used to prevent the liquid and dust ingress as required.

Tab. 3.6 Protection of electronic equipment, meaning of the first digit following the IP: protection from solid foreign objects

First digit	Level of protection	
	Short description	Definition
IP 0X	No protection	No protection
IP 1X	palm	Protected against solid objects up to $50mm^2$
IP 2X	finger	Protected against solid objects up to $12mm^2$
IP 3X	tool	Protected against solid objects up to $2.5mm^2$
IP 4X	tool	Protected against solid objects up to $1mm^2$
IP 5X	any device	Protected against dust, limited ingress (no harmful deposit)
IP 6X	any device	Totally protected against dust (dust must not disturb the operation of electrical equipment)

Tab. 3.7 Protection of electrical equipment, meaning of the second digit following the IP: protection from liquid ingress

Second digit	Level of protection	
	Short description	Definition
IP X0	No protection	No protection
IP X1	Dripping	Protection against vertically falling drops of water (e.g. condensation)
IP X2	Dripping with a 15° tilt	Protection against direct sprays of water up to 15° from vertical axis
IP X3	Obliquely incident	Protection against direct sprays of water up to 60° from vertical axis
IP X4	Splashing water	Protection against water splashed from all directions - limited ingress permitted
IP X5	Water jets	Protected against low pressure jets of water from all directions- limited ingress permitted
IP X6	Powerful water jets	Protected against powerful jets of water or heavy seas - limited ingress permitted
IP X7	Immersion	Protected against the effect of immersion-between 15cm and 1m for 30 minutes
IP X8	Permanent submersion	Protected against long periods of immersion under pressure - user stated requirement (eventual water ingress must not disrupt the operation of electrical equipment).

Although cameras are heating during their operation, which is why the water condensation does not occur in the common environment, it may occur in humid environment. This is why the camera should be properly protected. Also, cameras often states humidity of operation, for example 20% to 80% (without condensation) [21].

Beside the necessity to protect the camera from dust and water, the system must sometimes be able to withstand exposure to other environmental influences, often vibrations that are directly affecting the imaging quality and reduce the equipment life. This issue is solved either by using tripod and camera mounts featuring anti-vibration elements separating the cameras from the manufacturing machine, or by manufacturing a separated frames and tripods that are not in contact with the manufacturing machine.

Another influence to withstand is the increased or decreased temperature. Cameras are designed for a certain range of temperatures; an operating temperature range (often from 0 to 60°C [25] or -5 to 55°C [21]) and storage and transportation temperature range (often from -20 to 70 °C [25] or from -25 to 60 °C [21]). Lower or higher temperatures may result in irreversible damage to the camera electronic systems. Therefore, it is necessary to protect the cameras from temperatures outside the given operating temperature range. In case of lower temperatures, the camera cases (boxes) are fitted with heating and temperature regulation. When using the camera in high temperature environment, for example glass industry, one shall consider significant increase of cost of the machine vision solution resulting from the necessity to have a cooling system for the camera. Heating of the camera itself during its operation should be considered as well, since the temperatures may reach the maximum operating temperature even when in 20°C environment (cameras are fitted with processor and other electronic systems that produce relatively high amounts of heat all by themselves).

Designing a camera cooling system for the camera requires determining the basic requirements on:

- Effectiveness – determine the required decrease of camera temperature,
- Humidity – determine conditions for preventing water condensation within the camera,
- Energy – optimisation of energy consumption,
- Regulation – determine requirements on temperature regulation within the box according to the ambient temperature,
- Weight – weight optimisation may not be required in all cases; it represents the ratio of weight of the cooling system and its effectiveness.

There are several main approaches to protection of camera from increased temperature, however, the most usual ones are:

1. Filtering the incoming heat radiation by means of an IR filter, thus decreasing the sensor heating. Unfortunately, this solution does not decrease the heat generated by the camera itself. This method is suitable with combination with other solutions.
2. Placing the camera away from the heat source (e.g. glass gob) and use lens with higher focal length (telephoto lens). This approach is vulnerable to

vibrations, even small ones, regardless of whether the vibrations are affecting the camera mount of the observed object, since logically, increasing distance results in increasing effect of such vibrations.

3. Decreasing the camera temperature by air blowing. This method requires either a source of air that is cooler than the temperature fed from close range, or a cooling system able to create a certain volume of air with required temperature.

4. Place the camera in a cooling box. This solution is the most expensive one and requires proper calculation of thermal balance and either active cooling elements inside the camera, or coolant. Peltier thermoelectric module or Heatpipes can be used as coolant elements. Absorption cooling using air or water can be used as well [33].

5. Combination of the aforementioned methods.

Among other indicators to consider when selecting the device is the sensor and camera manufacturer. Since it is a device that must operate for a long time without any malfunctions, it is wise to invest in renowned and/or proven manufacturer of the desired sensor as well as the camera. In case there are special requirements regarding the observed spectrum (UV and NIR), possibilities to modify the sensor or other parts of the camera should be considered as well. It may be necessary to specify additional parameters for the given applications.

3.2 Intelligent cameras

Intelligent, or smart, camera features control, evaluation and communication unit. For example, they allow to program relatively simple PLC-based tasks. Using intelligent camera is a compact solution allowing quick implementation, quire fast parameterisation (programming) and autonomous image processing as well as connection to a control system, Fig. 3.5. Unlike so called camera (video) sensors, the output is not a simple YES/NO information. These intelligent cameras offer extended output information, e.g. position, orientation, direction, speed, dimension, colour and shape of the object. In early days of intelligent cameras, they featured only basic tools for image analysis. However, the features are getting more and more complex. The resolution of the cameras is increasing as well, starting at VGA (0.3 Mpx), or now commonly available UXGA (1.9 Mpx). The state-of-the-art cameras however offer resolution of up to 5 Mpx (and will increase in the future). Intelligent camera manufacturers fit their cameras with integrated ring light (suitable only for absolutely basic image processing tasks) and allow interchanging lens, thus expending the application potential of these devices. These days, most applications in industrial practice can be solved by using these cameras. The advantage of these cameras is not being the least expensive solution, which is comparable with PC-based one, but in their compactness and simple implementation [34].

Fig. 3.5 SICK smart camera

Beside the image sensor itself and the systems for image processing, the cameras must feature other electronics. Intelligent cameras consist of several basic parts: elements for image capture and digitisation, computational part, inputs and outputs, communication interfaces.

Image capture and digitisation is performed by means of these sensors: CCD and CMOS APS – CMOS APS are prioritised due to their speed of reading image data. A computational part of an intelligent camera is basically a microprocessor optimised for working the image. Since image processing requires fast processing of large amount of data, the system must be fitted with high-performance microprocessors. The processor frequency in MHz is an important parameter as well. Mostly for reasons of advertising, some manufacturers state the processing power in millions of instructions per second – MIPS. However, this unit is not much suitable for comparing cameras of various manufacturer, since the processing time of a real task depends rather on the effectiveness of software tool, while each manufacturer and camera type may use different software. Modern intelligent cameras work with complete digital image stored in RAM. Capacity of this memory is selected according to the required complexity of the program that is copied from the program flash memory into RAM upon initiation and is executed there. Therefore, size of RAM and the flash memory (in MB) is an important parameter. In order to speed up some standard operations required for image analysis (filtration, searching for edges, etc.), very fast, single-purpose finite state machines based on gate arrays are sometimes used [34].

Nevertheless, it’s the digital inputs and outputs that makes the intelligent cameras usable for machine vision applications, and it is often necessary to have them, since perform the synchronisation with process state after the image is captured. The observed object must usually be captured in certain position and the camera is triggered for example by the auxiliary proximity sensor or a control system signal. The amount of data obtained by the camera is usually higher than the user requires, which is why the output data is usually reduced to data suitable for further processing.

Communication interface of intelligent cameras has several purposes. The primary purpose is usually the connection to HMI (Human-Machine-Interface) that allows to perform various settings of the camera or program and parametrize. Currently, HMIs are usually standard PCs. In order to facilitate a comfortable development of the given task, it is necessary to have the HMI stream the image captured by the camera in real time. Transmission speed must be quite high, which is why GigE Vision (Ethernet) is the most widely used communication interface. Communication interface may feature additional function. It can transmit data to the master control system, can be fitted with a module to increase the amount of inputs and outputs, can perform communication between several cameras in more complex tasks requiring coordinating more than one camera. The interface is often used for maintenance, camera firmware update and similar purposes. Beside the usual GigE interface, the intelligent cameras are often fitted with auxiliary RS-232/422/485 serial interface. The reason is that the interface is easily connectible to most PLCs.

Intelligent cameras are similar to industrial ones. However, in general, they are larger and offer more connection spots; additionally, the communication interface features inputs and outputs as well. Another frequent example is the lens placement which is not on the camera front, but its side, complemented by a ring light.

Intelligent cameras features own software, which is given by their tight connection with their subsystems – microprocessor, limited memory, connection methods of sensor chip, inputs and outputs. However, this segment of cameras is not yet governed by standards and manufacturers are not interested in portability of their software to other devices (the software is often provided for free with the camera). Software of intelligent cameras is currently able to perform the following tasks (according to [34]):

- Image processing: detection of edges and transformation to vector format (vectorization), searching for contrast objects, brightness analysis.
- Measurement within the image: determining distances in vector image, or calibration in engineering units.
- Reading text (OCR – optical character recognition) and identification codes (bar-codes and matrices),
- Operation of camera hardware: sensor chip control, handling of input, output and communication interface.
- Data processing required in more complex applications using for example mathematical or statistical procedures.
- Control of evaluation procedures using an algorithm: a possibility to create a user program.

Manufacturer often provides their programming application along with the intelligent camera, offering a user-friendly graphic interface running on the connected PC. The interface cooperates with the camera firmware and in common and simpler tasks, the user merely parametrize the camera. The provided software often allows programming other tasks that require knowledge of the given development environment (for example C programming language). Another case, although less usual, is an intelligent camera delivered only with its operating system (special OS or a Linux-based version). Tools are provided as library functions for general programming language, usually C. Using these functions, the user can create a program and compile it for the specific processor used by the intelligent camera and implements the program into the camera memory.

This type of camera offers a great potential. It can be used in situations that used to require (with regard to the necessary performance) classic interconnection of the camera and PC or a compact evaluation unit with a control system (Compact vision). It can also be used as a substitution of several sensors by means of a single camera, creating a simpler and less expensive solution. Intelligent cameras are used in industry, medicine, security, transportation, entertainment and many other industries.

3.3 Camera sensors

There is no definition separating camera sensors (also called video sensors) and intelligent cameras. The boundaries are unknown even among the manufacturer, therefore, only the basic properties of this type of cameras can be defined. Camera sensors are basically simplified intelligent cameras in terms of software and hardware. Camera sensors offer a limited amount of tasks they can solve, while the output is usually a YES/NO or RIGHT/WRONG information. However, this simplification makes the control simple as well. Manufacturers are often boasting with the minimum necessity of knowledge of installation and parametrization of the sensor. The extent of required knowledge is not, in fact, too high, however, basic understanding of image analysis and camera hardware is important.

A typical example of camera sensor is compact, simple to control sensor with VGA resolution and integrated lighting as well as basic tools for image analysis allowing inspection of dimensions, position, colour scale and comparison of samples. The tools are based on principle of storing reference image or size combined with the feature of calculating and deciding about right/wrong.

There are also camera sensors with small integrated screen (often touch screen), which displays a real-time scene and allows setting parameters without the necessity to use PC. Another option is to connect an intuitive console that serves as a built-in screen and allows setting parameters of the camera sensor.

Camera sensors can be used for example for label detection, presence of parts and their orientation, detection of data and manufacturing codes, inspection caps, inspection of moulded products, inspection of packaging completeness, inspection of gaps, etc.

3.4 Line scan cameras

Line scan cameras (also linear cameras or simply line cameras, Fig. 3.4) are used to inspect continuously moving objects on production lines. As described in Chapters 2.2 and 2.4, line cameras features one to three lines (4 or 5 in exceptional cases). The resulting 2D image is obtained by moving the camera or the scanned object. Office scanners are working on a similar principle. A typical example of this application is a conveyor with the product that is transversely scanned by the line scan camera (Fig. 3.6). Based on synchronization of the camera with conveyor speed, the process creates a 2D image from assembling individual lines.

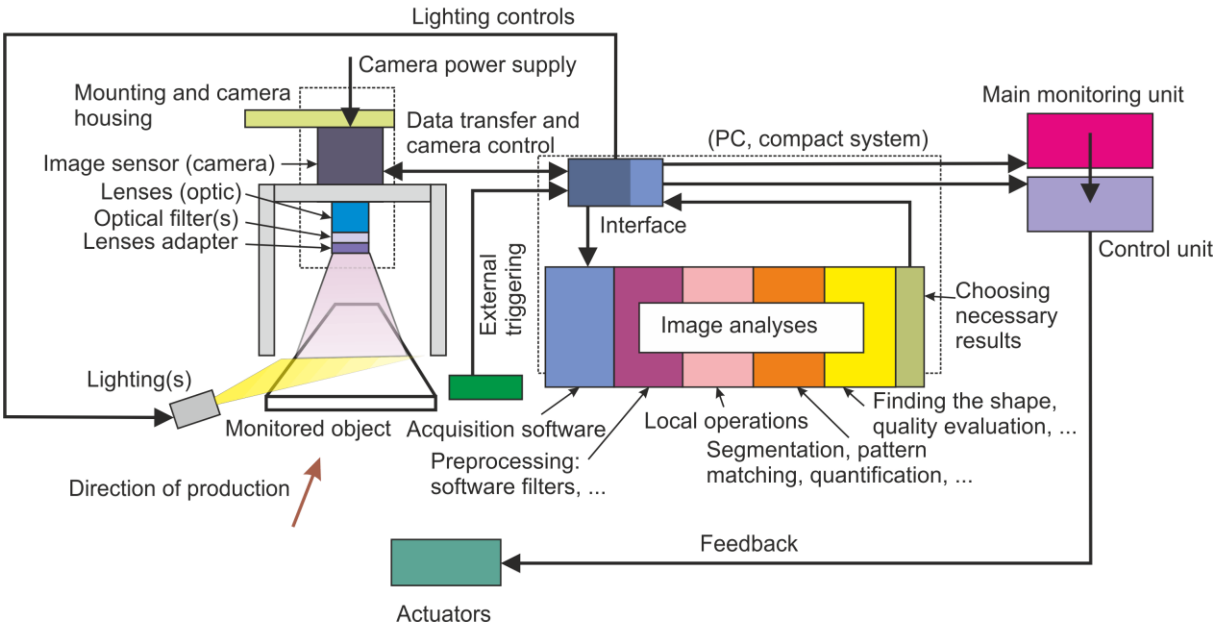

Fig. 3.6 Diagram of a system with line scan camera

Line sensors offer high Fill Factor, since whole area of the pixel is filled by a photosensitive cell, while the electronic parts are next to it, Fig. 3.7. Which is why line scan cameras features two times higher sensitivity than area sensors. These sensors are able to capture technological scenes with very short exposure times and high shutter speed (frame rates are between 4 and 140 kHz). Resolution of a single line ranges from 512 pixels to more than 16 000 pixels [25]. Therefore, line scan cameras offer resolution parameters that are not offered by any other available area sensor camera. Horizontal resolution is given by the physical resolution of the camera (amount of pixels) or the set area of interest. Vertical resolution depends of the captured amount of lines per second and movement speed of the object or camera. Using intensive linear lighting focusing high luminous flux to a thin line is sufficient for lighting the observed spot.

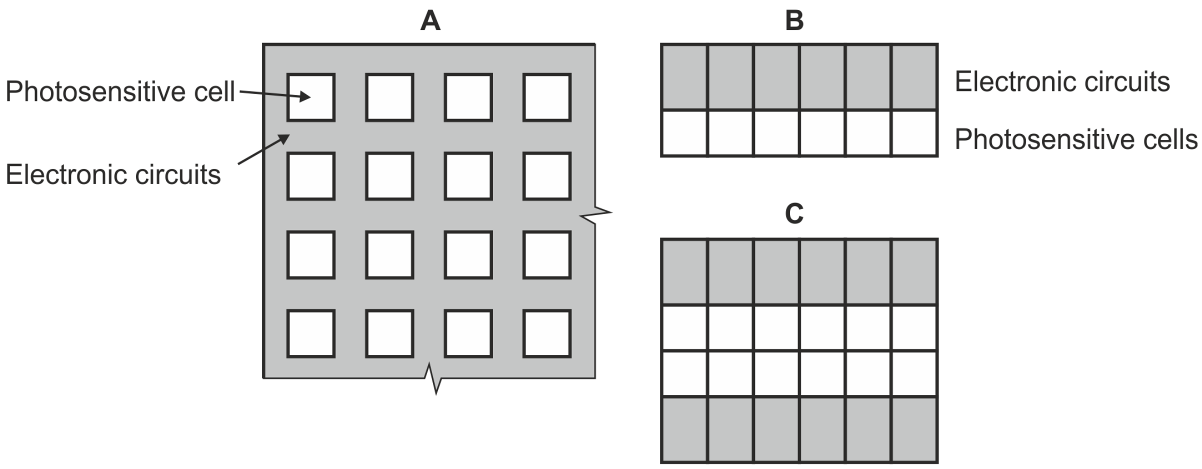

Fig. 3.7 Configuration of area and line sensor, C – dual-line sensor

The sensors can use CCD as well as CMOS APS and may scan in both, colour and monochromatic scale. Colour scale sensors usually use Bayer filter (Fig. 2.8 A), three sensors placed above each other (Fig. 2.8 B) or decomposition of light by means of an optical prisms to three sensors (Fig. 2.8 D). The difference between the methods lies in resolution (higher in three-chip configuration without interpolation), precision of synchronization with the captured process (necessary in three-chip configuration above each other), price (three-chip configurations are more expensive) and more (Chapter 2.4).

Using a line scan camera is logically limited to applications with constant movement speed of the scanned object as well as sufficient possibilities of synchronization with such speed. The synchronization (triggering is usually performed by the following methods):

- Fixed frame rate not linked to object movement. After the frame is captured, the image data are sent to a control unit. This method can be used only with constant movement speed of objects,

- Variable frame rate linked to movement of objects, where cameras allow connection of sensor (trigger) that triggers the camera via sent impulses that changes with movement speed of the object, conveyor speed can be variable (more usual case).

The camera submits data via communication interface to the control unit (computer) when a determined amount of lines (frames) is captured. The resulting image is then process similarly to area sensor. Some cameras also allow using another digital input to determine start of the frame, i.e. a whole image (for example starting upon present of product on the conveyor).

Standard line scan cameras use high-speed communication interfaces, mostly GigE, Camera Link or CoaXPress, which is now on the rise. Camera Link in Full configuration is usually used in high-frequency systems (total bandwidth of 680 MB/s).

Line scan cameras often require special lens with as linear MTF (modular transfer function) characteristics as possible and small relative exposure change, Chapter. 4.3. Purpose of these measures is to minimize the image dimming towards the border of field of vision.

The enhanced variant is a dual-line monochromatic camera. In this case, the sensor consists of two lines fit closely to each other, Fig. 3.7. That allows increasing frame rate up to twice (e.g. From 70 kHz to 140 kHz), while keeping the accuracy and sensitivity as the one-line camera. The camera captures the image alternately via the two lines with frequency of 70 kHz each, thus reaching double speed. Camera also allows increasing sensitivity by decreasing area resolution, which is performed by means of so called pixel binning, Fig. 2.9. This method connects two adjacent pixels (vertical pixel binning) or 4 adjacent pixels (horizontal and vertical pixel binning), Chapter 2.5.

Line scan cameras can be divided as shown in Tab. 3.8. These cameras are usable in many branches of industry: for example in inspection of endless conveyor (woodworking, paper, foil, textiles, etc.), capturing of textures, finishes and varnishes, inspection of surfaces (installation of printed circuit boards, manufacturing flat panel and displays, semiconductors, etc.), detection of steel rope burrs and textile fibre burrs, document scanning and letter sorting at post offices, food inspection, checking the freshness of fruit, legumes and other food, etc. [35]. In glass industry, these cameras can be used to observe continuous production of flat glass, pipes, and in theory, to observe endless fibres, and in other branches. On the other hand, line scan cameras cannot be used in cases, where the scanned objects move in manner other than continuous movement perpendicular to the sensor, for example rolling object.

Tab. 3.8 Overview of common use of line scan cameras

Monochromatic camera	One-line
	Dual-line
Colour camera	One-line camera with Bayer filter
	Dual-line camera with Bayer filter
	Triple-line camera
	Triple chip (3CCD/3CMOS APS)
	Quadra-line – three colour components (RGB) are fitted with additional possibility to scan in Near Infra-red Radiation (NIR).

4 Objective lens

Objective lens is an integral part of common industrial application, and a proper attention should be paid to its selection. Their purpose is to project light to the sensor. Objective lens imitate the perspective view (also projective and centreline); expressed by a relation (1,1), using the pin hole camera model, Fig. 1.5.

Objective lens (Fig. 4.1) may consists of a single lens or the whole system and focuses the image to the sensor, and by means of lens ring, this image can be focused. The objective lenses are usually fitted with a mechanical aperture allowing to regulate the amount of light passing through the lens. All parts of this system are governed by the laws of optics [36].

The main parameter is focal length that determines a field of view for the given chip size. Most objectives for industrial applications have a fixed focal length. Objective lenses can be divided to three basic groups:

- **Common objective lens** – field of view is approximately 50°, which is similar to field of view of the human eye; images captured by means of this lens offer a perspective that is most natural to human.
- **Wide angle objective lens** – focal length is shorter then in common lens, therefore the image is wider; "fisheye" lens are an extreme case and their field of vision is commonly up to 180°, sometimes even more.
- **Telephoto objective lens** – their angular field of view is narrower (focal length is higher) allows filling the whole image with an object that is quite distant.

There are other special objective lenses:

- Macro objective lens – objective lens used for macro photography; is able to focus on distances shorter than common objective lenses, thus allowing display of up to 1:1 ration; offers optimised optical properties (distortion, MTFs).
- Telecentric objective lens – objective lens able to eliminate perspective; is used in environments requiring precise measurement of dimensions and shape.

- Mirror objective lens – a telephoto objective lens that uses a tilted mirror instead of classic lens (design is similar to modern astronomical telescopes); mirror objective lens are used rather scarcely, since they are usually not fitted with interchangeable aperture and creates an unusual bokeh[7] (these objective lenses are not commonly used in industrial practice).
- Tilt-shift objective lenses – objective lenses allowing shifting or tilting the optical axis, thus corrective the perspective distortion; are used in architecture photography (these lenses are not commonly used in industrial practice).

Fig. 4.1 Illustration of objective lenses, C-mount thread, focal lengths: 12,5; 50; 35; 4,8; 75; 35; 25 (from the left)

[7] Expression Bokeh (from *boke*, “blur”) is a term for aesthetic quality of parts of image located outside the plane of sharpness. Mostly, it is a case of defocused background with unsharp (blurry) edges. In case of mirror objective lens, the unsharp areas of an image are displayed in an annulus shape.

4.1 Focal length

Definition of focal length: focal length of object (image) space is a ratio of linear size of the image (object) within the focal plane and a perceived size of the object (image) with infinite distance (definition by Carl Friedrich Gauss). That says that in general, there may be two different values of focal length for the given optical system, the first one in the space where the captured object is located (in front of the system), while the second one is in space where the image is created (behind the system). The |FO| distance (Fog. 4.2) is called *Front focal length* (FFL) f, while $|OF^I|$ distance is *Back focal length* (BFL) f^I. If there is the same environment behind and in front of the thin lens, the following is applicable: $f = f^I$. In such case, we use lens focal length f.

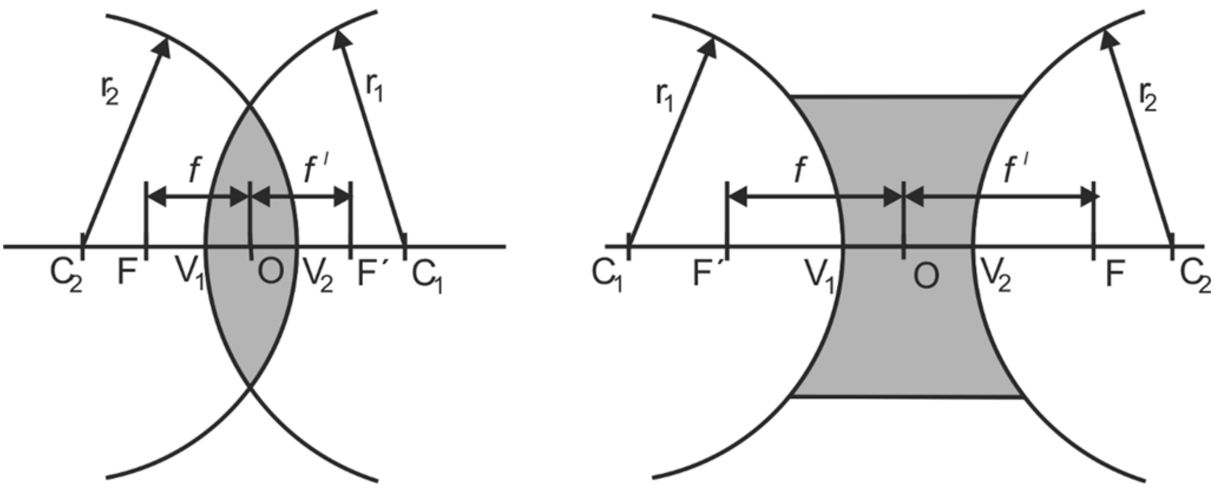

Fig. 4.2 Parameters of negative lens and positive lens

Focal length depends on the n_2 refractive index of glass, from which the lens is made, on n_1 refractive index of the environment and on r_1 and r_2 radii of curvature of optical areas according to:

$$\frac{1}{f} = \left(\frac{n_2}{n_1} - 1\right)\left(\frac{1}{r_1} + \frac{1}{r_2}\right). \qquad (4.1)$$

According to sign of the focal length, we distinguish positive (convex) lens lenses $f > 0^-$, focal lengths are real and negative (concave) lens $f < 0^+$, where the focal lengths are “unreal”. Reciprocal of focal length is called optical power and is measured in dioptres.

Ratio of sensor size and focal length determines the angular field of view of the captured scene area, Fig. 4.3. Based on calculation of right-angled triangle $\frac{\alpha}{2}$ angle, whose opposite leg is $\frac{u_S}{2}$ and adjacent leg is f, an angular field of view α can be derived:

$$\alpha = 2.arctg\left(\frac{u_S}{2f}\right), \qquad (4.2)$$

Where the considered sensor size is u_S and f is the focal length. When changing the sensor size, the angular field of view changes as well if the focal length u_S of the given optical system remains unchanged. Size of the sensor can be measured diagonally (by the sensor diagonal line), horizontally (by the longer side of the sensor) or vertically (by the shorter side of the sensor), therefore, the angular field of view will correspond to the given type of dimension and will vary.

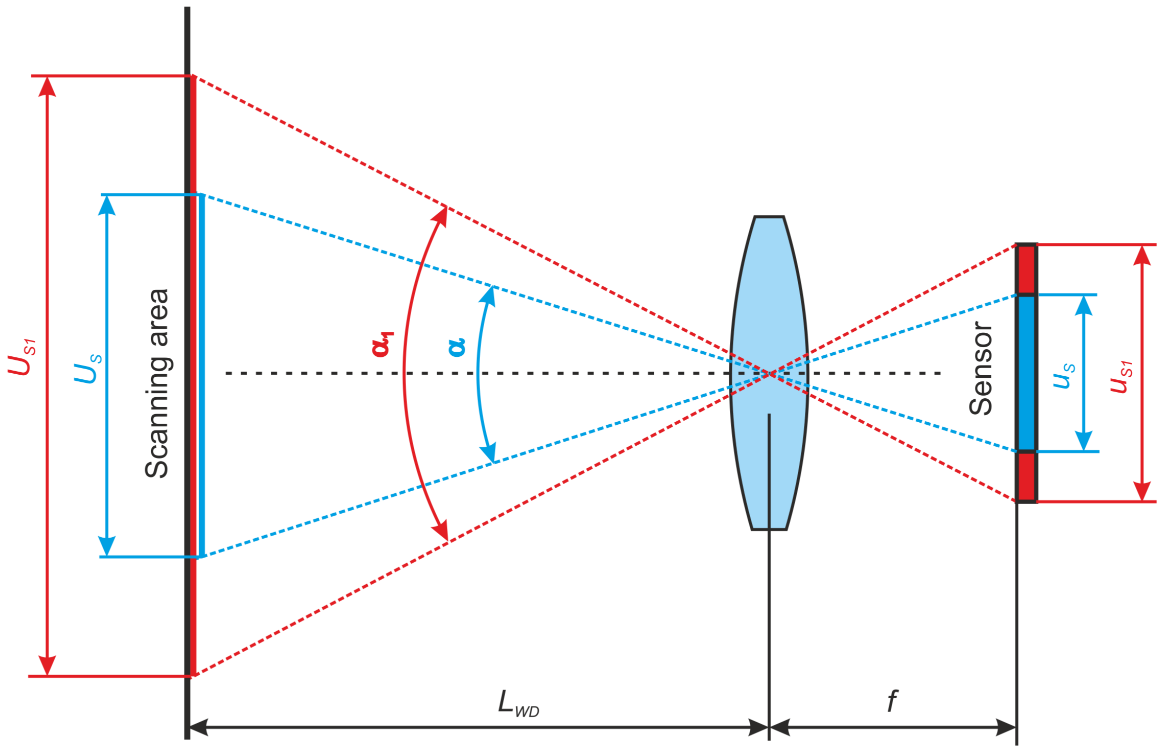

Fig. 4.3 Relation between angular field of view, focal length and sensor size

The aforementioned imply that an objective lens must be selected with regard to the camera type and given sensor size in order to have a proper angular field of view, allowing capturing the whole scanned scene. The selection of scene size is given by the maximum assumed size of the captured object (without the necessity to change objective lens). However, the resulting scene size should be larger due to decrease of resolution and contrast of the objective lens toward the edges.

Additionally, it is necessary to respect the shape of the area of interest in relation to the sensor shape. If the area of interest is of circular or square shape, then the shorter side of the sensor must be regarded. Having a rectangular area makes the situation more complicated, since it is requires considering the ratio between the shorter and the longer side of the area of interest and ration of sides for the sensor. If the ration between the shorter and longer side of the area of interest is lower than the ratio between the shorter and longer side of the sensor, the longer side of the sensor must be regarded. Vice versa, when the ratio of the area of interest is higher than the sensor ration, it is necessary to regard the shorter side of the sensor.

Sensor size is provided by the camera manufacturer sometimes in mm, but more often in a form of diagonal line size given in inches. The size however, is not directly corresponding to the sensor diagonal line (Chapter 3.1). It is necessary to use a table listing various sizes of optical sensors (the common ones are listed in Tab. 3.1).

With regard to geometry and relation between the observed object and displayed object, the calculation of objective lens with suitable focal length can be performed using one of multiple ways. For example, the basic magnification M_{PMGA} is defined by the following formula:

$$M_{PMGA} = \frac{u_S}{U_S} \quad , \qquad (4.3)$$

where u is the selected dimension of the sensor in mm and U_S is the selected dimension of the area of interest in mm (Fig. 4.3). Focal length of the objective lens f in mm is given by:

$$f = L_{WD} . M_{PMGA}, \qquad (4.4)$$

where L_{WD} is the working distance in mm. Equations (4.3 and 4.4) are based on:

$$\frac{u_S}{U_S} = \frac{f}{L_{WD}}, \qquad (4.5)$$

which is apparent from Fig. 4.3.

Example:

The area of interest has diameter D = 60 mm, while the distance from lens (working distance) L_{WD}= 230 mm. Camera is fitted with 1/2" sensor.

Calculation:

Tab. 3.1 shows that sensor size is 6.4 x 4.8 mm. It is necessary to consider the shorter side of sensor – *u_S=4.8 mm. Area of interest is defined by diameter U_S=60 mm.*

The basic magnification of the objective lens M_{PMGA} is:

$$M_{PMGA} = \frac{u_S}{U_S} = \frac{4{,}8}{60} = 0{,}08. \qquad (4.6)$$

Focal length is given by:

$$f = L_{WD}.M_{PMGA} = 230 * 0.08 = 18.4\ mm \quad (4.7)$$

In many cases, the initial focal length is different than the market offers. In such case, it is necessary to round the length down and select the nearest available objective lens with lower focal length than the calculated one, in this case for example with focal length of 18 mm or lower. Usually, the real field of vision is recalculated for the specific objective lens.

4.2 Vignetting

Vignetting is an optical flaw resulting from the laws of optics. The result is a reduction of an image brightness or saturation from the centre to edges. In optics, vignetting can be divided to three types.

The first one is ***natural vignetting***, mentioned in Chapter 1.3.1. Natural vignetting is caused by laws of optics applicable for the camera lens system. The falloff is approximated by the $\cos^4$ or "cosine fourth" law of illumination falloff (equation 1.3). A maximum amount of light is incident on the sensor in perpendicular direction. If the light is incidence is tilted, its effectiveness is reduced. This phenomenon is manifested by gradual dimming towards the edges. This vignetting is more visible with wide angle objective lens than with telephoto objective lenses. Since this is a systematic optical flow, it is possible to suppress it.

There is also ***optical vignetting*** manifested with objective lens system consisting of multiple lenses. Each lens has its thickness causing that the rear lenses are shadowed by the front ones. That results in decreased effective aperture of the objective lens, thus in decreased intensity of light incident under a tilted angle. Vignetting increases with rising aperture value. Natural and optical vignetting are the properties of objective lens.

Additionally, vignetting may be caused by the objective lens ring, attachments or filter; this vignetting is called ***mechanical vignetting***.

In general, vignetting is affected by objective lens structure, its length, amount of lenses (more lenses means stronger vignetting), curvature of lenses, etc. However, the method of micro-lens usage in chips themselves with the purpose to increase the fill factor are increasing the vignetting effect as well. Vignetting can be partially eliminated by increasing the aperture value, by optical elements, by shifting the optical system further away from the sensor, or it can be removed by means of software filters prior to image analysis (via PC software or directly in the camera).

4.3 Selecting objective lens

Tab. 4.1 lists terms taken from [31, 37] and the reader may find them in image capture operations.

Tab. 4.1 Basic values used in image capturing by camera

Term	Abbreviation
Primary Magnification	PMAG
Field of View	FOV
Angular Field of View	AFOV
Object Space Resolution	
Image Space Resolution	
Camera Resolution	
Focal Length	
Working Distance	WD
Depth of Field	DOF

The objective lens is selected according to the calculated focal length as well as the task for which the lens is being selected. High quality objective lenses are necessary in applications for precise measurement of dimensions, display of small details, inspection of print, etc. These applications require much more precise display of the observed object for the purpose of further processing than offered by common digital images. Quality of objective lens is based on these basic parameters:

- Lens speed.
- Geometrical distortion of the lens.
- Rate of decline in contrast from the centre to the edges of the field of view.
- Chromatic aberration (in colour cameras only, the aberration of monochromatic cameras is manifested by image blur).

Lens speed refers to how much light will the objective lens permeate into the camera if the aperture is fully open. It is labelled f/X, where X is the lowest possible aperture number – f-stop (represented the maximum aperture diameter) as well as lens speed. The lower the value, the higher the lens speed, which provides wider scanning possibilities in deteriorated light conditions. However, objective lenses designed for high lens speed do not offer good MTF characteristics. Industrial practice applications use objective lenses with lens speed ranging from f/0.95 to f/2.8.

Significant limitations of image measurement accuracy may cause geometrical distortion of objective lens image area. In real conditions, the conversion of angle (or distance) of the displayed object from the optical axis and distance of this object within the image area is not entirely linear. Angle-to-distance conversion is a quadratic, often a cubic polynomial. Even high quality lens used on cameras offering images with resolution of several megapixels, offer radial distortion in units and tens of pixels, Fig. 4.4. The fault is manifested by barrel distortion, pincushion distortion or their combination – mustache distortion. This fault often manifests in wide angle objective lenses. The distortion is not so relevant in some applications such as reading of texts and codes or counting of parts. However, it becomes a fundamental criterion in applications that require accurate measurement of dimensions. The problem can be solved by well performed calibration of the camera using software equipment of the control unit rather than buying an expensive, high quality objective lens [38]. However, calibration makes the measurement linear, it does not increase its accuracy.

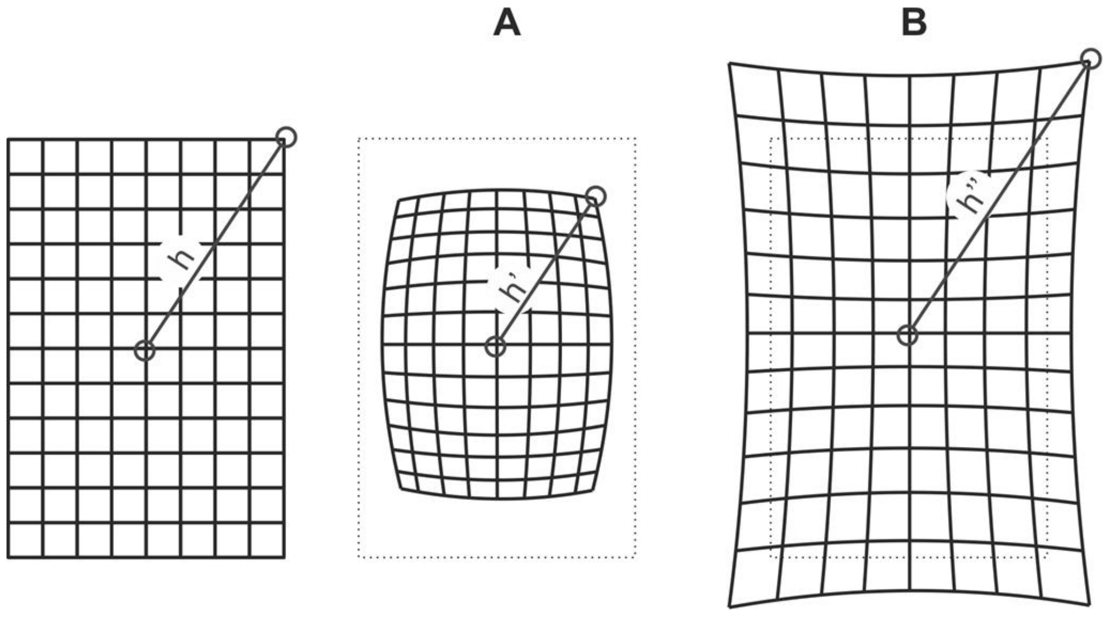

Fig. 4.4 Basic types of radial distortion objective lens, A – barrel distortion, B – pincushion distortion

Line scane cameras often require objective lenses with high quality and as linear MTF characteristics (modular transfer function – directly related to resolution and image contrast, Fig. 4.5) and relative exposure as possible. Due to their physical principles, the resolution and relative exposure decreases in common objective lenses, causing image dimming towards the edges of the field of view. This problem is often insignificant in area sensor cameras, since the edges are usually not very important during the image processing. However, line scan cameras usually use the whole field of view.

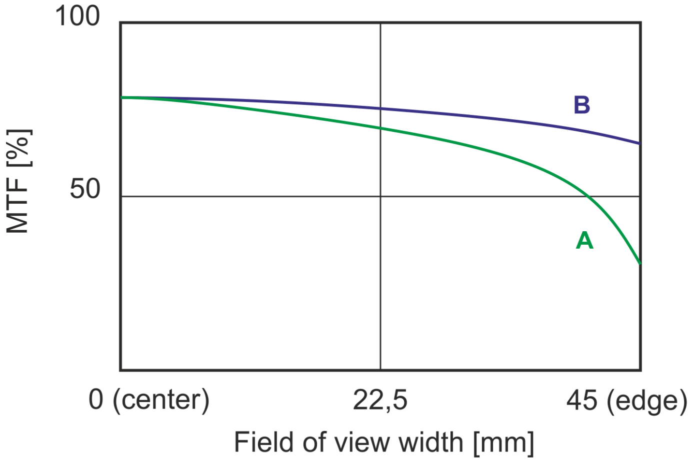

Fig. 4.5 The MTF (modular transfer function) characteristics, A – common objective lenses, B – example of special high quality objective lenses; accurate MTF curve also depends on working distance and aperture value; the parameters are distinguished separately for radial and tangential direction

When selecting objective lens, the chromatic aberration (colour aberration) must be considered as well. Aberration is a result of different refractive index of individual colour components of the given spectrum. The aberration is most manifested on sharp contrast edges that will be displayed as defocused and in a form of colour spectrum. The largest difference in the refractive index is between colours from the opposite end of spectrum (red-violet). Those colours are the image edges. Wave lengths for which the objective lens is not corrected pose a risk. The chromatic aberration is corrected (especially with telephoto lenses, where the beams are projected under a low angle) by means of optical elements manufactures from special glass and combination of their various types with different refractive index.

Lens mount of the given camera must be respected as well as objective lens type, its quality and focal length. Individual types of lens mounts (threads) are described in Chapter. 3.1.10. Although in some cases, you may use adapters enabling using different type of lens mount, it is not recommended to use them. The reason is that the lenses within the objective are of different diameter than in the given threaded tube, shift of lens distance from the sensor, and much more. Therefore, using the adapter may cause complications that are best eliminated by using the same lens mount as the camera.

Common types of objective lenses project the image to the area with so called perspective projection. This must be considered when designing the image processing system. In principle, using line scan cameras and area sensors means that the observed 3D object is transformed to a 2D object, Chapter. 1.2. That is why the properties of projective display of a three-dimensional scene into a two-dimensional one must be considered. In this case, field of view of the objective lens has a shape of truncated cone. Rectangular area of the image sensor reduces the cone to a pyramid, whose vertex is called the focal point of projection. When transforming the scene image in the pyramid into an image plane, loss of significant amount of information occurs. Each half-line going through the focal point of projection is represented by a single point in the image plane. When scanning surface under a certain angle, the perspective is formed by a single point, scanning 3D objects uses two-point or three-point perspective.

Loss of space information during perspective projection for example complicates accurate measurement of 3D objects. Without any prior knowledge regarding shape of captured objects, it is not possible to sufficiently correct these errors. Even when we know the object shapes, correcting projection errors requires identifying these objects by means of machine vision software equipment. Therefore, a high level of image understanding is required. The calculation of real distances then requires using appropriate methods of “perspective calibration”.

Beside the standard objective lenses with perspective projection, there are also special objective lenses with orthographic projection, so called telecentric objective lens. These lenses do not display the scene in focal point perspective, but as a perpendicular parallel projection. This means that the size of displayed objects is always the same, regardless of their distance. The principle is based on using an aperture stop at the plane of the image main point (i.e. lens’s focal point) that will shield all other beams coming from directions other than parallel to optical axis, Fig. 4.6. This solution is suitable, but very expensive. This type of projection requires the size of the objective lens input area to be identical with the area of the scanned scene; therefore, these objective lenses are very large. Further shift of the aperture towards the sensor creates a hypercentric objective lens that “flips” the perspective.

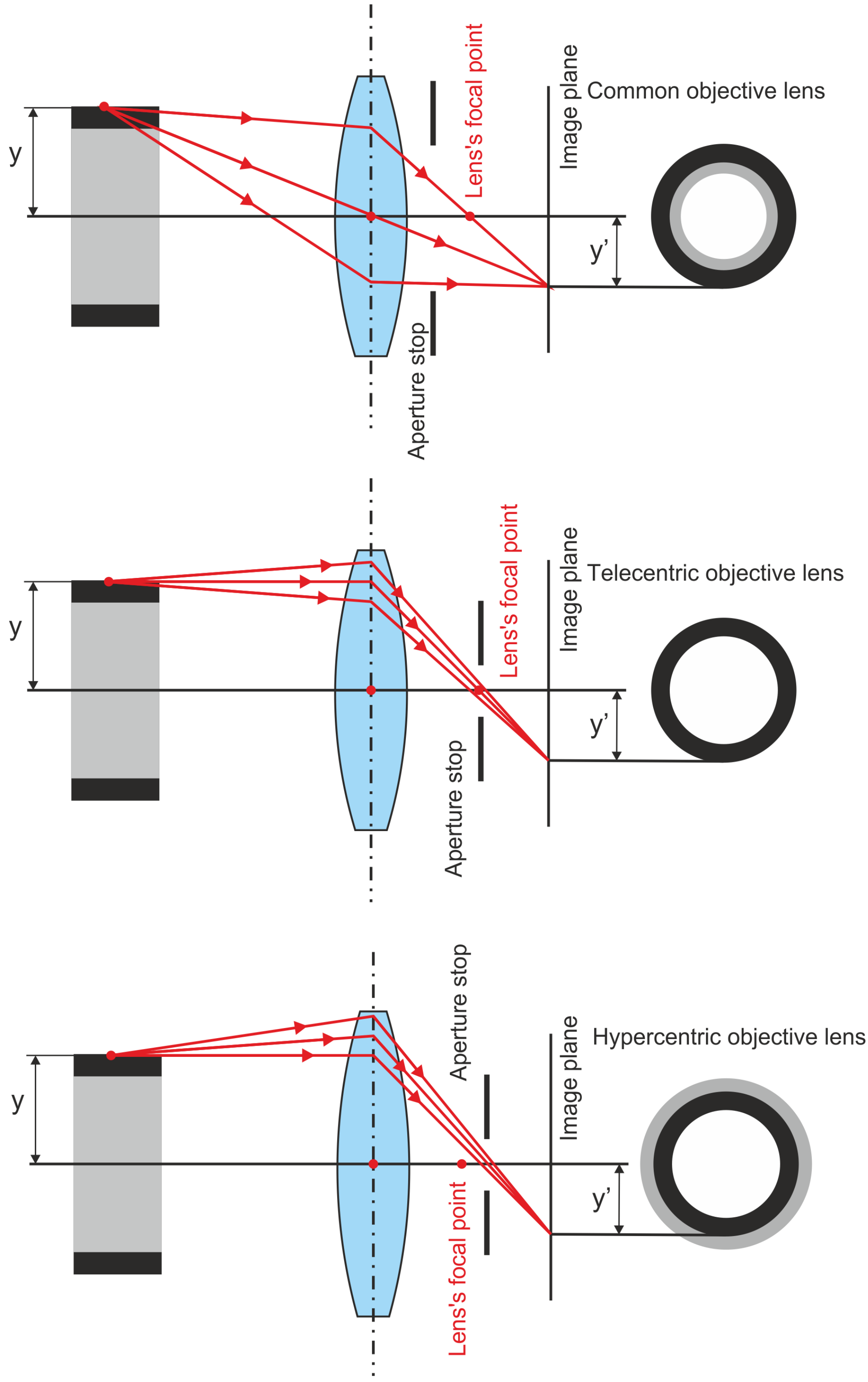

Fig. 4.6 Difference in scanning by means of normal, telecentric and hypercentric objective lens

5 Exposure – basic parameters in sensing

In order to obtain the required image, the camera must be properly set to allow obtaining the required information about the observed object. Beside the correct focus, scene illumination, selection of appropriate objective lens, sensor and other factors such as exposing the sensor to the required amount of light and perform appropriate amplification. That is determined by exposure that depends on three main factors: duration of exposure (exposure time, shutter speed), aperture and gain (contrast, called ISO in commercial cameras).

Sutter sped (exposure time) determines for how long the shutter is open, resulting in exposure of sensor to light coming from the scene, Chapter. 2. Duration of exposure can be set to values rounded to squares of 2:

$$T_E \approx 2^n \approx \frac{1}{1000}, \frac{1}{500}, \frac{1}{250}, \frac{1}{125}, \frac{1}{60}, \frac{1}{30}, \frac{1}{15}, \frac{1}{8}, \frac{1}{4}, \frac{1}{2}, 1,2,4,8,15,30,60,\ldots [s] \,, \qquad (5.1)$$

where n is a positive integer exposure parameter. Sometimes, we also use intermediate values to allow setting more combinations. The value of industrial cameras may be given in milliseconds.

Aperture is a diameter of a hole that permeates light into the camera, therefore it is a physical limitation of light coming to a sensor. Aperture is often a part of objective lens (Chapter 4). We often use f-number N and is given by:

$$N = \frac{f}{C} \,, \qquad (5.2)$$

where f is the focal length [mm] and C is the diameter of aperture hole [mm]. Clearly, the f-number depends on the diameter of the whole and the focal length . That also affects the amount of incident light coming through the camera lenses, therefore it affects the exposure. By doubling the aperture – the f-number N, the exposure (amount of light incident on the sensor) rises four times, since the double diameter of aperture C results in double area S, through which the light is coming ($S = \pi . \frac{C^2}{4}$). In order to double the exposure, the aperture diameter must be increased by $\sqrt{2}$ (≈ 1.4). The focal length affects the decrease of the amount of light incident on the sensor as well. Doubling the distance decreases the amount of incident light four times. That is due to the fact that light is then distributed to larger area. Aperture numbers (f-numbers) are often in series respecting the aforementioned dependency, changing by one doubles the exposure (amount of light):

$$N = \sqrt{2^n} = 1; 1.4; 2; 2{,}8; 4; 5.6; 8; 11; 16; 22; \dots, \qquad (5.3)$$

where n is a positive integer number. (The numbers are often named f-stops and labelled: f/1.0, f/5.6, f/8, etc.)

Amplification of the sensor signal – ***gain*** directly affects the noise of sensor. In film cameras, the gain is called ISO sensitivity. In practice, there is no point for creating standards for camera gains, since sensitivity of sensors to light varies. Standardization is performed in common and industrial cameras, where the total sensor sensitivity is standardized as well. The sensitivity is usually given in units of ISO and should correspond to sensitivity of classic cine-film. Each adjacent ISO value changes the sensitivity two times. Therefore, the typical ISO scale is:

$$ISO = \dots 50; 100; 200; 400; 800; 1600; \dots \qquad (5.4)$$

If the ISO scale gain increases twice (e.g. from ISO 100 to ISO 200), reaching the same exposure required only half the light-induced energy (very simply speaking, half the incident photons is required).

Exposure can be created by means of three aforementioned parameters. The same exposure can be achieved by combining all three factors, however, the obtained image is not the same. Scene, that changes quickly, must be captured in a short time. That is achieved by opening the aperture (setting small f-number) or increases the gain of sensor output. Opening the aperture decreases the depth of field. Parts of the scene that more and less distant from the object in focus are blurry, which may be undesirable in some tasks of the industrial practice. Another problem of objective lenses with lesser quality is the deterioration of quality in edge parts of the image. The other way, to boost the output signal, increases the relative noise rate of the obtained image.

Proper settings of aperture, gain and duration is performed by an exposure meter (also called light meter), which also measures the *EV* (Exposure Value), a combination of duration and aperture in absolute values. *EV* is calculated for the given combination of shutter speed and f-number (aperture) by means of the following formula:

$$EV = \log(N^2/T_E) / \log 2, \qquad (5.5)$$

where N is the f-number and T_E is the shutter speed - duration of exposure (applicable for ISO 100). Combination of shutter speed and f-number is equivalent since the amount of light incident onto a chip of digital camera remains the same and have the same *EV* value.

E.g.:
shutter speed 1/60s with f-number 16
shutter speed 1/250s with f-number 8
shutter speed 1/1000s with f-number 4

EV of all these combinations is 14. That means that the sample can be captured under the same light condition in shorter time but lower depth of field, which affects the proposed system of automatic inspection. EV 0 on this absolute scale means that the shutter speed is 1s with f-number of 1.0 (f-stop f/1.0) and ISO 100.

EV can be used to express light intensity E_f in luxes, since with given chip sensitivity (equivalent to sensitivity of a common cine-film), a certain level of lighting corresponds to certain proper exposure. Since EV 0 represents an exposure level of 2.69 lux, the formula for conversion to luxes is

$$E_f\ [lux] = 2.69 \ .\ 2^{EV}. \ (5.6)$$

Formula applies to ISO 100. If we measure a film with ISO sensitivity, *N* f-number and *t* shutter speed (in seconds), the lighting is

$$E_f\ [lux] = 2{,}69\ (100/ISO)\ N^2/t \qquad (5.7)$$

This value can be used to set the lighting.

Exposure time (shutter speed) must be selected with regard to light conditions, required depth of field and the type of light. In case a cheap type of fluorescence tube lighting is used, the duration of exposure must be longer than the tube’s refresh rate that is usually 1/50 second. Exposure times can vary in three-chip configuration for each colour component, allowing performance of colour balance without using gain that may generate noise.

Sometimes, industrial applications use automatic exposure that is based on calculation of EV values and settings of gain, shutter and aperture of the objective lens. This application is suitable mainly for varying light conditions.

Ways of exposure control are as follows:

- **mechanically,** using the aperture as in camera objective lenses (change of f-number). To perform automatic setting, the objective lens of the camera must be fitted with a mechanical aperture and servomechanism that will control it; delay of aperture's reaction to the new settings must be considered when working with the camera;
- **the amount of light** incident on the captured scene upon exposure – if there is not enough light, the scene is lighted by an additional external light source with higher intensity than the other sources, while the lighting period controls the shutter speed; by using LED, it is possible to generate pulses shorter than 1 ms and capture fast moving objects without having the resulting image blurry;
- **electronic shutter**, where the shutter speed is governed by changing timing of CCD or CMOS substrate hours; the shutter speed with electronic shutter ranges from micro-seconds to tens of seconds;
- **changing the gain (ISO) in camera,** gain can be set by the software provided with the camera or by some types of software such as LabView, Matlab, etc.;
- or combining the mentioned methods.

6 Basic principles of lighting and light types

Although the issues of lighting were not presented until now, they are one of the first elements in the image processing chain. It is one of the key parts of the system that must be selected properly. These issues are very difficult to solve, since there are no simple rules and lighting used for specific type of technological scene and the observed object may not be suitable for other.

Logically, the radiation incident on the sensor is important for the following image display and processing. However, the settings and possibilities of obtaining a high quality image of means of the sensor together with the optical system have their limits (Chapter 5). The amount of radiation incident on the sensor can be quite effectively be affected by lighting the object. Appropriate lighting conditions are determined by means of three factors used in machine vision:

1. Sensor properties (range of electromagnetic radiation wavelengths that the sensor is able to sense, sensor size, possibilities of amplifying the signal, ...).

2. Properties of the captured technological scene (reflection of certain wave lengths from the surface, absorbed and emitted irradiation, transmitted irradiation, Chapter 1, position of the captured objects).

3. Radiation properties (light intensity, position of the source, wavelength of the emitted radiation, diffusion rate of the radiation and its directionality, ...).

Sensor properties were described in the previous chapters. Properties of the observed technological scene consisting of the observed object and its background should be limited to visible spectrum radiation (approx. 380 to 750 nm) or ultraviolet and infra-red radiation. When selecting lighting, you should mainly know the following:

- What wavelength will be reflected from the object surface? Part of radiation will be absorbed upon incidence. Absorption can be identical throughout the whole spectrum or higher for certain wavelengths (forming colours if the radiation is within the visible spectrum).

- While the surface of the observed object ranges between the extremes (perfect mirror surface and Lambert surface, Chapter 1.3.2)? Meaning that the reflection will be rather mirror or diffuse within the observed band of electromagnetic radiation.
- In case the object is transparent, what wavelength is more absorbed and what wavelength passes through the object (causing decrease of intensity of the passing light and eventually a change of colour of the passing radiation that is within the visible spectrum), if the radiation becomes diffused upon passing (e.g. strong diffusion of the "exiting" light)?
- Whether the object emits radiation to which the used sensor is sensitive (e.g. melted glass, where the glass emits visible radiation).
- What is the position of the observed object and the background objects towards the sensor and source of radiation?

Properties of radiation (lighting) will be described further and with regard to properties of the observed technological scene.

Properly selected lighting of the technological scene may speed up and simplify the image analysis and decide whether it is possible to capture the scene at all. While in previous points, the proper solution can be quite well estimated, often precisely calculated, selecting proper lighting requires a lot of experience and usually a lot of experimenting. The design of lighting is a key to success, especially if the scene consists of transparent, glossy, or featureless object in terms of relief. Which is why this matter must have the attention and investment it requires. It can be said that application with high quality and professional solution will require lighting with similar price as the price of camera itself.

When selecting lighting, it is necessary to determine the type of source, amount of sources as well as their location, configuration, colour of the light (wavelengths) and many more. The solution must often feature shielding of undesired ambient light by means of shields or colour filters on the camera. Polarization filters may greatly participate on reducing undesirable reflections.

6.1 Light sources

A light source may be for example a daylight, fluorescent tube, light bulb, gas tube, professional solutions use LEDs, laser, etc. Daylight is problematic in industrial applications, since its direction and intensity is changing with time. Using ambient light only, e.g. room lighting, is not suitable in most applications, since it can change regardless of the observed manufacturing process. It is therefore common and appropriate to use separate lighting for the technological scene. In simple tasks, it is logically better to use less expensive lighting such as fluorescent tube or gas tubes (even without the electronic stabilizer). When using this relatively inexpensive solution, it is necessary to respect certain limits of these sources. For example, the duration of exposure must be longer than the refresh rate of the light.

Lighting system for industrial applications is usually fitted with LED or lasers. Both types offer better quality and better parametrization (can be even directly controlled by the camera in some cases). ***LEDs*** do not require high voltage and their luminous flux can be regulated. Properties of the light remain almost unchanged with time and the lighting has long service life.

Light sources can be distributed by their **wave length**, usually the radiated colour:

- ultraviolet (light sources using wavelengths ranging from 360 to 395 nm, UV radiation below 315 nm are proven to be harmful to health),
- blue (approx. 470 nm),
- green (approx. 525 nm),
- red (approx. 630 nm),
- infrared (light sources using wavelengths ranging from 850 nm to 950 nm),
- white (characteristics of which light is stated as colour temperature or chromatic temperature),
- RGB.

Beside the common light, the light can be modified to polarized, coherent or monochromatic.

One of the aforementioned light sources are ***lasers*** (Light Amplification by of Radiation). Unlike LED sources, fluorescent tubes, gas tubes and light bulbs, the emitted light is coherent, monochromatic and emitted as a single beam. Lasers are classified by their type to solid-state, gas and semiconductor. The term “laser” includes a whole scale of equipment for various purposes. Lasers are used in biology, medicine, and are finding their way into industrial applications when performing controlled thermonuclear reaction or particle acceleration. Only some of them are suitable for lighting of a technological scene. In general, lasers are divided to four basic safety classes with subclasses. Dividing lasers to classes is decided by the **maximum permissible exposure** (MPE) providing the level of laser radiation to which a person can be exposed under normal circumstances without occurrence of adverse effects of the radiation. MPE levels are the maximum levels of exposure that will not result in damage to skin or eyes immediately of after a longer time period. The MPE levels are dependent on wavelength, length of impulse or duration or radiation, type of irradiated tissue, and size of the image on retina for wavelengths ranging from 400 nm to 1400 nm. Lasers are divided to the following safety classes:

Class 1 - small capacity lasers that are safe under all condition. In these lasers, there is no risk of exceeding the MPE upon direct eye contact or even when using optical accessories (e.g. microscopes, binoculars, etc.). This class includes high capacity lasers, where the beam is shielded from the environment and will automatically shut down upon opening the cover.

Class 1M – lasers are safe when used in the usual ways. The laser poses a risk only when the beam is amplified by means of an optical system (microscope, binoculars). The MPE for class 1M cannot be exceeded, unless the beam is optically focused to a very small cross-section.

Class 2 - lasers with low capacity that emit visible radiation (light) – safety is ensured by physiological reactions of the eye, including the blink reflex. Class 2 lasers can be dangerous upon direct eye contact with the beam for more than 0.25 s. That applies only to lasers working on visible wavelength (400-700nm). These lasers have a limited output power of 1 mW in the continuous mode. Class 2 lasers are for example laser pointers.

Class 2M – lasers are safe – eyesight is protected by the blink reflex. That applies only in case when the beam is not amplified by an optical system. This class includes lasers with high cross-section of the beam or its large divergence. Light coming through the eye pupil must not exceed the Class 2 capacity.

Class 3A - safe lasers when observing with unprotected eyesight; direct eye contact via optical element may pose a risk. The MPE can be exceeded, however, there is a slight risk of damage. Class of lasers operating in visible wavelengths and continuous mode must not exceed 5 mW. Limits for pulse modes and other wavelengths are different.

Class 3B - direct eye contact with the beam is always dangerous; mirror reflection may pose a risk as well. However, reflection from other materials is not dangerous. Lasers operating in continuous mode and with wavelengths from 315 nm to far IR must not exceed the capacity of 0.5W. The limit for pulse lasers with wavelength of 400 to 700 nm (visible spectrum) is 30mJ per pulse.

Class 4 - high capacity lasers are dangerous not only to the eye, but to skin as well, mirror and diffuse reflections are dangerous as well if the Nominal Hazard Zone is not maintained. Lasers within this class are able to severely cut or burn skin. There is also a risk of ignition upon interaction with flammable material.

Lighting is performed only by lasers Class 2, 2M, 3A and 3B. As is apparent from the aforementioned laser safety, it is necessary to ensure own safety system for Class 3 lasers as well as ensure safety if the operator that may be in contact with the laser.

6.2 Lighting configuration

Lighting configuration is different for LED sources (always partially diffused) and lasers (always a beam concentrated to points or curves).

LEDs (even sources based on fluorescent and has tubes) can be divided by their basic geometry to (Fig. 6.1), [38]:

- front bright-field lighting – light is reflecting from the surface of the observed object right into the camera objective lens,
- front dark-field lighting – light is reflecting from the surface structure of the observed object,
- back lighting – or backlight, lighting "goes" around the object (form silhouettes), or in case the object is transparent, it goes through the object itself.

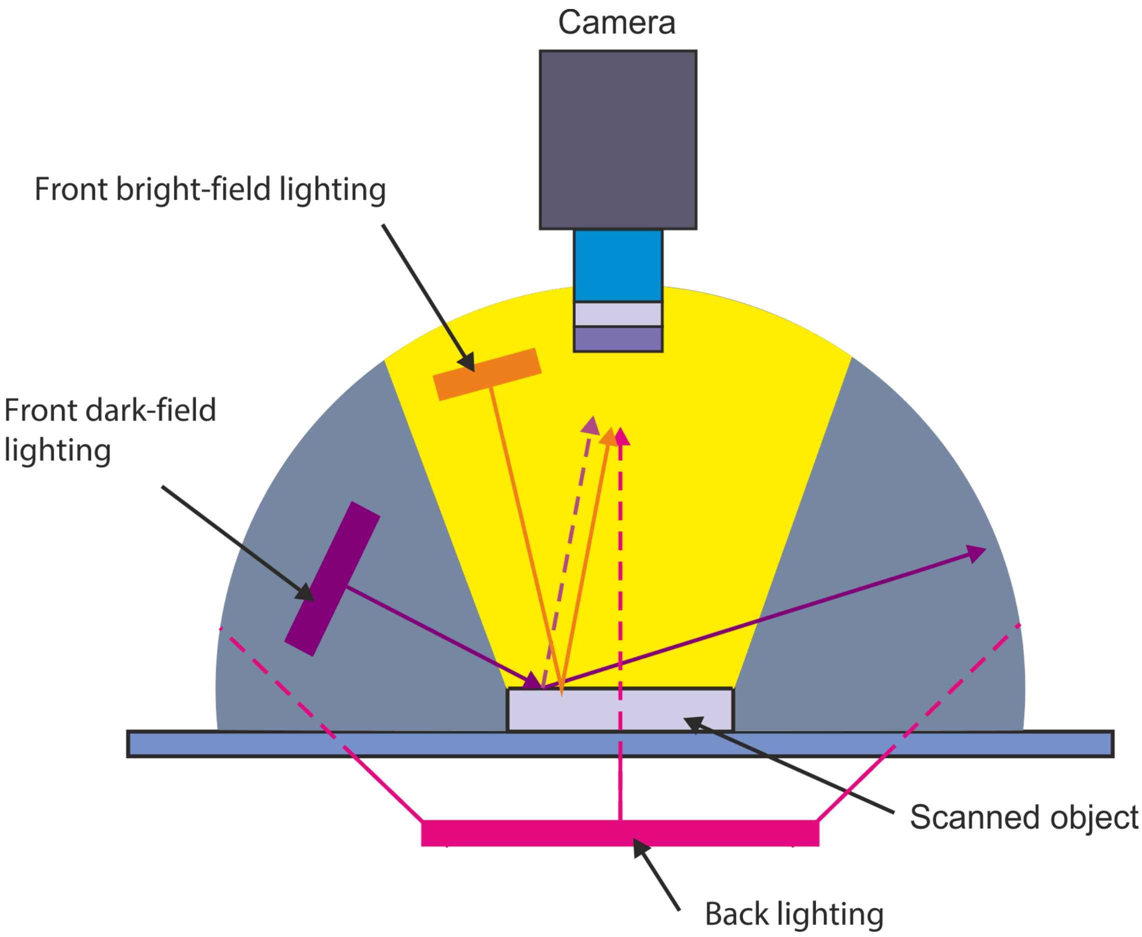

Fig. 6.1 Basic geometry of light used when designing machine vision solution

Light from the light sources can be either diffuse, or direct, or also concentrated (often using Fresnel[8] lenses). Generally, diffuse light decreases the contrasts caused by articulation and structure of the object and highlights the contrasts caused by the absorption. Conversely, when trying to highlight the articulation, directional light is used.

Front bright-field lighting is used in combination with diffuse light in order to create contrasts based on varying absorption. Directional light is used as well, although much less. It is used to create a contrast of glossy surfaces.

Front dark-light lighting creates a contrast image of articulation of surface under the dark-light – this property was a basis for the name; directional light is used.

Back light is used to create a silhouette of an object, typically when measuring dimensions. This type of light is also used to create a contrast silhouette behind a transparent glossy obstacle. Diffuse light is often used.

6.3 Basic types of light

Basic types of light (illuminator) can be divided to:

- area lighting,
- linear lighting,
- ring lighting,
- backlight,
- diffuse dome lighting ,
- diffused on-axial lighting,
- ring lighting with dark-field.

[8] Fresnel lens is named after its inventor, a French physicist Augustin-Jeana Fresnel. Fresnel lens is significantly lighter than lens made from common materials (glass, plastics), while having similar parameters. That is due to the fact that it does not contain parts that are not directly participate on refraction. It's lesser thickness and therefore weight is suitable not only for displaying (where its optical flaws manifests), but for its application in lighting and signalling systems, using solar energy.

Special types of lighting are:

- **co-axial dome** - dome CDI,
- square continuous diffuse lighting (SCDI),
- multi-axial lighting,
- and more.

Lighting is manufactured in configurations for continuous mode as well as pulse mode, Chapter 6.4.

Area Lighting (Area Array) is often used as a directional light source, Fig. 6.2. The diffusion rate is given by the radiation angle of diodes or diffusion filter placed in front of the light source. In simpler applications, this light source can be used instead of expensive bright-field systems and diffused light. It is often used as a source of directional light for dark-field lighting. Small and special area lighting are used as auxiliary lighting or as light sources of complex lighting systems.

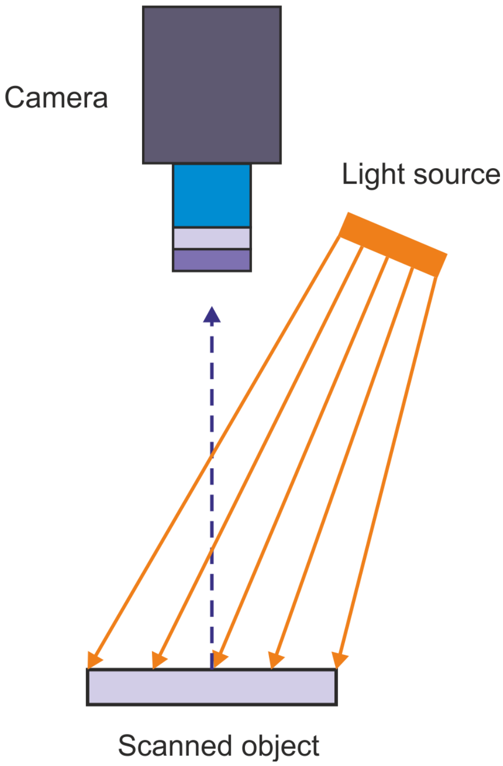

Fig. 6.2 Area lighting

Linear lighting is used mainly as directional lighting for line scan cameras, where light with cylindrical lens is used to create a thin line of light for camera's field of view. This type of lighting is used in other cases as well, e.g. when lighting edges of an object made of transparent material. In such case, light in considered to be the edge of an object and the edges can brightly glow under certain conditions (e.g. sheet of flat glass).

Ring lighting is used to light in the objective lens axis. It is a ring that is encircling the objective lens of the camera, Fig. 6.3. Illuminator provides diffused light coming from the objective lens area and is used for light-field lighting and diffused lighting. It is one of the most usual types of lighting and can be integrated into the camera sensors. The simple configuration is not ideal due to the fact that it is hard to achieve lighting evenness (light characteristics depends on the distance of illuminated scene and varies with the type, diameter and manufacturer). Can be fitted with various filters: Fresnel lens to focus, polarization filter, diffusion (milk) glass,...

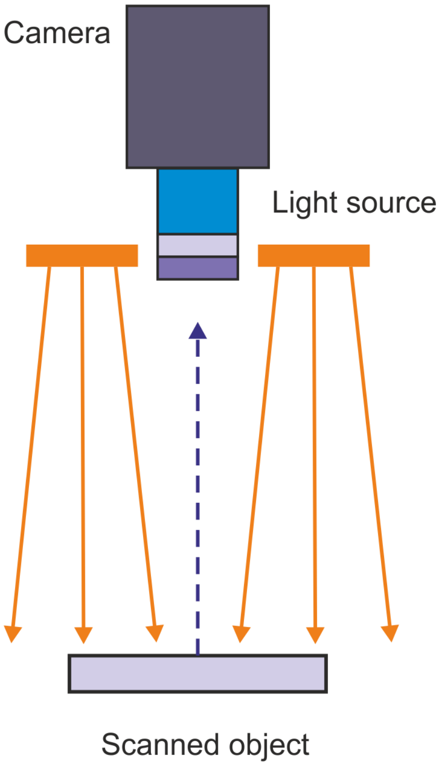

Fig. 6.3 Ring lighting

Backlight is the most usual type of back lighting, Fig 6.4. This type of illuminator is similar to area lighting formed by an array of LED, but is almost always fitted with a large-format diffuser. Diffusion quality (and its evenness) is important mainly when performing measurements where edges of an object must be displayed, or when identifying an object in a transparent case (e.g. inspection of light bulb filament).

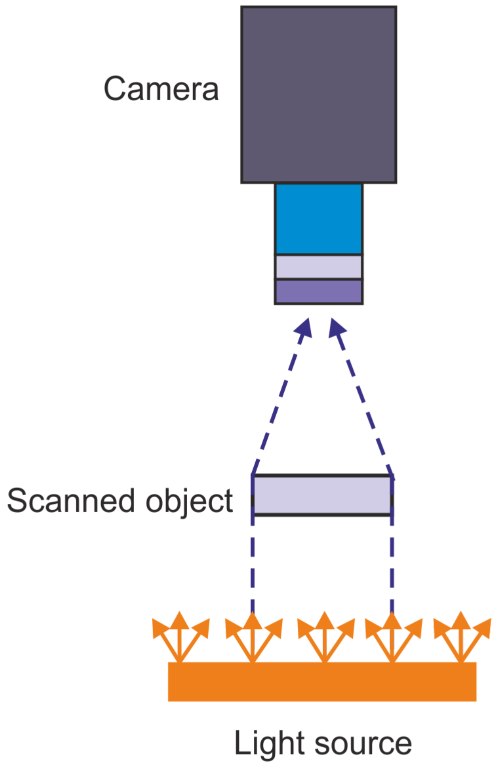

Fig. 6.4 Backlight

Dome Lighting, or shadowless illuminator is used to illuminated by bright-field light. It provides a very well diffused light formed by a diffusion reflection within the inner side of dome, Fig. 6.5. The diffusion may be performed by another shape than dome, e.g. cylinder. The lighting is used for even illumination of the captured object from all four sides if possible to avoid false detection due to shadow of the object or its part. Used for extreme homogeneous lighting of articulated objects and surface. It suppresses the disruptive elements of 3D structures and surface curvature. Using shadowless light is usually used when reading texts and codes on packaging. For example, it allows reading text on a crumpled shiny foil.

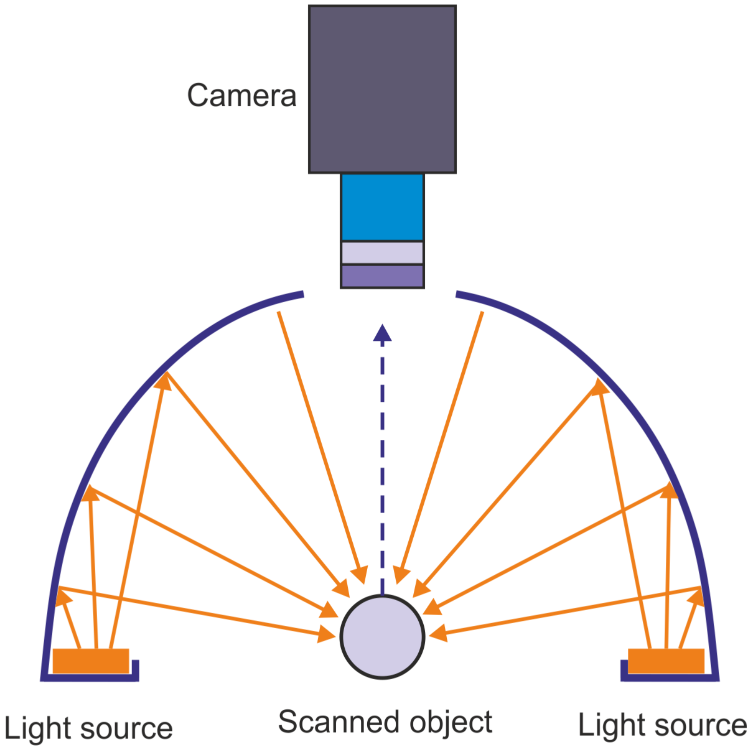

Fig. 6.5 Dome lighting

Diffused On Axis Lighting (DOAL) – illuminator provides lighting of very high quality and bright field as well as diffused light, Fig. 6.6. Diffuser ensures better diffusion of light and its structure with semi-transparent (dichroic) mirror, beam splitter allows directing beams in parallel to axis of the camera objective lens. Such structure allows homogeneous lighting of the captured surface. This system is used to capture glossy surfaces, reading of texts and sensing 1D and 2D codes. Using it with telecentric objective lens provides a very good tool for accurate measurements and display of contours and silhouettes as well as inspection of high reflective surfaces. Coaxial optics is optimised for visible spectrum ranging from 450 to 700 nm. It disadvantage lies in high price and limited field of view.

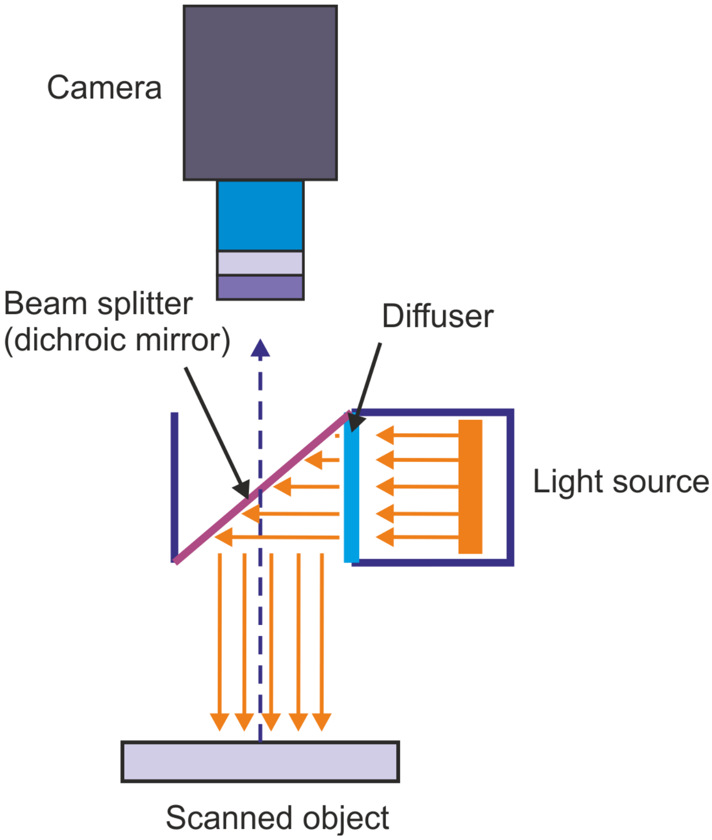

Fig. 6.6 Diffused on axis lighting

Ring Lighting with Dark-Field (Off-Axis Ring) offers front lighting with dark-field, Fig. 6.7. It is implemented as light directed in a plane that is perpendicular to the axis of camera lens. This radial lighting is used to accentuate outlines of shaped and engraved surfaces such as coins, engraved numbers or labels, laser burned labels, texts, etc. These lights are usually used for identifying scratches on smooth surfaces or transparent materials (perspex). Flat glossy surfaces will stay dark for the camera, while the cracks and scratches on the material reflect the light. A typical dark-field illuminator used for these purposes consists of a ring of LEDs with thin radiation characteristics lighting from the circular plane towards its centre.

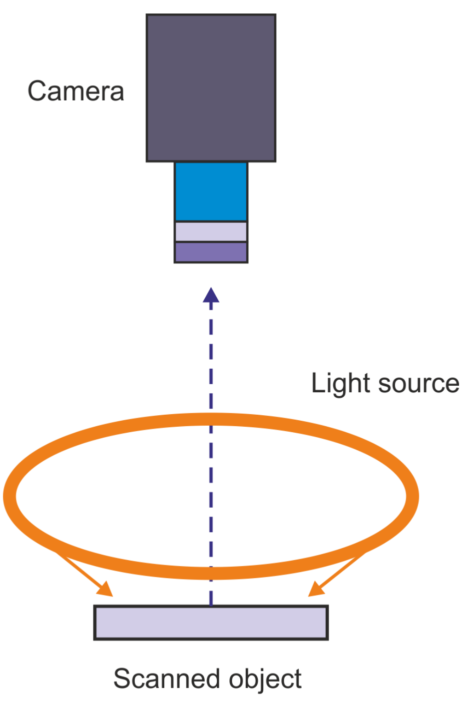

Fig. 6.7 Dark-field ring lighting – low angle lighting

Special types of lighting are a combination of the aforementioned lighting. For example, the dome lighting can be fitted with Diffusion On-Axis Light to direct the light to the direction of camera. That provides perfectly diffused light as in cloudy environment (***CDI- Cloudy Day Illuminators***, Fig. 6.8). **Multi-axial lighting** is very similar, Fig. 6.9. In this case, the diffusion on-axis lighting is fitted with low angle lighting, forming a combination of diffused bright-field light and dark-field light. Another example is the Square Continuous Diffuse Illuminators, (SCDI, Fig. 6.10). It is a modified version of diffusion on-axis lighting that offers better diffused light.

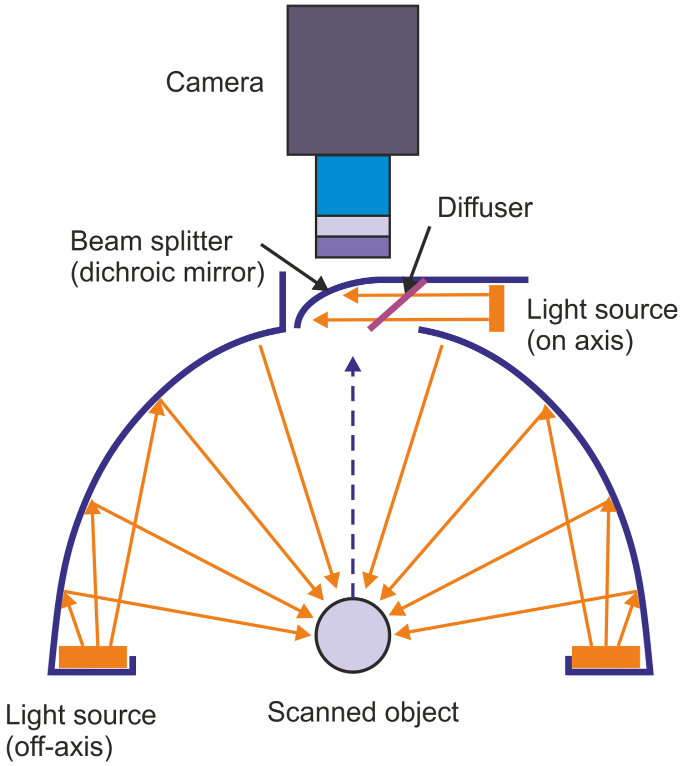

Fig. 6.8 CDI Lighting – Cloudy Day Illuminators

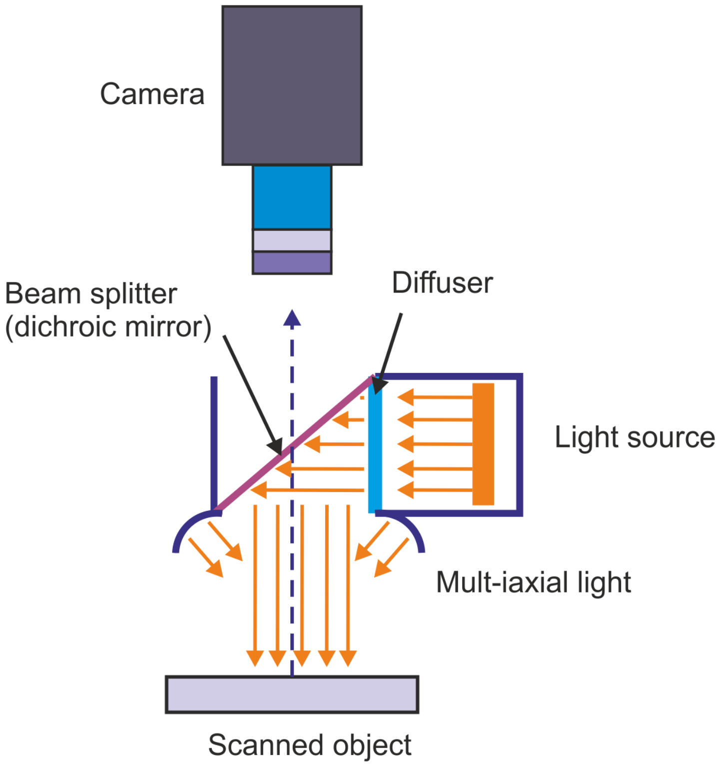

Fig. 6.9 Multi-axial lighting

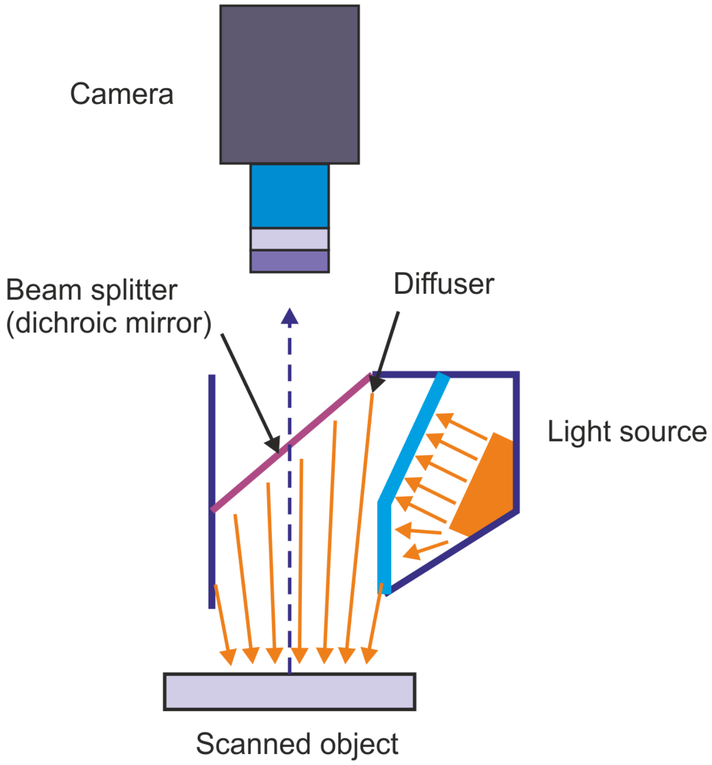

Fig. 6.10 Square Continuous Diffuse Lighting,

Advanced and demanding applications of lighting must respect the lighting characteristics and rate of lighting declination from centre to edges. The dependency also changes with distance of light from the observed technological scene. Such characteristic is provided in documentation of high-quality lighting. If it is necessary to control the lighting, e.g. set brightness, colours or flashing, this system offers a great advantage in a form of light control performed directly by the camera.

The lighting system can also be fitted with various patterns so it does not have to be diffused all around the technological scene. In glass industry, those are for example zebra stripes or other motives. Those are usually detection of optical faults on the products, generally made from transparent materials.

Lasers are used for point or line lighting. However, there are other types of lasers, Fig. 6.11. The wavelength usually corresponds to red colour.

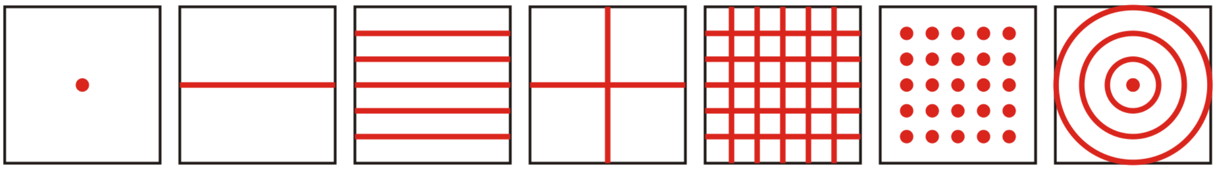

Fig. 6.11 Example of laser lighting types, from the left: point, line, 5 lines, cross, grid, matrix of point, set of circles

6.4 Lighting design process

Lighting must correspond with the goal of analysis. Unless it is an analysis of an object such as shape, measurement, calculation, determination of orientation, reading of text, etc., the main goal is to achieve maximum contrast of parts that are to be a subject of analysis and parts that are not to be inspected. In general, contrast of the part of interest and "distracting" part of an object can be achieved in two basic ways: taking advantage of various absorption (of emission) or of the difference in brightness generated by a properly designed directional lighting [38]. But it does not apply generally. When analyzing the surface it is necessary to obtain the relevant information. The surface structure is in 3D, but it is converted into a 2D representation. A type of lighting structure need to choose that adequately show the structure without distortion used illuminator. From this perspective, this task is one of the most difficult. The whole process of lighting design can be summarized in the following six points:

- analysis of properties of the captured technological scene with regard to the machine vision task,
- analysis of required properties of the light used,
- determining appropriate light geometry,
- selection of appropriate illuminator,
- selecting illuminator control,
- eliminating disturbing elements and restricting unauthorised access into the system.

In order to find an appropriate lighting, it is necessary to know ***the properties of the captured technological scene***. Properties of the background are important as well. The background is often selectable, and it is therefore wise to use this possibility when designing a machine vision system.

Additionally, one should focus primarily on:

- shape and size of the observed scene; this information is used to determine size and distance of lighting, including its characteristics with regard to decrease of light intensity towards the edges,
- properties of reflectance/absorption of light in individual parts of the observed technological scene with the aim to obtain optimal information about the captured object or its surface, therefore define properties of the captured object that must be in contrast with the objects in background,
- to decide whether the structure of the observed object should be accentuated or suppressed,
- to define whether the articulation of the object (protrusions and gaps) are to be suppressed or made a point of interest,
- in case an object is transparent (translucent or even transparent), it is necessary to determine whether this property should be used or suppressed,
- determine the way of position change of the object (in what position will the objects be during capturing), and will the object move during the scanning process?

When ***analysing the necessary properties of the light used***, it is necessary to respect its interaction with the observed object and the background as well as its following effect on the image sensor. When performing the analysis, basic light interaction properties of an object and background can be used. Based on laws of physics, the light is reflected, absorbed, transmitted, part of the light may even cause secondary emission in the material, Chapter 1, Fig. 1.3. Each of these phenomena can be used in machine vision. Additionally, the interaction of light also depends on the colour (wavelength) of the light and the object. The amount of excited electrons (amount of accumulated charge), therefore the camera sensitivity, depends on the wavelength as well. In machine vision, the most common assembly is a combination of monochromatic camera together with a red light illuminator. This combination is good due to spectral sensitivity of most cameras, however, in many cases, it may not provide with optimal results, since the contrasts created on the object by absorption of monochromatic red light may significantly different from contrasts visible in white light.

In case of glass industry, it is possible to use radiation emission (in visible spectrum) of heated glass. Naturally, this can be used in other materials. The main issues are obtaining dark background, protection of visualisation system from increased temperature, Chapter 3.1.13.

Determination of basic appropriate light geometry is based on basic geometry shown in Fig. 6.1. With regard to information in Chapter 6.2, it is necessary to decide if it is better to use front bright-field lighting (usually suitable for detection and display of the object followed by detection of shape, object counting, identifying orientation, ...), front dark-field lighting (used mostly when determining surface quality or structure, ...) and back lighting (for example for precise measurement or display of opaque parts in transparent material). Furthermore, it is necessary to decide whether the light will be directional (accentuating articulation and structure of an object) or diffuse (increasing contrasts by absorption).

Appropriate illuminator can be selected from the basic types mentioned in Chapter. 6.3. There are many manufacturers, but most professional solutions use LEDs. As mentioned above, there are alternative types available, where certain limitations must be respected.

Illuminator control is necessary if the light is turned on upon exposure (e.g. when controlling exposure by duration of light – scanning swift movement or if higher intensity is required – flash is more intensive than continuous operation, up to 10 times). However, the illuminator control is usually performed only with a control unit provided by the manufacturer of illuminator. Practically, all machine vision systems feature a digital output allowing synchronising light with the image that can be directly connected to the illuminator control unit. However, flashing of the source may negatively affect the human operator, since such flashing causes fatigue, and health problems upon long-term exposure.

It is not necessary to control the illuminator when continuous lighting is required. Nevertheless, it is possible to use power control (luminous flux, changes of light intensity). This way of control is good for setting lighting, lighting experiments in a technological scene and when eliminating the eventual decrease of luminous flux due to source weariness by increasing the power. When using illuminator control, the set control values must not be change by an unauthorised person.

Elimination of disruptive effects and unauthorised interventions into the system are a very important aspect of the illumination design. Ideally, repeatable lighting conditions can be achieved by having an appropriate lighting system as the only light source. Real industrial implementation require considering changes of day-light and artificial interior lighting. Change of these parameters in time (depends on time of day and year, weather and many more) is often hard to define and may cause undesired influence on lighting of the technological scene. That may lead to errors in image processing, even malfunction of the whole system. Basically, there are two ways of eliminating the undesired light sources:

1. Shielding – the more suitable variant, which is also less expensive and should be preferred.

2. Using illuminators with much higher luminous flux; this method is used when it is not possible to shield the scene. The effect of disruptive lighting is decreased, but not eliminated. Another problem lies in higher power, which is however compensated by controlled (flashing) illuminator at the time of exposition. Unfortunately, this solution may have negative effect on human health. Additionally, strong illumination often requires aperture of the objective lens, which may be undesirable in some cases.

Another problem in industrial practice lies in the necessity to eliminate parasitic reflection on the captured object and background of the technological scene, since these reflection may lead to blooming. These reflections may not be in visible spectrum, but may occur in near infra-red radiation. Due to the fact that some cameras are able to operate in this spectrum, some reflection may not be visible by eye. In such case, use a camera with an IR cut filter or use such filter on objective lens.

Equipment for image processing, that is the camera and its cover and the lighting system, should be regularly cleaned. Impurities (often dust) deposits on the lighting as well. In dusty environment, the illuminator as well as the camera must be regularly cleaned by blowing compressed air on each part. When the structure of illuminator case is designed badly, dust may get inside and deposit on diffusers, mirrors and other optical elements. Decrease of luminosity may not be noticeable by eye, but may cause unexplainable increase of measurement error.

7 Conclusion

This text is the first part of set for studying subjects related to machine vision problems. Contains information about basic principles of machine vision, cameras, lighting and other chapters that are important for image interpreting.

Application possibilities of machine vision are very extensive. From common identification and interpretation of object in 2D, quality assessment, through description of structure to applications for 3D image modelling.

This publication is followed by another one: Introduction to Problems of Machine Vision, part 2: Image processing, [1]. Part two shall be studied as well in order to have a comprehensive overview about this matter.

Literature

[1] HOTAŘ, V. *Introduction to Machine Vision, part 2: Image processing basics.* Separate materials, 2017

[2] BLANCHET, G., CHARBIT, M. *Digital signal and image processing using MATLAB*. Volume 1, Fundamentals. Revised and updated 2nd edition. London: ISTE, 2014. ISBN 978-1-84821-640-2.

[3] SONKA, M., HLAVAC, V., BOYLE, R. *Image Processing, Analysis, and Machine Vision.* Pacific Grove: Books/Cole Publishing Company, 1998. 770 s.

[4] HARTLEY, R., ZISSERMAN, A. *Multiple view geometry in computer vision.* 2cn Edition. Cambridge; New York : Cambridge University Press, 2003. 655 s. ISBN 978-0-521-54051-3

[5] GONZALEZ RC, WOODS RE. *Digital Image Processing.* 3rd edition. Prentice-Hall, New Jersey. 2010.

[6] WOHLER, Ch. *3d computer vision: efficient methods and applications.* 2nd. London : Springer, 2013. str. 381. ISBN 978-1-4471-4149-5.

[7] TYLER, Ch., W. *Computer vision*: *from surfaces to 3D objects*. Boca Raton: CRC Press, c2011, 250 s. ISBN 978-1-4398-1712-4.

[8] FOLEY, James D. *Computer graphics: principles and practice*. 2nd ed. Boston: Addison-Wesley, 1997. ISBN 0-321-21056-5.

[9] Eck, J., D. *Hobart and William Smith Colleges.* [online]. Hobart and William Smith Colleges. 2010. [accessed 25/05/2017]. Available at: http://math.hws.edu/eck/cs424/downloads/graphicsnotes2010-linked.pdf

[10] GRUM, Franc a Richard J BECHERER. *Radiometry*: *Optical Radiation Measurements*. 1. vyd. New York: Academic Press, 1979, 335 s.

[11] IKEUCHI, K. *Surface reflection mechanism. Handbook of Pattern Recognition and Image Processing: Computer Vision*. Red. Young, T. Y. Vol. 2, San Diego (USA): Academic Press, 1994, pp. 131-160.

[12] HOLST, G.C., LOMHEIM, T.S. *CMOS/CCD Sensors and Camera Systems*. J C D Publishing, 2011.

[13] ADEPT TURNKEY. *Sensor Comparison II: Interline Scan, Frame Transfer & Full Frame*. [online]. [accessed 25/05/2017]. Available at: http://www.adept.net.au/news/newsletter/200810-oct/sensors.shtml

[14] COMPUTER HISTORY MUSEUM. *1963 - Complementary MOS Circuit Configuration is Invented* [online]. [accessed 25/05/2017]. Available at: http://www.computerhistory.org/semiconductor/timeline/1963-CMOS.html

[15] MEROLI, S. *CMOS Active Pixel Sensor Vs CCD. Performance comparison*. [online]. [accessed 25/05/2017]. Available at: http://meroli.web.cern.ch/meroli/lecture_cmos_vs_ccd_pixel_sensor.html

[16] Teledyne DALSA*. CCD vs. CMOS. Which is better? It's complicated...* [online]. [accessed 25/05/2017]. Available at: http://www.teledynedalsa.com/imaging/knowledge-center/appnotes/ccd-vs-cmos/

[17] R.L. BAER, W.D. HOLLAND, J. HOLM, P. VORA, *A comparison of Primary and Complementary Color Filters for CCD-based Digital Photography*, Proceedings IS&T/ SPIE Symposium on Electronic Imaging, San Jose, CA, pp. 16-25, 1999. [online]. [accessed 25/05/2017]. Available at: http://www.seas.gwu.edu/~poorvi/EI99.pdf

[18] THE PHOBLOGRAPHER. *An Introduction to and Brief History of Digital Imaging Sensor Technologies.* [online]. [accessed 25/05/2017]. Available at: http://www.thephoblographer.com/2013/07/31/an-introduction-to-and-brief-history-of-digital-imaging-sensor-technologies/

[19] BASLER. *Products – Cameras, Powerpacks, Accessories, Software.* [online]. [accessed 25/05/2017]. Available at: https://www.baslerweb.com/en/products/

[20] WIKIPEDIA. *Fresnel lens.* [online]. Wikipedia Foundation. [accessed 25/05/2017]. Available at: https://en.wikipedia.org/wiki/Fresnel_lens

[21] JAI. *Industrial CCD and CMOS cameras.* [online]. [accessed 25/05/2017]. Available at: http://www.jai.com/

[22] PIXELINK. *Machine Vision Camera.* [online]. [accessed 17/05/2014]. Available at: http://www.pixelink.com/

[23] IIDC,1394 Trade Association. *1394-based Digital Camera Specification.* [online]. 2. 12. 2004 [cit. 17/05/2014]. Available at:

[24] PHYS.ORG. *The trillion-frame-per-second camera*. [online]. [accessed 25/05/2017]. Available at: https://phys.org/news/2015-04-trillion-frame-per-second-camera.html

[25] BASLER. *NIR Cameras* [online]. [accessed 17/05/2014]. Available at: https://www.baslerweb.com/en/vision-campus/camera-technology/nir-cameras/

[26] NATIONAL INSTRUMENTS CORPORATION. *Choosing the Right Camera Bus.* [online]. [accessed 25/05/2017]. Available at: http://www.ni.com/white-paper/5386/en/

[27] WIKIPEDIA. *IEEE 1394.* [online]. Wikipedia Foundation. [accessed 25/05/2017]. Available at: https://en.wikipedia.org/wiki/IEEE_1394

[28] NATIONAL INSTRUMENTS. *What Are The Differences Between Base, Medium, and Full Camera Link Configurations?* [online]. [accessed 25/05/2017]. Available at: http://digital.ni.com/public.nsf/allkb/2EE80B8C381D61D286257B1F005674A5

[29] TEXAS INSTRUMENTS INCORPORATED. *LVDS Owner's Manual. Including High-Speed CML and Signal Conditioning.* [online]. [accessed 25/05/2017]. Available at: http://www.ti.com/lit/ml/snla187/snla187.pdf

[30] COAXPRESS. *The next generation digital interface.* [online]. [accessed 25/05/2017]. Available at: http://www.coaxpress.com

[31] EDMUND OPTICS. *Understanding Focal Length and Field of View.* [online]. [accessed 26/05/2017]. Available at: https://www.edmundoptics.com/resources/application-notes/imaging/understanding-focal-length-and-field-of-view/

[32] WIKIPEDIA. *Lens mount.* [online]. Wikipedia Foundation. [accessed 17/05/2014]. Available at: http://en.wikipedia.org/wiki/Lens_mount

[33] FIKAR, D*. Thermovision Camera Cooling.* [Bachelor thesis, in Czech]. Liberec: Technical University of Liberec. 2010.

[34] RINNER, B. *Distributed Smart Cameras.* [online]. Rice University, Department of Electrical and Computer Engineering. [accessed 26/05/2017]. Available at: http://www.ece.rice.edu/~av21/Documents/pre2011/Distributed%20Smart%20Cameras.pdf

[35] Teledyne DALSA. *Understanding Line Scan Camera Applications.* [online]. Edmund Optics. [accessed 26/05/2017]. Available at: http://leadwise.mediadroit.com/redirect/9883/29207/

[36] HECHT, E. *Optics.* 4th Edition. San Francisco: Pearson Education, Inc., 2002. ISBN 0-8053-8566-5

[37] HOLLOWS, G. Imaging Optics Fundamentals. [online]. Edmund Optics. [accessed 26/05/2017]. Available at: https://www.visiononline.org/userAssets/aiaUploads/file/T4-Beginning-Optics-for-Machine-Vision.pdf

[38] DARYL M. *A Practical Guide to Machine Vision Lighting.* [online]. Advanced illumination. 2007. [accessed 26/05/2017]. Available at: www.graftek.com/pdf/Marketing/MachineVisionLighting.pdf

List of used symbols (in both parts of textbook)

A	point set in the n-dimensional Euclidean space, morphological transformation
B	structural element for morphological transformation
$\mathbf{B}$	binary matrix
$\breve{B}$	symmetrical point set, morphological transformation
c	image of constant brightness (grey etalon)
C	aperture diameter [mm]
$\mathbf{C}$	general matrix of complex number
d	lens diameter [mm]
D_4	distance in city blocks
D_8	chessboard distance
D_E	Euclidean distance
d_b	original (ideal) brightness (image) function
E	radiation intensity (irradiance, radiance) [W m^{-2}]
E_β	radiation intensity under an angle [W m^{-2}]
E_f	light intensity [lux, lm m^{-2}]
$E_i(\lambda)$	intensity of radiation incident on the surface area [W m^{-2}]
$E_r(\lambda)$	intensity of radiation reflected back to space [W m^{-2}]
E_0	radiation intensity in optical axis [W m^{-2}]
$\mathbb{E}_n$	n-dimensional Euclidean space
f	focal length [mm]
f_b	brightness (image) function
$\mathbf{f}_b$	matrix of brightness (image) function
f_c	brightness (image) function of captured image with constant brightness (etalon)
f_n	interpolated brightness (image) function
f_r	two-way distribution function of BRDF reflection [sr^{-1}]
F_b	transformed brightness (image) function
$\mathbf{F}_b$	matrix of transformed brightness (image) function
g_b	sampled version of continuous brightness function f_b (output image)
h	convolution mask (core)
h_f	brightness histogram

$\mathbf{I}$ general unitary matrix

j imaginary unit

K structural element for grey-scale morphological transformation

$\mathbf{L}$ source of light

L radiance (specific luminosity) [W m^{-2} sr^{-1}]

L_f amount of brightness levels of a histogram

L_{WD} working distance between a camera and technological scene [mm]

m characteristic rates in granulometry

$\mathbf{M}$ general matrix of real numbers

$\mathbf{M}^T$ Transposed matrix of a matrix $\mathbf{M}$

M_{PMGA} basic magnification [-]

$\mathbf{n}$,n normal

n_i refractive index of surrounding environment [-]

n_t refractive index of transparent material (glass) [-]

N aperture value

$\mathbb{N}$ set of natural numbers

$\mathbf{P}$ transformation matrix with size $U \times U$

O set of all pixel area

O_4 circle of radius 1 in the grid of 4 neighbours

O_8 circle of radius 1 in the grid of 8 neighbours

O_E circle of radius 1 in Euclidean space

O_H circle of radius 1 for hexagonal grid

P background of an area

$\mathbf{Q}$ transformation matrix with size $V \times V$

$\mathbb{Q}$ set of rational numbers

r *radius*

r_1 and r_2 curvature radii of optical surface [mm]

R reflectance function

$\mathbb{R}$ set of real numbers

s coefficient of scale change

$T[A]$ top value of set A, grey-scale morphology

T_h threshold

T_E	duration of exposure
u_S	selected dimension of camera sensor [mm]
U_S	selected dimension of area of interest of technological scene [mm]
U	point set, morphological transformation
U^C	complement of point sets, morphological transformation
$\mathbf{V}$	observer
$V[f]$	shadow function f, grey-scale morphology
$\mathbb{Z}$	set of integers
α	angular field of view [°]
β	angle between an optical axis and a beam incident on the sensor [sr]
δO	elementary surface on the object
δI	display of elementary surface in an image
Θ	angle between normal line of the surface and connection with the initial point of a coordinate system
Θ_i	angle of incidence (between the normal line and incident beam)
Θ_r	angle of reflection (between the normal line and reflected beam)
Θ_t	angle of refraction (between normal line and refracted beam)
λ	wavelength of the electromagnetic radiation [m]
ρ	refraction coefficient or albedo (describes the ratio of incident energy and is reflected back to space)
σ	standard deviation
ϕ	radiant flux [W]
Φ	edge direction
$\mathbf{\Phi}_{JJ}$	transformation dimension matrix $J \times J$
ψ	direction of brightness function gradient
ψ_T	morphological transformation
${\psi_T}^*$	dual morphological transformation
∇	gradient (vector function)
$\lvert\nabla f_b(u,v)\rvert$	brightness function gradient module
∇^2	omnidirectional linear Laplace operator - Laplacian
*	convolution

List of abbreviations used (for both textbooks)

CCD charge-coupled device, a sensor

CMOS APS Active Pixel Sensor, manufactured by means of Complementary Metal–Oxide–Semiconductor

CMY(K) subtractive colour mixing model: cyan–magenta–yellow(–key)

DFT discrete Fourier Transformation

EV Exposure Value as a combination of time and aperture in absolute values

FF image data loading from Full Frame sensor (requires mechanical aperture)

FFD distance of flange at the end of thread from image focus, which is an active surface of the image sensor (Flange Focal Distance)

FFT Fast Fourier Transformation

fps frames per second

FT Fourier Transformation

FTr Frame Transfer Electronic Shutter

HMI Human-Machine-Interface, e.g. intelligent cameras

HSV colour model: hue – saturation – value

IA Imaging Area (used with chips fitted with electronic shutter)

IP International Protection, standardized according to ČSN EN 60529

IR infra-red radiation

IR cut filter filter for eliminating near infra-red radiation (NEAR IR)

IT Interline Transfer electronic shutter

MTF Modulation Transfer Function

NIR near infra-red (approx from 400 to 1000 nm)

RGB additive colour model: red-green-blue

SA Storage Area (used with chips fitted with electronic shutter)

SNR Signal Noise Ratio, ratio of undistorted signal and noise

UV ultraviolet radiation

UXGA general name for resolution of sensor or screen of 1600 x 1200 (1.9 MPx)

VGA general name for resolution of sensor or screen of 640x480 (0.31 MPx)

VIS sensor working in visible spectrum of radiation (approximately 380 to 780 nm)

INTRODUCTION TO MACHINE VISION
PART 1: FUNDAMENTAL PRINCIPLES AND HARDWARE

Vlastimil Hotař
Translation: Adam Babka

Reviewers	Assoc. Prof. Maroš Tunák
Published by	Technical University of Liberec
Approved by	The Rector's Office at TUL on 21st of December 2017
Ref. No.: RE 72/17	
Published on	December 2017
Pages	126
Issue	1st
Publication No.	55-072-17
ISBN	978-80-7494-391-1
Electronic version	