普通高等教育“十二五”规划双语系列教材

Intelligent Visual Perception Tutorial

智能视感学

（英文版）

Xiubin Zhang （张秀彬）
Muhammad Mansoor Khan （曼苏乐） 著

中国水利水电出版社
www.waterpub.com.cn

内 容 提 要

本书从计算机视感及其信号处理的基本概念与基础理论出发，阐述了基于图像信息的识别、理解与检测技术原理与方法。

本书根据作者多年来从事智能视感理论与技术研究的成果，结合研究性本科与研究生教学特点编撰而成。全书分为基础篇与应用篇两大部分，其中，基础篇系统地介绍了智能视感的基本原理、方法、关键技术及其算法；应用篇则由配合主要基础理论和方法的应用技术实例所组成。全书遵循理论知识与实用技术的紧密结合、数学方法与实用效果的相互映证等编写原则。本书涉及的教学内容主要包括：图像处理基础、摄像机数学模型、视感识别与检测原理、智能视感实用技术等。

本书可以作为检测与控制、自动化、计算机、机器人及人工智能等专业的高年级本科生和研究生的教材，也可作为专业技术人员的参考工具书。

图书在版编目（CIP）数据

智能视感学 = Intelligent Visual Perception Tutorial : 英文 / 张秀彬，曼苏乐著. -- 北京 : 中国水利水电出版社，2012.8
普通高等教育“十二五”规划双语系列教材
ISBN 978-7-5170-0090-7

Ⅰ. ①智… Ⅱ. ①张… ②曼… Ⅲ. ①计算机视觉一高等学校一双语教学一教材一英文 Ⅳ. ①TP391.41

中国版本图书馆CIP数据核字(2012)第195290号

书　名	普通高等教育“十二五”规划双语系列教材 **Intelligent Visual Perception Tutorial 智能视感学**（英文版）
作　者	Xiubin Zhang（张秀彬） Muhammad Mansoor Khan（曼苏乐） 著
出版发行	中国水利水电出版社 （北京市海淀区玉渊潭南路1号D座　100038） 网址：www.waterpub.com.cn E-mail：sales@waterpub.com.cn 电话：(010) 68367658（发行部）
经　售	北京科水图书销售中心（零售） 电话：(010) 88383994、63202643、68545874 全国各地新华书店和相关出版物销售网点
排　版	中国水利水电出版社微机排版中心
印　刷	北京瑞斯通印务发展有限公司
规　格	184mm×260mm　16开本　20印张　698千字
版　次	2012年8月第1版　2012年8月第1次印刷
印　数	0001—2000册
定　价	**39.00**元

Contents

Application article

Appendix

Base article

Chapter 1 Introduction

1.1 Overview

The concept of Intelligent Visual Perception (IVP) can be considered as a non-contact sensor. It is based on the new scientific concept aimed to achieve scenery image cognition and understanding . The inspect sense technology belongs to a more generalized engineering application field call detection technology and hence is often called inspect sense detection.

1.1.1 Concept about the Visual Perception

From the 1960s onward till the beginning of the 21st century, the image sensor and signal processing theory has attracted worldwide attention. The academic community has also taken a keen interest in this area not only because it is a new technology but it also act as a way to provide the relation between observed natural phenomenon and the associated scientific laws.

When the real-world objects which are often three-dimensional (3D) and even sometime are multi-dimensional, are viewed through an image sensor, a 2D image of the scenery is obtained. The basic problem is to restore the 3D information to get a perception of a 3D world using the 2D image obtained from the sensor. The problem comes directly in the domain of computer science. The solution to the problem is still in the domain of research and need further exploration. The basic idea involves extraction of image information features and using signal analysis tools to regenerate the 3D perspective.

The perception of the image of scenery in brain using image of an object on retina of eyes is very similar to a process involving the image sensor and signal processing. The concept of "visual perception" is based on the following three aspects of consideration:

First, both "computer vision" and "machine vision" are just simple extensions of the terminology "vision" originating from eyes, which obviously cannot embody the essence of physics and biology. Although the "visual sense" and "vision" makes little difference, but the former is already fully displayed to pick the brain organ that play the important information and the role that cannot be replaced.

Secondly, the terminology "visual perception" divides the image processing into two parts: image sensing and image perception. Image sensing is equivalent to the human eye, retina rod and vertebral shape cells. The rapid contact between image sensing with surrounding environment of outside world is directed through the light. The image perception is made on brain through some part of eye called retina that consists of retinal cells that send signals to the brain as the sensing area. It is in the induction zone reading. The analysis of light signal helps to identify object distance, size, color, shape and other specific information. It can also be perform associative, flip and so on imaging. Undoubtedly, neither the image sensing nor the image perception alone could finish the whole job. In other words, to realize the non-contact detection based on machine vision, we should combine the two together, which originated the terminology "*visual perception*" . Based on the idea of "visual perception" an intelligent machine for image understanding and comprehension can be integrates two scientific areas of "machine vision" and "image analysis" . Although machine vision and image analysis are two related topics but are quite different research fields. In some sense, image analysis is a part of machine vision, but in a wider sense, it has a much broader domain.

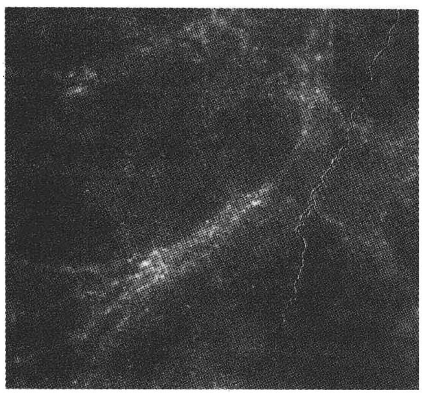
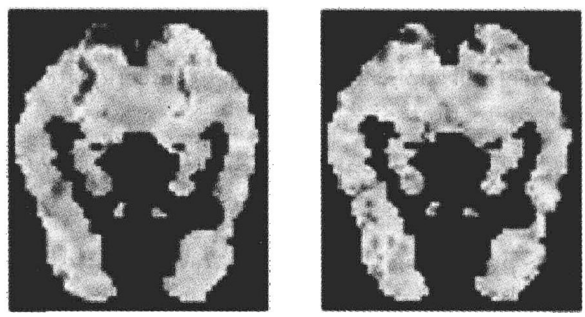


Fig. 1 - 1 A mammalian brain operation microscopic image



(a) (b)

Fig. 1 - 2 Two human brain thinking microscopic images

Fig. 1 - 1 gives a frame of mammal' s cerebral motion under microscope, which is used in research on clinical pathology mechanism and representation of mammals through analysis, transfer, display and other means. "Machine vision" obviously does not have this kind of function, and in other words, "machine perception" has expanded itself into the field of "image analysis" .

Fig. 1 - 2 gives 2 frames of computed tomography of human' s brains. The Fig. 1 - 2 (a), shows a plane image of brain scan that represents imagination of a person. And in Fig. 1 - 2 (b) is another observation which indicates another plane image of brain scan representing imagination of the person. Obviously, it is impossible to read and interpret the underlying information with "machine vision" only. The image analysis algorithms can be used to understand the difference between the two. This is the value of "perception" .

1.1.2 The Development of Visual Perception Technology

The creative work about Visual Perception technology was started in the mid

1960s. But from the beginning until now, its theory and method has been used to explain the terms and concepts in a Vision. The paper "Machine Perception of Three-Dimensional Solids" proposed in 1965 by Lawrence G. Roberts (1937-, an American scientist) is a typical representative in which Roberts proposed block theory. David Courtnay Marr the, British neurologist a who is considered to be the founder of visual processing and computational neuroscience, put forward a set of visual calculations to describe the visual process in his book "Vision: Computational Investigation into the Human Representation and Processing of Visual Information". Its core content is the method to recover the 3D shape of objects from a computer imaging. The evolution of this field can be broadly divided into 3 stages:

- Early visual technology.
- 2-dimensional half diagram (Intrinsic image).
- 3-dimensional model.

The early visual technology emphasized on the physical characteristics, such as distance, surface direction color and texture, etc.

The idea in 2-dimensional half diagram (Intrinsic image) is to reconstruct the boundary, surface and volume segmentation of an object present on the surface using description of various characteristics defined by an observer sitting in the center of the coordinate system. However, during this stage, the blocked surface could not be described like the unblocked surface, which makes it "2-dimensional half diagram".

The stage of 3-dimensional model is the highest level of image processing to date. In this stage 3-dimensional model is reconstructed for image recognition using the information extracted from 2-dimensional half diagram.

In fact, none of the above stages could meet the accuracy requirement of recognition and measurement for assorted disturbances and inverse imaging, although all of them are mathematically feasible.

In the 1980s, the 3-dimensional inspect sense technology research contributed to larger development, where geometric method and physical knowledge were used to study problems of stereo vision. Active vision was introduced during this period and the technology was combined with proximity sensing (laser ranging, microwave ranging, ultrasonic ranging, etc.) and view sensing.

In the field of detection technology, a great success and progress is achieved through study of 30 to 40 years. Considering the current level of technological development, the future trends of development of this field can be identified as:

(1) Image sensing would be merged together with other kind of sensing techniques to dig inner information, which is known as *Multi-sensor Information Fusion Technology*.

(2) More image information would be extracted using dynamic stereo perception with decreased computational difficulty using multi-viewpoint instead of traditional two-viewpoint.

(3) More Intelligent technology will be introduced, and new stereo perception research method based on knowledge, model and rule will be proposed.

(4) Recognizing and detecting technology of object and user oriented visual percep-

tion are under way.

In comparison with the natural visual perception function, the present theory and technology of machine vision are however quite imperfect and have a lots of difficulties like:

(1) Stereo matching is required in both analysis and comprehension of images, which indeed is not well developed. If the image is distorted by the introduction of grey scaled distortion, geometrical distortion, perspective distortion, rotation distortion, scaled distortion or any other noise, it would be quite difficult to recognize and match image. There is still no perfect solution for recognizing and matching images, especially when big distortion and noise are introduced. At the same time, sheltering between objects and from itself makes the object even more difficult to recognize and match.

(2) Human beings have amazing ability of analysis and comprehension of images, but this procedure remains a myth in science. Visual sampling has a large redundancy of information, but scientists still could not extract the right features to reproduce the images.

(3) The amount of computation to sense the 3-dimensional image information of the face is very large. The computational speed of current technology is still limited and cannot fulfill the real time requirements to develop a universal sense system.

1.1.3 Classification of Visual Perception System

The Visual Perception System (VPS) could be classified into monocular system, binocular system and multi-vision system. This classification is made on the basis of cameras used to construct the image sampling and processing system.

In monocular system, one camera is used as image sensor to get image. The complete perception detecting system consists of one camera, single input channel with one image acquisition card and one computer. Mathematical model of a single camera image sensor can also help in research of multi-purpose image sensor. Thus, the research of mathematical model of monocular system provides the foundation for multi-vision system.

On many occasions, monocular system alone is enough for recognizing and measuring 1-dimensional moving objects. One such example is the detection of objects on assembly line. The theoretical foundation of monocular system is formed on monocular vision geometry that recognizes morphological characteristics, measures geometrical scale with affine geometry, projective geometry and the information contained in just one image. The method is more flexible and convenient for buildings, accident scene and the detection of multiple rules for the shape, when the priori information known to the system is rich.

Recognition and detection based on one image have attracted much attention in the past years. However, this method needs systematical research. Since information contained by just one image is quite limited, we must rely on *a priori* knowledge during detecting procedure. Generally, monocular system applies to object recognition, object tracking, displacement measure etc.

Binocular system is an imitation of human eyes. It uses two image sensors of same performance to acquire two images from different but fixed angles. Then the depth of different points is computed and 3-dimensional scene is built.

In binocular system, usually two image sensors are used. The sampling devices (camera) are connected to computer by 2-input-channel image acquisition card. In a typical binocular system, the optical center of two cameras should be placed along the baseline, and two optical axes should be at an angle each other. Points in the 3-dimensional space will image on both image sensors and distance between image points is called optical parallax. It is quite easy to compute the depth if parallax is available. The computation of parallax is however a big problem, as it requires model analysis, camera calibration, image processing, feature selection, feature match, etc. The greatest strength of binocular system is that it could find the public feature points with the help of the constrain relationship of outside-epipolar line, to get the 3-dimensional coordinate of corresponding feature point.

Binocular system is popular in robot navigation, parameter detection of micromanipulation system, 3-dimensional measurement and virtual construction etc. Binocular visual sense test is also known as detection by double machine eyes.

Multi-vision system is an extension of binocular system. It uses several cameras or one camera from different viewpoints to observe 3-dimensional scene. The technology of perception, recognition and comprehension of images sampled in multi-vision system is named multi-vision system technology.

The multi-vision system is much more complicated than binocular system due to more visual sense. As the information pick up in multi-vision system much richer than that in binocular visual system, so it can be used to make up the binocular visual sense by identifying technical defects. Indeed, the vision system of human beings and insects are both instances of multi-vision system.

The theoretical foundation of multi-vision system is laid on multiple view geometry. Multiple view geometry is a recent field developed in the past ten years. It is similar to the traditional vision measurement but some significant differences exist. Tensor algebra provides a convenient representation and operation function for multi-vision system. A novel and significant work in this area can be created by studying how to introduce multiple view geometry into traditional vision measurement and how to complement the defect limitation with multiple view geometry.

To sum up the discussion we conclude by defining *visual perception* as a technology that perceives 3-dimensional information in monocular system, binocular system or multi-vision system by processing 2-dimensional image with computer.

In an image processing system, image sensor and signal processor are the hardware basis and software with algorithm is the system core. The software consists of operating system, image processing software, control software, etc. The image processing software is the key area that reflects different mathematical models for detecting objects. Fig. 1 – 3 shows the generalized flow chart of a perception system.

The block representing *Image Preprocessing* consists of algorithms of image enhancement, filtering, background fading etc. The block representing *Main Operations* depend on the desired objectives. For instance, if the main aim is to measure volume, it

Fig. 1 - 3 A signal flow diagram of the Visual Perception System

generally consists of perspective transformation, aberration correction, coordinate data transformation, corresponding point allocation and matching, 3-dimensional coordinate computation, volume computation etc. The block representing *Data Output* provides the conclusive output of recognition and comprehension e. g. it provide display of results of object's parameters and output of controlling instructions.

In short, the field of development of image processing algorithm is a very wide research area. It consists of camera calibration algorithm, image input processing, image filter, edge detection, feature extraction, image matching, depth recognition, depth map generating, and an understanding of different types of sensor signals and a priori knowledge for integration.

1.2 A Visual Perception Hardware-base

Generally, image sensor (e. g. , camera) is responsible for the acquisition of images. The image acquisition card should preprocess the images, which act as a bridge between image sensor and the host processor (an image processor like computer) . The host processor is the core arithmetic device to realize algorithms.

1.2.1 Image Sensing

Solid-state sensory devices are generally classified into three types:

- Charge Coupled Device (CCD).
- Complementary Metal-Oxide Semiconductor (CMOS), also known as Self Scanned Photodiode Array (SSPA).
- Charge Injection Device (CID).

All of them are semiconductor device, which uses photodiodes to realize the conversion between light and electricity.

1.2.1.1 Common image sensors

Currently CCD and CMOS are most widely used in image sensor all over the world.

1. *CCD image sensors*

The elementary cells of CCD are metal-oxide semiconductor. It has two parts: the photosensitive unit and shift register unit. The photosensitive unit receives the incident light of the effective frequency and generates the signal charge by stimuli photoelectric effect. A clock signal is used to transfer the charge to an analog shift register unit with the junction capacitance of the gate voltage. The signal charge transfer from the former to the latter (coupling), rely on the formation of potential wells.

In the CCD image sensor, the signal charge of the photosensitive unit is moved to the neighboring position on the shift register unit, and then in turn transmitted to the video output line through a shift.

To make sure that signal charge would be shifted along the true path, impulse voltage (clock impulse) of all the cells should meet the requirements of phase. Then the changing direction of potential well would be in a fixed direction.

Through the permutations and combinations of rules, the basic unit can be made linear or area array CCD image sensors with various specifications.

Some major advantages of CCD sensors are:

- It has a low fixed pattern noise and readout noise.
- It's video output capacitance is smaller.
- For the same pixel size and the same light-sensitive chip, CCD sensitivity and resolution is higher than CMOS.
- It is conducive to maintaining the integrity of the image information, because CCD sensor charge is concentrated in a single amplifier centrally.

Some disadvantages of CCD sensors are:

- As the CCD photosensitive surface is covered with oxide or polysilicon electrode in a multi-layer structure, it will cause greater reflection losses.
- As clock of CCD has more types, while the production process is more demanding, so its external circuit is too complicated.
- Polysilicon thickness of CCD will result in the loss to the blue and in the lack of response to ultraviolet absorption.

It must be pointed out that for CCD as the image sensor, the first output mode to be pursued is fixed and to ensure that the clear quality of video output, low readout noise is desired in the signal followed by its video output capacitor. The structural features of CCD enable to fully maintain the signal distortion during transmission. The transfer each pixel's charge is carried out centrally by a single amplifier and hence it can maintain the integrity of information. With the development of technology of semiconductor, new generation of CCD is produced, which is named CCPD (Charge Coupled Photodiode Devices) image sensor. It is structurally optimized as compare to the previous generation. It uses photodiode as light-sensitive cell, and use CCD analog shift register as readout device. These changes keep the advantages of CCD and in addition it exhibit good spectral response characteristic, photoelectrical converting characteristic as SSPA.

2. *CMOS image sensors*

CMOS image sensors are the self-scanning optical array image sensors. CMOS and CCD image sensors are different in structure. The CMOS image sensor has advantage of production process of integrated MOS semiconductor industry where all external facilities are integrated on a single chip at a time, which save the processing required in fabrication of chips hence reducing the cost of manufacturing the chip.

CMOS image sensor unit and storage unit are photosensitive photodiode. Digital shift register is the read out charge structure, which can be controlled by a group of multi-switch in order to put the charge on each photosensitive unit removed and sent to the video output cable (also known as video output bus) . In CCD image sensor, light-sensitive cells and storage cells transfer charges through surface depletion layer. Moreover, charges in each light-sensitive cells are transferred to corresponding bit of shift registers in storage cells, and then in turn transmitted to the output shift line.

The main advantages of CMOS image sensors are:

- CMOS exhibits nearly ideal properties i. e. high quantum efficiency, wide spectral response and a small dark current.
- A flexible CMOS design of shape and size can be made into the ring plane array to facilitate detection.
- CMOS anti-radiation is many times greater than the CCD.
- As the photosensitive surface is covered by a layer of transparent SiO_2, so the CMOS reflection losses and absorption losses are small.

The main disadvantages of CMOS image sensors are:

- As the photodiode readout signal is passed through a separate MOSFET switch, thus the switching noise of the gate-source capacitance is relatively large. The slight differences in switching characteristics will be generated in the dark level fixed pattern noise.
- As the video output capacitor is relatively large, its capacity is proportional to the number of the photodiode. Due to large output capacitor, not only the photodiode output signal will get weakened but the thermal noise would also become large.
- As the CMOS production process is relatively simple, it does not have an exclusive channel design. Therefore it must first amplify the signal, and then integrate the various pixels. It will therefore cause increased noise and defect information.
- Since each pixel of CMOS channel contains the amplifier and A/D converter circuit, the available light-sensitive surface area for a single pixel will be reduced too much. For the same pixel size sensitivity of CMOS is lower than that of CCD and for same size of the light-sensitive chip, resolution of CMOS is usually not as CCD.

However, if the constrain size limit is not present, CMOS technology can offer advantage on the pixel resolution and large sensor sizes. CMOS technology will overcome the manufacturing difficulties to produce large-size light-sensitive device, especially for the full-frame chip. CMOS sensor devices are already designed with 14 million pixels/full-frame design. It can be anticipated that as CMOS manufacturing process will develop further, performance of the CMOS image sensor will become better than CCD sensors in future.

CMOS line array is used as a direct image sensor and is used in equipments such as desktop scanners and bar code reading device etc. CMOS array can also be implemented along the column to implement pinhole cameras.

Traditionally, the camera implemented using CCD sensor is called CCD cameras and the camera containing CMOS sensor is commonly known as CMOS cameras. A number of operating parameters are involved in both of them. On the current technology status, the sensitivity and resolution of CCD camera is higher than that of CMOS camera. As in image processing applications it is necessary to ensure the accuracy of detection, a CCD camera image sensor is a more appropriate choice.

1.2.1.2 Physical characteristics of CCD image sensors

Physical characteristics of CCD image sensors include spectral response characteristics, photoelectric conversion characteristics, sensitivity, and uniformity, dynamic response of dark signal resolution, frequency and transfer efficiency.

1. *Spectral response characteristics*

The typical spectral response characteristic of CCD is similar to that of photodiodes. CCD spectral response can be used to indicate the relative degree of response, its theoretical formula is

$$R_\lambda = \frac{q\eta(\lambda)}{hc/\lambda} u(\lambda_g - \lambda) \tag{1-1}$$

Where q denotes electronic charge, h denotes the Planck's constant, c denotes speed of light, λ denotes wavelength, $\lambda_g = hc/E_g$ is the cut-off wavelength of intrinsic band gap E_g (Such as silicon, $\lambda_g = 1.1\mu m$), $u(\lambda_g - \lambda)$ denotes step function.

$u(\lambda_g - \lambda) = 1$ for $\lambda \leqslant \lambda_g$, $u(\lambda_g - \lambda) = 0$ for $\lambda \geqslant \lambda_g$. This means that, when $\lambda > \lambda_g$, incident photon energy is not enough to cause a transitions of electron from valence band to the conduction band and generate a charge carriers. The incident photon will not contribute to the generation of photocurrent. In other words, to generate photoelectric effect, frequency of incident light must be greater than the cut-off frequency and energy of photon must be greater than characteristic energy band gap E_g. The initial kinetic energy of photoelectrons increases with frequency of incident light and has nothing to do with intensity.

When incident light reaches the characteristic energy band gap E_g, the photoelectron emission is almost instantaneous, usually no more than 10^{-9}s. When light electron acquires maximum initial kinetic energy, the number of emitted photoelectron from characteristic energy band gap in unit time (e. g. 1 second) is proportional to the incident light intensity.

$\eta(\lambda)$ represent effective quantum efficiency. It depends on the silicon surface reflectivity, the thickness of consumption, light energy absorption coefficient of silicon materials to different wavelengths and other factors.

For example, $0.4\mu m < \lambda < 0.8\mu m$, $\eta(\lambda) \approx 0.7$. If $\lambda < 0.4\mu m$ or $\lambda > 0.8\mu m$, effective quantum efficiency decreases and spectral response also decreases. The reason is reflection. The absorption loss and interference effect caused by the structure of CCD would lead to low efficiency and poor spectral response, especially for blue ray when ultraviolet response is poor.

Fig. 1 - 4 shows two curves of spectral response characteristics. Fig. 1 - 4 (a) gives the spectral relative responsiveness from 320nm to 1100nm of Si PIN photodiodes. Fig. 1 - 4 (b) gives the spectral relative responsiveness from 900nm to 1800nm of InGaAs PIN photodiodes. It can be seen that for Si PIN photodiodes, responsiveness reaches the peak when $R_\lambda = 0.58$, and for InGaAs PIN photodiodes, the peak value is $R_\lambda = 1.49$.

Formula (1 - 1) shows that we must choose the effective quantum efficiency $\eta(\lambda)$ of products in order to ensure excellent performance CCD image sensor resistance.

2. *Photoelectric conversion characteristic*

The operation of the CCD charge storage device is based on the reverse action of photodiodes. The PN junction of photodiodes is first reversed biased to a fixed bias voltage V (usually several voltages) and then the circuit is disconnected. Therefore, the de-

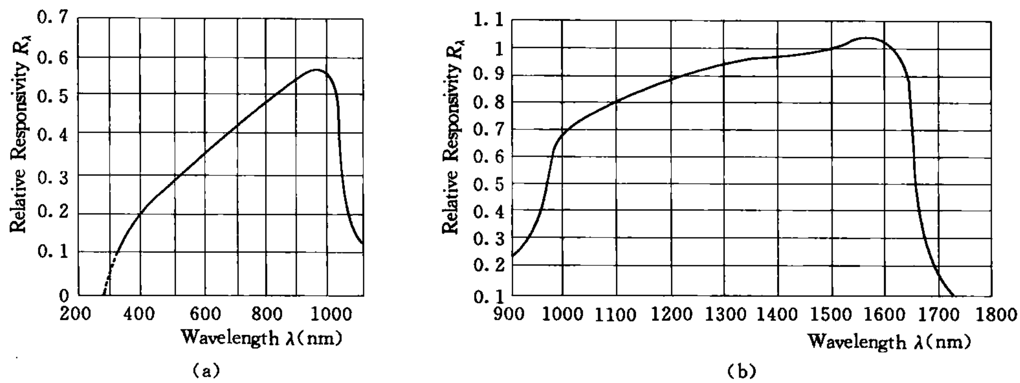


Fig. 1 - 4 Photodiode's spectral response curve

cay rate of the charge stored on the capacitor in the diode is proportional to the incident light intensity.

After some integral time T_i and diode capacitor is resumed to original bias voltage V by switching. The supplementary charge will be in proportion to the product of intensity of light and integral time of light. At this moment, the output charge Q of CCD capacitor-diode is proportional to unsaturated exposure value E equal to the product of the light intensity and light integration time.

When exposure value E increases to a specified value E_S, the output charge Q reaches to the maximum Q_m. E_S is called the saturation exposure, Q_m is called the saturation charge. The irradiation H_S that result in the generation of saturated exposure E_S is called saturation irradiance that is described by the equation:

$$E_S = H_S T_i \tag{1-2}$$

Where T_i denotes light integral time and is also known as initial impulse period.

The output charge Q is directly related to exposure value E_S. Formula (1 - 2) tells us that Q also depends on irradiance H_S and light integral time T_i. When the product of the light irradiance and the light integration time is greater, the output potential is higher from CCD photoelectric sensor.

3. *Sensitivity and inhomogeneity*

The sensitivity of CCD image sensor is a parameter that describes the photoelectric conversion characteristic between exposure value and output signal of light-sensitive device. It also indicates the photoelectric conversion in photosensitive area. According to standards proposed by International Commission on Illumination (CIE), the sensitivity of image sensor S is defined as follow:

$$S = Q_m / E_S \tag{1-3}$$

It means that under illumination of standard illuminant, the sensitivity of CCD is the ratio of saturated charge Q_m (maximum output charge) to saturated exposure value E_S in linear response region.

Ideally under uniform illumination, the image sensor produces constant output magnitude on each unit. Due to uneven production of the device semiconductor material, the image sensor output amplitude of each unit will however be uneven under uniform illumination. The uneven production of device semiconductor material is due to impurities concentrations, uneven distribution of crystal defects, mask error, process conditions and other factors. It is usually described by the non-uniformity parameter NU defined as

$$NU=\frac{V_{0max}-V_{0min}}{\overline{V}}\times 100\% \qquad (1-4)$$

Where V_{0max}、V_{0min} and $\overline{V}$ denote maximum, minimum and average value of video output voltage respectively.

The NU value changes with the light intensity. When the light intensity is close to saturation, the value of NU is small. To reduce the affects of non-uniformity, the image sensor is operated in a region near saturation of the light intensity to increase uniformity.

Heterogeneity has a relationship with the wavelength λ of light. Small λ leads to good uniformity and vice versa. This is because the greater penetration depth of near-infrared light causes a crosstalk between pixels. For general visual perception detection system, $|NU| \leqslant 10\%$ is enough. For low digit linear array, we have $|NU| \leqslant 5\%$ and for high digit linear array, we have $|NU| \leqslant 15\%$.

4. *Dark signal and dynamic response characteristic of sensors*

Some signal is present at the output of CCD image sensor even when it neither received light nor any injected power. This signal is called the dark signal which is caused by the dark current.

The root cause of dark current generation is thermal excitation of semiconductors including:

- Thermal excitation produced by recombination centers at depletion region.
- Thermal diffusion generated by small number charge carrier on the edge of depletion region.
- Thermal excitation generated in interface center.

The most important cause of dark current is the thermal excitation produced by recombination centers at depletion region. The dark current is therefore significantly influenced by the temperature, and is proportional to light integral time.

The dark current continuously joins signal charge packet and generates image of dark current signal. A fixed image noise is therefore added to optical signal image due to dark current. The presence of semiconductor defects and manufacture problems also cause high-density recombination center thus generating dark current peaks at the defective region. These defects pollute images with some bright streaks and bright spots.

Also dark current occupies the capacity of CCD potential well and narrow down device's dynamic range. To diminish the influence of dark current, we should reduce the integral and transfer time of signal charge. The operating frequency limit of CCD is therefore limited by dark current.

The dynamic response range means the ratio of output saturated signal to the dark signal, and it is decided by the quality of signal processing circuit.

5. *Resolution*

The image resolution represents capabilities to distinguish between light and dark