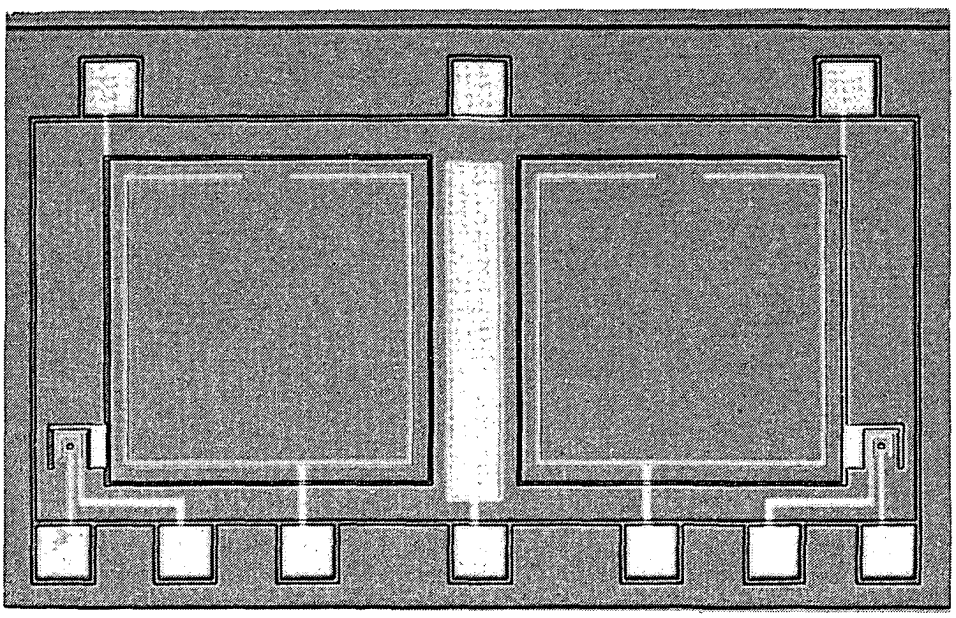
INTEGRATED SILICON COLOUR SENSORS



Reinoud Wolffenbuttel

INTEGRATED SILICON COLOUR SENSORS

INTEGRATED SILICON COLOUR SENSORS

Geintegreerde silicium kleurensensoren

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof.drs. P.A. Schenck,
in het openbaar te verdedigen ten overstaan van een commissie
aangewezen door het College van Dekanen, op
donderdag 27 oktober 1988, te 16.00 uur

door

Reinoud Felix Wolffenbuttel

elektrotechnisch ingenieur
geboren te Abcoude

Dit proefschrift is goedgekeurd door de promotor,
Prof.dr.ir. S. Middelhoek

CONTENTS.

1 INTRODUCTION.

Optical sensors and especially the wavelength dependence of the response therein have often been the subject of extensive research. Around the turn of the century experimental data on the pronounced wavelength dependence of the photoelectric effect greatly enhanced the understanding of physics, which, in turn, led to a maturing of the detector technology. In 1887, while investigating the electric discharge between two electrodes in a vacuum tube, Heinrich Herz observed an increasing electron emission at an increasing illumination of the electrodes. A year later Wilhelm Hallwachs confirmed the increasing discharge of a negatively charged zinc plate at an increasing radiant intensity of ultraviolet illumination. Shortly after Philipp Lenard discovered the effect of the electrode material work function; regardless of the intensity, beyond a certain maximum long-wavelength limit, depending on the electrode metal, no photocurrent was observed [1.1]. This experiment, which is to be found in any physics textbook [1.2], enabled Albert Einstein in 1905 to assign corpuscular properties to electromagnetic radiation and to assume that light consists of a stream of photons each carrying an amount of energy proportional to the frequency. This causes the radiation in the interaction with electrons to be able to transfer only an energy proportional to the frequency of the electromagnetic radiation. This treatment was evoked by and consistent with the conclusions drawn by Max Planck in 1900 with respect to blackbody radiation and implied a quantization of energy. This rather revolutionary concept led to a drastic change in perception among physicists and finally resulted in the quantum theory as it is known today [1.3].

During those years different kinds of radiant detectors, such as the detectors based on thermal transduction principles viz. thermopiles and pyrodetectors, entered a period of growth [1.4]. However, for the quantum detector a special role was reserved because of its short response delay and high sensitivity. By the time the basic vacuum photodetector was mature and the mere physical significance faded away, the conviction that television was likely to become a new mass medium gradually gained ground. This resulted in a continuation of the developments in a more commercially-oriented framework. The right timing contributed to the continuity in photodetector research.

Television served as the prime vehicle up to the stormy introduction of semiconductor materials as replacements of the vacuum tube in many applications. Solid-state microelectronics gave rise to new concepts of photodetection, because the underlying principles are quite different from the operating mechanisms of the vacuum tube detector. When looking back

superficially over the past 30 years of microelectronic revolution, it seems that silicon optical sensors have benefited to a far larger extent compared to sensors in other signal domains. Silicon has been the material that paves the way in microelectronic developments, because it is intrinsically easier to process with a small number of dislocations compared to compounds. The main promotors of optical silicon devices are the solar-cell and solid-state camera market and the telephone manufacturers. In the latter application area, one is primarily interested in high-speed detectors for glassfibre communication with a high sensitivity in the part of the spectrum where the glassfibre is optimally transparent. The attenuation of light in a glass fibre decreases from 2 dB/km when operating at the low-wavelength transparent window at 830 nm down to 0.25 dB/km when operating at the high-wavelength window at 1550 nm [1.5]. The silicon bandgap prevents photodetection at the high-wavelength window, whereas indium-gallium-arsenide performs quite well at such wavelengths. This property explains the current shift of interest from silicon PIN photodiodes to detectors based on III-V compounds.

The main feature currently pursued in solid-state cameras is a spatial resolution comparable to that of professional camera tubes, while maintaining an industrially acceptable yield [1.6][1.7]. This requirement can only be fulfilled in an unaltered process when reducing the dimensions of a picture element. Recent developments have resulted in CCD cameras consisting of 1280×970 picture elements with element sizes equal to 9.9×9.8 um including the CCD readout mechanism [1.7]. Basic limitations are imposed by design rules, dark current and noise. As the signal is proportional to the detector area it reduces when the dimensions are scaled down.

In detectors for optical communication purposes, silicon optical sensor developments appear to be of a declining importance, because of the obvious consequences resulting from the choice of a different material. In imaging sensors, development efforts are also not likely to contribute much to the future state-of-the-art in optical silicon sensors beyond the rather narrow scope of these imagers themselves, because the solutions found for specific problems, such as smear and blooming as well as the enormous reduction in element size realized recently, will not give rise to much spin-off in other optical silicon sensors.

For solar-cells the situation is slightly different. Although solar-cell research is sometimes referred to as a 'fashionable phenomenon' the striving for an energy-conversion efficiency close to the theoretical maximum has led to an accurate modelling of the non-idealities in the opto-electrical conversion in the siliconoxide-silicon detector system [1.8][1.9]. All efforts are aiming at the elimination of wavelength-dependent effects that would

otherwise deteriorate the response in a certain part of the spectrum and thus affect the total efficiency in a negative sense. Somewhat apart from this trend is the amorphous silicon solar-cell, which is subject of a great deal of interest and is primarily prompted by economic motives. This type of solar-cell exhibits a reduced efficiency combined with a strongly reduced price and implies that an important step towards economically accountable terrestrial application of solar energy has been made. Another development, with a high affinity to solar-cell research, concerns efforts to allow silicon photodetectors to be used for radiometric applications [1.10]. Also for these devices is a thorough understanding of all cross-effects is required and satisfactory results are obtained. This knowledge is directly applicable in related optical silicon devices.

The above brief explanation shows that a first glance at optical sensor activities give a slightly overestimated impression and a closer look at the research efforts in optical silicon sensors with respect to their individual merits in a wider perspective is required. Nevertheless, it seems indeed justified to consider the optical transduction domain to be in a convenient situation due to its history, which contrasts favourably with international research efforts performed on integrated silicon sensors in other signal domains. Therefore, the 'sensor lag' applies in a far lesser extent to optical sensors. This threat refers to an arreage in integrated silicon sensor developments compared to integrated analog and digital circuits and the extra growth opportunity the microelectronic industry would miss if this lag is not catched up with in time [1.11].

1.2 Motivation and objectives.

Despite its successful evolution the historical course of optical sensor developments can hardly be referred to as 'colourful'. The main characteristics of the radiant signal domain are the intensity and the spectral distribution. In the visible part of the spectrum the latter is usually referred to as colour. Radiant sensor research is focussed on the property of intensity leaving the spectral sensitivity as an undesirable side effect. In solar-cell research such an attitude is not surprising as the main objective is the maximizing of the energy conversion efficiency. In CCD cameras it is conceivable to apply a spectral-dependence of the response. However, in literature, the practical use of such an effect is rarely reported. Colour imaging sensors are usually composed of an array of photodiodes covered, generally in a particular checkerboard-pattern, by three different types of dyed polymer filters of either the primary or the complementary colours [1.12][1.13]. From the photocurrent detected by the three, clustered, integrated photodetectors, the colour of the incident spectrum at the respective spatial position of the cluster on the imaging surface can be reproduced unambiguously [1.14].

Despite the industrially acceptable control of the dye-deposition processes involved and the corresponding satisfactory yield, the remaining intrinsic disadvantage of having to deal with a non-semiconductor material in a silicon foundry is not merely of an aesthetic nature. Therefore, this thesis intends to break a lance for the intrinsic colour filtering properties of silicon, the material in which the basic photodetectors are integrated, as its application necessarily implies compatibility with the silicon process. Early material research on silicon has already shown beyond any doubt that the wavelength-dependence of the photodiode response is caused by a reproducible effect [1.15]. This effect is in this case due to a basic physical property in silicon; the wavelength-dependence of the penetration depth in the optical part of the spectrum due to the indirect bandgap, and can be manipulated, to a certain extent, by device design parameters such as oxide thickness and junction impurity profile. The effectiveness of those measures have already been investigated extensively for solar-cell purposes and have attributed to substantial improvements in efficiency by adjusting the structure in such a way that the impact of this wavelength-dependent material-property on the performance is minimized.

This thesis is based on the complementary starting-point and claims that a reproducible cross-effect, although up to now considered as undesirable, might reveal interesting sensor properties and is thus worthwhile investigating. Especially in the case of the wavelength-dependent response of photodiodes we are dealing with a significant effect with a magnitude comparable to the prime optical quantity, the intensity. Using the known instrumentation techniques, such as compensation and correction, a maximizing of the effect of this property is pursued to obtain photodetectors in the same silicon wafer with significantly different spectral sensitivities to allow a successful supersedence of colour dyes in colour imaging sensors. Under certain special conditions the effect of this wavelength-dependence is even electronically tunable, which would allow the disclosure of applications in which a simple one-point colour indicator would provide a cost-effective solution. This circumstance appears especially in areas where a colour imager is highly redundant, as is often the case in robotic applications [1.16].

In this thesis the available model for the opto-electrical conversion in a silicon photodiode as well as the associated mathematical background developed for solar-cells, is unscrupulously and enthousiastically adopted and applied to realize the opposite objective; optical devices featuring electronically tunable response curves with a maximum spectral separation over the tuning range of the electrical control parameter by maximizing the effect of the wavelength dependence of the absorption coefficient in silicon for visible radiation [1.17]. This category of optical sensors will loosely be referred to as 'colour sensors', which is basically incorrect, as the term colour refers to a physiological quantity and is related to human

perception rather than the more appropriate physical counterpart, the spectral distribution. No fundamental consequences should be drawn from the mixed terminology used in this thesis, because it merely implies a conformation with a present status quo and, as will become apparent when trying to rewrite the title in a more physically-accountable form, it has a much better ring.

1.3 Implementation in an electrical information processing system.

An electrical measurement system is usually composed of the three components shown in Figure 1. The input transducer selectively extracts the input signal from the non-electrical signal domain of interest and transforms this into a signal suitable for processing in the modifier [1.18].

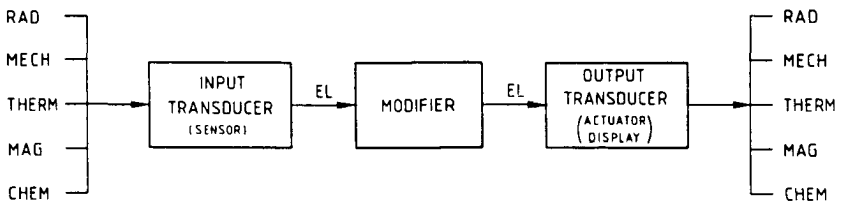


Figure 1 Functional block diagram of an electrical measurement system.

Environmental signals presented to the input transducer and supplied by the output transducer are of a radiant, mechanical, thermal, magnetic or chemical nature. Due to the widespread availability and increasing performance of electrical information processing equipment, the modifier usually operates in the electrical signal domain. The modifier circuits are usually integrated in silicon, making the advantages of silicon sensors quite obvious. Silicon sensors are compatible with read-out circuits and can therefore be integrated with these circuits on a single chip, which leads the distinction between the sensor component and the modifier unit fading away. The usual reasons for combining the sensor with the on-chip signal processing to an 'intelligent transducer' are both technical and economic. In most sensors a signal processing is required, which might involve a signal-to-noise ratio improvement technique or an enhancement of sensor characteristics such as offset or non-linearity. Moreover, the magnitude of noise and offset can be reduced by short leads and a symmetry of components with good thermal coupling, which also favours the integration on a single chip. The required signal processing is strongly sensor-type oriented and by integrating the sensor with these circuits the smallest accessable unit is now an intelligent sensor and the user, who is only concerned with the external specifications, has at his disposal of a sensor with apparently improved specifications. A similar feature, equally important for users of sensor systems, is the increasing versatility offered

by the addition of integrated electronics to equip the sensor with a standardized electrical output or, although even more demanding, a connection to a standard digital sensor bus. Sensors complying with such characteristics have a number of strong marketing advantages and at a price comparable to available discrete-component sensors, which is quite well possible in mass-production, such developments are bound to become successful in the end. A complication in integrated sensors is the possible interference of the performance of the signal-processing circuits with the nonelectrical signal. Also the process compatibility of the sensor with the circuits might impose limitations. If proper operation of an integrated sensor can only be maintained by using extra processing steps for shielding the electronics, or when serious compromises in the operation of the sensor or the circuit elements are inevitable to obtain this compatibility, it might be opportune to resort to a hybrid realisation.

In the sensor now examined, the radiant signal domain or, more precisely, the visual part of the radiant spectrum is being discussed. As there is no physical effect in silicon able to provide a predominant colour-determined electrical output, one of the prime quantities, the intensity, has to be compensated for. Somehow in the sensor the relation i=f(intensity, wavelength) has to be solved. Thus, rather than merely improving the sensor specifications by adding signal conditioning circuits, the silicon colour sensor exists by virtue of electronic circuits for the manipulation of a photodiode, which, of course, makes it even more attractive to implement the colour sensor as an 'intelligent' transducer.

Such a colour sensor can be employed successfully in numerous applications. Apart from the possible supersedence of colour dyes in solid-state imagers, which is the most demanding objective, there is a wide potential range of domestic, medical, agricultural and robotic applications where only a significant shift in the optical spectrum has to be detected. No image information is required and no tristimulus colorimetric information need be reproduced so a simple colour indicator can be used.

To conclude this introduction a few typical applications briefly pass in review starting with the automatic monitoring of a flame. At an increasing carbondioxide content and decreasing oxide concentration in the air surrounding a flame the flame colour changes from blue to yellow-reddish. The output of the colour sensor can be used to trigger an alarm when the colour becomes too red, which is associated with a certain threshold setting of the oxygen concentration [1.19]. The combination flame-colour sensor can thus be considered to be an oxygen sensor. A simple colour indicator can also serve as an aid to the visually handicapped. Using such a sensor would allow them to identify an object by its colour. An interesting application would be the selection of colour matched clothing from a wardrobe using a pocket-size integrated colour indicator.

Another application would involve the automatic inspection of the colour of car-lacquer after paint-spraying by an industrial robot, in order to equip the robot with an error indicator when the colour fails to meet the specification. An additional requirement in such an industrial application is usually a relatively high independence from environmental light intensity, which favours the application of a sensor where the intensity is already compensated for. Using such a sensor greatly enhances the versatility of the robot. A colour sensor can also be applied in a fruit-picking machine to allow automatic detection of the degree of ripeness, which is usually associated with a colour change. Again a certain threshold level can be set according to a picking criterion. This survey of applications is by no means exhaustive and is merely intended to provide an impression of the range of possibilities.

1.4 Aim and organisation.

The aim of this thesis is the realisation of different types of bipolar process compatible colour sensors in silicon based on the wavelength dependence of the absorption therein. Two different categories of colour sensors will be discussed. The first group involves devices in which a photodiode is used to obtain clearly different optical response curves by switching between several values of the reverse voltage. Under certain conditions, the colour-matching functions can be derived from these responses. These elements are designed for implementation in colour imagers and the response should, therefore, be compared with the performance of the conventional dyed colour filters.

The colour sensors in the second category are merely colour indicators in which a particular technique is used for obtaining a colour-determined output signal not affected by the optical intensity. The sensors in this group reveal interesting properties for use in the application areas mentioned above. These sensors only provide a, not necessarily unambigous, indication of colour and do not have the pretension of superseding colour dyes in solid-state imagers. In this category two different types of colour indicators are discussed.

For a proper understanding of the colour-sensing principle an extensive survey of the physical background of both the wavelength-dependent absorption in silicon and the interfering wavelength-dependencies in the $Si\text{-}SiO_2$ structure are discussed in chapter 2. In chapter 3 the basic techniques used for the colour indicators are discussed along with the colour sensors described in literature. In chapter 4 the measurements on the detectivity limits, due to dark current and noise, as well as the performance of both the basic colour element and the colour indicators are presented. Finally, the conclusions are drawn in the fifth chapter.

1.5 References.

1.1 N.H. de V. Heathcote, Nobel prize winners in physics 1901-1950, Schuman, New york, 1953, pp.34-40 and pp. 180-197.

1.2 M.Alonzo and E.J.Finn, Fundamental university physics, Vol 3, Addison-Wesley, 1980, pp.11-14.

1.3 M.Jammer, The conceptual development of quantum mechanics, McGraw-Hill, 1966, pp.28-46.

1.4 E.H.Putley, Thermal detectors, in 'Optical and infrared detectors', R.J.Keyes Edr., Springer-Verlag, Berlin, 1977, 71-100.

1.5 S.R.Forrest, Optical detectors: three contenders, IEEE Spectrum, Vol 23 No 5, May 1986, pp. 76-84.

1.6 P.K.Weimer, Image sensors for solid-state cameras, Adv. Electronics and Electron Physics, Vol 37, 1975, pp. 181-262.

1.7 H.Shiraki, Recent progress in CCD image sensors, Proc. of the 6th sensor symposium, 1986, pp.153-159.

1.8 H.J.Hovel, Solar cells, Semiconductors and Semimetals, vol 11, Eds. R.Willardson and A.Beer, Academic Press, 1975.

1.9 R.N.Hall, Silicon photovoltaic cells, Solid-state Electronics, Vol 24, 1981, pp.595-616.

1.10 J.Geist, E.F.Zalewski and A.R.Schaefer, Spectral response self-calibration and interpolation of silicon photodiodes, Appl. Optics, Vol 19 No 22, 1980, pp. 3795-3799.

1.11 S.Middelhoek, The sensor lag: a threat to the electronics industry?, Solid-state devices 1982, pp.73-95.

1.12 P.L.P.Dillon, A.T.Brault, J.R.Horak, E.Garcia, T.W.Martin and W.A. Light, Fabrication and performance of color filter arrays for solid-state imagers, IEEE Tr. on Electr. Dev., Vol ED-25 No 2, 1984, pp. 97-101.

1.13 D.Manabe, T.Ohita and Y.Shimidzu, Color filter array for IC image sensor, Proc. of the 3rd sensor symposium, 1983, pp. 103-105.

1.14 R.W.G.Hunt, The reproduction of colour, Fountain Press, New York, 3rd edn. 1975, pp.70-95.

1.15 H.R.Philipp and E.A.Taft, Optical constants of silicon in the region 1 to 10 eV, Phys. Rev., Vol 120, Oct. 1960, pp. 37-38.

1.16 P.P.L.Regtien and R.F.Wolffenbuttel, A novel solid-state colour sensor suitable for robotic applications, Proc. 5th Int. Conf. on Robot Vision and Sensory Controls, 1985, pp. 259-266.

1.17 R.F.Wolffenbuttel and P.P.L.Regtien, A novel approach to solid-state colour sensing, Sensors and Actuators, Vol 9, 1986, pp. 199-211.

1.18 S.Middelhoek, Integrated sensors, Proc. of the 3rd sensor symposium, 1983, pp.1-10.

1.19 N.Kako, N.Tanaka and C.Suzuki, Combustion detection with a semi-conductor color sensor, Sensors and Act., Vol 4, 1983, pp. 655-660.

2 WAVELENGTH DEPENDENCES OF THE OPTO-ELECTRICAL CONVERSION IN SILICON.

2.1 Introduction.

The basic colour sensor in monocrystalline silicon is composed of a boron implanted layer in an n-epitaxial layer, which is grown on a p-substrate. This simple structure, shown in figure 2.1, is identical with a conventional photodiode structure. The colour sensing is based on a controlled biasing of the photodiode rather than on a sophisticated device structure. The photodiode is alternatingly biased by different reverse voltages in such a way that a maximum distinction between the photodiode responses is obtained during the succeeding measurements by taking optimum advantage of the wavelength-dependent properties of silicon in the optical part of the spectrum. The characteristic course of the curve of the intrinsic absorption coefficient at different wavelengths is shown in figure 2.2. The indirect bandgap in silicon permits incident photons with an energy in excess of 3.5 eV (equivalent to a wavelength in vacuum smaller than 350 nm) to initiate a direct transition of electrons from the valence band to the conduction band, which results in a maximum absorption. In the case of incident radiation with an energy between the bandgap at 1.12 eV and the energy sufficient to allow a direct transition, an indirect transition could occur, provided energy and momentum are preserved. Incident photons are not able to provide a momentum, so the probability of such a transition taking place depends on lattice vibrations. The absorption increases at an increasing photon energy, since less change in momentum is required to generate an electron-hole pair at greater energies, resulting in a larger chance of an indirect transition occurring. This wavelength dependence causes a very shallow absorption of blue light and enables red light to penetrate deeply into the silicon.

In a conventional photodiode or solar-cell a photocurrent is generated that is, at a certain intensity, in first approximation proportional to the sensitive area as almost all generated charge-carriers are collected. In the colour sensor an additional mechanism is implemented that allows the collection of only the charge carriers generated in a layer extending from the surface down to an adjustable boundary, x_b. As shown in figure 2.1, the biasing with a certain reverse voltage across the shallow p^+ n junction for depleting the lighter doped epilayer down from the junction, and concurrently applying a reverse voltage across the epilayer-substrate junction for depleting the complementary part of the epilayer, permits the selective detection of only the charge carriers generated in the upper part of the silicon. As the short-wavelength components in the spectrum are

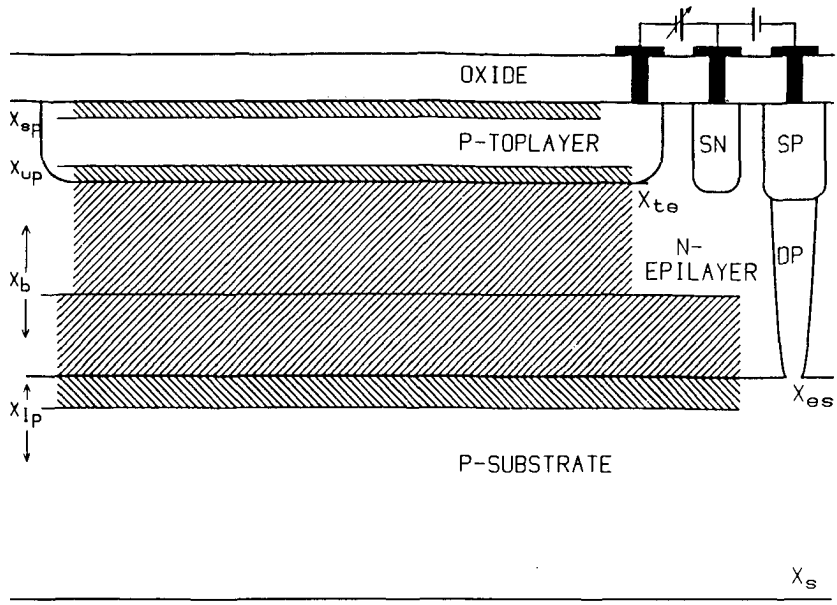


Figure 2.1 Schematic diagram of the basic colour sensor structure in silicon.

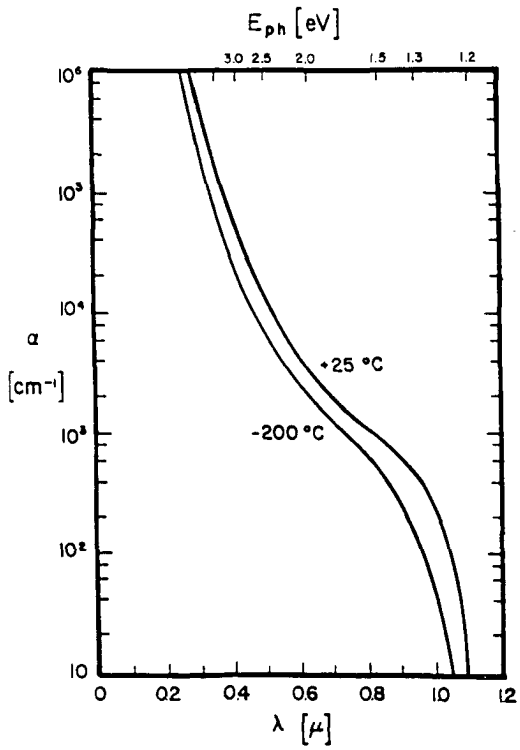


Figure 2.2 Wavelength-dependence of the absorption coefficient in silicon.

absorbed shallowly, all the blue light has already been absorbed at very thin layers. Therefore, when illuminated with light with predominantly short-wavelength components, the perceived photocurrent remains almost constant at the increasing width of this upper depleted part of the epilayer associated with an increasing reverse voltage. However, when illuminating with long-wavelength light the detected photocurrent increases with the layer width and thus with the reverse voltage. The depletion of the lower part of the epilayer prevents the existence of a neutral layer in which charge carriers generated beyond the depleted region could otherwise diffuse upwards and contribute to the photocurrent and thus avoids an impediment in the operation of the colour sensor. The total effect on the response can be represented by an electronically tunable optical filter.

The process from incident radiation on a silicon wafer covered with a thin oxide layer up to the collection of photon-generated charge carriers in these junctions can be described by a number of successive steps. In the intermediate stages parasitic wavelength dependences occur. The first wavelength dependence encountered by incident photons is that of the reflection due to the interference between the transmitted optical radiation at the air-SiO_2 interface and the radiation reflected at the SiO_2-Si interface. This interference depends largely on the thickness of the SiO_2 layer relative to the wavelength in the part of the spectrum under consideration. Moreover, the reflection coefficient at the SiO_2-Si interface is coupled to the absorption coefficient through the index of refraction and the extinction coefficient in silicon and is therefore also wavelength dependent.

The photons that are transmitted through the SiO_2 layer into the silicon are basically able to generate electron-hole pairs. The charge generation efficiency or internal quantum efficiency can be defined as the average number of electron-hole pairs generated by the absorption of one photon in the semiconductor. As will be shown, this internal quantum efficiency can be assumed unity within the part of the spectrum being considered and therefore reveals no wavelength dependence.

As has already been mentioned, the quintessence of the colour sensor has to do with the collection of generated charge carriers in a well-defined layer of silicon. To enable an estimation of the extent in which this wavelength dependence of the absorption in silicon causes a reproducible and measurable variation in the response of a reverse-biased photodiode at different reverse voltages, this chapter provides an outline of the underlying physical properties. Moreover, all the mentioned parasitic wavelength dependences are reviewed and methods for minimizing their effect are discussed.

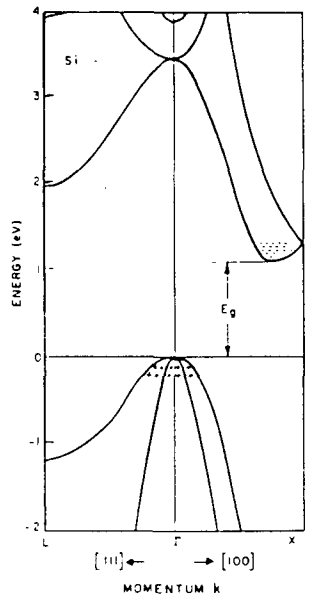


Figure 2.3 Energy band structure of silicon.

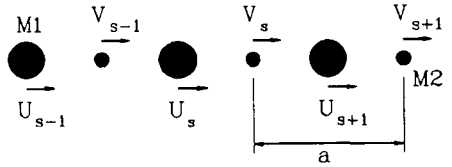


Figure 2.4 One-dimensional model used for determining the lattice vibrations.

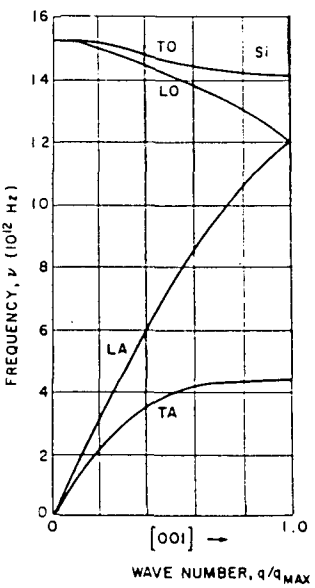


Figure 2.5 Phonon spectra in silicon.

2.2 Optical properties of silicon.

The optical properties of silicon are a direct consequence of its energy-band structure as shown in figure 2.3. A characteristic feature in this particular energy-momentum relationship or dispersion diagram is the difference between the wave vectors at the highest energy state in the valence band, k_v, and that at the lowest energy state in the conduction band, k_c. Such a dispersion diagram is typical of indirect bandgap materials such as germanium and silicon [2.1]. Incident photons carrying an energy equal to the bandgap energy, E_g, are only able to excite electrons from the valence to the conduction band at a simultaneous change in momentum equal to $k = k_c - k_v$. Such a photon carries a momentum $P = h\,k/2\pi$ and is thus able to give rise to a $\Delta k = 2\pi E_g/hc = 5.7\times10^7\ m^{-1}$. In silicon $k_v = 0$ and k_c occurs at about $0.8\,\pi/a$, where π/a denotes the first Brillouin zone edge with a as the lattice constant of the silicon diamond structure equal to $a = 5.43\times10^{10}$ m. Therefore, the required change in momentum for absorption of a photon carrying a bandgap energy, E_g in silicon is 1.12 eV, is equal to $\Delta k = 0.8.\pi/a = 5\times10^9\ m^{-1}$. This value exceeds the supplied momentum by about two orders of magnitude. The momentum supplied by the absorption of a photon can thus be disregarded and a momentum P of amount $h.k_c/2\pi$ must be supplied by the lattice. This can be achieved in two ways viz. either by absorption of a phonon of momentum $-h.k_c/2\pi$ or by emission of a phonon of momentum $+h.k_c/2\pi$.

The energy-momentum relationship of a phonon can be obtained from three-dimensional lattice considerations. The crystal structure of silicon consists of a diamond structure, which implies a tetrahedral bonding of two atoms per primitive unit cell each with a similar face-centered cubic lattice. A one-dimensional classical model of crystal vibrations with only nearest-neighbour coupling of two masses M_1 and M_2, which are placed alternatingly and representing the two atoms in a primitive cell, yields two coupled differential motion equations given by:

$$M_1 \frac{d^2 U_s}{d t^2} = C\,(V_s + V_{s-1} - 2U_s)$$

$$M_2 \frac{d^2 V_s}{d t^2} = C\,(U_{s+1} + U_s - 2V_s) \qquad (2.1)$$

where C denotes the force constant and U_s and V_s the displacements of atoms M_1 and M_2 respectively in directions shown in figure 2.4 [2.2].

Solving this set of differential equations by looking for a solution in the form $U_s = U\,exp(\,j\,k\,a)\,exp(-j\omega\,t\,)$ and $V_s = V\,exp(\,j\,k\,a\,)\,exp(-j\omega\,t\,)$ results in the dispersion relation.

$$\omega^2 = \frac{C}{M_1\,M_2}\left\{(M_1 + M_2) \pm \sqrt{(M_1 + M_2)^2 - 4\,M_1 \times M_2 sin^2(k\,a/2)}\right\} \quad (2.2)$$

From this dispersion relation it becomes evident that there are two different vibration modes viz. the acoustic mode in which the adjacent masses vibrate with a small phase difference resulting in a lattice frequency corresponding to that of sound, and an optical mode in which the adjacent masses vibrate in opposite direction resulting in a frequency in the optical range. For a three-dimensional lattice with two atoms per primitive cell there are three acoustical and three optical modes possible viz. one longitudinal acoustic mode (LA), one longitudinal optical mode (LO), two transverse acoustic modes (TA) and two transverse optical modes (TA). The typical phonon spectra measured for silicon are shown in figure 2.5. From this figure the phonon energy associated with the required change in momentum can be derived for all vibration modes. In silicon both optical and acoustic phonons make an appreciable contribution to the absorption and the corresponding phonon energies, E_p, can be obtained from this dispersion relation at $k = 0.8\,\pi\,/a$.

As the absorption in an indirect bandgap material, such as silicon, is phonon-controlled, the minimum frequency, ν, of the incident radiation for which such a transition might occur is now given by $h\,\nu = E_g - E_p$ for absorption of a phonon and $h\,\nu = E_g + E_p$ for emission of a phonon. Since E_p is generally quite small compared to E_g all vibration modes will affect the shape of the absorption curve for $h\,\nu$ near E_g, the absorption edge. The expected shape of this curve due to absorption or emission of one type of phonon can be derived from relatively simple theoretical concepts. The final course of the absorption coefficient near the absorption edge due to the combined effect of all types of phonons is a superposition of the individual effects.

The transition probability between a state near the top of the valence band, associated with a wave vector $k = k'$ and a state near the bottom of the conduction band with $k = k_c - k''$ will not differ much from the transition probability between the top of the valence band to the bottom of the conduction band if k', $k'' << k_c$, and can, in a first approximation be taken as constant near the absorption edge. Taking this assumption into consideration makes the absorption coefficient dependent only on the density of states from which the transitions can take place and on the probability of emission and absorption of phonons.

The conservation of energy yield: $h\nu = E_c(k_c) - E_v(k_v) = E_g \pm E_p + E + E'$ where E' denotes the energy of an electron in the initial state in the valence band and E the energy in the final state in the conduction band as shown in figure 2.6. Starting from a certain fixed state in the valence band with a value E' and incident radiation with a frequency within the interval ν and $\nu + d\nu$ results in an attainable interval of energy states in the conduction band of size $dE = h\,d\nu$. The number of states within this interval is equal to:

$$N_c(E)\,dE = b_1\sqrt{E}\,dE = b_1\sqrt{h\nu - E_g \mp E_p - E'}\,dE, \qquad (2.3)$$

where b_1 is constant. Similarly, the interval of energy states in the valence band excitable to a fixed energy state in the conduction band, E, by incident radiation of frequency between ν and $\nu + d\nu$ is equal to:

$$N_v(E')\,dE' = b_2\sqrt{E'}\,dE' \qquad (2.4)$$

To find the total number of possible transition pairs an integration over the number of states in the valence band from which energy level E in the conduction band can be reached must be performed. The upper integration limit E_m' of E' is given by $E_m' = h\nu \mp E_g + E_p$, which implies two different expressions for the total number of transition pairs and thus also leads to two different contributions to the absorption coefficient associated to absorption (+) and emission (-) of a phonon.

The number of transition pairs is thus given by the equation:

$$N(\nu)\,d\nu = h\,d\nu \int_0^{E_m'} N_c(E)N_v(E')\,dE' = b_1 b_2 h\,d\nu \int_0^{E_m'} \sqrt{(E_m - E')E'}\,dE' = b_3 E_m'^{\,2}\,d\nu \qquad (2.5)$$

The optical absorption coefficient due to absorption of a phonon is thus proportional to both $(h\nu - E_g + E_p)^2$ and the available number of phonons N_p of energy E_p, which is described by a Bose-Einstein distribution given by $N_p = 1/(\exp(E_p/kT) - 1)$. Therefore, the contribution by absorption of a phonon of momentum k_c, corresponding to a phonon energy E_p according to the phonon dispersion relation, to the absorption coefficient is equal to:

$$\alpha_a = A\,N(\nu)\,N_p = \frac{A(h\nu - E_g + E_p)^2}{\exp(E_p/kT) - 1} \qquad h\nu > E_g - E_p$$

$$= 0 \qquad h\nu \le E_g - E_p, \qquad (2.6)$$

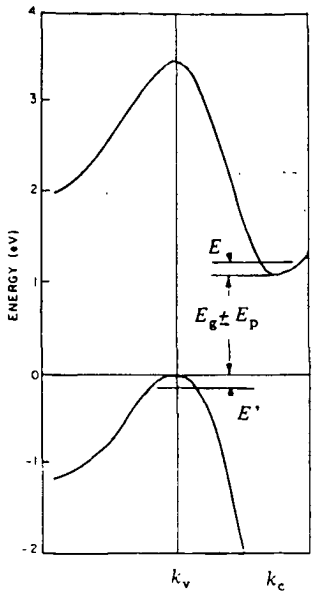


Figure 2.6 Indirect transition across the silicon bandgap.

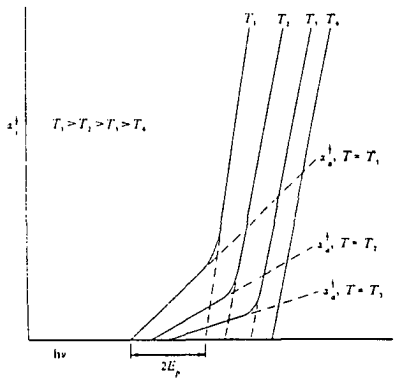


Figure 2.7 The two components of the absorption coefficient near the bandgap and the temperature-dependence thereof.

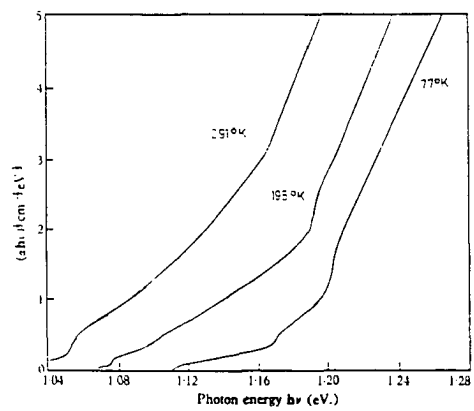


Figure 2.8 Measured absorption coefficient near the silicon bandgap.

where A is a slowly varying function of ν and is usually taken as constant. Similarly, the absorption coefficient due to emission of a phonon is proportional to $(h\nu - E_g - E_p)^2$. Inserting the ratio of the probabilities for emission or absorption of a phonon, which is equal to $(N_p + 1)/N_p$ results in:

$$\alpha_e = A\,N(\nu)\,(1 + N_p) = \frac{A\,(h\nu - E_g - E_p)^2}{1 - exp(-E_p/kT)} \qquad h\nu > E_g + E_p$$

$$= 0 \qquad h\nu \leq E_g + E_p \qquad (2.7)$$

Therefore, for $h\nu > E_g + E_p$ the total absorption coefficient is proportional to the sum of the two individual contributions; for $E_g - E_p < h\nu \leq E_g + E_p$ only the absorption of a phonon has to be taken into account and for $h\nu \leq E_g - E_p$ no absorption of the incident radiation occurs. The effect of the two individual responses becomes apparent when plotting $\sqrt{\alpha}$ against $h\nu$ as shown in figure 2.7. At a certain temperature $\sqrt{\alpha}$ vs. $h\nu$ reveals a straight line intersecting the $h\nu$ axis at $h\nu = E_g - E_p$ and with a slope proportional to $1/\sqrt{(\exp(E_p/kT) - 1)}$, which tends to zero at very low temperatures. The contribution from α_a therefore also tends to zero at a very low temperature. On the other hand, if $\sqrt{\alpha_e}$ is plotted against $h\nu$ straight lines are observed having a slope proportional to $1/\sqrt{(1 - \exp(-E_p/kT))}$. The slope tends to increase only slightly with increasing temperature at very low temperatures and intersects the $h\nu$ axis at $h\nu = E_g + E_p$. In early measurements on the absorption coefficient near the bandgap wavelength a curve-fitting technique was generally applied based on only one phonon energy. This resulted in a dominant E_p at 0.0515 eV. The value of the coefficient A can be found experimentally at very low temperatures, where the contribution of α_a can be disregarded, and was established at $5.0\times10^3 / E_g^2$ cm^{-1} ev^{-2}.

As has already been mentioned both optical and acoustic modes have to be taken into account, which results in an absorption coefficient near the absorption edge composed of 8 components ((TA + LA + TO + LO)× (phonon abs + phonon emission)). A more accurate modelling needs to include the multi-phonon energies. More precise measurements on silicon allowed a curve fitting in which the energies of all four phonon modes corresponding to $k=k_c$ could be derived and are found to agree quite accurately with the required values already obtained from the dispersion relation. Therefore, optical absorption measurements in the range between 1 eV and 1.3 eV, as shown in figure 2.8, indeed confirm a course of the absorption coefficient in conformation with the theory.

For incident radiation with a sufficiently high frequency to satisfy $h\nu >> E_g$ a more complicated situation appears. As the curvature of the conduction band for $(k - k_c)/k_c < 1$ can be approximated by $E(k) = E(k_c) + \left(h^2/(8\pi^2 m_e)\right)(k - k_c)^2$ an excess energy equal to $\Delta E = E(k) - E(k_c) = h\nu - E_g$ results in a decreasing momentum required for transitions from the top of the valence band to the conduction band. The reduction of the associated minimum wave-vector is proportional to $k_1 = k_c - k = (2\pi/h)\sqrt{2 m_e \Delta E'}$, where m_e denotes the effective electron mass. Similar reasoning applies for the holes in the filled energy states underneath the top of the valence band, which are, at an exccess radiant energy, also liable to make a transition to the bottom of the conduction band. The curvature of the valence band in silicon can also be expressed by a hyperbolic relation and is equal to $E(k) = E(k_v) - \left(h^2/(8\pi^2 m_h)\right) k^2$. The occurrence of these transitions also implies a reduction of the minimum wave-vector required, $k_2 = (2\pi/h)\sqrt{2 m_h \Delta E}$. Using the dispersion relations for the four different phonon modes in silicon, shown in figure 2.5, a band of phonon energies can be derived, which are associated with the band of phonon momenta appearing directly at the left of $k = 0.8\,\pi/a$. This implies an increase in the number of available phonons and thus in an increasing transition probability. Therefore, this very crude approximation shows that, at an increasing incident radiant energy exceeding the bandgap, both the number of phonons with sufficient momentum to support the transition and the number of participating energy states increases, which implies a strongly increasing absorption coefficient.

At a radiant energy beyond 3.4 eV vertical transitions can occur. Such a transition, requiring no change in momentum, is typical for a direct bandgap material and only appears in indirect bandgap materials at radiant energy levels significantly exceeding the absorption edge. This type of transition is not phonon-controlled so only the photon energy has to be taken into account when deriving an expression for the absorption coefficient in this part of the spectrum. The required photon energy for allowing such a transition is equal to $h\nu = E_c(k_v) - E_v(k_v) = E_c(0) - E_v(0)$ as determined by the conservation of energy. Rather like the case of indirect transitions, the interval of energy states able to perform a direct transition is related to the interval of radiant energy in between ν and $\nu + d\nu$ by the equation $dE' = h\,d\nu$, which results in the number of energy states in the interval between E' and E'+ dE' being equal to $N(E')\,dE' = A\sqrt{E'}\,dE'$. Quite different, though, is the expression for the transition probability, which remains constant and maximum due to the independence from the number of phonons present. As the absorption coefficient beyond this frequency is thus proportional to the number of states it can be described by:

$$\alpha_d = B\sqrt{h\nu - \left(E_c(0) - E_v(0)\right)} \qquad h\nu \geq E_c(0) - E_v(0), \quad (2.8)$$

where B is constant. The overall curve of the intrinsic absorption coefficient in silicon, being a combination of the mentioned absorption mechanisms, has already been depicted in figure 2.2 and clearly confirms the onset of indirect transitions at about 1.1 eV for room temperature, whereas the bending at about 2.5 eV indicates the threshold of direct transitions.

Another absoption mechanism in doped silicon devices, which could affect the absorption of incident optical radiation, arises from transitions within a band. This "free-carrier" absorption does not attribute to a photocurrent. However, it reduces the optical intensity at high-wavelength radiation by absorbing photons before being able to give rise to an indirect transition. Obviously, in a practical silicon quantum detector a minimizing of the influence of this absorption mechanism has to be pursued. As the free-carrier absorption is also phonon-controlled (the simultaneous absorption of a photon and absorption or emission of a phonon is involved) it can be treated in a similar way as in the indirect transitions and the absorption at room temperature is given for n-type material by:

$$\alpha_f = 2\times10^{-18}\ \lambda^2\ n \qquad [cm^{-1}], \quad (2.9)$$

where n denotes the free-carrier concentration as determined by the doping concentration and λ the free-space wavelength [2.3]. Interesting

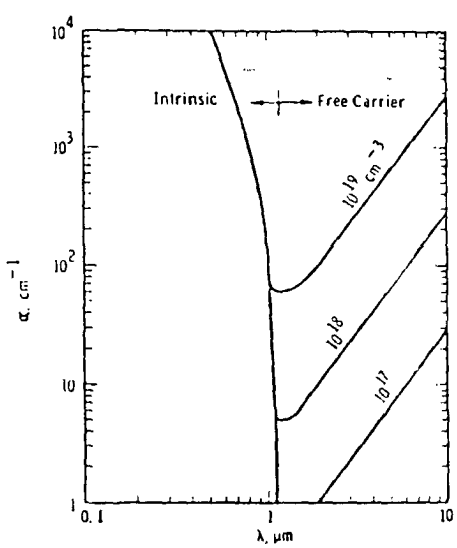


Figure 2.9, Free-carrier absorption in silicon.

properties of this free-carrier absorption are the linear proportionality on the concentration and the square dependence on the wavelength of incident radiation. These properties determine the part of the spectrum in which this effect becomes noticable. As shown in figure 2.9 this effect determines the infrared transmission of silicon, but does not seriously affect the performance of the detector in the visible part of the spectrum [2.4].

In this discussion all mechanisms affecting the shape of the absorption coefficient are mentioned. However the bandgap in silicon also affects the value of other optical properties such as the refractive index. The frequency dependence of the real and imaginary part of the refractive index near the bandgap can be described reasonably accurate by using Lorenz's simple classical treatment, which assumes the silicon lattice to consist of a system of oscillators which are set in forced vibration by the incident radiation. The electric field impressed by the radiation is equal to $-eE\exp(j\omega t)$ and results in a displacement x of an electron which in turn causes a restoring force equal to $m\omega_o^2 x$ [2.5]. Then the differential equation describing the motion of an electron is

$$m\frac{d^2x}{dt^2} + mxf\frac{dx}{dt} + m\omega_o^2 x = -eE\exp(j\omega t), \qquad (2.10)$$

where f denotes the damping coefficient, ω_o the resonance frequency positioned at the direct transition bandgap and m the classical electron mass. Solving this differential equation by looking for a solution in the form $x = x_o \exp(jwt)$ results in an amplitude given by:

$$x_o = \frac{-(e/m)E}{\omega_o^2 - \omega^2 + jf\omega_o} \qquad (2.11)$$

This amplitude, x_o, gives rise to a polarization, P, equal to $P = -Nex_o$ where N denotes the electron density. Therefore, this equation can be rewritten as

$$\frac{P}{E} = \frac{N(e^2/m)}{\omega_o^2 - \omega^2 + jf\omega} \qquad (2.12)$$

From the Maxwell equations the dielectric displacement can be derived as $D = \epsilon_o\epsilon^* E = \epsilon_o E + P$, which gives $\epsilon^* = 1 + P/\epsilon_o E$. Hence

$$\epsilon^* = 1 + \frac{N e^2 / (m \epsilon_o)}{\omega_o^2 - \omega^2 + j f \omega} \qquad (2.13)$$

The relation between the permittivity, ϵ^*, and the refractive index, n^*, can be obtained by solving the Maxwell equation for the electric component of the incident electromagnetic field. Considering only one direction of the electric field leads to the differential equation given by:

$$\frac{d^2 E}{d x^2} - \sigma \mu_o \frac{d E}{d t} - \mu_o \epsilon \epsilon_o \frac{d^2 E}{d t^2} = 0 \qquad (2.14)$$

This yiels for the y-component of the electric field:

$$E_y = E \exp(j\omega (t - n^* x / c) \qquad (2.15)$$

This solution satisfies the differential equation 2.14 provided that $(n^*)^2 = c^2(\epsilon \epsilon_o \mu_o - j \sigma \mu_o / \omega)$. Since $c^2 \mu_o \epsilon_o = 1$ this yields $(n^*)^2 = (n - jn')^2 = \epsilon - j \sigma / \omega \epsilon_o = \epsilon^*$. Separating real and imaginary parts gives $n - n' = \epsilon$ and $2 n n' = \sigma / \omega \epsilon_o$. Insertion in equation 2.13 yields the dispersion relations

$$\left(n^2 - n'^2\right) - 1 = \frac{N e}{m \epsilon_o}^2 \frac{\omega_o - \omega^2}{(\omega_o - \omega^2)^2 + (f \omega)^2}$$

$$2 nn' = \frac{N e}{m \epsilon_o}^2 \frac{f \omega}{(\omega_o - \omega^2)^2 + (f \omega)^2} \qquad (2.16)$$

Although this approach does not take into account the indirect bandgap it demonstrates reasonably accurately the wavelength dependence of the refractive index. Also a direct relation between the absorption coefficient and the imaginary part of the refractive index, the extinction coefficient, can be derived by interpreting the absorption coefficient as the reciprocal value of the penetration depth, which refers to the depth after which the energy of the incident radiation has fallen to exp(- 1) of its surface value. The radiant energy is proportional to the product of the amplitudes of the electric and magnetic field vectors. As the magnetic field can be described by a similar expression as the electric field, according to the Maxwell

equations, the penetration depth can also be defined as the depth at which the amplitude of the electric field has fallen to exp(-0.5) of its original value. Insertion in equation 2.15 yields $\exp(j\omega n^* / (\alpha c)) = \exp(-0.5)$. Hence $\alpha = 4 \pi n' / \lambda$. Figure 2.10 shows the measurements of the refractive index in silicon reported in literature. These results fit the equation 2.16 reasonably well after selection of $\omega_0 / f = 3/2$ and a refractive index at a wavelength well above the absorption edge equal to $n_0 = 3.45$.

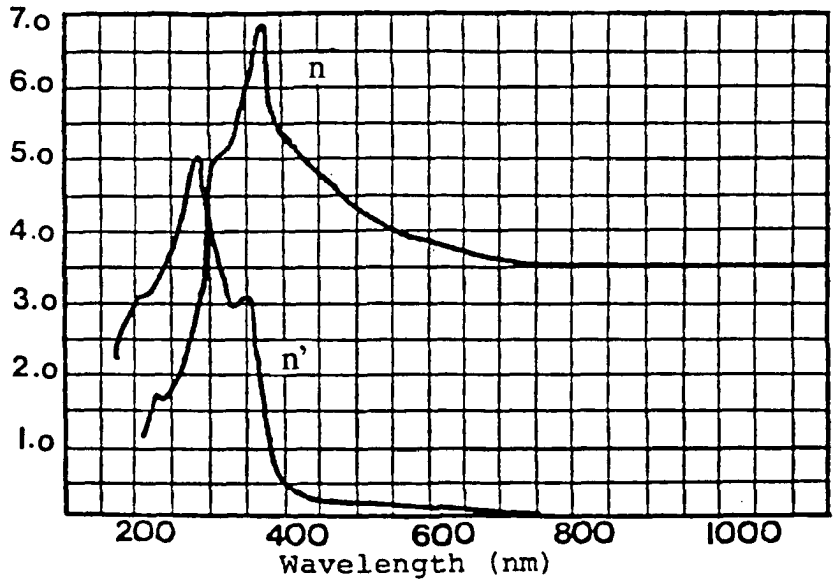


Figure 2.10 Real, n, and imaginary, n', part of the index of refraction.

The optical constants in silicon, α and n^*, described in this section are the prime optical parameters and their wavelength dependences are closely related to the magnitude and shape of the bandgap of the material. The operation of the colour sensor is dependent on the reproducibility and measurability of these parameters and to the extent to which parasitic wavelength dependences can be surpressed. In this context it has to be emphasised that the course of the absorption coefficient establishes the colour sensing operation rather than a direct optical property of a transmission medium. Therefore, any cross-effects changing the value of the bandgap energy cannot be distinguished from a spectral change in the incident radiation and thus have to be avoided. Some of these detrimental effects arise from the temperature and the pressure dependence of the bandgap, -2.8×10^{-4} eV / K and -1.5×10^{-6} eV / bar respectively, and the

bandgap narrowing due to high-doping effects. The latter can be disregarded when realizing moderately-doped, $N < 10^{18}$ cm^{-3}, p-type shallow layers in the basic device shown in figure 2.1. As will be shown in the following section this measure also optimises the collection efficiency for short-wavelength visible radiation. The pressure dependence can usually be disregarded, yet the temperature dependence limits the reproducibility and thus imposes the limitation on reliable operation within a specified temperature range.

2.3 Optical properties of the air-SiO_2-Si system.

The optical properties of the bulk single-crystal silicon are usually described by its primary parameters such as the absorption coefficient and its index of refraction. In an actual device, such as the colour sensor in question, a more observation-oriented terminology in derivative parameters is usually preferred. The main optical parameter in this category is the reflection coefficient. Rather than describing the optical performance of a material this parameter expresses the optical consequences of the interface between two different media and can be expressed in the prime optical parameters of the respective media using the Fresnel equations. The reflection and refraction of a plane incident wave, travelling from a medium having a complex refactive index n_1^*, on the interface with a medium having a refractive index n_2^* can be found after resolving the electric and magnetic components, associated with this incident electromagnetic wave, into their normal and parallel components with respect to the plane of incidence as shown in figure 2.11.

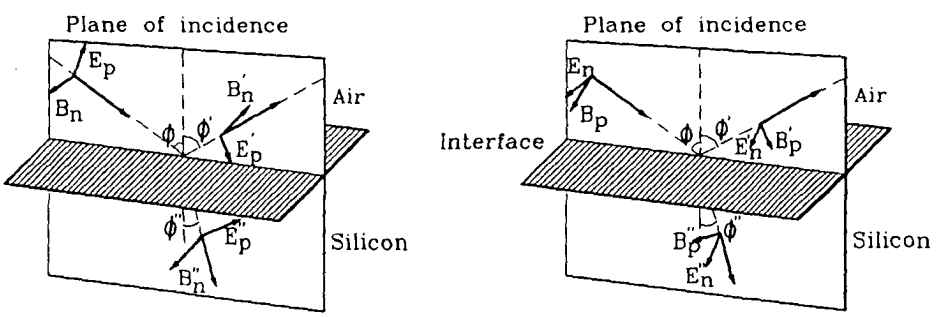


Figure 2.11 Normal and perpendicular components of incident, refracted and reflected electric and magnetic fields at an air-silicon interface.

Using Snell's laws and applying the boundary conditions at the interface, which imply that the tangential components of the electric and magnetic field reveal no discontinuities when passing through the plane x=0, Fresnell's laws on the amplitudes of the reflected and refracted waves can be found. Applying this to the electric field yields [2.5]:

$E_p \cos\phi + E_p' \cos\phi' = E_p'' \cos\phi''$, hence $(E_p - E_p')\cos\phi = E_p'' \cos\phi''$

and

$$E_n + E_n' = E_n'' \qquad (2.17)$$

From the Maxwell equations the relation between the amplitude of the magnetic field, H_p, and the electric field, E_n, in a medium can easily be found to be equal to $H_p = (n^*/\mu_o c) E_n$. Therefore, similar equations can be obtained from the boundary conditions imposed on the magnetic field:

$$(E_n - E_n')\, n_1^* \cos\phi = E_n''\, n_2^* \cos\phi'' \text{ and } (E_p + E_p')\, n_1^* = E_p''\, n_2^* \qquad (2.18)$$

Solving these four equations yields the amplitudes of the reflected and refracted waves.

$$E_p' = E_p \frac{n_2^* \cos\phi - n_1^* \cos\phi''}{n_2^* \cos\phi + n_1^* \cos\phi''} = E_p \frac{\tan(\phi - \phi'')}{\tan(\phi + \phi'')}$$

$$E_n' = E_n \frac{n_1^* \cos\phi - n_2^* \cos\phi''}{n_1^* \cos\phi + n_2^* \cos\phi''} = E_n \frac{\sin(\phi'' - \phi)}{\sin(\phi + \phi'')}$$

$$E_p'' = E_p \frac{2 n_1^* \cos\phi}{n_1^* \cos\phi'' + n_2^* \cos\phi}$$

$$E_p'' = E_p \frac{2 n_1^* \cos\phi}{n_2^* \cos\phi'' + n_1^* \cos\phi} \qquad (2.19)$$

Hence, the reflection coefficient for both the parallel and normal electric components of incident radiation are equal to

$$R_p = \left(\frac{E_p'}{E_p}\right)^2 = \left(\frac{\tan(\phi - \phi'')}{\tan(\phi + \phi'')}\right)^2 \text{ and } R_n = \left(\frac{E_n'}{E_n}\right)^2 = \left(\frac{\sin(\phi'' + \phi)}{\sin(\phi + \phi'')}\right)^2 \qquad (2.20)$$

A general expression for reflection at the silicon surface at normal incidence can now be derived. Insertion of the complex index of refraction $n_2^*(\lambda) = n_2(\lambda) - j\,n_2'(\lambda)$ yields

$$R = \frac{E_p'}{E_p}\,\frac{E_p'}{E_p} = \frac{\left(1-n_2(\lambda)+jn_2'(\lambda)\right)\left(1-n_2(\lambda)-jn_2'(\lambda)\right)}{\left(1+n_2(\lambda)-jn_2'(\lambda)\right)\left(1+n_2(\lambda)+jn_2'(\lambda)\right)} = \frac{(n_2(\lambda)-1)^2+n_2'(\lambda)^2}{(n_2(\lambda)+1)^2+n_2'(\lambda)^2} \qquad (2.21)$$

Therefore, the wavelength dependence of the reflection coefficient in an air-Si system as shown by curve a in figure 2.12 is entirely determined by the wavelength dependence of the refractive index.

In an actual device such a neat air-Si interface is usually not available. Even the exposure of a sample to an environment outside the high-vacuum for a short period of time results in a native oxide layer of about 1-2 nm thickness. In the early reflectance studies in silicon performed to obtain absorption data observers were unaware of its presence and thus a minor correction of these data is required [2.6]. As will be shown this native oxide layer width exerts an appreciable effect on the reflection of radiation at energy levels beyond 4 eV, which is well into the ultraviolet range and is thus not of particular importance for the colour sensor. However, the colour sensor is manufactured in a standard bipolar process in which oxide layers act as masking and passivation layers. In the finished device there will be an oxide layer on top of a photodiode with such a thickness that the reflection in the visible part of the spectrum is appreciably affected by its presence. For this reason the transmission of radiation through an air-SiO_2-Si system has to be discussed.

In a bulk material covered by a thin film, multiple reflections will occur, which will give rise to constructive or destructive interference at certain wavelengths. The oxide layer is assumed to be non-absorbing with a thickness d_{ox} and a real index of refraction equal to $n_1 = 1.46$. The net transmission of radiation into the bulk silicon can be derived using the characteristic matrix for normal incidence [2.7].

$$\begin{bmatrix} E \\ H \end{bmatrix}_{surf.} = M \begin{bmatrix} E \\ H \end{bmatrix}_{sil.} = \begin{bmatrix} \cos\delta & -j\sin\delta/n_1 \\ -jn_1\sin\delta & \cos\delta \end{bmatrix} \mu_o c \begin{bmatrix} 1 \\ n_2^*(\lambda) \end{bmatrix} = \mu_o c \begin{bmatrix} \cos\delta - j(n_2^*(\lambda)/n_1)\sin\delta \\ n_2^*(\lambda)\cos\delta - j\,n_1\sin\delta \end{bmatrix} \qquad (2.22)$$

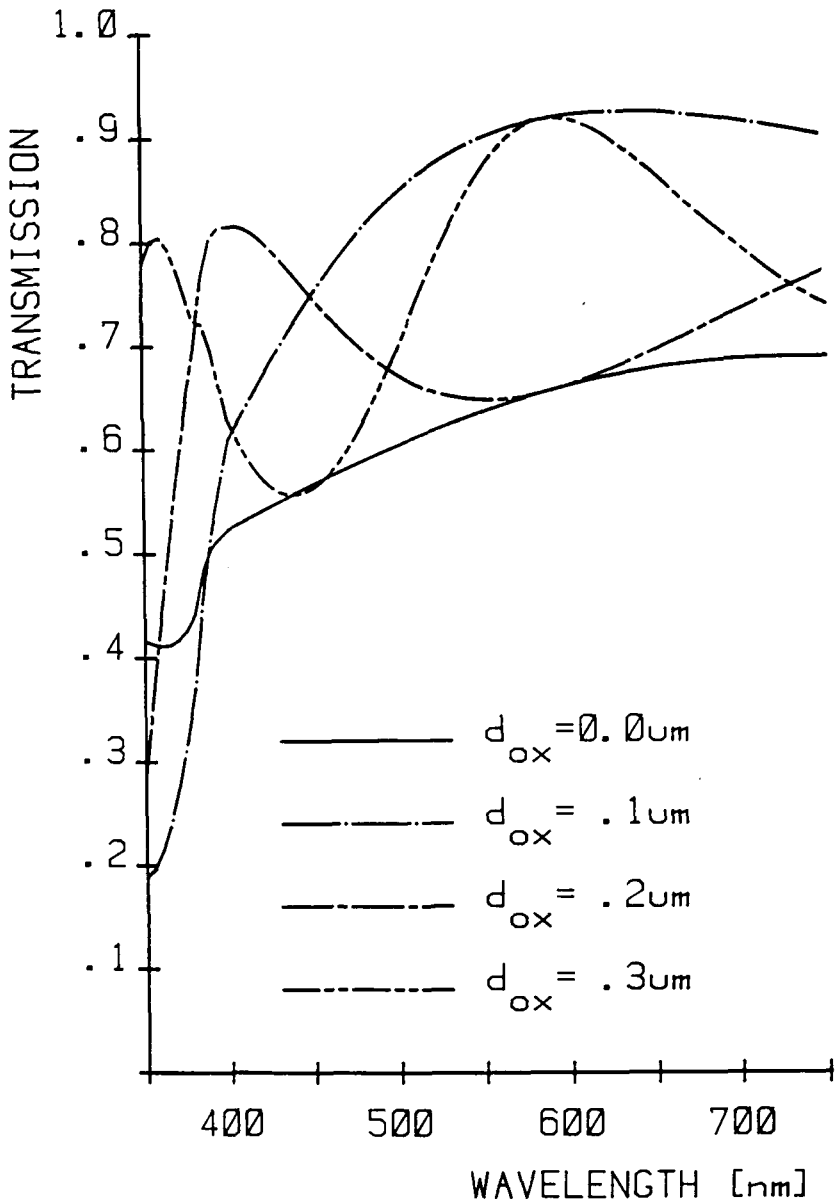


Figure 2.12 Wavelength dependence of the transmission of optical radiation in the visible part of the spectrum, calculated for an air-silicon interface (a) and for an air-silicondioxide-silicon system for three different values of the oxide thickness (b)-(d).

Where $\delta = (2\pi / \lambda) n_1 . d_{ox} \cos \phi''$ denotes the phase change on one traversal through the film and M the characteristic matrix. Applying the relation between the amplitude of the reflection and the admittance, H/E, yields:

$$r = \frac{\mu_o c (H / E) - 1}{\mu_o c (H / E) + 1} = \frac{(n_2^*(\lambda) - 1) \cos\delta - j \left[n_1 - (n_2^*(\lambda)/n_1) \right] \sin\delta}{(n_2^*(\lambda) + 1) \cos\delta - j \left[n_1 + (n_2^*(\lambda)/n_1) \right] \sin\delta} \qquad (2.23)$$

Hence after insertion of $n_2^*(\lambda) = n_2(\lambda) - j\, n_2'(\lambda)$,

$$R = \frac{\left[(n_2(\lambda)-1)\cos\delta + (n_2'(\lambda)/n_1)\sin\delta \right]^2 + \left[(n_2(\lambda)/n_1 - n_1)\sin\delta - n_2'(\lambda)\cos\delta \right]^2}{\left[(n_2(\lambda)+1)\cos\delta - (n_2'(\lambda)/n_1)\sin\delta \right]^2 + \left[(n_1 + n_2(\lambda))/n_1)\sin\delta + n_2'(\lambda)\cos\delta \right]^2} \qquad (2.24)$$

This rather unwieldy expression does not only describe the wavelength dependence of the reflection due to interference between incident radiation and the once or multiple times reflected radiation in the oxide but also reveals the complication due to wavelength dependence of the refractive index in silicon. The advantage of the characteristic matrix approach is its general applicability and simple expansion to a multi-film structure such as an air-SiO_2-polySi-SiO_2-Si system using the matrix multiplication of three different characteristic matrices. The wavelength dependence of the index of refraction makes computer analysis of equation 2.24 mandatory. Figure 2.12 shows the transmission vs. the wavelength for normal radiation at three different values of the oxide thickness.

An important property of the pure air-silicon interface is the independence of the net reflection of radiant energy on the angle of incidence up to high angles of incidence. Assuming the wavelength of the incident visible light to be larger than 500 nm allows the extinction coefficient to be disregarded so that the refractive index is real and equal to $n_2 = 3.42$. If unpolarised radiation falls on the air-silicon interface the net reflection will be $R = 0.5 \times (R_p + R_n)$. Applying equation 2.22 gives a reflectance remaining remarkably constant up to the Brewster angle at about 73°. However, in the air-oxide-silicon system the phase change on one traversal through the oxide film also depends on the angle of refraction, ϕ'', resulting in an extra wavelength dependence originating from a possible non-normal incidence, $\phi \neq 0$. Using Snell's law and the expression for the phase change, δ, results in an apparent increase in the oxide thickness of less than 3% at an angle of incidence exceeding 20°, which is well within the tolerances.

The transmitted radiant energy with an energy beyond the indirect bandgap is basically able to generate electron-hole pairs. The figure of merit on the efficiency of this generation process is the internal quantum efficiency, η, which can be defined as the average number of electron-hole pairs generated by one photon entering a quantum detector with a finite thickness and lifetime of excess minority charge carriers. Any detrimental effect due to optical causes, such as the reflection, or electrical causes, such as non-ideal collection efficiencies, are accounted for in the external quantum efficiency. The latter effect is due to limitations in the electrical performance of the opto-electrical conversion and will be reported in the next section.

Measurements reported in literature on the internal quantum efficiency are quite uniform in their conclusions with respect to the visible part of the spectrum, however some discrepances are noticeable for higher energy levels [2.8]. The measurements demonstrate a constant internal quantum efficiency at radiant energy levels between 1.2 eV and 3.3 eV equal to $1.05 \pm 8\%$. As impact ionisation is not possible for radiant energies in between the indirect and direct bandgap the actual value is generally believed to be equal to 1.0, which is within the specified accuracy range. At energies exceeding 3.5 eV impact ionisation causes the occurrence of internal quantum efficiencies larger than unity. In this process an electron-hole pair is initially generated whereby the photoelectron is excited from valence to conduction band and where the photohole stands a reasonable chance, depending on the phonon scattering, on transferring its surplus kinetic energy to another electron thereby exciting the latter to the

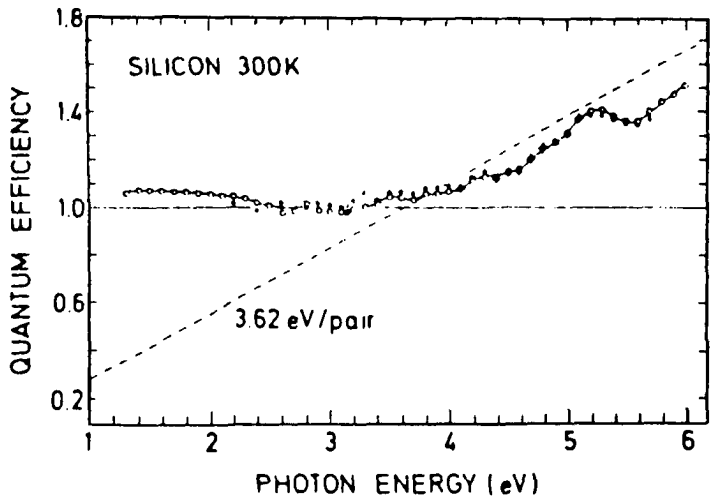


Figure 2.13 Results of quantum efficiency measurements [2.8[.

conduction band. Results of measurements are shown in figure 2.13 and clearly reveal the linear increase beyond 3.6 eV and the plateau from about 4.7 eV to 5.6 eV in which 4.7 eV corresponds to 3.6 + 1.1 eV, being the indirect + direct bandgap energy. In the visible part of the spectrum such an effect does not occur and the internal quantum efficiency of the colour sensor is considered wavelength-independent and equal to unity.

2.4 Properties affecting the collection efficiency in the colour sensor.

The final stage in the opto-electrical conversion process concerns the collection of generated charge carriers. This part offers the most pronounced point of application for flexibly influencing the conversion process, as the control of the oxide thickness is restricted by technological considerations and therefore exerts only a marginal contribution to the programmability of the response. The basic device structure, characteristic for a silicon colour sensor, consists of a layered doping profile as shown in figure 2.14. This figure also indicates the different current components and will, therefore, be used as a reference in the equations derived in this chapter. In such a device three depleted regions arise in which generated charge carriers are separated due to a high electric field and, in between, quasi-neutral layers exist in which diffusion of charge carriers is predominant. The upper depleted region is due to charge in the oxide layer and in the silicon space charge regions occur on either side of a pn junction due to the built in voltage, possibly enhanced by a reverse voltage applied across the junction. The sensor structure depicted in figure 2.14 anticipates a colour sensor structure consisting of a shallow p-type layer implanted in an n-epilayer on a p- substrate and thus compatible with a standard bipolar process. Although such an approach reveals obvious advantages, alternatives are quite possible and are indeed discussed in the literature survey in the next chapter. This restriction of the theoretical discussion to a particular configuration offers the possibility of simplifying the problem, since the epilayer and substrate doping can be considered to be constant and only the doping gradient of the shallow layer has to be taken into account, however, the results reveal only a limited validity in differing device structures. The charge collecting properties will now be discussed for a pnp structure, but the results are, apart from the obvious adaptations, also applicable in a npn device structure.

The steady-state response of a photodetector on the generation of optically induced electron-hole pairs can be described by the current density equations, under the boundary conditions valid at that particular depth, and the continuity relations. The boundaries of the different layers can be derived by solving the Poisson equations for both the surface space-charge region and the two junction space-charge regions.

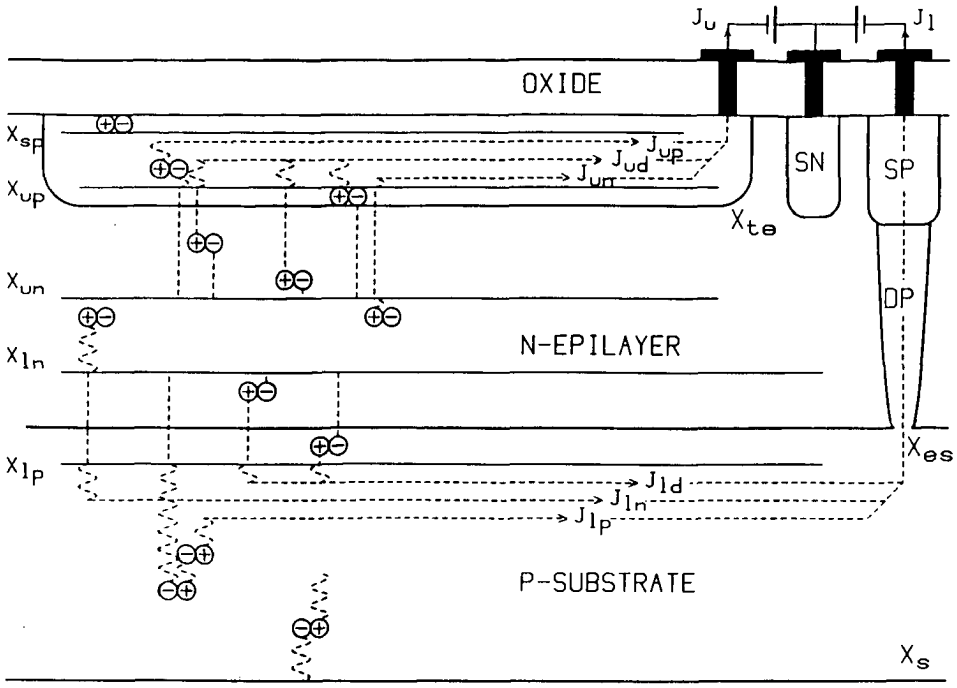


Figure 2.14 Schematic cross-sectional view of a dual-junction optical sensor structure indicating the photocurrent components and the layers in which the associated charge carriers are generated. The subscripts in the variable x denote the layer boundary, with x_{sp} that of the surface depletion layer, x_{up} and x_{un} that of the upper junction depletion layer in the top layer and epilayer respectively and, similarly, x_{ln} and x_{lp} the boundaries of the lower junction depletion layer in the epilayer and the substrate respectively.

The current density equation is generally composed of a drift current density component, proportional to the amount of charge traversing per unit of time a layer governed by an electric field, and of a diffusion component, proportional to the diffusion coefficient and the carrier concentration gradient. For electrons respectively holes in the one-dimensional problem these are formulated by:

$$J_n = q\,\mu_n\,n(x)\,E + q\,D_n\,\frac{\mathrm{d}\,n(x)}{\mathrm{d}\,x} \qquad (a)$$

$$J_p = q\,\mu_p\,p(x)\,E - q\,D_p\,\frac{d\,p(x)}{d\,x} \qquad (b) \qquad (2.25)$$

in which μ_n and μ_p denote the electron respectively hole mobilities connected to the respective diffusion coefficients, D_n and D_p, through the Einstein relation [2.9].

The continuity equations describe the balancing process of excess minority charge carriers towards the thermal equilibrium density. The total rate of change is, therefore, determined by the generation rate G and the recombination rate U [cm^{-3}/s] and the change in the minority current density according to:

$$\frac{\partial\,n_p(x)}{\partial\,t} = G_n - U_n + \frac{1}{q}\,\frac{\partial\,J_n(x)}{\partial\,x} =$$

$$G_n - \frac{n_p - n_{po}}{\tau_n} + n_p(x)\mu_n\,\frac{\partial\,E(x)}{\partial\,x} + \mu_n E(x)\,\frac{\partial n_p(x)}{\partial\,x} + D_n\,\frac{\partial^2 n_p(x)}{\partial\,x^2} \qquad (2.26)$$

which results from the insertion of equation 2.25a for J_n under low injection conditions. In equation 2.26, $n_p(x)$ stands for the actual electron density in a p-type layer and n_{po} for the thermal equilibrium electron density in that layer. The time constant τ_n refers to the electron lifetime in the p-type layer. A similar expression can be derived with respect to the holes in the n-doped layer.

$$\frac{\partial\,p_n(x)}{\partial\,t} = G_p - \frac{p_n - p_{no}}{\tau_p} - p_n(x)\mu_p\,\frac{\partial\,E(x)}{\partial\,x} - \mu_p E(x)\,\frac{\partial p_n(x)}{\partial\,x} + D_p\,\frac{\partial^2 p_n(x)}{\partial\,x^2} \qquad (2.27)$$

in which the coefficients have the same meaning apart from the obvious adaptations. The electron and hole generation rates at depth x, $G_n(x)$ and $G_p(x)$ respectively, are caused by either external effects, such as incident photons or high-energy nuclear particles or by internal causes, such as impact ionisation due to high electric fields. The discussion, naturally focusses on optically generated charge carriers.

The absorption process of incident photons can be modelled as a Poisson probability distribution with an average penetration depth equal to $1/\alpha(\lambda)$. Therefore, the probability of the absorption of a photon at depth x is equal to $\alpha(\lambda)\exp(-\alpha(\lambda)x)$. The incident phonon flux entering the silicon is equal to $F_o(\lambda)\left(1-R(\lambda)\right)$, where $F_o(\lambda)$ denotes the number of incident photons of wavelength λ. This leads to a generation rate at depth x equal to $G(x) = F_o(\lambda)\left(1-R(\lambda)\right)\eta\alpha(\lambda)\exp(-\alpha(\lambda)x)$, where the optical parameters have the meaning indicated in the previous section. The photocurrents in the two junctions of the device shown in figure 2.14 can now be calculated by analyzing the individual contributions of the depleted and the neutral layers.

The surface space-charge region originates from either mobile positive charges, such as sodium ions which might have entered the oxide during processing, or from surface traps induced by high energy radiation, or from the surface-state charge, or from fast surface states that originate from the disruption of the silicon lattice at the interface [2.10]. In the colour sensor the oxide layer is also applied as a masking material. This measure allows the matching of the oxide layer to the projected range of the B^+ implant energy for realizing a maximum in the doping profile at the surface [2.11]. However, this method also adds a B^+ ion concentration to the oxide. The net result of all these charges is a surface electric field in the p-type toplayer attracting minority charge carriers to the interface. Assuming a device structure having both a p-type top layer with an abrupt impurity concentration change from a constant acceptor impurity N_{a1} to a constant donor impurity N_d in the underlying n-type layer at x_{te} as well as a constant oxide charge concentration N_{ox} extending in the oxide from the silicon surface up to a distance x_{ox}, allows the calculation of the surface potential and the associated surface depletion layer width using the Poisson equation. This equation states:

$$-\frac{\partial^2 V}{\partial x^2} = \frac{\partial E}{\partial x} = \frac{\rho(x)}{\epsilon} \qquad (2.28)$$

Application of this equation in both the oxide and the upper layer yields two expressions for the surface potential:

$$\Phi_s = \frac{q}{\epsilon_o \epsilon_{SiO_2}} \int_{-x_{ox}}^{0}\int N_{ox}\,dx\,dx = \frac{q}{\epsilon_o \epsilon_{Si}} \int_{0}^{x_{sp}}\int N_{a1}\,dx\,dx, \qquad (2.29)$$

Hence:

$$x_{sp} = x_{ox} \sqrt{\frac{N_{ox}}{N_{a1}} \frac{\epsilon_{Si}}{\epsilon_{SiO_2}}} \qquad (2.30)$$

Minority charge carriers captured within this surface depleted region will be swept to the silicon surface and will recombine there, because of the high density of surface states. This simplified model shows one of the reasons for the poorer quantum efficiency at short wavelengths where the penetration depth is minimum. A possible improvement would involve the implantation of negative charge in the oxide to compensate for the positive oxide charge. Also a transparent gate applied to a negative voltage sufficient to cause flatband at the surface will yield a similar improvement. In the latter solution care must be taken on the absorption in the gate. In particular the thoughtless applications of relatively thick polysilicon gates will prove counterproductive. Another conclusion arising from equation 2.30 concerns the advantage of having a high surface acceptor doping density. This conclusion is only valid up to a concentration of about 10^{18} cm^{-3}, as it is a consequence of the introduced assumptions with respect to doping profile and the ignorance of high doping effects. The actual implanted layer will reveal a maximum surface concentration at a projected range determined by the implant dose and a decreasing density at increasing depth. The internal field induced by this doping gradient affects the motion of minority charge carriers and can be determined using the current density equations. In thermodynamic equilibrium conditions the majority current density in the toplayer must equal zero as charge carriers have nowhere to go, which yields:

$$J_p = q \mu_p N_a(x) E(x) - qD_p \frac{\partial N_a(x)}{\partial x} = 0 \qquad (2.31)$$

Hence,

$$E(x) = (D_p/\mu_p) \frac{1}{N_a(x)} \frac{\partial N_a(x)}{\partial x} = \frac{kT}{q} \frac{1}{N_a(x)} \frac{\partial N_a(x)}{\partial x} \qquad (2.32)$$

This effect would result in an electric field repelling the minority charge carriers from the interface as the doping concentration decreases at greater depths and would therefore enhance the spectral reponse in the blue-UV part of the spectrum. However, two different effects oppose this

conclusion. The first arises from the redistribution of Boron into the oxide during annealing due to its segregation coefficient being smaller than unity, resulting in a decreased doping towards the surface and a subsequent dragging, majority-carrier field near the surface. Therefore, an implantation energy with an associated projected range slightly greater than the oxide thickness and a relatively low-temperature annealing must be applied. Another effect originates from the occurence of a different recombination mechanism at high impurity concentrations, the Auger process. For lightly doped silicon crystals the energy levels of the impurity atoms can be considered discrete as their number is small enough to avoid overlap and to maintain perfect periodicity of the crystal. Therefore, the edges of the valence and conduction band are well defined. An increased doping concentration causes the formation of impurity bands and a tail instead of an edge in the valence and conduction bands and thus leads to a narrowing of the bandgap. This effect does not only affect the optical properties, as the wavelength dependence of the absorption is directly related to the size and shape of the bandgap, but also results in a higher intrinsic carrier concentration, $n_i = \sqrt{N_c N_v} \exp(-E_g/kT)$, at increasing doping density and constant temperature. The latter is usually accounted for by introducing a net intrinsic carrier concentration, $n_{ie} = n_i \exp(\Delta E_g / kT)$, which subsequently leads to a decreased net doping, $N_{eff}(x) = N(x) / (n_{ie} / n_i)^2$, at a projected increased profile and becomes predominant at doping levels exceeding 10^{19} cm^{-3}. Therefore, an extra retarding field can occur in the p-type top layer, driving the minority charge carriers towards the surface space charge region, which might subsequently lead to a further reduction of the low-wavelength quantum efficiency. This detrimental effect can be avoided by limiting the surface concentration to the already prescribed $N_a(0) = 10^{18}$ cm^{-3}. Under this restriction equation 2.30 can be applied for obtaining the boundary of the shallow layer in which generated charge carriers will be lost.

The total photocurrent density generated in this surface depleted layer can be calculated by integration of the generation rate over the entire layer thickness, which yields:

$$J_{sd} = F_o(\lambda)\left(1 - R(\lambda)\right) \eta q \alpha(\lambda) \int_0^{x_{sp}} exp(-\alpha(\lambda)\, x)\, dx =$$

$$F_o(\lambda)\left(1-R(\lambda)\right) \eta q \left(1-exp(-\alpha(\lambda)\, x_{sp})\right) \qquad (2.33)$$

This component does not contribute to the detected photocurrent as the charge carriers recombine at the surface and consequently a poor quantum efficiency is generally observed at the low wavelength part of the spectrum due to the high corresponding absorption coefficient. Much

research has been performed to improve the properties of silicon photodiodes in this respect.

The width of the space charge region extending on either side of a pn junction can also easily be derived using the Poisson equation. For a one-dimensional abrupt junction approximation one obtains for the upper junction [2.9]:

$$-\frac{\partial^2 V}{\partial x^2} = \frac{\partial E}{\partial x} = \frac{\rho(x)}{\epsilon} = \frac{-q}{\epsilon_o \epsilon_{Si}} N_{a1} \quad for\ x_{up} \le x < x_{te}$$

$$= \frac{q}{\epsilon_o \epsilon_{Si}} N_d \quad for\ x_{te} \le x < x_{un} \qquad (2.34)$$

Hence,

$$x_{te} - x_{up} = \frac{N_d}{N_d + N_{a1}} \sqrt{\frac{2\epsilon_o \epsilon_{Si}}{q} \left(V_{bi1} + V_{rev1}\right) \left(\frac{1}{N_{a1}} + \frac{1}{N_d}\right)}$$

$$x_{un} - x_{te} = \frac{N_{a1}}{N_d + N_{a1}} \sqrt{\frac{2\epsilon_o \epsilon_{Si}}{q} \left(V_{bi1} + V_{rev1}\right) \left(\frac{1}{N_{a1}} + \frac{1}{N_d}\right)} \qquad (2.35)$$

where V_{bi1} denotes the built-in diffusion potential equal to $(kT/q) \ln(N_{a1} N_d / n_i^2)$ and V_{rev1} the reverse voltage applied across the upper junction. Similarly, the boundaries of the depleted regions on either side of the lower junction can be found to be equal to:

$$x_{es} - x_{ln} = \frac{N_{a2}}{N_d + N_{a2}} \sqrt{\frac{2\epsilon_o \epsilon_{Si}}{q} \left(V_{bi2} + V_{rev2}\right) \left(\frac{1}{N_{a2}} + \frac{1}{N_d}\right)}$$

$$x_{lp} - x_{es} = \frac{N_d}{N_d + N_{a2}} \sqrt{\frac{2\epsilon_o \epsilon_{Si}}{q} \left(V_{bi1} + V_{rev1}\right) \left(\frac{1}{N_{a1}} + \frac{1}{N_d}\right)} \qquad (2.36)$$

All charge carriers generated within a junction depletion layer are collected in the respective junction as the traverse time is much smaller than the carrier lifetime due to the electric field. The photocurrent in the upper junction depleted region is therefore equal to:

$$J_{ud} = F_o(\lambda)\left(1 - R(\lambda)\right)\eta\, q\, \alpha(\lambda) \int_{x_{un}}^{x_{up}} \exp(-\alpha(\lambda)\, x)\, dx =$$

$$F_o(\lambda)\left(1-R(\lambda)\right)\eta\, q\, \left(\exp(-\alpha(\lambda)\, x_{up}) - \exp(-\alpha(\lambda)\, x_{un}\right) \qquad (2.37)$$

where the subscripts have the meaning already depicted in figure 2.14. Similarly, the photocurrent in the lower junction depleted region can be described by:

$$J_{ld} = F_o(\lambda)\left(1-R(\lambda)\right)\eta\, q\, \left(\exp(-\alpha(\lambda)\, x_{ln}) - \exp(-\alpha(\lambda)\, x_{lp}\right) \qquad (2.38)$$

Minority charge carriers generated within the non-depleted part of the p-type implanted layer are assumed to drift towards x_{up} due to the majority carrier built in field, described by equation 2.32, resulting in an extra current density component at x_{up} equal to;

$$J_{up} = F_o(\lambda)\left(1-R(\lambda)\right)\eta\, q\, \left(\exp(-\alpha(\lambda)\, x_{sp}) - \exp(-\alpha(\lambda)\, x_{up}\right) \qquad (2.39)$$

The final issue that has to be considered in respect to the collection efficiency is the extent to which charge carriers generated in a quasi-neutral layer in between depleted layers contribute to the photocurrent and, moreover, to which of the respective junctions will the diffusion current prevail.

The diffusion in the neutral part of the epilayer and substrate can be calculated by solving the continuity relation in a neutral layer, $E(x) = 0$, under the appropriate boundary conditions and the insertion of the result in the current density equation at both the adjacent lower boundary of the upper junction and upper boundary of the lower junction. The continuity relation for holes in the neutral part of the epilayer states:

$$D_p \frac{\partial^2 p_n(x)}{\partial x^2} - \frac{p_n(x) - p_{no}}{\tau_p} + F_o(\lambda)\left(1-R(\lambda)\right)\eta\, \alpha(\lambda) \exp(-\alpha(\lambda)\, x) = 0 \qquad (2.40)$$

Solving this expression and applying the boundary conditions $n_p(x_{un}) = n_{po}$ and $n_p(x_{ln}) = n_{po}$ yields for $x_{un} < x < x_{ln}$:

$p_n(x)=$

$$\frac{F_o(\lambda)\left(1-R(\lambda)\right)\eta\alpha(\lambda)}{D_p\left((1/L_p)^2-\alpha(\lambda)^2\right)}\left\{\frac{\exp\left[-\alpha(\lambda)x_{ln}-\frac{x_{ln}}{L_p}\right]-\exp\left[-\alpha(\lambda)x_{un}-\frac{x_{un}}{L_p}\right]}{\exp(-2\,x_{un}/L_p)-\exp(-2\,x_{ln}/L_p)}\exp(-\frac{x}{L_p})+\right.$$

$$\left.\frac{\exp\left[-\alpha(\lambda)x_{ln}+\frac{x_{ln}}{L_p}\right]-\exp\left[-\alpha(\lambda)x_{un}+\frac{x_{un}}{L_p}\right]}{\exp(2\,x_{un}/L_p)-\exp(2\,x_{ln}/L_p)}\exp(\frac{x}{L_p})+\exp(-\alpha(\lambda)\,x)\right\} \qquad (2.41)$$

where L_p denotes the diffusion length of holes in the n-type layer: $L_p=\sqrt{D_p\tau_p}$. The current density equation describing the photocurrent entering the upper junction depleted region at $x = x_{un}$ can be obtained by insertion of the continuity relation in:

$$J_{un}=J_p|_{x_{un}}=-q\,D_p\left.\frac{d\,p(x)}{d\,x}\right|_{x_{un}}=\frac{F_o(\lambda)\left(1-R(\lambda)\right)q\,\eta\,\alpha(\lambda)\,L_p}{1-\left(\alpha(\lambda)\,L_p\right)^2}\times$$

$$\left\{\frac{\exp\left[-\alpha(\lambda)x_{ln}-\frac{x_{ln}+x_{un}}{L_p}\right]-\exp\left[-\alpha(\lambda)x_{un}-\frac{2x_{un}}{L_p}\right]}{\exp(-2\,x_{un}/L_p)-\exp(-2\,x_{ln}/L_p)}-\right.$$

$$\left.\frac{\exp\left[-\alpha(\lambda)x_{ln}+\frac{x_{ln}+x_{un}}{L_p}\right]-\exp\left[-\alpha(\lambda)x_{un}+\frac{2\,x_{un}}{L_p}\right]}{\exp(2\,x_{un}/L_p)-\exp(2\,x_{ln}/L_p)}+\alpha(\lambda)\,L_p\exp(-\alpha(\lambda)\,x_{un})\right\}$$

$$(2.42)$$

In a similar manner the current density equation describing the photocurrent entering the lower junction depleted region at $x = x_{ln}$ can be obtained.

$$J_{ln} = J_p|_{x_{ln}} = -q\,D_p \left.\frac{d\,p(x)}{d\,x}\right|_{x_{ln}} = \frac{F_o(\lambda)\left(1-R(\lambda)\right) q\,\eta\,\alpha(\lambda)\,L_p}{1-\left(\alpha(\lambda)\,L_p\right)^2} \times$$

$$\left\{\frac{\exp\left(-\alpha(\lambda)x_{ln}-\frac{2\,x_{ln}}{L_p}\right)-\exp\left(-\alpha(\lambda)x_{un}-\frac{x_{ln}+x_{un}}{L_p}\right)}{\exp(-2\,x_{un}/L_p)-\exp(-2\,x_{ln}/L_p)}-\right.$$

$$\left.\frac{\exp\left(-\alpha(\lambda)x_{ln}+\frac{2\,x_{ln}}{L_p}\right)-\exp\left(-\alpha(\lambda)x_{un}+\frac{x_{ln}+x_{un}}{L_p}\right)}{\exp(2\,x_{un}/L_p)-\exp(2\,x_{ln}/L_p)}+\alpha(\lambda)\,L_p\exp(-\alpha(\lambda)\,x_{ln})\right\}$$

(2.43)

The photocurrent detected in the lower junction also contains a component due to charge carriers generated beyond the lower boundary of its space charge region and able to diffuse towards this region, which can be described by the current density equation at $x = x_{lp}$. In this case, however, different boundary conditions are required for solving the continuity relation, as this quasi-neutral layer is not enclosed by two depleted layers, but extends from the lower boundary of the lower junction space-charge region down to the back contacting at x_s. The actual boundary conditions in this problem are therefore described by $n_p(x_{lp}) = n_{po}$ and $n_p(x_s) = n_{po} + (D_n/S_n)(\partial n_p(x)/\partial x)$, where S_n denotes the recombination velocity of electrons at the rear surface of this p-type layer. Applying these conditions to the continuity equation for electrons yields for $x_{lp} < x < x_s$:

$$n_p(x) = \frac{F_o(\lambda)\left(1-R(\lambda)\right)\eta\,\alpha(\lambda)}{D_n\left((1/L_n)^2-\alpha(\lambda)^2\right)} \times$$

$$\left\{\frac{\left(1-\frac{D_n}{S_n L_n}\right)\left(\exp(-\alpha(\lambda)x_{lp}-\frac{x_{lp}}{L_n})\right)-\left(1+\frac{\alpha(\lambda)\,D_n}{S_n}\right)\left(\exp(-\alpha(\lambda)x_s-\frac{x_s}{L_n}\right)}{\left(1+\frac{D_n}{S_n L_n}\right)\left(\exp(-2\,x_s/L_n)\right)-\left(1-\frac{D_n}{S_n L_n}\right)\left(\exp(-2\,x_{lp}/L_n)\right)}\exp\left(\frac{-x}{L_n}\right)+\right.$$

$$\frac{\left(1+\frac{D_n}{S_n L_n}\right)\left(\exp(-\alpha(\lambda)x_{lp}+\frac{x_{lp}}{L_n})\right)-\left(1+\frac{\alpha(\lambda)D_n}{S_n}\right)\left(\exp(-\alpha(\lambda)x_s+\frac{x_s}{L_n}\right)}{\left(1-\frac{D_n}{S_n L_n}\right)\left(\exp(2\,x_s/L_n)\right)-\left(1+\frac{D_n}{S_n L_n}\right)\left(\exp(2\,x_{lp}/L_n)\right)}\exp(-\frac{x}{L_n})-$$

$$\exp(-\alpha(\lambda)\,x)\Bigg\} \qquad (2.44)$$

The current density equation describing the photocurrent entering the lower junction depleted region at $x = x_{lp}$ can be obtained by insertion of the continuity relation in:

$$J_{lp} = J_n|_{x_{lp}} = q\,D_n \frac{d\,n(x)}{d\,x}\Bigg|_{x_{lp}} = \frac{F_o(\lambda)\left(1-R(\lambda)\right)\,q\,\eta\,\alpha(\lambda)\,L_n}{1-\left(\alpha(\lambda)\,L_n\right)^2} \times$$

$$\Bigg\{\frac{\left(1+\frac{D_n}{S_n L_n}\right)\left(\exp(-\alpha(\lambda)x_{lp}+\frac{x_{lp}}{L_n})\right)-\left(1+\frac{\alpha(\lambda)\,D_n}{S_n}\right)\left(\exp(-\alpha(\lambda)x_s+\frac{x_s}{L_n}\right)}{\left(1-\frac{D_n}{S_n L_n}\right)\left(\exp(2\,x_s/L_n)\right)-\left(1+\frac{D_n}{S_n L_n}\right)\left(\exp(2\,x_{lp}/L_n)\right)}-$$

$$\frac{\left(1-\frac{D_n}{S_n L_n}\right)\left(\exp(-\alpha(\lambda)x_{lp}-\frac{x_{lp}}{L_n})\right)-\left(1+\frac{\alpha(\lambda)D_n}{S_n}\right)\left(\exp(-\alpha(\lambda)x_s-\frac{x_s}{L_n}\right)}{\left(1+\frac{D_n}{S_n L_n}\right)\left(\exp(-2\,x_s/L_n)\right)-\left(1-\frac{D_n}{S_n L_n}\right)\left(\exp(-2\,x_{lp}/L_n)\right)}-$$

$$\alpha(\lambda)\,L_n \exp(-\alpha(\lambda)\,x_{lp})\Bigg\} \qquad (2.45)$$

The subscripts have a meaning as indicated in figure 2.14. In literature on similar structures, equations 2.41 through 2.45 are sometimes expressed in hyperbolic functions, which is a result of having made a different choice for the homogenous solution of the continuity relation [2.12] [2.13]. As $\cosh(x)$ and $\sinh(x)$ can be expressed in $\exp(x)$ and $\exp(-x)$, the solutions are interchangable.

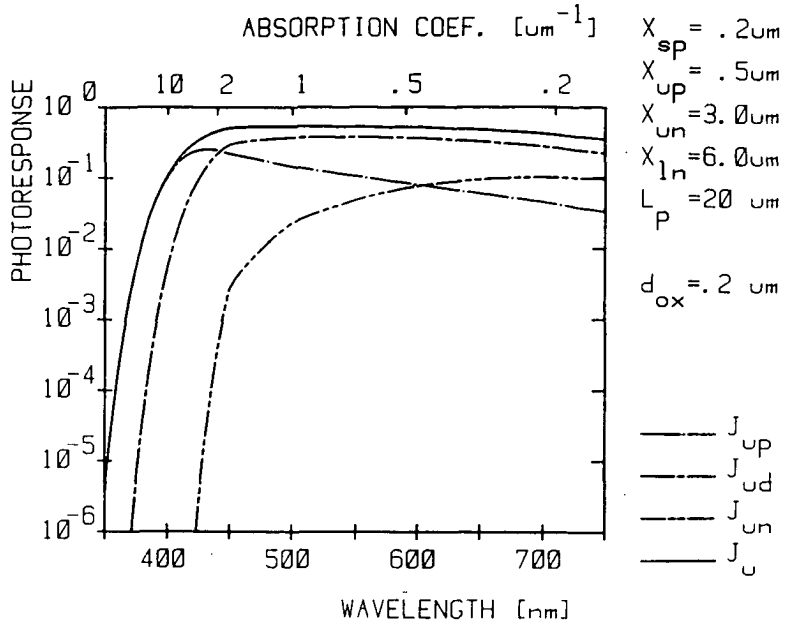


Figure 2.15 Calculated photocurrent components vs. wavelength in the upper junction that are generated in the different silicon layers.

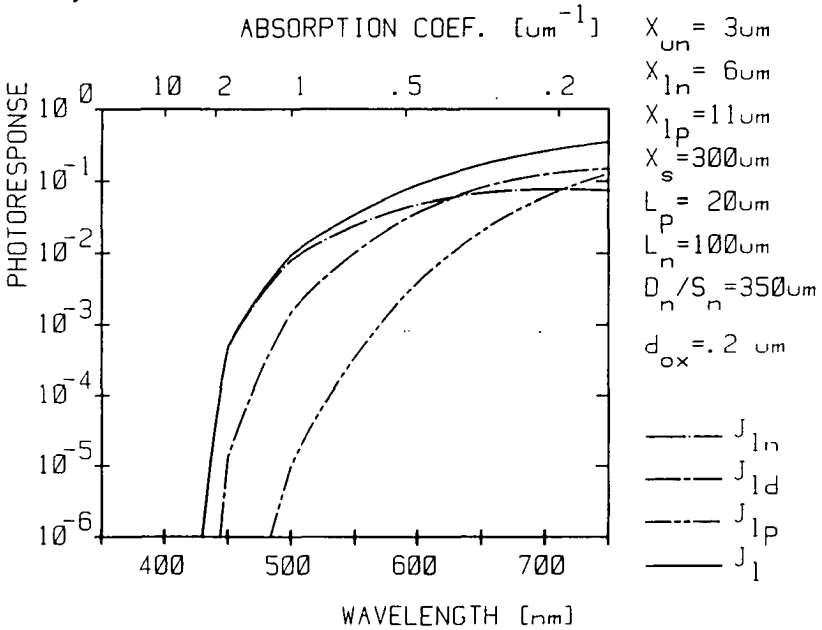


Figure 2.16 Calculated photocurrent components vs. wavelength in the lower junction that are generated in the different silicon layers.

2.5 Conclusions.

The value of the various photocurrent components and the total photocurrent detected in both the upper junction, $J_u = J_{ud}+J_{up}+J_{un}$, and the lower junction, $J_l = J_{ld}+J_{ln}+J_{lp}$, is, for a certain sensor configuration, depicted as a function of λ in figures 2.15 and 2.16. For the given dimensions these figures clearly indicate a response of the upper diode that is dominated by surface effects for wavelengths up to 400 nm. Beyond this wavelength the spectral response is primarily determined by the drift-photocurrent in the upper junction space-charge region. The effect of the diffusion current from the neutral part of the epilayer is negligible. The short-wavelength response of the lower junction is primarily determined by the diffusion current from the epilayer. The contribution of the drift current in the lower space-charge region becomes significant at wavelengths beyond 600 nm. Therefore, if J_{ud} is reduced the response of the upper junction is that of a short-pass filter and, similarly, if J_{ln} is reduced an effective long-pass filter is realised in the lower junction. The total photocurrent in the upper and lower junction for a known spectral distribution of incident light can now be calculated for a given three-layered colour sensor structure. In practical colour sensors described in literature, a particular relation between these photocurrents, such as the ratio, is taken as the characteristic colour-determined output signal. In the following chapter the merits and drawbacks of colour sensors reported in literature are addressed.

2.6 References.

2.1 R.A.Smith, Semiconductors, Cambridge University Press, 1959, pp. 189-214.

2.2 C.Kittel, Introduction to solid-state physics, Wiley, New York, 1976, pp. 107-124.

2.3 D.K.Schroder, R.NoelThomas and J.C.Schwartz, Free carrier absorption in silicon, IEEE Tr. Electron Devices, Vol ED-25 No2, 1978, pp. 254-261.

2.4 E.Barta, Optical constants of various heavily-doped p- and n-type silicon crystals obtained by Kramers-Kronig analysis, Infrared physics, Vol 17, 1977, pp. 319-329.

2.5 T.S.Moss, G.J.Burrell and B.Ellis, Semiconductor optoelectronics, Butterworths, London, 1973.

2.6 H.R. Philipp, Influence of oxide layers on the determination of the optical properties of silicon, J. Appl. Phys., Vol 43 No 6, June 1972, pp. 2835-2839.

2.7 O.S.Heavens, Optical properties of thin solid films, Butterworths, London, 1955.

2.8 O.Christensen, Quantum efficiency of the internal photoelectric effect in silicon and germanium, J. Appl. Phys., Vol 47 No 2, Febr. 1976, pp. 689-695.

2.9 S.M.Sze, Physics of semiconductor devices, Wiley, New York, 2nd edn, 1981.

2.10 A.S.Grove, Physics and technology of semiconductor devices, Wiley, New York, 1967, pp. 334-346.

2.11 A.S.Silard, Using silicon dioxide as device engineering parameter in the fabrication of photovoltaic cells through masked ion implantation, Solar energy materials, Vol 14, 1986, pp. 95-105.

2.12 S.G.Chamberlain, D.J.Roulston and S.P.Desai, spectral response limitation mechanisms of a shallow junction n^+p photodiode, IEEE Tr. Electron Devices, Vol ED-25 No2, Febr. 1978, pp. 241-246.

2.13 H.J.Hovel, Solar cells; carrier collection, spectral response and photocurrent, Semiconductors and Semimetals, Vol 11, pp. 8-47.

3 APPLICATION OF THE WAVELENGTH-DEPENDENT PROPERTIES IN SILICON FOR COLOUR SENSING.

3.1 Introduction.

In the previous chapter the different physical mechanisms that contribute towards the wavelength dependence of the response of silicon photodetectors were outlined. This response has been shown to be determined by both the intensity of the incident radiation, or the brillance of the observed object and the spectral distribution, or the colour of the object. For improving the flatness of the detector response the effect of the spectral dependence has to be reduced and for operation as a colour sensor this effect has to be enhanced. Therefore, the two application areas have in common a particular treatment of the wavelength-dependent absorption in silicon; favouring the effect for colour sensing and reducing it in optical radiometric applications.

As has already been indicated a number of design parameters such as oxide thickness and impurity profile can be optimised for a desired response. Although the study and subsequent optimising of these parameters strongly improved the solar-cell performance, their impact is not pronounced enough for obtaining clearly distinct responses and can not be used for realising colour sensors. Moreover, these measures will not lead to an electronic programming of the response. Therefore, the colour sensing will be incorporated in the analog signal conditioning circuits rather than in the device structure. The key point in such a circuit will be the determination of the colour irrespective of the interfering optical quantity, the intensity.

If no extra information is added to a sensor output signal, for example by chopping of the input or an averaging over previous measurements, there are three fundamental instrumentation techniques for improving the desired quantity as against an interfering quantity; viz. isolation, correction and compensation. This chapter starts with a discussion of the relation between the optical spectrum and human colour perception and what this means in relation to the colour sensor, followed by a brief description of the three basic signal-enhancement techniques and a classification of colour sensors as reported in literature. It will appear that these colour sensors all utilize a similar correction technique. After analysing the merits and drawbacks of these colour sensors the properties and advantages of the colour sensors actually developed within the context of this study will be discussed.

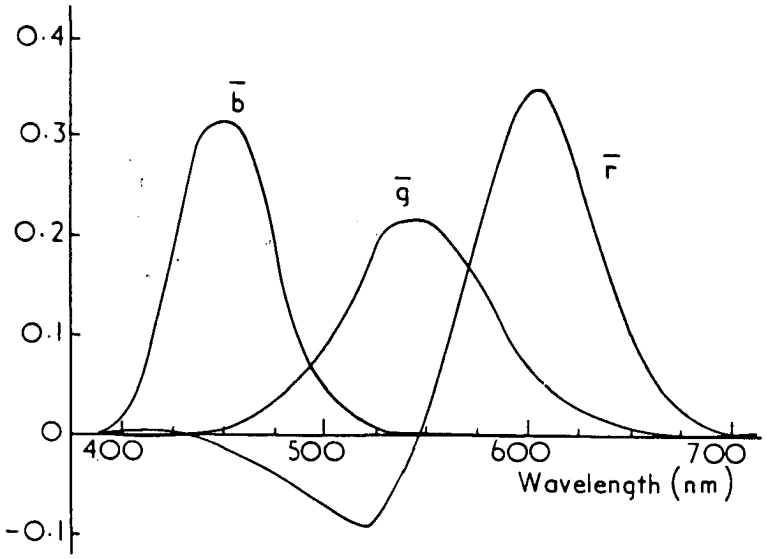


Figure 3.1 Colour-matching functions for the 2° standard observer [3.1].

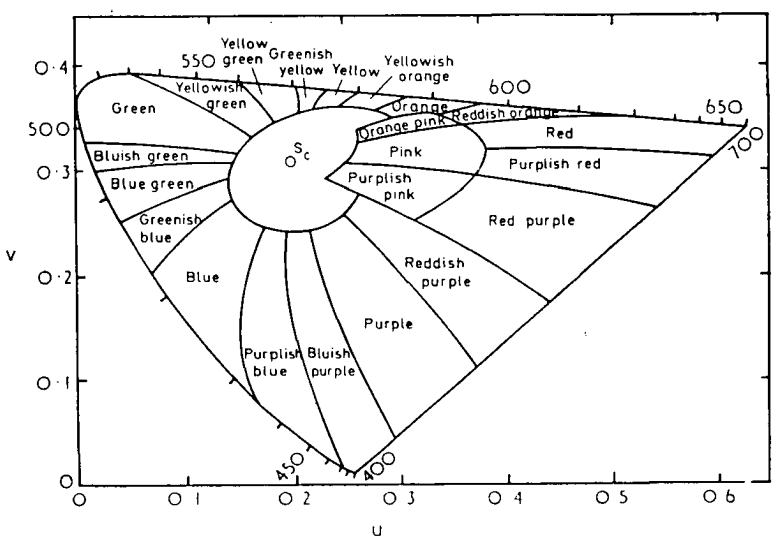


Figure 3.2 Colour triangle.

3.2 Relation between optical spectrum and colour sensation.

In the colour sensor a response to a certain optical spectrum is generated. The same spectrum causes a certain colour sensation in a human observer. To justify the term 'colour sensor' the sensor should provide an output signal accordingly. Extensive experimental research has shown that the human colour perception can be represented by an additive mixture of three monochromatic wavelengths. This observation is in accordance with the fact that there are three different types of receptors in the eye plus one additional type for night vision. The corresponding stimuli are choosen at 435.8 nm, 546.1 nm and 700 nm [3.1]. The relative intensities of each of these stimuli, required for bringing about the same colour impression in a standard observer as that of a monochromatic wavelength, are shown in figure 3.1 vs. this wavelength. These functions are generally referred to as colour-matching functions or eye sensitivity curves and can also be interpreted as optical filter curves. Placing such a filter in front of an ideal detector would give an output response identical to that of the corresponding receptor in the human eye. When applying a certain matrix transformation to this set of three functions, two modified colour-matching functions, with positive values only, plus a luminance signal can be obtained. The latter represents the physiological counterpart of the optical intensity. Figure 3.2 shows the required value of each of the normalised colour-matching functions u and v for obtaining a particular colour at a constant luminance level. The outline of this colour triangle determines the range of monochromatic wavelengths. Due to the additive mixing property any colour inside this triangle can be represented by the centre of gravity of the intensities of the monochromatic wavelengths situated on the outline. Therefore, the reproduction of a colour is possible if three different response curves are realised in a photodetector with a transformation matrix to the colour-matching functions.

Applying several different reverse voltages to a silicon photodiode makes it possible to obtain three different responses and thus allows the reproduction of the colour of incident light. In the simple colour indicators only an indication of colour is obtained. However, in the more elaborated versions of the colour indicator described in this chapter, a tuning parameter is controlled in such a way that the unambiguous detection of colour can be achieved. Therefore, the unambiguous reconstruction of a colour is possible using the proposed silicon colour sensor. However, several effects need to be investigated as they might deteriorate the response. As will be discussed in the following chapter, the values of the nondiagonal elements in the transformation matrix arising from the actual responses to the normalised eye-sensitivity curves have to be minimised. This requirement implies a maximum distinction between the response curves and thus a maximum enhancement of the wavelength-dependent effect in silicon is required.

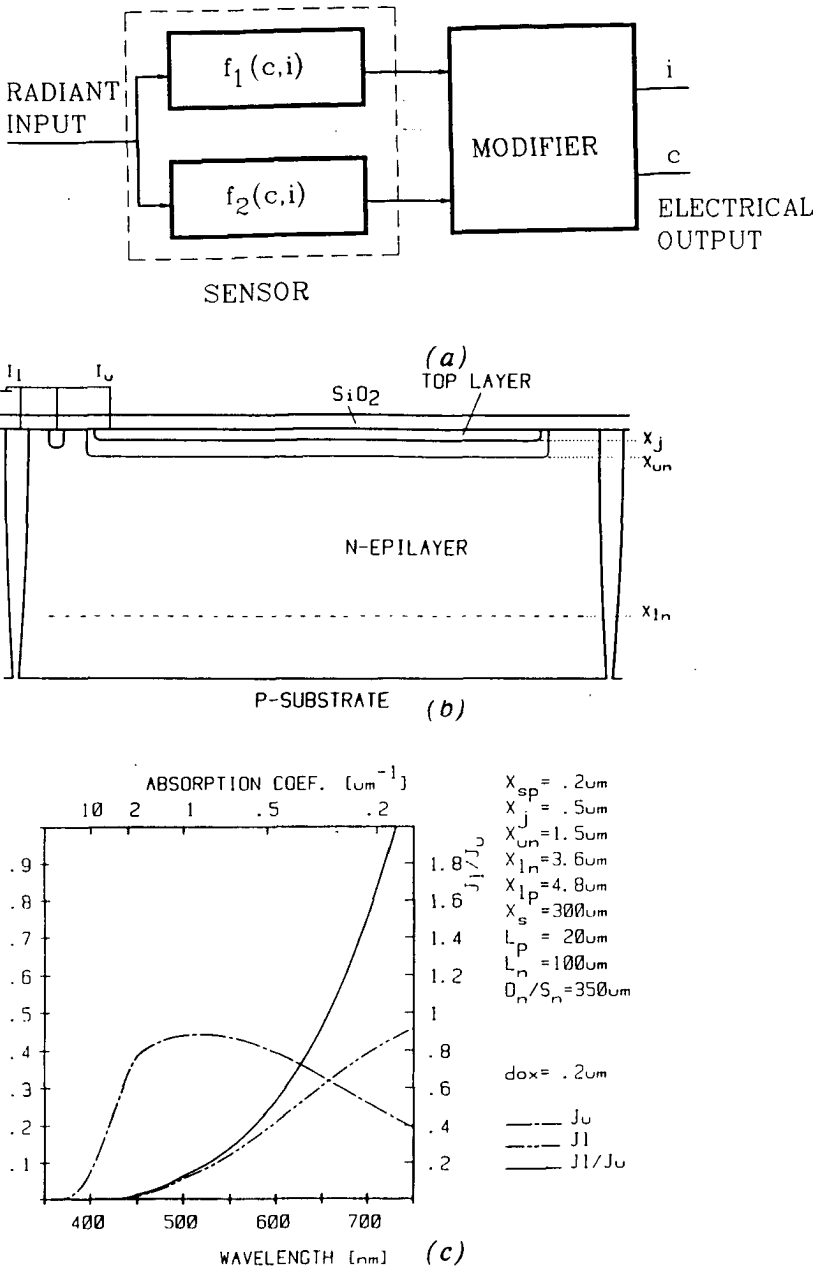


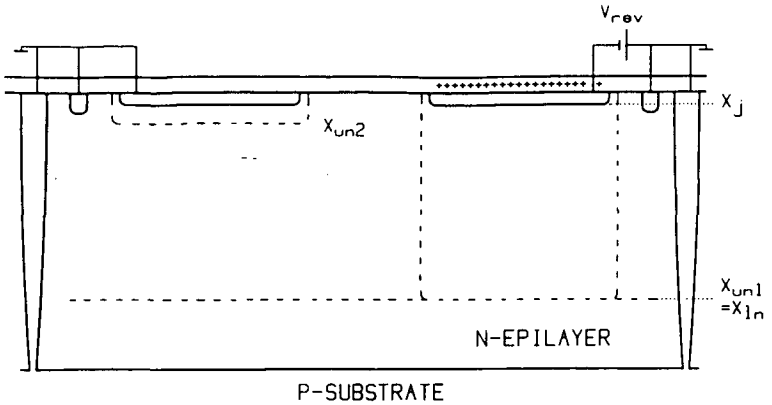
Figure 3.3. (a) Schematic diagram of the correction method.
(b) Diagram of the implementation of the correction technique in the superposed dual-junction colour sensor.
(c) Calculated response of the dual-junction ratio-type of colour indicator with dimensions as shown.

3.3 Signal enhancement techniques and the application in colour sensing.

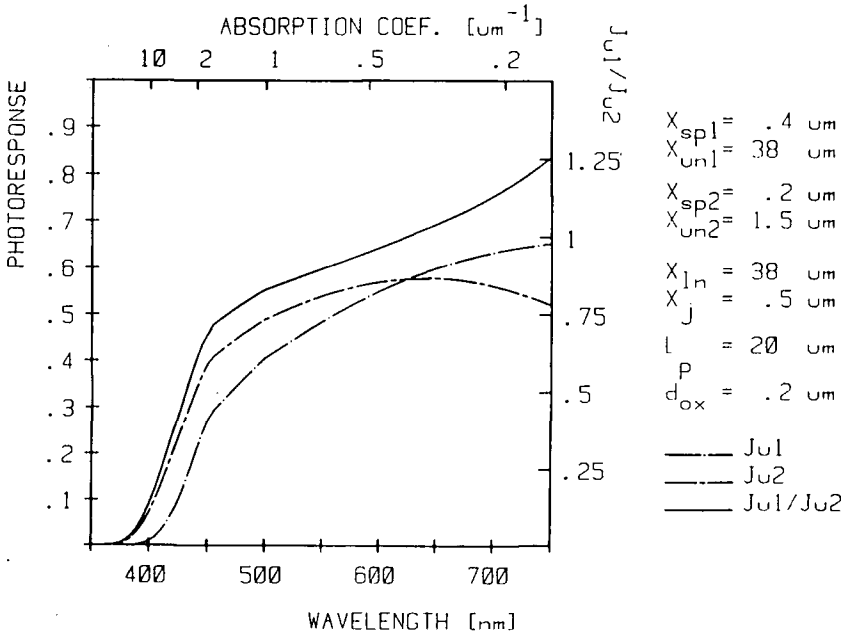
A crucial issue in instrumentation for transducer read-out has always been the extent to which the desired signal can be separated from an interfering signal and subjected to a treatment which maximally favours this signal. An enhanced discrimination between these signals automatically yields a reduced detection threshold and thus leads to an improved meaningful sensitivity of the measurement system. The methods developed to improve this threshold are usually related to noise as the undesired signal, however these general techniques can also be applied in colour sensors to reduce the effect of the optical intensity.

The most straightforward technique involves the isolation of the signals, where the additive interference is prevented from affecting the desired signal. This method is not applicable in colour sensing as it assumes an effective shielding of an originally undisturbed signal. The optical intensity and the colour are, however, two basic optical quantities that together determine the spectrum of the incident radiation. Therefore, these signals are intermixed to start with and can not be separated. Thus, an isolation technique, which prevents an increase in the undesired signal level, does not provide an adequate means of extracting colour information.

Another, frequently applied, technique is correction. In this method the modifier performs a certain operation on the signals supplied by several sensors, each with a different sensitivity for both the desired quantity and the interfering signal. This method is very suitable for colour sensing and is, indeed, the method usually encountered in colour sensors described in literature. A general schematic diagram of this method is depicted in figure 3.3a for two sensors. The two transducers reveal a sensitivity to the input quantities, colour c and intensity i, described by $f_1(c,i)$ and $f_2(c,i)$ respectively. The two unequal functions of the two measurands can be solved, either analog or digital, in the modifier, so an unambiguous modifier output is possible, giving both the colour and the intensity, provided that the sensitivities of the transducers for both the desired and interfering signal are reproducible and not suffering from another cross-effect; such as temperature. In silicon photodetectors the sensitivity to the intensity, as determined by the response at ideal charge-collection, is solely determined by the detector sensitive area, whereas all wavelength dependencies are accounted for in the colour response. This property permits considerable simplification, because the function $f_j(c,i)$ can now be separated into $g_j(c)\times h_j(i)$. The assumption of two superposed detectors in silicon with equal area permits the extraction of a predominantly colour-determined output signal by simply taking the ratio between the photocurrents in the short-circuited photodiodes as shown in figure 3.3b. Due to this normalisation the response is not affected by the intensity and



(a)



(b)

Figure 3.4 *(a) Schematic diagram of the lateral ratio-type of colour indicator using a surface charge and a junction reverse voltage.*
(b) Calculated response of the photodiodes and the ratio between these currents in a sensor with dimensions as shown.

thus the smearing of the sensor surface does not reduce the sensor performance, provided that it does not introduce a major wavelength dependency.

For this sensor structure the responses of both the upper and lower short-circuited photodiode can be calculated using the equations 2.37-2.39, 2.42, 2.43 and 2.45 and result in a response for these diodes as shown in figure 3.3c. Taking the ratio J_l / J_u results in the calculated intensity-independent colour response. This curve clearly reveals a monotonical course as a function of wavelength. For wavelengths smaller than 400 nm the response of a practical device will become inaccurate due to the division between two very small numbers.

Another implementation of the correction technique involves the normalisation of the photocurrent supplied by a photodiode subjected to an extra As^+ implantation in the oxide and also connected to a high reverse voltage on the response of an identical short-circuited photodiode without an oxide charge. The extra positive oxide charge in the former diode deteriorates the the blue sensitivity, as the boundary of the surface depletion layer, x_{sp}, becomes larger, and the reverse voltage enhances the long-wavelength sensitivity, as the value of x_{un} increases. The depletion layer down from the shallow junction is sufficiently larger than $1 / \alpha(\lambda)$ for the entire range of wavelengths in the spectrum. The sensor structure is shown in figure 3.4a. The responses of the individual photodiodes can, again, be calculated from the equations derived in the previous chapter. For the reverse-biased diode a large epilayer is entirely depleted down from the shallow junction to the upper boundary of the built-in depletion layer of the short-circuited epilayer-substrate junction. Moreover, surface depletion is obtained which causes $x_{sp} = x_{up} = 0.4$ μm. The epilayer-substrate junction is at $x_{es} = 40$ μm. At the appropriate reverse voltage the upper depleted layer coincides with the epilayer-substrate built-in depletion layer at $x_{un} = x_{ln} = 38$ μm. This also causes $J_{un} = J_{ln} = 0$ in this detector. The shallow short-circuited photodiode reveals $x_{sp} = 0.2$ μm and only a built-in depletion layer is present. The responses of these two photodetectors are shown in figure 3.4b. The external response remains, of course, affected by the wavelength-dependence of the reflection due to the presence of the 0.2 μm thick oxide layer on top of the detector. Taking the ratio between these two currents results in a colour-determined output signal as shown by the solid line in the figure.

The third technique mentioned is the compensation of the undesired quantity by controlling certain parameters in two identical transducers in such a way that the interference is cancelled out in the output signal after subtraction of the signals supplied by the transducers as shown schematically in figure 3.5. Rather than processing sensor information,

which is a characteristic of the correction technique, in this method at least two different parameters, both having a significant and unequal effect on the sensor response for the desired and the interfering quantity, are set in such a way that an equal response to the interfering signal is obtained, whereas unequal responses for the desired signal are observed. Therefore, after subtraction of the sensor responses the interference is compensated for and the desired signal remains; preferably not attenuated.

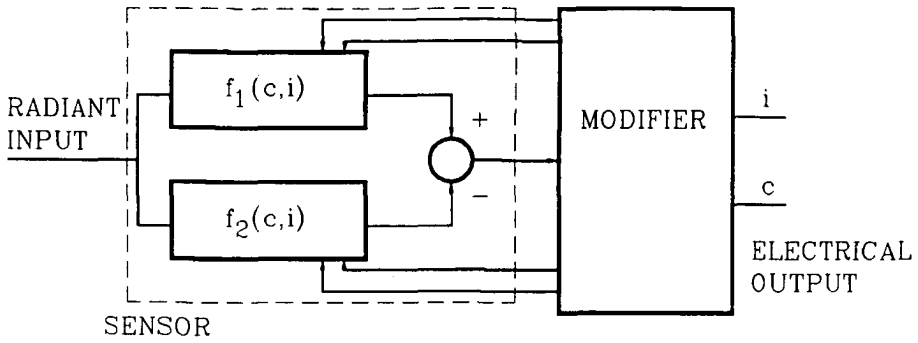


Figure 3.5. Schematic diagram of the compensation method.

As mentioned previously, the photocurrent in a photodiode is determined by the sensitive area and the collection depth, giving two different methods for affecting the response. One of the controllable parameters is, therefore, the sensitive area and another the reverse voltage applied across a shallow junction in combination with the depletion of remaining quasi-neutral parts of the epilayer using a JFET-controlled epilayer-substrate reverse voltage. The epilayer-substrate reverse voltage is, for that purpose, controlled by the pinch-off in the integrated dual-gate JFET with an upper gate identical to a colour element and the substrate acting as the lower gate. The resulting controllable boundary in the epilayer, beyond which no charge carriers are collected in the upper junction, establishes the collecting depth. In a sensor structure composed of two shallow junctions with unequal sensitive areas, equal photocurrents can nevertheless be obtained with this method when using a smaller reverse voltage across the larger photodiode. At a certain fixed reverse reference voltage across the smaller diode (the reference diode) a large difference in depletion depth is required at predominantly short-wavelength incident light, as almost all radiation is shallowly absorbed, whereas at predominantly long-wavelength light a reverse voltage only slightly smaller than the reference voltage is required across the larger diode (the compensating diode) for obtaining equal photocurrents. Therefore, the reverse voltage required across the larger photodiode of the dual-diode structure can also serve as the colour-determined output signal.

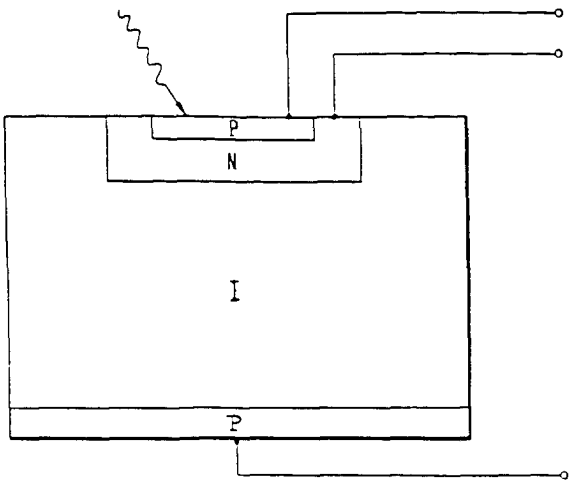
This version of the colour sensor reveals a great degree of flexibility since the actual response can be programmed by the actual difference between the sensitive areas of the photodiodes and the fixed reference voltage. The difference in effective area can be programmed electronically by a controlled attenuation in the photocurrent supplied by one of two equal photodiodes. Rather than producing a straightforward function of sensor currents this method is based on a feedback configuration resulting generally in a more complex sensor read-out. This property combined with the fact that industry has been reluctant to adopt sensors with integrated read-out electronics is the main cause to the rather one-sided attention for the ratio type of colour sensor in literature.

3.4 Survey of silicon colour sensors.

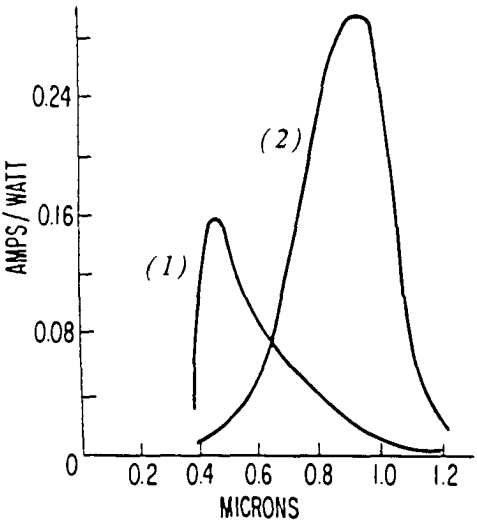
The research effort on monocrystalline silicon colour sensors as reported in literature is very limited and is focussed on the application of the dual-junction ratio-type of colour sensors. The operation generally involves the determination of the photocurrent ratio of two short-circuited photodiodes at different depths. Apart from this category a small but increasing attention to amorphous silicon colour sensors is noticable. This effort mainly stems from the extremely small value of diffusion length in this material. This property, although it is a disadvantage in active devices, makes direct control of the photocurrent in a simple photodiode possible, using the reverse voltage across the junction without having to add a mechanism for depleting the remaining silicon layer underneath the space-charge region. The photocurrent is almost entirely composed of drift current so a very simple device structure can be used. However, the poorer performance and reproducibility of active components in amorphous silicon will make the realisation of amorphous silicon colour sensors with integrated signal conditioning unlikely in the near future.

The early reports on the application of semiconductor materials for colour sensors are mainly to be found in the patent literature. The first reports are, more or less, of curiosity value rather than of direct practical importance [3.2][3.3]. These devices are based on the etching of a slice of semiconductor material, usually Ga As, down to a thickness equal to $1/\alpha(\lambda)$, giving a maximum response at the designed wavelength. Especially for Ga As where the absorption coefficient changes rapidly between 870 nm and 920 nm, due to the direct bandgap at 1.4 eV, allows a very sharp filtering of a wavelength lying in this range.

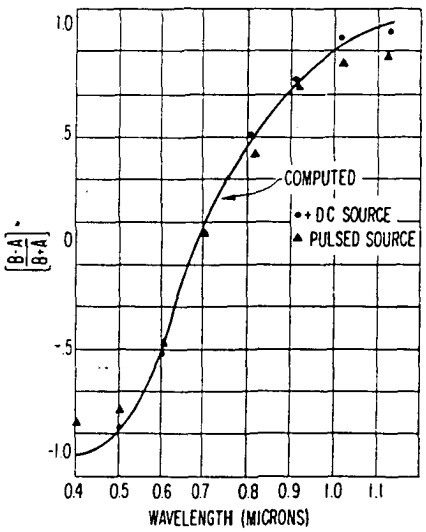
One of the first reports on the basic ratio-type of colour sensors in literature dates from 1977 [3.4]. This report describes the familiar three-layered structure, however, as a refinement a PNIP structure is used for



(a)

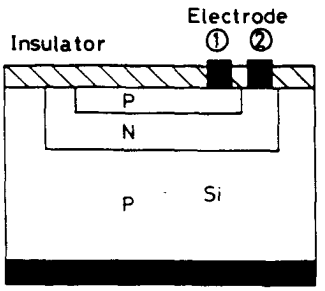


(b)

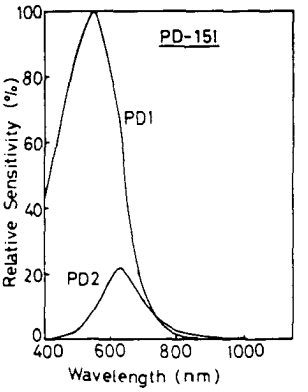


(c)

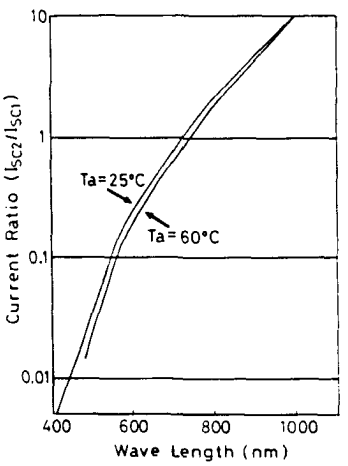
Figure 3.6 *(a) Schematic diagram of the PNIP ratio-type of colour indicator.*
(b) Responses of the upper (1) and lower junction (2).
(c) Colour response as determined by the ratio between the difference and the sum in the photocurrents.



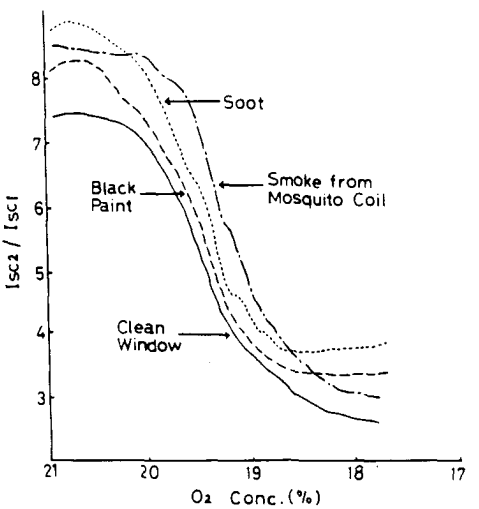
(a)



(b)



(c)



(d)

Figure 3.7 *(a) Cross-sectional view of the colour sensor.*
(b) Photoresponse of the upper and lower junction.
(c) Relation between shortcircuited photocurrent ratio and wavelength.
(d) Influence of smearing of the sensor window on the colour shift at decreasing oxygen concentration.

allowing the efficient collection of charge carriers generated in a thick intrinsic layer, I, in the N I junction as shown in figure 3.6a. Also a criterion which differs slightly from the direct ratio has been used for the read-out. The value $(I_u-I_l)/(I_u+I_l)$ is used in order to obtain a normalisation between -1 and 1. The response of the individual diodes is shown in figure 3.6b and in figure 3.6c the calculated and measured colour output signal is shown as a function of wavelength.

Another colour sensor which exactly fits the category of the ratio type is shown in figure 3.7a [3.5][3.6]. The upper and lower photodiodes reveal a reponse as shown in figure 3.7b from which the colour-determined output signal, I_u / I_l, can be derived. The course of the output signal vs. the wavelength is shown in figure 3.7c. This particular sensor has been developed for oxygen sensing in close proximity of a flame and seems to serve its purpose quite well as is confirmed by the measurements in figure 3.7d [3.7]. The colour-shift of the flame from blue to reddish at incomplete combustion, associated with a decreasing oxygen content, is detected almost independently from the degree of smearing of the sensor window.

Another realisation of a superposed dual-junction colour sensor is distinguished from the general configuration by the fabrication method [3.8]. The usual fabrication process involves either the growth of a n-type epilayer on a p-type substrate followed by a low-energy boron implantation or a n-type diffusion on the substrate follwed by a p-type diffusion or implantation. The complementary polarities with the obvious adaptations are also possible. A common disadvantage of such devices is an increased doping concentration towards the surface. As mentioned in the previous chapter, a high top layer concentration gives rise to a high surface recombination and thus leads to a poor short-wavelength response of the photodiode. In the alternative device a NPN structure is realised using a high-energy boron implantation step, resulting in a p-type layer in the n-type epilayer or substrate. A 1 MeV implantation gives a relatively high-doped p-type layer with lower doped n-type layers on either side and junctions, depending on the implant dose and annealing conditions, at about 1.2 μm and 2.1 μm. The low-doped surface layer results in an improved blue sensitivity of the upper junction [3.9] and thus to an enhanced low-wavelength spectral discrimination of the colour sensor. An additional p-type diffusion is required for contacting.

So far only the developments in simple silicon colour indicators have been outlined. However, also in the solid-state imager research a noteworthy attempt has been made to utilize the wavelength-dependent properties of silicon for loosening the coupling between resolution and imager

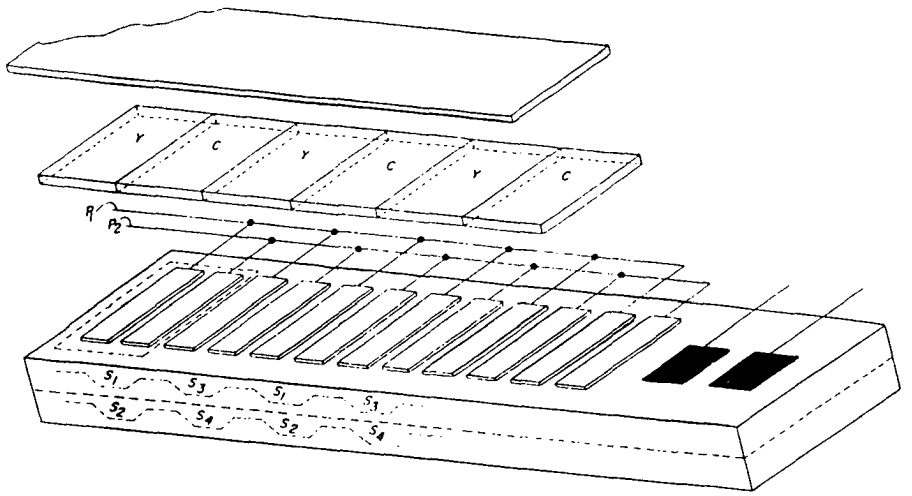
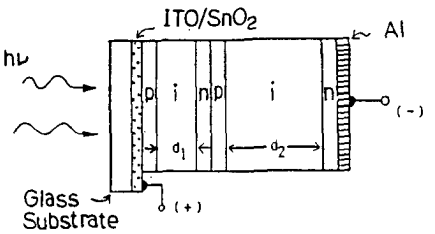
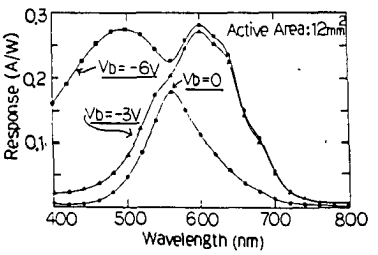


Figure 3.8 Exploded view of a colour imager with only two dyes per colour element.





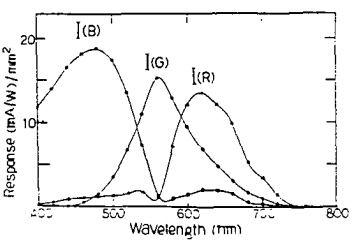


Figure 3.9

(a) Diagram of the a:Si colour sensor.

(b) Responses to three values of the reverse voltage.

(c) Response curves after:
I(G)=I(0)
I(B)=I(-6)-I(-3)
I(R)=I(-3)-I(0)

conditions, doping concentration and annealing conditions [3.13]. In the device under consideration the upper I-type layer has a width, d_1, equal to 0.1 μm and the lower I-type layer, d_2, is 0.5 μm wide and both reveal an indirect optical bandgap at 1.73 eV. At short-circuiting of the terminals only the charge carriers generated within the built-in depletion layer of the n p junction at 0.1 μm depth are efficiently collected, giving a peak response at a wavelength of about 580 nm. At an increasing reverse voltage first the lower I-type layer is depleted, giving an enhanced red response and at even larger values of this reverse voltage also the upper I-type is depleted, giving a maximum overall response. This is confirmed by the measurements shown in figure 3.9b for three values of the reverse voltage. From these responses reasonable approximations of the positive parts of the eye sensitivity curves can be obtained as shown in figure 3.9c. Similar bias-controlled colour sensors can be made using a: Si / SiC devices [3.14] [3.15].

3.5 The ratio type of colour sensor that results from this study.

In the colour indicators reported in literature the objective was usually the realisation of a colour sensor structure in silicon, which directly supplied a predominantly colour-determined output signal with as little additional signal conditioning circuits as possible. In the present study somewhat different priorities are set and flexibility and compatibility with a standard bipolar process are considered of more importance than a minimum system complexity. This is not an ad-hoc decision as the compatibility with a standard bipolar process implies a smooth integration of sensor with read-out electronics, making an increased complexity of the latter of lesser significance.

Two different colour sensors are studied in greater detail. The first can be classified among the correction type of sensor in which provisions are made for ensuring a maximum flexibility. This sensor with the read-out is shown schematically in figure 3.10. The ratio $I_u / (I_u+I_l)$ is choosen as the colour-determined output signal. The reverse voltage across the shallow junction is implemented as an extra design parameter. The current I_u flows from the upper p-type layer and the current $I_u+ I_l$ flows to the epilayer contact. Also in this device the entire epilayer is depleted using an JFET controlled substrate voltage for ensuring $J_{un} = J_{ln} = 0$. The operation and limitations of this JFET-controlled depletion mechanism are discussed in the next paragraph.

Using this method with a small value of V_{ds} and assuming an uniform epilayer doping and width precludes the existence of a neutral layer in the epilayer. Applying equations 2.37-2.39, 2.42, 2.43 and 2.45 yields for the

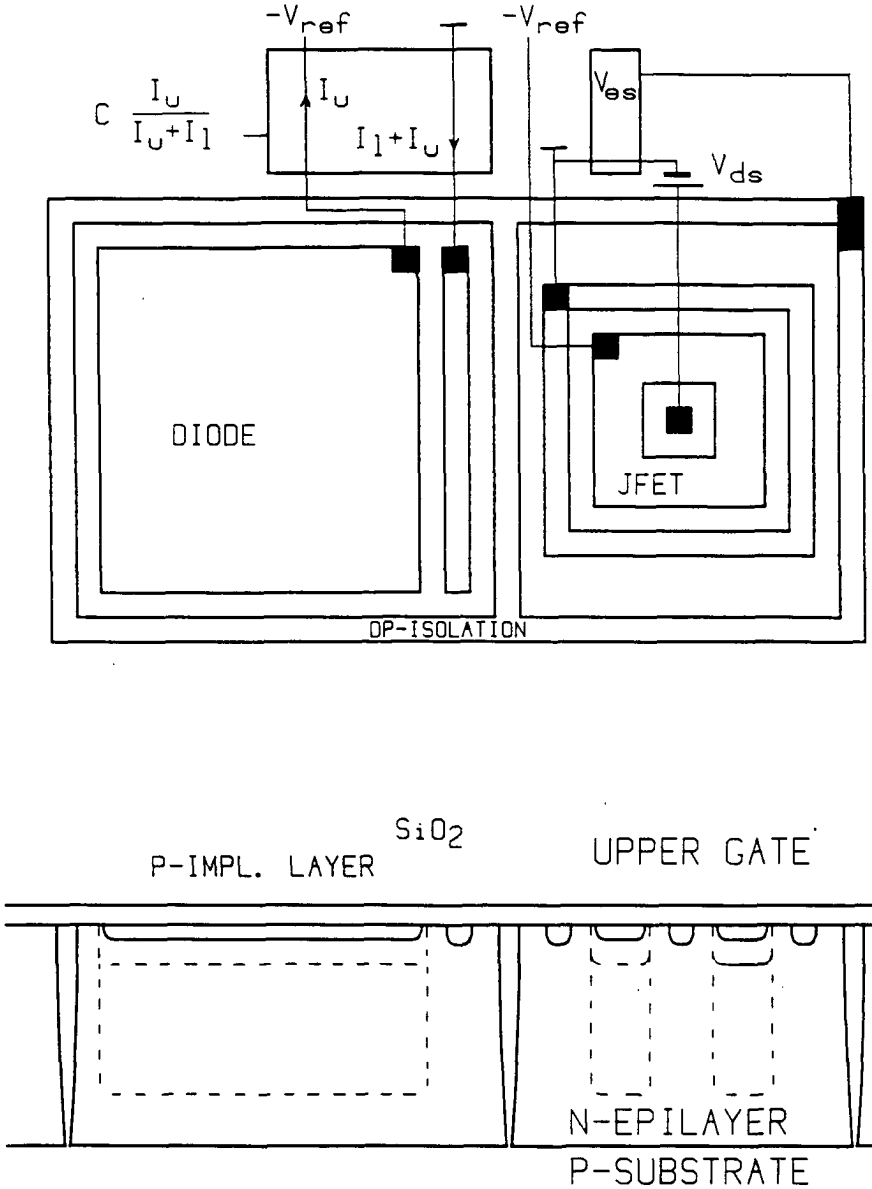


Figure 3.10 Plan and cross-sectional view of the ratio type of colour indicator including a simplified schematic of the read-out circuits.

output signal:

$$\frac{J_u}{J_u+J_l} = \frac{J_{up}+J_{ud}}{J_{ld}+J_{lp}+J_{up}+J_{ud}} \simeq \tag{3.1}$$

$$\frac{\exp(-\alpha(\lambda)\,x_{sp}) - \exp(-\alpha(\lambda)\,x_e)}{\exp(-\alpha(\lambda)\,x_{sp}) - \exp(-\alpha(\lambda)\,x_{lp}) - \dfrac{\alpha(\lambda)L_n}{1-(\alpha(\lambda)L_n)^2}\left(1-\exp(-\alpha(\lambda)\,x_{lp}\right)}$$

where x_e denotes the boundary of the upper depletion layer. For very short wavelengths where $\alpha(\lambda)\,x_{lp} > \alpha(\lambda)\,x_e > \alpha(\lambda)\,x_{sp} >> 0$ this equation reduces to:

$$\frac{J_u}{J_u+J_l} = 1 \tag{3.2}$$

For long wavelengths where $\alpha(\lambda)\,x_{sp} < \alpha(\lambda)\,x_e < \alpha(\lambda)\,x_{lp} << 1$, yet with $\alpha(\lambda)\,L_n > 1$, this equation yields:

$$\frac{J_u}{J_u+J_l} = \frac{x_e - x_{sp}}{2\,x_{lp} - x_{sp}}\,Ju \tag{3.3}$$

Hence, it appears that for a maximum range of the output signal a small value of the reverse voltage across the shallow junction is advisable. The output criterion is solved numerically using the nonappproximated equations derived in the previous chapter. The resulting response curves are shown for three different values of the upper depleted layer in figure 3.11. The flexibility introduced by the reverse voltage gives an advantage in this set-up as it serves as an electronically tunable means of obtaining more spectral information. For narrowband radiation this sensor is well suited for directly detecting the colour. However, for broadband incident radiation ambiguities can occur. Figure 3.11 shows that for small values of the reverse voltage the long wavelength part of the spectrum does not give an appreciable change in the output signal, making the distinction between red and green colours rather difficult, whereas for large values of this voltage the poor discrimination between the blue and the green becomes

apparent. This problem, obviously, becomes even more severe when dealing with mixed colours with several spectral peaks. This implies that for a single measurement, performed at a particular setting of the tuning parameter, insufficient spectral information will be supplied to ensure efficient and unambiguous colour determination. At this stage the advantage of the tuning parameter becomes apparent, because it can be used for determining the current ratio at several values of the reverse voltage. The information supplied by these successive measurements gives much more spectral information, making the detection of the above-mentioned colours possible at the expense of a more complex sensor system. Another practical problem in this method arises from the infrared component in the incident spectrum, which will give rise to a large J_{lp}. In case of an infrared peak a ratio close to zero will be observed irrespective of the colour in the visible part of the spectrum, making the use of an infrared blocking filter in front of the sensor mandatory.

3.6 Depletion of the epilayer from junctions on either side.

A general feature of the silicon colour sensors presented here is the absence of a neutral epilayer between the upper and lower depletion layers that result from the toplayer-epilayer and epilayer-substrate reverse voltage respectively. As will be shown, this measure ensures maximum separation between the response curves at different values of the reverse voltage, as no diffusion currents are present ($J_{un} = J_{ln} = 0$) that obscure the actual boundary depth. The effective boundary line will be in the middle of the neutral layer if such a neutral layer is present. So far, an ideal mechanism that controls the depletion of the epilayer underneath the upper depleted layer using the substrate voltage is assumed. This measure should result in a controllable abrupt straight-lined boundary in the epilayer parallel to the junction. In a practical device a number of limitations give rise to departures from the boundary-line model. If such closely-joined space charge regions are actually realised the lateral photocurrents will cause considerable lateral potential drops and an associated bending of the minimum equipotential curve. An accurate analysis of this effect requires a two-dimensional Poisson modeling. A simplified analysis on a square-shaped junction of $600 \times 600\ \mu m^2$ in a $6\ \Omega cm$ epilayer with a $0.1\ \mu m$ wide neutral layer between the two space charge regions, and an all around contacting, results in a worst-case voltage drop smaller than 100 mV at a current density of $5\ A/m^2$. The effect of the lateral voltage drop can, therefore, be disregarded provided that a small, however nonzero, neutral layer thickness is present.

This requirement favours the application of a JFET as the control mechanism as shown in figure 3.10. Applying a small drain-source

voltage, V_{ds}, across the JFET and simultaneously monitoring the drain current allows the detection of pinch-off in this JFET structure. If the source is at the same potential as the epilayer around the photodiode and the substrate voltage is controlled up to pinch-off, then a small neutral layer exists between the depleted layers, because the pinch-off occurs at the drain of the JFET. This mechanism operates correctly if the thickness and doping of the epilayer are constant in the part of the wafer where the sensor and the JFET are located. Therefore, serious problems are to be expected in large-area devices. These problems will be discussed in the following chapter. The thickness of the neutral layer is small compared to the epilayer thickness, so the effect on the spectral response can be disregarded. A similar reasoning can be applied with respect to the resistivity of the toplayer and the substrate. Lateral currents will give rise to lateral voltage drops that can only be disregarded if the layer resistivity is small enough. In the colour sensors realised this requirement has been fulfilled.

3.7 The compensating type of colour indicator.

The alternative colour sensor studied in detail can be classified among the compensation type. As this sensor type neccessarily has a feedback configuration in which a control parameter is tuned until a current equality criterion is met, despite the unequal setting of a second control parameter, it can also be referred to as a feedback-type of colour sensor. The sensor and the simplified schematic diagram of the read-out is depicted in figure 3.12. The sensor is composed of two identical photodiodes with one of the diodes connected to a reverse reference voltage and the other to an adjustable compensating voltage. As shown in the figure, a difference in sensitive areas is established in spite of equal physical diode areas using a controllable attenuation, k, in the photocurrent supplied by one of the diodes. The method used to apply this sensor to colour sensing is based on finding the appropriate voltage required across the compensating diode to compensate for the loss in effective sensitive area in the reference diode due to the tunable attenuation in its photocurrent. This imbalance in photocurrent can be reduced by decreasing the reverse voltage across the compensating diode, which reduces the width of the space charge region of this diode in which charge carriers can be collected and thus approaches a current balance. If the stationary condition is reached, the current supplied by the compensating diode will be equal to the attenuated current in the reference diode. The required compensating voltage will be determined by the colour, the attenuation setting and the reference voltage. In case of short-wavelength light, the boundary of the collecting layer in the compensating diode has to move upwards close to the surface, as all the

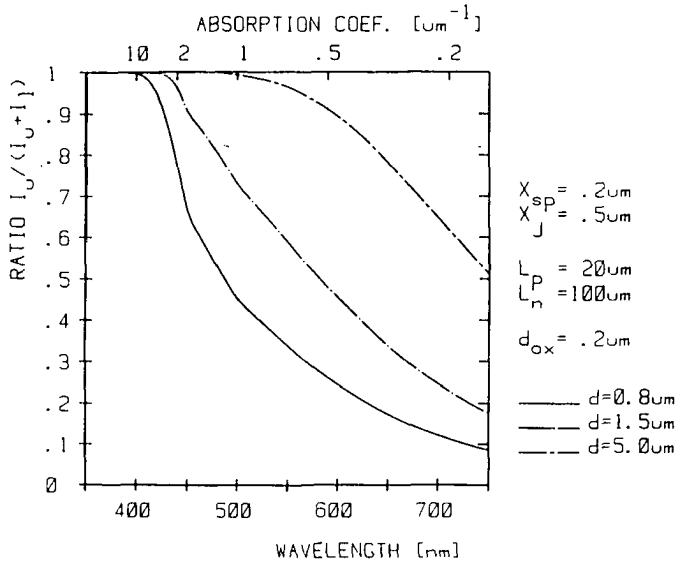


Figure 3.11 Calculated response of the ratio type of colour indicator vs. wavelength for three different values of the upper collecting layer and an epilayer-substrate junction at 7 μm.

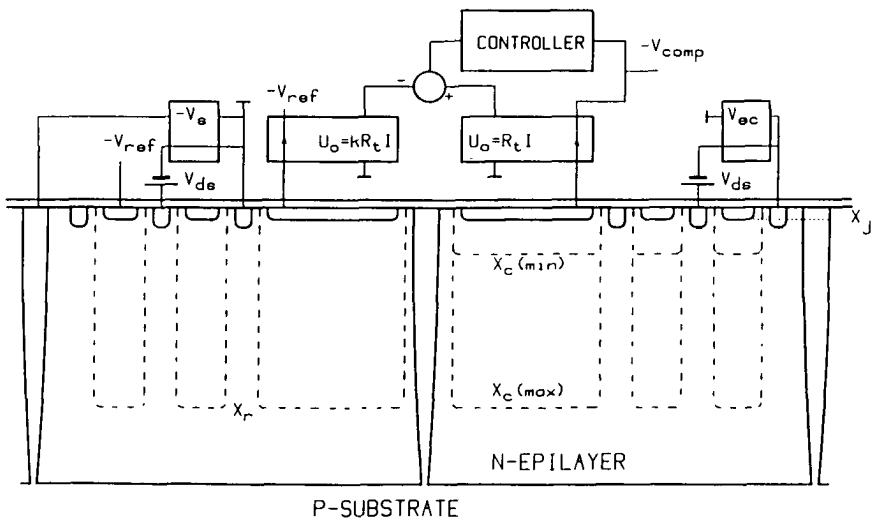


Figure 3.12 Cross-sectional view of the compensating type of colour indicator including a simplified block-diagram of the read-out circuits.

absorption of this light takes place very shallowly, so the difference in reverse voltages required for realising a certain current ratio is much more than for long-wavelength light. A controllable fraction, k, ranging from 0.5 to close to unity, of the photocurrent supplied by the reference diode is for this purpose subtracted from the compensating diode photocurrent and the remaining current is fed to a controller to generate the proper biasing across the compensating diode. At a value of the factor k much smaller than unity a larger difference in collecting depths is required for obtaining the balance. For short-wavelength light, where the difference is already large to begin with, due to the basic operating mechanism described already, there is a certain wavelength below which no balance can be maintained as the built-in depletion layer remains at a reverse voltage equal to zero. At a decreasing value of k this limit will shift to longer wavelengths. As will be shown later, this property provides an interesting means of obtaining more spectral information. An extra difficulty is introduced by the dual diode structure applied in this version of the colour sensor. The remaining part of the epilayer underneath each of the photodiodes has to be depleted up from the epilayer-substrate junction for ensuring optimum sensor operation at $J_{un1} = J_{un2} = 0$. As the two photodiodes are generally biased by different reverse voltages, their space-charge regions have a different depth and a special mechanism has to be implemented for depleting the epilayer underneath the compensating diode. This can be achieved by increasing the epilayer voltage of this compensating diode simultaneously with the operation of the feedback loop up to pinch-off of a second JFET integrated in that part of the epilayer. The two photodiodes have to be separated using a DP isolation or an extra junction, which is connected to a high reverse voltage, in between.

A more quantitative analysis of the sensor response in the stationary condition results in the following expression:

$$J_{comp} = J_{up1} + J_{ud1} = k \times J_{ref} = k \left(J_{up2} + J_{ud2} \right) \qquad (3.4)$$

Assuming $N_a >> N_d$ in the upper junction yields $J_{ud1} = J_{ud2} = J_u$ irrespective of the reverse voltage and results, after insertion of the equations derived in the previous chapter, in:

$$\exp(-\alpha(\lambda)\, x_c) = (1-k) \exp(-\alpha(\lambda)\, x_{sp}) + k \times \exp(-\alpha(\lambda)\, x_r) \qquad (3.5)$$

where x_c denotes the boundary of the depletion layer of the compensating diode associated with the compensating voltage and, likewise, x_r the

boundary in the reference diode. For $k \neq 1$ and short wavelengths, where $\alpha(\lambda)\, x_r > \alpha(\lambda)\, x_{sp} >> 0$, an x_c smaller than x_{sp} is required which is physically not possible. This confirms the inability of the feedback loop to obtain a balance for large values of the attenuation as already mentioned.

For $k \neq 1$ and long wavelength light, where $\alpha(\lambda)\, x_{sp} < \alpha(\lambda)\, x_r << 1$, the criterion reduces to:

$$x_c = k \times x_r + (1-k) \times x_{sp} \qquad (3.6)$$

This property confirms the fact that at a larger attenuation also a larger difference in reverse voltages is required and also that an attenuation close to unity is preferred for obtaining a maximum range of the output signal. The results of accurate numeric simulations of the response are shown in figure 3.13 for three different values of the area ratio. The shift of the minimum detectable wavelengths to longer wavelength at an increasing attenuation, k, can be used to solve the ambiguity which could arise at the perception of a multicoloured object as already mentioned. A blue-reddish object, at an attenuation of the reference-diode photocurrent close to unity, gives the same colour response as a uniformly greenish object.

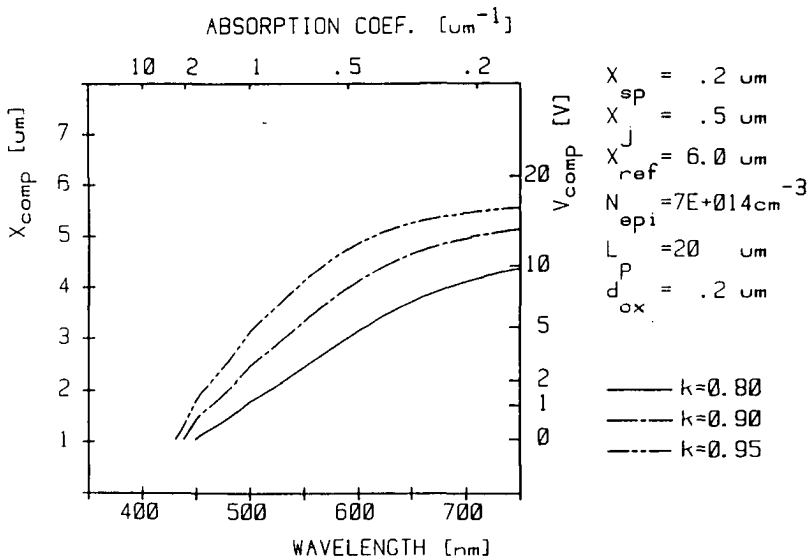


Figure 3.13 *Calculated response vs. wavelength of the compensating type of colour indicator for three values of the attenuation, k, and full depletion of the epilayer underneath the compensating diode.*

A larger attenuation emphasises the long-wavelength components of the incident spectrum and thus, in the case of multi-coloured objects the detected colour shifts to the longer wavelengths. This sensor is therefore suitable for preventing ambiguities by determining the colour for several values of the electronically tunable attenuation. In fact, it is possible to determine a kind of incremental spectrum of the incident radiation when increasing the attenuation from close to unity up to 0.5 and simultaneously monitoring the output signal.

The flexibility realised in this colour sensor is also reflected in the complexity of the read-out circuits. The feedback loop has to be implemented and also the circuits for controlling the continuous depletion of the entire epilayer underneath the compensating diode have to be realised. Moreover, the epilayer voltage has to be subtracted again from the steady-state compensating-diode reverse voltage to provide the correct colour output. These requirements seriously complicate the integrated colour sensor design, which evokes questions about the inevitability of the measure.

These complications could be circumvented if, simply, only the epilayer below both the photodiodes from the substrate-epilayer junction is depleted up to the boundary of the depletion layer of the reference diode at x_r. The width of the depleted layer of the compensating diode would be smaller, so a neutral layer would exist underneath the compensating diode in between the depleted regions as shown in figure 3.14. In this simplified arrangement the minority charge carriers generated in the neutral layer in between x_c and x_r would diffuse towards both the substrate and the upper junction as current components J_{ln} and J_{un} respectively and would, therefore, also partly contribute to the photocurrent detected in the compensating diode depletion layer. As a result a greater distance between the boundaries of the space-charge regions of the reference and compensating diode would be necessary to achieve the selected current ratio and consequently the spectral range in which the sensor gives a response would decrease. The effect would be hardly affected by the diffusion length in the neutral layer as this length is substantially larger than the width of this layer.

This rather intuitive arguing can be supported by a piece of more concrete evidence, by taking the extra diffusion photocurrent component, J_{un}, as already derived in equation 2.42 into account in the compensating diode. Adding this component to the current equality criterion for $\alpha(\lambda) L_p >> 1$ and $x_c / L_p < x_r / L_p << 1$ yields:

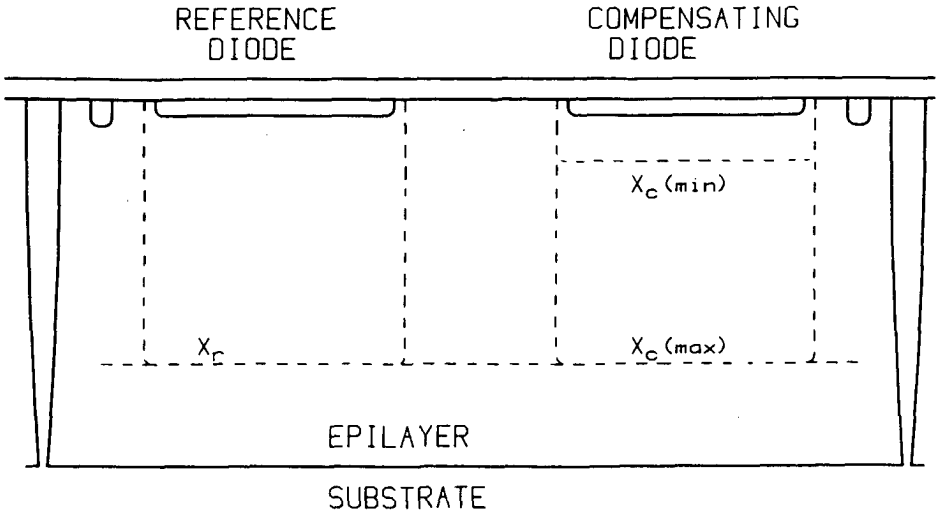


Figure 3.14 Operation of the simplified compensating type of colour indicator.

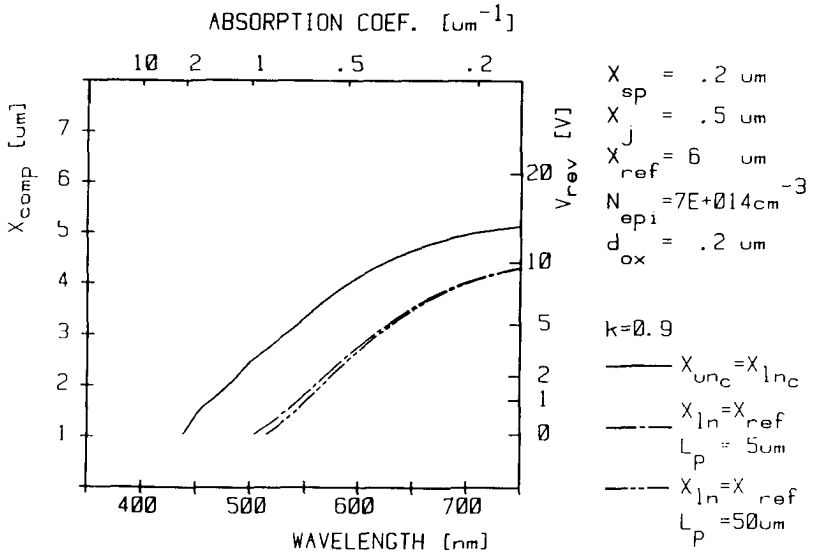


Figure 3.15 Calculated response of the simplified version of the compensating colour indicator for k=0.9 and two values of the diffusion length of holes in the quasi-neutral part of the epilayer as compared to the response in the case of full depletion of the epilayer underneath the compensating diode (solid line).

$$2\exp(-\alpha(\lambda)x_c) - \frac{\exp(-\alpha(\lambda)x_c) - \exp(-\alpha(\lambda)x_r)}{\alpha(\lambda)\left(x_r - x_c\right)} = \tag{3.7}$$

$$(1-k)\times\exp(-\alpha(\lambda)x_{sp}) - k\times\exp(-\alpha(\lambda)x_r)$$

The results of accurate numerical calculations on this criterion are depicted in figure 3.15 and clearly reveal the reduced operating range compared to the solid line and the independence of the actual value of the diffusion length in the epilayer. Even at an attenuation close to unity the operating range is restricted to wavelengths longer than 600 nm, making the full depletion of the epilayer underneath the compensating diode inevitable.

The properties of these two particular colour sensors seem to be sufficient to allow an unambiguous colour determination and, therefore, these sensors are selected for further research and realisation in a standard bipolar process. For practical verification of the dependence of the response on the parameters mentioned, a number of measurements are performed on colour sensor structures. In the following chapter the actual colour response along with the measured influence of several practical effects such as dark current and noise will be presented. The colour sensor is supposed to have a colour response unaffected by the optical intensity. However, the effects mentioned above result in a limit of detectivity which determines both the colour resolution and the minimum intensity for which the colour can be detected. After the discussion of the properties of the colour sensor the signal conditioning circuits designed for the sensor read-out will be described.

3.8 References.

3.1 R.W.Hunt, The reproduction of colour, Fountain press, 3rd edn, New York, pp. 70-95, 1975.

3.2 G.Racette, Absorption edge modulator utilizing a PN junction, Proc. IEEE, Vol 52, June 1964, pp.716.

3.3 J.A.Love and J.R.Sizelove, Low ohmic semiconductor tuned narrow bandpass barrier photodiode, US Patent 3, 452,204, June 24 1969.

3.4 C.R.Laine and W.Ford, Semiconductor radiation wavelength detector, U.S. Patent 4,011,016, March 8 1977.

3.5 N.Kako, N.Tanaka and C.Suzuki, Cumbustion detection with a semiconductor color sensor, Sensors and Actuators, Vol 4, pp.655-660, 1983.

3.6 T.Yoshikawa, Z.Tani, A.Aso and H.Kawanabe, Apparatus for sensing the wavelength and intensity of light, US Patent 4,309,604, Jan 5 1982.

3.7 N.Tanaka, N.Kako and C.Suzuki, An application of a semiconductor color sensor III, Proc. of the 3rd sensor symposium, pp. 93-96, 1983.

3.8 J.Middelhoek, S.Oosterhoff en T.A.Scholten, Halfgeleiderinrichting, N.L. Patent 8202761, 8 Juli 1982.

3.9 S.Oosterhoff and T.A.Scholten, A blue/UV sensitive silicon dual photodiode, Proc 1st Sensors and Actuators Symposium, pp. 137-142, 1982.

3.10 J.S.Hayward, B.B.Snavely and P.S.Tschang, Colour image sensing, EUR. Patent 0015711, Febr. 27 1980.

3.11 Anon, Fotodiode, NL Patent 7907416, 5 Oct 1979.

3.12 F.Koike, H.Okamoto and Y.Hamakawa, A new type amorphous silicon full color sensor, Proc. Transducers 87, Tokyo, Japan, pp. 223-226, 1987.

3.13 G.D.Cody, The optical absorption edge of a:Si:H, In: R.K.Willardson et. al.; Semiconductors and Semimetals, Vol 21B, Chap 2, pp.11-82, 1984.

3.14 K.C.Chang, C.Y.Chang, Y.K.Fang and S.C.Jwo, The amorphous Si/SiC heterojunction color-sensitive phototransistor, IEEE Electron Dev. Letters, Vol. EDL-8 No 2, pp. 64-65, Febr. 1987.

3.15 H.K.Tsai, S.C.Lee and W.L.Lin, An amorphous SiC/Si two-color detector, IEEE Electron Dev. Letters, Vol. EDL-8 No 8, pp. 365-367, Aug. 1987.

4 PRACTICAL ASPECTS AND MEASUREMENTS.

4.1 Introduction.

In the previous chapters, the basic operating principle of the colour sensor and its link with the wavelength dependent absorption coefficient in silicon has been discussed theoretically, along with several speculative device structures which are expected to reveal particular operating advantages. The performances are anticipated using a reasonably accurate modeling of the parasitic wavelength dependences in the air-Si-SiO_2 structure and the charge collecting mechanisms. Comparative computer simulations led to optimum values of the operating parameters in the sensor structure. The constraints imposed by the bipolar process were the boundary conditions in the optimisation procedure. The final design was realised in a modified bipolar process and the resulting devices were subjected to various measurements. These measurements fall into three categories: viz. those performed for confirming the anticipated effectiveness of the tunability of the charge-collecting mechanism, those for determining the magnitude of interfering quantities, such as leakage current and noise, in order to allow an estimation of the lower threshold of detectivity and colour resolution and, finally, those needed for specifying the actual performance with respect to the response; its sensitivity to processing tolerances and the remaining dependence on the optical intensity. Sufficient measurements were performed on the first two issues to justify a number of conclusions drawn in the following chapter, whereas not enough data was accumulated for setting up a statistics from which a conclusion could be drawn on the nonuniformity of the critical parameters over a wafer and on the reproducibility of the responses of devices not processed simultaneously. Therefore, only a typical response curve and a few additional remarks are made on the third characteristic.

4.2 Processing aspects.

The colour sensors were manufactured at the IC workshop of the Delft University of Technology in a modified version of the TUD-06 process [4.1]. The operation in the visible part of the spectrum, between 380 nm and 740 nm, imposes a minimum on the width of the epilayer. The response to monochromatic light with a wavelength of 740 nm, which is associated with a penetration depth of $1/\alpha(\lambda) = 6\ \mu m$, has to be maximised at the maximum size of the collection layer, so a minimum is imposed upon the size of the epilayer, in which the collecting layer can be created. Setting of the maximum response for 740 nm at maximum reverse voltage to $1 - \exp(-\alpha(740nm)(x_{un}(max) - x_{un}(min))) = 1 - \exp(-1)$ results in a

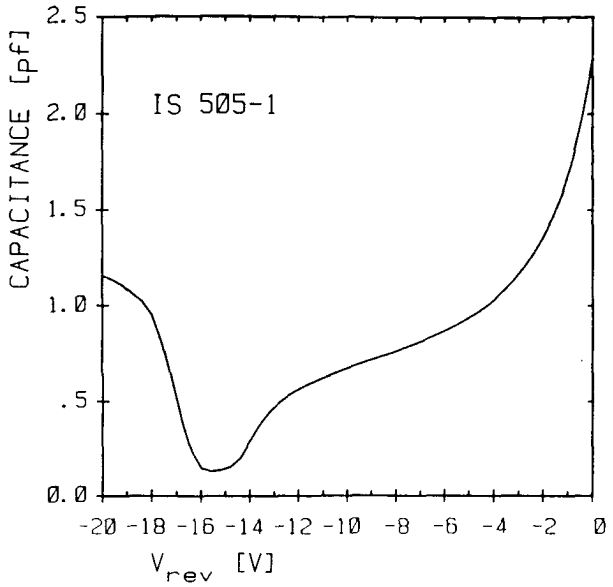


Figure 4.1 Typical C-V plot of a photodiode used in the colour sensor.

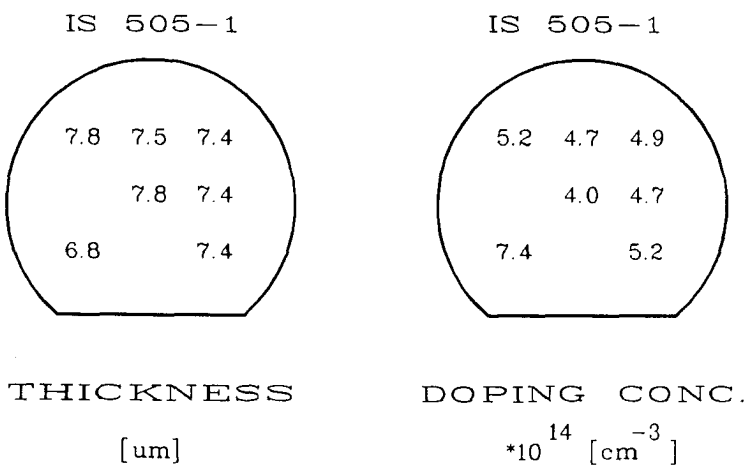


Figure 4.2 Thickness and doping concentration variations of the epilayer over wafer IS 505 as derived from C-V measurements.

minimum epilayer width equal to the penetration depth of 6 μm. The width of the built-in depletion layer of the epilayer-substrate junction has to be added to avoid reach-through. The shallow junction and the oxide growth during processing also consumes a part of the initial epilayer, so a 7-8 μm thick epilayer is grown on a standard <100> 3.6-5.4 Ωcm p-type substrate. The required epilayer doping concentration directly results from device operating conditions. For anticipating a smart sensor realisation the allowed maximum of the reverse voltage for depleting 6 μm of this epilayer is set to about 20V. Using the abrupt junction approximation for the shallow pn junction results in:

$$N_d = 2\,\epsilon_o\,\epsilon_{Si}\,V / (e\,W_{depl}^{\,2}) \approx 7.5\times10^{14}\ cm,^{-3} \qquad (4.1)$$

which results in a resistivity, ρ, equal to:

$$\rho = \frac{1}{\mu_n\,N_d\,e} = 6\ \Omega\,cm. \qquad (4.2)$$

Such epilayers can be grown reasonably well using $SiHCl_3$ at 1200 °C. The epilayer doping and width is measured using CV plots verified by sheet resistivity measurements. A typical sensor CV plot is shown in figure 4.1. The doping concentration can be derived from the slope of $1/C(V_{rev})^2$ according to:

$$N_{epi} = \frac{2}{q\,\epsilon_o\epsilon_{Si}\,\partial(1/C^2)/\partial V} \qquad (4.3)$$

The thickness can be calculated from the sharp fall of the capacitance at the complete depletion of the epilayer, when exceeding a certain reverse voltage, after correction for the built-in depletion layer at the epilayer-substrate junction [4.2]. The effect of the series resistance is disregarded. The capacitance was measured across a 620×620 μm^2 junction using a HP 4280 CV meter, which operates at a frequency of 1 MHz. The substrate is connected to ground. Up to a reverse voltage of 15 V a decreasing capacitance is observed that is due to the increased width of the upper depletion layer. At the minimum capacitance the epilayer is fully depleted and the measured capacitance is determined by the sideway capacitance. A further increase in the reverse voltage results in the epilayer-substrate junction to be pulled in forward conduction and a capacitance between epilayer terminal and top layer is measured that is determined by the epilayer-substrate capacitance, in the part of the epilayer that is not covered by the upper junction, in series with the

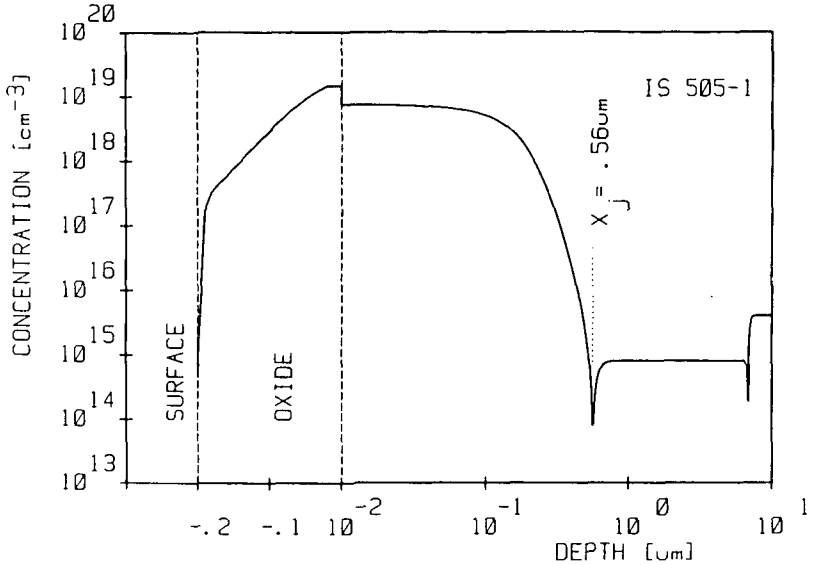


Figure 4.3 SUPREM simulation of the doping profile in IS 505-1; 60 keV / 2×10^{14}cm^{-2} B$^+$ implantation in a 6 Ω cm 7 μm thick epilayer after 30 min. anneal at 1000 °C. (Note the lin-xaxis in the oxide and the log-xaxis in the silicon)

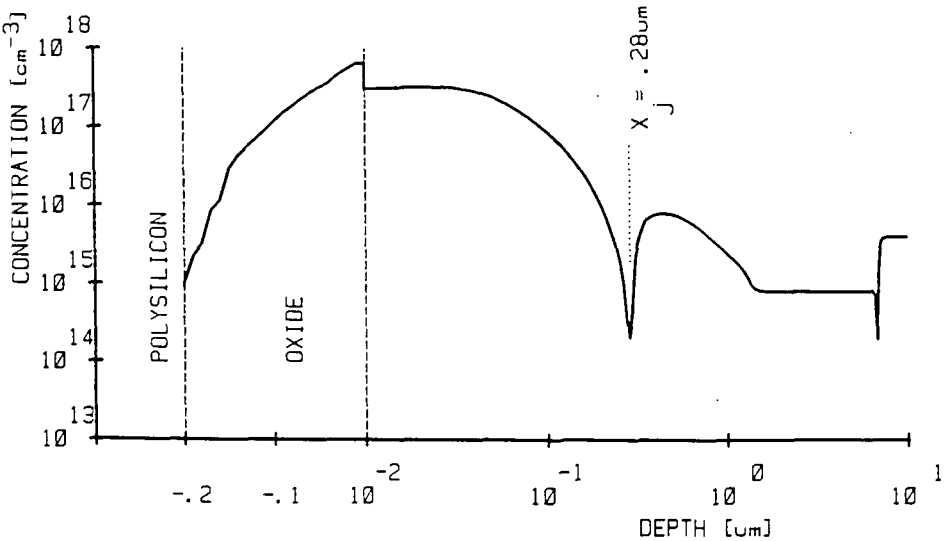


Figure 4.4 SUPREM simulation of the doping profile in IS 505-2; First a buffer layer is formed by a 150 keV / 2×10^{12} cm^{-2} As$^+$ implantation followed by a drive-in at 1200 °C for 60 min. and concluded by a 60 keV/5×10^{12} cm^{-2} B$^+$ implantation and an anneal for 30 min. at 800 °C.

depletion-layer capacitance. The measurements indicate a distribution of the thickness and doping concentration as shown in figure 4.2. The actual doping concentration will be slightly larger due to the abrupt junction approximation. However, the importance of this measurement lies in the nonuniformity of the epilayer characteristics rather than in the absolute value. As shown, the non- uniformity over the wafer is considerable. The nonuniformity of the doping concentration directly affects the relation between the reverse voltage and the size of the depletion layer, whereas the maximum size of the space-charge region is limited by the actual width of the epilayer. The doping gradients in a wafer lead to unequal collecting layers in different devices at equal reverse voltages and thus lead to nonuniformity in the spectral response, whereas tolerances in the epilayer width determine the maximum size of the collecting layer. Therefore the epilayer nonuniformities will have a great effect on the performance of the colour sensor.

For obtaining an optimum response at short-wavelength visible light the maximum in the doping concentration of the shallow boron layer should occur at the surface. This requirement would favour the use of a diffused layer, however after deposition, during the subsequent drive-in and oxidation, a large amount of the boron would redistribute into the oxide, making it difficult to realise shallow junctions. Moreover, a boron implantation step reveals a better compatibility with the process used. For this reason a boron implantation through an 0.2 μm oxide was performed with an implantation energy associated with a projected range slightly smaller than the oxide thickness and a dose sufficiently high to ensure reliable contacting while low enough to avoid high-doping effects. The most suitable profile emerges after a 60 keV 2×10^{14} cm^{-2} B^+ implantation followed by a 30 min anneal at 1000 °C. A SUPREM simulation [4.3] of the resulting doping profile is shown in figure 4.3 and reveals a surface concentration equal to 5×10^{18} cm^{-3} and a junction at 0.56 μm. These devices were realised in design IS 505-1. The log-x axis is rather unconventional, however, it gives a better presentation of the impact on the wavelength, because of the nearly exponential relation between absoprtion coefficient and wavelength.

The low epilayer doping results from a long-wavelength constraint that again results from the depletion of the required layer of silicon, using values of the reverse voltage as mentioned before. However, the lower limit of the tunability of the upper junction space-charge region is determined by the built-in depletion layer, which increases with decreasing epilayer doping. The 7.5×10^{14} cm^{-3} epilayer results in a built-in depletion layer of almost 1 μm width, giving a total non-controllable upper limit of the tunable collecting layer of 1.5 μm. A

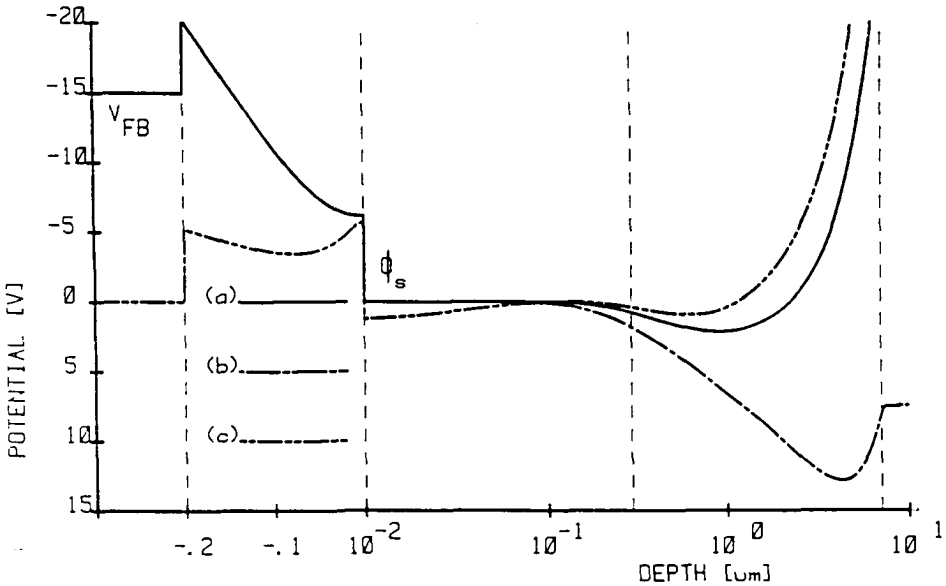


Figure 4.5 Electrostatic potential for (a): a gate voltage equal to the flatband voltage and 1V across the junction, (b) Flatband and 10V across the junction and (c): zero gate voltage and shortcircuited junction.

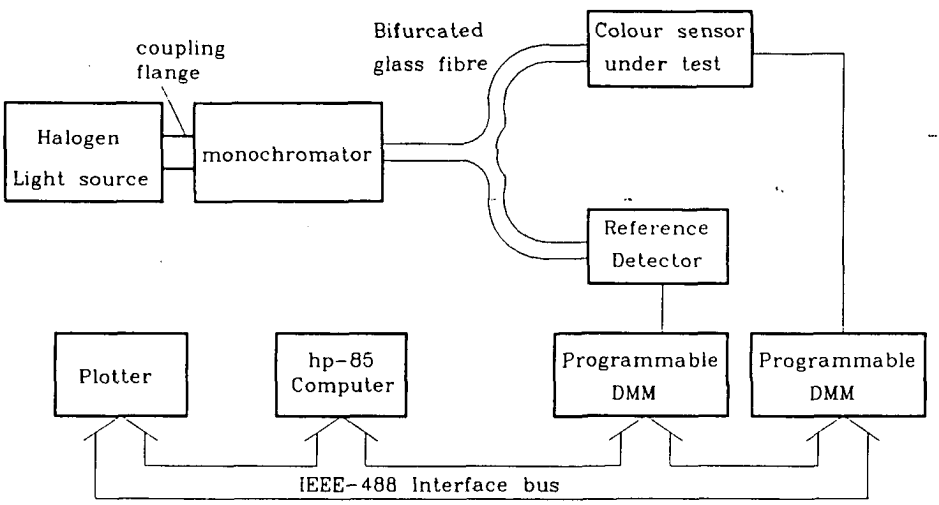


Figure 4.6 Diagram of the measurement set-up.

penetration depth of 1.5 μm in silicon is associated with a wavelength of about 550 nm, so this short wavelength constraint would prevent the distinction between colours of lower wavelength.

This serious impediment can be circumvented by locally increasing the epilayer doping concentration just underneath the junction using a As^+ implantation before a B^+ implantation. In the processing sequence of wafer IS 505-2 first a 150 keV 2×10^{12} cm^{-2} As^+ implantation is performed followed by a high-temperature step to drive the As away from the surface before oxidation and a 60 keV 5×10^{12} cm^{-2} B^+ implantation. After a 30 min anneal at 800 °C a doping profile will emerge as shown by the Suprem simulation in figure 4.4. The junction is now at 0.28 μm and the width of the built-in depletion layer is reduced to about 0.4 μm giving a total noncontrollable upper layer size of 0.7 μm. This upper limit of the tunable collecting layer corresponds to the penetration depth at 480 nm, which gives a much better perspective for practical application of the silicon colour sensor.

Apart from the maximising of the effect of this single tuning parameter by pushing back the limits of the upper collecting layer, a different strategy can be applied. This alternative approach involves the introduction of a second tuning parameter that is pre- eminently effective in the short-wavelength range. Promising possibilities lie in the control of the size of the surface space charge region using either a polysilicon gate on top of the 0.2 μm oxide or an extra As^+ implantation in this oxide. Applying a negative gate voltage exceeding a certain threshold value brings about the flatband condition at the $Si-SiO_2$ interface, which results in surface accumulation. The absence of the surface space-charge region will result in an improved short-wavelength response. However, switching the gate to zero voltage results in surface depletion and a poorer response. Similarly, implanting an extra oxide charge in the oxide results in an increased surface space-charge region and a reduced short-wavelength response. Therefore, the short-wavelength response is effectively tuned by the control of the surface charge. These improvements have been studied theoretically, however, no reliable measurements are available because of a strongly increased leakage current in the shallow junctions in devices realised according to the doping profile in figure 4.4 (IS 505-2). The advantage of the control of the surface depletion layer using a gate voltage connected to the polysilicon gate, rather than using an extra ion implantation, lies in the flexibility of the electronic control, however, the operation in the short-wavelength range is seriously hampered by the absorption in the gate.

The effectiveness of the short-wavelength tuning parameter can easily be confirmed from SUPREM simulations of the implanted impurity profile.

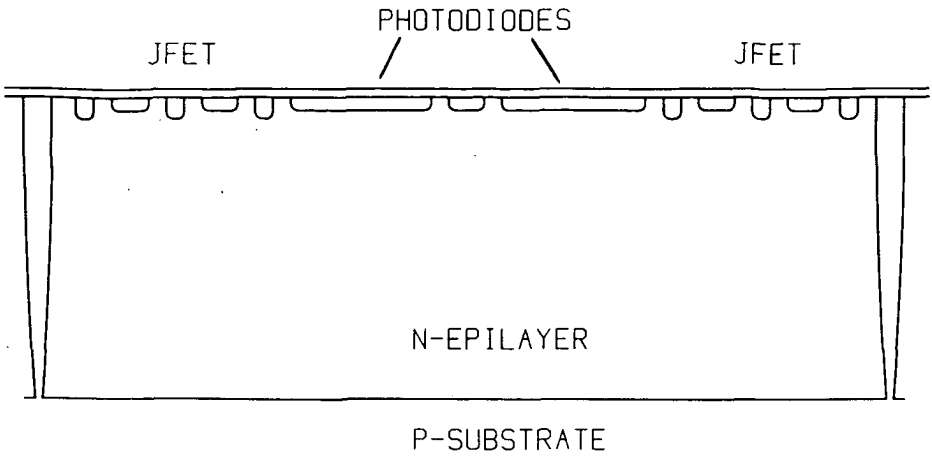


Figure 4.7 Cross-sectional view of the device used for measurements on the basic sensor operation (note that the horizontal and vertical axes have a different scale).

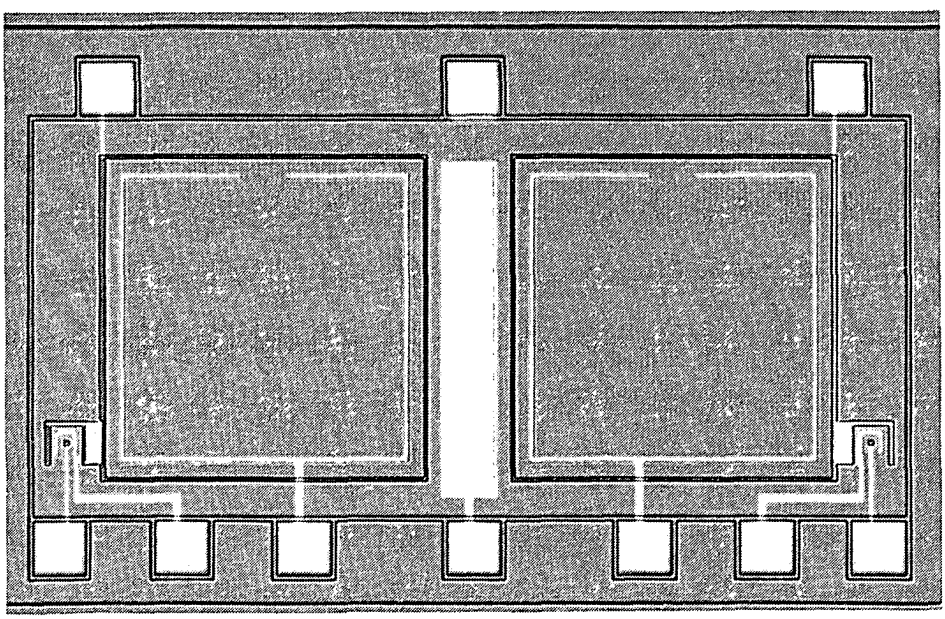


Figure 4.8 Photograph of the basic colour sensor structure.

The electrostatic potential can be obtained from the impurity profile using the Poisson's equations and the result is shown in figure 4.5. This figure shows the electrostatic potential vs. log. depth for three different values of the reverse voltage across the junction and two different gate voltages. The potential minimum in the shallow p-type layer corresponds to x_{sp} and the reverse-voltage controlled maximum in the n-type epilayer with $x_{un} = x_{ln}$ and the potential at the end point with the substrate potential. The position of x_{sp} is shifted towards the junction at an increasing oxide charge or a less negative gate voltage, whereas the position of $x_{un} = x_{ln}$ remains affected only by the junction reverse voltage and shifts towards the substrate at higher values. The control of the surface space-charge region is, therefore, expected to provide a designer with an additional independent short-wavelength tuning parameter. The realisation of the shallow junctions using an implantation through the oxide for optimising the original control parameter was an unfortunate choice with respect to the short-wavelength control parameter, as a relatively large positive oxide charge is already present, resulting in a reduced effect of the extra positive oxide charge or a high value of the flatband voltage. Due to the leakage current in the alternative devices, the measurements presented in the following sections will always refer to a bonded device realised in wafer IS 505-1.

4.3 Verification of the basic operating mechanism.

The first measurements are performed to confirm the basic colour sensing operation. A researcher is usually anxious to obtain such information in an early stage of the research for releasing him of the agony about a possible error in the initial idea. This idea is, at that stage, usually not free from a considerable amount of optimism which is demonstrated by the relatively large number of asssumptions and ommissions; resulting in a temporarily strained relationship with scientific caution.

The optical measurement set-up used for all measurements was composed of a monochromator-illuminator with a halogen light source, a bifurcated glass fiber and a reference detector as shown in figure 4.6. The bifurcated silica glass fiber guides the light to both the reference detector and the device under test. In the measurements a calibrated silicon photodiode was used as a reference. The measurements were controlled by a HP 85 computer via an IEEE-488 instrumentation bus. The inaccuracy is mainly determined by the optical reference, which is rated at 2%.

The measurements were performed on the basic colour sensor structure shown schematically in figure 4.7 and also in the photograph in figure 4.8. This structure consists of two photodiodes each with an integrated JFET

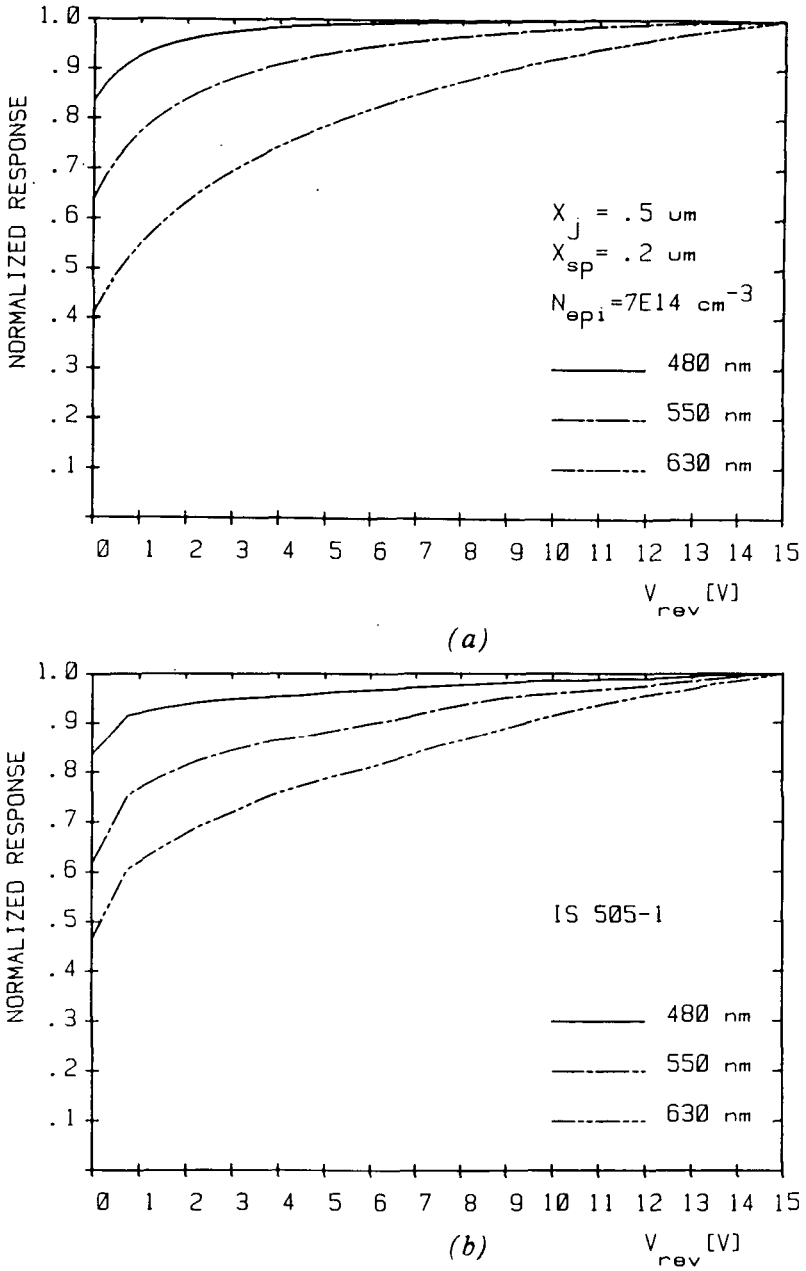


Figure 4.9 Normalised response vs. reverse voltage and complete depletion of the epilayer at three monochromatic wavelengths, (a) calculated, (b) Measured for IS 505-1.

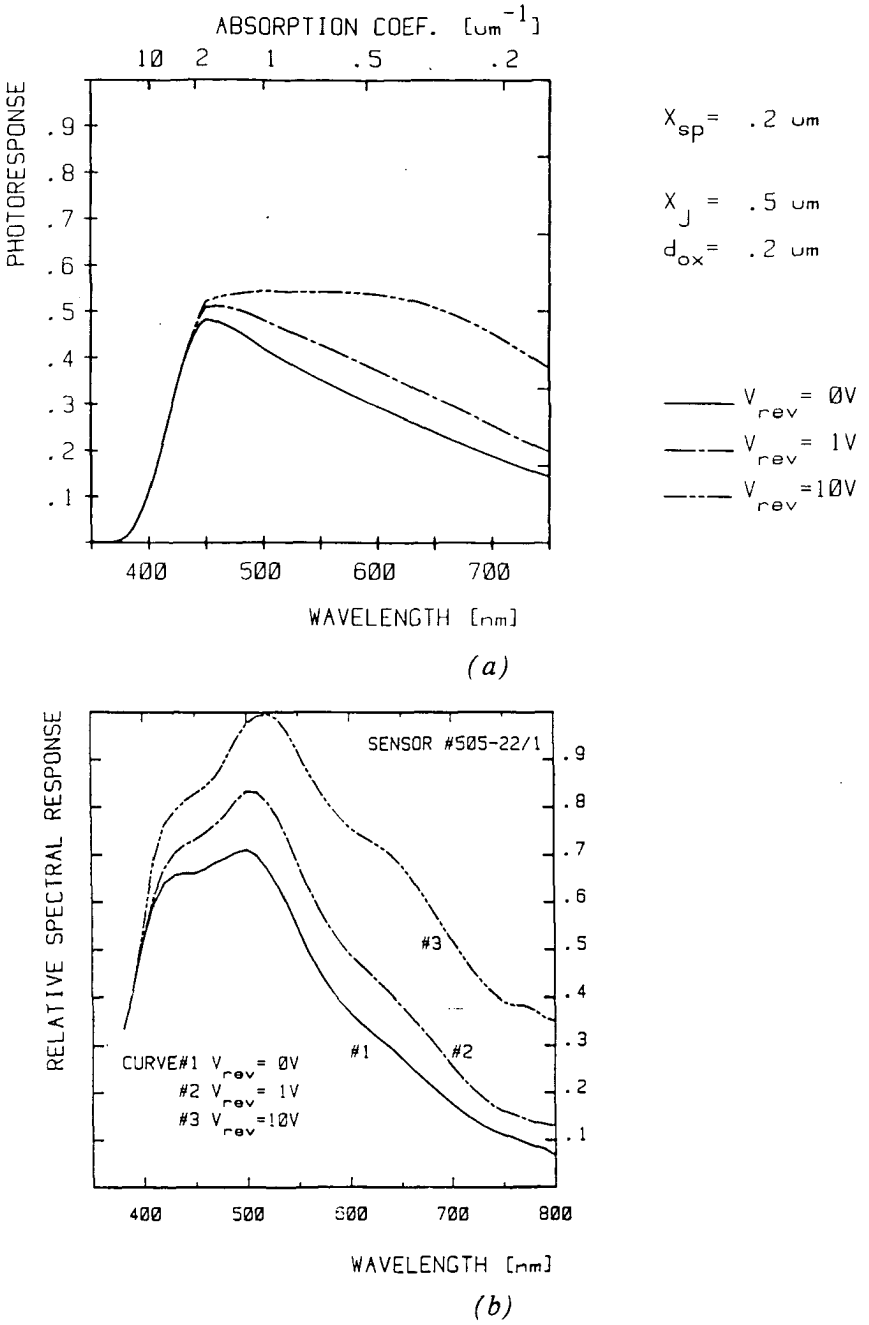


Figure 4.10 Spectral responses for reverse voltages equal to 0 V, 1 V and 12 V of (a) Calculated and (b) Measured on IS 505-1.

and separated by a junction in the middle. The latter junction is realised with the same implantation as that used for the photodiodes and, when applied to a high reverse voltage, ensures an effective isolation between the photodiodes. The dimensions of a junction are 620 x 620 μm, giving a sensitive area of the photodiode equal to 600 x 600 μm. For a given epilayer doping ($N_d = 7\times10^{14}$ cm^{-3}), junction depth ($x_j = 0.56$ μm) and oxide charge (giving $x_{sp} = 0.2$ μm) the response vs. the reverse voltage can be calculated for complete depletion of the epilayer using the JFET-controlled substrate voltage and the results are depicted in figure 4.9a for three different monochromatic wavelengths. Devices realised in wafer IS 505-1 were manufactured according to these specifications and measurements on the normalized response vs. the applied voltage are shown in figure 4.9b for three different monochromatic wavelengths. The curves are normalised on the response at 15 V reverse voltage and the measurements are in reasonable accordance with the theory. The lower x-axis indicates the applied reverse voltage and the upper x-axis the required substrate voltage for a drain current equal to 1 μA at $V_{ds} = 1$V. The effect of epilayer noninhomogeneities becomes particularly apparent at larger values of the reverse voltage. An epilayer doping smaller than the nominal value results in a larger depletion layer thickness and a quicker saturation of the response at increasing reverse voltage. Also the dynamic range of the tuning parameter is reduced due to the enlarged built-in depletion layer, which causes the normalized response vs. wavelength to start at a higher value.

These measurements confirm the predicted effect of the tuning voltage on the spectral response. For practical colour filtering the response to a spectral scan is important. Figure 4.10 a shows the result of calculations on the response to monochromatic wavelengths in between 350 nm and 750 nm for three different values of the reverse voltage using the equations derived in chapter 2 with $x_{un} = x_{ln}$ and an epilayer doping equal to 7×10^{14} cm^{-3}. Figure 10 b shows the results of measurements on IS 505-1 using the same values of the reverse voltage. The increasing long-wavelength response at a higher value of the reverse voltage is clearly demonstrated. IS 505-2 was also subjected to the same measurement and, although the results are unreliable as a result of the higher leakage current, the enlarged epilayer doping beyond the junction and the oxide charge indeed confirm the anticipated higher degree of distinction between the curves measured at a reverse voltage of 0V and 1V.

4.4 Leakage current and noise.

The reproducibility of the measured photocurrent, and thus the lower limit of the useful sensitivity, is determined by non-photon generated

charge carriers and the fluctuations therein. The two main causes for the uncertainty in the response of a photodiode are thus the leakage current and the noise. The leakage current gives an offset and the noise causes random fluctuations in the photocurrent. The dominating noise source in a photodiode is the shot noise, which is due to the random fluctuations in the number of photons arriving at a photodetector per unit of time interval. The mean square noise current is proportional to the diode current and, thus, the lower threshold of detectivity is determined by the leakage current and the noise it generates. Any noise generated in the signal conditioning circuits or added by operation on sensor signals will reduce the signal-to-noise ratio.

The leakage current flowing through the upper junction of the basic colour sensor structure is determined by the thermal generation current in the upper depleted part of the epilayer. Due to the depletion of the remaining epilayer no diffusion of thermally generated minority charge carriers need be taken into account. The leakage current density is equal to:

$$J_d = q G \left(x_{un} - x_{sp} \right), \qquad (4.4)$$

where G denotes the generation rate that is related to the the intrinsic carrier concentration, n_i, and the generation lifetime, τ_e, by $G = n_i / \tau_e$. As the intrinsic carrier concentration is determined by the bandgap the generation rate can be obtained from lifetime measurements. Equation 4.4 shows that the leakage current depends in a first approximation linearly on the width of the depletion layer. From this conclusion the course of the leakage current vs. the reverse voltage can be anticipated. When short-circuiting the upper junction a leakage current is measured depending on the width of the built-in depletion layer and increases proportionally with the root of the reverse voltage applied.

The results of a leakage current measurement is shown in figure 4.11 for two different operating conditions. The leakage current is measured for both the regular colour sensor operation, in which closely joined space-charge regions deplete the entire epilayer with a boundary at $x_{un} = x_{ln}$, and for a shortcircuited epilayer-substrate junction, $V_{es} = 0$. For a shortcircuited shallow junction a small leakage current of opposite polarity is observed, which is due to reach-through effects from the substrate. The high substrate reverse voltage required for depleting the thick remaining epilayer results in a potential maximum in the epilayer close to the shallow junction and a potential barrier equal to the built-in voltage. The difference between the two curves at low values of the reverse voltage

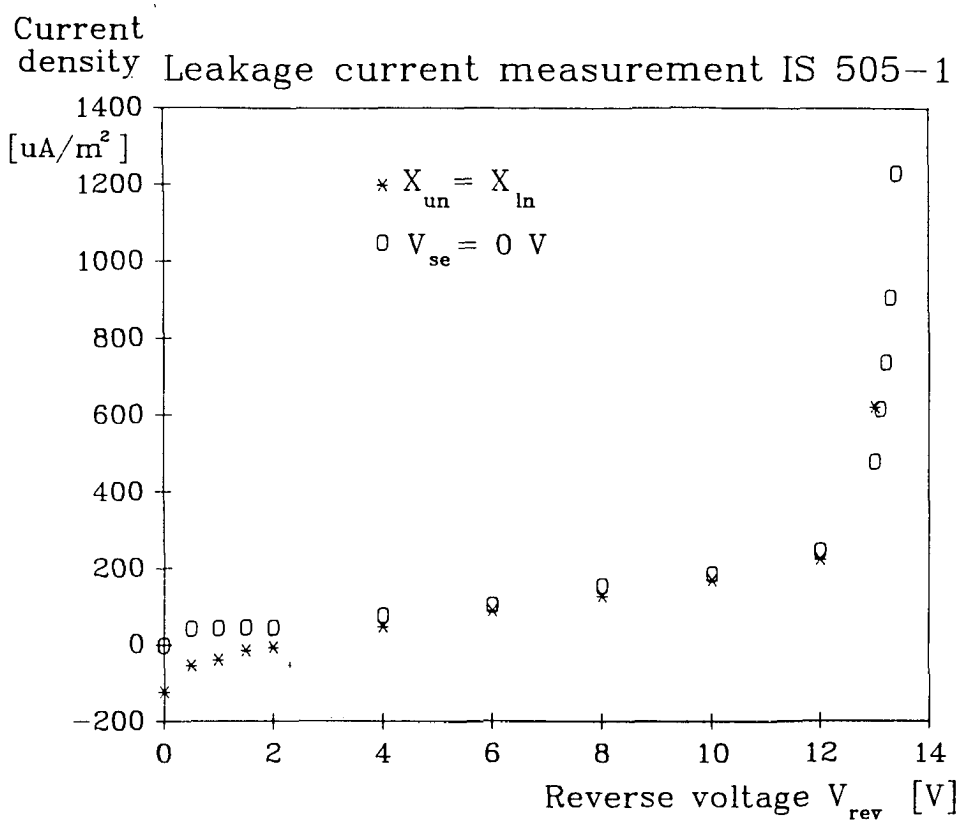


Figure 4.11 Leakage current vs. reverse voltage in IS 505-1.

indicates the extra current from the diffusion of thermally generated holes in the quasi-neutral part of the epilayer to the upper depletion layer. The width of this neutral layer, obviously, reduces with an increasing upper depletion layer and the curves will coincide at full depletion of the epilayer. The increasing leakage current proportional to the square root of the increasing reverse voltage has indeed been confirmed up to a reverse voltage of about 12V. When biasing the shallow junction with a reverse voltage beyond this value flatband occurs at the epilayer- substrate junction. This implies that the latter junction is pulled into forward biasing and holes generated in the substrate that are able to diffuse to this junction will also reach the top layer resulting in a sudden increase in the leakage current when exceeding this reach-through voltage.

The shot noise in a photodiode is related to the diode current, I_p, by the relation:

$$i_n^2 = 2 e I_p B, \qquad (4.5)$$

where B denotes the bandwidth. Usually the noise equivalent power is defined as the optical intensity required for obtaining an output voltage equal to the rms shot noise per unit bandwidth with a dark current I_d and can thus be formulated by:

$$NEP = \sqrt{2 e I_d / R} \quad (W / \sqrt{Hz}), \qquad (4.6)$$

where R denotes the current responsivity of the photodiode (A / W) that is related to the quantum efficiency by the relation $R = (e / h \nu) \times \eta$. The responsivity R is usually wavelength-dependent and in the photodiode under consideration also a wavelength-dependent NEP can be expected. Moreover, since the control of the collecting layer between a minimum and maximum, lying as far as possible apart for proper colour sensor operation, a NEP is expected that also depends on the value of the reverse voltage. The actual $NEP(\lambda, V_{rev})$ can be obtained by combining the spectral responses vs. the wavelength at the various settings of the reverse voltage. For short wavelenghts the NEP increases with the reverse voltage across the junction, due to the increasing leakage current and the almost constant response. For longer wavelengths the NEP will decrease at an increasing reverse voltage, due to the increased long-wavelength response. The signal- to-noise ratio is equal to $I_{ph} / \sqrt{2e I_{ph} B} = \sqrt{I_{ph} / 2e B}$ resulting in a reduction in the SNR at a decreasing response. In the simple colour indicator a bandwidth, B, smaller than 1 kHz is more than adequate, giving a threshold of optical power at about 1 nW. However, for colour imaging applications a much wider bandwidth is required resulting in a higher noise level.

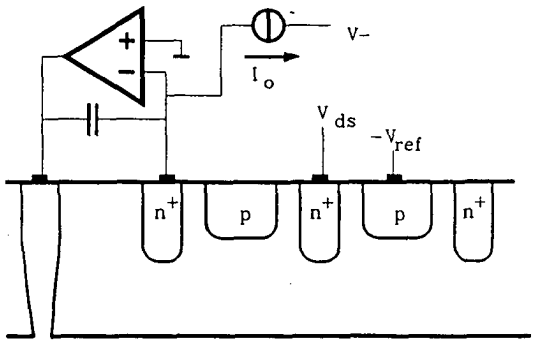


Figure 4.12 Simplified diagram of the JFET-controlled feedback loop for generating the appropriate substrate voltage.

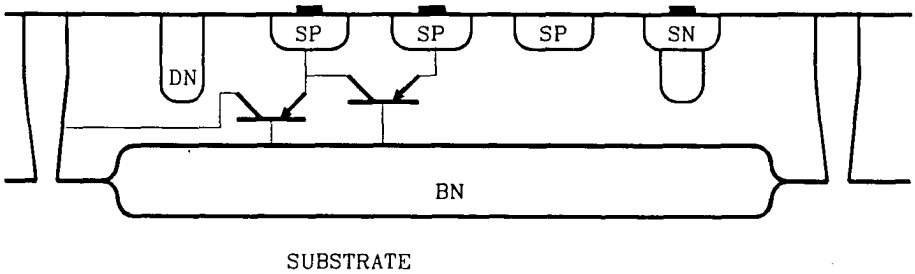


Figure 4,13 Lateral-pnp with parasitic substrate-pnp.

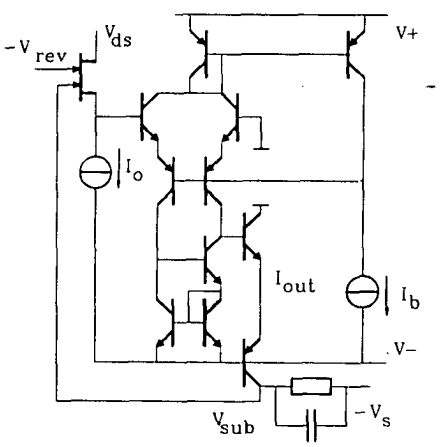


Figure 4.14 Actual circuit of the JFET-controlled feedback loop.

4.5 Colour indicator circuits.

A unifying characteristic of both the colour indicators mentioned is the necessity for a circuit for depleting the remaining part of the epilayer using a JFET-controlled feedback loop for providing the substrate voltage. The general schematic diagram of this circuit is depicted in figure 4.12. When assuming ideal opamp characteristics, the magnitude of the substrate voltage is increased until the drain current becomes equal to the offset current I_o. The use of this offset current is required to be able to reduce the absolute value of the negative substrate voltage again, at an increasing reverse biasing of the shallow junction, as otherwise reach-through might occur. Preventing the reverse voltage from adopting a value beyond the reach-through voltage is a general operating constraint. At a reverse voltage exceeding this value a large fraction of the charge carriers generated in the substrate are collected in the upper junction, resulting in a loss of the colour information. Moreover, the response should be fast enough to avoid the occurence of this condition, which is another reason for imposing a lower limit to the drain current in the integrated JFET.

The actual circuit is slightly complicated by the fact that the substrate should be at the most negative potential. Therefore, direct driving from the output of an operational amplifier is not possible. Moreover, special care has to be taken to avoid lateral-pnp transistor in critical places in the design. The reduced epilayer doping concentration results in a reduced maximum collector-base reverse voltage and a higher current gain. The usual lateral-pnp specifications can be obtained by increasing the distance between the collector and emitter SP implants to 10-12 μm, which results in an increased maximum collector base voltage at the expense of a reduced forward current gain, β_F. A more severe problem results from the increased epilayer width combined with the decreased doping concentration compared to original bipolar process. This results in a nonnegligible current gain in the parasitic pnp transistor composed of SP (emitter), epilayer (base) and substrate (collector). The improved lifetime of minority charge carriers in the lower doped base region results in larger currents from the collector of the lateral pnp to the substrate. As the DN diffusion reaches down 4 μm, the encompassing of the lateral pnp collector using this diffusion together with a far-extending BN layer, as shown in figure 4.13 only partly solves the problem.

The actual circuit is shown in figure 4.14 and consists of a standard opamp input stage in which the output current, I_{out}, drives a substrate-pnp for obtaining the desired voltage drop across the external resistor. The absolute value of the negative voltage V_{sub} should be large enough to allow the full depletion of the epilayer under all conditions.

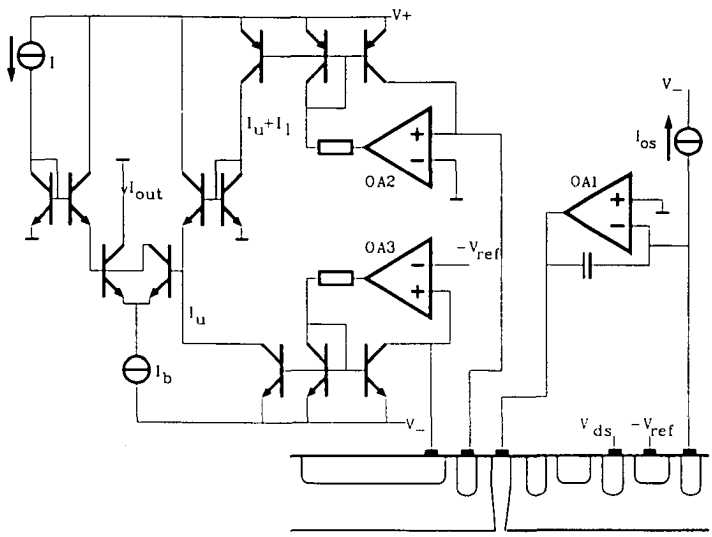


Figure 4.15 Simplified diagram of the correction type of integrated colour indicator.

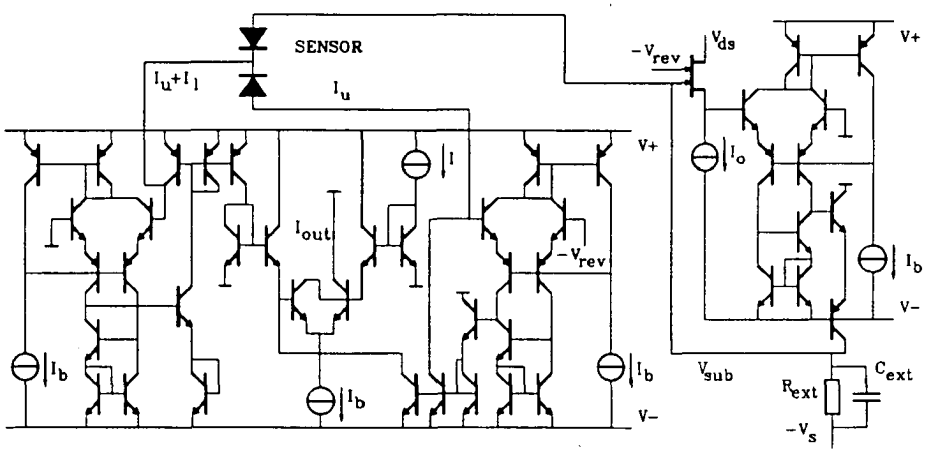


Figure 4.16 Circuit diagram of the actual integrated ratio type of colour indicator realised in design IS 505.

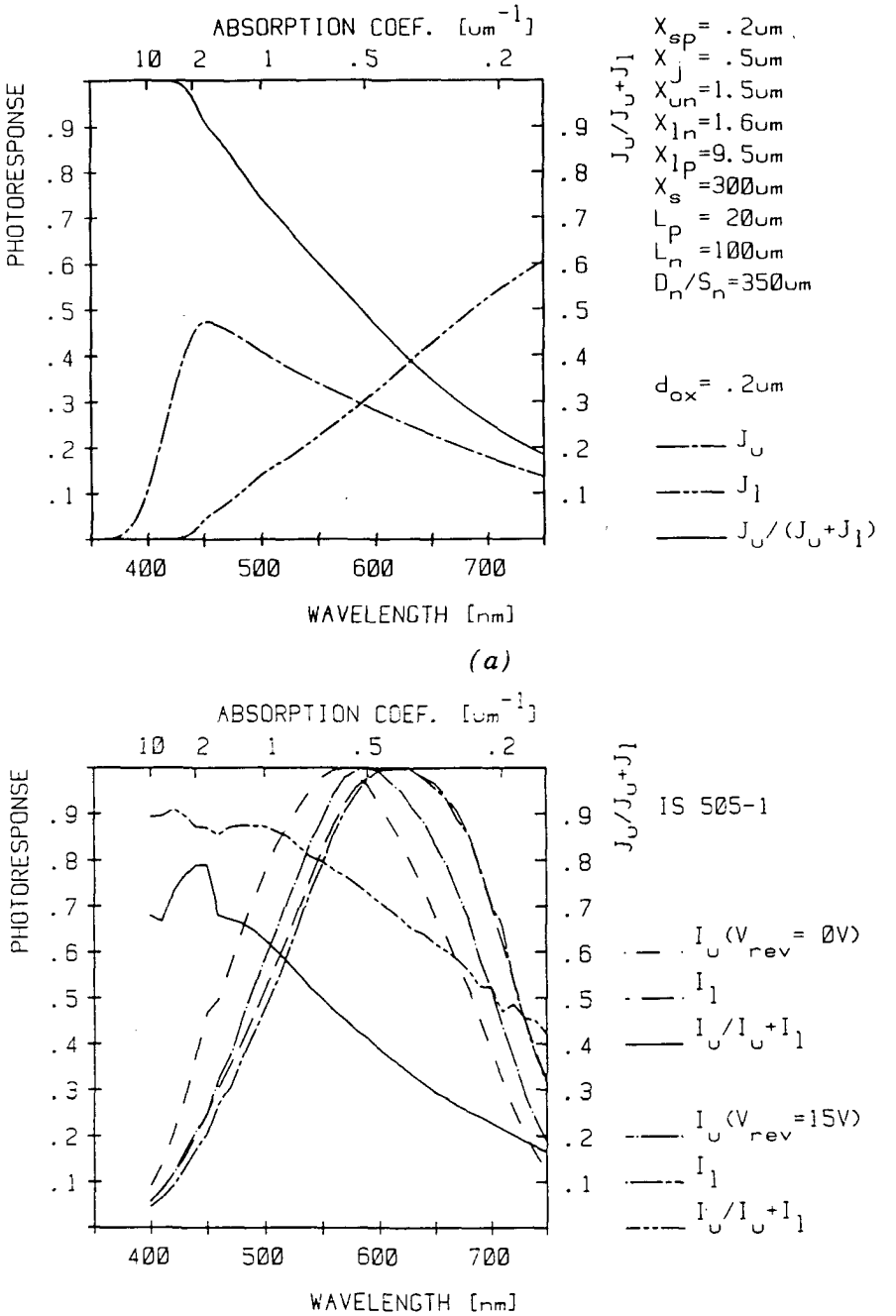
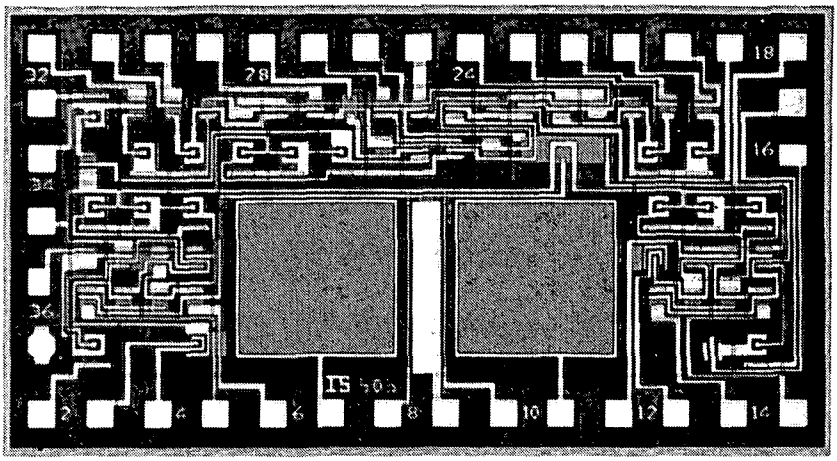
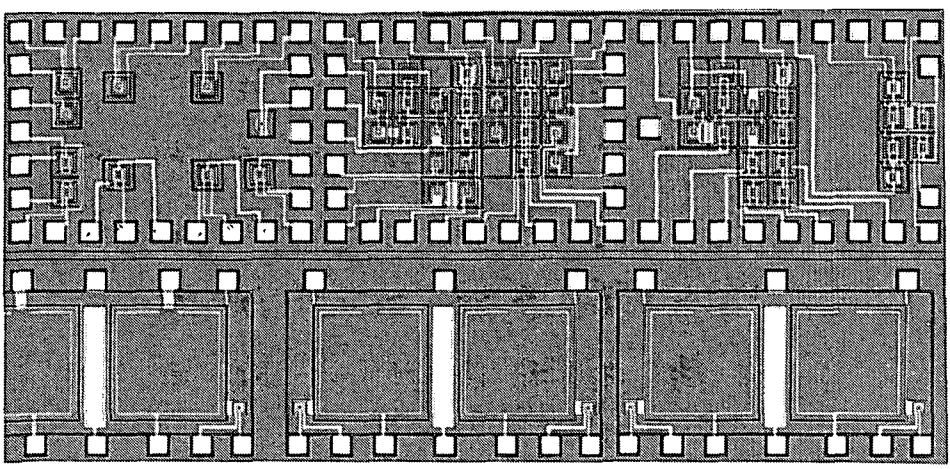


Figure 4.17 calculated (a) and measured (b) spectral responses of the ratio type of colour indicator with device dimensions as shown.



(a)



(b)

Figure 4.18 Microphotographs of (a) The ratio type of colour indicator and (b) of the main modules.

The simplified schematic diagram of the integrated correction-type of colour indicator is shown in figure 4.15. Operational amplifier, OA1, denotes the circuit used for generating the correct substrate voltage. Opamp OA2 pulls a current from the pnp current mirror sufficient for compensating for the photocurrent flowing into the epilayer in such a way that no current flows to the noninverting input of the opamp. A similar operation applies to the opamp OA3 with respect to the photocurrent flowing from the shallow junction. When there are identical transistors in the current mirrors these currents will also flow to the translinear circuit in which the division $I_u / (I_l + I_u)$ is performed. The translinear circuit is composed of a lefthand and a righthand core that both consist of three stacked transistors. The emitters of the outer core transistors are both connected to ground and the emitters of the inner transistors are connected to a common current source. Therefore, the sum of the base-emitter voltages across the lefthand and righthand core are equal. Due to the exponential relation between the base-emitter voltage and the collector current, the products of the collector corrents in the cores are equal. A feedback between the lefthand and righthand core elliminates the effect of the biasing current I_b, and when disregarding the base currents an output current equal to $I_u / (I_l + I_u)$ is realised.

The actual circuit is shown in figure 4.16 and clearly shows the identical input stages used for the operational amplifiers. The pushing and pulling of a current to a npn or pnp current mirror is implemented using a pnp or npn transistor connented to the npn active load. The response of this colour indicator can be anticipated using the expressions for the current density components as derived in chapter 2 with $J_{un} = J_{ln} = 0$. Insertion of the actual dimensions give the simulation results depicted in figure 4.17 a. Figure 4.17 b shows the actual response of the indicator to a monochromator scan measurement for two different values of the reverse voltage across the shallow junction. In this measurement the monochromator is driven by an uncalibrated halogen light source resulting in normalized responses in the upper and lower junctions as shown in the figure. The ratio is not affected by the spectral distribution of the illuminating source. The deviation of the ratio from unity at short wavelengths is due to the substrate leakage current and to the fact that the epilayer-substrate junction is not entirely covered by the shallow layer. Also the high infrared sensitivity of the epilayer substrate junction results in a response at short wavelengths smaller than unity, due to spurous infrared radiation. A microphotograph of the design is shown in figure 4.18 a. A two-diode design is realised in order to be able to test the lateral ratio type of colour indicator, shown in figure 3.4 a, as well. For testing purposes the design has also been split into a few main modules shown in the photograph in figure 4.18 b.

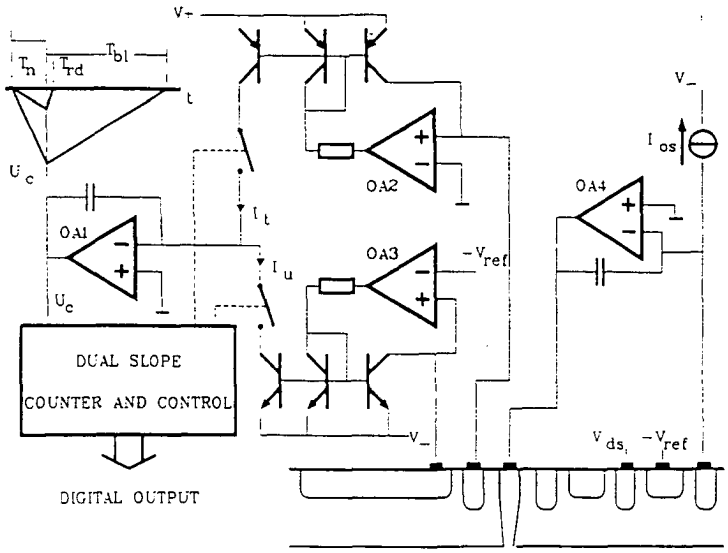


Figure 4.19 Simplified diagram indicating how the operation of the ratio type of colour indicator can be combined with a dual-slope AD converter.

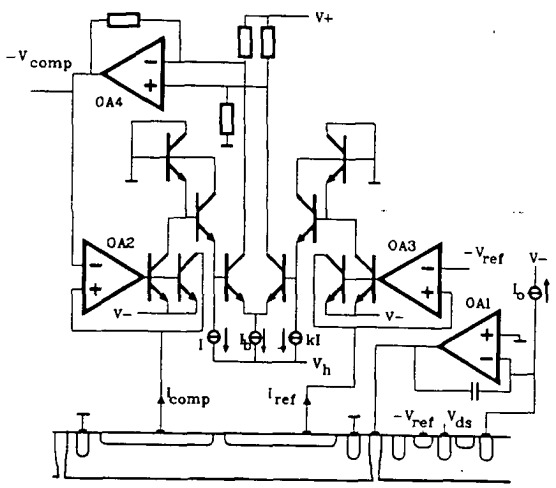


Figure 4.20 Schematic circuit diagram of the simplified compensating-type of colour indicator.

The operation of this colour indicator can conveniently be combined with dual-slope analog-to-digital converter as shown in figure 4.19. In the first part of the conversion cycle the switch that supplies the current $I_t = I_u + I_l$ to the integrator is grounded and the integrating capacitor is charged with I_u. After a fixed time interval, T_n, this switch is connected to the integrator input and the integrating capacitor is discharged with the photocurrent $I_u + I_l$ from the lower junction. After a complete measurement the dual-slope counter content will display the fraction $I_u / I_l + I_u$ and a colour-determined digital output is available.

Also the simplified version of the compensation-type of colour sensor has been elaborated to an integrated colour indicator. The simplified schematic is depicted in figure 4.20. Operational amplifier OA1 is for generating the substrate voltage. OA2 and OA3 are used for the read-out of I_{comp} and I_{ref} respectively. The inverting inputs are connected to the correct reverse biasing voltage, whereas the output drives a npn transistor in such a way that the corresponding photocurrent is compensated for and no current flows into the noninverting input. At equal transistors these photocurrents are also fed to the translinear circuit in which a differential voltage proportional to $k \times I_{ref} - I_{comp}$ is formed. The lefthand and righthand cores are stacked slightly different than in the ratio type of colour indicator and, when disregarding the base current, the product relation between the collector currents yields: $I_{comp} \times I \times I_{c1} = I_{c2} \times kI \times I_{ref}$. This results in $\Delta I_c / I_c = (I_{comp} - k I_{ref}) / (I_{comp} + k I_{ref})$. Under steady-state conditions and in case of sufficient loop gain, this relation results in such an output voltage, V_{comp}, that the input voltage of the differential amplifier, realised in OA 4, becomes zero; thus $I_{comp} = k.I_{ref}$.

As mentioned before, the operation can be improved by applying a positive voltage, sufficient to deplete the entire epilayer, to the epilayer underneath the the compensating diode. However, as indicated by the simulations in the previous chapter, the compensating type of colour indicator reveals no real advantage over the ratio type. As a disadvantage the increased complexity remains. Therefore, no extra efforts are made to realise the improved version of the compensating type of colour indicator.

The schematic diagram of the actual circuit is shown in figure 4.21. The operational amplifiers are again implemented as a standard input stage connected to a npn or pnp transistor to drive the output current. The integrated circuit is shown on the photograph in figure 4.22. The chip area is 3000 x 1500 μm. Due to the poor performance of the lateral pnp transisitors, measurements were made using the modules shown in figure 4.18 b. The additional problem in this design was the limited loop gain and the long integration time constant required for obtaining a sufficient phase

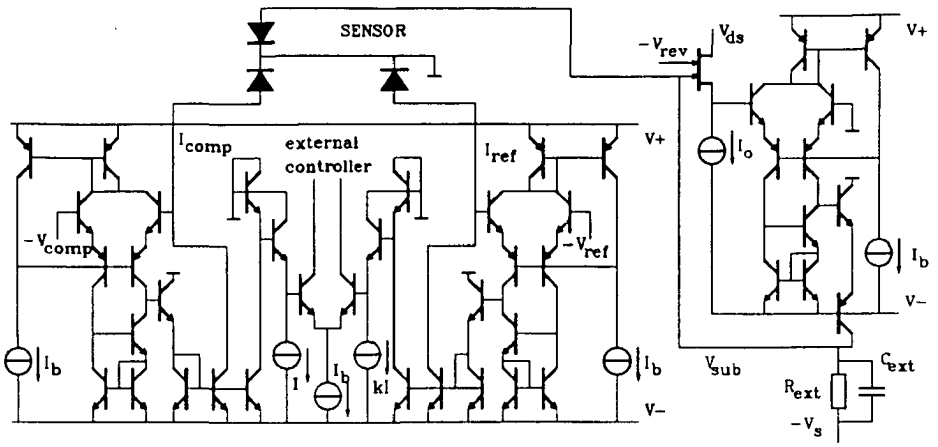


Figure 4.21 Actual circuit diagram of the compensating type of colour indicator realised in design IS 505.

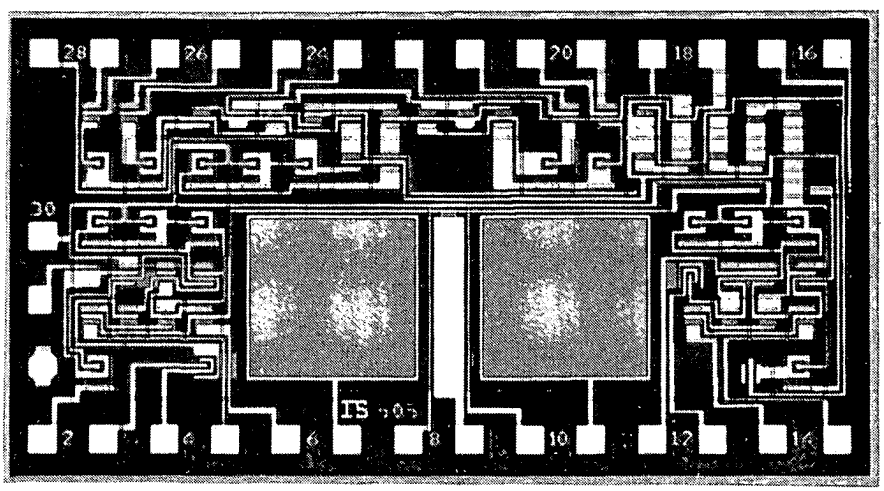


Figure 4.22 Microphotograph of the realised compensating type of colour indicator.

margin for maintaining stability in the overall feedback loop. The integrating action is obtained by placing capacitors in parallel to the resistors in OA4 and simultaneously the loop gain is controlled using I_b.

The theoretical aspects of the compensating type of colour indicator have already been discussed in the previous chapter along with the calculated responses for different values of the attenuation, k, of the photocurrent in the reference diode. The actual responses to a monochromator scan for three different values of the tuning current $k \times I$ and a reference voltage of 10 V are shown in figure 4.23 and the results are in reasonable accordance with the responses anticipated in figure 3.13. The minimum operating wavelength is at slightly larger wavelengths due to the somewhat lower epilayer doping and the moderate loop gain.

This compensation type of colour indicator lends itself quite well to combination with a special type of analog-to-digital converter. Replacing the operational amplifier OA4, which is basically the proportionally-integrating loop controller, for a comparator, a counter and a digital-to-analog converter, results in a servo analog-to-digital converter as shown in the simplified diagram in figure 4.24. The current $I_{comp} - k \times I_{ref}$ is fed to the comparator, which controls the counter in such a way that a larger photocurrent in the compensating diode causes the counter to count down. The counter content is available in the form of a digital output and also drives the DAC to provide the reverse voltage across the compensating diode. The epilayer voltage underneath the compensating diode can be generated using the operational amplifier OA5, which pulls a current from a pnp mirror until the channel resistance equals the value of an integrated resistor. This epilayer voltage, V_{ec}, has to be added to $-V_{comp}$ before using the latter for biasing the compensating diode in order to obtain a collecting layer in the shallow junction solely determined by the counter content.

4.6 Conclusions.

In this chapter measurements are described that clearly demonstrate the effectiveness of the basic colour-sensing principle. The curve of the photocurrent vs. the reverse voltage at different monochromatic wavelengths satisfies the expected course as is indeed confirmed by figure 4.9, and also the results of measurements on the devices with the different impurity profiles meet the expectations as is indicated in figure 4.10. The magnitude of the leakage and noise currents is low enough to ensure proper operation of the two colour indicator versions, however it impedes the competitive application of this principle in CCD colour imagers. The reduction of the chip area for one colour element by using one diode per

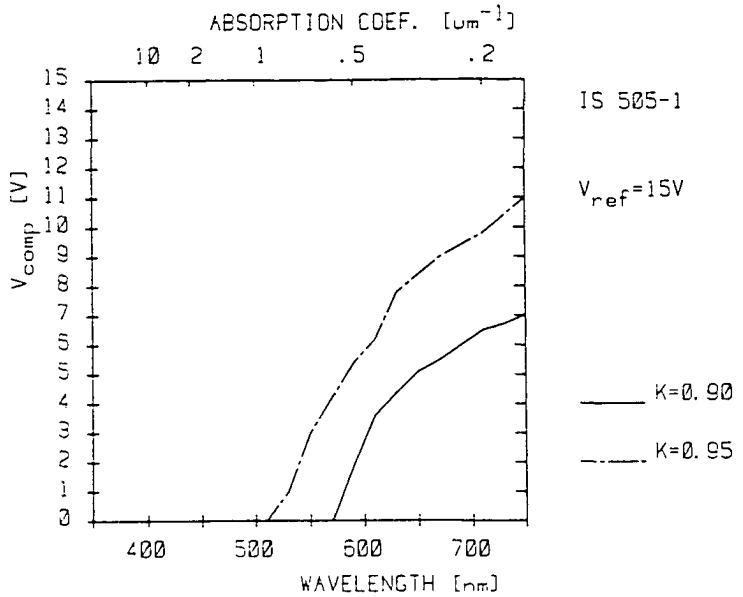


Figure 4.23 Spectral response of the simplified compensating type of colour indicator to a monochromator scan between 350 nm and 750 nm for two different values of the programmable attenuation.

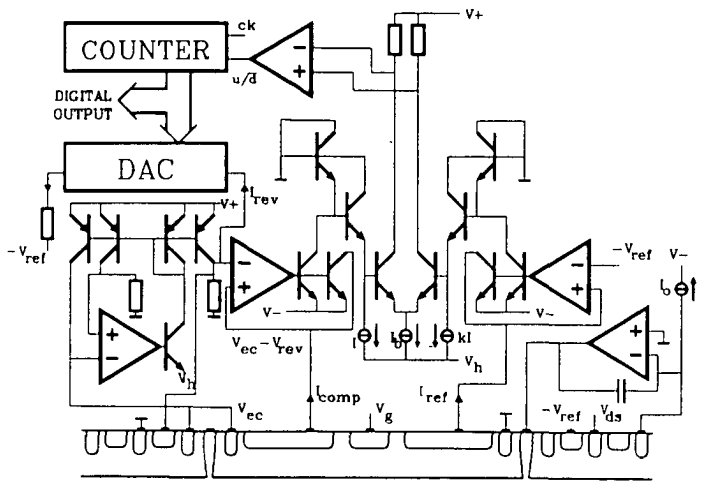


Figure 4.24 Simplified diagram indicating how the operation of the compensating type of colour indicator can be combined with a servo AD conversion.

colour element instead of three has to be sacrificed for obtaining the SNR customary in such imagers. Other effects that seriously hamper the use of this principle in two-dimensional colour imagers are the noise added by the differential operator that converts the response curves to the normalized eye-sensitivity curves, the epilayer inhomogeneity and the adaptive substrate potential. The nonuniformity of the epilayer doping concentration would cause a different collecting depth in different colour elements distributed over the large area of such an imager and thus to a different response of colour elements over the imager area.

When taking all merits and disadvantages of the different types of colour indicators into account, it becomes apparent that the compensating type of colour indicator does not offer sufficient extra flexibility over the ratio-type to justify the greater complexity. Both types of indicators reveal interesting properties for the realisation of smart sensors with a digital output.

4.7 References.

4.1 L.K.Nanver, High-performance bifet process for analog integrated circuits, Thesis TU Delft, 1987.

4.2 S.M.Sze, Physics of semiconductor devices, Wiley, 1981.

4.3 D.A.Antoniadis, S.E.Hansen and R.W.Dutton, Suprem 2; A program for IC process modeling and simulation, Stanford internal report 5019-2, June 1978.

dimensions in an unaltered process by reducing the number of photodetectors per colour element from three to two and maintaining an optimum image sharpness [3.10]. This imager consists of two types of dyed colour filters deposited on a silicon wafer, each covering two superposed photodiodes as shown in figure 3.8. The subtractive colour filter of the first type is yellow, so when considering the three primary colours, the blue part of the incident spectrum is blocked. Due to the properties of silicon the upper junction is sensitive to the shorter wavelength part of the remaining spectrum. As the blue has already been removed this upper junction is sensitive to the green part of the spectrum and the lower to the green and red part. The other filter type is cyan and removes the red part of the incident spectrum. Therefore, the upper junction covered by this filter gives a response for the blue and green part of the spectrum and the lower only for the green part. If the upper and lower junction are designed to possess equal sensitivities to the green, each spatial element provides the correct green information. The image sharpness is determined by the high-frequency luminance signal, which is determined for 59% by the green response due to the high green-sensitivity of the human eye. As every spatial element supplies the green information a maximum image sharpness is maintained while only two spatial positions are required for obtaining the tristimulus colour information for unambiguous colour reconstruction. By subtraction of the responses of the photodetectors covered by the yellow and cyan filter the red and respectively the blue signal can be derived. The read-out is performed using a lower and upper charge transfer channel.

A similar image sensor structure, also composed of an array of superposed diodes, however, without dyed colour filters, has been developed for simultaneously providing a visible and an infrared image of an observed object [3.11]. The array of shallow diodes generates the visible black-and-white image, whereas the lower array allows the detection of the near-infrared image.

Not only monocrystalline silicon has been found suitable for colour sensing, amorphous silicon also reveals promising properties for such an application. Especially the absence of minority carrier diffusion currents permits colour detection using a simple $p^+ n$ photodiode. In a practical device usually a PIN diode structure is used [3.12]. For the detection of all three primary colours a dual stacked PIN diode structure can be used as shown in figure 3.9a. A reverse voltage is applied across the upper p-type layer and the n-type back layer, so the two diodes are connected in series. In amorphous silicon the actual width of the bandgap, and thus the course of the absorption coefficient at different wavelengths, is not as well determined as in monocrystalline silicon and depends on deposition

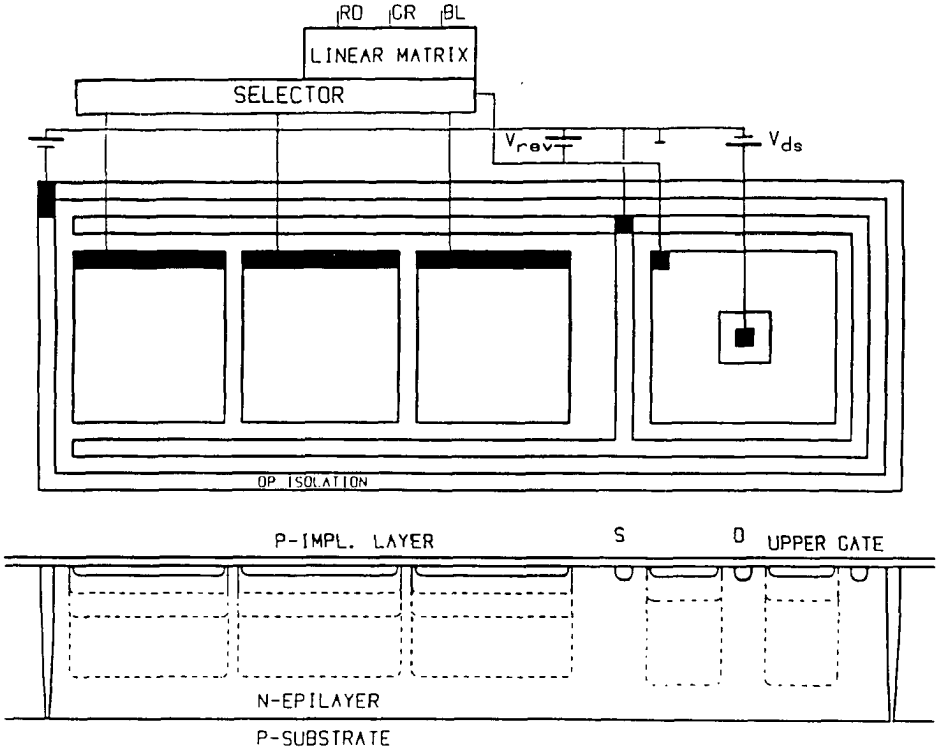
5 DISCUSSION AND PERSPECTIVES.

The aim of the research presented in this thesis is the investigation of the feasibility of using the wavelength-dependent absorption in silicon for colour sensing. Various devices have been discussed ranging from a simple colour indicator up to a colour element suitable for implementation in a two-dimensional colour imager. The basic operating mechanism involves the tunable allocation of an absorbing layer of silicon to a collecting junction using the reverse voltage applied across this junction. The allowed range of the critical operating parameter for proper operation of the sensor was derived using accurate simulations. The performance of devices manufactured according to the results of the simulations was in accordance with the expectations. Reproducible and clearly distinguishable responses were obtained.

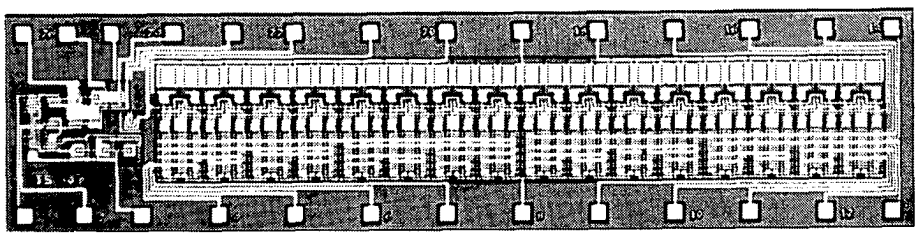
The basic operating mechanism also implies that charge carriers generated beyond the boundary of the upper space charge region are not collected and are thus wasted. This inefficient conversion of the available optical signal results in a relatively poor signal-to-noise ratio, which seriously impedes the application of the single diode silicon colour element in high-density CCD colour imagers. Another disadvantage that hampers the application of the silicon colour sensing technique in this application area is the required complete depletion of the epilayer underneath each colour element. This constraint forces the imager into a mode as shown in figure 5.1 a, where the reverse voltage of all colour elements is switched sequentially to different values and the substrate is controlled accordingly. In this time-multiplex operation each element determines the particular colour component simultaneously, which gives rise to a kind of 'colour flicker' in the image. Moreover, the complete depletion of the epilayer underneath each of the elements in the large-area sensor is not possible due to the nonuniformity of the epilayer thickness.

The colour indicators, however, consume much less chip area and these problems apply to a far lesser extent. The inefficient optical energy conversion and the subsequent poorer SNR at an equal detector area is a direct consequence of the operating mechanism and can, therefore, not be avoided. For maintaining a certain SNR the area has to be enlarged, compared to that of a diode in a conventional colour imager, and the question that remains to be answered is whether there is a net gain in chip area. A definite answer would require further research.

A 32-element colour line sensor has been realised, which indeed confirms the limitations mentioned. Photograph 5.1 b shows the actual design that consumes a chip area of 1000×6000 μm. The selection of a colour element



(a)



(b)

Figure 5.1 (a) diagram showing the implementation of the silicon colour sensor in a colour imaging sensor.
(b) Photograph of a 32-element colour line sensor based on the silicon colour sensing mechanism. Chip dimensions are 6000×1000 μm.

takes place using a multi-emitter selection transistor, which drives a JFET switch. The density is not impressive compared to conventional CCD imagers, however a number of improvements are possible to circumvent the second disadvantage making the application in moderate-density colour imagers possible. These improvements are mainly technological; for example, by locally replacing the epilayer for polysilicon layers, a well-defined depth of absorbing silicon can be realised and no depletion of the remaining polysilicon layer is required for the proper operation of the colour sensor because of the extremely short lifetime in this layer.

For the colour indicator the two constraints mentioned are of less importance. The results of the modelling are remarkably consistent with the measurements on the devices manufactured subsequently. The colour indicators are applicable in the application areas mentioned in the introduction of this thesis. The compensating type of colour indicator seems to reveal a smaller performance/complexity ratio compared to the ratio type of indicator. The extra flexibility offered in the compensating indicator, using the controllable attenuation of the photocurrent in the reference diode, does not justify the increased complexity of the circuitry needed for implementation of the feedback loop and the extra provisions required for adaptive depletion of the epilayer underneath the compensating diode. The ambiguity that can arise when observing multicolored objects can be reduced using the electrical control parameter to perform several measurements at different settings.

The required compatibility with a standard bipolar process does not lead to serious compromises that reduce the sensor performance. Some problems with respect to the lateral pnp transistor are discussed and effective measures are presented. Therefore, the realisation of colour sensors integrated in silicon is a promising concept. Both types of colour indicators also reveal interesting properties for combining the colour sensing operation with an analog-to-digital conversion. The exploration of the full potential of the colour sensing principle will, therefore, require a further Ph.D. research.

SUMMARY.

The aim of the research presented in this thesis is to investigate the possibility of applying the wavelength dependence of the absorption coefficient in silicon for radiation in the visible part of the spectrum for colour sensing. This effect is due to the indirect bandgap in silicon which results in a larger change in the momentum required for the generation of an electron-hole pair at an energy of incident photons closer to the indirect bandgap at 1.12 eV. The absorption probability is related to the required change in momentum that has to be supplied by the lattice, resulting in a small absorption coefficient for long-wavelength visible light and a short penetration depth for short-wavelength light. This effect is usually considered harmful in silicon optical devices as it reduces the performance of solar cells. However, electronic tuning of the spectral response of a silicon photodiode will make the realisation of integrated colour sensors possible without having to resort to the current customary technique of depositing polymer colour dyes on a photodiode.

There are a number of application areas where the silicon colour-sensing technique can be applied viz. in solid-state colour imagers and in silicon colour indicators. The latter category is particularly useful in applications where no image information need be obtained and no reproduction of the colour in colorimetric terms is required. This situation often occurs in robotics where, up to now, a colour imager is used in applications where its implementation is superfluous.

An additional advantage of the silicon colour-sensing principle is the compatibility with standard silicon processing, which makes the integration of the sensor with the read-out circuits possible. Such a 'smart sensor' will provide a direct microprocessor-compatible output signal. The advantages of such an approach along with several examples of application areas are discussed in the introduction. After the introduction, the basic operating principle is discussed followed by an outline of all the wavelength dependences encountered when light impinges on an air-siliconoxide-silicon system. An accurate modelling of the effects that affect the transmission of incident radiation and the collection of the generated charge carriers, as well as the wavelength dependences therein, enables the prediction of the effectiveness of the electical control parameter on the response.

The evaluation of the parasitic wavelength dependences has shown that colour detectors in silicon with a well-defined, distinct and reproducible response can be obtained by using an electrical control parameter in a single photodiode. The basic device structure is composed of a shallow p^+n junction in a n-type epilayer that is grown on a p-type substrate.

Applying a certain reverse voltage across the shallow junction for depleting a part of the epilayer down from the junction, and concurrently applying a reverse voltage across the epilayer-substrate junction for depleting the remaining epilayer, permits the selective detection of only the charge carriers generated in the upper space-charge layer. When increasing the reverse voltage across the shallow junction and adjusting the substrate voltage, an increased response is observed in the case of long-wavelength light, whereas the response remains constant in the case of incident short-wavelength light, as the incident radiation is entirely absorbed when very thin silicon layers are used. Therefore, this voltage can be used as an electrical control parameter.

Several implementations of the basic colour-sensing mechanism are evaluated. The applicability of this sensor in a high-density CCD colour imager is restricted by the signal-to-noise ratio. For small values of the reverse voltage across the junction especially, a large amount of the available optical signal is wasted. This low conversion efficiency is a direct result of the basic operating mechanism. The chip area gained by using only one diode per colour element has to be sacrificed for the SNR.

The values of the epilayer width and doping concentration directly result from the operating conditions. Operation in the entire visible part of the spectrum, using reverse voltages not exceeding 20 V, results in a 6 Ω cm 7 - 8 μm thick epilayer. The tuning range of the collecting layer extends from the width of the shallow junction plus the built-in depletion layer (1.5 μm in IS 505-1) up to the width of the epilayer minus the epilayer-substrate built-in depletion layer (6 μm). The small value of the epilayer doping concentration and the finite size of the shallow junction limit the tuning range at short wavelengths, as a penetration depth of 1.5 μm is associated with a wavelength of incident light of 550 nm. To improve the short-wavelength performance, a higher doped buffer layer just underneath the shallow junction can be used to reduce the built-in depletion layer, giving an upper limit in the tunability of the collecting layer at 0.7 μm (IS 505-2), which is associated with a penetration depth at 480 nm. The short-wavelength response is also affected by the surface charge, which can be controlled by implanting an extra oxide charge or by realising a polysilicon gate connected to a voltage between flatband and full depletion of the top layer. Measurements on this device provide a qualitative confirmation of the effectiveness of the control of the surface space-charge region due to high leakage currents.

The properties of the two different versions of the colour indicator are also discussed. The measured responses are consistent with the results of the simulations. The ratio-type of colour indicator seems to reveal the best

performance/complexity ratio. The flexibility makes the colour indicators suitable for implementation in the application areas outlined in the introduction. The operation of the colour indicator can conveniently be combined with an analog-to-digital conversion, making the realisation of an integrated smart colour sensor equipped with a digital output possible.

SAMENVATTING.

In dit proefschrift wordt een onderzoeken beschreven naar de mogelijkheid om de golflengteafhankelijkheid van de absorptiecoefficient in silicium in het zichtbare deel van het optische spectrum toe te passen voor de realisatie van geintegreerde kleurensensoren. Dit golflengteafhankelijke effect is het gevolg van de indirecte bandafstand in silicium, waardoor een grotere verandering in impulsmoment nodig is voor de generatie van een electron-gat paar in het geval van een inkomend foton met een energie dichter bij de indirecte bandafstand van 1.12 eV. De absorptiekans is gerelateerd aan de benodigde verandering in dit impulsmoment, dat vrijwel geheel door het kristalrooster geleverd moet worden, wat resulteert in een kleine waarde van de absorptiecoefficient voor invallend licht van lange golflengte en een korte indringdiepte voor kortgolvig licht. Dit effect wordt in het algemeen als schadelijk beschouwd in optische detectoren en zonnecellen. Toepassing van dit effect is echter ook mogelijk door, via een electronische beinvloeding van de responsie, geintegreerde kleurensensoren te realiseren zonder dat daartoe polymeer kleurenfilters behoeven te worden aangebracht.

Er zijn diverse toepassingen denkbaar waarin een dergelijke silicium kleurensensor kan worden gebruikt: n.l. in halfgeleider kleurenkameras en in eenvoudige kleurenindicatoren. De laatstgenoemde toepassing is met name interessant ter vervanging van de kleurenkamera in situaties waar een dergelijke beeldvormende sensor aanzienlijk meer informatie verstrekt dan voor de oplossing van het meetprobleem nodig is.

Een bijkomend voordeel van de silicium kleurensensor is de kompatibiliteit met een standaard fabricageproces van geintegreerde schakelingen in silicium, waardoor de integratie van sensor met uitleeselektronica mogelijk wordt. Een dergelijke 'slimme sensor' kan een uitgangssignaal leveren dat direct door een microprocessor systeem verwerkt kan worden. Het voordeel van een dergelijke aanpak en een aantal toepassingsvoorbeelden van de onderhavige kleurensensor worden in de inleiding besproken. Daarna wordt het onderliggende fysische principe uiteengezet gevolgd door een inventarisatie van de golflengteafhankelijkheden, welke in een lucht-oxide-silicium systeem optreden en in een modellering meegenomen moeten worden, om een realistische voorspelling te kunnen maken met betrekking tot de effectiviteit van de elektrische stuurparameter op de responsie.

Een evaluatie van deze effecten heeft uitgewezen dat kleurendetectoren in silicium met verschillende, eenduidige en reproduceerbare responsies

mogelijk zijn door gebruik te maken van eenzelfde diode aangestuurd door een elektrisch stuursignaal. De kleurensensor is in principe opgebouwd uit een dunne p-type laag in een n-type epitaxiale laag, welke op een p-type substraat is gegroeid. Door de bovenste junctie op een sperspanning aan te sluiten en tevens het substraat op een zodanige potentiaal te brengen dat de gehele epitaxiale laag gedepleteerd is, wordt het mogelijk om in de bovenste junctie selectief alleen de ladingsdragers te verzamelen welke in de bovenste depletielaag gegenereerd zijn. Indien de sperspanning over de bovenste junctie verhoogd wordt, waarbij tevens de substraatpotentiaal wordt aangepast, dan zal de gedetecteerde fotostroom toenemen in het geval van invallend licht met een lange golflengte en vrijwel constant blijven in het geval van kortgolvig licht, aangezien al het licht reeds is geabsorbeerd in een zeer dunne laag silicium. De sperspanning fungeert dus als sturende parameter.

Een aantal varianten op de elementaire kleurensensor zijn nader onderzocht. De toepassing van de sensor in een CCD kamera wordt belemmerd door de signaal-ruis verhouding. Met name bij kleine waarden van de sperspanning over de bovenste junctie zal een groot deel van het beschikbare optische signaal verloren gaan. Deze lage efficientie resulteert in een lage signaal-ruis verhouding, waardoor een deel van de winst in chip oppervlakte, dat behaald is doordat slechts een diode per kleurelement nodig is, weer teniet gedaan wordt.

De vereiste waarden van de epitaxiale dikte en dotering volgen uit de optische randvoorwaarden en de noodzaak om de maximale sperspanning voor depletering van de totale epitaxiale laag te begrenzen tot 20 V. Dit resulteert in een 6 Ωcm 7 - 8 μm epitaxiale laag. Het bereik waarbinnen de depleteringslaag kan worden gevarieerd is aan de onderzijde begrensd door de dikte van de bovenste dunne laag vermeerderd met de depleteringslaag bij kortgesloten junctie. Deze is in sensor IS 505-1 gelijk aan 1.5 μm en komt overeen met een indringdiepte bij een golflengte van 550 nm. Aan de bovenzijde wordt het bereik begrensd door de dikte van de epitaxiale laag verminderd met de dikte van de depleteringslaag aan substraatzijde bij kortgesloten epilaag-substraat junctie. Voor vergroting van het stuurbereik wordt in IS 505-2 een bufferlaag aangebracht direct onder de bovenste junctie, waardoor de ondergrens kan worden teruggebracht tot 0.7 μm, wat overeenkomt met een indringdiepte bij een golflengte van 480 nm. De responsie bij korte golflengten wordt ook beinvloed door de oppervlaktelading, welke bestuurd kan worden via een extra oxidelading of door sturing van een, extra aan te brengen, polysilicium gate boven de diode. Metingen aan deze sensoren geven slechts een kwalitatieve bevestiging van de werking van de stuurparameter in het kortgolvige deel van het spectrum ten gevolge van de grote lekstroom.

De eigenschappen van twee verschillende typen kleurenindicatoren zijn aangegeven. De metingen komen overeen met de simulaties. De kleurenindicator van het ratio-type heeft een betere prestatie/complexiteit verhouding. De flexibiliteit van de kleurenindicator maakt deze geschikt voor de toepassingen welke in de inleiding zijn genoemd. De werkwijze van de verschillende indicatoren is geschikt voor combinatie met een analoog naar digitaal omzetting, waardoor geintegreerde kleurensensoren met een digitaal uitgangssignaal ook tot de mogelijkheden behoren.

PUBLICATIONS AND PRESENTATIONS RELATED TO THIS THESIS;

P.P.L. Regtien and R.F. Wolffenbuttel, A novel solid-state colour sensor, 15th ESSDERC, Aachen, West-Germany, 9-12 September 1985, Proc. pp. 80-81.

P.P.L. Regtien and R.F. Wolffenbuttel, A novel solid-state colour sensor suitable for robotic applications, 5th International Conference on Robot Vision and Sensory Controls, Amsterdam, The Netherlands, 29-31 October 1985, Proc. pp. 259-266.

R.F. Wolffenbuttel, Non-conventional smart colour sensing in robotics, NATO Advanced Research Workshop on Sensors and Sensory Systems for Advanced Robots, Maratea, Italy, 28 April-3 May 1986, proc. pp. 571-580.

R.F. Wolffenbuttel and P.P.L. Regtien, A novel approach to solid-state colour sensing, Sensors and Actuators, Vol 9(1986) pp. 199-211.

R.F. Wolffenbuttel, Colour filtering in silicon, 16th ESSDERC, Cambridge, UK, 8-11 September 1986, Proc. pp. 101-102.

R.F. Wolffenbuttel, Smart colour sensor, 12th ESSCIRC, Delft, The Netherlands, 16-18 September 1985, Proc. pp. 190-192.

R.F. Wolffenbuttel, Method and circuit for determining the wavelength of light, US patent nr. 4,749,851, Filed 29 August 1986 Granted 7 June 1988.

R.F. Wolffenbuttel, Color filters integrated with the detector in silicon, IEEE EDL, Vol 8 No 1 (1987) pp. 13-15.

R.F. Wolffenbuttel, A simple integrated color indicator, IEEE J. SSC, Vol 22 No 3 (1987) pp. 350-356.

R.F. Wolffenbuttel, Photodiodes in silicon with an intrinsic colour filtering capability, Transducers 87, Tokyo, Japan, 2-5 June 1987, Proc. pp. 219-222.

R.F. Wolffenbuttel, Integrated all-silicon color filtering element with an enhanced wavelength tunability, IEEE EDL, Vol 9 No 7 (1988) pp. 337-339.

R.F. Wolffenbuttel, Operation of the silicon colour filtering element, Sensors and Actuators (Accepted for publication).

BIOGRAPHY.

Reinoud F. Wolffenbuttel was born in Abcoude, The Netherlands, in 1958. He received the M.Sc. degree in Electrical Engineering in 1984 from the Delft University of Technology, Delft, The Netherlands. From October 1984 to October 1988 he has been working towards his Ph.D. on integrated silicon colour sensors at the Laboratory for Electronic Instrumentation, Department of Electrical Engineering, Delft University of Technology. Since 1986 he is a staff member of the Laboratory for Electronic Instrumentation and is involved in digital instrumentation and smart sensors.

ACKNOWLEDGEMENTS.

The author is indebted to all persons, who have, directly or indirectly, contributed to the realisation of this thesis. Especially,

my parents, for their love and moral support.
P.P.L. Regtien, for his guidance
the members of the Delft IC Workshop, in particular E. Goudena, L. Nanver, J. Groeneweg and J. Teven, for processing the devices.
the members of the Laboratory for Electronic Instrumentation for maintaining a stimulating working atmosphere, in particular G. de Graaf for his assistance.
J.B. Zaat-Jones for debugging the English.