BASIC OPERATION OF THE CHARGE COUPLED DEVICE

D.J. Burt*

INTRODUCTION

The Charge Coupled Device (CCD) is a new type of silicon integrated circuit which is fabricated using conventional MOS processes. In operation, information is represented by a quantity of electric charge, as distinct from conventional circuits where current and voltage levels are generally used. The device is basically a shift register; signal charges are stored and transferred in clocked shift register fashion under an array of closely spaced control electrodes. However, a unique feature of the CCD is that it will transfer analogue signals, which can be introduced either electrically or optically. Thus, by providing suitable input/output circuitry and clocking waveforms, the CCD can be employed to advantage in the following three main application areas:

i) As an analogue shift register or delay line
ii) As a serial memory for binary data storage
iii) As a solid state imaging device.

The following sections describe the basic operation of the charge coupled device.

BASIC OPERATION

The CCD operates by a mechanism of charge storage and transfer under an array of MOS *control electrodes* or gates. Information in the form of electric charge is transferred along the silicon surface in clocked shift register fashion by sequential manipulation of the voltages on the control electrodes that constrain this charge. There are basically two types of CCD, *surface channel* and *buried channel*. In the surface channel devices the charge is stored and transferred at the silicon surface whereas in buried channel devices the doping of the silicon substrate is modified such that the storage and transfer of charge takes place in the bulk silicon just beneath the silicon surface. There are also several different types of electrode structure and clocking techniques commonly used to realise a practical charge coupled device. Initially, however, the basic operation of the CCD is described for surface channel operation with *three phase* electrodes. Buried channel operation and alternative electrode structures and clocking techniques are described later.

1 CHARGE STORAGE

The structure of a basic MOS CCD element (surface channel) is shown in Figure 1(a). Figure 1(b) illustrates the manner in which the element stores charge. The silicon substrate is shown in the figures as being p-type, but obviously n-type devices are also possible.

Application of a positive voltage $+V_G$ to the electrode has the effect of repelling the positively charged majority carriers in the silicon (i.e.

* GEC Hirst Research Centre, Wembley, England.

holes in our case) away from the vicinity of the electrode. This region beneath the electrode that becomes depleted of holes is called a *depletion region*. For a given electrode structure and substrate doping concentration, the extent of the depletion region into the silicon is a function of the applied electrode voltage. Information is stored in the depletion region in the form of minority carriers (i.e. electrons in our case). Later sections describe how this charge may be introduced to the structure by either electrical or optical means.

The stored electrons are localised at the Si/SiO_2 interface because they are attracted to the positive charge on the control electrode. The magnitude of the charge which may be stored under a given electrode is variable up to a maximum value that is dependent upon the electrode size and bias voltage. As the amount of charge stored is increased, the extent of the depletion region decreases in order to preserve overall charge neutrality in the system.

The CCD is inherently a *dynamic memory* since the stored information disappears with increasing time. The mechanism for this is the thermal generation of electron-hole pairs which takes place in any semiconductor (commonly called *dark current*). This causes the depletion regions to be slowly filled with minority carriers which gradually mask the stored information.

1.1 POTENTIAL WELL MODEL

An alternative model, which is sometimes more useful in describing the operation of a CCD, is to consider that the electrons are filling a *potential well* formed by the potential minimum in the silicon which constrains the electrons to remain under the electrode. This is evident from the energy-band diagrams shown in Figure 2 which represent the two conditions shown in Figures 1(a) and (b). The potential minimum at the silicon interface is generally referred to as the *surface potential*, ϕ_S. The surface potential for an empty well, ϕ_{SO} (i.e. the *depth* of the potential well) is easily derived from well known MOS equations:-

$$V_G - V_{FB} = \phi_{SO} + B\,\phi_{SO} \qquad (1)$$

where $B = \sqrt{2\varepsilon\varepsilon_o qN}/C_{ox}$ (2)

and V_{FB} = flat-band voltage

$\varepsilon\varepsilon_o$ = silicon dielectric constant

q = electronic charge

N = substrate doping concentration

C_{ox} = oxide capacitance per unit area

When carriers are present in the well, the surface potential decreases, as shown in Figure 3, and a larger voltage appears across the insulator.

2 BASIC STRUCTURE OF A CCD

A cross section of a typical three phase CCD is shown schematically in Figure 4. The basic device consists of a linear array of closely spaced control electrodes on a continuous silicon dioxide dielectric layer which covers the single-crystal silicon substrate material. Charge storage and transfer takes place in the *channel region* of the device, which is generally bounded by high concentration *channel-stop diffusions*. The charge constrained in the depletion region beneath a given electrode is called a *charge packet*. For analogue and memory devices, charge packets are introduced by applying suitable voltages to a p-n junction at the input of the CCD. For optical imaging applications they are formed as a result of electron-hole pair generation caused by light energy incident on the silicon substrate.

3 CHARGE TRANSFER

Once a charge packet has been introduced to the CCD it may be moved substantially intact through the structure in the manner illustrated by Figure 5. The figure also illustrates the required *driving waveforms*, or *clocking pulses* as they are sometimes called.

At time t_1, Figure 5(a), a charge packet is held under electrode ϕ_2, the voltage on which is $+V_{CC}$. Electrodes ϕ_1 and ϕ_3 are held at a resting potential, $+V_{SS}$.

At a later time t_2, Figure 5(b), electrode ϕ_3 is pulsed to $+V_{CC}$. This produces a depletion region under ϕ_3 which, because the electrodes are closely spaced, couples with that under ϕ_2, with the result that charge begins to move from under ϕ_2 to under ϕ_3. The voltage on ϕ_2 is then reduced to $+V_{SS}$ with a *slowly falling edge*. The voltage on ϕ_2 is not reduced to $+V_{SS}$ instantly because the charge carriers require a finite time to diffuse across the width of the electrode. Figure 5(c) shows the charge transfer complete at time t_3, with the charge now stored in the depletion region under ϕ_3.

Note that ϕ_1 has to be kept at a low potential throughout to prevent backflow of charge. Thus three electrodes are required to store and transfer one charge packet and, in consequence, are usually referred to as one *...ment* of the CCD. Thus the electrodes of a CCD array are arranged in ...nd are connected sequentially to the drive lines carrying the ...ing waveforms. Hence, continuation of the clocking se... ...re 5(d), would result in the movement of the ...o under the next ϕ_1, then to ϕ_2 and so on to ...te particularly that there can be a charge ...ctrode in an array and that application of ...harges to be moved simultaneously. A zero ...ble.

...the CCD it is generally necessary to maintain ...pletion at all times; this is the reason for ...e $+V_{SS}$. If V_{SS} is applied as a bias to the sub... to all electrodes, the generation of the clock ...ified.

4 OUTPUT FROM THE CCD

...output circuit for a CCD is shown schematically in ...s of a reverse biased p-n junction positioned in such ...tion region couples with the depletion region under ...des. An additional electrode at fixed bias is generally ...the last transfer electrode and the diode to minimise capacitive voltage pick-up from the clock pulses on the last transfer electrode. Thus, as ϕ_3 goes to V_{SS}, any charge present in the last potential well will be collected by the output diode and will appear as a small current pulse in the output circuit. A voltage output is simply obtained by using a load resistor, R_L. More sophisticated charge detecting circuitry is also possible, but in general the principle of detecting charge in a p-n diode remains universal.

5 INPUT TO THE CCD

Methods for injecting charge into a CCD array depend largely on application.

For electrical inputs, a p-n diode and control electrode or gate similar to those used at the output are generally used, Figure 7. The input diode acts as an infinite source of minority carrier. Charge injection from the input diode is similar to the supply of channel current from the source of an MOS transistor, with the first potential well acting as a virtual drain. The actual amount of charge injected may be controlled by voltage signals applied to either the input diode or the gate. Various techniques are available for analogue and digital signals.

For imaging, an optical system focusses the light onto the CCD. The incident light quanta enter the substrate and impart their energy to the silicon causing the generation of free carriers. Depending on device structure, the quanta may enter between the control electrodes, through transparent electrodes or, in specially thinned devices, through the back face of the substrate. The generated carriers collect as a charge pattern under the array of electrodes. This pattern is an analogue replica of the variation of light intensity across the original image. The charge pattern may be extracted from the CCD using the previously described sequential clocking technique; it appears as a train of pulses whose amplitudes vary with the grey scale of the image. The array is thus self scanning.

6 SIGNAL SIZE

The maximum amount of signal charge, Q, which can be stored under any one control electrode and transferred within the CCD is given by:

$$Q = kCV \tag{3}$$

where C is the total oxide capacitance of the storage electrode and V is the voltage swing on the control electrode. The charge under the electrode is actually both minority carrier charge (i.e. signal) and depletion layer charge. At full well a constant fraction k of this is minority carrier charge, where $k \sim 0.5$ for typical surface channel structures.

As to orders of magnitude; an element size of about 30 μm square is commonly used, together with 10 V drive pulses. Hence, with a 0.1 μm oxide thickness, C is about 0.1 pF and the maximum amount of signal charge transferred is of the order of 0.5 pC. For a given clock rate, the output signal may be calculated as a mean current level from the basic equation $i = dQ/dt$. With the above example, at 1 MHz clock rate the maximum mean output signal would be around 0.5 μA.

7 POWER DISSIPATION

Power is dissipated during charge transfer because the carriers fall through a potential drop per transfer whose magnitude is approximately equal to the voltage amplitude of the clock pulses. This power is given by

$$P = mf_o VQ \tag{4}$$

for m-phases. Again, with the above example, at 1 MHz with a three phase clock the power consumed is about 15 μW, which compared to other semiconductor devices is a relatively small value. Power is of course also dissipated in the generation of drive pulses, the extent of which is determined by the circuit techniques used.

8 TRANSFER INEFFICIENCY

In general there is no loss of charge during the transfer process. However, a small fraction of a given charge packet is "left-behind" at each transfer and emerges at the CCD output at a time later than the bulk of the packet which has been transferred correctly. *Transfer efficiency*, η, is defined as the fraction of a charge packet transferred correctly per transfer and is typically in the range 99.9% - 99.99% for three phase devices. It is often more convenient, however, to use *transfer inefficiency*, ε, where ε = 1 - η. Note that the effect of transfer inefficiency through an array of n electrodes is cumulative, i.e.

$$Q_n \simeq Q_o \eta^n \simeq Q_o (1 - n\varepsilon) \tag{5}$$

where Q_n is the charge remaining under the nth electrode after n transfers and Q_o is the initial charge. This equation is a good approximation only as η can vary somewhat with signal size.

There are two major causes of transfer inefficiency:

i) Incomplete charge packet transfer due to the fact that the carriers require a finite time to diffuse from under one electrode to the next. As the spacing between electrodes is reduced the effect becomes smaller but even so this mechanism imposes an upper limit on the speed at which the device can be operated. Note that the lower speed limit is set by the thermal carrier generation effect as described previously.

ii) The effect of carrier *trapping states* which are largely localised at the Si/SiO_2 interface. As a charge packet arrives at a storage electrode, a small portion of it is instantly "trapped" by the interface states under that electrode. When the packet is moved on, the states will empty (with various values of time constant) into trailing packets, thus giving rise to small residual charges trailing behind the main packet. The effect of interface states is naturally more pronounced in surface channel devices.

9 "FAT ZERO"

Of the factors which contribute to transfer inefficiency in surface channel devices, the effect of interface states is most significant at all but the highest operating frequencies. The effect of these states may be reduced by passing a constant background charge or *fat zero* through the array. This ensures that, to a first order, the interface states remain permanently filled and in consequence react with the desired signal to a minimal extent. The fat zero charges are in practice about 10 - 25% of the full well capacity. Some interface states are in general still active, however, due to edge effects in the channel. The main disadvantage of fat zero is the reduction in dynamic range.

10 THE BURIED CHANNEL CONCEPT

In the surface channel CCD, the potential minimum for minority carriers is at the Si/SiO_2 interface. If however an appropriate doping distribution of polarity opposite to the substrate is introduced over a small region adjacent to the surface, the potential minimum is moved away from the surface to a location within the impurity layer, as shown in Figure 8. For a p-type substrate, an n-type layer would be used. This is the basis of the buried channel CCD, where charge transfer takes place in the bulk silicon.

The buried channel CCD structure prevents contact of the signal carriers with the interface during normal operation and essentially eliminates surface state trapping. A transfer inefficiency of 10^{-4} - 10^{-5} is readily achieved. The speed of charge transfer is also enhanced because the carriers are kept further from the electrodes and are thus subject to more fringing field effects.

The surface layer inherent with the buried channel structure can be produced using either epitaxy or ion-implantation. Thicknesses of the order of 1 μm are typically used. The drive waveforms and input/output structures of a buried channel device can be similar to those of surface channel dedevices, although voltage levels may differ somewhat. The electrode structure can also be similar provided that the electrodes are very closely spaced.

Detailed theoretical analysis of the reshaping of the charge packet during transfer has shown an action similar to that of the oesophagus in swallowing. For this reason the buried channel CCD is sometimes known as the *peristaltic* CCD.

11 OTHER CCD STRUCTURES

CCD electrode structures other than three-phase are also possible. A two-level metallisation process may be used to fabricate either two-phase or four-phase devices as shown in Figure 9. In the two-phase device the direction of charge transfer is governed by asymmetry, the smaller depletion layer depth under the thicker oxide always causes charge to spill in one direction. By splitting each electrode into separate thick- and thin-oxide electrodes, four-phase clocking becomes possible where the direction of charge transfer may be in either direction depending on the pulse sequence. Asymmetry to obtain transfer directionality can also be achieved by increasing the substrate doping concentration under part of the control electrode using ion-implantation. These are called *implanted barrier* CCDs.

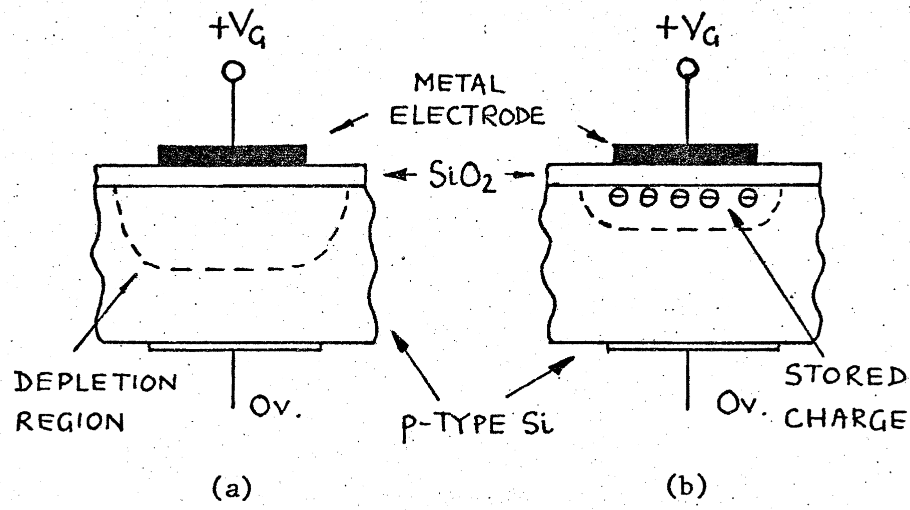


Figure 1 CHARGE STORAGE

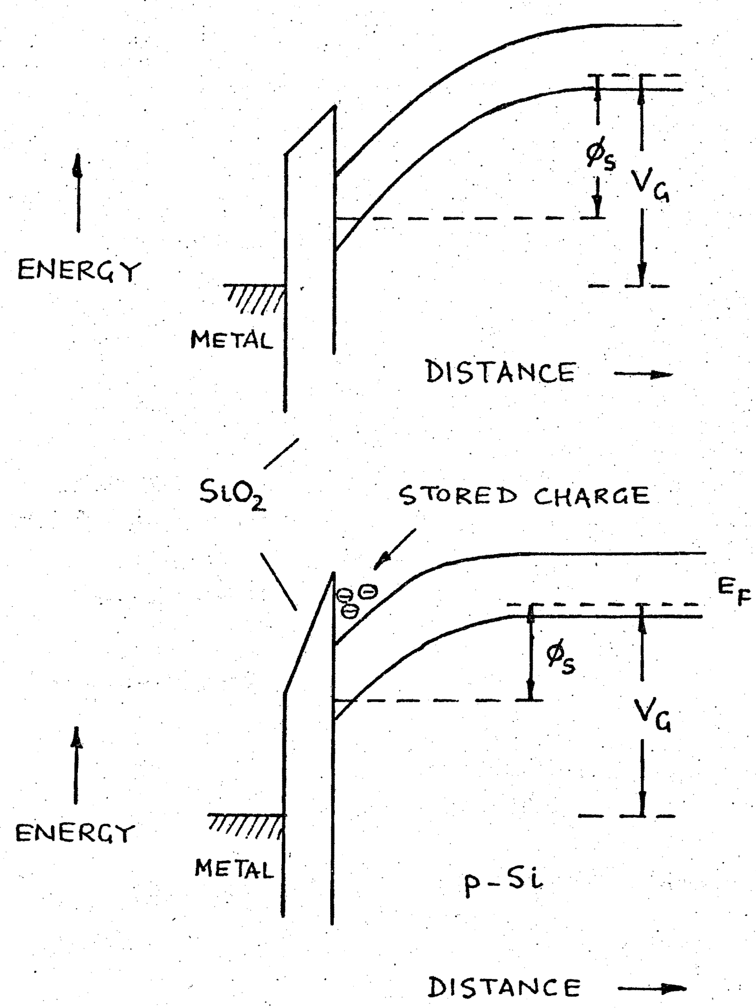


Figure 2 ENERGY-BAND DIAGRAMS

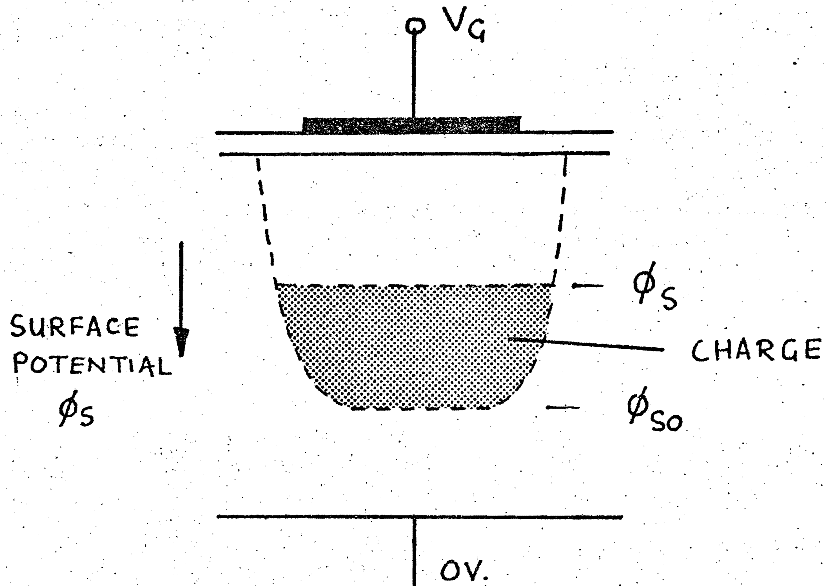


Figure 3 POTENTIAL WELL MODEL

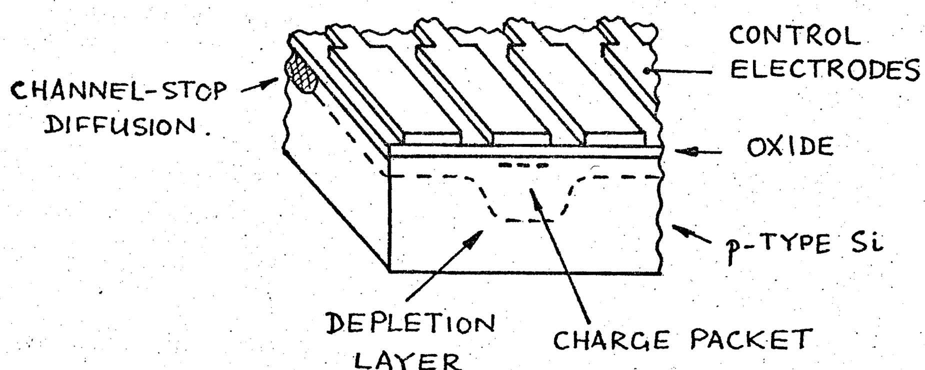
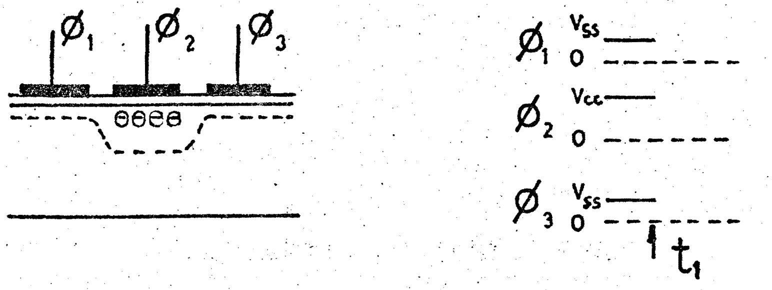
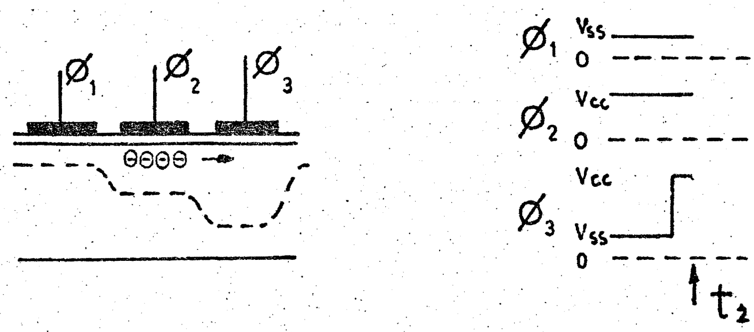


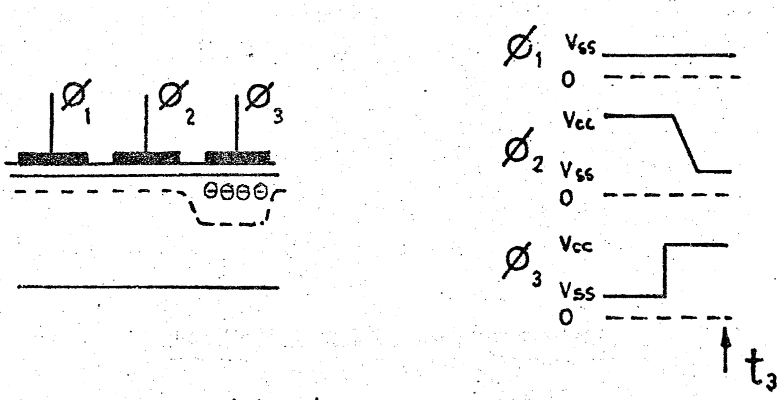
Figure 4 SECTION OF A THREE-PHASE CCD



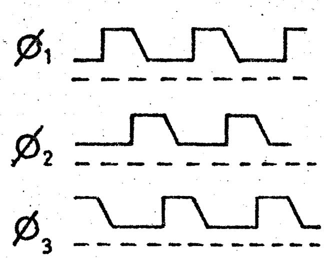
(a) Charge stored under ϕ_2



(b) ϕ_3 comes on



(c) ϕ_2 goes off



(d) Complete clocking waveforms

Figure 5 CHARGE TRANSFER

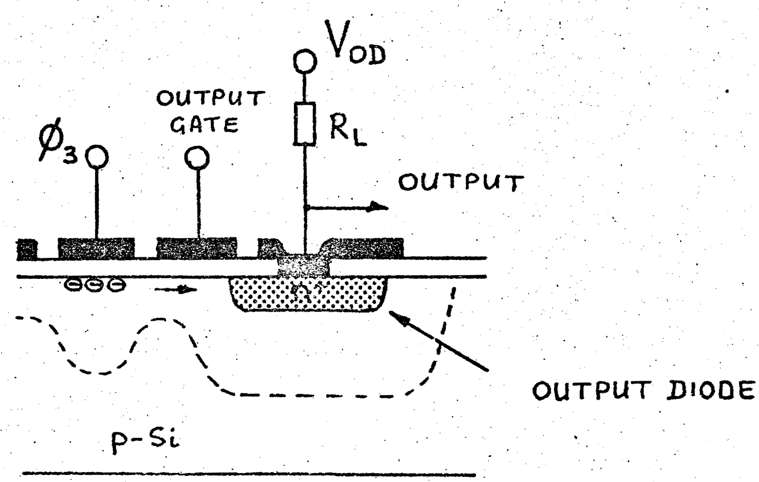


Figure 6 OUTPUT FROM A CCD

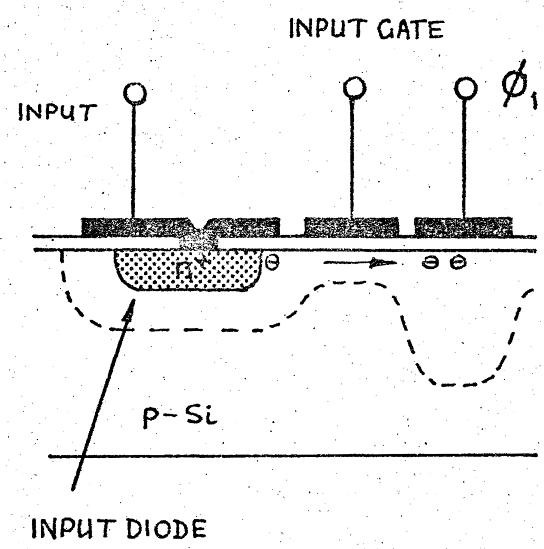
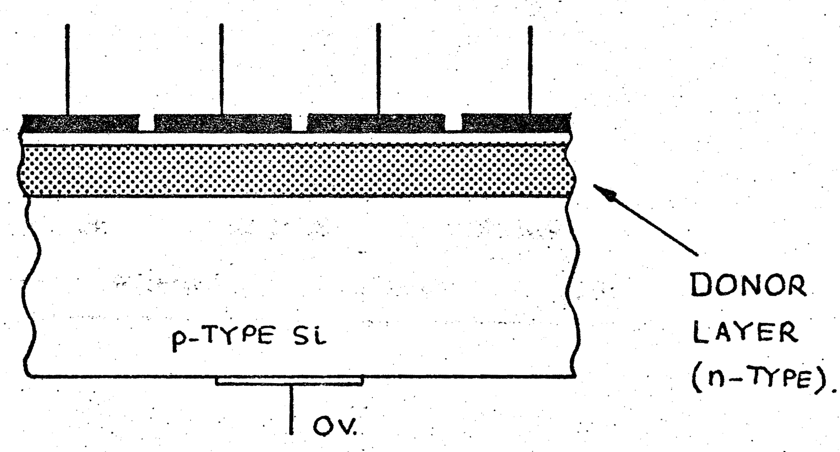
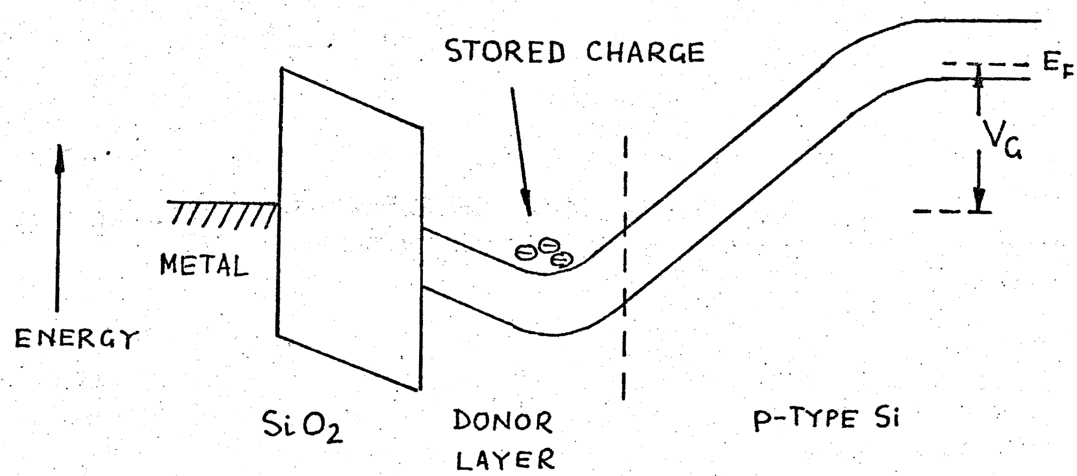


Figure 7 ELECTRICAL INPUT

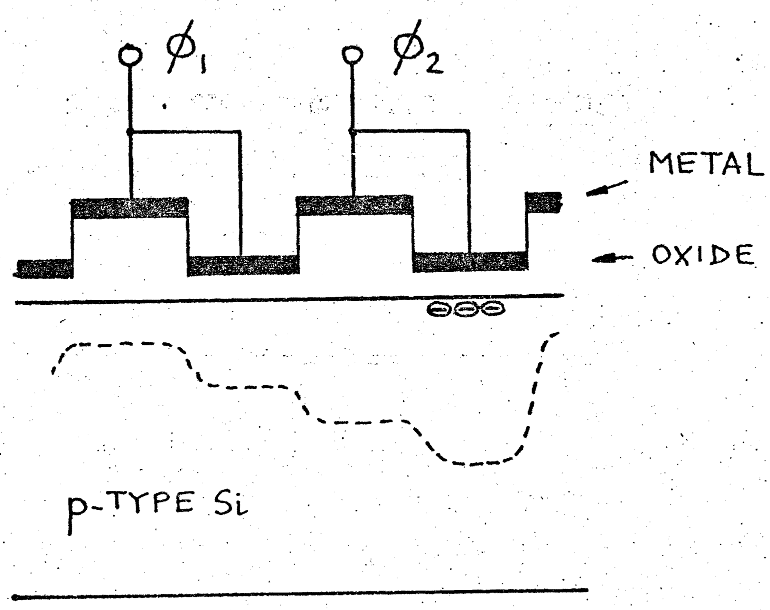


(a) structure



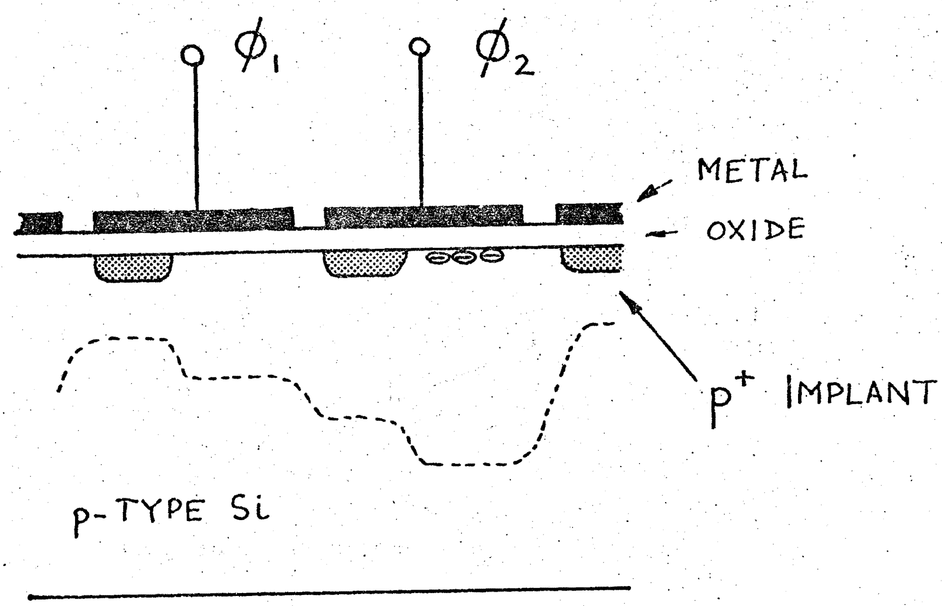
(b) energy-band diagram

Figure 8 BURIED CHANNEL CCD



(a) two-level metallisation

$V_G(\phi_1) > V_G(\phi_2)$



(b) implanted barrier

Figure 9 TWO-PHASE CCD STRUCTURES