Founded 1925
Incorporated
by Royal Charter 1961
To promote the advancement of radio, electronics and kindred subjects by the exchange of information in these branches of engineering

Volume 50, Number 5 May 1980

The Radio and Electronic Engineer

The Journal of the Institution of Electronic and Radio Engineers

Charge-Coupled Devices

THE FUTURE OF C.C.D.s

THIS year the charge-coupled device—or c.c.d. as it is generally known—celebrates its tenth anniversary. In so far as several c.c.d. structures have been available commercially for several years, the device has already passed from infancy to at least adolescence, and maturity, in the form of complex structures, is not far off. When viewed against the time scale for the evolution from laboratory to market place of the bipolar and m.o.s. transistors, this rate of development is truly remarkable. One has to recognize, however, that whereas in the case of the transistors mentioned, a new *technology* was needed, charge-coupled structures have been realized using an available (m.o.s.) technology into which enormous developmental effort had already been put. It is this, coupled with the inherent basic simplicity of the charge-coupling concept, that has enabled ideas to be translated so rapidly into structures of quite staggering complexity.

The c.c.d. was conceived as a semiconductor alternative to the magnetic-bubble memory. Ironically , however, its major impact has been, and is likely to remain, in two quite different areas. Bearing in mind that the device is, at its simplest, an analogue shift register, and that silicon is almost ideal as a photosensitive material in the visible and near infra-red, one has only to expose the device to light to form an exceptionally elegant image sensor. This fact was quickly appreciated and explains why over the first four years the major development effort on c.c.d.s was for imaging applications—first on one-dimensional, but soon after on two-dimensional, structures. The ultimate goal has naturally been the replacement of the conventional television camera tube; many problems remain but, as indicated in the paper by Burt, c.c.d. workers have achieved outstanding results and the goal is well within sight. More recently the attractiveness of c.c.d.s for infra-red image sensors has been increasingly appreciated; progress in this important area is reviewed in the paper by Lamb and Foss.

In the past five years increasing interest has been shown in capitalizing on the c.c.d.'s analogue capabilities for a quite different application—signal processing. The storage of analogue information has been something of a cinderella of electronics, generally requiring large and cumbersome realizations. The notable exception has been surface acoustic wave devices, but these have been limited to very short ($\lesssim 10\,\mu s$) delays, making them most useful for u.h.f. and microwave frequencies. The c.c.d., however, is capable of storing analogue data for at least milliseconds, with seconds—and even minutes—in prospect in cooled, high-quality silicon. This fact, coupled with circuit developments making possible the detection of very small ($\ll 1\,pC$) charges, has opened up a quite new field whereby complex signal processing operations such as correlation, convolution, signal averaging and adaptive filtering, can be carried out with elegant simplicity using charge-coupled structures. Some examples of recent applications are given in the papers by White *et al.*, Mavor *et al.*, and by Traynar and Beynon. Chamberlain's paper contains some interesting examples of how c.c.d. structures are beginning to combine image sensing and signal processing functions.

Interest in c.c.d.s for semiconductor memories has not been totally overlooked; indeed a few years ago it did seem as if c.c.d.s would have a significant edge over dynamic m.o.s. circuits and considerable efforts were being expended by several large companies to develop

large (up to 64 and 128 Kbit) c.c.d. memories. More recently the outlook for c.c.d.s for large memories has become less favourable, and interest in this application has waned considerably. One digital area which is, however, attracting interest is that of logic: it is possible to carry out various operations such as AND, OR, etc., very simply by means of c.c.d. structures employing some of the sensitive charge-detection circuits referred to earlier. The paper by Montgomery and Gamble discusses some recent work in this area.

What will the future hold for c.c.d.s? Although in some areas of application it has been said that 'c.c.d.s offer all the advantages of an untried device', there can be little doubt that c.c.d.s are likely to make a long-term contribution to image sensing. The basic simplicity of the optical sensing and electrical read-out processes makes the charge-coupled structure an almost-ideal transducer in the visible region. The precise way in which electrical output can be related to geometric location is something which is likely to be increasingly used, as also is the combination of the basic detector structure with on-chip signal processing for such things as automatic exposure control. Compared to the visible region, infra-red sensitive structures employing c.c.d.s are in their infancy. Nevertheless the increasing importance of infra-red imaging coupled with the serious disadvantages (complex read-out, limited resolution, etc.) of existing approaches, suggests that c.c.d.s are likely to play an important part in this field, if only for their elegant multiplexing in hybrid structures.

Although in the area of analogue signal processing c.c.d.s are also likely to play an important part, the precise nature of their role here is less clear. Many of the companies currently working on c.c.d.s for analogue signal processing are in fact equipment manufacturers who have in mind the inclusion of c.c.d.s in their products. Such chips as they may produce are likely to be 'specials' designed for their specific applications and are unlikely to become commercially available. There are, of course, some signal processing functions (e.g. delay lines, chirp filters) which are sufficiently widely needed to justify the production of commercial chips; the success of c.c.d.s for such applications will ultimately depend on whether their electrical performance (in terms of dynamic range, linearity, etc.) and their drive requirements are acceptable to would-be users. The c.c.d. programmable filter (and especially the adaptive filter) is of course a particularly flexible tool which could well find a substantial market (subject again to acceptable performance). An important attribute of the c.c.d. is that it employs the same (m.o.s.) technology as that for digital circuits. Already there are experimental circuits in which the c.c.d. represents only part of a larger signal-processing chip containing much digital as well as analogue circuitry; it is likely that this trend will continue, with distinction between the digital and analogue parts of the overall circuit becoming increasingly blurred.

It is in the purely digital area that the future of c.c.d.s is least clear, an ironic situation when one recalls the original motivation behind the c.c.d. concept. As pointed out previously, interest in c.c.d.s for large-memories has waned, although it has not vanished completely. Their application in logic circuits still awaits serious development.

But after ten years, one thing is certain: in one form or another c.c.d.s are here to stay.

J. D. E. BEYNON

GUEST EDITOR

Professor John Beynon (Fellow 1977) graduated in physics in 1960 from the University of Wales, and after obtaining an M.Sc. in electronics at Southampton University, worked as a Scientific Officer in the plasma physics group at the Radio and Space Research Laboratory (now the Appleton Laboratory), Slough. He joined the staff of the Electronics Department at Southampton University in 1964 where he continued research first in plasma physics and subsequently in microelectronics. He served a three-year term as Assistant Dean of the Faculty of Engineering and Applied Science at Southampton and in 1974 was appointed to a Readership in Electronics. During 1974 and 1975, Dr Beynon was on sabbatical leave in the Department of Electronics, Carleton University, Ottawa, and he has visited Cairo University and the Indian Institute of Technology, Delhi, to assist in setting up microelectronics laboratories. In 1977 Professor Beynon was appointed to the Chair of Electronics at the University of Wales Institute of Science and Technology and in July 1979 he moved to the University of Surrey as Head of the Department of Electronic and Electrical Engineering.

The Finniston Inquiry—The IERE's Reply

Following the publication on 9th January 1980 of the Report from Sir Monty Finniston's Committee of Inquiry into the Engineering Profession,* the President was invited by the Secretary of State for the Department of Industry to put forward the Institution's views on the Report and the recommendations made by the Committee.

Comment was particularly invited on a number of key issues and whether the Institution agreed with:

the Committee's analysis of the problems to be remedied;
their objectives and directions for changes;
their particular recommendations (in general terms) for the changes needed;

and on issues where the Institution was not in agreement, its own recommendations.

The invitations and the questions posed were considered by the Council when it discussed the Report at its meeting on 28th February and the Secretary was instructed to reply in the following terms.

1 The Institution of Electronic and Radio Engineers welcomed the report of the Finniston Committee of Inquiry into the Engineering Profession, the more so because many of its findings and recommendations were fundamentally in agreement with the evidence initially submitted by this Institution in March 1978.†

2 Noting from the preamble to the Summary Report and the Preface to the full Report, however, that the Finniston Committee chose to interpret its remit concerning British manufacturing industry widely in the broad context of the nation's economic needs, the IERE was disappointed to find that the Committee had little to say about the root cause of the inadequate performance of the nation's manufacturing industry, namely the general lack of enthusiasm for work at non-professional levels and the consequent low standard of industrial relations within many areas of British industry. This omission is all the more regrettable because it could give the impression that implementation of the Finniston Committee's recommendations concerning the education, training, employment and general organization of professional engineers would alone transform the performance of British manufacturing industry which, in our judgement, it most certainly would not.

3 The IERE also regrets, in this same context, that the Report fails to give adequate credit to those elements of the engineering profession where credit is manifestly due, notably engineers concerned with design and manufacture in the electronic equipment and system engineering business which consistently achieves high export success and tends to indicate thereby that the required quality of professional engineering expertise. drive and business acumen is not lacking in those areas. Nor does the Report adequately indicate the very high regard in which the British electronic and radio engineer is held overseas: a point which it is understood was repeatedly made to the Finniston team during its overseas visits, and particularly in the USA.

4 Notwithstanding these important factors, however, the IERE welcomes the main conclusions of the Finniston Report and particularly endorses its recommendations concerning the need to improve the balance of theory and practice in the education and training pattern for the engineer of the future at both main professional levels, to give statutory authority to the body responsible for the registration of British engineers, to carefully limit the placing of restrictions on the freedom of qualified engineers to practice their profession, and to avoid extending the requirement for licensing of engineers beyond those areas of work affected by public health and safety considerations: all of which points were featured in the IERE evidence to the Committee.

5 Concerning the proposed British Engineering Authority, to which the Committee quite rightly attaches the utmost importance, the IERE particularly welcomes the proposal that 'the Authority will work to draw together groups with common interests who now tend to act in relative isolation from each other because there is no active mechanism for linking them. These include the working engineer, employers, engineering teachers, public agencies and the Government itself'. The IERE was concerned to see, however, that it is proposed that the profession as such will not be represented among the 15/20 members of the Authority to be appointed by the Secretary of State for Industry. Similar concern is felt about the proposal that 'The Authority will maintain an expert staff to implement its policies'. These last two proposals, it would seem to the IERE, could deny the Authority direct access to the Institutions which are the focal points of the best expertise available in each of the engineering disciplines at both Board and working levels, whilst at the same time creating an unnecessarily expensive new area of bureaucracy for the registration of engineers in place of the present self-financed specialist professional resources available in and through the engineering Institutions. The IERE hopes, therefore, that this area of the Report will be given further careful study in the public interest lest concern about the likely cost of the British Engineering Authority be allowed to overshadow its true importance and perhaps even deny it the wholehearted Government approval it will need if it is to succeed.

6 As to its own role, as seen by the Report, the IERE was particularly pleased to note the recommendation that the learned society task of the Institutions should be maintained and strengthened wherever possible, since the primary role of the IERE under the terms of its Royal Charter is 'to advance

* 'Engineering Our Future', Cmnd 7794. (HMSO, 1980, price £5.00.) Summary Report (HMSO, 1980, price £2.00).

† 'A Memorandum to the Committee of Inquiry into the Engineering Profession'. *The Radio and Electronic Engineer*, **48**, no. 5, pp. 255–6, May 1978.

the art, science and practice of electronic and radio engineering'—the qualifying role always having been in support of, and subordinate to, that primary learned society task. And in this context the IERE welcomes the Report's suggestion that some re-grouping of the Institutions might be advantageous to the profession as a whole, reflecting as it does the point made by our President (Professor William Gosling) in his 1979 Inaugural Address* that it is perhaps time that we gave careful consideration to whether the engineering profession is not now two professions—the old with its scientific basis of Newtonian Mechanics and the new, as represented by the IERE, whose business is founded on Quantum Mechanics which gave rise to the technical basis of the transistor and later of micro-electronics. It also has its own new science of network, control theory, and information, which has resulted in all the present day advances in computer science, telecommunications and automation of industry.

7 Lastly in the general summary of the IERE reaction to the Finniston Report it must be said that the Institution was greatly relieved to see that the Committee was unanimous in the view that the new statutory register of engineers must embrace the nation's current stock of engineers and those at present under training as well as the future products of the formation plans recommended by the Committee. This point is clearly most important since the nation's engineering achievement in the immediate future must depend on the goodwill, enterprise and general performance of the existing members of the profession; and it was for this reason that the IERE was most concerned to note that the Committee had been unable to agree on precisely how the integration of current and future registers of engineers was to be achieved. In this context we would stress the over-riding importance of avoiding any move to 'grade' the nation's existing stock of engineers according to their academic achievement alone, since these people are what they are as a result of the sum total of their education, training, experience and achievements. For this reason the IERE stands convinced that they must be judged 'in the round' rather than on any one of the facets contributing to their present status in, and value to, the profession.

8 Having thus established the necessary general backcloth for our reply we turn now to the key issues and specific questions posed in your letter.

(a) The establishment of a new Engineering Authority (Recommendations 77–80)

The setting up of a statutory body to promote and strengthen the engineering dimension is wholeheartedly supported, though it is considered to be essential that some means be found by which the Secretary of State could ensure that the several specialized engineering disciplines were always well represented in its composition at both Board and working levels. The most cost-effective way of achieving this would appear to be the direct involvement of the highly motivated voluntary expertise already available on the Councils and Standing Committees of the individual specialized engineering Institutions.

* 'Electronics—a profession in its golden age', *The Radio and Electronic Engineer*, **50**, no. 1/2, pp. 1–8, January/February 1980.

(b) The establishment of a statutory register of qualified engineers (38, 60–62)

A statutory register of qualified engineers would be welcomed, though how this could be set up and maintained cost effectively except by absorbing the existing ERB registers and utilizing (as mentioned at para. 8(a)) the undoubted expertise of the specialized professional engineering Institutions cannot be envisaged. Indeed, it has been widely suggested that the Report grossly underestimates the registration function and the IERE would welcome the opportunity to participate in the discussions which must surely soon be initiated to resolve the dilemma stressed in para. 5.24 of the Report.

(c) Accreditation of engineering degree courses, training programmes, and individual applicants for registration (38, 43, 44)

The IERE supports these proposals, but again would emphasize its conviction that the accumulated expertise of the specialized Institutions should be utilized for this purpose, not only on economic grounds but because of their invaluable experience. As to the general approach to, and pattern of, future education and training for engineers as proposed by the Finniston Committee, the Institution will be submitting its considered views on these proposals in the form of position papers for the National Conference on Engineering Education and Training which is to be held under the auspices of the Department of Education and Science in October 1980.

(d) Controls over engineering practice (63–67)

The recommendations are supported.

(e) Measures to encourage the continuing formation of engineers (53–59)

These recommendations are supported, although the IERE would again wish to stress the contribution which the specialized Institutions can make since they alone are qualified in their own disciplines to continue the 'formation' of engineers; and there would need to be some provision for existing qualified engineers.

(f) Future roles and activities of the Institutions (68–74)

Although the present roles and activities of the Institutions are seen to be underestimated in the Report, the recommendations are welcomed, most particularly in their reference to the expansion of learned society activities, and provision of specialized support and assistance to the Engineering Authority.

9 Finally, I am asked by the IERE Council to stress that the Institution stands ready to help in any way it can to ensure the smooth and speedy implementation of all or any of the recommendations of the Finniston Inquiry that will serve to bring about the nation's economic recovery or, at worst, prevent its further decline. And it is in this context that the IERE would press for a harder look at the root cause of the problem underlying the nation's inadequate industrial performance (para. 2 of this letter); the need for urgent and sensitive resolution of the dilemma regarding the criteria to be adopted for registration of existing engineers and those currently undergoing training (paras. 7 and 8 of this letter); and the necessity to make good the shortage of technicians (para. 3.31 of the Report).

Applicants for Election and Transfer

THE MEMBERSHIP COMMITTEE at its meeting on 11th March 1980 recommended to the Council the elections and transfers of the following candidates. In accordance with Bye-law 23, the Council has directed that the names of the following candidates shall be published under the grade of membership to which election or transfer is proposed by the Council. Any communication from Corporate Members concerning the proposed elections must be addressed by letter to the Secretary within twenty-eight days after publication of these details.

March Meeting (Membership Approval List No. 270)

GREAT BRITAIN AND IRELAND

CORPORATE MEMBERS

Transfer from Graduate to Member

DEEDMAN, David Kenneth. *Hockley, Essex.*
McDONALD, Gerard. *Clontarf, Dublin.*
MAINWARING, Harry. *Warrington, Cheshire.*
MORTIMER, Colin Lee. *Camberley, Surrey.*

Direct Election to Member

FITTON, Dennis. *Rickmansworth, Herts.*
GIBB, James Walker. *Dunblane, Perthshire.*

NON-CORPORATE MEMBERS

Transfer from Student to Graduate

ATKINSON, Julia. *Runcorn, Cheshire.*
COLE, Christopher James. *Dartford, Kent.*
GERMER, Robert Ian. *Reading, Berks.*
HOLLAND, Stephen William. *Newcastle-upon-Tyne.*
SCOTLAND, Ronald Ian Henry. *Slough, Berks.*

Direct Election to Graduate

CURD, Colin. *Sunderland, Tyne & Wear.*

Direct Election to Associate Member

*PARNABY, Stanley Allan. *Plymouth, Devon.*
WATT, Brian. *Malvern Link, Worcestershire.*

Direct Election to Associate

WINFIELD, Brian George. *Long Eaton, Nottingham.*

Direct Election to Student

DAVID, Lee John. *Glynneath, West Glamorgan.*
JELF, Alan John. *Swansea.*
KENYON, Paul Douglas. *Old Coulsdon, Surrey.*
THYER, Stephen William John. *Swansea.*

OVERSEAS

CORPORATE MEMBERS

Transfer from Student to Member

FOONG, Chee Hong. *Alor Setar, Kedah, W. Malaysia.*

Direct Election to Member

FERNANDO, G. R. P. *Colombo, Sri Lanka.*
JACOBS, David Adolph. *San Angel Inn, Mexico*
NG, Kwong Lam. *Hong Kong.*

NON-CORPORATE MEMBERS

Direct Election to Graduate

MASOOD, Tariq Khan. *Faisalabad, Pakistan.*

Direct Election to Associate Member

DARAMOLA, Olufemi Olalekan. *Lagos, Nigeria.*
McGARRY, Vincent George. *B.F.P.O.*
RUTHERFORD, Frank Douglas. *Berlin*

Direct Election to Associate

ZUZARTE, Charles Herbert. *Rawalpindi, Pakistan.*

Direct Election to Student

LEE, Wee Cheong. *Singapore.*
LIM, Min Kiong. *Singapore.*
MAN, Wai Choi Kieran. *Shaukiwan, Hong Kong.*
WONG, Chun Kau Jolly. *Kowloon, Hong Kong.*

* Subject to Mature Candidate Regulations.

Members' Appointments

FELLOWSHIP OF ENGINEERING

Among the elections to Fellowship which were made at the Annual General Meeting of the Fellowship of Engineering, held at St James's Palace on February 15th under the Chairmanship of HRH The Duke of Edinburgh, the Senior Fellow, are the following members of the Institution:

Dr P. A. Allaway, C.B.E., D.Tech. (President 1976/77) who is a member of the Main Board of EMI and of the Executive Board of Thorn Electrical Industries, and Chairman of EMI Electronics. He is currently Chairman of the Council of Engineering Institutions.

Professor G. D. Sims, O.B.E., M.Sc., Ph.D. (Fellow 1966) has been Vice Chancellor of the University of Sheffield since 1974; he was previously Head of the Department of Electronics at the University of Southampton. Professor Sims has contributed several papers to the Institution on microwave techniques and on educational matters. He was a member of the Annan Committee on the Future of Broadcasting.

CORPORATE MEMBERS

Sir Ieuan Maddock, C.B., O.B.E., F.R.S., F.Eng. (Fellow 1955, Member 1943), President of the Institution in 1974–75, is to be awarded the honorary degree of Doctor of Technology by the Council of National Academic Awards on May 19th. Sir Ieuan, who was Chief Scientist of the Department of Industry prior to his retirement from Government Service in 1977, was appointed Principal of St Edmund Hall, Oxford, last year; he is Secretary of the British Association for the Advancement of Science.

Dr P. K. Patwardhan, M.Sc., Ph.D. (Fellow 1969, Member 1959, Graduate 1952) has recently received from the Federation of Indian Chamber of Commerce and Industries an Individual Scientist/Technologist Award for his outstanding contributions to technology with particular reference to its interaction with industry. The FICCI's awards of Rs 10,000 are made annually to institutions and individual scientists and technologists. Dr Patwardhan who is a Vice President of the Institution and National Representative of the Council in India, is a senior scientist at the Bhabha Atomic Research Centre, Trombay and Head of its Computer Section. He has previously been honoured both nationally and internationally for his contributions to nuclear electronics, data handling systems and instrumentation.

P. Sample, B.Sc. (Member 1970) has joined the Tecalemit Group as Managing Director of their electronics activities. Since 1977 he had been Divisional Manager of the Radar Division of Microwave and Electronic Systems (now Racal-MESL).

N. A. Vahidy, B.Sc., M.Sc. (Member 1973, Graduate 1969) is Technical Director of the International Consulting Division of Duna Corporation, Riyadh, Saudi Arabia. He was previously Chief Engineer of Riyadh Region, RTN, with Faat Engineering International, Jeddah, Saudi Arabia.


Dr P. A. Allaway


Prof. G. D. Sims


Sir Ieuan Maddock


Dr P. K. Patwardhan

New and Revised British Standards

VARIABLE CAPACITORS

There is a little doubt that variable capacitors are subject to greater misuse and ill-treatment than any other electronic component, probably because of their more robust construction. This has prompted BSI to publish two new British Standards which aim to maintain the performance and extend the useful life of these components by ensuring they are handled correctly during production and throughout their operational service.

The first of these documents is BS 5787 **Guide to the use of variable capacitors in electronic equipment** (£2.60), which emphasizes some of the many abuses to which these items may be exposed during packaging and storage, mounting and assembly. Advice is also given on the soldering and connection of leads, adjustment and measurement of capacitance. The standard is technically identical with IEC Publication 612.

The second standard in this issue is BS 5786 **Variable capacitors** Part 1 **Methods of test** (£7.50), which describes relevant procedures for these components according to their type, dielectric, style and application. The object is to establish, as far as possible, under laboratory conditions, their suitability for use over stated ranges of temperature, air pressure and humidity; also their ability to withstand specified conditions of mechanical shock and the rigours of normal assembly procedures. The standard is identical with IEC Publication 418.

AUDIO-VISUAL EQUIPMENT AND SYSTEMS

BSI has just published BS 5817 **Audio-visual, video and television equipment and systems** Part 10 **Audio cassette systems** (£4.50) and BS 5818 **Helical scan video-tape cassette system using 0·5 in (12·70 mm) magnetic tape (50 Hz–625 line).** (£7.50).

BS 5817 refers to the application for educational and training purposes of magnetic tapes specified in BS 1568 Part 2, which is concerned with dimensions and characteristics of cassettes for commercial tape records and domestic use. It deviates from BS 1568 Part 2 in that it provides for utilization of the tracks in order to secure student response facilities and fully compatible pre-recorded information as well as the safeguarding of this information. The standard applies to cassettes used in audio-active-comparative, audio-passive-visual and audio-active-visual systems. It has five sections covering general topics, magnetic tracks, cue tones, performance requirements and program identification. Appendices refer to AA–AP–AAC systems and AAV–APV systems. This standard is identical with IEC 57–10.

BS 5818 specifies dimensional and other characteristics necessary to ensure interchangeability of cassettes. Requirements given are related to monochrome CCIR systems B, C, D, G, H, I, K, K_1 and L and to the CCIR, PAL and SECAM colour coding systems. It is identical with IEC 511.

FIRE DETECTION AND ALARM SYSTEMS

BSI has recently completed a major revision of its code of practice for electrical alarm systems which has just been published as BS 5839 **Fire detection and alarm systems in buildings** Part 1 **Code of practice for installation and servicing** (£10.50). This replaces the earlier document, CP 1019.

Part 1 contains recommendations for the planning, installation and servicing of alarm systems in and around buildings and deals with a variety of types ranging from simple installations with one or two manual call points, to complex systems incorporating automatic detectors, control and indicating equipment, direct connection to the public fire service and means of signalling to ancillary services such as fixed fire extinguishing devices and fire doors. Ancillary services, however, are not themselves covered.

There are a number of significant changes from the 1972 edition. For example, attention is drawn to the differences between detection and alarm systems intended to safeguard life and those designed to protect property. Guidance is given on methods of zoning protected areas to ensure that in an emergency the precise location of a particular alarm may quickly be established. Another departure from the earlier document is the advice given on the selection of detectors employing different operating principles, such as optical, optical beam, ultraviolet and infra-red beam detectors. In addition, electromagnetic radiation instead of wiring is now accepted for transmitting signals and the use of multiplex systems has also been recognized. A further important change is that the new edition includes recommendations for buildings in multiple occupancy.

Copies of the above British Standards may be obtained from BSI Sales Department, 101 Pentonville Road, London N1 9ND. (BSI subscribing members receive 40% discount.)

Standard Frequency Transmissions

(Communication from the National Physical Laboratory)

Relative Phase Readings in microseconds
NPL—Station
(Readings at 1500 UT)

February 1980	MSF 60 kHz	GBR 16 kHz	Droitwich 200 kHz
1	−4·6	9·1	29·4
2	−4·8	8·6	29·2
3	−4·6	7·7	29·1
4	−4·7	8·1	29·0
5	−4·9	7·1	28·8
6	−4·7	8·2	28·6
7	−4·7	7·0	28·3
8	−4·7	7·9	28·2
9	−4·5	7·4	27·9
10	−4·3	7·7	27·7
11	−4·3	8·0	27·5
12	−4·1	8·4	27·2
13	−4·1	9·3	27·0
14	−4·1	9·4	26·8
15	−4·0	8·4	26·6
16	−3·8	11·0	26·3
17	−3·6	9·2	26·2
18	−3·4	9·2	26·0
19	−3·4	11·9	25·8
20	−3·4	13·7	25·7
21	−3·4	13·2	25·5
22	−3·4	13·6	25·3
23	−3·2	13·0	25·1
24	−3·0	13·4	24·9
25	−3·0	14·2	24·7
26	−3·0	12·8	24·6
27	−2·9	12·4	24·4
28	−3·0	13·7	24·2
29	−3·0	12·2	24·0

Notes: (a) Relative to UTC scale (UTC_{NPL}–Station) = +10 at 1500 UT, 1st January 1977.
(b) The convention followed is that a decrease in phase reading represents an increase in frequency.
(c) Phase differences may be converted to frequency differences by using the fact that 1 μs represents a frequency change of 1 part in 10^{11} per day.

Introductory/Tutorial Paper

Charge-coupled Devices: Concepts, Technology and Limitations

Professor J. D. E. BEYNON, M.Sc., Ph.D., C.Eng., F.I.E.E., F.I.E.R.E.*

Principle of Operation

The structure of a charge-coupled device is, in many respects, rather like a multi-gate m.o.s. transistor, and its operation—with one important difference—is based on m.o.s.t. principles.

To understand the operation of a c.c.d. consider first what happens when an increasingly positive voltage is applied to the gate of the m.o.s. structure shown in Fig. 1(a). Prior to the application of a bias there is a uniform distribution of holes (majority carriers) in the p-type semiconductor;† as the gate is made positive, however, the holes are repelled from the semiconductor immediately beneath the gate, thereby creating a 'depletion layer' (Fig. 1(b)). As the gate bias is increased, the depletion layer extends further into the bulk semiconductor until a point is reached when the potential at the semiconductor/insulator interface (the surface potential, ϕ_s) becomes so positive that electrons (i.e. minority carriers) are attracted to the surface to form an extremely thin ($\sim 10^{-2}$ μm) 'inversion layer' (Fig. 1(c)). The creation of the inversion layer corresponds to the establishing of a conducting channel in an m.o.s. transistor (m.o.s.t.) and the c.c.d. gate potential corresponding to the onset of inversion is called the threshold voltage, V_{th}. By analogy with an m.o.s.t., a c.c.d. in which the inversion charge is negative is often called an n-channel c.c.d.; likewise a device in which the inversion charge comprises holes is called a p-channel c.c.d.

There is one important difference, however, between the m.o.s. capacitor structure of Fig. 1 and an m.o.s. transistor. In an m.o.s.t., the application of a gate bias in excess of the threshold voltage results in the immediate formation of a channel because a copious supply of minority carriers is available from the transistor's source. But in the structure of Fig. 1 no ready supply of minority carriers exists and so, even if the gate voltage is suddenly pulsed well beyond the threshold voltage, an inversion layer cannot form immediately. In the absence of an inversion layer the depletion region extends much further into the bulk semiconductor and most of the potential difference between the gate and the substrate is dropped across this depletion layer. Take, for example, an oxide of 0·1 μm grown on a substrate with 10^{15} acceptor atoms per cm^3; if the gate were suddenly pulsed to 15 V and no minority carriers were available, 13 V would be dropped across the depletion layer and only 2 V across the oxide. If minority carriers are subsequently made available the surface potential will fall as the charge in the inversion layer increases and an increasing fraction of the 15 V will be dropped across the oxide. Given a sufficient supply of minority carriers an equilibrium situation will be established; the c.c.d. is, however, inherently a non-equilibrium device and equilibrium is a condition which rarely occurs.

† An n-type semiconductor could also be used but in this case the gate voltage would be negative.

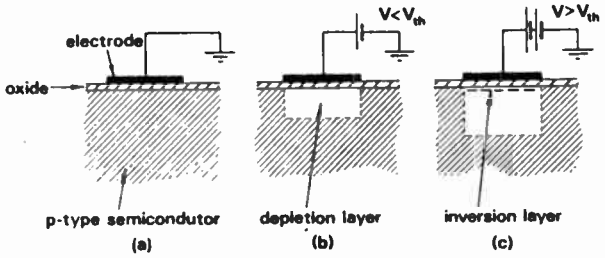


Fig. 1. Single c.c.d. electrode showing the creation of depletion and inversion layers under the influence of an increasingly positive electrode voltage.

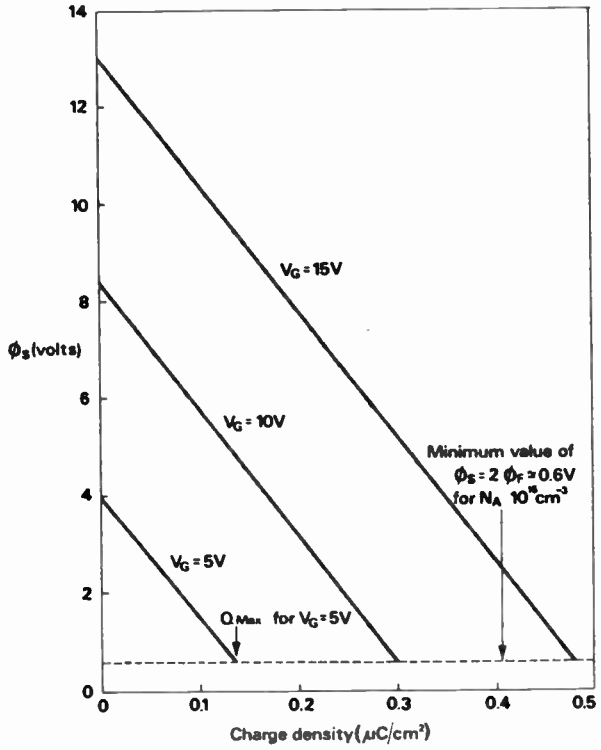


Fig. 2. Variation of surface potential with charge density and gate voltage for the structure referred to in the text.

An analytical expression for the variation in surface potential with the inversion charge, Q, can readily be established using standard m.o.s. theory. Some curves relating ϕ_s, Q and V_G for the structure referred to above are shown in Fig. 2. The almost-linear relationships between these parameters suggest a simple hydraulic model for the charge-storage mechanism, a model which is extremely useful in picturing the operation of all c.c.d. structures: the application of a gate voltage (in excess of V_{th})

* *Department of Electronic and Electrical Engineering, University of Guildford, Surrey GU2 6XH.*

The Radio and Electronic Engineer, Vol. 50, No. 5, pp. 201–204, May 1980

0033-7722/80/050201+04 $1.50/0

creates a 'potential well' the depth of which is related to the magnitude of V_G (see Fig. 3). The introduction of minority charges which bring about a reduction in the magnitude of ϕ_s is analogous to liquid being poured into the well, the depth of which (measured from the top of the well to the surface of the liquid) decreases (Figs. 4(b) and (c)). The well clearly has a maximum capacity Q_{MAX} governed by the magnitude of the gate voltage and the oxide thickness (both of which determine the depth of the well) and the area of the electrode (which determines the area of the well's cross-section); for a given structure and applied voltage, however, the well can contain any amount of charge from zero up to Q_{MAX}.

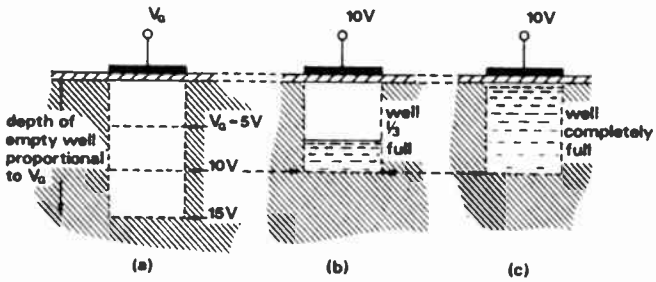


Fig. 3. The potential well concept: the depth of an empty well is approximately proportional to the gate voltage. For a given gate voltage, the depth of the well (measured to the surface of the liquid) decreases linearly with increasing charge.

To understand how a potential well can be moved from one location to another in a c.c.d. structure, consider the arrangement of four closely-spaced electrodes shown in Fig. 4(a); charge is stored initially in the potential well under the second electrode (counting from the left), which is biased to 10 V. All other electrodes are at a potential of say 2 V (which is assumed to be in excess of V_{th} but which gives rise to much 'shallower' potential wells). Now suppose the third electrode is also biased to 10 V. If the two 10 V-electrodes are very close together, their respective depletion layers will merge (i.e. the walls 'overlap' or are 'coupled') and so the charge packet originally under the second electrode becomes shared between the wells under the second and third electrodes (Fig. 4(b)). If now the potential on the second electrode is reduced to 2 V, the remaining contents of its potential well will empty into the third well and we end up with a situation (Fig. 4(c)) similar to that in Fig. 4(a) but with the potential well and its associated charge packet moved one place to the right.

By applying a succession of varying voltages to the c.c.d. electrodes, a charge packet can be propagated just beneath the surface of the semiconductor. To do this using the electrode arrangement shown a three-phase clocking system is required to make the charge move unambiguously in one direction.

It is the c.c.d.'s ability to store and move around a variable amount of charge that gives the device its analogue (as well as its digital) processing capabilities.

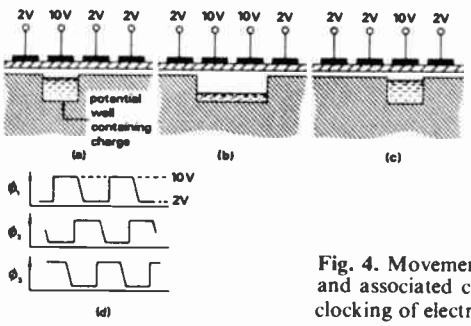


Fig. 4. Movement of potential well and associated charge packet by clocking of electrode voltages.

Charge Inputting and Sensing

Having considered how a charge packet can be stored and moved around in a multi-electrode m.o.s. structure we now consider the problem of introducing charge packets at one end of a c.c.d. shift register and subsequently detecting the emergence of such packets at the other end of the shift register. Although the details of the techniques used to input and detect charge packets may differ according to the application, the principles on which most techniques depend are those shown in Fig. 5.

Charge is introduced into the potential well created periodically under the first ϕ_1 electrode by pulsing the input gate beyond threshold at the appropriate point in the c.c.d. electrode clocking cycle. The potentials on the input diffusion (usually held at a slight reverse bias) and on the input gate control the amount of charge allowed to flow into the first ϕ_1 well. The linearity of the inputting (and output sensing) circuitry is very important in analogue applications. Of course, if the c.c.d. is to be used as an image sensor the inputting circuitry on the left-hand side of Fig. 5 is not needed. The minority carriers are now generated (within a micron or so of the surface of the semiconductor) by the visible light incident on the structure. Thus if the clocking of the electrodes is halted for a short while with say, the ϕ_1 phase, only, being ON, optically generated minority carriers will accumulate in the

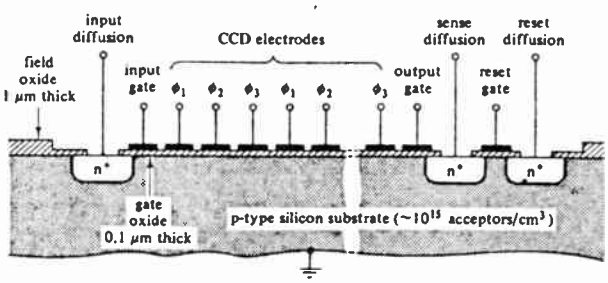


Fig. 5. Diagrammatic representation of a complete c.c.d. system, including input and output circuitry (n^+ denotes heavily doped n-region).

potential wells under the ϕ_1 electrodes at rates proportional to the local light intensity. When the clocking is subsequently resumed, the charge packets delivered to the output will have magnitudes which depend on the light intensity over the linear array. The principle of the c.c.d. linear image sensor can be readily extended to an area sensor.

As with the electrical inputting circuitry, a reverse-biased diode is generally used for detecting charge packets from the c.c.d. shift register. This 'sense' diffusion is, however, strongly reversed biased so that it acts as a sink for any electrons arriving in the potential well associated with the last electrode of the c.c.d. structure. In principle, this electrode could be the last ϕ_3 electrode but in practice it is usual to have an additional electrode—the output gate; this is normally held at a fixed bias (just beyond threshold) and minimizes the electrostatic pick-up by the sense diffusion of the clock pulses on the last ϕ_3 electrode. The simplest way to detect the arrival of charge packets at the sense diffusion is to measure the voltage fluctuations across a series resistor connecting the diffusion to the bias supply. Due to the very small charges that are generally associated with c.c.d.s, however, some form of on-chip amplifier is usually desirable to minimize capacitive loading. For example, the charge in a 'full' packet under an electrode 20 μm × 10 μm formed on 0·1 μm thick silicon oxide insulator is generally only a fraction of a picocoulomb; thus to obtain a reasonable output voltage using a sense capacitance of

much less than 1 pF is required. The simplest form of low-capacitance on-chip amplifier is an m.o.s.t., the gate of which is connected directly to the sense diffusion. In this case, however, a technique must be incorporated in the output circuitry for resetting the potential of the sense diffusion after the arrival of each charge packet. This is achieved by adding a 'reset' gate and 'reset' diffusion as shown in Fig. 5. The reset diffusion is kept at the most positive potential of all and the reset gate is pulsed momentarily after each charge packet has been detected. This resets the potential of the sense diffusion ready for the detection of the next charge packet.

C.C.D. Structures

Almost all the c.c.d.s made so far have been fabricated in silicon using standard planar technology. The c.c.d. structure shown in Fig. 5, for example, could be fabricated using the aluminium-gate m.o.s. transistor process; indeed this was the technology used for all early c.c.d.s. One of the major problems with these early structures, however, was the reliable fabrication of the interelectrode gaps. We assumed in our earlier discussion of charge transfer that the potential wells of adjacent electrodes 'overlap' so that charge moves easily from one well to the next. But consider what this requires in practice: for lightly-doped substrate material ($\sim 10^{15}$ dopant atoms/cm^3) and moderate clock voltages (~ 10 V), the depletion layer beneath the oxide extends ~ 1 μm into the semiconductor; it follows that the interelectrode gaps must be no more than 2 or 3 μm which is considerably smaller than the dimensions used in making conventional integrated circuits.

The technique now almost universally used to make c.c.d. structures is a variant of the silicon-gate m.o.s.t. process (see Fig. 6). The first step in the process is to grow the gate oxide; this is followed by the deposition of the silicon nitride and then a layer of polycrystalline silicon, the latter subsequently being patterned to form the first set of electrodes. A layer of oxide is then formed on the surface of these electrodes by thermal oxidation to provide electrical insulation from the second layer of polycrystalline silicon which is then deposited. This layer is subsequently patterned and oxide insulated, and the process is repeated once more to produce the third-phase electrodes. The final structure allows the fabrication of minimum dimension cells because the inter-electrode gaps are reduced to the thickness of the inter-electrode oxide layers, i.e. of the order of 0·1 μm.

So far we have considered only three-phase structures. Consider, however, the structure shown in Fig. 7 in which a slight addition doping has been introduced (by ion implantation) in the substrate on the left hand side of each electrode. It can be shown that, for a given electrode voltage, the depth of the potential well decreases with increasing doping level. Thus the potential wells under the structure of Fig. 7(a) will be shallower to the left and so, provided too much charge is not introduced into each well, all the charge will be stored on

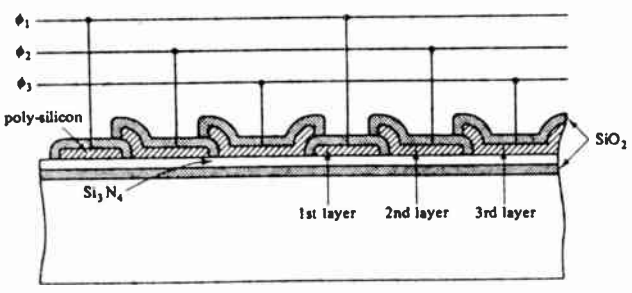


Fig. 6. Three-phase structure employing three overlapping levels of oxidized polycrystalline silicon.

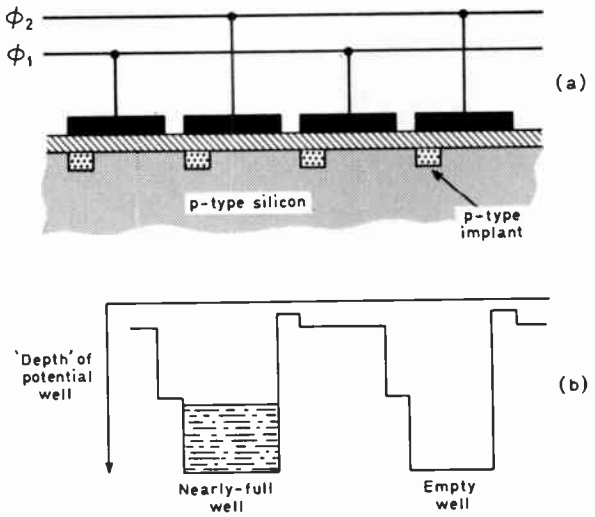


Fig. 7. (a) Two-phase c.c.d. structure. (b) 'Profile' of associated potential wells.

the right-hand side of the well (Fig. 7(b)). Due to this inbuilt asymmetry, it is possible to move the charge from one well to the next using only two clock pulses.

Speed Limitations

In our earlier discussions of charge transfer we did not consider the factors which limit the speed at which c.c.d.s. can be operated. There are, in fact, upper and lower limits on the operating frequency of c.c.d.s directly attributable to physical process occurring within the semiconductor.

Consider first the low-frequency limit. Hole-electron pairs are constantly being generated in the bulk semiconductor by thermal vibration of the crystal lattice. In c.c.d. structures those minority carriers that are generated close to the silicon surface stand a good chance of 'falling' into the potential wells where, of course, they become indistinguishable from the minority carriers conveying information. Thus, irrespective of their initial charge, potential wells which are maintained for long periods of time eventually fill with charge. Clearly then, in operating a c.c.d. shift register each potential well must be clocked through the structure in a time so short that the amount of additional charge picked up en route is small. Typically the maximum transit time is tens or hundreds of milliseconds, although storage times of several seconds are possible in high quality silicon. For applications such as low-light image sensing where even longer storage periods may be necessary, cooling of the silicon crystal, which greatly reduces the generation rate of hole-electron pairs, can be employed to give storage times of many minutes.

The upper frequency limit at which c.c.d.s may be operated is directly related to the processes by which charge transfers from one potential well to the next. The details of the charge transfer process are extremely complex, but the following description, although grossly oversimplified, will serve to indicate the main features of the transfer process. Consider once again the transfer of charge from well 2 to 3 in Fig. 4. We suppose that initially well 2 is full and well 3 has just been created by the application of a clock pulse to the gate. Because well 3 is empty, the surface potential under electrode 3 is much higher than that under electrode 2 (see Fig. 2) and so electrons will move to the right under the influence of a lateral electric field. This drift process occurs relatively quickly but, of course, as the charges move across, the drift field falls. In normal operation, the potential on electrode 2 is reduced shortly after well 3 is

created; this tends to maintain the drift field and so encourage the rapid movement of charge into well 3. There is, however, a danger of reducing the potential on electrode 2 too quickly; if this happens the depletion layer beneath the inversion layer will collapse and cause majority holes to rush up to the surface where they very rapidly recombine with the minority electrons. Clock waveforms typically have a sharply rising leading edge and a linearly falling trailing edge as depicted in Fig. 4(d) which also indicates the phase relationship between the clocks for a 3-phase c.c.d.

Whatever the driving clock waveform, however, the drift field must become very small when nearly all of a full charge packet has been transferred. The remaining charge, therefore, crosses by diffusion which is a relatively slow process characterized by a time-constant $\tau = L^2/D$ where L is the electrode length and D the diffusion coefficient of the charge carriers. D is typically $\sim 10\,\mathrm{cm}^2/\mathrm{s}$ for silicon and so if $L = 10\,\mu\mathrm{m}$ (a reasonable minimum), $\tau \sim 10^{-7}$ s, this implies a maximum clock rate $\sim 10^7$ Hz, a figure borne out in practical structures of the type shown in Fig. 5. Clearly, however, at frequencies of this order there is a trade-off between transfer time and transfer efficiency: the more rapidly the electrode voltages are switched, the more likely it is that some charge will be left behind. Such charge usually falls into the succeeding potential well, giving rise to a 'residual' charge packet.

The 'transfer inefficiency', ε, of a c.c.d. (defined as the fraction of charge left behind when a packet is transferred between adjacent wells) is of paramount importance because it sets a limit on the number of transfers that can be carried out before excessive signal degradation occurs. For surface channel structures like Fig. 5, ε is typically $10^{-3} \sim 10^{-4}$. This may seem to be so small as to be insignificant; bearing in mind, however, that a charge packet in a 300-bit, 3-phase structure will experience nearly 10^3 transfers, one can appreciate the need for a very low value of ε. The degree of signal degradation is in fact determined by the *transfer inefficiency product*, $n\varepsilon$, where n is the number of charge transfers through the device. The significance of this quantity for digital applications can be seen in Fig. 8, where, when $n\varepsilon = 1$, the signal charge packet cannot be distinguished from the first trailing 'zero'. When considering analogue applications, the $n\varepsilon$ product influences the gain and phase shift experienced by the signal in passing through the c.c.d. Thus, if a sinusoidal signal of frequency f is transferred through the device, the gain and phase shift are given by

$$G = \exp\left\{-n\varepsilon\left[1-\cos\frac{2\pi f}{f_c}\right]\right\}$$

$$\Delta\phi = -n\varepsilon\left\{\frac{2\pi f}{f_c} - \sin\left(\frac{2\pi f}{f_c}\right)\right\}$$

where f_c is the clocking frequency. (See Fig. 9.)

Buried-channel C.C.D.s

It was to reduce the charge transfer inefficiency and to improve the high-frequency performance that a variant of the structure shown in Fig. 5 was developed. In the so-called buried (or bulk)-channel c.c.d. (b.c.c.d.), a cross-section through which is shown in Fig. 10, the potential minimum for electrons occurs about half-way through the thin n-type layer (produced by ion implantation). The electrons thus move laterally in the bulk rather than at the surface of the silicon and this results in two beneficial effects. Firstly the electrons do not get caught up in the 'traps' which are characteristic of silicon–silicon dioxide interfaces; this decreases ε to about $10^{-4} \sim 10^{-5}$. Secondly, the fringing electric field between adjacent electrodes now aids

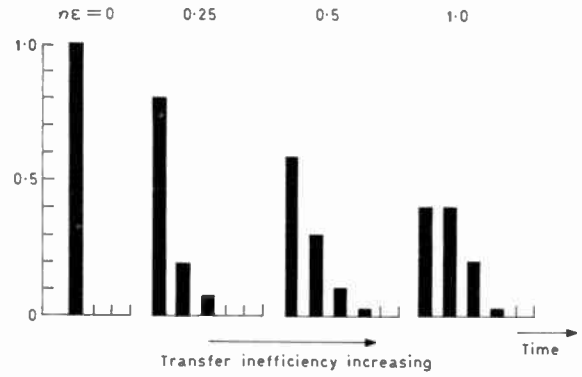


Fig. 8. C.c.d. output for input signal of single 'one' followed by a stream of 'zeros' showing effect of increasing transfer inefficiency.

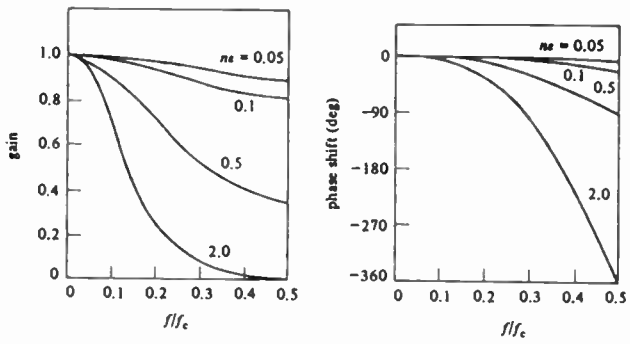


Fig. 9. Effect of transfer inefficiency on propagation of sinusoids through c.c.d. shift register. Left curves show gain versus signal frequency (normalized to clock frequency); right curves show phase shift versus signal frequency. The parameter $n\varepsilon$ is the product of number of transfers and transfer inefficiency.

charge transfer between wells, resulting in a much better high frequency performance, such devices having been operated at hundreds of megahertz. The charge-carrying capability of b.c.c.d.s is slightly inferior ($\sim 50\%$) to that of surface channel structures but this is not a serious disadvantage. Rather more significant, particularly in low-frequency applications, are the higher leakage effects in b.c.c.d.s due to the larger volume of depleted silicon.

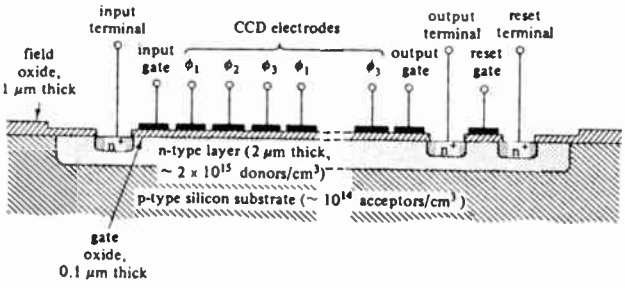


Fig. 10. Diagrammatic representation of a typical buried-channel c.c.d. fabricated in silicon (n^+ denotes heavily doped n-type regions).

Bibliography

'Charge-Coupled Devices and their Applications', Beynon, J. D. E. and Lamb, D. R., (McGraw-Hill, Maidenhead, 1980).
'Charge-Coupled Devices', Barbe, D. F. (Springer-Verlag, Berlin & Heidelberg, 1980).
'Charge-Transfer Devices', Hobson, G. S. (Arnold, London, 1978).
'Charge-Transfer Devices', Sequin, C. H. and Tomsett, M. F. (Academic Press, New York, 1975).

(*Paper No. 1929*)

UDC 621.382: 621.397.331

Indexing Terms: Semiconductor devices, Integrated circuits, Charge-coupled devices, Television cameras

Development of c.c.d. area image sensors for 625-line television applications

D. J. BURT, B.A.*

SUMMARY

The development of large area c.c.d. image sensors is described and illustrated with the example of a frame-transfer array with 385 × 576 image elements intended for 625-line television applications. The problems that still exist in meeting more demanding applications such as broadcast television are discussed and the various possible solutions are described.

* *GEC Hirst Research Centre, East Lane, Wembley, Middlesex HA9 7PP*

1 Introduction

A major objective for the semiconductor industry has been to produce a solid-state equivalent to the conventional electron-beam scanned television camera tube. Such a device would offer important features such as small camera size, low voltage operation, extreme ruggedness and very long operating life. However, in spite of the major advances that have been made in semiconductor technology, it has been found difficult to design a solid-state image scanning mechanism that could match the elegant simplicity and performance of the electron beam. It is only comparatively recently with the advent of the charge coupled device that satisfactory means have been available to overcome the various practical problems.

2 Basic Array Types

There are many possibilities for the design of practical solid-state area image sensors using c.c.d. technology. Ideally the design should employ a read-out sequence that matches the format of a conventional line-scanned television display with all of the signal charges transferred to a single low capacitance output amplifier to maximize the signal-to-noise ratio. The most successful approaches have been:

(a) the frame-transfer array[1] (f.t.);
(b) the inter-line transfer array[2] (i.l.t.);
(c) the charge injection device[3] (c.i.d.);

and variations thereof. These arrays are illustrated schematically in Fig. 1. The basic operation of each is as follows. The f.t. array collects a 'frame' of photogenerated charge in the image section which is transferred down into the shielded store for subsequent sequential line-by-line read-out through the charge detection amplifier. A second frame is being collected in the image section whilst the first is being read out through the store, and so on. The i.l.t. array is similar to the f.t. array, but the shielded store is now spaced between the columns of the image section. Transfer of a column of photogenerated charge from image to store sections now takes place in parallel, rather than serially as in the f.t. array. The c.i.d. differs from both f.t. and i.l.t. arrays in that charge transfer is not employed to read out signals from the array. Instead, an X–Y addressed array of m.o.s. capacitor pairs is used to store photogenerated charge. The signal charge is detected sequentially by detecting voltage changes on the capacitance of the address lines, and then cleared by injection into the substrate.

Table 1 lists some of the principal applications of solid-state area sensors, together with an indication of the more important operating parameters and the relative usefulness of the various array types. The choice of an array type for any given application will generally be a compromise based on consideration of the major

The Radio and Electronic Engineer, Vol. 50, No. 5, pp. 205–212, May 1980

0033-7722/80/050205+08 $1.50/0

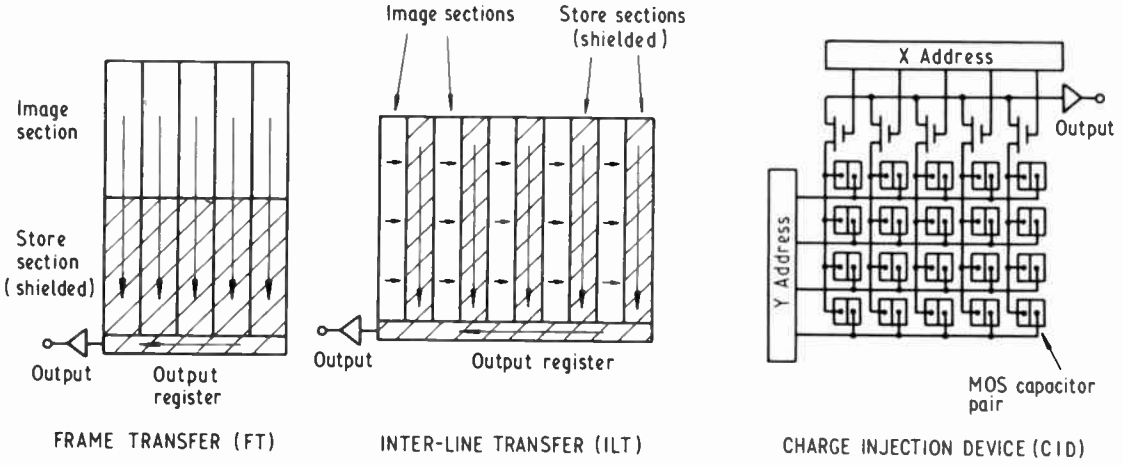


Fig. 1. Frame transfer, inter-line transfer and charge injection device array schematics.

performance parameters and characteristics, such as:

spectral response and optical sensitivity;
dynamic range;
resolution;
ease of operation;

and other factors such as commercial availability. It may be noted, for example, that the opaque storage areas in the i.l.t. array will mask small image detail, making the arrays unsuitable for applications such as tracking, astronomy and star sensing where point images are encountered. The c.i.d. has a poorer signal-to-noise ratio than the c.c.d. types since the charges are detected on the relatively higher capacitance of the address lines. At normal operating temperatures, however, the dark current non-uniformity tends to dominate the temporal noise sources (i.e. as a fixed pattern noise) and the noise advantage of the c.c.d.s is lost. Overall it is clear that the f.t. array is the most versatile and, as such, has been the type chosen by this laboratory for development. This work has led to the design and fabrication of an array having 385 × 576 photosensitive elements on a chip 1·0 cm × 1·4 cm with an image area equivalent to that of a $\frac{2}{3}$ in (1·7 cm) vidicon and an output that matches the European 625-line television format.

Table 1

Choice of array types for various applications

Application	Requirements	F.T.	I.L.T.	C.I.D.
industrial	simplicity	G	G	G
inspection	small size camera	G	E	G
surveillance	high sensitivity and dynamic range	G	M	G
home video	low cost colour	G	G	M
astronomy	high quantum efficiency, low noise	E	P	G
star sensing	precise sensor geometry	E	P	M
tracking	fast read-out	M	P	E
broadcast	very high image quality, colour	G	M	P

E = excellent, G = good, M = medium, P = poor.

3 Design of the 385 × 576 Element Array

The organization of the array with 385 × 576 photosensitive image elements is shown schematically in Fig. 2. The design makes use of three separate levels of semi-transparent polysilicon to form a three-phase overlapping electrode structure,[4] as shown in Fig. 3. The three electrodes (ϕ_1, ϕ_2, ϕ_3) required to store one charge 'packet' (i.e. signal charge) are referred to as one 'element' of the c.c.d. Although the fabrication sequence for such a structure is relatively complex, this technology has two considerable advantages for the production of large area image sensors. Firstly, the layout dimensions can be quite coarse (i.e. the electrode width and spacing) in comparison with other technologies such as two or four-phase, a factor that is beneficial in achieving the necessary perfection in the photolithography. Secondly, there is a reduced probability of inter-phase short circuits fatal to device operation since all the electrodes in one level of polysilicon are connected to the same phase and the three levels are separated by oxide insulation of high integrity formed by thermal oxidation of the polysilicon itself.

The array design has an image section that comprises 288 lines (i.e. 288 vertical c.c.d. elements or 864 phase electrodes) each of 385 horizontal photosensitive elements. Separation between the horizontal elements or columns is by means of vertical strips of channel-stop diffusion. Separation between vertical elements is achieved by appropriate biasing of the phase electrodes. Both the horizontal and vertical spacings between photosensitive elements are 22 μm, the vertical spacing being achieved with two phase electrodes of 7 μm active length and one of 8 μm, similar dimensions being employed for the read-out section. The image area is therefore approximately 8·5 mm horizontally by 6·4 mm vertically in a standard 4:3 aspect ratio and is of a size roughly equivalent to that obtained with a $\frac{2}{3}$-in vidicon.

For television applications the number of picture lines can be effectively doubled to 576 by displaying two

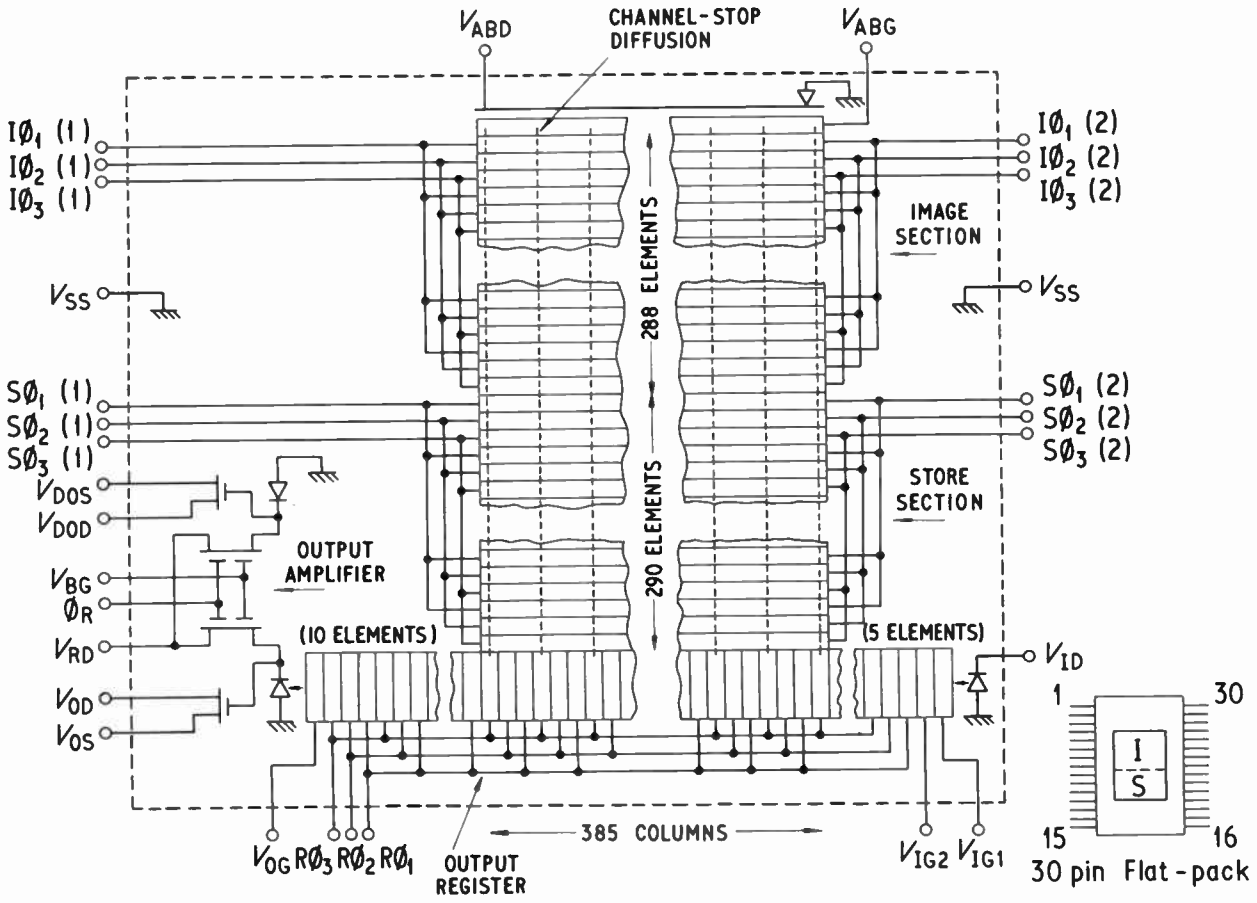


Fig. 2. Organization of 385 × 576 element array.

interlaced fields each of 288 lines, as described later. Thus the device is fully compatible with European 625-line television systems, where 575 active lines are usually displayed and the remaining 50 are used for the field blanking periods.

The storage section consists of 290 vertical by 385 horizontal c.c.d. storage elements, again of size 22 μm × 22 μm. The two extra vertical elements are to contain any charge residues that might arise through inefficient charge transfer out of the image section. The element dimensions of image and store sections are identical to allow the whole array to be operated in the 'full-frame' mode if so desired. In this mode, common connections are made to the image and store sections such that the whole chip is used for imaging. However, this is only feasible with long integration times under cooled conditions or with shuttering such that the relatively long read-out period of this approach does not give rise to significant frame-shift smear (i.e. spurious charge picked up during read-out).

The output register has a total of 400 elements. One register element is associated with each of the horizontal elements in the array, plus ten extra at the output end of the register (and five at its input) which may be used to establish a black reference level at the start of each line.

The charge detection amplifier is of the conventional reset-diode plus source follower type, a second 'dummy' amplifier (i.e. it receives no signal charge) being incorporated to enable residual pulse feedthrough in the output (mainly from the reset pulse) to be cancelled by a differential amplifier in the external electronics. The output register also has an electrical input for test purposes.

The overall chip size for the array is 1·0 cm × 1·4 cm. The device design is such that no single mask dimension smaller than 8 μm is employed, a feature that is

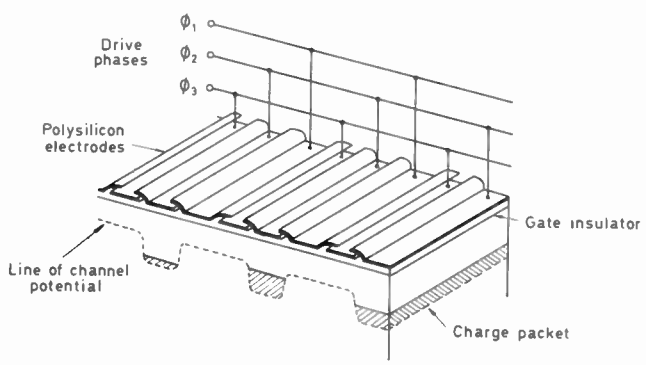


Fig. 3. Three-phase, three-level polysilicon electrode structure.

advantageous in achieving a high manufacturing yield with devices of this physical size. The exception is that the channel stop isolation between columns is reduced to 4 μm to maximize the charge storage area of each c.c.d. element. The channel stop region is not totally insensitive to light and the optically 'dead' area between elements is significantly smaller than the physical dimensions alone suggest.

4 Device Fabrication

Photomasks for fabricating the 385 × 576 element array were produced using an optical pattern generator to form the reticle plates at ×4 magnification. The reticle plates were then stepped-and-repeated to produce the master arrays from which copies can be taken for use in device fabrication.

The devices are fabricated using buried n-channel m.o.s. technology with the buried channel incorporated via phosphorus ion-implantation at an early stage of the process. A p-type substrate is employed with p-type channel-stop and n-type source/drain diffusions. After formation of the gate insulator, the three polysilicon layers are sequentially deposited, doped, patterned and thermally oxidized. Fabrication is completed by etching contact windows through the oxide layers and the provision of aluminium metallization for inter-connections and the wire bonding pads. The processing sequence also includes various hydrogen annealing steps aimed at reducing the density of fast surface states and hence the level of background dark current.

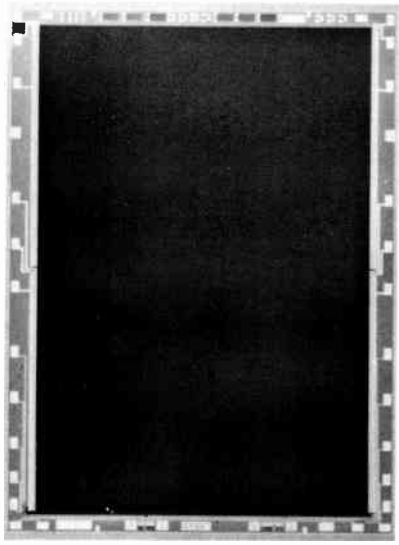


Fig. 4. Micrograph of 385 × 576 element array.

A photomicrograph of a completed array is shown in Fig. 4. Devices in silicon wafer form are first operated on a probe-test station and functioning devices are mounted in a 30-pin flat-pack.

5 Device Operation

Operation of the 385 × 576 element array in the frame-transfer mode requires application of three sets of three phase drive pulses ($I\phi_{123}$, $S\phi_{123}$, $R\phi_{123}$), an output reset pulse (ϕ_R) and a number of d.c. bias levels (24 V max.). A pulse timing diagram for operation of the array in the 625-line television compatible mode is shown in Fig. 5. Interlace (as required for a conventional television display) is achieved[5] as follows. A first field is obtained by collecting photogenerated charge under the ϕ_1 electrodes of the image section. These charges are read out and the drive pulse sequence altered such that the next field is collected under the electrodes of the other two phases, $I\phi_2$ and $I\phi_3$, and then read out (after charge amalgamation under $I\phi_2$) as before. In this way the centres of charge collection are shifted back and forth to give a 2 : 1 interlace with each c.c.d. element behaving as though it were two sensing elements and the number of television lines is effectively double the number of c.c.d. elements in the image section of the array. Thus 288 vertical c.c.d. elements are adequate to achieve full 625-line television capability.

The field time is 20 ms and the line time 64 μs. Of the 20 ms field time, 25 line times (1·6 ms) are conventionally used for the field blanking period, which thus sets a maximum time for the frame-transfer to take place. As there are 290 lines in the storage section, the minimum clock frequency for frame transfer is thus 180 kHz. A higher frequency is however beneficial for reducing the effects of frame-shift smear, but this can be difficult to achieve in practice as the phase electrode capacitances are quite high (~ 5 nF). A frequency of ~ 400 kHz is found to be possible without over-complex electronics. After frame transfer the store section is pulsed at 15·6 kHz, with the last line of the store being transferred to the output register during the 12 μs non-active part of the line time. Since there are 385 horizontal elements in the array, and these must all be read out during the 52 μs active line time, this necessitates an output clock frequency of 7·4 MHz. In practice it is found more convenient to read out all 400 elements in 52 μs, in which case the clock frequency is 7·7 MHz and the television display is slightly compressed in the horizontal direction.

Prototype drive electronics have been assembled using a commercial sync. pulse generator chip and other TTL logic to produce the basic timing waveforms. Pulse buffers are then employed to translate between the TTL levels and the typically 10 V levels required to operate the array. The optimum pulse shape for driving the array has been determined by experiment. The distributed *RC* time-constant effect of the resistive polysilicon electrodes in combination with the capacitance of the gate insulator has been simulated with a representative chain of discrete resistors and capacitors. Pulses with symmetrical rise and fall time were found to be optimum. In this way one

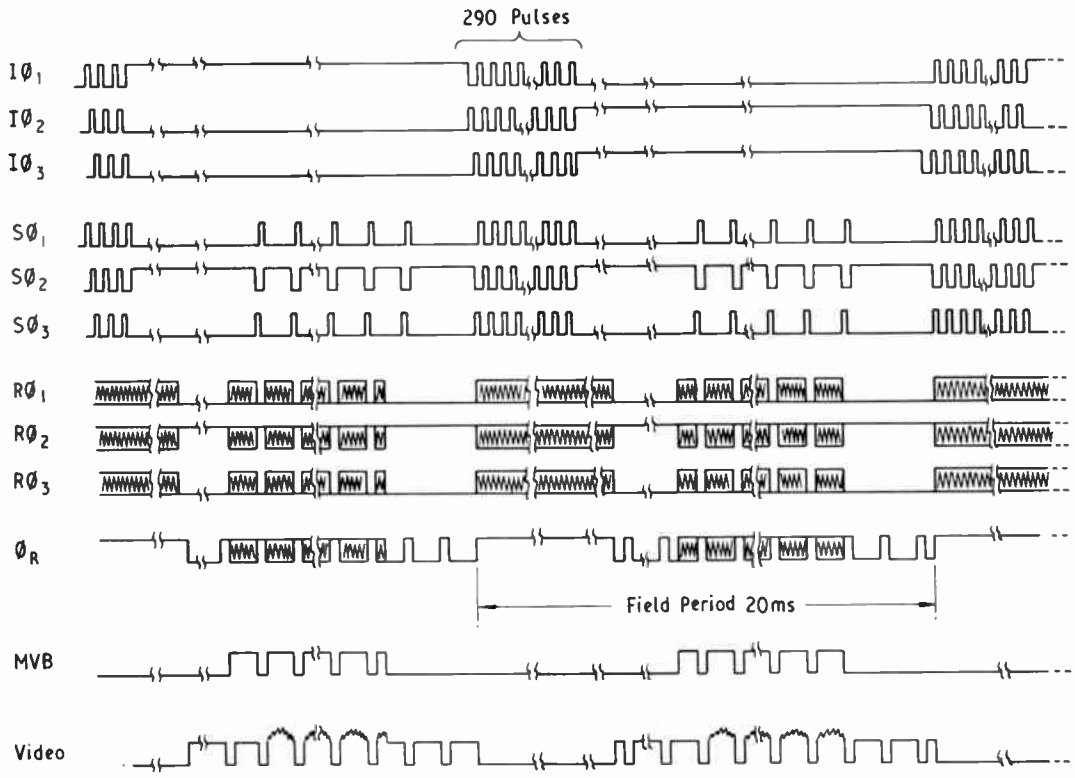


Fig. 5. Pulse timing diagram for 625-line television operation.

pulse turns off as the next turns on and cross coupling between waveforms is minimized. Such waveforms are also ideal for efficient charge transfer in buried channel structures.

6 Results

A video display of the output from one of the 385 × 576 element arrays is shown in Fig. 6. Some picture blemishes are evident, but in general the image quality is

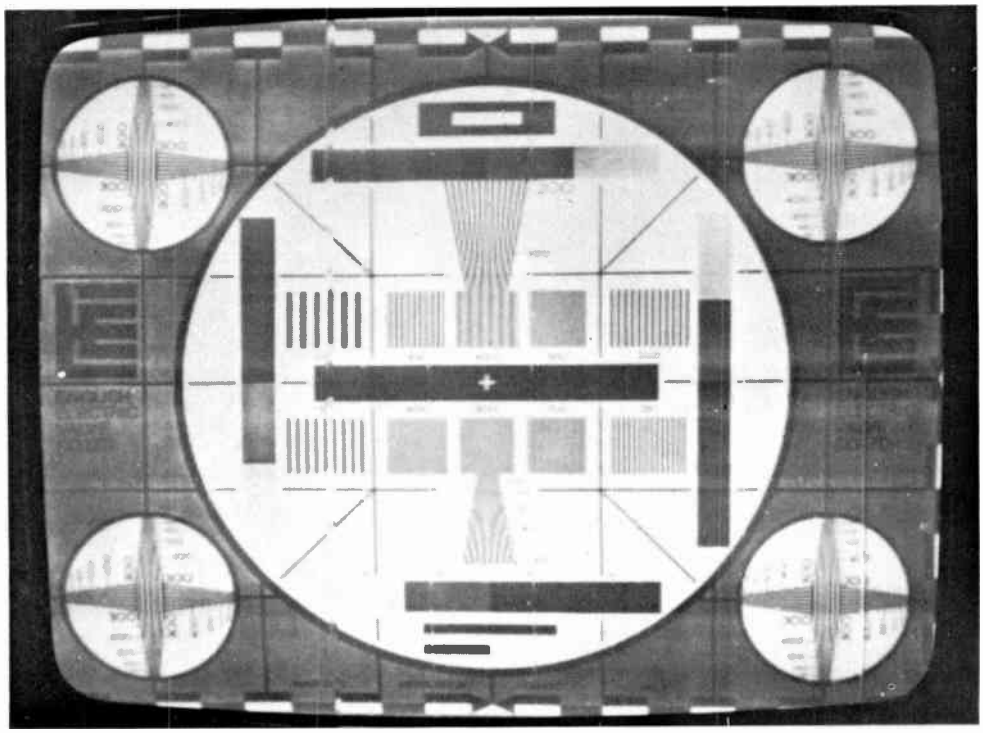


Fig. 6. Video display from a 385 × 576 element array.

good. The relevant electrical characteristics of the array are:

Peak signal current (measured at V_{RD} with uniform illumination at saturation)	600 nA
Background dark current (25°C)	$\lesssim$ 6 nA
Output amplifier noise	$\sim$ 0·5 nA r.m.s.

The peak signal current is equivalent to approximately 1·1 μA cm^{-2} of image area, in good agreement with the charge handling calculated from the buried channel layer parameters, and corresponds to about 6×10^5 electrons per photosensitive element. Although the output noise is low, at 25°C the dark current non-uniformities tend to limit the effective dynamic range to about 100 : 1.

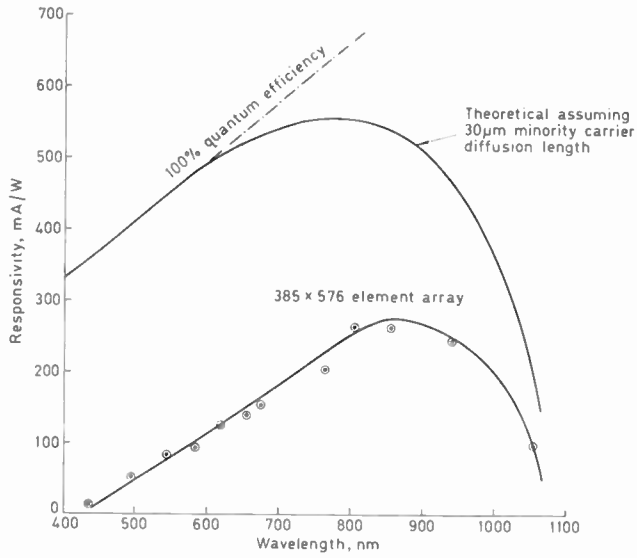


Fig. 7. Spectral response curve.

The spectral response curve is shown in Fig. 7. The mean quantum efficiency is about 25%. A response lower than ideal occurs through interference and absorption of light in the polysilicon electrode layers. Normally, using polysilicon electrodes, interference causes a severe ripple to be superimposed on the spectral response curve with the position of the peaks and troughs dependent on the thickness of the polysilicon. However, in the case of a three-level electrode structure, the thicknesses of the three levels will tend to differ slightly, thus the peaks and troughs will occur at different wavelengths in each layer and tend to average out overall. This results in a relatively smooth spectral response curve. In cases such as star sensing where the image may be smaller than the width of an electrode, the ripple in response due to one electrode layer will be observed.

With white light (3000 K) the device responsivity is $\sim$ 2 mA/lumen and the illumination on the image section for saturation is $\sim$ 5 lux. Thus with f2 optics, the device can be expected to operate at scene illumination levels of about 100 lux, assuming 0·5 scene reflectance and 0·8 lens transmission.

The resolution of the array may be expressed as the modulation depth in the output signal that arises from a black–white bar-pattern imaged on the device, commonly called the modulation transfer function or m.t.f. It is convenient to normalize the spatial periodicity of the bar-pattern, f_0 line pairs per mm, to the spatial pitch of the sensor elements, f_s elements per mm. Unambiguous resolution of image detail having a spatial periodicity higher than $f_s/2$ is not possible because this would contravene the Nyquist sampling theorem. Using visible light, the measured m.t.f. for the 385×576 element array approaches 100% for relatively coarse image detail, i.e. $f_0 \ll f_s$, and decreases as f_0 is increased, reaching about 65% at the horizontal Nyquist limit ($f_0 = f_s/2$) and 0% at the vertical Nyquist limit. The behaviour in the vertical direction is to be expected since the method of interlace makes use of one c.c.d. element to form two photosites. Hence, at the Nyquist limit with a black–white bar-pair imaged on each c.c.d. element, the net output will be an average grey level in each field. At longer wavelengths there is a reduction in the output modulation depth (i.e. a loss of resolution) due to the effects of minority carrier diffusion in the substrate.[6]

7 Discussion

The realization of a 385×576 element array is an important step in the development of solid-state image sensors. The performance is adequate for a variety of industrial, professional and military applications not requiring the highest resolution but where considerations of long life, small size, precise sensor geometry and low voltage operation are important. Further improvement will be necessary to meet more demanding applications such as broadcast television. The major problem areas and the various practical solutions that will be the subject of further development are now described.

7.1 Resolution

The normal video bandwidth for 625-line television is 5·5 MHz. Since the active line time is 52 μs, the Nyquist sampling theorem suggest that at least 572 horizontal photosites are required for full television resolution. However, achieving this number of elements will require the dimensions of the c.c.d. elements in the read-out register to be reduced considerably (i.e. to reduce the column spacing) and/or the chip size to be increased, both of which will tend to affect the manufacturing yield adversely. A possible method[7] for increasing the horizontal resolution without increasing the demands on the device fabrication technology is to use 'horizontal interlace' in the same way that the vertical interlace in a frame transfer array effectively doubles the number of sensor elements.

7.2 Spectral Response

Various methods have been proposed for improving the spectral response of the c.c.d., particularly in the blue. Other materials more transparent than polysilicon, e.g. tin oxide, are currently under research,[8] but serious problems are being encountered regarding the compatibility with the other fabrication processes in the silicon chip. An alternative approach[9] to obtain highest sensitivity and good blue response is to thin the silicon chip behind the image section and to image the device from the back-face. Light then passes directly into the silicon substrate without being attenuated by an electrode layer. This technique is however very difficult to implement since the final silicon thickness must be about 10 μm for good optical response. Packaging is also a problem. It is unlikely, therefore, that devices of this type will become commercially available in the near future.

A compromise solution that appears to be gaining popularity is to design the c.c.d. electrode structure such that part of the image element is free of polysilicon.[10] Such structures are sometimes referred to as 'photodiode sensors'. Although not ideal, an improved blue response is obtained without special fabrication techniques. Furthermore, the simplicity of the front-illuminated array is maintained.

7.3 Dark Current

Dark current arises from thermal generation of minority carrier in the array. Unfortunately, the dark current generation tends not to be uniform over a whole array and the fixed pattern noise introduced by non-uniformity is often the major factor in determining sensor sensitivity and dynamic range. In addition, many devices show exceptionally high dark current generation for some elements in the array—the so-called dark current 'spikes'—probably through contamination picked up during manufacture. These appear as white spot defects in the display. An important part of process development is reduction of dark current and a level of ~ 1 nA cm^{-2} (25°C) can be predicted for the near future.

Dark current is highly temperature dependent, decreasing by a factor of approximately 2 for every 10 degC fall in temperature. Cooling is thus a possibility for reducing dark current and its non-uniformity, thereby extending the dynamic range of an array. Such a procedure also makes possible the long charge collection times required for astronomical observation. Concealment of a few spike blemishes may also be a practical proposition since the output data from each c.c.d. element is precisely located in the read-out sequence. Impaired video information at these locations can thus be removed electronically and substituted with values interpolated from the outputs of adjacent picture elements.

7.4 Blooming Control

Operating a c.c.d. image sensor with illumination levels above saturation gives rise to 'blooming' as excess-generated carriers spread sideways from the area of overload into adjacent picture elements in the array. Blooming thus results in loss of picture information and is subjectively unpleasant to the viewer. On the other hand, 'image burn' from an optical overload as is found with vidicons does not occur with the c.c.d.

The basic approach to providing anti-blooming for c.c.d.s is that excess charge must be soaked up before it can spread to other picture elements in the array. One technique[11] is the electrical analogue of the overflow pipe of the domestic bath. Diffused drain structures are provided in the array in place of the channel stop in the image section and biased so that excess charge flows into these rather than into adjacent c.c.d. elements. An m.o.s. gate or implanted barrier region is necessary next to the drain to effect the necessary control. Images from an experimental 120×150 element array with anti-

(a) (b)

Fig. 8. Images from a 120×150 element array with anti-blooming structures.

blooming drains are shown in Fig. 8. In (a) the candle is imaged on the unprotected part of the array included as a control and blooming from this bright light, approximately 300 times saturation, has completely obliterated the picture. By comparison, (b) shows the candle on the protected part of the array with no noticeable charge spreading.

Anti-blooming structures are considered to be necessary for most imaging applications, but complicate the device structure and take up chip area that is optically dead since charge photogenerated in a drain will be lost. Some further development to optimize the structure is therefore indicated.

7.5 Colour Cameras

Colour television cameras generally employ three separate sensors for the red, green and blue colour channels, each with the appropriate colour separation filter. A similar approach is possible with c.c.d. sensors, provided of course that they have an acceptable spectral response. The sensitivity of a colour camera is usually less than a monochrome unit using the same sensors because the colour separation filters are quite narrow in wavelength interval (about 40 nm) and only a small fraction of the total incident light power is available for imaging. The theoretical sensitivity of an ideal c.c.d. colour camera is not high, mainly because of the low response of silicon in the blue, with scene illumination levels of 3000 lux and more predicted for device operation at normal signal levels. For the c.c.d. camera to be competitive with tube cameras, which give useful operation at a few hundred lux, the c.c.d. sensors will have to operate at far lower signal levels. Thus low dark currents are essential to make this possible and it may even be necessary to cool the blue channel device.

A potential problem with solid-state colour cameras is that there can be no equivalent to the adjustment of scan waveforms in tubes to compensate for aberrations in the colour-splitting optics. A far higher precision will thus be required from the optical components in the system, especially with regard to the variations with temperature. As it may be difficult, therefore, to maintain the optical alignment between the three colour channels, there is growing interest in the possibility of single sensor colour cameras. The use of coloured dyes or filters deposited directly on the array[12] is a practical possibility in this respect.

7.6 Scan Formats

Arrays will need to be designed for a particular line standard, i.e. separate devices and drive circuitry will be required for 625-line and 525-line cameras. This is in contrast to tubes where the scan can be externally controlled as required. The design of a dual standard c.c.d. would be too complex to be commercially viable.

8 Conclusions

The realization of an array with 385×576 image elements has demonstrated that charge coupled image sensors can be produced with many of the performance characteristics necessary for 625-line television. The achievement of a significant yield of low picture-blemish devices of this complexity will be extremely difficult, however, and the initial cost of a high quality c.c.d. sensor will probably be greater than that of a tube. It is likely, therefore, that initial applications will be the less demanding, such as home video and possibly electronic news gathering (ENG). Several more years of development will then be necessary before the very demanding requirements of studio quality television can be satisfied.

9 References

1 Séquin, C. H., *et al.*, 'A charge-coupled area image sensor and frame store', *IEEE Trans. Electron Devices*, **ED-20**, pp. 244–52, 1973.

2 Steffe, W., Walsh, L. and Kim, C. K., 'A high performance 190×244 CCD area image sensor array', Proc. 1975 CCD Applications Conference, San Diego, pp. 101–8.

3 Burke, H. K. and Michon, G. J., 'Charge-injection imaging: operating techniques and performances characteristics', *IEEE J. Solid-State Circuits*, **SC-11**, no. 1, pp. 121–28, February 1976.

4 Bertram, W. J., *et al.*, 'A three-level metallization three phase CCD', *IEEE Trans. Electron Devices*, **ED-21**, no. 12, pp. 758–67, December 1974.

5 Séquin, C. H., 'Interlacing in charge-coupled imaging devices', *IEEE Trans. Electron Devices*, **ED-20**, no. 6, pp. 535–41, June 1973.

6 Seib, D. H., 'Carrier diffusion degradation of mtf in charge coupled imagers', *IEEE Trans. Electron Devices*, **ED-21**, no. 3, pp. 210–17, March 1974.

7 Le Couter, G. M., 'Solid-state image sensors, improvements in signal processing', BBC Research Report No. 1976/4.

8 Schroder, D. K., 'Transparent gate silicon photodetectors', *IEEE J. Solid-State Circuits*, **SC-13**, no. 1, pp. 16–23, February 1978.

9 Shortes, S. R., *et al.*, 'Characteristics of thinned backside-illuminated ccd imagers', *Appl. Phys. Letters*, **24**, no. 11, pp. 565–7, 1974.

10 Ochi, S., Yamanaka, S., Kanoh, Y. and Nishimura, T., 'A device structure and spatial spectrum for checker-pattern ccd colour camera', *IEEE J. Solid-State Circuits*, **SC-13**, no. 1, pp. 187–92, February 1978.

11 Kosonocky, W. F., *et al.*, 'Control of blooming in charge-coupled imagers', *RCA Review*, **35**, no. 1, pp. 3–24, March 1974.

12 Dillon, P. L. P., Lewis, D. M. and Kaspar, F. G., 'Colour imaging system using a single CCD area array', *IEEE J. Solid-State Circuits*, **SC-13**, no. 1, pp. 28–33, February 1978.

Manuscript first received by the Institution on 28th November 1979 and in final form on 14th February 1980.
(Paper No. 1930/CC 322)

UDC 621.382:621.372.54

Indexing Terms: Semiconductor devices, Large scale integrated circuits, Digital filters, Programmable transversal filters, Least-mean-square algorithm

A monolithic c.c.d. programmable transversal filter for analogue signal processing

J. MAVOR, B.Sc., Ph.D., C.Eng., M.I.E.E., M.I.E.R.E.*

P. B. DENYER, B.Sc.†,

J. W. ARTHUR, B.Sc., Ph.D., M.Inst.P.‡

and

C. F. N. COWAN, B.Sc.*

SUMMARY

This paper describes the operational features and performance of a fully-integrated programmable transversal filter (p.t.f.), using c.c.d./m.o.s.t. technology. The choice of filter architecture for a prototype realization is discussed with particular reference to a novel multiplier array implementation using a single, time-multiplexed m.o.s. transistor. The performance characteristics of a prototype, 64-point filter design based on this approach are detailed with reference to frequency- and matched-filtering. Techniques for optimizing the performance of this analogue filter structure under microprocessor control are suggested, through the iterative adaption of the filter impulse response, and equivalent results are given to show the improvement gained. An alternative technique for improving the filter characteristics which enables it to optimize the processing of signals under certain conditions has also been demonstrated. This adaptive filter configuration is based on the linear Widrow least-mean-square (W.l.m.s.) algorithm, and has been realized using the p.t.f. with minimal additional circuitry, without the requirement for a microprocessor.

A general signal-processing module of 256-points using four cascaded filters is described; and results are presented when it is used in a sonar, matched-filtering experiment. Also a 64-point adaptive filter based on a prototype p.t.f. is described and its application to inverse filtering and self-tuning filtering is demonstrated.

Finally, the potential of this miniature integrated filter for sonar-type applications is reviewed against new developments. In particular, a 256-point monolithic p.t.f. currently in development, and the concept of a dedicated adaptive filter in single chip form.

* *Department of Electrical Engineering, University of Edinburgh, Mayfield Road, Edinburgh EH9 3JL, Scotland.*

† *Wolfson Microelectronics Institute, University of Edinburgh, Mayfield Road, Edinburgh EH9 3JL, Scotland; now with Denyer-Walmsley Microelectronics, Ltd., Edinburgh, Scotland.*

‡ *Racal-MESL Ltd., Microwave Division, Newbridge, Edinburgh; formerly with WMI.*

1 Introduction

The development of charge-coupled devices[1] in complex configurations has permitted many electronic systems to be produced in compact, low-power forms particularly for mobile equipments. These performance attributes can be achieved over conventional digital hardware solutions when the c.c.d. is operated as a sampled-data, analogue signal processor. However, when the c.c.d. is used as an analogue device it has to be associated with other peripheral circuits such as sample-and-holds, buffers, etc. If the complete processor is to be realized in monolithic form then it is necessary to realize all of the circuitry in a common l.s.i. technology: fortunately combined c.c.d./m.o.s.t. integrated circuit processes permit circuit design in analogue and digital forms. Many developments have been reported[2, 3] in linear m.o.s. peripheral circuitry specifically for integration with c.c.d.s.

The suitability of c.c.d.s for compact signal processing based on the transversal filter concept is now established. Such filter realizations have usually been based on either resistively-weighted tapped delay lines[4] or using split-gate delay lines.[5] With both of these approaches the transversal filter impulse and frequency responses are fixed, owing to the inflexible nature of the weighting coefficient programme.

The concept of a transversal filter with electronically variable or 'programmable' impulse responses, however, has two main attractions:[6] Firstly, a flexible filter structure results which may be controlled remotely; when associated with a microprocessor the programmable transversal filter (p.t.f.) may be considered as a powerful peripheral function which can perform signal computation in hardware form at high speed due to its parallel nature. Secondly, when combined with a permanent reference memory (r.o.m.) which is user programmable this filtering sub-system forms an economical, versatile alternative to split-gate devices for dedicated, stand-alone applications, especially when a small number of different filters are required. This enables, for example, a filter manufacturer to realize a wide repertoire of filtering functions to individual customer specifications at final test, from a stock of p.t.f. devices.

Many different programmable filter architectures have been proposed although insufficient results are yet available for direct performance comparison. Generally two filter types may be identified; those which employ analogue weighting coefficients and multipliers, and those which combine digital weighting coefficients with some form of multiplying, digital/analogue converter at every filter point. Even within these classes different realizations are possible depending upon the form of signal and reference storage and the manner in which the necessary time shift operation is achieved. Practical attempts at realizing programmable filters—certainly in

0033–7722/80/050213+13 $1.50/0

digital form—have generally resulted in large, high-power, low-bandwidth processors with their performance usually restricted by the speed and power of the multipliers. Until recently, however, the complexity of integrated filter structures has limited practical realizations, essentially because of the problem of miniaturizing the multipliers. Currently, several viable approaches have been demonstrated in monolithic form; an analogue coefficient filter[7] and a digital coefficient filter[8] have been reported. Other devices have been described[9,10] which incorporate a single-bit, digital reference memory and multipliers. These devices may be paralleled and their outputs suitably weighted, to achieve equivalent multi-bit references.[11]

In this paper we present the latest results and applications potential of an analogue coefficient 64-point p.t.f.[6,12,13] It has performance parameters that make it particularly suitable for sonar, instrumentation, biomedical and communications applications. When cascaded or multiplexed the performance of the basic filter may be extended, and it may be remotely controlled from a digital source. We will present here results for the p.t.f. when used for matched- and frequency-filtering, and also demonstrate a prototype adaptive filter configured with a p.t.f. and additional circuits. Consideration is given to a single chip adaptive filter which has potential application in several signal processing situations. Finally, the future of this analogue p.t.f. approach is discussed with reference to a 256-point filter currently in development.

2 P.T.F. Design

2.1 P.T.F. Architecture

The transversal filtering function may be realized in serial or parallel forms and operate on analogue or digital signals. In previous discussion[14] we have shown the advantages of a parallel configuration filter operating on all-analogue signals.

Specifically, the transversal filter, shown schematically in Fig. 1, is a general-purpose sampled-data, signal processing element useful for matched filtering operations and for realizing all-zero responses in the frequency domain. In operation, the filter is conceptually simple; input samples are successively delayed and multiplied by a set of weighting coefficients with all products summed within each time period to form the output samples. Mathematically, we may write:

$$c(n) = \sum_{m=1}^{N} s(n-m)r_m \qquad (1)$$

where s is the serial signal sequence and r is the filter impulse response sequence of N elements. This is often referred to as the convolution sum because it represents a convolution of the input sequence with the weighting coefficients which form the impulse response of the device. The chosen transversal filter architecture shown in Fig. 2 is both simple and compact, involving the minimum of signal manipulation and requiring a minimum of silicon area in integrated circuit form. It is a direct realization of the block diagram of Fig. 1, and has been implemented in this work using linear c.c.d. and m.o.s. component technology.

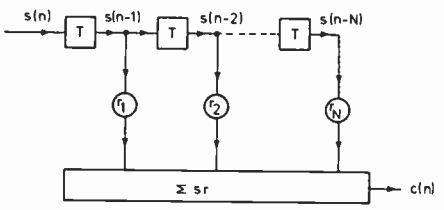


Fig. 1. A transversal filter.

In our design we have formed the signal register from a tapped, analogue, c.c.d. delay line[15] which simultaneously realizes the signal (eqn. (1)) storage and time-shift operations required in the convolution sum. C.c.d. realization of this element is optimum for these applications in that the analogue time-delay and shift processes are achieved inherently with the most economical use of silicon area. In a prototype design, we have employed a single, three-phase c.c.d. register with delay outputs (taps) implemented using the floating-gate, reset sensing technique.[13]

Because the c.c.d. signal register provides the necessary time-shift process, a stationary† analogue reference register is sufficient, supplying the weighting values to the multipliers in parallel form. An electrically simple and physically compact realization of this element uses discrete m.o.s. capacitors for analogue voltage memory. These feed the multiplier reference terminals, via buffer amplifiers, and reference values are updated individually via a single, digitally-multiplexed, analogue input bus.

Unfortunately, analogue coefficient storage is by nature volatile and thus may require some form of dedicated, external memory for refresh purposes, as shown in Fig. 2. Despite this digital memory requirement, the arrangement is still optimally compact for filters of more than (approximately) 32 points, because of the relative simplicity of the analogue multipliers compared with the multiplying d.a.c. structures implied with a single, digital reference memory architecture.

2.2 Multipliers

The disadvantage of parallel realizations of transversal filter over serial designs is that N multipliers are required for an N-point filter, whereas serial configurations employ only one. When the objective is a fully integrated p.t.f. then the requirement for many integrated multipliers, one at each filter point, is in itself a formidable problem.[16] The prototype p.t.f. described in

† Spatially 'stationary' relative to the signal information.

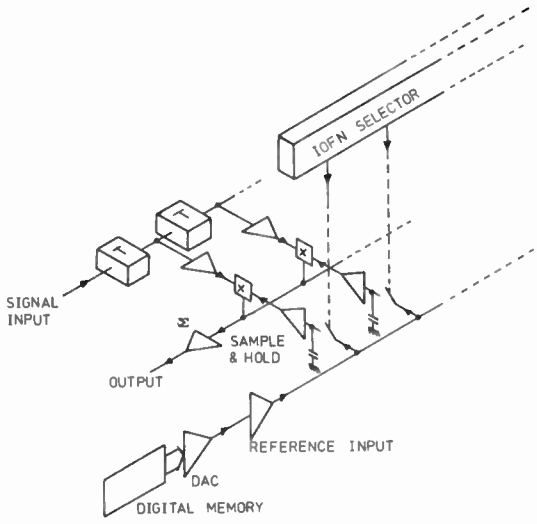


Fig. 2. The programmable transversal filter.

this paper relies on a novel analogue multiplier arrangement[6, 12] that has been developed in this work to provide accurate multiplication of the signal and reference (weighting coefficient) samples.

An economical multiplication technique based upon the essentially linear transconductance of an m.o.s. transistor operating in the 'triode' (pre-saturation of drain current) region is adopted here. A first-order expression for the drain current of such a transistor is given by:

$$I_{DS} = \beta_M \left[(V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right] \qquad (2)$$

where β_M is a process dependent gain constant and the other symbols have their usual meaning. Note that this expression does contain a potential linear multiplication term $V_{GS}V_{DS}$ but this must be isolated from the other undesirable terms.

For a constant drain-source voltage, V_{DS}, any change in gate voltage, ΔV_{GS}, stimulates a change in drain current given by:

$$\Delta I_{DS} = \beta_M V_{DS} \Delta V_{GS} \qquad (3)$$

for all sign combinations of V_{DS} and V_{GS}.† Clearly this is the desired product term. This change in drain current appears on a quiescent current given by:

$$I_0 = \beta_M \left[(V_{GS_0} - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right] \qquad (4)$$

where V_{GS_0} is the gate voltage corresponding to 'zero' signal. All terms on the right-hand side of equation (4) are constant for a given reference coefficient, thus I_0 remains constant. The following multiplication techniques rely on cancelling I_0 to leave the pure product term given by equation (3) as output.

† If the drain and source potentials remain constant so does the threshold voltage; thus equation (3) remains true for more complex drain current expressions than equation (2) which involves threshold bias dependence.

Previous realizations of multiplier have been synthesized using two identical transistors with a common diffused terminal, which have suffered from poor accuracy and dynamic range in addition to long-term stability and drift problems. The technique has relied on the two transistor gains, β_M, being matched so that I_0 may be cancelled adequately. For this reason, in our prototype p.t.f. a multiplication arrangement has been developed around a single m.o.s. transistor which gives much improved performance over contemporary realizations. In relation to the filter circuit, we may multiply the signal and reference samples at each filter point using the m.o.s. transistor as follows. The reference sample, which remains constant, is applied to the transistor drain, whilst the source is held at a quiescent level equivalent to the reference-zero, r_0. The signal sample is applied to the transistor gate, and stimulates the required change in drain current, from equation (3). Currents from all the multiplier transistors may be summed on a common source busbar Σ, shown in Fig. 2, which is held at the required reference potential by the output summing amplifier (off-chip).

The unwanted quiescent component, I_0, may be cancelled from the output by alternatively switching zero-samples (V_{GS_0}) onto all of the multiplier gates. I_0, which contains all of the unwanted terms, can then be detected and cancelled at the output. As signal samples are subsequently switched onto the multiplier gates, via the c.c.d. taps, the resultant output corresponds to the change in drain current given by equation (3), which is the sum-of-products term required; the quiescent current level, I_0 in equation (4), remains cancelled. Provided that the reference voltages applied to the drains of the multiplying transistors remain constant during this switching operation, all of the unwanted linear and quadratic terms are exactly cancelled.

An advantage of this time-multiplexed multiplication scheme is that the desired and the unwanted output components are generated by the same m.o.s. transistors, thus obviating the mis-matching errors associated with the dual-transistor balanced multiplier. Also variations in the transistor threshold affect only I_0 via equation (4) which is cancelled, thus they also contribute no error term.

2.3 P.T.F. Implementation

A 64-point, prototype p.t.f. has been designed based on the principles discussed in Sections 2.1 and 2.2 and is illustrated in Fig. 3. Each filter point required 14 m.o.s. transistors and 2 c.c.d. bits of storage (6 gates for a 3-phase c.c.d. process). The m.o.s. peripheral circuitry required a 56 μm cell pitch when laid out in integrated circuit form and permitted two c.c.d. shift bits in this length (the tapped c.c.d. was thus 128 bits long). The device was fabricated in an n-channel metal-gate process[17] having a chip size of 4·5 mm × 3·3 mm,

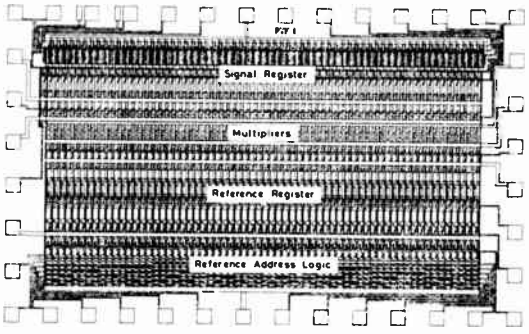


Fig. 3. The prototype 64-point p.t.f. device.

demonstrating the high packing density achievable with this design technique.

The summing amplifier was not integrated with the p.t.f. for three main reasons; although we have had experience[2] in integrating m.o.s. linear circuits to meet this requirement. Firstly, only a simple summing amplifier is sufficient because the single multiplier m.o.s.t. array requires only a single summing busbar (Σ). We have used a single, off-chip bipolar transistor stage in our experimental work which can easily be mounted near to each p.t.f. on a printed-circuit board. (Previous realizations of multiplier incorporating two matched m.o.s.t.s have required cancellation circuitry to remove undesired d.c. offsets by using three differential amplifier stages.) Secondly, we had a design priority for optimizing the number of filter points at the expense of any strictly unnecessary peripheral circuitry. Thirdly, there was a strong requirement to reduce on-chip power dissipation to minimize dark current generation, which is an exponential function of temperature, caused by chip power dissipation. Power dissipation in the sense and buffer amplifiers, and in the address logic, was set at 300 mW in the prototype 64-point p.t.f. (that is a value of approximately 5 mW per filter point).

2.4 P.T.F. Error Analysis

The prototype filter of Fig. 2 is a complex analogue structure, consequently, its error analysis is extremely involved. Here we will only identify the main imperfections in the p.t.f. and discuss their combined effect on the overall filtering function. In addition, in Section 3, where we present results for the prototype p.t.f., some discussion on the overall effect of device error will be given. However, unless precise information about both the signal and reference waveforms is available then a quantitative error analysis for a particular device is not feasible.

2.4.1 Error sources

Charge transfer inefficiency (ε) in the c.c.d. register is known to degrade signal information at high frequencies, this being a cumulative effect with increasing $N\varepsilon$ products. In terms of frequency filtering applications, however, the net effect at low frequencies is a shift in the transition edges between pass- and stop-bands of a factor $(1-\varepsilon)$, regardless of the number of filter stages.[18] The effect upon p.t.f. performance is somewhat more signal dependent but charge-transfer inefficiency is known to cause a slump in the correlation peak and a corresponding increase in sidelobe significance. It has been suggested[5] that in certain correlation applications an $N\varepsilon$ product of 2 is tolerable.

Random gain errors in the multiplication process are analogous to tap weight errors in split-gate filters and as such may be expected to impair stop-band suppression in frequency filtering applications. Correlation applications are more tolerant to these errors which become attenuated by the 'processing gain' ($N^{\frac{1}{2}}$) of the filter. The requirements placed on filter accuracy are dealt with in more detail in Section 2.4.2.

Quiescent offset errors in both the signal and reference channels appear as such at the convolver output for suitably large time-bandwidth products. Taking for example quiescent reference errors, Δr, the net output error will be $\bar{s} \cdot \overline{\Delta r}$, where the bar denotes an average value. Clearly, such errors may be reduced or eliminated where either sequence can be chosen to have zero mean value.

The finite output admittances of the buffer stages which drive the multipliers result in a signal distortion term which unfortunately precludes a general linear analysis of the resultant error. However, in correlator applications additional correlation peaks are to be expected whenever a match is detected between the generated harmonic signal component and the reference waveform.

The refreshing and subsequent decay of reference values results in a modulation of the output waveform. As this process is not generally synchronized with the incoming data sequences it appears as a form of noise at the output. Naturally, the magnitude of this noise is directly related to the decay rate and update frequency of the reference samples; these factors determine the maximum number of reference points which may be updated sequentially. Thereafter reference refreshing must take place in parallel blocks. In our experiments with the prototype filter we have found this refresh noise to be insignificant.

Although several potential error sources have been identified here it is clear that not all of them will be applicable to any one application of the device. In general, frequency filtering applications requiring large stop-band attenuations are more sensitive to these errors. Matched filtering (correlation) applications, however, are considerably more tolerant to random errors which become attenuated by the 'processing gain' of the filter; generally, the only significant errors are those which correlate with either the signal or reference sequences.

For frequency filtering applications accuracy of tap weight setting is crucial. Normally, the maximum stop-band attenuation that can be achieved in an ideal filter increases with the number of stages used. However, if tap weight inaccuracies are present, an average normalized output error, E, results which also increases with N and is given by:

$$E^2 = \frac{T^2}{3} \cdot \frac{N}{(\Sigma r)^2} \qquad (5)$$

where T is the tolerance on the maximum tap weight. This average, or expected, output error will mask the ideal attenuation in the stop-bands and in practice imposes a limit on realizable stop-band levels: greater attenuation can then only be achieved by cascading filter sections. In Section 3.1 we will discuss the effective tap weight tolerance for our p.t.f. by estimating the stop-band attenuation of a low-pass filter.

2.4.2 Error correction

The inherent inaccuracies in the individual multiplier elements can be compensated for by employing an external control loop. With this technique, which we have called iterative reference adaption, the actual response of the filter can be compared to its desired response and then corrected for using an external store. Programmable transversal filters employing a static reference are here at a great advantage in that each multiplier is associated uniquely with one reference coefficient and any weighting errors in the multiplier may be corrected by adjusting that coefficient.

The means for generating and storing the error 'signature' for an individual p.t.f. can be performed in several ways. Basically the impulse response could be adjusted manually by comparing the experimentally observed impulse response with a theoretical value and then compensating the reference input to allow for errors in a p.t.f. In Fig. 5 we have in fact applied this procedure to obtain the result (5(b)). However, in practice it would be desirable to compensate automatically for such errors in the impulse response. When applied continuously, this would have the added advantage of tracking device errors with ageing and temperature variation.

Two forms of automatic error compensation have been applied in our work: microprocessor control and

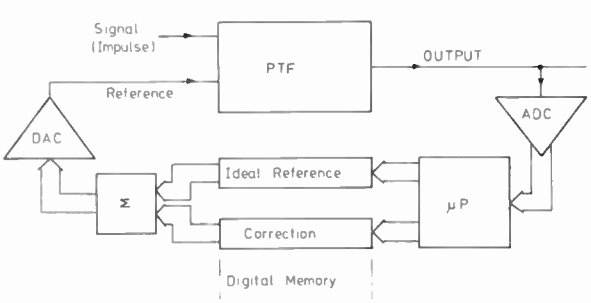


Fig. 4. A microprocessor controlled reference adaption loop.

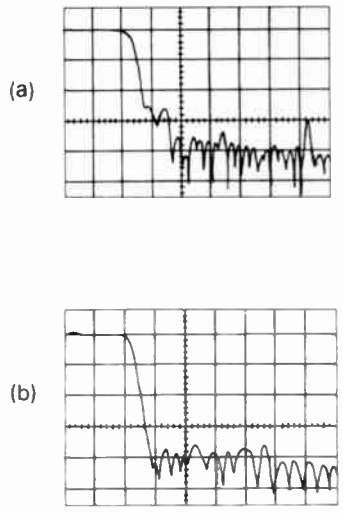


Fig. 5. (a) Frequency response of the filter when it is set up as a low-pass filter with no error compensation. (b) Frequency response when filter errors are compensated.

adaptive control using the convergence properties of the least-mean-square (l.m.s.) algorithm.[19] Both techniques employ feedback compensation; however, the adaptive filter approach does not in fact require a microprocessor. When a microprocessor is available with the p.t.f., then its software programmability makes the p.t.f.-based system very powerful and error compensation can be regarded as an inherent bonus.

A block diagram of the reference adaptation loop using microprocessor control is shown in Fig. 4. The analogue-digital converter (a.d.c.) converts the impulse response to the required digital form and, after taking a suitable number of averages, the microprocessor compares this with the desired value and then applies the necessary correction. The programmable filter is now updated with the ideal reference coefficients plus corrections, via the d.a.c., and the process is iterated as required. We have calculated[6] that to improve the impulse response of our 64-point p.t.f. to the equivalent of 8-bits digital accuracy with an initial dynamic range of 34 dB requires an average to be taken of 400 responses. This requires approximately 250 ms at a 100 kHz sample rate. The number of iterations required depends upon the form of error to be corrected but four interations appear significant in simulations involving gain errors up to 10% and second harmonic distortion to 3%. Thus, using this system, we may expect to achieve tap weight accuracies of 8 bits over an adaptation period of one second.

The alternative technique for improving the p.t.f. accuracy using the l.m.s. algorithm is discussed with reference to results in Section 3.3.

2.5 P.T.F. Performance Summary

Characterization of a general signal processing sub-system such as a p.t.f. integrated circuit is extremely complex and is dependent upon an appropriate choice of

Table 1
Performance summary of c.c.d./m.o.s. monolithic transversal filter

Parameter	Prototype result	Future device
Filter points per chip	64	32/64 (frequency filter) 256 (matched filter)
Total potential cascadable points (matched filtering)	1000	2000
Power consumption	300 mW/chip (5 mW/point)	1 mW/point
Dynamic range per filter point D_{fp}	52 dB	up to 60 dB
Weighting accuracy with reference adaptation	2% better than 1%	2% better than 1%
Harmonic (signal) distortion	−34 dB	−40 dB
Storage time (reference)	4 s	up to 10 s (dependent on process)
Signal bandwidth	50 kHz	>1 MHz
Chip area	4·5 × 3·3 mm (0·055 × 3 mm/point)	(0·0275 × 3 mm/point)

signal and reference test signals. In Table 1 we present a performance summary of results for the prototype 64-point p.t.f., and results for future device design.

The present 64-point device can process 50 kHz signals within a total chip power dissipation of 300 mW. The reference signal can be loaded into the p.t.f. at this rate and stored dynamically for up to about 4 s (thus complete reference refresh must occur on this timescale). The equivalent weighting accuracy of the basic p.t.f. is 2%, but with reference adaption as described in Section 2.4.2. this figure may be improved to better than 1%.

The measured dynamic range of the impulse response for the 64-point filter is 34 dB. Clearly though, this parameter is a direct function of the number of filter points used, as each additional point contributes noise (originating in the quiescent current of each multiplier transistor). A means of making a direct comparison between filters of any number of points is to refer the dynamic range to a filter containing a single point; thus the dynamic range per filter point, D_{fp}, is given by:

$$D_{fp} = D_{ir} + 10 \log_{10} N \qquad (6)$$

where D_{ir} is the measured dynamic range of the impulse response and N is the total number of filter points. For this device then, $D_{fp} = 52$ dB. Over this dynamic range the harmonic distortion in the signal register is better than −34 dB.

Further, for a given signal processing application, the dynamic range of the filter is a function of the processing gain for that application. Referred to peak output amplitudes with an impulse response at 0 dB, the processing gain for unweighted sine-wave autocorrelation or linear f.m. matched filtering is $20 \log_{10} N - 6{\cdot}02$ dB, and the maximum gain is $20 \log_{10} N$ dB for the autocorrelation of squarewaves. For linear f.m. matched filtering applications, the dynamic range of this device may therefore be quoted as 64 dB.

3 P.T.F. Applications

The applications potential of the prototype p.t.f. described in Section 2 is very extensive and already we have demonstrated a number of key filtering functions in compact, low-power form. The flexible nature of the p.t.f. makes it invaluable in many signal processing applications, especially for mobile equipment. We will discuss here three significant applications in which we have put it to use; although of course many others exist and are currently under investigation.

Before we turn our attention to the programmable nature of the p.t.f. it is worth discussing its use as an individual, accurately weighted transversal filter, equivalent in many aspects to a split-gate device.[5] This can be achieved by associating one or more p.t.f. chips with the 'overhead' of a dedicated digital reference store (p.r.o.m.) and d.a.c. circuit. Having adapted the filter response as required, for example using the microprocessor compensation scheme of Section 2.4.2., the contents of the digital reference and correction memories may be combined and transferred to p.r.o.m. The reference channel of the filter may then be driven directly from the p.r.o.m. plus d.a.c. to form a filter module with one or more dedicated impulse responses dependent upon the size of the memory. Clearly the processor and adaptive loop form a single laboratory facility for servicing programmable-filter-plus-p.r.o.m. modules, which might take circuit-board, hybrid, or fully integrated form. No additional processing is required and turn-around can be fast and economical for users requiring a dedicated filter response to a specified weighting accuracy. The overhead of the p.r.o.m. may be insignificant in complex signal processing situations, e.g.

a chirp-Z transform realization, where considerable memory would normally be associated with the transversal filters to perform the overall filtering function.

3.1 Frequency Filtering

By adjusting the impulse response of a p.t.f. appropriately, a frequency filter of desired response can be achieved within certain accuracy limits. For example, it is possible to design an ideal low-pass transversal filter having defined stop-band characteristics using windowing techniques or 'optimal' filter design programs.[20] However, we have seen in Section 2.4.2 that tap weight inaccuracies in practical filters cause an average output error that will mask the ideal attenuation in the stop-bands and in practice impose a limit on realizable stop-band levels. Greater attenuation can then only be achieved by cascading filter sections.[21] The effective tap weight tolerance may thus be estimated by measuring the performance of a realized low-pass filter having a large ideal stop-band attenuation.

The frequency response of such a realization, using the prototype 64-point device, is shown in Fig. 5(a); the theoretical stop-band level in this example is approximately -50 dB. From this figure the average stop-band attenuation achieved is about 34 dB and has a minimum value of 26 dB. From equation (5) we calculate the effective tap weight tolerance to be 2% and attribute this to individual multiplier gain errors caused by oxide thickness variation, and to random offset errors in the reference coefficient buffers caused by threshold variation. By visually adjusting the reference coefficients—displayed on an oscilloscope—to correct for these multiplier errors, as suggested in Section 2.4.2, an improvement in frequency response, given in Fig. 5(b), was achieved, which demonstrates an average stop-band attenuation of 40 dB (minimum value 34 dB), with a corresponding tap weight tolerance of 1%.

The advantage over split-gate filters is, of course, that filters of an experimental nature, or for applications for which no hard and fast design rules exist, can be individually specified and produced. Once a filter design is established, there is no severe obstacle to reasonable production volumes of individually optimized filters.

3.2 Matched Filtering

A prime application of the programmable device is as a matched filter, or correlation detector. Essentially, the impulse response of the filter is chosen to be the time-inverse of the waveform to be detected. Suitable waveforms may be chosen such that the signal energy is time compressed into a single output peak. Waveforms commonly employed in sonar and radar equipments are called 'chirps' (linearly modulated f.m. waveforms) and are used to maximize transmitted signal energy whilst retaining range definition at the detector output. Chirp waveforms generate a correlation peak of the sinc x form having a compression factor which is directly proportional to the time-bandwidth product of the chirp and, for a filter of N points, the maximum TB product that may be achieved is:

$$TB_{max} = \frac{N}{2} \tag{7}$$

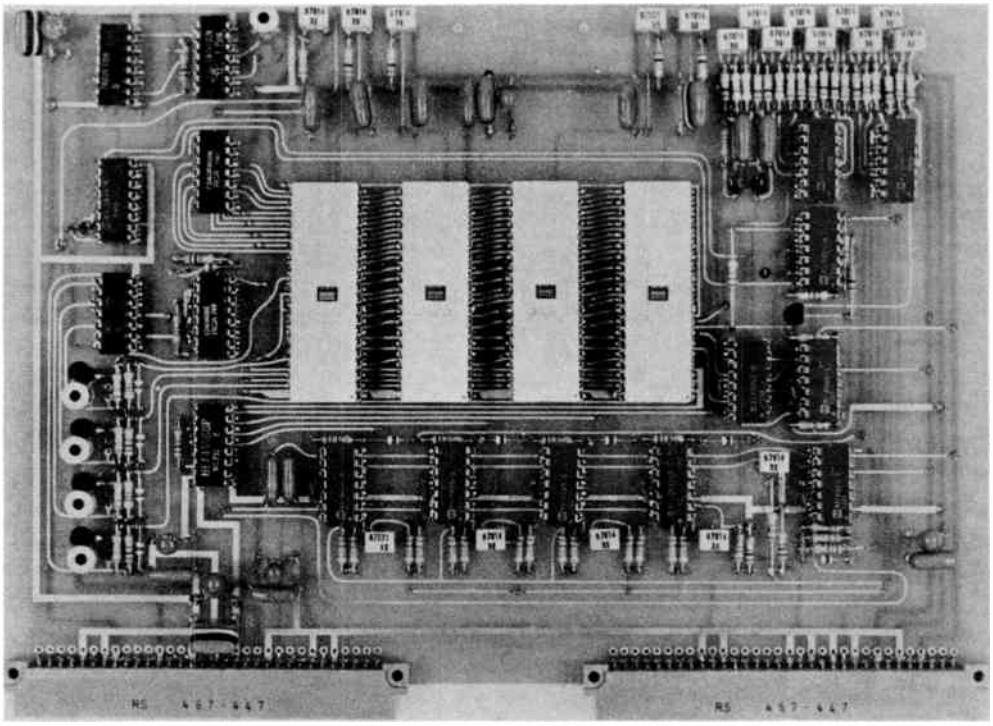


Fig. 6. A 256-point correlator system constructed by using four cascaded 64-point p.t.f. devices.

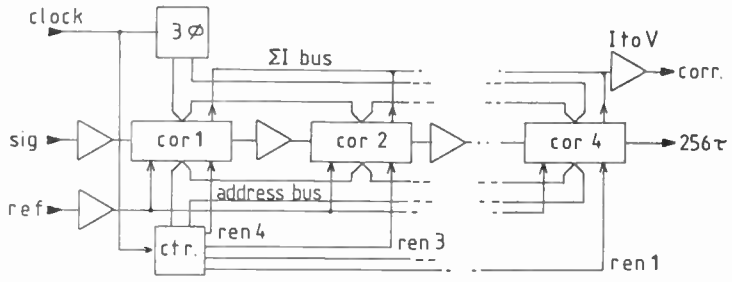


Fig. 7. The 256-point correlator system.

Clearly high *TB* figures are desirable and may be achieved by cascading p.t.f.s to increase the number of filter points.

To demonstrate the potential for the application of our c.c.d. p.t.f. to high time-bandwidth product, low-power, sonar matched-filter signal processing systems, a printed-circuit board containing four cascaded correlator i.c.s and virtually all the ancillary circuitry required to generate a complete 256-point correlator sub-system, has been produced and is shown in Fig. 6. The system configuration is given in Fig. 7. A single interstage sample-hold amplifier is required to cascade the c.c.d. delay line stages, and the reference is cascaded just as easily through a common address bus and individual r.en. (write enable) strobes. All clocking and timing circuity is provided so that the correlator samples in synchronism with an externally applied TTL clock. Externally, the correlator sub-system appears as a simple three-port element, with signal reference as inputs and correlation as output.

The complete p.t.f. sub-system is mounted on a double-sized Eurocard printed-circuit board and dissipates about 3 W, of which 1·2 W is contributed by the p.t.f. chips. Signals from d.c. to above the audio range are correlated very adequately over the full bandwidth available at a given sampling rate, giving a very versatile system. In particular f.m. matched filtering suitable to sonar signal processing may be readily accomplished. This is demonstrated in Fig. 8 which shows the performance of such a matched filter utilizing a linear f.m. sweep of maximum bandwidth, i.e. d.c. to the Nyquist limit. Performance is very acceptable considering the greater size, power consumption and cost of comparable digital processors.

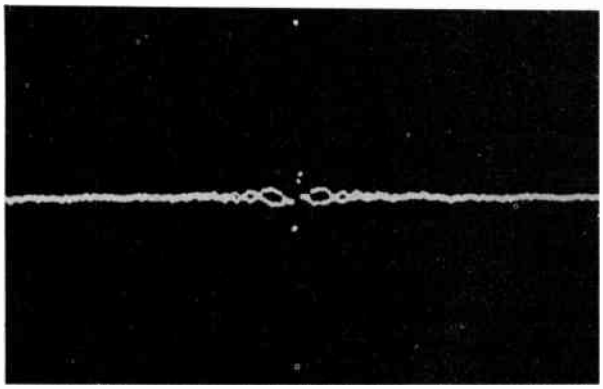


Fig. 8. Output of the 256-point correlator system configured as a matched filter to a linear frequency modulated signal swept from d.c. to the Nyquist frequency.

3.2.1 Sonar application

Two of the above 256-point correlator sub-systems have been configured into a low-i.f., 2-channel quadrature sonar processing system, Fig. 9(a), and this has been demonstrated against a submarine target in very shallow water (approximately 10 m depth), as illustrated in Fig. 9(b). The processor output shows a high resolution return in which the target is clearly visible against a background of strong reverberation.

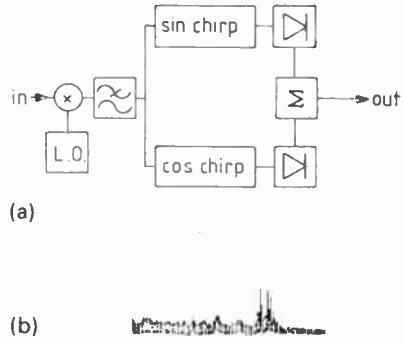


Fig. 9. (a) Low-i.f., two-channel quadrature sonar processing system using two 256-point correlator systems. (b) High-resolution sonar return from a submarine target.

3.3 Adaptive Filtering

A useful device for many applications in signal processing would be a filter needing the minimum *a priori* information about an incoming signal immersed in noise to detect and reproduce it. The adaptive filter is probably the optimum form for this type of application and it has been extensively studied in theory and modelled on computers. However, little is known about the actual physical implementation of this class of filter. The purpose of our work here is to demonstrate the feasibility of employing a p.t.f. as a central element with ancillary circuitry to illustrate basic adaptive filter operation. The eventual goal of this approach is a fully-

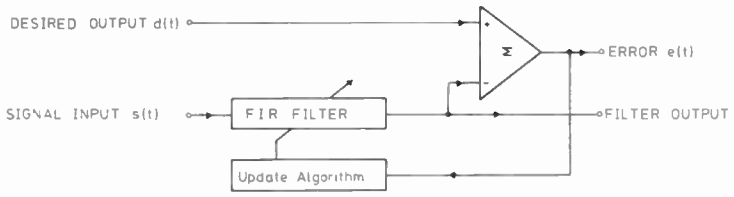


Fig. 10. Basic block diagram of an adaptive filter.

integrated adaptive filtering system for use in a wide range of applications.

3.3.1. Implementation of l.m.s. algorithm

A block diagram of the basic adaptive filter element is shown in Fig. 10. The system is supplied with two inputs; the input signal $s(t)$ and the desired filter response to this input, $d(t)$. The aim is to force the output of the filter to resemble, as closely as possible, the desired response $d(t)$. To achieve this the filter weight vector **H** is updated continuously using the following algorithm:[19]

$$h_k(t+1) = h_k(t) + 2\mu\varepsilon(t)s(t-k) \qquad (8)$$

where $\varepsilon(t) = d(t) - r(t)$ is the output error, $r(t)$ is the actual filter output and μ is a selected convergence factor. The subscript k refers to the weight or tap position in the filter, the terms in brackets are the time or delay indices, and t is in normalized units of one delay time. This is known as the Widrow least-mean-square adaption algorithm, and a detailed derivation and discussion may be found in Reference 19.

Figure 11 shows a block diagram of the system used to implement this algorithm. The circuit function within the dotted box was achieved simply by utilizing a prototype 64-point p.t.f. described in Section 2.

Since with this filter it is impossible to update all 64 weights in the time between successive output data points, the output error $\varepsilon(t)$ is sampled at only one point in 65 and the weights are updated sequentially. In addition, it is not possible to read individual weight values once they have been stored in the filter holding register. It was therefore necessary to provide a separate memory (external to the p.t.f. reference analogue store), which could be both read from, and written to, at any time to provide storage of weight values during the weight-vector updating cycle.

In the circuit implementation shown in Fig. 11 the weight store and updating loop was configured using digital circuitry with appropriate analogue-digital conversion (a.d.c.) after the convergence multiplier, and a digital-analogue converter (d.a.c.) at the input to the p.t.f. analogue store. An alternative implementation involved off-chip analogue storage based on 64 sample-and-hold stages as a precursor to a fully-integrated version based on all-analogue circuitry. This latter arrangement gave the adaptive filter an inferior performance over the digital update version essentially because it suffered from weight value decay. For certain filtering functions, in particular inverse filtering, permanent storage of the reference is required and thus digital memory would be mandatory. However, in monolithic adaptive filter designs, now at the planning stage, the all-analogue approach will be pursued further because of the improved characteristics of on-chip analogue storage (by a factor of 100 to $\sim$1 V/s).

3.3.2 Adaptive filter results

We report here results for the prototype adaptive filter based on Fig. 11 using (a) an external digital update memory[22] (as shown), and (b) an analogue memory[23] (by omitting the a.d.c. and d.a.c. and replacing the r.a.m. with an analogue unit). In Fig. 12 we show a photograph of the completed filter with an analogue update store which has been assembled on one double Eurocard-sized printed-circuit board and consumes less than 10 W of power. An additional board containing the a.d.c., d.a.c.

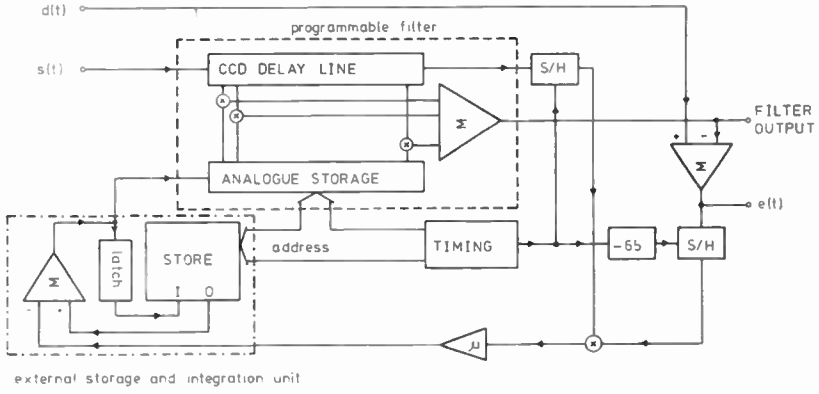


Fig. 11. Adaptive filter system using one 64-point p.t.f.

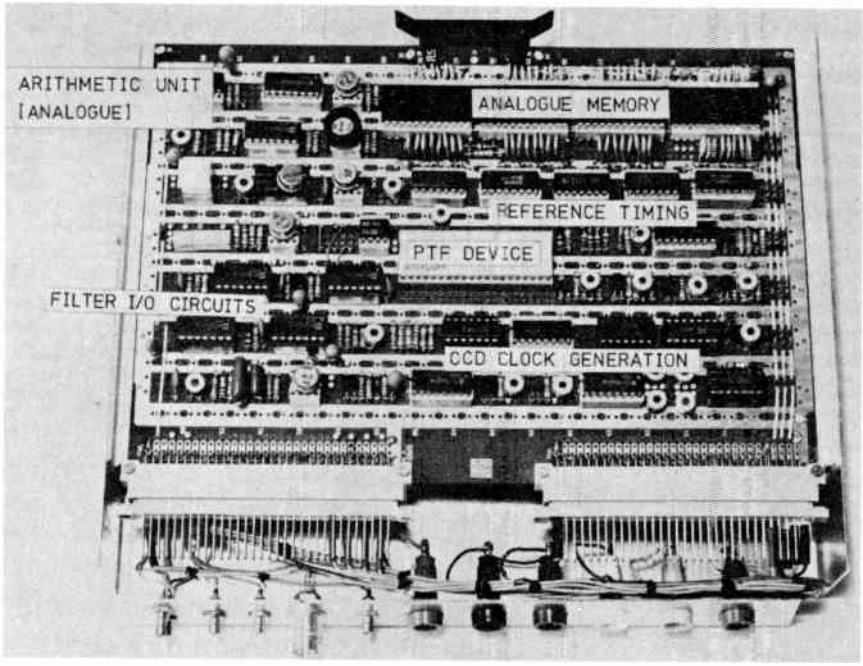


Fig. 12. The analogue adaptive filter module.

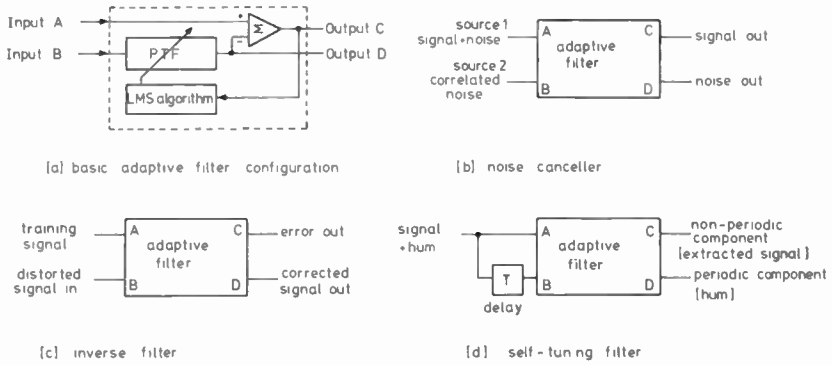
and memory is required for the analogue/digital hybrid version.

The adaptive filter unit may be used in a number of ways dependent upon the configuration of the two input ports and which of the two output ports is used. Figure 13 shows the three main filter configurations which have been used with our prototype system. Figure 13(b) shows the necessary input conditions for a noise canceller where the A input is a signal with additive noise and the B input is a correlated (but not identical) version of the noise. The filter converges to reconstitute the interfering noise at the filter output therefore subtracting coherently from A and leaving the uncorrupted signal at the error or canceller output C. One example of this type of operation would be where input A is supplied from a microphone which picks up a speech signal and some arbitrary background noise. Input B would be supplied from a second microphone which picks up only the background noise in the room. Output C should then be the uncontaminated speech signal.

As an inverse filter it is desired that the system should reconstitute a signal which has been subjected to distortion by some intervening medium, such as a transmission line. Figure 13(c) shows the adaptive filter configured as such a system. The B input is the distorted signal input and the A input is a training signal which is the required filter output. The output used in this case is D which should converge to the same waveform as A, after an initial training period during which the input signal B is known. In this time the filter impulse response converges to the inverse of the transmission line impulse response. The adaption algorithm is then switched off, freezing the weight vector, and the system is thereafter used as a straight forward transversal filter equalizer. Such a system has been described by Corl,[25] although his implementation uses a computer to update the weight vector using a zero-forcing algorithm.

The last configuration to be considered here is that of a self-tuning filter shown in Fig. 13(d); this structure is similar in operation to the noise canceller. The signal input will commonly comprise two components, (i) a narrow-band periodic component and (ii) a broadband non-periodic component. The delay T causes the non-periodic component to be decorrelated between A and B so that the filter converges to form a band-pass impulse response which produces the narrow-band periodic component of A at the filter output D (rejecting the broadband non-periodic component). Since this output is coherent in phase and amplitude with the periodic component of A, the C output then consists only of the broadband non-periodic component (periodic noise cancellation). The system effectively separates periodic and non-periodic signal components and may be used, for example, to cancel unwanted hum in speech signals and also as a self-tuning filter to reduce unwanted broadband noise on a required periodic signal. This structure has been described in detail by Widrow *et al.*[19]

As a significant and illustrative demonstration of the operation of our prototype adaptive filter with digital update, we here describe its performance when it is desired to cancel a very strong first harmonic interference on a wanted sinusoidal signal. This represents operation in the noise cancellation mode (c.f. Fig. 13(b)). Figure 14(a) shows the $s(t)$ input, i.e. the correlated interference; Fig. 14(b) is the conditioning signal input $d(t)$, in this case the input signal with interference; Fig. 14(c) is the



[a] basic adaptive filter configuration

[b] noise canceller

[c] inverse filter

[d] self-tuning filter

Fig. 13. Adaptive filter configurations in various application areas.

filter output; and Fig. 14(d) is the canceller output. The frequency amplitude spectrum of $d(t)$ is given in Fig. 14(e) with a corresponding spectrum for the error output in Fig. 14(f), which indicates cancellation of the first harmonic by 50 dB. The cancellation for the analogue update unit for the same test yields a value of about 25 dB: the poorer cancellation is attributed to the leaky nature of an analogue store produced with discrete components which results in excess error in the converged weight vector. A fully integrated analogue store should have improved characteristics and thus yield an improved cancellation result.

A general figure of merit for an adaptive filter is usually taken as the adaptivity,[24] χ, where:

$$\chi = 20 \log_{10}\left[\frac{\varepsilon(t)}{d(t)}\right] \qquad (9)$$

For the digital reference system described here χ was measured to be -50 dB.

Results for a simulated inverse filter and a self-tuning filter are now presented to demonstrate the versatility of an adaptive filter based on a p.t.f.

Figure 15 shows results obtained for a simulated inverse filter structure using the digital store version of the adaptive filter. The input signal $s(t)$ in this case was a linear f.m. signal swept from d.c. to the Nyquist frequency (shown in Fig. 15(a)). The training signal $d(t)$ was a single pulse having a width of one sample period, positioned at the end of the input signal sequence. The filter output after training is given in Fig. 15(b) showing good correspondence to the training signal. In this case the weight vector should match the time reverse of the input signal and the impulse response shown in Fig. 15(c) confirms this. A linear f.m. signal may be considered as a pulse with quadratic phase distortion, and is therefore a valid test signal to demonstrate inverse filter operation.

Tests carried out on the self-tuning filter showed that the system would tune accurately to a wide range of signals including monochromatic and wideband periodic signals. Figure 16 shows some typical results for the self-tuning filter where the input signal $s(t)$ (Fig. 16(a)) was a sinusoid contaminated by broadband noise. For the case shown, using a sinusoid at approximately 1 kHz, the broadband signal source was in fact a white noise generator which was limited in bandwidth between 1·5 kHz and 7 kHz. It is possible for the filter to converge when the noise spectrum is truly white from d.c. to Nyquist but this admits noise in the output in the region of the pass-band. Figure 16(b) shows the filter output D for this case and spectral analysis shows noise rejection to be between 25 and 30 dB after convergence. Figure 16(c) shows the cancellation output C where the sinusoidal component has been cancelled to a significant degree. The weight vector (shown in Fig. 16(d) is the expected matching sine-wave.

It should be noted that the weight vector shown in Fig. 16(d) is the signal which appears at the reference input of the p.t.f. and therefore contains all the gain and offset errors associated with the filter. Further, the impulse response cannot be perfectly predicted since the bandwidth of the input signals does not cover the entire system bandwidth; this means that a certain degree of latitude is available in the choice of weight values.

4 Conclusions

We have demonstrated in this paper how innovative c.c.d. and m.o.s. design techniques may be combined to realize an electronically-programmable transversal filter in compact form and having low power consumption. This filter configuration offers powerful signal processing capability and its programmable aspect allows flexible and fast control of the filter function, either remotely or internally, via a system-based microprocessor. The number of filter stages to be included on one chip must be a decision based upon the eventual application. For frequency filtering, relatively few filter points are required and must, in any case, be minimized to avoid dynamic range restrictions—possibly a 32- or 64-stage cascadable block is optimum. For matched filtering applications, however, many filter points may be necessary and it becomes desirable to realize as many points on one chip as possible.

In this paper, we have demonstrated the extremely

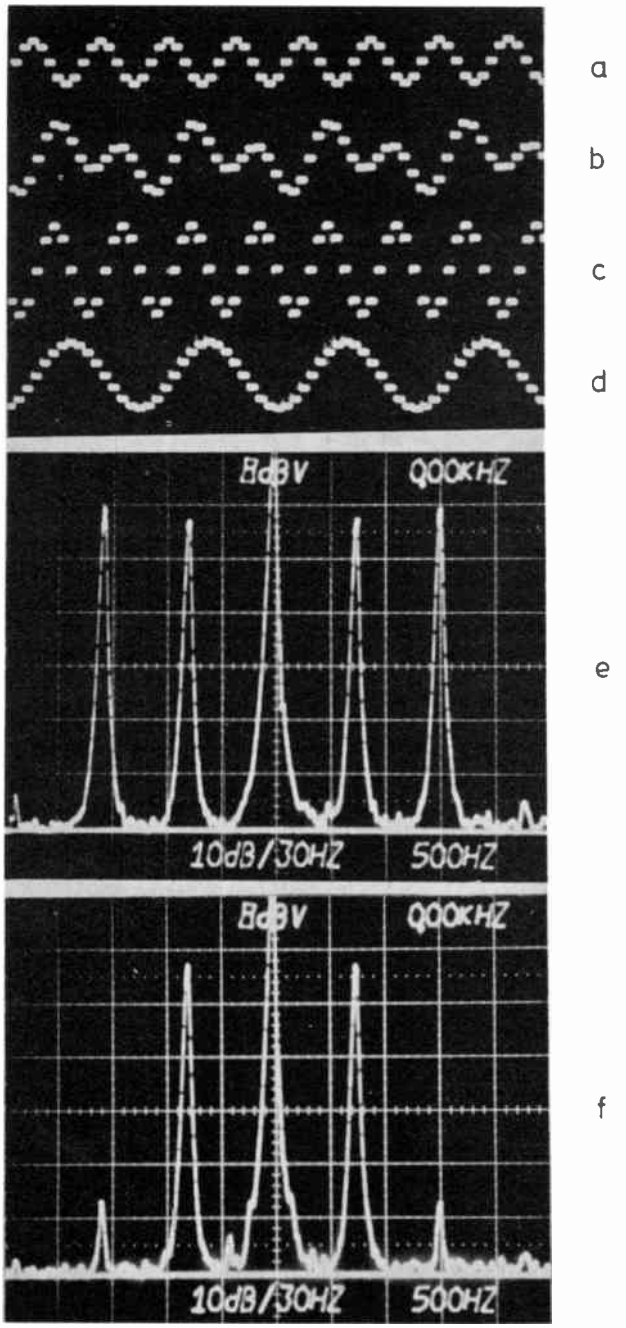


Fig. 14. Demonstration of the basic cancelling performance of the adaptive filter. (a) signal input $s(t)$. (b) target signal $d(t)$, desired sinusoid with added second harmonic distortion. (c) filter output, a phase and amplitude corrected version of the distortion. (d) canceller or error output. (e) spectrum of $d(t)$. (f) spectrum of canceller output.

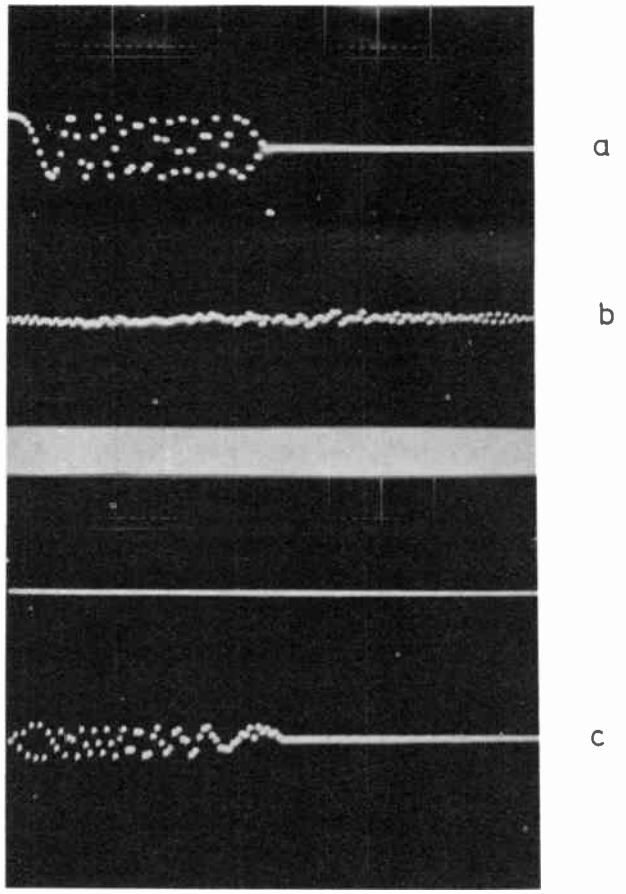


Fig. 15. Inverse filter operation. (a) input signal $s(t)$, dispersed pulse. (b) filter output, reconstituted pulse. (c) filter impulse response.

flexible performance of a prototype, 64-point electronically-programmable transversal filter for frequency- and matched-filtering applications. These results indicate that compact, low-power analogue c.c.d. p.t.f. sub-systems with TB products in excess of 1000 will be a reality in the near future, and will have an impact on such applications as mobile, high-resolution sonar signal processing. This is particularly evident in view of a 256-point p.t.f. chip already in design for high time-bandwidth applications. These filter structures are thus far only prototypes, but the 64-point device reported here and its contemporaries are proving that good performance parameters may be achieved in miniature form and that, with careful design, production devices must now be feasible.

We have also shown here that a 64-point adaptive filter may be realized using an analogue p.t.f. resulting in a compact unit of low-power dissipation (about 10 W for the complete system described). This type of processor could be used in a number of applications such as telephone line equalization, speech processing, cancellation of noise in transmitted speech, and a number of noise cancellation tasks in medical electronics. This is thought to be the first realistic demonstration of such a complex filtering system which could be integrated monolithically. Indeed, this work indicates that a dedicated adaptive filter with all necessary peripheral circuits, e.g. convergence multipliers and integrators, etc., could be produced in single chip form. However, such a dedicated design would not have the flexibility of a 'hybrid' unit as described here which could also be configured conveniently for other filtering functions such as spectrum analysis, matched filtering, etc. Also it would

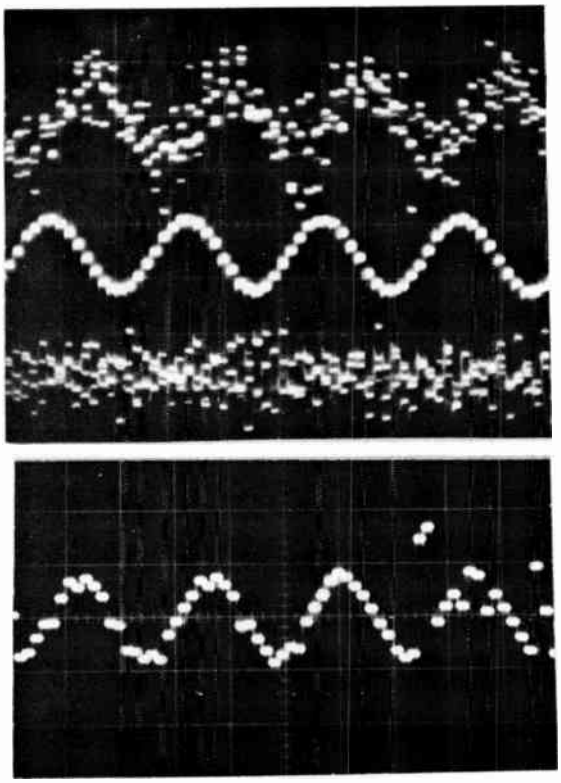


Fig. 16. Self-tuning filter performance. (a) sinusoid in bandlimited noise constituting the input signal. (b) filter output, noise rejection is approximately 25 dB. (c) error output showing cancellation of the periodic component. (d) converged weight vector.

not be able to achieve the same number of filter points because of the chip area needed for the peripheral circuitry.

5 Acknowledgements

This work has been carried out with the support of the Procurement Executive, Ministry of Defence: support from the Science Research Council is also acknowledged. Design and evaluation of the devices was undertaken in the Wolfson Microelectronic Institute, University of Edinburgh. Devices were processed by the Plessey Company. C. F. N. Cowan was supported by the Department of Education, Northern Ireland.

6 References

1 Boyle, W. S. and Smith, G. E., 'Charge-coupled semiconductor devices', *Bell Syst. Tech. J.*, **49**, pp. 587–93, April 1970.
2 Weste, N. and Mavor, J., 'M.o.s.t. amplifiers for performing peripheral integrated circuit functions', *IEE J. Electronic Circuits and Systems*, **1**, pp. 165–72, September 1977.
3 Hodges, D. A., Gray, P. R. and Brodersen, R. W., 'Potential of m.o.s. technologies for analog integrated circuits', *IEEE J. Solid-State Circuits*, **SC-13**, pp. 285–94, June 1978.
4 See, for example, RETICON TAD-32 Application Note, Reticon Corp., 910 Benicia Ave., Sunnyvale, California 94086.
5 Buss, D. D., Collins, D. R., Bailey, W. H. and Reeves, C. R., 'Transversal filtering using charge-transfer devices', *IEEE J. Solid-State Circuits*, **SC-8**, pp. 138–46, April 1973.
6 Denyer, P. B., Mavor, J. and Arthur, J. W., 'Miniature programmable transversal filter using c.c.d./m.o.s.t. technology', *Proc. IEEE*, **67**, pp. 42–50, January 1979.
7 Bosshart, P., 'An integrated analogue correlator using charge-coupled devices', Proc. of IEEE Int. Solid State Circuits Conf., 1976, pp. 198–9.
8 Haque, Y. A. and Copeland, M. A., 'Design and characterization of a real-time correlator', *IEEE J. Solid-State Circuits*, **SC-12**, pp. 642–9, December 1977.
9 Tiemann, J. J., Engeler, W. E. and Baertsch, R. D., 'A surface charge correlator', *IEEE J. Solid-State Circuits*, **SC-9**, pp. 403–9, December 1974.
10 Burke, B. E. and Lindley, W. T., 'New c.c.d. programmable transversal filter', *Electronics Letters*, **13**, pp. 521–3, September 1977.
11 Buss, D. D., Bailey, W. H. and Tasch, Jr., A. F., 'Signal processing applications of charge-coupled devices', Proceedings CCD 74, Edinburgh, pp. 179–97, September 1974.
12 Mavor, J., Arthur, J. W. and Denyer, P. B., 'Analogue c.c.d. correlator using monolithic m.o.s. multipliers', *Electronics Letters*, **13**, pp. 373–4, June 1977.
13 Denyer, P. B., Arthur, J. W. and Mavor, J., 'Monolithic programmable analogue c.c.d. transversal filter', *Electronics Letters*, **13**, pp. 373–4, November 1977.
14 Harp, J. G., Vanstone, G. F., MacLennan, D. J. and Mavor, J., 'Analogue correlators using charge coupled devices', Proceedings of CCD 75, pp. 229–35, September 1975.
15 Denyer, P. B. and Mavor, J., 'Design of c.c.d. delay lines with floating gate taps', *IEE J. Solid-State and Electron Devices*, **1**, pp. 121–9, July 1977.
16 Mavor, J. and Denyer, P. B., 'The design and development of c.c.d. programmable transversal filters and correlators', AGARD Conference Proceedings No. 230, pp. 2.8-1/15, October 1977.
17 Perkins, K. D. and Browne, V. A., 'Sub-micron gap metal gate technology for c.c.d.s', *Microelectronics J.*, **7**, pp. 14–22, January 1975.
18 Buss, D. D., Bailey, W. H., Holmes, J. D. and Hite, L. R., 'Charge transfer device transversal filters for communication systems', *Microelectronics J.*, **7**, pp. 46–53, January 1975.
19 Widrow, B. *et al.*, 'Adaptive noise cancelling: principles and applications', *Proc. IEEE*, **63**, pp. 1692–1716, December 1975.
20 Rabiner, L. R. and Gold B., 'Theory and application of digital signal processing' (Prentice-Hall, Englewood Cliffs, NJ, 1975).
21 Puckette, C. M., Butler, W. J. and Smith, D. A., 'Bucket-brigade transversal filters', *IEEE Trans. on Communications*, **22**, pp. 926–34, July 1974.
22 Cowan, C. F. N., Arthur, J. W. and Mavor, J., 'Noise cancellation and inverse filtering using a compact high-performance c.c.d. adaptive filter', *Electronics Letters*, **15**, pp. 35–7, January 1979.
23 Cowan, C. F. N., Mavor, J. and Arthur, J. W., 'Implementation of a 64-point adaptive filter using an analogue c.c.d. programmable filter', *Electronics Letters*, **14**, pp. 568–9, August 1978.
24 White, M. H., Mack, I. A., Borsuk, G. H., Lampe, D. R. and Kub, F., 'C.c.d. analogue adaptive signal processing', Proc. International Conference on 'The Technology and Applications of CCDs', San Diego, pp. 3A-1/14, October 1978.
25 Corl, D., 'A c.t.d. adaptive inverse filter', *Electronics Letters*, **14**, pp. 60–2, February 1978

Manuscript first received by the Institution on 25th June 1979 and in final form on 10th October 1979
(Paper No. 1931/CC 323)

UDC 621.382:535-1:621.383.8

Indexing Terms: Semiconductor devices, Integrated circuits, Charge-coupled devices, Pyroelectric detectors, Infra-red sensing

The applications of charge-coupled devices to infra-red image sensing systems

D. R. LAMB, Ph.D.*

and

N. A. FOSS, B.Sc., M.Sc.*

SUMMARY

The paper reviews the various ways in which c.c.d.s can be employed in i.r. sensing systems. These include: (i) monolithic structures fabricated using narrow band semiconductors such as HgCdTe or InSb, extrinsic silicon structures doped with deep-level impurities, and silicon Schottky barrier devices; (ii) hybrid structures in which the c.c.d. is used as the read-out mechanism from an array of, for example HgCdTe, PbTe, or pyroelectric detectors. The relative merits of these different approaches are compared and recent experiment results for many structures are quoted.

* *Honeywell Systems and Research Center, 2600 Ridgway Parkway, Minneapolis, Minn. 55413*

The Radio and Electronic Engineer, Vol. 50, No. 5, pp. 226–236, May 1980

1 Introduction

Visible image sensors detect radiation reflected from objects of the scene and deal with situations where there is a low background flux and high scene contrast. Infra-red image sensors detect radiation emitted by objects in the scene and thus, in general terrestrial applications, have to detect very small temperature differences, frequently of the order of 0·1 K, against a background of around 300 K. As a result of this the two forms of imaging have developed in different ways.

In a visible imaging system, such as the silicon c.c.d. sensor, the scene is focused onto a large two-dimensional detector array having one detector per picture element. Small differences in responsivity of adjacent detectors are not important because the scene contrast is high. Also the photo-generated charge due to the low background photon flux is easily accommodated within the dynamic range of the c.c.d. for integration times compatible with standard television frame rates.

In i.r. imagers the detectors often have to be cooled to very low temperatures necessitating encapsulation of the sensor in a Dewar which leads to problems with the read-out interconnections from the i.r. detectors forming a thermal leakage path. This problem increases as the number of detectors is increased. Additionally, because of the low scene contrast, fixed-pattern noise in the detector array is a serious problem, (see Sect. 3) and in many cases some form of background suppression must be used because of the high background photon flux. For these reasons i.r. imaging is currently based on linear, or small two-dimensional arrays of detectors rather than large two dimensional arrays. To generate the number of picture elements for a complete picture field the image is mechanically raster-scanned over the detector array by means of rotating mirrors.

The initial concept of applying c.c.d.s to i.r. imaging systems arose from their ability to receive, store and manipulate charge packets to produce a multiplexed output. Thus one can envisage the use of a c.c.d. on the focal plane which will scan and read out the information from the detector array. In such a system a hybrid structure with separate detector and c.c.d. chips is used, thus taking advantage of both existing detector technology and the silicon integrated circuit technology.

With more ambitious programmes the aim is to produce monolithic i.r. c.c.d. imaging chips in which the photodetectors and c.c.d.s are fabricated in the same semiconductor material. Here the choice lies between producing silicon photodetectors by a technique which is compatible with the existing c.c.d. technology, or fabricating c.c.d.s in the narrow-band-gap semiconductors already used for i.r. detectors.

Both of these approaches are in fact being investigated; the former technique is, however, potentially of greater importance because it allows the possible use of silicon l.s.i. technology to make large two-

0033–7722/80/050226+11 $1.50/0

Table 1.

Possible I.R. C.C.D. Technologies

MONOLITHIC		HYBRID
I.R.-sensitive c.c.d. structures	I.R. detector with integrated c.c.d. read-out	Independently fabricated detector and c.c.d. structures
Inversion mode Extrinsic silicon	Extrinsic silicon detectors Silicon Schottky barrier detectors	Narrow band gap semiconductor detectors
Accumulation mode Intrinsic narrow band gap semiconductors		Pyroelectric detectors

dimensional detector arrays.

Table 1 summarizes the major approaches that are being taken towards producing i.r. imaging systems employing c.c.d. concepts and we will now discuss each of these techniques in more detail.

2 Monolithic C.C.D. Structures

2.1 Intrinsic Narrow-band-gap Semiconductors

For these structures the basic concept is to produce a device in which the c.c.d. both forms the imaging sites and provides the read-out registers as in c.c.d. visible imaging devices. The most direct approach would thus be to fabricate a normal inversion mode surface channel c.c.d. in a narrow-band-gap semiconductor where the intrinsic optical absorption matches the wave-length region of interest. For the spectrum from $\lambda = 3{\cdot}0\,\mu m$ to $\lambda = 13{\cdot}0\,\mu m$ this requires the band gap to be in the range 0·1 eV to 0·4 eV. Semiconductors with such narrow band-gaps include InAs, InSb, HgCdTe and PbSnTe, with the latter two being examples of materials in which the band gap can be tailored to the wavelength of interest by altering the composition of the alloy. Unfortunately, fabricating high performance c.c.d.s in a material other than silicon is extremely difficult because of three basic performance limitations, namely: (1) low breakdown potential, (2) high dark currents, and (3) poor insulator-semiconductor interface characteristics.

The bulk breakdown voltage of a semiconductor is of course related to the magnitude of its energy band gap and the breakdown voltage decreases as the energy gap decreases. For detectors with energy gaps in the range of 0·1 eV to 0·4 eV this limits the effective gate voltage, V_G, that can be applied to m.i.s. structures to around 1–2 volts as compared to 10–20 V for silicon devices. Since the maximum charge handling capability of a surface channel c.c.d. is given by $Q = C_G V_G$ this seriously limits the compound semiconductor devices, unless an insulator with a high dielectric constant can be used to increase the gate capacitance, C_G.

The dark current in a c.c.d. limits the maximum storage or integration time and introduces a noise component. In a narrow-band-gap material the dark current results from thermal generation, which occurs in all semiconductors, and from direct band-to-band tunnel currents which can occur when the gate voltage causes significant band bending at the semiconductor–insulator interface. The magnitude of the thermal generation rate, whether it be taking place through the presence of bulk recombination-generation centres or surface states at the semiconductor–insulator interface, is proportional to the intrinsic carrier concentration n_i. Since n_i is proportional to $\exp(-E_g/KT)$ the thermal generation rate for the narrow-band-gap materials at a given temperature is much higher than, for example, that in silicon. Thus to achieve reasonable storage times and reduce the dark current noise, a c.c.d. fabricated in a narrow-band-gap material must be operated at relatively low temperatures.

In order to minimize the tunnel current contribution to the dark current the substrate doping level must be low and the surface potential, i.e. surface band bending, must be kept as small as possible. This latter condition again restricts the c.c.d. charge handling capability. For the 8–13 μm wavelength region with band gaps of the order of 0·1 eV the tunnel currents may be sufficiently large to make fabrication of c.c.d.s in these materials difficult.[1]

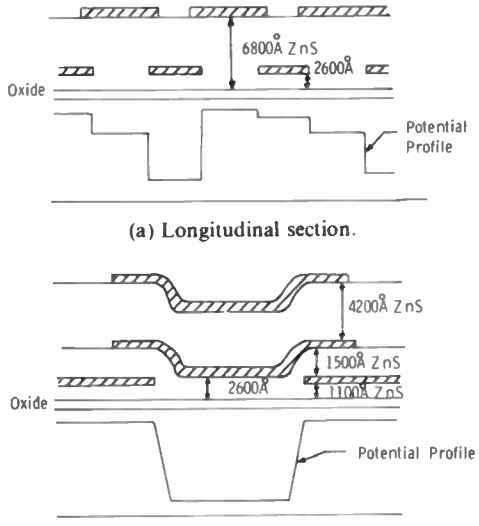


Fig. 1. One type of c.c.d. shift register geometry (After Ref. 3.)

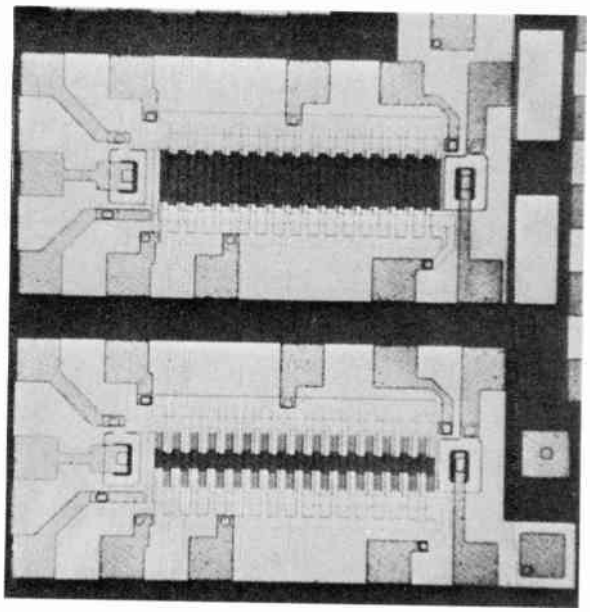


Fig. 2. Two 16-stage HgCdTe c.c.d. shift registers with transparent electrodes. (After Ref. 3.)

The third performance limitation on c.c.d.s fabricated in narrow-band-gap materials is the difficulty of obtaining a low surface state density at the semiconductor–insulator interface. As discussed above the surface state density influences the dark current in the device; it also limits the transfer efficiency that can be obtained and gives rise to trapping noise. Typical surface state densities measured with compound semiconductor substrates are $>10^{11}\ cm^{-2}\ eV^{-1}$ with the lowest value so far reported[2] as $\sim 4\times10^{10}\ cm^{-2}\ eV^{-1}$. This indicates a maximum transfer efficiency at 77 K of the order of 0·999. This could be improved by the introduction of a background charge; however this reduces the dynamic range of the device. A solution to the interface problem must be found before large arrays can be developed.

Imagers so far planned using compound semiconductors are designed to alleviate the low transfer efficiency by using the c.c.d.s to perform the sensing and time delay and integrate (t.d.i.) function only in small block-scanned systems. This type of approach has recently been reported by Chapman *et al.*[3] who fabricated 8, 16 and 32 bit c.c.d.s in HgCdTe utilizing a four-phase structure with a tunnel/avalanche breakdown input and a floating gate output. The structure of the device is depicted in Fig. 1, and Fig. 2 shows a photograph of two 16-stage devices. These c.c.d.s were operated at 77 K and at frequencies of up to 50 kHz. Transfer efficiencies ranged from 0·996 to 0·995 and t.d.i. operation was demonstrated as in Fig. 3.

Operation of an InSb linear c.c.d. imaging array has also been reported by Thom *et al.*[4] In this device the c.c.d. was used to read out the signals from a series of 20 InSb m.o.s. detector elements integrated on the same chip. In this case the c.c.d. was a $20\frac{1}{2}$-bit four-phase

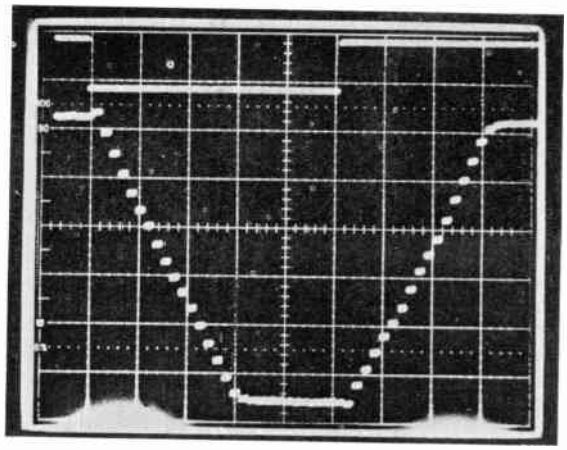


Fig. 3. Time-delay-and-integrate operation demonstrated for 16-stage c.c.d.-12AT3 with 50 kHz clocks using a GaAs near-i.r. diode emitter. (After Ref. 3.)

device with p^+-n junctions formed by beryllium ion implantation to form the input and output stages. Transfer inefficiencies ranged from 0·991 to 0·996 at clock frequencies of up to 200 kHz. Figure 4 is a photomicrograph of the device structure and Fig. 5 shows the read-out of the 20 detectors with i.r. signal input.

2.2 Extrinsic Silicon Structures

One way of making silicon sensitive to infra-red radiation is by introducing a deep level impurity to produce absorption in the appropriate wavelength region. The major potential advantage of this approach is that it offers the possibility of using silicon l.s.i. technology to produce large two-dimensional detector arrays with, say, 10^6 elements which would give adequate resolution without the need for mechanical scanning.

The major disadvantages of extrinsic silicon detectors are the need for very low operating temperatures and the difficulty of finding a suitable deep level impurity which will give adequate detectivity in the appropriate

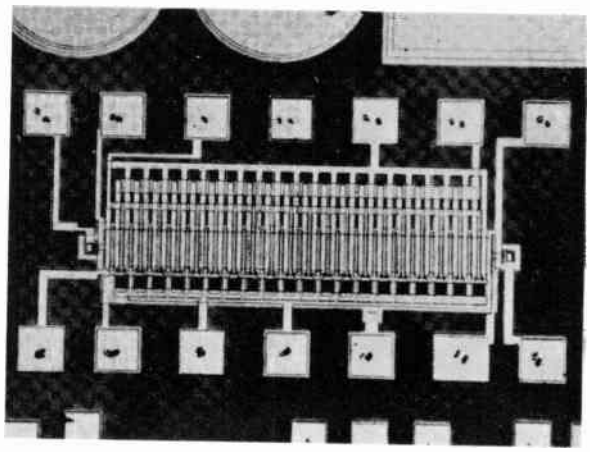


Fig. 4. 20-element InSb c.c.i.r.i.d. (After Ref. 4.)

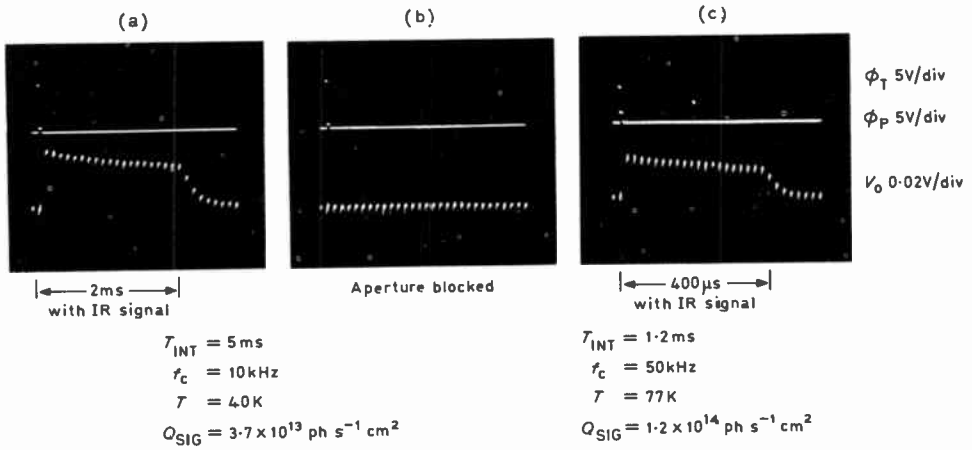


Fig. 5. 20-element InSb c.c.i.r.i.d. operation showing read-out of twenty detectors with i.r. signal input. (After Ref. 4.)

wavelength regions.

The signal-to-noise ratio for i.r. detectors is characterized by the detectivity D^* which, for a p type photoconductor, is given by

$$D^* = \frac{\eta\lambda}{2hc}\sqrt{\frac{\tau}{(P_{th}+P_{op})l}}$$

where

P_{th} = thermally-generated hole concentration
P_{op} = optically-generated hole concentration
η = quantum efficiency
τ = lifetime
l = sample thickness

The values of P_{th} and P_{op} represent the competition between thermal and optical ionization of the deep level impurity. For a high D^* it is desirable to have $P_{th} \ll P_{op}$ for the performance to be dominated by the optically-generated carriers. Since P_{th} is proportional to $\exp(-E_A/KT)$ where E_A is the energy level of the extrinsic impurity, this sets the basic limit on the operating temperature. In practice things are not quite this simple because all silicon contains residual boron atoms, typically in the range 10^{12}–10^{14} atoms cm^{-3}. For an extrinsic silicon detector operating at around 50 K the boron impurities would still be thermally ionized and give rise to a high P_{th} value which would seriously degrade the detector. There are two possible techniques of reducing this P_{th}. One is to reduce the temperature even more to freeze out the boron atoms and is clearly not desirable; the second is to add compensating donors. This latter technique is difficult to achieve with traditional doping methods; recently however the use of transmutation doping (where the interaction of thermal neutrons with the silicon atoms forms phosphorus atoms) has been used very successfully to achieve accurate compensation.[5]

In the expression for the detectivity D^*, the quantum efficiency, η, is given by

$$\eta = \frac{(1-R_1)[1-\exp(-\alpha l)][1+R_2\exp(-\alpha l)]}{[1-R_1R_2\exp(-\alpha l)]}$$

where R_1 and R_2 are the reflectivities of the front and back surfaces of the sample and α, the absorption coefficient, is given by

$$\alpha = \sigma N_A^\circ$$

where σ is the photoionization cross-section and N_A° is the neutral deep level acceptor concentration.

If R_1 and R_2 are assumed to be negligibly small then

$$\eta = [1-\exp(-\alpha l)] \simeq \alpha l = \sigma N_A^\circ l$$

The photoionization cross-section of suitable dopants is typically of the order of 10^{-16} cm^2; thus assuming an acceptor concentration of 10^{16} cm^{-3} to achieve a quantum efficiency of even 1% requires a detector thickness of 100 μm. Using detectors of this thickness or more in large two-dimensional arrays where the detector spacing may be also of the order of 100 μm or less can lead to serious optical crosstalk problems between adjacent detectors.[6] The other possible way of increasing η is to raise N_A°; unfortunately this is also difficult because the maximum electronically active concentration of most deep level impurities in silicon is low, being typically less than 10^{16} cm^{-3} and special doping or crystal growth techniques may be required to reach the solubility limit. This has led in recent years to an extensive search for suitable deep level impurities. Initially indium and gallium, because of their higher solubilities in silicon, were considered favourably for the 3 μm–5 μm and 8 μm–13 μm wavelength regions.[7,8] Unfortunately neither is particularly well matched to its respective wavelength region and both require cooling to much less than 77 K. More recently other dopants such

as copper, thallium, selenium and sulphur have been investigated[9,10]; however many of these dopants have high diffusion coefficients in silicon and tend to diffuse through the entire silicon wafer during processing.

2.3 Read-out Techniques for Extrinsic Silicon Structures

The most direct approach to building an extrinsic silicon c.c.d. focal plane array would be to use a device in which the c.c.d. storage sites are used to collect the photogenerated charge carriers and such a device was reported by Nelson in 1974.[11] Although this device, known as an 'accumulation mode c.c.d.', resembles a c.c.d. visible image sensor its mode of operation is rather different because, although it is a surface channel structure, the photogenerated charge carriers are majority carriers rather than minority carriers and to achieve this collection of majority carriers the silicon surface is biased into accumulation rather than inversion. The wavelength region of operation is determined by the ionization energy of the deep level impurity and the device must be operated at a temperature low enough to cause the impurity level to freeze-out. Because of this low operating temperature the silicon substrate acts as an insulator and the surface potential appears across the whole of the substrate and the silicon energy bands vary linearly with distance as shown in Fig. 6. Holes generated by the incident photons drift to the oxide-silicon interface and are then shifted out of the device along that interface.

The second and more common approach is to use separate extrinsic detectors and c.c.d. read-out registers fabricated in the same chip. This is achieved by using extrinsic silicon substrate and growing on it an epitaxial layer of opposite conductivity type silicon in which the c.c.d. registers are fabricated. The energy band structure of this device is shown in Fig. 7. In this case the c.c.d. is a conventional minority carrier device and the photogenerated majority carriers are either injected directly across the p(substrate)−n(epitaxial) junction to be collected directly under a c.c.d. storage gate or are collected via a p^+ diffusion at the substrate–epitaxial interface and then injected into the c.c.d.

Few performance data are currently available from extrinsic silicon structures although both line and area arrays have been fabricated with up to 32 × 32 detectors. Indium and gallium have been the principal dopants and responsivity variations from element to element are typically ±15%.[12]

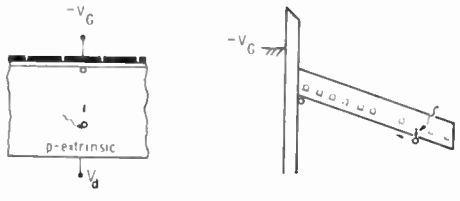


Fig. 6. Accumulation mode extrinsic silicon c.c.d.

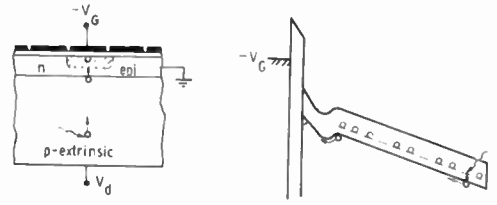


Fig. 7. Extrinsic silicon detector with minority carrier c.c.d. read-out.

2.4 Silicon Schottky Barrier Structures

With these structures the basic sensor cell is a metal electrode deposited directly on to a silicon substrate to form a Schottky barrier with the energy band structure shown in Fig. 9. Like the extrinsic silicon devices described above, this structure is amenable to large scale integration. In operation the Schottky barrier is first reverse biased and then disconnected from the bias circuit. Photons impinging on the metal electrode are absorbed and generate 'hot' electrons. Those electrons whose energy exceeds the metal-silicon barrier height are injected into the silicon and neutralize some of the ionized donors in the depletion region. This results in a reduction of the depletion width and a corresponding reduction in the voltage across the structure which is proportional to the intensity of the optical signal. At the end of the frame time the cell is reset to its original depletion level, the current required to do this providing the video signal. The Schottky barrier height determines the wavelength range of maximum sensitivity and can be selected for the 3 μm–5 μm range; the quantum efficiency increases as the barrier height decreases; so does the thermally-generated dark current, however, and this determines the maximum operating temperature. Since the free carriers are generated in the metal rather than in the silicon substrate the photogeneration rate is independent of the silicon doping concentration and lifetime, thus the photoresponse is very uniform and the sensor resolution is limited by optics and cell dimensions.

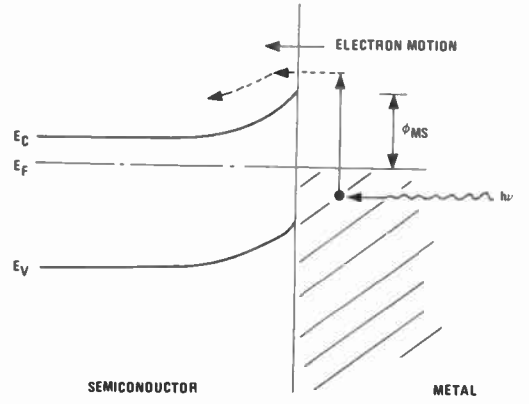


Fig. 8. Energy band structure for a silicon Schottky barrier.

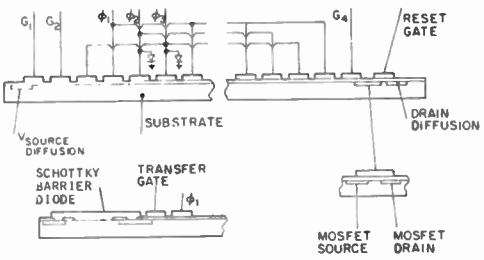


Fig. 9. Circuit diagram of the i.r.-c.c.d. chip. (After Ref. 15.)

The basic structure of the Schottky barrier/c.c.d. combination has been described by Kohn *et al.*[13] and Shepherd *et al.*[14] and is shown in Fig. 9. Both Pd_2Si and Pt_3Si detectors have been used, which have long wavelength cut-offs of 3·5 μm and 4·6 μm respectively, together with three phase surface channel c.c.d.s. Sixty-four-element linear arrays have demonstrated overall uniformity of $<0{\cdot}5\%$ r.m.s.[15]

3 Hybrid Structures

Hybrid structures consist of a detector array coupled to c.c.d. read-out electronics which has been fabricated in a separate structure or material. The advantages of this approach are that the detector and the c.c.d. read-out electronics structure can be developed separately and optimized independently.

One of the primary problems for this type of structure then becomes that of electrically and mechanically interfacing the detector and c.c.d. arrays for optimum performance and reliability. With a general push toward the development of ever higher packing densities and larger number of pixels, (picture elements), the use of photovoltaic detector elements is favoured. While better developed, photoconductive detectors generally suffer from power dissipation problems in these high density configurations. The photovoltaic detectors are readily fabricated in either front or backside-illuminated structures and also exhibit a theoretical D^* that is $\sqrt{2}$ better.[16]

Two general approaches are commonly used in fabricating hybrid arrays: (1) planar processing, frontside-illuminated structures and (2) flip-chip interconnected backside-illuminated configurations. Figure 10 shows a diagram of the planar array structure. In this case the detector base material is attached (via, e.g. epoxy) to the passivated c.c.d. electronics chip. The detector layer is then thinned and the stripes of the detector material delineated by etching techniques. Following the junction formation the detector is passivated and individual leads delineated to connect each detector to each of the c.c.d. input nodes. This type of configuration is generally limited to linear scan, t.d.i. configurations and relatively low packing density 2-dimensional structures.

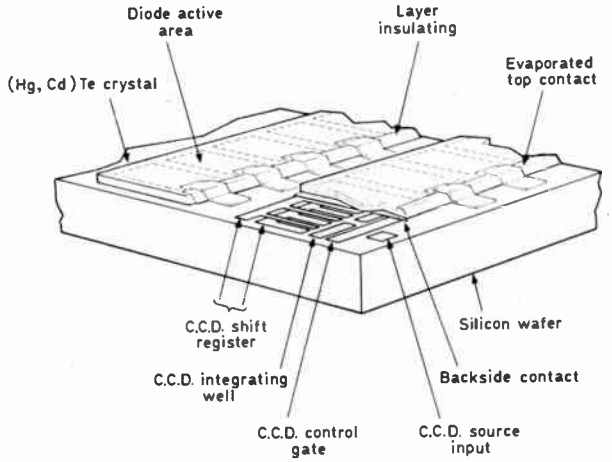


Fig. 10. Hybrid planar array structure.

For focal plane arrays with cell sizes less than about $0{\cdot}127 \times 0{\cdot}127$ mm (5×5 mils) the frontside-illuminated structure becomes inefficient since a larger fraction of the scene radiation will fall on the detector/c.c.d. interconnect areas, rather than on the active areas, as the cell size decreases. The ratio of active area to interconnect area is known as the optical fill factor and to maintain a 100% optical fill factor with these very high density arrays it is necessary to utilize the back-side illuminated flip-chip bump interconnect technique shown in Fig. 11.

In this case, both the detector and c.c.d. arrays have solder bumps formed at the diode junction and c.c.d. input node respectively. After the c.c.d. structure and detector arrays have been characterized individually, the arrays are mated together with the solder bumps providing both the electrical path and mechanical support for the detector/c.c.d. structure. The detector material can be either attached to the support substrate by infra-red transparent epoxy or by, for example,

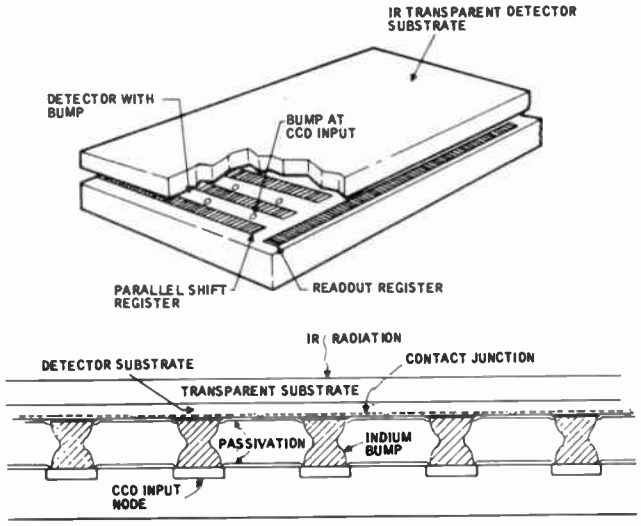


Fig. 11. Backside-illuminated flip-chip bump-interconnect structure.

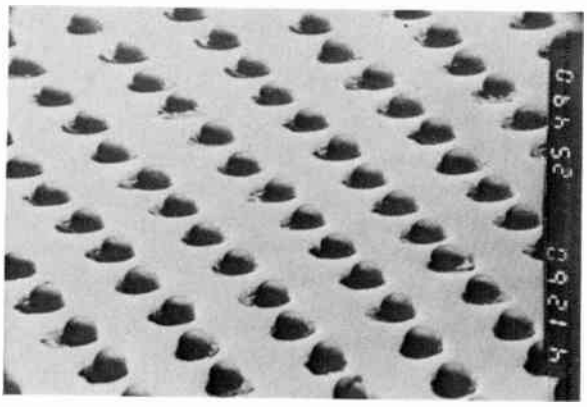


Fig. 12(a) S.e.m. photograph of 32 × 32 element backside-illuminated PbTe array with 1 × 1 mil indium bumps on 4 mil centres.

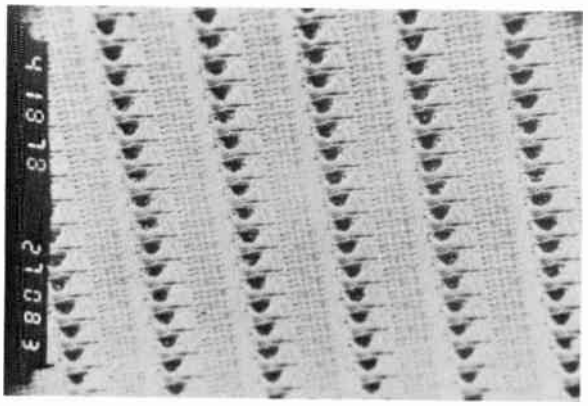


Fig. 12(b) S.e.m. of c.c.d. read-out electronics chip with indium bumps ready for flip-chip interfacing to detector array.

epitaxial growth of the detector onto an appropriate transparent substrate. Figure 12(a) shows a s.e.m. photograph of a PbTe 32 × 32 element detector array with indium bumps ready for interfacing with the c.c.d. readout electronics shown in Fig. 12(b). The detector diodes are diffused junctions and are not evident on the s.e.m. The centre-to-centre spacing for this array is 0·102 × 0·102 mm (4 × 4 mils) with 0·0254 × 0·0254 mm (1 × 1 mil) solder bump interconnects.

Since the detectors are backside-illuminated there is a trade-off between the external quantum efficiency and pixel-to-pixel crosstalk which is determined by the detector thickness and minority carrier diffusion length. Crosstalk levels on the order of 5–10% are generally attainable even for quite high packing density arrays.

To realize the full performance capability of any of the hybrid focal plane structures, it is desirable to have the c.c.d. read-out electronics provide the following functions:

Detector multiplexing
Time-delay and integration (scanned arrays)
Background suppression
Blooming control
Unity noise figure

The problem of background suppression relates to the fact that the limited charge-handling capability of the c.c.d. requires special circuitry to suppress the background-generated current from the detector. This is particularly true for source-coupled input circuits in which the detector current is directly injected into the c.c.d. through a source input diffusion. This is a frequently used coupling technique since it can be designed to minimize c.c.d. input gate $1/f$ noise effects and results in a unity input noise figure. Figure 13 shows a schematic diagram for the source-coupled input with a spill-gate background suppression and anti-blooming control. In operation a charge packet is introduced into the potential well under gate V_w from the diode. This contains both signal information and the large background charge. ϕ_T is then turned on to allow a fraction of this charge packet into the shift register. The remaining charge is then dumped onto drain diffusion V_D by turning on gate ϕ_{BS}. Thus the charge handling requirement of the c.c.d. shift register can be reduced. ϕ_{BS} can also be used as an anti-blooming control in that if its off-level is set just below that of ϕ_T, a large input charge that would overfill well V_W spills over ϕ_{BS} into drain V_D rather than over ϕ_T and into the shift register. In this case the signal injection efficiency, N_s, (fraction of detector signal injected into c.c.d.) can be expressed as:

$$N_s = \frac{g_m R_D}{1 + g_m R_D}$$

where g_m = c.c.d. input transconductance
and R_D = detector a.c. impedance at the detector operating point.

The $1/f$ noise and thermal noise injection efficiency N_{CCD}, can be shown to equal:

$$N_{CCD} = \frac{1}{1 + g_m R_D}$$

Comparison of these equations shows that the noise injection efficiency is strongly attenuated for large $g_m R_D$

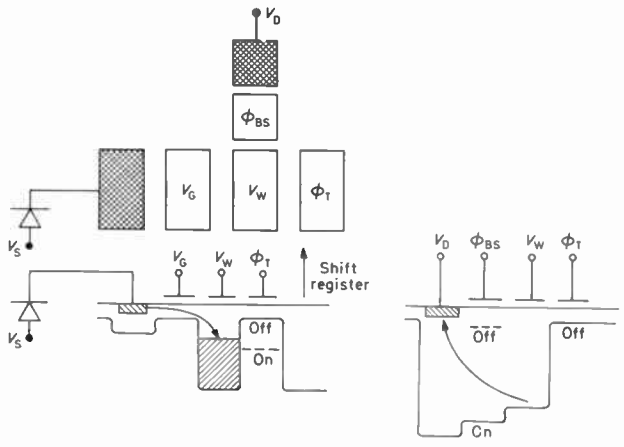


Fig. 13. Source-coupled input coupling technique with anti-blooming and background suppression functions.

products while the detector signal injection efficiency tends to unity. For the case where the input m.o.s. structure is operated in the sub-threshold mode,

$$g_m = \frac{qI_{CCD}}{nKT}$$

where I_{CCD} = c.c.d. input current from the detector, and the condition $g_m R_D \gg 1$ can be written

$$g_m R_D = \frac{qI_{CCD} R_D}{nKT} \gg 1$$

Rearranging this equation gives

$$qI_{CCD} \gg \frac{nKT}{R_D}$$

which is basically the same equation as the definition for b.l.i.p. detectors† i.e. the background generated shot noise, $\sqrt{(2qI)}$, is greater than the detector thermal noise $\sqrt{(4KT/R_D)}$. Thus source-coupled b.l.i.p. detector focal planes result in detector-limited, rather than c.c.d.-limited, performance characteristics.[17]

4 Pyroelectric Detector/C.C.D. Hybrids

A pyroelectric detector provides a means of detecting infra-red signals with the focal plane operating at room temperature. This is a major advantage over all the other hybrid and monolithic systems being investigated which require cooling to at least 180 K and in many cases to 77 K or below.

Basically, a pyroelectric detector is a parallel plate capacitor in which the dielectric has a permanent electric polarization. This polarization is temperature-dependent and when radiation falls on the device and is absorbed, the resultant change in temperature produces a change in charge on the plates of the capacitor and thus a voltage is developed across it. In the steady state, stray charges are attracted to, and trapped on, the surface thereby neutralizing the charge associated with the polarization. However, the surface charges are relatively stable and unable to respond quickly; the detector therefore responds to changes in the radiation incident on it. Because of this dynamic nature pyroelectric detectors are not applicable to imagers which continually survey the scene in a 'staring' mode but are used in a 'blinking' mode with a chopper interposed between them and the scene or in a scanned array.

As the pyroelectric detector produces a voltage output the most direct way of interfacing the detector and c.c.d. is to couple the detector directly to the input gate of the c.c.d. If a d.c. bias is now applied to the detector it is resistively divided between the pyroelectric leakage resistance and the c.c.d. input gate leakage resistance. By adjusting this bias the c.c.d. input gate can thus be biased so that a background level of half a full charge packet can be continuously injected into the c.c.d. If now a fluctuating i.r. signal is incident on the pyroelectric detector, it generates an alternating voltage across the detector capacitance which is capacitively divided between the detector and the c.c.d. input gate capacitance, thus modulating the bias on the c.c.d. input and hence the charge packet size. With a background level of a half-full charge packet being continuously injected into the c.c.d. this charge packet can be increased or decreased by the signal and bipolar signal detection is achieved.

Although Steckl *et al.*[18] initially suggested depositing the pyroelectric material directly on the gate region of the input m.o.s. structure, this has not yet been achieved. Thus the detectors are either simply wire bonded to the c.c.d. inputs for small linear arrays or in 2-dimensional arrays the detector chip and the c.c.d. chip are connected together with an array of solder bumps as discussed previously.

This mechanical interface gives rise to one of the major problems of these hybrid structures, namely that of the heat-sinking of the pyroelectric material. Since the detector relies on temperature changes for its operation this heat-sinking reduces the sensitivity; for example, with a 20 μm thick layer of triglycine-sulphate on silicon and with radiation modulated at 25 Hz, Logan[19] has calculated that the detector responsivity is reduced by a factor of 30. This effect can be reduced by 'chopping' the radiation at a higher frequency, say 2·5 kHz, at which the thermal diffusion length is about 10 μm and very little signal degradation due to heat-sinking will occur in a 20 μm thick device. Some penalty in performance does, however, occur at higher frequencies because of the roll-

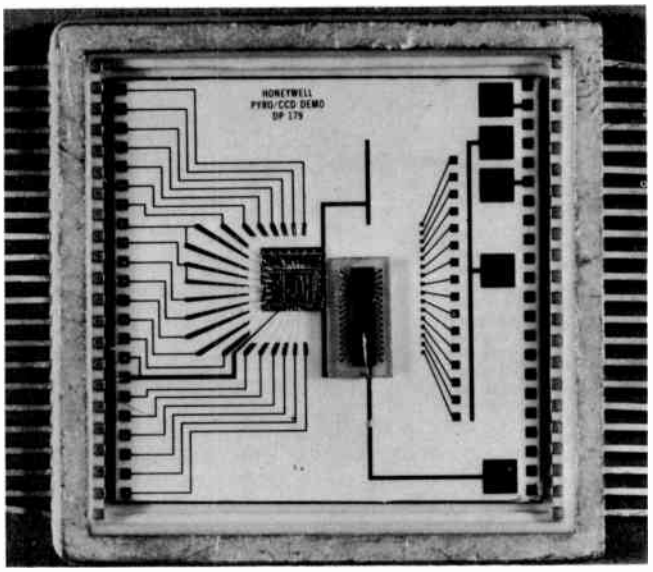


Fig. 14. Photograph of pyro/c.c.d. test structure.

† In detector systems there are numerous noise sources over and above the shot noise inherent in the incident photon flux. The best attainable detector performance will be obtained if the noise introduced by the detector system is less than the photon shot noise. Detectors operating in this condition are said to be background limited infra-red photon or b.l.i.p. detectors.

off of the detector voltage responsivity with frequency. (For a review of pyroelectric detectors and materials see reference 20.)

A further major problem is that of maintaining low noise in the c.c.d. read-out registers. Ideally the c.c.d. noise should be lower than the detector noise and then the D^* of the detector would not be degraded. Because pyroelectric detectors have such a low frequency response, and because the detector is coupled to an m.o.s. gate, as well as the conventional c.c.d. noise contributions, $1/f$ noise in the input m.o.s. structure is important and tends to be one of the major noise sources.

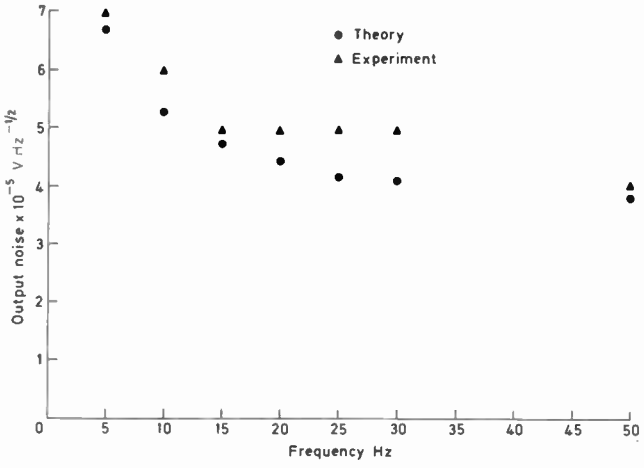


Fig. 15. Noise at pyro/c.c.d. output.

The feasibility of using pyroelectric detectors with c.c.d. read-out has been demonstrated by Iwasa *et al.*[21] with test structures as shown in Fig. 14. The pyroelectric detectors are connected directly to the input gates of a two-phase buried channel c.c.d. multiplexer. Figure 15 shows a comparison of the calculated and measured noise at the c.c.d. output as a function of frequency, and Fig. 16 shows the resultant detectivity of the system as a function of frequency.

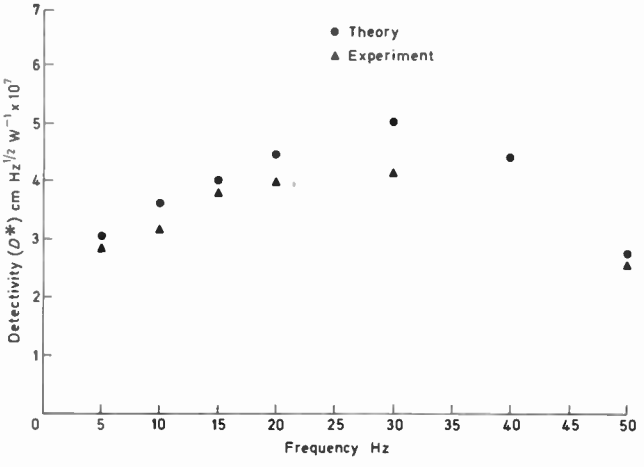


Fig. 16. Detectivity of pyro/c.c.d. structure.

5 Non-uniformity Effects

Infra-red focal planes utilizing c.c.d. read-out electronics generally exhibit a characteristic fixed pattern noise at the output due to variations in: (1) detector quantum efficiency, (2) detector responsivity, (3) detector active area, (4) cut-off wavelength, (5) c.c.d. threshold variations, and (6) channel-to-channel transfer inefficiency effects.

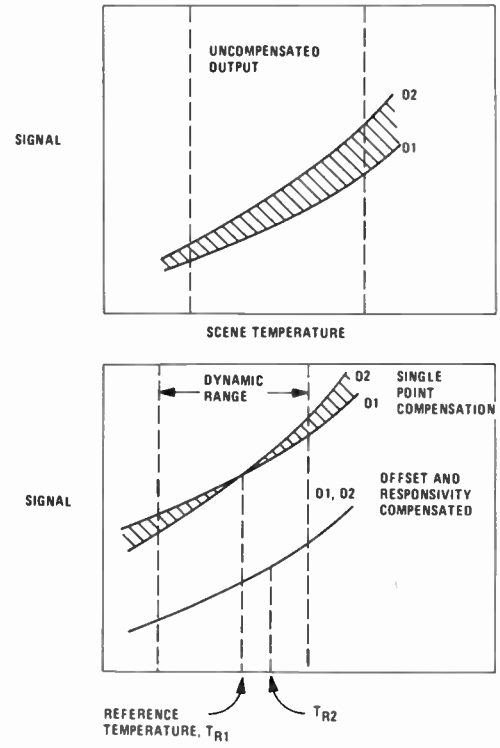


Fig. 17. Effects of non-uniformity compensation on pixel-to-pixel signal variations.

Figure 17 shows a diagram of the channel-to-channel variations for a b.l.i.p. array. Depending upon the focal plane operation, these fixed pattern noise levels can completely mask any low contrast signals (particularly with non-b.l.i.p. systems) and must be normalized (or compensated) in order to realize optimum focal plane performance.

Figure 18 shows an example of the fixed pattern noise from a detector/c.c.d. i.r. focal plane. The d.c. offset levels are about 25% of maximum signal level for this array. Figure 19 shows a video display of a 32×32 element staring focal plane fixed pattern noise for uniform background radiation with the focal plane operating under non-b.l.i.p. conditions. In this case the offset fixed pattern noise is about 50% of the maximum output level and thus reduces the focal plane dynamic range by 6 dB.

Automatic electronic compensation of the responsivity and offset non-uniformities generally requires the use of a thermal reference source for calibration. For linear scanned systems the reference source is generally placed just outside the scene scan area so that the thermal

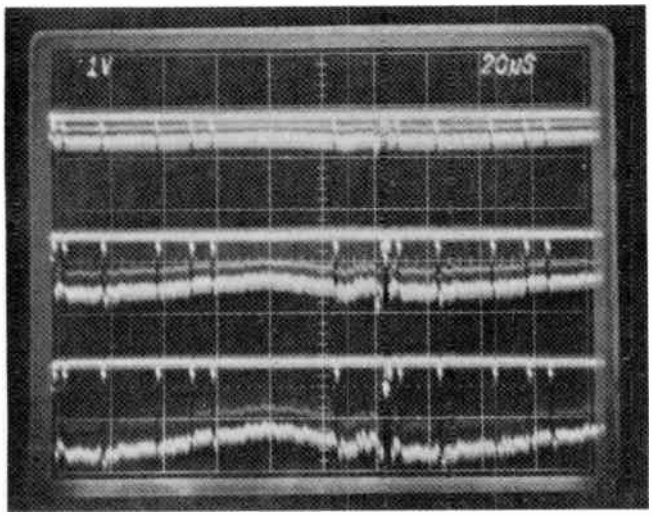


Fig. 18. Characteristic fixed pattern noise from detector/c.c.d. focal plane at three different signal injection levels.

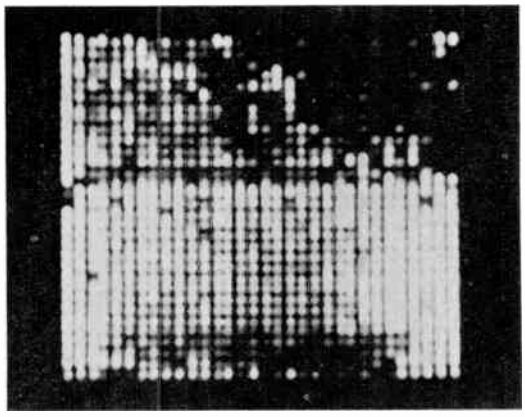
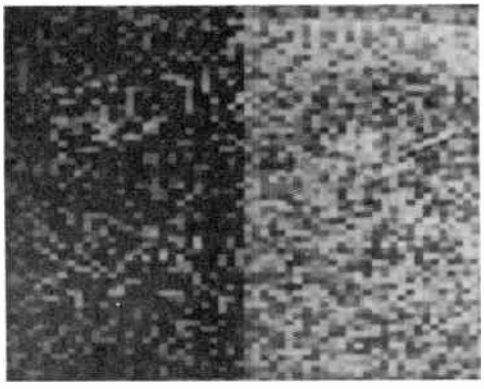


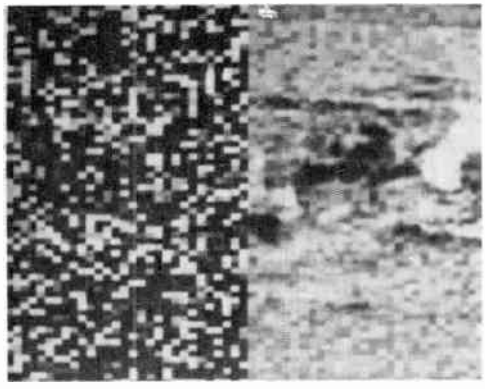
Fig. 19. Video display output from non-b.l.i.p. 32×32 element PbTe/c.c.d. focal plane showing characteristic fixed pattern noise from uniform scene radiation.

reference source is imaged on the array during the vertical retrace time. For staring arrays either a mechanical chopper or electro-optical shutter is used for the non-uniformity thermal reference source.

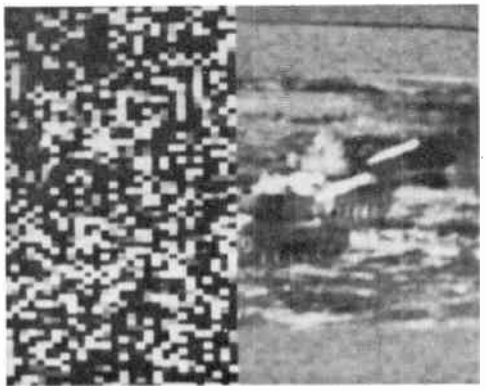
In non-b.l.i.p. imagers in particular, the fixed pattern noise level may be equivalent to a scene temperature variation, ΔT, of 20–50°C (dependent upon frame rate, wavelength and system optics). These high level non-uniformities can be compensated using a variety of algorithms, one of which is demonstrated in Fig. 20. This shows a split-screen video display with the uncompensated f.l.i.r. imagery plus computer-generated offsets (left side of screen) and the corrected f.l.i.r. video on the right side of the display. The algorithm used in



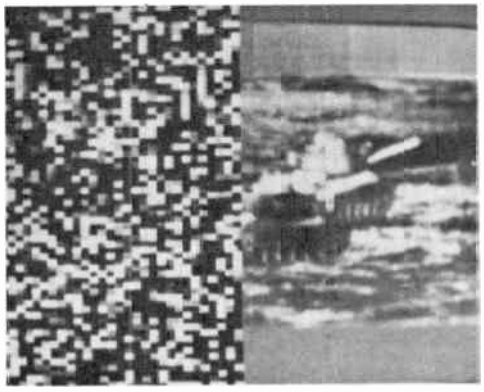
1. Initial, uncorrected frames



3. After 200 sampled frames



2. After 150 sampled frames



4. After 250 sampled frames (full compensation achieved)

Fig. 20. Sequence showing split-screen display of corrected and uncorrected f.l.i.r. imagery containing computer generated fixed pattern noise.

this case employs the scene dynamics to calculate the non-uniformity levels and eventually leads to the full-compensated image (lower right photograph).

6 Conclusion

Through the use of c.c.d. technology, the field of infrared imaging has taken a great stride forward in the development of improved performance i.r. focal planes. The c.c.d. technology has led to new trends in focal plane development which include: (a) more detector elements on the focal plane, (b) increased signal processing functions, (c) increased packing densities, and (d) development of fully populated (i.e. staring) focal planes. It is evident that c.c.d. technology has expanded the earlier perceived performance limitations on i.r. imaging systems by allowing thousands or even millions of detectors to be addressed by monolithic and hybrid structures.

Initial expectations of the performance of monolithic extrinsic silicon focal planes have not yet been fully accomplished. This, along with limited spectral band operation and extremely low temperature operation, has resulted more recently in an emphasis being placed on other detector materials such as HgCdTe and InAsSb operating in hybrid configurations whose spectral band can be accurately tailored to the application and where elevated temperatures operation can also be achieved.

Initial work on monolithic c.c.d. structures (i.e. monolithic HgCdTe) has shown very promising results but this development appears to lag behind hybrid development. This type of structure may also have wavelength limitations, however, due to tunnelling problems in the longer wavelength materials.

Regardless of whether it is the monolithic or hybrid structure which results in the optimum imaging systems, it is clear that c.c.d. technology is crucial to the development of any of these advanced i.r. systems.

7 References

1 Anderson, W. W., 'Tunnel current limitations of narrow bandgap infrared charge coupled devices', *Infrared Physics* **17**, No. 2, p. 147, 1977.

2 Kim, J. C., 'InSb charge injection device imaging array', *IEEE J. Solid State Circuits*, SC-**13**, no. 1, p. 232, 1978.

3 Chapman, R. A., Kinch, M. A., Borello, S. R., Simmons, A., Roberts, C. G. and Buss, D. D., 'HgCdTe charge coupled devices', AIAA/NASA Conference on 'Smart' Sensors, Hampton, Va, 1978.

4 Thom, R. D., Koch, T. L. and Parrish, W. J., 'Indium antimonide infrared c.c.d. linear imaging arrays with on-chip preprocessing', AIAA/NASA Conference on 'Smart' Sensors, Hampton, Va, 1978.

5 Thomas, R. N., Braggins, T. T., Hobgood, H. M., Takei, W. J. and Nathanson, H. C., 'Compensation of extrinsic Si: In detector material by neutron transmutation of silicon', IRIS Detector Specialty Conf., Colorado Springs, Co, 1977.

6 Nummedal, K., Fraser, J. C., Su, S. C., Boron, R. and Fimmila, R. M. 'Extrinsic silicon monolithic focal plane array technology and applications', Proc. Int. Conf. on the Applications of CCDs, San Diego, p. 19, 1975.

7 Pines, M. Y. and Baron, R., 'Characteristics of indium doped silicon i.r. detectors', Proc. Int. Electron Devices Meeting, Washington, p. 446, 1974.

8 Pines, M. Y., Murphy, D., Alexander, D., Baron, R. and Young, M., 'Characteristics of gallium doped silicon i.r. detectors', Proc. Int. Electron Devices Meeting, Washington, p. 502, 1975.

9 Sclar, N., 'Extrinsic silicon detectors for 3–5 μm and 8–14 μm', *Infrared Physics*, **16**, no. 4, p. 435, 1976.

10 Sclar, N., 'Survey of dopants in silicon for 2–2·7 μm and 3–5 μm infrared detector applications', *Infrared Physics*, **17**, no. 1, p. 71, 1977.

11 Nelson, R. D., 'Accumulation-mode charge-coupled devices', *Appl. Physics Lett.*, **25**, no. 10, p. 568, 1974.

12 McLean, E. O. and Merriam, J. D., 'Evaluation of 2 × 32 Si: In mfpa', IRIS Detector Specialty Conf., Colorado Springs, Co., 1977.

13 Kohn, E. S., Roosild, S. A., Shepherd, F. D. Jr. and Yang, A. C., 'Infrared imaging with monolithic, c.c.d.-addressed Schottky barrier detector arrays: theoretical and experimental results', Proc. Int. Conf. on CCD Applications, San Diego, 1975.

14 Shepherd, F. D., Yang, A. C., Roosild, S. A., Bloom, J. H., Capone, B. R., Ludington, C. E. and Taylor, R. W., 'Silicon Schottky barrier monolithic irtv focal planes', *Adv. in Electronics and Electron Physics*, **40B**, p. 981, 1976.

15 Taylor, R. W., Shepherd, F. D., Roosild, S. A., Yang, A. C. and Kohn, E. S., 'Schottky irccds', IRIS Detector Speciality Conf., Colorado Springs, Co., 1977.

16 Foss, M. A., Gurnee, M. N. and Marshall, D., 'Focal plane ccd for t.e. cooled detector arrays', Contract No. DAAKO-76-C-0250, Army Night Vision Laboratory, 1977.

17 Foss, N. A., Gregorich, D. and Marshall, D., 'Direct coupled pv detector/ccd interfacing for non-blip detector focal planes', IRIS Detector Specialty Conf., Annapolis, Maryland, 1978.

18 Steckl, A. J., Nelson, R. D., French, B. T., Gudnumdsen, R. A. and Schechter, D., 'Application of ccds to ir detection and imaging', *Proc. IEEE* (Special issue on IR Technology), **63**, p. 67, 1975.

19 Logan, R. M., 1977 (Unpublished work at RSRE, Malvern, Worcs., England.)

20 Liu, S. T. and Long, D., 'Pyroelectric detectors and materials', *Proc. IEEE*, **66**, p. 14, 1978.

21 Iwasa, S., Gelpey, J., Marciniec, J., Marshall, D., White, W., Lamb, D., Liu, S. T. and Paffel, D., 'Pyro/ccd direct signal injection; theory and experiment', IEDM, Washington D.C., 1978.

Manuscript first received by the Institution on 31st August 1979 and in final form on 7th January 1980.
(Paper No. 1932/CC 324)

UDC 621.382.049.771 : 621.395.664

Indexing Terms: Semiconductor devices, Integrated circuits, Charge-coupled devices, Bandwidth compression

Multiplexed c.c.d.s for bandwidth compression applications

J. C. WHITE, Ph.D.,*

J. G. HARP, B.A.,*

J. R. HILL, B.Sc., Ph.D.,*

D. V. McCAUGHAN,
B.Sc., Ph.D., F.Inst.P.,*

J. M. KEEN, M.Sc.*

and

Professor J. D. E. BEYNON,
M.Sc., Ph.D., C.Eng., F.I.E.E., F.I.E.R.E.†

SUMMARY

A 321-cell c.c.d. bandwidth compressor capable of analogue sampling at 100 MHz is described. This system is designed to accept data with a bandwidth of 50 MHz and subsequently to clock the data out at rates up to 2 MHz for recording on magnetic tape via a cheap, low speed a–d converter. Since power requirements are at a premium, low overall power consumption and high speed performance were essential design goals. Details of the chip architecture are given and associated driver circuitry and preliminary experimental results described.

* *Royal Signals and Radar Establishment, St. Andrews Road, Malvern, Worcestershire.*
† *Department of Electronic and Electrical Engineering, University of Surrey.*

1 Introduction

C.c.d. analogue shift registers are ideal for bandwidth compression (time expansion) or bandwidth expansion (time compression), principally because the data clocking rate can be easily varied. For the former, a constant high-speed clocking rate provides continuous fast sampling of data until the c.c.d. is filled; the clocking rate may then be reduced and the data read out at a slower, more manageable rate. A simple and common type of bandwidth compressor can be implemented with a single c.c.d. channel.

When a charge packet is clocked forward, charge transfer inefficiency (c.t.i.) causes a small fraction (typically 1×10^{-5} to 1×10^{-3}) to be left behind so corrupting the information in the following cell. The amount of this 'smearing' depends on the total number of transfers and the clocking rate, more charge generally being left behind at higher clock speeds. For the particular case of a bandwidth compressor the samples read in first undergo a larger number of high speed transfers than the samples read in last. This results in the first bits having a greater smearing than the last.

The approach we have adopted to reduce the effects of c.t.i. is to employ three parallel multiplexed c.c.d. channels. Such an arrangement alleviates the problems caused by c.t.i. because

(a) multiplexing reduces the total number of transfers undergone by a particular charge packet and
(b) a given sampling rate requires a lower c.c.d. clock frequency.

Thus a 3-phase c.c.d., multiplexed three ways, reduces the total number of transfers by a factor of 3 and also allows a three-fold reduction in clocking rate.

A potential difficulty associated with any multiplexed structure is that of non-uniformity between channels, If, for example, the gain associated with channel 1 of a 3-channel system were slightly lower than that of the other two channels, every third pulse in the output waveform would be attenuated. Although the problem is not insurmountable, the added circuitry needed to overcome it detracts from the advantages of multiplexing. Non-uniformity is less of a problem in a single-chip design than would be the case if discrete devices were multiplexed; nevertheless in designing the integrated multiplexed c.c.d.s described in the next Section, particular attention was paid in the layout to minimizing differences between channels. In particular, the channels are close together and are driven by common clocks; the structure was made as simple as possible by making the degree of multiplexing equal to the number of phases, (three and four phases respectively in our structures).

In the next two Sections we consider in more detail points (a) and (b) above.

0033–7722/80/050237+06 $1.50/0

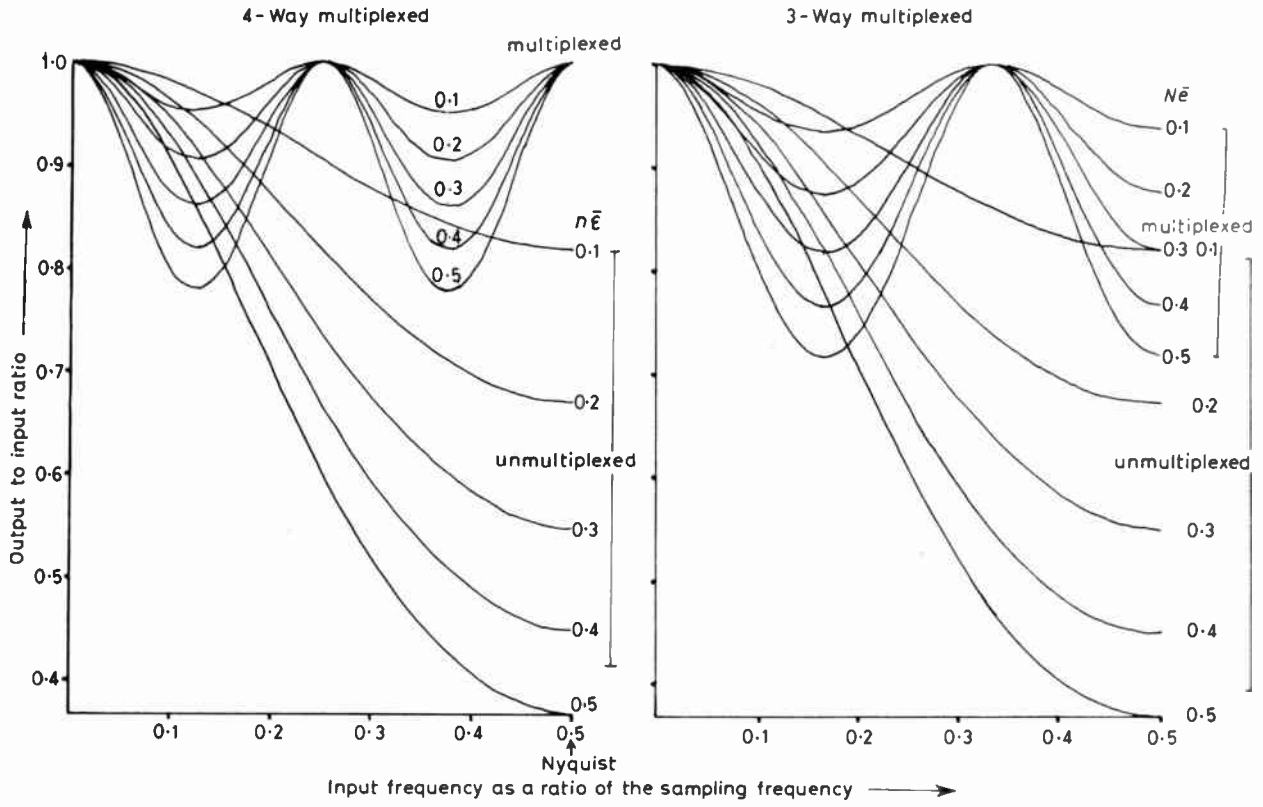


Fig. 1. Frequency response of multiplexed c.c.d.s.

2 Comparison of C.T.I. Effects in Multiplexed and Un-multiplexed Structures

For a single channel c.c.d. with no multiplexing, the transfer function (in z notation) is given by (1)

$$H(z) = \left[\frac{(1-\bar{\varepsilon})}{1-\bar{\varepsilon}z^{-1}}\right]^n z^{-n} \qquad (1)$$

where n is the total number of c.c.d. stages, and $\bar{\varepsilon}$ is the transfer inefficiency per stage.† If $\bar{\varepsilon}$ is very small (typically 10^{-4}) then (1) may be approximated as follows:

$$H(z) \simeq \{\exp n[\ln(1-\bar{\varepsilon}) - \ln(1-\bar{\varepsilon}z^{-1})]\}z^{-n}$$
$$= \{\exp[-n\bar{\varepsilon}(1-z^{-1})]\}z^{-n}.$$

Thus, in addition to the required time delay (represented by z^{-n}), we have a dispersion $D(z)$ given by

$$D(z) = \exp[-n\bar{\varepsilon}(1-z^{-1})]$$

caused by c.t.i. Substituting $z = \exp(j2\pi f/f_s)$, where f_s is the c.c.d. clocking frequency, we can translate this dispersion into a frequency response

$$|D(f)| = \exp[-n\bar{\varepsilon}(1-\cos 2\pi f/f_s)] \qquad (2)$$

For a k-way multiplexed device a similar expression to (1) may be written down

$$H(z)_{(\text{mult})} = \left[\frac{(1-\bar{\varepsilon})^{n/k}}{(1-\bar{\varepsilon}z^{-k})}\right] z^{-n} \qquad (3)$$

which leads to

$$|D(f)| = \exp\left[-\frac{n\bar{\varepsilon}}{k}(1-\cos k \,.\, 2\pi f/f_s)\right]. \qquad (4)$$

A comparison between (1) and (3) or (2) and (4) shows that the multiplexed device exhibits considerably less dispersion than the normal single c.c.d. delay line. Although this gives less smearing of the signal, due to the fewer number of transfers in the multiplexed case, the smearing will occur over a larger time period, i.e. instead of the charge smearing into the adjacent time slot it spreads into the kth time slot when demultiplexed at the device output.

The effect on the c.c.d. transfer function is shown in Fig. 1. Multiplexing decreases the magnitude of the ripple in the c.c.d. passband but its periodicity is increased.

The above analysis applies for a c.c.d. clocked at a constant rate. Clearly if the clocking speed is varied, as in bandwidth compression, the analysis becomes slightly more complex and will depend on the relative transfer inefficiencies at the two clock rates. However the basic consideration outlined above will still apply.

† Note: In Beynon's companion paper,[6] ε was defined as the transfer inefficiency *per charge transfer*; here, $\bar{\varepsilon}$ is defined as the transfer inefficiency *per stage*. Thus $\bar{\varepsilon} = p\varepsilon$ where p is the number of phases.

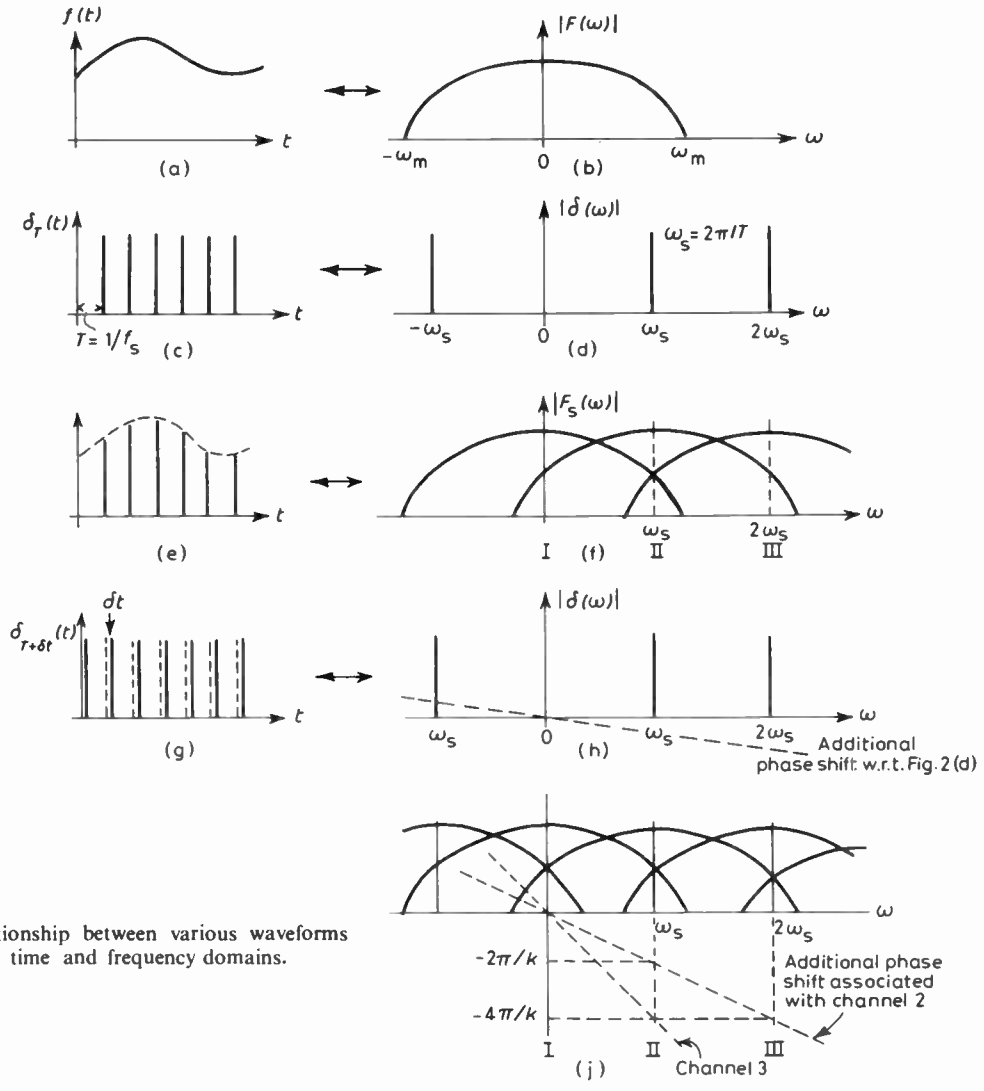


Fig. 2. Relationship between various waveforms in the time and frequency domains.

3 Aliasing in Multiplexed Channels

Problems of aliasing arise in any sampled-data system if the sampling rate is less than twice that of the highest frequency component in the signal—the so-called Nyquist criterion. In a k-channel multiplexed system, the signal is sampled by each channel at $1/k$ of the overall sampling rate; it follows, therefore, that if the overall sampling rate just satisfies the Nyquist criterion, this condition is violated for the individual channels, i.e. substantial aliasing will occur within each channel. It is interesting to consider this situation in more detail.

Firstly consider a signal $f(t)$ (Fig. 2(a)) of bandwidth ω_m (Fig. 2(b)) which is sampled at intervals T (Fig. 2(c)). The spectrum of the sampled signal (Fig. 2(e)) is shown in Fig. 2(f), this being the convolution of Fig. 2(b) and the spectrum of the sampling signal itself (Fig. 2(d). Bands II, III, etc. are replicas of the original signal spectrum multiplexed at frequencies ω_s, $2\omega_s$, etc., where $\omega_s = 2\pi/T$.

Suppose now that $f(t)$ is sampled by a train of pulses identical to Fig. 2(c) except that the train (Fig. 2(g)) is delayed by δt with respect to Fig. 2(c). The spectrum of Fig. 2(g) is identical in magnitude to Fig. 2(d) but has an additional phase delay of $\exp(-j\omega\delta t)$ (Fig. 2(h)). It follows that the spectrum of $f(t)$ sampled by this delayed signal is identical in magnitude to Fig. 2(e); band II frequencies will, however, have an additional phase lag of $\exp(-j\omega\,\delta t)$, band III frequencies an additional phase lag of $\exp(-2j\omega\,\delta t)$, etc.

In the case of a k-way multiplexed system in which each channel is clocked with a periodicity T, channel 2 can be thought of as sampling the signal at times T/k after channel 1, channel 3 samples will be at times $2T/k$ behind channel 1, and so on. It follows that Band II of the spectrum of the channel 2 samples has a phase lag of $\omega_s T/k = 2\pi/k$ (since $T = 2\pi/\omega_s$) with respect to Band I; Band III of the channel 2 spectrum has a phase lag of $4\pi/k$ with respect to Band I, as does Band II of channel 3; this situation is depicted in Fig. 2(j).

At the output of the multiplexer, all the channels are summed, i.e. the k spectra of Fig. 2(j) with their different phase shifts are, in effect, superimposed. Referring to Fig. 3(a) and considering Band 1, we see that the output is the sum of the individual spectra, i.e. we have a resultant band I spectrum of k times the strength of the spectra of the individual channels. When we sum the band II components, however, we have k spectra each having a phase shift of $2\pi/k$ with respect to each other; the resultant spectrum is therefore zero (by 'vector' addition we have a closed k-sided polygon). This is also true of bands III, IV, . . ., k; this situation is depicted in Fig. 3(a) for $k = 3$. Band $k+1$ components, however, are effectively all in phase (since they differ by 2π) and so, as in the case of Band I, the spectra of the individual channels add. The spectrum of the output is therefore as depicted in Fig. 3(b). It follows that the overall system response is identical to that of a single-channel c.c.d. which samples the incoming signal at a frequency $k/T = kf_s$. Thus the signal may contain frequencies up to $kf_s/2$ without causing aliasing in the final output, even though substantial aliasing occurs in the spectra of the individual channels.

4 Multiplexed C.C.D. Implementation—Chip Architecture

A 3-phase–3-way multiplexed c.c.d. was designed and manufactured at RSRE for an initial investigation into multiplexed structures. The first device, shown in Fig. 4 has 321 samples, but longer devices (4 phase, 1024 bits, and 3 phase, 1026 bits) have also been designed.

The device shown in Fig. 4 is an n-channel structure fabricated on a p-type substrate and has overall

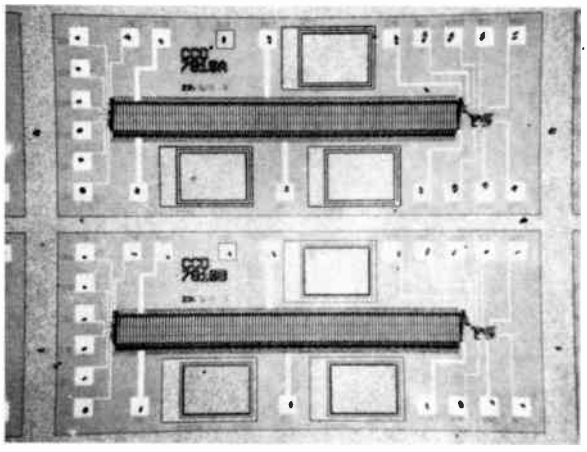


Fig. 4. 321-stage 3-phase c.c.d. bandwidth compressor layout.

dimensions of 3·1 mm × 1·3 mm and a cell size of 21 μm × 50 μm. It employs three parallel channels 50 μm wide delineated by p^+ channel stops. Automatic input and output demultiplexing is provided by staggering the three inputs and outputs such that each input is sampled in turn during each appropriate clock phase. Figure 5 shows the input structure in more detail. The two-gate input structures on each channel have a common reference electrode (i.p.g.), each of the other input gates (IG1–IG3) being fabricated on the same level of polysilicon. Since each phase electrode uses a different polysilicon level, a small diffusion region separates IG1–3 from the first phase electrode in each channel; threshold voltage variations from one level of polysilicon to the other are thus avoided. There will still be a slight threshold voltage variation (of the order of 10 mV) from channel to channel, however, even under the same level of polysilicon, due to substrate doping fluctuations; therefore each input gate (IG1–3) is separately addressable to enable fine tuning of the balance between channels.

Various input techniques for high speed c.c.d.s have been discussed by other authors.[2, 3] Many have chosen a charge partition or pulsed input gate technique where an input gate is pulsed rapidly off, leaving under the adjacent gate a charge packet, the magnitude of which is determined by the potentials of the gate and the input source diffusion. This should be appreciably faster than input charge equilibration techniques such as phase-referred[4] or fill-and-spill.[5] However considerable additional high-speed circuitry would be necessary for

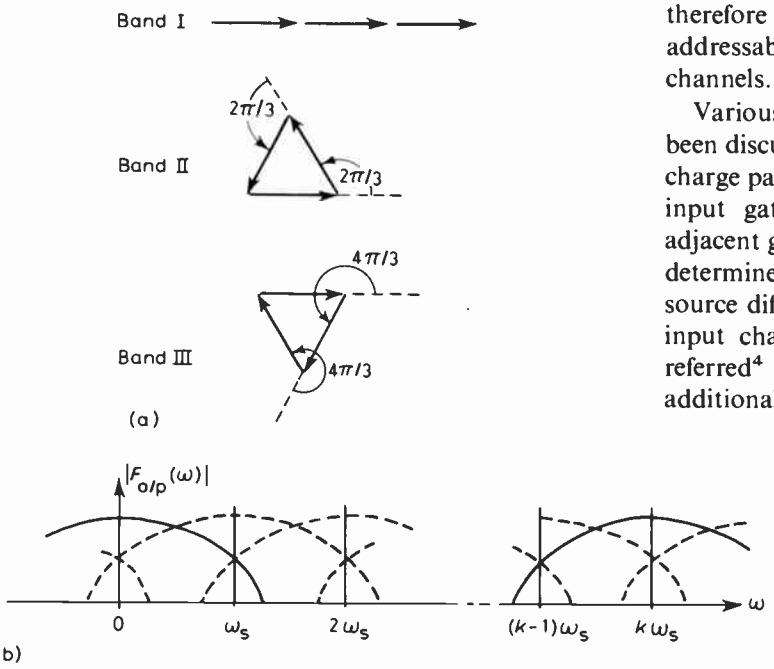


Fig. 3. Summation of outputs of signals from all k channels. (a) 'Vector' addition of components (for $k = 3$). (b) Resultant spectrum.

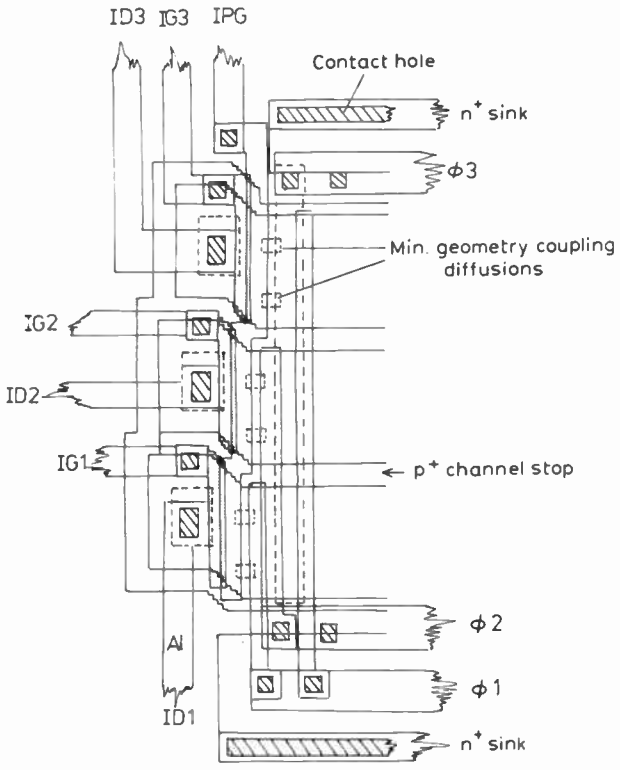


Fig. 5. Input structure of multiplexed 3-phase c.c.d. bandwidth compressor.

the generation of the gating pulse. To avoid the need for additional circuitry, we have employed a simple phase-referred input method instead. This is an easy technique to implement, the appropriate input diode being connected to the phase immediately following the metering well in each channel, e.g. phase 1 in the lower channel, phase 2 in the second and so on. Experimental work, to be reported elsewhere, shows that the harmonic distortion associated with this inputting arrangement is better than −40 dB. At the output a demultiplexing structure similar to the input is employed so that each charge packet is fed through a common amplifier/reset system.

In summary, channel-to-channel uniformity is thus ensured by

(a) making each channel of identical dimensions
(b) giving each channel an identical input structure
(c) feeding the outputs into a common amplifier/reset system and
(d) putting all channels close together on the same chip to minimize threshold and dark current variations.

Note that n^+ diode sinks along the periphery of the channels are provided to reduce dark current effects associated with the adjoining substrate.

5 Driver Circuitry

The necessary driver circuitry for the system is provided off-chip; a block diagram is shown in Fig. 6. A start pulse initiates the control logic to read in the input data by switching the 100 MHz oscillator through the multiplexer and drivers to the c.c.d. A divide-by-107 counter stops the high speed clocks when the device is full and data are then read out at the slower 2 MHz rate. In the change-over period, from the 100 MHz to the 2 MHz rate, it is essential that one of the phases is held high with the others off. If this is not the case, and the 2 MHz clock does not start up in a controllable manner, then data may be lost. When the data have been read out the control logic is enabled to accept another start pulse and the cycle repeated. The entire board including the c.c.d. occupies an area 22 cm × 12 cm and employs MECL 10 K for the high-speed logic. Total power dissipation at 100 MHz is 7·5 W.

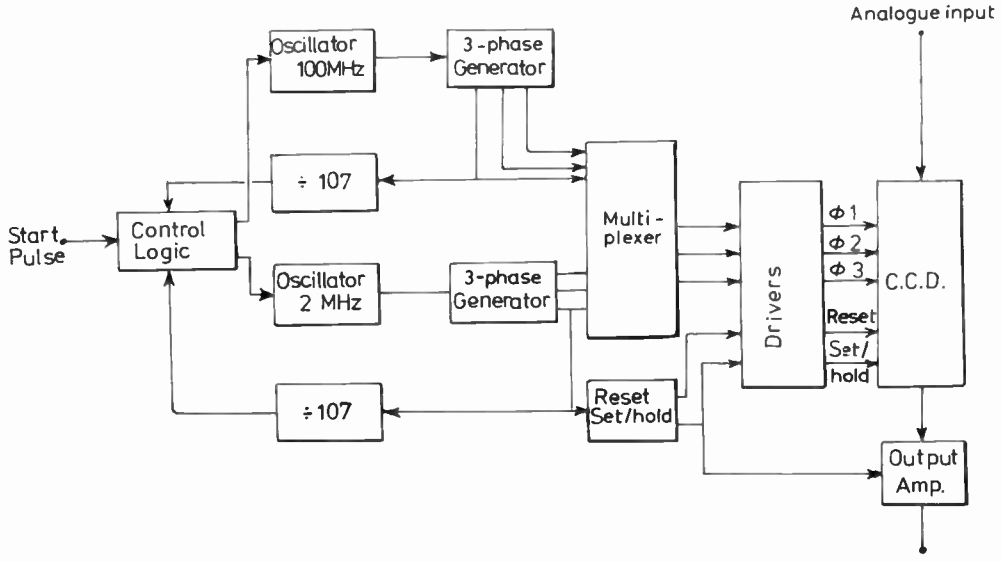
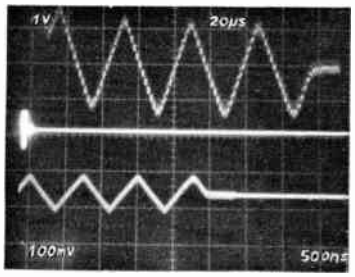


Fig. 6. Block diagram of c.c.d. driver and control circuitry.

Table 1
Device Characteristics

Size	3·1 mm × 1·3 mm
Record length	321
Sampling frequency	100 MHz
Output data rate	2 MHz
Transfer inefficiency (max)	$\sim 10^{-3}$ per stage
Dynamic range	47 dB
Linearity	>30 dB
Channel-to-channel non-uniformity	1%
P.c.b. size	220 mm × 120 mm
Power dissipation	7·5 W

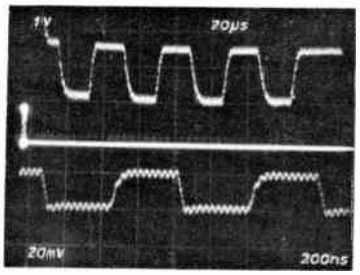


Output

Input (on same time axis) 20 μs/div

Input (expanded time axis) 500ns/div

(a)

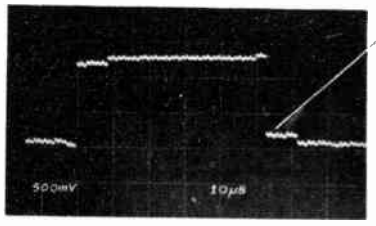


Output

Input (on same time axis) 20 μs/div

Input (expanded time axis) 200ns/div

(b)



(c)

Fig. 7. (a) and (b) Results for triangular and square wave input waveforms, respectively. (c) The device being used as a straightforward delay line; note how multiplexing makes c.t.i. effects more apparent.

6 Results

A number of simple surface channel versions of the device shown in Fig. 4 have been tested. Buried channel versions which should exhibit better transfer efficiency are awaiting testing.

Figures 7(a) and (b) show some results obtained at an input sampling rate of 100 MHz, the output being clocked at approximately 2 MHz. The photograph of the triangular waveform (Fig. 7(a)) gives an indication of the linearity and the degree of channel-to-channel matching obtainable with the device. The finite slopes on the output waveform of the supposedly rectangular pulses (Fig. 7(b)) are due partly to the poor input waveform and partly to transfer inefficiency effects. The effects of transfer inefficiency are more easily observed in Fig. 7(c) where the device is clocked at a constant frequency, i.e. it is being operated as a straightforward delay line rather than as a bandwidth compressor. More detailed results of the characteristics of the device are given in Table 1. Reasonable linearity and dynamic range are obtainable even with a surface-channel c.c.d. operating at such high sampling frequencies. A buried-channel version will clearly give enhanced transfer efficiency and linearity and should be capable of operating at several hundred megahertz. More detailed results including an analysis of the charge input and output will be given in a later paper.

7 Conclusion

Various structures for high-speed transient recording have been discussed. A simple surface-channel multiplexed c.c.d. has been shown to be capable of sampling data at 100 MHz while buried-channel four-phase versions should bring 200 MHz easily within reach.

8 References

1 Sequin, C. H. and Tompsett, M. F., 'Change Transfer Devices' (Academic Press, New York, 1975).
2 Gradl, D. A., Groenwald, R. A. and Flanagan, T. J., 'High speed operation of c.c.d.s', Proc. 1975 Int. Conf. on the Applications of C.C.D.s, pp. 389–98, 1975.
3 Chan, Y. T., French, B. T., Hughes, A. J., Lin, W. N., McNutt, M. J. and Meyer, W. E., 'A 1024-cell c.t.d. shift register capable of digital operation at 50 MHz', Proc. 1978 Int. Conf. on the Applications of C.C.D.s, pp. 471–83, 1978.
4 McCaughan, D. V. and Harp, J. G., 'Phase-referred input: a simple new linear c.c.d. input method', *Electronics Letters*, **12**, no. 25, pp. 682–3, 1976.
5 Sequin, C. H. and Mohsen, A. M., 'Linearity of electrical charge injection into charge-coupled devices', International Elect. Dev. Meeting, Washington D.C. December 1974.
6 Beynon, J. D. E., 'Charge coupled devices: concepts, technology and limitations', *The Radio and Electronic Engineer*, **50**, no. 5, pp. 201–4, May 1980.

Manuscript first received by the Institution on 29th November 1979 and in revised form on 18th February 1980.
(Paper No. 1933/CC 325)

UDC 621.382: 621.391.826.2: 621.396.963.3
Indexing Terms: Semiconductor devices, Integrated circuits, Charge-coupled devices, Radar interference, clutter

C.c.d. integration techniques for clutter reduction in radar systems

C. P. TRAYNAR, B.Sc., Ph.D.*

and

Professor J. D. E. BEYNON,
M.Sc., Ph.D., C.Eng., F.I.E.E., F.I.E.R.E.†

SUMMARY

Substantial signal-to-clutter improvements are theoretically achievable using c.c.d.s to integrate a large number of returns. In practice, however, the performance of simple c.c.d. implementations is severely limited by charge transfer inefficiency, particularly if large range and high resolution are required. The paper describes several ways in which the effects of transfer inefficiency can be overcome or avoided.

* *Plessey Semiconductors Ltd., Cheney Manor, Swindon, Wiltshire SN2 2QW.*
† *Department of Electronic and Electrical Engineering, University of Surrey, Guildford, Surrey GU2 5XH.*

1 Introduction

In many radar situations, the detection of targets is seriously limited by the presence of unwanted reflections (clutter) and noise. As a result many signal processing systems have been devised in order to increase the probability of detecting genuine targets. The invention of the charge-coupled device (c.c.d.) has enabled virtually all of these schemes to be implemented with, in many cases, a substantial reduction in hardware, size, weight and power consumption. The reason for the swift acceptance of c.c.d.s in this area is the ease with which many of the complicated mathematical operations necessary in signal processing can be performed with these devices. This results from their ability accurately to manipulate, spatially and temporally, discrete charge packets representing a sampled analogue signal.

The proliferation of systems has occurred because no single processor can cope adequately with the wide variety of situations encountered. Thus, for example, when both the target and the clutter are moving with different velocities relative to the radar platform, the target and clutter signals are most effectively separated by spectral analysis of the Doppler-shifted return waveform. Charge-coupled devices permit spectral analysis to be elegantly effected by using transversal filters to perform the many complex multiplications encountered in the discrete Fourier Transform; several implementations have been described.[1-7]

In the less complex radar situation where there is relative motion only between the target and the transmitter, delay line canceller techniques can be used to increase the signal-to-clutter ratio. Being basically an analogue delay line, a simple c.c.d. can readily be used for this application. In those cases where a more uniform detection probability is required, a c.c.d. multiple-pole recursive filter can be used.[8]

A radar situation commonly encountered is one where there is little or no relative motion between the target and the radar platform as, for example, in the detection of navigation buoys and small boats with land-based radars. One technique which may be employed in such circumstances is matched filtering which enables the signal-to-clutter ratio of a radar return to be optimized. Several c.c.d. matched filters have been constructed, some of which can be 'adapted' to a specific radar environment.[9-12]

A conceptually-simpler technique for processing radar returns of this type, however, is that of signal integration. The simplicity and effectiveness of this approach has been well established in the past using long-persistence c.r.t.s and photographic techniques although these have generally been inconvenient; digital integration techniques are feasible but are, of course, cumbersome and expensive. The advent of the c.c.d., however, has made possible extremely simple but much more versatile approaches based on delay-line storage techniques and it

0033–7722/80/050243+06 $1.50/0

is with various realizations of c.c.d. integrators that we shall be concerned in this paper.

2 Principles of Signal Integration

The principle of signal integration is based upon the periodic nature of a target signal in successive returns from a single bearing and the randomness of the clutter signals. Summation of m successive returns leads to an m-fold increase in target signal while the clutter signal increases by only $\sqrt{m}$ (providing the sweep period is longer than the clutter decorrelation time which for sea clutter is typically ~ 100 ms).

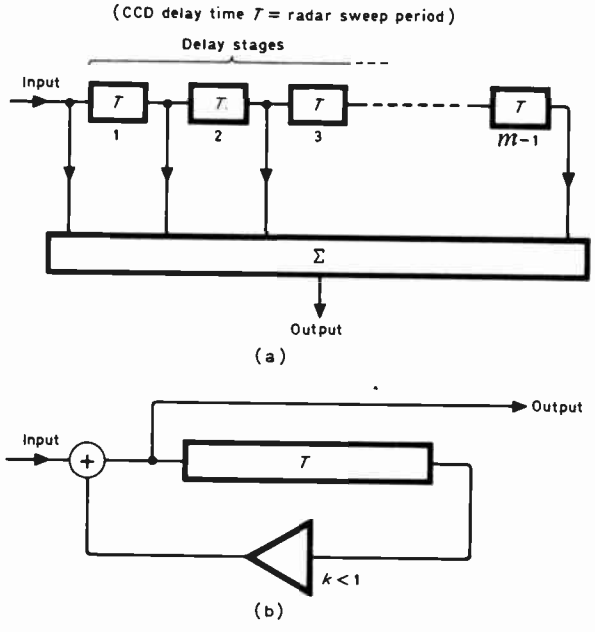


Fig. 1. (a) A non-recursive integrator. (b) A recursive integrator.

Conceptually, the simplest c.c.d. integrator would consist of a series of cascaded delay lines (Fig. 1(a)), the output of each delay line being summed. This approach (called the non-recursive integrator) clearly forms the sum of m equally weighted returns. The number of delay lines can be reduced to one if the recursive technique of Fig. 1(b) is used. In this case simple algebra shows that the sum is formed from exponentially weighted returns; the weighting and hence the effective number, m, of returns forming the sum is determined by the feedback coefficient k, m being approximately $(1-k)^{-1}$. If the hardware is not a limiting factor, the choice of integration technique depends upon whether or not the number of returns available is limited (as would be the case, for example, if a target is moving slowly through a range bin). It has been shown[13] that, for an infinite number of returns, the recursive approach is best; if, however, only a finite number is available, and if this number is known, the non-recursive approach provides the best performance. The majority of c.c.d. integrators described in the literature employ weighted summation.[14, 15]

Theoretically, large signal-to-clutter improvements are possible with signal integrators simply by summing a large number of returns. In practice, however, the performance of simple c.c.d. implementations is severely limited by charge transfer inefficiency (c.t.i.), particularly if large range and high resolution are required. This is because a large number of charge transfers results in the residual signals becoming comparable to the clutter level after which further integration is clearly detrimental.

Several techniques have been proposed for reducing the sensitivity of c.c.d. integrators to the effects of c.t.i.; they can conveniently be divided into two categories: those that attempt to overcome the problem (i.e. by preventing the residual charges from degrading the performance) and those that avoid the problem (i.e. by minimizing the number of charge transfers and hence the generation of residuals).

One approach is to sample the input signal on alternate clock cycles so that the data in the c.c.d. are separated by initially empty potential wells. As the integration proceeds, some charge will be lost from each 'data' well but will, to a first approximation, collect in the 'empty' wells. At the output the signal may be substantially restored by summing the charges in the 'data' wells with their respective 'residuals'.[16] The disadvantages of this scheme are that the c.c.d. must have twice the number of storage elements and must operate at twice the speed of a c.c.d. not employing this technique.

Another approach is to delay the output signal for a clock period and subtract a fraction of it from the following sample. By careful choice of the fraction, the residual can be partially cancelled.[16]

3 Overcoming C.T.I. Problems

A variation of this approach has been described by Cooper *et al.*[14] In this scheme the input to an n-bit c.c.d. is multiplied by a code which cycles through three values every $(n+1)$ clock cycles. By correctly decoding the output with a similar code, it is possible to cancel the residuals. Although this works reasonably well it has some disadvantages. The radar signal normally applied to the integrator is unipolar; however the code generates a bipolar signal which must be handled by the c.c.d. thus necessitating a reduction in dynamic range. Furthermore, the coding and decoding circuitry must be stable and all units must be d.c. coupled.

4 Avoiding C.T.I. Problems

In the conventional c.c.d. the signal undergoes many transfers that do not contribute directly to the processing function as such; thus in Fig. 1(a) and (b) for example, the transfers through the many stages of each c.c.d. are not in themselves vital—each c.c.d. is merely required to store analogue information for a finite time. These redundant transfers may be eliminated in several ways.

4.1 'Accumulator' Approach

Lobenstein's approach[17] is to use a serial c.c.d. to position the signals from the various range bins adjacent to a series of large potential wells or 'accumulators' (see Fig. 2). The charges associated with a particular radar return are allowed to equilibrate with the corresponding charges already in the accumulators, the charges remaining in the c.c.d. subsequently being transferred out. Considering the sequence of charge packets associated with any particular range bin it will be clear that the charge read out will be the exponentially weighted sum of all charges previously read in. The weighting factor is determined by the ratio of the area of

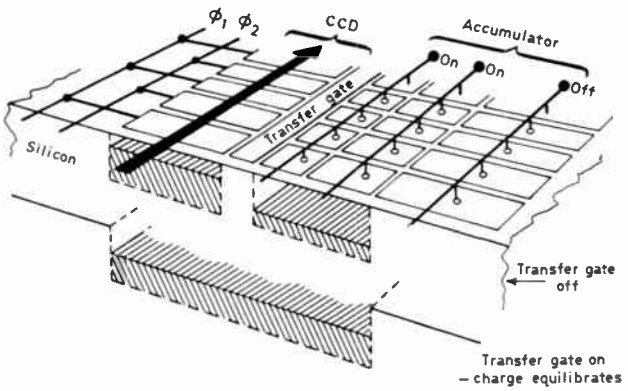


Fig. 2. Lobenstein's weighted integrator.

the accumulator cell to that of the c.c.d. wells; by constructing the accumulators using several adjacent electrodes, the effective accumulator areas can be altered as more or less electrodes are switched on. Lobenstein's approach clearly reduces c.t.i. problems substantially since any incoming signal undergoes only a single pass along the c.c.d. This rather elegant technique also overcomes the problem of feedback instability which is generally a problem associated with recursive integrators of the type depicted in Fig. 1(b). The disadvantage of this technique is that it takes a relatively long time for charge equilibration to occur in the accumulator and so operating speed is limited to less than 500 kHz.

4.2 Recursive Parallel Transfer Integrator

An alternative approach to eliminate redundant charge transfers on which we have been working is to use a new c.c.d. architecture based on parallel processing (see Fig. 3). This has two important advantages over Lobenstein's approach. Firstly, the residual charges due to c.t.i. are independent of the number of range bins; secondly, since no equilibration of charge over large areas is necessary, the full speed capability of the c.c.d. technique may be realized.

Referring to Fig. 3, the input range gate register (r.g.r.) is a digital shift register along which a single '1' is propagated. This pulse is used to gate each return

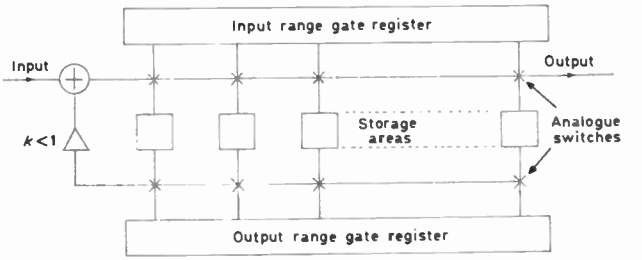


Fig. 3. The recursive parallel-transfer integrator.

sequentially into a series of c.c.d. storage areas (one for each range bin). The output r.g.r. (which can, in fact, be combined with the first) gates the stored samples on to an output bus; the signals are then recirculated in the conventional manner. Although this approach requires a larger chip area than the conventional serial device, it has the significant advantage that residuals are maintained at a very low level since each signal sample undergoes a minimum number of transfers (two in our case) per circulation. Furthermore, since each sample is processed individually, any residual that forms is confined to its 'parent' range bin and so does not 'smear' into other range bins. Clearly, the number of bits in this system can be increased indefinitely with no degradation in performance.

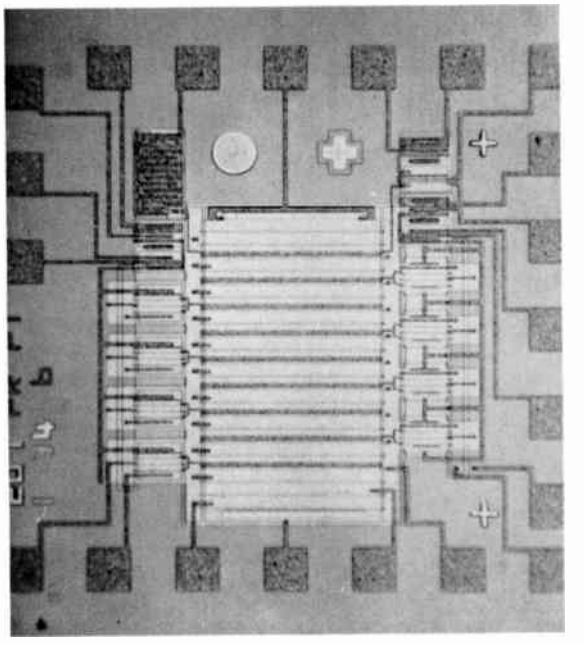
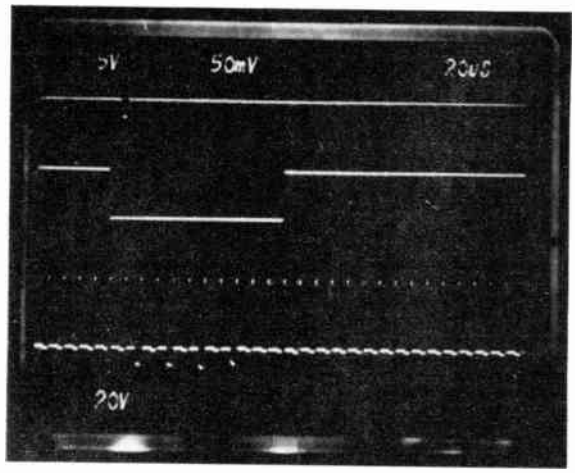
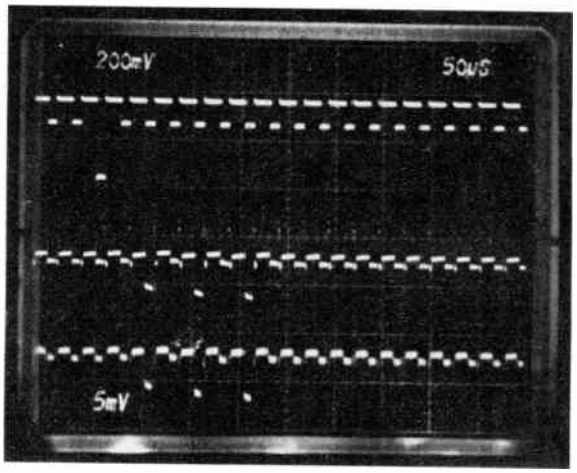


Fig. 4. Photomicrograph of chip implementing function of Fig. 3.

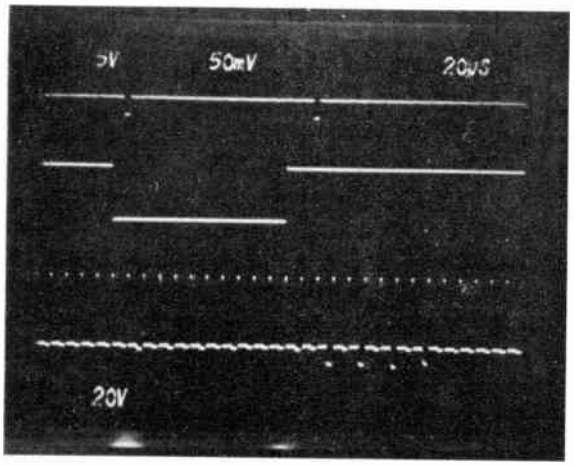
The r.g.r. itself has been implemented in the form of a floating-gate tapped c.c.d. fabricated on the same chip as the storage areas. The storage sites, are in effect, single-bit c.c.d.s that incorporate the input and output gating for each cell.[18, 19] The feasibility of the c.c.d. addressing technique has been demonstrated both theoretically and



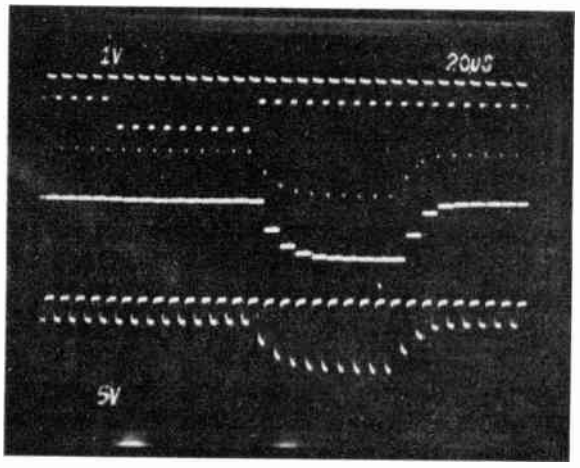
(a) An oscillograph showing the storage of a rectangular waveform in the upper set of storage sites.



(c) The recirculating digital storage capability of the integrator. The top trace is the r.g.r. input; 2·5 V/vert. div. The storage site input and output are shown respectively below. The middle trace is 200 mV/vert. div. and the bottom trace is 5 mV/vert. div.



(b) The storage and read-out of a rectangular waveform with two consecutive r.g.r. pulses. The scales are as shown in Fig. 5(a).



(d) The operation of the sample/hold circuit. The top trace is the r.g.r. input; 1 V/vert. div. The middle trace is the sample and held output and the bottom trace is the normal floating gate voltage. Both traces are 5 V/vert. div

Fig. 5.

experimentally[19] and a test structure comprising eight storage areas has been fabricated. Figure 4 is a photomicrograph of the chip which measures 1·5 mm × 1·5 mm; the storage areas are positioned on each side of the r.g.r. Also fabricated on the chip is a floating-gate sample/hold circuit to simplify the interfacing of r.g.r.s on several chips thereby enabling as many range bins as necessary to be assembled. An oscillogram of a device storing a digital signal is shown in Fig. 5(a); the upper trace is the address pulse applied to the r.g.r. input—this both stores the signal input (middle trace) and outputs the previously stored information (lower trace). Only the upper set of four storage areas was used for this photograph so only four output pulses are visible. Figure 5(b) demonstrates the ability of this device to give a variable delay, independent of data rate, simply by applying two address pulses per return to the r.g.r.: the lefthand pulse of the upper trace is used to store the waveform of the middle trace. As the righthand pulse traverses the r.g.r. the signal is read out (lower trace). Figure 5(c) shows the device operating as a recirculating digital shift register (only three pulses are visible since the fourth storage area had failed). The successful operation of the sample/hold circuit is shown in Fig. 5(d).

At the time of these experiments, devices fabricated for analogue operation were not available and so a recursive integrator could not be demonstrated. These preliminary experiments do, however, establish the feasibility of parallel-transfer c.c.d. structures for signal processing.

4.3 Non-recursive Parallel Transfer Integrator

Reference was made earlier to the fact that, under certain circumstances, non-recursive integration gave superior

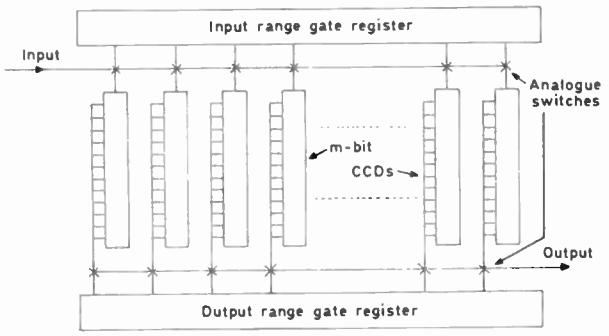


Fig. 6. The non-recursive parallel transfer integrator.

performance. We have extended the parallel transfer technique described above to enable a non-recursive integrator which is insensitive to c.t.i. effects to be built.[20] A schematic of the implementation is shown in Fig. 6. An r.g.r. is again used to address sequentially a series of storage areas; in this system, however, each storage area consists of an m-bit floating gate tapped c.c.d. Summation is performed using standard transversal filter techniques, each tap weight being unity in this case. Although c.t.i. effects are present, residuals associated with each c.c.d. are confined to that c.c.d. and are not 'smeared' between range bins. In common with the parallel recursive structure described above, this arrangement can be extended to as many range bins as required without degradation of performance. Clearly, a separate c.c.d. is required for each range bin; on the other hand the arrangement has the considerable advantage over most c.c.d. signal processors that commercially available devices can be used to realize the technique, obviating the considerable cost and time involved in producing a custom-designed chip. For example, it would be possible to construct this type of non-recursive integrator using a simple c.c.d. delay line for each range bin and implementing the r.g.r. with c.m.o.s. or TTL logic. One phase of each c.c.d. may be used to provide the integrated signal which may be detected with either voltage or current sensing techniques.[21] A suitable circuit is shown in Fig. 7; in this arrangement four-phase c.c.d.s driven from standard c.m.o.s. packages are used in conjunction with current sensing.

The feasibility of this approach has been demonstrated using a simulation requiring only one c.c.d.[20] In this demonstration a 50-bit two-phase c.c.d. was used to process a complete period of a signal by sweeping the periodic sample point in the waveform slowly with respect to the waveform period. Figure 8 shows the expected signal-to-noise improvement of 17 dB obtained by integrating 50 waveforms with this system.

Although in general only stationary or slowly moving targets may be processed with a signal integrator, a simple modification to the clocking sequence would allow effective signal processing on fast-moving platforms such as a hovercraft. For example, by adding or subtracting an appropriate number of clock pulses per sweep period, successive radar returns can be effectively shifted with respect to the individual c.c.d. range bins.

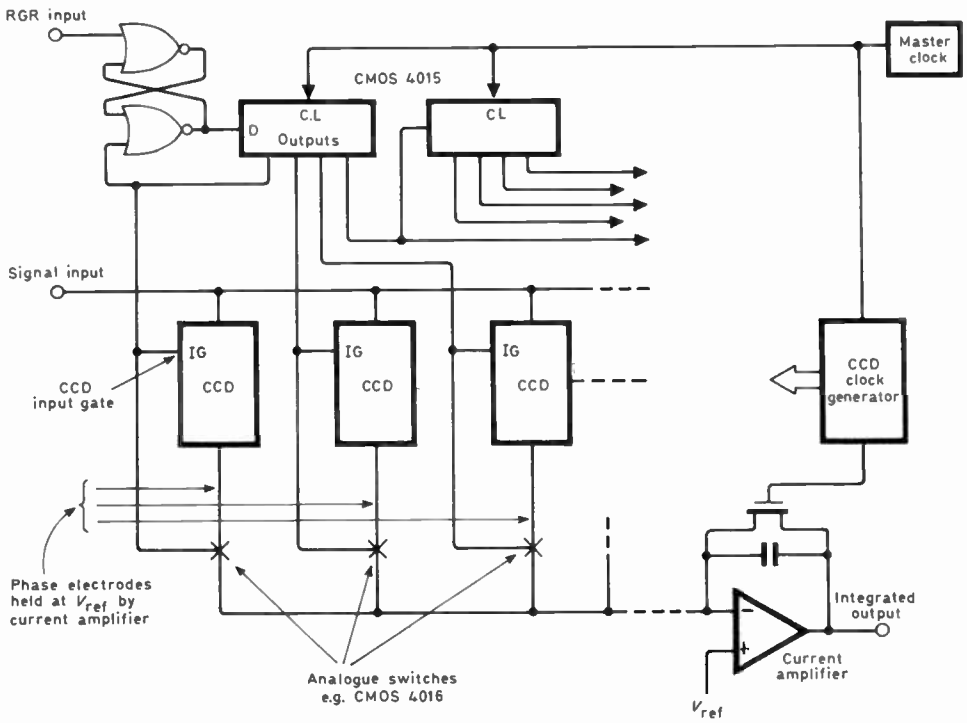


Fig. 7. A possible implementation of the non-recursive parallel transfer integrator using commercially available devices.

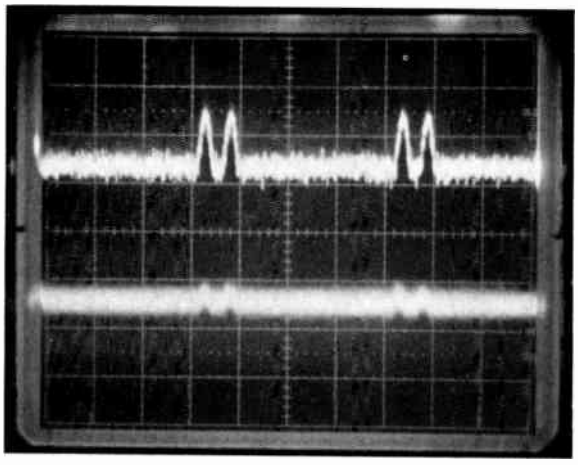


Fig. 8. The performance of a 50-bit integrator of the form shown in Fig. 6. The lower trace is the input wave form and the upper trace is the output; the latter exhibits the expected 17 dB improvement in signal-to-noise ratio. The horizontal scale is 5 μs/div.

5 Conclusions

We have shown how c.c.d.s can provide elegant solutions to integration techniques for improving signal-to-clutter ratios in radar systems. Charge transfer inefficiency effects rule out the most straightforward approaches but the alternative ways are still very much simpler than non-c.c.d. techniques, especially in regard to flexibility, compactness, etc. These general comments also apply to the use of c.c.d.s in other radar signal processing approaches. This should lend to an increasing use of such signal processing techniques on systems where cost, weight, etc. have ruled them out hitherto.

6 References

1 Rabiner, L. R., Schafer, R. W. and Rader, C. M., 'The chirp Z-transform algorithm', *IEEE Trans. Audio & Electroacoustics*, **AU-17**, pp. 86–92, 1969.

2 Mayer, G. J., 'The chirp Z-transform—a c.c.d. implementation', *RCA Review*, **36**, pp. 759–73, 1975.

3 Wardrop, B. and Bull, E., 'A discrete Fourier transform processor using charge-coupled devices', *Marconi Rev.*, **40**, pp. 1–41, 1977.

4 Brodersen, R. W., Hewes, C. R. and Buss, D. D., 'A 500 stage c.c.d. transversal filter for spectral analysis', *IEEE Trans. Electron Devices*, **ED-23**, pp. 143–52, 1976.

5 Eversole, W. C., Mayer, D. J., Bosshart, P. W., de Wit, M., Hewes, C. R. and Buss, D. D., 'A completely integrated thirty-two point chirp Z-transform', *IEEE J. Solid State Circuits*, **SC-13**, pp. 822–31, 1978.

6 Jack, M. A., Park, D. G. and Grant, P. M., 'C.c.d. spectrum analysis using prime transform algorithm', *Electronics Letters*, **13**, pp. 431–2, 1977.

7 Roder, C. M., 'Discrete Fourier transforms when the number of data samples is prime', *Proc. IEEE*, **56**, pp. 1107–8, 1968.

8 Bounden, J. G. and Tomlinson, M. J., 'A c.c.d. Chebychev filter for radar m.t.i. applications', *Electronics Letters*, **10**, pp. 89–90, 1974.

9 Herrmann, G. P., Gandolfo, D. A., Bommard, A. and Stepps, D. B., 'C.c.d. programmable correlator', Proc. Int. Conf. on Tech. and Applications of C.C.D.s, Edinburgh, 1976.

10 Tiemann, J. J., Baertsch, R. D. and Engeler, W. E., 'A surface charge correlator', *IEEE J. Solid State Circuits*, SC-9, pp. 403–10, 1974.

11 Lagnado, I., Steigerwalt, S., Boud, J., Speiser, J. and Whitehouse, H. W., 'A c.c.d. analogue and digital correlator', Proc. Int. Conf. on Tech. and Appl. of C.C.D.s, Edinburgh, 1976.

12 Harp, J. G., Vanstone, G. F., MacLennan, D. J. and Mavor, J., 'Analogue correlators using charge-coupled devices', Proc. Int. Conf. on Appl. of C.C.D.s, San Diego, pp. 229–36, 1975.

13 Beynon, J. D. E. and Traynar, C. P., 'Some considerations of signal/noise improvements using recursive and non-recursive integrators', *Electronics Letters*, **14**, pp. 554–6, 1978.

14 Cooper, D. C., Roberts, J. B. G., Akham, B. and Callan, R. D., 'The performance of a c.c.d. signal averager coded to reduce transfer inefficiency effects', Proc. Int. Conf. on Tech. and Appl. of C.C.D.s, Edinburgh, 1976.

15 Roberts, J. B. G., Chesswel, M. and Eames, R., 'Video integration using c.c.d.s', *Electronics Letters*, **10**, pp. 169–71, 1974.

16 Chowaniec, A. and Hobson, G. S., 'A time domain analysis of video integrators', Proc. Int. Conf. on the Appl. of C.C.D.s, San Diego, pp. 323–29, 1975.

17 Lobenstein, H., Quick, M., Roeder, A., Engeler, W. E., Tiemann, J. and Baertsch, R. D., 'C.c.d. accumulators in spread spectrum signal identification', Proc. Int. Conf. on Tech. and Appl. of C.C.D.s, Edinburgh, 1976.

18 Traynar, C. P., Roberts, P. C. T., Bloodworth, G. G. and Taylor, R. G., 'A novel c.c.d. signal integrator', *Electronics Letters*, **13**, pp. 456–8, 1977.

19 Traynar, C. P., Ph.D. Thesis, Southampton University, 1978.

20 Traynar, C. P. and Beynon, J. D. E., 'A non-recursive signal integrator using a parallel c.c.d. structure', *IEEE J. Solid-State Circuits*, **SC-14**, pp. 742–6, 1979.

21 Traynar, C. P., Beynon, J. D. E., Roberts, P. C. T. and Pennock, J., 'Discussion of depletion-capacitance induced distortion in surface channel c.c.d. transversed filters', *IEE J. Solid State & Electron Devices*, **1**, pp. 73–80, 1977.

Manuscript first received by the Institution on 9th August 1979 and in final form on 19th February 1980.
(Paper No. 1934/AMMS 100)

UDC 621.382:621.397.33

Indexing Terms: Semiconductor devices, Large-scale integration, Charge-coupled devices, Signal processing, Imaging

Advances in c.c.d. scanners with on-chip signal processing for electronic imaging

Professor SAVVAS G. CHAMBERLAIN, M.Sc., Ph.D.*

SUMMARY

The paper deals briefly with a few widely-used image signal processing algorithms and discusses how these can be incorporated on the same silicon chip as that of the c.c.d. scanner. Recent work on c.c.d. scanners is reviewed and solid-state scanners which include on-chip signal processing functions are described.

Future trends are towards 'smart' scanners; these are scanners with on-chip real-time processing functions, such as analogue-to-digital conversion, thresholding, data compaction, edge enhancement and other real-time image processing functions.

* *Electrical Engineering Department, University of Waterloo, Waterloo, Ontario, Canada.*

1 Introduction

C.c.d. imagers are a dominant subject of research and development in solid-state input technology today, as is indicated by published papers in this field.[1-11] Large c.c.d. imagers such as 512×320 (Ref. 12), 1728×1, 500×1, 256×1, 2048×1 (Refs. 13 and 14) and others are now commercially available. These devices are being used at the front end of optical readers, in facsimile machines, in surveillance and monitoring systems,[36] rapid-scan spectrometers, industrial measurement and inspection systems, in specialized military systems and in black and white cameras for industrial television.[34,35]

Recent developments in c.c.d. area arrays will not be discussed in the present paper since they are dealt with in the paper of D. J. Burt[40] which is given in this issue.

The trend in the new generation electronic imaging systems for business, satellite communications, and for advance facsimile machines is to have output data rates well in excess of 5 Mz. Most of the present electronic imaging systems employ a c.c.d. solid-state scanner at the front end; the analogue video output signal is then fed into another sub-system for thresholding, data compression, and for other necessary signal processing function. One way of achieving overall system high speed output data rates is to include real-time signal processing functions on the same silicon chip as the scanner. Such real-time, on-chip, signal processing functions could include an analogue-to-digital converter, thresholding, picture segmentation, picture edge enhancement, data compression and other signal processing functions. C.c.d. solid-state scanners which incorporate one or more of the above functions are being developed and are referred to in the literature as 'smart-scanners'.

In this paper we review the state of the art of c.c.d. solid-state scanners with on-chip signal processing functions. Before we deal with this topic, however, a brief discussion on signal processing with charge coupled devices which is compatible with c.c.d. scanners is given in Section 2. Also a brief description of some widely-used signal processing techniques is given in Section 3.

2 Signal Processing with Charge-coupled Devices

In charge-coupled devices the signal charge is stored on a gate capacitor in analogue form.[16] In c.c.d. shift registers when the signal is shifted its analogue form is faithfully maintained. Since the analogue shift register is run by digital clock pulses, a precise digital delay in addition to the memory is available along with the analogue signal. Such delay is very useful in the realization of transversal filters and other signal processing functions.[17] With the present state of the art, signal distortion due to charge transfer inefficiency is low since charge transfer inefficiencies of 10^{-5} to 10^{-4} are easily achieved.

C.c.d.s are well suited for low-light level imaging and signal processing functions because of their good signal-

The Radio and Electronic Engineer, Vol. 50, No. 5, pp. 249–257, May 1980

0033–7722/80/050249+09 $1.50/0

to-noise ratio. This good signal-to-noise ratio is an inherent property of c.c.d.s and arises because the signal charge is always stored on low capacitance nodes, thus the kTC noise is small compared to the incoming photon shot noise.[18]

Presently available c.c.d. imagers such as 1728×1 and 2048×1 (Ref. 13) where two c.c.d. shift registers are used for reading out the video signal, offer r.m.s. signal-to-noise ratios of greater than 500:1. In most facsimile applications a peak-to-peak signal-to-noise ratio of 160:1 is adequate; in such a case the image could have up to 60 grey levels. With such a large signal-to-noise ratio it is becoming feasible to compensate for the dark leakage current of the photoelement array and also compensate for the variations of the responsivity of the individual photoelements without sacrificing grey levels and hence picture quality.

When spatial resolution is considered, present c.c.d. scanners offer comparable resolution to some widely used lenses.[6,39]

2.1 The Split Electrode Transversal Filter

The split electrode transversal filter forms the basis of most signal processing c.c.d. circuits and units. The split electrode technique[19] provides an easy method to implement the multiply-and-add functions required in digital transversal filters. The principle of such split electrode techniques is that, as the signal charge is being transferred in the c.c.d. shift register channel under the gate electrode, in order to maintain overall charge neutrality an opposite charge must flow onto the electrode from the clock line. On one phase, the electrodes are all split into two sections of different sizes as shown in Fig. 1. This diagram shows ϕ_2 being split, but any other phase may be used. One side of each split electrode is connected to the ϕ_2^+ clock line and the other to the ϕ_2^- clock line. The two ϕ_2 lines are clocked simultaneously with the same clock phase. The signal dependent charge which flows into each portion of a split electrode is proportional to the area of that part of the electrode and to the signal charge flowing into the region under the electrode. By measuring the difference in the charge in the two sections of a split electrode, the non-destructive sampling and weighting operations are performed. Since the ϕ_2^+ electrodes are connected together, and the ϕ_2^- are also connected together, the summation occurs automatically, and the output signal at each clock period is proportional to the difference in current flowing in the two lines of the split phase. If the split in the kth electrode occurs exactly in the centre of the c.c.d. channel the tap weight $h_k = 0$. If the electrode is split at a position such that all of the charge flows into ϕ_2^+ or ϕ_2^- clock lines, the resulting weighting coefficient is $h_k = +1$ or $h_k = -1$. Intermediate values are possible, the resolution being limited by tolerances of photomasks and photolithography. The filter output is obtained by integrating the difference current which flows in the ϕ_2^+ and ϕ_2^- clock lines. In terms of the clock charge, V_{out} is proportional to

$$V_{out} \propto \sum_{k=1}^{m} \tfrac{1}{2}(1+h_k)Q_k^c - \sum_{k=1}^{m} \tfrac{1}{2}(1-h_k)Q_k^c$$

$$= \sum_{k=1}^{m} h_k Q_k^c$$

where Q_k^c is the clock line charge.

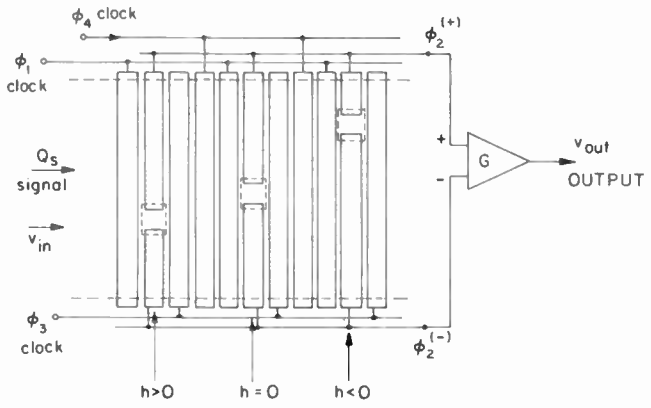


Fig. 1. Schematic diagram of analogue c.c.d. split-electrode transversal filter. Q_s is the signal charge and h represents the tap weight coefficient.

Experimental performance and limitations of these c.c.d. transversal filters are covered extensively in the literature.[20]

By using the basic transversal filters other complex arithmetic functions, such as the chirp-Z transform (CZT) algorithm,[21] for computing the power spectral density of signals have been realized and reported in the literature.[20] Other more sophisticated arithmetic units and digital-to-analogue converters can also be realized with c.c.d.s.[17]

3 Algorithms for Image Signal Processing

Modern satellite and other communication systems include a sending imaging unit, a communications data channel, and a receiving data system. Most of these systems incorporate algorithms in order to improve the received picture. Further, some of these systems are capable of recognizing patterns in an image. The algorithms necessary to do all the above are basically based on the application of the Laplacian linear derivative operator. Data compression is also included with present electronic systems.

In this Section we discuss a few of the widely used algorithms which can be incorporated (integrated) on the silicon scanner chip.

3.1 Relevance of the Laplacian to Image Processing

The Laplacian is the linear derivative operator

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

where f is the picture function and x and y are spatial

variables. This operator is rotation invariant, i.e. rotating f and then applying the operator gives the same result as applying the operator to f and then rotating the output. In order to see the usefulness of this operator consider the following examples.[22,23]

3.2 Sharpening of the Picture

Picture degradation generally involves blurring, the extent of which is described by the spread functions of the degradation operation. Simple methods can be used for counteracting blur by 'crispening' or 'sharpening' the picture.

Blurring is an averaging, or integration operation; this suggests that we may be able to sharpen by performing differentiation operations. Blurring also weakens high spatial frequencies more than low ones; this suggests that pictures can be sharpened by emphasizing their high spatial frequencies. Differentiation operations such as the 'Laplacian' can be used to sharpen the picture.

When a picture is noisy as well as blurred, differentiation and high emphasis filtering cannot be used indiscriminately to sharpen it, since the noise generally involves high rates of change of grey level, and it usually becomes stronger than the picture signal at high frequencies. Sharpening methods should be restricted if possible to frequency ranges where the picture is stronger than the noise. Alternatively, one should attempt to remove or reduce the noise before attempting to sharpen the picture.

Let a blur in a picture be the result of a diffusion process that satisfies the partial differential equation:

$$\frac{\partial g}{\partial t} = k\nabla^2 g$$

where g is the blurred picture and is a function of x, y and t. At $t = 0$, $g(x, y, 0)$ is the unblurred picture $f(x, y)$. k is a constant and greater than zero.

At τ ($\tau > 0$) we have the blurred picture $g(x, y, \tau)$. If we expand $g(x, y, t)$ in a Taylor series around $t = \tau$ we have

$$g(x,y,0) = g(x,y,\tau) - \tau\frac{\partial g}{\partial t}(x,y,\tau) + \frac{\tau^2}{2}\frac{\partial^2 g}{\partial t^2}(x,y,\tau) + \ldots$$

Neglecting the quadratic and higher-order terms and substituting f for $g(x, y, 0)$ and $k\nabla^2 g$ for $\partial g/\partial t$ we get

$$f = g - k\tau\nabla^2 g.$$

Hence to a first approximation we can restore the unblurred picture f by subtracting from g a positive multiple of its Laplacian.

If the blur is due to a diffusion process the brightness of the spot will follow a Gaussian distribution. The variance of this distribution is proportional to $k\tau$. Thus $k\tau$ can be estimated by fitting a Gaussian to the point spread function.

3.3 Edge Extraction using Linear Operations

Discontinuities in an image where there is a more or less abrupt change in grey levels indicating the end of one region and the beginning of another are called edges.

The gradient and the Laplacian operators are often used in extracting edges in an image. With digital computing, these operators are approximated by finite differences. Operators built by finite differences cannot be invariant under rotation, though they can be approximated arbitrarily closely. The squared gradient operator may be approximated by

$$|\nabla f|^2 \simeq (f(i+1,j+1)-f(i,j))^2 + (f(i+1,j)-f(i,j+1))^2$$

and the Laplacian by

$$\nabla^2 f \simeq f(i+1,j)+f(i,j+1)+f(i-1,j)+f(i,j-1)-4f(i,j).$$

These forms have a high degree of symmetry.

It is possible to transform differential operators to discrete operators by starting with the definition of a derivative given by

$$\nabla_x f(i,j) = f_x \equiv f(i+1,j)-f(i,j)$$

$$\nabla_y f(i,j) = f_y \equiv f(i,j+1)-f(i,j)$$

and to obtain higher-order derivatives by applying this definition recursively. The operators obtained in this way tend to be not as symmetric as those defined earlier.

To extract edges, we set a threshold and call any point an edge point if the squared gradient (or the absolute value of the Laplacian) at that point exceeds the threshold.

3.4 Signal Processing Technique for Dark Noise and Switching Noise Compensation

One method which compensates for the dark response and switching noise associated with the photoelement of the imaging device is the 'focus/defocus' algorithm.[24] As can be seen from the following outline, this algorithm can easily be implemented and realized on the same chip as that of the scanner, for real-time signal processing. This algorithm utilizes a focused and a defocused image. It can best be understood if we let the image be viewed through a haze or fog or through a frosted glass and finally detect it by a scanner; the latter process will result in a defocused image. Let

G_i be the responsivity of the ith element of the scanner,
D_i be the dark signal generated by the ith element.

In practical scanners the responsivity can vary by as much as 10%. This variation can be represented by Δ, hence

$$G_i' = 1 + \Delta G_i.$$

Let H represent the uniform brightness of the haze, i.e. this will represent the average brightness of the defocused

image. Also let A_i be the brightness variation attributable to the part of the object imaged on the ith element. If the focused and defocused response is represented by R_F and R_D respectively then we have

$$R_F = (H + A_i)G_i' + D_i$$

therefore

$$R_F = (H + A_i)(1 + \Delta G_i) + D_i$$
$$= A_i(1 + \Delta G_i) + H(1 + \Delta G_i) + D_i.$$

If now the image is defocused, the brightness variation attributable to the part of the object imaged on the element i becomes zero and

$$R_D = H(1 + \Delta G_i) + D_i.$$

If we then take the difference between the focused, and defocused responses we get

$$R_F - R_D = A_i(1 + \Delta G_i) + H(1 + \Delta G_i) + D_i - [H(1 + \Delta G_i) + D_i]$$
$$= A_i(1 + \Delta G_i).$$

Thus the difference between the focused and defocused response is approximately equal to the brightness variation attributable to the part of the object imaged on the element i. It can be seen that the additive dark response of the sensor has been removed by the subtraction process. Thus this algorithm automatically compensates each element of the sensor for its dark response and for any switching noise that may be associated with that photoelement.

One drawback which is inherent in this scheme is the analogue memory which is necessary to store the focused and defocused images and then take their difference.

The traditional method of obtaining the focused and defocused images is to use a mechanically focusing and defocusing optical system. However, with the development of c.c.d. imagers it is feasible to obtain the defocused image electronically by averaging a number of neighbouring photoelements together. Such a technique eliminates the need for a mechanical defocusing system, and it can easily be implemented with c.c.d. area arrays.

3.5 Removal of Shading Effects from an Image

Uneven object illumination produces shading effects in the image. In many instances we would want to remove such shading effects from the image without degrading the detail in the picture. It is known that the illumination incident on an object usually varies more slowly than does the object's reflectance. Thus uneven illumination has more significant effects at low than at high spatial frequencies.

One way to sharpen ('crispen') the image is to pass the image through a high-pass filter and thereby emphasize the high-frequency components of the image. This of course can only be done with imagers where the signal-to-noise ratio is high, otherwise the image will deteriorate further.

It has been shown experimentally by Stockham[25] that an image can be crispened much better if one compresses the dynamic range of the image logarithmically before the high-pass filter and then expands it exponentially afterwards. This can be explained by noting that the brightness of an image is approximately the product of the illumination $I(x, y)$ and the reflectance of the object $R(x, y)$. Taking the logarithm makes the two factors additive

$$\log IR = \log I + \log R.$$

Therefore, the details of the object can be crispened more or less independently of the illumination; since the Fourier transform is a linear operation it is possible to de-emphasize selectively the effects of $I(x, y)$ by high-emphasis filtering.

With the present-day technology the above algorithm can easily be implemented on-chip with c.c.d. solid-state image scanners; such on-chip signal processing hardware can be externally programmable and used at will whenever it is necessary.

3.6 Transformational Coding

A digitized image may be characterized by a sequence of messages. The messages can be, for example, the brightness levels of each individual sample. Each message may contain the brightness levels of a pair of neighbouring samples. Another example is that the messages may be first differences of adjacent samples along each horizontal line. There are many ways to choose the messages, the only requirement being that we should be able to reconstruct the digitized image from the sequence of messages.

3.7 Hadamard Transform Coding

To understand the Hadamard transform, consider a sampled image which is divided into 2×2 blocks. Let the intensities of the four samples in the block be represented by x_1, x_2, x_3, and x_4. These intensities are transformed into y_i by a 4×4 Hadamard matrix:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}.$$

Since in a typical image, neighbouring samples tend to have equal intensities, $y_i (i \neq 1)$ tend to be very small. Thus, when digitally representing y_i, more bits may be used for y_1 and fewer bits for y_2, y_3, and y_4, expecting that we may end up with a small average number of bits per sample and yet get a good quality reconstructed image. Quantification of images using the Hadamard matrices was studied by Huang and Woods.[33] This scheme was applied to several images and various block sizes were tried. It was found that, for a given average bit rate, the use of a large block size tended to make the

degradation in the reconstructed image appear as random noise, while the use of a small block size made the degradation appear in the form of discontinuities at block boundaries.[23] Huang *et al.*[23] found that with three bits per sample, the picture quality becomes as good as that of 6-bit originals.

For implementation of image coding schemes in real time, it is easier to work along a scanline rather than in two dimensions. The Hadamard scheme can also be used with one-dimensional blocks. Huang *et al.*[23] using a block size of 1×256 and 3 bits, obtained a picture which was as good as a 3-bit picture using 16×16 blocks, but the one-dimensional block 2-bit and 1-bit pictures had inferior quality when compared to the corresponding pictures using 16×16 blocks.

In the next Section we review solid-state c.c.d. silicon scanners which include on-chip signal processing functions.

4 Examples of Realized C.C.D. Imagers with On-chip Signal Processing: 'Smart Scanners'

Over the past three years silicon solid-state-scanners with on-chip signal processing capabilities have been reported. These include scanners with electronically-variable photoelement area, a scanner with neighbourhood Laplacian operators, time-delay-and-integration imagers, a Hadamard transform coding imager, imager with multiple-gate photoelement and others. In this Section we review the properties of these devices.

4.1 Imager with Electronically-variable Photoelement Area

Recently, Blouke, Hall and Beitzmann[15] reported an 800×800 element c.c.d. silicon imager built for the Jet Propulsion Laboratory for use in the Galileo and Space Telescope programs. This imager was intended to operate from $-100°C$ to room temperature. It was fabricated with a three-level polysilicon gate technology on a 10 Ω-cm *p*-type silicon substrate and had an ion-implanted *n*-type buried layer. Its photoelement size was $15\ \mu m \times 15\ \mu m$ with an imaging area of $1{\cdot}4\ cm^2$; the overall silicon chip dimension was $17{\cdot}8\ mm \times 17{\cdot}8\ mm$. This device was backside illuminated; to improve the spectral response at short visible wavelengths the substrate was thinned in the active imaging area down to $8{\cdot}3\ \mu m$.

This device included new features aimed towards video signal processing. A serial-parallel-serial configuration was used, with the 800×800 array divided into four sections, each 800×200. An integration structure prior to the output gate was incorporated on chip. Summation of the video charge prior to the read-out varied electronically the effective photoelement area up to four times, thus combining at will the 2×2 contiguous photoelements. A dual-gate voltage-sensing amplifier helped to minimize resetting noise.

The device was operated successfully at room temperature with an output video rate of 1·25 MHz. It required 0·545 seconds to read out all the video data. Its dark leakage current at 300 K was $6\ nA/cm^2$, which is typical for silicon c.c.d. imaging technology.

4.2 Imager with Neighbourhood Laplacian Operators

A 324×340 c.c.d. area imager has been designed and fabricated to provide nine simultaneous analogue video outputs which represent a 3×3 pixels element that scans the imaging array during a read-out of the device.[27] A schematic showing the organization of this scanner is shown in Fig. 2. The purpose of this array is to facilitate the use of 3×3 'neighbourhood operators' to perform various image processing functions. Specifically, this imager was intended for applications in areas of defect cancellation, edge detection and enhancement including point source location by application of centroiding algorithms.

In the usual processing applications of such neighbourhood operators, video data from the individual pixel is converted to digital information and stored in a memory. This image with nine individual outputs allows such an operation to be performed during the image-read-out period with no appreciable time penalty.

Three-level polysilicon-gate buried-channel technology was used to realize this imager. The photoelement size was $25 \times 25\ \mu m$ with an overall chip size of $1{\cdot}27 \times 1{\cdot}27$ cm. There were nine individual video

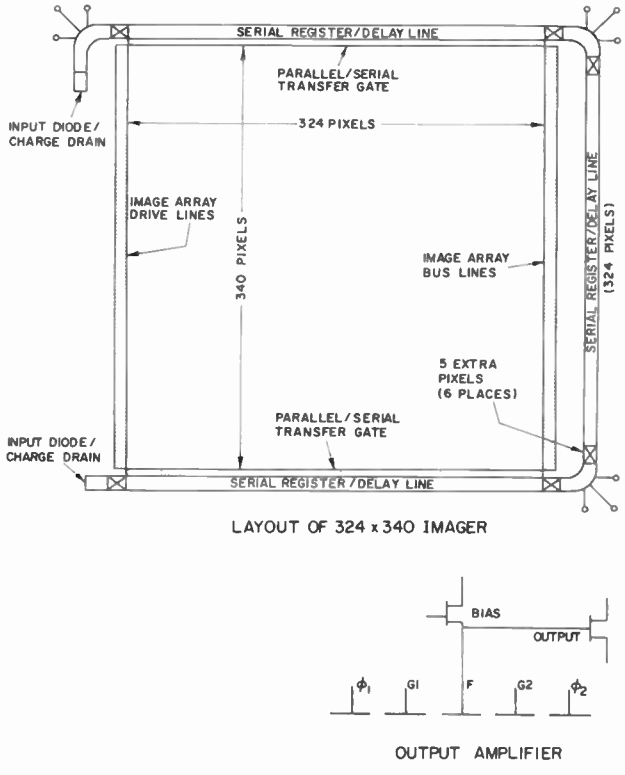


Fig. 2. Schematic of the organization of the 324×340 photoelement array and circuit diagram of the floating gate amplifier (from Ref. 27).

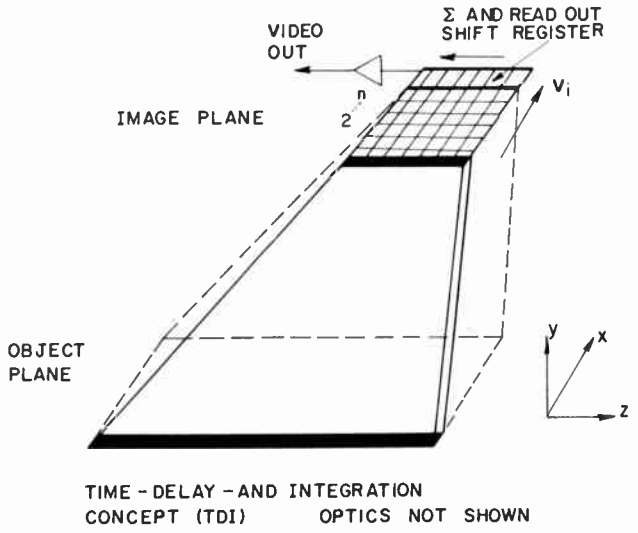
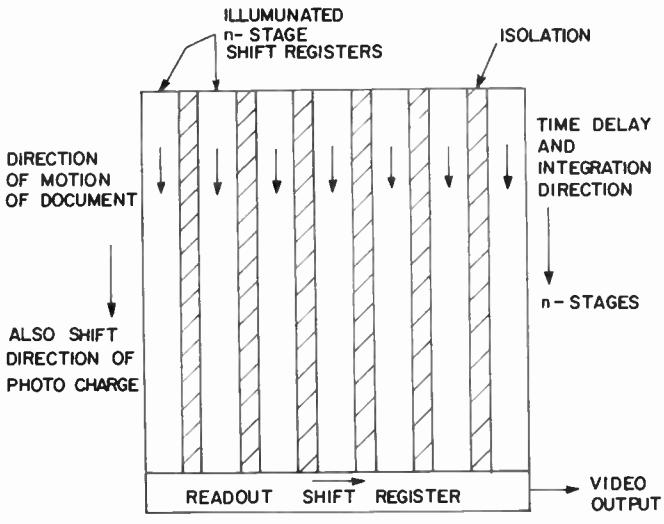


Fig. 3. (a) Schematic illustrating the principle of the time-delay-and-integration concept. Optics are not shown.



(b). Schematic of the time-delay-and-integration scanner. The illuminated shift registers function as linear imaging arrays.

channel outputs, and the sensing was achieved by nine individual floating gate amplifiers.

This device has novel properties which enable a user to implement signal processing such as picture sharpening and other signal processing functions which involve a Laplacian Operator. The floating-gate amplifiers, apart from sensing the video charge non-destructively, offer low noise. The disadvantage of this imager is that it utilizes three-level polysilicon, which is a complex technology. However, the same device can be realized with a two-level polysilicon gate technology. A natural addition to this device would be an analogue-to-digital converter attached to each video output.

4.3 Real-time Image Transform Coding with Charge Transfer Devices

Recently Michon, Burke, Vogelsong and Merola[26] developed a charge-injection[32] solid-state video sensor device capable of producing in real time a Hadamard transform of the incident optical image. This 128×128 photoelement array with on-chip signal processing was used as a Hadamard focal plane processor for image bandwidth compression. In one mode of operation, four on-chip, real-time 1×4 Hadamard transforms in parallel were obtained. Minimal off-chip hardware was required to produce a 4×4 real-time two-dimensional transform. These authors reported that the performance of the device was verified by computing the two-dimensional transform of various images, performing the inverse transform and displaying the resultant images.

Such video signal processing and real-time transform coding of an optical image directly on charge transfer video sensor chips is an important first step toward the realization of miniature, low power, bandwidth compression systems suitable for many applications.

4.4 Time Delay and Integration Imager

One method for improving the low-light level sensitivity of an area or a line imager is to use the time delay and integration principle (t.d.i.).[18] In the case of a linear line imager, instead of a horizontal single line of photoelements looking at a line on the object, a large vertical number (n) of photoelement lines are used to detect the single line on the object (Fig. 3). To achieve this, the document speed is synchronized with the clock speed of the video charge which is being shifted in the time delay and integration direction (Fig. 3(a)). In this way the n charge packets are summed together in the final horizontal read-out shift register; this read-out shift register contains the video information of a single horizontal line of the document. In such a scheme the signal photocharge increases proportionally to the number of delay stages (n) while the noise is proportional to $\sqrt{n}$; thus the t.d.i. is good for low-light level facsimile imaging applications. In the above example we considered the detection of a single line from the object; however, when the relative speed of an area to be imaged is synchronized with the clock speed of the photocharge of the t.d.i. in Fig. 3(b), the video output would correspond to the image of the area.

For high-resolution, low-light level, high-speed facsimile applications the time-delay-and-integration principle offers a number of potential advantages over the conventional integration mode.

Angle, Carnes, Kosonocky and Sauer have fabricated a 748×96 element t.d.i. chip aimed at electronic message systems applications.[28] The requirements of such an electronic message service system include the reading of an $8{\cdot}5 \times 11$ in ($21{\cdot}6 \times 28$ cm) page size at a rate of 20 pages per second at a paper subcombustion level of illumination with an object resolution of 0·125 mm in

each direction. For 8·5 × 11 in paper size this resolution requirement corresponds to a 2200 photoelement linear array. To achieve the required low-level illumination, 96 photoelements are used in the t.d.i. direction; thus a 2200 × 96 area scanner operated in the t.d.i. mode is needed. Further, to satisfy the above objectives, such imagers should have an output video rate of 84 MHz with a dynamic range of one hundred to one.

As a test vehicle Angle *et al.*[28] designed and fabricated a 748 × 96 photoelement array which was operated in the t.d.i. mode. On the chip, a one-horizontal-line-storage was also incorporated. The photoelement size was 15 μm × 15 μm which satisfied the demanded spatial resolution. A two-phase technology with two-level polysilicon and an electrode-per-bit[10] in the t.d.i. direction was used. Four-way multiplexing enabled them to achieve an 84 MHz output video rate. This multiplexing, however, made the layout of the experimental array very complex. In applications where the video output rate is required to be so high, it could be better to use the profiled bulk channel c.c.d. technology;[29,37,38] this technology offers c.c.d. shift registers with data rates in excess of 100 MHz.

Another time-delay-and-integration imager was designed and tested by Thompson *et al.*[5] This c.c.d. area array consisted of 20 channels each of nine t.d.i. stages. A read-out shift register terminated both ends of each channel so that by manipulation of the clock phases that make up the t.d.i. stages, the array could be made to scan forward or reverse. The channels were further subdivided such that ten were read out by each of the two registers on either side of the array. The nine t.d.i. stages were subdivided into three subsections of three cells each such that three, six, or nine stages could be clocked into the output register. Such a scheme provided electronic photosensitivity control.

Four-phase buried channel construction was used with front surface imaging through the visible transparent tin oxide gates. A 65% quantum efficiency with smooth spectral response was achieved. Each photoelement of the imaging section was 76 μm × 76 μm. It was formed by four 19 μm electrodes of transparent tin oxide of 400 Ω/square and 1500 Å thick layers. The c.c.d. channels were separated by ion implanted channel stop regions.

4.5 Multiple-gate C.C.D.-photodiode Sensor Element for Imaging Arrays

Polysilicon photogates absorb and attenuate short wavelength visible light, thus front illuminated imagers with polysilicon photogates offer poor blue spectral response.[31] One way of improving the blue response of linear line imagers is to use $n^+ - p$ photodiodes as the light sensors scanned by a c.c.d. read-out shift register. However, scanners with photodiodes such as bucket-brigade imagers can suffer from incomplete photocharge transfer between the photoelement and the read-out shift register. White and Chamberlain[3] have reported results on an experimental line imager which overcomes this problem. Their device also incorporates other novel features: the addition of barrier and storage gates between the photodiodes and transfer gates of a c.c.d. imaging array provide improved and versatile operating characteristics compared to the available silicon scanners.

In order to realize the full advantages of the multi-gate structure, a slight addition to the conventional clocking sequence is required. Figure 4 shows the timing diagram for the clock waveforms with cross-sectional diagrams of the surface potential at selected times. The photocharges are collected by the diffused photodiode, which is prevented from being strongly reverse-biased since the adjacent polysilicon barrier gate (BAR) never has a large voltage applied. The collected charges spill over the barrier during integration and are stored under the next polysilicon electrode (STORE) as indicated in Fig. 4(a). When the desired exposure is reached, integration is stopped by decreasing the barrier gate voltage so that the subsequently collected photocharge is taken away via the blooming gate (BL.G.) and the sink diffusion (DIFF) as shown in Fig. 4(b). Meanwhile, charges integrated during the previous exposure cycle are being moved along the shift-register channel as indicated by the dashed potential wells containing various amounts of charge.

Eventually, the charges in the shift register are all clocked to the output at the end of the array, and the shift-register wells are empty (Fig. 4(c)). With the appropriate shift-register phase voltage (S.R.) high, the transfer gate (XFER) is enabled and charge begins to move from the storage site into the shift-register channel (Fig. 4(d)). Transfer is completed by squeezing the charge in sequence from under the storage and transfer electrodes into the shift-register (Fig. 4(e)).

An empty potential well is recreated under the storage electrode (Fig. 4 (f)) and the new sequence of charges in the shift-register begin their transfer to the output. Integration of the next sequence of charges begins as soon as the barrier voltage is applied again.

The blooming gate and diffusion can be used to add a fat-zero to the integrated charge in the storage gate. Alternatively, the blooming gate voltage can be readjusted to prevent the photodiode surface potential from changing while the barrier voltage is low.

The experimental scanner, made with multiple-gate photoelements, has new functional capabilities resulting from the on-chip structure. These include exposure control in real time, which can compensate for temporal illumination variations during the integration cycle, linearization of the output signal with respect to light intensity, and adaptive level setting to normalize the output based on the whitest portion of the image. An

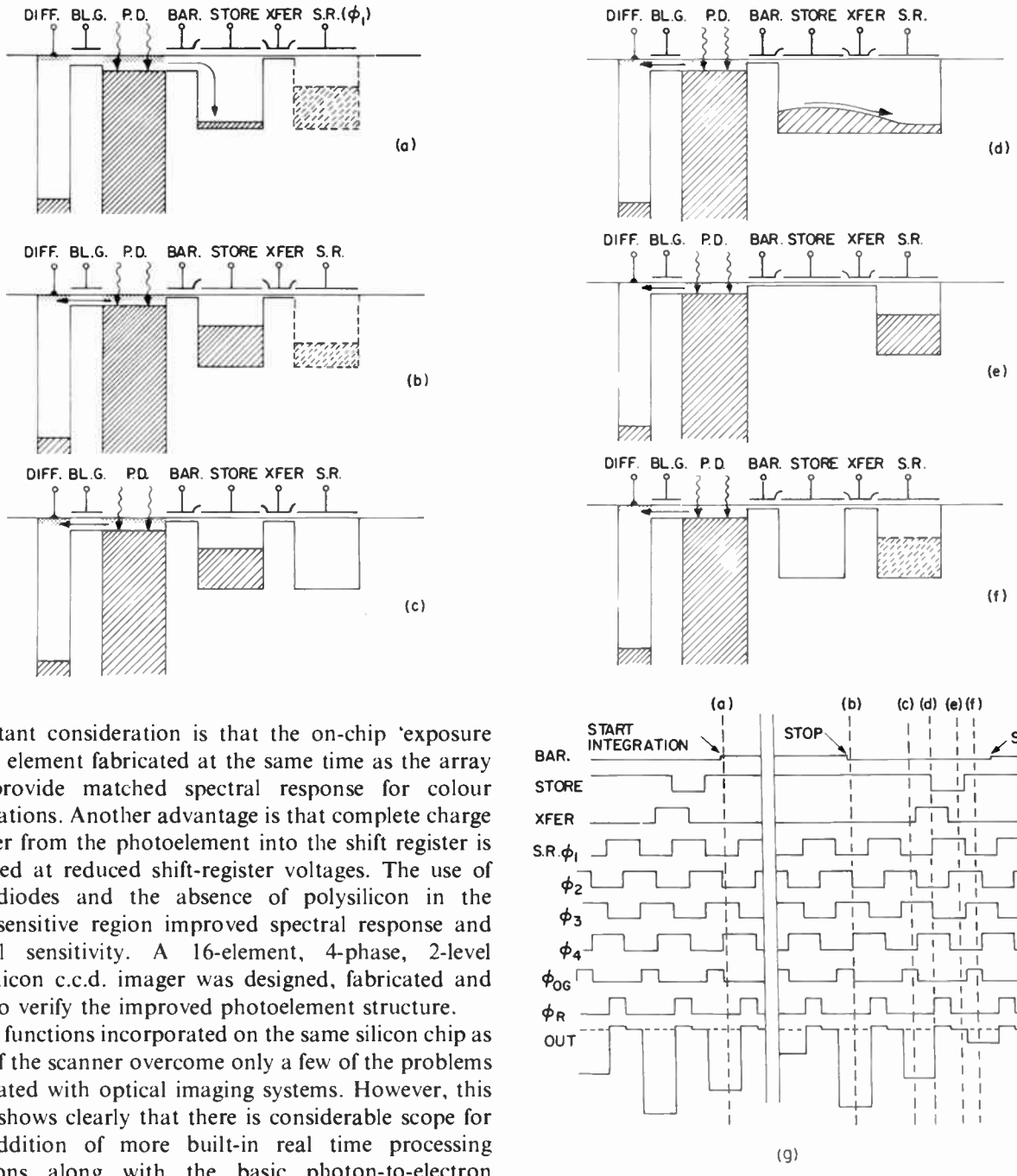


Fig. 4. Cross-sectional potential well diagrams with clocking sequence. For most experiments a 1 μs rise and fall time was used on the clocks. Dotted areas at the surface indicate region of n^+ diffusions.

important consideration is that the on-chip 'exposure meter' element fabricated at the same time as the array can provide matched spectral response for colour applications. Another advantage is that complete charge transfer from the photoelement into the shift register is achieved at reduced shift-register voltages. The use of photodiodes and the absence of polysilicon in the photosensitive region improved spectral response and overall sensitivity. A 16-element, 4-phase, 2-level polysilicon c.c.d. imager was designed, fabricated and used to verify the improved photoelement structure.

The functions incorporated on the same silicon chip as part of the scanner overcome only a few of the problems associated with optical imaging systems. However, this work shows clearly that there is considerable scope for the addition of more built-in real time processing functions along with the basic photon-to-electron transducer requirement of the imager scanner. This developed scheme can also be used for colour scanners.[30]

5 Conclusions and Projections

In this paper we have discussed briefly some widely used image signal processing algorithms and how these can be incorporated on the same silicon chip as the scanner. We have reviewed recent important work on c.c.d. scanners and have described some scanners which presently include on-chip image signal processing functions.

From the present review we can see that the future trends are towards 'smart' scanners, i.e. silicon scanners with various on-chip signal processing functions, such as analogue-to-digital-conversion, digital memory, Laplacian operator-based signal processing algorithms, edge detection, thresholding, image segmentation, image contrast enhancement algorithms and other image signal processing functions.

6 Acknowledgments

This work was supported in part from a research grant of the Natural Sciences and Engineering Research Council

of Canada, Grant No. A7384.

I would like to thank Dr. Robert A. Myers of the IBM Thomas J. Watson Research Center and Dr. James M. White of IBM, Charlotte, for reading the manuscript and suggesting improvements. Discussion on the subject of c.c.d. scanners with Gene Schlig and Keith Pennington of IBM, Thomas J. Watson Research Center, Yorktown Heights, were also very helpful.

7 References

1 Dillon, P. L. P., Lewis, D. M. and Kaspar, F. G., 'Colour imaging system using a single ccd area array', *IEEE J. Solid-State Circuits*, **SC-13**, pp. 28–33, February 1978.†
2 Dillon, P. L. P., Brault, A. T., Horak, J. R., Garcia, E., Martin, T. W. and Light, W. A, 'Fabrication and performance of colour filter arrays for solid-state imagers', *IEEE Trans. Electron Devices*, **ED-25**, pp. 97–101, February 1978.†
3 White, J. M. and Chamberlain, S. G., 'A multiple-gate ccd-photodiode sensor element for imaging arrays', *IEEE Trans. Electron Devices*, **ED-25**, pp. 145–53, February 1978.
4 Anagnostopoulos, C., 'Signal readout in cid image sensors', *IEEE Trans. Electron Devices*, **ED-25**, pp. 85–9, February 1978.
5 Thompson, L. L., McCann, D. H., Tracy, R. A., Kub, F. J. and White, M. H., 'Time-delay-and-integration charge-coupled devices using tin oxide gate technology', *IEEE Trans. Electron Devices*, **ED-25**, pp. 132–4, February 1978.
6 Chamberlain, S. G. and Harper, D. H., 'MTF simulation including transmittance effects and experimental results of charge-coupled imagers', *IEEE Trans. Electron Devices*, **ED-25**, pp. 145–54, February 1978.
7 Ochi, S., Yamanaka, S., Yamanaka, Y., Kanoh, Y. and Nishimura, T., 'A device structure and spatial spectrum for checker-pattern CCD colour camera', *IEEE J. Solid-state Circuits*, **SC-13**, pp. 187–92, February 1978.
8 Brown, D. M., Ghezzo, M. and Sargent, P. L., 'High density imagers', *IEEE J. Solid-State Circuits*, **SC-13**, pp. 5–10, February 1978.
9 Chamberlain, S. G., Roulston, D. J. and Desai, S. P., 'Spectral response limitation mechanisms of a shallow junction $n^+ - p$ photodiode', *IEEE Trans. Electron Devices*, **ED-25**, pp. 241–6, February 1978.
10 Heyns, H and Santen, J. G. van, 'The resistive gate ctd area image sensor', *IEEE Trans. Electron Devices*, **ED-25**, pp. 135–9, February 1978.
11 Sei, H., Miyamoto, Y., Arakawa, I., Minra, S., Sakai, H. and Ohtsuki, O., 'A meander channel ccd linear image sensor', *IEEE Trans. Electron Devices*, **ED-25**, pp. 140 4, February 1978.
12 Rodgers, R. L., 'A 512 × 320 element silicon imaging device', in IEEE International Solid-State Circuits Conference Digest, pp. 188–9, February 1975.
13 Fairchild Technical data sheet, 'CCD 122/142, 1728/2048 element linear image sensor charge couple device', Mountain View, CA. April 1979.
14 Reticon product summary 'Solid-state image sensors, 64 to 1872 elements', RC 1728H, RC 1024H, Reticon Corp., Sunnyvale, CA, 1976.
15 Blouke, M. M., Hall, J. E. and Beitzman, J. F., '640 kilopixel ccd imager for space applications', IEDM Technical Digest, Washington, D.C., December 1978.
16 Sequin, C. H. and Tompsett, M. F., 'Charge Transfer Devices' (Academic Press, New York, 1975).
17 Barbe, D. F., Baker, W. D. and Davis, K. L., 'Signal processing with charge-coupled devices'. *IEEE J. Solid-State Circuits*, **SC-13**, pp. 34–51, February 1978.
18 Barbe, D. F., 'Imaging devices using the charge-coupled concept', *Proc. IEEE*, **63**, pp. 38–67, January 1975.
19 Buss, D. D., Collins, D. R., Bailey, W. H. and Reeves, C. R., 'Transversal filtering using charge-transfer devices'. *IEEE J. Solid-State Circuits*, **SC-8**, pp. 138–46, 1973.
20 Brodersen, R. W., Hewes, C. R. and Buss, D. D., 'A 500-stage ccd transversal filter for spectral analysis', *IEEE J. Solid-State Circuits*, **SC-11**, pp. 75–84, 1976.
21 Rabiner, L. R., Schafer, R. W. and Rader, C. M., 'The chirp-z transform algorithm', IEEE Trans. Audio and Electroacoustics, **AU-7**, pp. 86–92, 1969.
22 Rosenfeld, A. and Kak, A. C., 'Digital Picture Processing' (Academic Press, New York, 1976).
23 Huang, T. S., Schreiber, W. F. and Tretiak, O. J., 'Image processing', *Proc. IEEE*, **59**, pp. 1586–1609, November 1971.
24 Meier, M. J., Adams, E. C. and Lindberg, R. W., 'Real time image processor', U.S. patent number 4,065,788, December 27th, 1977.
25 Stockman, T. G., in 1966 Spring Joint Computer Conf., AFIPS Conf., Proc., Vol. 28.
26 Michon, G. J., Burke, H. K., Vogelsong, T. L. and Merola, P. A., 'Charge injection device (cid) Hadamard focal plane processor for image bandwidth compression', Proc. 34th Tech. Meeting of the Avionics Panel on Impact CCD's and SAW's on Signal Processing and Imagery in Advanced Systems, Ottawa, Ont., Canada, October 10th–14th, 1977.
27 Hall, J. E., Breitzmann, J. F., Blouke, M. M. and Carlo, J. T., 'A multiple output ccd imager for image processing applications', IEDM Technical Digest, pp. 415–418, Washington, D.C., December 1978.
28 Angle, R. L., Carnes, J. E., Kosonocky, W. F. and Sauer, D. J., 'Techniques for the design of high-density high speed tdi–ccd image sensors', CCD 1978 Conference Proceedings, San Diego, CA, pp. 1–13, October 1978.
29 Esser, L. J. M., 'The peristaltic charge-couple device: a new type of charge-transfer device', *Electronics Letters*, **8**, pp. 620–1, December 1972.
30 Chamberlain, S. G. and Critchlow, D. L., 'Solid-state fet colour scanner', *IBM Technical Disclosure Bulletin*, **20**, pp. 436–437, June 1977.
31 Brown, R. W. and Chamberlain, S. G., 'Quantum efficiency of a silicon-gate charge coupled optical imaging array', *Physica Stat. Sol. (a)*, **20**, pp. 675–85, 1973.
32 Burke, H. K. and Michon, G. J., 'Charge-injection imaging: operating techniques and performance characteristics', *IEEE Trans. Electron Devices*, **ED-23**, pp. 186–96, February 1976.
33 Huang, T. S. and Woods, J. W., 'Picture bandwidth compression by block quantification', presented at the Int. Symp. Information Theory Ellenville, N.Y., January 1969; abstract in *IEEE Trans. Information Theory*, **IT-16**, p. 119, January 1970.
34 Younse, J. M., Mercer, H. D. and Gooch, R. W., 'Miniature ccd cameras', CCD 1978 Conference Proceedings, San Diego, CA, pp. 63–70, October 1978.
35 Monro, C. R., 'Performance evaluation of ccd imagers', CCD 1978 Conference Proceedings, San Diego, CA, pp. 49–54, October 1978.
36 Wight, R. H., 'Low contrast imaging', Airborne Reconnaissance Proceedings of the Society of Photo-Optical Instrumentation Engineers, Vol. 101, Bellingham, Washington, 1977.
37 Esser, L. J. M., 'The peristaltic charge-coupled device for high speed charge transfer', in ISSCC Dig. Tech. Papers, p. 28, February 1974.
38 Peek, H. L., 'Twin-layer PCCD performance for different doping levels of the surface layer', *IEEE J. Solid-State Circuits*, **SC-11**, pp. 167–70, February 1976.
39 Kamins, T. I. and Fong, G. T., 'Photosensing arrays with improved spatial resolution', *IEEE Trans. Electron Devices*, **ED-25**, pp. 154–9, February 1978.
40 Burt, D. J., 'Development of c.c.d. area image sensors for 625-line television applications', *The Radio and Electronic Engineer*, **50**, pp. 205–12, May 1980.

† The issues for February 1978 of *IEEE Journal of Solid-State Circuits* and *IEEE Transactions on Electron Devices* contain the same papers.

Manuscript first received by the Institution on 31st August 1979 and in final form on 20th February 1980.
(Paper No. 1935/CC 326)

UDC 621.382: 681.325.65

Indexing Terms: Semiconductor devices, Integrated circuits, Charge-coupled devices, Logic devices, gates

Basic c.c.d. logic gates

J. H. MONTGOMERY, B.Sc.*

and

H. S. GAMBLE, B.Sc., Ph.D.*

SUMMARY

The basic logic elements of AND, OR and INVERTER gates as implemented in c.c.d. logic are described and experimentally evaluated. The AND and OR gates are incorporated in one simple structure using the mechanism of charge spillage over a d.c. control gate from a common input well. Design equations are presented for the floating-gate master-slave sensing structure which is the basis of the inverter and are shown to be in good agreement with practical results. The problems associated with the devices, namely speed of operation and charge storage under the slave, are discussed in detail.

* *Department of Electrical and Electronic Engineering, The Queen's University of Belfast, Ashby Building, Stranmillis Road, Belfast BT9 5AH.*

1 Introduction

Since their invention in 1970 by Boyle and Smith[1] charge-coupled devices (c.c.d.s) have found applications in the areas of imaging and analogue signal processing. Also because of their low power requirements and high packing density considerable effort has been expended in the design and manufacture of c.c.d. digital memories with very high storage capacity. These same features of low power and small size together with the simplicity of construction and operation make c.c.d.s attractive for digital logic implementation. As in the digital memory case the binary '1' and '0' levels are represented by approximately full and empty charge packets respectively.

The development of c.c.d. logic functions would be useful not only for digital systems, but also in the area of analogue signal processing and imaging. Such development would enable correlations, digital control and logic functions to be implemented on the same chip with a common technology. Recently Zimmerman[2] and Handy[3] have been investigating the feasibility of implementing logic functions using c.c.d. technology.

The basic building blocks required for logic implementation are AND, OR and INVERTER gates. This paper reports the results of a study into the operation of these basic gates. In the first part of the paper the design and operation of an AND/OR gate are discussed and results presented. Following this, the sensing mechanism of the inverter structure is analysed both qualitatively and quantitatively. Finally the design of a c.c.d. logic inverter incorporating this structure and the results obtained from the device in various modes of operation are presented.

2 The AND/OR Gate

The output of a 2-input c.c.d. AND gate must yield a full charge packet (i.e. a '1') when two full charge packets are presented to the inputs but provide empty packets when only one or zero full packets are present. Similarly a 2-input OR gate must provide a '1' output when one or two charge packets are present at the inputs and form an empty packet when both inputs are logic '0'. In c.c.d. logic both of these functions can be accomplished by one simple structure obviating the need for either the destruction or generation of charge packets.

Essentially the structure consists of an electrode which is served by two input shift registers (Fig. 1). Adjacent to this common electrode is a d.c. control gate over which charge is allowed to spill into another shift register, to generate the AND function. The common electrode itself has a shift register output allowing the formation of the OR function.

A voltage is applied to the d.c. control gate such that the charge capacity of the common electrode, D, is equivalent to that of a full 1. Any charge present under A and B will be dumped under D when D and C are

0033–7722/80/050258+11 $1.50/0

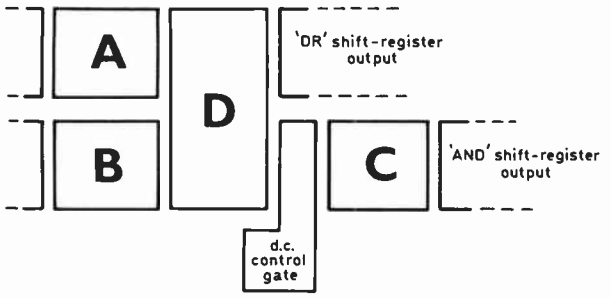


Fig. 1. Block diagram of 2-input AND/OR gate.

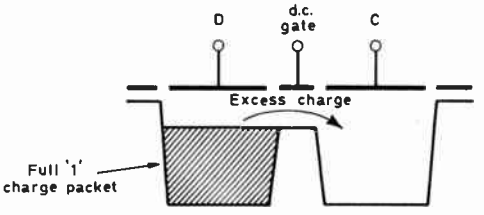


Fig. 2. Trapping of charge behind d.c. gate. Charge in excess of that allowed by d.c. gate (i.e., greater than full '1') will spill into C.

clocked on and A and B are turned off. The limited capacity of D will cause charge in excess of a full 1 charge packet to be spilled over the d.c. control gate into C (Fig. 2). Thus if both A and B contain full packets of charge, a packet of charge will be formed under C and the other packet will be trapped behind the d.c. barrier. A single charge packet in either A or B will not cause spillage into C but remain under D. Thus

$$D = A \text{ OR } B$$

$$\text{and} \quad C = A \text{ AND } B$$

The simplicity of the structure allows the design of the AND/OR gate to be based on geometrical considerations if a simple clocking technique is used.

It should also be noted that, if the two inputs are in the binary 0 state but have some charge present (due to thermal generation or transfer inefficiency) the AND output will be a clean 0 but the OR output will contain both inputs' residual charge. For the case of two 1's, degraded by transfer inefficiency, at the input, the OR output will be a full 1 and the AND charge packet will contain the cumulative effect of the inputs' transfer inefficiency.

2.1 Design of AND/OR Gate

To test the operation of the c.c.d. AND/OR gate, a two-input structure was designed and fabricated. Figure 3 shows the schematic layout. The inputs to the actual AND/OR gate consist of two parallel shift registers separated by a channel stop diffusion which is terminated inside the ϕ_4 polysilicon electrode. This allows the inputs to mix before being dumped under the adjacent common electrode. Another channel stop separates the OR and AND shift register outputs.

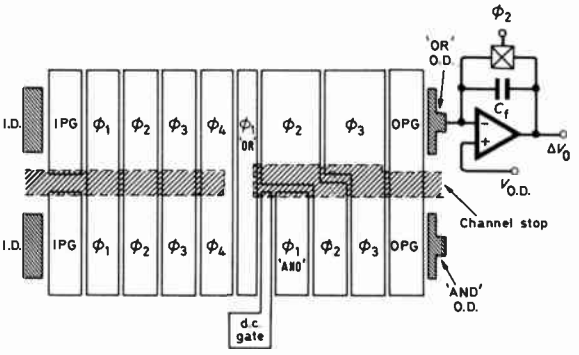


Fig. 3. Schematic diagram of experimental AND/OR gate showing off-chip output stage.

To accommodate large charge packets enabling easy detection of the signal and hence highlight the problems of speed at low clock frequencies, the standard electrode dimension was made $300 \times 100\ \mu m^2$. Since the common ϕ_1 OR electrode was required to hold the same amount of charge as the standard electrodes it was designed to have the same active area. The d.c. control gate was made $300 \times 17\ \mu m^2$ and, for ease of manufacture, the OR output electrodes were made larger than the standard dimension.

The device was fabricated on a 25 Ω-cm p-type substrate with an oxide thickness of 1400 Å. Double layer polysilicon was used to provide the electrodes for 4-phase operation.

Off-chip output stages were used to perform the charge-to-voltage conversion. These consisted of charge integrators reset by c.m.o.s. switches during the ϕ_2 clock on period. This provided a linear method of sensing the

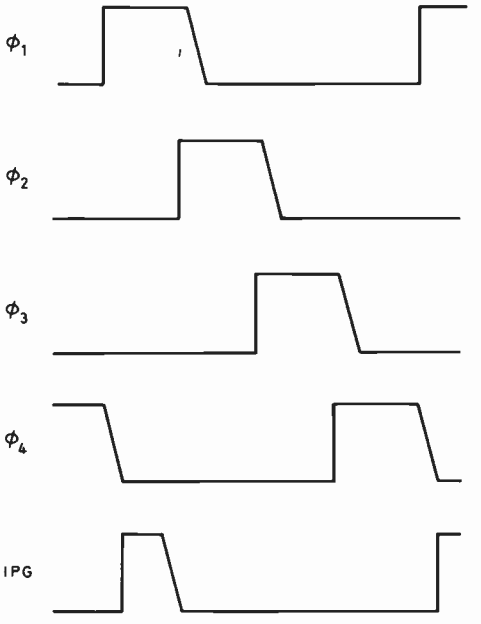


Fig. 4. Four-phase clock and input gate pulse required for diode cut-off input mechanism.

charge packets with a voltage output given by[4]

$$\Delta V_0 = -\frac{Q_{SIG}}{C_F}$$

where C_F is the storage capacitance.

2.2 Testing of AND/OR Gate

The operating frequency of the 4-phase clock was set at 12 kHz on account of the large dimensions of the device. The clock amplitudes of the phases were all made 11 V (Fig. 4); due to the measured threshold voltage of −1·3 V no back bias was required.

Prior to operation of the device in a digital manner, analogue signals were applied to both inputs of the c.c.d. to check the spilling performance of the d.c. gate. The input gates are pulsed on during the ϕ_1 on period to allow sampling of the diffusion potentials. The charge packets thus formed in the two shift registers are clocked along their respective channels and dumped under the common storage electrode.

As charge flows in, the silicon surface potential of the common storage electrode falls. If the incoming charge is sufficient to lower the surface potential below that of the d.c. gate then the rest of the charge is spilled over into the AND shift register. The charge remaining behind the d.c. barrier is moved into the OR register channel during the next clock pulse.

Constant and variable charge packets are introduced into the two input channels by applying a d.c. level and a triangular waveform to the respective source diffusions. Since a low input voltage introduces a large charge packet and vice-versa, setting the d.c. gate at 1·2 V effectively clips off the bottom of the input triangular wave. Thus for charge inputs greater than that allowed by the set d.c. gate a constant charge packet appears at the OR output whilst the excess charge appears at the AND output (Fig. 5).

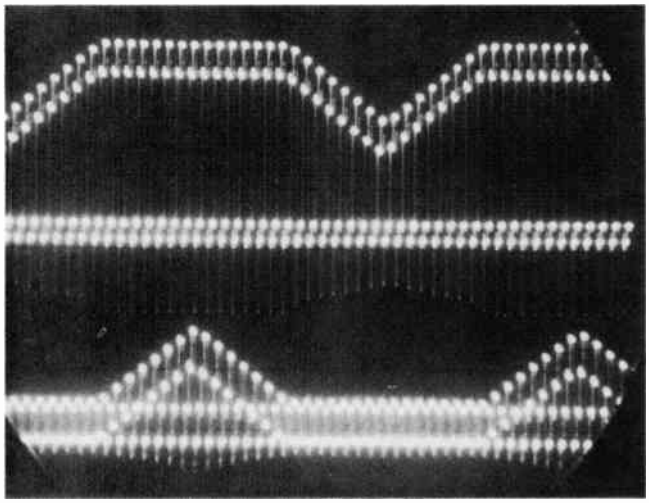


Fig. 5. Output waveforms for triangular wave and d.c. level inputs. Top trace: OR output. Bottom trace: AND output. Horizontal scale: 0·5 ms/cm. Vertical scale: 1 V/cm.

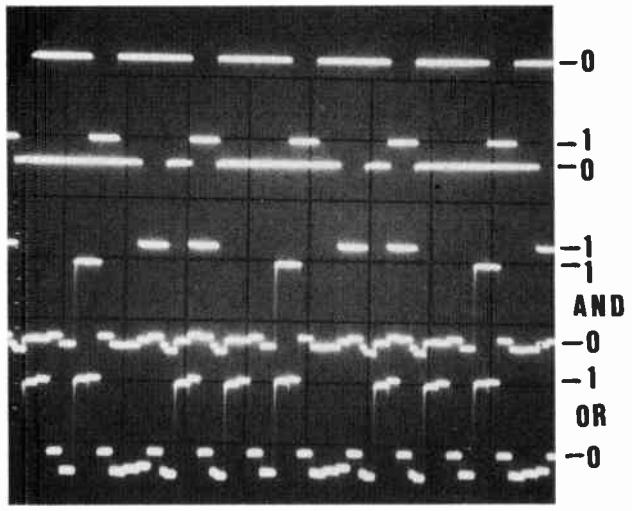


Fig. 6. Digital operation of 2-input AND/OR gate. Top traces: Digital input waveforms (low level ≡ '1' input). Upper trace input 0101010 etc. Lower trace input 00110011 etc. Vertical scale: 5 V/cm. Bottom traces: Output waveforms (high level ≡ '1' output). Upper trace output of AND channel 000100010 etc. Lower trace output of OR channel 110111011 etc. Vertical scale: 2 V/cm. Horizontal scale: 50 μs/cm. The output for any input appears after a delay of $1\frac{3}{4}$ clock periods.

If all of the mobile charge from the part of the common electrode adjacent to the d.c. control gate is not removed before the ϕ_1 clock amplitude falls below that of the d.c. gate it will be dumped into the AND shift register instead, due to the shorter path length. This causes a small but noticeable output from the AND shift register which is an attenuated replica of the OR output. Increasing the fall time of the ϕ_1 clock to beyond 2 μs partially alleviates the situation.

Figure 6 shows the 2-input gate operating in the digital mode with all possible input combinations. The output for any set of input conditions appears $1\frac{3}{4}$ clock periods later.

The above problem of long charge path lengths may be avoided by utilizing an approximately square geometry system whereby two sides of the common electrode are used to input charge and the other two are used to form the AND and OR outputs. Under this regime the slowest mechanism would be the spill from the common OR electrode into the AND channel but which could be compensated for by allowing a longer clock on-pulse for this process.

The application of square electrodes would preclude the use of 3-input AND/OR gates but these can be made from two 2-input gates allowing a higher clock frequency for that operation though at the expense of an extra clock period.

3 The Inverter

The essential requirements for an inverter are a method of detecting the signal charge present and a means of using this information to provide the logic inversion by creating either a full or empty charge packet as

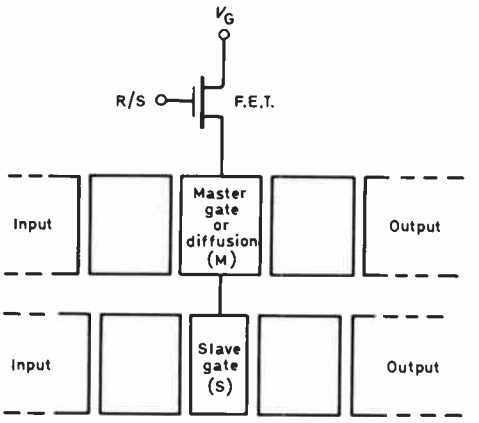
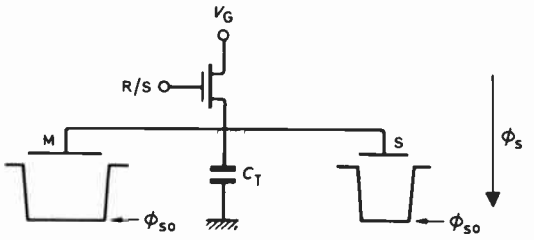


Fig. 7. Block diagram of master-slave sensing structure.

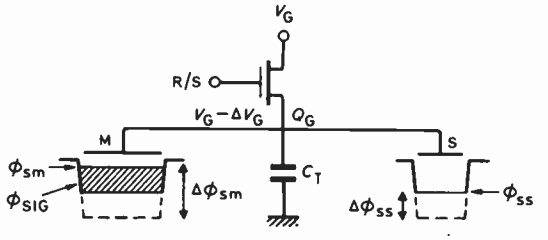
appropriate elsewhere. One way of achieving this is to use a master-slave structure.

Basically the structure consists of a common electrode which crosses two c.c.d. channels and which can be isolated from its drive voltage by an f.e.t. (Fig. 7). The part of the electrode which senses the signal charge is termed the master whilst the portion over the other channel is called the slave. Whilst floating, the master non-destructively detects the presence or absence of signal charge and this information is used by the slave to control the charge in the second channel. A floating master diffusion would give greater sensitivity but the floating gate version is preferred because of its greater versatility.

To understand the operation of the master-slave structure it is necessary to include the parasitic capacitance, C_T, as shown in Fig. 8(a). This capacitance includes all the capacitances of the floating gate structure other than those of the actual master and slave active areas, and is due mainly to the diffusion of the f.e.t. and the overlap area between the floating-gate electrode and the adjacent ones.



(a) Floating with no charge under master.



(b) Charge, Q_{SIG}, present under master.

Fig. 8. Master-slave floating-gate structure with parasitic capacitance.

Initially the f.e.t. is clocked on, allowing master, slave and parasitic capacitors to be precharged to a voltage, V_G. When the f.e.t. is clocked off the structure is left floating with a charge, Q_G, trapped on it. This trapped charge is distributed over the electrode so that:

(i) the charge on the master section of the electrode balances the charge of the master silicon depletion region;
(ii) the charge on the slave section of the electrode balances the charge of the slave silicon depletion region;
(iii) the charge on the other areas of the electrode satisfy the parasitic capacitance.

When the signal charge, Q_{SIG}, in the master channel is moved under the master electrode (Fig. 8(b)), charge is induced from the slave and parasitic capacitors to help the master electrode balance the new silicon charge. Since the charge on the parasitic and slave capacitors is reduced the electrode voltage must also fall. In response to the decrease in electrode potential the silicon surface potential under the slave electrode also drops. Thus the presence of a packet of charge underneath the master causes the silicon surface potential underneath the slave to decrease. If the supply of charge to the slave channel (whether from a diffusion or an adjacent electrode) is at a higher surface potential then no charge will flow past the slave gate. However, with no charge underneath the master electrode the slave silicon surface potential remains high and charge can flow under the slave gate from a suitably biased source to provide a full packet of charge underneath the next storage electrode. A packet of charge can therefore be placed in the slave channel for zero charge in the master channel and zero charge placed in the slave when a packet of charge is present in the master. The signal therefore supplied to the slave channel is the logical inverse to that in the master.

It should be noted that since the presence of charge in the master reduces the electrode potential, this also lowers the total amount of charge the master can store and so the dynamic range of the master is significantly reduced.

A quantitative analysis of the floating gate mechanism to determine the surface potential swing under the slave as a function of the charge packet size under the master and the physical constants of the device is presented in the Appendix. This is based on the combination of the standard c.c.d. equation and the charge balance equation before and after the signal charge enters the master.

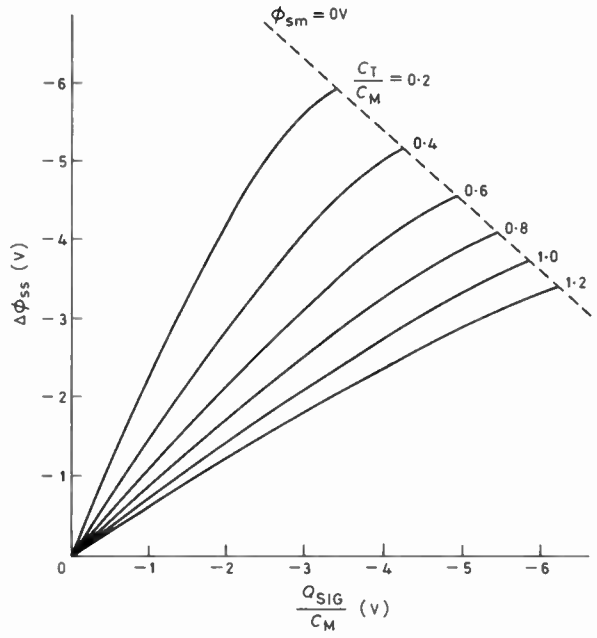
Design curves are shown in Figs. 9(a) and (b) exhibiting the effect of varying device parameters on the slave sensitivity. Normalization of the parasitic capacitance and the signal charge to master capacitance has been used to make the curves independent of absolute master and slave dimensions. The heavy

damping effect of the parasitic capacitance on the slave silicon surface potential swing for a given master signal charge in Fig. 9(a) is exhibited, for a master to slave capacitance ratio of $C_S/C_M = 1{\cdot}0$. The dashed diagonal line represents the limit of charge handling of the numerical solution, i.e. $\phi_{sm} = 0$ V. The practical limit would be slightly less than this since a surface potential of at least $2\phi_F$ is required to support the inversion layer. It can be seen therefore that, as the parasitic capacitance is increased, the signal charge handling capacity is increased and the relationship between the slave silicon potential swing and signal charge size becomes more linear.

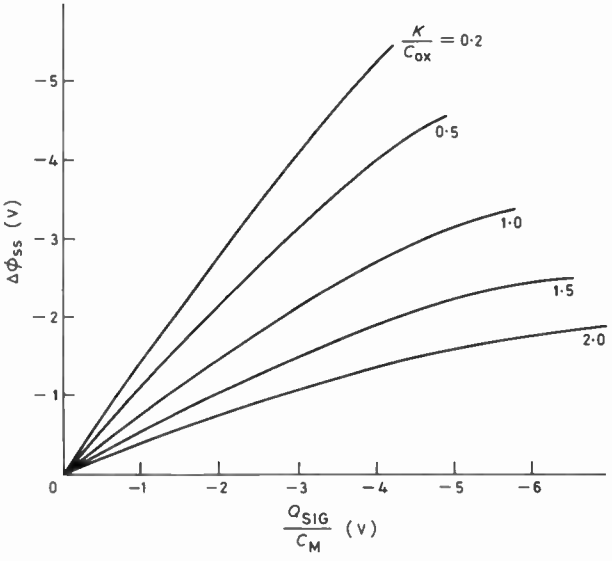
A trade-off also exists between the doping density, N_A, and the oxide capacitance, C_{ox}, due to the relationship

$$\frac{K}{C_{ox}} = \frac{\sqrt{2\varepsilon_s\varepsilon_0 qN_A}}{C_{ox}}.$$

Figure 9(a) was obtained for a K to C_{ox} ratio of 0·5. The dependence of the sensitivity of the structure on the K/C_{ox} ratio is given in Fig. 9(b). It can be seen that maximum sensitivity and linearity are obtained for low values of substrate doping and oxide thickness.



(a) Effect of parasitic capacitance on master-slave structure.
$V_G = 10$ V, $V_{FB} = 0$ V, $\frac{K}{C_{ox}} = 0{\cdot}5$, $\frac{C_S}{C_M} = 1{\cdot}0$.



(b) $\frac{K}{C_{ox}}$ variations on slave sensitivity.
$V_G = 10$ V, $V_{FB} = 0$ V, $\frac{C_S}{C_M} = 1{\cdot}0$, $\frac{C_T}{C_M} = 0{\cdot}6$.

Fig. 9.

3.1 Design of Master-slave Inverter

A simple device structure was designed and fabricated to evaluate the master-slave sensing amplifier. Figure 10 shows the schematic diagram of the structure which consists of two adjacent short charge-coupled registers separated by a channel stop diffusion. The master sensing electrode is situated in the middle of the larger register while the slave acts as a second input gate to the shorter c.c.d. Thus the slave controls the flow of charge from the source diffusion into the slave register.

As for the AND/OR gate the standard electrode dimension was made 300 × 100 μm except for the slave which was 300 × 30 μm. The same fabrication technique was also used, again yielding four-phase operation and an oxide thickness of 1400 Å. Due to the constraints of the process the master and slave electrodes were formed on different polysilicon layers and were interconnected by aluminium.

No attempt was made to minimize the parasitic capacitance because of the large dimensions of the electrodes. The electrode overlaps were measured as 4·5 μm which, along with the other non-active regions of the master-slave yielded a capacitance of 3·35 pF. The drain diffusion of the isolating transistor was 60 × 40 μm giving an estimated capacitance of 0·05 pF. Thus the total parasitic capacitance of the master-slave gate is 3·4 pF.

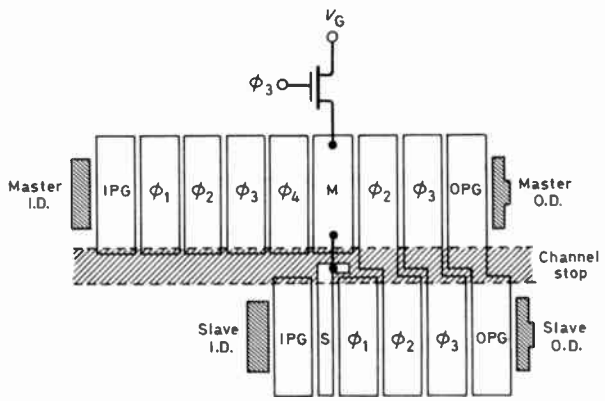


Fig. 10. Design of c.c.d. master-slave inverter.

The inverter is designed to be driven from a four-phase clock with the amplitudes of phases two, three and four set at a voltage, V_p, whilst that for phase one is approximately $\frac{2}{3}V_p$. The pre-set gate voltage for the master-slave structure is made approximately 0·5 V_p.

A charge packet is introduced into the master by lowering its source diffusion potential and pulsing on the input gate during ϕ_1. This allows charge to flow from the source to underneath ϕ_1 during the interval the input gate is on. The source diffusion is so biased that the ϕ_1 electrode is only half filled even for a full 1 logic level. The charge packet is then clocked along the master channel under the ϕ_2 electrode and subsequently the ϕ_3 electrode. Whilst the charge resides under the ϕ_3 electrode the f.e.t. is pulsed on setting the master and slave gates to $V_p/2$. Isolation of the electrode structure is achieved when ϕ_3 turns off, also transferring the signal charge to underneath the ϕ_4 electrode. Since the charge packet is less than half the total capacity of the ϕ_4 electrode and the floating master is at half the ϕ_4 electrode potential, all of the signal charge resides under ϕ_4.

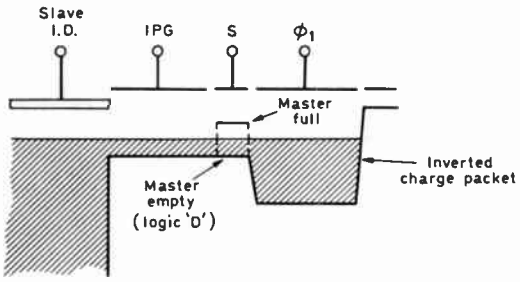


Fig. 11. Slave surface potentials required for inversion of digital information.

When ϕ_1 is turned on and ϕ_4 turned off the signal charge is dumped under the master electrode thus causing the electrode potential to fall. This in turn causes the slave silicon surface potential to fall (Fig. 11). During this ϕ_1 on period the slave input gate is pulsed to $V_p/2$ giving access to the slave source diffusion which is held at a constant surface potential of $\phi_{SO}-\frac{1}{2}\Delta\phi_{SSF}$. The term $\Delta\phi_{SSF}$ is the slave silicon surface potential swing for a full 1 logic level of signal charge under the master electrode.

If the charge under the master is insufficient to produce a change in the slave surface potential of greater than $\Delta\phi_{SSF}/2$, charge from the slave source is injected over the slave to fill ϕ_1 to the same potential as the source thereby producing a logic 1 at ϕ_1. Conversely if the charge packet in the master does produce a slave swing greater than $\Delta\phi_{SSF}/2$, the slave will act as a barrier stopping charge flow from the source to ϕ_1 thereby producing a logic 0 at ϕ_1.

To isolate the regenerated charge packet in the slave the input gate is turned off before the ϕ_2 electrode comes on. During ϕ_2 all the signal charge is drained from the master since the electrode potential of the master is only half that of ϕ_2. Also, as ϕ_1 turns off, no charge is left under the slave electrode. The charge in both channels is then clocked to the respective output diodes where the output of the slave channel is the inverse of that obtained from the master channel.

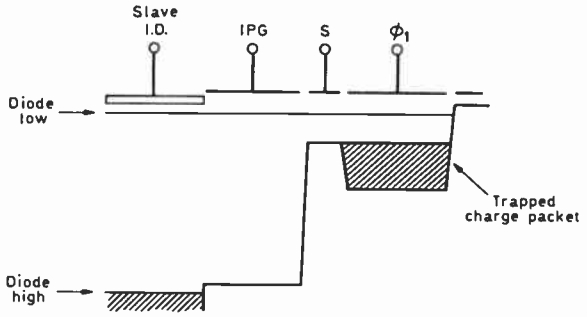


Fig. 12. Fill-and-spill mechanism for determining slave surface potential swings.

3.2 Determination of Master-slave Sensitivity

The clock frequency was again set at 12 kHz while the clock voltage amplitude, V_p, used was 11 V.

The variation in slave silicon surface potential with signal charge in the master was determined by using the fill-and-spill technique for inputing charge into the slave and driving the master channel as described previously. During the ϕ_1 clock on period when the signal charge is residing under the master sense electrode, the slave input gate is pulsed on to +15 V followed by the diode pulsed low. This causes charge to be injected over any barrier at the slave into the ϕ_1 well until its surface potential equals that of the source (Fig. 12). The diffusion is then returned to a high potential allowing charge to drain back into the diffusion. The amount of charge still trapped in the ϕ_1 well is determined by the silicon surface potential of the slave. If the silicon surface potentials underneath the slave and ϕ_1 electrodes are equal in the absence of any master charge, the amount of charge trapped after the fill-and-spill sequence is proportional to the slave swing caused by the signal charge in the master. Thus the slave swing can be calculated from the charge generated in the slave from the expression

$$Q_{SLAVE} = A_{\phi_1}(C_{ox}+C_s)\Delta\phi_{ss}$$

where A_{ϕ_1} is the area of the ϕ_1 electrode and C_s is the depletion capacitance/unit area of the ϕ_1 electrode. The charge in the master channel is sensed at the output so that the sensitivity of slave surface potential to master signal charge can be obtained. A comparison of the experimental results obtained with those calculated by the numerical method for the actual device parameters is given in Fig. 13.

Since the slave silicon surface potential swing is approximately proportional to the master charge packet

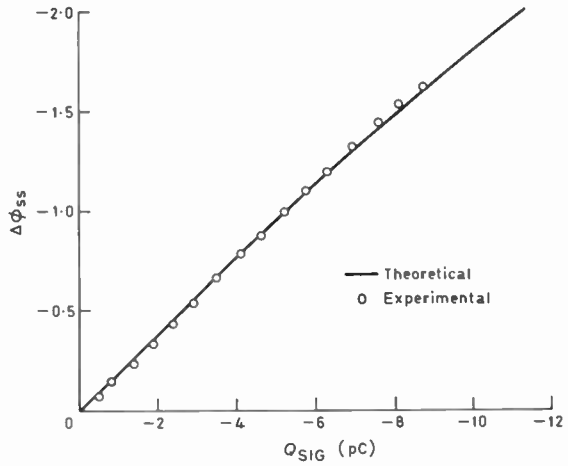


Fig. 13. Measured slave surface potential swings for charge in the master. Theoretical graph is calculated for

$$\frac{C_S}{C_M} = 0{\cdot}3, \quad \frac{K}{C_{ox}} = 0{\cdot}515, \quad \frac{C_T}{C_M} = 0{\cdot}45.$$

size the output of the slave channel for the fill-and-spill input is a replica of the master signal. Figure 14 shows the master and slave outputs in response to a triangular wave input under this mode of operation. The slave signal output can be an amplification of the master signal provided the reduced slave surface potential swing compared to that of the master is more than compensated for by the geometry of the slave ϕ_1 electrode.

During the preceding testing of the device a 1 V increase in the floating electrode voltage was apparent, independent of the charge in the master. This was caused by clock and reset noise but since this occurred irrespective of master charge packet size the operation of the floating gate structure was not significantly affected.

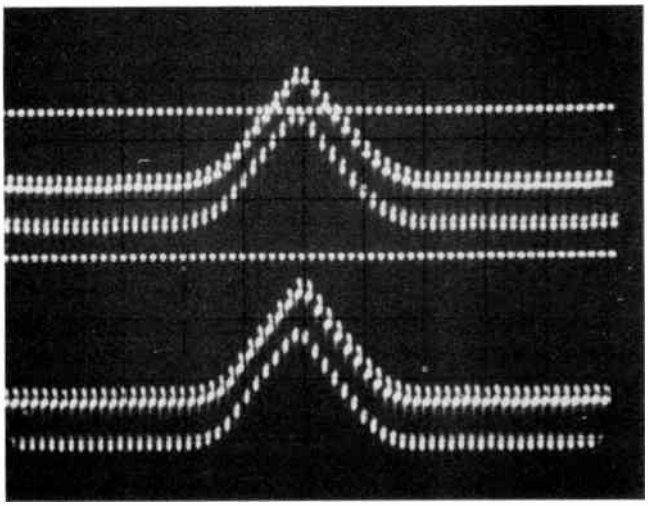


Fig. 14. Master and slave output waveforms for triangular waveform input to master during slave sensitivity measurements. Top trace: Master output. Bottom trace: Slave output. Horizontal scale: 0·5 ms/cm. Vertical scale: 0·2 V/cm.

3.3 Operation as an Inverter

With the device driven as outlined previously for inverter operation the output from the slave channel, for a triangular wave input to the master, is shown in Fig. 15. It can be seen that some modulation of the size of a logic 1 charge packet is present. This is due to the charge stored under the slave which becomes part of the slave signal charge when the slave input gate is turned off. Since the slave surface potential varies with the magnitude of the master signal charge, the amount of charge stored underneath the slave will also vary with master signal charge.

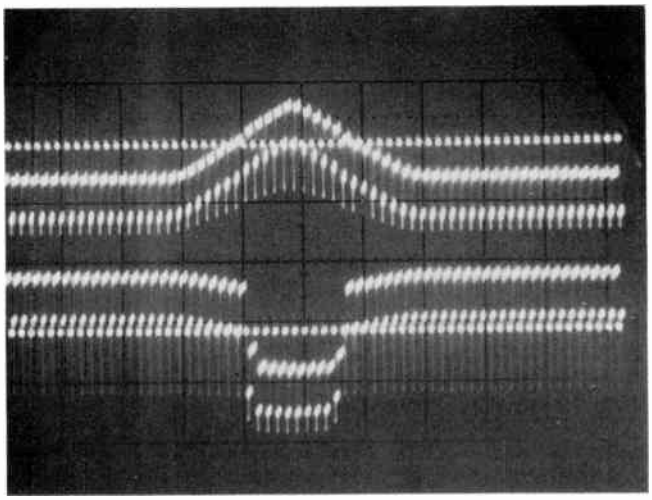


Fig. 15. Inversion of triangular master signal. Top trace: Master output. Bottom trace: Digital slave output (high level ≡ 1 output). Horizontal scale: 0·5 ms/cm. Vertical scale: 0·2 V/cm.

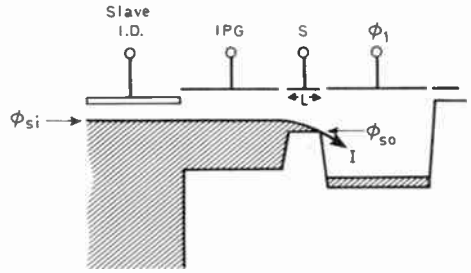


Fig. 16. Finite charge flow rate over slave.

If the slave silicon surface potential is slightly below that of the source, partitioning of charge over the slave can occur as the input gate turns off, yielding partial charge packets in the slave register when none is desired. The major mechanism which results in partially filled charge packets, however, is the finite rate of charge flow across the slave into the ϕ_1 well when the slave surface potential is slightly greater than that of the source (Fig. 16). After the initial transient, the flow of charge into the ϕ_1 well is constant until its surface potential falls to that of the slave surface potential when the flow begins to decrease. Ignoring the influence of the input gate the

flow of charge in the constant current mode can be approximated by the equation[5]

$$I = \frac{Z\mu C_G}{2L}(\phi_{SO} - \phi_{SI})^2 \qquad (1)$$

where μ is electron mobility, C_G is the capacitance/unit area of the slave gate, Z is the slave electrode width, L the slave electrode length, ϕ_{SO} is the slave surface potential for no charge present and ϕ_{SI} is the potential of the source. A voltage transition region thus exists between the point where the slave silicon surface potential is low enough to just fill the ϕ_1 well in the time that the input gate is on, and the point where the slave surface potential is equal to the source potential. Rearrangement of equation (1) yields the width of this transition region (ΔV_T) assuming the constant current mode

$$\Delta V_T = \sqrt{\frac{-2Q_1 L}{Z\mu C_G \tau}} \qquad (2)$$

where Q_1 is the charge packet size for a full 1 and τ is approximately equal to the 'on' period of the input gate.

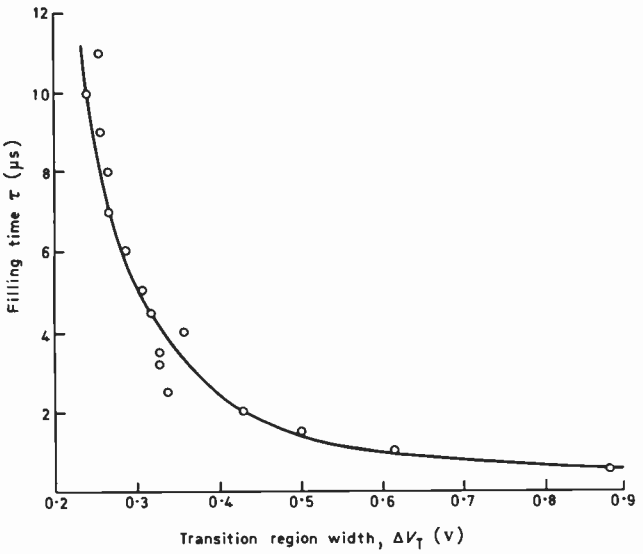


Fig. 17. Experimental determination of transition region voltage, ΔV_T, for various filling times, τ.

Figure 17 shows the results of the experimental measurement of the transition region width for a slave charge packet size of −16 pC. The c.c.d. is driven in a similar manner to that used to find the sensitivity of the slave with the exception that the source low potential is variable. The difference in the value of the diffusion potential required to just fill the ϕ_1 well up to the level of the slave (i.e. input −16 pC) and the value required to just stop any charge flow yields the transition region width for the time allowed for filling, τ. A driven slave electrode produced similar results to those obtained for the floating electrode case thus implying that the floating slave is not a significant factor.

The finite charge flow rate and input gate charge partitioning combine to produce a threshold between an inverted 1 and 0 which is not sharp but rather occurs over a range of master charge packet size. Since it is desirous to operate the inverter well clear of this transition region, regeneration of the deteriorating charge packets must be performed periodically. However, as the output of the inverter itself is a regenerated signal, extra regenerations may not be required above those needed to form the digital function.

Care must be taken in the design of any c.c.d. logic device to ensure that no charge is trapped under the slave or master electrodes. With the present device this can occur if the ϕ_2 clock amplitude is insufficient to drain all the charge out of the slave ϕ_1 well or the master respectively before ϕ_1 goes off.

The trapped charge will remain under the master or slave whilst the electrode is reset to V_G causing an extra induced charge equal to the total trapped charge to be stored on the electrode. Charge trapping under the master does not significantly affect the sensing operation of the floating-gate structure because this charge remains trapped throughout the complete sensing cycle and causes the extra electrode charge to be immobile. Charge trapped under the slave is removed when the ϕ_1 clock comes on with the result that some of the extra electrode charge is redistributed over the whole electrode surface effectively causing an increase in the floating electrode voltage. Thus if ϕ_4 is turned off and a charge packet is dumped into the master, sensing will proceed normally with the exception that the effective reset voltage is higher. This shift in operating conditions would pose serious problems for c.c.d. logic where the relative surface potentials are critical and so charge trapping must be avoided.

The combined effect of charge trapping under the slave and the finite charge flow rate into the slave can be seen

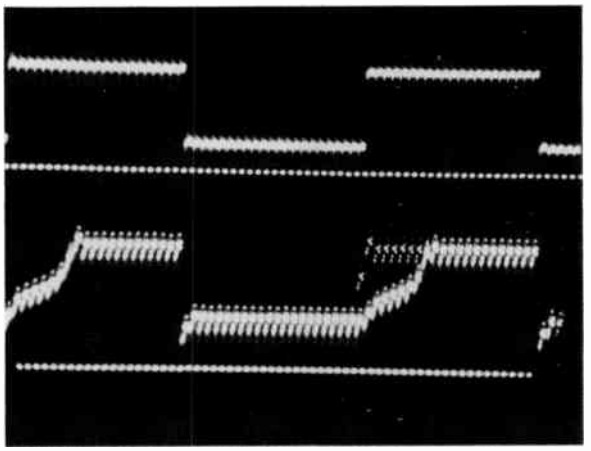


Fig. 18. Response of master and slave to square wave input showing transition region effect on storage mode. Top trace: Master output. Bottom trace: Slave output. Horizontal scale: 0·5 ms/cm. Vertical scale: 0·2 V/cm.

in Fig. 18. Here the device was operated in the manner used to find the slave sensitivity with the exception that the ϕ_2 clock amplitude was reduced to permit charge trapping under the slave. An asynchronous square wave was applied to the master channel. After a series of 0's down the master channel, no charge is left trapped under the slave since the slave signal follows that in the master register.

Upon receipt of the first full packet in the master the slave surface potential swing is sufficient to bring it close to that of the diode low potential restricting charge flow over the slave into the ϕ_1 well. A small packet is formed in the ϕ_1 well and some of this is trapped under the slave when the ϕ_1 clock is turned off. This causes a slight upward shift in the operating conditions of the floating electrode so that though the subsequent full charge packet produces the same slave potential swing, the final slave surface potential is slightly higher than for the previous sample. The diffusion is therefore able to inject more charge into the slave register thus allowing greater charge trapping and further shift in the effective reset voltage. This process continues till the slave ϕ_1 well can be filled up to its required level and the rest of the master square wave is transmitted faithfully by the slave.

In the centre of the photograph the ramping-up of the slave signal can be clearly seen. In the background of this portion of the trace, the slave output achieves its final value quickly. The above procedure in this case has been short-circuited by a master charge packet which has sampled on the edge of the asynchronous square wave. This reduced charge packet causes a smaller slave swing, avoiding the transition region effect and permitting sufficient charge storage to allow the subsequent slave output to follow the master.

4 Discussion

In the operation of the AND/OR gate the limiting speed factors are the long path lengths presented to the signal charge, both for the spilling process and for removing all the OR charge from the common electrode. The maximum device operating speed could be considerably enhanced by using a nearly square common electrode, permitting inputs and outputs from two sides each, and reducing the electrode dimensions to less than 10 μm.

The AND/OR gate also suffers from the cumulative effects in the two channels of either transfer inefficiency or thermal generation. The former determines the highest operating speed and the latter the slowest speed. Natural refresh in a logic function occurs by virtue of the inversion operation, and by careful design such operations may be sufficient to ensure error-free performance. If more stages are deemed necessary NAND and NOR gates, with built-in refresh using dual master-slave gate structures, could be incorporated in the function.

The master-slave inverter structure exhibits a transition region voltage which is dependent on device geometry and speed of operation. Provided adequate slave silicon surface potential swings can be obtained then it is possible to avoid operation near this region. As long as the parasitic-to-master gate capacitance ratio can be kept small the master-slave structure can be scaled down whilst still maintaining satisfactory sensitivity. Special fabrication techniques can be employed to minimize the overlap capacitance. Further reduction could also be achieved by allowing the master width/length ratio to approach unity. This would have a severe limitation in speed unless the electrode dimensions were again reduced to say 10 μm square, at which point the AND/OR gate would provide the limit to the maximum operating speed.

Under the present system charge must be allowed to settle under the master electrode before the slave input gate is pulsed on. An improved mode of operation would allow one complete clock-phase on-pulse for the charge transfer into the master and the subsequent phase pulse for the slave to perform its function. The timing for the slave input gate pulse would thus become less critical, thereby increasing device operating speed. The use of several clock pulse on-periods to achieve the completion of the slower charge transfers and functions therefore appears to be one answer for improving the performance.

It follows from the preceding work that c.c.d. logic circuits require the charge packets to be presented to the inputs of the AND/OR gate or inverter during the correct clock period, and the operations themselves take several fractions of a clock period to be completed. Thus for a more complex logic function with multiple parallel inputs—such as the addition of two 4-bit words—the input charge packets must be suitably delayed to allow sufficient time for the carry-out of the previous stage to be formed and presented at the next stage of the adder. The output charge packets must also be delayed appropriately to allow a parallel output to be formed if required. Some logic functions such as those required for performing digital correlations, fast Fourier transforms, etc., can operate with skewed arithmetic thus minimizing the need for synchronizing shift registers between the various arithmetic sub-units.

Comparing d.c.c.l. (digital charge coupled logic) with I^2L—its nearest rival—Zimmerman[6] has shown that d.c.c.l. has a power consumption an order of magnitude lower if both are based on 5 μm design rules. Power consumption in charge coupled devices is mainly due to the charging and discharging of the electrode capacitances. Size-wise too, d.c.c.l. has a five to one advantage over I^2L if device active areas only are considered. Digital charge coupled logic, however, appears to be limited to a maximum clock speed of around 10 MHz for surface channel devices.

Whilst the master-slave structure in this paper performed a digital function it is also suitable for use

with purely analogue or mixed analogue and digital c.c.d.s. An example of the latter could be in its use to multiplex several analogue signals into a common serial shift register. Here the master channel would be driven with a circulating 0 charge packet with the slave gates controlling the analogue inputs. Another example of mixed analogue and digital c.c.d.s is to be found in Traynar and Beynon's paper in this issue.

5 Conclusion

Digital logic can be implemented in c.c.d. technology with simple gate structures which are easily fabricated. Consideration must be taken in the design of the 2-input AND/OR gate to minimize the charge transfer distances in the common OR well. For this reason it would appear to be better to cascade several 2-input structures rather than increase the number of inputs to a single structure.

The design curves produced for the sensitivity of the master-slave structure are shown to be in good agreement with the experimental results. In the inversion mode the master-slave inverter exhibited a threshold band between the 0 and 1 logic levels. This band is mainly caused by the finite rate of charge flow from the diffusion into the storage well and is dependent on the dimensions of the structure and the time allowed for filling. If the electrode lengths and filling time are reduced *pro rata* the transition region width will remain the same. Further increase in operating speed will result in a widening of this threshold band. The parasitic capacitance is the most important damping factor on the sensitivity of the master-slave structure and should be minimized. Reduction of the standard electrode size would improve the performance of the inverter structure in terms of speed, provided the parasitic capacitance can be lowered accordingly.

6 Acknowledgments

The authors wish to thank Dr S. H. Raza for the fabrication of the devices and the Science Research Council for the financial support provided.

7 References

1 Boyle, W. S. and Smith, G. E., 'Charge-coupled semiconductor devices', *Bell Syst. Tech. J.*, **49**, no. 4, pp. 588–93, April 1970.
2 Zimmerman, T. A., 'The digital approach to charge-coupled device signal processing', Circuits and Systems Symposium, Boston, April 1975.
3 Handy, R. J., 'Use of c.c.d.s in the development of digital logic', *IEEE Trans. Electron Devices*, **ED-24**, no. 8, pp. 1049–61, August 1977.
4 Baertsch, R. D. *et al.*, 'The design and operation of practical charge transfer transversal filters', *IEEE Trans. Electron Devices*, **ED-23**, no. 2, pp. 133–41, February 1976.
5 Chamberlain, S. G. and Woo, B. Y., 'Transient analysis of electrical charge injection into charge-coupled devices', *Int. J. Electronics*, **40**, no. 6, pp. 569–86, 1976.
6 Zimmerman, T. A., Allen, R. A. and Jacobs, R. W., 'Digital charge-coupled logic (d.c.c.l.)', *IEEE J. Solid-State Circuits*, **SC-12**, no. 5, pp. 473–85, October 1977.

8 Appendix: Theoretical analysis of the master-slave gate

The analysis below determines the surface potential swing under the slave as a function of the charge packet size under the master and the physical constants of the device.

The relationship between the silicon surface potential, ϕ_s, and the electrode voltage, V_G, of an m.o.s. structure is given by

$$V_G = V_{FB} + V_{ox} + \phi_s \tag{3}$$

where V_{FB} is the voltage required to obtain flat band conditions at the oxide-silicon interface and V_{ox} is the voltage drop across the oxide.

Now

$$V_{ox} = \frac{Q_G}{C_{ox}} = -\frac{Q_{SI}}{C_{ox}}$$

where Q_{SI} is the total charge in the silicon per unit area consisting of the depletion layer charge, Q_D, and the inversion layer charge, Q_I. The depletion layer charge can be expressed as

$$Q_D = -\sqrt{2q\varepsilon_s\varepsilon_o N_A}\,\sqrt{\phi_s}. \tag{4}$$

Considering the case where the inversion layer $Q_I = 0$ and $\phi_s = \phi_{so}$ then putting

$$V_o = \frac{q\varepsilon_s\varepsilon_o N_A}{C_{ox}^2}$$

and

$$V_G' = V_G - V_{FB}$$

then

$$V_G' = \sqrt{2V_o}\,\sqrt{\phi_{so}} + \phi_{so}$$

or

$$(V_G' - \phi_{so})^2 = 2V_o\phi_{so}$$

giving

$$\phi_{so} - 2\phi_{so}(V_o + V_G'^2) = 0.$$

Thus in the absence of signal charge

$$\phi_{so} = V_G' + V_o - \sqrt{2V_oV_G' + V_o^2}. \tag{5}$$

This equation therefore allows the silicon surface potential, ϕ_{so}, under both the floating master and slave electrodes to be calculated before the signal charge is dumped under the master (Fig. 8(a)). From this value of surface potential the depletion layer charge per unit area under the master and slave electrodes can be found using equation (4). The total charge on the master-slave electrode, Q_G, is numerically equal to the sum of the two-depletion layer charges and the charge on the parasitic capacitance, C_T. Therefore

$$Q_G = A_mK\sqrt{\phi_{so}} + A_sK\sqrt{\phi_{so}} + C_TV_G \tag{6}$$

where

A_m = area of master electrode
A_s = area of slave electrode
$K = \sqrt{2\varepsilon_s\varepsilon_o qN_a}$.

When the signal charge, Q_{SIG}, is dumped under the master, both the master and slave surface potentials change to new (and different) values (Fig. 8(b)). Poisson's law can still be applied individually to the master and slave electrodes. For the master

$$V'_G+\Delta V_G = \sqrt{2V_o\phi_{sm}}+\phi_{sm}-\frac{Q_{SIG}}{A_mC_{ox}}. \quad (7)$$

The slave still has no inversion layer

$$V'_G+\Delta V_G = \sqrt{2V_o\phi_{ss}}+\phi_{ss}. \quad (8)$$

where

ϕ_{sm} = final master surface potential

ϕ_{ss} = final slave surface potential

ΔV_G = change in the electrode potential.

Since the electrode is floating the total electrode charge remains constant though redistributed. The charge balance equation yields

$$Q_G = A_mK\sqrt{\phi_{sm}}+A_sK\sqrt{\phi_{ss}}+C_T(V_G+\Delta V_G)-Q_{SIG}. \quad (9)$$

Equations (7), (8) and (9) may be solved numerically on a computer to determine the change in slave surface potential ($\Delta\phi_{ss}$) for a given signal charge, Q_{SIG}, in the master.

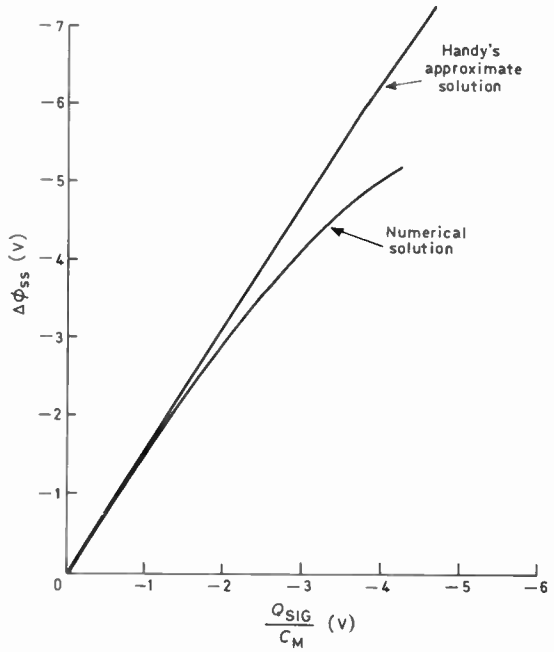


Fig. 19. Comparison of numerical and Handy's analytical solution.

$V_G = 10$ V, $V_{FB} = 0$ V, $\frac{K}{C_{ox}} = 0{\cdot}5$, $\frac{C_S}{C_M} = 1{\cdot}0$, $\frac{C_T}{C_M} = 0{\cdot}4$

An alternative manual approach, still using the same equations, is to calculate the initial gate charge and silicon surface potential of the slave from equations (5) and (6). Then by choosing a slave surface potential variation, the new electrode voltage can be calculated using equation (8). The new charge on the master, Q_M, can then be found

$$Q_m = Q_G-A_sK\sqrt{\phi_{ss}}-C_T(V_G+\Delta V_G). \quad (10)$$

This charge is balanced by an equal amount in the silicon

$$Q_m = A_mK\sqrt{\phi_{sm}}-Q_{SIG}. \quad (11)$$

Equations (7) and (11) can then be combined to find the master signal charge required for the specified swing in slave surface potential

$$Q_{SIG} = A_mK\left(V'_G+\Delta V_G-\frac{Q_m}{A_mC_{ox}}\right)^{\frac{1}{2}}-Q_m. \quad (12)$$

If design curves are required the above exercise can be repeated for various changes in the slave surface potential.

An approximate closed form analytical solution has also been derived by Handy but the presented result contained some minor errors. His solution is presented below for the case where the master and slave gates have the same m.o.s. parameters per unit area

$$\Delta\phi_{ss} = \int_{\phi so}^{\phi ss} d\phi_s = \frac{V_oA_m}{2M^3}\left\{1-\frac{1}{\sqrt{K_1}}\right\}\times$$
$$\times\left[M^2(K_2-K_1)-2M(M+A_m)(\sqrt{K_2}-\sqrt{K_1})+\right.$$
$$\left.+2A_m(M+A_m)\ln\left(\frac{A_m+M\sqrt{K_2}}{A_m+M\sqrt{K_1}}\right)\right] \quad (13)$$

where

$$K_1 = 1+\frac{2V'_G}{V_o}$$

$$K_2 = K_1+\frac{2Q_{SIG}}{V_oC_{ox}A_m}$$

$$M = A_s\left(\frac{1}{\sqrt{K_1}}\right)+\frac{C_T}{C_{ox}}.$$

A comparison of the numerical solution and Handy's analytical solution for the dependence of the change in slave surface potential, $\Delta\phi_{ss}$, on the size of the signal charge is shown in Fig. 19 for a typical set of operating conditions. The difference between the two solutions increases with increasing signal charge as the assumption in the approximate analytical proof becomes less valid. This also causes the analytical solution to predict an over optimistic value for the dynamic range of the gate. The numerical solution has therefore been used throughout to obtain the design curves of Figs. 9(a) and (b).

Manuscript first received by the Institution on 25th June 1979 and in final form on 28th December 1979
(Paper No. 1936/CC 327)

Contributors to this issue

D. J. Burt read natural sciences at Cambridge University and on receiving his degree in 1966 he joined Hirst Research Centre, where he worked on m.o.s. device technology. He has been particularly concerned with charge coupled devices since 1971 and he is at present a member of the Leading Scientific Staff and Department Head responsible for c.c.d. research and development.

Peter Denyer graduated with first-class honours in electronic and electrical engineering at Loughborough University of Technology in 1975. Since that time he has spent four years designing analogue and digital m.o.s. large-scale integrated circuits at the Wolfson Microelectronics Institute. He is now a Director of Denyer-Walmsley Microelectronics, Edinburgh.

Colin Cowan received the B.Sc. degree in electrical and electronic engineering from the University of Edinburgh, in 1977. He is now working on the optimization of c.c.d.-based systems at the University of Edinburgh as a postgraduate research student.

John Mavor (Member 1966) received the B.Sc. degree from City University in 1964 and the Ph.D. degree from the University of London in 1968 for research into m.o.s. transistor noise mechanisms. He initially spent several years in the semiconductor industry with periods at Texas Instruments and Hughes Microelectronics in the UK. Then, in 1971, he was appointed Lecturer and in 1979 Reader in the Department of Electrical Engineering at the University of Edinburgh. Dr Mavor has had over fifteen years' experience in c.c.d./m.o.s. integrated circuit technology and device design with special applications interests of memory and signal processing, and he has been author or co-author of over 50 papers in this area. He is a Joint Honorary Editor of the new IEE Journal on Solid-State and Electronic Devices.

Dr Mavor was Chairman of the 1974 and 1976 International Conferences on the Technology and Applications of Charge-Coupled Devices, held at the University of Edinburgh, and was Technical Programme Co-Chairman of CCD 79.

John Arthur received the B.Sc. degree in physics from the University of Toronto in 1971 and the Ph.D. degree from the University of Edinburgh for work on light scattering in crystals. This was followed by post-doctorate work in the same field on the computer control and analysis of experiments. In 1976 he joined the Wolfson Microelectronics Institute at Edinburgh University as a Research Fellow working on c.c.d. systems design. In 1979 Dr Arthur went to Racal-MESL as a Group Leader in the Microwave Division.

Professor Savvas Chamberlain studied electronics at the Northern Polytechnic, London, and received M.Sc. and Ph.D. degrees from the University of Southampton in 1965 and 1967 respectively. He then joined the Allen Clark Research Centre of the Plessey Company at Caswell where he worked on device physics, design of integrated circuits, high-frequency bipolar transistors, and m.o.s.f.e.t. self-scanned optical imaging arrays. In 1969 he went to the University of Waterloo, Ontario, as an Assistant Professor in the Department of Electrical Engineering. He was promoted to Associate Professor in July 1972 and he has held a chair since July 1977. In August 1971 Professor Chamberlain had leave from the University to work for nine months as a full-time consultant at Bell-Northern Research, Ottawa, on charge-coupled devices, and he continued as a part-time consultant to the company on charge-coupled devices and integrated circuits until August 1974. From September 1974 he spent a sabbatical year at IBM's Thomas J. Watson Research Center, at Yorktown Heights, as a Visiting Scientist, where he worked on c.c.d. scanners, photo-detectors, imaging devices and integrated circuits. Professor Chamberlain has published numerous papers on modern semiconductor device design, analysis, and modelling. In addition, he has a number of patents and patents pending on integrated circuits, optical imaging devices, and solid-state scanners.

Paul Traynar entered Southampton University in 1971 to study electronics. He received the B.Sc. degree in 1974 and remained at Southampton to do research in the Microelectronics Group on the application of charge-coupled devices to radar signal processing. He was awarded the degree of Ph.D. in 1978 and continued research into c.c.d. techniques and processing technology as a research fellow. At the end of 1978 Dr Traynar joined Plessey Semiconductors, Swindon, and he is now involved in the design of integrated circuits for television digital tuning systems.

Harold Gamble received the B.Sc. degree in electrical engineering from The Queen's University of Belfast in 1966 and his Ph.D. degree in 1979. From 1970 to 1972 he worked at Standard Telecommunications Laboratory at Harlow in Essex where he was involved in the technology of silicon gate m.o.s.t.s. He returned to the Department of Electrical and Electronic Engineering at the University as an SRC-supported research fellow to work on baritt microwave diodes. In May 1973 he was appointed to a lectureship in electronics and has been in charge of the microelectronics fabrication laboratory since 1977. For the past six years Dr Gamble has worked on various aspects of charge-coupled devices with the support of the Science Research Council.

John Montgomery received a first class honours B.Sc. degree in electrical engineering from The Queen's University of Belfast in 1977. He worked for a short period with the Wolfson Signal Processing Unit at Belfast before returning to Queen's to work for his Ph.D. degree. He is supported by a Northern Ireland Department of Education postgraduate studentship and a University Foundation studentship. His research topic for the past two years has been the operation of charge-coupled devices for digital logic implementation.

Gordon Harp studied at Worcester Technical College and was awarded an H.N.C. in physics in 1962 and an H.N.C. in electronics in 1964; he received the degree of B.A. in mathematics in 1978 from the Open University. Mr Harp joined RSRE in 1959 and worked on signal processing research for airborne radar systems until 1976; he has had responsibility for applications of c.c.d.s in systems since then. He is currently involved in image processing research.

John White received the B.Sc. degree in electronics in 1974 and the Ph.D. degree in 1977 for his investigation into infra-red emission from semiconductors, both from Southampton University. Since 1977 he has been employed at RSRE and has had technical responsibility for the design and device physics of a number of m.o.s. and charge-coupled devices. Dr White's present principal research interests are in the field of small m.o.s. devices.

John Keen was awarded a State Scholarship and read physical metallurgy at the University of Birmingham, graduating with a B.Sc. degree in 1958. In 1960 he qualified for an M.Sc. in industrial metallurgy with a thesis on the electrodeposition of zinc. He joined RSRE from University, working initially on the growth of magnetic thin films for computer storage memories. Subsequent research was on process-induced defects in silicon integrated circuits and on the fabrication of high-power, solid-state microwave devices in gallium arsenide. He is currently a Principal Scientific Officer with responsibility for fabrication and technology research on silicon devices and circuits and for running the silicon i.c. process facility.

At the 1973 IEEE International Solid State Circuits Conference, Mr Keen received an outstanding contributed paper award for a paper on 'Nondestructive techniques for electrically-testing dielectric layers on integrated circuits'. From 1975 to 1977 he was a Visiting Fellow in the Department of Electrical and Electronic Engineering at the University of Newcastle and since 1976 has been a member of the Editorial Board of the journal *Surface Technology*.

Jethro Hill read physics at the University of Nottingham and obtained his B.Sc. in 1970 and a Ph.D. in 1973. His thesis, and three years subsequent fellowship work, was devoted to the study of the tunnelling rotation of the methyl group and its interaction with lattice vibrations using n.m.r. and e.s.r. and thermal relaxation measurements. After a short period studying the motion of yarn on knitting machines, Dr Hill joined RSRE in 1977. Here he took responsibility for the physics and design of charge-coupled devices, taking particular interest in non-linear transient charge injection. His present work involves the physics and development of novel submicron devices.

Daniel McCaughan graduated with a B.Sc. degree in physics from the Queen's University of Belfast in 1964, and was awarded his Ph.D. degree in 1968 from the same University for a thesis on ion bombardment of metal surfaces. He was a Member of Technical Staff at Bell Laboratories, Murray Hill, from 1968 to 1974, where he worked on problems related to silicon device technology. From 1974–1976 he was a Principal Research Fellow at RSRE, Malvern, where he was engaged in research into the physics and technology of charge-coupled devices. Since 1976 Dr McCaughan has been leader of the c.c.d. and silicon devices group at RSRE. His research interests include ion transport phenomena in solids, plasma interactions with materials, silicon device technology, charge-coupled devices, and m.o.s. structures and devices. He has published over 50 papers and has five patents in these fields.