UDK 621.3:(53+54+621+66)(05)(497.1)=00 ISSN 0352-9045

INFORMACIJE MIDEM

Strokovno društvo za mikroelektroniko elektronske sestavne dele in materiale

1◦2001

Strokovna revija za mikroelektroniko, elektronske sestavne dele in materiale
Journal of Microelectronics, Electronic Components and Materials

INFORMACIJE MIDEM, LETNIK 31, ŠT. 1(97), LJUBLJANA, marec 2001

UDK 621.3:(53+54+621+66)(05)(497.1)=00 ISSN 0352-9045

INFORMACIJE MIDEM 1 ◦ 2001

INFORMACIJE MIDEM LETNIK 31, ŠT. 1(97), LJUBLJANA, MAREC 2001
INFORMACIJE MIDEM VOLUME 31, NO. 1(97), LJUBLJANA, MARCH 2001

Revija izhaja trimesečno (marec, junij, september, december). Izdaja strokovno društvo za mikroelektroniko, elektronske sestavne dele in materiale - MIDEM.
Published quarterly (march, june, september, december) by Society for Microelectronics, Electronic Components and Materials - MIDEM.

Naslov uredništva
Headquarters
Uredništvo Informacije MIDEM
Elektrotehniška zveza Slovenije
Dunajska 10, 1000 Ljubljana, Slovenija
tel.: + 386 (0)1 51 12 221
fax: + 386 (0)1 51 12 217
e-mail: Iztok.Sorli@guest.arnes.si
http://paris.fe.uni-lj.si/midem/journal.htm

Letna naročnina znaša 12.000,00 SIT, cena posamezne številke je 3000,00 SIT. Člani in sponzorji MIDEM prejemajo Informacije MIDEM brezplačno.
Annual subscription rate is DEM 200, separate issue is DEM 50. MIDEM members and Society sponsors receive Informacije MIDEM for free.

Znanstveni svet za tehnične vede I je podal pozitivno mnenje o reviji kot znanstveno strokovni reviji za mikroelektroniko, elektronske sestavne dele in materiale. Izdajo revije sofinanci rajo Ministrstvo za znanost in tehnologijo in sponzorji društva.
Scientific Council for Technical Sciences of Slovene Ministry of Science and Technology has recognized Informacije MIDEM as scientific Journal for microelectronics, electronic components and materials.
Publishing of the Journal is financed by Slovene Ministry of Science and Technology and by Society sponsors.

Znanstveno strokovne prispevke objavljene v Informacijah MIDEM zajemamo v podatkovne baze COBISS in INSPEC.

Prispevke iz revije zajema ISI® v naslednje svoje produkte: Sci Search®, Research Alert® in Materials Science Citation Index™

Scientific and professional papers published in Informacije MIDEM are assessed into COBISS and INSPEC databases.

The Journal is indexed by ISI® for Sci Search®, Research Alert® and Material Science Citation Index™

Po mnenju Ministrstva za informiranje št.23/300-92 šteje glasilo Informacije MIDEM med proizvode informativnega značaja.

Grafična priprava in tisk
Printed by
BIRO M, Ljubljana

Naklada
Circulation
1000 izvodov
1000 issues

Poštnina plačana pri pošti 1102 Ljubljana
Slovenia Taxe Percue

UDK621.3:(53+53+621+66), ISSN0352-9045 Informacije MIDEM 31(2001)1, Ljubljana

VSEBINA **CONTENT**

37th INTERNATIONAL CONFERENCE ON MICROELECTRONICS, DEVICES AND MATERIALS

and the WORKSHOP on

OPTOELECTRONIC DEVICES AND APPLICATIONS

MIDEM

CONFERENCE 2001

IMAPS Slovenia Chapter

Announcement and Call for Papers

October 10. - 12. 2001
Bohinj, SLOVENIA

http://paris.fe.uni-lj.si/midem/conf2001/

Elektrotehniška Zveza Slovenije

IEEE Slovenia Section

PERCOLATION THEORY AND ITS APPLICATION IN MATERIALS SCIENCE AND MICROELECTRONICS (Part I – Theoretical description)

Andrzej DZIEDZIC[1)], Andrei A. SNARSKII[2)]

1) Institute of Microsystem Technology, Wroclaw University of Technology, Poland
2) Department of General and Theoretical Physics, National Technical University of Ukraine, Kiev, Ukraine

Keywords: percolation, percolation system, percolation threshold, critical exponent, effective conductivity, 1/f noise, effective noise intensity, weak nonlinearity, strong nonlinearity, voltage susceptibility, third harmonic, normalised amplitude of third harmonic, temperature dependence of resistance, percolationlike system, exponentially wide spectrum of resistances, continuum percolation, Swiss-cheese model

Abstract: Percolation theory permits to characterise (calculate) the effective properties of random inhomogeneous two-phase systems with comparable concentration of both phases (near the percolation threshold) but with significant difference of their properties. This paper presents the critical behaviour of various kinetic phenomena (conductivity, 1/f noise, weak and strong nonlinearity, third harmonic generation, and temperature dependence of resistivity). These quantities can be described analytically using for example hierarchical model of percolation structure. The characteristic critical indexes are dependent on conductivity and correlation length critical exponents. Possible application of percolative theoretical description for systems with exponentially broad or disordered continuum spectrum of properties is presented, too. The nonelectrical effective properties could be analysed by methods of percolation theory because of analogy between the quasistatic electrical and other physical fields.

Teorija perkolacije in njena uporaba v znanosti o materialih in mikroelektroniki (Prvi del - Teorija)

Ključne besede: fizika, kemija, perkolacija, prag perkolacije, sistem perkolacije, teorija perkolacije, znanost o materialih, mikroelektronika, eksponenti kritični, prevodnost električna efektivna, šum 1/f, intenzivnost šuma efektivnega, nelinearnost šibka, nelinearnost močna, susceptibilnost napetostna, harmonske tretje, amplitude harmonskih tretjih normalizirane, odvisnost temperaturna upornosti električne, sistem podoben perkolacijskemu, spekter upornosti električnih širok eksponencialno, perkolacija neprekinjena, model sira švicarskega

Izvleček: Teorija perkolacije dovoljuje izračun lastnosti naključnih dvofaznih sistemov s primerljivima koncentracijama obeh faz (blizu perkolacijskega praga), pri čemer imata obe fazi vsaka zase različne lastnosti. V prispevku prikazujemo vedenje različnih kinetičnih parametrov, kot so prevodnost, 1/f šum, nelinearnost, generacija tretje harmonske frekvence in temperaturna odvisnost upornosti. Omenjene količine lahko predstavimo v analitični obliki z uporabo hierarhičnega modela perkolacijske strukture. Predstavimo tudi možno uporabo teorije perkolacije pri opisu sistemov s širokopasovnim eksponentnim ali neurejenim kontinuiranim spektrom lastnosti. Neelektrične lastnosti lahko analiziramo s pomočjo metod perkolacijske teorije zaradi analogije med kvazistatičnim električnim poljem in drugimi fizikalnimi polji.

Introduction

The percolation problem was formulated for the first time almost 45 years ago by Broadbent and Hammersley /1/. Since that time the idea and methods of percolation theory were applied into many areas of physics, chemistry as well as other basic and applied sciences. The original results based on percolation theory can be found in numerous papers. Therefore preparation of a complete bibliography devoted to this topics seems almost unrealisable. However beginner in such area could find some interesting books or review papers, for example /2-9/.

The so-called hierarchical model of percolation structure (HMPS) appeared during recent years. This model permits to describe analytically various properties of macroscopically disordered media near the percolation threshold – for example resistivity (also Hall effect), 1/f noise, electrical breakdown, nonlinear properties of composites and many others. This review will be devoted to the above mentioned phenomena. One should note that we will discuss experimental, analytical and numerical results received very recently – it means that they were not summarised in books and papers mentioned above.

1. Effective conductivity near p_c

Experimental and numerical investigations have shown, that effective conductivity σ_e is an analogous of order parameter in theory of phase transitions where temperature T is replaced by concentration of well-conducting phase – p and critical temperature T_c is replaced by percolation threshold - p_c. Based on the above analogy Efros and Shklovskii /10,11/ used scaling formula for σ_e

$$\sigma_e(\tau,h) = \sigma_1 h^s F(\tau/h^u) \qquad (1.1)$$

where $h = \sigma_2/\sigma_1$ - distance from percolation threshold, $\sigma_2 << \sigma_1, \sigma_e$ - local conductivity, and F(z) - scaling function

$$F(z \to \infty) \propto z^t, \quad F(z \to -\infty) \propto z^{-q}$$

$$F(z \to 0) \propto 1 \qquad (1.2)$$

where t and q – critical conductivity exponents and only the basic (single) components of sequence decomposition in relation to scaling are given in (1.2).

According to (1.1) and (1.2) there are three ranges of universal behaviour of effective conductivity, where separate equations describe an universal behaviour of effective conductivity – above ($p > p_c$), below ($p < p_c$) and in the vicinity ($p \approx p_c$) of percolation threshold

$$\sigma_e = \sigma_1\sigma_2(D_0 + D_1 h^{-\frac{1}{t+q}}|\tau| + \ldots), |\tau| \le \Delta$$

$$\sigma_e = \sigma_1\tau^t(A_0 + A_1 h\tau^{-(t+q)} + \ldots), p > p_c, \tau >> \Delta$$

$$\sigma_e = \sigma_2\tau^{-q}(B_0 + B_1 h|\tau|^{-(t+q)} + \ldots), p < p_c, |\tau| >> \Delta \qquad (1.3)$$

As it is visible the above equations consist on only the basic components but also smaller ones. A_i, B_i and D_i indicate constants, which according to absolute value are almost equal to 1. It is interesting to become familiar with conception of smearing region, $\Delta = (\sigma_1^q\sigma_2^t)^{1/t+q}$ - it is such $|\tau|$, where good and bad conductive phase possesses the same contribution into the effective conductivity $\sigma_1\Delta^t = \sigma_2\Delta^{-q}$ (qualitative behaviour of σ_e is shown in Fig. 1).

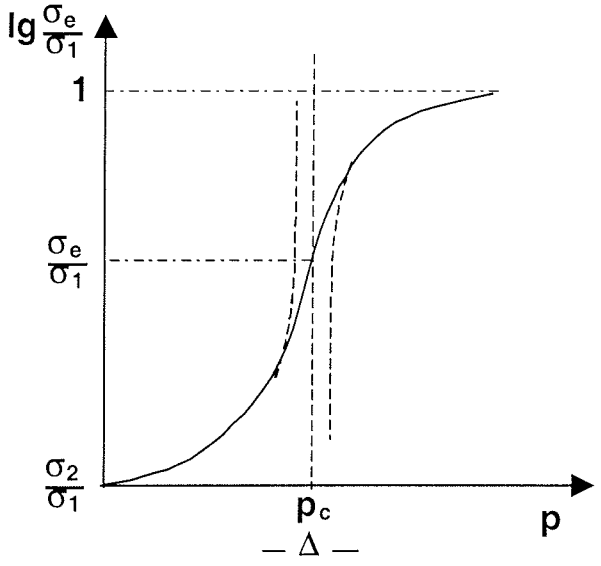


Fig. 1. Normalised conductivity versus good conductor concentration in two-phase percolation system

Many various models, in general based on percolation idea, have been used for explanation of σ_e shape. In the first one (which by occasion is the simplest) it is assumed that for $p > p_c$ case it is enough to consider number of single connected bonds (SCB) at the correlation length ξ ($\xi \propto |\tau|^{-\nu}$) /12,13/. This is so-called "bridge" with resistance R_1 consisting of seriously connected unit resistances from the first phase r_1, where $r_1 = (1/\sigma_1)a_0^{d-2}$, a_0 – minimal dimension in the system (for example mean size of composite grains or connection length in bond problem), $d = 2, 3$ – Fig. 2 (left). For analogous model, but below the percolation threshold ($p < p_c$) /14/, it was assumed that number of single disconnected bonds (SDCB) i.e. so-called interlayer (with resistance R_2) consisted of parallel connected unit resistances r_2 made from the second phase is the basic element – Fig. 2 (right).

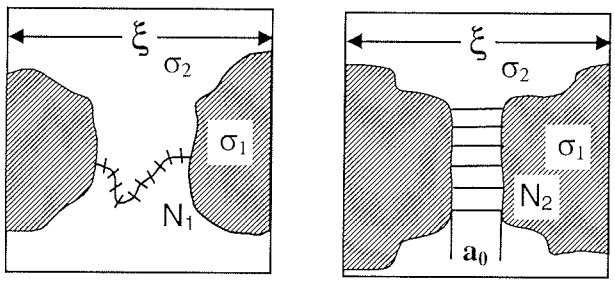


Fig. 2. Graphical representation of bridge ($p > p_c$) and interlayer ($p < p_c$) in hierarchical model of percolation structure (HPMS)

The metallic bridge (R_1) and dielectric interlayer (R_2) resistances are dependent on number of unit resistances of metallic and dielectric phases

$$R_1 = r_1 N^{\alpha_1}, \qquad R_2 = r_2 N^{\alpha_2} \qquad (1.4)$$

Based on theoretical and probability analyses it has been assumed in many papers (for example in /10-17/) that $\alpha_1 = \zeta_R = 1, \alpha_2 = \zeta_G = 1$ .However the critical conductivity exponents t and q, calculated in this model, have been almost equal to each other but their values do not agree with results of numerical calculations. Moreover it is possible to find some cases where such a simple model leads to contradictory results. For example it is shown in /2/ that the correlation length is increased faster than bridge length ($a_0 N_1$) in 2D system when $\alpha = 1$ at $\tau \to 0$.

Much more reliable results could be obtained based on so-called HMPS. Its idea has been presented in /18-20/. According to this approach the values of α_1 and α_2 are calculated based on t and q values (these quantities are considered as known) both below and above percolation threshold. It is assumed in this process that conduction process takes part both in good

as well as bad conducting phases – Fig. 3. When $\sigma_2/\sigma_1 \to 0$ HPMS is transferred into the standard model, discussed a little earlier.

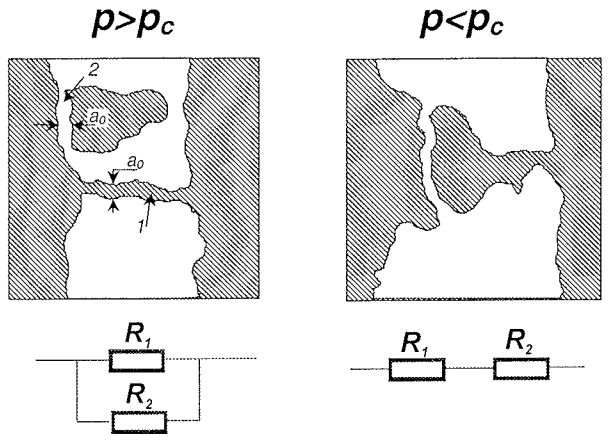


Fig. 3. Second step in the hierarchy of HPMS above (left) and below (right) percolation threshold together with proper electrical equivalent circuit (hatched area – good conductor)

It has been shown in /18-20/ that it is possible on the basis of this hierarchical model to write down a self-consistent equation. In the case of effective conductivity this equation in symbolic representation has the following form – Fig. 4.

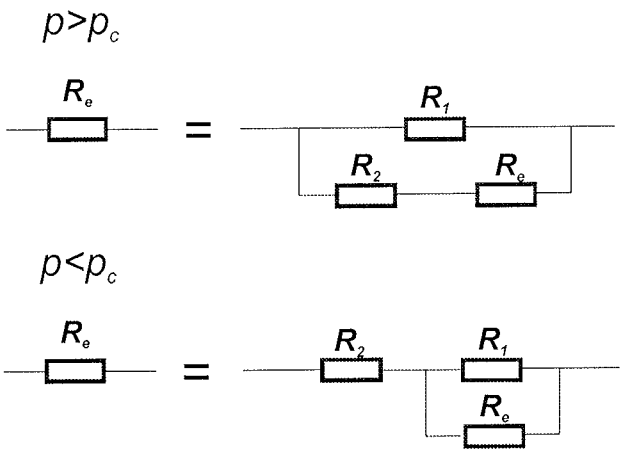


Fig. 4. Analogues of Dyson equation above (top) and below (bottom) percolation threshold

During closing to the percolation threshold ($|\tau| \to 0$) the bridge resistance R_1 is increased whereas interlayer resistance R_2 is decreased. Both resistances are equal in the smearing region but when Eq. (1.3a) or (1.3b) are obligatory then $R_1 << R_2$. It means that dielectric interlayer gives small contribution into the effective conductivity σ_e above p_c whereas the bridge - below p_c. Therefore it could seem that counting and regarding of so small contributions, especially outside the smearing region is needless because this does not lead to important properties of percolation system. However below the readers will find some examples in which manner elements affecting σ_e only insignificantly can decide about other properties of percolation systems.

2. 1/f noise

1/f noise is an universal phenomenon. It is characteristic for many physical (but not only) processes. The amplitude of that noise has especially large importance for composites /21,22/. The quantity of 1/f noise is characterised usually by effective noise intensity

$$C_e = \Omega S \tag{2.1}$$

where Ω - volume of analysed pattern, S - relative power spectral density

$$S = \frac{S_R}{R^2} \equiv \frac{\{\delta R \delta R\}}{R^2} \tag{2.2}$$

$S_R = \{\delta R \delta R\}$ – power spectral density, $\{\ldots\}$ – denotes the Fourier transform of the time correlation function.

Based on the situation that time fluctuations of resistance δR are spatially uncorrelated it is possible to describe (present) the effective noise intensity in terms of the Joule power dissipated in the inhomogeneous media

$$C_e = \frac{\langle C(Ej)^2 \rangle}{\langle \langle E \rangle \langle j \rangle \rangle^2} \tag{2.3}$$

where $\langle \ldots \rangle$ denotes volume averages.

The beginning of 1/f noise investigations in percolation systems is connected with scientific activity of Rammal /23/ (the reader interested in this topic can find more detailed bibliography of papers dealt with 1/f noise in /24/). We can tell that for the case of finite conductivity of both phases ($h = \sigma_2/\sigma_1 \neq 0$) C_e near percolation threshold can be written as

$$C_e(\tau > 0, \tau >> \Delta) = C_1 \tau^{-k} + C_2 h^2 \tau^{-w} \tag{2.4a}$$

$$C_e(|\tau| << \Delta) = C_1 h^{-k/(t+q)} + C_2 h^{-k'/(t+q)} \tag{2.4b}$$

$$C_e(\tau < 0, |\tau| >> \Delta) = C_2 |\tau|^{-k'} + C_2 |\tau|^{-w'} \tag{2.4c}$$

where C_1 and C_2 - noise intensities of first and second phase and values of k and k' i.e. critical exponents of 1/f noise are given in Table 1.

Table 1. Numerical estimates of the noise critical exponents κ and κ' (/23/ and references herein)

Critical index	Numerical simulations	Rigorous bounds
κ	1.47÷1.58	1.53÷1.60
κ'	0.55÷0.74	0.38÷1.02

According to HMPS the critical exponents w and w' can be expressed by k and k' in a simple manner as

$$w = k' + 2(t + q),\ w' = k + 2(t + q) \qquad (2.5)$$

For example, it is directly visible from (2.4) that above the percolation threshold but in the smearing region ($|\tau| \le \Delta$) the second phase could give higher income into the total 1/f noise of composite when C_2 is higher than C_1 /18/.

3. Weak nonlinearity

The deviation from linear Ohm's law is possible for large current densities. In the case of so-called weak nonlinearity (or weak cubic nonlinearity) the dependence between current density and electric field is given by the following formula

$$j = \sigma(r)E + \chi(r)|E|^2 E \qquad (3.1)$$

where χ - local nonlinear susceptibility. Of course (3.1) presents polynomial description of the field where the second constituent is significantly smaller than first one. The effective properties are used for description of weakly nonlinear system in the same manner as for linear system, this is

$$\langle j \rangle = \sigma_e \langle E \rangle + \chi_e \langle |E^2| \rangle \cdot \langle E \rangle \qquad (3.2)$$

As has been shown in /25,26/ there is analogy between behaviour of effective noise intensity C_e and effective nonlinear susceptibility. Problem becomes mathematically equivalent to the estimation of effective 1/f noise intensity, $\chi_e \propto C_e \sigma_e^2$ for the system with the local noise intensity $C(r) = \chi(r)/\sigma^2(r)$. Thus the critical behaviour of χ_e is given immediately from the equation describing the behaviours of the effective noise intensity and the effective conductivity

$$\chi_e(\tau > 0) = C_e(\tau > 0)\sigma_e^2(\tau > 0) = \chi_1 \tau^{2t-k} + \chi_2 h^4 \tau^{-2q-k'}$$

$$\chi_e(\tau < 0) = C_e(\tau < 0)\sigma_e^2(\tau < 0) = \chi_2 |\tau|^{-2q-k'} + \chi_1 h^4 |\tau|^{-w'-2q}$$

$$\chi_e(|\tau| \le \Delta) = C_e(|\tau| \le \Delta)\sigma_e^2(|\tau| \le \Delta) = \chi_1 h^{(2t-k)/(t+q)} + \chi_2 h^{-(2q+k')/(t+q)} \qquad (3.3)$$

The important question in analysis of nonlinear media effective properties is connected with Eq. (3.2) application range. Most often it is assumed /27-30/ that formula (3.2) is proper for

$$\langle j \rangle << \langle j \rangle_c \text{ and } \langle E \rangle << \langle E \rangle_c \qquad (3.4)$$

where so-called critical electric field $\langle E \rangle_c$ and critical current density $\langle j \rangle_c$ are defined as the value of field or current at which linear contribution (first constituent of (3.2)) is equal to nonlinear one, i.e. $\langle E \rangle_c = \sqrt{\sigma_e / \chi_e}$, $\langle j \rangle_c = \sqrt{\sigma_e^3 / \chi_e}$. Moreover the local criterion of Eq. (3.1) usability has been introduced in /31/. According to this attempt not only average but also local fields and currents (both in bridge and interlayer) should not exceed proper critical values

$$E_{loc} << E_c = \sqrt{\sigma_i / \chi_i},\ j_{loc} << j_c = \sqrt{\sigma_i^3 / \chi_i} \qquad (3.5)$$

where $i = 1, 2$ is related to first and second phase.

4. Third-harmonic generation

If a pure sinusoidal current (with frequency ω) flows through the symmetrical nonlinear medium, then the voltage that appears across the medium will contain odd harmonics (with frequencies 3ω, 5ω, ...). It has been shown that their amplitude is especially large in strongly nonlinear systems /32-36/. However the small amount of nonlinearity also affects this phenomenon, which appears for example due to local Joule heating. In this approach it is assumed that both components of the composite have finite temperature coefficient of resistance. The dissipated power (Joule heat) caused by current $j = j_0 \cos \omega t$ modulates medium conductivity with 2ω frequency and phase shift. It is well known that flow of pure sinusoidal current with frequency ω through the sample with resistance modulated with 2ω frequency results in odd harmonics generation. The amplitude of third harmonic $\langle E \rangle_{3\omega}$

$$\langle E \rangle = \rho_e \langle j_0 \rangle \cos \omega t + \langle E \rangle_{3\omega} \cos(3\omega t + \phi) \qquad (4.1)$$

can be expressed with the aid of 1/f noise amplitude C_e. Normalised amplitude of third harmonic $B_{3\omega}$

$$B_{3\omega} = \frac{\langle E \rangle_{3\omega}}{\langle j_0 \rangle^3} \qquad (4.2)$$

agrees with $\rho_e^2 C_e$ (with accuracy to inessential numerical multipliers) when - in formula for C_e - factor C_i is changed by temperature coefficient of resistivity of *i*-th phase - β_i.

Generalisation of expression for $B_{3\omega}$ given in /32-35/ for the case $h = \sigma_2/\sigma_1 \neq 0$ is presented in /36/ and we obtain the following Equation

$$B_{3\omega}(\tau > 0) \propto \beta_1 \left(\frac{\rho_e}{\rho_1}\right)^{\frac{\kappa}{t}+2} + \beta_2 \left(\frac{\rho_1}{\rho_2}\right)^2 \left(\frac{\rho_e}{\rho_1}\right)^{\frac{\kappa'+2(t+q)}{t}+2},$$

$$B_{3\omega}(|\tau| \leq \Delta) \propto \beta_1 \left(\frac{\rho_e}{\rho_1}\right)^{\frac{\kappa}{t}+2} + \beta_2 \left(\frac{\rho_e}{\rho_2}\right)^{\frac{\kappa'}{t}+2},$$

$$B_{3\omega}(\tau < 0) \propto \beta_2 \left(\frac{\rho_e}{\rho_1}\right)^{\frac{\kappa'}{q}+2} + \beta_1 \left(\frac{\rho_1}{\rho_2}\right)^2 \left(\frac{\rho_e}{\rho_2}\right)^{-\frac{\kappa+2(t+q)}{q}+2} \quad (4.3)$$

where the dependence from τ in (4.3) is connected with effective resistivity $\rho_e = 1/\sigma_e$ which of course is different for various regions.

5. Strong nonlinearity

Contrary to weak nonlinearity case the current-voltage characteristics of strongly nonlinear medium are not linear even for very weak fields. The medium with the following current-voltage relation

$$j = \chi |E|^{\beta-1} E \quad (5.1)$$

has been analysed in /38,39/ where the behaviour of effective nonlinear susceptibility χ_e near percolation threshold has been described for the case of ideal insulator ($\chi_2 = 0$). The opposite case, *i.e.* $1/\chi_1 = 0$. The calculus of the χ_2/χ_1 ratio when both phases exhibit identical nonlinearity relation (the same parameter β) has been presented in [40]. And the most general case ($\gamma = \beta_1 \neq \beta_2 = \beta$) has been analysed in /41/. When the current-voltage characteristics are described by the following formulas

$$E = \rho_1 |j|^{\gamma-1} j \qquad j = \sigma_1 |E|^{\frac{1-\gamma}{\gamma}} E$$

$$j = \sigma_2 |E|^{\beta-1} E \qquad E = \rho_2 |j|^{\frac{1-\beta}{\beta}} j \quad (5.2)$$

then three field regions can be distinguished for appropriate β and γ (Fig. 5).

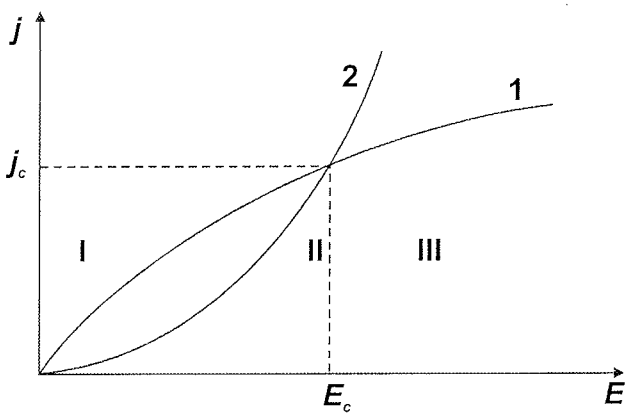


Fig. 5. Current-voltage characteristic of first (1) and second (2) phase in strongly nonlinear system

The percolation treatment is possible in region I ("strongly nonuniform" medium) where

$$\langle j \rangle = \sigma_1 \tau^{\tilde{t}} |\langle E \rangle|^{\frac{1}{\gamma}} \frac{\langle E \rangle}{|\langle E \rangle|} \left[1 + \frac{\sigma_2}{\sigma_1} \tau^{-\tilde{\varphi}} |\langle E \rangle|^{\beta-1/\gamma}\right], \quad p > p_c$$

$$\langle j \rangle = \left(\sigma_1^{\tilde{q}} \sigma_2^{\tilde{t}}\right)^{\frac{1}{\tilde{\varphi}}} |\langle E \rangle|^{\frac{1}{\tilde{\varphi}}\left(\beta\tilde{t}+\frac{\tilde{q}}{\gamma}\right)} \frac{\langle E \rangle}{|\langle E \rangle|}, \quad |\tau| \leq \Delta$$

$$\langle j \rangle = \sigma_2 \tau^{-\tilde{q}} |\langle E \rangle|^{\beta} \frac{\langle E \rangle}{|\langle E \rangle|} \left[1 - \left(\frac{\sigma_2}{\sigma_1} \tau^{-\tilde{\varphi}} |\langle E \rangle|^{\beta-1/\gamma}\right)^{\gamma}\right], \quad p < p_c \quad (5.3)$$

where $\tilde{\varphi} = \tilde{q} + \tilde{t}$ and

$$\tilde{t} = \frac{t + \nu(d-1)(\gamma-1)}{\gamma}, \quad \tilde{q} = q + \nu(\beta-1) \quad (5.4)$$

The size of smearing region Δ in strongly nonlinear system is field-dependent

$$\Delta = \left(\frac{\sigma_2}{\sigma_1} |\langle E \rangle|^{\beta-1/\gamma}\right)^{\frac{1}{\tilde{\varphi}}} \quad (5.5)$$

Moreover, let's note that current-voltage characteristic of system, composed of strongly nonlinear phases, becomes linear for specified values of β and γ ($(\beta\tilde{t} + \tilde{q}/\gamma)/\tilde{\varphi} = 1$).

6. Temperature dependence of resistance

The temperature dependence of resistivity (resistance), usually characterised by means of differential temperature coefficient of resistivity (*TC*ρ) or resistance (*TCR*), is one of the most important features of composite materials or devices based on them.

The effective temperature coefficient of resistivity (resistance), i.e. $TC\rho_e = \frac{1}{\rho_e}\frac{d\rho_e}{dT}$ or $TCR_e = \frac{1}{R_e}\frac{dR_e}{dT}$ ($\rho_e = 1/\sigma_e$), for percolation system created by two-phase medium with finite conductivity ratio ($h = \rho_1/\rho_2 \neq 0$) has been found in /42/. Similarly as in the case of the other above-considered properties the analytical formulas have been worked out for three concentration subranges.

$$TC\rho_e = TC\rho_1 + (\rho_1/\rho_2)\tau^{-(t+q)}TC\rho_2, p > p_c \tag{6.1}$$

$$TC\rho_e = A\cdot TC\rho_1 + B\cdot TC\rho_2 - D(TC\rho_1 - TC\rho_2)(\frac{\rho_1}{\rho_2})^{\frac{1}{t+q}}\tau, \quad |\tau| \leq \Delta \tag{6.2}$$

$$TC\rho_1 = TC\rho_2 + (\rho_1/\rho_2)\tau^{-(t+q)}TC\rho_1, p < p_c \tag{6.3}$$

In the above equations we have $TC\rho_i = \frac{1}{\rho_i}\frac{d\rho_i}{dT}$ (i = 1, 2) – temperature coefficient of resistivity of i-th phase and A, B, D – constants (equal to about 1).

7. Continuum problems

It has been assumed for all so-far analysed cases, that the problem of current distribution in system can be transferred to model, where the random resistance distribution of first and second phase r_1 and r_2 is given as

$$f(r) = p\delta(r - r_1) + (1-p)\delta(r - r_2) \tag{7.1}$$

(p – concentration of first phase, $\delta(...)$ - Dirac function.

But the case, where distribution function can be written as

$$f(r) = p(1-\alpha)r^{\alpha}, 1/r_2 = 0 \tag{7.2}$$

has been examined in /43/. It has been shown that critical index stops to be universal (it is said it goes to the second universality class)

$$t = t_0 + \alpha/(1-\alpha) \tag{7.3}$$

where t_0 – standard critical conductivity index above percolation threshold.

The case when the spectrum of resistances is continuous and exponentially broad /44/

$$r = r_0 e^{-\lambda x}, \lambda >> 1, \tag{7.4}$$

where $x \in (0, 1)$ is a random variable with smooth probability distribution $D(x)$, is no less interesting. The problem with a continuous spectrum of resistance distribution is not a straightforward percolation problem – it does not exhibit the percolation threshold at which one of the two phases forms an infinite percolating cluster because the phases themselves do not exist. However, there is a method which simplifies the exponential distributed resistances problem to the standard two-phase percolation problem /45-47/ and makes it possible to determine the principal system regularity, this is to find a critical index of effective percolation conductance. The general assumption of this method is that all resistances with a random variable between x and 1 are considered as one phase. In a crude approximation the network effective conductivity is described by the largest resistance, at which this phase becomes infinite. This is related to the percolation threshold in a classical percolation, i.e. from

$$\int_{x_c}^{1} D(x)dx = p_c \tag{7.5}$$

it is possible to calculate x_c and next to find the largest resistance, which defines (with accuracy to the preexponential factor in σ_e) the resistance of the whole system,

$$r \propto r_0 e^{-\lambda x_c} \tag{7.6}$$

It is possible to consider the above problem analogously but to start from the reverse side. Lets take a system with an exponential broad spectrum of resistance and keep in mind site of particular resistances in the network. Then we replace them in the network by "zero-resistivity" connection, and again put resistances into their previous position in the network but according to proper sequence starting from the smallest one. This process is carried on till appearance the resistance, which disconnect the current flow through the "zero-resistivity" phase. We can tell that this critical resistance specifies the resistance of such system (with accuracy to preexponential factor). The details of such

treatment are presented in /48-51/. It could seem that applied critical resistance search methods give opposing results /50,51/. However this contradiction is removed by assumption that the system is in smearing region just as the critical resistance is included in the network. Generalisation of two-phase percolation model in smearing region /52/ for systems with exponentially broad resistance spectrum lead to the following expression for effective conductivity (for simplicity it has been assumed, that $D(x) = 1$)

$$\sigma_e = \frac{A}{a_0 r(x_c)} \lambda^{-y} \tag{7.7}$$

where a_0 – minimal characteristic dimension in the system (of order of lattice cell), A – variable with a weak dependence on λ ($A \propto (\ln\lambda)^{\alpha_1+\alpha_2+\nu(d-2)}$), d – dimensionality of the problem, and critical exponent y is equal

$$y = \frac{\alpha_1 - \alpha_2 + 2\nu(d-2)}{2} \tag{7.8}$$

In terms of widely accepted values of $\alpha_1 = \varsigma_R = 1$ and $\alpha_2 = \varsigma_G = 1$ Eq. (7.8) reduces to

$$y = \nu(d-2) \tag{7.9}$$

The above result has been shown for the first time in [48]. Choice $\alpha_1 = \varsigma_R = t - \nu(d-2)$ and $\alpha_2 = \varsigma_G = q + \nu(d-2)$ gives very similar numerical results; for more details please see /53,54/)

The model described in /51/, using network with exponential distribution of properties, permits to find the behaviour of many other physical quantities. Moreover, even if for example resistance distribution is not exponential but power one $r = r_0 x^{-\lambda}$ and we have somewhat different formula for effective conductivity

$$\sigma_e = \frac{A}{a_0 r_0} x_c \lambda^{-y} \tag{7.10}$$

The critical index y from Eq. (7.10) is still given by Eq. (7.9).

There are no basic troubles in characterisation of more complex quantities than effective conductivity using percolationlike model. However it is necessary to make supplementary assumption related to local properties of these quantities. For example, calculation of 1/f noise in exponentially distributed systems demands generalisation of Hooge hypothesis /21/, according to which $C = \alpha / \sigma$ (α - so-called Hooge parameter). It is logically to assume, that for considered system with local conductivity $\sigma(x) \propto e^{-\lambda x}$

$$C(x) = \alpha / \sigma(x), \tag{7.11}$$

This is in agreement with empirical Hooge law – system (device) with higher resistivity (more precisely with lower concentration of charge carriers) is characterised by larger noise intensity.

The effective noise intensity of system with an exponentially wide spectrum of resistances obeys the form

$$C_e \propto \lambda^m e^{-\lambda x_c} \tag{7.12}$$

where exponent m is given as

$$m = y + 2\nu \tag{7.13}$$

As has been mentioned earlier the exponent y is related to the correlation length exponent ν by Eq. (7.9). Therefore

$$m = d\nu \tag{7.14}$$

The above calculations have been generalised in /24,54,55/ for situation when

$$C(x) = \alpha / \sigma^{\theta}(x) \tag{7.15}$$

(for $\theta = 1$ we have standard Hooge formula (7.11)). Very interesting feature of the exponent m has been observed for $0 < \theta < 2$; m is independent on θ parameter, this is

$$C_e \propto \sigma_0 e^{-\lambda\theta x_c} \lambda^{m_\theta}, \quad m_\theta = \nu d \tag{7.16}$$

Lets note that even if phenomenological Hooge formula is locally true, i.e. $C(x)\sigma^{\theta}(x) = \text{const}$, it is broken for the whole system, that is $C_e \sigma_e \neq const$.

Except of effective conductivity and noise intensity investigations of temperature behaviour /42/ or third harmonic generation /36/ also have been analysed in systems with exponentially broad spectrum of resistances. It has been shown that normalised amplitude of third harmonic $B_{3\omega}$ for such systems is related very simply to its effective conductivity ρ_e

$$B_{3\omega} \propto \rho_e^3 \tag{7.17}$$

The successive model with disordered continuum spectrum of resistances has been presented for the first time in /56,57/. This is so-called Swiss-cheese

i.e. a disordered continuum system where spherical holes are randomly placed in a uniform transport medium. The distance between spherical voids is unrestrictedly small. This means that so-called microgeometry, in other words current distribution in narrow necks between mentioned spherical holes becomes very important. Such a model has been analysed based on percolation approach and it has been proved in already mentioned papers /56,57/ that critical conductivity exponents for Swiss-cheese model and corresponding indices in a discrete lattice differ in value and depend on microgeometry details (shape of inclusions). For example, when $p > p_c$ and $\sigma_2 = 0$ then

$$t = t_{standard} + y \qquad (7.18)$$

where exponent y is dependent on kind of voids. For random-void model and $\sigma_2 = 0$ we have $y = 0$ in the case of 2D system and $y = ½$ for 3D medium.

Moreover there are yet other classes of continuum model, namely potential model (space between voids is not limited by spherical area but by hyperboloidal one /57/, blue-cheese model /58/ and so on. Microstructure in fact affects not only effective conductivity but also other properties such as dielectric (e.g. in $d = 3$ critical exponent of effective dielectric constant in Swiss-cheese model differs in standard one by 5/2 /56/), electrical and mechanical destruction /58/ and the like. It appears, that microgeometry influences behaviour of 1/f noise near percolation threshold /59/. Last but not least matter of this paper is that we have to be conscious of analogies between various physical fields (presented for example in /60/). Therefore percolation theory and analysis can be applied not only in calculation of electrical effective properties but also electrostatic, magnetic, thermal, fluidic and mechanical ones. Chosen examples of percolation or percolationlike systems, which have been studied experimentally as well as some numerical simulations performed with the aid of approaches given in this paper will be presented and discussed in second part of this article /61/.

Acknowledgement

This work was supported in part by Polish State Committee for Scientific Research, Grant no 8T11B 055 19.

References

/1/ S.R. Broadbent, J.M. Hammersley, Percolation processes. I. Crystals and mazes, Proc. Cambridge Philos. Soc., vol.53 (1957), p.629-641

/2/ B.I. Shklovskii and A.L. Efros, Electronic Properties of Doped Semiconductors, Springer-Verlag, Berlin, 1984

/3/ A.L. Efros, Fizika i gieometria biezporiadka, Moscow, Nauka 1982 (in Russian)

/4/ D. Stauffer, A. Aharony, Introduction to Percolation Theory, Taylor and Francis, London-Washington, 1992

/5/ M. Sahimi, Application of Percolation Theory, Taylor and Francis, London-Washington 1994

/6/ M. Sahimi, Flow phenomena in rocks: from continuum models to fractals, percolation, cellular automata, and simulated annealing, Rev. Mod. Phys., vol.65 (1993), 1395

/7/ D.J. Bergman, D. Stroud, Physical properties of macroscopically inhomogeneous media, Sol. St. Phys., vol.46 (1992), p.147-269

/8/ J.P. Clerc, G. Giraud, J.M. Laugier, J.M. Luck, The electrical conductivity of binary disordered systems, percolation clusters, fractals and related models, Adv. Physics, vol.39 (1990), p.191

/9/ T. Nakayama, K. Yakubo, R.L. Orbach, Dynamical properties of fractal networks: scaling, numerical simulation, and physical realizations, Rev. Mod. Phys., vol.66 (1994), p.381

/10/ A.L. Efros, B.I. Shklovskii, Critical behaviour of conductivity and dielectric constant near the metal-non-metal transition threshold, Phys. Stat. Sol. (b), vol.76 (1976), p.475-485

/11/ J.P. Straley, Critical exponents for the conductivity of random resistor lattices; Phys. Rev. B, vol.15 (1977), p.5733-5737

/12/ A.S. Skal, B.I. Shklovskii, Topology of infinite cluster in percolation theory and hopping conductivity theory, Sov. Phys. Semicond., vol.8 (1974), 1029

/13/ P.G. de Gennes, On a relation between percolation theory and the elasticity of gels, J. de Physique, vol.37 (1976), p.L1

/14/ T. Ohtsuki, T. Keyes, Conduction in random networks of supernormal conductors: Geometrical interpretation and enhancement of nonlinearity, J. Phys. A, vol.17 (1984),p.L559-L563

/15/ A. Coniglio, Thermal phase transition of the dilute s-state Potts and n-vector models at the percolation threshold, Phys. Rev. Lett., vol.46 (1981), p.250-253

/16/ A. Coniglio, H.E. Stanley, Screening of the deeply invaginated clusters and the critical behaviour of the random superconductivity network, Phys. Rev. Lett., vol.52 (1984), p.1069-71

/17/ D.C. Wright, D.J. Bergman, Y. Kantor, Resistance fluctuations in random resistor networks above and below the percolation threshold. Phys. Rev. B, vol.33 (1986), p.396

/18/ A.E. Morozovsky, A.A. Snarskii, Critical behavior of the 1/f noise in percolation systems, Sov. Phys. JETP, vol.68 (1989), p.1066

/19/ A.E. Morozovsky, A.A. Snarskii, Critical behavior of fracture stress in randomly inhomogeneous composites near the percolation threshold, JETP Lett., vol.52 (1990), p.244

/20/ A.E. Morozovsky, A.A. Snarskii, Multiscaling in randomly inhomogeneous media: effective conductivity, relative spectral density of 1/f noise, and higher-order moments, Sov. Phys. JETP, vol.75 (1992), p.366

/21/ S.M. Kogan, Nizkochastotnyj tokovyj shum so spektrom tipa 1/f v tviordych tielach, Uspiechi Fizicheskich Nauk, vol.145 (1985), p.285-328 (in Russian)

/22/ M.B. Weissman, 1/f noise and other slow, nonexponential kinetics in condensed matter, Rev. Mod. Phys., vol.60 (1987), p.537

/23/ R. Rammal, C. Tannous, P. Brenton, A.-M.S. Tremblay, Flicker (1/f) noise in percolation networks: a new hierarchy of exponents, Phys. Rev. Lett., vol.54 (1985), p.1718; 1/f noise in random resistor networks: fractals and percolating systems, Phys. Rev. A, vol.31 (1985), p.2662

/24/ A.A. Snarskii, A.E. Morozovsky, A. Kolek, A. Kusy, 1/f noise in percolation and percolation-like systems, Phys. Rev. E, vol.53 (1996), p.5596-5605

/25/ D. Stroud, P.M. Hui, Nonlinear susceptibilities of granular matter, Phys. Rev. B, vol.37 (1988), p.8719-8724

/26/ A. Aharony, Crossover from linear to nonlinear resistance near percolation, Phys. Rev. Lett., vol.58 (1987), p.2726

/27/ R. Blumenfeld, D.J. Bergman, Comment on "Nonlinear susceptibilities of granular matter", Phys. Rev. B, vol.43 (1991), p.13 682

/28/ P.M. Hui, Crossover electric field in percolating perfect-conductor – nonlinear-normal metal composites, Phys. Rev. B, vol. 49 (1994), p.15344

/29/ K.W. Yu, P.M. Hui, Percolation effects in two-component nonlinear composites: crossover from linear to nonlinear behavior, Phys. Rev. B, vol.50 (1994), p.13327

/30/ G.M. Zhang, Critical behavior of non-linear susceptibility in ran-

dom non-linear resistor networks, J. Phys. C, vol.8 (1996), p.6933
/31/ A.A. Snarskii, S.I. Buda, Critical fields and currents in a weakly nonlinear medium near the percolation threshold, Technical Physics, vol.43 (1998), p.619-621
/32/ M.B. Weissman, C.D. Dollinger, J.C. Garland, J. Appl. Phys., vol.52 (1981), p.3059
/33/ M.A. Dubson, Y.C. Hui, M.B. Weissman, J.C. Garland, Measurement of the fourth moment of the current distribution in two-dimensional random resistor networks, Phys. Rev. B, vol.39 (1989), p.6807-6815
/34/ Y. Yagil, G. Deutscher, Third-harmonic generation in semicontinuous metal films, Phys. Rev. B, vol.46 (1992), p.16115-16121
/35/ Y. Yagil, G. Deutscher, D.J. Bergman, The role of microgeometry in the electrical breakdown of metal-insulator mixtures, Int. J. Mod. Phys. B, vol.7 (1993), p.3353-3374
/36/ A.A. Snarskii, Generation of third harmonics in highly heterogeneous composites near flowing threshold, Tech. Phys. Lett., vol.21 (1995), p.1-3
/37/ A.A. Snarskii, K.V. Slipchenko, A.M. Satanin, The generation of the third harmonic in highly anisotropic media near the percolation threshold, Technical Physics, vol.45 (2000), p.194-198
/38/ S.W. Kenkel, J.P. Straley, Percolation theory of nonlinear circuit elements, Phys. Rev. Lett., vol.49 (1982), p.767-770
/39/ J.P. Straley, S.W. Kenkel, Percolation theory for nonlinear conductors, Phys. Rev. B, vol.29 (1984), p.6299-6305
/40/ A.E. Morozovsky, A.A. Snarskii, Effective properties of high anisotropic and non-linear metal/non-metal thin films, Int. J. Electr., vol.73 (1992), p.925-927
/41/ A.A. Snarskii, K.V. Slipchenko, V.A. Sevrukov, Critical behavior in two-phase, highly inhomogeneous composites, JETP, vol.89 (1999), p.788-796
/42/ A.A. Snarskii, A. Dziedzic, B. Licznerski, Temperature behaviour of percolation and percolation-like systems, Int. J. Electr., vol.81 (1996), p.363-370
/43/ P.M. Kogut, J. Straley, J. Phys. C, vol.12 (1979), p.2151
/44/ A. Miller, E. Abrahams, Impurity conduction of low concentrations, Phys. Rev., vol.120 (1960), p.745
/45/ V. Ambegaokar, B.I. Halperin, J.S. Langer, Hopping conductivity in disordered systems, Phys. Rev. B, vol.4 (1971), p.2612
/46/ M. Pollak, A percolation threshold of d.c. hopping conduction, J. Non-Cryst. Solids, vol.11 (1972) p.1
/47/ B.I. Shklovskii, A.L. Efros, Primiesnaja zona i provodimost kompensirovanych poluprovodnikov, Zh. Eksperimentalnoi i Teoreticheskoi Fiziki, vol.60 (1971), p.867 (in Russian)
/48/ P. Le Doussal, Percolationlike exponent for the conductivity of highly disordered resistor networks, Phys. Rev. B, vol.39 (1989), p.881-884
/49/ S. Tyc, B.I. Halperin, Random resistor network with an exponentially wide distribution of bond conductances, Phys. Rev. B, vol.39 (1989), p.877-880
50. A.E. Morozovskii, A.A. Snarskii, Conductivity critical exponent of exponentially distributed resistances, JETP Lett., vol.56 (1992), p.268-272
/51/ A.E. Morozovskii, A.A. Snarskii, Percolation description of the conductivity of random networks with a broad spectrum of the distribution of resistances, JETP, vol.77 (1993), p.959-965,
/52/ A.A. Snarskii, A.E. Morozovskii, Percolation structure model in the smearing region, Int. J. Electronics, vol.78 (1995), p.135-137
/53/ A. Kolek, A.A. Snarskii, A.E. Morozovskii, Structure of the percolation cluster and excess 1/f noise in systems with an exponentially broad spectrum of resistances, JETP, vol.81 (1995), p.490-495
/54/ A.A. Snarskii, A. Kolek, Excess 1/f noise in system with an exponentially wide spectrum of resistance and dual universality of the percolation-like noise exponent, JETP Lett., vol.63 (1996), p.651-656
/55/ A.A. Snarskii, A. Kolek, Double universality of 1/f noise percolation-like exponent in systems with exponentially wide spectrum of resistances, Physica A, vol.241 (1997), p.355-359
/56/ B.I. Halperin, S. Feng, P.N. Sen, Difference between lattice and continuum percolation transport exponents, Phys. Rev. Lett., vol.54 (1985), p.2391-2394
/57/ S. Feng, B.I. Halperin, P.N. Sen, Transport properties of continuum systems near the percolation threshold, Phys. Rev. B, vol.35 (1987), p.197-214
/58/ D. Sornette, Critical transport and failure in continuum crack percolation, J. Phys. France, vol.46 (1988), p.1365-1377
/59/ A.-M.S. Tremblay, S. Feng, P. Brenton, Exponents for 1/f noise near the a continuum percolation threshold, Phys. Rev. B, vol.33 (1986), p.2077-2080
/60/ A. Dziedzic, B.W. Licznerski, A.A. Snarskii, Application of percolation theory in description of composite materials, Proc. Conf. "Kompozyty`98 – Teoria i praktyka", Warszawa 1998, p.25-35 (in Polish)
/61/ A. Dziedzic, A.A. Snarski; Percolation theory and its application in materials science and microelectronics (Part II – Experiments and numerical simulations), to be published in Informacije MIDEM

Andrzej Dziedzic
Institute of Microsystem Technology, Wroclaw University of Technology,
Wybrzeze Wyspianskiego 27,
50-370 Wroclaw,Poland,
e-mail: adziedzic@pwr.wroc.pl

Andrei A. Snarskii
Department of General and Theoretical Physics,
National Technical University of Ukraine,
Prospect Peremoga 37, 252 056 Kiev, Ukraine,
e-mail: asnar@phys.carrier.kiev.ua

Prispelo (Arrived): 05.02.01 *Sprejeto (Accepted): 07.03.01*

SHIFTER DESIGNS FOR ASICs

Dalibor Grgec* and Željko Butković
Faculty of Electrical Engineering and Computing
University of Zagreb, Croatia

Keywords: computer science, shifters, phase shifters, logarithmic shifters, data buses, data paths, arithmetic operations, decoders, ALU, Arithmetic-Logic Unit, FPU, Fast Processing Units, ASIC, Application Specific Integrated Circuits, VLSI circuits, Very Large Scale of Integration circuits, CAD, Computer Aided Design, logical simulations, electrical simulations, MAGIC computer tools, IRSIM computer tools, SPICE computer tools, Simulation Program with Integrated Circuit Emphasis computer tools, transmission gates, restoring buffers, propagation delay, power dissipation

Abstract: This paper presents four versions of 32 bit shifter designs that can be used in ASICs, namely: barrel shifter and logarithmic shifter, each implemented with pass transistors and transmission gates. The circuits are designed in the standard MOSIS scalable CMOS n-well technology with the 0.8 μm minimal feature size fabrication process. The design procedure is thoroughly explained. The designs are logically and electrically simulated. They are compared according to the functionality, wafer area, number of transistors, delay and power dissipation. The usage and optimization guidelines are given.

Načrtovanje premikalnih podsklopov za integrirana vezja po naročilu

Ključne besede: računalništvo, pomikalniki, pomikalniki fazni, pomikalniki logaritmični, vodila podatkovna, operacije aritmetične, dekoderji, ALU enota aritmetično-logična, FPU enote obdelave hitre, ASIC vezja integrirana za aplikacije specifične, VLSI vezja integracije zelo visoke stopnje, CAD snovanje računalniško podprto, simulacije logične, simulacije električne, MAGIC orodja računalniška, IRSIM orodja računalniška, SPICE orodja računalniška, vrata prenosna, bufferji obnovitveni, zakasnitev razširjanja, stresanje moči

Izvleček : V prispevku predstavljamo štiri tipe 32-bitnih pomikalnih vezij, ki jih lahko uporabimo pri načrtovanju integriranih vezij po naročilu. To sta matrični in logaritemski premikalni vezji, pri čemer je vsako lahko izvedeno s prehodnimi tranzistorji, oz. prenosnimi vrati. Vezja so načrtovana v standarni MOSIS CMOS tehnologiji z n otokom in minimalnimi risanimi dimenzijami 0.8 μm. Podrobno smo razložili postopek načrtovanja. Opravili smo logično in električno simulacijo vezij ter jih med seboj primerjali glede na funkcionalnost, površino čipa, število tranzistorjev, hitrost in porabo. Podali smo napotke za uporabo in optimizacijo vezij.

1. INTRODUCTION

Shifting of binary numbers is an arithmetic operation required in many operations such as multiplication, division and bit-manipulation. Shifting is performed in specially designed circuits. Shifters are part of every contemporary datapath, usually located at the output of the Arithmetic-Logical Unit (ALU).

Shift operations can be classified into left-right, logical, arithmetical or circular shift (rotating). Usually shifting is implemented only in one direction, i.e., right. Shifting left by m bits is realized with a shift right of $n–m$ bits in an n-bit machine /1/. During the logical shift the LSB takes the value of a predefined input (usually 0/1 or bit-stream from an outer source). Arithmetical shift is a shift operation where the MSB, which represents the sign of the binary number, is preserved. Circular shift puts the LSB in the place of MSB and vice-versa. All sorts of shift operations are required in modern processing units /2, 3/.

According to implementation, shifters can be classified into shift-register (sequential logic) and flow-through (combinatorial logic) types. In the shift-register type, shift by one bit requires at least one machine cycle. In the flow-through type, the time required for shifting depends only on the circuit combinatorial delay and it is usually shorter than one machine cycle. The dominant shifter implementation in modern datapaths is the flow-through type. The flow-through shifters can further be classified into /4,6/:

- binary shifter,
- crossbar switch,
- barrel shifter,
- logarithmic shifter,
- other shifter implementations.

The common requirements set upon shifter implementations in modern datapaths are:

1. bus width of 32 or 64 bits (n bits in general),
2. performing $n \times n$ shift in one clock cycle,
3. performing many types of shifter operations according to control signals (left-right, logical, arithmetic or circular shift, masking, etc.)
4. separate control signals, usually perpendicular to direction of data,
5. coded control signals,
6. low propagation delay and no degradation of output signal electrical characteristics,
7. compatibility with the rest of the datapath.

Barrel and logarithmic shifter implementations satisfy these requirements best, and thus are two most frequently used shifters /4, 5, 6/.

* Now with Institute for Theoretical Electrical Engineering and Microelectronics, University of Bremen, Germany

Binary shifter only performs a one-place left-right shift, and crossbar switch is a universal circuit and a basis for a barrel shifter design.

In this paper, designs of barrel and logarithmic shifter in most common logic styles and comparisons of their performance are presented. The general block schematic of the designed shifters is shown in Fig. 1.

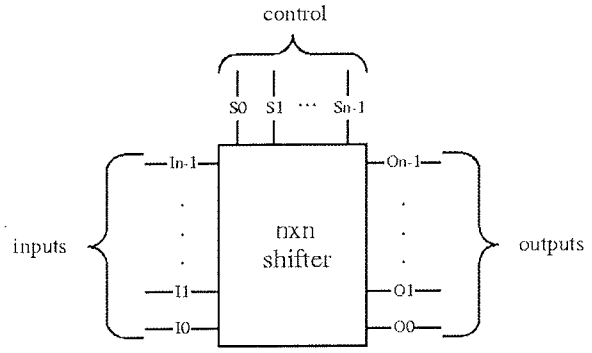


Fig. 1. General shifter block.

The design requirements set upon shifter designs are as noted above except for the requirement no. 3. For the purpose of comparison only the circular right shift as the most common shift operation is implemented. Other shift operations can be easily implemented by adding some peripheral circuitry /2, 3/ and by modifying control signals. Furthermore, the layouts of the shifters are designed in standard MOSIS /7/ scalable CMOS (SCMOS) n-well technology using 2 metallization layers, in such a way that they follow both the standard SCMOS, as well as submicron SCMOS (SCMOS_SUBM) set of design rules. They are, therefore, without any modification of the circuit or layout design, realizable in a whole range of different fabrication processes offered by MOSIS Service, starting with the 2 μm minimal feature process down to the more recent 0.35 μm minimal feature fabrication process.

Given the above mentioned requirements, the shifters are designed with the goal of a minimal layout area on the chip. Circuit and layout designs are modified and optimized in order to satisfy these requirements and the functionality of the shifters or, in some cases, to lower the circuit power dissipation to the acceptable level. No attempts to optimize the circuits according to other criteria (e.g. delay, current drive, etc.) are undertaken. This is to insure the impartiality of the shifter performance comparison.

Next, a detailed overview of the shifter circuit and layout design is given. The shifter performance is evaluated through simulations, which are described in the following section. At the end of the paper, the performance of shifter designs and some usage and optimization guidelines are presented.

2. Shifter design

2.1. Circuit design

Barrel shifters

All versions of barrel shifters are based on the crossbar switch. A version of crossbar switch with pass transistors is shown in Fig. 2.

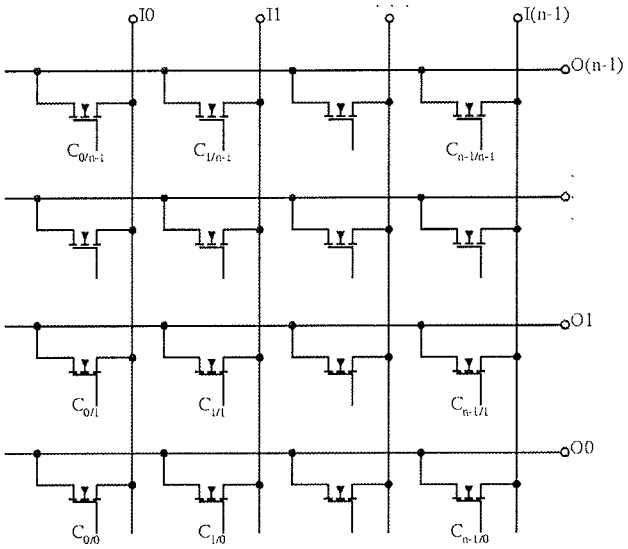


Fig. 2. Crossbar switch with pass transistors.

The crossbar switch can in theory emulate any multi-input multi-output logical function. By careful design of wiring one can readily obtain the main field of the barrel shifter shown in Fig. 3a. Control inputs to the pass transistors are designated *Si* and only one of them is high at a time, defining the shift value. If we replace the pass transistors in the circuit in Fig. 3a with transmission gates, we obtain the main field of the barrel shifter with transmission gates shown in Fig. 3b. As usual in circuits with transmission gates, this shifter requires both inverted and non-inverted control signals designated *Si'* and *Si,* respectively. By examination of both schematics, one can easily deduce that they perform the required function: circular right shift.

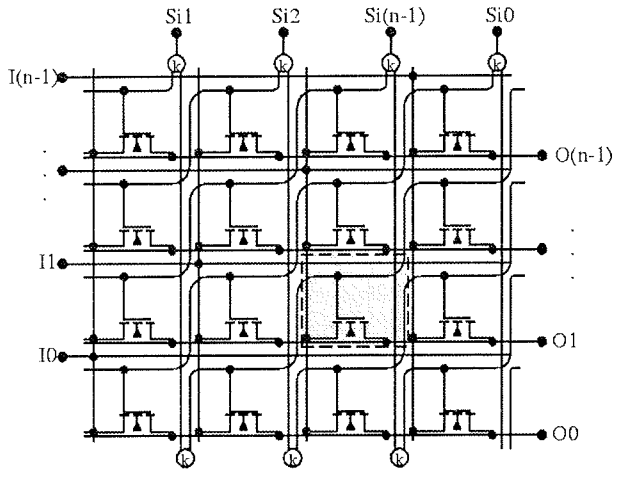


Fig. 3a. Main field of the barrel shifter: with pass transistors

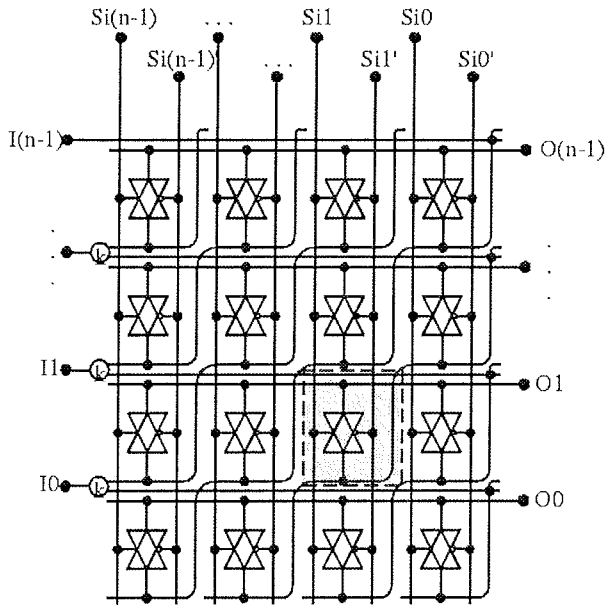


Fig. 3b. Main field of the barrel shifter: with transmission gates.

In order to comply with the requirements set upon the shifter designs, and to have coded control input signals, a decoder must be added to the circuit. In order to minimize the layout area, a NOR decoder in dynamic logic with p-type precharge transistors is chosen /4, 6/. The schematic of this decoder is shown in Fig. 4.

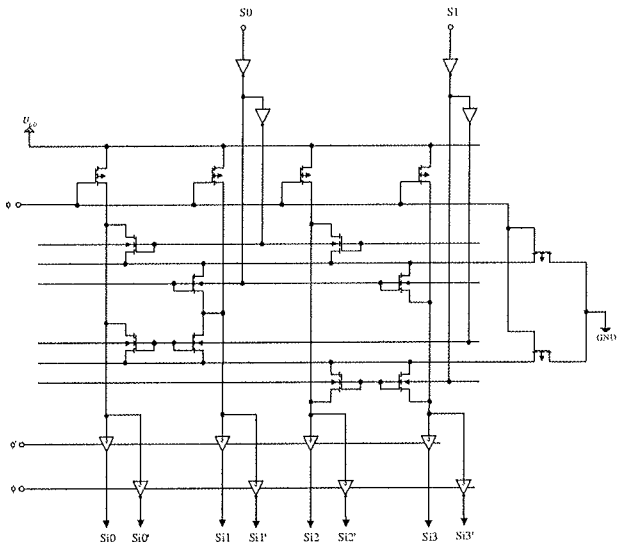


Fig. 4. NOR decoder in dynamic logic (4-bit example).

During the precharge phase the clock ϕ is low, and during the evaluation phase the clock is high. Outputs of the decoder are control inputs $Si_{(0\div n-1)}$ to the barrel shifter main field.

In order to ensure the correct logic value of control inputs *Si* and *Si'* during the precharge phase /4/, interface clocked buffers, specially devised for this purpose, are used. The circuitry is shown in Fig. 5. The buffer exists in two versions: inverting and non-inverting, depending on the used output. Both buffer versions are used in the decoder for the barrel shifter with transmission gates and can be seen in Fig 4. Only the non-inverting buffer version is used in decoder for the barrel shifter with pass transistors. Both decoders require an inverted clock ϕ'. It should be noted that the shifters are fully functional even without the interface buffers /8/, but the power dissipation is too large, due to the current leakage during the evaluation phase.

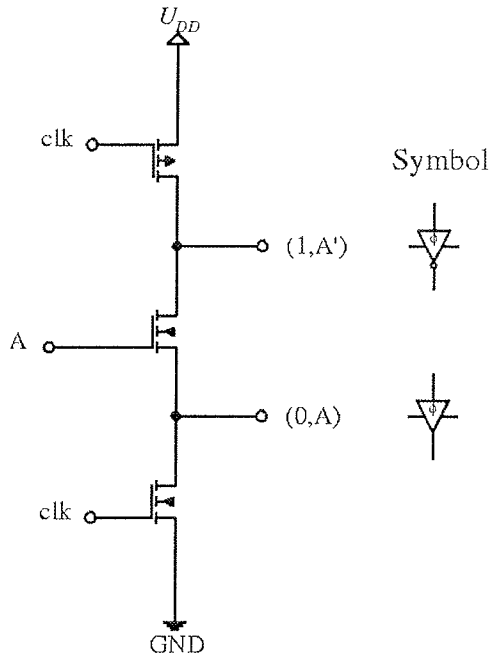


Fig. 5. Interface clocked buffers.

An interface is required to connect the barrel shifter main field with the rest of the datapath. A column of inverters at the input and at the output of the barrel shifter main field is used for this purpouse. Being in pair, these inverters do not influence the logic function of the shifter. When integrating these shifters in an actual datapath, the inverters could be replaced by latches or some other form of interface.
Degradation of electrical characteristics of output signals is a common problem in pass transistor logic. In the nMOS pass transistor the output high voltage level is lowered by the amount of the transistor threshold voltage. This effect can lead to increased power dissipation and lower noise margins of the output signals /4/. The signal level can be restored using the buffer shown in Fig. 6. The feedback transistor *Tr* pulls the inverter input high when the output of the buffer goes low, and prevents the lasting intermediate voltage value in buffer. This buffer is used at the outputs of the barrel shifter with pass transistors instead of regular inverters.

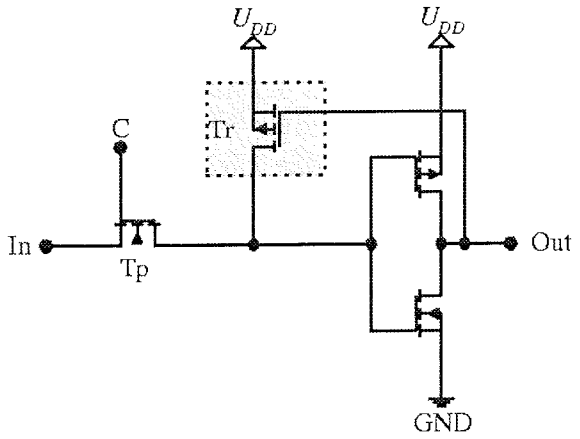


Fig. 6. Output level-restoring buffer with adjacent nMOS pass transistor /4/.

Logarithmic shifters

While the barrel shifter implements the whole shifter as a single switch field, the logarithmic shifter uses a multistage approach. The total shift value is decomposed into stages. Each stage shifts the data by some fixed amount, usually by power-of-two /4/. The shift stages can be represented with columns of multiplexers. The basis for such logarithmic shifter in base 2 is shown in Fig. 7. Such a logarithmic shifter has $\log_2 n$ stages for n-bit data bus and is usually smaller than the equivalent barrel shifter /4/. Note that the control signals S for this shifter are already encoded, i.e. no decoder is required.

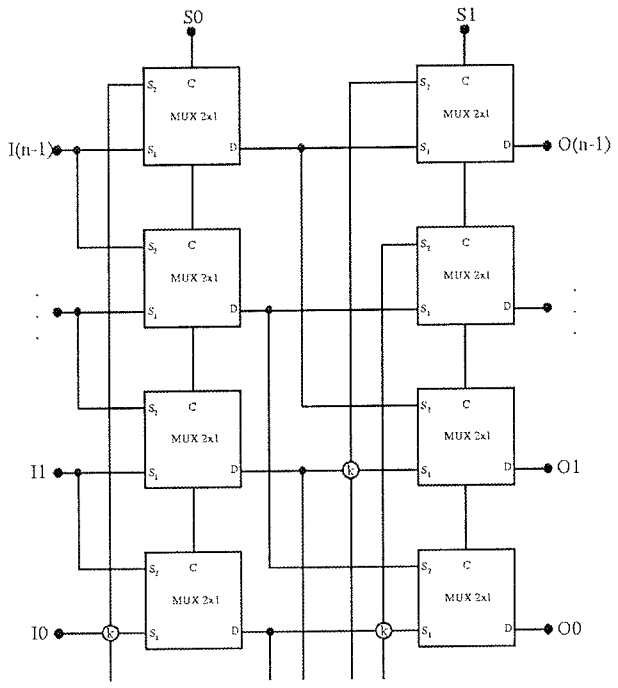


Fig. 7: Logarithmic shifter basic schematic (only first two stages).

The design of the multiplexer determines the logarithmic shifter type. Two multiplexer designs are implemented: with nMOS pass transistors and with transmission gates. The logarithmic shifter basic cells are shown in Fig. 8. An inverting buffer is placed at the output of the multiplexer. The odd number of stages in 32-bit shifter ($\log_2 32 = 5$), combined with the input inverter column, ensures that the output signal is in-phase with the input signal. As in the case of a barrel shifter, the restoring buffer is used at the output of the logarithmic shifter cell with pass transistors.

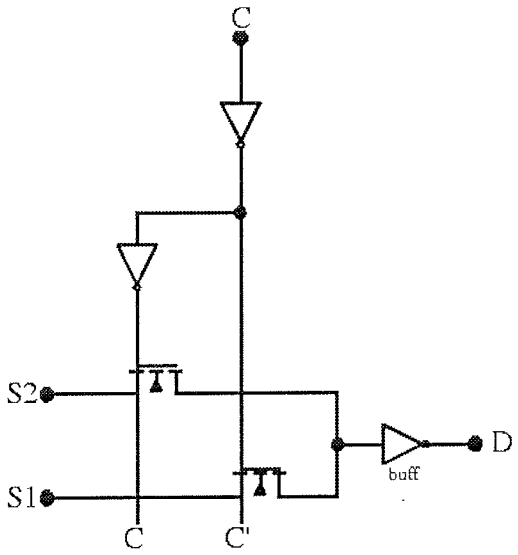


Fig. 8a. Logarithmic shifter basic cells: multiplexer with pass transistors.

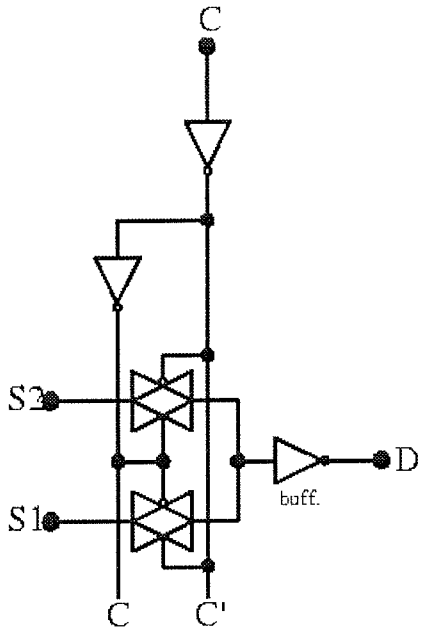


Fig. 8b Logarithmic shifter basic cells: multiplexer with transmission gates.

2.2. Layout design

The shifter layouts are designed by MAGIC, a widely accepted software tool /9/. The layouts are designed so that they follow both the standard scalable CMOS and submicron scalable CMOS set of design rules from MOSIS /7/. The chosen fabrication process is HP CMOS26G 0.8 μm n-well process with 3 metallization layers (of which only 2 are used in the design). This fabrication process was chosen because it supports both the standard and submicron set of scalable CMOS design rules, as well as a supply voltage of 5V, and because of the availability of parameter files for many computer tools /10/. The main features of this fabrication process are listed in Table I. The ratio of *p* to *n* channel widths in a layout is usually determined from the requirement for symmetrical output characteristic, equal rise and fall times or minimal propagation delay /4, 6/. This ratio is chosen to be 4, which, according to Table I, corresponds to the ratio of MOSFET *K* constants. Beside this requirement, the layouts are designed with the primary goal of achieving minimal chip area.

Table I Parameters of MOS fabrication process HP CMOS26G, run N68O /10/.

Minimal Feature Size	Supported Set of Design Rules	Lambda	# of Gates per mm²	Operating Voltage
0.8 μm	SCMOS (_TM) SCMOS_SUB	0.5 μm 0.4 μm	6200	5.0 V
Normalized MOSFET *K* Constants	**Ratios of MOSFET's *K* Constants**	**Threshold Voltages**	**Minimum Inverter Delay**	
$K'_n = 1.2704 \cdot 10^{-4} A/V^2$ $K'_p = 3.2077 \cdot 10^{-5} A/V^2$	$K'_n/K'_p = 3.960$ $\sqrt{K'_n/K'_p} = 1.990$	$U_{GS0n} = 0.7086V$ $U_{GS0p} = -0.8446V$	290 ps	

The shifter layouts are compiled of many basic shifter cell layouts. A basic cell is designed in such a way that it can be easily connected with the equal adjoining cells. After arranging arrays of basic cells, the final layout is obtained by adding contacts, external connections and power supplies.

Barrel shifters

The layout of the barrel shifter consists of two main parts: main field and decoder. Both are composed of basic cells.

The layouts of the barrel shifter basic cells are shown in Fig. 9. The schematics of the same basic cells can be seen shaded in Fig. 3 (the contacts *k* are implemented in only some cells when the layout is complete).

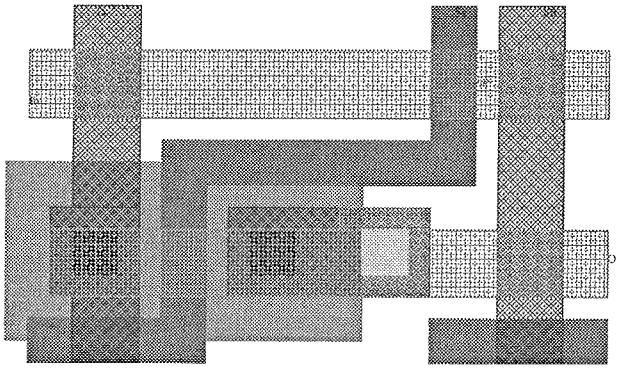


Fig. 9a. Barrel shifter basic cells layouts: cell with pass transistors.

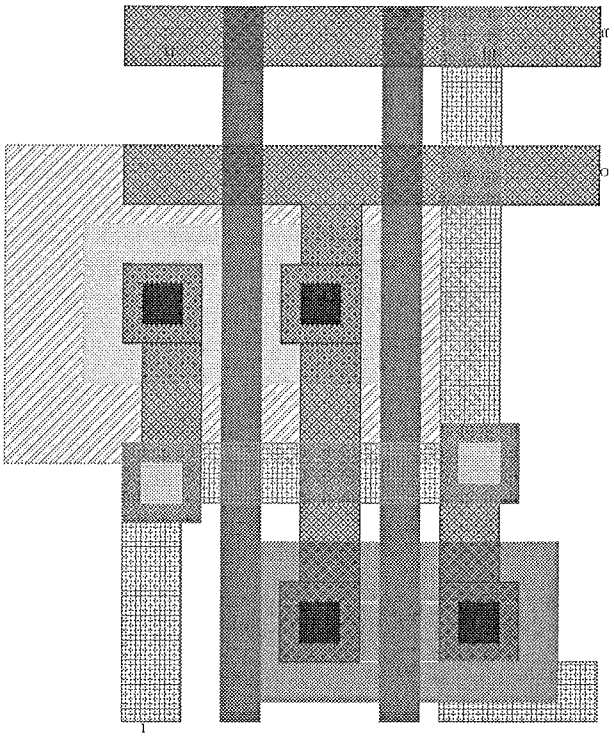


Fig. 9b Barrel shifter basic cells layouts: cell with transmission gates.

The same decoder is used for both shifters. The layout of its basic cell is shown in Fig. 10.

The layout of the output level-restoring buffer (Fig. 6) is adjusted to the vertical pitch of the basic shifter cell in Fig. 9a and shown in Fig. 11.

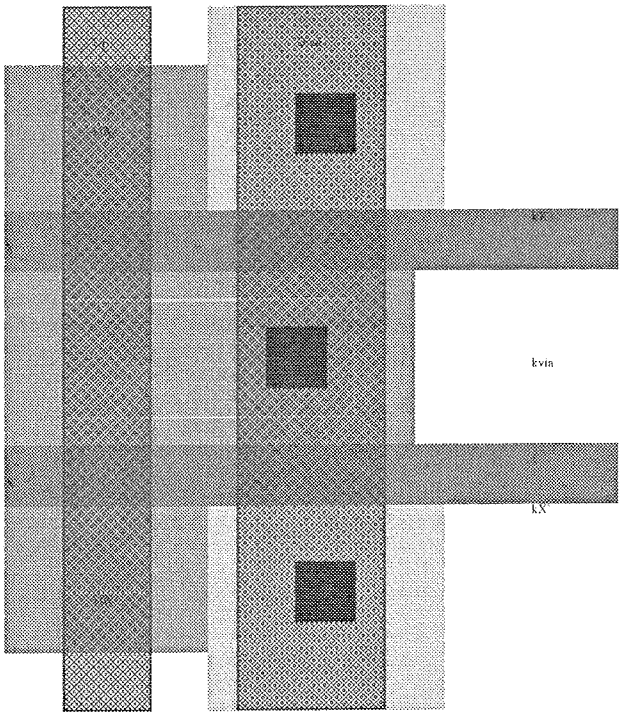


Fig. 10. Decoder basic cell layout.

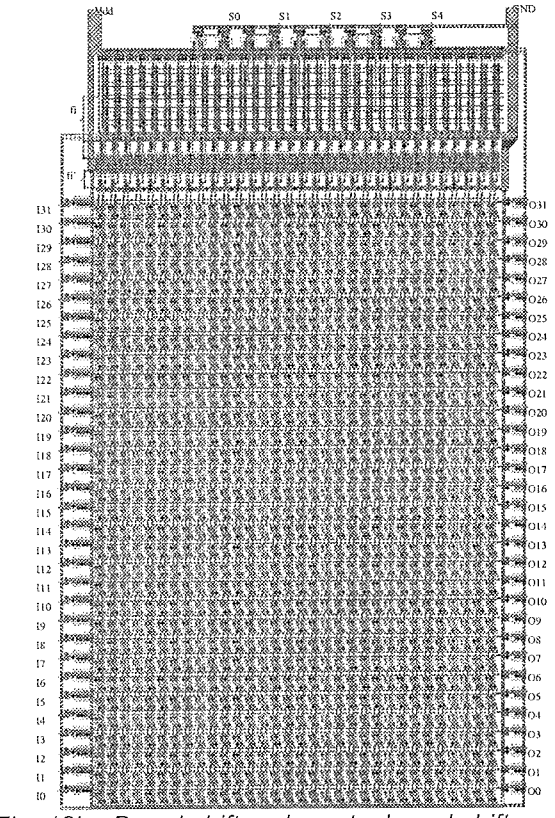


Fig. 12b. Barrel shifters layouts: barrel shifter with transmission gates.

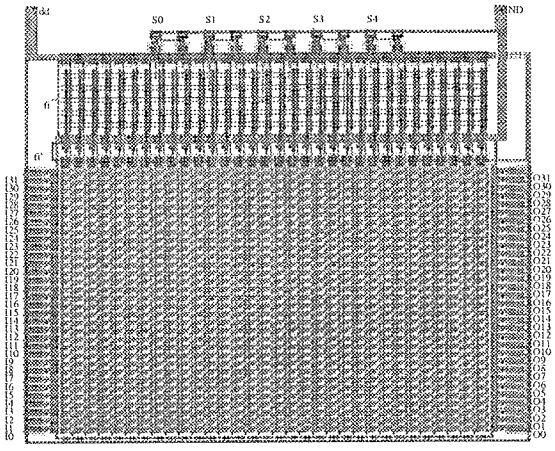


Fig. 12a. Barrel shifters layouts: barrel shifter with pass transistors

The layouts of other inverters and buffers are not shown separately, as their design is traditional. The vertical buffers, which drive the control signals into the decoder field (Fig. 4), have channel widths of twice the minimal size, because of the large capacitive load of polysilicon control lines in decoder field. The precharge *p*-type transistors in the NOR decoder have also larger widths, determined by the horizontal pitch of the basic decoder cell.

The final layouts of the barrel shifters are shown in Fig. 12. Note the different aspect ratio of these two layouts (the barrel shifter layout with transmission gates is much higher due to the difference in basic cell size).

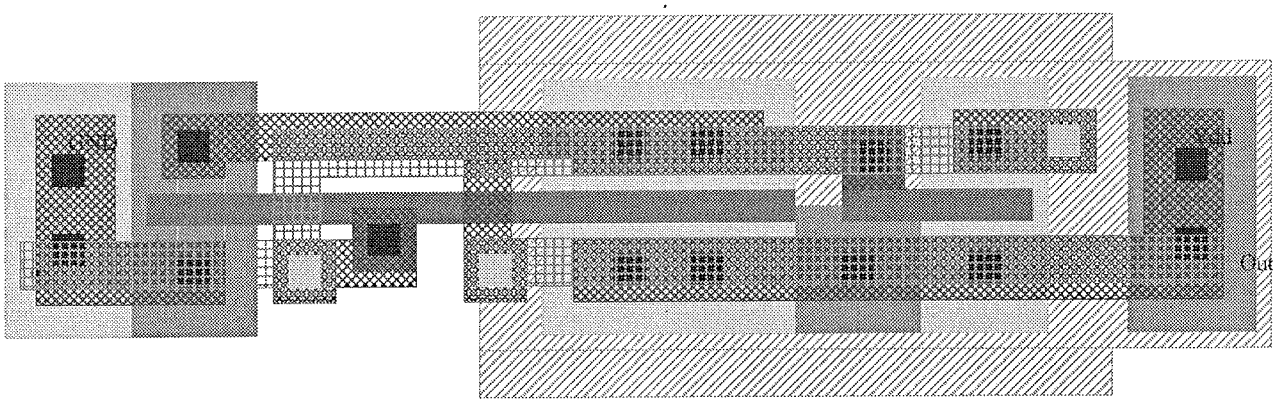


Fig. 11. The output level-restoring buffer layout.

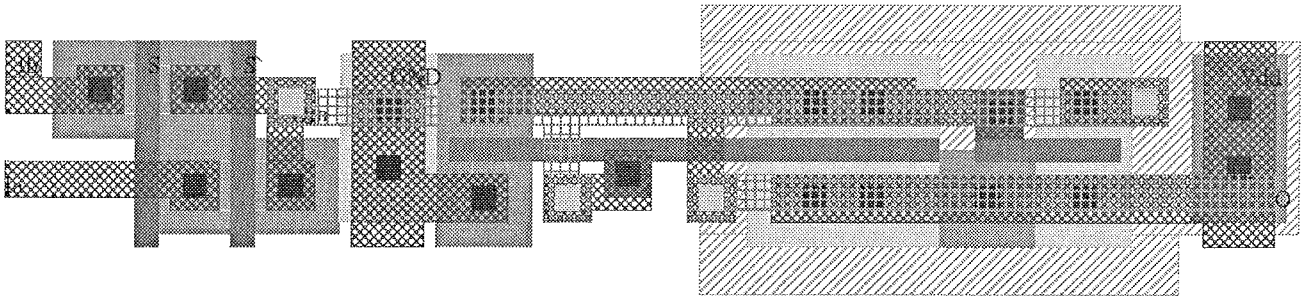


Fig. 13a. Logarithmic shifter basic cells layouts: multiplexer with pass transistors.

Logarithmic shifters

The layout of a logarithmic shifter consists of two parts: columns of multiplexers and switch field between the columns. The layouts of logarithmic shifter basic cells are shown in Fig. 13.

The vertical buffers, which are twice the minimal size, are added in the path of control signals at the top of each multiplexer column. These buffers are shown together with multiplexer cells in Fig. 8. This is necessary due to the large capacitance of polysilicon multiplexer control lines. The switch field connects the output of the cell m in the stage k with the input of the cell $(m\text{-}2^k)$ modulo 2^n in the stage $k+1$, and is typical of logarithmic shifters. It is made in two metal layers. The minimal distance between the vertical metal lines determines the size of the switch field. The width of the switch field increases exponentially with the stage shift value – 2^k /4/.

The complete layouts of the logarithmic shifters are shown in Fig. 14. The difference in layout aspect ratio is also visible in this case.

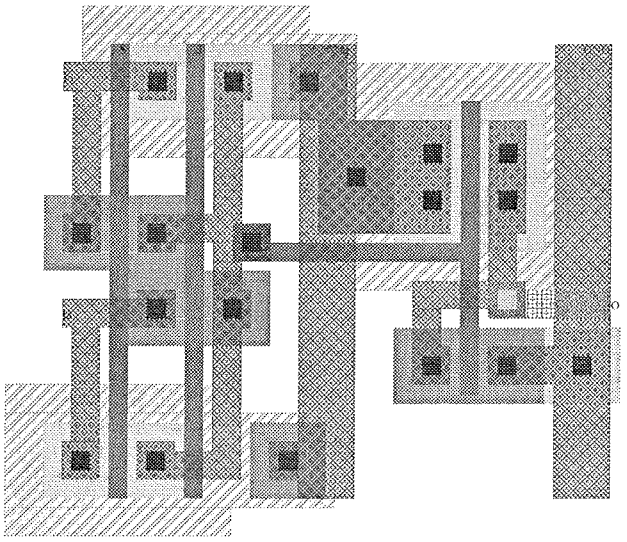


Fig. 13b. Logarithmic shifter basic cells layouts: multiplexer with stransmission gates.

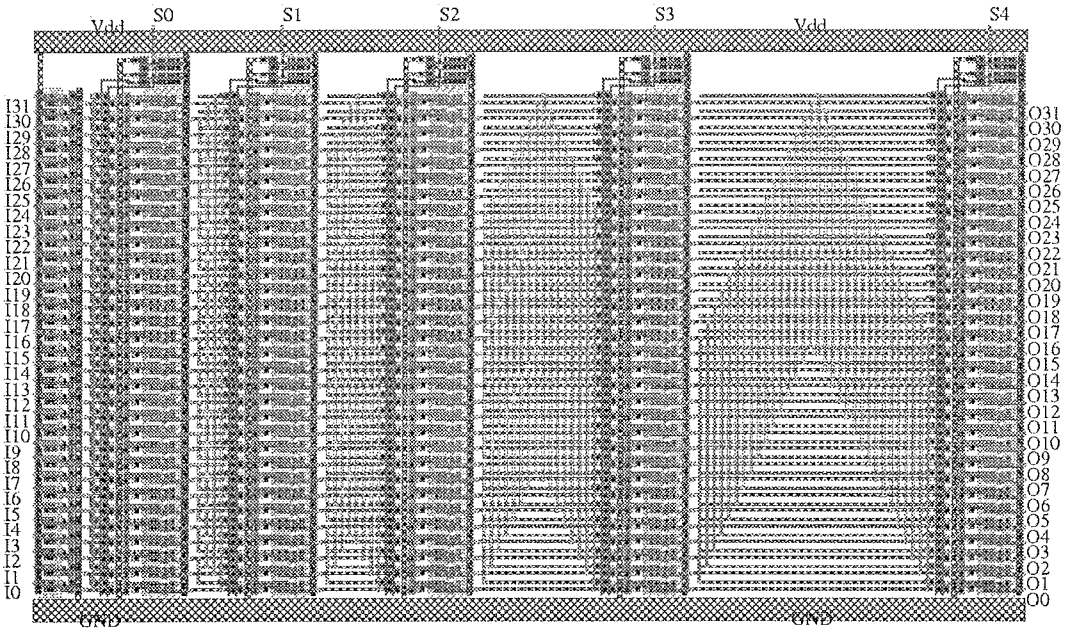


Fig. 14a. Logarithmic shifters layouts: logarithmic shifter with pass transistors,

Finally, the layout complexity of the designed circuits is presented in Table II by comparing: the number of transistors, linear dimension in units of λ, and chip area for implementation in the HP CMOS26G 0.8 μm fabrication process.

Table II. Layout parameters of designed circuits.

Shifter circuit layout*	Number of Transistors			Linear size	Chip Area, S (λ=0.4 μm), mm²
	nMOS	pMOS	Total	height × width, λ	
Barrel 1	1364	180	1544	848 × 989	0.134
Barrel 2	1428	1204	2632	1522 × 894	0.218
Logarithmic 1	522	362	884	645 × 1080	0.111
Logarithmic 2	522	522	1024	1635 × 829	0.217

* 1 corresponds to circuit version with nMOS pass transistors
2 corresponds to circuit version with transmission gates

3. SIMULATIONS

In order to evaluate circuit performance, logical and electrical simulations are used. The simulations are performed with circuit models extracted from the layout by using MAGIC and additional tools that come with the software /9/. All circuit models are extracted with the parameters of HP CMOS26G 0.8 μm *n*-well fabrication process.

For logic validation, debugging of design and initial electrical simulation, the event-driven electrical simulator IRSIM /11/ is used. It enables easy handling of wide data buses and logical states. The results obtained with this program include logical values, propagation delays and power dissipation /8/. However, due to the simple MOS transistor model based on resistance, this program could only give the approximate values of electrical parameters; the range of values and their relationships.

Simulations in IRSIM showed that all designed shifters were fully functional at a chosen clock frequency of 50 MHz. The results of one typical simulation are shown in Fig. 15.

IRSIM simulations determine the combinations of input and control signals for which the propagation delay is the largest. After initial simulations in IRSIM, additional electrical simulations for chosen input and control signal combinations are performed by SPICE /12/, in order to determine the circuit electrical parameters: voltage levels, propagation delay and power dissipation. The MOS transistors are modeled with the SPICE Level 3 MOSFET model obtained from MOSIS /10/. This model offers the advantage of faster convergence compared to the more sophisticated BSIM (Level 4 and 5) SPICE MOSFET models also available from MOSIS. The accuracy of the model is satisfying for the used minimal feature size of 0.8 μm. Summarized results of SPICE simulations are given in Table III.

Two types of propagation delays are determined from SPICE simulations: the delay from control input to output ($S \rightarrow O$, t_{DSO}) and delay from input to output ($I \rightarrow O$, t_{DIO}). The control input to output delay for barrel shifters, which have a decoder in dynamic logic, and require clock signals,

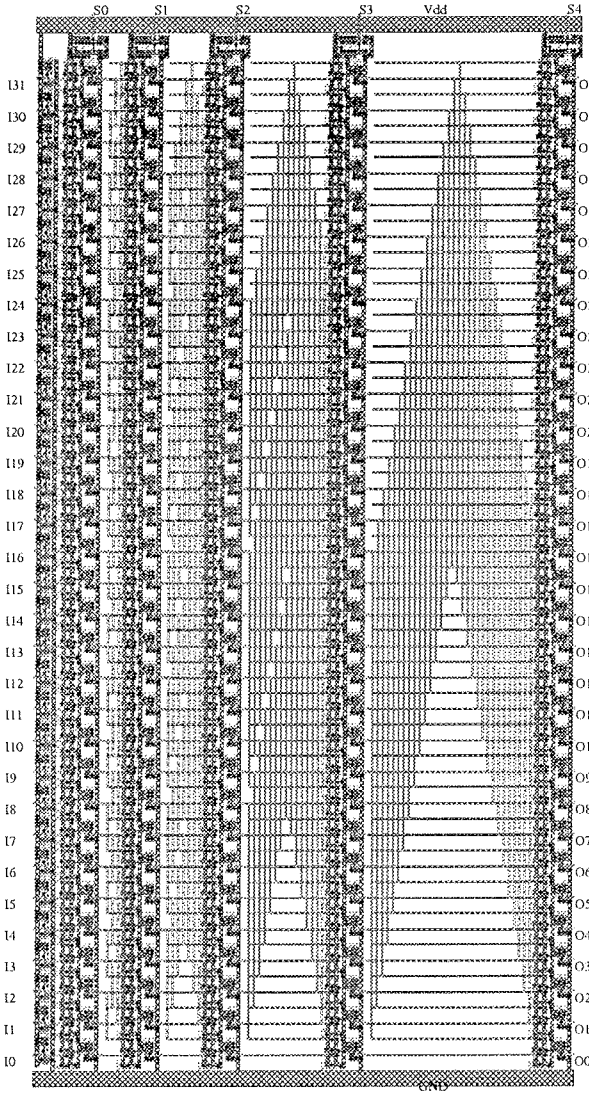


Fig. 14b. Logarithmic shifters layouts: logarithmic shifter with transmission gates.

Table III. Summarized results of SPICE simulations.

Simulated Shifter Circuit	Propagation delay, ns			Average Power Dissipation P_t, mW	Specific Power Dissipation P/S, W/cm²
	t_{DSO}	t_{DIO} H→L	t_{DIO} L→H		
Barrel 1	7.1	1.7	1.1	6.4	4.8
Barrel 2	1.4	0.6	0.5	6.0	2.8
Logarithmic 1	2.6	1.8	1.6	5.8	5.2
Logarithmic 2	2.4	1.2	1.1	5.2	2.4

is defined starting from the rising edge of clock signal ϕ, i.e., from the beginning of the circuit evaluation phase. Only the longest delay t_{DSO} is shown in Table III regardless of the transition direction. The input to output delays are evaluated for both high to low (HL) and low to high (LH) transitions of the output signal for various input and output signals. The longest delays t_{DIO} are given separately for HL and LH transitions. Fig. 16 shows the simulated electric signals from a typical SPICE simulation.

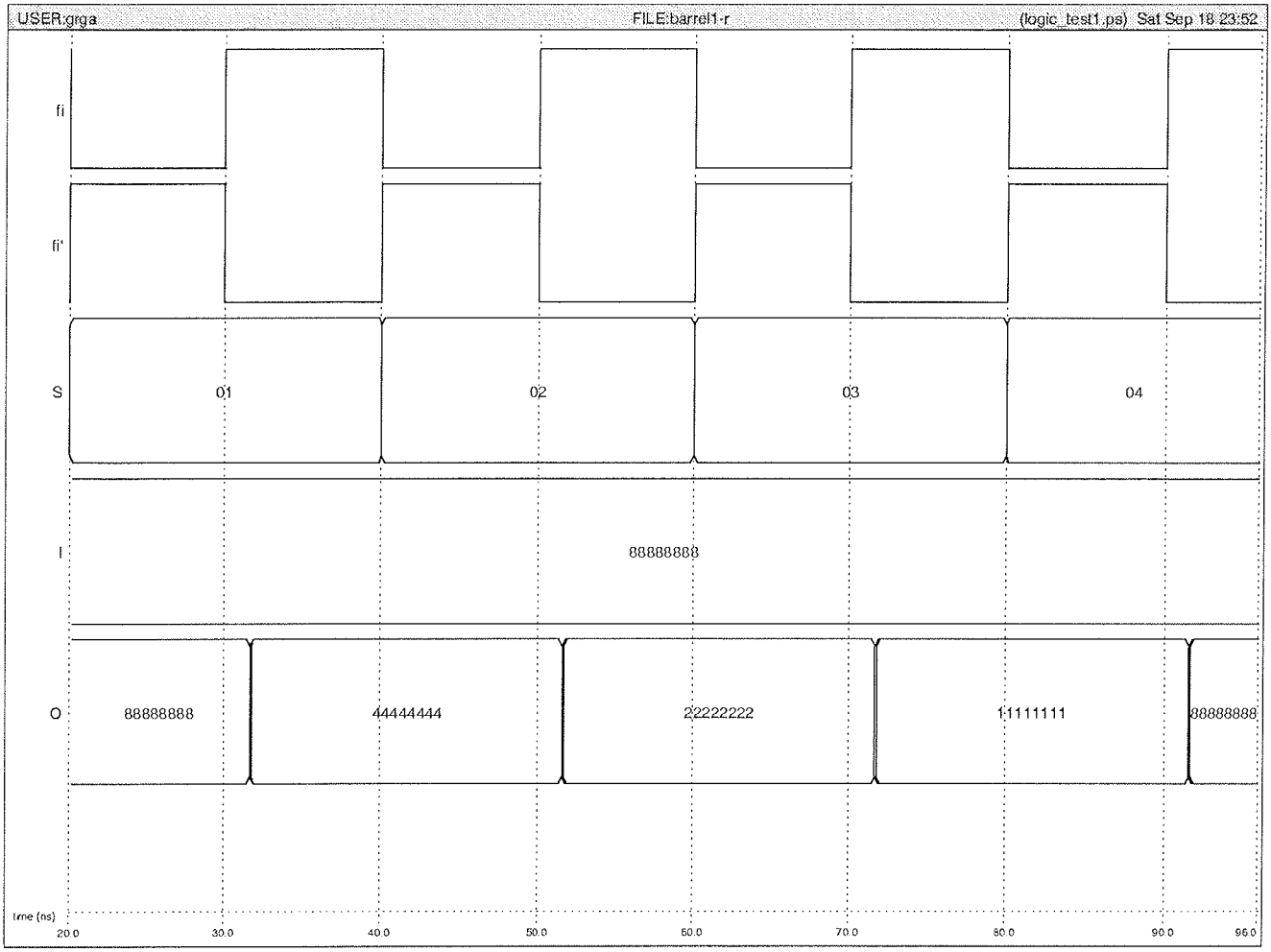


Fig. 15 Signal simulated by IRSIM (logic validation for barrel shifter with pass transistors).

Note the difference in the propagation delays for HL and LH transitions. High to low transition time is always longer. This is caused with the lower output voltage of the nMOS pass transistor and lower current drive of pMOS transistor in the transmission gate. Both of these effects cause slower rising of the signal at the output of the pass logic (which is than inverted in the output buffer). The level-restoring buffer, shown in Fig. 6, can only partially decrease this problem. The difference in delay times is however smaller for the circuits with transmission gates, which are therefore more often used in present pass logic circuits /4, 5/.

The total power dissipation of the circuits is evaluated at the clock frequency of 50 MHz (and corresponding input signal frequency of 25 MHz, see Fig. 15) for selected signal pattern by simulations over a longer time period. The input and control signal patterns were chosen to achieve the maximum dynamic power dissipation by permanent change of logic states at every clock cycle at as many circuit nodes as possible. The final signal pattern is determined by test simulations. The calculated specific power dissipation, dissipated power to chip area ratio, is given in Table III. This parameter is an indicator of thermal flux in VLSI chips and limits the integration density as well as the clock frequency for the circuit implemented in the chosen fabrication process.

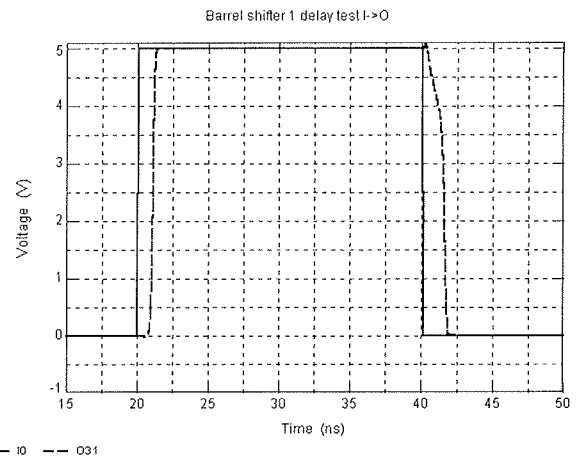


Fig. 16 Signals simulated by SPICE (propagation delay t_{DIO} for barrel shifter with pass transistors).

4. COMPARISONS AND CONCLUSIONS

The comparison of designed circuits can be made according to many criteria: number of transistors, chip area, speed, power dissipation. All designed circuits offer the same functionality; they perform circular right shift at the clock frequency of at least 50 MHz and satisfy all the requirements listed in the introductory section. One must also note, according to delay times shown in Table III, that the majority of the circuits could support significantly higher clock frequencies.

The main design difference is the existence of the dynamic logic decoder in the barrel shifters, and therefore the requirement for independently generated non-overlapping clock pair ϕ and ϕ' in these shifters. This requirement must not be considered as a deficiency, since in the datapath where this shifter would be integrated, a clock would be present anyway. By using the decoder in dynamic logic, the chip area and transistor count are minimized. Dynamic logic circuits usually have a higher speed due to the lower capacitive load /4/. In the case where the decoded control signals are already available in the chip, one would implement only the barrel shifter's main field, which would lead to significant circuit simplification, decreased chip area and increased speed.

The logarithmic shifters presented in this work are designed in combinatorial pass and static logic and do not require clock signals. When integrating these shifters in the datapath, one can add latches at the input and output of the shifter. Pipelined versions of logarithmic shifters with multiplexers in dynamic logic and with latches appeared in the literature, and are used mostly in high-performance chips /3/.

The comparison of the main shifter parameters is given in Table IV. The comparison is performed according to three criteria: complexity, delay and power dissipation.

According to Table IV, no shifter comes as a clean winner. The application of a particular shifter will depend on the requirements set upon the chip as a whole. The shifters with nMOS pass transistors have in general lower complexity. The logarithmic shifter has lower complexity than the barrel shifter in equivalent logic style. In particular, the logarithmic shifter with pass transistors has the lowest transistor count and smallest chip area. The lowest propagation delay is achieved in the barrel shifter with transmission gates. The logarithmic shifter with transmission gates has the smallest power dissipation. However the total power dissipated in other circuits is not substantially larger. The difference in the specific power dissipation is much larger and determined by the difference in the used chip area. The specific power is not a deciding design criterion except in some special cases (low-power electronics with reduced cooling possibilities).

The barrel shifter with transmission gates, due to its high speed and low power demands, seems to be the best choice for implementation in general purpose ASIC designs if no limits to the chip area are set. In the applications where a circuit with low power dissipation is needed, the best choice is the logarithmic shifter with transmission gates. The shifters with pass transistors should be preferred if the chip area is limited. The barrel shifter has the disadvantage of larger power dissipation and, especially, substantially larger control signal delay t_{DSO}.

Further improvements of the shifters are possible. For example, it is possible to enlarge or replace interface buffers or decoders. With none or small gain in chip area, one can also enlarge the otherwise minimal channel width of pass transistors. With such modification one could equalize the LH and HL propagation delays in shifters with transmission gates or lower the propagation delay in shifters with pass transistors. By adding additional peripheral and control circuitry one can enhance the shifter functionality to support more shift operations.

Finally, it is worth noting that the shifter design procedures presented in this work can be used as application guidelines for integrated circuits in general.

Acknowledgments

This paper is supported by *"Ministry for Science and Technology"* of the Republic of Croatia within the sci-

Table IV. The comparison of shifter designs and performance.

Designed Shifter Circuit	Complexity		Delay			Power	
	Tot. # Tran.	Chip Area	t_{DSO}	t_{DIO} HL	t_{DIO} LH	Tot. Power	Spec. Power
Barrel 1	1.747	1.207	5.071	2.833	2.200	1.231	2.000
Barrel 2	2.977	1.964	1.000	1.000	1.000	1.154	1.167
Logarithmic 1	1.000	1.000	1.857	3.000	3.200	1.115	2.167
Logarithmic 2	1.158	1.955	1.714	2.000	2.200	1.000	1.000

entific project *"036001 Research on VLSI/ULSI semiconductor structures"*. We are grateful to Dr. Adrijan Baric for many useful suggestions regarding the paper.

REFERENCES

/1/ R.S. Lim, "A Barrel Switch Design", Computer Design, August 1972, pp. 76-79
/2/ S.M. Kang, "Domino-CMOS Barrel Switch for 32-Bit VLSI Processors", IEEE Circuits and Devices Magazine, May 1987, pp. 3-8
/3/ R. Pereira, J.A. Michell, J.M. Solana, "Fully Pipelined TSPC Barrel Shifter for High-Speed Applications", IEEE Journal of Solid-State Circuits, Vol. 30, No. 6, June 1995, pp. 686-690
/4/ J.M. Rabaey, "Digital Integrated Circuits - a Design Perspective", Prentice Hall, 1996
/5/ N.H.E. Weste, K. Eshraghian, "Principles of CMOS VLSI Design", Addison Wesley, 1993
/6/ E.D. Fabricius, "Introduction to VLSI Design", McGraw-Hill, 1990
/7/ The MOSIS Service, "MOSIS Scalable CMOS (SCMOS) Design Rules, Revision 7.2, 1996
/8/ D. Grgec, Ž. Butković, "Barrel Shifter Design and Simulations", Proceedings of MIPRO '99 Conference, Opatija, Croatia, pp. 62-65, 1999
/9/ Digital Western Research Laboratory: "WRL Research Report 90/7 -1990 DECWRL/Livermore Magic Release", September 1990, "Magic Addendum: Version 6.5 differences", 1994
/10/ MOSIS Service Web Site: http://www.mosis.org
/11/ University of California at Berkeley, IRSIM ver. 9.4, "IRSIM User's Manual", 1993
/12/ MicroSim Corporation, "MicroSim PSpice A/D & Basics+ Circuit Analysis Software User's Guide", Version 6.3, April 1996

Dalibor Grgec
Institute for Electromagnetic Theory and Microelectronics
University of Bremen
Kufsteiner Str. , Postfach 33 04 40
28334 Bremen, GERMANY
Phone: +49 421 218 2204
Fax: +49 421 218 4434
E-mail: grgec@item.uni-bremen.de

Željko Butković
Department of Electronics, Microelectronics, Computer and Intelligent Systems
Faculty of Electrical Engineering and Computing
Unska 3, HR –10000 Zagreb, CROATIA
Phone. +385 1 6129 924
Fax. +385 1 6129 653
E-mail: Zeljko.Butkovic@fer.hr

Prispelo (Arrived): 17.07.00 *Sprejeto (Accepted): 22.11.00*

CAPACITOR-AREA AND POWER-CONSUMPTION OPTIMIZATION OF HIGH ORDER Δ - Σ MODULATORS

Drago Strle
University of Ljubljana, Faculty for electrical engineering, Slovenia.

Keywords: computer science, microelectronics, SP, Signal Processing, DELTA-SIGMA modulators, Δ - Σ modulators, noise reduction, area optimization, power consumption optimization, design

Abstract: A design methodology for the power consumption and capacitor area optimisation of high-order high-resolution single-bit switched-capacitor type Δ-Σ modulators is described. The main reasons for the power consumption are determined and appropriate steps for its reduction are proposed. The algorithm is coded in MATLAB and gives the unit capacitor size of each integrator stage for an arbitrary topology and the required specification of each opamp used in the integrator to achieve the required S/N ratio, minimize the power consumption and the silicon area and preserve matching accuracy. The 5th order modulator was built and the results prove the effectiveness of the approach.

Optimizacija moči in površine Δ-Σ modulatorjev visoke stopnje

Ključne besede: racunalništvo,mikroelektronika, SP obdelave signalov, DELTA-SIGMA modulatorji , Δ - Σ modulatorji, zmanjšanje šuma, optimiranje površine, optimiranje porabe energije, snovanje

Povzetek: V članku obravnavamo metodologijo načrtovanja eno in večbitnih S-C Δ-Σ modulatorjev, ki omogoča optimizacijo površine silicija in porabo moči. Prikazani so glavni razlogi za porabo moči in potrebni koraki za zmanjšanje. Algoritem smo realizirali v programu MATLAB. Rezultat je tabela kondenzatorjev in specifikacije ojačevalnikov vsake integracijske stopnje za poljubno topologijo modulatorja. Cilja optimizacije sta zmanjšanje površine polja kondenzatorjev ter porabe moči ob upoštevanju robnih pogojev: zahtevano razmerje S/N, točnost razmerij kondenzatorjev itd. Primer optimizacije modulatorja 5. reda dokazuje učinkovitost algoritma in metodologije.

1. INTRODUCTION

The power-consumption optimisation methodology for high-order high-resolution Δ-Σ S-C modulators is presented. The optimisation procedure tries to minimize the unit capacitances of the integrators in a loop filter in such a way that the noise requirements are fulfilled, area is minimized and matching accuracy is maintained.

The contribution of each noise source is dependent on the architecture, coefficients, capacitances and noise generated in the opamps. Usually this is calculated by using the linear model of the modulator /3/, which is adequate method if one is satisfied with approximate results. Since modulators are non-linear systems, linear model is not good enough and we need real non-linear time-domain simulation and optimisation. The kT/C noise simulation principle was presented in /5/ assuming the unit capacitor sizes are given. In this work the unit capacitor sizes are optimised in such a way that contributions to the total noise are approximately the same. At the same time the requirements for the opamps regarding 1/f and thermal noise as well as slew-rate are determined and can be used as specifications for the circuit design. The model of chopping and more realistic models for the opamps are added to the existing algorithm given in /5/. The algorithm is coded in MATLAB and is based on state-space description of an arbitrary topology S-C Δ-Σ modulator. The constraints are minimum possible noise and at the same time appropriate accuracy of the capacitor ratios which are the area and the technology dependent. To illustrate the methodology one example is given. It presents power consumption optimisation of the 5th-order single-loop ML-FT (multiple-loop feedback topology) with poles optimised for stability, having S/N ≥ 110dB. The improvement of the power efficiency and area of the modulator prove the correctness of the approach.

In section 2 the reasons for power consumption in an arbitrary S-C modulator are analysed and the algorithm for noise optimisation is presented. Since most of the power is consumed by the opamps (≥80%) driving the capacitive loads, they must be minimized, taking the area and matching of the capacitor ratio as a constraint. The algorithm also defines the required noise level of the opamps, while the design of the opamps is beyond the scope of this article and the algorithm. Short example illustrates the procedure and gives some additional information.

Section 3 briefly repeats the published state-space description of an arbitrary modulator with an additional noise model of the S-C stages and opamps and gives some optimisation results as a table of capacitances and noise requirements. In section 4 the conclusions are presented.

2. REASONS FOR POWER CONSUMPTION

The main reason for power consumption in any switched-capacitor-type modulator is the power needed by the opamps to drive capacitive loads and to maintain the required noise level. To simplify the description of the power-optimisation procedure, the 5th-order modulator is taken as an example, for which the power-optimisation results are presented at the end. The procedure works well for any S-C-type modulator. The system design considerations for the 5th-order modulator have been presented in an internal report /1/ (figure 1).

The coefficients are realised with appropriate capacitor ratios according to figure 3. The coefficient a_1 is given by $a_1 = \frac{C_{u_1}}{C_{I_1}}$. In reality, S-C integrators are fully differential to maintain good rejection against noise coming from the substrate and to gain 3dB in S/N ratio.

Most of the power in a modulator is consumed by the opamps, which must be capable of driving capacitive loads and have a sufficiently low noise to maintain the S/N ratio of the modulator. The power consumption of the transconductance amplifier, with its simplified AC model shown in figure 2 is proportional to the equation (1) /2/:

$$P \propto kTBw(DR)^2 \frac{V_{GS} - V_{TH}}{V_{DD}} \qquad (1)$$

Where: k is Boltzman's constant, T is the absolute temperature, DR is the dynamic range, Bw is the bandwidth, V_{GS}–V_{TH} is the gate-source over-drive voltage and V_{DD} is the supply voltage.

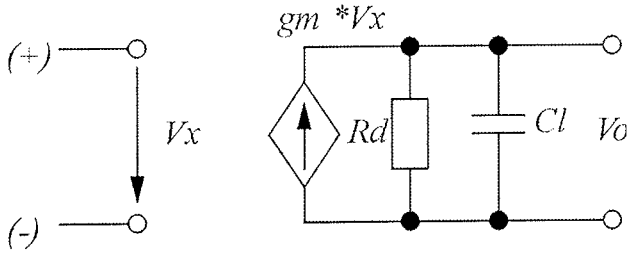


Figure 2: Simplified model of the opamp

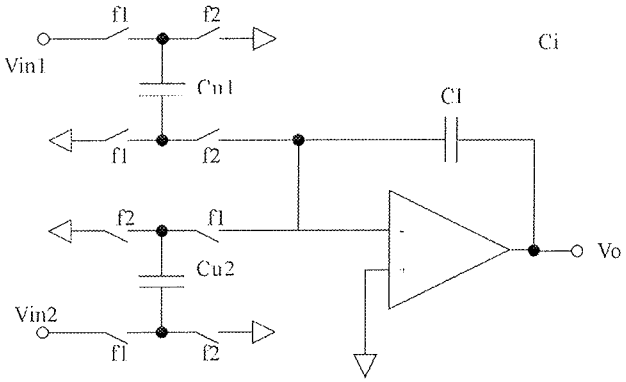


Figure 3: Single-ended S-C integrator

The capacitive loads of the opamps are defined by the coefficients of the loop-filter characteristics, stability constraints /1/ and by the thermal *kT/C* noise generated by the switched capacitors. For higher S/N-ratio requirements the noise must be smaller at the same supply voltage. This can be realised by increasing the unit (usually sampling) capacitance of the integrator; thus the capacitive load of the opamp is increased and more power is required to maintain the bandwidth (Bw). At the same time, for a larger S/N ratio a smaller thermal noise is required for the amplifier, which can only be realised by an increased current and/or increased area of the differential stage transistors. Fortunately only the first amplifier is critical for the low-pass modulator because all the other contributions are attenuated by **noise-transfer functions** of the loop filter.

The slew-rate must be larger than the maximum slope of the settling of v_{out} at maximum input signal to maintain linear settling behavior /3/. This requires the tail current of the differential stage to be greater than: $I_{tail} \geq S_r C_{load}$ /4/. If this limit is not respected the distortion and in-band noise will increase.

Chopping translates 1/f noise and the DC offset around multiples of the chopping frequency ($f_{ch} = f_s / 2$), which is later attenuated by a decimation filter together with shaped quantization noise. The only important parameter concerning 1/f noise is its corner frequency, usually between 10kHz and 100kHz, which in our case is out of the signal bandwidth after chopping.

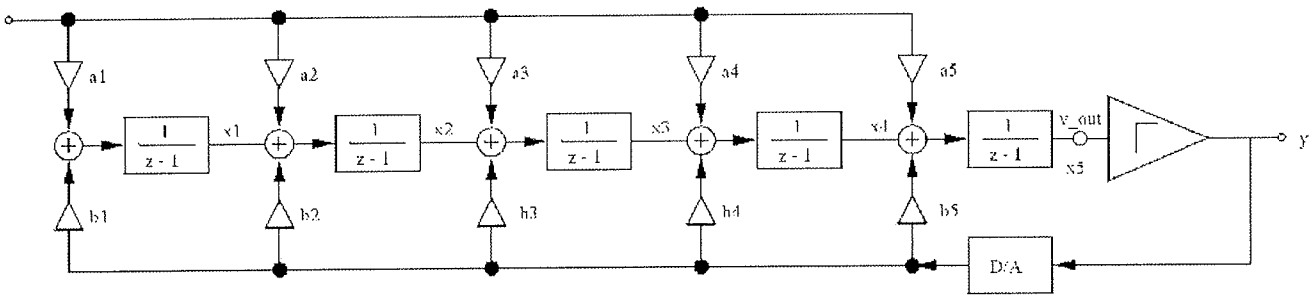


Figure 1: Simplified 5th-order modulator model

In any S-C circuit the coefficients are defined by capacitor ratios, and calculated from the loop-filter poles and zeroes. Each switched capacitor produces a noise power proportional to *kT/C*, where C is the capacitance, T is the absolute temperature and k is Boltzman's constant /4/. According to this model the noise power spectrum density is flat from 0 to fs/2. Every *kT/C* noise source contributes a certain amount of thermal noise which depends on the capacitance value and **noise-transfer functions** g_i. Different noise sources are uncorrelated, so noise contributions at the output of the loop filter can be added according to the equation (2). This is only possible for a simplified linear model of the whole modulator.

$$N_{loop}\left[\frac{W}{Hz}\right] = \frac{1}{f_N}\left[\sum_{i=1}^{i=M} N_{sc_i}\int_0^{f_N} g_{sc_i}^2(f)df + \sum_{j=1}^{j=N} N_{op_i}\int_0^{f_N} g_{op_j}^2(f)df\right] \quad (2)$$

Where: M is the number of S-C stages, N is the number of integrator stages, which is equal to the order of the loop filter, N_{sci} is the noise-power density of S-C stage i (independent of the frequency), N_{opi} is the noise-power density of the opamp j (independent of the frequency), $g_{sci}(f)$is the frequency-dependent **noise-transfer function** for each of the S-C stages, $g_{opj}(f)$ is the frequency dependent **noise-transfer function** for each opamp, N_{loop} is the noise-power density at the output of the loop filter and f_N is the Nyquist rate after decimation.

We can see that assuming uncorrelated noise sources the noise powers of the switched capacitor stages and opamps multiplied by the corresponding noise-transfer functions are added at the output of the loop filter. In addition to this the quantization noise power (which is assumed to be random) is also added. The summation is possible because at the beginning we assumed a linear model for the modulator, and so the theorem of superposition holds up to the output of the modulator. This simplified linear model is used just to get an insight into the behavior and is later replaced by a time-domain method.

1/f noise powers of all the integrator stages, except the first, are negligible because of the high-pass characteristics of the noise-transfer functions g_{opj} (figure 4).

Figure 5 shows AC noise-simulation results for the first opamp. Its thermal and 1/f noise components can be easily recognised. AC simulation results were taken as the input for the noise-generation procedure /5/ and then time-domain simulations were performed, including chopping by $f_s/2$. The 1/f noise component has been translated around $f_s/2$. The thermal noise at the output of the loop filter is composed of contributions from all S-C stages and opamps. The contributions are different for every S-C stage and opamp, and are dependent on the **noise-transfer functions** and the noise powers of the corresponding source; the total noise can be optimised by calculating appropriate capacitor sizes for the switched capacitors and the thermal noise of the opamps. The optimisation is achieved when the contributions at the output of the loop filter are the same.

Figure 6 shows the input-referred power-spectrum noise density of the first integrator: (a) with and (b) without chopping. It is thus possible to reduce capacitor sizes and noise requirements for the opamps, except for the first one. Unfortunately the unit capacitor size can not be made smaller than, for example, C_{unit}=0.3pF; the smallest dimension is constrained by the capacitor-ratio mismatch requirements, which are obtained from the stability and gain accuracy requirements of the modulator. The limit is a result of the synthesis procedure defined in /1/. The unit capacitor size of the first integrator must be bigger than $C_u \geq 15$pF to achieve the S/N ≥ 110dB at an oversampling ratio of D ≥ 256, taking into consideration all noise sources. The unit capacitor size in the following stages drops drastically (in the third and following stages the unit capacitor size is $C_u \geq 0.3$pF(see table on figure 9); the lower limit is calculated from the matching-accuracy parameters for a particular process and its requirements.

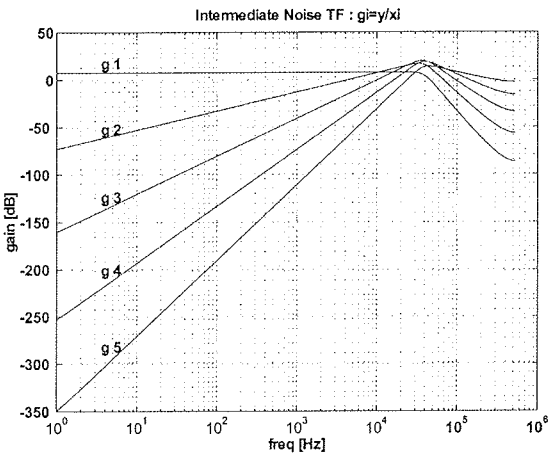


Figure 4: Noise transfer function of the 5th-order modulator

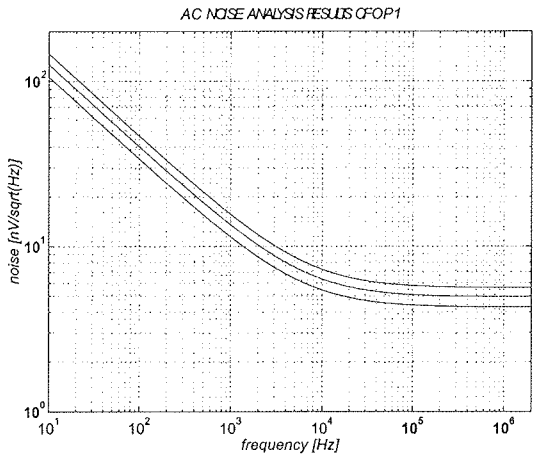


Figure 5: AC noise characteristics of the first opamp

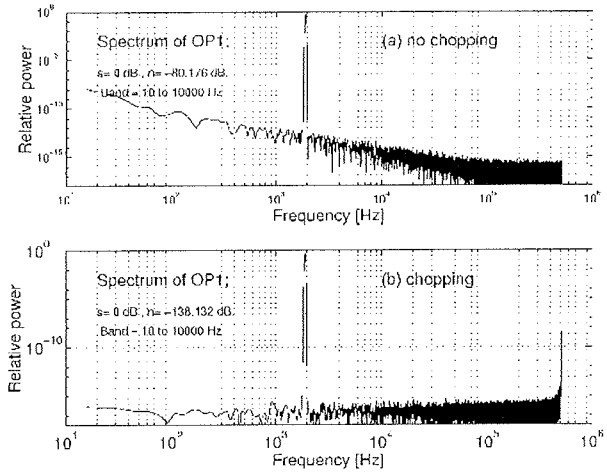


Figure 6: Spectrum at the output of the first integrator a) no chopping; b) with chopping

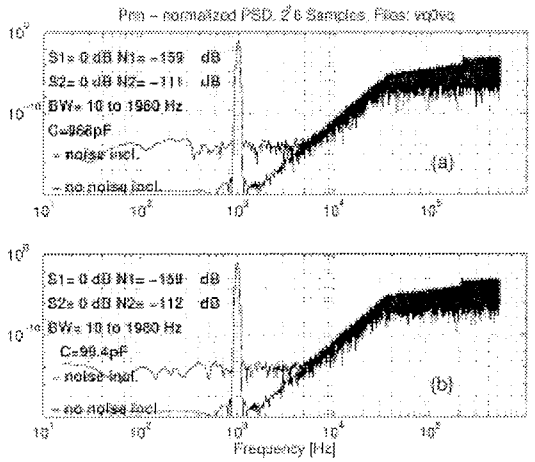


Figure 7: Spectrum of the bit-stream for 5th-order modulator a) no optimisation, no chopping; b) optimisation with chopping

Despite the fact that the loop filter is linear, the whole modulator is a highly non-linear device because of the comparator and 1 bit D/A converter. The quantization noise is assumed to be random but in reality it is not, and also the signal is not small as predicted by the AC small-signal-analysis method. For this reason the Monte-Carlo time-domain method is used in our simulations.

Figure 7 shows the noise simulation result for the 5th-order modulator: (a) before optimisation (without chopping) and (b) after optimisation (with chopping). The unit capacitor sizes and the thermal noise requirement for the opamps after optimisation if the S/N ratio of 110dB is required, are given in figure 9. By optimising capacitor sizes, the area of the capacitor arrays can be reduces by almost 10 times and the power consumption can be drastically reduced because the capacitive loads are lower, which is also shown on figure 8, which represents the layout of the realized modulator. It is evident that the area of the capacitor array is the biggest for the first integrator and than smaller for the second while the third and the following integrators have unit capacitor size of 0.3pF and thus need very small area.

Integ	Cu[pF]	ΣC[pF]	$V_n\left[\frac{nV}{\sqrt{Hz}}\right]$
1	15	66	≤9
2	1.2	24	≤85
3	0.3	4.4	≤1000
4	0.3	2.8	≤1000
5	0.3	2.0	≤1000

Figure 9: Table of capacitors before and after optimisation and opamp noise requirements

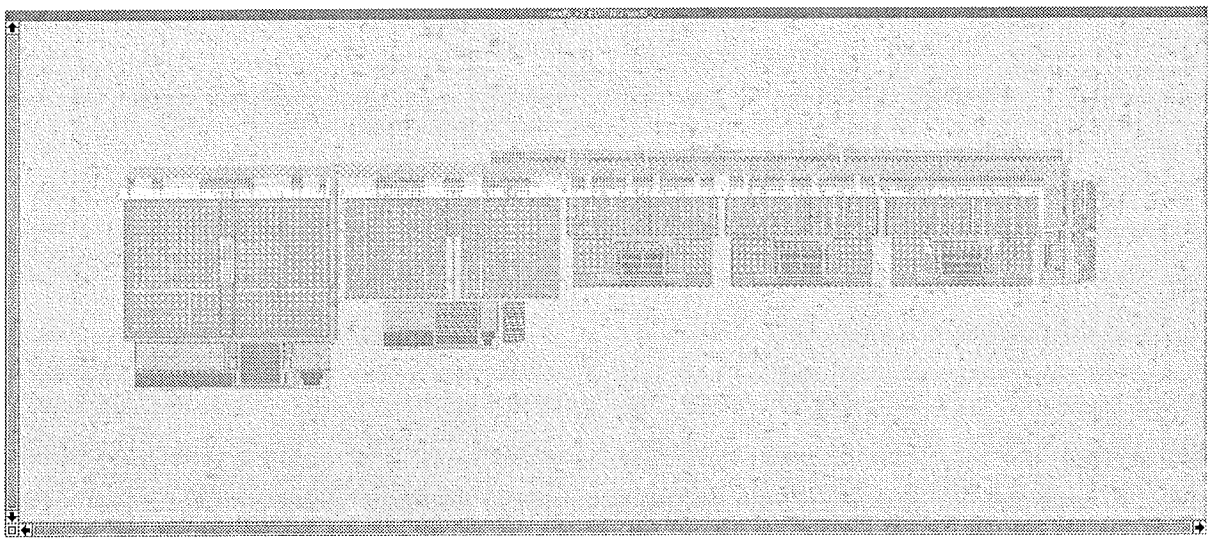


Figure 8: Layout of the realized 5th-order modulator

3. OPTIMIZATION ALGORITHM

The block diagram of a general Δ - Σ modulator is presented in figure 10, /3/. Its loop filter is generally an n^{th}-order FIR or IIR analog or sampled data filter. The behavior of the whole modulator is most generally and efficiently described by a combination of a linear time-domain multi-input single-output state-space description of the loop filter (state variables are outputs of the integrators) and of non-linear part, which describes single-bit or multi-bit quantizer (equation 3).

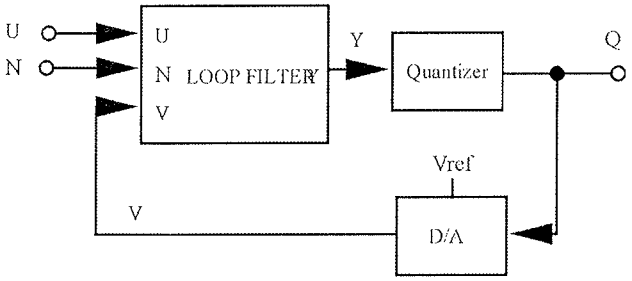


Figure 10: Block diagram of Δ-Σ modulator

$$x(n) = Ax(n-1) + Bv(n-1) + Nn(n-1)$$
$$y(n) = c^T x(n-1) + Dv(n-1)$$
$$v(n) = f_Q(y(n-1))V_{ref} \qquad (3)$$

The relation between the matrix elements and capacitors is defined for an arbitrary S-C loop filter in /5/. This model is improved by additional noise and linear and non-linear settling parameters. For convenience the simulation can be performed by turning on or off the chopping effects.
The optimisation algorithm gets its input data from the synthesis procedure. All the capacitor ratios are known, as is the architecture of the modulator. At the beginning the program assumes equal capacitors for all integrator stages. They are calculated from the S/N and oversampling-ratio requirements. The same thermal noise requirement for the opamps is taken as a basis. The time-domain Monte-Carlo simulation is then performed for each switched capacitor and each opamp in a modulator according to /5/. The contribution of each element is calculated in a band 0Hz to f_n and saved, then the capacitor sizes and the opamps' thermal noise parameters are adjusted according to these contributions in such a way that each noise source adds approximately the same contribution. The result is a table of unit capacitor sizes, the sum of all the capacitors for each integrator and the thermal noise requirements for the opamps. For our example the result is given in figure 9. We assumed that chopping of the first amplifier is performed, so the 1/f noise is translated around $f_s/2$ and later removed by the decimation filter. At the end a complete simulation of the modulator with optimised elements is performed. The result is presented in figure 7 (b).

4. CONCLUSIONS

A methodology for power-consumption optimisation for an S-C high order, high-resolution modulator is presented. The main reasons for power consumption are explained and a method to reduce the effects is proposed. An algorithm coded in MATLAB is developed and the results of an optimisation for a practical example of the 5^{th}-order single-bit Σ-Δ modulator are given. The required power consumptions of two modulators with equal architecture and different capacitors and opamps are compared. The power consumption for an optimised structure is reduced by more than 5 times compared to the non-optimised, while the capacitor array area is reduced by almost 10 times.

The advantage of the proposed algorithm is the speed and the effectiveness of the optimisation procedure and the possibility of performing the optimisation and the noise simulation in a reasonable time. It is only possible to use a simplified mathematical model for the circuits involved, which is the main limitation of the approach.

References

/1/ D.Strle, "A system design considerations of high resoltion, l ow power Δ-Σ A/D converter", tech. rep., University of Ljubljana, Faculty of Electrical Engineering, Tržaška 25, Ljubljana, 2000.

/2/ V.Peluso, P.Vancorenland, A.M.Marques, M.Steyaert, and W.Sansen, "A 900-mv low power Δ-Σ A/D converter with 77-db dynamic range", IEEE JSSC, vol.33, pp1887-1897, Dec 1998.

/3/ J.C.Candy and G.Temes, "Oversampling methods for A/D and D/A conversion in Oversampling Delta-Sigma Converters", IEEE Press, 1991.

/4/ R.Gregorian and G.Temes, "Analog MOS integrated Circuits for signal processing", John Wiley and Sons, 1986.

/5/ D.Strle, "Noise modelling and simulation of high resolution Δ-Σ A/D converters", Proceedings of the 1997 International MIDEM Conference, pp148-154, 1997.

Dr. Drago Strle dipl. ing.
University of Ljubljana,
Faculty for electrical engineering,
Tržaška 25, 1000 Ljubljana, Slovenia.
e-mail: drago.strle@fe.uni-lj.si

Prispelo (Arrived): 05.02.01 *Sprejeto (Accepted): 07.03.01*

FREQUENCY WARPING AND CHAOTIC BEHAVIOUR GENERATED BY SPICE

Boštjan Peršič, Niko Basarič
University of Ljubljana, Slovenia

Keywords: physics, electrotechnics, circuit analysis, SPICE computer tools, DSP, Digital Signal Processing, frequency warping, numerical integrations, Z-transform, LAPLACE transform, chaos

Abstract: The numerical procedure of the SPICE simulator is a compromise between accuracy and speed, so that due to efficiency, the simulator can yield errors. One source is frequency warping, which is discussed in the article. The simulator is presented as a digital signal processor. General terms describing the distortion emanating from a time discrete treatment of the continuous signal are achieved in the frequency domain by comparing the *Z*-transform of the simulator model to the Laplace transform of the model of the actual circuit. The theoretical derivations are illustrated by examples treating autonomous circuits. Within the first example, the SPICE transient analysis is analysed and the discrepancy between actual and calculated responses is presented in the time domain. In addition to frequency warping, the discrete treatment can generate chaotic behaviour. This is presented in the second example, an analysis of a CL oscillator. This example demonstrates the influence of magnitude of the numerical integration step. It is shown that if a step is not limited, the simulation does not follow the behaviour of the circuit and an unpredictable shape is output. The calculated response seems to be chaotic, despite the actual circuit having a closed limit cycle.

Frekvenčno izkrivljanje in kaotično obnašanje simulatorja SPICE

Ključne besede: fizika, elektrotehnika, analize vezij, SPICE orodja računalniška, DSP obdelava signalov digitalna, izkrivljanje frekvenčno, integracije numerične, Z-transformacija, LAPLACE transformacije, kaos

Povzetek: Numerični postopki, ki jih uporablja simulator SPICE, so kompromis med učinkovitostjo in točnostjo. Članek opisuje nekaj vidikov frekvenčnega izkrivljanja in kaotičnega obnašanja CL oscilatorja pri analizi s simulatorjem SPICE. Simulator smo predstavili kot digitalni procesor signala. Splošne analitične izraze za popačenja, ki nastanejo zaradi diskretnega obravnavanja zveznega signala, smo izpeljali v frekvenčnem prostoru prek primerjave rezultatov *Z*-transformacije odziva modela simulatorja in Laplaceove transformacije odziva dejanskega vezja. Teoretične izpeljave ilustrirata dva zgleda analize avtonomnih vezij. Prvi zgled obdeluje linearno vezje. Napake tranzientne analize simulatorja SPICE so prikazane v prostoru stanj in časovnem prostoru. Drugi zgled je nelinearno vezje, pri katerem je poudarjen vpliv velikosti koraka numerične integracije. Če je maksimalni dopustni korak numerične integracije prevelik, analiza vezja s simulatorjem ne sledi obnašanju dejanskega vezja, temveč generira nepredvidljive rezultate. Izračunani izhodni signal je tedaj videti kaotične oblike, čeprav ima obravnavano vezje zaključen limitni cikel.

1. INTRODUCTION

Implementation in microelectronics technology is impossible without the use of a wide palette of CAD tools. The basic tool, which yields circuit responses at the level of elements, is the SPICE simulator (Simulation Program with Integrated Circuits Emphasis) /1/. The simulation is performed on a digital computer, so its behaviour is similar to digital signal processing. Digital signal processing does not encompass the entire signal; it treats a chain of discrete samples only, and the samples have a limited number of values. This discrete approach introduces some inevitable impairment of the signal. The transformation of the magnitude from continuous to discrete values is a distortion that is presented as quantization noise. The amount of the noise is inversely proportional to the distance between two neighbouring levels. Similar distortion is caused by errors emanating from unavoidable truncations and by rounding off the intermediate results of mathematical operations /2,3/. Naturally, the noise impairs the signals calculated. In addition, manipulation of a signal in discrete instants of time is the second source of errors /2/. The sampling theorem states that an original signal can be reconstructed by its samples if they are nearer each other than half a period of the highest spectral component. If this condition is violated, the spectrum transposed around multiples of the sampling frequency is added to the baseband spectrum, which causes irreparable corruption of the original signal.

A detailed analysis reveals that the treatment of a signal in discrete instants introduces distortion, even if they are close enough to fulfil the requirement of the theorem. This distortion is called frequency warping. Its size depends on the ratio between the frequency of the signal and the sampling frequency. Assuming that an exact result would be obtained incorporating an infinite number of samples, the difference between the actual result, acquired by a limited amount of samples, and the exact one can be defined as the error. The quantitative estimate of frequency warping is normally done in the frequency domain by comparing a *Z*-transform of the discrete system to a Laplace transform of the continuous counterpart /3, 4/. The error caused by the discrete approach comes into play as the observed system is excited by a different frequency, which changes the shape of the response /3/. Within some applications, analog circuit simulators

evaluate the response in discrete points of time; therefore, they have all the essential properties of a digital signal processor. The feature denoted is incorporated into the SPICE simulator. The result of SPICE transient analysis can be presented as an output of a digital integrator /5, 6/ whereby sampling time, numerical accuracy and the integration method can be selected. Due to its discrete nature, numerical integration causes the distortion mentioned above.
Contrary to standard digital signal processing, SPICE does not maintain a constant distance between samples. In order to accelerate analysis, the integration step of the simulator continues until an initially controlled error surpasses the threshold value chosen. Interrupting a lengthy step by an abrupt change in a signal can be another source of errors. The error can even transform an undoubtedly unstable response into a stable one /7/. In addition, it can cause completely irregular results /8/. Despite the possibility of chaotic behaviour in a Colpitts oscillator /9, 10/, a circuit which has a simple closed limit cycle with no bifurcation can yield a strange response with no steady state. It has been confirmed that the unexpected results are consequences of the imprecise numerical integration of the simulator.

2. MODEL OF SPICE TRANSIENT ANALYSIS

SPICE transient analysis has been created to be a numerical solver of differential equations. In fact, the analysis has been adapted to cover non-linear circuits and uses implicit integration methods /1/, where the magnitudes of signals and their derivatives are simultaneously calculated by previous values. However, since we wish to avoid an overly intricate explanation by omitting facts not essential to our topic, the circuit analysed is assumed to be linear and is presented by a system of equations in a normal form. The excitation vector, the state vector, and its time derivative are depicted by $\mathbf{u}(t)$, $\mathbf{y}(t)$ and $\dot{\mathbf{y}}(t)$ respectively. **A** and **B** are corresponding matrices.

$$\dot{y}(t) = Ay(t) + Bu(t) \qquad (1)$$

The explicit form of the signal $y(t)$ can be obtained by integrating the differential equation.

$$y(t) = \int_{-\infty}^{t} \dot{y}(t)dt = \int_{-\infty}^{t} (Ay(t) + Bu(t))dt \qquad (2)$$

The simulator captures and manipulates signals at discrete time instants only. Suppose that these instants are equidistant and that they are h units of time apart. If this is so, vectors can be replaced by their samples $u(t) \rightarrow u(kh) = u_k$, $y(t) \rightarrow y(kh) = y_k$, $\dot{y}(t) \rightarrow \dot{y}(kh) = \dot{y}_k$. The usual numerical integration technique exploited by the simulator is the trapezoidal method. The procedure for one integration step is presented by (3) where h depicts the time step.

$$y_{k+1} = y_k + \frac{h}{2}\dot{y}_k + \frac{h}{2}\dot{y}_{k+1} \qquad (3)$$

Derivatives of the state vector in the equation above can be replaced by the right side of the equation (1) for matching instants. After some rearrangement, (4) is obtained:

$$y_{k+1} = y_k + h\frac{A}{2}(y_k + y_{k+1}) + h\frac{B}{2}(u_k + u_{k+1}) \qquad (4)$$

A system operating according to equation (4) is sketched in Figure 1. The block z^{-1} depicts a time delay lasting one integration step. The input signals of the middle summator are formed by the mean value of the present and past samples of the output signal and the excitation, multiplied by **A** and **B** respectively. The integrator is presented as an accumulator.

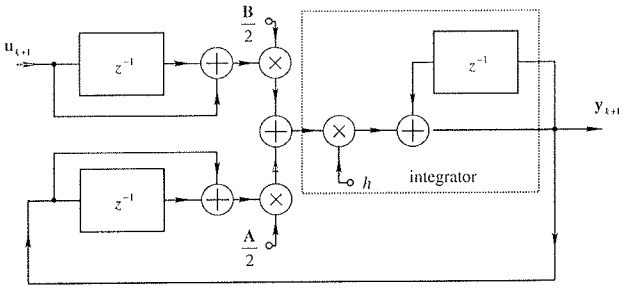


Fig. 1. Model of the trapezoid method

Figure 1 confirms that numerical integration can be treated similarly to digital processing of signals, and that consequently the SPICE transient analysis can be regarded as a digital signal processor. The logical conclusion is that this task of the simulator has all the main features, including deficiencies, of a processor of this kind.

3. FREQUENCY WARPING IN THE FREUENCY DOMAIN

Frequency warping is the phenomenon where a discrete system corresponds to its continuous counterpart at a distinct frequency. This discrepancy increases significantly if the frequency of the signal approaches a value limited by the sampling theorem. An error can be caused if this fact is neglected when analysing an oscillator /4/. Our quantitative description of the phe-

nomenon inside the SPICE analysis is divided into steps, as follows.

3.1. Comparison of Transfer Functions

The transfer function of the continuous integrator can be obtained by performing a Laplace transform on basic equations (1) or (2) that describe the behaviour of the system.

$$H(s) = \frac{Y(s)}{U(s)} = \frac{B}{s - A} \tag{5}$$

Equation (3) presents the signal in discrete time instants; therefore, the natural approach is a *Z*-transform. The transfer function of a numerical integrator which exploits the trapezoidal method is given by (6).

$$H(z) = \frac{B}{\frac{2}{h} \cdot \frac{z-1}{z+1} - A} \tag{6}$$

A comparison of equations (6) and (5) confirms that the discrete approach alters the transfer function of a system. Consequently, a change in the transfer function results in a difference between the responses of the discrete system and its continuous (e.g. actual) counterpart. Nevertheless, it can be noticed that the structure of equation (5) is equal to the structure of (6), except that the variable *s* is replaced by the fraction. The transfer functions are identical if mapping, defined by equation (7), is introduced.

$$s = \frac{2}{h} \cdot \frac{z-1}{z+1} \quad \Rightarrow \quad z = \frac{2+hs}{2-hs} \tag{7}$$

As the variables of the transforms have both real and imaginary parts, the last equation describes two-dimensional mapping. The map of a significant curve is additionally highlighted in the next subsection.

3.2. Map of *s*-Plane Ordinate to *z*-Plane

The continuous complex frequency *s* is expressed as a sum of the real and imaginary parts ($s = \alpha + j\omega$), and variable *z* by its polar co-ordinates ($z = r \cdot e^{j\omega h}$). To distinguish the frequencies, the upper index in the parenthesis is added at the discrete signal ($\omega_{discrete} \Rightarrow \omega^{(d)}$). The portrayal of harmonic signals in the *s*-plane is its ordinate. As our main concern is studying relations between the frequencies of continuous and discrete systems, a map of this curve into the *z*-plane can explain the point. After replacing the variable *s* in equation (7) by $j\omega$ and after some manipulation, the relation presented by (8) can be found.

$$r = 1; \quad \omega^{(d)} = \frac{2}{h} \arctan \frac{\omega h}{2} \tag{8}$$

The map of the ordinate of the *s*-plane is a circle in the *z*-plane, as shown in Figure 2. The frequency of the discrete system corresponds to the phase in the *z*-plane.

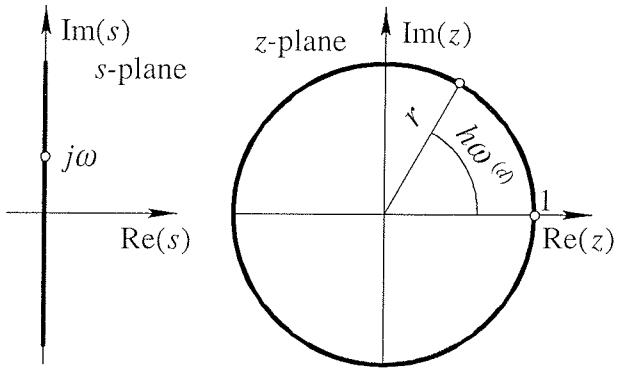


*Fig. 2. Map of the ordinate of the **s**-plane into the **z**-plane according to the trapezoidal method.*

The relation between continuous and discrete frequencies, dictated by equation (8), is presented in Figure 3. The abscissa is proportional to the frequency in a continuous (actual) system, and the ordinate to the frequency of its discrete counterpart (simulated system). The scales of both axes are normalised by the sampling frequency, which is calculated from the time step (f_s=1/*h*).

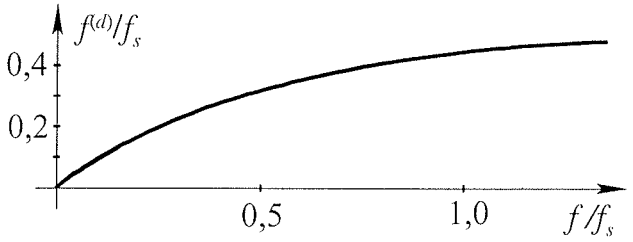


Fig. 3. Relations between frequencies of the continuous system and the discrete system if equal responses are required.

The curve reveals the relation between the systems. It can be inferred that the test tone frequency inside the simulator ($f^{(d)}$) has to differ from the actual frequency (f) if the same magnitude of response is required. If the ratio between the test and the sampling frequency is small, the value of the function arc tan is almost equal to its argument and the differences between the responses are hardly observed. When the ratio approaches or even exceeds one half, the corresponding discrete frequency is significantly distinct.

This non-linear relation emanates from the treatment of the signal in discrete time instants.

3.3. Approximation

Equation (8) is the exact description of the distortion of a signal due to treatment of a continuous system by the simulator. The error can be easily estimated if a non-linear tangent function is expanded into a truncated Taylor series. The result of the procedure is equation (9). A detailed analysis reveals that with ten samples per period, the first neglected term is approximately 4% of the last included term.

$$\omega \approx \omega^{(d)}\left(1+\frac{(\omega^{(d)}h)^2}{12}\right) \qquad (9)$$

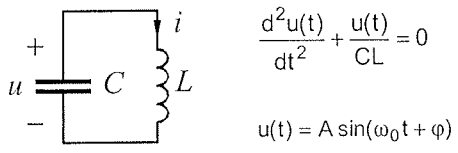
As the integration step of the SPICE transient analysis is adapted automatically to changes in the signal [11], the number of steps in a period is normally greater than ten, especially if the signal is formed by a single spectral component. With this in mind, the approximations of the distortion seem acceptable for obtaining a fair estimate of an error of frequency obtained by the simulator.

4. EXAMPLE I - FREQUENCY WARPING

This section illustrates the subject through an example. It aims to display the typical effect of frequency warping and to clarify the relation (9).

4.1. The circuit

The simplest form of an autonomous circuit, producing a response that incorporates only the imaginary part of the complex frequency ($s = j\omega_0$), is a CL combination with an initial state ($u(0)$, $i(0)$) different from zero. The circuit is presented in Figure 4. It can be described by an homogeneous second order linear differential equation. Its solution is undamped oscillation .



$$\frac{d^2u(t)}{dt^2}+\frac{u(t)}{CL}=0$$

$$u(t) = A\sin(\omega_0 t+\varphi)$$

Fig. 4. The circuit in Example 1

4.2. Numerical Integration in the Time Domain

Frequency warping of the example is demonstrated in the time domain. The trapezoidal method has no amplitude distortion /3/; therefore, any point of the signals acquired by numerical integration defines a sine curve with given constant amplitude.

The trapezoidal algorithm (3) includes the initial and the final derivatives of the signal within each integration step. For convenience, the step should be divided into two equal parts. During the first part, the signal of the discrete integrator is assumed to be changing according to the slope through the initial point. This slope can be unambiguously found from the known point of the signal. During the second half of the interval, the simulator follows the tangent of the final point. Presentation of the outcome without calculation is rather intricate, as the final point is unknown. For a particular case, the obscurity can be bypassed by choosing an initial point on the abscissa (y_0=0) and by setting the magnitude of the integration step to the value where the signal reaches amplitude during its first half. With this intermediate result, the only solution at the end of the integration step is the slope parallel to the abscissa, which holds only at the maximum value of the harmonic signal. The point after the second integration step is obtained similarly. The first move within the step is parallel to the abscissa, and the second has maximum slope: therefore the abscissa is reached again, as shown in Figure 5.

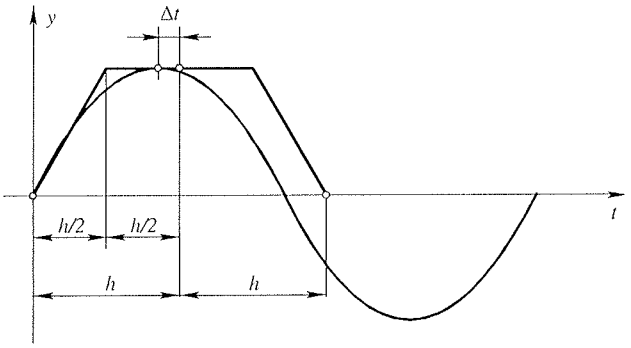


Fig. 5. Frequency warping demonstrated in time domain.

The three points describe half a period of the calculated signal. It can be observed in Figure 5 that the peak of the signal obtained is shifted rightwards compared to the actual signal. The delay is depicted by Δt. Sequential integration steps add the same delay and the period is prolonged, which means that the frequency of the response is lowered. The difference can be approximated by (9).

5. EXAMPLE II - CHAOS

As noted, transient analysis of the simulator consists of numerical integration with discrete time increments. A circuit has to be solved inside each step, which results in a considerable amount of calculations if the increments are tiny. To output a result in a reasonable

time, the standard SPICE simulator tends to enlarge the increment towards $(T_{stop}-T_{start})/50$ or T_{max}, whichever is smaller /11/. The times mentioned are the second, third and fourth parameters in the .TRAN statement respectively. Another limitation stems from the slopes of signals. Slowly varying signals enlarge the time increment, and vice versa. If an abrupt change occurs in a signal, the integration step shrinks gradually. The gradual change and asynchronism between the integration instants and the abrupt change can be another source of errors.

5.1. The circuit

The active element of the circuit in Figure 7 is a general purpose NPN transistor in a common emitter orientation, loaded by a resistor. Two capacitors and an inductor form the feedback. The latter provides the quiescent current into the base, so that the transistor operates near saturation. Note that the ratio of capacitances causes a considerable signal at the base; therefore, the transistor is forced to traverse a strongly non-linear region. The non-linearity yields significant digress of actual frequency of oscillation (18.5 kHz) from the resonant frequency of the passive π feedback circuit (8.9 kHz).

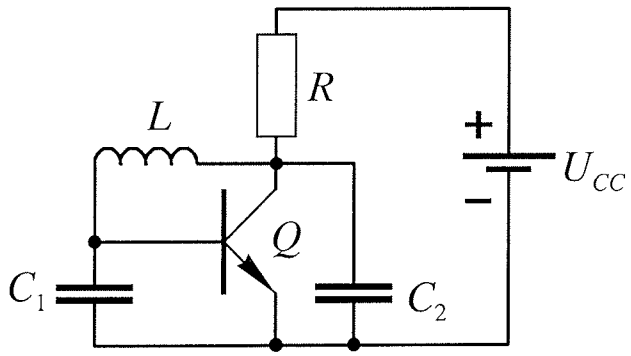


Fig. 7. The oscillator analysed (U_{CC} = 5 V, Q = 2N2222, R = 1 k, C_1 = 47 nF, IC = 0.6074422 V, C_2 = 100 nF, IC=1.361650 V, L = 10 mH, IC = 28.98162E-6 A)

5.2. Results of simulations

The first solution presented is referred to as exact. Setting a small enough time between outputted results (T_{step}) and the maximum time increment of the numerical integration (T_{max}), and starting the simulation with adequate initial conditions, an actual steady-state response is obtained. Observing the actual time increment of the integration in the raw file, it was confirmed that T_{max} chosen defines the time increment inside the entire period. Figure 8 displays the current into the base, the collector-emitter voltage in the time domain and the current through the inductor in the frequency domain.

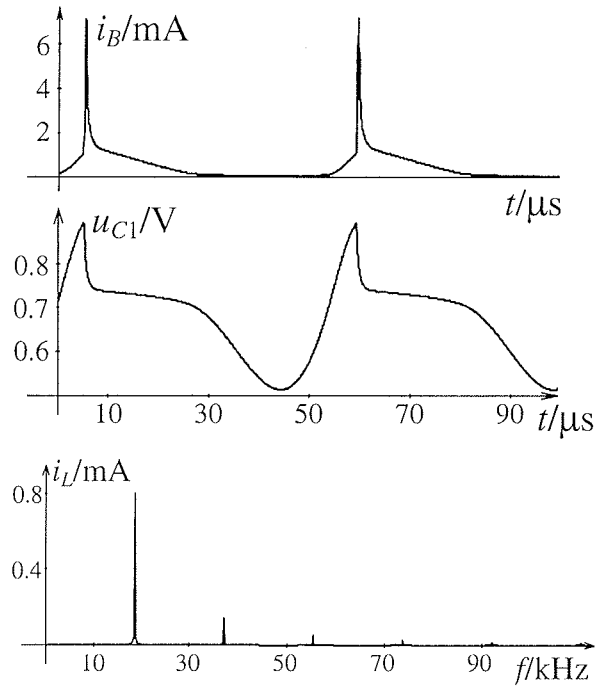


Fig. 8. Base current, base-emitter voltage and the spectrum of the inductor current ($T_{step} = T_{max}$ = 50 ns)

Leaving T_{max} unchanged and T_{step} significantly increased, the simulation is repeated. Figure 9 shows the results in the state domain formed by the voltage across the capacitor at the collector side and the inductor current. The upper part of the trajectory is spitted. The inset explains that the vertices of curves are merely different points of the original trajectory. Thus, presenting the signals using a small number of points, the distortion appears.

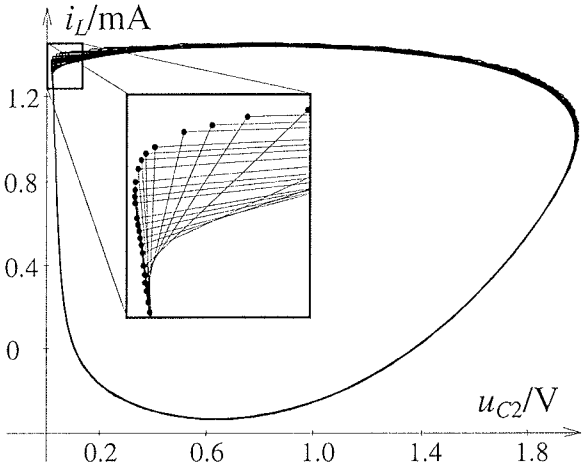


Fig. 9. The trajectory with increased T_{step} (T_{step} = 1,5 μs, T_{max} = 50 ns)

A simulation with slightly increased T_{max} causes the spectral lines to be lower and wider and the noise floor raised. In the state domain, the points calculated randomly diverge from the original trajectory, but the limit cycle can be unambiguously confirmed. Omitting T_{max} leads to an unpredictable yield of the simulator (Figure 10). The spectrum has some peaks around the original frequency of oscillation. The spikes are less than 6dB above the vicinity, so all spectral components are significant inside a limited interval. In addition, the trajectory gained is not a closed curve; moreover, it seems chaotic.

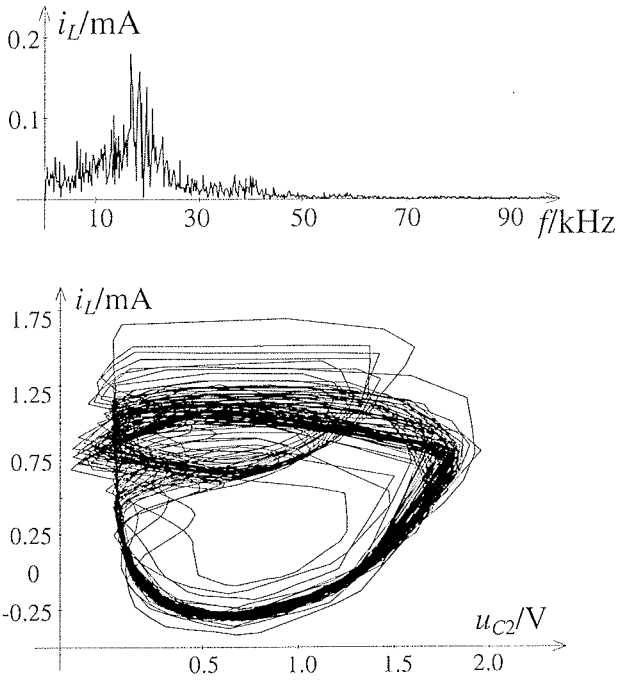


Fig. 10. The spectrum and the trajectory (T_{step} = 1,5 μs, T_{max} = omitted).

5.3. Comments

Observing the current at the base side (Figure 8), it can be inferred that significant changes occur when the transistor passes from the cut-off to the active region. As the amplifier is idle for almost the entire period except inside this short interval, the slopes of signals are defined mainly by passive feedback. Thus, all signals are smooth before the spike of the base current occurs. In observing the raw results, it can be found that the integration increment has its maximum allowed value at the beginning of the abrupt change. Overly large steps cause alternation of the calculated base current. If this numerically provoked alternation is not damped enough, the process can continue through the entire period. Not being synchronised with the signal, the moments of calculation are not congruent to the response. Given this, the calculated amount of the charge injected into the base randomly varies from cycle to cycle, so the response is aperiodic.

2. CONCLUSION

The aim of the contribution is to highlight some deficiencies of the SPICE simulator. In addition, it explains the origin of frequency warping and the irregular behaviour due to manipulation of a signal in discrete time instants. Considering the equidistant time steps, a similar approach is found within digital signal processing, so the transient analysis of the SPICE simulator is presented as a digital signal processor. Expressions for the distortion caused by numerical integration are derived in the frequency domain. It is shown that the discrepancy between a continuous system and its discrete counterpart increases if a time steps approach is made to periods of signals. A simple incorporated electronic circuit, the solution to which is harmonic oscillation, is presented to illustrate these features. The circuit is simulated exploiting numerical integration. Within this example, the distortion is additionally surveyed in the time domain. In addition, the SPICE simulator varies the time step of numerical integration, which introduces another error, treated in the second example. The circuit used in the example is a Colpitts oscillator. It has been confirmed that the simulator outputs the periodic signals if adequate controls are set. However, the automatic adaptation of the time increment leads to an aperiodic response.

It must be stressed that the SPICE simulator is an excellent tool for circuit analyses, and usually accomplishes its job requiring no additional actions by its users. Sometimes the facts mentioned above prevail, so outcomes contain errors. We hope that the paper clearly elucidates some unpleasant features of the SPICE simulator, and trust that consideration of the features and their consequences will facilitate detection of false results. We advise that a survey of the simulator's outputs should be made and if they seen doubtful, additional controls must be utilised. As the examples illustrate, the most powerful control is the limitation of the time step.

References

/1/ L.W. Nagel, A Computer Program to Simulate Semiconductor Circuits. Memorandum No. ERL-M520, Berkeley, 1975.

/2/ James V. Candy, Signal Processing: the Modern Aproach. McGraw-Hill, NY, 1988.

/3/ Andreas Antoniu, Digital Filters Analysis, Design and Applications. McGraw-Hill, NY, 1993.

/4/ B. Peršič, "SPICE Transient Analysis Errors Estimated in Frequency Space", IEE Electronics Letters, Vol. 30, No. 8, pp. 617-8, April 1994.

/5/ A. Brambila, D. D'Amore, "The Simulation Errors Introduced by the SPICE Transient Analysis," IEEE Trans. CAS I, Vol. 40 No. 1, pp. 57-60, Jan. 1993.

/6/ K.G. Nichols, T.J. Kazimerski, M. Zwolinski, A.D. Brown, "Overview of SPICE-Like Circuit Simulations Algorithms," IEE Proc.: Circuits Devices Systems, Vol. 141, No. 4, pp. 242-250, August 1994.

/7/ P. Kinget, J. Crols, M. Ingles, E. Peluso, "Are Circuit Simulators Becoming Too Stable", IEEE Circuits and Devices Magazine, Vol. 10, No. 3, pp. 50, May 1994.

/8/ B. Peršič, I. Medič, "Chaotic Results of the SPICE Simulator," Proceedings of ECCTD '97, pp. 1226-1230, Budapest 1997.

/9/ G. Sarafin, B. Kaplan, "Is the Colpitts Oscillator Chua's Circuit", IEEE Trans. CAS-I, Vol. 42, No. 6, pp. 373-6, 1995.

/10/ M.P. Kennedy, "On the Relationship Between Chaotic Colpitts Oscillator and Chua's Circuit", IEEE Trans. CAS-I, Vol. 42, No. 6, pp. 376-9, 1995.

/11/ SPICE 3F4 online help, Berkeley, 1993.

Boštjan Peršič
University of Ljubljana
Tržaška 25, 1000 Ljubljana, Slovenia
bostjan.persic@fe.uni-lj.si

Niko Basarič
University of Ljubljana
Tržaška 25, 1000 Ljubljana, Slovenia
niko.basaric@fe.uni-lj.si

Prispelo (Arrived): 23.06.00 *Sprejeto (Accepted): 22.11.00*

VITERBIJEV ALGORITEM ZA DSP PROCESORJE

**Srečo Plevel, Tomaž Javornik, Igor Ozimek,
Roman Trobec and Gorazd Kandus
Institut Jožef Stefan, Ljubljana, Slovenia**

Ključne besede: sistemi komunikacijski digitalni, naprave radijske mobilne, kodiranje, dekodiranje, VITERBI algoritem, DSP obdelava signalov digitalna, DSP procesorji, kodiranje konvolucijsko, dekodiranje trdo, dekodiranje mehko, TMS320C4x Texas Instruments DSP procesorji, TMS320C6x Texas Instruments DSP procesorji

Povzetek: Sodobni digitalni komunikacijski sistemi, zlasti mobilne radijske naprave, so v vse širši uporabi. Za kvaliteten prenos podatkov in učinkovito izrabo razpoložljivega radijskega frekvenčnega spektra so potrebni kvalitetni modulacijski in kodirni postopki. Ena od pomembnih tehnik kodiranja je konvolucijsko kodiranje in dekodiranje z uporabo Viterbijevega algoritma. Modulacijski in kodirni postopki so relativno zahtevni za obdelavo v realnem času in so se običajno izvajali v specializiranih vezjih. Razvoj vedno bolj zmogljivih univerzalnih signalnih procesorjev pa omogoča programsko izvedbo teh postopkov in s tem koncept t.i. programljivega radia. Ta prinaša razne prednosti, med drugim enostavno spreminjanje in s tem prilagodljivost raznim komunikacijskim sistemom in standardom. V članku je predstavljena izvedba Viterbijevega algoritma za dve družini univerzalnih signalnih procesorjev TMS320C4x in TMS320C6x. Opisani so razni načini optimizacije in pohitritve Viterbijevega algoritma ter predstavljene dosežene zmogljivosti algoritma pri izvajanju na teh dveh procesorjih tako za trdo kot mehko dekodiranje.

Viterbi Algorithm for DSP Processors

Keywords: digital communication systems, mobile wireless devices, coding, decoding, VITERBI algorithm, DSP, Digital Signal Processing, DSP processors, convolution coding, hard decoding, soft decoding, TMS320C4x Texas Instruments DSP processors, TMS320C6x Texas Instruments DSP processors

Abstract: Modern digital communication systems, especially mobile wireless devices, are playing more and more important role in everyday life. For reliable data transmission and efficient utilisation of limited frequency spectrum resources, bandwidth-efficient modulation and coding must be employed. An important coding technique is convolutional coding and decoding with Viterbi algorithm. Modulation and coding procedures are relatively demanding for digital real-time processing and until recently they were mostly implemented in specialised integrated circuits. The development of ever more powerful general-purpose signal processors came to the point, where software implementation of these procedures is viable. This approach, known as software radio, has a number of advantages. It allows easy modification and adaptation to various communication systems and standards. In the article, the implementation of the Viterbi algorithm is described for two families of general-purpose digital signal processors, TMS320C4x and TMS320C6x. Various optimisations are presented together with the resulting performance in terms of achievable bit rates and error bit rates for hard and soft Viterbi decoding.

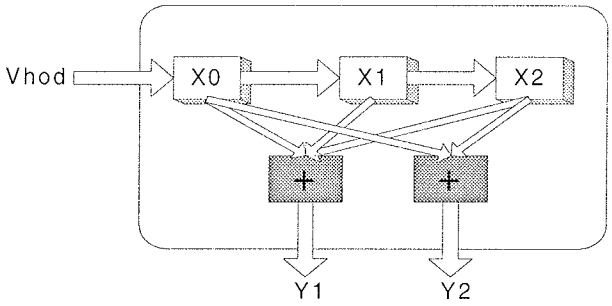
1. Uvod

V sodobnih digitalnih telekomunikacijskih sistemih se uporabljajo za zmanjšanje verjetnosti bitne napake kanalni kodirni postopki. Ti vnašajo v podatkovni tok dodatne bite, ki sprejemniku omogočajo popravljanje napak pri prenosu. Kanalne kodirnike delimo v bločne in konvolucijske. Pri bločnem kodiranju razdelimo bitni tok na posamezne bloke in vsakemu dodamo dodatne (redundantne) bite. Dekodiranje (s popravljanjem napak) se izvaja neodvisno na vsakem bloku posebej, nakar se izhodni biti sestavijo nazaj v neprekinjen bitni tok, seveda z zakasnitvijo, ki je posledica sprejemanja in obdelave posameznega bloka. Pri konvolucijskih kodirnih postopkih poteka kodiranje in dekodiranje sproti na neprekinjenem bitnem toku. Posamezen simbol na izhodu kodirnika je odvisen od trenutnega vhodnega simbola in od stanja kodirnika, to pa od določenega števila predhodnih vhodnih simbolov. Število izhodnih bitov je večje od vhodnih za nek faktor, kar zagotavlja redundanco in s tem možnost odpravljanja napak. Za dekodiranje konvolucijsko kodiranih signalov se uporabljajo različni algoritmi, najpogosteje pa Viterbijev algoritem. Na tržišču obstaja vrsta integriranih vezij za dekodiranje konvolucijsko kodiranih signalov, vendar je večina od njih vezana na določen konvolucijski kodirnik, torej na določen komunikacijski sistem. V zadnjih nekaj letih se je na področju mobilnih komunikacij pojavil koncept programljivega radija, za katerega je značilno, da se radijski vmesnik menja glede na storitve, ki jih zahteva uporabnik, in stanje radijskega kanala med sprejemnikom in oddajnikom. Eden glavnih gradnikov v sistemu programljivega radija so digitalni signalni procesorji - DSP. Viterbijev algoritem smo priredili tako, da ga je mogoče izvajati v digitalnih procesorjih družine Texas Instruments TMS320C4x in TMS320C6x.

2. Opis konvolucijskega kodiranja

Konvolucijsko kodiranje bomo prikazali na preprostem primeru kodirnika z enim vhodom in dvema izhodoma, pri katerem je izhodni 2-bitni simbol odvisen od trenutnega vhodnega bita in dveh predhodno poslanih bitov. To pomeni, da ima kodirnik dva registra, torej $2^2 = 4$ možna notranja stanja. Bločna shema kodirnika je prikazana na sliki 1. Konvolucijske kodirnike pogosto označujemo z zapisom (št. izhodnih bitov, št. vhodnih

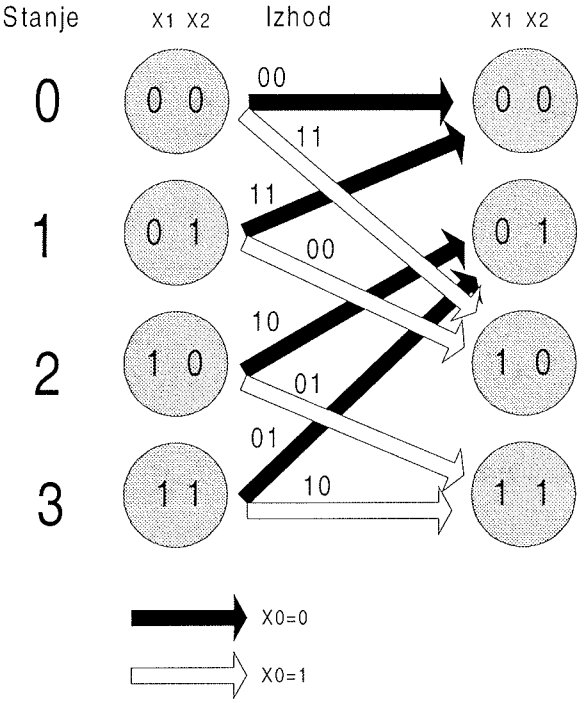
bitov, št. registrov), v našem primeru je oznaka (2,1,2). Ta zapis ne navaja povezav med zakasnilnimi členi in izhodi, zato z njo kodirnik še ni enolično določen.



Slika 1: Konvolucijski kodirnik (2,1,2)

Registra X1 in X2 predstavljata notranje stanje kodirnika, register X0 pa je vhod v kodirnik. Ko je na vhodu v kodirnik nov bit, se shrani v skrajno levi register (na sliki X0), vsi ostali pa se pomaknejo za eno mesto v desno. Indeks notranjega stanja predstavlja kar binarni zapis registrov X1 in X2.

Diagram prehajanja stanj prikazuje slika 2. Črne puščice pomenijo prehod stanja pri vhodnem bitu 0, bele pa pri vhodnem bitu 1. Binarni zapis v krogih pomeni stanja binarnih registrov pomnilnika konvolucijskega kodirnika. Iz diagrama prehajanja stanj je razvidno, da na vsakem koraku nov vhodni bit pride na najpomembnejše binarno mesto (z leve strani), najmanj pomemben bit (desni) pa odpade.



Slika 2: Diagram prehajanja stanj konvolucijskega kodirnika (2,1,2)

Če označimo bita notranjega stanja *x1* in *x2*, bit na vhodu pa *x0*, potem izhod kodirnika zapišemo z naslednjima enačbama:

$$y1 = x0 + x1 + x2$$
$$y2 = x0 + x2$$

Pri tem pomeni znak + vsoto po modulu 2 oziroma operacijo *izključujoči ali* (*xor*). Pri vsakem prehodu iz trenutnega v naslednje stanje sta izhodna bita kodirnika enolično določena s prehodom, njuna vrednost je v sliki 2 zapisana nad vsakim prehodom. Izhodni bitni tok opisanega konvolucijskega kodirnika ima dvojno število bitov v primerjavi z originalnim nekodiranim podatkovnim bitnim tokom na vhodu in zahteva zato dvakrat večjo hitrost prenosa v zameno za večjo zanesljivost.

3. Viterbijev dekodirnik

Za dekodiranje konvolucijsko kodiranega signala se običajno uporablja Viterbijev algoritem. (Ta se uporablja tudi v primerih, ko podatki sicer niso konvolucijsko kodirani, pač pa neidealna prenosna karakteristika komunikacijskega kanala povzroča enak učinek, t.j. medsebojni vpliv med podatkovnimi biti oz. simboli.) Na prenosnih poteh prihaja do napak. Če so napake na prenosni poti statistično neodvisne in se ne pojavljajo v rafalih (burstih), je postopek dekodiranja konvolucijsko kodiranega signala s pomočjo Viterbijevega algoritma optimalen. Rafalom napak, ki pogosto nastopajo v telekomunikacijskih sistemih, se izognemo s premešanjem (scrambling) poslanih bitov v daljšem časovnem intervalu.

Dekodirnik opravi svojo nalogo, če ugotovi, kako so prehajala notranja stanja v kodirniku, ko je kodiral podatke za prenos. Po dogovoru se kodirnik na začetku nahaja v stanju 0 (vsi biti notranjega pomnilnika so nič).

Viterbijev algoritem gradi usmerjen graf vseh možnih poti v mrežnem diagramu iz začetnega stanja 0. Mrežni diagram dobimo, če prehajanje po diagramu stanj raztegnemo po časovni osi. Vsako vozlišče predstavlja notranje stanje ob določenem času, vsaka povezava pa predstavlja možen prehod med notranjimi stanji kodirnika. Vsaki povezavi izračunamo Hammingovo razdaljo med sprejetim simbolom in izhodnim simbolom kodirnika pri izbrani povezavi. (Hammingova razdalja je preprosta aritmetična vsota napak, v binarnem primeru je to kar število napačnih bitov.) Minimalna pot v grafu predstavlja najbolj verjetno prehajanje stanj kodirnika. Dolžina vsake poti pomeni natanko število napačnih bitov.

Viterbijev algoritem uporablja princip dinamičnega programiranja, ki pravi, da si je potrebno za vsako vozlišče v grafu zapomniti le najboljšo pot do njega.

Druge poti v vozlišče se zavrže. Viterbijev algoritem za vsako vozlišče izračuna kumulativno metriko vseh poti, ki vodijo v vozlišče. V vsakem vozlišču ohrani le najboljšo pot (pot z najmanjšo metriko). Na koncu izmed vseh vozlišč poišče vozlišče z minimalno metriko in iz opisa poti v to vozlišče določi izhodni simbol dekodirnika.

Čeprav ima konvolucijski kodirnik končno dolžino (število registrov), bi moral dekodirnik, ki opravlja recipročno funkcijo, za teoretično optimalno dekodiranje v vsakem trenutku računati metriko celotne poti od začetka oddajanja signala. To v praksi seveda ni mogoče, saj mora imeti dekodirnik končno dolžino. Napaka glede na optimalno rešitev postane zanemarljiva, če je dolžina dekodirnika (število upoštevanih vozlišč na poti pri vsakokratnem izračunu) enaka ali večja od približno štirikratne vrednosti števila registrov kodirnika.

4. Uporaba tabel za hitrejše delovanje Viterbijevega algoritma

Delovanje Viterbijevega algoritma pohitrimo z uporabo tabel. Število in vrsta tabel je odvisna od velikosti pomnilnika, ki je na razpolago, in velikosti konvolucijskega kodirnika.

Za začetek smo se odločili, da implementiramo Viterbijev algoritem v C-ju na enem procesorju. V članku /1/ je predstavljen pristop s pomočjo vnaprej izračunanih tabel, ki jih algoritem uporablja za hitrejše računanje. Optimizacija je usmerjena v hitrost izvajanja, velikost porabljenega pomnilnika pa ni pomembna. Velikost tabel eksponentno narašča s številom registrov kodirnika, ki pa je relativno majhno (tipično manjše od 10).

Uporabljene so sledeče tabele:

Tabela vseh možnih prehodov med stanji *MOrder* (Mesh Order):
Prva pomembna tabela je tabela vseh možnih prehodov med stanji za vsak vhodni simbol. Vrednost tabele predstavlja stanja, iz katerih lahko pridemo v dano stanje. S pomočjo te tabele v trenutku dobimo vsa stanja, iz katerih je možen prehod v neko stanje, kar je ključnega pomena pri hitrem iskanju najbolj verjetne poti.

Tabela vnaprej izračunanih vseh možnih Hammingovih razdalj - *Distan*:
To je najpomembnejša tabela, v kateri so shranjene vse možne Hammingove razdalje za vsako pot pri vsakem vhodnem simbolu. Tabela je tri-dimenzionalna. Prva dimenzija pove, za kateri vhodni simbol gre, druga pove, za katero pot po vrsti gre, tretja pa, za katero končno stanje. Branje iz te tabele nadomesti relativno počasno sprotno računanje razdalj, kar pomeni večkratno pohitritev izvajanja dekodirnega algoritma.

Tabeli trenutnih minimalnih razdalj do vozlišč - *OldMetric* in *NewMetric*:
Potrebujemo še dve tabeli, ki vsebujeta trenutno metriko (minimalno Hammingovo razdaljo) do vsakega notranjega stanja.

Tabela optimalnih prehodov med vozlišči - *NewSurviv*:
V tabeli je shranjenih zadnjih 32 poti, ki so preživele za vsako vozlišče. Vsakič pomaknemo bite v tabeli za eno mesto v levo, če preživi druga pot, besedo zapolnimo z desne z 1, sicer pa z 0. Tako se po 32 korakih v njej nahaja 32 bitov, ki po vrsti hranijo informacijo, katera pot je preživela.

Delovanje dekodirnika, ki ga ponazarja diagram na sliki 3, je sledeče:

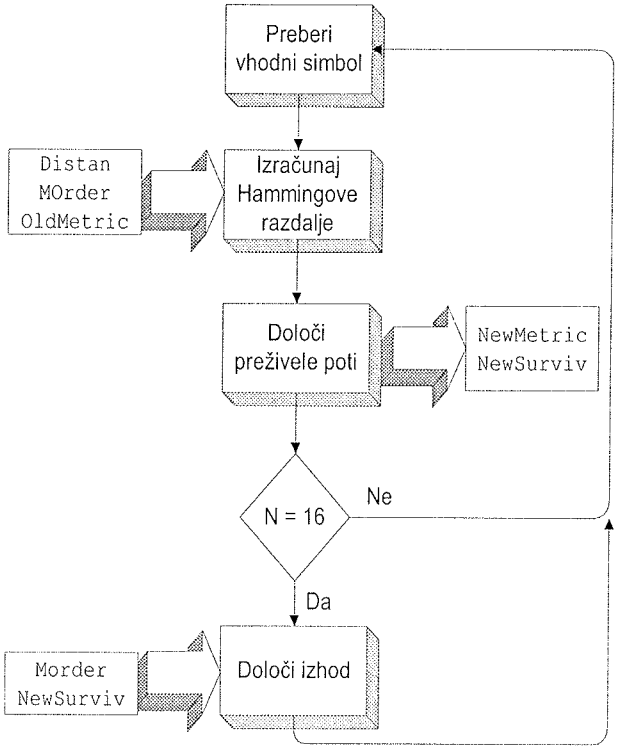
Preberemo vhodni simbol, imenujmo ga *Symb*.

Hammingove razdalje računamo preko vseh notranjih stanj. Na vsakem koraku prideta v vsako vozlišče dve poti, od katerih ena preživi, druga pa odpade. Če računamo stanje *state*, potem nam *MOrder[state][0]* ter *MOrder[state][1]* povesta obe stanji, iz katerih vodi pot v diagramu prehajanja stanj v končno stanje *state*. Imenujmo ju *s0* in *s1*. Stari metriki stanja *OldMetric[s0]* je potrebno prišteti še Hammingovo razdaljo za zadnjo pot *Distan[Symb][0][state]*. Isto ponovimo za drugo pot (*OldMetric[s1]*, pri kateri prištejemo *Distan[Symb][1][state]*).

Preživelo pot in s tem novo metriko do vozlišča *state* dobimo tako, da manjšo vsoto zapišemo v *NewMetric[state]*. Bite v *NewSurviv[state]* vsakič premaknemo za eno mesto v levo, najbolj levi bit odpade, z desne pa se zapiše nova vrednost, ki pove, katera pot je preživela na zadnjem koraku.

Določanje izhoda se izvede na vsakih 16 sprejetih simbolov. Algoritem najprej poišče stanje z minimalno metriko, nato pa s pomočjo tabele *NewSurviv* izračuna najbolj verjetno poslane podatke. Izpisuje se le 16 podatkov, ki so bili poslani pred 32 simbolnimi intervali. To pomeni, da vedno izpisujemo dekodirano zgodovino za 16 do 32 bitov nazaj, kar ustreza 32-bitni besedi procesorja. Celotna potrebna zgodovina preživelih poti za vsako stanje je shranjena v eni besedi tabele *Newsurviv*. Dekodirnik se mora najprej premakniti 16 prehodov stanj v zgodovino brez izpisovanja izhodnega simbola, nato pa še 16 z izpisovanjem.

Koda je napisana splošno za poljubne konvolucijske kodirnike z enim vhodom in dvema izhodoma. Pri povečevanju števila registrov kodirnika se eksponentno veča število notranjih stanj, s tem pa velikosti tabel in čas izvajanja.



Slika 3: Viterbijev dekodirnik

3.2. DSP algoritem prirejen za TMS320C44

Hitrost Viterbijevega algoritma smo testirali na procesorju Texas Instruments TMS320C44. Kot testni primer smo izbrali konvolucijski kodirnik z enim vhodom, dvema izhodoma in petimi zakasnilnimi celicami (2,1,5), ki se uporablja za kanalno kodiranje pri prenosu govora ali podatkov v GSM telekomunikacijskih sistemih. Izhodna hitrost vokoderja uporabljenega v GSM mobilnih telefonih je 13 kbit/s, (blok 260 bitov, dolžina bloka je 20 ms). Blok 260 bitov je razdeljen na 3 dele, od katerih sta konvolucijsko kodirana le prva dva podbloka dolžine 182 bitov. Prvemu podbloku dodamo tri paritetne bite. Zaradi lastnosti konvolucijskega kodirnika je potrebno celotnemu bloku dodati še štiri ničle na koncu prvih dveh podblokov, tako da je podatkovna hitrost na vhodu konvolucijskega kodirnika 9,4 kbit/s za en govorni kanal. Pri prenosu podatkov s prenosno hitrostjo 9,6 kbit/s uporabljamo enak konvolucijski kodirnik, vendar je v tem primeru vhodna podatkovna hitrost v konvolucijski kodirnik 12,2 kb/s.
Pri taktu DSP procesorja 50 MHz in algoritmu napisanem v programskem jeziku C je procesor potreboval za izpis 16 bitov informacije 0,305 ms, kar pomeni podatkovno hitrost 52,4 kbit/s.

3.2.1 Uporaba hitrega notranjega pomnilnika

V naslednjem koraku smo vse tabele prestavili v hitri notranji pomnilnik, ki ga imajo procesorji TMS320C44 na naslovih od 0x002FF800 do 0x002FFFFF. Velikost naslovnega prostora je 8 kB, oziroma 2048 lokacij. Branje in pisanje v notranjem pomnilniku traja samo en strojni cikel, v običajnem pomnilniku pa dva strojna cikla. V tem primeru dobimo hitrost 72,9 kbit/s.

3.2.2 Uporaba naslavljanja »post incremented«

Dostop do tabel preko kazalcev predstavlja naslednjo izboljšavo. Uporabili smo posebno naslavljanje (*post incremented addressing*), pri katerem se pri linearnem prehodu skozi tabelo ne izgubi nič časa za naslavljanje tabele, saj se register (ki predstavlja kazalec na tabelo) ob branju vrednosti, kamor kaže, samodejno poveča za ena. RISC procesorji za opisani ukaz porabijo le en cikel. Vsaka od vrstic v spodnjem primeru sešteje staro metriko stanja in metriko dane poti, hkrati v istem ukazu pa še poveča oba kazalca za 1, tako da zaradi posebne razvrstitve elementov tabel v naslednjem koraku že dostopa do naslednjega stanja in poti:

```
ADDI   *AR2++,*AR7++,R10
ADDI   *AR2++,*AR7++,R9
```

V dveh urinih periodah procesorja dobimo dolžini obeh možnih poti v končno stanje. Podatkovna hitrost se je povečala na 78,3 kbit/s.

3.2.3 Implicitna uporaba vrednosti *MOrder*

Dodatno povečanje hitrosti dosežemo z implicitno uporabo vrednosti v tabeli *MOrder*. Za izbrani konvolucijski kodirnik opazimo, da se vrednosti v tabeli *MOrder* za prvo polovico stanj in prvo ter drugo vhodno vejo v stanje povečujejo linearno s korakom 1. Enako velja za drugi del stanj konvolucijskega kodirnika.

Tab. 1: Tabela MOrder

Pot\stanje	0	1	2	3	4	5	6	7
0	0	2	4	6	8	10	12	14
1	1	3	5	7	9	11	13	15

Pot\stanje	8	9	10	11	12	13	14	15
0	0	2	4	6	8	10	12	14
1	1	3	5	7	9	11	13	15

Lastnost tabele smo izkoristili tako, da vrednosti kazalca nismo brali iz tabele *MOrder*, temveč smo ga linearno povečali. Z opisano izboljšavo smo dosegli za 10% hitrejše delovanje algoritma in s tem podatkovno hitrost 86,7 kbit/s.

3.2.4 Uporaba razporeditve dimenzij v tabeli *Distan*

V tabeli *Distan* je v prvi dimenziji vhodni simbol, tako da so v pomnilniku vse vrednosti za isti vhodni simbol skupaj (32 vrednosti = 16 stanj × 2 poti). Algoritem v enem koraku vse razdalje po poteh računa za isti

vhodni simbol, ki ga sprejme kot parameter. Na začetku algoritem nastavi odmik kazalca na pravo vrednost, nato se samodejno ob vsakem naslavljanju poveča za 1. Zaradi omenjenega načina naslavljanja ne porabimo dodatnega časa. Dosegli smo propustnost 93 kbit/s.

3.3 Mehki Viterbijev dekodirnik

Opisani dekodirniki so trdi dekodirniki, ker obdelujejo le vrednosti 0 in 1. V vseh komunikacijskih sistemih, tudi digitalnih, se podatki v komunikacijskih kanalih prenašajo s pomočjo signalov, ki so analogne veličine. S pretvorbo signala v logične vrednosti (0,1) izgubimo del informacije, ki bi lahko izboljšal postopek dekodiranja. Če ničlo in enico predstavimo z vrednostjo, ki pove, kako blizu je sprejeti podatek logični enici ali ničli, in te vrednosti obdelujemo v dekodirniku, govorimo o mehkem Viterbijevem dekodirniku. Pri 8-bitni A/D pretvorbi sprejme dekodirnik za vsak prenesen simbol vrednost med 0 in 255. Vrednost 0 pomeni čisto ničlo, vrednost 255 pa čisto enico. Dekodirnik (2, 1, *n*) mora sprejeti vrednosti za dva sprejeta bita (t.j. skupno 16 bitov) za vsak dekodirani bit. Računanje razdalj je bolj zamudno, saj bi bilo za enak način uporabe tabel kot pri prejšnjem dekodirniku potrebno imeti zelo velike tabele. Tabela *Distan* bi bila velika kar $16\times2\times2^{16}$ = 2M naslovov, saj je sedaj 2^{16} možnih vhodnih simbolov. Če pa bi npr. uporabili le vrednosti od 0 do 15 (4-bitna A/D pretvorba), potem bi bila velikost 16×2×256 = 8K naslovov.

Pri mehkem Viterbijevem dekodirniku se Hammingova razdalja nadomesti s kvadratom evklidske razdalje (razlike sprejetega simbola in pravilnega simbola) za dano pot. Minimalna vsota po poti pomeni najbolj verjetno oddano zaporedje simbolov, torej je to Viterbijevo dekodiranje na podlagi maksimalne verjetnosti (*MLVD - Maximum likelihood Viterbi Decoding*, /7/, poglavje 5-1-4).

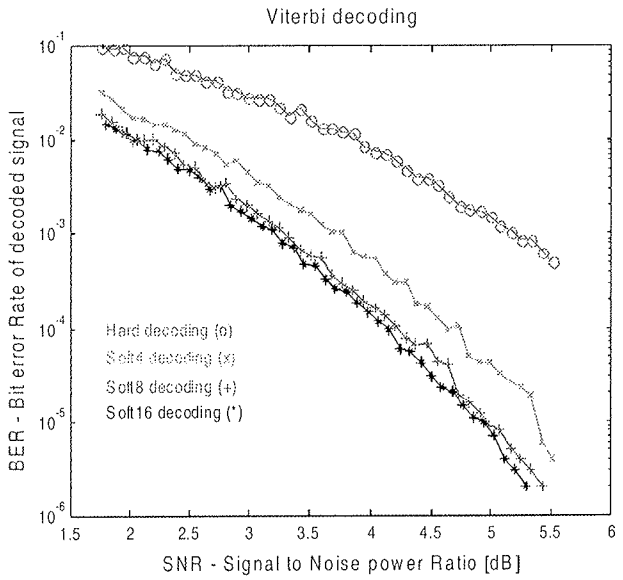
3.3.1 Hitro dekodiranje s 16 nivoji

Napisali smo program za hitro dekodiranje za 16 nivojski dekodirnik (N=16) z uporabo dveh novih tabel *Dist1[SIMBOL][STANJE][POT]* in *Dist2[SIMBOL][STANJE][POT]*. Tabeli imata enako vlogo kot jo je imela tabela *Distan* pri trdem dekodiranju, le da se sedaj prva tabela nanaša na prvi, druga pa na drugi sprejeti simbol. Vsaka ima dimenzijo vhodnega simbola enako 16, tako da sta velikosti 16×16×2 = 512 lokacij. Skupaj zasedeta 1024 lokacij, kar je že polovico hitrega pomnilnika. Ostale tabele so v delu druge polovice hitrega pomnilnika, nekaj pomnilnika pa mora ostati prostega, saj se tam po privzetem nahaja sklad. Funkcija za računanje razdalje je sedaj rahlo spremenjena, saj je potrebno prišteti stari metriki razdalji za oba sprejeta simbola.

Skupaj z vsemi opisanimi časovnimi optimizacijami in uporabo opisanih tabel preko kazalcev in hitrega pomnilnika smo dosegli hitrost mehkega dekodirnika 73 kbit/s, kar je več kot dvakratna pohitritev zaradi uporabe tabel in le 20% slabše od trdega dekodirnika.

3.4 Primerjava odpornosti proti šumu med mehkim in trdim dekodiranjem

Z manjšanjem šuma se razmerje med učinkovitostjo mehkega in trdega dekodiranja močno povečuje. Mehko dekodiranje je veliko učinkovitejše kot trdo dekodiranje.



Slika 4: BER diagram za različne Viterbijeve dekodirnike

Slika 4 prikazuje odvisnost pogostosti napak (BER – *Bit Error Rate*) od razmerja moči signala in šuma (SNR). Štiri krivulje prikazujejo: trdi dekodirnik, mehki dekodirnik s štirimi nivoji, mehki dekodirnik z osmimi nivoji ter mehki dekodirnik s šestnajstimi nivoji. Na sliki se lepo vidi, da je mehki dekodirnik veliko boljši od trdega. Z večanjem razmerja signal/šum pa pride razlika še bolj do izraza. Vsa merjenja so zaradi omejenega časa približna.

Mehki dekodirnik je za več razredov boljši od trdega dekodirnika, razmerje se s kvaliteto prenosnega kanala še povečuje. Vendar pa večanje natančnosti na več kot 16 nivojev (en prenesen bit predstavljen s šestnajst vrednostmi) ne prinese vidne izboljšave h kvaliteti dekodiranja. Tako med rezultati dekodirnika s 16 nivoji in dekodirnika z 256 nivoji pri Gaussovem šumu nismo zasledili razlike. Več nivojev pa pomeni mnogo večje tabele, kar lahko posledično na večini sistemov pomeni tudi počasnejše izvajanje zaradi omejene velikosti najhitrejšega pomnilnika (npr. notranji pomnilnik pri DSP, predpomnilnik pri PC).

Pri mehkem dekodiranju dobimo malo boljše rezultate, če simbolov 0 in 1 pri A/D pretvorbi ne postavimo čisto na meje amplitudnega območja pretvornika. Tako pri 16 nivojskem mehkem dekodirniku dobimo najboljše rezultate, če vrednost 0 predstavimo z vrednostjo 2, vrednost 1 pa z vrednostjo 13. Na ta način ostane šum znotraj amplitudnega območja pretvornika in ohrani svoje lastnosti, t.j. Gaussovo porazdelitev. V nasprotnem primeru bi bil šum na eni ali drugi strani porezan, ne bi bil več Gaussov, in zato tudi dekodiranje ne bi bilo več optimalno.

4. Zaključek

Hitrost trdega dekodiranja na 50 MHz procesorju zadošča za bazno postajo GSM sistema, ki mora delovati z 8-kratno hitrostjo mobilnih GSM terminalov. Za mehko dekodiranje pa je potrebno imeti približno 20% hitrejši procesor. Pri opisanem algoritmu se ne uporablja operacij s plavajočo vejico. Zaradi tega je mogoče uporabiti hitrejše in cenejše celoštevilčne DSP procesorje za izvedbo Viterbijevega algoritma v realnem času. Opisani algoritmi omogočajo realizacijo programljivega radijskega vmesnika, ki se prilagaja zahtevam uporabnika in pogojem prenosa radijskega signala med sprejemnikom in oddajnikom.

5. Literatura

/1/ S. M. Said, K. R. Dimond, »Realtime implementation of the Viterbi decoding algorithem on a high-performance microprocessor«, *Microprocessors and microsystems*, vol. 10, no. 1, january/february 1986

/2/ N. Pavešič, *Informacija in kodi*, Fakulteta za elektrotehniko, Ljubljana, 1997

/3/ Domen Šuligoj, Roman Trobec, Borut Robič, »Vzporedna izvedba Viterbijevega algoritma«, v tisku - *Elektrotehniški vestnik*, Ljubljana

/4/ G. Feygin, P. G. Gulak, P. Chow, »A Multi processor Architecture for Viterbi Decoders with Linear Speedup«, *IEEE Transactions on Signal Processing*, Vol. 41, No. 9, September 1993

/5/ I. N. Bronstein, K. A. Semendjajew, G. Musiol, H. Mühlig, *Matematični priročnik*, Tehniška založba Slovenije, 2. Predelana izdaja, Ljubljana, 1997

/6/ Texas Instruments, *TMS320C4x User's guide*, Digital Signal Processing Products, 1993

/7/ John G. Proakis, *Digital Communications*, McGraw-Hill International Editions, Electrical Engineering Series, 1995

Srečo Plevel
študent 4. letnika FRI
Institut Jožef Stefan, Jamova 39, Ljubljana
tel.: (1) 477-3900, fax.: (1) 251-9385, (1) 426-2102
email: sreco.plevel@campus.fri.uni-lj.si

Dr. Tomaž Javornik
Institut Jožef Stefan, Jamova 39, Ljubljana
tel.: (1) 477-3900, fax.: (1) 251-9385, (1) 426-2102
email: tomaz.javornik@ijs.si

Dr. Igor Ozimek
Institut Jožef Stefan, Jamova 39, Ljubljana
tel.: (1) 477-3900, fax.: (1) 251-9385, (1) 426-2102
email: igor.ozimek@ijs.si

Dr. Roman Trobec
Institut Jožef Stefan, Jamova 39, Ljubljana
tel.: (1) 477-3900, fax.: (1) 251-9385, (1) 426-2102
email: roman.trobec@ijs.si

Prof. Gorazd Kandus
Institut Jožef Stefan, Jamova 39, Ljubljana
tel.: (1) 477-3900, fax.: (1) 251-9385, (1) 426-2102
email: gorazd.kandus@ijs.si

Prispelo (Arrived): 20.10.00 *Sprejeto (Accepted): 22.11.00*

ZASNOVA MOBILNE MULTIMODALNE KOMUNIKACIJSKE NAPRAVE - OSEBNI NAVIGATOR

Bojan Kotnik, Tomaž Rotovnik, Zdravko Kačič, Bogomir Horvat, Boštjan Horvat, Iztok Kramberger
Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko, Maribor, Slovenija

Ključne besede: telekomunikacije, komunikacije podatkovne, komunikacije mobilne, naprave komunikacijske multimodalne, navigatorji osebni, GPS sistem pozicioniranja globalni, GSM sistemi komunikacij mobilnih globalnih, strežniki navigacijski, kartiranje

Izvleček: Osebni navigator predstavlja zasnovo multimodalne komunikacijske naprave, ki uporabniku omogoča vizualno, taktilno in govorno-slušno komunikacijo. Glavni vmesnik med uporabnikom in sistemom je LCD zaslon občutljiv na dotik, ki tako hkrati služi kot vhodna in izhodna naprava. S pomočjo GSM podatkovne komunikacije, je osebni navigator povezan z navigacijskim strežnikom, ki teče na osebnem računalniku in je prav tako sestavni del celotnega sistema osebnega navigatorja. Osebni navigator ima vgrajeno podporo za digitalni signalni procesor, s pomočjo katerega bosta izvedeni avtomatska sinteza in analiza govora. Trenutna zasnova osebnega navigatorja omogoča široko paleto aplikacij, kot so funkcije turističnega vodnika (obveščanje uporabnika kje se nahaja, prikazovanje znamenitosti in vodenje), informatorja (posredovanje dnevnih novic in brskalnik po internetu) in komunikacijske naprave (beležnica, pošiljanje SMS sporočil in elektronske pošte).

The Design of Mobile Multimodal Communication Device - Personal Navigator

Keywords: telecommunications, data communications, mobile communications, multimodal communication devices, personal navigators, GPS, Global Positioning System, GSM, Global System for Mobile communications, navigation servers, mapping

Abstract: In today's world there are increasing demands for products that would enable a palette of services and would be portable, light, small and ergonomically designed. Their handling, however, should be as simple as possible, e.g. voice driven. Demands like that have brought about an integration of different technologies. The concept of the device which we describe in this article is called personal navigator (Pic.1). Personal navigator presents a multimodal communication device. It enables the user visual, tactile and voice communication. Visual and tactile communication between the user and the navigator runs over a graphical LCD display with touch-panel. Speech communication is supported by a powerfull digital signal processor, which makes automatic recognition and the synthesis of speech possible. That way, the personal navigator can be controled by speech. In that case ergonomically designed headphones are connected to the device, together with a combined microphone. The concept alone also includes image transmission, connection to the Internet and positioning - using Global Positioning System. Personal navigator also needs a navigation server (Pic.2), which runs on a PC. Combined GPS/GSM module /3/ is connected to the server. The GPS module in the navigation server is needed for implementation of differential positioning; GSM module serves for establishing data connection with Personal navigator.

The current testing implementation of personal navigator (Pic.3), combines GPS unit for establishing position of the user. It consists of GPS receiver module and a GPS unit processor — microcontroler AT90S2313. The microcontroler communicates with the central processor Atmega103 through a parallel data connection. The processor, which presents the centre of the navigator, is connected to the GSM and LCD modules. They interact through a serial data communication (UART). Both these modules share a joint interface by a special switch unit, which is also controlled by the central processor. The unit with the digital signal processor is connected to the central processor through a serial peripheral interface (SPI). Because of the quantity of the processed data there is also an external RAM connected to the central processor. The navigator is designed modularly, which enables easier and faster upgrading and optimisation. The Personal navigator currently performs the functions of a tourist guide, an information and a communication device. As a tourist guide it informs the user about his position, displays the interesting tourist sights, tourist information and tourist guiding. Besides the above-mentioned, it serves as an information source; it presents the daily news and offers the possibility of browsing the Internet. It's communicational services include notebook, sending SMS messages and electronic mail.

The purpose of "the server - navigator system" is to enable a mobile device (personal navigator) to perform more demanding functions as are available in other mobile devices today. This mobile device however does not need a very powerfull processing unit. The data, which demand more power to process, are dealt with by the server. The server keeps a lot of useful information, needed for operating, in the storage unit. Therefore the navigator requires as much memory, as is needed for performing basic operations (displaying the data received from the server). Smaller RAM and a less capable processing unit consume less energy and for mobile devices that presents a compromise between their weight, size, capability and price. The efficiency of this system will in future increase by upgrading the existing GSM network with the mobile network of the third generation, called UMTS. This network will enable a faster transporting of multimedia data.

1. Uvod

V zadnjem času smo priča skokovitemu porastu sodobnih tehnologij, kot so mobilna telefonija, internet, brezžični podatkovni prenosi, globalni pozicionirni sistem in jezikovne tehnologije. Sprva so se ta področja razvijala samostojno, vsako zase in se kot taka tudi ločeno predstavljala na svetovnem tržišču v obliki naprav, ki so tako bolj ali manj služile le enemu namenu. Satelitski navigacijski sistem GPS /1/, ki je bil prvotno namenjen izključno v vojaške namene ZDA, se je v zadnjih letih nezadržno razširil tudi na civilno

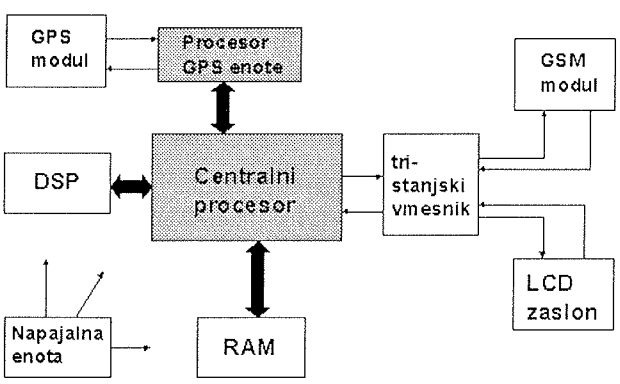
področje. Največji razcvet je doživela digitalna mobilna telefonija GSM /2/. Število uporabnikov GSM telefonije po svetu namreč eksponentno raste. Na področje mobilne telefonije je že pričel prodirati tudi internet v obliki WAP protokola, ki je nekako optimiran na relativno majhno velikost zaslona na mobilnem terminalu.

V sodobnem času se kažejo čedalje večje zahteve po napravah, ki bi omogočale celo paleto najrazličnejših storitev in bi ob vsem tem bile še prenosne, lahke, majhne, čim dalj časa avtonomne, ergonomsko oblikovane, upravljanje z njimi pa bi naj bilo čim bolj enostavno, na primer z govorom. Takšne zahteve so privedle k združevanju oziroma integriranju zgoraj omenjenih tehnologij. V članku predstavljamo zasnovo mobilne komunikacijske naprave, ki omogoča multimodalno komunikacijo z uporabnikom in združuje zgoraj omenjene tehnologije. Napravo smo poimenovali osebni navigator.

2. Koncept osebnega navigatorja

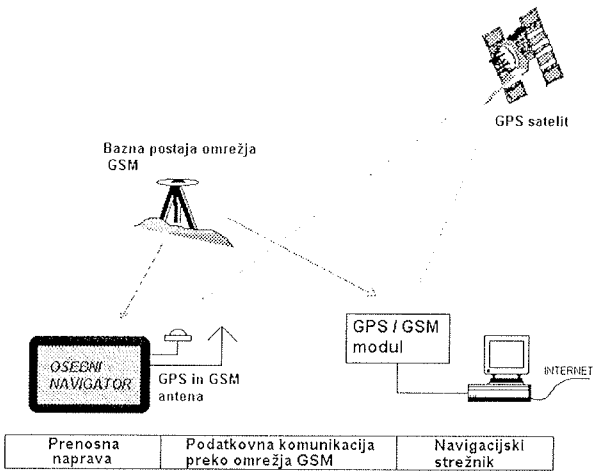
Osebni navigator predstavlja multimodalno komunikacijsko napravo. Z uporabnikom omogoča vizualno, taktilno in govorno-slušno komunikacijo. Sam koncept zajema še prenos slike, povezavo z internetom ter pozicioniranje s pomočjo globalnega pozicionirnega sistema.

Osebni navigator (slika 1) združuje GPS enoto za določanje pozicije uporabnika, GSM enoto za telefoniranje in omogočanje brezžičnega podatkovnega prenosa z navigacijskim strežnikom ter grafični, na dotik občutljiv zaslon iz tekočih kristalov. Za avtonomijo celotne naprave skrbi napajalna enota. Zaradi vse večjega pomena in porasta multimodalnih komunikacij, je za osebni navigator predvidena še enota z zmogljivim digitalnim signalnim procesorjem, ki bo omogočala govorno komunikacijo z navigatorjem. Tako bo mogoče osebni navigator upravljati tudi z govorom. V tem primeru bodo na napravo priključene ergonomsko oblikovane slušalke s kombiniranim mikrofonom. Naprava je zasnovana modularno, saj s tem omogočimo lažje in hitrejše kasnejše nadgrajevanje in optimiranje.



Slika 1: Blokovna shema osebnega navigatorja.

Osebni navigator potrebuje za svoje delovanje tudi navigacijski strežnik (slika 2), ki teče na osebnem računalniku PC. Nanj je priključen kombiniran GPS/GSM modul /3/. GPS modul je v navigacijskem strežniku potreben za implementacijo diferenčnega pozicioniranja, za dosego večje točnosti, GSM modul pa služi za zagotavljanje podatkovne povezave z osebnim navigatorjem.



Slika 2: Blokovna shema celotnega sistema osebnega navigatorja

3. Zgradba in delovanje osebnega navigatorja

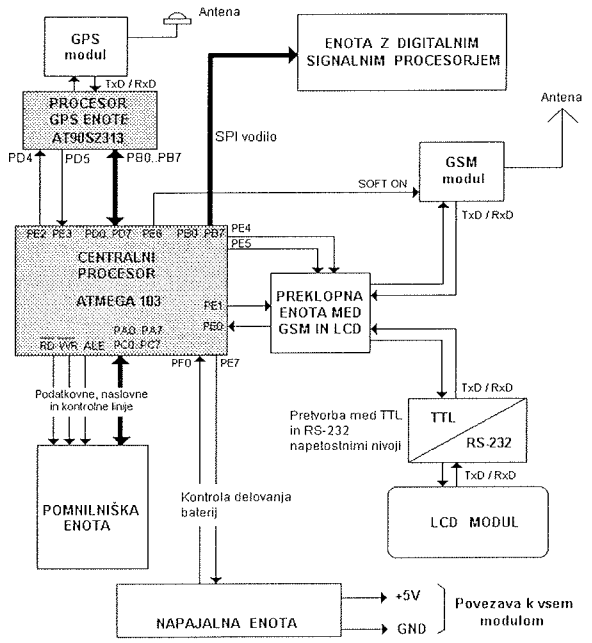
Na sliki 3 je prikazana električna shema osebnega navigatorja. V nadaljevanju bomo opisali naloge in delovanje posameznih sklopov in enot.

3.1 Centralni procesor

Naloga centralnega procesorja je krmiljenje vseh enot osebnega navigatorja. Za centralni procesor smo izbrali 8-bitni mikrokrmilnik ATmega103 ameriške firme Atmel /4/. Za ta mikrokrmilnik smo se odločili zato, ker že imamo preizkušena razvojna orodja zanj, glavni razlog pa so bile njegove dobre tehnične lastnosti:

- RISC arhitektura,
- 121 učinkovitih ukazov, od katerih se večina izvede v samo enem urnem ciklu,
- velika hitrost delovanja do 6 MIPS,
- 128KB notranjega FLASH programskega pomnilnika,
- 4KB notranjega statičnega RAM pomnilnika,
- vgrajen SPI vmesnik, ki bo služil za povezavo z digitalnim signalnim procesorjem,
- vgrajen serijski vmesnik (UART - Universal Asynchronous Receiver Transmitter), ki služi za krmiljenje GSM in LCD modula,
- vgrajen analogno-digitalni pretvornik s pomočjo katerega nadziramo stanje akumulatorskih baterij,

- možnost priključitve zunanjega dodatnega pomnilnika,
- majhna poraba.



Slika 3: Električna shema osebnega navigatorja

ATmega103 je 8 bitni mikrokrmilnik narejen v CMOS tehnologiji z izboljšano RISC (Reduced Instruction Set Computer) arhitekturo, kar mu omogoča, da lahko izvrši en ukaz v enem urnem ciklu. Pozna 12 načinov naslavljanja. Poleg že naštetih lastnosti ima ATmega103 še 32 splošno namenskih delovnih registrov, ki so neposredno povezani z aritmetično logično enoto (ALE). Le-ta lahko ob izvršitvi določenega ukaza hkrati dostopa do dveh registrov v le enem urnem ciklu. Zato je omenjeni mikrokrmilnik lahko tudi do 10 krat hitrejši od podobnega mikrokrmilnika CISC (Complex Instruction Set Computer) arhitekture. Slika 4 prikazuje RISC arhitekturo mikrokrmilnika ATmega103. Centralni procesor izvaja naslednje glavne naloge:

- **Krmiljenje preklopne enote med LCD modulom in GSM modulom**

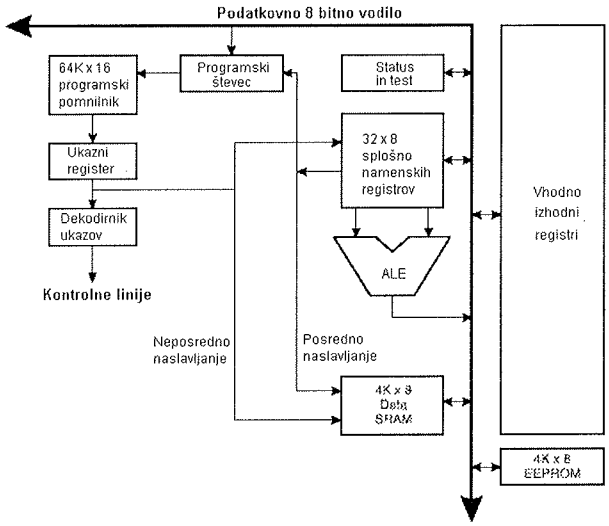
ATmega103 ima en serijski vmesnik UART (Priključka PE0 in PE1), ki si ga delita tako GSM kot LCD modul, zato je potrebna posebna preklopna enota, ki jo bomo opisali pozneje. Ob vklopu osebnega navigatorja se UART mikrokrmilnika najprej poveže z GSM modulom. Sledi vklop in inicializacija GSM modula. Po uspešno opravljeni inicializaciji se UART mikrokrmilnika poveže s serijskim priključkom LCD modula. Mikrokrmilnik sedaj čaka na ukaze poslane z LCD modula in se po potrebi poveže z GSM modulom, da opravi zahtevano nalogo, kot je prenos slike, pošiljanje SMS sporočila,... Po opravljeni nalogi se UART mikrokrmilnika zopet poveže z LCD modulom ter mu posreduje rezultate zahtevane naloge, oziroma pričakuje nov ukaz.

- **Vklop, inicializacija ter krmiljenje GSM modula:**

Vklop GSM modula se izvede tako, da mikrokrmilnik za 3 sekunde postavi pin PE6 na visok nivo (logična enica, oziroma 5V). PE6 je povezan na priključek SOFT ON na GSM modulu. Sledi vpis PIN kode in njena potrditev preko serijskih priključkov UARTa PE1 (TxD, oziroma oddaja podatka) in PE0 (RxD, oziroma sprejem podatka). Glede na zahtevan ukaz z LCD modula, mikrokrmilnik vzpostavi s pomočjo GSM modula podatkovno povezavo z navigacijskim strežnikom, opravi telefonski klic ali pošlje SMS sporočilo.

- **Sprejemanje in izvrševanje ukazov ż LCD modula:**

LCD prikazovalnik ima preko svojega zaslona nalepljeno posebno, na dotik občutljivo folijo s pomočjo katere uporabnik upravlja z osebnim navigatorjem. Ob pritisku na določeno polje, kot je na primer *"Pošlji"*, LCD po serijskem vmesniku pošlje mikrokrmilniku posebno, tako imenovano povratno kodo (Return Code), ki jo mikrokrmilnik dekodira in izvede zahtevano akcijo, ki je v tem primeru pošiljanje SMS sporočila. Mikrokrmilnik iz svojega zunanjega RAM pomnilnika posreduje LCD modulu tudi podatke za izpis na zaslon, kot so tekoča ura, datum, trenutna pozicija, slika zemljevida in podobno.



Slika 4: Arhitektura mikrokrmilnika ATMEGA103

- **Sprejemanje GPS podatkov od GPS enote:**

GPS enota, ki jo sestavljata GPS modul in mikrokrmilnik AT90S2313, neodvisno od centralnega procesorja sprejema in obdeluje podatke o poziciji osebnega navigatorja ter o uri in datumu. Na zahtevo

centralnega procesorja se ti podatki prenesejo v RAM centralnega procesorja in le-ta jih posreduje LCD modulu za izpis na zaslon ali GSM modulu, ki jih po podatkovni povezavi pošlje navigacijskemu strežniku.

- **Upravljanje z zunanjim statičnim RAM pomnilnikom:**

Ker ima centralni procesor premalo že vgrajenega notranjega RAM pomnilnika, smo dodali še večji zunanji delovni pomnilnik.

- **Nadzor stanja akumulatorskih baterij:**

S pomočjo vgrajenega analogno digitalnega pretvornika mikrokrmilnik spremlja stanje napolnjenosti akumulatorskih baterij.

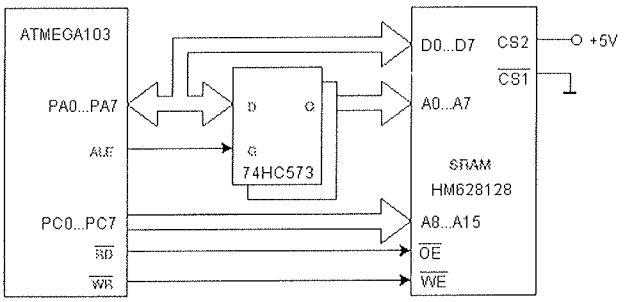
- **Pošiljanje podatkov digitalnemu signalnemu procesorju:**

Centralni procesor posreduje glede na zahtevani ukaz podatke digitalnemu signalnemu procesorju preko posebnega serijskega perifernega vmesnika SPI (Serial Peripheral Interface).

3.2 Pomnilniška enota

Centralni procesor ATmega103 vsebuje 4KB notranjega RAM pomnilnika. Sama aplikacija zahteva večji pomnilnik, zato smo dodali še zunanji RAM pomnilnik velikosti 128KB z dostopnim časom 70ns. Odločili smo se za RAM HM628128, japonske firme Hitachi /5/.

Zaradi varčevanja s številom vhodno izhodnih priključkov, se je proizvajalec mikrokrmilnika ATmega103 odločil za multipleksirano naslovno-podatkovno vodilo. Zato moramo med mikrokrmilnik in pomnilnik priključiti še zadrževalnik (latch) 74HC573, kot je prikazano na sliki 5.



Slika 5: Priključitev zunanjega pomnilnika na mikrokrmilnik Atmega103

3.3 GPS enota

GPS enota je sestavljena iz GPS modula in mikrokrmilnika AT90S2313. V osebnem navigatorju predstavlja glavni del GPS enote GPS sprejemni modul GN-74, ki je produkt ameriške firme Furuno Electric Co., Ltd. /6/. Gre za 8 kanalni GPS sprejemnik, to pomeni, da lahko sočasno spremlja in sprejema signale z osmih satelitov ter z njimi določa svojo pozicijo. Njegova deklarirana točnost znaša 26 metrov. Zaradi svoje relativno majhne porabe (5V, 130mA), male velikosti in teže je zelo primeren za baterijsko napajane prenosne naprave. Slika 8a prikazuje način priključitve modula.

Modul ima tri priključke za priključitev napajalne napetosti +5V. V_{CC} je glavno napajanje, V_{ANT} je priključek za dovajanje napetosti aktivni anteni, V_{RTC} pa je priključek za napajane vgrajene ure realnega časa, zato je ta priključek preko diode D2 povezan na rezervno baterijsko napajanje V_{BAT}. V samem sprejemniku se nahaja tudi statični RAM, v katerem je zapisan almanah (podatki o tem, kje se posamezni sateliti nahajajo). Zelo pomembno je, da imamo napajanje V_{RTC} vedno prisotno, ker se tako vsebina pomnilnika lahko ohrani do naslednjega vklopa sprejemnika in sprejemniku potem ni potrebno začeti iskati novega almanaha, kar je lahko zelo zamuden postopek. Omenimo naj, da dokler sprejemnik nima almanaha ne more pričeti s pozicioniranjem. Aktivno anteno priključimo preko koaksialnega kabla na BNC konektor. Ojačane signale z antene vodimo do sprejemnika po koaksialnem kablu, ki hkrati služi še za dovod napajanja aktivni anteni. Komunikacija z modulom poteka po standardnem serijskem protokolu NMEA-0183 /7/ z bitno hitrostjo 4800bps. Slika 6 prikazuje opis protokola NMEA-0183.

Protokol	Tip	BPS	Podatkovna dolžina	Stop bit	Pariteta
NMEA-0183	ASCII	4800	8	1	NE

Slika 6: Opis protokola NMEA-0183

NMEA-0183 (National Marine Electronic Association) je mednarodni standard za medsebojno povezovanje navtičnih elektronskih naprav.
Asinhrona serijska komunikacija po linijah RD (Receive Data) in TD (Transmit Data) uporablja polnodupleksni način, kar pomeni, da lahko modul sočasno oddaja podatke in sprejema nove ukaze iz krmilnega procesorja AT90S2313. Mikrokrmilnik komunicira z GPS modulom v obliki ukazov in odzivov, ki se v NMEA terminologiji imenujejo *vhodni* in *izhodni stavki* (input, output sentences). Le-ti imajo točno predpisano obliko, ki je prikazana na sliki 7.

$<naslovno polje>	,<podatkovno polje>	[*<nadzorna vsota>]	<CR><LF>

Slika 7: Osnovna zgradba vhodnih in izhodnih stavkov

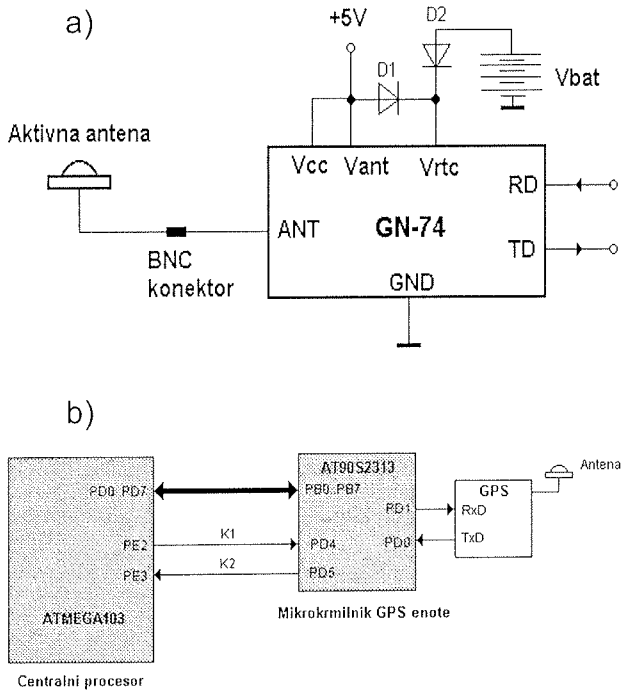
Modul GN-74 je z mikrokrmilnikom AT90S2313 povezan preko vgrajenega serijskega vmesnika UART. Mikrokrmilnik AT90S2313 /8/ ima zelo podobno arhitekturo kot že opisani centralni procesor ATmega103, s tem da razpolaga z manj programskega (2KB) in delovnega (128B) pomnilnika. Ima tudi samo 15 vhodno izhodnih priključkov. Mikrokrmilnik AT90S2313 opravlja naslednji dve nalogi:

- **Upravljanje z GPS modulom GN-74:**

Mikrokrmilnik ob vklopu napajanja najprej inicializira GPS modul tako, da mu poda podatek o približni lokaciji, uri in datumu. S temi podatki prične GPS modul mnogo hitreje samostojno iskati svojo pozicijo. Ko GPS modul prične s pozicioniranjem in oddajanjem stavkov $GPGGA in $GPZDA pa mikrokrmilnik dekodira njuno vsebino ter si podatke o poziciji, uri, datumu in statusu zapisuje v svoj notranji RAM. Osveževanje vsebine pomnilnika se izvrši vsako sekundo, saj je to tudi interval v katerem GPS modul oddaja izhodna stavka.

- **Pošiljanje GPS podatkov centralnemu procesorju:**

Ko pride od centralnega procesorja zahteva po GPS podatkih, mikrokrmilnik AT90S2313 prekine z dekodiranjem stavkov, ter vsebino svojega pomnilnika posreduje centralnemu procesorju. Uporabili smo paralelno podatkovno povezavo z osmimi podatkovnimi in dvema kontrolnima linijama, tako kot je prikazano na sliki 8b.



Slika 8 a): Priključitev modula GN-74
b): Podatkovna povezava med centralnim procesorjem in GPS enoto

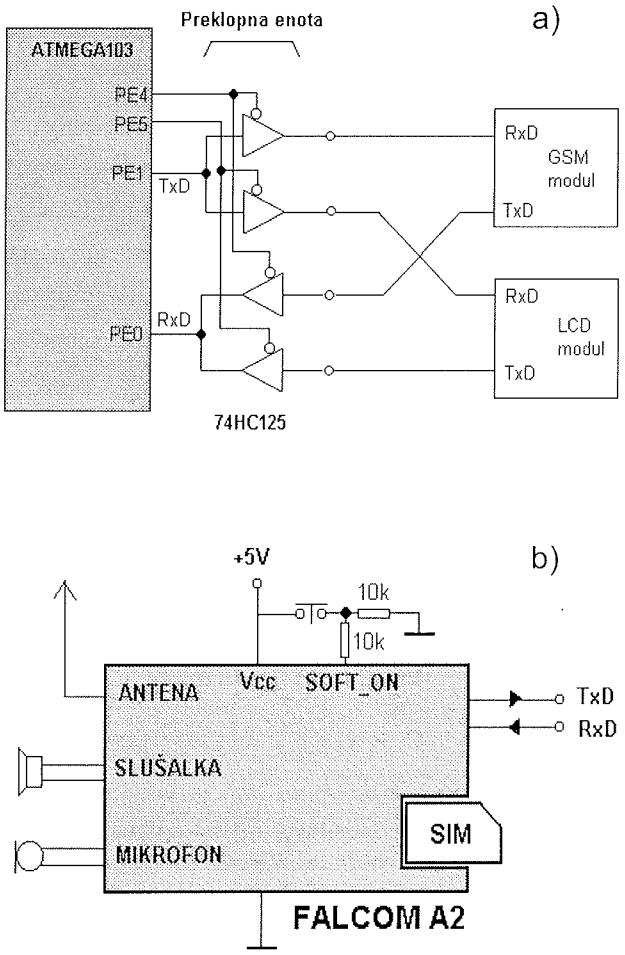
3.4 Preklopna enota med GSM in LCD modulom

Kot smo že omenili, ima centralni procesor le en serijski vmesnik UART, zato ga moramo deliti med GSM in LCD modulom. Izbiro, komu nameniti UART, opravlja program v centralnem procesorju. V stanju mirovanja je UART centralnega procesorja povezan z LCD modulom, tako da lahko procesor sprejema ukaze, ki jih uporabnik osebnega navigatorja pošilja preko zaslona, občutljivega na dotik. Ko se pojavi potreba po delu z GSM modulom, pa se UART centralnega procesorja s pomočjo preklopne enote poveže na GSM modul in izvrši zahtevano opravilo. Nato spet sledi preklop nazaj na LCD modul. Slika 9a prikazuje izvedbo preklopne enote s pomočjo štirih tristanjskih vmesnikov. S slike lahko tudi vidimo, da predstavlja priključek centralnega procesorja PE4 (aktiven nizek nivo) omogočitveni signal za GSM modul, PE5 (aktiven nizek nivo) pa omogočitveni signal LCD modula. Programsko je obvezno potrebno poskrbeti, da izhoda PE4 in PE5 nikoli ne bosta hkrati aktivna.

3.5 GSM enota

V osebnem navigatorju predstavlja GSM modul A2, nemške firme Falcom GmbH, glavni del GSM enote, katere naloga je vzpostavljanje podatkovne povezave z navigacijskim strežnikom, ki prav tako vsebuje podoben GSM modul. Sicer pa je s pomočjo tega modula možno tudi telefoniranje in pošiljanje kratkih SMS sporočil. S stališča uporabe, se obnaša GSM modul A2 v bistvu kot modem s klasičnim, tako imenovanim Hayesovim naborom *AT* ukazov, ki jih dandanes poznajo vsi modemi. Na tem mestu pa velja opozoriti, da GSM modul ni modem v pravem pomenu besede. Izraz *modem* namreč pomeni MOdulator-DEModulator, saj je njegova naloga modulirati digitalne podatke na analogni nosilni signal in potem na drugi strani prenosne PSTN telefonske linije demodulirati analogni signal tako, da spet dobimo prvotni, odposlani digitalni signal. Glede na to, da je omrežje GSM že v svoji osnovi digitalno, se podatki kot taki tudi prenašajo digitalno, torej ni med dvema komunicirajočima GSM moduloma nobenega procesa modulacije oziroma demodulacije več. Zato tudi ni smiselno uporabljati izraza *modem*. Modul A2 je prirejen za priključitev na omrežje GSM 900 v asinhronem transparentnem ali netransparentnem načinu s standardnimi bitnimi hitrostmi od 300bps do 9600bps. Kompatibilen je s standardi CCITT V.21, V.22, V.22BIS, V.32 in V110.

Na sliki 9b je prikazan način priključitve modula FALCOM A2. Priključek SOFT ON je z dvema uporoma 10kΩ priključen na maso ter preko tipke na napajalno napetost 5V. Če tipko pritisnemo za 3 sekunde, modul vklopimo. Tipka je na sliki prikazana le simbolično. V osebnem navigatorju izvrši vklop modula centralni procesor. Če želimo modul uporabljati kot telefon, nanj priključimo še slušalko in mikrofon, medtem ko ju za podatkovno povezavo ne potrebujemo.



Slika 9 a): Izvedba preklopne enote med GSM in LCD modulom
b): Priključitev GSM modula FALCOM A2

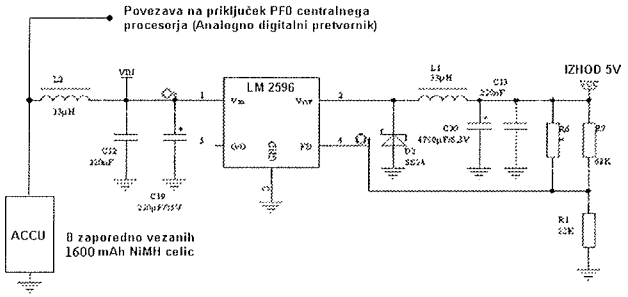
LCD modul EA KIT240-7CTP predstavlja vmesnik med uporabnikom in sistemom osebnega navigatorja, saj preko njega uporabnik zahteva izvršitev določenih akcij, kot sta na primer telefoniranje in pošiljanje SMS sporočil. Po drugi strani pa na LCD modulu sistem posreduje uporabniku željene informacije, kot so izris zemljevida, izpis željenih internetnih strani in podobno. Velikost uporabljenega monokromatskega modro-belega LCD zaslona znaša 240 x 128 slikovnih pik ali pikslov. Ima lastno osvetlitev, kar nam omogoča, da lahko osebni navigator uporabljamo tudi v popolni temi. Že sam modul vsebuje 64KB pomnilnika, in ima zmožnost, da vanj zapišemo praktično celotni zunanji vmesnik z vsemi menuji vred. Celotni zaslon je prekrit s 60 na dotik občutljivimi celicami, kar zadostuje tudi za implementacijo običajne tipkovnice. Glavna prednost na dotik občutljivega zaslona je ta, da tako na sami napravi ne potrebujemo več nobenih drugih tipk. S tem se velikost prednje plošče osebnega navigatorja omeji le na velikost zaslona.

3.7 Enota z digitalnim signalnim procesorjem

V osebnem navigatorju smo predvideli tudi podporo za digitalni signalni procesor (DSP), ki bo lahko opravljal avtomatsko sintezo in analizo govora. Pri avtomatski sintezi govora bo generiral govor iz posebej opremljenega teksta, ki se bo preko GSM podatkovne povezave prenesel iz navigacijskega strežnika. S pomočjo avtomatske analize govora bo mogoče osebni navigator upravljati tudi z govorom. Predvideli smo podporo za DSP s celoštevilsko aritmetiko serije C5000, ameriške firme Texas Instruments. DSP bo s centralnim procesorjem komuniciral preko serijskega perifernega vmesnika SPI. Trenutno je enota v fazi razvoja.

3.8 Napajalna enota

Napajalna enota skrbi za energetsko oskrbo osebnega navigatorja. Ocenjena tokovna poraba celotnega vezja znaša pri napetosti 5V v povprečju 800mA, zato je zelo pomembna izbira načina stabiliziranja napetosti, saj mora napetostni stabilizator delovati s čim večjim izkoristkom. Upoštevaje zahteve, je najprimernejši DC/DC stikalni stabilizator (switcher). Stikalni stabilizator je možno izvesti z MOS FET tranzistorjem, vendar smo se raje odločili za integrirano izvedbo stabilizatorja LM2596 firme National /9/, saj ima možnost povratne vezave z vgrajeno notranjo regulacijo. Poleg tega je taka izbira tudi cenovno dokaj ugodna. Slika 10 prikazuje celotno električno shemo napajalne enote. Za napajanje služi 8 zaporedno povezanih NiMH akumulatorskih celic. Stanje njihove napolnjenosti kontrolira preko priključka PF0 centralni procesor s pomočjo internega analogno digitalnega pretvornika.



Slika 10: Električna shema napajalne enote

4. Podatkovna povezava z navigacijskim strežnikom

Za komunikacijo osebnega navigatorja s strežnikom smo uporabili ISO referenčni model za povezovanje odprtih sistemov (OSI) /10/. Komunikacija poteka med dvema GSM enotama. Osebni navigator ima vlogo stranke, strežnik pa se odziva na njegove zahteve. Uporabili smo tri sloje:

- **Fizični sloj**

Fizični sloj vsebuje mehanične, električno-proceduralne vmesnike in prenosno sredstvo. Do sloja podatkovne povezave je uporabljen serijski prenos podatkov.

- **Sloj podatkovne povezave**

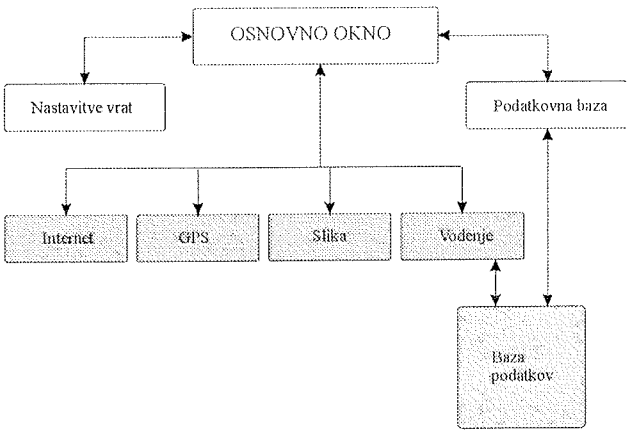
Sloj podatkovne povezave uporablja storitve fizičnega sloja. Ta sprejema golo bitno zaporedje in ga skuša prenesti na sprejemno stran. Ukvarja se z uokvirjanjem, nadzorom napak, krmiljenjem pretoka in dodeljevanjem dostopa. Zaradi uporabljenega protokola za oddajo z mirovanjem v drugem sloju ni potrebno krmiljenje pretoka, ker se naslednji okvir pošlje šele po potrditvi predhodnega. Na sliki 11 je prikazana osnovna sestava podatkovnega okvirja.

VELIKOST	1 ZLOG	1 ZLOG	4 ZLOGI	N ZLOGOV	2 ZLOGA
IME POLJA	ZAČETNI ZNAK	STATUS	DOLŽINA	PODATKI	CRC-16
KODIRANJE	NEKODIRANO		KODIRANO		

Slika 11: Sestava podatkovnega okvirja

- **Aplikacijski sloj**

Aplikacijski sloj predstavlja osnovni uporabniški vmesnik strežnika (Slika 12).



Slika 12: Funkcionalna shema strežnika

5. Funkcionalni opis osebnega navigatorja

Opisana zasnova strojne opreme osebnega navigatorja omogoča široko paleto aplikacij. Trenutna testna implementacija zajema funkcije turističnega vodnika, informatorja ter komunikacijske naprave. V ta namen strežnik vsebuje podatke, ki omogočajo vodenje in posredovanje informacij za središče mesta Maribor. V nadaljevanju opisana funkcionalnost pa je splošno uporabna in ni vezana na trenutno izvedbo. Pri tem seveda mora strežnik vsebovati ustrezne baza podatkov. Storitve, ki jih navigator opravlja kot turistični vodnik so:

- **Obveščanje uporabnika kje se nahaja.** Na zaslonu se prikaže ime ulice ali ime znanega objekta v neposredni bližini, poleg tega pa še celotna zgodovina in sedanji namen objekta ter vse ostale objekte, ki so povezani z opisanim. Na zaslonu se izpiše tudi zemljepisna širina ter dolžina.
- **Prikazovanje znamenitosti.** Na zaslonu se izpišejo vse pomembnejše znamenitosti v središču mesta. Vsaka znamenitost ima še kratek opis, ki se izpiše po želji. Za vsako izbrano znamenitost je možno vodenje od trenutne lokacije do lokacije, kjer se nahaja znamenitost. Na zaslonu se izriše zemljevid, kjer je ta pot označena. Hkrati se izračuna dolžina poti in potreben čas za pot, podatke o hitrosti potovanja ter podatek o preostalem času potrebnem do izbranega objekta pa uporabnik dobiva tudi med samim potovanjem.
- **Turistična pot.** Izriše se načrt turistične poti. Uporabnik lahko poveča ali zmanjša obseg poti (večje ali manjše število znamenitosti), odvisno od razpoložljivega časa.
- **Informacije in vodenje** do nakupovalnih središč, avtobusne in železniške postaje, do posameznega naslova v mestu...

Storitve osebnega navigatorja kot informatorja zajemajo poleg prej naštetih storitev še naslednje:

- **Posredovanje dnevnih novic.** Uporabniku se na zaslon izpišejo naslovi trenutno najbolj svežih novic, ter povezava za podrobnejši opis, ki ga navigacijski strežnik dobi z interneta.
- **Brskalnik po internetu.** Uporabnik lahko izbere naslove, ki vodijo do posameznih spletnih povezav. Lahko tudi sam vnese naslov izbrane spletne strani, kar da uporabniku neomejene možnosti pregledovanja.

Kot komunikacijska naprava pa osebni navigator zagotavlja naslednje:

- **Beležnica.** Možen je vnos lastnih zapiskov preko na dotik občutljivega zaslona (tipkovnica), ki se lahko preko elektronske pošte prenesejo na uporabnikov naslov ali pa jih uporabnik shrani, če jih bo potreboval za kasnejša opravila.
- **Pošiljanje SMS sporočil.**
- **Pošiljanje elektronske pošte.** Je zelo podobno pošiljanju SMS sporočil.

V trenutni testni izvedbi se ob vklopu naprave na LCD zaslonu najprej pojavi začetni pozdravni menu in sporočilo o inicializaciji. Po uspešno opravljeni inicializaciji, ki traja približno 30 sekund (inicializacija GPS modula ter vpis PIN kode in ostalih parametrov v GSM modul), se na zaslonu prikaže glavni izbirni menu, kot ga prikazuje slika 13. Glavni menu vsebuje v desnem zgornjem kotu podatke o tekoči uri in datumu, v njegovem srednjem delu pa so štiri velika, na dotik občutljiva izbirna polja:

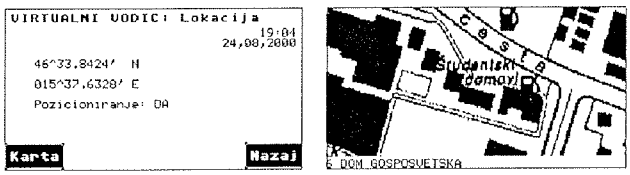
- Trenutna lokacija,
- Telefon,
- WWW,
- Pomoč.

Željeno možnost izberemo s pritiskom na izbrano polje.

Slika 13: Glavni izbirni menu osebnega navigatorja

- **Trenutna lokacija**

V tem podmenuju, ki ga prikazuje slika 14a, ima uporabnik na voljo informacijo o tekoči uri in datumu ter o svoji poziciji. Pozicija je podana v dveh koordinatah, prva je zemljepisna širina (severno, južno) druga pa zemljepisna dolžina (vzhodno, zahodno). V zadnji vrstici je prikazan status, ki pove ali GPS pozicioniranje v sprejemniku že teče. S pritiskom na polje "*Karta*" se podatki o uporabnikovi poziciji prenesejo v navigacijski strežnik. Strežnik na osnovi teh podatkov posreduje osebnemu navigatorju sliko zemljevida. Hkrati iz baze podatkov določi objekte, ki se nahajajo v neposredni bližini in ulico v kateri se osebni navigator trenutno nahaja (slika 14b). Na uporabnikovem zaslonu se tako prikaže zemljevid (del načrta mesta) z označeno uporabnikovo trenutno pozicijo (utripajoč znak).

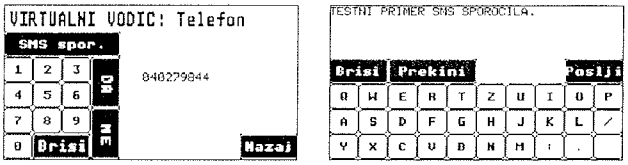


a) b)

Slika 14 a): Podmenu "Trenutna lokacija"
b): Prikaz izrisa zemljevida in izpisa ulice

- **Telefon**

Podmenu telefon, prikazuje ga slika 15a, omogoča uporabniku telefoniranje in pošiljanje SMS sporočil. Če želimo telefonirati, enostavno odtipkamo željeno telefonsko številko skupaj z omrežno skupino, ter pritisnemo "*DA*". Klic prekinemo s pritiskom na "*NE*". Zadnja izbrana telefonska številka vedno ostane v pomnilniku, četudi podmenu zapustimo in jo lahko ponovno pokličemo, če pritisnemo "*DA*".



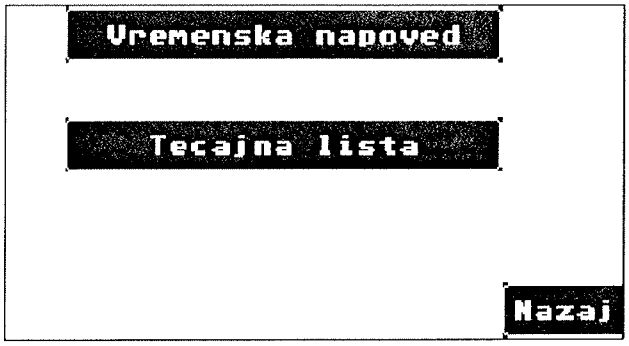
a) b)

Slika 15 a): Podmenu telefon
b): Podmenu za pisanje SMS sporočil

Ko uporabnik želi poslati SMS sporočilo, mora najprej vpisati telefonsko številko prejemnika, nato pa pritisniti polje "*SMS spor.*". Odpre se nov podmenu, ki je prikazan na sliki 15b. V tem menuju pišemo SMS sporočilo s pomočjo klasične, za naš prikazovalnik nekoliko prirejene, "QWERTZ" tipkovnice. Znak zbrišemo s pritiskom na polje "*Brisi*", SMS odpošljemo s "*Poslji*", pisanje SMS sporočila pa prekinemo s "*Prekini*".

- **WWW**

Ta podmenu uporabniku omogoča dostop do podatkov na nekaterih spletnih straneh na internetu. Takoj po izbiri tega podmenuja v glavnem izbirnem menuju, navigator pokliče navigacijski strežnik. Le-ta obdela vsebino izbranih www strani in njihovo vsebino posreduje osebnemu navigatorju. Ko so podatki prenešeni, lahko uporabnik pregleda željene informacije (slika 16).



Slika 16: Podmenu WWW, podatki so na razpolago

Kot je prikazano na sliki 17a, ima trenutno uporabnik na voljo tekoče vremenske podatke in tekočo tečajno listo Banke Slovenije (slika 17b).

a) b)

Slika 17 a): Prikaz vremenske napovedi
b): Prikaz tečajne liste

Če se zapis nahaja na več straneh, uporabimo gumb ">>". Informacija se krožno ponavlja. S pomočjo gumba "*Nazaj*" se vrnemo v podmenu WWW. Od tam lahko izberemo pregled drugih podatkov, ki so na voljo, lahko pa se vrnemo v glavni izbirni menu. Kadarkoli v glavnem izbirnem menuju izberemo WWW, se vedno zahteva prenos svežih podatkov z navigacijskega strežnika.

- **Pomoč**

Ta podmenu nudi uporabniku razne nasvete ter izčrpna navodila kako uporabljati osebni navigator. V določenih primerih lahko uporabnik zahteva še dodatna navodila.

6. Zaključek

Namen sistema strežnik - osebni navigator je omogočiti mobilni napravi (osebnemu navigatorju) opravljanje več zahtevnejših funkcij, kot jih trenutno opravljajo današnje mobilne naprave. Pri tem mobilna naprava ne potrebuje zelo zmogljive procesne enote. Podatke, pri katerih je za obdelavo potrebno zelo veliko procesne moči, posreduje strežnik. Strežnik hrani v pomnilniku veliko uporabnih podatkov, ki jih potrebuje pri opravljanju že prej imenovanih nalog. Torej potrebuje navigator toliko pomnilnika, kot ga je potrebno za opravljanje osnovnih operacij (prikaza podatkov posredovanih od strežnika). Manjši pomnilnik in manj zmogljiva procesna enota porabita manj energije, kar predstavlja pri mobilnih napravah kompromis med težo, velikostjo, zmogljivostjo in ceno. V nadaljnjem razvoju osebnega navigatorja želimo razširiti njegove funkcije ter uporabnost z dodajanjem govorne vhodno/izhodne enote podprte z DSP (Digital Signal Processing) procesorjem. Tako bomo omogočili govorno posredovanje informacije z možnostjo razpoznavanja in sinteze govora. Z uporabo tehnologije za sintezo in analizo govora se bo navigator približal multimodalni komunikacijski napravi, saj bo za komunikacijo s človekom uporabljal več vhodno/izhodnih kanalov, kar bo povečalo hitrost in učinkovitost komunikacije med osebnim navigatorjem in uporabnikom. Tako bosta na primer SMS sporočilo ali elektronska pošta hitreje sestavljena in poslana.

Uporabljeno GSM omrežje je omejeno na relativno počasni podatkovni prenos s hitrostjo 9600bps, kar je ena njegovih največjih omejitev. Prav tako je s tem omejen tudi danes implementirani WAP protokol. Slabost protokola je posredovanje le tekstovnih informacij z omejenimi grafičnimi zmožnostmi. Učinkovitost samega sistema bomo v prihodnosti povečali z načrtovano nadomestitvijo obstoječega GSM omrežja druge generacije z mobilnim omrežjem tretje generacije (UMTS). Storitev GPRS in omrežje UMTS bosta omogočali hitrejši prenos multimedijskih podatkov. Skupaj s standardom IPv6 bo prišlo do velikega porasta tako imenovanega mobilnega interneta (mobile IP). To napoveduje tudi Nokia, eno izmed vodilnih svetovnih podjetij za mobilne telekomunikacijske sisteme, s svojo novo tehnologijo IP-RAN (IP-Radio Access Network). /12/

V članku podana zasnova navigatorja predstavlja analizo možnosti uporabe multimodalne komunikacijska naprave za učinkovito povezavo uporabnika s sodobnim komunikacijskim omrežjem in možnostjo nudenja množice obstoječih ter tudi povsem novih storitev, ki jih omogoča združevanje sodobnih tehnologij.

7. Literatura

/1/ T. Logsdon: *The Navstar global positioning system*, Van Nostrand Reinhold, New York, 1992.
/2/ J. Eberspaecher, H. J. Voegel: *GSM Switching, Services and Protocols*, Willey, Chichester, 1997.
/3/ Funkanlagen Leipoldt OHG: *Falcom A2 User Manual and Command List*, Langewiesen, 1999.
/4/ Atmel Corporation: *ATmega103 Datasheet*, San Jose, 2000.
/5/ Hitachi LTD: *HM628128DI Series*, Tokyo,1998.
/6/ Furuno Electric Co., LTD: *GPS Receiver Model GN-74 Technical Information*, 1995.
/7/ Furuno Electric Co., LTD: *GPS Receiver Model GN-74 Protocol Manual*, 1995.
/8/ Atmel Corporation: *AT90S2313 Datasheet*, San Jose, 2000.
/9/ National Semiconductor, *LM2596 Switcher*, 1997
/10/ A. S. Tanenbaum, *Computer Networks*, Prentice Hall, 1996.
/11/ Z. Kačič, *Komunikacija človek - stroj*, FERI, Maribor, 1995
/12/ www.nokia.com

Bojan Kotnik, univ. dipl. inž. el.,
Tomaž Rotovnik, univ. dipl. inž. el.,
izr. prof. dr. Zdravko Kačič,
red. prof. dr. Bogomir Horvat
Boštjan Horvat
Iztok Kramberger, univ. dipl. Inž. el.

Inštitut za elektroniko,
Fakulteta za elektrotehniko, računalništvo in informatiko, Maribor, Slovenija

Institute of Electronics,
Faculty of Electrical Engineering and Computer Science,
Smetanova 17, 2000 Maribor, Slovenia

Tel. +386 02 220 7000
Fax. +386 02 251 1178
E-mail: dsplab@uni-mb.si

Prispelo (Arrived): 20.11.00 *Sprejeto (Accepted): 07.03.01*

RAZVOJ RAZŠIRITVENE KARTICE S SIGNALNIM PROCESORJEM ZA PCI VODILO

Daniel Čeh-Ambruš, Iztok Kramberger, Zdravko Kačič
Fakulteta za elektrotehniko, računalništvo in informatiko
Univerza v Mariboru, Slovenija

Ključne besede: računalništvo, DSP obdelava signalov digitalna, kartice računalniške, kartice računalniške razširitvene, PCI povezava komponent perifernih, PCI vodila, PC računalniki osebni, flash pomnilniki, TMS320C31 Texas Instruments DSP procesorji

Izvleček: Članek obravnava razvoj in programiranje PCI razširitvene kartice za digitalno procesiranje signalov, zgrajene z DSP procesorjem s plavajočo vejico Texas Instruments TMS320C31. Razširitvena kartica predstavlja ustrezno razvojno okolje za razvoj zahtevnejših aplikacij, ki zahtevajo procesiranje velike količine podatkov in njihov prenos med pomnilnikom razširitvene kartice in gostiteljskega računalnika in za katere so vodila obstoječih sistemov prepočasna. Obstoječi sistemi s tem procesorjem na PCI vodilu so zelo redki in ne izpolnjujejo postavljenih zahtev. Razvit sistem je zasnovan tako, da v največji možni meri izkorišča zmožnosti signalnega procesorja in lahko brez razširjanja pomnilniških kapacitet na njem realiziramo aplikacije kot so spektralni analizator za nizkofrekvenčne signale (z analognim vmesnikom na sistemu do frekvenc okoli 20kHz), 32-kanalni logični analizator z vzorčenjem do 25MHz (pri uri procesorja 50MHz), generator 32-bitnih digitalnih signalov do 12.5MHz, FIR filter z okoli 500 koeficienti za signale vzorčene s frekvenco 48kHz, procesor zvočnih efektov in vrsto drugih. Sistem omogoča povezavo z gostiteljskim osebnim računalnikom preko PCI vodila in deluje v okolju Windows 95/98/Me/NT/2000 operacijskega sistema. Kartica vsebuje 1MB statičnega pomnilnika in 512kB Flash pomnilnika za zagon sistema, PCI vodilo za priključitev na gostiteljski računalnik, 20-bitni dvokanalni analogni vhodno/izhodni vmesnik s frekvenco vzorčenja do 48kHz, razširitveni vmesnik za priključitev uporabniških modulov ter programsko opremo, ki je potrebna za inicializacijo sistema, programiranje signalnega procesorja in komunikacijo s kartico. Programska oprema deluje v sistemu Windows 95/98/Me/NT/2000.

Development of Expansion Card with DSP Processor for PCI bus

Keywords: computer science, DSP, Digital Signal Processing, computer cards, computer expansion cards, PCI, Peripheral Component Interconnect, PCI buses, PC, Personal Computer, PCS, Personal ComputerS, flash memories, TMS320C31 Texas Instruments DSP processors

Abstract: This article treats the design and programming of a PCI expansion card, build around the Texas Instruments TMS320C31 floating point DSP. We encountered the need for a development system, which could be used for development of advanced applications in signal processing. The existing development kits usually use serial or parallel interface to the host PC, which are too slow for advanced signal processing applications. There are also some expansion cards using the ISA interface and very few PCI cards, but none of them suits the defined needs. The expansion card consist of several units: The first is the memory unit, which consist of 1MB static memory implemented in the form of two memory banks of 128k x 32bit. The schematic of the first memory bank is shown in figure 2. A 512kB Flash memory for booting and initialization is in-system programmable. The second unit is the PCI local bus interface for connecting to the host PC, which is realized with the PLX PCI9050 PCI bridge /9/. The PCI bridge provides low cost connectivity for PCI slave designs. It is specifically targeted at easing the transition of existing ISA designs to the more feature rich and performance oriented PCI bus. The PCI 9050 provides direct slave PCI functions by interfacing the adapter's I/O circuitry (control, address and data lines) to a host computer's microprocessor/memory architecture via the 32-bit PCI bus, which typically runs at 33 MHz. The behavior of the bridge is controled via the local configuration registers printed in table 1. The third unit is the two channel analog input/output interface, which consists of a 20-bit stereo audio CODEC from Crystal (CS4222 /8/), supporting sampling rates up to 48kHz, the corresponding input and output buffers and a microphone preamplifier. The fourth unit is presented with expansion connectors for connecting application specific hardware and an emulator connector for connection of a hardware emulator unit. The system data transfer rates between the DSP and the host PC vary between 50MB/s and 132MB/s. These figures are valid for the DSP running at 50MHz. This means that the DSP can read with a 25MHz clock (25MHz x 32bit = 100MB/s) and write with a 12,5MHz clock (50MB/s). In this way the whole 1MB RAM on the expansion card can be rewritten in 20ms. When accessing the memory via the PCI bus, the local bus operates with a 33MHz clock, which gives us the peak performance of 132MB/s. The development environment consists of the device drivers for Windows 95/98/NT/2000 and the utilities for initialization of the system, downloading of program code and communication with the card via the PCI bus. Without upgrades, the system is capable to run applications like a spectrum analyzer for low frequencies (up to 20kHz), 32 channel logic analyzer with a frequency up to 25MHz, a 32 bit generator of digital signals, with a frequency up to 12,5MHz, FIR filters with about 500 coefficients for filtering analogue signals, which are sampled with a frequency up to 48kHz, a sound effect processor, and many more.

Uvod

Pri delu z že obstoječimi sistemi (npr. DSK komplet podjetja Texas Instruments /2/) vedno znova naletimo na problem prenosa podatkov med signalnim procesorjem in osebnim računalnikom. Ta in temu podobni sistemi uporabljajo za komunikacijo z osebnim računalnikom standardni zaporedni ali vzporedni vmesnik. Slaba lastnost obeh vmesnikov je počasnost. Komunikacija preko teh vmesnikov sicer zadostuje za spoznavanje osnov digitalnega procesiranja signalov, odpove pa pri zahtevnejših aplikacijah, pri katerih imamo opravka z veliko količino podatkov. Da rešimo ta problem je smiselno počasni zaporedni ali vzporedni vmesnik nadomestiti s hitrim PCI vodilom, ki bo

omogočal razvoj kompleksnejših aplikacij (npr. razpoznavanje in sinteza govora, obdelava slike, celotni spekter avdio procesiranja, telekomunikacijski algoritmi,...). Ker pa je razvoj takšnih aplikacij pogosto povezan tudi z dodatno strojno opremo, mora sistem omogočati tudi nadgradnjo z dodatnimi vezji, kot so npr. digitalna sita, hitri A/D in D/A pretvorniki, MPEG kompresorji,... . Trenutno je na tržišču zelo malo proizvajalcev, ki bi ponujali takšno rešitev. Eno izmed njih ponuja podjetje Innovative Integration, Inc., ki pa ne izpolnjuje podanih zahtev, saj je kartica namenjena predvsem merjenju signalov v industriji in ima slabe možnosti razširitve. Večje število proizvajalcev pa ponuja kartice zasnovane na ISA vodilu, ki pa, upoštevaje prihodnji razvoj osebnih računalnikov ni več zanimiva rešitev, saj novejša računalniška oprema ne podpira več ISA vodila. Prav tako je ISA vodilo občutno počasnejše od PCI vodila. Obstoječe ISA kartice pa tudi ne vključujejo CODECov z več kot 16 bitno kvantizacijo, in tudi možnosti razširitev so precej skromnejše.

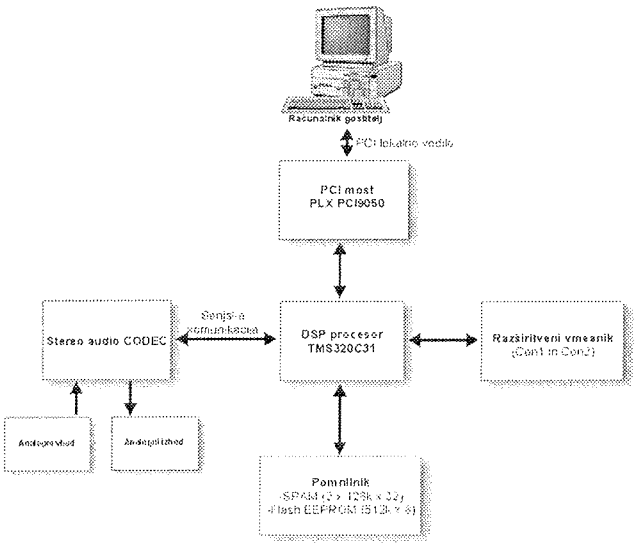
Zasnova razširitvene kartice

Osnovni deli strojnega dela sistema so: DSP procesor, pomnilniški podsistem, vhodno-izhodni vmesniki in razširitveni vmesnik.

Na sliki 1 je blokovna shema sistema. Osrednji del je digitalni signalni procesor (DSP) Texas Instruments TMS320C31. Naslednji večji del sistema je pomnilniški podsistem, ki je sestavljen iz 1MB statičnega pomnilnika in 512kB Flash pomnilnika. Slednji omogoča samostojno vzpostavitev in delovanje sistema brez priključitve na gostiteljski računalnik. Razširitveni vmesnik sestavljata dva konektorja, na katera je speljanih večina signalov. Tako imamo na razširitvenem vmesniku podatkovno in naslovno vodilo, napajanje, signale iz naslovnega dekodirnika (za enostavno dodajanje novih vezij) in prekinitvene linije. Ostali deli sistema so različni vhodni, izhodni in vhodno/izhodni vmesniki, namenjeni priključitvi perifernih naprav ter komunikaciji sistema z drugimi napravami.

DSP procesor Texas Instruments TMS320C31

Procesor TMS320C31 /1//5/ je eden od članov družine 32-bitnih procesorjev s plavajočo vejico Texas Instruments TMS320C3x. Značilnost te družine procesorjev je visoka stopnja vzporednosti; v enem ciklu lahko hkrati izvedejo množenje in aritmetično/logično operacijo nad celoštevilčnim ali realnim številom (številom s plavajočo vejico). Spremembe naslovnih registrov izvajata dve namenski aritmetično logični enoti (ARAU0 in ARAU1) in lahko zato potekajo vzporedno. Na procesorjih sta ob centralni procesni enoti (CPU-Central Processing Unit) tudi dva časovnika, zaporedni vmesnik (dva pri 'C30), DMA koprocesor (dva pri 'C32), dva bloka



Slika 1: Blokovna shema razširitvene kartice

pomnilnika, ROM pomnilnik (Boot Loader pri 'C31 in 'C32) in instrukcijski predpomnilnik (Cache).

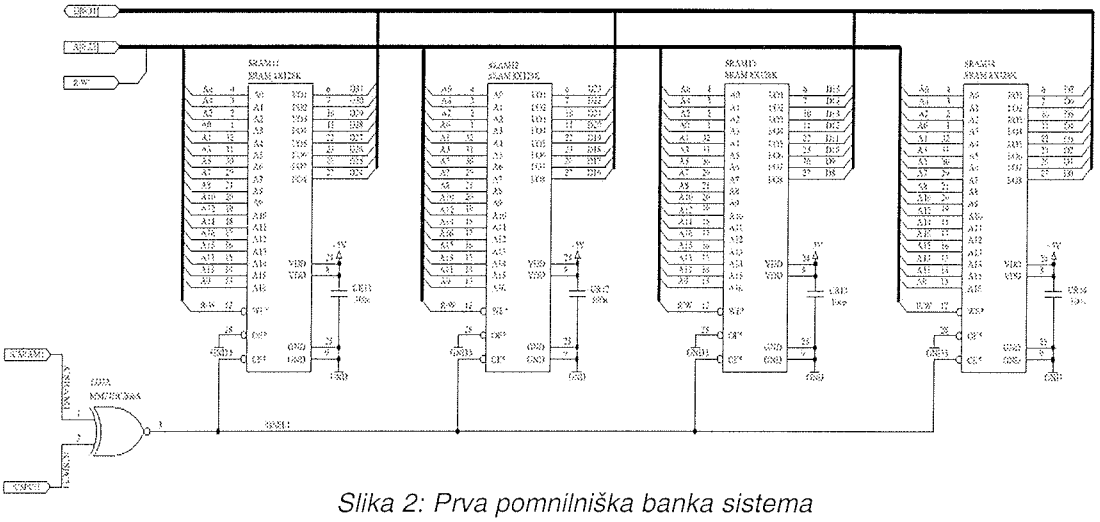
Pomnilniški podsistem

Pomnilniški sistem sestavljata statični pomnilnik /6/ velikosti 1MB in Flash pomnilnik /7/ velikosti 512kB, ki omogoča vzpostavitev in delovanje sistema brez posredovanja gostiteljskega računalnika.

Do statičnega pomnilnika lahko procesor dostopa brez čakalnih stanj, s tem popolnoma izkoriščamo procesorjevo hitrost tudi pri dostopanju do zunanjega pomnilnika. Flash pomnilnik lahko programiramo v samem sistemu, brez potrebe po programatorju. Lahko ga tudi zaščitimo pred pisanjem.

Statični pomnilnik /6/ je zgrajen iz dveh 32-bitnih bank velikosti 128k pomnilniških lokacij. Skupna velikost pomnilnika na sistemu je tako 2x128kx32=8Mbitov oziroma 1M zlogov (1MB). Podatkovne linije RAMov so povezane na podatkovne linije procesorja. Ker ima procesor 32 podatkovnih linij, vsak RAM pa osem, so za vsako banko potrebni štirje RAM-i. Na naslovne linije RAMov je vezanih spodnjih 17 naslovnih linij procesorja (2^17=131072 oziroma 128k naslovnih lokacij v eni banki). Prva banka je preslikana v naslovno območje od 820'000h do 83F'FFFh, druga pa v območje od 840'000h do 85F'FFFh. Zaradi izvedbe dekodirnika z vezjem 74F138 pride do večkratne preslikave fizičnega pomnilnika. Vezavo prve pomnilniške banke prikazuje slika 2.

Statični pomnilnik je dostopen tudi neposredno preko PCI vodila. Pri takšnem dostopu mora PCI most podati zahtevo po vodilu. Ko jo TMS odobri, lahko neposredno dostopamo do pomnilnika. Tak način komunikacije nam omogoča velik pretok podatkov in optimalno izkoriščenost DSP sistema. V našem sistemu smo uporabili DSP s frekvenco ure 50MHz, kar pomeni da



Slika 2: Prva pomnilniška banka sistema

lahko bere iz pomnilnika s frekvenco 25MHz. To nadalje pomeni da je prenos v tem primeru 100 MB/s. Ta vrednost je seveda najvišja meja, saj ne upošteva časa, ki je potreben za obdelavo podatka, prav tako so pisalni cikli enkrat počasnejši, kar pomeni da lahko zapisujemo z največ 50MB/s. Če posegamo do pomnilnika preko PCI vodila, je največja frekvenca vodila 33MHz, kar omogoča največji prenos do 132MB/s. Če upoštevamo dejstvo, da ima naš sistem 1MB spomina, pomeni to, da ga lahko v celoti prepišemo v 20ms, preberemo pa v 10ms.

Celotni Flash EEPROM je razdeljen na osem sektorjev, ki jih lahko neodvisno ali hkrati brišemo. V sistemu je Flash EEPROM preslikan v naslovno območje 400'000h-47F'FFFh in omogoča vzpostavitev sistema iz njega. Velikost naslovnega območja je 512k naslovov, kakršna je tudi zmogljivost EEPROMa, zato je v tem naslovnem območju le ena slika EEPROMa.

Povezava DSP sistema z osebnim računalnikom

Za povezavo signalnega procesorja z osebnim računalnikom smo uporabili PCI most podjetja PLX Technologies. Funkcija PCI mostu je, da povezuje PCI vodilo osebnega računalnika in lokalno vodilo DSP sistema. PCI most PCI 9050 /9/ je visokozmogljiv PCI vmesnik za razširitvene kartice. Načrtovan je tako, da omogoča priključitev širokega spektra sistemov z lastnim lokalnim vodilom na PCI vodilo osebnega računalnika in jim omogoča izmenjavo podatkov z osebnim računalnikom tudi do 132 MB/s. PCI 9050 lahko povežemo na multipleksirano ali nemultipleksirano lokalno vodilo, širine 8, 16 ali 32b. Predvsem možnost povezave 8b in 16b vodil na PCI vodilo omogoča enostavno predelavo starejših ISA kartic za uporabo na PCI vodilu. PCI 9050 popolnoma izpolnjuje zahteve PCI 2.1 specifikacije.

Delovanje PCI mostu PLX PCI9050 /9/ določamo in spremljamo s pomočjo lokalnih konfiguracijskih registrov. Spisek lokalnih konfiguracijskih registrov in njihovi odmiki od baznega naslova, so podani v tabeli 1.

Te registre lahko razdelimo v več skupin:

- registri za definiranje velikosti lokalnih naslovnih prostorov,
- registri za preslikavo PCI naslovnega prostora v lokalni naslovni prostor,
- registri za nastavitev lastnosti posameznih lokalnih naslovnih prostorov,
- registri za nastavitev izbirnih (chip select) linij,
- register za nastavitev in obravnavo prekinitev ter
- kontrolni register.

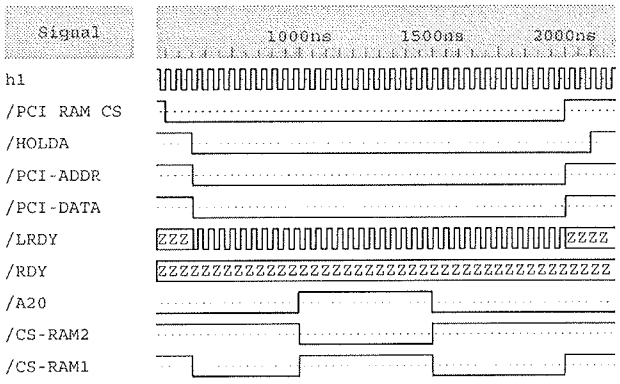
Tab. 1: Lokalni konfiguracijski registri PLX PCI9050 /9/.

PCI (Offset from Local Base Address)	To ensure software compatibility with other versions of the PCI 9050 family and to ensure compatibility with future enhancements, write "0" to all unused bits. 31 ... 0	PCI and EEPROM Writable
00h	Local Address Space 0 Range	Y
04h	Local Address Space 1Range	Y
08h	Local Address Space 2 Range	Y
0Ch	Local Address Space 3 Range	Y
10h	Local Expansion ROM Range	Y
14h	Local Address Space 0 Local Base Address (Remap)	Y
18h	Local Address Space 1 Local Base Address (Remap)	Y
1Ch	Local Address Space 2 Local Base Address (Remap)	Y
20h	Local Address Space 3 Local Base Address (Remap)	Y
24h	Expansion ROM Local Base Address (Remap)	Y
28h	Local Address Space 0 Bus Region Descriptors	Y
2Ch	Local Address Space 1 Bus Region Descriptors	Y
30h	Local Address Space 2 Bus Region Descriptors	Y
34h	Local Address Space 3 Bus Region Descriptors	Y
38h	Expansion ROM Bus Region Descriptors	Y
3Ch	Chip Select 0 Base Address	Y
40h	Chip Select 1 Base Address	Y
44h	Chip Select 2 Base Address	Y
48h	Chip Select 3 Base Address	Y
4Ch	Interrupt Control/Status	Y
50h	EEPROM Control, PCI Slave Response, User I/O Control, Int Control	Y

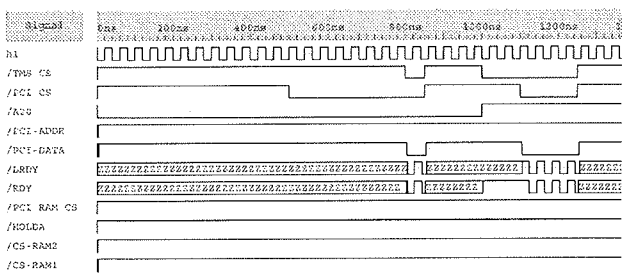
PCI most PLX PCI 9050 ima vgrajen vmesnik za priključitev zaporednega EEPROMA, ki ga lahko uporabimo za zapis začetnih nastavitev PCI mostu. PCI most podpira 1k bitne EEPROME, ki imajo MICROWIRE vmesnik in omogočajo sekvenčno branje. Takšni EEPROMI imajo ponavadi oznako 93C(S)46 (Pomen oznak: 93 - MICROWIRE vmesnik, C/CS – različne opcije, odvisno od proizvajalca in 46 - 1k bit spomina). V sistemu smo uporabili EEPROM proizvajalca ST Microelectronics z oznako M93C46 /10/.

Pri povezavi DSP sistema z osebnim računalnikom smo implementirali dva načina dostopa do DSP sistema. Prvi način je neposredni dostop do lokalnega pomnilnika DSP sistema, drugi način pa je neposredna komunikacija s signalnim procesorjem TMS320C31.

Neposredni dostop do lokalnega pomnilnika omogoča hitro branje in pisanje podatkov v lokalni pomnilnik DSP sistema. Pri tem načinu PCI 9050 da zahtevo po vodilu signalnemu procesorju. Ko ta zahtevo odobri, lahko PCI 9050 neposredno dostopa do lokalnega pomnilnika. Ker smo uporabili hiter statični pomnilnik, se ta komunikacija izvede v hitrem načinu. Hitrost ki jo dosežemo na ta način je odvisna od lokalne ure sistema in se giblje med 50MZlogov/s (12.5MHz x 32b) in 132MZlogov/s (33MHz x 32b), kar je tudi največja možna hitrost PCI 9050. Takšna komunikacija npr. omogoča prikaz rezultatov procesiranja v realnem času in popolno izkoriščanje DSP sistema. Signali, ki sodelujejo pri tej komunikaciji, so predstavljeni na sliki 3. Pri tem je h1 ura lokalnega vodila signalnega procesorja, /PCI RAM CS je signal s katerim procesorju signaliziramo, da želimo dostopati do lokalnega pomnilnika. Na to zahtevo DSP odgovori s postavitvijo signala /HOLDA, s čimer se odklopi od lokalnega vodila in prepusti pomnilnik PCI vodilu. PCI vodilo se priklopi na lokalno vodilo s pomočjo signalov /PCI-ADDR in / PCI-DATA. Pisalni in bralni cikel je določen s signalom /LRDY, pomnilniško banko pa izbiramo s signali /A20, /CS-RAM1 in /CS-RAM2.



Slika 3: Simulacija signalov za neposredni dostop do lokalnega pomnilnika.



Slika 4: Simulacija signalov za neposredno komunikacijo s TMS320C31.

Pri neposredni komunikaciji s TMS320C31 gre za neposredno komunikacijo s signalnim procesorjem TMS320C31. Pri tem izkoriščamo HOST-LOCKED način komunikacije signalnega procesorja TMS320C31. Ta način komunikacije uporabljamo večinoma za krmilno-statusno komunikacijo. Omogoča inicializacijo sistema, nalaganje programov v DSP ter krmilno - statusno komunikacijo z DSP procesorjem. Potek komunikacije prikazujejo signali na sliki 4. Če želimo podatke vpisati v DSP, postavimo signal /PCI CS na logično enico, DSP to zazna v obliki prekinitve in odgovori s postavitvijo signala /TMS CS. Ko sta oba signala aktivna se izvede pisalni cikel s pomočjo signalov /LRDY in /RDY.

Analogni vmesnik Crystal CS4222

Analogni vmesnik je namenjen zajemanju in predvajanju dveh kanalov analognih signalov, katerih frekvence so v območju človeku slišnih frekvenc (20Hz-20kHz). Kot osnovo analognega vmesnika smo uporabili stereo avdio kodek podjetja Crystal, z oznako CS4222 /8/. To je integrirano vezje, ki vsebuje 20 bitnidvokanalni analogno-digitalni (A/D) pretvornik in 20 bitni dvokanalni digitalno-analogni (D/A) pretvornik. Prenos digitalnih podatkov poteka preko zaporednega vmesnika, ki smo ga priključili na zaporedni vmesnik procesorja. A/D in D/A pretvornika delujeta po principu delta- sigma modulacije. Po podatkih proizvajalca je razmerje signal/šum 99dB, skupno harmonsko popačenje in šum pa sta tipično manjša od –90dB. Vezje podpira standardne frekvence vzorčenja 32kHz, 44.1kHz in 48kHz. S kakšno frekvenco vzorčenja bo vezje delovalo, je odvisno od sistemske ure vezja ter od urinih signalov na zaporednem vmesniku kodeka. Frekvenca sistemske ure je lahko 256-, 384- ali 512-kratnik frekvence vzorčenja. Avdio CODECu smo dodali tudi potrebno vhodno in izhodno prilagoditveno vezje, ter mikrofonski predojačevalec, ki omogoča neposredno priključitev mikrofona na kartico.

Emulatorski in razširitveni vmesnik

V procesor TMS320C31 je vgrajena tehnologija imenovana MPSD (Modular Port Scan Device), ki omogoča popolno emulacijo preko zaporednega XDS510 vmesnika.

Razviti sistem vsebuje tudi razširitveni vmesnik, ki služi priključevanju različnih perifernih naprav na DSP sistem. Razširitveni vmesnik je sestavljen iz dveh 50 pinskih ploščatih konektorjev. Na te konektorje je speljano celotno lokalno vodilo (naslovne in podatkovne linije, R/W', /READY in /STROBE signali), signali iz naslovnega dekodirnika, prekinitvene linije, prosti pini iz programirljive logike, linije za generiranje čakalnih stanj, liniji TCLK0 in TCLK1, ki jih lahko uporabljamo kot liniji internega časovnika TMS320C31 ali pa kot prosto programirljivi liniji (npr. za I2C) ter napajalne linije (+12V, -12V, +5V, +3.3V in GND).

Takšna zasnova rezširitvenih konektorjev omogoča enostavne nadgradnje sistema v obliki uporabniških modulov, ki se priključijo na te konektorje, saj imamo na voljo vse potrebne signale za krmiljenje in napajanje modulov ter za komunikacijo med procesorjem in moduli.

Programiranje sistema:

Programiranje sistema lahko razdelimo na dva dela. Prvi del zajema komunikacijo med osebnim računalnikom in PCI kartico. Programiranje tega dela izvajamo v okolju Windows 95/98/Me/NT/2000. Programiramo lahko v različnih programskih jezikih, kot so npr. C++, Visual Basic, Delphi,...

Drugi del programiranja zajema programiranje signalnega procesorja TMS320C31 in perifernih naprav (Avdio CODEC, I²C vodilo, razširitvene kartice, Flash pomnilnik). Ta del programiranja izvajamo v programskem razvojnem okolju, ki vsebuje tudi zbirnik in programski jezik C, ki ga ponuja podjetje Texas Instruments /2/ /3/ /4/.

Trenutno je na voljo naslednja programska oprema:

1. Gonilnik za operacijski sistem Win dows95/98/Me/NT/2000 /11/.
2. Dinamična knjižnica (DLL) z osnovnim naborom funkcij, ki omogočajo inicializacijo sistema, nalaganje programske kode v DSP in komunikacijo z DSP procesorjem preko PCI vodila.
3. Programi v zbirniku, ki omogočajo uporabo perifernih enot (avdio CODEC, FLASH pomnilnik ter I²C vmesnik).

Na voljo je tako vsa osnovna programska oprema, ki omogoča razvoj široke palete aplikacij digitalnega procesiranja signalov v realnem času.

Zaključek

Sistem je načrtovan zelo odprto, saj omogoča enostavno nadgrajevanje preko razširitvenih konektorjev in omogoča tudi spremembe znotraj samega sistema. Logiko v programirljivem logičnem vezju Lattice, v katerem je precejšen del sistema, lahko izkušen uporabnik po potrebi spremeni oziroma prilagodi lastnim zahtevam. V Flash pomnilnik, ki je del sistema, lahko enostavno vpišemo delujočo aplikacijo in tako dobimo namensko PCI kartico (npr. glasbena kartica, koprocesorska kartica,...). Sistem je zasnovan tako, da v največji možni meri izkorišča zmožnosti signalnega procesorja in lahko brez razširjanja pomnilniških kapacitet na njem izvedemo aplikacije kot so spektralni analizator za nizkofrekvenčne signale (z analognim vmesnikom na sistemu do frekvenc okoli 20kHz), 32-kanalni logični analizator z vzorčenjem do 25MHz (pri uri procesorja 50MHz), generator 32-bitnih digitalnih signalov do 12.5MHz, FIR filter z okoli 500 koeficienti za frekvenco vzorčenja 48kHz, procesor zvočnih efektov in podobno. Zamenjava vzporednega oz. zaporednega vmesnika s hitrim PCI vodilom, omogoča razvoj množice aplikacij, ki zahtevajo kompleksnejše obdelave signalov in predvsem hitre prenose podatkov med razširitveno kartico in gostiteljskim računalnikom.

Literatura:

/1/ Texas Instruments TMS320C3x User's Guide, SPRU031E Revision L. Texas Instruments Incorporated, U.S.A., Julij 1997.
/2/ Texas Instruments TMS320C3x DSP Starter Kit User's Guide, SPRU163A. Texas Instruments Incorporated, U.S.A., Junij 1996.
/3/ Texas Instruments TMS320C3x/C4x Assembly Language Tools User's Guide, SPRU035C. Texas Instruments Incorporated, U.S.A., Februar 1998.
/4/ Texas Instruments TMS320C3x/C4x Optimizing C Compiler User's Guide, SPRU034H. Texas Instruments Incorporated, U.S.A., Junij 1998.
/5/ TMS320C31, TMS320LC31 Digital Signal Processors, SPRS035A. Texas Instruments Incorporated, U.S.A., Julij 1997.
/6/ 131'072-Word by 8-bit CMOS Static Ram. Toshiba Corporation, Junij 1997.
/7/ Flash Memory MBM29F040A - 4M (512K x 8) BIT. Fujitsu Limited, Japan, 1994.
/8/ CS4222, 20-Bit Stereo Audio Codec with Volume Control. Crystal Semiconductor Corporation , U.S.A., Januar 1997.
/9/ PCI 9050-1 Data Book, V1.01, PLX Technology Inc., April 1997
/10/ M93C86, M93C76, M93C66, M93C56, M93C46, M93C06 16K/8K/4K/2K/1K/256 Serial Microwire Bus EEPROM, STMicroelectronics, 1999.
/11/ WinDriver for PLX I/O devices V5.0 Developer's Guide, KRF Tech LTD., 1999.

Daniel Čeh-Ambruš univ. dipl. inž., raziskovalec v laboratoriju za digitalno procesiranje signalov na fakulteti za elektrotehniko, računalništvo in informatiko v Mariboru,

Iztok Kramberger univ. dipl. inž., asistent na fakulteti za elektrotehniko, računalništvo in informatiko v Mariboru,

izr. prof. dr. Zdravko Kačič, izr. profesor na fakulteti za elektrotehniko, računalništvo in informatiko v Mariboru.

Fakulteta za elektrotehniko, računalništvo in informatiko v Mariboru
Smetanova 17, 2000 Maribor

Prispelo (Arrived): 17.11.00 *Sprejeto (Accepted): 07.03.01*

A STUDY OF HIGHLY DOPED LAYERS FOR BICMOS COLLECTOR INSERTS

Radko Osredkar, Faculty of Computer Sciences and Faculty of Electrical Eng., University of Ljubljana, Slovenia,
and
Boštjan Gspan, Repro MS, Ljubljana, Slovenia

Keywords: microelectronics, IC, Integrated Circuits, BiCMOS circuits, Bipolar CMOS circuits, circuit fabrication, dopant concentration profile, thin layers, collector insert layers, computer simulations, technological treatments

Abstract: In order to develop a BiCMOS IC fabrication module based on implantation and diffusion of dopants for the collector insert layers for the vertical npn bipolar transistors, we performed a detailed comparison of simulations of such a process and results of actual fabrication of the layers. We conclude that the thickness and half—width of the collector insert layers can be simulated accurately (within 4% of measured values, on average), which satisfies the requirements for developing the process module. However, for process control purposes simulations seem to have to be augmented by a direct measuring method.

Študija močno dopiranih plasti za BiCMOS kolektorske vložke

Ključne besede: mikroelektronika, IC vezja integrirana, BiCMOS CMOS vezja bipolarna, izdelava vezij, profil koncentracije dopantov, plasti tanke, vložki kolektorski transistorjev, simulacije računalniške, postopki tehnološki

Izvleček: Z namenom, da bi za razvili modularni tehnološki postopek za izdelavo integriranih vezji tipa BiCMOS, na osnovi implantirane in difundirane dopirane plasti za kolektorski vložek vertikalnih, bipolarnih npn tranzistorjev, smo izvedli natančno primerjavo računalniške simulacije izdelave takšnih plasti in fizikalnih karakteristik izdelanih plasti. Ugotovili smo, da se simulacije in meritve debeline plasti in njene polovične širine ujemajo v povprečju na 4 %. To zadošča za potrebe razvoja procesa, medtem ko je za procesno kontrolo simulacije dopolnjevati z neposrednimi meritvami fizikalnih lastnosti plasti.

1. Introduction

BiCMOS technology integrates both CMOS and bipolar device structures on the same chip. This capability can be exploited in a number of ways to produce integrated circuits (IC) with performances that exceed those that are possible when only one of the device types is used. These benefits are attained at the expense of a more complex fabrication technology, including development and chip manufacturing tasks, however, in many high performance digital applications and in mixed analog/digital systems the benefits often exceed these extra costs. The key feature of a BiCMOS technology is fabrication of the collector inserts for the bipolar npn vertical transistors. In high performance BiCMOS ICs the fabrication of the insert is usually based on epitaxial layers. However, if one wishes to augment an existing CMOS technology by adding bipolar devices (e.g. in I/O drivers), an improvement of IC performance can still be achieved by a simpler, and relatively low cost BiCMOS technology, involving essentially only additional implantation and diffusion processing steps /1/. A serious limitation in developing such a process is that there exists no reliable nondestructive method for characterizing the insert layers, in addition to the fact that lack of a reliable measuring method makes tight control the fabrication process difficult. However, computer process simulation is sometimes a viable alternative to characterization measurements required by the development and production. In this contribution we present an attempt to correlate a destructive layer characterization method, Spreading Resistance Analysis (SRA), with results of a detailed simulation of a 1.2 μm IC fabrication process, extended to BiCMOS capability, and comment on the applicability of simulation for development and production needs.

2. Experimental

Test processing in our study was designed as an extension (i.e. insertion module) in an existing 1.2 μm CMOS process. (This is a standard, although proprietary 17 layer process defined and used by the IMP, San Jose, Ca., USA.) /2, 8/ It was executed on p-type, <1-0-0> silicon wafers, nominally 25-50 Ωcm, selected for uniformity. The fabrication of the collector insert layer consisted of doping in a two-step process: a phosphorus implantation followed by a deposition of a layer of phosphorus-doped poly-Si. The implantation was performed at 150 keV, at the dose of $8 \cdot 10^{12}$ ions/cm^2. The poly-Si layer was the dopant source for diffusion of phosphorus into silicone during the deposition itself and in subsequent processing step. In a preliminary study it has been determined that the

required insert layer properties can be achieved only by a modification of the standard process: doped poly-Si layer was deposited at 1150 °C in N_2 atmosphere, followed by oxidation at 1130 °C. At this stage the poly-Si layer was removed from the wafer surface and further diffusion performed at 1000 °C for 240 min. This processing results in a thick (approximately 5 μm), low resistivity (1.5 to 2.5 Ωcm) layer that is adequate for fabrication of the collector inserts. For convenience, the details of the process parameters of short loop are summarized in Table 1.

Tab. 1: Process parameters of key process steps in fabrication of the collector insert layer.

process step	process parameters
P implantation	dose: 8 x 10^{12} ions/cm^2; voltage: 150 keV
poly-Si deposition	temp: 1150 °C dep. time (diffusion 1): 32, 24.5, 22, 20 min
oxidation	temp: 1130 °C time (diffusion 2): 14.5, 11, 8.5, 8.5 min
diffusion	temp: 1000 °C time: 240 min

Finally the experimental run was split in two parts: one was analyzed by the SRA method and the other further processed (metallization, etching) to allow standard C-V analysis. The poly-Si deposition and diffusion steps was performed in a Thermco diffusion furnace /3/.

A C-V characterization of the insert layers with a parametric tester, Hewlett Packard model HP 4062B, was attempted. From the numerical analysis of the C-V data the dopant profile can in principle be determined /4,5/ if certain limitations of the method, concerning the size of the capacitor, dopant concentration and inversion layer depth are not exceeded. However, in our study precisely this was the case; the method yielded reasonable data for surface dopant concentrations but the insert layer thickness could not be determined by this method.

Spreading Resistance Analysis (SRA) allows for an accurate and relatively simple determination of the dopant profile /6/ by a 2-point method. However, it is destructive. A bevel, inclined approximately 0.5 deg to the wafer surface, is ground into the wafer with 5 μm grit. Knowing the angle of the taper gives the distance from the surface as a function of displacement along the bevel. The shallow taper allows an accurate probing of the profile in over 100 steps of 10 μm each. The probe contact force was 7.5 x 10^{-2} N. SRA measurements were performed at Solecon Laboratories Inc, Ca., USA, according to their standard procedures. Having the SRA measurements performed at a commercial laboratory was deemed necessary as no reliable calibrated dopant profile standards are available which would insure the required precision of measurements.

For simulations of the collector insert fabrication steps the well known 2 dimensional program package SUPREM 4, with capability of generating data for the 2-D IC simulator PISCES 2B, (7), run on s Sun workstation, was used.

3. Results and discussion

The simulations program package that we have used is usually applied for simulations and analysis of standard IC fabrication steps for which processing parameters are, to a certain extend, generic. However, fabrication of collector inserts involves parameters that exceed the standard ones and it was not obvious from the outset that the mathematical models used in the simulations package are adequate for modeling the formation of very thick, highly doped films. A detailed, point by point comparison of simulation results and SRA profile measurements was therefore performed. 4 different processing sequences (each involving 5 wafers) were studied, the relevant details of which are given in Table 2, and the results of measurements and simulations in Table 3.

Tab. 2. Details of the 4 different processing sequences studied.

#	p_B (cm^{-3})	diffusion 1	diffusion 2	bevel angle (rad)
1	8 . 10^{14}	32 min	14.5	0.0099
2	8	24.5	11.0	0.0103
3	8	22	8.5	0.0107
4	8	20	8.5	0.0107

Tab. 3. Results of SRA measurements and simulations for the 4 different processing sequences studied

#	SRA measurements			simulation		
	n_S (cm^{-3})	$d_{½}$ (μm)	d_i (μm)	n_S (cm^{-3})	$d_{½}$ (μm)	d_i (μm)
1	9.0 10^{19}	3.7	4.8	4.5 10^{19}	3.6	4.8
2	9.5	3.8	4.5	3.8	3.8	4.9
3	7.0	3.8	4.6	4.0	3.7	4.8
4	6.5	3.6	4.5	4.3	3.9	4.7

It has been determined that some of the features of the profiles can be simulated quite accurately while the surface concentrations of the treated wafers less so. Even though over-all agreement was deemed sufficient for proces development purposes, it has

been attempted to reduce these differences by varying the simulation parameters, but no significant improvement could be achieved.

The resulting shape of the concentration profiles resulting from fabrication steps employed in our study is quite similar to those of the standard processing steps, which have been well studied and are understood /2, and References therein/. Therefore it is sufficient to characterize our profiles with 4 parameters only: concentration of dopants in untreated wafer p_B (bulk concentration; for our study wafers with the same $p_B = 8 \cdot 10^{14}$ cm^{-3} were selected), surface concentration of the treated wafers (i. e. on top of the fabricated collector insert) p_s, half-width of the distribution $d_{1/2}$, conveniently defined as the depth at concentration

$$c_{1/2} = \mathrm{sqrt}(n_s \times p_B),$$

and the depth of the inversion layer d_i. Comparison of measured and simulated values shows that the differences in $d_{1/2}$ and d_i in no case exceeds 8 %, the average vale of the difference being less than 4 %. Our data demonstrate that the doped poly-Si layer is not depleted as a source during the deposition and subsequent oxidation. Thus the details and possible variations during processing of these two steps are not critically reflected in the properties of the final insert layer. The layer profile and its thickness (depth of the inversion layer) are primarily determined by the final diffusion step, which insures reliability and repeatability of the fabrication module, as was intended.

The relatively large discrepancy between the measured and simulated values of n_S (largest difference 60 %, average over the 4 sequences 45 %) could not be traced to a single cause with any degree of certainty. The measured values are in all cases larger than the simulated ones and it seems probable that the actual thermal balance of the processing is not faithfully reflected by the simulation, and thus the effect of the depletion of dopants in the poly-Si exaggerated. However, the integrated concentrations of the dopants in a fabricated layer, which largely determine its bulk properties and have been calculated from the measuring and simulations data, differ only slightly and are not significantly effected by the differences in n_S obtaint via the two routes. From the production standpoint the discrepancy is a point of only minor concern as the over-all properties of the collector insert are satisfactory and the developed production module has been proven to yield working ICs with the required properties. However, it demonstrates the limitations of simulations for process control and the need for a physical measuring method for such purposes.

4. Conclusion

On the basis of over-all and reasonable agreement between the results of measurements and simulations, we conclude that computer simulations of processing steps required for fabrication of the collector inserts are a reliable tool in developing BiCMOS technologies in which such layers are fabricated by implantation and diffusion of dopants. In particular, the thickness of the collector insert layer and its resistance can be simulated accurately. It seems that with some care such an approach, combined with a direct measuring method (e.g. C-V measurements), can be used in process control, and possibilities in this direction are presently being explored in our laboratory.

Acknowledgements

The authors gratefully acknowledge the use of IMP, San Jose, Ca., USA, facilities. Special thanks are due to Mr. A. Belič of IMP for discussions regarding the utility of simulations in process control. The study has been supported by a grant from the Ministry of Science and Technology of the Republic of Slovenia.

References

/1/ S. Wolf, Silicon Processing for the VLSI Era, Vol. 2 – Process integration, Lattice Press, Sunset Beach, Ca. USA, (1990), pp. 529—558
/2/ B. Gspan, Magistrsko delo, Fakulteta za elektrotehniko, Univerza v Ljubljani, (1992)
/3/ Thermco Systems Specification No.: TXSN—8100—UL—CL—72
/4/ A. Belič, R. Osredkar, Avtomatizirano določanje implantacijskega profila s CU metodo, ELVEA 2, 55 (1988), pp. 21—23
/5/ Technical information on Application Software "Doping", HP, pp. 1—8, Aug. (1986)
/6/ S. Wolf, Silicon Processing for the VLSI Era, Vol. 1 – Process technology, Lattice Press, Sunset Beach, Ca. USA, (1990), p. 274
/7/ J. D. Plummer, R. W. Dutton, Process Simulators for Silicon VLSI and High Speed GaAs Devices, SRC Technical Report No.: T86085 – Integrated Circuits Laboratory, Stanford University, Ca., USA, Oct. (1986)
/8/ B. Gspan, Ph.D. Thesis, Faculty of Electrical Eng., University of Ljubljana, 1995

Prof. dr. Radko Osredkar
FRI in FE Univerze v Ljubljani
Tržaška 25
SI 1000, Ljubljana
Slovenia
e-mail: radko.osredkar@fri.uni-lj.si

Dr. Boštjan Gspan
Repro MS
Šmartinska 106
SI 1000, Ljubljana
Slovenia
e-mail: bostjan.gspan@reproms.si

Prispelo (Arrived): 17.02.01 *Sprejeto (Accepted): 07.03.01*

37th INTERNATIONAL CONFERENCE ON MICROELECTRONICS, DEVICES AND MATERIALS

and the WORKSHOP on

OPTOELECTRONIC DEVICES AND APPLICATIONS

Announcement and Call for Papers

October 10. - 12. 2001
Bohinj, SLOVENIA

http://paris.fe.uni-lj.si/midem/conf2001/

Elektrotehniška Zveza Slovenije

IEEE Slovenia Section

ORGANIZER

MIDEM - Society for Microelectronics, Electronic Components and Materials
Dunajska 10, 1000 Ljubljana, SLOVENIA

CONFERENCE SPONSORS

Ministry of Education, Science and Sport, Republic of Slovenia
IMAPS, Slovenia Chapter
IEEE, Slovenia Section

CONFERENCE PROGRAMME COMMITTEE

Franc Smole, Faculty of Electrical Engineering, Ljubljana, Slovenia, **chairperson**

Marko Topič, Faculty of Electrical Engineering, Ljubljana, Slovenia, **cochairperson**

Marija Kosec, Jožef Stefan Institute, Ljubljana, Slovenia
Marko Hrovat, Jožef Stefan Institute, Ljubljana, Slovenia
Slavko Amon, Faculty of Electrical Engineering, Ljubljana, Slovenia
Lojze Trontelj, Faculty of Electrical Engineering, Ljubljana, Slovenia
Cor Claeys, IMEC and KU Leuven, Belgium
Gerhard W.Herzog, Technische Universitaet, Graz, Austria
Bruno Cvikl, Faculty of Civil Engineering, Maribor, Slovenia
Miloš Komac, Ministry of Science and Technology, Republic of Slovenia
Wilhelm Kusian, SIEMENS Corporate R& D Department, Muenchen, Germany
Peter Panjan, Jožef Stefan Institute, Ljubljana, Slovenia
Stane Pejovnik, Faculty for Chemistry and Chemical Technology, Ljubljana, Slovenia
Wolfgang Pribyl, Austria Mikro Systeme Intl.AG, Graz, Austria
Nava Setter, Ecole Polytechnique Federal de Lausanne, Lausanne, Switzerland
Giovanni Soncini, University of Trento, Trento, Italy
Giorgio Pignatel, University of Trento, Trento, Italy
Iztok Šorli, MIKROIKS d.o.o., Ljubljana, Slovenia
Jiri Toušek, Charles University, Prague, Czech Republic
Anton Zalar, ITPO, Ljubljana, Slovenia
Miloš Somora, Technical University, Košice, Slovakia
Leszek J.Golonka, Technical University, Wroclaw, Poland
Zsolt Vitez, Technical University, Budapest, Hungary
Monika Jenko, Institute for Metals and Technology, Ljubljana, Slovenia

CONFERENCE ORGANIZING COMMITTEE

Meta Limpel, MIDEM, Ljubljana, Slovenia,
Uroš Aljančič, Faculty of Electrical Engineering, Ljubljana, Slovenia
Mišo Vukadinovič, Faculty of Electrical Engineering, Ljubljana, Slovenia
Iztok Šorli, MIKROIKS d.o.o., Ljubljana, Slovenia

GENERAL INFORMATION

The 37th International Conference on MIcroelectronics, DEvices and Materials, MIDEM 2001, continues the tradition of annual international conferences organized by the MIDEM Society. These conferences have always attracted a large number of Slovene and foreign experts working in these fields.

The topics covered by the conference are quite diverse, and presenting about 60 papers in five sessions over three days seems rather demanding. However, once a year scientists and engineers have the opportunity to present their work to the international public and to meet and discuss trends, news and problems related to their field of work. We believe that this at least balances the effort required by the attendees and the organizer.

The conference is well known in the electronics community. Hundreds of distinguished scientists from all over the world have taken part in previous MIDEM conferences. The goal of establishing contacts, collaboration and friendship among scientists and their companies remains the main aim for the organizer.

Therefore, you are kindly invited to take part in the forthcoming :

37th International Conference on Microelectronics, Devices and Materials - MIDEM 2001 Conference

The conference will be held in Hotel Zlatorog, Bohinj, Slovenia, October 10 – 12, 2001

ORIGINAL PAPERS RELATING TO THE FOLLOWING AREAS ARE INVITED FOR SUBMISSION :

- Novel monolithic and hybrid circuit processing techniques
- New device and circuit design
- Process and device modeling
- Semiconductor physics
- Sensors and detectors
- Electromechanical devices
- Microsystems
- Optoelectronics
- Photovoltaic devices
- New electronic materials and applications
- Electronic materials science and technology

- Materials characterization techniques
- Reliability and failure analysis
- Education in microelectronics, devices and materials

As with previous years each session will be introduced by distinguished invited speakers giving an overview presentation on a related field.

INVITED CONFERENCE PAPERS

Prof.Dr.Radivoje S.Popović
EPFL-Swiss Federal Institute of Technology Lausanne, Switzerland
Christian Schott and Robert Racz
SENTRON AG , Zug, Switzerland
Integrated Hall / Flux Concentrator Microsystems

Abstract: The speaker gives a survey of a new class of highly sensitive single-axis and two-axis integrated Hall magnetic sensor microsystems. They are based on an integrated combination of a CMOS integrated circuit and a planar magnetic flux concentrator. The magnetic flux concentrator is made of a thick ferromagnetic layer deposited on the CMOS wafer in a post-processing step. The CMOS part of the system contains two or more conventional Hall elements positioned under the peripheries of the flux concentrator. The flux concentrator converts locally a magnetic field parallel with the chip surface into a field perpendicular to the chip surface. Therefore, a conventional Hall element can detect an external magnetic field parallel with the chip surface. Moreover, the flux concentrator provides a magnetic gain of about 10.

WORKSHOP on OPTOELECTRONIC DEVICES AND APPLICATIONS

Begining in 1998, workshops dedicated to a special field were incorporated to the programme of the MIDEM Conferences. During the workshop, five to seven invited speakers present papers on the chosen topics from different aspects within their special field, thus offering the audience valuable information. Time for thorough discussions is provided between invited presentations. Conference attendees are encouraged to present their research results in the Conference session dealing with the dedicated topic. Attendance at the workshop is included in the Conference registration fee.

For the year 2001, we are pleased to announce a

Workshop on OPTOELECTRONIC DEVICES AND APPLICATIONS.

Selected topics associated with advanced aspects within lasers, light emitting devices, thin-film transistors, optical fiber systems, etc. will be presented, covering the basic physical principles, as well as actual and possible applications of these optelectronic devices and systems.

The workshop is organized by Laboratory of Semiconductor Devices at Faculty of Electrical Engineering.

The programme committee is pleased to announce the following invited speakers, who will give their presentations on the following subjects:

Prof. Dr. Martin Stutzmann
Walter Schottky Institut, Technische Universität München, Germany
The Status of GaN-based LEDs and Laser Diodes

Abstract: High brightness blue, green, and white light emitting diodes based on InGaN/GaN/AlGaN heterostructures have developed into a mass product, with more and more companies competing with each other in a rapidly expanding market. Full colour displays, traffic lights, car lighting, back lighting for LCD displays, and indoor spot light applications are the main present and emerging applications. Yet, the efficiency of GaN-based LEDs is still limited by poor materials quality due to the lack of a suitable substrate for homoepitaxy and by the problems with p-type doping, in particular in the AlGaN-cladding layers of a typical quantum well LED. Here, more research is necessary before the full potential of solid state lighting can be exploited also for high efficiency applications in office or home illumination. The basic materials problems are even more important in the development of blue laser diodes for data storage. At present, only Nichia Corp. has reported laser diodes emitting in the near UV and exhibiting a reasonable device lifetime of several thousand hours. Most other companies still are fighting severe degradation problems of their laser diodes, which are mainly attributed to defects caused by heteroepitaxy on sapphire or SiC substrates. Also, no real blue laser diodes with an acceptable lifetime have been reported. The purpose of this presentation is to summarize the present status concerning GaN-based LEDs and LDs, to discuss the basic materials physics behind the success and the limitations of these devices, and to point out present attempts and ideas to overcome these limitations in the future.

Prof. Dr. R. E. I. Schropp
Debye Institute, Physics of Devices, Utrecht University, The Netherlands
New Developments in Thin Film Transistor Techology

Abstract: Thin film transistors (TFTs) are currently used as the switching element in active matrix (AM) liquid crystal displays (LCD), such as the TFT display

used in lap top computers. Recently, page-size 2D matrix-addressed image sensors for application in digital copiers and X-ray imaging have been presented. Current trends in research and development are: 1) TFTs on plastic substrates (low T deposition is required); 2) The development of low temperature polysilicon for the pixel TFTs: polysilicon TFTs have a higher field-effect mobility and allow larger drive currents, so that the pixel aperture ratio can be increased and bright, low power LCDs can be obtained even when they have a large size, such as in the notebook and the CRT-monitor replacement market; 3) Integration of row and column drivers on the glass, so that panels can be made much thinner: by supporting both n-channel and p-channel operation, poly-Si TFTs also enable CMOS circuits for display drivers; 4) Addressing of OLEDs (Organic Light Emitting Diodes) by silicon TFTs. For such advanced applications of TFTs a few issues are relevant: (i) fast, defect free deposition of thin silicon films and gate dielectrics, (ii) higher electron mobility, (iii) stability. Whereas a high deposition rate is generally needed to reduce the production cost, for novel high current applications the latter two issues have recently become more essential.

Prof.Dr.Gil Rosenman
Department of Electrical Engineering-Physical Electronics, Tel Aviv University, Israel

Engineered Domain Configurations For Nonlinear Optical Devices

Abstract: New generation of nonlinear optical devices allows developing coherent light sources in spectral regions where conventional lasers are unavailable. It is based on engineered micrometer-scale ferroelectric domain configurations. Applied quasi-phase-matching technique exploits tensorial properties of ferroelectrics where specifically designed domain structure with alternative direction of spontaneous polarization provides needed variation of a sign of a nonlinear optical coefficient.

Ferroelectric $LiNbO_3$, $LiTaO_3$, KTP and isomorphic crystals are mainly exploited due to their high nonlinear optical coefficients and stable built-in domain structure. The conventional process of domain grating fabrication consists of application of an electric field to a photolithographically patterned electrode on a polar crystal surface made of a arrays of metallic and insulating strips. Detailed studies showed those fundamental processes of minimization of depolarization field, domain nucleation, anisotropy of domain walls velocity determine quality of tailored domain configurations and strongly affect parameters of optical converters.

Highly homogeneous periodic, aperiodic and multiple domain gratings with periods in the range 4-39 μm for diverse quasi-phase-matched nonlinear optical devices in UV, visible and infrared regions have been developed.

Dr. Christian Hanke
Corporate Research Photonics, Infineon Technologies, Munich, Germany

High Power Semiconductor Laser Diodes

Abstract: In the last decades the output power of semiconductor laser diodes has increased dramatically. Starting from a power range of several milliwatts, which is sufficient for a range of mass applications in the field of optical communication and optical storage systems, now semiconductor laser systems with an output power in the kilowatt range are available. The main advantages of diode lasers are the small volume, the high overall efficiency up to 60%, the availability of a wide spectral range and the high reliability. This combination together with the high output power opens a wide field of applications covering e. g. pumping of solid state lasers and amplifiers, transfer to printing plates, soldering and direct machining.

The high optical and thermal power densities in semiconductor laser diodes lead to limitations of the maximum useful output power. The approaches to improve the device performance will be described for single stripe laser diodes with diffraction limited beam quality and for laser arrays with multimode emission. To further increase the output power combinations of several emitters are used and beam combining techniques have been developed. The present status and future developments will be presented.

Dr. Helmut Stiebig
Institut für Photovoltaik, Forschungszentrum Jülich, Germany

Color Aliasing Free Detectors

Abstract: Color image processing is usually performed with the aid of color filter array (CFA) coated CCD or CMOS sensor arrays. However, color detection with CFA leads to the color moiré or color aliasing effect, which is observed when structures with high spatial frequencies are captured. Furthermore, traditional sensor systems exhibit a rather limited resolution and a low fill factor, because one color pixel is split into several chromatic sub pixels. In order to overcome the color aliasing effect, vertically integrated sensor structures have been proposed, which detect the color information in the depth of the structure. The spectral response, dynamic range and the transient behavior of the vertically integrated sensors based on amorphous and crystalline silicon are determined by the optoelectronic properties of the employed materials and the device design (two, three and four terminal devices). We will discuss different three-color detectors and compare the properties of the structures (e.g. p/n, pin, piiin, nipin) regarding their appropriated applications.

Additionally, three-color sensors are limited by the mismatch between the spectral sensitivity of the de-

tector and the human eye. This mismatch leads to a color error. For high quality and low color error imaging applications (e.g. picture archiving) typically multispectral technology is applied. In this case the sensor array is covered sequentially with different color filters and the individual spectrum of each color point is reconstructed afterwards. Due to the fact that several images of the same scene are taken sequentially real time imaging is prevented. To overcome this drawback of a multi shot mode we have developed vertically integrated 4 and 6 channel detectors based on amorphous silicon and its alloys, which can be read out with one and two shots, respectively.

The purpose of this presentation is to demonstrate the advantages of amorphous silicon based detectors in the field of color sensor technology and to point out the operation of the different device structures to fit the demand of various applications.

Dr. Matjaž Vidmar
Faculty of Electrical Engineering, University of Ljubljana, Slovenia

Optical-fiber Communications: Components and Systems

Abstract: Optical-fiber communications brought a revolution to communication technology, outperforming other communication systems by several orders of magnitude in transmission capacity, unrepeated and repeated communication range and decreasing installation and operating costs. The optical-fiber revolution started approximately 30 years ago, when technology improvements decreased the optical-fiber loss to less than 20dB/km. The theoretical loss limit for silica (SiO2) based fibers was reached only 10 years later, but the fiber handling and line-terminal technology was far from mature at that time. Even with primitive line-terminal technology, optical fibers immediately outplaced coaxial-cable systems and decreased the importance of microwave and satellite point-to-point radio links. In the last two decades, significant improvements have been made in the line-terminal technology, including narrow-spectrum solid-state lasers, wide-bandwidth modulators, laser optical amplifiers, fast and sensitive photodetectors and last but not least, high speed electronics. Although advanced laboratory experiments are quickly approaching the theoretical capacity offered by the >10 THz bandwidth of a single mode optical fiber, several problems have yet to be solved to make high-capacity systems viable, including linear and nonlinear propagation effects in the optical fiber itself, high performance electro/optical and opto/electric converters, efficient high speed electronics an all-optical signal-processing components. The purpose of this presentation is to summarize the present status of optical-fiber communication technology, to discuss the basic components and the limitations of these devices, and to present the requirements and proposals for future systems.

Dr. Gvido Bratina
Nova Gorica Polytechnic, Slovenia

Organic Semiconductors as Candidates for Advanced Optoelectronic Devices

Abstract: Organic semiconductors are rapidly emerging as promising candidates for the expansion of the optoelectronic devices to the field of flexible material systems. Relatively weak van der Waals intermolecular bonding allows the fabrication of organic semiconducting layers on a variety of substrates ranging from glass to thin polymer foils. The electronic and optical properties of thin organic semiconducting layers are strongly dependent on their chemical composition and structural parameters. Both features may be varied, making the synthesis of light emitting or light detecting devices with variable operating wavelength straightforward. Typical possible applications include multi-color light emitting diodes and flexible full color displays.

The level of understanding of the processes in the field of organic light emitting devices, however, is analogous to the status in the field of III-V semiconductor-based devices in the early seventies. The material purity is one of major concerns, and depends heavily on the method of preparation. The structural parameters crucially affect the mobility of charge carriers, and are a function of growth protocol as well as of the type of the substrate. Doping of organic semiconductors appears to be extremely uncontrollable. Metal-organic-semiconductor contacts are instrumental to the operation of every device, and as such are a subject of intensive research.

The first part of the lecture will give an overview of the demonstrated organic-semiconductor-based light emitting devices. In the second part the efforts to clarify electronic transport in thin organic-semiconductor layers and metallic contacts will be illustrated.

CONFERENCE PROCEEDINGS

Invited and contributed papers will be published in the Conference Proceedings and distributed at the Conference registration.

LANGUAGE

The official Conference language is English.

IMPORTANT DATES

Abstract deadline : **June 1st**
Notification of acceptance : **June 25th**
Preliminary Program : **September 25th**
Only on Conference Web page
http://paris.fe.uni-lj.si/midem/conf2001/
Paper deadline : **September 15th**
Final conference program : on registration, **October 10th**

Please, see instructions for ABSTRACT and PAPER submission on our WEB page !!

ACCOMODATION

The Conference will take place in the Hotel Zlatorog, Bohinj.

The hotel is mirrored in the crystal waters of Lake Bohinj in the heart of Triglav National Park and is surrounded by the peaks of the Julian Alps. A setting of scilence, tranquility, sunshine and greenery.

Please send your room reservations indicating " for MIDEM 2001 Conference", directly to :
Hotel Zlatorog
Ukanc 65, 4265 Bohinjsko jezero
Slovenia
Tel.+386 (0)4 5723 281, fax.+386 (0)4 5723 384
Hotel.zlatorog@bohinj.si

SOCIAL EVENTS

The Conference dinner will be held on Thursday, October 11 at 20:00.

REGISTRATION

The registration fees are as follows :

- FULL registration fee : **200 EUR (*)**
- MIDEM Society members, MIDEM sponsors : **150 EUR**

The fee includes Conference Proceedings and free access to all Conference events (Conference dinner).
(*) Full registration fee (applies to non MIDEM society members ONLY) includes a two year full MIDEM Society membership.

Undergraduate students have free access to all Conference sessions on submitting their study papers. For other Conference events there will be an additional charge.

Accompanying persons who will not take part in the conference may join other conference events at an extra charge (DINNER : 30 EUR).

Programme and Organizing Committee, MIDEM 2001 Conference

MIDEM at MIKROIKS
Mrs Meta Limpel
Stegne 11
1521 Ljubljana, SLOVENIA
tel.+386-(0)1-5112 221, fax.+386-(0)1-5112 217
email : Iztok.Sorli@guest.arnes.si

Conference Web page:
http://paris.fe.uni-lj.si/midem/conf2001/
MIDEM 2001 Conference official Email:
midem@paris.fe.uni-lj.si

Contact person for the Workshop on OPTOELECTRONIC DEVICES AND APPLICATIONS

Dr.Marko Topič, workshop chairperson
Faculty of Electrical Engineering in Ljubljana
Tržaška 25, 1000 Ljubljana, Slovenia
tel.+386 (0)1 4768 470 , fax.+386 (0)1 4264 630
Email : marko.topic@fe.uni-lj.si

ODPRTJE KOLARJEVE PREDAVALNICE

Zamisel o tem, da bi seminarsko sobo Odseka za keramiko poimenovali po prof. dr. Dragu Kolarju, se je porodila že februarja lani, potem ko nas je naš cenjeni kolega nepričakovano zapustil. 14. februarja je minilo leto od takrat, zato smo se tega dne na Odseku za keramiko poklonili spominu nanj. Nabito polna, od sedaj Kolarjeva, predavalnica, ki sta jo odprla prof. dr. Vito Turk in prof. dr. Marija Kosec, je zgovorno pričala o tem, kako živ je spomin na prof. Kolarja v vseh, ki smo ga imeli priložnost poznati in delati z njim in ob njem. Vabilu so se odzvali številni sodelavci Instituta ,"Jožef Stefan", profesorjevi sodelavci s fakultet, institutov, podjetij in tovarn, prišla je tudi njegova hči Jana. Prireditev je obogatila nadarjena mlada violinistka, dijakinja srednje glasbene šole, Živa Cigalenečki, ki je zaigrala nekaj skladb, s katerimi se uspešno uveljavlja na mednarodnih mladinskih tekmovanjih. Prvo predavanje v Kolarjevi predavalnici z naslovom Strukture v sistemu $TiO_2-Ga_2O_3-La_2O_3$ je pripravil prof. dr. Ljubo Golič, Kolarjev dolgoletni fakultetni in institutski sodelavec.

Sodelavci Odseka za keramiko se zahvaljujemo vsem, ki so z nami počastili spomin na prof. Kolarja, vsem, ki so poklicali, da ne bodo mogli priti in vsem, ki v sebi, kot mi, hranijo majhen prostorček za tega velikega človeka.

Govor ob odprtju Kolarjeve predavalnice je imela vodja Odseka za keramiko prof.dr. Marija Kosec.

Prispevek je pripravila Sanja Fidler in je bil objavljen v internem glasilu NOVICE IJS št. 87, marec 2001