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

INFORMACIJE MIDEM 3·2003

Strokovno društvo za mikroelektroniko elektronske sestavne dele in materiale

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

INFORMACIJE MIDEM, LETNIK 33, ŠT. 3(107), LJUBLJANA, september 2003

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

INFORMACIJE MIDEM 3 ◦ 2003

INFORMACIJE MIDEM LETNIK 33, ŠT. 3(107), LJUBLJANA, SEPTEMBER 2003
INFORMACIJE MIDEM VOLUME 33, NO. 3(107), LJUBLJANA, SEPTEMBER 2003

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
MIDEM pri MIKROIKS
Stegne 11, 1521 Ljubljana, Slovenija
tel.: + 386 (0)1 51 33 768
fax: + 386 (0)1 51 33 771
e-mail: Iztok.Sorli@guest.arnes.si
http://paris.fe.uni-lj.si/midem/

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 EUR 100, separate issue is EUR 25. 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+54+621+66), ISSN0352-9045 Informacije MIDEM 33(2003)3, Ljubljana

VSEBINA CONTENT

Obnovitev članstva v strokovnem društvu MIDEM in iz tega izhajajoče ugodnosti in obveznosti

Spoštovani,

V svojem več desetletij dolgem obstoju in delovanju smo si prizadevali narediti društvo privlačno in koristno vsem članom.Z delovanjem društva ste se srečali tudi vi in se odločili, da se v društvo včlanite. Življenske poti, zaposlitev in strokovno zanimanje pa se z leti spreminjajo, najrazličnejši dogodki, izzivi in odločitve so vas morda usmerili v povsem druga področja in vaš interes za delovanje ali članstvo v društvu se je z leti močno spremenil, morda izginil. Morda pa vas aktivnosti društva kljub temu še vedno zanimajo, če ne drugače, kot spomin na prijetne čase, ki smo jih skupaj preživeli. Spremenili so se tudi naslovi in način komuniciranja.

Ker je seznam članstva postal dolg, očitno pa je, da mnogi nekdanji člani nimajo več interesa za sodelovanje v društvu, se je Izvršilni odbor društva odločil, da stanje članstva uredi in **vas zato prosi, da izpolnite in nam pošljete obrazec priložen na koncu revije.**

Naj vas ponovno spomnimo na ugodnosti, ki izhajajo iz vašega članstva. Kot član strokovnega društva prejemate revijo »Informacije MIDEM«, povabljeni ste na strokovne konference, kjer lahko predstavite svoje raziskovalne in razvojne dosežke ali srečate stare znance in nove, povabljene predavatelje s področja, ki vas zanima. O svojih dosežkih in problemih lahko poročate v strokovni reviji, ki ima ugleden IMPACT faktor.S svojimi predlogi lahko usmerjate delovanje društva.

Vaša obveza je plačilo članarine 25 EUR na leto. Članarino lahko plačate na transakcijski račun društva pri A-banki : 051008010631192. Pri nakazilu ne pozabite navesti svojega imena!

Upamo, da vas delovanje društva še vedno zanima in da boste članstvo obnovili. Žal pa bomo morali dosedanje člane, ki članstva ne boste obnovili do konca leta 2003, brisati iz seznama članstva.

Prijavnice pošljite na naslov:

MIDEM pri MIKROIKS

Stegne 11

1521 Ljubljana

Ljubljana, april 2003

Izvršilni odbor društva

ALGORITEM ZA IZBIRO USTREZNEGA EMI FILTRA

[1]Marko Podberšič, [2]Matjaž Šegula, [3]Vojko Matko

[1]Ministrstvo za obrambo, Uprava RS za zaščito in reševanje, Ljubljana, Slovenija
[2]ISKRAEMECO d.d, Razvoj in raziskave, Kranj, Slovenija
[3]Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko, Maribor, Slovenija

Ključne besede: EMC, EMI filter, kritična dolžina linije, tipična frekvenca.

Izvleček: Članek opisuje način izbire ustreznega EMI filtra za določeno signalno linijo. Do sedaj smo EMI filtre izbirali na osnovi meritve sevanja celotnega CPU modula. V tem članku pa opisujemo izbiro filtra na osnovi meritve FFT signala. Takšna izbira je zagotovo optimalnejša od prve saj optimiziramo filter za vsako signalno linijo posebej. To smo pokazali na primeru CPU modula.

A Suitable EMI Filter Selection Alghoritm

Key words: EMC, EMI filter, electrically long trace, critical frequency.

Abstract: The article describes a new method of selection of a suitable EMI filter for a signal line. This method is based on measurement of FFT of a signal. The selection is rather good, because we optimise EMI filter for each signal line separately. This was presented on an example of a CPU module.

We know actually three EMI filter selection methods. The first EMI filter selection method is based on a realization of EMI filters with a help of ground planes. All signal lines must be surrounded with ground (fig. 1). We can get so rather good capacitive draining of high frequency noise to the ground. Such capacitive draining is some sort of EMI filter. This capacitive draining is better if the coupling path between the signal line and the ground plan is longer. Many times this filter is not good enough. In these cases real EMI filters must be used. The second EMI filter selection method is based on a measurement of a radiation of whole equipment. At this method, we measure the radiation of whole equipment and get a discrete frequency component with maximum amplitude. We select EMI filters with maximum insertion loss at frequency, which is close by this frequency. We use EMI filters with similar characteristics on the whole equipment. The third EMI filter selection method is based on a measurement of a FFT on a single signal line. This method is described in this article.

Some technical expressions are described in first chapters of this article. Such expressions are a critical line length, S parameters of two-input circuit, a typical frequency - observed as EMI, an input impedance of EMI filter, adjustment and filtering.

An expression "critical line length" is known in the high-speed transmission-line theory. We determine a critical line length with help of the frequency F_{knee} (equation 14). At this critical line length the rise-time, T_r, exactly matches the propagation delay time, T_{pd}. This means that the transient phenomenon formed by the low-to-high signal transition precisely fits the line length. For that reason, this distance is called the "length of the rising edge". We must stress that the critical line length l_{max} means two-way propagation delay (source-load-source). A line length equal to or longer as the critical length certainly behaves as a transmission-line. This means that you must consider characteristic impedance, delay and reflections in that case.

S parameters are almost always presented because EMI filters are usually two-input circuits. The letter S comes from an english word "Scattering". S parameters describes that an incoming power in one input is distributed among all inputs of multi-input linear circuit. We are using these S parameters for an input impedance of the EMI filters calculation.

Typical frequency - observed as EMI, depends upon used logic elements and microcontroller. Precisely, it depends upon rise time of signals, which are transmited by these elements (equation 35). We must be attentive when we select appropriate logic elements and microcontroller. If we use, for example, faster HCT instead of slower LS-TTL, the electromagnetic emission increases for up to three times. Of course this typical frequency finds expression at certain line length (emission radiation problem). These typical frequencies - observed as EMI are very important when we develop an electronic circuit.

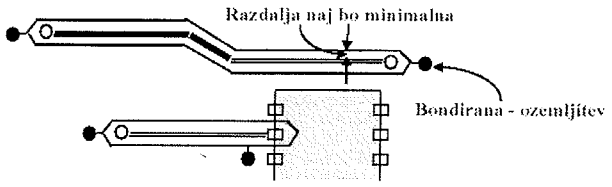
A procedure of the new EMI filter selection method is as following:
- Measuring the rise time t_r of the signal;
- Calculating (or measuring) the typical frequency - observed as EMI;
- Selecting a suitable EMI filter family with regard to an application;
- Selecting an EMI filter from the family. We select the filter with maximum insertion loss at frequency, which is close by the typical frequency.

The need of the EMI filters is conditional on the critical line length. This is verified by our experiences. If a two-way line length is shorter to the previously calculated l_{max} (critical line length) and there is no vias on the line, the usage of an EMI filter is not necessary.

We have shown on a example of a CPU module that this new EMI filter selection method based on FFT measurements of signals is better then the old one based on a measurement of a radiation of the whole CPU module.

1 Uvod

Zelo velikega pomena pri reševanju EMC problemov je pravilna izvedba ozemljitve. Pri tem je pomembno, da so posamezne signalne vezi tesno obdane z ozemljitveno površino (slika 1). Tako dobimo dokaj dober kapacitivni odvod visokofrekvenčnih motenj na zemljo /1/.

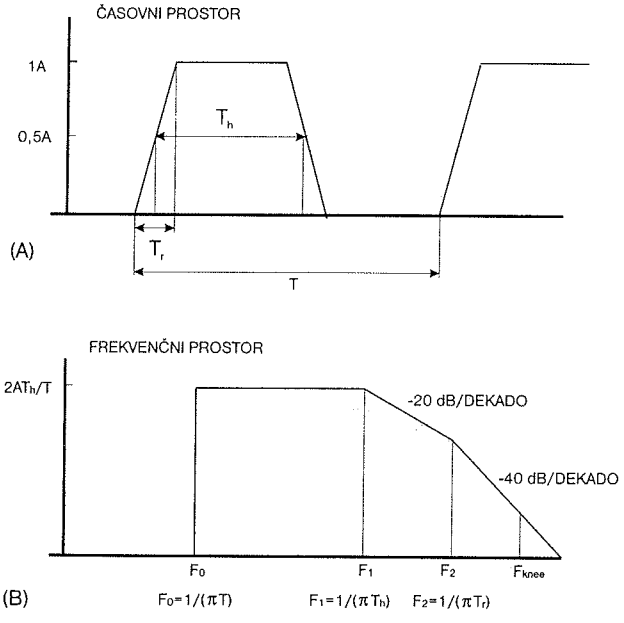


Slika 1: Pravilna izvedba ozemljitve

Takšen kapacitivni sklop predstavlja nekakšen EMI filter, ki pa seveda mnogokrat ne zadošča. V teh primerih uporabimo prave EMI filtre. Način izbire teh filtrov si bomo ogledali v nadaljevanju tega članka.

2 Kritična dolžina linije

V teoriji hitrih prenosnih poti poznamo izraz imenovan "kritična dolžina linije" /2/. Kaj je kritična dolžina linije? Odvisna je od frekvenc signalov, ki jih želimo prenašati preko linije. Spekter digitalnega periodičnega signala trapezoidne oblike prikazuje slika 2.



Slika 2: Spekter per. signala trapezoidne oblike

Ta signal lahko opišemo kot serijo spektralnih komponent:

$$f(t)=\frac{A_0}{2}+\sum_{n=1}^{\infty}\left(A_n\cos(n\omega_0 t)+B_n\sin(n\omega_0 t)\right) \qquad (1)$$

Koeficiente A_0, A_n in B_n dobimo s pomočjo sledečih enačb:

$$\omega_0=\frac{2\pi}{T} \qquad (2)$$

$$A_0=\frac{2}{T}\int_{t_0}^{t_0+T}f(t)dt \qquad (3)$$

$$A_n=\frac{2}{T}\int_{t_0}^{t_0+T}f(t)\cos(n\omega_0 t)dt \qquad (4)$$

$$B_n=\frac{2}{T}\int_{t_0}^{t_0+T}f(t)\sin(n\omega_0 t)dt \qquad (5)$$

kjer je:
ω_0 - osnovna krožna frekvenca,
T - perioda signala,
t_0 - čas v trenutku opazovanja.

Ko imamo enak dvižni čas T_r in čas spusta T_f trapezoidnega signala, dobimo poenostavljen primer Fourierove vrste. Tok ali napetost n-tega harmonika dobimo iz enačbe (6).

$$A_n=2A\frac{T_h}{T}\cdot\frac{\sin\left(n\pi\frac{T_h}{T}\right)}{n\pi\frac{T_h}{T}}\cdot\frac{\sin\left(n\pi\frac{T_r}{T}\right)}{n\pi\frac{T_r}{T}} \qquad (6)$$

kjer je:
A - amplituda signala od vrha do vrha,
T_h - širina pulza,
T_r - čas vzpona,
T_f - čas upada,
T - perioda signala,
n - številka harmonika.

Ta spekter je sestavljen iz diskretnih frekvenčnih komponent $f_n = nf_T$, kjer je $f_T = 1/T$. Če narišemo asimptote na ta spekter, dobimo horizontalno linijo do prve kolenske frekvence (sl. 2B):

$$F_1=\frac{1}{\pi T_h} \qquad (7)$$

Od tu naprej dobimo padajočo linijo, ki pada z naklonom 20 dB/dekado do druge kolenske frekvence (sl. 2B):

$$F_2=\frac{1}{\pi T_r} \qquad (8)$$

Od tu naprej dobimo padajočo linijo, ki pada z naklonom 40 dB/dekado (sl. 2B).

$$F_{knee}=\frac{1}{2T_r} \qquad (9)$$

Frekvenca F_{knee} je praktično maksimalna frekvenca in je približno 1,5 x F_2.

S pomočjo frekvence F_{knee} bomo določili kritično dolžino linije. Znano je, da sevanje narašča s frekvenco, dokler polovična valovna dolžina signala, ki se širi po liniji, ne doseže dolžine linije.

$$F_{knee} = \frac{1}{2T_r} \quad \text{[Hz] lastnost signala} \qquad (10)$$

$$\frac{\lambda_{knee}}{2} = \text{dolžina linije [m]} \quad \text{lastnost linije} \qquad (11)$$

Hitrost širjenja signala po liniji je:

$$V_{prop} = \frac{1}{T_{pd}} \text{ [m/s]} \qquad (12)$$

T_{pd} je zakasnitev signala. Sedaj lahko valovno dolžino (11) zapišemo kot:

$$\lambda_{knee} = \frac{V_{prop}}{F_{knee}} \text{ [m]} \qquad (13)$$

S pomočjo enačb (10) in (12) dobimo kritično dolžino linije l_{max}:

$$l_{\max} = \frac{\lambda_{knee}}{2} = \frac{T_r}{T_{pd}} \text{ [m]} \qquad (14)$$

Pri kritični dolžini linije, se prehodni pojav, ki nastane s prehodom signala iz nizkega v visoko stanje, natančno ujame z dolžino linije. Zato, to razdaljo imenujemo tudi "dolžino pozitivne flanke". Poudariti je potrebno, da je pri kritični dolžini linije l_{max} upoštevana celotna dolžina linije - potovanje signala od oddajnika do sprejemnika in nazaj.

Za poenostavitev enačbe (14), uporabimo dejansko vrednost časa zakasnitve signala za material FR-4. Tako dobimo enačbi (15) in (16), po katerih lahko izračunamo kritično dolžino linije. Ti izračuni veljajo za dielektrično konstanto ε_r=4,6. Ta dielektrična konstanta velja za material FR-4 in je dobljena na osnovi referenčnega signala 1MHz /3/.

$l_{max} = 9 \cdot T_r$ (T_r vstavimo v [ns]) (15)

(za mikrostrip tehnologijo - v cm)

$l_{max} = 7.T_r$ (T_r vstavimo v [ns]) (16)

(za stripline tehnologijo - v cm)

Linija, ki ima dolžino daljšo ali enako kritični dolžini se obnaša kot prenosna linija. To pomeni, da moramo upoštevati karakteristično impedanco, zakasnitev in odboje.

3 S parametri dvo-vhodnega vezja

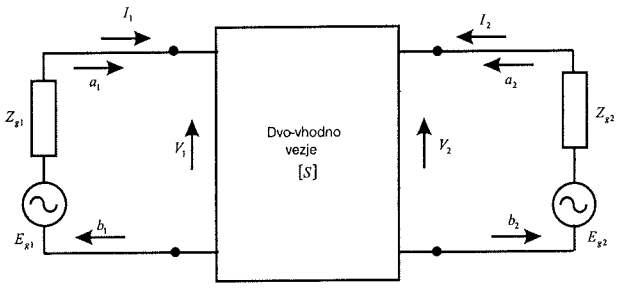
EMI filtri so običajno dvo-vhodna vezja. Skoraj vedno so za njih podani S parametri /4/.

Črka S izvira iz angleške besede "Scattering", kar pomeni "porazdeljeni". S parametri opisujejo, kako se dovedena moč na enem vhodu porazdeli med vse vhode več-vhodnega linearnega vezja.

Pri dvo-vhodnem vezju lahko normalizirana odbita valova b_1 in b_2 izrazimo z normaliziranima vpadnima valovoma a_1 in a_2 kot sledi:

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \qquad (17)$$

Takšno dvo-vhodno vezje prikazuje slika 3.



Slika 3: S parametri dvo-vhodnega vezja

Normalizirana vpadna vala sta podana z enačbama:

$$a_1 = \frac{V_1 + Z_{g1} I_1}{2\sqrt{\mathrm{Re}(Z_{g1})}} \qquad (18)$$

$$a_2 = \frac{V_2 + Z_{g2} I_2}{2\sqrt{\mathrm{Re}(Z_{g2})}} \qquad (19)$$

Normalizirana odbita vala pa sta podana z enačbama:

$$b_1 = \frac{V_1 - Z_{g1}^{*} I_1}{2\sqrt{\mathrm{Re}(Z_{g1})}} \qquad (20)$$

$$b_2 = \frac{V_2 - Z_{g2}^{*} I_2}{2\sqrt{\mathrm{Re}(Z_{g2})}} \qquad (21)$$

Spremenljivke S_{11}, S_{22}, S_{12} in S_{21} imenujemo S parametre dvo-vhodnega vezja. Iz zgornjih enačb lahko izrazimo S parametre kot sledi:

$$S_{11} = \frac{b_1}{a_1} \Big| a_2 = 0 \qquad (22)$$

S_{11} je odbojni koeficijent na vhodu 1, pri prilagoditvi na vhodu 2.

$$S_{21} = \frac{b_2}{a_1} \Big| a_2 = 0 \qquad (23)$$

S_{21} je prenosni koeficijent dvo-vhodnega vezja pri pogoju a_2 = 0 (prilagoditev na vhodu 2).

Podobno lahko zapišemo še:

$$S_{12} = \frac{b_1}{a_2} \Big| a_1 = 0 \qquad (24)$$

in

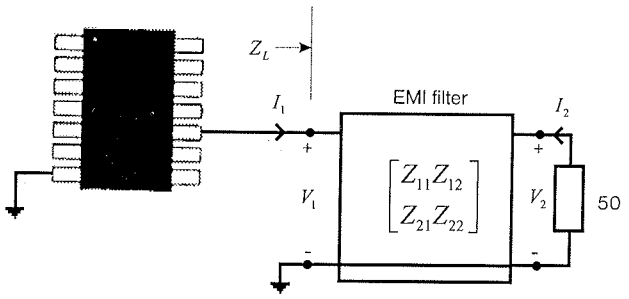
$$S_{22} = \frac{b_2}{a_2}\Big|a_1 = 0 \qquad (25)$$

S_{12} in S_{22} sta prenosni in odbojni koeficijent dvo-vhodnega vezja pri prilagoditvi na vhodu 1.

Te S parametre bomo uporabili za izračun vhodne impedance EMI filtrov.

4 Izračun vhodne impedance

Za izražanje vhodne impedance EMI filtrov so sicer bolj primerni Z parametri, saj Z_{11} že izraža vhodno impedanco vhoda 1 pri odprtih sponkah na vhodu 2. EMI filter je tu predstavljen kot klasično dvo-vhodno vezje, tako da izračun velja splošno za vsa dvo-vhodna vezja. Ekvivalentno vezje takšnega EMI filtra prikazuje slika 4. Na izhod filtra je priključeno 50 Ω breme /5/.



Slika 4: Ekvivalentno vezje EMI filtra

Od proizvajalca EMI filtrov smo za posamezne filtre dobili podane S parametre v odvisnosti od frekvence. S parameri so podani v obliki kompleksnih števil.

S parametre pretvorimo v Z parametre kot sledi:

$$\begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} = \frac{1}{(1-S_{11})\cdot(1-S_{22}) - S_{12}\cdot S_{21}} \cdot$$

$$\cdot \begin{bmatrix} (1+S_{11})\cdot(1-S_{22}) + S_{12}\cdot S_{21} & 2\cdot S_{12} \\ 2\cdot S_{21} & (1-S_{11})\cdot(1+S_{22}) + S_{12}\cdot S_{21} \end{bmatrix} \qquad (26)$$

Z Z parametri lahko zapišemo povezavo med tokovi in napetostmi na sponkah vezja na sledeči način:

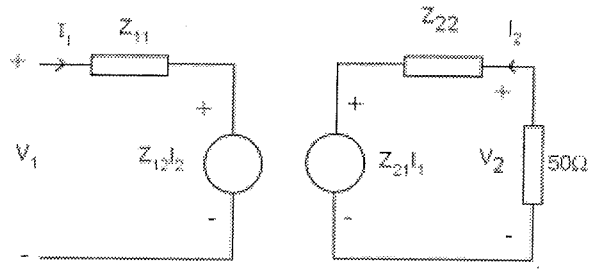
$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} \cdot \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \qquad (27)$$

Iz zgornje enačbe dobimo izraz za Z_{11} kot sledi:

$$Z_{11} = \frac{V_1}{I_1}\Big|I_2 = 0 \qquad (28)$$

Z_{11} torej predstavlja vhodno impedanco dvo-vhodnega vezja, pri pogoju, da so na vhodu 2 odprte sponke. Dejansko imamo redko na vhodu 2 odprte sponke, pač pa neko breme. V našem primeru je to 50 Ω breme. Zanima nas, kakšno impedanco Z_L čuti izhod integriranega vezja v tem primeru.

Impedančne enačbe dvo-vhodnega vezja lahko ponazorimo z nadomestim vezjem, ki ga prikazuje slika 5.



Slika 5: Nadomestno vezje

Velja:

$$Z_L = \frac{V_1}{I_1} \qquad (29)$$

Iz nadomestnega vezja na sliki 5 sledi:

$$V_1 = Z_{11}I_1 + Z_{12}I_2 \qquad (30)$$

$$I_2 = \frac{Z_{21}I_1}{Z_{22} + 50} \qquad (31)$$

$$V_1 = Z_{11}I_1 + \frac{Z_{12}Z_{21}I_1}{Z_{22} + 50} \qquad (32)$$

Impedanca Z_L je torej:

$$Z_L = \frac{V_1}{I_1} = Z_{11} + \frac{Z_{12}Z_{21}}{Z_{22} + 50} \qquad (33)$$

Impedanca Z_L nam predstavlja impedanco, ki jo čuti izhod integriranega vezja na svojih sponkah (slika 4). Če hočemo, da odda izhod integriranega vezja bremenu (filtru) največjo moč in da ne pride do odboja, mora biti vhodna impedanca filtra Z_L enaka konjugirano kompleksni impedanci izhoda:

$$Z_L = Z_{izh}^* \qquad (34)$$

V tem primeru pravimo, da sta izhod integriranega vezja in filter z bremenom prilagojena.

5 Tipična frekvenca opazovana kot EMI

Tipična frekvenca - opazovana kot EMI, je odvisna od uporabljenih logičnih elementov in mikrokrmilnika. Bolje rečeno, odvisna je od dvižnih časov signalov, ki jih ti elementi oddajajo (35). Pri razvoju strojne opreme je še kako pomembna pravilna odločitev pri izbiri ustrezne logične družine in mikrokrmilnika. Če npr. namesto LS-TTL vzamemo hitrejši HCT, bo narasla emisija električne poljske jakosti za trikratni faktor. Zato naj ne bi nikoli izdelali vezja v širšem pasovnem območju, kot je nujno. Seveda pa ta tip-

ična frekvenca pride do izraza (emisija sevalne narave) šele pri določeni dolžini vezi /3/.

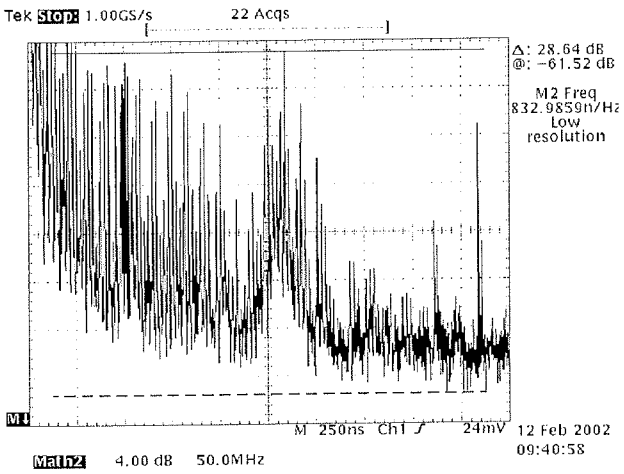
$$F_{\max} = \frac{10}{\pi T_r} \qquad (35)$$

Pri načrtovanju elektronskih vezij v smislu EMC-ja so zelo pomembne tipične frekvence, opazovane kot EMI, posameznih elementov.

Da bi preverili pravilnost enačbe (35), smo napravili tudi meritev harmonika pri tipični frekvenci opazovani kot EMI. Zaradi ne povsem ustrezne merilne opreme, smo morali meritev nivoja harmonika pri tipični frekvenci, opazovani kot EMI, izvesti na 74HC logiki, ki ima tipično frekvenco pri 270 MHz. Najboljši nam dostopen osciloskop, je imel frekvenčno omejitev pri 500 MHz, kar pa je za AC logiko prenizko. Tipična frekvenca AC logike se nahaja pri 1,6 GHz.

Izhodni signal iz 74HC245 logičnega vezja prikazuje slika 6. Tipična frekvenca je lepo vidna pri 265,258 MHz. Prepoznali smo jo po ponovnem porastu harmonikov nekje med 220 MHz in 300 MHz, z vrhom pri tipični frekvenci 265,3 MHz.

Na sliki 6 vidimo, da je nivo harmonika pri tipični frekvenci 28,64 dB (27,04 V).



Slika 6: FFT signala na izhodu iz 74HC245 logičnega vezja

Z meritvijo časa porasta signala t_r smo dobili za t_r vrednost 12 ns. Iz časa porasta signala t_r smo lahko izračunali tipično frekvenco, opazovano kot EMI, po enačbi (35). Dobili smo vrednost 265,258 MHz. Rezultat (36) se popolnoma sklada z meritvijo.

$$F_{\max} = \frac{10}{\pi T_r} = \frac{10}{\pi \cdot 12ns} = 265{,}258MHz \qquad (36)$$

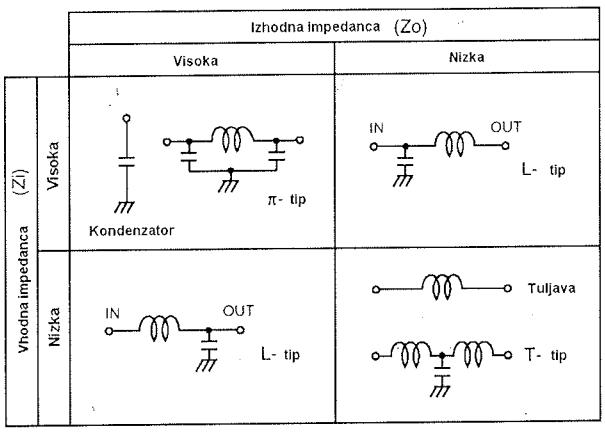
Tipično frekvenco, opazovano kot EMI lahko torej dokaj enostavno izračunamo. Pravilnost izračuna smo potrdili z meritvijo.

6 Prilagoditev in filtriranje

Do sedaj na prilagoditve nismo izrecno pazili, pa tudi vhodne impedance filtrov nismo preračunavali. To smo si nekako lahko privoščili, saj smo imeli opravka s sorazmerno počasnimi procesorji (do 24 MHz) in majhnimi ploščicami tiskanega vezja (kratke antene).

Izvajanja v poglavju 3, so torej plod razmišljanj za bližnjo prihodnost, saj se temu verjetno ne bomo mogli dolgo izogibati. Seveda pa smo na linije, ki so daljše od kritične dolžine linije (l_{max}) namestili EMI filtre. Postavili smo jih blizu izvora motenj. S filtri smo poleg filtriranja signala izboljšali tudi prilagoditev. Pri tem smo morali paziti na strukturo uporabljenih EMI filtrov.

V katalogih podajamo dušenje EMI filtrov pri vhodni in izhodni impedanci 50Ω /6/. Običajno pa v realnih vezjih nimamo take impedance. Znano je, da je učinkovitost filtrov močno odvisna od vhodne in izhodne impedance, to je od impedanc vezja, kamor je filter vgrajen. Veljajo neka splošna pravila, ki se jih običajno držimo /7/. Ta pravila prikazuje slika 7.



Slika 7: Izbira ustreznega EMI filtra

Vemo, da je kondenzator bolj učinkovit pri dušenju motenj v visoko impedančnih vezjih, tuljava pa je bolj učinkovita v nizko impedančnih vezjih. Slika 7 prikazuje tabelo, s pomočjo katere izberemo ustrezen filter glede na vhodno in izhodno impedanco.

7 Izbiranje ustreznih EMI filtrov

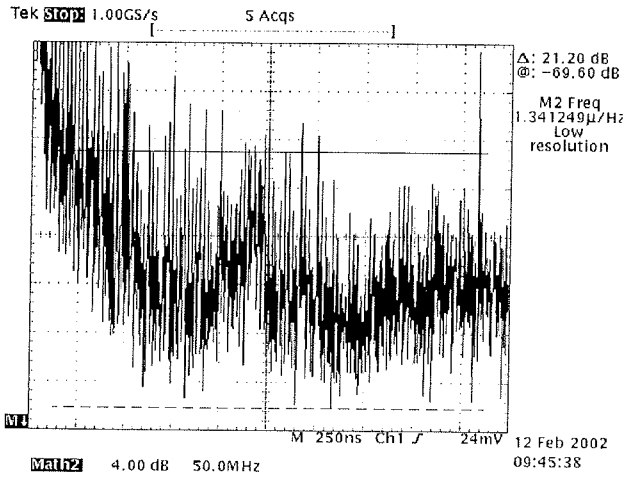
Postopek izbire optimalnega EMI filtra je sledeč:

- izmerimo čas porasta signala t_r;
- izračunamo (lahko tudi izmerimo) tipično frekvenco, opazovano kot EMI;
- glede na aplikacijo izberemo ustrezno družino EMI filtrov;
- iz izbrane družine EMI filtrov vzamemo tistega, ki ima maksimum dušenja čim bližje tipični frekvenci.

Potreba po EMI filtrih je pogojena s kritično dolžino linije. To potrjujejo tudi naše izkušnje /8/. Če je dvosmerna dolžina linije krajša od predhodno izračunane l_{max} (kritična dolžina linije) in ni na liniji nobene vije razen pri priključkih integriranega vezja (DIP ohišje), uporaba EMI filtra ni nujna.

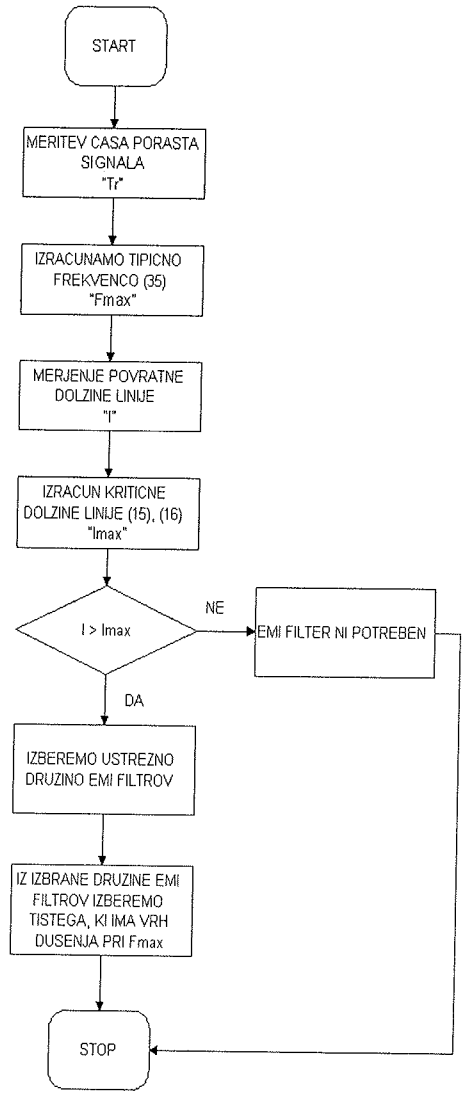
V poglavju 4 na sliki 6 vidimo FFT signala na izhodu iz 74HC245 integriranega vezja. V to signalno linijo smo vstavili EMI filter, ki smo ga izbrali s pomočjo zgoraj opisanega postopka. To je bil EMI filter NFW31SP506X1E4 firme Murata. Maksimalno dušenje ima pri približno 250 MHz.

Na sliki 8 vidimo, da se je nivo harmonika pri tipični frekvenci zmanjšal iz 28,64 dB (27,04 V) na 21,20 dB (11,48 V). Uporaba EMI filtra se je torej obrestovala.



Slika 8: FFT signala D/3/ za Muratinim EMI filtrom NFW31SP506X1E4 - izhod iz 74HC245 logike

Opisan postopek izbire EMI filtrov prikazuje slika 9.



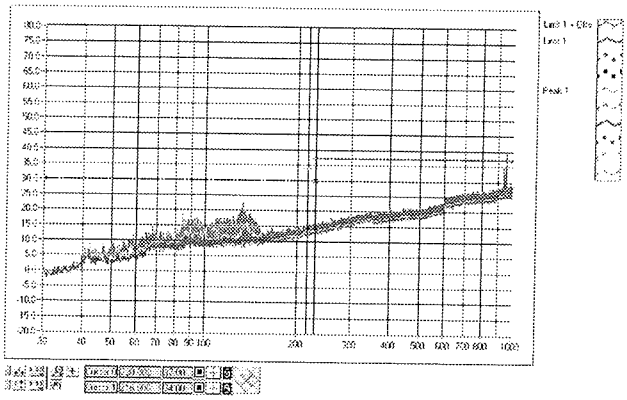
Slika 9: Diagram poteka izbire EMI filtra

8 Primer CPU modula

Na primeru CPU modula smo pokazali, da je metoda izbire EMI filtrov na osnovi meritve FFT signalov dejansko boljša od metode izbire na osnovi meritve sevanja celotnega CPU modula.

Najprej smo CPU modul opremili z EMI filtri, ki smo jih izbrali po metodi merjenja sevanja CPU modula. Pomerili smo sevanje tega modula. Rezultat meritve prikazuje višja krivulja na sliki 10. Sevanje je bilo sicer v dopustnih okvirih, vendar še vedno precej opazno.

Isti CPU modul smo opremili še z EMI filtri, ki smo jih izbrali po metodi meritve FFT signala. Pričakovali smo boljši rezultat, kot v prejšnjem primeru, saj smo problem sevanja reševali za vsako linijo posebej. Naša pričakovanja so se uresničila. Rezultat meritve prikazuje nižja - modra krivulja na sliki 10.



Slika 10:Sevanje CPU modula

Literatura

/1/ Dugald Campbell, Harald Kreidl, "Solving EMC Issues", Motorola, Workshop 36, EMV'03, Augsburg, 2003.

/2/ Dir. Frits J. K. Buesink, Thomson - CSF Signaal, Henglo, NI, "High Speed Digital Design Topics for Printed Circuit Boards", Workshop 24, EMV'01 Augsburg, 2001.

/3/ M. I. Montrose, "EMC and the printed circuit board: Design, Theory, and Layout Made Simple", 1. edition, IEEE Press Editorial Board, New York, 1999, vols 1, 3, 6, 7.

/4/ Dick Anderson, Lee Smith, Jeff Gruszynski, „S-Parameter Techniques for Faster, More Accurate Network Design", Hewlett-Packard Company, USA, 1996-1997.

/5/ S. Kazama, S. Shinohara, R. Sato, Evaluation of methods of measuring digital IC terminal output, EMC Research Laboratories Co., Ltd., Sendai, Japan, 2000, pp. 329-334.

/6/ Vladan Đ. Desnica, Ljiljana D. Živanov, Obrad S. Aleksić, Miloljub D. Luković and Miroslav D. Nimrihter, "Comparative Characteristics of Thick-Film Integrated LC Filters", IEEE transactions on instrumentation and measurement, vol.51, no. 4, august 2002, pp. 570-576.

/7/ muRata, "Noise Suppression by EMI Filtering: Basics of EMI Filters", No. TE04EA-1, 1998.

/8/ Šegula Matjaž, Podberšič Marko, "Reševanje EMC problematike", Kranj: ISKRAEMECO, d.d., 2002. 56 f., ilustr., graf. prikazi. [COBISS.SI-ID 512034160]

Marko Podberšič,
Ministrstvo za obrambo, Uprava RS za zaščito in reševanje, Kardeljeva ploščad 21, 1000 Ljubljana
Email: marko.podbersic@mors.si

Matjaž Šegula,
ISKRAEMECO d.d, Razvoj in raziskave,
Savska loka 4, 4000 Kranj
Email: matjaz@rd.iskraemeco.si

Vojko Matko,
Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko,
Smetanova ulica 17, 2000 Maribor
Email: vojko.matko@uni-mb.si

Prispelo (Arrived): 31.05.2003 *Sprejeto (Accepted): 26.08.2003*

UDK621.3:(53+54+621+66), ISSN0352-9045 Informacije MIDEM 33(2003)3, Ljubljana

ANALYSIS AND DESIGN OF COMBINED ELECTRONIC AND MICRO - MECHANICAL SYSTEM THROUGH MODELLING AND SIMULATION

[1]M. Atanasijević-Kunc, [1]V. Kunc, [2]J. Diaci, [1]J. Trontelj, [1]R. Karba

[1]Faculty for Electrical Engineering, Ljubljana, Slovenia
[2]Faculty of Mechanical Engineering, Ljubljana, Slovenia

Key words: micro-mechanical systems, smart sensors, modelling, simulation, control design

Abstract: In the paper an integrated micro-mechanical acceleration sensor is described through several stages, where open and closed loop dynamical properties are studied by modelling, simulation and animation in comparison with measurement data. The accurate and reliable model is not important only for realization of the system, but can be used also for development of the new generation of similar systems and for optimization of the existing devices. In this way a number of costly iterations of micro-mechanical prototypes can be greatly reduced and the design time and effort minimized.

Analiza in načrtovanje kombiniranega elektronskega in mikro-mehanskega sistema s pomočjo modeliranja in simulacije

Ključne besede: mikro-mehanični sistemi, pametni senzorji, modeliranje, simulacija, načrtovanje vodenja

Izvleček: V prispevku smo predstavili nekatere dinamične lastnosti mikro-mehaničnega sistema, ki deluje kot senzor pospeška. Omenjene lastnosti smo proučevali v večih korakih s pomočjo modeliranja, simulacije in animacije in sicer v primeru, ko sistem deluje kot odprtozančen, pa tudi pri realizaciji povratnozančnega vodenja. Rezultate modeliranja smo ovrednotili v primerjavi z meritvami na realnem sistemu. Tovrstni rezultati so pomembni tako v fazi samega načrtovanja sistema, pri optimiranju delovanja, uporabni pa so tudi za študij načrtovanja novih generacij podobnih sistemov, kar vse lahko pripomore k zmanjšanju stroškov in časa realizacije.

1. Introduction

Integrated micro-mechanical sensors are ideal for volume production of accurate, cheap and extremely small sensors for different physical quantities such as pressure, acceleration, rotation speed and many others. They comprise two main parts. The first part is the micro-mechanical device, which can be a self-standing element (in a joint package with the electronic) or an integrated element. The integrated micro- mechanical element as in our case is done during post-processing of the silicon wafer using technological procedures similar to the one used in microelectronic production. This usually limits the choice of material to silicon and silicon oxide and the geometry of the mechanical device is predominately two-dimensional realized using a limited number of stacked layers.

The second part is the electronic system. It is realized using standard micro-electronic production techniques and interacts with the mechanical part using capacitive position sensing principles that means it influences the mechanical part with electrostatic force.

To successfully join both parts into a measurement system the task of modelling the complete system is of vital importance. The stem level model must comprise the model of electronic part, the model of mechanical part and all the interactions between the two. In case of closed loop sensor realization also the controller functions must be integrated for correct dynamic interpretation.

2. System description

The mechanical part of the measurement system is realized as a poly-silicon cantilever, which is 490μm long, 211 μm wide and 0.9μm thick. On one side it is attached to the bulk silicon and the other side is free to move. The mechanical part is placed above the electronic part as is schematically presented in Fig. 1.

The electronic part comprises three electrodes made by top metalization layer forming a capacitance to the beam. The beam is electrically connected to the ground potential of the electronic system since such solution is most favorable from the mechanical point of view. So the electrodes of the electrical part are forming capacitors with the cantilever as a grounded electrodes. They are used to form distance measurement capacitors. In combination with extremely sensitive capacitance measurement electronic integrated in the electronic part these electrodes precisely detect the distance of the mechanical part to the sensing

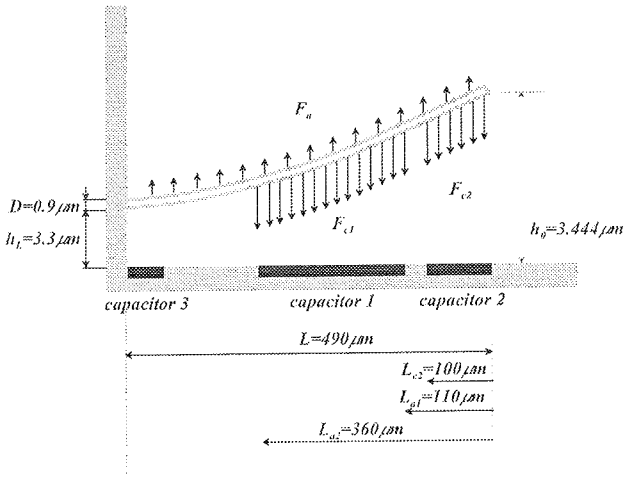


Fig. 1. Schematic system representation

electrodes. The electrode beneath the beam tip and the electrode in the middle of the beam are also used to exert electrostatic force to attract the beam (the actuation force).The third capacitor is used only to detect the starting beam distance from the silicon surface.

The described measurement system can operate in two modes, the open loop and closed loop mode. In the open loop mode the feedback part of the system is disabled and the electronic is only used to monitor the cantilever position and for testing the dynamic properties of the system itself. When the system is subjected to acceleration the corresponding force displaces the beam. The displacement magnitude is dependant on the beam stiffness and the dynamic is governed by the air damping. The displacement is detected by the beam position sensing electronic and conveyed to the system output.

Functional system properties can be significantly improved by the realization of the close loop. In this case the electrostatic actuation force counters the acceleration induced beam displacement. The resulting displacement is so much smaller and is dependant on the closed loop gain. The measure of the acceleration is in this case the amount of the actuation force applied. It is clear that such system is far more linear than the open loop one and its frequency behavior can be influenced by the closed loop gain.

3. System modelling

The model construction was done in two main phases. The first started during the initial stage of the project to help the design of the first prototypes of the measurement system. The construction of this initial model based on the theoretical model of the mechanical part of the system and on the results of the simulation of the electrical part, which was realized by SPICE simulator /1/.

After receiving the first prototypes of the micro-mechanical system a detailed evaluation phase followed. During this a number of measurement data enabled further model improvement. This enhanced model will serve for further system development and optimization.

The theoretical model is governed by the equilibrium of all the forces in the system. They can be described from the mechanical and electrical point of view.

It is obvious that for the mechanical part partial differential equations can be used for displacement description. Since our wish was the description of the system in time domain for simulation and animation purposes we decided to use for the mechanical part the finite element differential description /2/. For this the beam was divided into twenty segments each consisting of mass, spring and damping element, as is illustrated in Fig. 2. It seems that this number is a good compromise between the simulation accuracy and simulation time. Boundary conditions were satisfied from displacement and rotation angle calculations.

Here it is important to point out that damping is in our case highly nonlinear function of deflection and is heavily dependant on the air gap between the cantilever and the silicon surface. This can be explained by the fact that the distance between the silicon and the cantilever beam is extremely small. The consequence is that the air in the gap can not freely move /3,8/. So adapted air damping model was used as is described with the following equation:

$$f_i = \frac{\nu B^3 L \alpha \ (B/L)}{y_i^4} \tag{1}$$

where f_i denotes the damping of the *i*-th element, y_i is the distance of the *i*-th element from the surface, *L* and *B* are the length and width of the beam, n is viscosity and α is geometry dependant constant.

Model describing mechanical part has therefore 40-states (2 for each element) and is highly nonlinear.

As mentioned the cantilever beam is placed above the measurement electronic consisting of two main capacitor plates forming the capacitance to the beam as is illustrated in Fig. 3. Electrostatic force of each element of the corresponding capacitor to the corresponding part of the beam can be described with the following equation:

$$F_{ci} = \frac{1}{2} \varepsilon A_i \left(\frac{U_{in}}{y_i} \right)^2 \tag{2}$$

where ε is dielectric constant, A_i is the area of the *i*-th part of the capacitor and U_{in} is applied input voltage. It must be taken into account that the applied voltage on both capacitors was in open loop mode only up to 50% of the time due to the multiplexing of measurement and actuation. So for open loop purposes the equation (2) was replaced with:

$$F_{ci} = \frac{1}{8} \varepsilon A_i \left(\frac{U_{in}}{y_i} \right)^2 \tag{3}$$

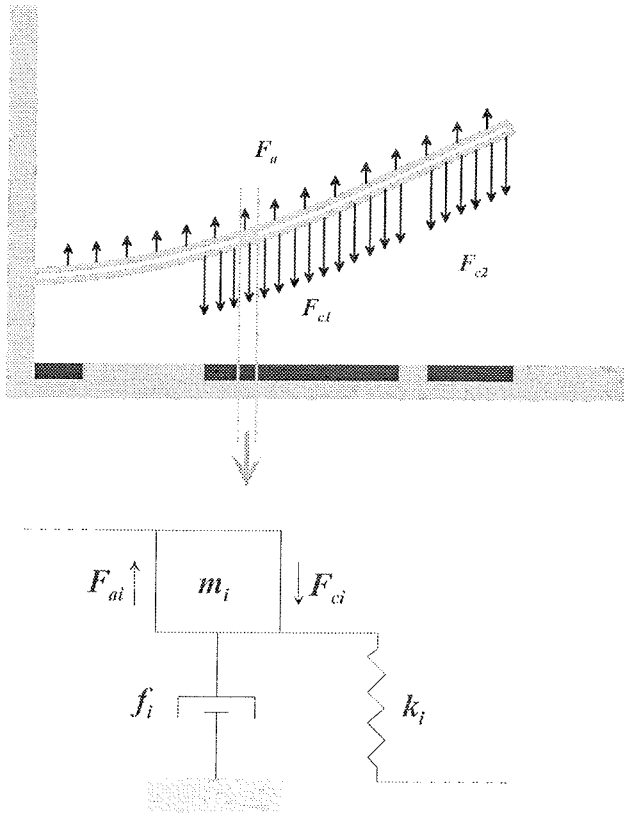


Fig. 2. Modelling of the mechanical part

as in our case only 25% of the time period was used for actuation purposes.

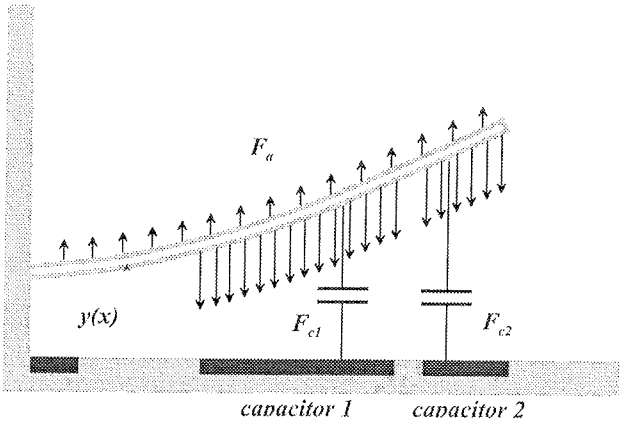


Fig. 3. Modelling of the electrical part

Deflection measurement property can be for each capacitor segment described with:

$$U_{ci} = K_{1i} \frac{1}{K_{2i} + K_{3i} \frac{1}{y_i}} \qquad (4)$$

where K_i are adjustable constants of electronic measurement system. In our case for measurement purposes only capacitor 1 was used and the output voltage was evaluated as the average value of all segments above the capacitor plate:

$$U_{out} = \frac{1}{n_j} \sum_{j=1}^{n_j} U_{cj} \qquad (5)$$

In the second phase of model construction the actual data gathered during the evaluation phase of the first prototypes was introduced to adjust the model parameters. The measurement results were divided into a static and dynamic data. The static data consisted of initial distance of the beam tip and core from the surface electrodes and the value of the capacitance at the beam tip. Also the static characteristic of the beam tip deflection as a function of actuation voltage was presented.

Data from a number of devices were entered as input data to the model simulation and were used to adapt the model. Once this was done and a reasonable matching between the simulation and measurement results was obtained the more demanding task of dynamic system model adjusting could begin.

The measurements of dynamic behavior of the prototype samples were done in both possible modes of operation. The open loop measurements consisted of time domain observation of the cantilever tip displacement resulting from step function changes in actuation voltage. The digital oscilloscope was used to track the beam tip movement as measured by the distance sensing electronic. System and model responses for different step changes are illustrated in Figs. 4 to 6. In all figures input voltage is presented first and followed by the beam tip displacement. In the third part measurement data are compared with model responses. All simulation results were obtained using Matlab/ Simulink program environment /4,7/.

From this it is possible to conclude that good open loop matching was obtained in different voltage ranges.

The closed loop measurements used the externally adjustable reference voltage as the input stimulus. This reference voltage is used in the normal system operation to position the beam tip at the required distance from the sensing electrode. The step-wise change of the reference voltage thus meant that the close loop system must change the actuation force to move the beam tip into the newly required position. The movement of cantilever tip was again monitored using a digital oscilloscope. Efficacy of the closed loop operation is illustrated in Fig. 7 where error signal is compared for the model response and measurement data.

4. Conclusion

In the paper modelling and simulation results are presented and compared with measurement data for the integrated micro-mechanical acceleration sensor. Presented work

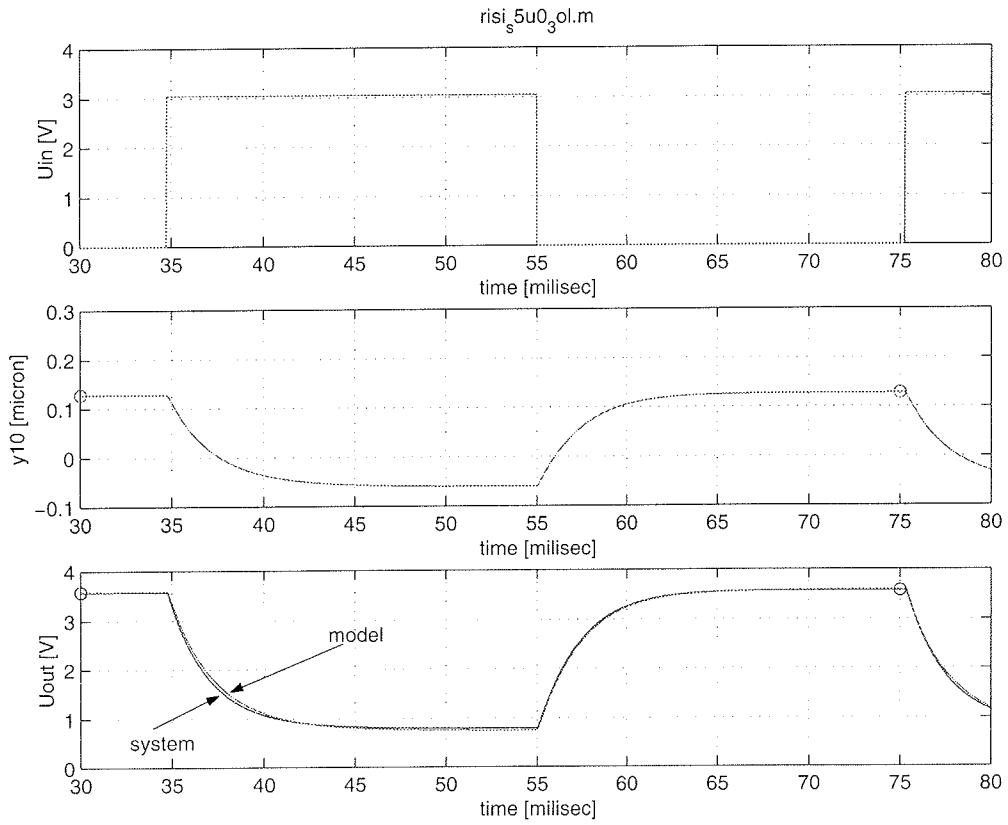


Fig. 4. Open loop system responses in comparison with measurement data (input voltage changes form 0 to 3 V)

is a res
chanic

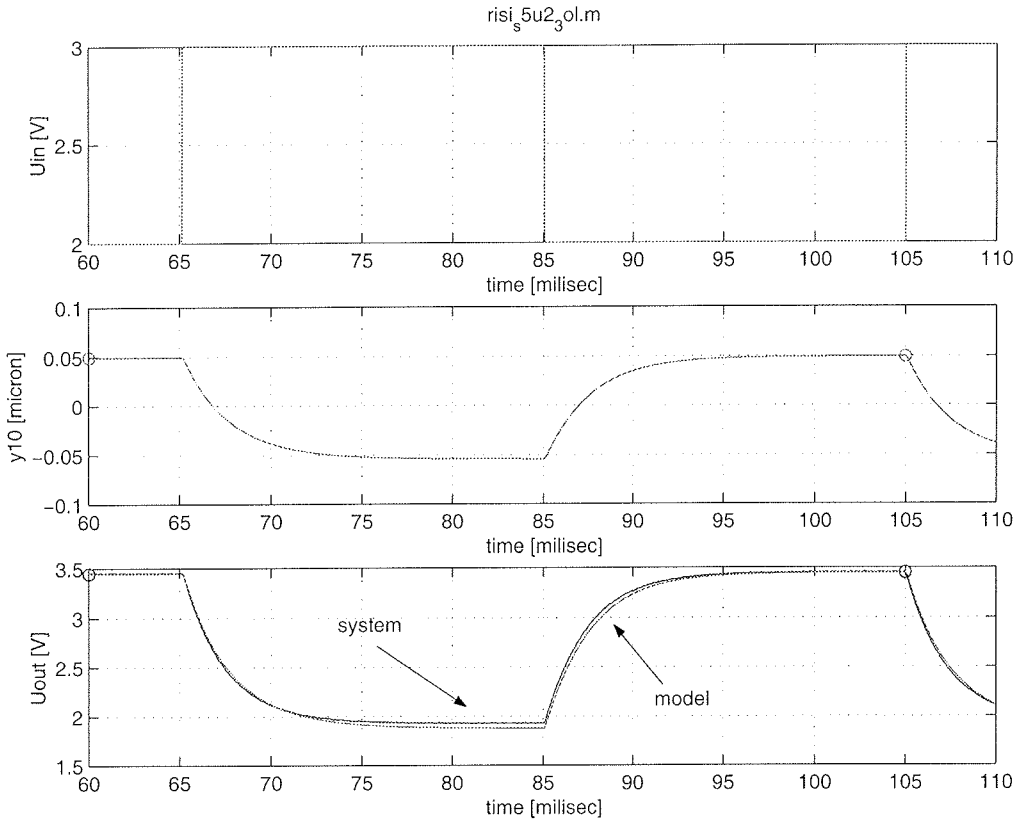


Fig. 5. Open loop system responses in comparison with measurement data (input voltage changes form 2 to 3 V)

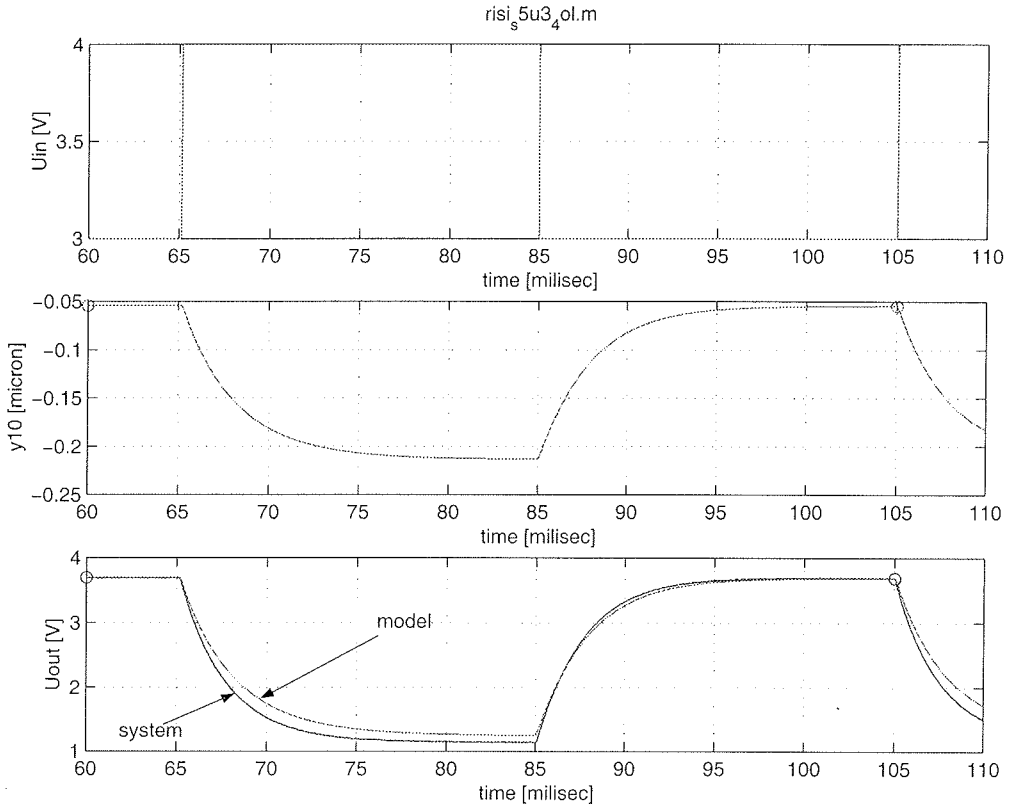


Fig. 6. Open loop system responses in comparison with measurement data (input voltage changes form 3 to 4 V)

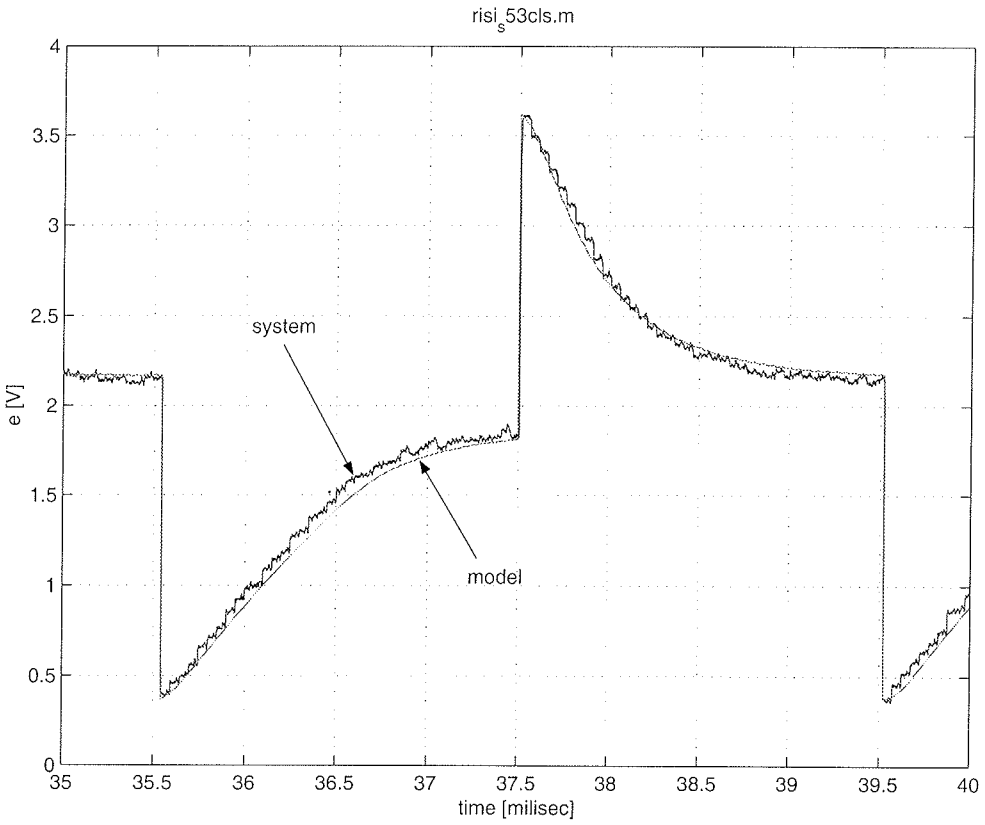


Fig. 7. Closed loop system response – error function - in comparison with measurement data

The starting model was based on theoretical equations and was further improved using measurement data taken from prototype devices. The model development was done in two major steps. The first step covered the static characteristic of the system resulting from the equilibrium of the electrical actuation force and cantilever spring force. The second step was the introduction of dynamic properties governed mainly by air damping mechanisms. When acceptable matching of open loop responses was obtained the model was tested also in the closed loop where the main controller functions are the adaptation of frequency range which can be satisfied by suitable chosen gain and of course linearization effect, needed because of highly nonlinear system properties.

The final results seems to have a high degree of compliance to the actual measurements and will be used for further optimization of the system, saving a lot of extremely costly experiments.

References

/1/ I. Zelinka, J. Diaci, V. Kunc, L. Trontelj, "Modeling and simulation of a microsystem with SPICE simulator", *Inf. MIDEM*, vol. 27, no. 1, pp. 16-22, 1997.

/2/ D. Matko, R. Karba, B. Zupančič, "Simulation and modeling of continuous systems, A case study approach", Prentice Hall, 1992.

/3/ R. L. Daugherty, J. B. Franzini, E. J. Finnemore, "Fluid Mechanics with Engineering Applications", McGraw - Hill, 8th ed., Singapore, 1985.

/4/ "*MATLAB*", The Language of Technical Computing, Version 5, The MathWorks Inc., 1999.

/5/ E. P. Popov, "Engineering Mechanics of Solids", Prentice Hall, New Jersey, 1990.

/6/ L. Ristic, (ed.), "*Sensor Technology and Devices*", Artech House, 1994.

/7/ "*Simulink*", User's Guide, The MathWorks Inc. 1999.

/8/ J. B. Starr, "Squeeze-film Damping in Solid-State Accelerometers", *IEEE Solid-State Sensor and Actuator Workshop*, Hilton Head Island, pp. 44-47, June 1990.

/9/ M. Atanasijević-Kunc, Maja, V. Kunc, J. Diaci, R. Karba, "Modelling and analysis of combined electronic and micro - mechanical system", I. Troch, F. Breitenecker, (ed.), 4th MATHMOD Vienna, 4th IMACS Symposium on Mathematical Modelling, *Proceedings*, Vienna, 2003.

Asst. Prof. Dr. Maja Atanasijević-Kunc
e-mail: maja.atanasijevic@fe.uni-lj.si
Tel.: +386 (01) 4768 314

Dr. Vinko Kunc
e-mail: vinko@kalvarija.fe.uni-lj.si
Tel.: +386 (01) 4768 337

Prof. Dr. Janez Trontelj
e-mail: janez@kalvarija.fe.uni-lj.si
Tel.: +386 (01) 4768 333

Prof. Dr. Rihard Karba
e-mail: rihard.karba@fe.uni-lj.si
Tel.: +386 (01) 4768 251

Faculty for Electrical Engineering
Tržaška 25, Ljubljana, Slovenia

Asst. Prof. Dr. Janez Diaci
Faculty of Mechanical Engineering,
Aškrčeva 25, Ljubljana, Slovenia
e-mail: janez.diaci@fs.uni-lj.si
Tel.: +386 (01) 4771 615

Prispelo (Arrived): 7.04.2003 *Sprejeto (Accepted): 26.08.2003*

AN EVOLUTIONARY APPROACH TO CHIP DESIGN: AN EMPIRICAL EVALUATION

Gregor Papa

"Jožef Stefan" Institute, Ljubljana, Slovenia

Key words: chip design, area and time optimization, low-power circuits, evolutionary algorithms

Abstract: This paper presents a new method with an evolutionary approach to some parts of integrated-circuit (IC) design. This study, however, is focused on application-specific integrated circuits (ASICs), which need an even more sophisticated design (in terms of size, speed and low-power) because of their specific uses.

Optimally scheduled operations are not necessarily optimally allocated to units. To enable optimal allocation we need to consider some allocation criteria while the scheduling is being done. Therefore, algorithms with concurrent scheduling and allocation produce the best results. It is obvious that we have to deal with a trade-off between the quality of the solution and its design time. The main part of the paper is a presentation of an improved method of the evolutionary search for the optimal design of ICs. The evolutionary approach considers scheduling and allocation constraints and ensures a globally optimal solution in a reasonable time.

The evaluation of our method shows that the evolutionary method is able to find a solution that is more appropriate in terms of all the considered and important objectives than is the case when working with classical deterministic methods.

Evolucijski pristop pri načrtovanju čipov: empirično ovrednotenje

Ključne besede: načrtovanje čipov, energijsko varčna vezja, optimizacija velikosti in časa, evolucijski algoritmi

Izvleček: V delu je predstavljena metoda sočasnega razvrščanja operacij in dodeljevanja enot, ki temelji na evolucijskem načinu in je uporabna v postopku načrtovanja digitalnih integriranih vezij. Delo obravnava predvsem načrtovanje namenskih vezij, ki potrebujejo zaradi svoje specifičnosti toliko bolj dodelano strukturo, tako s stališča velikosti, kakor tudi s stališča hitrosti delovanja in majhne porabe energije.

Optimalno razvrščenih operacij v splošnem ni mogoče tudi optimalno dodeliti posameznim enotam. Da bi lahko operacije dodelili optimalno, je treba pravila za dodeljevanje upoštevati že med razvrščanjem. Problema smo se lotili z evolucijsko tehniko, ki je pogosto uporabljena v metodah za iskanje rešitev na najrazličnejših področjih. Razvit je bil evolucijski algoritem, ki upošteva razvrščevalne in dodeljevalne zahteve, omogoča kratek načrtovalni čas ter globalno optimalne rešitve. Algoritem je bil tudi računalniško realiziran ter uporabljen pri množici preizkusnih vezij. Vezja so bila izbrana glede na pogostost pojavljanja v sorodni literaturi, in sicer vezja različnih velikosti z različnim številom tipov operacij.

Obravnava rezultatov je pokazala, da je opisani evolucijski algoritem, v primerjavi s klasičnimi determinističnimi metodami, sposoben poiskati rešitev, ki je v splošnem ugodnejša s stališča vseh obravnavanih in pomembnih parametrov.

1 Introduction

Whenever a new integrated circuit (IC) is designed the problem of selecting the best register-transfer level (RTL) specification has to be faced. And as circuits get bigger, so too does the problem: more combinations have to be examined before the optimal combination is found. Therefore, automatic circuit optimization is required, as this speeds up the whole design process and eliminates some of the errors.

High-level synthesis /2/, /4/ is an automatic design process that transforms the initial behavioral description into the final specification of the RTL. The process consists of the following tasks: compilation, transformation, scheduling, allocation and binding. Of these, the operation scheduling and the resource allocation are the most important subtasks of the high-level synthesis because they are at the core of the design and crucially influence both the design and the final layout. Due to the interdependence of these two tasks, the solution of one task depends on an estimation of the solution of the other task, which is not solved yet. The scheduling of the operation into different control steps therefore affects the allocation of operations to different units. The interaction of these two tasks presents formidable obstacles to the goal of optimization /1/. There are, however, some approaches to concurrent solving, but their solutions, to some extent, are less than optimal.

The evolutionary technique is used in various search methods for a range of different optimization areas /8/. Its undetermined approach, i.e., a probabilistic approach, reduces its popularity, but at the same time gives it an advantage when it comes to multicriteria problems and problems with more local optima. The genetic algorithm (GA), as a frequent implementation of evolutionary techniques, is an optimization method based on the mechanism of evolution and natural genetics.

2 Definitions

An IC described with a hardware description language, e.g. VHDL, can be presented as a control/data-flow graph (CDFG) /4/, and in our case it is enough to consider only the data-flow part, i.e., the data-flow graph (DFG). Each node *i* represents the operation to be executed, and the type of node determines the type of operation. The edges *e* represent dependencies between operations. An edge e_{ij}, between nodes *i* and *j*, represents the data produced by the node *i* and used by the node *j*. All edges are unidirectional.

In the final design there is a group of resources that define the implementation. Here we have different functional units (FUs) FU_i ($i=1..N$), where *N* is the number of different types of functional units. There are also storage (registers *Reg*) and connection (buses *Bus*) units, while *T* represents the execution time of the DFG. Functional units (adders, multipliers, ...) perform transformations on data values, connection units transport values from one unit to another, while storage units preserve those values over time.

The parameters are calculated as follows:

- the number of FU_i is the highest number of the *i*-th functional unit needed in a separate control step;
- the number of registers *Reg* is the highest number of variables needed in a separate control step. We consider variables that are needed by the functional unit as input data, variables that are returned as output data and variables that are not used at the moment but will be used in some of the later control steps or must be available until the end of the execution of all operations. Of course, more functional units can use the same data from the same register in a control step:
- the number of buses *Bus* is the highest number of data transmissions (into or from the functional units) in a separate moment;
- the execution time *T* is the time needed to execute all the operations of the schedule;
- the weights *w* are the weights of the IC parameters to be considered in the IC quality evaluation cost function. They depend on the ratio of different units' (functional, storage) sizes.

3 Concurrency

Optimally scheduled operations are not necessarily optimally allocated to units. To be also optimally allocated some allocation constraints should be considered during the scheduling. Therefore, algorithms that perform concurrent scheduling and allocation are better in terms of the optimality of the solutions. These algorithms are, however, very time consuming. Therefore, we have to deal with a trade-off between the quality of the final solution and the length of the design-time for it. There are some approaches to concurrent solving, but their solutions, to some extent, are less than optimal.

When tasks are performed separately (Figure 1a) the solution is not necessarily optimal; it is better to use an approach with iterative repetition of the scheduling and allocation (Figure 1b). Here, though, the problem of the next operation/unit to be changed appears, since the order of changes can influence the final solution. It is a similar situation with the approach that involves partitioning of the operations into small groups, within which there is an iterative repetition of the scheduling and the allocation (Figure 1c). Since there are fewer operations in the group there is no problem with the order of changes, but there is a problem with the appropriate partitioning of the operations. Obviously, the best approach is the one with purely concurrent scheduling and allocation (Figure 1d), where the iterative-refinement order does not influence the quality of the solution /7/. Concurrency is achieved through the use of algorithms that do not depend on the order of transformations. Therefore, there is no influence of the changed start time on the allocated unit, nor is there any influence of the allocated unit on the start time. When all the transformations are made, then the appropriateness of the changes is checked.

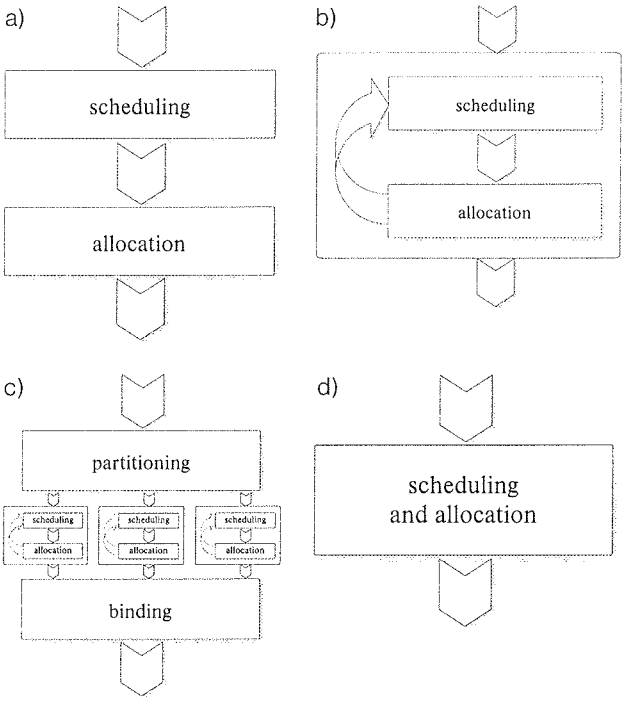


Figure 1: Scheduling and allocation concurrency

4 The ECSA algorithm

The facts presented in the introduction paragraphs and the promising results of different evaluations /8/, /9/ led us to the ECSA (Evolutionary Concurrent Scheduling and Allocation) design approach /7/. This approach considers scheduling and allocation constraints, allows a short design time and can find globally optimal solutions. The input description of the circuit is transformed into two basic (initial) schedules, obtained with the ASAP and ALAP algo-

rithms. The FUs used in the first case are those that are the fastest for each operation, and in the second case are those that are the slowest for each operation. These two schedules present some kind of boundary solutions, since all the other solutions are executed in between the time limits defined by these two schedules. In other words, no other solution can be faster or slower, irrespective of the combinations of used units.

Each solution has to be properly encoded (into the chromosome), i.e., each operation's start time and functional unit have to exist in the chromosome. The initial population is built upon the two initial solutions, which are multiplied to form the population with the so-called boundary solutions. The optimal solution has to be somewhere in between the boundaries, therefore genetic operators (crossover, mutation, variation) transform those encoded solutions. With the transformations their start times and allocated functional units are changed. Within that, also the multicycling solutions are supported by the approach. The final solution obtained using the genetic operators is also influenced by the simulated annealing algorithm, which improves the solution if it stopped somewhere near the globally optimal point.

4.1 Encoding

The performance of the algorithm depends on the proper encoding. In the ECSA algorithm integer encoding is used, i.e., in the chromosome string are the numbers that represent the starting time of each operation and the allocated unit for each operation, where the position in the string depends on the order of the operations in the input IC description. This means that the chromosome consists of pairs of time/unit information for each operation. And the genetic operators can influence both parts of that information, either together or separately.

Integer encoding was chosen, since it does not need any transformation (into binary values) at the beginning and at the end (back into decimal values). Also, the used implementation of genetic operators can check the changed values instantly, without any transformation. The correctness of the transformation, i.e., the crossover, the mutation or the variation, can therefore be checked within the function itself.

4.2 Cost function

One of the most important parts of the algorithm is its cost function. Our algorithm is multi-objective /3/, which means it takes control over more criteria or objectives. The cost function, represented by Eqn. 1, considers the number of functional units, the number of registers, the number of buses and the execution time of all the operations in the DFG.

$$
\begin{aligned}
Cost &= \sqrt[2]{\sum_{i=1}^{n}(cost_{FU_i})^2 + cost_{Reg}^2 + cost_{Bus}^2 + cost_T^2} \\
cost_{FU_i} &= w_{FU_i} \cdot FU_i \\
cost_{Reg} &= w_{Reg} \cdot Reg \\
cost_{Bus} &= w_{Bus} \cdot Bus \\
cost_T &= w_T \cdot T
\end{aligned}
\qquad (1)
$$

To obtain the cost of a certain DFG, the algorithm has to evaluate the required number of resources. In contrast to the other multi-objective functions that give more than one final solution, this one already includes the decision-making part, i.e., it chooses one solution from all the solutions on the Pareto front. The chosen solution has the shortest distance to the origin, where the origin represents the ideal, costless, solution and the axis represents the considered objectives.

The main weakness /3/ of this approach is the difficulty in determining the appropriate weights when there is not enough information about the problem. Since we are aware of the problem specifics and we know the cost weights of the resources being used, this weakness is not so significant.

4.3 Evolutionary operators

In each iteration (or generation) of the algorithm there are four genetic operators that transform the chromosome. They consider data dependencies and the given library of available functional units. Each time after the genetic operators transform the chromosome, the chromosome is checked to see if it meets all the constraints by considering data dependencies and unit types. Besides the basic implementation of the operators, we also applied the independent operators, which do not need any parameter value to be set in advance: they depend only on the progress of the search and on the size of the problem to be solved.

4.3.1 Basic operators

Selection. Based on the cost-function values the worst solutions are aborted in the selection step. And to ensure that the size of the population remains the same, these solutions are replaced with the best solutions. This ensures that the best solutions of a given generation are involved in the creation of the next generation (elitism).

Crossover. In a crossover task, two approaches are used, with each task expressing the dominancy of the characteristics. After two crossover points are determined, in the first case the unit information is changed between the two chromosomes and the start times are adapted, and in the second case the start times are changed and a suitable unit is allocated. So the dominancy is expressed either in functional units or the start times of operations.

Mutation. We also have two similar approaches for transforming the chromosome. In both cases the starting time

is changed. Either it is moved to later control steps, with the use of faster functional units, or it is moved to earlier control steps, if data dependencies allow this, with slower units.

Variation. After two operations are selected, and when they are of the same type, e.g. additions, their functional units are switched. If needed, their start times are also updated.

4.3.2 Independent operators

The advantage of the independent GA approach /12/ is that there is no need to preset some working parameters, e.g. the number of generations, the population size, and the probabilities of crossover, mutation and variation. These parameters are set automatically during the optimization phase, depending on the progress and the speed of the optimization.

Setup. If the chromosome that presents a solution is large, then the population size also has to be large enough to ensure that a lot of different chromosomes will be involved in a search. The population size therefore depends on the size of the chromosome or the complexity of the problem.

Crossover. Considering four candidates - two parents and their two offspring - only the first and the third, rated according to their fitness, pass to the next generation. This forces at least one of the offspring to be passed to the next generation in addition to the best candidate. Otherwise the offspring have only a small influence on new generations, since the crossing of two good parents probably produces offspring that are not so good; however, they might be good after a few more transformations.

Mutation. Chromosomes with low fitness are mostly exposed to mutation. Each bit in the chromosome is mutated if that position of the chromosome is of the same value in the majority of chromosomes in the population. This is the way to change the bad characteristics in "poorly fitted" chromosomes and to redirect the search to another direction. In the case of "well-fitted" chromosomes, bits are mutated if the value of the bit differs from majority of bits in other good chromosomes at the same position. This ensures faster convergence in the final stages of the optimization.

Variation. The interchange of the values of two bits, as described for the basic operators, is performed if the frequency of the value in that position in the population of one bit is high and the frequency of another bit is low.

4.3.3 Simulated annealing

When the GA finishes its work and the most appropriate solution is found, that solution is additionally influenced by the simulated annealing algorithm. It checks whether the search for the optimal solution stopped in the vicinity of the global optimum, and it changes the solution if needed.

5 Evaluation

The appropriateness of the proposed approach was tested by a computer implementation of the ECSA algorithm, which was used with test-bench ICs. The ICs used for the evaluation were chosen based on their appearance in the literature and similar studies. They differ in terms of size and the number of operation types.

5.1 Test-bench circuits

5.1.1 Differential equation

The relatively small circuit of differential equation /14/ has only 11 operations, but 4 different operation types (6 multiplications, 2 additions, 2 subtractions, 1 comparison). This circuit is useful when testing libraries with different implementations of the same operation types.

5.1.2 Elliptic filter

This filter /6/ consists of 34 operations, but only two operation types: 26 additions and 8 multiplications. The circuit is suitable for comparison, due to its size and operation dependencies, since they form two independent, similar critical paths, both influencing the circuit delay.

5.1.3 Bandpass filter

One of the implementations of the bandpass filter /5/ is the circuit used for our evaluation. It consists of 29 operations: 11 multiplications, 10 additions and 8 subtractions. Due to data dependencies, almost all the operations influence the circuit delay.

5.1.4 Least-mean-square filter

This filter for signal adaptation (noise reduction) is based on the least-mean-square method /2/. It consists of 47 operations: 24 multiplications and 23 additions. This test-bench circuit is useful due to its size and unique data dependencies.

5.2 Functional units

For an easier and more realistic comparison of different algorithms when testing the size and delay of implemented circuits we made a library of different functional units, which differ in terms of their sizes and delays. Table 1 shows the sizes and delays of various implementations of the arithmetic logic operations. Here, different types of logic were used to make the units with different delays and sizes. The values are based on an analysis of data on circuits and their complexities /11/. The number of gate transitions defines the delay, while the overall number of gates needed to implement the unit defines its size. These values are just for orientation, since the real numbers depend on the chosen technology /11/. The delays presented in Table 1 are relative, e.g., normalized to the fastest functional unit among all the operations. Most of the units are multifunctional, i.e., they can perform different types of operation.

Table 1. Technical characteristics of the functional units

functional unit {operations}	delay [No. of steps]	No. of gates
FE1 {+, -}	6, 6	370
FE2 {+, -}	1, 1	665
FE3 {<}	6	353
FE4 {+, -, <}	1, 1, 1	696
FE5 {x}	4	3040
FE6 {x}	2	7296
FE7 {÷}	21	3040
FE8 {÷}	9	7296
FE9 {x,÷}	4, 21	3344
FE10 {x,÷}	2, 9	8025
FE11 {+, -, <, x, ÷}	1, 1, 1, 4, 21	3692
FE12 {+, -, <, x, ÷}	1, 1, 1, 2, 9	8373

5.3 Parameters

By considering 18750 different schedules of each circuit with the ECSA algorithm and 3125 different combinations of the parameters, we statistically compared (using the procedure described in /10/) the results according to their cost function (Eq. 1). To ensure that most solutions were time-constrained, i.e., executed in shortest possible time, the weight w_T was set to an extremely high value.

As presented in *Table 2*, the high-quality solutions are mostly obtained with the following values of the parameters: probability of crossover, 0.7; probability of mutation, 0.04; and probability of variation, 0.03. In addition, considering the sizes of the circuits, the number of generations and the population size should be set to 3-times and 3.5-times the size of the circuit, respectively.

Table 2. Optimal values of the parameters for different testbench circuits

	Differential equation	Fifth-order elliptic filter	Bandpass filter	Least-mean-square filter	Average optimal values
number of generations	40	100	90	130	3 x DFG size
population size	55	120	110	160	3.5 x DFG size
probability of crossover	0.8	0.6	0.7	0.6	0.7
probability of mutation	0.04	0.05	0.04	0.05	0.04
probability of variation	0.04	0.02	0.02	0.05	0.03

The values of the parameters in this combination are referred to as the optimal values. These optimal values are determined on the basis of the percentage of solutions with certain parameters from among the good solutions. A parameter value that is to be considered as optimal should have at least a 25% share of the high-quality solutions, as well as having a less than 10% share of the low-quality solutions.

The ECSA algorithm was used with the values of the parameters as presented in *Table 2*. Other parameters needed to run the FDS and ECSA algorithms and the cost function depended on the sizes of the FUs.

5.4 Results

The ECSA algorithm was evaluated by a comparison with nearly optimal /13/ force-directed scheduling (FDS) /12/. FDS tries to optimally schedule the DFG considering a uniform distribution the operations of the same type over the available control steps.

Table 3 presents the results of the following evaluations: FDS with fast units, FDS with slow units, and ECSA with basic and independent genetic operators. There are two types of DFGs for each circuit. The first, or plain, is an ordinary data-flow graph with nodes that represent operations, as described in similar studies; and the second, or improved, considers the input variables (start registers) via some additional nodes to ensure a more accurate estimation of the registers and the buses needed to implement the circuit.

5.4.1 Differential equation

Because of the small circuit size there is no improvement in the solutions obtained with the ECSA algorithm (either basic or independent) when considering an ordinary DFG - all the solutions are of a larger size. But when we consider the start registers (input variables) there are some ECSA solutions with a slightly larger size and a smaller number of buses.

5.4.2 Fifth-order elliptic filter

The evolutionary method with a basic approach found a smaller circuit with a smaller number of buses and a slightly longer execution time for the ordinary DFG, while the independent approach could not find any improved solution. When dealing with the improved DFG, both approaches (basic and independent) found considerably smaller circuits with a slight increase in the execution time, while the independent approach also found the solution with a substantial decrease in the required number of registers and buses.

5.4.3 Bandpass filter

Both ECSA methods found, when dealing with the ordinary DFG, the solutions with a smaller number of registers and buses; the basic approach also found the smaller circuit, but with a slightly longer execution time. When dealing with the improved DFG, both approaches found the solutions with the same circuit size and execution time as the comparable FDS solution, but the required number of registers and buses was considerably smaller for the ECSA solutions.

5.4.4 Least-mean-square filter

At the expense of a small increase in the delay, the basic ECSA was able to decrease the size and lower the number of registers and buses of the ordinary DFG; but the independent ECSA was not able to improve any parameter. When dealing with the improved DFG, the basic ECSA was able to keep the initial delay, to decrease the circuit size and to lower the number of required registers and buses. The independent ECSA was only able to decrease the number of buses while increasing the circuit size.

Table 3. The evaluation results of the ECSA algorithm with different test-bench ICs

algorithm	functional units	size	registers	buses	delay	runtime [s]
differential equation						
FDS-fast	1xFE2 + 1xFE4 + 3xFE6	23249	17	6	6	0.01
FDS-slow	2xFE1 + 1xFE3 + 2xFE5	7173	18	6	20	0.01
ECSA-basic	2xFE2 + 1xFE4 + 3xFE6	23914	18	8	6	0.11
ECSA-independent	2xFE2 + 1xFE4 + 3xFE6	23914	17	6	6	0.09
differential equation with start registers						
FDS-fast	1xFE2 + 1xFE4 + 3xFE6	23249	10	9	6	0.01
FDS-slow	2xFE1 + 1xFE3 + 2xFE5	7173	11	9	20	0.01
ECSA-basic	2xFE2 + 1xFE4 + 4xFE6	31210	10	7	6	0.15
ECSA-independent	2xFE2 + 1xFE4 + 3xFE6	23914	10	7	6	0.35
fifth-order elliptic filter						
FDS-fast	3xFE2 + 3xFE6	23883	26	8	17	0.02
FDS-slow	5xFE1 + 3xFE5	10970	30	6	78	0.03
ECSA-basic	2xFE2 + 1xFE5 + 2xFE6	18962	29	4	21	3.80
ECSA-independent	4xFE2 + 4xFE6	31844	30	8	17	1.60
fifth-order elliptic filter with start registers						
FDS-fast	3xFE2 + 3xFE6	23883	21	16	17	0.02
FDS-slow	5xFE1 + 3xFE5	10970	25	16	78	0.04
ECSA-basic	2xFE2 + 1xFE5 + 2xFE6	18962	24	16	19	4.80
ECSA-independent	2xFE2 + 2xFE6	15922	18	9	21	3.40
bandpass filter						
FDS-fast	3xFE2 + 4xFE6	31179	34	10	10	0.01
FDS-slow	4xFE1 + 3xFE5	10600	35	8	44	0.04
ECSA-basic	3xFE2 + 3xFE6	23883	33	8	11	1.70
ECSA-independent	3xFE2 + 1xFE5 + 4xFE6	34219	33	8	10	1.30
bandpass filter with start registers						
FDS-fast	3xFE2 + 4xFE6	31179	25	23	10	0.02
FDS-slow	4xFE1 + 3xFE5	10600	26	23	44	0.04
ECSA-basic	3xFE2 + 4xFE6	31179	23	19	10	2.40
ECSA-independent	3xFE2 + 4xFE6	31179	23	19	10	2.90
least-mean-square filter						
FDS-fast	3xFE2 + 6xFE6	45771	68	12	13	0.40
FDS-slow	3xFE1 + 4xFE5	13270	72	8	70	2.79
ECSA-basic	3xFE2 + 6xFE5 + 3xFE6	42123	67	10	14	6.30
ECSA-independent	9xFE2 + 6xFE5 + 9xFE6	89889	69	30	15	8.05
least-mean-square filter with start registers						
FDS-fast	3xFE2 + 6xFE6	45771	33	29	13	0.48
FDS-slow	3xFE1 + 4xFE5	13270	37	27	70	3.52
ECSA-basic	4xFE2 + 2xFE5 + 5xFE6	45220	32	25	13	9.20
ECSA-independent	5xFE2 + 3xFE5 + 7xFE6	63517	33	25	13	12.30

6 Conclusions

Optimally scheduled operations are not necessarily optimally allocated to functional units. To enable optimal allocation we need to consider some allocation criteria while the scheduling is being done. This paper describes such an evolutionary approach that considers scheduling and allocation constraints and ensures a globally optimal solution in a reasonable time. To evaluate our method we built an algorithm and implemented it with a computer. It was used with a group of test-bench ICs. These circuits were chosen because the same type were used in similar studies. They differ in terms of their size and the number of operation types.

It turned out that the evolutionary method (either basic or independent) is able to find a solution that is more appropriate in terms of all the considered and important parameters than is the case when working with classical deterministic methods. There are slightly longer runtimes when the ECSA algorithm is used. But considering the speed (a few seconds) and the computational dependence, where the runtimes for larger circuits increase enormously (exponentially) when the FDS algorithm is used, we can conclude that small and large circuits can be designed and optimized with the use of the proposed evolution-based algorithm, which exhibits a linear increase in the design time with an increase in circuit size.

References

/1/ J. R. Armstrong, F. G. Gray, VHDL Design: Representation and Synthesis, Prentice Hall, Upper Saddle River, 2000.

/2/ J. Benesty, P. Duhamel, A Fast Exact Least Square Adaptive Algorithm, IEEE Transactions on Signal Processing 40, 1992, pp. 2904-2920.

/3/ C. A. Coello Coello, A Comprehensive Survey of Evolutionary-Based Multiobjective Optimization Techniques, Knowledge and Information Systems, Vol. 1, No. 3, pp. 269-308, August 1999.

/4/ D. Gajski, N. Dutt, A. Wu, S. Lin, High-Level Synthesis: Introduction to Chip and System Design, Norwell, Massachusetts, Kluwer Academic Publishers, 1992.

/5/ G. W. Grewal, T. C. Wilson, An Enhanced Genetic Algorithm for Solving the High-Level Synthesis Problems of Scheduling, Allocation, and Binding, Intl. Journal of Computational Intelligence and Applications. 1, 2001, pp. 91-110.

/6/ T. Kung, H. J. Whitehouse, T. Kailath, VLSI and Modern Signal Processing, Prentice Hall, 1985.

/7/ G. Papa, Concurrent operation scheduling and unit allocation with an evolutionary technique in the process of integrated-circuit design, Ph.D. Thesis, Faculty of Electrical Engineering, University of Ljubljana, Ljubljana, 2002.

/8/ G. Papa, B. Koroušić-Seljak, B. Benedičič, T. Kmecl, Universal motor efficiency improvement using evolutionary optimization, accepted for publication in IEEE Transactions on Industrial Electronics.

/9/ G. Papa, J. Šilc, Automatic Large-Scale Integrated Circuit Synthesis Using Allocation-Based Scheduling Algorithm, Microprocessors and Microsystems 26, 2002, pp. 139-147.

/10/ G. Papa, J. Šilc, Evolutionary Synthesis Algorithm - Genetic Operators Tuning, In: A. Grmela, N. Mastorakis (ed.) Advances in Intelligent Systems, Fuzzy Systems, Evolutionary Computation, WSEAS Press, 2002, pp. 256-261.

/11/ B. Parhami, Computer Arithmetic: Algorithms and Hardware Designs, Oxford University Press, New York, 2000.

/12/ P. G. Paulin, J. P. Knight, Force-directed Scheduling in Automatic Data Path Synthesis, Proc. 24th ACM/IEEE Design Automation Conference, Miami, USA, June 1987, pp. 195-202.

/13/ P. G. Paulin, J. P. Knight, Scheduling and Binding Algorithms for High-Level Synthesis, Proc. 26th ACM/IEEE Design Automation Conference, Las Vegas, NE, pp. 1-6, June 1989.

/14/ P. G. Paulin, J. P. Knight, E. F. Girczyc, HAL: A Multiparadigm Approach to Automatic Data Path Synthesis, Proc. 23rd ACM/IEEE Design Automation Conference, Las Vegas, USA, June 1986, pp. 263-270.

Gregor Papa
Institut "Jožef Stefan"
Computer Systems Department,
Jamova c. 39, 1000 Ljubljana, Slovenia
tel. +386 1 4773 514
fax. +386 1 4773 882
Email: gregor.papa@ijs.si

Prispelo (Arrived): 13.05.2003 *Sprejeto (Accepted): 26.08.2003*

HEURISTIC APPROACH TO CIRCUIT SIZING PROBLEM

Janez Puhan, Árpad Bűrmen and Tadej Tuma

University of Ljubljana, Faculty of Electrical Engineering, Ljubljana, Slovenia

Key words: computer aided design, integrated circuits, optimization algorithms, circuit sizing.

Abstract: Circuit sizing problem in application specific analog integrated circuit design is in most cases limited to setting MOSFET channel widths and lengths. It is usually performed by hand by an experienced human designer. As the circuit sizing is an optimization process by its nature, optimization methods could be used. They always lead to one of the minima of the cost function while eventual other minima stay unknown. To reveal different cost minima an optimisation process composed from many individual optimisation runs is proposed. Individual runs are started from various initial points in the parameter space. A particular initial point is determined by a heuristic method which maximises the probability of finding a new cost function minimum in the next run. The optimization process is demonstrated on several real operating amplifier designs.

Heuristični pristop k določevanju elementov v integriranih vezjih

Ključne besede: računalniško podprto načrtovanje, integrirana vezja, optimizacijski algoritmi, določitev elementov.

Izvleček: Določitev dimenzij polprevodniških komponent v analognem integriranem vezju se največkrat prevede na določevanje dolžin in širin kanalov MOSFET-ov. To delo navadno opravi izkušen načrtovalec. Ker je celoten proces določevanja dimenzij po svoji naravi optimizacijski postopek, bi lahko v ta namen uporabili optimizacijske metode. Le-te vedno vodijo k enemu izmed minimumov kriterijske funkcije, medtem ko morebitni ostali minimumi ostanejo skriti. V članku predlagamo optimizacijski proces, sestavljen iz več posameznih optimizacijskih tekov, katerih namen je najti več različnih minimumov kriterijske funkcije. Posamezni teki so sproženi iz različnih začetnih točk v parameterskem prostoru. Začetne točke določimo s pomočjo heuristične metode, ki maksimizira verjetnost odkritja novega minimuma v naslednjem teku. Celoten optimizacijski proces je predstavljen tudi na realnih primerih integriranih operacijskih ojačevalnikov.

1 Introduction

Creating a good analogue integrated circuit (or analogue part in a mixed circuit) design is still a hard task, which usually requires senior designer knowledge and skills. There are no predefined libraries of standard cells and networks as in the digital world. Therefore the design of an analogue circuit consisting of a few transistors can be more time consuming than designing a fairly complex digital circuit. Application specific integrated circuit (ASIC) designers also frequently reuse their previous solutions and adapt them to their current needs. A circuit simulator is indispensable in this development procedure. The computers are mainly used to analyze human designs.

Initially a suitable circuit configuration is required, which can potentially fulfil the given requirements. This task is mostly left to the designer although several tools partially automating the topology synthesis appeared in the past /1/-/4/. Then the circuit sizing problem has to be solved. One desires such element sizes (e.g. MOSFET channel widths and lengths, capacitors, resistors, etc.) that required circuit properties are met in the most robust manner. Circuit sizing is an optimization process by its nature and one can find quite an extensive literature in this area. Sizing of nominal circuits was considered in /5/-/6/, sizing problems accounting for parameter tolerances (parameter centering) were addressed in /7/-/9/, and worst-case optimization in /10/-/12/. Various optimization tools were developed, like equation based GPCAD /13/, which uses geometric programming formulation of an optimization problem /14/ on predefined posynomial equations, AMG /15/, utilising a symbolic simulator /16/ to obtain circuit equations, and the simulation based ASTRX/OBLX /17/. Recently numerous papers (e.g. /12/, /18/-/23/) are addressing the sizing problem from different aspects like process and operating tolerances, mismatch, yield and robustness.

Despite all the research efforts made circuit sizing is still a task that is addressed manually. New sizes for the next experiment are determined by a human designer and not automatically by the optimization method. In our opinion the automated optimization is rarely used because of three major reasons:

- there are no general optimization tools integrated into any of the most popular circuit simulators for ASIC design (optimization tools, e.g. /13/, /15/, /17/, are not integrated into commercial simulators and therefore offer only very limited capabilities),
- the mathematical formulation of the cost function, which would yield acceptable solutions, is rather complicated and demands an experienced user (optimization algorithms can get trapped in senseless regions of parameter space, resulting in degenerated solutions; searching for the minimum of the cost function can also result in circuits highly sensitive to manufacturing process and operating condition variations

/21/; a possible solution is the use of implicit constraints /14/, /20/, /23/), and

- the results of the optimization run are not to be unlimitedly trusted (in many cases the minimum found is not the global one, even if a global optimization method was used).

This paper focuses on the last of these three drawbacks. There exists many different gradient, quasi gradient, and direct search optimization algorithms. A good survey of the first family can be found in /24/. Gradient based methods are greedy by default and require the derivatives of the cost function to be calculated at each iteration. When applied to circuit sizing, the derivatives are usually calculated by a sensitivity analysis, meaning that the cost function can't be of arbitrary form. Those methods have a strong local nature and are therefore usually used for finetuning circuits /25/.

On the other hand direct search methods /26/-/28/ do not require additional gradient computations. Convergence properties for pattern search methods have been reported in /29/. These methods can be classified by their behaviour as local or global. Some global methods even guarantee to find the global minimum if certain conditions are fulfilled /30/-/31/.

Performance of an optimization method on cost functions depends on many parameters one of which is the initial point. The same method can lead to quite different results for different initial points. Local methods are more sensitive than global ones. The latter have always some randomness build into them, which at least partially neutralises the importance of the proper selection of the algorithm's initial point.

The selection of the initial point is usually left to the user, who relies upon knowledge and intuition. Usually a point is chosen where the circuit's best performance is expected. If the choice is right, the minimum of the cost function lies near and the optimization task turns to fine tuning of the circuit. But on the other hand no additional information is gained. The optimization process just confirms the expectations. A great part of the parameter space is left unexplored and the question of finding a better solution remains open.

If we want to be assured that no better point exists then the whole parameter space has to be explored. One way to do this is to optimize the circuit starting from several different initial points, and each optimization run has to cover a different part of the parameter space. The optimization process becomes a group of individual optimization runs.

Optimization methods have limited memory and therefore only a few points from previous iterations are used to determine the next step. Today computers easily store all the evaluated points, while the evaluation itself is still computationally expensive. Thus the initial point for the next optimization run should be determined using the information obtained from evaluated points. This paper proposes a heuristic method based on the probabilistic approach /32/-/33/. The method puts the new initial point in a part of the parameter space, where the probability of finding a new minimum is high. It can be applied to multidimensional parameter space and does not require significant computer effort.

Several minima are obtained in such an optimization process. The designer can decide, which one is most appealing and may even continue with the investigation of the unexplored parts of the parameter space. First the mathematical background of the assumptions used later in the heuristic algorithm are highlighted. Several optimization cases of CMOS integrated operational amplifiers are illustrated and the obtained results are commented.

2 Mathematical Background, One Dimensional Probabilistic Approach

Let $E(\mathbf{x})$, $\mathbf{x} \in A \subseteq \Re^n$, $E : \Re^n \rightarrow \Re$ denote the cost function where A denotes a feasible region. The purpose of every optimization process is to find a global minimum $\mathbf{x}_0$ of the cost function $E(\mathbf{x})$, $E(\mathbf{x}_0) \leq E(\mathbf{x})$, $\forall \mathbf{x} \in A$. In one dimension the feasible region of the parameter space is defined as an interval $A = [x_{low}, x_{high}]$. Let us define a continuous stochastic process $f(x, \omega)$. It assigns a function $f(x)$ to every outcome $\omega \in \Omega$ of experiment ζ. The domain of ω is the set of all experimental outcomes Ω, and the domain of x is a set of real numbers $\Re$. Let the one dimensional cost function $E(x)$ be equal to a realisation of the stochastic process $f(x, \omega)$ for an outcome ω_0 on the interval A.

$$E(x) = f(x, \omega_0) \quad \omega_0 \in \Omega \quad x \in A \tag{1}$$

Cost function $E(x)$ is an arbitrary real function on the interval A. By its definition the distribution function $G(f_0, x)$ gives the probability of an event $\{f(x, \omega) \leq f_0\}$ at a particular x. We assume normal distribution for $G(f_0, x)$ with variance $\sigma^2(x)$ and expected value $m(x)$.

$$G(f_0, x) = \mathrm{P}\{f(x, \omega) \leq f_0\} = \frac{1}{\sqrt{2\pi}\sigma(x)} \int_{-\infty}^{f_0} e^{-(f - m(x))^2 / 2\sigma^2(x)} df \tag{2}$$

After one or more optimization runs the cost function has been evaluated at several points. Lets say we have k such points $x_1, x_2, \ldots x_k$, and the corresponding cost function values $E(x_i)$, $i = 1, 2, \ldots k$, are known. An event Z_k is defined as $\{f(x_i, \omega) = E(x_i), i = 1, 2, \ldots k\}$. In other words, the event Z_k occurs, when the stochastic process function $f(x, \omega)$ is equal to the cost function $E(x)$ in all known points $x_1, x_2, \ldots x_k$, for outcome ω. The event Z_k becomes certain if the expected value $m(x)$ is equal to the cost function and if variance $\sigma^2(x)$ is zero at all known points. Therefore $m(x_i) =$

$E(x_i)$ and $\sigma(x_i) \rightarrow 0$ for $i = 1, 2, \ldots k$. When mean and variance have the above properties, the distribution $G(f_0, x)$ becomes the conditional probability of event $\{f(x, \omega) \leq f_0 / Z_k\}$.

Let *opt* be the index of a point with the lowest cost function value among known points. So the relation $E(x_{opt}) \leq E(x_i)$, $i = 1, 2, \ldots k$, is valid. We define a function $f_{min}(x, \omega)$. Its value is always lower than $E(x_{opt})$ for an arbitrary x and any outcome ω.

$$f_{min}(x,\omega) = \min(E(x_{opt}), f(x,\omega)) \tag{3}$$

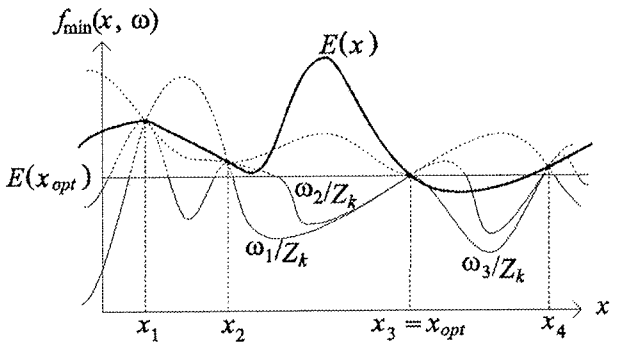


Figure 1: Functions $f_{min}(x, \omega)$ (solid) and realisations of a stochastic process $f(x, \omega)$ (dashed) for different outcomes ω. The event Z_k is certain, therefore $m(x_i) = E(x_i)$ and $\sigma^2(x_i) \rightarrow 0$, $i = 1, 2, \ldots k$, $k = 4$.

Distribution $G_{min}(f_0, x)$ of function $f_{min}(x, \omega)$ gives the probability of event $\{f_{min}(x, w) \leq f_0 / Z_k\}$, where Z_k represents a certain event as mentioned above. It can be obtained from the distribution $G(f_0, x)$ and the definition of $f_{min}(x, \omega)$. The probability density function $g_{min}(f_0, x)$ is the derivative of the distribution $G_{min}(f_0, x)$.

$$\begin{aligned} G_{min}(f_0,x) &= \mathrm{P}\{f_{min}(x,\omega) \leq f_0\} \\ &= G(f_0,x) + (1 - G(f_0,x))u(f_0 - E(x_{opt})) \end{aligned} \tag{4}$$

$$\begin{aligned} g_{min}(f_0,x) &= \frac{\partial G_{min}(g_{min}(f_0,x),x)}{\partial f_0} = \\ &= \frac{e^{(f_0 - m(x))^2 / 2\sigma^2(x)}}{\sqrt{2\pi}\sigma(x)}(1 - u(f_0 - E(x_{opt}))) + \\ &(1 - G(E(x_{opt}),x))\delta(f_0 - E(x_{opt})) \end{aligned} \tag{5}$$

Functions $u(f_0 - E(x_{opt}))$ and $\delta(f_0 - E(x_{opt}))$ in (4) and (5) represent a unit step function and its derivative, a unit Dirac impulse, respectively.

The expected value $E\{f_{min}(x, \omega) / Z_k\}$ is the mean of the function $f_{min}(x, \omega)$ at a particular x. Because of event Z_k it is equal to the cost function's value $E(x_{opt})$ in all k known points. The question is where to choose the new initial point for the next optimization run, if the cost function is already known in k points. A natural decision is to set it where the expected value $E\{f_{min}(x, \omega) / Z_k\}$ is minimal. To find out a new starting point x_0 a minimisation problem (6) has to be solved. The integral definition of the expected value expresses the minimisation problem with the density function $g_{min}(f_0, x)$. The upper bound of the integral can be set to $E(x_{opt})$ using equation (5).

$$\begin{aligned} x_0 &= \min_{x \in A}\left(\mathrm{E}\{f_{min}(x,\omega) / Z_k\}\right) \\ &= \min_{x \in A}\left(\int_{-\infty}^{\infty} f_0 g_{min}(f_0,x)df_0\right) \\ &= \min_{x \in A}\left(\int_{-\infty}^{E_{opt}} f_0 g_{min}(f_0,x)df_0\right) \end{aligned} \tag{6}$$

The minimisation problem (6) can be transformed into a maximisation problem (7) using the distribution function $G_{min}(f_0, x)$ instead of the probability density.

$$\begin{aligned} x_0 &= \max_{x \in A}\left(\int_{-\infty}^{E_{opt}} G_{min}(f_0,x)df_0\right) \\ &= \max_{x \in A}\left(\frac{\sigma(x)}{\sqrt{2\pi}} \int_{-\infty}^{E(x_{opt}) - m(x)/\sigma(x)} \int_{-\infty}^{u} e^{-t^2/2}dtdu\right) \end{aligned} \tag{7}$$

The probability distribution and the density function of a limited random walk, also known as Wiener process $w(t)$, are normal with constant mean and variance increasing with t. We also assume normal distribution for our process $f(x, \omega)$. Wiener process $w(t)$ is a continuous function of variable t. Suppose the cost function $E(x)$ is continuous in the vicinity of known points, so it can be a sample path of a Wiener process there. This assumption does not place any physically unrealistic limitations on types of cost functions, which take place in circuit design optimization problems. Therefore we can presume a constant expected value and a linearly increasing variance near known points. We set the mean and variance to $m(x) = E(x_i)$ and $\sigma^2(x) = a|x - x_i|$ around i^{th} point. Then event Z_k is certain as well. In the neighbourhood of every determined point equation (7) becomes

$$x_0 = \max_{x \in A}\left(\sqrt{\frac{\alpha|x - x_i|}{2\pi}} \int_{-\infty}^{E(x_{opt}) - E(x_i)/\sqrt{\alpha|x - x_i|}} \int_{-\infty}^{u} e^{-t^2/2}dtdu\right) \tag{8}$$

$i = 1,2,\ldots k.$

The expression in equation (8) is a monotonically decreasing function of cost value $E(x_i)$ and monotonically increasing function of distance $|x - x_i|$. This leads to two conclusions:

- first due to decrease with $E(x_i)$ the new initial point x_0 lies rather closer to the known points with lower cost function value, than to those with higher cost function value,
- due to the increase resulting from $|x - x_i|$ it lies away form all known points so the distance to the nearest one is as large as possible.

Both conclusions can be intuitively generalized to n dimensional parameter space. A simple heuristic method described in the following section is based on this generalisation.

3 A Heuristic Method for Finding New Initial Points

The second conclusion tells us, that a new initial point has to be somewhere in the parameter space, where the density of already evaluated points is low. If it is low, then we expect the average distance between two nearest points to be large in general. But we have to define how to measure the density of known points. Let us divide the parameter space into 2^n equal subspaces (2^n equal boxes). Let the density be equal to the number of known points in a particular subspace, and let it be constant across the whole subspace. A new initial point will be chosen in the subspace with the lowest density.

The first conclusion on the other hand tells us, that the contribution to the density is not always the same for all already evaluated points. Those with lower cost function values should contribute less, than the ones with higher cost function values. In the previous definition all of them contributed one unit, regardless of the cost function value. Therefore known points have to be weighted. Each point will contribute its weight, which has to be proportional to its cost. Let the weight u of a point with cost function value E be defined by equation (9).

$$u = \frac{(\beta - 1)E + E_{\max} - \beta E_{\min}}{E_{\max} - E_{\min}} \qquad (9)$$

E_{min} and E_{max} represent the lowest and the highest cost function value among already determined points, respectively. The point with the lowest cost function value has always weight one. The weight of the point with the highest cost function value is given by coefficient β, and now it contributes β times more to the density, than the lowest point.

So far all known points, for which we know, that they violate implicit constraints, are still not included in our definition of density. They lack a cost function value E, so their weight can not be calculated by equation (9). But those points give us some information about the cost function and therefore they have to be taken into account. We set their weight to 2β.

Finally the heuristic algorithm for determining a new initial point for the next optimization run is described in the repeat until loop (Fig. 2) below. The space is divided into 2^n equal subspaces, until we find a subspace with no points determined yet. A new initial point is selected there randomly. The algorithm is very simple, so it demands only a small amount of computational time.

```
calculate weights for all known points;
temporary space := explicitly constrained space;
repeat
  divide temporary space into 2^n equal subspaces;
  add up weights in particular subspaces;
  temporary space :=
    subspace with the lowest sum of weights;
until lowest sum ≠ 0
randomly pick new point in temporary space;
```

Figure 2: Symbolic algorithm of heuristic initial point determination for a new optimization run.

4 Sizing Problem Cases and Results

In this section three CMOS design cases are described to illustrate the capabilities of the proposed approach. Two simple two-stage operational amplifiers with *p* and *n*-channel differential pair (Figs. 3 and 4) and a telescopic cascode operational amplifier (Fig. 5) were optimized. Several versions of the above three sample circuits optimized to meet different requirements were used as a part of larger mixed signal integrated circuits. The amplifiers were designed for and produced in 0.3μm and 0.8μm technology. The parameters varied were all transistor channel dimensions (widths and lengths), MOS multiplier factors and also the resistances and the capacitances.

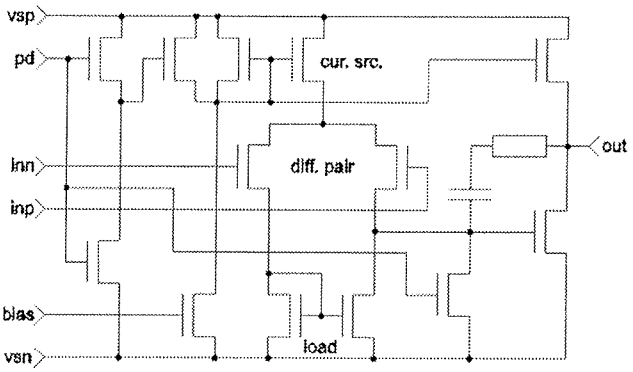


Figure 3: Operational amplifier with p-channel differential pair.

The circuit characteristics that take part in the cost function are listed in the upper part of Tables 1 and 2 . The cost function is a rather complicated mathematical formulation which combines results of several types of analyses in several different operating conditions (variable supply and reference voltages, variable bias current, variable temperature etc.) and manufacturing environments (variable production process conditions given with corner transistor models) /12/. Beside searching for an optimal nominal circuit the robustness is also taken into account. For the

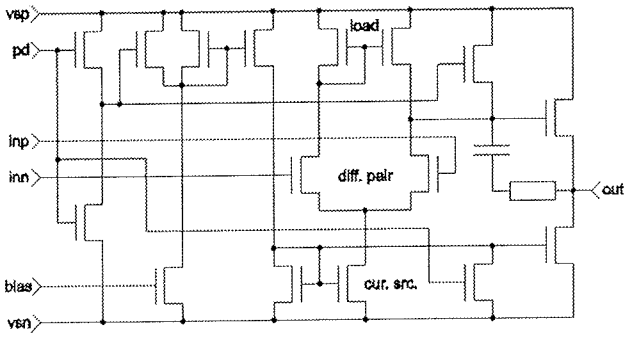


Figure 4: Operational amplifier with n-channel differential pair.

two-stage amplifiers mismatching is simulated by slight model variations of one of the matching transistors. The shape of such a complicated cost functions in multidimensional parameter space is completely unknown. Finding a global minimum is a difficult task for any optimization method and circuit simulator since it requires many circuit anal-

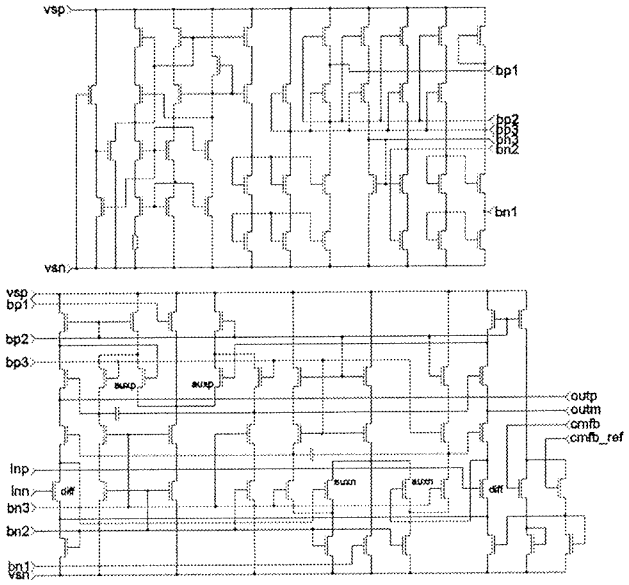


Figure 5: Telescopic cascode operational amplifier.

property		target	*p*-channel diff. pair				*n*-channel diff. pair			
A	μm^2	↓	11619	12289	12241	10105	14151	13521	17706	14286
v_{pp}	V	↑	3.7	3.7	3.8	3.6	3.8	3.9	3.8	3.8
v_{pp}/v_{inpp}		↑	2101	2937	2153	2159	4535	4246	4232	4741
v_{offset}	µV	↓	87	60	96	49	32	81	49	11
$v_{outoffset}$	mV	↓	201	199	198	199	99	101	100	100
i_p	µA	↓	727	636	674	559	689	828	754	659
f_{0dB}	MHz	↑	20	20	20	14	16	20	14	13
pm	°	↑	37	37	31	23	34	40	55	37
am	dB	↓	-39	-37	-24	-22	-40	-32	-38	-40
CMRR	dB	↓	-96	-100	-91	-97	-108	-106	-104	-102
PSRRp	dB	↓	-89	-90	-112	-101	-49	-50	-46	-48
PSRRn	dB	↓	-62	-62	-60	-58	-50	-51	-51	-52
$noise_{1/f}$	$nV/Hz^{1/2}$	↓	100	91	80	56	114	100	102	108
$noise_{term}$	$nV/Hz^{1/2}$	↓	9.0	9.5	9.0	10.3	8.6	9.4	9.0	8.6
t_{rise}	ns	↓	361	431	405	431	285	259	243	312
t_{fall}	ns	↓	174	134	171	216	479	426	440	582
transistor			*mw / l* ratio							
differential pair			173	141	200	151	130	117	41	154
active load			12	9	12	4	18	21	41	14
current source			18	24	18	7	19	10	31	39

Notations: A ... area, v_{pp} ... peak-to-peak voltage, v_{pp}/v_{inpp} ... dc gain, v_{offset} ... offset voltage, $v_{outoffset}$ ... symmetry, i_p ... current consumption, f_{0dB} ... frequency at 0dB gain, *pm* ... phase margin, *am* ... amplitude margin, *CMRR* ... common mode rejection ratio, *PSRRp* ... power supply rejection ratio to positive terminal, *PSRRn* ... power supply rejection ratio to negative terminal, $noise_{1/f}$ ... noise at low frequencies (at 100Hz), $noise_{term}$ ... thermal noise at higher frequencies (at 100kHz), t_{rise} ... rise time, t_{fall} ... fall time, *m* transistor multiplier, *w* channel width and *l* channel length. Symbols ↑ and ↓ indicate that the desired value is as high or as low as possible.

Table 1: Results of some succesfull optimization runs for both two-stage amplifiers (0.8µm technology)

yses. Nevertheless we expect that somewhere in the parameter space there is a global minimum which defines the optimal solution satisfying the given requirements.

The results for the two-stage operational amplifiers are summarised in Table 1 and for the telescopic cascode operational amplifier in Table 2. Only some of the optima found with the initial point set by the described heuristics are given because of the tables size. The upper part of both Tables contains nominal circuit performances. The lower part summarises parameter values in each minimum. Multiplying factor * channel width / channel length (*mw* = *l*) ratio is given for some transistors in all three cases. If short channel effects in submicron region are neglected then the ratio defines a transistor. Therefore it is convenient for estimating if two solutions are equivalent.

The optimization method used in a particular run is not essential. In fact any local method can be used since global methods tend to the global minimum regardless of the chosen initial point. Direct methods are preferable since the derivatives of the cost function are not required (often impossible to calculate without resorting to perturbation methods which are not accurate enough). So one can use any simplex, quasi gradient (metric matrix, trust region etc.), heuristic, etc. based method. In our experiments a heuristic simplex based method was used. The cost function was composed as a weighted sum of deviations from the target values for nominal and worst conditions. If a particular target is fulfilled the optimization process does not tend to improve it any further. Approximately 500 to 1000 circuit evaluations were needed for one run to converge and on the average every third run was successful. Thus the results in Table 2 were obtained in 30000 circuit evaluations. Comparing this result to a performance of well known global optimization methods like simulated annealing or genetic algorithms is encouraging since over 150000 circuit evaluations are needed to optimize a circuit like the telescopic cascode amplifier.

From all presented cases we can see that many different solutions of the circuit sizing problem exist. An interesting parallel can be drawn with /34/-/35/ where the entire circuit synthesis problem (topology and sizing) was addressed by genetic programming. Uncommon circuit topology solutions were found beside well known ones.

More or less the same circuit properties can be obtained with several different sets of circuit parameters. Two explanations are at hand: 1.) the target values are to loose for the used circuit configuration and for the given technology and are easily fulfilled, or 2.) the optimization run is stopped at different trade offs among given targets. Because all requirements are never fulfilled the second explanation is more probable. To confirm this, the same experiments were repeated with tighter targets. The requirements remained unfulfilled and individual solutions didn't merge.

A closer look at the Table 2 also confirms that the solutions represent trade offs among required targets. We can see for instance that the last two results have complementary properties. While the solution from column nine has low v_{pp}, *pm* and *am* it has high i_p and f_{0dB}. On the other hand the last circuit (column 10) has opposite properties. The same observations can be made in Table 1 .

5 Conclusion

A simple heuristic method for setting the initial points of individual optimization runs was described. The idea is based on a one dimensional probabilistic approach extended to multidimensional parameter space. The main objective is to uniformly search the parameter space with a sequence of optimization runs. Each run contributes some

property		target	telescopic coscode operational amplifier									
A	μm^2	↓	2795	2605	2688	2603	2735	2706	3000	2686	2905	2479
v_{pp}	V	↑	3.0	2.7	2.9	2.8	2.8	2.9	2.8	2.8	2.3	3.1
v_{pp}/v_{inpp}		↑	133	139	135	135	137	134	135	135	136	135
$cmfb_{offset}$	μV	↓	24	0.4	34	1	38	5	21	0.3	25	30
i_p	μA	↓	1.4	1.2	1.3	1.4	1.4	1.3	1.3	1.4	1.4	1.1
f_{0dB}	MHz	↓	242	260	269	263	250	261	268	273	305	171
pm	o	↑	74	73	73	75	76	70	65	73	66	79
am	dB	↑	-25	-25	-26	-25	-28	-25	-20	-25	-24	-28
transistor			*mw / l* ratio									
main differential pair			290	350	290	290	230	350	290	290	410	230
auxiliary *p* differential pair			28	22	22	16	16	16	28	28	22	28
auxiliary *n* differential pair			14	20	8	8	11	14	20	20	11	11

Notations: *A* ... area, v_{pp} ... peak-to-peak voltage, v_{pp}/v_{inpp} ... dc gain, $cmfb_{offset}$ ... common mode feedback offset, i_p ... current consumption, f_{0dB} ... frequency at 0dB gain, *pm* ... phase margin, *am* ... amplitude margin, *m* transistor multiplier, *w* channel width and *l* channel length. Symbols ↑ and ↓ indicate that the desired value is as high or as low as possible.

Table 2: Results of some successful optimization runs for telesopic cascode amplifier (0.3μm technology)

new information about the cost function shape in the multidimensional parameter space. Different local minima are found, if they are present. Multiple solutions are obtained providing additional insight into circuit behaviour. The designer can decide which one is the most appropriate and continues his/her work from there with finetuning. Finetuning is usually necessary since the obtained minimum of the cost function not necessarily satisfies the designer's expectations. A statistical model of the cost function was presented. The construction of cost function itself /12/ is beyond the scope of this paper.

The method takes into account all collected cost function data. Therefore all calculated points must be stored and some additional MBytes of RAM are occupied for that reason. But on the other hand it requires only a small computing effort and does not take a considerable amount of time. The optimization method used in the individual runs can be an arbitrary fast greedy (local) method. Fast convergence of such methods ensures short runtimes since global methods (like simulated annealing or genetic algorithms etc.) have in general slow convergence. More information is obtained instead of a single minimum. Our method can try several different initial points in the time needed by a global method to converge.

6 References

/1/ M.G.R. Degrauwe et al., "IDAC: An interactive design tool for analog CMOS circuits," *IEEE J. Solid-State Circuits*, vol. sc-22, no. 6, Dec. 1987, pp. 1106-1116

/2/ R. Harjani, R.A. Rutenbar and L.R. Carley, "OASYS: A framework for analog circuit synthesis," *IEEE Trans. Computer-Aided Design*, vol. 8, no. 12, Dec. 1989, pp. 1247-1266

/3/ H.Y. Koh, C.H. Séquin and P.R. Gray, "OPASYN: A compiler for CMOS operational amplifiers," *IEEE Trans. Computer-Aided Design*, vol. 9, no. 2, Feb. 1990, pp. 113-125

/4/ J.P. Harvey, M.I. Elmasry and B. Leung, "STAIC: An interactive framework for synthesizing CMOS and BiCMOS analog circuits," *IEEE Trans. Computer-Aided Design*, vol. 11, no. 11, Nov. 1992, pp. 1402-1417

/5/ R.K. Brayton, G.D. Hachtel and A.L. Sangiovanni-Vincentelli, "A survey of optimization techniques for integrated-circuit design," *Proc. IEEE*, vol. 69, no. 10, Oct. 1981, pp. 1334-1364

/6/ W. Nye et al., "DELIGHT.SPICE: An optimization-based system for the design of integrated circuits," *IEEE Trans. Computer-Aided Design*, vol. 7, no. 4, Apr. 1988, pp. 501-519

/7/ S.W. Director and G.D. Hachtel, "The simplicial approximation approach to design centering," *IEEE Trans. Circuits Syst.* I, vol. cas-24, no. 7, July 1977, pp. 363-372

/8/ K.J. Antreich and R.K. Koblitz, "Design centering by yield prediction," *IEEE Trans. Circuits Syst.* I, vol. cas-29, no. 2, Feb. 1982, pp. 88-96

/9/ P. Feldmann and S.W. Director, "Integrated circuit quality optimization using surface integrals," *IEEE Trans. Computer-Aided Design*, vol. 12, no. 12, Dec. 1993, pp. 1868-1879

/10/ K.J. Antreich, H.E. Hraeb and C.U. Wieser, "Circuit analysis and optimization driven by worst-case distances," *IEEE Trans. Computer-Aided Design*, vol. 13, no. 1, Jan. 1994, pp. 57-71

/11/ A. Dharchoudhury and S.M. Kang, "Worst-case analysis and optimization of VLSI circuit performances," *IEEE Trans. Computer-Aided Design*, vol. 14, no. 4, Apr. 1995, pp. 481-492

/12/ Á. Bürmen et al., "Automated robust design and optimization of integrated circuits by means of penalty functions," *Int. J. Electron. Comm.*, accepted for publication

/13/ M. del Mar Hershenson, S.P. Boyd and T.H. Lee, "GPCAD: A tool for CMOS op-amp synthesis," *1998 IEEE/ACM Int. Conf. Comput.-Aided Design*, New York, 1998, pp. 296-303

/14/ ——, "Optimal design of a CMOS op-amp via geometric programming," *IEEE Trans. Computer-Aided Design*, vol. 20, no. 1, Jan. 2001, pp. 1-21

/15/ G. Gielen et al., "An analogue module generator for mixed analogue/digital ASIC design," *Int. J. Circuit Theory and App.*, vol. 23, no. 4, July-Aug. 1995, pp. 269-283

/16/ G.G.E. Gielen, H.C.C. Walscharts and W.M.C. Sansen, "Analog circuit design optimization based on symbolic simulation and simulated annealing," *IEEE J. Solid-State Circuits*, vol. 25, no. 3, June 1990, pp. 707-713

/17/ E.S. Ochotta, R.A. Rutenbar and L.R. Carley, "Synthesis of high-performance analog circuits in ASTRX/OBLX," *IEEE Trans. Computer-Aided Design*, vol. 15, no. 3, Mar. 1996, pp. 273-294

/18/ G. Debyser and G. Gielen, "Efficient analog circuit synthesis with simultaneous yield and robustness optimization," *1998 IEEE/ACM Int. Conf. Comput.-Aided Design*, New York, 1998, pp. 308-311

/19/ R. Schwencker et al., "Automating the sizing of analog CMOS circuits by consideration of structural constraints," *DATE Conf. and Exhibition, 1999*, Los Alamitos, 1999, pp. 323-327

/20/ R. Phelps et al., "Anaconda: simulation-based synthesis of analog circuits via stochastic pattern search," *IEEE Trans. Computer-Aided Design*, vol. 19, no. 6, June 2000, pp. 703-717

/21/ T. Mukherjee, L.R. Carley and R.A. Rutenbar, "Efficient handling of operating range and manufacturing line variations in analog cell synthesis," *IEEE Trans. Computer-Aided Design*, vol. 19, no. 8, Aug. 2000, pp. 825-839

/22/ F. Schenkel et al., "Mismatch analysis and direct yield optimization by spec-wise linearization and feasibility-guided search," *Proc. 38th DAC*, New York, 2001, pp. 858-863

/23/ P. Mandal and V. Visvanathan, "CMOS op-amp sizing using a geometric programming formulation," *IEEE Trans. Computer-Aided Design*, vol. 20, no. 1, Jan. 2001, pp. 22-38

/24/ J. Nocedal, "Theory of algorithms for unconstrained optimization," *Acta Numerica*, vol. 1, 1992, pp. 199-242

/25/ A.R. Conn et al., "JiffyTune: circuit optimization using time-domain sensitivities," *IEEE Trans. Computer-Aided Design*, vol. 17, no. 12, Dec. 1998, pp. 1292-1309

/26/ M.H. Wright, "Direct search methods: once scorned, now respectable," *Proc. 1995 Dundee Biennial Conf. in Numerical Analysis*, 1995, pp. 191-208

/27/ M.J.D. Powell, "Direct search algorithms for optimization calculations," *Acta Numerica*, vol. 7, 1998, pp. 287-336

/28/ R.M. Lewis, V.Torczon and M.W. Trosset, "Direct search methods: then and now," *J. Computational and Applied Math.*, vol. 124, no. 1-2, Dec. 2000, pp. 191-207

/29/ V. Torczon, "On the convergence of pattern search algorithms," *SIAM J. Optimization*, vol. 7, no. 1, Feb. 1997, pp. 1-25

/30/ F. Romeo and A. Sangiovanni-Vincentelli, "A theoretical framework for simulated annealing," *Algorithmica*, vol. 6, no. 3, 1991, pp. 302-345

/31/ A.W. Johnson and S.H. Jacobson, "On the convergence of generalized hill climbing algorithms," *Discrete Applied Math.*, vol. 119, no. 1-2, June 2002, pp. 37-57

/32/ A.G. Zilinskas, "On statistical models of complex multimodal functions and their application to the design of optimization algorithms," *Problems of Control and Inform. Theory (English translation)*, vol. 10, no. 1, 1981, pp. 19-30

/33/ R. Buche and H.J. Kushner, "Rate of convergence for constrained stochastic approximation algorithms," *SIAM J. Control and Optimization*, vol. 40, no. 4, 2001, pp. 1011-1041

/34/ J.R. Koza et al., "Automated synthesis of analog electrical circuits by means of genetic programming," *IEEE Trans. Evol. Comput.*, vol. 1, no. 2, July 1997, pp. 109-128

/35/ ——, "Synthesis of topology and sizing of analog electrical circuits by means of genetic programming," *Comput. Methods in Applied Mechanics and Eng.*, vol. 186, no. 2-4, June 2000, pp. 459-482

Janez Puhan, Árpad Bűrmen and Tadej Tuma
University of Ljubljana
Faculty of Electrical Engineering
Tržaška 25, SI-1000 Ljubljana
E-mail: janez.puhan@fe.uni-lj.si

Prispelo (Arrived): 28.03.2003 *Sprejeto (Accepted): 26.08.2003*

FORMAL VERIFICATION OF DISTRIBUTED MUTUAL-EXCLUSION CIRCUITS

Robert Meolič, Tatjana Kapus, Bogdan Dugonik, Zmago Brezočnik

Faculty of Electrical Engineering and Computer Science, University of Maribor, Maribor, Slovenia

Key words: Asynchronous circuit, Fundamental mode, Process algebra, Model checking, ACTL

Abstract: Distributed mutual-exclusion (DME) circuits are an interesting example of asynchronous circuits. They are composed of identical DME cells connected in a ring of arbitrary size. Each DME cell provides a connection point for one user, and all users compete for exclusive access to a shared resource. This paper reports about formal verification of two well-known DME circuit implementations. External behaviour of the circuits is described with a simple process, whereas the required properties are expressed with temporal logic ACTL. We were able to detect hazards and verify correctness of external behaviour of the circuits under the fundamental mode of operation.

Formalna verifikacija vezij za porazdeljeno medsebojno izključevanje

Ključne besede: asinhrono vezje, fundamentalni način, procesna algebra, preverjanje modelov, ACTL

Izvleček: Vezja za porazdeljeno medsebojno izključevanje (DME) so zanimiv primer asinhronih vezij. Sestavljena so iz enakih celic DME, povezanih v obroč poljubne velikosti. Vsaka od celic DME ponuja priključno točko za enega uporabnika in vsi uporabniki med seboj tekmujejo za izključen dostop do skupnega vira. V članku obravnavamo formalno verifikacijo dveh znanih izvedb vezja DME. Obnašanje vezij opišemo s preprostim procesom, zahtevane lastnosti pa s temporalno logiko ACTL. Na ta način smo lahko odkrili hazarde ter verificirali pravilnost obnašanja vezij v fundamentalnem načinu delovanja.

1 Introduction

Asynchronous circuits have been built and used for decades, and nowadays, large and efficient circuits can be constructed /5, 7, 17, 19, 20/. Techniques and methodologies for designing asynchronous circuits differ from those used with the synchronous approach. An important issue in asynchronous design is hazard removal. Because synchronization is performed without a global clock, unwanted signal changes can kill the circuit. Well known techniques for hazard-free synthesis, decomposition and verification of asynchronous circuits are based on modelling with flow tables /4/, asynchronous finite state machines (AFSM), burstmode state machines (BM), signal transition graphs (STG), state graphs (SG) /17/, and also process algebrae. Some of algebraic approaches to verification of asynchronous circuits are Circal agents /2/, CCS-like burst-mode specification /20/, and DILL specification based on LOTOS /8/. An overview of the state-of-the-art in tools for asynchronous design can be found in /1/.

This paper describes an algebraic approach to detecting hazards and verifying correctness of asynchronous circuits. Muller's model is used for modelling, and fundamental mode of operation is assumed. Section 2 gives an overview of asynchronous design and introduces Muller's model and hazards. Section 3 describes a simple process algebra and shows how it can be used for modelling individual gates. Section 4 describes the procedure for verification of asynchronous circuits. Section 5 introduces ACTL model checking. Section 6 presents two well-known implementations of DME circuits and reports about the results of their verification. In the conclusion we evaluate our work.

2 Asynchronous design

Circuits are composed of *gates* and *wires*. In this paper, the term gate refers to simple or complex elements for which only external behaviour is considered, and the term wire refers to connections between gates carrying binary signals. Regarding their operation, circuits can be classified into *combinational* and *sequential*. In a combinational circuit, output values of all gates are logic functions of current circuit input values. In a sequential circuit, some gate outputs depend also on a history of circuit input values. The memory effect is achieved with feedback loops.

Fundamental to an asynchronous design are assumptions about gate and wire delays. If delays are overestimated, the resulting circuit is likely to be inefficient and expensive. If they are underestimated, the design may not guarantee correct circuit operation. Delays can be *bounded* or *unbounded*. For bounded delay the upper bound is given, while the magnitude of unbounded delay is only known to be positive and finite. The delays can also be *pure* or *inertial*. In the latter case, short pulses are filtered out.

With regard to assumptions about delays, there are two widely used models for designing asynchronous circuits. *Huffman's model* supposes that gate and wire delays are

bounded and known. Circuits designed with this model are called Huffman circuits. On the other hand, *Muller's model* supposes that wires have negligible delays in comparison to gates, which have inertial unbounded delays. Muller's model is typically used to design speed-independent circuits. Because of negligible wire delays, all *forks* in Muller's model are isochronic. This means that if a signal splits, all instances of this signal have equal delays. Thus, an output signal produced by a gate is equally delayed for all gates which consume it. By explicitly adding nonisochronic FORK elements, Muller's model can be extended to produce self-timed, delay-insensitive, and quasi-delayinsensitive circuits, where wire delays are important. Note that there are also other design methodologies for asynchronous circuits not based on the mentioned models (e.g. self-clocked circuits and micropipelines).

A simple Muller's model is presented in Figure 1. It represents the C-element, a standard building block used in many asynchronous systems, which was also introduced by Muller. The C-element changes its output only if both inputs are changed to 0 or 1. In the figure, the small boxes labeled with d_1, d_2, d_3 and d_4 are delay elements attached to the gate outputs, which are the only components in the circuit delaying signals. The figure clearly shows that in the Muller's model an output signal produced by a gate is equally delayed for all gates which consume it.

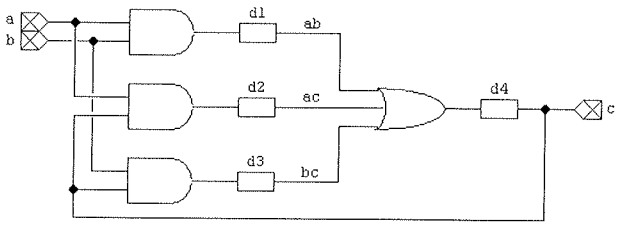


Figure 1: A gate-level implementation of the C-element

An important concept related to delays is circuit's mode of operation, which characterizes the interaction between a circuit and its environment. *Fundamental mode* of operation assumes that the environment will change the value of only one input signal at once and then wait until the circuit becomes stable. An asynchronous circuit is *stable* if no internal or output signal value can be changed without changing some input signal value. The opposite of fundamental mode is *input/output mode* of operation. In this mode, the environment can change values of input signals at any time. A third type of circuit's mode of operation is *generalized fundamental mode* or *burst mode*. There, signal changes are segregated in time forming input bursts (intervals where input signal values change) and output bursts (intervals where output signal values change). An output burst can be empty, whereas an input burst must contain at least one signal. Input and output bursts must alternate. Within a burst, the ordering of signal changes is not determined.

There are some anomalous types of asynchronous circuit behaviour, which the designers try to avoid. An example of an usually unwanted behaviour is a possibility that the circuit enters a closed loop of transitions without becoming stable. This can result in the circuit with oscilating outputs. A simple example of such a circuit and its simulation run are presented in Figure 2.

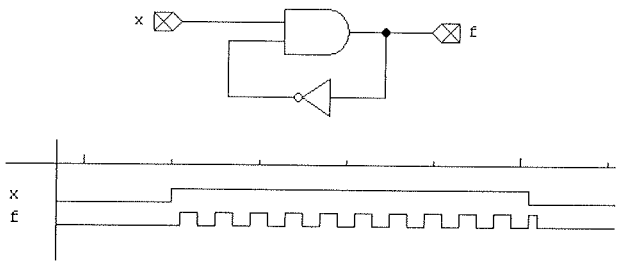


Figure 2: A circuit with oscilating output

If due to internal delays, a circuit can make an unwanted pulse called a *glitch* or can become stable with an unwanted combination of values on internal or output lines, we have a *hazard*. Hazards reflecting in glitches are classified with regard to their shape into *static* and *dynamic* hazards. Static hazard occurs when the signal is momentarily changed although it should remain the same. Dynamic hazard occurs if the signal oscillates before changing its value. A circuit which operates without hazards in the fundamental mode is called a *fundamental-mode circuit*.

For each hazard, there is a reason for its existence. In combinational circuits, three types of hazards are distinguished. *Logic hazards* are a property of particular implementation. A logic hazard exists in the circuit because for the same signal, two or more parallel paths through the circuit exist, which then reconverge. *Functional hazards* are a property of the logic functions which do not change monotonically during a sequence of particular input changes. The hazard arises when such inputs change simultaneously. Logic and functional hazards can both be either static or dynamic. A much different type of hazard is *delay hazard*. It occurs because a new input signal is applied before the circuit becomes stable. Logic hazards can always be avoided by redesigning the circuit. Functional and delay hazards can be removed only by engineering delays. Under fundamental and burst mode of operation, functional and delay hazards do not have an impact. In combinational circuits, hazardous behaviour is a transitory phenomenon, and if no new inputs are applied until the circuit stabilizes, then the correct outputs will be produced.

In sequential circuits, additional *sequential hazards* can exist as a consequence of the order in which input signals and feedback signals are considered. Sequential hazard which results in a glitch is *transient hazard*. On the other hand, if due to delays the circuit can become stable with incorrect values of internal or even output signals, we have *steady-state hazard*. Both types of sequential hazards are

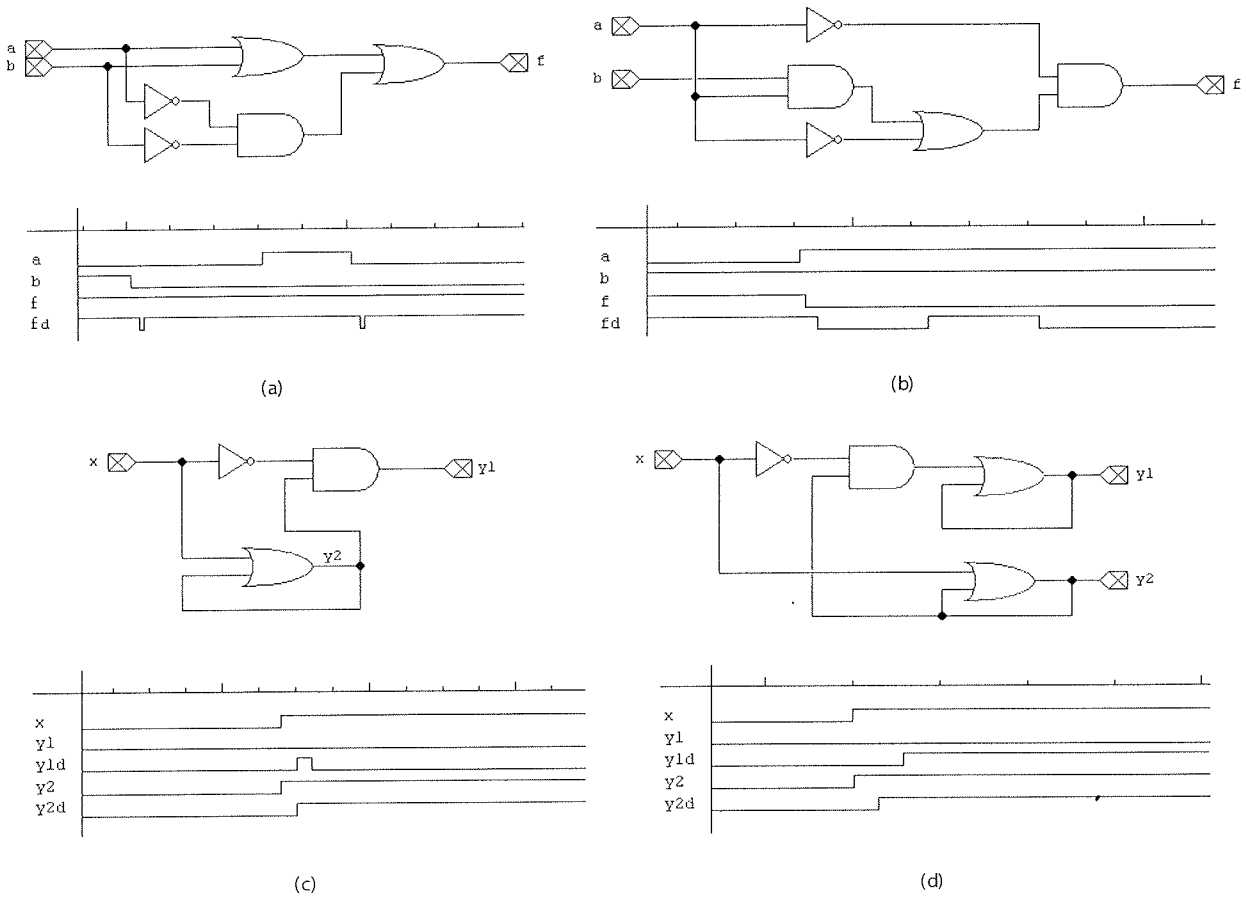


Figure 3: Simple circuits containing (a) static hazard, (b) dynamic hazard, (c) transient hazard, and (d) steady-state hazard.

an inherent property of sequential functions and not of the particular circuit implementation. Sequential hazards appear despite of fundamental and burst mode constraints.

Figure 3 shows circuits containing different types of hazards and their simulation runs. Signals f, y_1, and y_2 represent the behaviour of output signals without delays in the circuit, while signals fd, y1d, and y2d represent their behaviour after introducing significant delays. Figures 3(a) and 3(b) show logic hazards. The static hazard in Figure 3(a) is obtained by using delayed inverters. The dynamic hazard in Figure 3(b) appears when the left AND gate is delayed and the top inventer is even more delayed. Figure 3(c) represents a transient hazard which appears if the AND gate gets the new value from the feedback line earlier than the new input value. The circuit in Figure 3(d) has steady-state hazard because after changing input x, the circuit without delays produces outputs y1 = 0, y2 = 1, whereas using a delayed inverter, it produces outputs y1d = 1, y2d = 1.

3 Representing circuits with processes

Process algebrae are widely used formalisms for modelling and verification of concurrent systems, e.g. communication protocols. In a process algebra, a system is described as a set of communicating processes. Among others, well known process algebrae are CCS (*Calculus of Communicating Systems*) introduced by R. Milner in 1980 and CSP (*Communicating Sequential Processes*) introduced by C. A. R. Hoare in 1985. We will use an algebraic approach which is similar to CCS and has some notation from CSP.

In our approach, processes are labelled directed graphs. Graph nodes are called states. An edge from state p to state q labelled with action α is called an *α-transition* or shortly a *transition* from state p to state q. If there exists an *α-transition* from a given state, we say that in this state the process can *perform* α-transition or that it can *perform* action α The set of all actions which a process can perform is called the *alphabet* of the process. A sequence of transitions in the process is called a *path*. The alphabet of a process always includes a special action τ, which is used to model internal communications, not visible to an external observer. A transition with action τ is called *silent transition*. All actions others than τ are called *visible* actions. Visible actions are divided into input and output actions. The name of an *output action* always terminates by '!', e.g. a!, b!... The name of an *input action* always terminates by '?', e.g. a?, b?,... Two actions whose names differ only in

the last sign, e. g. a? and a!, are called *complementary actions*. An action complementary to the given action α is denoted by $\bar{\alpha}$. A sequence of visible actions is called a *trace*. Two states p and q have *equivalent traces* if from them the same traces can be performed. Processes are *trace-equivalent* if their initial states have equivalent traces. A state without outgoing transitions is a *deadlock* state. If there exists a sequence of transitions from the initial state of a process to a state p, then state p is *reachable* in this process. Otherwise, the state is *unreachable* in this process. Other important concepts in process algebrae are strong equivalence, observational equivalence, and determinacy of processes /16/.

Processes are used to represent external behaviour of circuits. Each transition in the process represents a change of a signal value and we do not distinguish whether the value changes from 0 to 1 or vice versa. A transition with an input action represents change of an input signal value, caused by the environment. A transition with an output action represents change of an output signal value, caused by the circuit. We will use interleaving semantics, i.e. a simultaneous occurence of two or more signal changes will be modeled by including multiple sequences of transitions, one for each permutation of changing signals.

The process represents an external behaviour of the circuit if each of its traces starting in the initial state of the process corresponds to a possible sequence of changes of input and output signal values during the operation of the circuit in the given environment considering given initial values of all input and output signals. We are not interested in exact timing when the value of signals changes. If two circuits have the same possible sequences of changes, they have the same external behaviour, although one of them is much faster than other. Therefore, two processes represent equal external behaviour if they are trace-equivalent. However, delays are important because different sequences of changes can be possible in the same asynchronous circuit if the delay of gates changes. Some traces in the process may correspond to sequences of changes possible only with particular arrangements of gate delays and not all of them. Note that we do not require fixed delays. Gate delays can change during the operation of the circuit. Because each action represents just a change of signal value, we cannot uniformly associate a circuit to the given process without knowing the initial value of all input and output signals. Namely, to determine whether performing a particular transition represents a rising or falling edge, one must know the initial value of that signal and follow its changes.

With regard to the level of abstraction, the process can represent more or less details on the circuit's internal behaviour and its connections with the enviroment. For the purpose of finding hazards and verification of asynchronous circuits under the fundamental mode of operation, we introduce a special form of processes called circuit models. A *circuit model* is determinate process without unreachable states and without silent transitions. All circuit models representing external behaviour of the same circuit in the same environment and with the same initial values of input and output signals are strongly equivalent. Among them, the one with minimal number of states will be identified and used for verification.

Circuit models in Figure 4 represent external behaviour of an AND gate, the C element, and the FORK element, respectively. Further, they will be used to build DME cells. The AND gate and the C element have two input signals (actions a? and b?) and one output signal (action c!). The FORK element is a frequently used gate in asynchronous circuits. It has one input signal (action a?) and two output signals (actions b! and c!). The value of both outputs changes simultaneously after the change of the input value. Simultaneous change of two outputs is modelled by the ability of performing corresponding output actions in both orders. The presented circuit models describe external behaviour of gates under the fundamental mode of operation with all input and output signals initially set to 0.

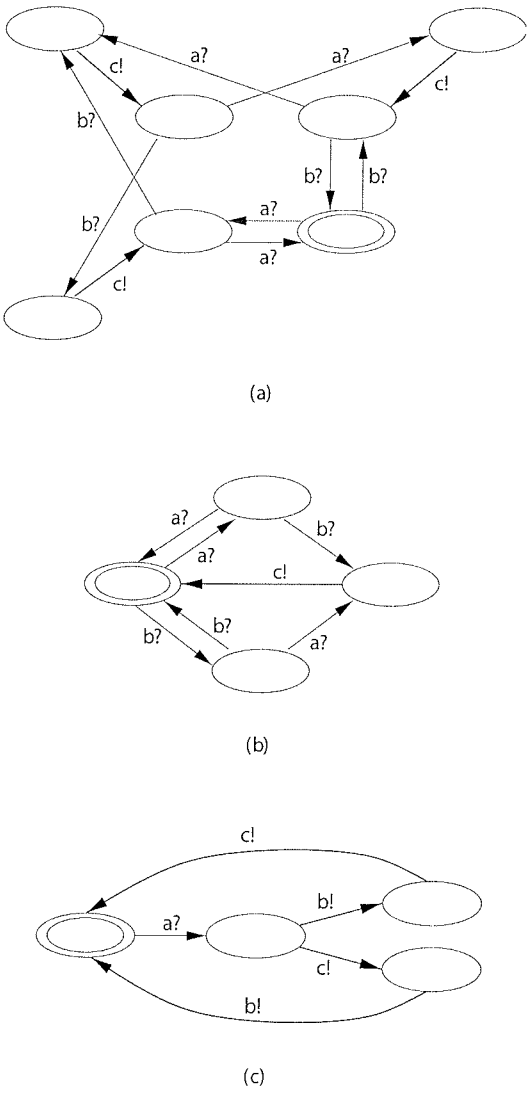


Figure 4: Circuit model of (a) two-input AND gate, (b) the C-element, and (c) the FORK element

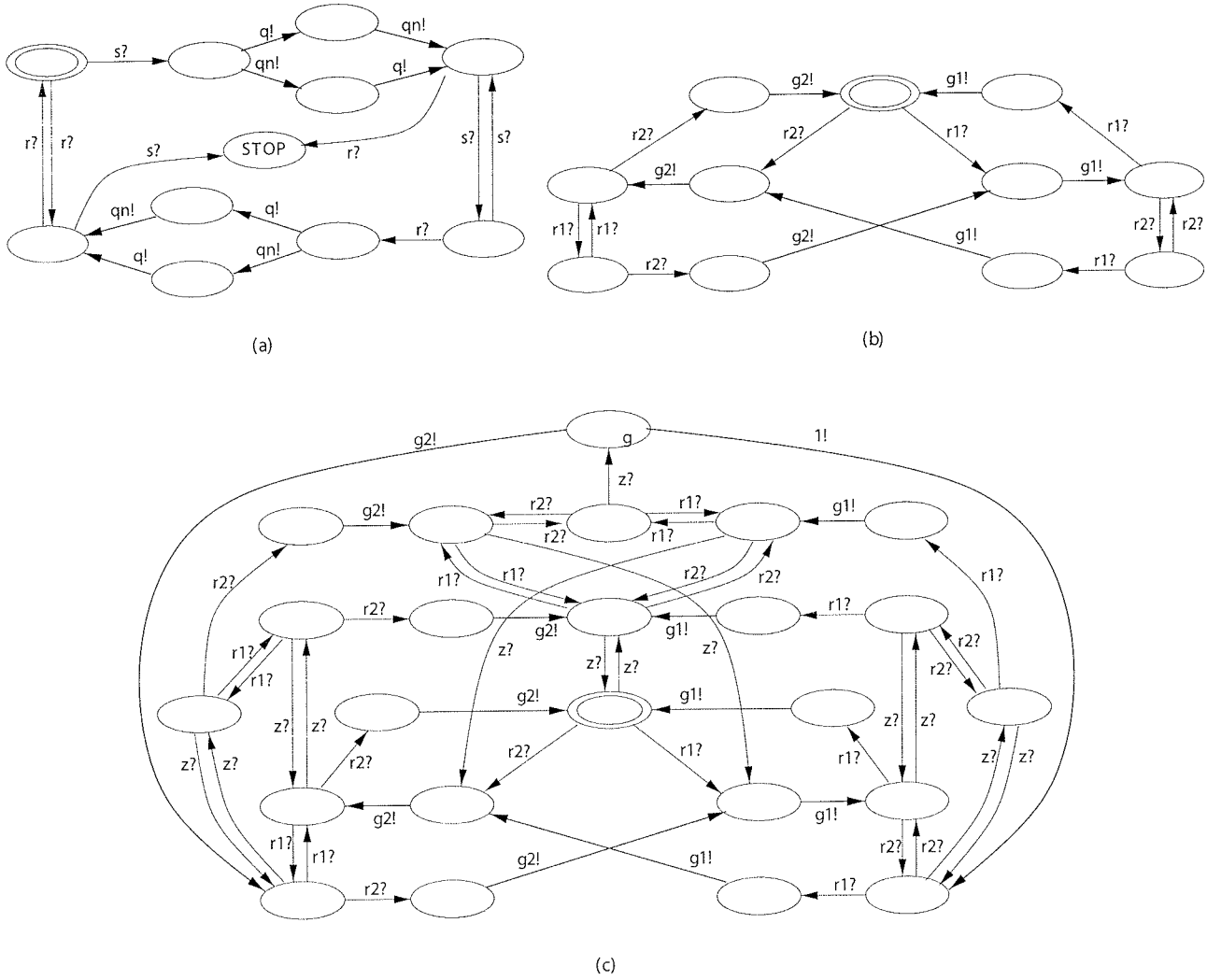


Figure 5: Circuit model of (a) RS flip-flop, (b) ME element, and (c) ME element with inhibit signal.

DME circuits contain some more complex gates, too. The circuit model in Figure 5(a) describes external behaviour of RS flip-flop with two inputs denoted by s (set) and r (reset) and two outputs denoted by q and qn, where qn is inverted q. Initially s, r, and q are equal to 0. When input signal s is set to 1, output q becomes 1. When input signal r is set to 1, then output q becomes 0. Inputs r and s must never be set to 1 at the same time. The circuit model contains a deadlock state, which is not reached if the RS flip-flop is properly driven by its environment.

Mutual exclusion (ME) element is a non-deterministic element used in asynchronous circuits as an arbiter. Two different versions of ME elements appear in DME circuits. In the first one, the ME element has two inputs denoted by r1 and r2 and two outputs denoted by g1 and g2. If input signal r1 is set to 1 and output g2 is not set to 1, then output g1 becomes 1. When input signal r2 is set to 1 and output g1 is not set to 1, then output g2 becomes 1. If inputs r1 and r2 are simultaneously set to 1, then the ME element non-deterministically chooses one output and sets it to 1. In another version, the ME element has an additional input z intended for an inhibit signal, which prevents the grant to new requests until the request that caused the previous grant is removed. The circuit models we used to describe external behaviour of the ME element without and with inhibit signal, where all input and output signals are initially equal to 0, are presented in Figures 5(b) and 5(c), respectively.

4 Verification of asynchronous circuits

The verification of an asynchronous circuit starts by representing all necessary gates with circuit models. Then, they are composed together using *parallel composition with multi-way synchronisation* /8, 20/. The compound process represents external behaviour of the circuit. The behaviour of gates in the circuit is asynchronous, but they are not completely independent. Complementary actions in different circuit models must be performed simultaneously as the value of a signal cannot change only in one part of a wire. Simultaneous performance of complementary ac-

tions is called synchronisation between circuit models. During the composition, two types of transitions with visible actions are distinguished. The transitions used for synchronisation with the system's environment are *external transitions*, whereas the transitions serving for synchronisation between processes in the system are *internal transitions*.

The parallel composition of circuit models will not return a meaningful result if:

- signals do not have unique names,
- the outputs of two or more gates are connected together,
- there exists a gate directly driving itself,
- an output signal observable by the environment is used as a feedback signal into the circuit.

The first two situations are straighforward. A gate is not allowed to drive itself because it is expected that at least two different circuit models cooperate in synchronisation. An output signal observable from the environment is not allowed to be a feedback signal into the circuit because the same visible action cannot be used in internal and external transitions. The last two requirements have an impact on the preciseness of the verification. For example, circuits in Fig. 3(c) and Fig. 3(d) cannot be verified in their original form without adding FORK elements after the OR gates which produce signals y1 and y2.

The next verification step is a transformation of the obtained compound process into a circuit model which represents the external behaviour of the assembled circuit under the fundamental mode of operation. This transformation consists of two steps, removal of redundant traces and determinization of the process. We call the first operation *fundamental-mode reduction* and it removes:

- all transitions with an input action from states where a silent transition can be performed,
- all transitions with an input action from states where a transition with an output action can be performed.
- all silent transitions from states where a transition with an output action can be performed.

An example of the circuit model obtained by composing circuit models of individual gates is presented in Figure 6. It is a circuit model which represents the external behaviour of the circuit with oscilating output under the fundamental mode of operation. In the initial state, it can perform input action x?. After changing input x, output f begins to oscilate. Afterwards, the circuit cannot change the value of input signal again because it never becomes stable.

One of the goals of our approach was hazard detection. It can be done by finding particular patterns in the circuit models which represent unwanted external behaviour. This patterns are simpler for the circuits containing only one output signal. Therefore, in the case of a circuit with many outputs we made the verification separately for each output. A hazard resulting in a glitch is present in the circuit if in its circuit model, after performing an output action, the same output action can be performed again before any input action is performed. Hence, most hazards can be easily detected by looking for such sequences of transitions, although in this way they cannot be classified into logic, functional and transient hazards. However, static and dynamic hazards can be distinguished, as a static hazard results in an output action successively repeated an even number of times, whereas in the case of a dynamic hazard, the output action is repeated an odd number of times.

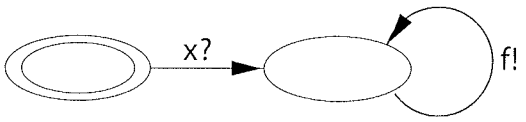


Figure 6: Circuit model of the circuit with oscilat. output

It is more complicated to detect steady-state hazards. A steady-state hazard is present in the circuit if after a particular sequence of input signals with some arrangements of delays an output signal appears, but with other arrangements of delays it does not. This always results in a state in the circuit model where both input and output actions can be performed. However, not all such states are a consequence of steady-state hazards:

- it can also indicate a glitch which appears with some arrangements of delays, but not with all of them,
- if some actions representing output signals have been abstracted from the model, then it can also indicate a non-deterministic behaviour of the circuit, which in particular situations produces the retained output signal and sometimes the abstracted one.

Glitches which appear only with some arrangements of delays and not with all of them are avoidable by engineering delays. We will call them *avoidable* hazards. Not all hazards are avoidable (i.e. they are *unavoidable*) because delays can be set only to gates and not to wires. An example of avoidable and unavoidable static logic hazard is shown in Figure 7. An example of a circuit with non-deterministic behaviour is the ME element with inhibit signal in Figure 5(c).

Circuit models in Figure 8 represent the external behaviour of circuits with hazards. They were obtained by composing circuits in Figure 3. The circuit model in Figure 8(a) shows that in the circuit in Figure 3(a), if the value of a and b is initially assumed to be 0, after changing any of them, output f may change two times consecutively. This represents a static hazard. Afterwards, if the same input is changed again, another static hazard may appear. All hazards in this circuit are avoidable. The circuit model in Figure 8(b) shows that in the circuit in Figure 3(b), if a and b have initial value 0 and the value of a changes, no hazards will occur. However, the situation is quite different after b

is changed to 1. Then, each change of a may be followed by three consecutive changes of output f, which is a dynamic hazard. This hazard is avoidable, too. The circuit model in Figure 8(c) indicates a transient hazard in the circuit in Figure 3(c). If x is initially assumed to be 0 and then changes to 1, a static hazard may appear on ouput y1. The hazard is avoidable and it is possible only after the first change of x. The circuit models in Figure 8(d), 8(e), and 8(f) represent the behaviour of the circuit in Figure 3(d). The initial value of all signals is assumed to be 0. After x changes to 1, the value of y1 with some arrangements of delays changes and with some arrangements it does not. In both cases, further changes of x do not affect y1 anymore. There are no hazards on line y2. Figure 8(d) shows the circuit model containing both output signals, while in the other two circuit models one output signal is abstracted away.

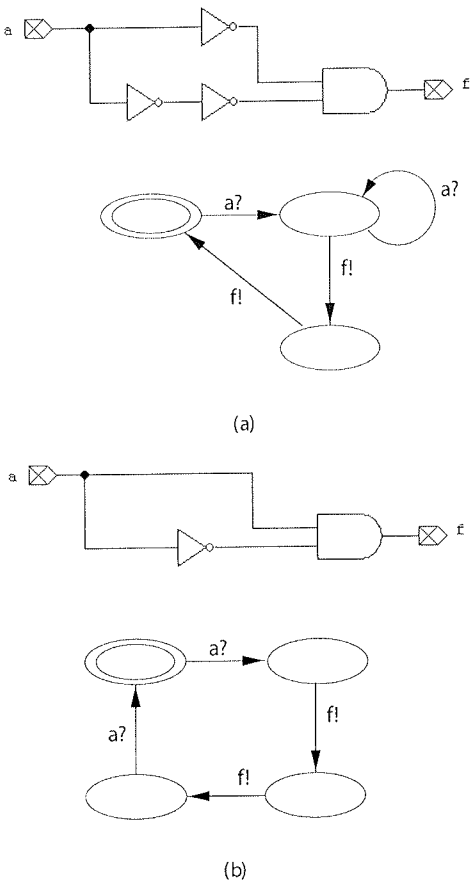


Figure 7: A circuit with (a) avoidable and (b) unavoidable static logic hazard

5 ACTL model checking

In Section 4 we detected hazards in asynchronous circuits by looking for particular patterns in the circuit models. This step of verification can be effectively done by *model checking*, which is a powerful technique for checking properties of processes. It is also used for verification of liveness and safety properties of asynchronous circuits.

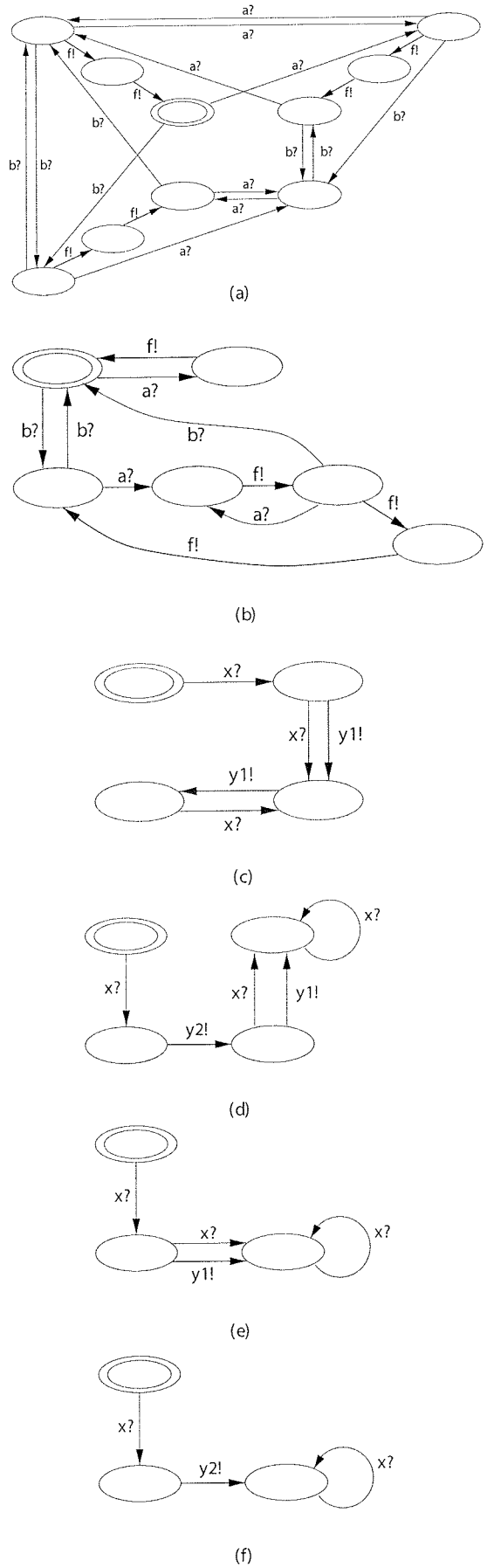


Figure 8: External behaviour of the circuits in Figure 3

We will specify circuit properties with *action computation tree logic* (ACTL), which is a propositional branching time temporal logic /6, 14/. The syntax of ACTL formulae includes constants true and false, standard Boolean operators **NOT, AND, OR**, *path quantifiers* **E** ("there exists a path") and **A** ("for all paths"), and *temporal operators* **U** ("until"),**W**("unless"), **X** ("for the next transition"), **F** ("for some transition in the future"), and **G** ("for all transitions in the future").

ACTL formulae are state formulae. A state where ACTL formula φ is valid will be called φ-state. ACTL formulae are constructed from action and path formulae. An action for which action formula χ is valid will be called χ-action. A transition from state p to state q where action formula χ is valid for the action executed during this transition and ACTL formula φ is valid in state q will be called (χ, φ)-*transition*. In a process, path formulae are evaluated as follows:

- Path formula **X** {χ} φ is valid on path π if the first transition on this path is a (χ, φ)-transition.
- Path formula **F** {χ} φ is valid on path π if there exists a (χ, φ)-transition on this path.
- Path formula **G** φ{χ} is valid on path π if ACTL formula φ is valid in the first state of this path and all transitions on the path are (χ, φ)-transitions.
- Path formula [φ {χ} **U** {χ'} φ'] is valid on path π if ACTL formula φ is valid in the first state of this path and the path begins with a finite sequence of (χ, φ)-transitions followed by a (χ', φ')-transition.
- Path formula [φ {χ} **W** {χ'} φ'] is valid on path π if formula [φ {χ} **U** {χ'} φ'] is valid on this path or formula **G** φ {χ} is valid on this path.

The given rules are used for finite and infinite paths. In ACTL, each path formula is always immediately preceded by a path quantifier. Path quantifier **E** requires that the property expressed by the path formula is valid for at least one path starting in the given state. On the other hand, path quantifier **A** requires that the property expressed by the path formula is valid for all paths starting in the given state. In a deadlock state, formulae **EG** φ {χ}, **AG** φ {χ}, **E** [φ {χ} **W** {χ'} φ'], and **A** [φ {χ} **W** {χ'} φ"] are valid only if the state is a φ-state. Formulae **EX** {χ} φ, **AX** {χ} φ, **EF** {χ} φ, **AF** {χ} φ, **E** [φ {χ} **U** {χ'} φ'] and **A** [φ {χ} **U** {χ'} φ'] are invalid in all deadlock states. We will take that an ACTL formula is valid in process P if it is valid in its initial state.

In ACTL formulae, the constant *true* can be omitted in many cases, for example:

E [true{χ} **U** {χ'} φ'] = **E** [{χ} **U** {χ'} φ']
A [φ {true} **U** {true} φ'] = **A** [φ **U** φ']

There are also two widely accepted abbreviations of ACTL operators:

<χ>φ = {χ} φ
[χ]φ = ¬**EX**{χ} ¬φ

ACTL formula <α>φ is valid in the given state if there exists a transition with action α from that state to a state where ACTL formula φ is valid. ACTL formula [α]φ is valid in the given state if all transitions with action α from that state lead to a state where ACTL formula φ is valid.

Suppose that the alphabet of the process contains actions a?, b? and f?. Here are some ACTL formulae which can be used for checking properties of this process:

- There is no deadlock state: **AG AF** {true}
- At any moment, ouput action f! will be performed in the future: **AG AF** {f!}
- There is no state where output action f! can be performed succesively two times: **NOT EF** {f!} <f!> true
- There is no state where output action f! and also input action a? or b? can be performed: **NOT EF** ((<f!> true) **AND** (<a? **OR** b?>true))

6 Results from verification of DME circuits

We verified *distributed mutual-exclusion* (DME) circuits. They are composed of DME cells connected in a ring. DME cells work by passing a token around the ring. The ownership of the token is determined by output signal q of the RS flip-flop. The token is exchanged via the request and acknowledge signals with the left cell (LR and LA) and with the right cell (RR and RA). The users gain exclusive access to the resource via the request and acknowledge signals UR and UA. The DME circuit was originally proposed by Martin in 1985 /11/. Martin's design does not work correctly under the input/output mode of operation /10/. In 1988, Burns gave a simpler design of the DME cell /3/. It was later slightly modified by McMillan and became a standard benchmark for asynchronous design verification tools /12, 13, 18/. The DME cells from Martin and McMillan are presented in Figure 9.

In Martin's design, the token indicates which user has last accessed the shared resource. If a DME cell receives a request but does not have the token, it notices this to its right neighbour via the RR signal. When a DME cell gets a request, either from the user via the UR signal or from its left neighbour via the LR signal, the DME cell attempts to satisfy the demand. If a DME cell has the token and no granted request is outstanding, then it sends an acknowledgement, either via the UA or LA signal, as appropriate. When a DME cell establishes it can approve user access, it immediately sends the UA signal to the user. If this DME cell does not have the token, then the token is transferred to it after the user removes the request. In McMillan's design, a user request is never acknowledged by a DME cell which does not possess the token. The token is transferred first, which makes the response to a user request slower. Moreover, McMillan's design has slower response times regardless of the token position because there the request signal has a longer path through the decision logic.

The verification was done with *Efficient Symbolic Tools* (EST), our BDD-based tool for symbolic verification of con-

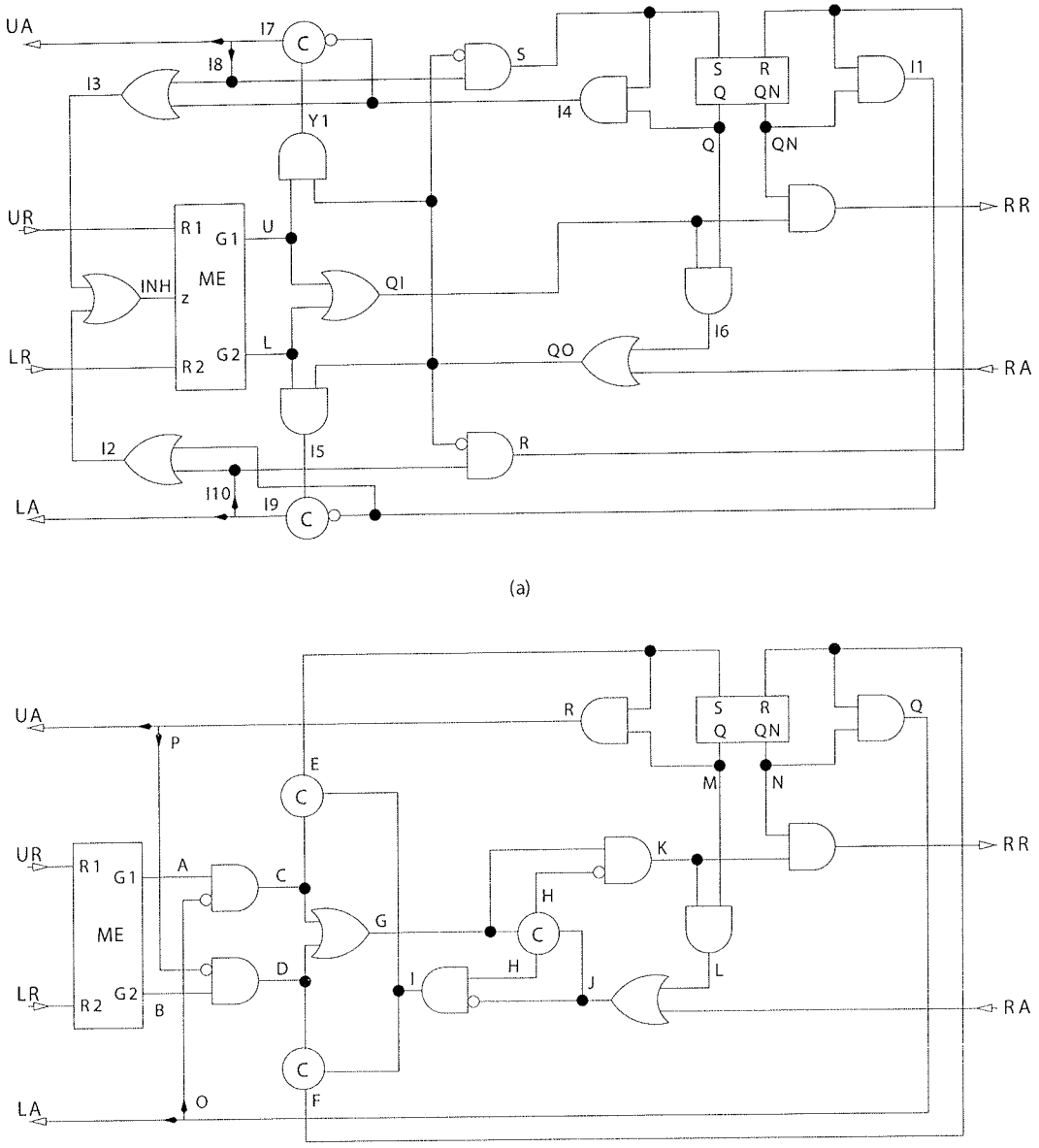
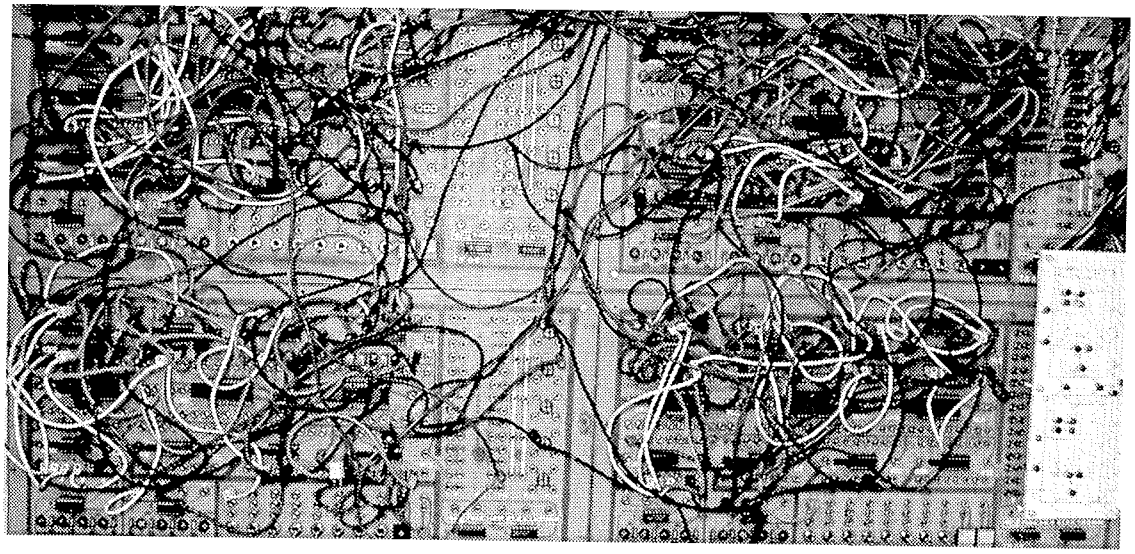


Figure 9: A cell of the DME circuit (a) as proposed by Martin /11/, and (b) as proposed by McMillan /12/

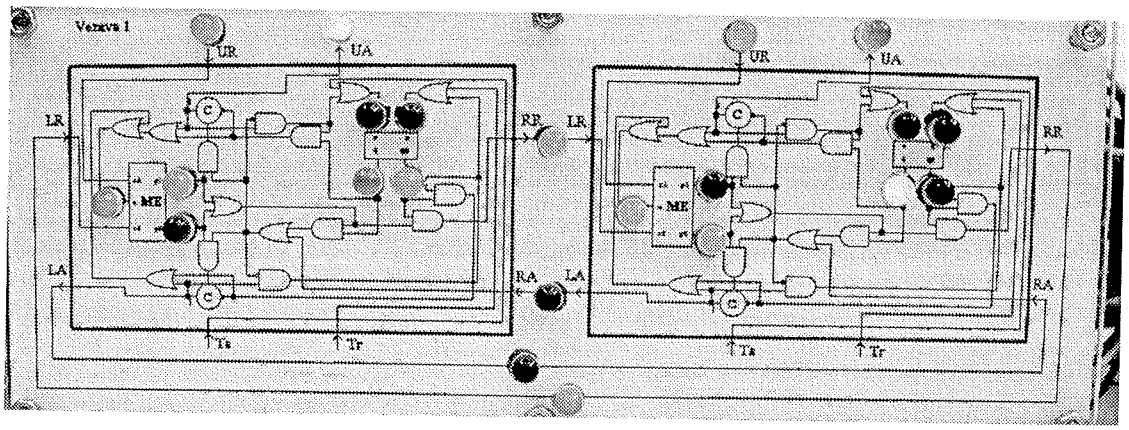
current systems /15/. We started by modelling all necessary gates and composing them in DME cells. Afterwards, possible hazards in each DME cell were examined. Finally, we composed DME cells into rings of different sizes and checked correctness of external behaviour of the obtained circuits. To confirm the results of formal verification we also implemented the circuits on the prototype board (Figure 10) and tested their behaviour by measurements with HP 1652B Logic Analyzer. In Figure 10(b), the reader may notice that some logic for initialisation of the RS flip-flops was added for testing. The test runs obtained for the ring composed of two DME cells are given in Figure 11. Signals UR, UA, Q, RR, LA, Z, S, R, G1, and G2 belong to the first DME cell, while others belong to the second one.

In the circuit model of DME cell, input signal changes were represented with actions ur?, lr?, and ra?, and output signal changes were represented with actions ua!, la!, and rr!. We created 3 different circuit models for each DME cell. Each circuit model represents the behaviour of one output signal, whereas the other two are abstracted away. This makes the verification simpler. To find hazards, we used the following ACTL formulae, which are for simplicity here presented using macros, although EST does not support them yet:

```
\define IN (ur? OR lr? OR ra?)
\define OUT (ua! OR la! OR rr!)

# Static hazards
NOT EF {IN} <OUT> <OUT> <IN> true

# Dynamic hazards
NOT EF {IN} <OUT> <OUT> <OUT> <IN> true

# Steady-state hazards
NOT EF {IN}
((<IN> true) AND (<OUT> <IN> true))
```



(a)

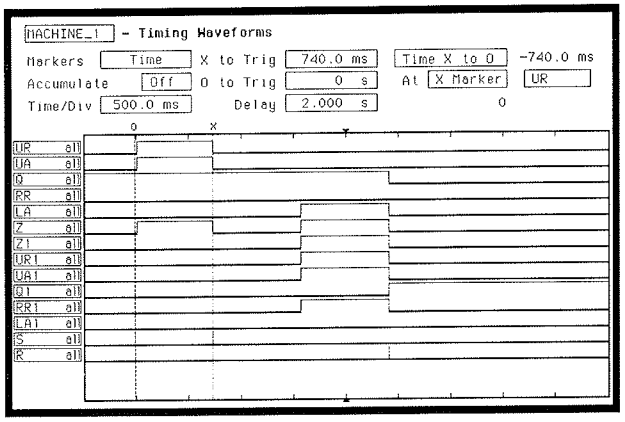


(b)

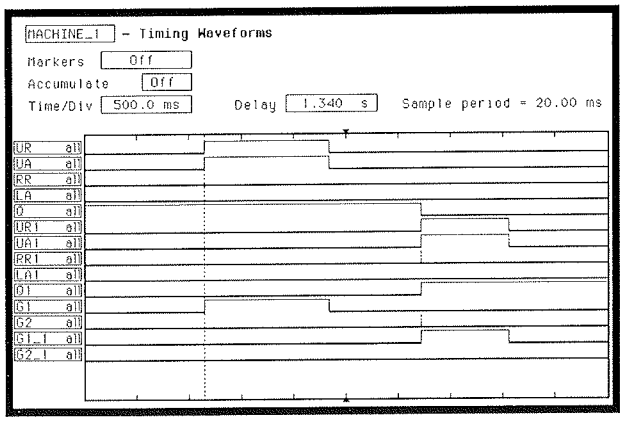
Figure 10: The ring of two Martin's DME cells: (a) implementation with gates and wires, (b) LED indicators for testing

For any formula which is invalid for a given model, EST shows a counterexample. To find all hazards on the particular output signal effectively, model checker should find all counterexamples (e.g. tree-like counterexamples /9/). Unfortunately, EST is not capable of that, and therefore we helped us with an iteration method. For each hazard found, we deleted outgoing transitions from the state where the hazard began and then checked the same ACTL formula again. We repeated this until the formula became valid. With the presented formulae only hazards containing two or three successive changes of an output signal can be detected, but we also verified that the model contains no other hazards. In Martin's DME cell, which initially does not possess the token, we found static hazards on output signals UA (1-4) and RR (5-8) and steady-state hazards on output signals UA (9), LA (10-11), and RR (12-20):

1. ra?,ur?,ua!,ra?,ua!,ua!
2. ra?,ur?,ua!,lr?,ra?,ua!,ua!
3. ra?,ur?,ua!,ur?,ur?,ra?,ua!,ua!
4. ra?,ur?,ua!,ur?,ur?,lr?,ra?,ua!,ua!
5. ur?,rr!,lr?,ur?,rr!,rr!
6. lr?,rr!,ur?,lr?,rr!,rr!
7. ra?,lr?,rr!,ur?,lr?,rr!,lr?,ra?,rr!,ur?,rr!,rr!
8. ra?,lr?,rr!,ur?,lr?,rr!,lr?,ra?,rr!,lr?,rr!,rr!
9. ra?,lr?,ur?,lr?,lr?,ra?,ra?,[ua!]
10. ra?,ur?,ur?,ur?,lr?,ra?,[la!]
11. ra?,lr?,la!,ur?,lr?,lr?,ra?,la!,ra?,[la!]
12. ra?,ur?,rr!,ur?,rr!,ur?,lr?,ra?,lr?,[rr!]
13. ra?,lr?,rr!,ur?,lr?,rr!,lr?,ra?,rr!,ra?,ra?,[rr!]
14. ra?,lr?,rr!,ur?,lr?,rr!,lr?,ra?,rr!,ra?,ur?,[rr!]
15. ra?,lr?,rr!,ur?,lr?,rr!,lr?,ra?,rr!,ra?,lr?,[rr!]
16. ra?,ur?,rr!,ur?,rr!,ur?,lr?,ra?,ra?,lr?,lr?,ra?,[rr!]
17. ra?,ur?,rr!,ur?,rr!,ur?,lr?,ra?,ra?,ur?,lr?,ra?,lr?,[rr!]
18. ra?,ur?,rr!,ur?,rr!,ur?,lr?,ra?,ra?,ur?,ur?,lr?,lr?,ra?,[rr!]
19. ra?,ur?,rr!,ur?,rr!,ur?,lr?,ra?,ra?,ur?,lr?,ra?,ra?,ur?,[rr!]
20. ra?,ur?,rr!,ur?,rr!,ur?,lr?,ra?,ra?,ur?,lr?,ra?,ra?,lr?,[rr!]



(a)



(b)

Figure 11: Test run of the ring composed of (a) two Martin's DME cells, (b) two McMillan's DME cells

In McMillan's DME cell, which initially does not possess the token, we found only static hazards on output signal RR:

1. ra?,ur?,rr!,rr!
2. ra?,lr?,rr!,rr!
3. ur?,rr!,lr?,ur?,rr!,rr!
4. lr?,rr!,ur?,lr?,rr!,rr!

Due to the separate verification for each output signal, counterexamples do not describe the behaviour of all output signals, which can be a drawback. A test run corresponding to the hazard no. 3 in McMillan's DME cell is given in Figure 12. Initially, all input and ouput signals are set to 0 and the DME cell does not possess the token. When the DME cell gets a user request, it immediately sends signal RR to its right neighbour. Afterwards, if LR and UR change before other signals, we get static hazard on signal RR.

The results of the verification show that both designs of DME cell contain hazards even under the fundamental mode of operation. However, McMillan's design is much cleaner. Table 1 gives the number of hazards found. Note that we distinguish two hazards only if they occur in different situation in the circuit. A DME cell can first acquire the token and then deliver it forward, and afterwards it is found in the same situation as in the beginning. Moreover, static hazard in Martin's DME cell described with the sequence ur?,ra?,ua!,ra?,ua!,ua! is treated to be the same as the first one in the given list because the situation in the circuit is the same regardless of the order in which the value of input signals ur? and ra? changes.

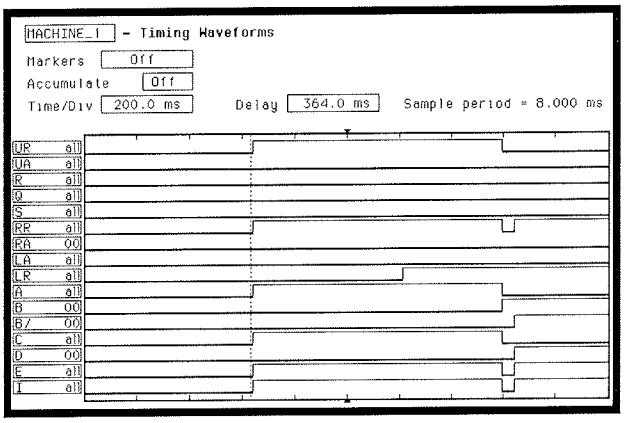


Figure 12: Hazard on signal RR in McMillan's DME cell

Table 1: Hazards in DME cells

	Martin [11]			McMillan [12]		
	UA	LA	RR	UA	LA	RR
Static hazards	4	0	4	0	0	4
Dynamic hazards	0	0	0	0	0	0
Steady-state hazards	1	2	9	0	0	0

In the next step of the verification we checked, whether the DME circuits of different sizes satisfy safety and liveness properties proposed in /13/:

1. An acknowledgement is not given without a request.
2. An acknowledgement is not removed while a request persists.
3. All requests are eventually acknowledged.
4. No two users are acknowledged simultaneously.

After composing DME cells, only signals UR and UA from and, respectively, to different users remain in the model. They were represented with actions ur1?, ua1!, ur2?, ua2!, etc. Because in a circuit model the same action represents either the change of signal value from 0 to 1 or vice versa, the first two properties can be verified with the same ACTL formulae. On the other hand, the last and the most important property cannot be directly expressed with one ACTL formula. We can only express mutual exclusion after a user gets the acknowledgement for a given number of times.

Here are the formulae used for verification of the DME circuit composed of two DME cells:

```
# After an acknowledgement is sent
# (removed), it will not be removed
# (sent) before the user requests this.

AG [ua1!] A[{NOT ua1!} UU {ur1?}]
AG [ua2!] A[{NOT ua2!} UU {ur2?}]

# All requests will be acknowledged.

AG [ur1?] AF {ua1!}
AG [ur2?] AF {ua2!}

# After a user gets the acknowledgement
# for the first time (second time etc.),
# other users will not get an
# acknowledgement until his acknowledgement
# is removed.

\define UA1 {NOT ua1!} UU {ua1!}
\define UA2 {NOT ua2!} UU {ua2!}

A[UA1 A[{NOT ua2!} UU {ua1!}]]
A[UA2 A[{NOT ua1!} UU {ua2!}]]

A[UA1 A[UA1 A[UA1 A[{NOT ua2!} UU {ua1!}]]]]
A[UA2 A[UA2 A[UA2 A[{NOT ua1!} UU {ua2!}]]]]
```

All listed formulae were valid for the circuit composed of two Martin's DME cell and also for the circuit composed of two McMillan's DME cell. Afterwards, we verified DME circuits composed of three and more DME cells, too. To do this, ACTL formulae must have been adequately adapted. Because no incorrect behaviour was detected, we may conclude that both designs of DME cell operate correctly under the fundamental mode of operation. Evidently, with the presented approach, we were not able to detect malfunction of DME circuits composed of Martin's cells because it is the result of delay hazards.

The size of circuit models used during the verification is given in Table 2. The most complex task during the verification was the composition of processes. Table 3 and Table 4 give some statistics about the complexity of parallel composition obtained on a system composed of 800 MHz Athlon processor, 512 MB RAM and Linux OS. The size of DME cells in Table 2 refers to the circuit model describing external behaviour of DME cells which initially do not possess the token. The sizes reported in the last two tables refer to the circuit model describing external behaviour of rings where internal behaviour of DME cells is abstracted away. We were not able to compose more than 4 Martin's and 5 McMillan's DME cells although the resulting process is supposed not to be enormous. Hence, an "on the fly" model checker would be of great interest here.

Table 2: The size of circuit models

Circuit	Inputs / Outputs	States / Transitions	BDD nodes
C element	2/1	4/7	32
RS flip-flop	2/2	11/16	75
ME element [11]	3/2	24/51	164
ME element [12]	2/2	11/16	73
DME cell [11]	3/3	72/148	474
DME cell [12]	3/3	48/104	304

Table 3: Parallel composition of Martin's DME cells

Circuit	States / Transitions	BDD nodes	Time for composition
2 DME cells	11/16	62	0.1s
3 DME cells	57/117	365	2.4s
4 DME cells	236/632	1752	100.2s

Table 4: Parallel composition of McMillan's DME cells

Circuit	States / Transitions	BDD nodes	Time for composition
2 DME cells	11/16	66	0.1s
3 DME cells	40/72	251	0.8s
4 DME cells	145/316	974	12.9s
5 DME cells	596/1545	4444	1299.1s

7 Conclusion

Asychronous circuits are an important class of digital circuits used as autonomous devices or just a part of otherwise synchronous circuits. They have many nice properties, but they are also difficult to design and verify, especially in an ad hoc fashion. Emerging tools for formal verification promise a solution for many problems in this area.

This paper introduces a method for formal verification of asychronous circuits modelled with Muller's model under

the fundamental mode of operation. It is suitable for detecting hazards in the given circuit and for verification of safety and liveness properties. The method is based on the representation of external behaviour of the circuit with processes. A special form of processes called circuit models is used to describe gates. A parallel composition with the ability of multi-way synchronisation and an operation called fundamental-mode reduction serve for construction of circuits from single gates. Determinization of processes assures a canonical form of specifications. The properties of circuits can be checked by ACTL model checking. The results obtained by verification and measurements on real circuits convince us about the correctness of our approach.

We used presented method for the verification of DME circuits. We were able to verify the external behaviour of DME circuits composed of up to 5 DME cells. For larger circuits, BDD-based tool EST exceeded memory limits of 512 MB during the parallel composition of circuit models. Maybe, the impact of state explosion could be moderated by using "on the fly" model checking. Alternatively, parallel composition and fundamental-mode reduction could be united in a more efficient operation.

There are many topics for further research. The ability of model checking to find tree-like counterexamples was mentioned in the paper. We are also interested in experiments with *input-quasi-receptive* circuit models, which describe the behaviour of the circuit under input/output mode of operation. They enable checking semi-modularity. Input-quasireceptive models accept input signals in all states and this leads to more complex specifications. It is a challenge how to apply such an approach to larger asynchronous circuits.

References

/1/ ACiD-WG. Report on Design, Automation and Test for Asynchronous Circuits and Systems, January 2002.

/2/ A. Bailey. Automatic Verification of Speed-Independent Circuit Designs Using the Circal System. In *Correct Hardware Design and Verification Methods (CHARME '93)*, volume 683 of *LNCS*, pages 167-178. Springer-Verlag, May 1993.

/3/ S. M. Burns. Automated Compilation of Concurrent Programs into Self-timed Circuits. Master's thesis, California Institute of Technology, 1988.

/4/ F.-C. Cheng. Exact Essential-Hazard-Free State Minimization of Incompletely Specified Asynchronous Sequential Machines. Technical report CUCS-033-94.

/5/ A. Davis and S. M. Nowick. An Introduction to Asynchronous Circuit Design. Technical Report UUCS-97- 013, University of Utah, September 1997. /6/ A. Fantechi, S. Gnesi, F. Mazzanti, R. Pugliese, and E. Tronci. A Symbolic Model Checker for ACTL. In *Proceedings of FM-Trends'98*, volume 1641 of *LNCS*, pages 228-242. Springer-Verlag, October 1998.

/7/ S. Hauck. Asynchronous Design Methodologies: An Overview. *Proceedings of the IEEE*, 83(1):69-93, January 1995.

/8/ J. He. Formal Specification and Analysis of Digital Hardware Circuits in LOTOS, August 2000. Technical Report CSM-158. University of Stirling.

/9/ Y. Lu. *Automatic Abstraction in Model Checking*. PhD thesis, Department of Electrical and Computer Engineering, Carnegie Mellon University, 2000.

/10/ A. R. Martello. *Temporal Analysis for Time-Bounded Causal Digital Systems*. PhD thesis, University of Pittsburgh, April 1993.

/11/ A. J. Martin. The Design of a Self-timed Circuit for Distributed Mutual Exclusion. In *Proceedings of the 1985 Chapel Hill Conference on VLSI*, pages 245-260, 1985.

/12/ K. L. McMillan. *Symbolic Model Checking*. PhD thesis, Carnegie Mellon University, May 1992. Technical report CMU-CS-92-131.

/13/ K. L. McMillan. The SMV system, November 2000. http://www-2.cs.cmu.edu/_modelcheck/smv.html.

/14/ R. Meolic. Checking correctness of concurrent systems behaviour. Master's thesis, University of Maribor, 1999. In Slovene.

/15/ R. Meolic, T. Kapus, and Z. Brezo˙cnik. The Efficient Symbolic Tools Package. In *Proceedings of the Soft- COM 2000*, volume I, pages 147-156, Split, Croatia, October 2000. http://www.el.feri.uni-mb.si/est/.

/16/ R. Milner. *Communication and Concurrency*. International Series in Computer Science. Prentice Hall, 1989.

/17/ C. Myers. *Asynchronous Circuit Design*. John Wiley & Sons, 2001.

/18/ O. Roig, J. Cortadella, and E. Pastor. Conservative Symbolic Model-Checking of Petri Nets for Speedindependent Circuit Verification, 1994. DAC/UPC Technical Report No. RR-94.

/19/ M. Shams, J. C. Ebergen, and M. I. Elmasry. Asynchronous Circuits. JohnWiley's Encyclopedia of Electrical Engineering. /20/ K. S. Stevens. *Practical Verification and Synthesis of Low Latency Asynchronous Systems*. PhD thesis, Dept. of Computer Science, University of Calgary, Canada, September 1994.

Robert Meolič, Tatjana Kapus,
Bogdan Dugonik, Zmago Brezočnik

Faculty of Electrical Engineering and Computer Science, University of Maribor, Smetanova ulica 17, 2000 Maribor, Slovenia

Prispelo (Arrived): 26.02.2003 *Sprejeto (Accepted): 26.08.2003*

ADAPTIVNA STRUKTURA S POLJI PROGRAMIRNIH VEZIJ ZA IZVEDBO NEREKURZIVNIH DIGITALNIH SIT

Davorin Osebik, Rudolf Babič, Mitja Solar

Fakulteta za elektrotehniko računalništvo in informatiko, Univerza v Mariboru, Maribor, Slovenija

Ključne besede: asinhrono vezje, fundamentalni način, procesna algebra, preverjanje modelov, ACTL

Izvleček: V članku so opisane izvedbe digitalnih FIR sit s polji programirnih FPGA vezjih, ki so primerna za uporabo v adaptivnih aplikacijah digitalnih sistemov. Prikazana je analiza kompleksnosti aparaturne izvedbe adaptivnih struktur digitalnih FIR sit. Podali smo primerjavo vzporedne oblike digitalnega FIR sita v strukturi porazdeljene aritmetike z digitalnim FIR sitom v strukturi koncentrirane aritmetike izvedenim z uporabo zaporedne logike za izvajanje aritmetično logičnih operacij. Za obe strukturi smo opravili analizo naraščanja aparaturne kompleksnosti v odvisnosti od stopnje sita. Z dobljenimi rezultati smo pokazali primernost izvedbe digitalnih FIR sit v strukturi koncentrirane aritmetike z uporabo zaporedne logike za izvajanje aritmetično logičnih operacij.

The Adaptive Structure with FPGA Circuits for Adaptive FIR Digital Filter Realization

Key words: digital signal processing, FIR digital filter, distributed arithmetic, direct calculation of partial products, bit serial multiplier, MAC - Multiply and Accumulate, digit serial FIR filter, FPGA implementation

Abstract: In this article the two mode of hardware implementation of digital FIR filter in adaptive structure implemented with field programmable gate arrays XC4000 is presented. Implementations of digital FIR filter with programmable logic cell array circuits can be realized using different structures. Adaptive application of digital FIR filters requires low complexity and quick enough entry of coefficients and calculation of output word $y(k)$.

These conditions are satisfied by implementation of digital FIR filter in the structure of concentrated arithmetic with distributed adders and implementation in the structure of distributed arithmetic in fully parallel form. The structure of distributed arithmetic in fully parallel form is shown in block diagram in figure 3 and digital FIR filter in the structure of concentrated arithmetic with distributed adders is shown in figure 1.

Fully parallel form of digital FIR filter in the structure of distributed arithmetic

The vector of directly calculated partial sums of coefficients $\mathbf{v}(k\text{-}1)$ is calculated in circuit's unit for up-to-date calculation of partial sum of coefficients. Equation (10) describes calculation of the vector of partial sums of coefficients $\mathbf{v}(k\text{-}1)$. Equations 5, 7 and 9 describe the relationship between input signal $u(k)$, vector of history of input signal $\mathbf{u}(k)$ and vector $\mathbf{b}_u(k\text{-}1)$, where the $b_{u(k-1),i}$ are the bits with values 0 or 1, $b_{u(k-1),B_U}$ is the sign bit and $b_{u(k-1),0}$ is the last significant bit (LSB). The vector $\mathbf{b}_u(k\text{-}1)$ has dimension $B_U \times N$, where B_U is the number of bits of input signal $u(k)$ and N is the number of coefficients of FIR filter. The output word is calculated by equation (13). Tables 1 and 2 show the increase of the number of configuration logic blocks. The increase of configuration logic block depends on number of taps digital FIR filter and number of bits B_H. B_H describes quantization of the taps of digital FIR filter. The hardware complexity of the structure of fully parallel form of digital FIR filter increases with $(\text{const.} \times N \times B_H)^2$.

Digital FIR filter in the concentrated arithmetic

In this chapter the two mode hardware implementation of digital FIR filter in concentrated arithmetic is presented. This implementation of FIR filter uses parallel multipliers or serial multipliers. We constructed serial multiplier based on parallel multiplier. Figure 4 shows a diagram of parallel multiplier. Implementation of the parallel multiplier with two 16 bits long words needs 528 configuration logic blocks. Preliminary estimation of device utilization for part XC4013 is shown on table 3. The parallel multipliers take up the disproportionally large amount of the configuration logic blocks.

The serial multiplier is composed of one B_{U+2} -input adder, one B_U -bit multiplexer, and two B_U -bit register. B_U is the number of bits of input word u(k). One input h_i of this multiplier is in parallel form while other $u(k)$ is bit serial with the least significant bit $b_{u(k-1),0}$ presented first. h_i are the coefficients of digital FIR filter. The output is bit serial, with the least significant bit first. Figure 5 shows the structure of serial multiplier. B_U -bit multiplexer can be replaced by B_U -bit register. Figure 6 shows the structure of serial multiplier without multiplexers. Preliminary estimation of device utilization for part XC4013 of one is shown on table 4 and Preliminary estimation of device utilization for part XC4013 of sixteenth is shown on table 5.

All these hardware structures of digital FIR filter were constructed with OrCAD Express and Xilinx XACT 5.2. Fully parallel form of digital FIR filter with 16-tabs in the structure of distributed arithmetic obtained equally hardware complexity than digital FIR filter with 16-tabs in concentrated arithmetic.

1. Uvod

Digitalna FIR sita, ki jih uporabljamo v adaptivnih sistemih, morajo imeti takšno strukturo, da je možno spremeniti vse njihove koeficiente v času enega otipka vhodnega signala. Pri večini aplikacij digitalnega procesiranja signalov, kakor tudi pri digitalnih FIR sitih, so njihovi osnovni gradniki množilniki. Žal so množilniki za aparaturno izvedbo najkompleksnejši elementi. Pri digitalnih sitih z nespremenljivimi koeficienti učinkovito in enostavno rešujejo problem aparaturne digitalnega FIR sita strukture porazdeljene aritmetike (DA, Distributed Arithmetic) /1, 2/. Pri izračun izhodnega signala po postopku porazdeljene aritmetike uporabljamo tabele, v katerih so zapisane predhodno izračunane

delne vsote koeficientov (LUT, look up table). Algoritem, kjer poteka množenje po postopku porazdeljene aritmetike (DA) s pomočjo tabel (LUT) s predhodno izračunanimi delnimi vsotami koeficientov, se imenuje DALUT. Izhodna vrednost digitalnega FIR sita se izračunava s pomočjo algoritma množenja in akumuliranja (MAC, Multiply and Accumulate) /3/ izhodne vrednosti. Običajno delne vsote koeficientov zapišemo v zunanji pomnilnik

Opisan postopek je uporaben le za sita s približno N=20 koeficienti, saj potrebujemo za sito z N koeficienti 2^N pomnilniških lokacij.

Za izvedbo sit v porazdeljeni aritmetiki višjih stopenj obstaja več načinov za zmanjševanje aparaturne kompleksnosti vezja. Dosedanje raziskave so potekale v smeri zmanjšanja potrebnega pomnilnika za zapis predhodno izračunanih delnih vsot koeficientov. Pomnilnik za zapis delnih vsot koeficientov je možno zmanjšati z 2^N na $2^{N/2}$ pomnilniških lokacij z upoštevanjem simetričnosti koeficientov /4/. Na osnovi simetričnosti koeficientov smo razvili FIR sito z 32 koeficienti s 16 bitno vhodno besedo, dolžina aritmetične enote je bila med 16 in 24 bit. Izvedba in rezultati tako napravljenega sita so prikazani v /5/.

Z razdelitvijo FIR sita na kaskade dosežemo zmanjšanje potrebnih pomnilniških lokacij za zapis delnih vsot koeficientov z 2^N na $2^{N1}+2^{N2}+...+2^{Nm}$, pri tem je N_1 N_2...Nm zapisano število koeficientov posamezni kaskadi. Pri tem mora biti vsota posameznih koeficientov kaskad enaka vsoti vseh koeficientov. Izdelano kaskadno obliko sita z 58 koeficienti, kjer smo še dodatno zmanjšala aparaturno kompleksnost sita z upoštevanjem simetričnosti koeficientov, smo skupaj z dobljenimi rezultati meritev predstavili v /6/.

Z razvojem zmogljivejših programirnih FPGA vezij se je pokazalo, da se je zunanjemu pomnilniku za zapis delnih vsot koeficientov možno izogniti /12/. Za ta namen lahko uporabimo strukture digitalnih FIR sit v porazdeljeni aritmetiki, kjer v več pomnilniške lokacije, ki se nahajajo v FPGA vezju, vpišemo le karakteristične delne vsote. Ostale delne vsote aritmetika digitalnega FIR sita izračunava sproti /8/. Izdelano digitalno FIR sito, ki je temeljilo na karakterističnih delnih vsota koeficientov in dobljene rezultate smo prikazali v /9/. Takšna oblika digitalnega FIR sita je uporabna tudi v aplikacijah adaptivnih digitalnih FIR sit.

Za digitalna FIR sita v adaptivnih aplikacijah digitalnih sistemov, ki temeljijo na porazdeljeni aritmetiki, je uporaben postopek sprotnega izračuna delnih vsot koeficientov /10,11/. Ta oblika digitalnega FIR sita se imenuje tudi paralelna oblika digitalnega FIR sita v porazdeljeni aritmetiki (ang. fully parallel DA mode).

Pri adaptivnih digitalnih FIR sitih, kjer se koeficienti sita spremenijo v vsakem otipku, so zanimive strukture digitalnih FIR sit v koncentrirani aritmetiki, ki vsebujejo množilnike. V članku bosta predstavljena dva pristopa izvedbe digitalnega FIR sita s programirnimi FPGA vezji: digitalna FIR sita v strukturi porazdeljene aritmetike s sprotnim izračunom delnih vsot koeficientov in digitalna FIR sita v strukturi koncentrirane aritmetike. Pri predstavitvi FIR sit v koncentrirani aritmetiki bomo posebej nakazali možnosti izvedene z zaporedno logiko za izvajanje aritmetično logičnih operacij.

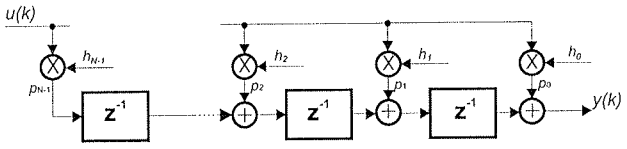
2. Digitalna FIR sita

Osnovno strukturo digitalnega nerekurzivnega digitalnega sita opisuje konvolucijska enačba,

$$y(k)=\sum_{i=0}^{N-1} h_i u(k-i). \qquad (1)$$

Za izvedbo digitalnih FIR sit s programirnimi FPGA vezji imamo na razpolago več različnih struktur. Digitalna FIR sita, ki jih želimo uporabljati v adaptivnih aplikacijah, morajo ustrezati naslednjim zahtevam: imeti morajo dovolj majhno aparaturno kompleksnost, dovolj hiter vpis koeficientov v strukturo FIR sita, dovolj hiter izračun izhodne besede.

Poznani sta dve osnovni obliki nerekurzivnih digitalnih FIR sit, kjer izračun izhodne vrednosti $y(k)$ poteka po konvolucijski enačbi (1). Prva je struktura s porazdeljenim seštevalnikom, ki jo prikazuje slika 1.



Slika 1: Digitalno FIR sito v strukturi koncentrirane aritmetike s porazdeljenimi seštevalniki

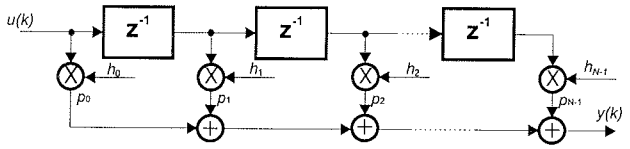
Fig. 1: The digital FIR filter in the structure concentrated arithmetic structure with distributed adders

Pri tej strukturi digitalnega FIR sita s porazdeljenim seštevalnikom običajno uporabljamo takšne algoritme za izvedbo aritmetično logičnih operacij, ki temeljijo na množenju in akumuliranju izračunane vrednosti (MAC, Multiply and Accumulate). Dobljene vrednosti posameznih produktov p_m (m=0, 1 ... N-1) zakasnimo z uporabo zakasnilnih členov z^{-1}. Zaradi tega je za izračun vsake vsote posameznega produkta na razpolago celotna perioda vzorčenja vhodne besede. Skupaj z zakasnilnimi členi, je pri izračunu izhodne besede možno enostavno uporabiti postopek cevljenja.

Druga struktura je digitalno FIR sito s skupnim globalnim seštevalnikom na izhodu, ki jo prikazuje slika 2.

Pri tej strukturi digitalnega FIR sita s skupnim globalnim seštevalnikom je potrebno vse dobljene produkte p_i iz množilnikov v enem taktu sešteti, kar pri sitih višjih stopenj za aparaturno izvedbo ni enostavno.

Obe predstavljeni strukturi digitalnega FIR sita s slik 1 in 2 vsebujeta množilnike. Digitalna FIR sita potrebujejo za veliko slabljenje v zapornem pasu veliko število koeficientov,



Slika 2: Digitalno sito realizirano v strukturi koncentrirane aritmetike s skupnim globalnim seštevalnikom

Fig. 2: The digital FIR filter in the structure concentrated arithmetic with global adder

s tem pa posledično tudi veliko število množilnikov. Aparaturna kompleksnost digitalnega FIR sita v koncentrirani aritmetiki se zaradi množilnikov zelo poveča. Za digitalna FIR sita uporabljena v adaptivnih aplikacija je možno zmanjšanje aparaturne kompleksnosti doseči z:

- izvedbo digitalnega FIR sita v strukturi porazdeljene aritmetike s sprotnim izračunom delnih vsot koeficientov in,
- z izvedbo digitalnega FIR sita v strukturi koncentrirane aritmetike z uporabo zaporedne logike za izvajanje aritmetično logičnih operacij /3/.

3. Digitalno FIR sito v strukturi porazdeljene aritmetike s sprotnim izračunom delnih vsot koeficientov

Digitalno FIR sito, ki se uporablja v adaptivnih aplikacijah digitalnih sistemov je možno izvesti v strukturi porazdeljene aritmetike s sprotnim izračunom delnih vsot koeficientov. Takšna oblika digitalnega FIR sita je poznana tudi kot paralelna oblika digitalnega FIR sita v strukturi porazdeljene aritmetike. To obliko FIR digitalnega sita smo že uspešno uporabili v adaptivnem digitalnem FIR situ s 16 koeficienti, s 16 bitno kvantizacijo vhodne-izhodne besede in med 16 in 24 bitno širino notranje aritmetične enote.

3.1. Matematični opis izračuna izhodne besede pri FIR situ v porazdeljeni aritmetiki

Pri aparaturnih izvedbah digitalnih FIR je vrednost izhodne besede $y(k)$ zakasnjena zaradi časa, ki ga zahteva digitalno procesiranje signala zakasnjena za en otipek. To zakasnitev opišemo z enačbo (2),

$$y(k)=\sum_{i=0}^{N-1} h_i \cdot u(k-1-i) \qquad (2)$$

V enačbi (1) so h_i koeficienti digitalnega FIR sita, N je število vseh koeficientov in $u(k)$ je vhodna vrednost FIR sita. Konvolucijsko enačbo (2) , ki podaja zvezo med izhodno vrednostjo $y(k)$, koeficienti FIR sita h_i in vhodno vrednostjo $u(k)$, lahko predstavimo tudi v vektorski obliki z enačbo (3),

$$y(k)=\mathbf{h}^T\mathbf{u}(k-1) \qquad (3)$$

V enačbi (3) je **h** vektor koeficientov digitalnega FIR sita in **u**(k-1) je vektor predhodnih vrednosti otipkov vhodnega signala $u(k)$. Pri digitalnem FIR situ z N koeficienti enačba (4) so komponente obeh vektorjev podane z:

$$\mathbf{h}=\begin{bmatrix} h_0 \\ h_1 \\ \vdots \\ h_{N-1} \end{bmatrix} \text{ in } \mathbf{u}(k-1)=\begin{bmatrix} u(k-1) \\ u(k-2) \\ \vdots \\ u(k-N) \end{bmatrix}. \qquad (4)$$

Pri izračunu izhodne vrednosti $y(k)$ digitalnega FIR sita v porazdeljeni aritmetiki s sprotnim izračunom delnih vsot koeficientov izhajamo iz enačbe (3). Po postopku porazdeljene aritmetike je potrebno opraviti množenje vektorja koeficientov digitalnega FIR sita **h** in vektorja koeficientov vhoda **u**(k-1). Pri tem postopku vrednosti vhodnega signala $u(k)$ zapišemo v bitni obliki z dvojiškim komplementom. Zapis za B_U bitno kvantizacijo vhodnega signala $u(k)$ podaja enačba (5),

$$u(k)=-b_{u(k),0}+\sum_{i=1}^{B_u-1} b_{u(k),i}\,2^{-i} \qquad (5)$$

V enačbi (5) najbolj utežni bit $b_{u(k-1),0}$ predstavlja predznak k-tega otipka vhodne besede. Zaradi časa, ki je potreben za postopek digitalnega procesiranje izhodne besede digitalnega FIR sita $y(k)$, je potrebno pri opisu z enačbama upoštevati zakasnitev za vrednost enega otipka. Zakasnitev je predstavljena v enačbi (2). Z upoštevanjem te zakasnitve obravnavamo vrednost vhodnega signala digitalnega FIR sita $u(k)$ zakasnjenega za en otipek. Zato enačba (5) za zapis vhodnega signala v bitni obliki preide v enačbo (6),

$$u(k-1)=-b_{u(k-1),0}+\sum_{i=1}^{B_u-1} b_{u(k-1),i}\,2^{-i} \qquad (6)$$

Trenutni otipek vhodnega signala $u(k)$ in njegove predhodne vrednosti $u(k$-1), $u(k$-2)... lahko predstavimo z vektorjem vhodnega signala **u**(k) Vektor vhodnega signala **u**(k) v bitni obliki z dvojiškim komplementom predstavimo z naslednjimi komponentami, ki jih podaja enačba (7),

$$\mathbf{u}(k)=\begin{bmatrix} b_{u(k),0}+\sum_{i=1}^{Bu-1} b_{u(k),i}\,2^{i} \\ b_{u(k-1),0}+\sum_{i=1}^{Bu-1} b_{u(k-1),i}\,2^{i} \\ \vdots \\ b_{u(k-N-1),0}+\sum_{i=1}^{Bu-1} b_{u(k-N-1),i}\,2^{i} \end{bmatrix} \qquad (7)$$

V enačbi (7) predstavlja zapis $b_{u(k),i}$ i-ti bit vhodne besede $u(k)$ digitalnega FIR sita. Pri tem je število bitov $b_{u,i}$ enako širini vhodne besede $u(k)$. V našem primeru jo označuje spremenljivka B_u. Z upoštevanjem zakasnitve vhodnega signala zaradi zaporedno vzporedne pretvorbe dobimo vektor vhodnega signala zakasnjen za en otipek **u**(k-1). Ta

$$\mathbf{u}(k-1)=\begin{bmatrix} b_{u(k-1),0}+\sum_{i=1}^{Bu-1} b_{u(k-1),i}2^{i} \\ b_{u(k-2),0}+\sum_{i=1}^{Bu-1} b_{u(k-2),i}2^{i} \\ \vdots \\ b_{u(k-N),0}+\sum_{i=1}^{Bu-1} b_{u(k-N),i}2^{i} \end{bmatrix} \quad (8)$$

vektor vhodnega signala zapišemo prav tako v bitni obliki z dvojiškim komplementom z enačbo (8),

Iz enačbe (8) izpišimo vrednosti bitov $b_{u,i}(k)$ za vsak otipek vhodne besede $u(k)$ in njihove predhodne vrednosti. Dobljene vrednosti bitov predstavimo z vektorjem bitnega zapisa $\mathbf{b}_u(k\text{-}1)$ vhodnega signala $u(k)$ in predhodnih vrednosti vhodnega signala $u(k\text{-}1)$.

$$\mathbf{b}_u(k-1)=\begin{bmatrix} b_{u(k-1),0} & b_{u(k-1),1} & \cdots & b_{u(k-1),Bu-1} \\ b_{u(k-2),0} & b_{u(k-2),1} & \cdots & b_{u(k-2),Bu-1} \\ \vdots & \vdots & \vdots & \vdots \\ b_{u(k-N),0} & b_{u(k-N),1} & \cdots & b_{u(k-N),Bu-1} \end{bmatrix} \quad (9)$$

Pri tem dimenzija vektorja bitnega zapisa $\mathbf{b}_u(k\text{-}1)$ predhodnih vrednosti vhodnega signala znaša $Bu \times N$. B_u je število bitov za zapis vhodne besede $u(k)$, N je število koeficientov digitalnega FIR sita. S produktom vektorja bitnega zapisa zgodovine vhodnega signala $\mathbf{b}_u(k\text{-}1)$ in vektorja koeficientov $\mathbf{h}$ dobimo vektor delnih vsot koeficientov $\mathbf{v}(k\text{-}1)$, ki ga podaja enačba (10),

$$\mathbf{v}(k-1)=\mathbf{b}_{u(k-1)}^{T}\mathbf{h} \quad (10)$$

Dobljeni vektor delnih vsot koeficientov $\mathbf{v}(k\text{-}1)$ ima N komponent, ki so predstavljene z enačbo (11),

$$\mathbf{v}(k-1)=\begin{bmatrix} v_0(k-1) \\ v_1(k-1) \\ \vdots \\ v_{N-1}(k-1) \end{bmatrix} \quad (11)$$

Vektor dobljenih delnih vsot $\mathbf{v}(k\text{-}1)$ je potrebno izračunati v vsaki k-ti taktu periodi vzorčenja vhodnega signala. Izhodno vrednost digitalnega FIR sita dobimo s pomočjo bitnega zapisa $\mathbf{b}_u(k\text{-}1)$ trenutnih in predhodnih vrednosti vhodnega signala $u(k)$. Pri tem lahko vektor $\mathbf{b}_u(k\text{-}1)$, ki ga opisuje enačba (9) razdelimo na posamezne vrstice i. Vsaka i-ta vrstica vsebuje vektor bitov $\mathbf{b}_u(k\text{-}i)$, ki je potreben za zapis otipka vhodne besede $u(k)$. Zapis tako razdeljenega vektorja $\mathbf{b}_u(k\text{-}1)$ na vektorje v posameznih i-tih vrsticah $\mathbf{b}_{u(k-i),bu}$ podaja enačba (12),

$$\mathbf{b}_{u(k-i)}=\begin{bmatrix} b_{u(k-i),0} & b_{u(k-i),1} & \cdots & b_{u(k-i),Bu} \end{bmatrix} \quad (12)$$

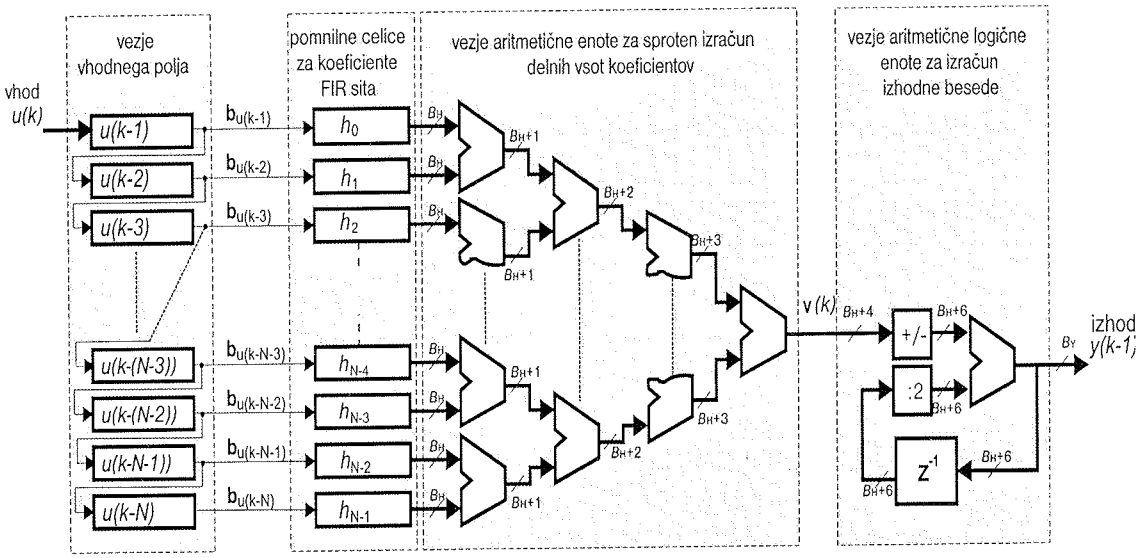
V enačbi (12) i teče od $i=1\ldots N$ in označuje i-te otipke vhodnega signala $u(k)$. Dobljeni vektor $\mathbf{b}_{u(k-i)}$ je osnova za izračun izhodne vrednosti digitalnega FIR sita $y(k)$. Izračun izhodne vrednosti digitalnega FIR sita poteka po enačbi (13).

$$y(k)=\sum_{i=1}^{Bu-1} v_i(k-1)2^{-i}-v_{Bu}(k-1) \quad (13)$$

Enačba (13) opisuje izračun vseh delnih vsot koeficientov. Za izračun izhodne besede po enačbi (13) je potrebno B_u iteracij.

3.2. Izvedba in delovanje paralelne oblike digitalnega FIR sita v strukturi porazdeljene aritmetike

S programirnim FPGA vezjem smo realizirali paralelno obliko digitalnega FIR sita v strukturi porazdeljene aritmetike. Blokovno shemo prikazuje slika 3.



Slika 3: Vzporedna oblika digitalnega FIR sita v strukturi porazdeljene aritmetike

Fig. 3: Fully parallel form of digital FIR filter in the structure of distributed arithmetic

Digitalno FIR sito, ki ga prikazuje slika 3, smo izdelali s programirnim FPGA vezjem družine XC4000E. Predstavljeno vezje digitalnega FIR sita ima N=16 koeficientov, vhodna beseda $u(k)$ je dolžine B_U=16 bitov izhodne beseda $y(k)$ je dolžine B_Y=16 bitov, koeficienti digitalnega FIR sita so dolžine B_H=16 bitov in dolžina notranje aritmetične enote se zaradi robustnosti giblje med 16 in 22 biti.

Vzporedno obliko digitalnega FIR sita s slike 3 sestavljajo štiri različna vezja: vezje vhodnega polja, vezje pomnilnih celic za koeficiente FIR sita, vezje aritmetične enote za sprotni izračun delnih vsot koeficientov in vezje aritmetične enote za izračun izhodne besede. Tabela 1 prikazuje naraščanje kompleksnosti posameznih vezij glede na število koeficientov digitalnega FIR sita N, kvantizacijo vhodne besede B_U oziroma kvantizacijo zapisa koeficientov digitalnega FIR sita B_H.

Tabela 1: Kompleksnost posameznih vezij v vzporedni obliki FIR sita v strukturi porazdeljene aritmetike

Table 1: The complexity of particular circuits in the fully parallel form of digital FIR filter realization in distributed arithmetic

Ime vezja	Kompleksnost vezja
vezje vhodnega polja	konst. xNxB_U
vezje pomnilnih celic koeficientov sita	konst. xNxB_H
vezje aritmetične enote za sproten izračun delnih vsot koeficientov	(konst.xNxB_H)^2
vezje aritmetične lo	konst. xB_H

Iz tabele 1 je razvidno, da kompleksnost vezja aritmetične enote za sprotni izračun koeficientov narašča eksponencialno, zaradi tega je takšna oblika digitalnega FIR sita primerna le za aparaturne izvedbe sit z manj kot N=20 koeficientov.

Pri predstavljeni vzporedni obliki digitalnega FIR sita v strukturi porazdeljene aritmetike je dosežen čas izračuna izhodne besede $y(k)$ pri frekvenci osnovne ure 32 MHz znašal 1 μs. V tabeli 2 podajamo kompleksnost predstavljene vzporednega digitalnega FIR sita v porazdeljeni aritmetiki izvedenega s programirnim FPGA vezjem družine XC4000E.

Tabela 2: Zasedenost programirnega FPGA vezja pri izvedbi vzporedne oblike digitalnega FIR sita v strukturi porazdeljene aritmetike

Table 2: The programmable FPGA device utilization for fully parallel form of digital FIR filter in the structure of distributed arithmetic

Preliminary estimate of device utilization for part 4013EPG223:

55% utilization of I/O pins.	(105 of 192)
82% utilization of CLB FG function generators.	(946 of 1152)
3% utilization of CLB H function generators.	(20 of 576)
92% utilization of CLB flip-flops.	(1058 of 1152)
17% utilization of bus resources.	(16 of 96)

4. Digitalna FIR sita v strukturi koncentrirane aritmetike z uporabo zaporedne logike za izvajanje aritmetično logičnih operacij

Množilniki so eden najpomembnejših elementov v algoritmih za digitalno procesiranje signalov. Pri digitalnih FIR sitih s konstantnimi koeficienti, postopek porazdeljene aritmetike elegantno reši izvedbo produkta dveh matrik. Za digitalna FIR sita višjih stopenj v adaptivnih sistemih je uporaba porazdeljene aritmetike zaradi naraščajoče aparaturne kompleksnosti zahtevna. Zato so zanimiva v adaptivnih aplikacijah za digitalno procesiranje signalov digitalna FIR sita izvedena v strukturi koncentrirane aritmetike z uporabo množilnikov.

Pri digitalnih FIR sitih izvedenih v strukturi koncentrirane aritmetike je potrebno pri vsakem otipku opraviti produkt vektorjev koeficientov digitalnega FIR sita $\mathbf{h}^T$ in vektorja koeficientov vhodnega signala $\mathbf{u}(k)$, po enačbi (3). Za izvedbo tega produkta potrebujemo pri FIR digitalnem situ z N koeficienti N množilnikov. Za izvedbo teh množilnikov je možno uporabiti zaporedno logiko za izvajanje aritmetično logičnih operacij /1/. S tem močno zmanjšamo aparaturno kompleksnost izvedbe takšnega sita, pri tem se čas izračuna izhodne besede bistveno ne spremeni. Pri uporabi zaporedne logike za izvajanje aritmetično logičnih operacij, ločimo različne izvedbe množilnikov glede na obliko zapisa množenca in množitelja: oba množenec in množitelj sta zapisana vzporedni obliki in množenec je zapisan v zaporedni obliki, množitelj je zapisan v vzporedni obliki.

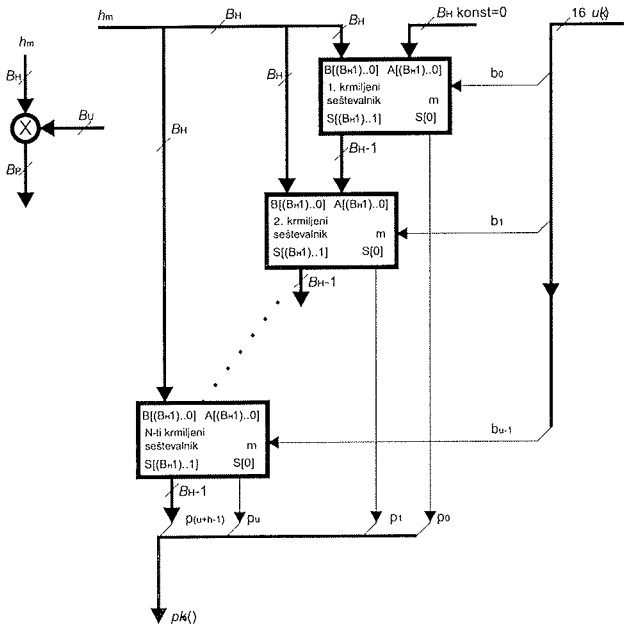
Uporaba zaporedne logike za izvajanje aritmetično logičnih operacij v digitalnem FIR situ ima naslednje prednosti pred paralelno obliko digitalnega FIR sita v strukturi porazdeljene aritmetike:

- kompleksnost aparaturne opreme narašča linearno s številom koeficientov sita,
- zaradi narave seštevalnikov v programirnih FPGA vezjih je čas izračuna izhodne besede $y(k)$ digitalnega FIR sita primerljiv s časom izračuna v vzporedni obliki digitalnega FIR sita v strukturi porazdeljene aritmetike in
- v primeru uporabe digitalnih FIR sit v adaptivnih aplikacijah je ugoden zaporedni prenos koeficientov v strukturo digitalnega FIR sita. Zato z uporabo zaporedne logike za izvajanje aritmetično logičnih operacij ne potrebujemo pretvornikov za pretvorbo zaporedne oblike koeficientov FIR sita v vzporedno obliko koeficientov FIR sita, kar še dodatno zmanjšuje aparaturna kompleksnost vezja.

Predstavili bomo dve obliki množilnikov, ki jih je možno izvesti s programirnimi FPGA vezji: vzporedni množilnik, pri katerem sta množenec in množitelj zapisana v vzporedni obliki in zaporedni množilnik, kjer je množenec zapisan v vzporedni obliki množitelj pa je podan v zaporedni obliki.

4.1. Vzporedni množilnik pri digitalnem FIR situ v koncentrirani aritmetiki

Naša izvedba zaporednega množilnika.je temeljila na vzporednem množilniku. Pri tej obliki množilnika sta množenec, v našem primeru h_i in množitelj $u(k)$ zapisana v vzporedni obliki. Dobljena izhodna vrednost p_i je prav tako v vzporedni obliki dolžine B_H+B_U bitov. Pri tem je množenec dolžine B_H bitov, množitelj pa dolžine B_U bitov. Blokovno shemo takšnega množilnika prikazuje slika 4.



Slika 4: Blokovna shema vzporednega množilnika
Fig. 4: The schematic of parallel multiplier

Za aparaturno izvedbo predstavljenega množilnika je potrebno pri digitalnem FIR situ z N koeficienti, zagotoviti $2\times N$ vhodnih vodil dolžine B_U-bitov oz B_H bitov in N izhodnih vodil dolžine (B_H+B_U)-bitov. Za sito z N=16 koeficienti in dolžinami množitelja B_U=16 bitov in množenca B_H=16 bitov to znaša okoli 768 povezav, kar je za aparaturno izvedbo s FPGA vezjih precej. Čas izračuna zmnožka p_i je odvisen od hitrosti seštevalnikov v FPGA strukturi. V programirnih FPGA vezjih družine XC4000E so večbitni seštevalniki izvedeni v zaporedni obliki, zato je potrebni čas izračuna vsote dveh besed dolžine16-bitov 12,5ns. Pri množilniku, kjer opravimo produkt dveh 16-bitnih števil, je uporabljenih 16 seštevalnikov. Čas izračuna produkta znaša 16×12,5ns. Aparaturna izvedba vzporednega množilnika v programirnem FPGA vezju je zahtevna, saj zahteva B_U multipleksorjev z dolžino B_H bitov. Multipleksorji so krmiljen z enim izmed bitov množitelja $u(k)$. Delovanje vzporednega množilnika za m-ti koeficient opisuje enačba (14),

$$p_m(k) = h_m u(k-m) \qquad m = 0,1,2...N-1 \tag{14}$$

V enačbi (14), je vrednost množitelja $u(k)$ zapisana v dvojiški obliki z enačbo (5). Posamezne dobljene produkte p_m zapišimo v vektorski obliki z enačbo (15),

$$p(k) = \begin{bmatrix} p_0(k) \\ p_1(k) \\ \vdots \\ p_{N-1}(k) \end{bmatrix} \tag{15}$$

Z upoštevanjem enačbe (14), zapišemo enačbo (16) v vektorski obliki. Zapis podaja enačba (16),

$$\mathbf{p}(k) = \mathbf{h}\mathbf{u}(k) \tag{16}$$

V enačbi (16) imata vektorja koeficientov FIR sita **h** in vektor predhodnih vrednosti vhodnega signala **u**(k) komponente, ki so podane z izrazom (4). Vrednost produkta nastaja v krmiljenih seštevalnikih. Delovanje m-tega seštevalnika s slike 4 izmed N seštevalnikov opisuje enačba (17),

$$S_{n,m} = \begin{cases} A_n + 0 & b_{u(k),m} = 0 \\ A_n + B_n & b_{u(k),m} = 1 \end{cases} \tag{17}$$

V enačbi (17) predstavlja spremenljivka $S_{n,m}$ dobljeno vsoto za posamezni bit $b_{u(k),m}$ vhodne besede $u(k)$. Pri tem posamezni bit $b_{u(k),m}$ krmili posamezne seštevalnike. Za vsak krmiljen seštevalnik potrebujemo krmilno vezje sestavljeno iz polja B_U IN vrat.

Dobljene vrednosti produktov p_m iz vzporednih množilnikov je potrebno sešteti v N-tih vzporednih seštevalnikih. Pri tem imamo dvoje možnosti: uporabiti strukturo s porazdeljenimi seštevalniki (slika , ki je primernejša za aparaturno izvedbo ali uporabiti strukturo z enim globalnim seštevalnikom (slika 2). V tabeli podajamo zasedenost programirnega FPGA vezja XC4013 pri izvedbi množilnika s 16. bitnima vhodnimi besedama.

Tabela 3: Zasedenost programirnega vezja pri izvedbi vzporednega množilnika s 16 bitnima vhodnima besedama
Table 3: The programmable FPGA device utilization for 16-bit parallel multiplier

Preliminary estimate of device utilization for part 4013PG223:	
42% utilization of I/O pins.	(80 of 192)
46% utilization of CLB FG function generators.	(528 of 1152)
0% utilization of CLB H function generators.	(0 of 576)
0% utilization of CLB flip-flops.	(0 of 1152)

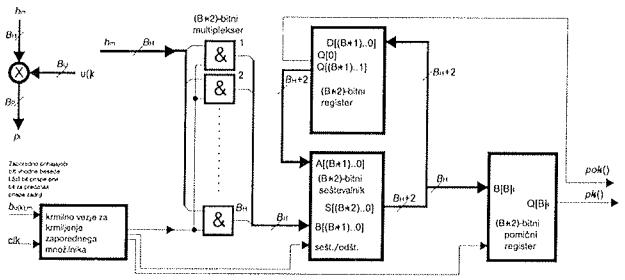
Zaradi zaporedne narave seštevalnikov v programirnih FPGA vezjih traja izračun produkta dveh 16 bitnih števil 170ns. Izvedba enega množilnika v programirnem FPGA vezju zaseda približno polovico logičnih blokov, ki jih zaseda vzporedna oblika digitalnega FIR sito v strukturi porazdeljene aritmetike, ki je predstavljena na sliki 3. Zaradi takšne aparaturne kompleksnosti je smiselno uporabiti zaporedne množilnike, ki temeljijo na MAC strukturah.

4.2. Zaporedni množilnik pri digitalnem FIR situ v koncentrirani aritmetiki

Izvedba množilnikov v zaporedni obliki močno zmanjša aparaturno kompleksnost vezja. Čas izračuna izhodnega produkta p_m je odvisen od frekvence osnovne ure s katero deluje zaporedne množilnik in od kvantizacije B_U množitelja $u(k)$. S primerno izbiro ure in kvantizacije vhodne besede, se čas izračuna izhodnega produkta p_m v zaporednem množilniku bistveno ne razlikuje od časa izračuna produkta p_m v vzporednem množilniku. Pri vzporednem množilniku poteka izračun produkta p_m po enačbi (14). Če v enačbi (14) zamenjamo zapis vhodnega signala $u(k)$ z bitno obliko zapisa vhodnega signala $u(k)$, ki ga podaja enačba (5) potem lahko zapišemo s končno enačbo (18) izračuna produkta p_m z zaporednim seštevalnikom.

$$p_m(k) = h_m\left(-b_{u(k),0} + \sum_{i=1}^{B_u-1} b_{u(k),i}\, 2^{-i}\right) \qquad m = 0,1,2...N-1 \qquad (18)$$

V enačbi (18) je množenec h_i zapisan v vzporedni obliki z B_H biti, množitelj $u(k)$ je zapisan v zaporedni obliki z B_U biti. Dobljena izhodna vrednost produkta p_m je v zaporedni obliki predstavljena z $B_H + B_U$ biti. Blokovno shemo množilnika v zaporedni obliki prikazuje slika 5.

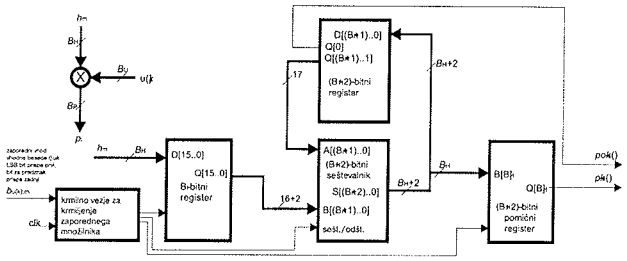


Slika 5: Blokovna shema zaporednega množilnika

Fig. 5: A plan of serial multiplier

Pri deljenju z 2 ostane na izhodu Q[0] 18-bitnega registra vrednost ostanka $po_m(k)$, ki je dolžine B_U bitov. Ta ostanek je v zaporedni obliki, zato ima prikazani zaporedni množilnik na sliki 5 ima za izhodno vrednost produkta pm dva izhoda: izhod zgornjega dela besede $pc_m(k)$ dolžine B_H bitov in izhod spodnjega dela besede $po_m(k)$ dolžine B_U bitov. Takšna oblika zaporednega množilnika potrebuje le en seštevalnik, multiplekser dolžine B_H bito, ki ga sestavlja B_H dvovhodnih IN vrat in zadrževalnik vmesnih vsot v_i. Slika 5 prikazuje m-ti množilnik za n-ti koeficient digitalnega FIR sita. Izhodni produkt $p_m(k)$ je vsota obeh produktov pom(k) in pcm(k) iz predstavljenega zaporednega množilnika v zaporedni obliki dolžine B_U+B_H bitov.

Pri izvedbi zaporednega množilnika je smiselno nadomestiti multiplekser dolžine B_H bitov s krmiljenimi pomnilniškimi celicami. Izvedba takšnega zaporednega množilnika je prikazana na sliki 6.



Slika 6: Blokovna shema zaporednega množilnika brez uporabe multipleksorjev

Fig. 6: A plan of serial multiplier without multiplexers

Na osnovi blokovne sheme prikazane na sliki 6 smo v programirnem FPGA vezju realizirali zaporedni množilnik. Zasedenost programirnega FPGA vezja za aparaturno izvedbo enega zaporednega množilnika prikazuje tabela 4.

Tabela 4: Zasedenost programirnega FPGA vezja pri implementaciji zaporednega množilnika s 16 bitnima vhodnima besedama.

Table 4: The programmable FPGA device utilization for 16-bit serial multiplier

Preliminary estimate of device utilization for part 4013EPG223:	
11% utilization of I/O pins.	(21 of 192)
8% utilization of CLB FG function generators.	(94 of 1152)
1% utilization of CLB H function generators.	(5 of 576)
8% utilization of CLB flip-flops.	(88 of 1152)

Realiziran množilnik ima dva 16 bitna vhoda, kjer je množitelj $u(k)$ dolžine B_U=16 bitov in množenec h_i je dolžine B_H=16 bitov, ter dvoje izhodov za zmnožka $pc_m(k)$ dolžine B_U=16 bitov in $po_m(k)$ dolžine B_H=16 bitov. Del logičnih konfiguracijskih blokov v FPGA vezju zahteva krmilno logiko zaporednega množilnika. Pri implementaciji več zaporednih množilnikov v isto FPGA vezje lahko uporabimo skupno krmilno vezje. Na osnovi razlike, ki nastane med implementacijo enega množilnika in šestnajstih množilnikov v istem programirnem FPGA vezju, lahko ugotovimo kolikšen del konfiguracijskih logičnih blokov je namenjen krmilni logiki in kolikšen del jih je namenjen množilnikom. V tabeli 5 je podana zasedenost enakega programirnega FPGA vezja v katerem je implementiranih 16 zaporednih množilnikov s 16-bitnima vhodnima besedama in 16 bitno izhodno besede.

Tabela 5: Zasedenost programirnega FPGA pri implementaciji šestnajstih zaporednih množilnikov s 16 bitnimi vhodnima besedama

Table 5: The programmable FPGA device utilization for sixteen 16-bit serial multiplier

Preliminary estimate of device utilization for part 4013PG223:	
19% utilization of I/O pins.	(36 of 192)
97% utilization of CLB FG function generators.	(1122 of 1152)
1% utilization of CLB H function generators.	(5 of 576)
73% utilization of CLB flip-flops.	(838 of 1152)

Iz tabel 4 in 5 je razvidno, da potrebujemo za implementacija samo enega zaporednega množilnika v programirnem vezju 67 konfiguracijskih logičnih blokov. Krmilna logika potrebuje 30 konfiguracijskih logičnih blokov. Implementaciji šestnajstih množilnikov v eno programirno FPGA vezje doseže enako aparaturno kompleksnost, kot implementacija vzporedne oblike digitalnega FIR sita v strukturi porazdeljene aritmetike. Za digitalna FIR sita z adaptivno strukturo, ki imajo več kot 16 koeficientov, je za aparaturno izvedbo primernejša izvedba digitalnega FIR sita v strukturi koncentrirane aritmetike z uporabo zaporednih množilnikov.

5. Zaključek

V prispevku smo podali pregled načinov izvedbe digitalnih FIR sit s polji programirnih logičnih FPGA vezij. Pri tem smo podrobno opisali strukturi, ki sta primerni za uporabo digitalnih FIR sit v aplikacijah adaptivnih digitalnih sistemov. Aparaturna izvedba digitalnih FIR v teh aplikacijah s programirnimi logičnimi FPGA vezji je zaradi vsebovanih množilnikov precej zahtevna, zato je potrebno uporabiti takšne strukture digitalnih FIR sit, ki omogočajo v času enega otipka vhodnega signala zamenjavo vseh njegovih koeficientov. Takšne strukture digitalnih FIR sit imenujemo tudi adaptivne strukture digitalnih FIR sit. Izvedbe adaptivnih struktur digitalnih FIR sit v programirnih logičnih vezjih so zanimive zaradi možnosti enostavnega povečanja stopnje sita z uporabo večjih programirnih logičnih vezij ali z dodajanjem novih programirnih logičnih vezij.

V prispevku je opisana vzporedna oblika digitalnega FIR sita v strukturi porazdeljene aritmetike. Podane so njene prednosti in slabosti glede na strukture digitalnih FIR sit v koncentrirani aritmetiki.

Pri digitalnih FIR sitih v strukturi koncentrirane aritmetike smo nakazali možnosti njihove izvedbe z uporabo množilnikov. Predstavili smo dve vrsti izvedb množilnikov s programirnimi logičnimi vezji, ki jih je možno uporabiti v digitalnem FIR situ z adaptivno strukturo: vzporedni množilnik, ker sta množenec in množitelj zapisana v vzporedni obliki in zaporedni množilnik, kjer je množitelj podan v zaporedni obliki, množenec pa v vzporedni obliki.

S pomočjo dobljenih rezultatov smo podali analizo aparaturne kompleksnosti izvedbe digitalnih sit z adaptivno strukturo. Analizo smo opravili za primer, kjer je bila vhodna beseda dolžine B_U=16, koeficienti sita so bili dolžine B_H=16 bitov in izhod digitalnega FIR sita je bil dolžine B_Y=16. Pri tem smo notranje aritmetične enote v FIR sitih zaradi zahtev izvedli s 16 do 24 bitno dolžino. Vsa sita smo načrtali s programskim paketom OrCAD 9.0 implementacijo v programirna FPGA vezja družine XC4000E smo izvedli s programskim paketom XACT 5.0 firme Xilinx.

Z rezultati smo pokazali, da izvedba digitalnih FIR sit v koncentrirani aritmetiki narašča linearno s številom koeficientov pri tem je aparaturna kompleksnost za sita z N=16 koeficienti po aparaturni kompleksnosti enaka vzporedni obliki digitalnega FIR sita v strukturi porazdeljene aritmetike.

6. Literatura

/1/ R. S. Grover, W. Shang, Q. Li, A Faster Distributed Arithmetic Architecture for FPGAs, Tenth ACM International Symposium on Field Programmable Gate Arrays Monterey, California, USA February 24-26, 2002

/2/ Martinez-Peiro, M.; Valls, J.; Sansaloni, T.; Pascual, A.P.; Boemo, E.I., A comparison between lattice, cascade and direct form FIR filter structures by using a FPGA bit-serial distributed arithmetic implementation, Electronics, Circuits and Systems, 1999. Proceedings of ICECS '99. The 6th IEEE International Conference on , Volume: 1 , 5-8 Sept. 1999 Page(s): 241 -244 vol.1

/3/ Ohsang Kwon, Kevin Nowka, Earl E., A 16-Bit by 16-Bit MAC Design Using Fast 5:3 Compressor Cells, The Journal of VLSI Signal Processing—Systems for Signal, Image, and Video Technology, June 2002

/4/ Les Mintzer, "FIR Filters with Field-Programmable Gate Arrays", Journal of VLSI Signal
Processing, vol. 6, pp. 119-127, 1993

/5/ OSEBIK, Davorin, JARC, Bojan, SOLAR, Mitja, BABIČ, Rudolf. A 30 tap FIR filter realization with FPGA circuits. V: 5th International Workshop on Systems, Signals and Image Processing, June 3-5, 1998, Zagreb. Proceedings IWSSIP'98. Zagreb: University of Zagreb, Faculty of Electrical Engineering and Computing, 1998, str. 86-89.

/6/ OSEBIK, Davorin, SOLAR, Mitja, BABIČ, Rudolf. Kaskadna izvedba nerekurzivnega digitalnega sita z 58 koeficienti s programirljivim poljem logičnih vezij, Zbornik šeste Elektrotehniške in računalniške konference ERK '97, 25. - 27. september 1997, Portorož, Slovenija. Ljubljana: IEEE Region 8, Slovenska sekcija IEEE, 1997, str. A/69-72

/7/ Kaluri, K.; Wen Fung Leong; Kah-Howe Tan; Johnson, L.; Soderstrand, M.; Signals, Systems and Computers, 2001. Conference Record of the Thirty-Fifth Asilomar Conference on , Volume: 2 , 4-7 Nov. 2001 Page(s): 1340 -1344 vol.2

/8/ Bill Allaire, Bud Fischer, Block Adaptivr Filter, XAPP 055, January 9, 1997 (Vrsion 1.1)

/9/ OSEBIK, Davorin, BABIČ, Rudolf. Izvedba FIR digitalnega sita v porazdeljeni aritmetiki z adaptivno strukturo, Zbornik osme Elektrotehniške in računalniške konference ERK '99, 23. - 25. september 1999, Portorož, Slovenija. Ljubljana: IEEE Region 8, Slovenska sekcija IEEE, 1999, zv. A, str. 59-62.

/10/ Steve Knapp, FPGAs furnish fast, furious FIR filters, Personal engineering, December 1998

/11/ Rolf Enzler, Tobias Jeger, Didier Cottet and Gerhard Troster, High-Level Area and Performance Estimation of Hardware Building Blocks on FPGAs, R. W. Hartenstein and H. Grunbacher (Eds.) FPL 2000, pp. 512-534 2000, Springer-Verlag Berlin Heidelberg 2000

/12/ Kaluri, K.; Wen Fung Leong; Kah-Howe Tan; Johnson, L.; Soderstrand, M.; Signals, Systems and Computers, 2001. Conference Record of the Thirty-Fifth Asilomar Conference on , Volume: 2 , 4-7 Nov. 2001 Page(s): 1340 -1344 vol.2

mag. Davorin Osebik, tel.: (02) 220-7238, e-mail: davorin.osebik@uni-mb.si
izr. prof. dr. Rudolf Babič, tel.: (02) 220-7230, e-mail: rudolf.babic@uni-mb.si
doc. dr. Mitja Solar, tel.: (02) 220-7232, e-mail: mitja.solar@uni-mb.si

Univerza v Mariboru
Fakulteta za elektrotehniko, računalništvo in informatiko
Smetanova 17, 2000 Maribor
tel.: (02) 220-7000, Fax: (02) 251-1178

Prispelo (Arrived): 25.06.2003 Sprejeto (Accepted): 26.08.2003

STEREO CAPTURE UNIT FOR REAL-TIME COMPUTER VISION, FEATURING HARDWARE-ACCELERATED DIGITAL FILTER DESIGN

Iztok Kramberger, Zdravko Kačič

University of Maribor, Faculty of Electrical Engineering and Computer Science, Maribor, Slovenia

Key words: computer vision, image processing, color filtering, spatial filtering, color space, skin color, user interface, interaction, hand tracking, head tracking, gesture recognition, skin formation, bit mask

Abstract: Real-time processing plays an important role in the achievement of more natural and physically intuitive ways for user-machine interaction. The main scope of this article is the development of a computer vision stereo capture unit that would fulfill more natural user-machine interaction within contemporary mobile telecommunication solutions. One of the main problems with current mobile devices is the restricted size of the showing area and the data input methods as a manner of device navigation. One of the most convenient ways to solve this problem is the use of gesture recognition within virtual or augmented reality applications.

In such cases the skin color can be a very comprehensive feature and so color tracking has been used to achieve the required goal. We have developed a method of color and spatial filtering for the purpose of indicating skin features and constituting skin formations within the level of the binary mask stream. A new stereo capture unit hardware structure with an accelerated digital nonlinear parametrical filter is shown. This features real-time searching for skin colored regions in a two-dimensional image stream where the parametrical design of the color filter also offers color detection in a comprehensive way. This parametrical scheme for the color filter enables automatic simultaneous adaptation of parameters due to changes in scene luminance and automatic adaptation possibilities of the color filter to the different types of skin.

Enota za zajemanje stereo slike za delovanje v realnem času s strojno izvedenim digitalnim filtrom

Ključne besede: računalniški vid, procesiranje slike, barvno filtriranje, prostorsko filtriranje, barvni prostor, barva kože, uporabniški vmesnik, interakcija, sledenje roke, sledenje glave, prepoznavanje kretenj, kožna formacija, bitna maska

Izvleček: Pomembno vlogo za čim bolj intuitiven način interakcije med uporabnikom in strojem ima procesiranje v realnem času. Članek prikazuje razvoj na računalniškem vidu zasnovanega vmesnika za zajemanje stereoskopskega video signala, ki omogoča poglobljeno interakcijo med uporabnikom in strojem znotraj aplikacij sodobnih telekomunikacijskih tehnologij. Glavni problem sodobnih mobilnih telekomunikacijskih tehnologij predstavlja omejena velikost prikazovalnega polja in načina vnosa podatkov oziroma navigacije same naprave. Eden izmed najprikladnejših načinov reševanja tega problema je uporaba razpoznavanja kretenj znotraj aplikacij navidezne in razširjene resničnosti.

Predstavljena je metoda barvnega sledenja, kjer smo razvili pristop barvnega in prostorskega filtriranja za označevanje kožnih značnic in tvorbe kožnih formacij na nivoju bitnih mask. Prikazana je struktura enote za zajemanje stereoskopskega video signala z uporabo strojnega pospeševanja v obliki digitalnega nelinearnega parametričnega filtra, ki omogoča iskanje kožno obarvanih regij v dvodimenzionalnem slikovnem zaporedju v realnem času. Parametrična zasnova predstavljenega pristopa omogoča samodejno sprotno prilagajanje na spremembe osvetlitve v sceni, kjer obstaja tudi možnost samodejne adaptacije na različne tipe kože.

1. Introduction

Most common intelligent environments suffer from the lack of natural and intuitive interactive devices. This can be especially felt within modern mobile telecommunication devices such as cellular phones, digital personal assistants, tablet computers and similar. The main problems for most of these mobile telecommunication devices that are becoming more and more intelligent and feature even more attractive services, is the lack of display size and interaction methods like data input or simply navigation and control. As a result conventional interaction devices such as the mouse and keyboard turn out to be unsuitable, which motivates the development of a vision-based tactile interface.

It can be seen on the basis of the human machine interaction that most communication is based on dynamic happening as the vision-based interface creates new content for each subject movement or gesture regardless of whether the movement or gesture was an expression of desired activity, selection, manipulation or just navigation. Dynamic happening within the captured scene is closely connected with any motion within the captured image sequence. When no subject movement is present at a given moment then the users have no requests to express themselves at the time.

It is reasonable to define user area for interaction within all three spatial dimensions because humans have a highly developed feeling of space. This enables the user more

natural and intuitive ways of communication. The user area for communication is defined as the area where the vision-based system is capable of tracking the movements and gesture of the navigational subject.

The basic idea of this work was to design an interface system that would enable tactile interaction within virtual or augmented reality user environments. As a control subject of tactile communication the user hand with stretched forefinger has been introduced, where the finger-tip represents the reference or navigation point of the interface.

From the beginning of the system design we had in mind that the interface should be able to gain real-time image processing, as this represents the basis for intuitive and more natural interaction with the machine. At this point the idea of interface usage was the ability to adapt the system into contemporary mobile telecommunication technology because the main problem lies within insufficient display size and user interaction methods. However, this is not the only problem concerning mobile technology because it suffers from available processing power and energy consumption. Real-time processing therefore was not the only concern to be considered although lower available processing power and data bandwidth was expected.

2. Stereo capture unit

In order to gain more natural and human-like interaction within the interface, a computer vision has been included that features the defining spatial position in all three dimensions. The use of stereoscopic vision requires two spatially and parametrically aligned image sensors shifted side by side, for which two analogue cameras have been used, as they offer an effective way of achieving the desired goal.

A stereo capture unit presents a special hardware design that can be used in fusion with digital signal processors from Texas Instruments. The testing hardware release was build for floating-point *DSP TMS320C6711* /19, 20/. This system design enables the capturing and preprocessing of two independent analogue video streams in real time. In addition stereoscopic vision requires the simultaneous capturing of two image sequences, thus the interface system should utilize two separate capturing units. As a matter of fact this approach would require more data bandwidth. Instead of utilizing two capturing units, a time-sequential video system is purposed where only one capturing unit is required. To achieve the stereoscopic cue, left and right images are spatially superimposed and temporally interleaved within one time-sequential image stream. It is reasonable to suppose that this approach lowers the quality of service, as the number of frames for a particular image sensor is halved but the interaction refresh rate still satisfies an adequate level of immersive communication. However, the drop in refresh rate is not the only concern of a time-sequential video system, as a time delay appears between the pair of captured stereo images. Consequently this time delay becomes a motion parallax problem when motion is present within the captured scene.

2.1. Skin color

Color is one of the most expressive visual features. Considerable work has been done on designing efficient descriptors for these features covering many applications. The color histogram is one of the most frequently used color descriptors that characterize the color distribution of an image. Common color description is based on dominant color, variable color and color structure within a color scheme.

Skin color has been chosen as the main feature in two-dimensional image, since this is the primary feature of the navigational subject. We have designed a hardware accelerated preprocessing stage on the level of programmable logic because of the pretension of color in spatial filtering for real-time. In several applications we encountered previous spatial filtering (low-pass filters) as it gives better results within the color segmentation stage. This requires an appropriate amount of free memory to store parts of the colored image. In our case the first stage of filtering is color feature extraction as a way of skin feature labeling, where in the next stage (spatial filtering) these features are spatially linked to skin formations, and noise singularities are removed. This approach for spatial filtering requires less memory requirements as the skin features exist within the binary mask, and gives adequate uniform color segmentation results. Color filter development has been accomplished through an appropriate mathematical model /34/ that issues the statistically acquired skin-color description in the *HSV* color space.

$$p(k)=\begin{cases}1, & h_{\min}\geq H(k)\leq h_{\max}, S(k)\geq s_{\max}, V(k)\geq v_{\max}\\ 0, & otherwise\end{cases} \quad (2.1)$$

The given distribution has been equipped with appropriate parameters that feature adaptation to different luminosity conditions within the captured scene, and adaptation to different skin types.

2.2. Motion detection

Motion detection within the motion analysis has been introduced using deterministic algorithms thus the processing time does not depend on the complexity of the captured scene, respectively the number of skin formations in the input image sequence. The motion detection core is based on a logical operator that performs its operation on three sequential binary masks. Continuous memory is needed for the storage of binary masks and so the motion detector algorithm is implemented within the digital signal processor.

$$LO=\left\{\left[m_i(T),m_i(T-1),m_i(T-2)\right]\mapsto mm_i(T)\right\} \quad (2.2)$$

The logical operator *LO* offers a fast and effective way of processing where its results are presented as a dynamic binary mask. The dynamic binary mask elements mm_i are those skin features m_i that are in motion. As in other similar methods for motion extraction using color filtering without any additional features like image edges, the given motion detector is liable to difficulties due to the mutual occlusion of individual skin formations that are projections of static or dynamic skin-colored subjects, and objects within the captured scene.

3. Hardware design for real-time processing

Appropriate hardware design has been suggested in order to achieve real-time processing of the given algorithms. The given hardware architecture features hardware acceleration using programmable logic that is implemented within the dynamically programmable *FPGA* circuits. The system's data path is divided into several processing stages.

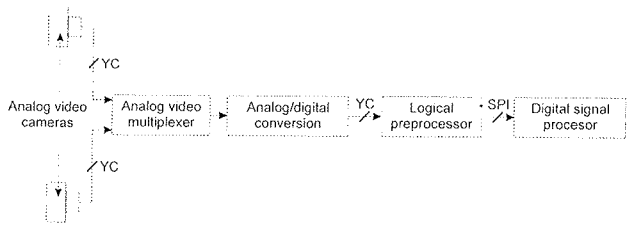


Figure 1: Hardware design for computer vision interface.

The first stage of the data path begins at the analog video multiplexer which combines two analog video streams from the employed analog cameras. Cameras are mounted in parallel configuration where the distance between them coincide with the average distance between human eyes. The parallel camera setup is advantageous over the toed-in approach since there is no vertical disparity introduced and the governing disparity equation is straightforward. However, the common field of view between the left and right acquired images becomes small and the correspondence problem becomes impossible for tokens at the outer extremes of the images. The analog video multiplexer creates a time-sequential video stream, where left and right video streams are spatially superimposed and temporary interleaved.

At the next processing stage the time-segmental video stream is captured and converted into a digital video stream which is described using a *YCbCr* color scheme. The analog video multiplexer and cameras are synchronized for capture period *T* depending on the used video standard. If the suggested *PAL* video standard is used then the capture period *T* is *12.5* frames per second, as the *PAL* video standard for the frame rate is defined at *25* frames per second. The processing requirements are not only defined by the frame rate but also by the employed video stream spatial resolution. The image resolution is defined at *625* lines for *PAL* video standard where each of them consists of *944* pixels. Of the given number of pixels, *768* are active within each line where *576* active lines are present. Thus the pixel frequency within the *PAL* video stream is *14.75 MHz*.

The high-quality single-chip digital video decoder *TVP5040* from Texas Instruments /17, 18/ was used as the video capture unit. This converts base-band analog *NTSC* and *PAL* video into digital component video. Both composite and S-video inputs are supported. The *TVP5040* includes two 10-bit A/D converters with 2x sampling. The output formats can be 8-bit, 10-bit, 16-bit, or 20-bit 4:2:2.

3.1. Logical preprocessor

The logical preprocessor is fully designed within the programmable logic devices and features hardware accelerated color and spatial filtering algorithms. The *FPSLIC* and *FPGA* circuits from Atmel were used as programmable logic devices. A logic device that interacts directly with the digital video decoder the *AT94K40 FPSLIC* device was used first /13/. This integrated circuit combines an *AT40K* dynamically configurable *SRAM FPGA* and high-performance 8-bit *AVR RISC* microcontroller with standard peripheral interfaces. The second logic device was a standard *SRAM* based *FPGA* integrated circuit AT40K40 /15/. Both devices can be efficiently used as coprocessors for high-speed (DSP/processor-based) designs by implementing a variety of computation intensive, arithmetic functions /9, 10/. These include adaptive finite impulse response (FIR) filters, fast Fourier transforms (FFT), convolvers, interpolators and discrete-cosine transforms (DCT) that are required for video compression and decompression, encryption, convolution and other multimedia applications. Both devices are capable of implementing Cache Logic (dynamic full/partial logic reconfiguration, without loss of data, on-the-fly) for building adaptive logic and systems. As new logic functions are required, they can be loaded into the logic cache without losing the data already there or disrupting

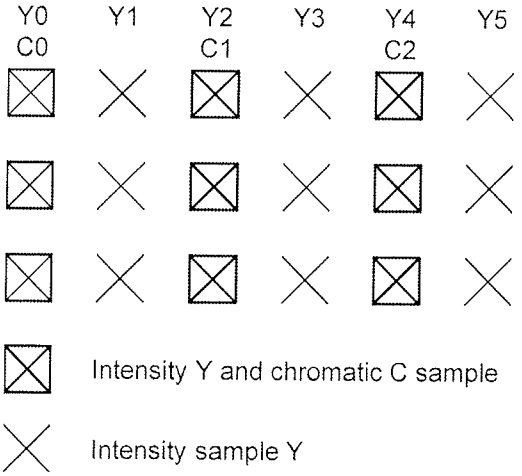


Figure 2: YCbCr 4:2:2 digital video format.

the operation remaining of the chip; replacing or complementing the active logic. The AT40KAL can act as a reconfigurable coprocessor /7/.

The captured digital video stream at the output of the conversion stage is given within the 4:2:2 *YCbCr* format. This means that the digital video stream includes twice the intensity information *Y* than chromatic information *Cb* and *Cr*. Prior to color filtering the employed sample format has to be converted into a 4:4:4 sampling scheme /26/. In such a way each of sampled image pixels has full color information. The given conversion is included within the preprocessing stage. The preprocessing stage can be gathered as a logical preprocessor (figure 4).

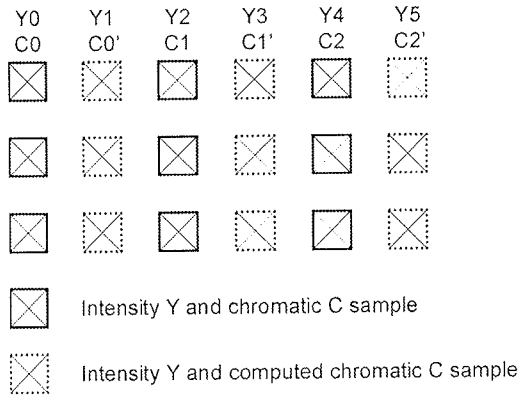


Figure 3: YCbCr *4:4:4 digital video format.*

The logical preprocessor input signal is the captured digital video stream within the *YCbCr* 4:2:2 sample format. The logical preprocessor output signal is represented as a binary mask stream. Thus it is obvious that less data bandwidth is required to transmit the binary mask stream as for *YCbCr* digital video format.

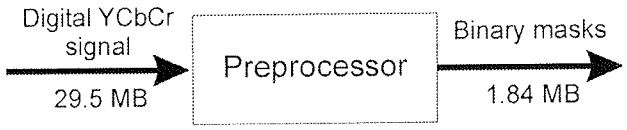


Figure 4: Color and spatial filtering stage.

The preprocessing stage includes the already mentioned digital video format conversion, and the color and spatial filter.

3.2. Skin color filter

An employed nonlinear parametrical digital color filter is based on the given mathematical model /35/ where the suggested logical units were used. In order to lower the logical and computing complexity of the digital filter the given structure of common logical unit *LU* was optimized for each of the given threshold constraints within the mathematical model in association with their computing complexity /35/.

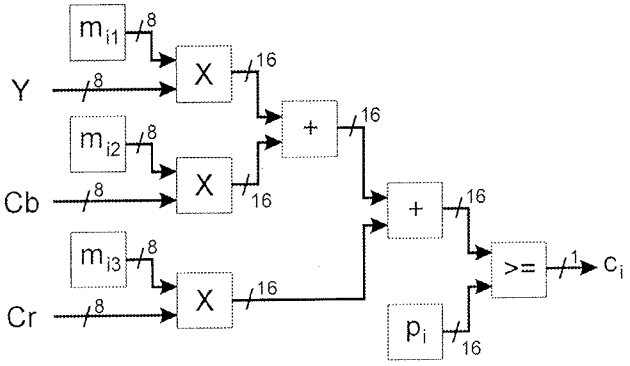


Figure 5: Structure of common logical unit LU.

Figure 5 presents a general design of a logical unit *LU* that features computing for all conditions c_i within the mathematical model of the filter /35/. Mathematically the logical unit expression can be written as a scalar product of the two vectors

$$c_i \Rightarrow [m_{i1} \quad m_{i2} \quad m_{i3}] \cdot [Y \quad Cb \quad Cr]^T \geq p_i. \quad (3.1)$$

The values m_{ij} present the parameters of the digital nonlinear filter and the *YCbCr* triple the input pixel sample where p_i is the given threshold value for the current condition c_i. On its input the logical unit accepts values of the individual image points within the *YCbCr* color space where each particular component of the *YCbCr* triple is given as an 8-bit value. Thus the color of each input point is given as 24-bit color depth. The *Y* component is given as value within the interval [0, 255] while chrominance components *Cb* and *Cr* are given as values within the interval [-128, 127].

Each individual input data component is multiplied by the appropriate parameter m_{ij} which is given as an 8-bit value. Used multipliers perform signed integer multiplication where, in the next step, the adders are forming partial sums that are compared to an adequate parameter p_i. The comparison is carried out by a 16-bit comparator that passes the binary output information to the output logical function of the digital filter. The parameters m_{ij} and p_i are implemented as registers within the digital filter and are able to memorize current setup values. These values can dynamically change during runtime.

The architecture of the given digital color filter is based on parallel configuration where appropriate condition values are computed in a parallel fashion. There are six conditions needed for the mathematical model. Therefore for each individual image pixel within the input digital video stream given conditions are computed where the output function of the color filter defines the logical state that represents the skin feature within the binary mask stream.

3.3. Automatic adjustment to scene luminance

The described digital nonlinear color filter has the ability to modify its parameters dynamically. The given parameters

are definable using the shown mathematical digital filter model. Thus it is possible to automatically adapt the parameters regarding scene luminance within the input image stream *i(t)*. It is reasonable that the parameters of the filter have to be corrected for larger changes in scene luminance, to achieve regular skin features extraction. It was found that within the changes in scene luminance it is necessary to modify the threshold values for saturation s_{max} and value v_{max} as the hue in such cases shows more or less constant value except when the brightness is very high.

Within the input sequence image, represented by the input image stream *i(t)* the color stimulus, in general, is non-uniform over the entire image area. Thus it is reasonable to obtain a general estimation for the brightness and saturation of the entire image area. As the input stream is represented within the *YCbCr* color scheme it is possible to use the *Y* component for an adequate estimation of brightness, whilst chrominance components *Cb* and *Cr* can be used for an adequate estimation of saturation. The average values for image brightness $\bar{v}$ and saturation $\bar{s}$ can, therefore, be computed as

$$\bar{v} = \frac{1}{N}\sum_{i}^{N} Y(i) \text{ and } \bar{s} = \frac{1}{N}\sum_{i}^{N} S(Cb[i], Cr[i]) \quad ,(3.2)$$

where *N* stands for the number of image points, *i* denotes the current coordinate of a point and *S(Cb,Cr)* is the saturation function indirect to chrominance.

The estimation for saturation $\bar{s}$ remains constant for global changes of scene luminance as a reflection of changes in the brightness of the light sources whilst the input image stream shows changes in estimation for value $\bar{v}$. The average estimation for saturation $\bar{s}$ mostly depends on the spectrum of the light stimulus within the scene. This shows that the average saturation estimation can be used to correct the saturation threshold of the digital filter and furthermore the average estimation value can be used to correct the value threshold.

The digital filter's threshold values can be modified dynamically in real time regarding the capture interval *T*. Thus both estimations become time-dependent as their values within the previous capture interval *T-1* can be used to correct the current threshold values at interval T. Consequently the threshold values v_{max} and s_{max} become time-dependent as their values are altered at the beginning of the capture interval *T*. The given implication can be written as

$$\{\bar{v}(T-1), \bar{s}(T-1)\} \Rightarrow \{v_{\max}(T), s_{\max}(T)\}. \quad (3.3)$$

The functional dependence of a given implication can be gathered using large sample set of test images under different lightning conditions where the appropriate values can be estimated to assure constant filtering quality.

3.4. Spatial filter

The introduced spatial filter features a median translation function within the binary mask stream. Spatial filtering is employed in order to remove singularities within the binary mask stream, caused by image noise, and to link small homogenous regions of skin features into larger skin formations. Skin formation *F* is the union of those skin features or elements of binary mask that are combined by spatial connectivity. Two elements of binary mask are spatially connected if a path between them exists where the characteristic function stays constant. If spatial connectivity is expressed as:

$$c(x,y,T) = \sum_{i=-1}^{i=1}\sum_{j=-1}^{j=1} m([x+i],[y+j],T), \quad (3.4)$$

then the element *m(x,y,T)* is spatially connected with one of its eight neighbors if *c(x,y,T)>1*. As noise within the binary mask not only exists as a singular skin feature but also as small skin formations it is reasonable to remove small skin formations or spatially link them to larger skin formation with the use of a spatial median filter *MF*. The translation function of the median filter within the level of binary masks can be expressed as:

$$MF = \begin{cases} 1, & \sum_{i=-\frac{(k-1)}{2}}^{i=+\frac{(k-1)}{2}} \sum_{j=-\frac{(l-1)}{2}}^{j=+\frac{(l-1)}{2}} m([x+i],[y+j],T) \geq med \\ 0, & \sum_{i=-\frac{(k-1)}{2}}^{i=+\frac{(k-1)}{2}} \sum_{j=-\frac{(l-1)}{2}}^{j=+\frac{(l-1)}{2}} m([x+i],[y+j],T) < med \end{cases}, \quad (3.5)$$

where *k* and *l* are positive odd numbers that define the size of the treated region within the binary mask. The adequate median value *med* can be expressed as

$$med = \frac{k \cdot l}{2} + 1. \quad (3.6)$$

As an example a digital spatial filter with size of translation matrix 3 x 3 is shown in Figure 6.

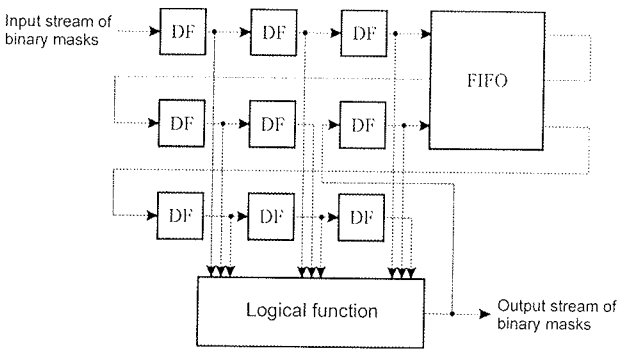


Figure 6: Architecture of spatial filter with size of translation matrix 3 x 3.

The logical function of a given spatial filter can be expressed in terms of partial sums by column or rows. The FIFO memory is employed in order to memorize those particular binary mask lines needed to gain proper spatial filter operation.

4. Initialization, synchronization and data path

Because *FPSLIC* and *FPGA* logical devices are *SRAM* based they need to be initialized with the appropriate configuration bitstream after power-up or master reset signal has been received. Therefore a serial *EEPROM0* configuration memory *AT17LV010* with capacity of *128 kB* was used /16/. Both devices are configured from the same configuration memory in master/slave fashion. The bitstream of the *FPSLIC* device also includes the firmware or the program memory of the embedded *AVR* microcontroller.

The firmware of the embedded AVR microcontroller is started after successful configuration of the logical devices. The first step of the *AVR* program is to configure the digital video decoder. This is done by uploading the digital video decoder firmware which is written in the second serial *EEPROM1* configuration memory *AT17C256* with a capacity of *32 Kb* /16/. The host peripheral interface *HPI* /17/ connected to the embedded *AVR* core is used for loading the program data and appropriate digital video decoder configuration. As the digital video decoder is being configured, the *AVR* program configures the time-sequential analog video multiplexer. The connection between the embedded *AVR* core and the video multiplexer is done by using one of two asynchronous serial ports. The second asynchronous serial port is used for direct communication between the embedded *AVR* core and a standard personal computer. This makes it possible to dynamically change capture parameters, digital color filter parameters or gather status values in runtime during operation. This hardware feature is useful during application development and debugging, or similar.

The master synchronization signal is generated by a controllable analog video multiplexer where both analog cameras are synchronized with each other. The analog multiplexer can receive control commands such as full/half frame auto-switching where an analog time-sequential and spatially superimposed video signal is generated. The video signal can also be configured as a single left/right camera signal. All commands to the video multiplexer synchronization controller are executed within the blanking of the video signal so that no synchronization errors can occur. Each time the video signal is switched between left and right camera signals, a special synchronization signal is send to the embedded *AVR* core over the serial link. This signal informs the *AVR* core to add framing to the data path.

As already stated the data path begins at the analog video multiplexer where the time-sequential video signal is generated. This signal is captured by the digital video decoder and passed to the digital preprocessor. The logical circuit within the *FPSLIC* device receives the digital video signal and extracts skin features. Skin features in the form of a binary mask stream are transferred to the second logical circuit within the *FPGA* device where spatial filtering is performed. The skin formations are created in such a way and then transferred to the *DSP* system.

There are three different communication channels implemented between the *DSP* and digital preprocessor. Two of them use *McBSP* (Multi-channel Buffered Serial Port) interfaces of the *DSP* /19/. These two ports need syn-

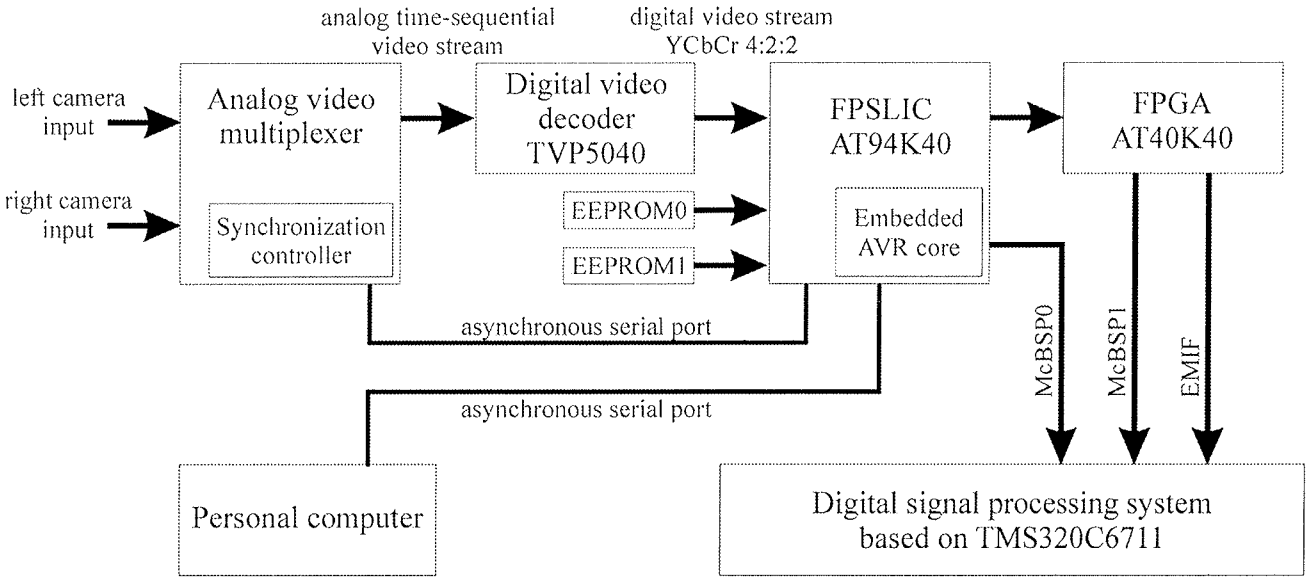


Figure 7: Block scheme of the capture unit.

chronous serial transmitter/receiver logic which is implemented within the *FPGA* and *FPSLIC* logical circuits. The *McBSP0* port is used for the transfer of the binary mask stream. This serial port is configured to transfer 32-bit long words where *24* framing signals are required to transfer *768* bits of the complete binary mask horizontal line. The *DMA* (Direct Memory Access) of this serial communication channel is configured to receive and transfer the binary mask stream to the main memory of the *DSP* system where complete framing of the binary mask is added. Motion extraction has been added within each of the DMA cycles where implication of the logical operator *LO* has been used. Each of the 32-bit long words received for the current binary mask *M(T)* is used to compute the current dynamic binary mask *MM(T)*. This means that, the dynamic binary masks are computed in-time with binary mask stream capturing.

The second *McBSP1* port of the *DSP* is used to send/receive configuration data as, for example, the parameters of skin feature extraction within the digital color filter. The second serial communication channel is bound to the embedded *AVR* core where the embedded firmware receives the configuration data and applies those configuration changes into logical units *LU* registers used to store the parameter values of the skin feature extraction process.

Automatic adjustment to scene luminance is solved over the second serial communication channel. The embedded *AVR* core accumulates the current average values for value $\bar{v}$ and saturation $\bar{s}$ within the current captured frame and passes them to the *DSP* system. The *DSP* program calculates new color filter parameters and sends them back to the *AVR* core where these new calculated values are applied to the color filter logic.

The third communication channel between the logical preprocessor and the *DSP* can be configured using the *EMIF* /19/ (Extended Memory Interface) of the *DSP*. The *EMIF* interface features 32-bit parallel access to the logic circuit within the *FPGA* device for read/write operation. This feature of the hardware design can be used within applications where larger data bandwidth is required. Within the *DSP* system the *EMIF* interface can also be configured for *DMA* cycles and can be used for non-filtered image transfers.

5. Efficiency estimation for given hardware structure

In order to gain real-time processing within the given system, the available time to perform all required algorithms is *80 ms* when *PAL* video standard is employed. A time delay of *40 ms* is present between the left and right captured images due to the employment of a time-sequential video system. To achieve efficiency estimation for the given hardware structure, a test bed was created that runs on a standard personal computer within the MS Windows operating system and has been developed with MS Visual Studio 6.

Some additional processes have been added to estimate time performance requirements for a particular algorithm within the given functionality. The current test bed features definition for the spatial positions of those skin formations stereo pairs that represent the nearest skin-colored subject within the captured scene. The algorithm that defines the spatial position uses the energy projection approach within the left and right dynamic binary mask. As energy projections are computed the set of potential stereo skin formation pairs is generated where the best match is chosen by the stereo correspondence algorithm that uses the stereoscopic and epipolar constraints of the stereo parallel camera setup /3/.

In such a way the combined processing path of the complete test bed functionality was divided into particular processes:

- Process of color and spatial filtering that includes skin feature labeling and the creation of homogenous skin formations.
- Process of dynamic binary mask creation with employment of a logical operator.
- Process of energy projections computation.
- Process of defining the spatial positions of skin formations that includes motion analysis and solving stereo correspondence problem.
- Process of frame backup that is required to gain dynamic binary mask creation.

Table 1 shows the gathered performance requirements measured using a personal computer where the times are given as a proportion of the disposable processing power.

Table 1: Relative times required to perform specific process.

Process	Required time
Color and spatial filtering	0.6386
Dynamic binary masks creation	0.3456
Energy projection computation	0.0005
Definition of spatial position	0.0004
Backup current frames	0.0007

The time required to define spatial position depends of the number of skin formations found or present within the captured scene. It can be seen (table 1) that most of the available processing time is required for color and spatial filter. About *35 %* of available time is required for the creation of dynamic binary masks, where other *1.6 %* is required for other operations that include energy projection computation, spatial position definition and backup of current frames.

Color and spatial filtering is performed using the suggested hardware architecture within the preprocessing stage, thus available processing time is increased by about *60 %*.

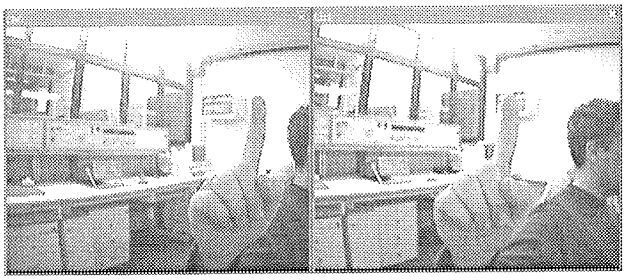


Figure 8: Non-filtered stereo image of input sequence.

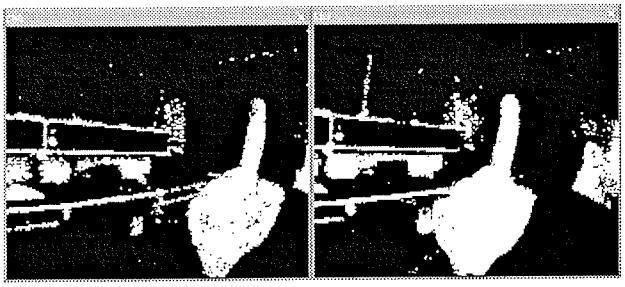


Figure 9: Appurtenant binary masks without employment of spatial filtering.

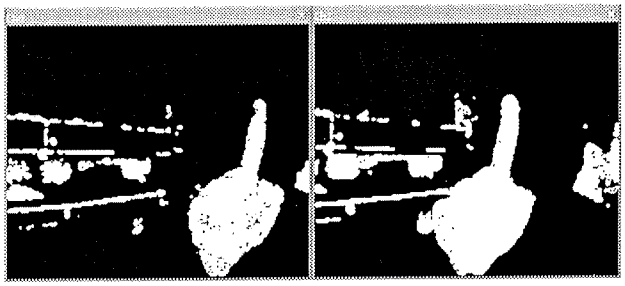


Figure 10: Appurtenant binary masks with employment of spatial filtering.

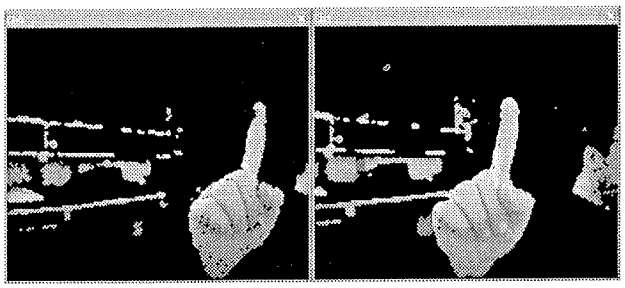


Figure 11: Placing the binary masks onto the original input stereo image.

The original stereo image of the input sequence is given in figure 8. Figure 9 shows the appurtenant binary masks where no spatial filter has been employed. In contrast, figure 10 shows the same binary masks with the employment of spatial filtering with size of translation matrix 5 x 5. Figure 11 presents the implication of the binary masks onto the original input stereo image where they are spatially superimposed.

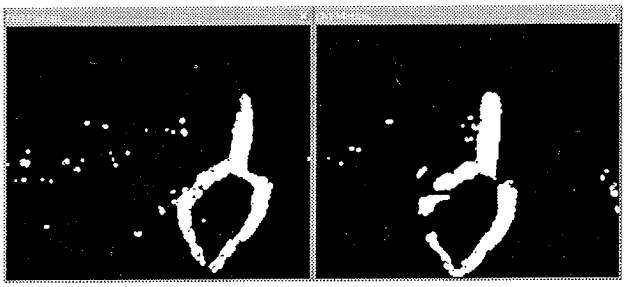


Figure 12: Dynamic binary masks $MM_L(T)$, $MM_R(T)$ *where the input sets of their elements are the binary masks* $M_L(T-1)$, $M_R(T-1)$.

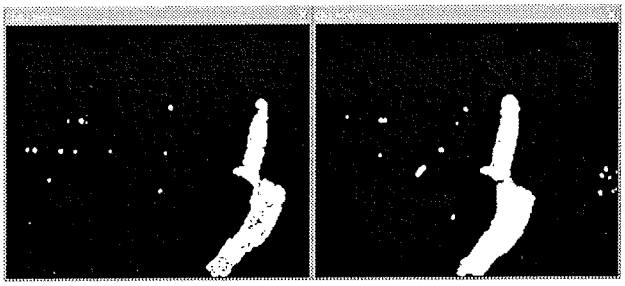


Figure 13: Dynamic binary masks $MM_L(T)$, $MM_R(T)$ *where the input sets of their elements are the binary masks* $M_L(T)$, $M_R(T)$.

Figures 12, 13 show two different approaches for motion extraction using the same logical operator. It can be seen that the skin formations are not fully extracted due to the slow-motion of the controlled subject regarding capture period *T*. Thus the quality of motion detection mostly depends on the velocity of the subjects in motion within the scene and the capture period *T*. The capture period was 6 stereo frames per second for the given results as they were captured within the test bed software.

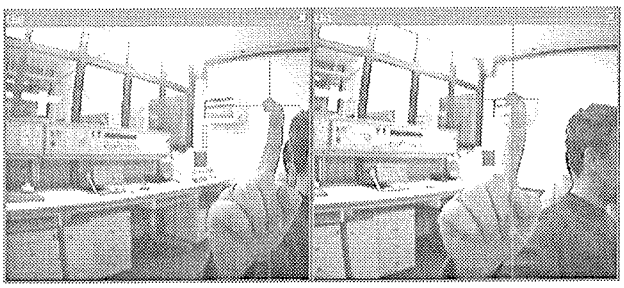


Figure 14: Definition of the navigation point's spatial position.

The definition of the navigation point's spatial position is shown in figure 14 where the vertical lines denote the position of extreme value within the appropriate energy projections whilst the horizontal lines denote the boundary frames B_L, B_R for given skin formations (figure 15).

Figure 15 shows vertical errors λ_L, λ_R of the adequate boundary frames B_L, B_R for given skin formations caused by motion parallax due to time delay in time-sequential ster-

eo capturing. In such cases the navigation point coordinates in the vertical direction can be computed as the average value between appropriate navigation point projections to the left and right image planes.

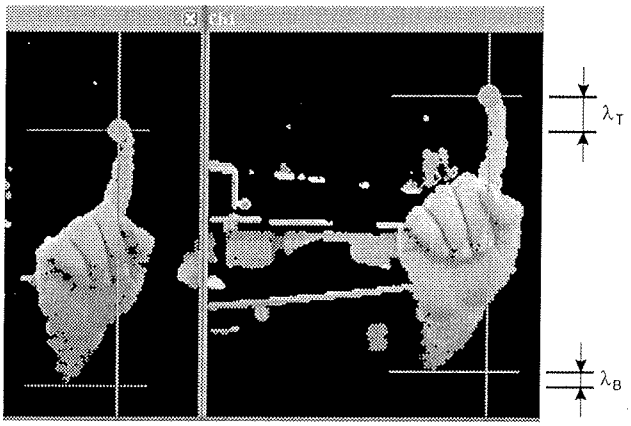


Figure 15: Vertical errors within the boundary frames B_L, B_R caused by motion parallax.

6. Conclusion

It has been shown that with the suggested time-optimized labeling or skin feature extraction within the captured stereo input image sequence performed by nonlinear parametrical digital color filter, it is possible to create binary masks in real-time. It is possible to create skin formations and remove noise singularities by spatial dependence examination between skin features within a small region of the binary mask. It has been shown that with dynamic binary mask computation with a suggested logical operator, it is possible to define those regions within the image sequence that are exposed to changes due to the motion of skin-colored subjects and objects within the captured scene.

The presented color and spatial filtering can also be foreseen within other vision-based applications that require the extraction of human parts from real-time image sequence such as a human face for example. In further work we suggest detailed analysis of automatic color filter adaptation to different types of skin. A more exact definition of skin-color area that would be more specific to user's skin types would gain more accurate extraction of skin features and thus, more precise definition of skin formations within the image sequence. Regions could be predicted where occlusion of dynamic and static skin formations could occur using appropriate scene complexity estimation. Therefore the employment of a two-dimensional model of the control subject could solve the problems of spatial positional definition within those regions.

Literature

/1/ G. Sharma. Ed. Digital Color Imaging Handbook. Xerox Corporation, CRC Press LCC, New York, 2003.

/2/ S. W. Perry, H. S. Wong, L. Guan. Adaptive Image Processing. A computational Intelligence Perspective. CRC Press LCC, New York, 2002.

/3/ R. Hartley, A. Zisserman. Multiple View Geometry in computer vision. Cambridge University Press, UK, 2000.

/4/ L. Guan, Ed. Multimedia Image and Video Processing. Boca Raton: CRC Press LCC, New York, 2001.

/5/ J. C. Russ. The Image Processing Handbook - Fourth Edition. Materials Science and Engineering Department, North Carolina State University, CRC Press LCC, North Carolina 2002.

/6/ Palmer, Steven. Vision Science. The MIT Press, Cambridge, MA, 1999.

/7/ I. Kramberger, M. Solar. DSP Acceleration Using a Reconfigurable FPGA. Proc. of IEEE International Symposium on Industrial Electronics - ISIE '99, Bled - Slovenija, 1999.

/8/ New, Bernie. A distributed arithmetic approach to designing scaleable DSP chips. EDN, pp. 107-114, 1995.

/9/ Atmel. Recommended Design Methods. Atmel Corp., September, 1997.

/10/ Atmel. Implementing Cache Logic with FPGAs. Atmel Corp., September, 1997.

/11/ Atmel. Implementing Bit-Serial Digital Filters. Atmel Corp., September, 1997.

/12/ Atmel. Implementing FreeRAM inside the FPGA or AT94K Series FPSLIC Using VHDL with IP Core Generator. Atmel Corp., 2001

/13/ Atmel. AT94K Series Field Programmable System Level Integrated Circuit. Atmel Corp., 2001.

/14/ Atmel. 8-bit AVR Microcontroller with 8K Bytes In-System Programmable Flash - AT90S8515. Atmel Corp., 2001.

/15/ Atmel. 5K - 50K Gates Coprocessor FPGA with FreeRAM™. Atmel Corp., 2002.

/16/ Atmel. FPGA Configuration EEPROM Memory. Atmel Corp., 2003.

/17/ Texas Instruments. TVP5040 NTSC/PAL Digital Video Decoder With Macrovision (TM) Detection. Data Manual, DAV Digital Video/Imaging, May 2001.

/18/ Texas Instruments. TVP5031/40/5145EVM User's Guide. DAV Digital Video/Imaging, September 2001.

/19/ Texas Instruments. TMS320C6711 Floating-Point Digital Signal Processors. Texas Instruments Incorporated, 1999, Revised 2003.

/20/ D. Bell. TMS320 Cross-Platform Daughtercard Specification - Revision 1.0. Texas Instruments Incorporated, Application report, 2000.

/21/ MathWorks, Inc.. Image Processing Toolbox User's Guide. Mathworks, 2001.

/22/ MathWorks, Inc.. Signal Processing Toolbox User's Guide. Mathworks, 2001.

/23/ MathWorks, Inc.. Statistics Toolbox User's Guide. Mathworks, 2001.

/24/ Y. Wu, Y. Shan, Z. Zhang, S. Shafer. Visual Panel: From an ordinary paper to a wireless and mobile input device. Microsoft Research, October 2000.

/25/ J. R. Ohm, K. Grunberg, E. M. Izquiredo, M. Karl. A realtime hardware system for stereoscopic videoconferencing with viewpoint adaptation. Heinrich Hertz Institute, Image Processing Department, Germany, 2000.

/26/ J. Keith. YCbCr to RGB Considerations, Converting 4:2:2 to 4:4:4 YCbCr. Intersil Application Note March 1997 AN9717.

/27/ M. Suen, R. Kleihorst, A. Abbo, E. C. Solal. Real time skin-tone detection with a single digital camera. Philips Research Laboratories, Eindhoven, Netherland. 2001.

/28/ S. Ahmad. A usable real-time 3D hand tracker. In proc. IEEE Asilomar Conf., 1994

/29/ J. Crowley, F. Berard, J. Coutaz. Finger tracking as as input device for augmented reality. In Proc. Int'l Workshop on Atomatic Face and Gesture Recognition, pages 195-200, Zurich, 1995.

/30/ R. Kjeldsen, J. Kender. Finding skin in color images. In Proc of Second International Conference on Automatic Face and Gesture Recognition, pages 312-317, 1996.

/31/ M. Jones, J. Rehg. Statistical color models with application to skin detection. Technical Report CRL 98/11, Compaq Cambridge Research Lab., 1998.

/32/ Raja, S. McKenna, S. Gong. Colour model selection and adaptation in dynamic scenes. In Proc. European Conf. Computer Vison, pages 460-475, 1998.

/33/ A. Wu, M. Shah, N. V. Lobo. A Virtual 3D Blackboard: 3D Finger Tracking using single Camera. Department of Computer Science, University of Illinois, Urbana, School of Computer Science, University of Central Florida, 1999.

/34/ N. Herodotu, K. N. Plataniotis, A. N. Venetsanopoulos. Automatic location and tracking of the facial region in color video sequences. Signal Processing: Image Communication, 14, 359, 1999.

/35/ I. Kramberger, Z. Kačič. Implementation of parametrical nonlinear digital filter for skin features identification in digital image using FPSLIC technology. Informacije MIDEM, št.: 3, 2003.

Dr. Iztok Kramberger, Faculty of Electrical Engineering and Computer Science, University of Maribor, Slovenia.

prof. Dr. Zdravko Kačič, Faculty of Electrical Engineering and Computer Science, University of Maribor, Slovenia.

University of Maribor, Faculty of Electrical Engineering and Computer Science Smetanova ulica 17, 2000 Maribor, Slovenia

Prispelo (Arrived): 21.06.2003 Sprejeto (Accepted): 26.08.2003

UDK621.3:(53+54+621+66), ISSN0352-9045 Informacije MIDEM 33(2003)3, Ljubljana

EKSPERIMENTALNO TESTNO OKOLJE ZA DRUŽINO STANDARDOV IEEE 1149.X

Robert Sedevčič, Uroš Kač, Franc Novak

Institut Jožef Stefan, Ljubljana, Slovenija

Ključne besede: robna testna linija, IEEE 1149.1 testno vodilo, krmilnik testnega vodila, Linux gonilnik, SVF prevajalnik

Izvleček: V članku je predstavljen projekt EBS (ang. Experimental Boundary-Scan), katerega osnovni namen je vzpostaviti preprost laboratorijski testni sistem za IEEE 1149.x združljiva vezja. EBS temelji na široko dostopnih elektronskih komponentah, ki podpirajo tehniko robne testne linije, ter prosto dostopnem operacijskem sistemu Linux. Projekt sestoji iz strojnega ter programskega dela. Razvit je bil računalniški vmesnik med sistemskim vodilom ISA ter IEEE 1149.1 testnim vodilom, ki temelji na standardnem krmilniku SN74ACT8990. Vmesnik je podprt z ustreznim gonilnikom za Linux. Realiziran je bil tudi splošen prevajalnik za jezik SVF (ang. Serial Vector Format), ki sodi med standardne formate opisa testnih postopkov na osnovi robne testne linije. Prevajalnik je preko vtičnih programskih modulov povezan z gonilnikom oz. strojnim vmesnikom ter skupaj s pripadajočima grafičnima vmesnikoma predstavlja funkcionalno zaključeno testno okolje. Izdelan testni sistem je bil uspešno uporabljen za izvedbo različnih laboratorijskih eksperimentov z IEEE 1149.4 združljivimi vezji.

Experimental test environment for IEEE 1149.x standards

Key words: boundary-scan, IEEE 1149.1 test bus, test bus controller, Linux device driver, SVF compiler

Abstract: The development of complex, multi-layer PCBs, associated with the miniaturisation of electronic device packages and new assembly methods (BGA, SMT, COB,...), made physical access, required by the traditional in-circuit test methods, increasingly difficult. The IEEE 1149.1 standard test access port and boundary-scan architecture presented a solution to the limited access circuit testing problem and since its adoption in 1990 became an important design-for-testability (DFT) technique in the electronic industry.

In order to efficiently use the boundary-scan (BS) infrastructure for testing and other purposes, adequate tester hardware as well as test development software tools are required. These can greatly facilitate the generation and application of BS based test or in-system programming (ISP) procedures in most modern complex devices. Although there is a number of adequate BS test solutions available on the market, these primarily target industrial applications where robust operation in medium or high-volume production testing is required. On the other hand, purchase of expensive professional BS test equipment can represent a major obstacle for academic institutions involved in research or educational activities regarding IEEE 1149.1 and related standards. Furthermore, most commercially available systems are relatively complex and do not provide the transparency, which is required for a thorough understanding of the boundary-scan test technique. Only a fully open and custom configurable platform could provide the necessary freedom to users with specific requirements as well as an efficent educational tool for teaching the basic principles of BS testing. The absence of similar solutions motivated the implementation of our Experimental Boundary-Scan (EBS) platform, which is intended as a suite of BS test tools, based on the GNU/Linux operating system.

The EBS environment was conceived as a simple laboratory test system for IEEE 1149.x compliant circuits. The project is based on widely available hardware supporting boundary-scan test techniques and on the open source Linux operating system. The project is divided into hardware and software related parts. A simple ISA-bus PC adapter featuring the SN74ACT8990 test bus controller was implemented along with the appropriate Linux device driver. The SN74ACT8990 performs transformation of the test data supplied by the host processor and generates adequate data and control IEEE 1149.1 test bus signals (TDO, TDI, TMS, TCK). The SN74ACT8990 data, control and status registers are mapped into the host processor I/O space through the 16-bit ISA bus interface. On-board glue logic provides a configurable I/O base address and IRQ level as well as a hardwired test clock frequency divider. Although no additional processing or storage capability is available on-board, the implemented test bus adapter can represent a sufficient and cost effective solution comparable to many commercially available adapters.

A generic, plug-in based compiler for the widely supported Serial Vector Format (SVF) test description language and basic graphical user interfaces were also developed. SVF is an ASCII format used for describing test patterns that represent stimulus, expected response and mask data according to the IEEE 1149.1 standard. Along with the implemented hardware platform these tools form a functional boundary-scan test environment. Since the software tools are conceived as independent modules featuring well-defined application interfaces, support for alternative BS hardware as well as additional software utilities can be easily included into the system. So far the existing test system was succesfully applied in a number of laboratory experiments concerning prototype IEEE 1149.4 mixed-signal test bus compliant devices. The complete EBS project is freely distributable and has been made available through the Sourceforge software development network service.

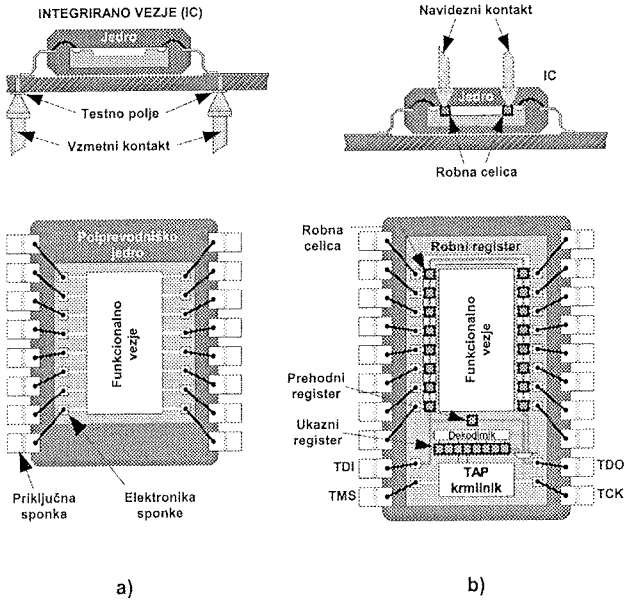
1 Uvod

Z uvajanjem novih tehnologij izdelave se povečuje gostota integriranih vezij, veča se število priključnih sponk posameznih komponent ter manjšajo razdalje med njimi. Posledično se zmanjšujejo tudi razdalje med bakrenimi povezavami, kar skupaj z večanjem števila plasti v tiskanih vezjih in novimi načini pritrjevanja komponent (BGA, SMT, COB, ...) izredno otežuje fizični dostop do posameznih komponent tiskanega vezja /1/ in s tem izvedbo klasičnih testnih postopkov z uporabo vmesnikov z vzmetnimi kontakti (slika 1a). Kot rešitev problema omejenega dostopa so proizvajalci komponent v osemdesetih letih razvili postopek t.i. robne testne linije (ang. boundary-scan). Ta nadomešča fizični dostop do sponk posameznih komponent vezja z dostopom preko celic pomikalnega registra,

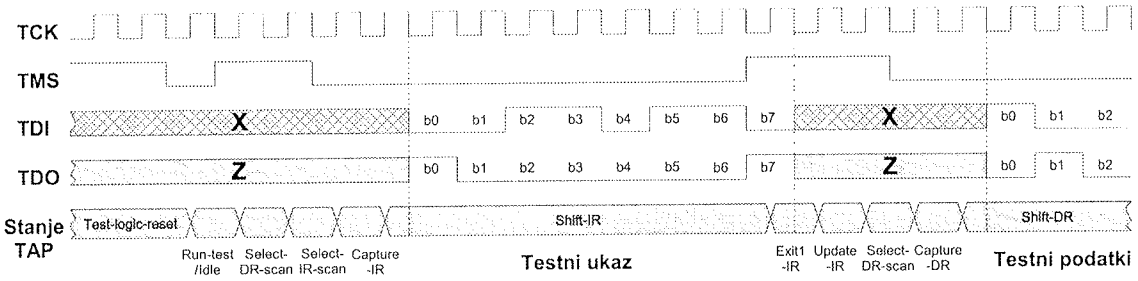
nameščenega okoli jedra integriranega vezja (slika 1b). Leta 1990 je bil postopek robne testne linije na pobudo združenja Joint Test Action Group (JTAG) sprejet kot mednarodni standard IEEE 1149.1 /2/.

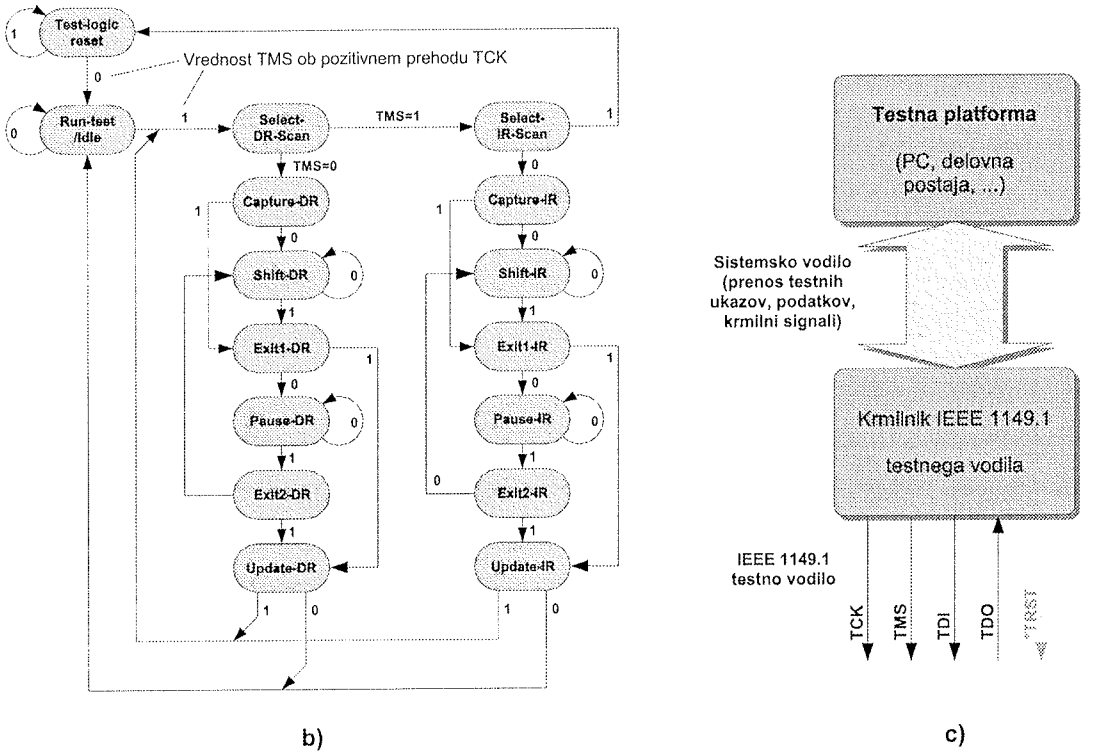
Uporaba robne testne linije zahteva ustrezno strojno ter programsko opremo /3/. IEEE 1149.1 združljiva integrirana vezja uporabljajo specifičen serijski protokol (slika 2a), ki omogoča prenos testnih ukazov in podatkov preko skupnega vodila in je določen z avtomatom končnih stanj v krmilniku testnega vmesnika (TAP) komponente (slika 2b). Arhitektura tipičnih testnih platform zahteva pretvorbo testnih podatkov in ukazov v ustrezno bitno sekvenco na IEEE 1149.1 testnem vodilu, za kar lahko uporabimo poseben krmilnik testnega vodila (slika 2c).

Namenska orodja nam lahko močno poenostavijo izvedbo ter skrajšajo čas, ki je potreben za realizacijo testnega postopka /4-7/. Kljub temu, da lahko na trgu srečamo več strojnih ter programskih rešitev različnih proizvajalcev, so tovrstna orodja v prvi vrsti namenjena industrijskim aplikac-



Slika 1: Osnovni koncept robne testne linije





Slika 2: Prikaz protokola IEEE 1149.1 testnega vodila in tipične povezave s testno napravo

ijam, ki zahtevajo predvsem preprosto uporabo in robustno delovanje. Temu je primeren tudi cenovni razred teh sistemov, ki lahko za akademske in raziskovalne ustanove predstavlja precejšnjo oviro pri vzpostavitvi laboratorijskega okolja, potrebnega za preučevanje, raziskave in razvoj novih komponent in ustreznih testnih postopkov na osnovi 1149.1 infrastrukture /8-11/.

Večina komercialnih testnih sistemov je relativno kompleksnih, hkrati pa uporabniku ne nudijo potrebne transparentnosti, ki je nujna za razumevanje tehnike robne testne linije. Zahtevnejši uporabnik za realizacijo specifičnih rešitev pogosto potrebuje popolnoma odprto, prosto dostopno in nadgradljivo platformo. Pomanjkanje tovrstnih rešitev je bil eden od glavnih razlogov za nastanek projekta EBS, ki združuje vsa osnovna orodja, potrebna za izvedbo testnega postopka v skladu s standardom IEEE 1149.1.

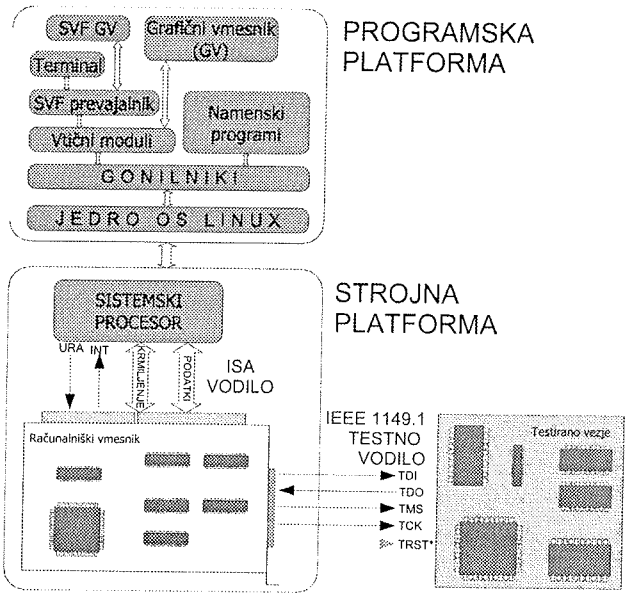
Projekt EBS nudi popolnoma odprto ter prosto razširljivo platformo, zasnovano na operacijskem sistemu (OS) GNU/Linux /12/. V nadaljevanju je predstavljen razvoj strojne in programske opreme, ki tvori preprost testni sistem. Strojni del temelji na ISA razširitveni kartici za osebni računalnik, ki omogoča nadzor IEEE 1149.1 testnega vodila. Vmesnik je ustrezno podprt z gonilnikom za OS Linux ter višje nivojskimi namenskimi programi. Programski del je zasnovan kot skupek neodvisnih modulov z dobro definiranimi medsebojnimi povezavami. To omogoča enostavno nadgradnjo ter razvoj dodatnih testnih orodij, neodvisno od nižje nivojske strojne opreme. Delo na projektu EBS še ni zaključeno, saj je v načrtu razvoj izboljšanega vmesnika kot tudi nekaterih dodatnih programskih orodij. Celoten projekt je dostopen preko internetne spletne strani /13/ kot del zbirke prostega programja Sourceforge /14/.

2 Struktura projekta EBS

Projekt EBS je sestavljen iz strojnega ter programskega dela. Kot osnovna platforma je privzeta PC arhitektura i386 z OS Linux. V programskem delu želimo realizirati univerzalno platformo, ki bi podpirala krmilnike testnega vodila različnih proizvajalcev. Ti krmilniki /15/ se medsebojno razlikujejo v hitrosti delovanja (od 20 do 65 MHz), širini podatkovne besede (8/16 bitni), velikosti medpomnilnika (od 16b do 8Kb) itd. Njihova osnovna lastnost je, da bistveno poenostavijo povezavo med klasičnim vzporednim vodilom in enim ali več neodvisnimi IEEE 1149.1 testnimi vodili. Za učinkovitejši razvoj in izvedbo testnega postopka je potrebno obstoječi strojni opremi zagotoviti ustrezno programsko podporo. Načrtovano programsko platformo projekta EBS smo razdelili na več medsebojno neodvisnih modulov: gonilniki, vtični moduli, SVF prevajalnik ter grafični vmesniki. Struktura projekta EBS je shematično prikazana na sliki 3.

2.1 Strojna platforma

V okviru projekta smo implementirali preprost računalniški (PC) vmesnik za ISA vodilo, ki temelji na krmilniku testnega

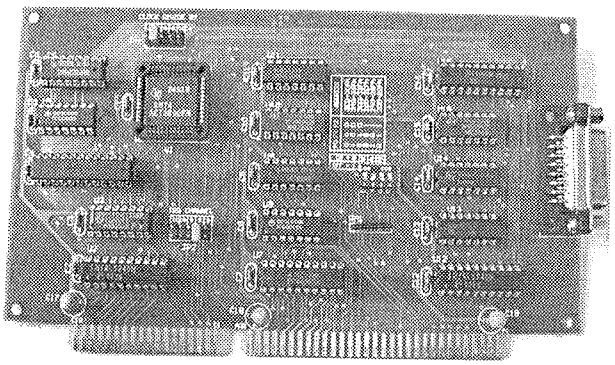


Slika 3: Strojna ter programska struktura projekta EBS

vodila SN74ACT8990 /16/ proizvajalca Texas Instruments. Njegova naloga je generiranje ustreznih signalov, ki osebnemu računalniku omogočajo dostop do testne infrastrukture v IEEE 1149.1 združljivih vezjih preko testnega vodila. Standard določa, da sestavljajo vodilo štiri obvezne linije (TMS, TCK, TDO, TDI) ter ena neobvezna linija (TRST*). Krmilni liniji TCK in TMS skrbita za pravilen prenos podatkov preko linij TDI in TDO. SN74ACT8990 omogoča vzporedno krmiljenje največ šestih IEEE 1149.1 testnih vodil, ki uporabljajo ločene TMS linije, medsebojno pa si delijo TDI, TDO in TCK linije.

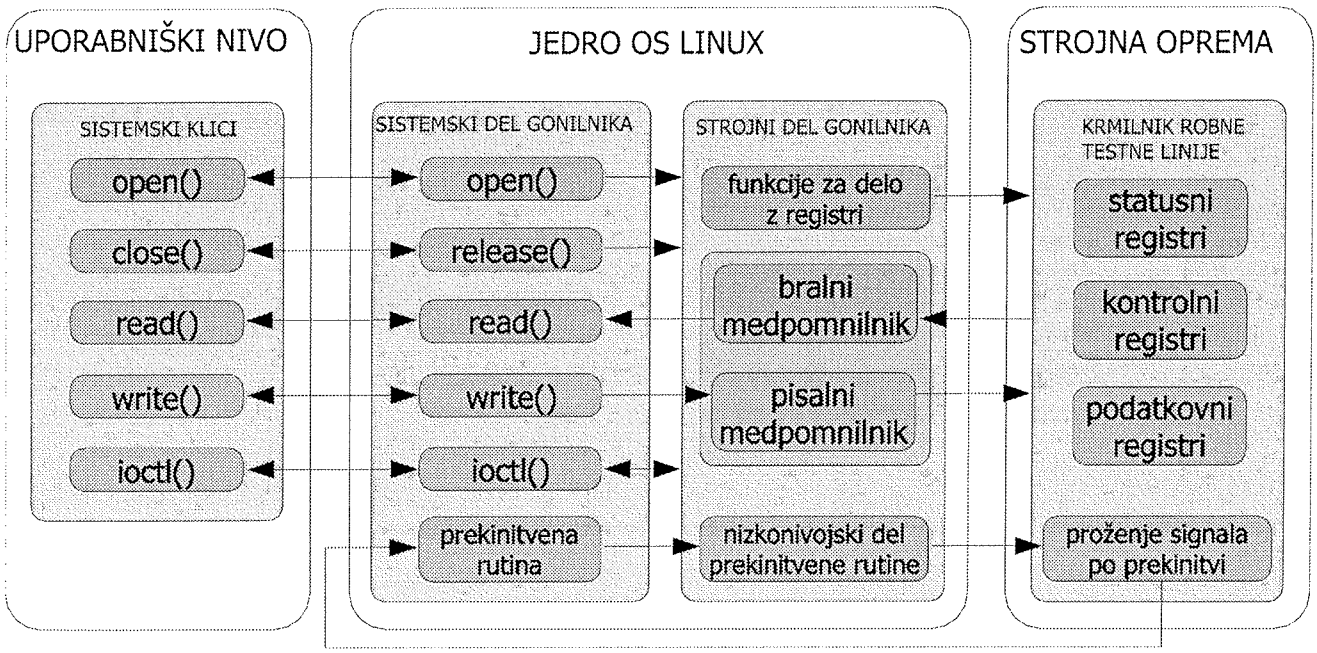
PC vmesnik, ki je prikazan na sliki 4, sestoji iz naslednjih sklopov:

- logike 16-bitnega ISA vodila;
- delilnika testne ure TCK;
- krmilnika testnega vodila SN74ACT8990;
- logike testnega vodila.



Slika 4: Računalniški vmesnik med vodilom ISA ter testnim vodilom

Podatkovni, kontrolni ter statusni registri krmilnika SN74ACT8990 so preko vodila ISA prezrcaljeni v vhod-



Slika 5: Shematičen prikaz zgradbe gonilnika za OS Linux

no/izhodni prostor sistemskega procesorja. Na razširitveni kartici je mogoče preko stikal izbrati osnovni V/I naslov ter ustrezen IRQ kanal. SN74ACT8990 deluje do frekvence 30 MHz, vendar pa obstoječa izvedba uporablja sistemsko uro vodila ISA (8,33 MHz). Frekvenco urinega takta lahko dodatno znižamo za 2x, 4x, 8x ali 16x z ustrezno nastavitvijo delilnika testne ure. Tri nivojski gonilniki, ki povezujejo krmilnik robne testne linije s testnim vodilom, so namenjeni zaščiti krmilnika ter izboljšanju kvalitete signalov na testnem vodilu. Kljub temu, da razširitvena kartica ne omogoča nobene nadaljnje obdelave oz. shranjevanja testnih podatkov, predstavlja uporabno in predvsem ceneno rešitev v primerjavi s komercialno dostopnimi razširitvenimi karticami. V prihodnosti načrtujemo realizacijo nekoliko zmogljivejšega vmesnika za vodilo PCI.

2.2 Programska struktura

Operacijski sistem Linux ima zaradi svoje odprtosti in zanesljivosti širok krog uporabnikov v raziskovalnih ter akademskih ustanovah. V postopku razvoja programske opreme, se je izkazal kot zelo solidno programsko okolje, ki uporabniku omogoča dober nadzor tudi nad najnižjimi programskimi nivoji. Programska platforma projekta EBS je razdeljena na dva osnovna modula (slika 3) in sicer:

- gonilnik krmilnika za OS Linux;
- okolje za pripravo testnih postopkov.

2.2.1 Gonilnik krmilnika za OS Linux

Gonilniki so programski moduli, ki so povezani v jedro OS z namenom, da ločeno nadzorujejo delovanje pripadajoče strojne opreme ter zakrijejo podrobnosti le-te pred uporabnikom /17/. Tako namenski programi iz uporabniškega nivoja dostopajo do jedra sistema in s tem do različnih funkcij strojne opreme s pomočjo standardnih sistemskih klicev. Jedro OS obravnava klice teh funkcij kot posebne zahteve in jih prenaša na ustrezne funkcije-metode znotraj programske kode gonilnikov /18/. Realizirani gonilnik krmilnika SN74ACT8990 podpira pet sistemskih klicev: *open(), close(), read(), write()* ter *ioctl()* /19/. Uporabnik pri pisanju testnih programov posredno preko sistemskih klicev oz. metod upravlja s krmilnikom testnega vodila. Gonilnik je zasnovan tako, da podpira nastavitev vseh možnih konfiguracij krmilnika ter tako omogoča izrabo vseh njegovih zmogljivosti. SN74ACT8990 lahko deluje v prekinitvenem načinu, zato gonilnik vključuje ustrezno prekinitveno funkcijo, ki skrbi za praznjenje pisalnega ter polnjenje bralnega medpomnilnika (slika 5). Prednost takšnega delovanja je, da uporabniku na višjih nivojih programiranja ni potrebno skrbeti za sprotno prenašanje testnih podatkov med sistemskim pomnilnikom in bralnim oz. pisalnim registrom krmilnika.

Programska koda gonilnika je razdeljena na strojni ter sistemski del. Prvi dostopa neposredno do registrov krmilnika ter tako skrbi za komunikacijo med krmilnikom ter gonilnikom (programiranje na nivoju registrov), drugi del pa tvori povezavo med jedrom OS Linux ter gonilnikom (slika 5). Takšna zasnova nam omogoča, da lahko v primeru spremembe strojne opreme, uporabimo isti gonilnik, saj je potrebno prilagoditi zgolj nižje nivojski del programske kode.

2.2.2 Okolje za pripravo testnih postopkov

Testni postopek lahko izvedemo neposredno z uporabo sistemskih klicev, kar sicer ponuja popoln nadzor nad izvajanjem testne procedure, vendar pa je za testiranje kompleksnih vezij takšen postopek preveč zamuden. V ta namen smo razvili prevajalnik za jezik SVF (ang. Serial Vector Format). SVF je standardiziran jezik /20/ za opis testnih

postopkov na osnovi robne testne linije. Struktura SVF je dokaj preprosta, saj ne vključuje nobenih odločitvenih in zančnih struktur. Definira štirinajst različnih stavkov, katerih parametri so v šestnajstiškem številskem sistemu zapisane logične vrednosti vhodnih in izhodnih vektorjev ter bitnih mask. Odzive vezja na vhodne vektorje je mogoče preko izhodne maske primerjati s pričakovanimi vrednostmi. Slika 6 podaja primer testnega postopka v jeziku SVF, ki v testirano vezje najprej prenese 8 bitov dolg ukaz (41) ter nato še 32 bitov dolg vhodni testni vektor (ABCD1234). Pričakovan odziv na vhodni vektor (11112222) primerjamo z dejanskim odzivom na vsakem bitu (maska FFFFFFFF).

```
.
.
.
SIR 8  TDI (41);
SDR 32 TDI (ABCD1234) TDO (11112222) MASK (FFFFFFFF);
.
.
.
```

Slika 6: Zapis v jeziku SVF

Prevajalnik je bil implementiran s pomočjo programskih orodij Flex ter Bison, postopek prevajanja pa združuje tri vrste analiz /21/:

- leksikalno analizo - iz vhodnih simbolov tvori nosilce pomena;
- sintaktično analizo - preveri strukturo stavkov;
- semantično analizo - določi pomen posameznim stavkom.

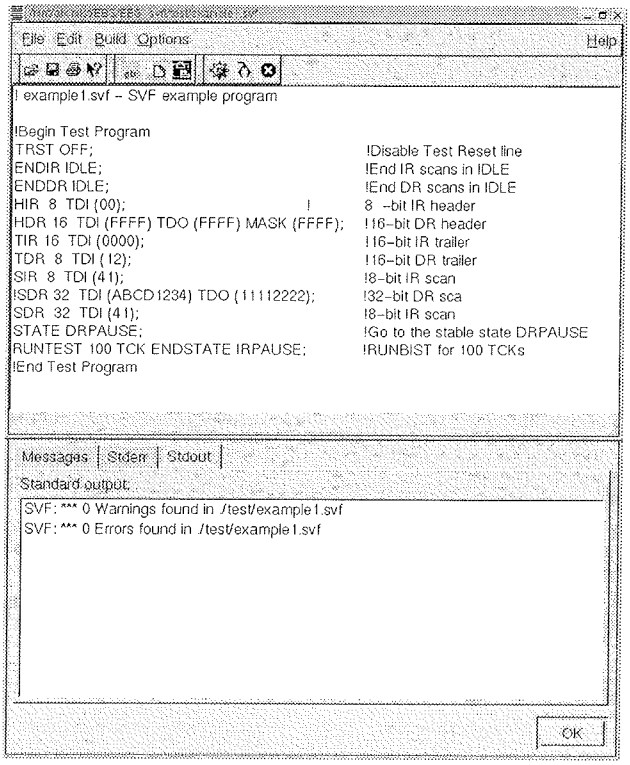
Prevajalnik deluje v t.i. interpretativnem načinu, kar pomeni da mora vsak stavek jezika SVF najprej uspešno prestati vse tri faze prevajanja, šele nato lahko temu sledi ustrezna akcija na strojni platformi. Prevajalnik je neodvisen od strojne platforme, saj so vse funkcije, ki izvajajo operacije nad specifičnim krmilnikom, vključene v posebno dinamično povezljivo knjižnico /22,23/. Splošen SVF prevajalnik lahko tako uporabimo za različne izvedbe krmilnikov testnega vodila, pri čemer je potrebno za vsakega zagotoviti le ustrezen vtični programski modul (ang. plug-in).

Predstavljen prevajalnik je sprva deloval zgolj v terminalskem načinu, kasneje pa je bil razširjen z ustreznim grafičnim vmesnikom (GV). Ta omogoča enostavnejše upravljanje tistim uporabnikom, ki jim je grafično okolje bolj domače. Slika 7 prikazuje realizirani grafični vmesnik.

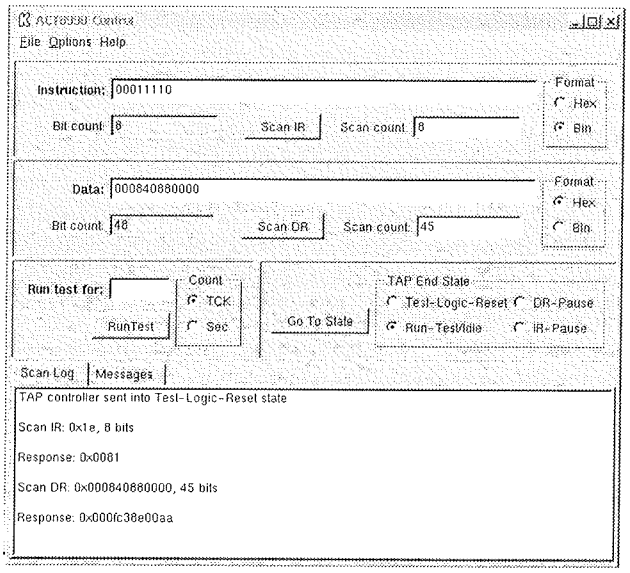
Poleg tega je bil razvit tudi preprost grafični vmesnik, s pomočjo katerega lahko uporabnik izvaja osnovne testne procedure združljive s standardom IEEE 1149.1, kot so pošiljanje posameznih ukaznih ter podatkovnih nizov ter upravljanje s testno infrastrukturo naprave. Ta grafični vmesnik (slika 8) dostopa do naprave neposredno preko sistemskih klicev.

3 Uporaba orodij ter razvoj testa

Za boljše razumevanje opisanih orodij bomo celoten testni postopek predstavili na preprostem zgledu. Vezje na sliki



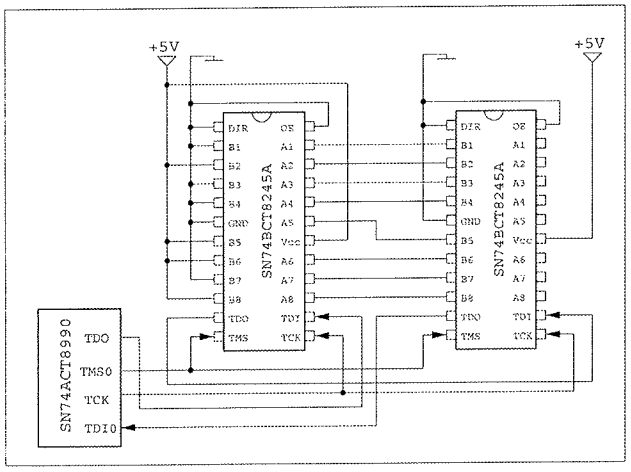
Slika 7: Grafični vmesnik za izvedbo testa opisanega s SVF jezikom



Slika 8: Preprost grafični vmesnik za izvajanje osnovnih IEEE 1149.1 procedur

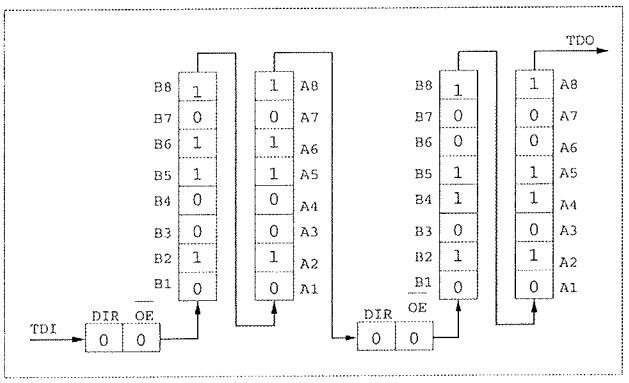
9 je sestavljeno iz dveh integriranih vezij SN74BCT8245A /24/, ki sta združljivi s standardom IEEE 1149.1. S pomočjo robne testne linije želimo odkriti morebitne napake na povezavah med priključki A1-A8 prve ter priključki B1-B8 druge komponente.

Povezave med posameznimi komponentami tiskanega vezja testiramo s pomočjo IEEE 1149.1 ukaza EXTEST. Komponente z vgrajeno robno testno linijo imajo na vseh digital-



Slika 9: Zgled preprostega vezja z vgrajeno robno testno linijo

nih vhodih in izhodih vgrajene posebne spominske celice, ki so povezane v robni pomikalni register /25/. Na vezje s slike 9 lahko gledamo kot na dve medsebojno povezani verigi spominskih celic, ki tvorita enoten pomikalni register. Testni postopek je sestavljen iz niza ukazov in testnih vektorjev. V 8-bitna ukazna registra obeh vezij je potrebno najprej naložiti kodo ukaza EXTEST, ki je v primeru SN74BCT8245A enaka vrednosti 00h (v šestnajstiškem zapisu). Vsak robni pomikalni register ima 18 celic, torej je testna linija vezja na sliki 8 dolga 36 bitov. V robna pomikalna registra vpišemo ustrezen testni vektor (001640000h), s katerim vzbujamo izhodne sponke prve komponente (A1-A8). Logično stanje posameznih spominskih celic po vpisu prvega vektorja je prikazano na sliki 10.



Slika 10:Logične vrednosti v pomikalnih registrih po vpisu prvega testnega vektorja - vzbujanja

```
!Begin Test Program
ENDIR IDLE;                                  !End IR scans in IDLE
ENDDR IDLE;                                  !End DR scans in IDLE
SIR 16 TDI (0000) TDO (8181) MASK (FFFF); !16-bit instruction scan
SDR 36 TDI (001640000);                      !36-bit data scan,first vector
SDR 36 TDI (000000000) TDO (000005959) MASK (00000FFFF); !36-bit data
!scan, second vector and response to first vector
!End Test Program
```

Slika 11:SVF zapis testnega postopka

Z vpisom naslednjega vektorja zajamemo in prenesemo odziv na vhodnih sponkah druge komponente (B1-B8) iz celic pomikalnega registra ter ga primerjamo s pričakovano vrednostjo (000005959h). Slika11 podaja SVF zapis opisane testne procedure.

4 Zaključek

Načrtovalci tiskanih vezij se pogosto ne zavedajo dovolj potrebe po vnaprejšnjem načrtovanju zmožnosti testiranja vezja, ki pa ob čedalje večji kompleksnosti postaja ena izmed osnovnih zahtev za uspešno realizacijo končnega proizvoda. Uporaba tehnike robne testne linije lahko občutno poenostavi odkrivanje morebitnih proizvodnih napak kot tudi vzdrževanje sistema med njegovo celotno življenjsko dobo.

Osnovni namen projekta EBS je realizacija preprostega a uporabnega orodja za izvajanje testnih postopkov v skladu s standardom IEEE 1149.1. Pomembni lastnosti predstavljenega sistema sta njegova odprtost in fleksibilnost, ki omogočata poseganje v vse nivoje delovanja. Namenjen je predvsem raziskovalnim ter akademskim ustanovam, saj omogoča enostavno nadgradnjo ter nadaljnji razvoj v skladu s specifičnimi potrebami uporabnika, obenem pa upamo, da bo naš projekt pripomogel tudi k širši uveljavitvi predstavljene tehnike testiranja med načrtovalci in razvojnimi inženirji v domači industriji.

Literatura:

/1/ K.P. Parker, *The Boundary-Scan Handbook*, 2nd edition, Kluwer Academic Publishers, Dordrecht, 1998, pp. 1-7.

/2/ IEEE, Standard test access port and boundary-scan architecture, IEEE Standard 1149.1a-1993 (1993).

/3/ R. Nelson, "Boundary-scan software aids PCB evaluation", Test & Measurement World, October (1999).

/4/ R. Raina, R. Bailey, D. Belete, V. Khosa, R. Molyneaux, J. Prado, A. Razdan, "DFT advances in Motorola's next-generation 74xx PowerPC microprocessor", Prodeedings on ITC 2000, pp. 131-140.

/5/ F. Golshan, "Test and on-line debug capabilities of IEEE Std 1149.1 in UltraSPARC-III microprocessor", Proceedings on ITC 2000, pp. 141-150.

/6/ B. Kerridge, "Flash stretches boundary-scan limits", Test & Measurement Europe, June (2002).

/7/ R. Nelson, "Systems expand IEEE 1149.1 test", Test and Measurement World, February (2000).

/8/ K. Lofstrom, "Early capture for boundary scan timing measurements", Proceedings on ITC 1996, pp. 417-422.

/9/ M. Santo-Zarnik, F. Novak, U. Kač, S. Maček, "Experiments with IEEE 1149.4 KLIC test chip: a case study", Proceedings on IMSTW 1999, pp. 131-135.

/10/ U. Kač, F. Novak, S. Maček, M. Santo-Zarnik, "Alternative test methods using IEEE 1149.4", Proceedings on DATE 2000, pp. 463-467.

/11/ U. Kač, F. Novak, Florence Azaďs, Pascal Nouet, Michel Renovell, "Implementation of an experimental IEEE 1149.4 mixed-signal test chip", 1st IEEE International Workshop on Board Test, October 2002.

/12/ Free Software Foundation, spletna stran http://www.fsf.org.

/13/ Spletna stran projekta EBS: http://ebsp.sourceforge.net.

/14/ Spletna stran projekta Sourceforge: http://sourceforge.net.

/15/ U. Kač, R.Sedevčič, F. Novak, A. Biasizzo, "Linux-based experimental boundary scan environment", Microprocessors and Microsystems, Vol. 26, No. 5, 2002, pp. 199-206.

/16/ Test-Bus Controller SN74ACT8990, Application Report, Texas Instruments, 2000, spletna stran http://www.ti.com/sc/docs/jtag/jtaghome.html.

/17/ R.Sedevčič, Izvedba gonilnikov v OS Linux, Monitor, januar 2002.

/18/ A.Rubini, J. Corbet, Linux Device Drivers, 2nd edition, O'Reilly & Associates, Inc., 2001.

/19/ U.Kač, R.Sedevčič, F.Novak, A.Biasizzo, Eksperimentalno okolje za vezja z IEEE 1149.1 testno infrastrukturo, Elektrotehniška konferenca, Portorož, 2001.

/20/ Serial Vector Format Specification, Revision E, Asset InterTech, 1999, spletna stran http://www.asset-intertech.com/support/svf.pdf.

/21/ N. Wirth, Compiler Construction, Addison Wesley, 1996.

/22/ R.Sedevčič, Programske knjižnice v Linuxu, Monitor, januar 2003.

/23/ David A. Wheeler, Program Library HOWTO, spletna stran http://www.dwheeler.com/program-library/.

/24/ Texas Instruments Inc., SN74BCT8245A Scan Test Device with Octal Bus Transceivers, data sheet, http://www-s.ti.com/sc/psheets/scbs043e/scbs043e.pdf.

/25/ U. Kač, R.Sedevčič, Testiranje vezij z vgrajeno IEEE 1149.1 testno infrastrukturo, Svet elektronike, letnik 9, januar 2002.

Robert Sedevčič
Institut "Jožef Stefan"
Jamova 39, 1001 Ljubljana, Slovenia
tel.: +386 (0)1 477 3550, fax: +386 (0)1 251 9385
email: robert.sedevcic@ijs.si

mag. Uroš Kač
Institut "Jožef Stefan"
Jamova 39, 1001 Ljubljana, Slovenia
tel.: +386 (0)1 477 3550, fax: +386 (0)1 251 9385
email: uros.kac@ijs.si

prof. dr. Franc Novak
Institut "Jožef Stefan"
Jamova 39, 1001 Ljubljana, Slovenia
tel.: +386 (0)1 477 3386, fax: +386 (0)1 251 9385
email: franc.novak@ijs.si

Prispelo (Arrived): 27.01.2003 *Sprejeto (Accepted): 26.08.2003*

MIKROVALOVNI FERITI

Darja Lisjak

Institut Jožef Stefan, Odsek za sodobne materiale, Ljubljana, Slovenija

Ključne besede: feriti, elektromagnetne lastnosti, kristalna strukutra, mikrostruktura

Izvleček: V dobi vedno večjega izkoriščanja elektromagnetnega valovanja povezanega z napredkom v informacijsko-telekomunikacijski tehnologiji se povečuje potreba po boljših ali novih elektronskih komponentah ter po zaščiti pred nezaželenimi elektromagnetnimi motnjami. Celoten razvoj je odvisen od razvoja ustreznih materialov in merilnih tehnik. V prispevku je zbran pregled magnetnih feritnih materialov, t.i. mikrovalovnih feritov, primernih za uporabo v radio frekvenčnem območju. Povzete so njihove elektromagnetne lastnosti, kristalna struktura, mikrostruktura ter soodvisnost le-teh. Podana sta tudi dva primera raziskav na Odseku za sodobne materiale Instituta Jožef Stefan.

Microwave Ferrites

Key words: ferrites, electromagnetic behaviour, crystal structure, microstructure

Abstract: The demand for new and better electronic components is increasing in line with the exploitation of the electromagnetic spectrum as a consequence of the developments in information and telecommunications. These developments are basically dependent on the introduction of improved materials and measuring techniques. Magnetic ferrite materials, i.e. microwave ferrites, suitable for applications in radio-frequency range are reviewed.

The applications of microwave ferrites are based on the interactions between an electromagnetic field and the ferrites. This interaction is determined by the frequency of the applied electromagnetic field, the strength of the static external field and on the ferrites' dielectric properties, magnetization, ferromagnetic resonance, permeability and magnetic losses.

Ferrites are ferrimagnetic oxides. The source of the ferrimagnetism is the antiferromagnetically coupled ferromagnetic sublattices with different magnetic moments. The superexchange interaction between the magnetic cations strongly depends on the type of magnetic ion, the bonding length and angle between the cation and the oxygen. Therefore, it is possible to tailor the ferrites' magnetic properties (i.e. magnetization, Curie temperature, anisotropy, ferromagnetic frequency) with their composition and crystal structure. Three types of microwave ferrites can be distinguished with respect to their crystal structure: spinel, garnet and hexagonal ferrites. Among these the spinel ferrites are the most widely used. They possess the highest saturation magnetization and they are applied at 3-30 GHz. The main advantage of garnets is their low losses, which makes them the best microwave material at 1-10 GHz. Hexagonal ferrites or hexaferrites possess a very high magnetocrystalline anisotropy, which makes them suitable for applications at higher frequencies, up to 100 GHz. Due to the variety of structural modifications the greatest variety of properties among ferrites can be achieved in the hexaferrite family.

The microstructure is definitely a parameter to be considered when we discuss the electromagnetic properties of ferrites. The preparation has the greatest influence on the microstructure of the ferrite with a particular composition. Therefore, magnetic properties, like permeability, anisotropy, dielectric and magnetic losses, ferromagnetic resonance frequency, can be tailored by the preparation conditions. A survey of the microstructural parameters influencing a particular electromagnetic property is included in the paper.

Additionally, two examples of the research at the Advanced Materials Department, Jožef Stefan Institute, are summarized. The research was focused on hexaferrites as one of the most promising materials for mm-wave applications. The first example is related to the M-hexaferrites with composition $BaFe_{12-2x}A_xSn_xO_{19}$, where A = Co, Ni, Zn and x = 0.1-2.5. Both, the saturation magnetization and the coercivity of the samples varied with the composition and the preparation method. The influence of the composition was more pronounced on the saturation magnetization, while the opposite was true for the coercivity. The second example is related to the synthesis of U-hexaferrites with the composition $Ba_4A_2Fe_{36}O_{60}$, where A = Co, Ni, Zn. The preparation of single-phase U-hexaferrites is very difficult, due to their complex crystal structure. A specially modified solid-state synthesis was used for this purpose. The saturation magnetization and Curie temperature varied only with the composition and not with the preparation method. The sample with composition $Ba_4Zn_2Fe_{36}O_{60}$ may be suitable for mm-wave applications.

1. Uvod

Feriti so magnetni oksidi Fe, ki združujejo najboljšo možno kombinacijo lastnosti električnih izolatorjev in magnetnih materialov z izredno fleksibilnostjo pri kontroli magnetnih in prevodniških lastnosti ter mrežnih parametrov. Mikrovalovni ferit je magneten material z visoko upornostjo, ki se uporablja pri 100 MHz - 300 GHz. Med mikrovalovne ferite štejemo garnete, spinelne in heksagonalne ferite. Njihova uporaba je odvisna od sestave, kristalne strukture in mikrostrukture. Uporabljajo se lahko kot del brezžičnih telekomunikacijskih sistemov (mobilna telefonija, satelitski sistemi, avtomobilska in vojaška industrija) ali samostojno kot absorberji elektromagnetnega valovanja. Njihovo uporabo bi lahko razdelili v naslednje kategorije: /1-3/

- Nerecipročne naprave, za katere so feriti nenadomestljivi. Naprave so izolatorji in cirkulatorji, naprave za fazni zamik (phase shifter), filtri...
- Recipročne naprave so električno kontrolirane naprave za fazni zamik, stikala, spremenljivi atenuatorji, katere lahko nadomestimo tudi s polprevodniškimi elementi.
- Nelinearne naprave, ki izkoriščajo nelinearni odziv materiala, npr. omejevalniki moči.
- Mikrovalovni absorberji, ki absorbirajo elektromagnetno valovanje v določenem frekvenčnem območju.

2. Značilnosti polikristaliničnih feritov v radio frekvenčnem območju

Območje elektromagnetnega (EM) spektra uporabe mikrovalovnih feritov do 300 GHz imenujemo tudi radio frekvenčno (RF) območje, kar ustreza valovni dolžini v vakuumu do najmanj 1 mm. Za interakcijo EM valovanja z medijem je pomembno, da EM polje prodira v material. Interakcija med materialom in RF poljem je osnova za delovanje mikrovalovnih feritnih naprav. Širjenje EM valovanja v feritu je popolnoma definirano z Maxwell-ovimi enačbami, dielektričnimi in magnetnimi lastnostmi materiala (dielektrična konstanta, dielektrične izgube, magnetizacija, permeabilnost, magnetne izgube) ter robnimi pogoji, ki jih določa geometrija sistema.

2.1 Dielektričnost magnetnega materiala izvira iz elektronske, ionske, lastne dipolne in medploskovne (interface) polarizacije. Dielektrična konstanta (ε) se manjša s frekvenco in se veča s temperaturo do maksimalne vrednosti ter se z nadaljnjim večanjem temperature manjša. Do temperature maksimalne ε se povečuje ionska polarizacija, z nadaljnjim večanjem temperature prevladajo termične oscilacije molekul, kar zmanjša stopnjo urejenosti in s tem ε. Sprememba ε s temperaturo je večja pri nižjih frekvencah. /4, 5/

2.2 Dielektrične izgube v mikrovalovnih feritih izvirajo v elektronskem "hopping-u" med Fe^{2+} in Fe^{3+}. Za zmanjšanje izgub se je potrebno znebiti Fe^{2+}, kar pomeni ustrezno pripravo stehiometričnega produkta iz homogenega stehiometričnega izhodnega material. Merilo za dielektrične izgube je tan δ_ε, ki predstavlja fazni zaostanek dipolnih oscilacij glede na zunanje električno polje. Na to vpliva število in vrsta prisotnih ionov. /2, 4/

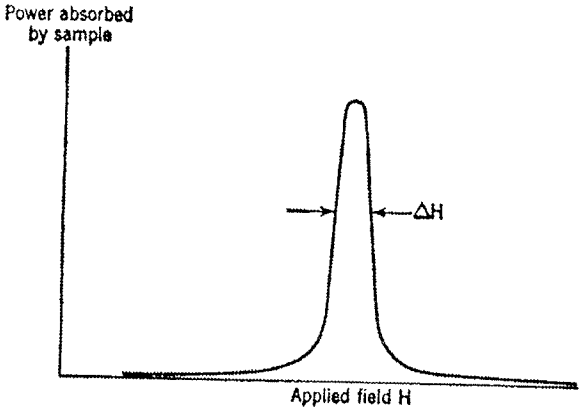
2.3 Magnetizacija direktno določa učinkovitost materiala. Magnetizacija feritov je posledica spinskih momentov elektronov. Če postavimo elektron v usmerjeno magnetno polje, se bo njegov magnetni moment usmeril s poljem tako, da minimizira svojo potencialno energijo. Če je RF polje pravokotno na usmerjeno polje, magnetizacija precesira okrog ravnotežne smeri. /2/

2.4 Feromagnetna resonanca (FMR): Obstajata dva osnovna načina prenašanja EM valovanja, ki potuje skozi ferit. Načina prenašanja imata nasprotni predznak polarizacije. V feritu končne dolžine ena polarizacija bolj interagira s feritom kot druga, zardi česar je vhodno valovanje drugačno kot izhodno. Če je frekvenca tega cirkularno polariziranega polja enaka precesiji magnetnih momentov, pride do posebej močne absorpcije RF polja v feritu, kar se imenuje feromagnetna resonanca (FMR). V feritu pride do FMR, ko frekvenca spreminjajočega se polja ustreza pogojem podanim z enačbo (1)/1/, pri čemer so ω kotna hitrost ($\omega=2\pi f$), γ giromagnetno razmerje, μ_0 permeabilnost vakuuma, M_s nasičena magnetizacija, H_r resonačno polje ter N_x, N_y in N_z faktorji demagnetizacije elipsoida v smereh x, y in z ($N_x + N_y + N_z = 1$).V materialu samem se notranje polje spreminja od ene do druge točke in zato naravna resonanca obsega neko frekvenčno območje.

$$\omega = \gamma\mu_0\{[H_r - (N_z - N_x)M_s][H_r - (N_z - N_y)M_s]\}^{1/2} \quad (1)$$

Interakcija s poljem polariziranim v nasprotni smeri je precej šibkejša. Smer rotacije magnetnih momentov je določena s smerjo zunanjega statičnega magnetnega polja. S spremembo usmeritve zunanjega polja lahko kontroliramo interakcijo med RF poljem in materialom. Ta efekt je osnova za nerecipročne naprave. /2/

2.5 Magnetne izgube so izražene v imaginarnemu delu permeabilnosti μ". Magnetne izgube so posledica premika domenskih sten in rotacije magnetizacije. μ" je največji pri FMR. Energija, ki jo absorbira material iz RF polja, se zaradi izgub sčasoma pretvori v toploto. Izgube izraža termin dušenja, ki opisuje relaksacijski čas potreben, da magnetizacija doseže ravnotežno stanje. Dušenje je odvisno od frekvence RF polja, temperature in kristalografske smeri. Eksperimentalno izmerimo dušenje z resonančno širino (ΔH), ki je širina črte FMR absorpcijske krivulje pri resonanci na polovični višini absorpcijskega vrha. FMR absorpcijska krivulja prikazuje μ" v odvisnosti od magnetnega polja, kar je prikazano na sliki 1. /2, 6/



Slika 1: Absorpcija ferita v RF polju [6]

3. Struktura in ferimagnetizem feritov

Feriti so ferimagnetni oksidi. Kisikovi ioni so razporejeni v osnovni celici okrog kovinskih v obliki tetraedra, oktaedra, dodekaedra, pentagonalne bipiramide. Med magnetnimi ioni preko kisikovih poteka superizmenjalna magnetna interakcija. Magnetni ioni v feritni mreži tvorijo magnetno podmrežo. Struktura posamezne magnetne podmreže je ponavadi kolinearna feromagnetna, različne podmreže pa so med sabo sklopljene antiferomagnetno. Rezultat različnega števila magnentih ionov v različnih podmrežah je skupni magnetni moment, ki je izvor ferimagnetizma. Ker je superizmenjalna interakcija odvisna od vrste magnetnega iona, dolžine vezi in veznega kota, je s substitucijo z različno velikimi ioni mogoče vplivati na magnetne lastnosti

kot so magnetizacija, Curijeva temperatura Tc, anizotropija, frekvenca FMR. /2, 7/

3.1 Spinelni feriti so najbolj pogosto uporabljeni mikrovalovni feriti. Uporabljajo se pri 3-30 GHz. Od vseh mikrovalovnih feritov imajo največjo nasičeno magnetizacijo, do približno 5500 G. /3/

Kristalna struktura spinelov je izomorfna z mineralom spinel $MgAl_2O_4$. Pri substituciji Al^{3+} z Fe^{3+} in Mg^{2+} z Fe^{2+} dobimo magnetit, Fe_3O_4. V spinelnih feritih je možna substitucija Fe^{2+} z različnimi dvovalentnimi kationi (Ni, Co, Mn, Cu...). Spinelno kubično osnovno celico sestavlja osem enot s 16 M_1^{3+} in 8 M_2^{2+} kationi. V normalnem spinelu 8 M_2^{2+} zaseda osem tetraedrskih A mest in 16 M_1^{3+} zaseda 16 oktaedrskih B mest. V inverznem spinelu osem od 16 M_1^{3+} zaseda vsa tetraedrska A mesta. V spinelih so tri vrste negativne - antiferomagnetne superizmenjalne interakcije: A-A, B-B in A-B. Najmočnejša je A-B interakcija, zaradi česar sta dve podmreži usmerjeni v nasprotni smeri - antiferomagnetno.

Ena izmed boljših lastnosti feritov je možnost različnih substitucij. Na ta način lahko spreminjamo magnetni moment, jakost superizmenjave, Tc, stopnjo inverzije. Efekt substitucije je osnova za pripravo mešanih feritov za mikrovalovno tehnologijo. Najbolj tipične substitucije za spremembo lastnosti spinelnih feritov: /2, 7/

- Al^{3+} zmanjša magnetizacijo;
- Co^{2+} zmanjša anizotropijo;
- Mn^{2+} zmanjša dielektrične izgube;
- Zn^{2+} poveča magnetizacijo, a zniža Tc.

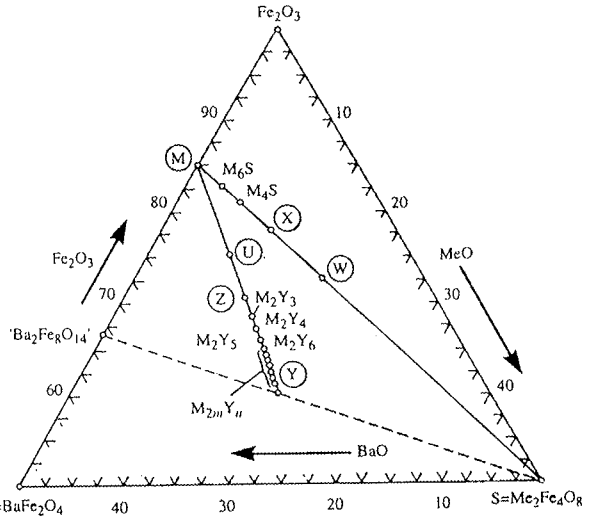
3.2 Garneti imajo manjše izgube kot spineli in so pri več aplikacijah boljši, so pa bolj občutljivi na mehanske napetosti kot spineli. Itrij železov garnet (YIG) je od svojega odkritja najboljši mikrovalovni material v območju 1-10 GHz. Za mikrovalovno napravo je potrebna majhna resonančna širina. Najboljši material je monokristal YIG, ki ima $\Delta H \leq$ 0.1G pri 10 GHz. Polikristalinični feriti imajo ΔH reda velikosti 10-100 G. /2, 8/

Ferimagnetni garneti so izomorfni z mineralom garnet $Ca_3Fe_2(SiO_4)_3$. Prvi mikrovalovni garnet je bil YIG, $Y_3Fe_5O_{12}$. V osnovni celici YIG je osem enot formule s skupno 160 ioni, torej 24 Y^{3+}, 40 Fe^{3+} in 96 O^{2-}. Največje nemagnetne katione Y^{3+} obdaja osem kisikovih ionov v popačenem dodekaedričnem okolju (mesta c). Pet Fe^{3+} je razdeljenih med tri tetraedrična mesta d in dve oktaedrični mesti a. Glavna superizmenjalna interakcija med mesti a in d je antiferomagnetna. Pri absolutni ničli je nasičena magnetizacija na molekulo garneta razlika med magnetizacijo ionov redkih zemelj na mestih c in vsoto magnetizacije Fe^{3+} ionov na mestih a in d. Interakcija med mesti c in d je precej šibkejša. Podmreža c je pri in nad sobno temperaturo šibko namagnetena. Pri nižji temperaturi sklopitev c-d prevlada nad termičnim efektom in podmreža c je z nižanjem temperature vedno bolj usmerjena, ima vedno večji prispevek k celotni magnetizaciji. /2, 7/

Ker je v garnetih Fe samo v trivalentnem stanju, imajo garneti manjše dielektrične izgube kot spineli. Magnetne lastnosti YIG lahko zelo spremenimo z različnimi substitucijami. Na magnetizacijo lahko vplivamo s substitucijo na tetra- in oktaedrskih mestih. S substitucijo na dodekaedrskih mestih z redkimi zemljami povečamo anizotropijo. Substituenti s 4f orbitalami, kot sta Ho^{3+} in Dy^{3+}, povečajo mikrovalovne izgube v obliki spinskega valovanja. S substitucijo na mestih a in d lahko vplivamo na Tc. Tako lahko načrtujemo material z ustreznimi magnetnimi lastnostmi. /3/

3. 3 Heksagonalni feriti se uporabljajo pri 1-100 GHz. So trdo magnetni materiali, z veliko koercitivnostjo - so permanentni magneti. Imajo polje anizotropije do 35 kG, magnetizacijo do 5 kG in Tc okrog 500°C. Imajo pravokotno histerezno zanko. Ker je Fe samo v trivalentnem stanju, so dobri izolatorji in imajo majhne dielektrične izgube. So temperaturno stabilni. Najbolj znan heksaferit je M-tip $BaFe_{12}O_{19}$ (Ferroxdure) in njegov analog $SrFe_{12}O_{19}$. /2, 9/

Heksagonalni feriti ali krajše heksaferiti so skupina feritov s heksagonalno ali romboedrično kristalno strukturo v sistemu B-A-Fe^{3+}-O. B predstavlja velik dvovaletni kation, npr. Ba, Sr, Pb, Ca ali kombinacijo le-teh. Ker se največ uporabljajo Ba-heksaferiti, od tu dalje izraz heksaferit označuje Ba-heksaferit. A predstavlja majhen dvovalentni kation, npr. Mn, Fe, Co, Ni, Cu, Zn ali kombinacijo le-teh. Že tako veliko možnih sestav lahko povečamo z delno substitucijo Fe^{3+} s trivalentnimi kationi (Bi, In) ali kombinacijo dvo- in štirivalentnih ionov (A-Ti, A-Ru, A-Ir, A-Sn).



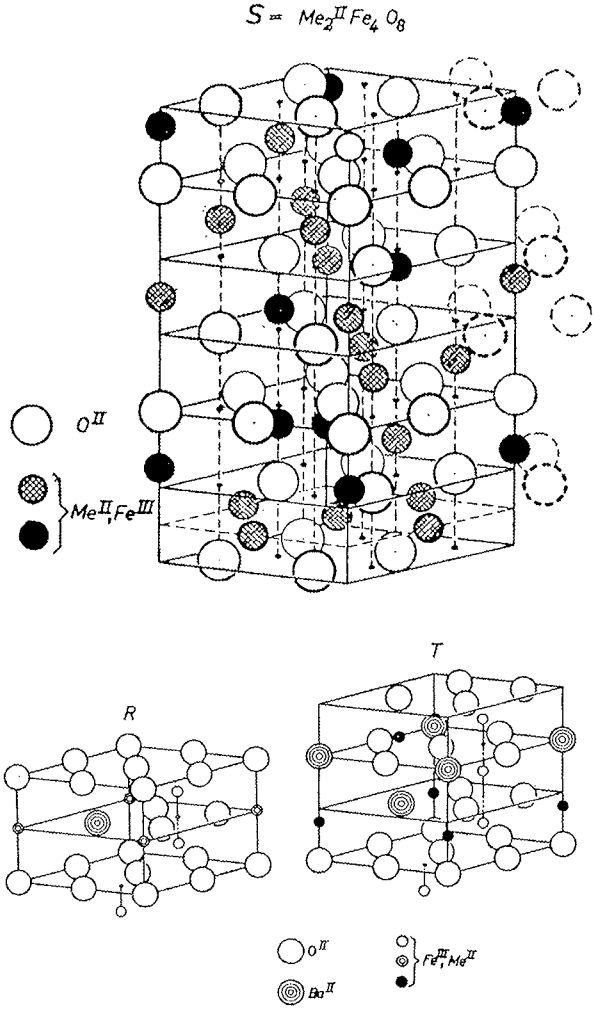
Slika 2: Fazni diagram Ba-heksaferitov /10/

Heksaferiti imajo različne kristalne strukture, odvisno od njihove sestave, kar je razvidno iz faznega diagrama Ba-heksaferitov prikazanega na sliki 2. /7, 10/ Za heksaferite je značilno da so sestavljenih iz treh osnovnih kristalnih blokov, ki se v različnem zaporedju nalagajo v smeri heksagonalne osi c in tako tvorijo različne osnovne celice. Trije osnovni kristalni bloki, S, R in T, so prikazani na sliki 3. Najbolj znani Ba-heksaferiti in njihova struktura so podani v tabeli 1. M-tip ima strukturo magnetoplumbita in je sestavljen iz R in S blokov. Na veznici M-S na sliki 2 se na-

hajata še W- in X- hexaferit, ki sta torej tudi kombinacija R in S blokov oz. M in S. Y-heksaferit je sestavljen iz T in S blokov. Na veznici M-Y najdemo še Z- in U-heksaferit, ki sta različni kombinaciji R, S in T blokov oz. M in Y. Obstaja 61 različnih strukturnih tipov heksaferitov na veznicah M-S in M-Y. Največji ($Ba_{62}A_{54}Fe_{420}O_{746}$) ima osnovno celico s konstantami a = 5.88 Å in c = 1577Å, ki je največja znana anorganska osnovna celica. /11/

*Tabela 1: Najbolj znani tipi heksaferitov in njihova kristalna struktura. Z * je označena rotacija strukturnega bloka za 180°. /10/*

tip	kristalna struktura	sestava
M	RSR*S*	$BaFe_{12}O_{19}$
Y	$(TS)_3$	$Ba_2A_2Fe_{12}O_{22}$
W	$R(S)_2R^*(S^*)_2 = MS$	$BaA_2Fe_{16}O_{27}$
X	$(RSR^*S^*_2)_3 = M_2S$	$Ba_2A_2Fe_{28}O_{46}$
Z	$RSTSR^*S^*T^*S^* = M_2Y_2$	$Ba_6A_4Fe_{48}O_{82} = Ba_3A_2Fe_{24}O_{41}$
U	$RSR^*S^*T^*S^* = M_2Y$	$Ba_4A_2Fe_{36}O_{60}$



Slika 3: Osnovni strukturni bloki heksaferitov: S, R in T /7/

V heksaferitni osnovni celici so ioni Fe^{3+} na različnih mestih z različno usmerjenimi magnetnimi momenti. Npr. v najenostavnejšem M-heksaferitu $BaFe_{12}O_{19}$ sta dve nasprotno usmerjeni feromagnetni podmreži. Eno sestavlja sedem Fe^{3+} na oktaedrskih mestih in dva na bipiramidalnih, drugo pa dva Fe^{3+} na oktaedrskih in dva na teatraedrskih mestih. Pri substituciji Fe^{3+} z drugimi kationi tako neposredno vplivamo na skupni magnetni moment in s tem na magnetizacijo materiala. Poleg magnetnega momenta substituenta je zelo pomembna tudi njegova preferenčna zasedba kristalografskih mest. /12-14/

Ferimagnetna resonanca spinelov in garnetov se pojavi pri frekvenci do nekaj GHz. V mm-območju jo lahko dosežemo z dodatnim zunanjim poljem 20 kG. Potrebi po velikem zunanjem magnetnem polju se lahko izognemo z uporabo permanentnih magnetih materialov, ki so zaradi tega bolj primerni za uporabo v mm-območju. Taki so heksaferiti z enoosno kristalno anizotropijo, ki imajo smer magnetizacije vzdolž c osi heksagona. Največjo enoosno anizotropijo imajo M-heksaferiti, katerega FMR je nad 30 GHz. /15/ Izvor velike magnetne anizotropije so Fe^{3+} na bipiramidalnih mestih s petštevno koordinacijo. /2, 3, 16-19/ S substitucijo kombinacije štirivalentnih kationov (Ti, Sn, Ir, Ru) skupaj z dvovalentnimi (Zn, Co, Ni, Mn) na mesta Fe^{3+} v M heksaferitu se pri kritičnem razmerju enoosna anizotropija spremeni v planarno. Planarni heksaferiti imajo magnetizacijo v smeri osnovne ravnine heksagona. V ostalih tipih heksaferitov so visoka anizotropna mesta "razredčena" v primerjavi z M-heksaferitom. Npr. v M-tipu je eno od 12 mest Fe bipiramidalno, v Z-tipu pa eno od 24, zato se Z-heksaferitom z ustrezno substitucijo anizotropija manj zmanjša in druge lastnosti manj spremenijo kot M-heksaferitom. Z velikostjo anizotropije v heksaferitih lahko vplivamo na frekvenco FMR, ki je tem večja, čim večja je aniztropija. Y-heksaferiti pa so že v osnovi planarno anizotropni in zato uporabni pri manjših frekvencah kot ostali tipi.

4. Mikrostruktura in elektromagnetne lastnosti feritov

Mikrostruktura feritov za uporabo v RF območju in tudi sicer zelo pomembno vpliva na njihove elektromagnetne lastnosti. Pri določeni sestavi na mikrostrukturo najbolj vpliva priprava. S spreminjanjem mikrostrukture oz. s spreminjanjem pogojev priprave lahko pripravimo material z različnimi elektromagnetnimi lastnostmi za različno uporabo. V tabeli 2 so zbrane pomembnejše elektromagnetne lastnosti za uporabo v RF območju in njihova odvisnost od posameznih mikrostrukturnih parametrov.

Za uporabo v RF območju morajo imeti feriti ustrezno permeabilnost pri frekvenci uporabe. Za frekvenco do približno 24 GHz so boljši spineli in garneti, pri večji frekvenci pa heksaferiti. Kot vidimo iz tabele 2, imajo visoko permeabilnost kemijsko homogeni grobozrnati feriti s čim večjo stopnjo orientacije zrn, čim manjšo poroznostjo, brez strukturnih in mikrostruktrunih defektov ter brez notranjih napetosti. FMR lahko povečamo z večanjem anizotropije, torej s stopnjo orientacije in z majhnimi zrni oz. s čim večjim razmerjem premer/debelina pri heksaferitih. Material, ki se uporablja za mikrovalovne naprave, mora imeti čim manjše

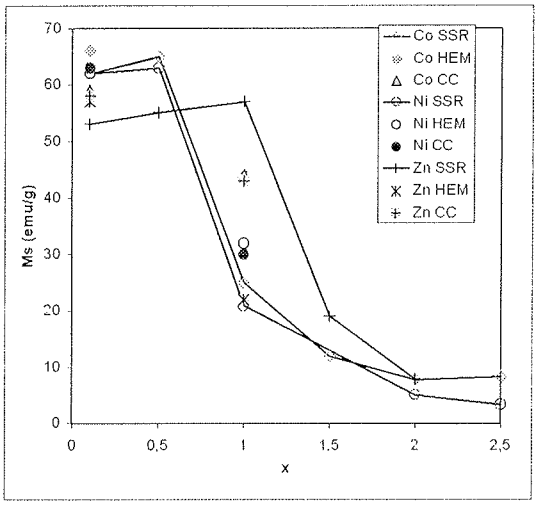
magnetne izgube s čim manjšo širino resonančne črte in čim manjše dielektrične izgube. Kot je razvidno iz tabele 2, tem pogojem ustreza kemijsko homogen material z minimalno poroznostjo, brez strukturnih in mikrostrukturnih defektov, brez notranjih napetosti. Po drugi strani pa mora imeti material za mikrovalovne absorberje čim večje magnetne in dielektrične izgube s čim večjo širino resonančne črte. Material ne sme imeti preveč finih zrn in mora biti čim bolje orientiran. Ker mikrostrukturni parametri različno vplivajo na različne lastnosti, je pogosto potrebno sprejeti kompromis med različnimi parametri, da bi pripravili ustrezen material za neko uporabo.

Tabela 2: Odvisnost elektromagnetnih lastnosti od mikrostrukturnih parametrov

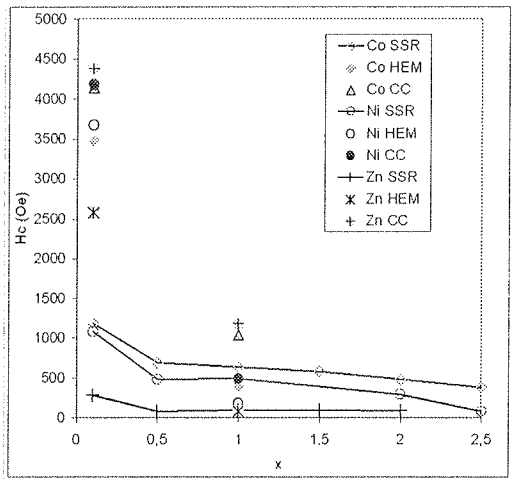
elektromagnetna lastnost	mikrostukturni parameter, ki jo poveča	mikrostrukturni parameter, ki jo zmanjša
permeabilnost [7, 16, 20-25]	velikost zrn stopnja orientacije	notranje napetosti poroznost lokalne kemijske, mikrostrukturne in kristalografske nehomogenosti ter napake
magnetne izgube [1, 22, 25]	poroznost stopnja orientacije velikost zrn	
anizotropija [26]	razmerje premer/debelina zrn	
notranja demagnetizacija [7]	poroznost	
koercitivnost [7, 14, 20, 27]	poroznost	velikost zrn
resonančna frekvenca [23, 28]	poroznost	velikost zrn
širina resonančne frekvence [7, 20]	notranje napetosti mikrostrukturne napake	velikost zrn
absorpcija elektromagnetnega valovanja [15, 21]	velikost zrn stopnja orientacije	
dielektričnost [25]		poroznost
dielektrične izgube [24, 25]	velikost zrn napetosti mikrostrukturne napake poroznost	

5. Raziskave na Odseku za sodobne materiale, Institut Jožef Stefan

Raziskave na Odseku za sodobne materiale so osredotočene na heksaferite. Prvi sklop raziskav obsega ASn-substituirane M-heksaferite s splošno kemijsko formulo $BaFe_{12-2x}A_xSn_xO_{19}$, pri čemer A = Co, Ni, Zn. /14/ Na sliki 4 je prikazana odvisnost nasičene magnetizacije (Ms) od sestave (x) in načina priprave (reakcija v trdnem - SSR, visokoenergetsko mletje - HEM in preobarjanje - CC). Vidimo, da Ms pada s stopnjo substitucije (z x) za CoSn- in NiSn-substitutirane vzorce. Pri ZnSn-substitutiranih vzorcih pa smo najprej izmerili povečanje magnetizacije za $x \leq 1$ ter padanje z nadaljnjo substitucijo. Tako odvisnost se da razložiti s preferenčno zasedbo mest Fe^{3+} s substituenti (Glej del 3.3.). Glede na možno preferenčno zasedbo mest Co^{2+} in Sn^{4+} /13/ se magnetni moment substituiranega heksaferita zmanjša v primerjavi z nesubstituiranim. Posledica tega je zmanjšanje magnetizacije. V primeru ZnSn-substituiranih vzorcev je odvisnost magnetizacije od sestave drugačna zaradi preference Zn^{2+} za zasedbo tetraedrskih mest v drugi podmreži, zaradi česar se poveča skupni magnetni moment. Pri povečanju substitucije Zn^{2+} zasede tudi druga mesta v prvi podmreži, zaradi česar se magnetni moment in s tem nasičena magnetizacija zmanjšata. Na sliki 4 vidimo tudi, da se magnetizacija vzorcev z x = 0.1 in 1.0 nekoliko razlikuje glede na način priprave. Sama priprava tudi vpliva na preferenčno zasedbo mest Fe^{3+}, zaradi česar se spremeni magnetizacija.



Slika 4: Nasičena magnetizacija (Ms) M-heksaferitov sestave $BaFe_{12-2x}A_xSn_xO_{19}$, A = Co, Ni, Zn, v odvisnosti od sestave (x) in načina priprave: sinteza v trdnem (SSR), visokoenergetsko mletje (HEM) in preobarjanje (CC).



Slika 5: Koercitivnost (Hc) M-heksaferitov sestave $BaFe_{12-2x}A_xSn_xO_{19}$, A = Co, Ni, Zn, v odvisnosti od sestave (x) in načina priprave: sinteza v trdnem (SSR), visokoenergetsko mletje (HEM) in preobarjanje (CC).

Na sliki 5 je prikazana odvisnost koercitivnosti (Hc) od sestave (x) in načina priprave (reakcija v trdnem - SSR, visokoenergetsko mletje - HEM in preobarjanje - CC). Hc po predvidevanjih pada z x. Hc je odvisna od magne-

tokristalne in anizotropije zaradi oblike zrn. Magnetnokristalna anizotropija se zmanjša s stopnjo substitucije in iz enoosne preide v planarno. Poleg tega je Hc precej večja za vzorce pripravljene s HEM in CC kot za vzorce pripravljene s SSR. Razlog je različna mikrostruktura različno pripravljenih vzorcev, kar vpliva na anizotropijo oblike zrn. Delci vzorcev SSR so precej večji (z razmerjem premer/debelina 10/2 μm) od delcev HEM (z razmerjem premer/debelina 2/0.2 μm) in delcev CC (z razmerjem premer/debelina 1/0.05 μm). Iz primera je razvidno, da se lahko s pravo izbiro sestave in načina priprave pripravimo material z željenimi lastnostmi.

Drugi del raziskav se nanaša na U-heksaferite s kemijsko formulo $Ba_4A_2Fe_{36}O_{60}$, A = Co, Ni, Zn, ali krajše A_2U. U-heksaferiti kristalizirajo v romboedrični prostorski skupini *R3m*. Dva bloka M in en blok Y se ponavljata v smeri c. V literaturi je le nekaj podatkov o lastnostih monokristalov Co_2U in Zn_2U /29, 30/ ter predvidoma enofaznem prahu in vlaknih /31, 32/. Zaradi kompleksne kristalne strukture (Glej poglavje 3.3.) heksaferitov, ki omogoča tvorbo različnih politipov je izredno težko kontrolirati vse pogoje priprave, ki vplivajo na nastanek določenih politipov. Z modificirano sintezo v trdnem nam je uspelo pripraviti enofazne polikristalinične heksaferite sestav A_2U. /27, 33/ Z običajno sintezo v trdnem je temperatura nastanka U-heksaferita previsoka, da bi lahko pripravili enofazen vzorec, ker le-ta že prične razpadat. Če reakcijsko mešanico aktiviramo, je temperatura nastanka U-heksaferita nižja in lahko pripravimo enofazen prah. Reakcijsko mešanico smo aktivirali na dva načina: z visokoenergetskim mletjem (HEM) in s topotaktično reakcijo (TR). Vzrok aktivacije pri metodi HEM je manjša velikost delcev ter nastanek notranjih napetosti in defektov. Pri TR pa izhajamo iz predreagiranih M in A_2Y sestav, ki jih stisnemo v magnetnem polju. Na ta način preskočimo nastanek prekurzorja $BaFe_2O_4$, katerega prisotnost otežuje nadaljno reakcijo, ter omogočimo usmerjeno (topotaktično) rast kristalov U-heksaferita. V tabeli 3 so zbrane osnovne lastnosti vzorcev A_2U. Tc je podobno kot v primeru ostalih tipov heksaferitov /7/ največja za Ni_2U in najmanjša za Zn_2U. Obratno pa je Ms največja za Zn_2U in najmanjša za Ni_2U. V območju do 18 GHz FMR za Zn_2U še nismo opazili, zaradi česar je μ še nizka. Na osnovi tega lahko sklepamo, da je ta material je potencialno uporaben v mm-območju.

Tabela 3: Lastnosti A_2U, pripravljenih z modificirano reakcijo v trdnem

sestava	priprava	Tc (°C)	Ms (emu/g)	lastnosti pri 0.5-18GHz
Co_2U	HEM	434 ± 4	50 ± 1	/
Ni_2U	HEM	454 ± 2	46 ± 0.4	/
Zn_2U	HEM, TR	404 ± 4	55 ± 3	FMR >> 18GHz ε = 5.5-4.5 μ = 2-1

7. Zaključek

Za civilne namene se izkorišča mikrovalovno frekvenčno območje, večje frekvence v območju milimetrskih valov pa so običajno namenjene vojaški uporabi. Zaradi vse večjega izkoriščanja elektromagnetnega valovanja za brezžično telekomunikacijo bo v prihodnosti frekvenčno območje za civilno uporabo potrebno razširiti tudi v milimetrski pas, kar pomeni potrebo po novih materialih, napravah in merilnih sistemih. Čeprav so garneti zaradi daleč najmanjših magnetnih izgub oz. spinelni feriti zaradi najnižje cene najboljša izbira do približno 24 GHz, pa so pri večji frekvenci najboljša možna izbira heksaferiti. Njihove tri najpomembnejše prednosti so neprimerljivo večja anizotropija, zadovoljiva permeabilnost in velika specifična upornost. Heksaferite tako lahko smatramo kot material prihodnosti. Ugotovili smo, da so za ta namen poleg M-tipa zanimivi tudi U-heksaferiti.

Zahvala

Zahvaljujem se prof. Mihi Drofeniku za podporo in mag. Vladimirju Boštjanu Bregarju za koristne nasvete pri pripravi tega prispevka.

Literatura

/1./ J. Nicholas, v Ferromagnetic Materials, A handbook on the properties of magnetically ordered substances, vol. 2, Ed. E. P. Wohlfarth, North-Holland Publishing Company, Amsterdam, 1980, poglavje 4

/2./ M. Pardavi-Horvath, Microwave Applications of Soft Ferrites, J. Magn. Magn. Mater., **215-216**, (2000), 171-183

/3./ G. P. Rodrigue, A Generation of Microwave Ferrite Devices, Proc. IEEE, **76** (2), (1988), 121-137

/4./ P. Singh, V. K. Babbar, A. Razdan, R. K. Puri, T. C. Goel, Complex Permittivity, Permeability, and X-band Microwave Absorption of CaCoTi Ferrite Composites, J. Appl. Phys., **87** (9), (2000), 4362-4366

/5./ M. R. Anatharaman, S. Sindhu, S. Jagatheesan, K. A. Malini, P. Kurian, Dielectric Properties of Robber Ferrite Composites Containing Mixed Ferrites, J. Phys. D: Appl. Phys., **32**, (1999), 1801-1810

/6./ M. Sparks, v Ferromagnetic Relaxation Theory, Mcgraw-Hill Inc., New York, 1964, str. 20

/7./ J. Smit, H. P. J. Wijn, v Ferrites, Philips' Technical Library, Eidenhoven, 1959, poglavja VII, IX, XIII, XIV, XV

/8./ K. Ishino, Y. Narumiya, Development of Magnetic Ferrites: Control and Application of Losses, Ceram. Bull., **66** (10), 1987, 1469-1474

/9./ G. H. Jonker, H. P. J. Wijn, P. B. Braun, »Ferroxplana, hexagonale ferromagnetische Eisenoxydverbindungen fuer sehr hohe Frequenzen«, Philips' Technische Rundschau, **18** (9), (1956/57), 249-276

/10./ Valenzuela, v Magnetic Ceramics, Cambridge University Press, Cambridge, 1994, poglavje 2.3

/11./ J. A. Kohn, D. W. Eckart, C. F. Cook, Jr., "Crystallography of the Hexagonal Ferrites", Science, **172** (3983), (1971), 519-525

/12./ P. Wartewig, M. K. Krause, P. Esquinazi, S. Roesler, R. Sonntag, Magnetic Properties of Zn-and Ti-substituted Barium Hexaferitte, J. Magn. Magn. Mater., **192**, (1999), 83-99

/13./ D. H. Han, Z. Yang, H. X. Zeng, X. Z. Zhou, A. H. Morrish, Cation Site Preference and Magnetic Properties of Co-Sn-Substi-

tuted Ba Ferrite Particles, J. Magn. Magn. Mater., **137**, (1994), 191-196

/14./ D. Lisjak, M. Drofenik, Synthesis and characterization of A-Sn-substituted (A = Co, Ni, Zn) BaM-hexaferrite powders and ceramics, J. Eur. Ceram, Soc., to be published

/15./ C. Suerig, K. A. Hempel, Magnetic anisotropy of chemically co-precipitated Zn_2W ferrite, IEEE Trans. Magn., **30** (2), (1994), 997-999

/16./ E. Brando, H. Vincet, O. Dubrinfaut, A. Fourrier-Lamer, R. Lebourgeois, Microwave Electromagnetic Characteristics of New Substituted M-Hexaferrites $BaFe_{12-2x}A_xMe_xO_{19}$ (A = Ru, Ir; Me = Co, Zn), J. Phys. IV France, **7**, (1997), C1-421 - C1-422

/17./ P. Lubitz, New Substitutions in Hexagonal Ferrites to Reduce Anisotropy without Using Co, J. Appl. Phys., **87** (9), (2000), 4978-4980

/18./ D. Autissier, A. Podembski, C. Jacquiod, Microwave Properties of M and Z type Hexaferrites, J. Phys. IV France, **7**, (1997), C1-409 - C1-412

/19./ H. Vincent, E. Brando, B. Sugg, Cationic Distribution in Relation to the Magnetic Properties of New M-Hexaferrites with Planar Magnetic Anisotropy $BaFe_{12-2x}Ir_xMe_xO_{19}$ (Me = Zn, Co, x » 0.85 and x » 0.50), J. Solid State Chem., **120**, 1995, 17-22

/20./ T. Inui, N. Ogasawara, Grain-Size Effect on Microwave Ferrite Magnetic Properties, IEEE Trans. Magn., **MAG-13** (6), 1977, 1729-1744

/21./ H. Takabayashi, Y. Kato, T. Kagotani, D. Book, S. Sugimoto, M. Homma, H. Ota, Y. Houjou, Effect of Crystallographic Alignement on the Microwave Absorption Properties of $BaFe_{12-x}(Ti_{0.5}Mn_{0.5})O_{19}$, Ferrites, Proceedings of The Eight International Conference on Ferrites (ICF 8), Kyoto and Tokyo, Japan 2000, 985-987

/22./ H. Marusawa, C. Kato, T. Konoike, K. Tomono, Magnetic Properties of Y-Type Hexagonal Ferrites at Ultarhigh Frequency, Ferrites, Proceedings of The Eight International Conference on Ferrites (ICF 8), Kyoto and Tokyo, Japan 2000, 942-944

/23./ H. Zhang, Z. Liu, X. Yao, L. Zhang, M. Wu, Dielectric and magnetic properties of ZnCo-substituted X hexaferrites prepared by citrate sol-gel process, Mat. Res. Bull., **38**, (2003), 363-372

/24./ S. Ram, Observation of enhanced dielectric permittivity in Bi^3 doped $BaFe_{12}O_{19}$ ferrite, J. Magn. Magn., Mater., **80**, (1989), 241-245

/25./ H. Zhang, J. Zhou, Z. Yue, P. Wu, Z. Gui, L. Li, Synthesis of Co_2Z hexagonal ferrite with planar structure by gel self-propagating method, Mater. Lett., **43**, (2000), 62-65

/26./ C. Suerig, K. A. Hempel, D. Bonnenberg, Hexaferrite particles prepared by sol-gel technique, IEEE Trans. Magn., **30** (6), (1994), 4092-4094

/27./ D. Lisjak, M. Drofenik, The synthesis and characterization of Zn_2U ($Ba_4Zn_2Fe_{36}O_{60}$) hexaferrite powder, J. Appl. Phys., **93**, (2003), 8011-8013

/28./ Gordon, R. L. Harvey, R. A. Braden, Preparation and magnetic properties of some hexagonal magnetic oxides, J. Am. Ceram. Soc., **45** (6), (1962), 297-301

/29./ J. Kerecman, A. Tauber, T. R. AuCoin, R. O. Savage, Magnetic properties of $Ba_4Zn_2Fe_{36}O_{60}$ single crystals, J. Appl. Phys., **39** (2), (1968), 726-727

/30./ J. Kerecman, T. R. AuCoin, W. P. Dattilo, Ferromagnetic Resonance in $Ba_4Zn_2Fe_{36}O_{60}$ (ZnU) and Mn-substituted ZnU single crystals, J. Appl. Phys., **40** (3), (1969), 1416-1417

/31./ R. C. Pullar, A. K. Bhattacharya, The Synthesis and Characterization of Co_2X ($Ba_2Co_2Fe_{28}O_{46}$) and Co_2U ($Ba_4Co_2Fe_{36}O_{60}$) ferrite fibres, manufactured from a sol-gel process, J. Mater. Sci., **36**, (2001), 4805-4812

/32./ G. Xiong, M. Xu, Z. Mai, Magnetic properties of $Ba_4Co_2Fe_{36}O_{60}$ nanocrystals prepared through a sol-gel method, Solid State Com., **118**, (2001), 53-58

/33./ D. Lisjak, M. Drofenik, The thermal stability range and magnetic properties of U-type hexaferrites, J. Magn. Magn. Mater., to be published

Dr. Darja Lisjak
Institut Jožef Stefan
Odsek za sodobne materiale
Jamova 39, 1000 Ljubljana
Tel.: +386 1 4773 4872
Faks: +386 1 4773 4875
e-mail: darja.lisjak@ijs.si

Prispelo (Arrived): 06.06.2003 *Sprejeto (Accepted): 26.08.2003*

IZKLOP VAROVALKE PRI PODALJŠANI TALILNI FAZI

Martin Bizjak, Mitja Koprivšek, Franc Pikl

ETI d. d. Izlake

Ključne besede: talilni vložek, talilni element, izklopna karakteristika, napetostna razlika, talilna faza, talilna napetost, prekinitveno mesto talilnega elementa

Izvleček: Pri preskusu izklopa varovalke z naznačeno vrednostjo 63 A s tokom 100 A je dopustno trajanje izklopa do 1 ure. V tem času potekata zaporedoma faza segrevanja in talilna faza. Obe se razlikujeta po značilnem časovnem poteku napetosti med priključki talilnega elementa: v prvi fazi narašča napetost počasi in sorazmerno segrevanju talilnega elementa, v drugi pa precej hitreje, ker se začne proces raztapljanja zožitev na segmentu z nanosom spajke. V talilni fazi sproščena toplote znatno segreje cel talilni vložek. Opaženo je bilo, da se lahko včasih talilna faza podaljša tudi na dvojno vrednost, sproščena toplota pa je bila do petkrat večja. Pri pojavu *podaljšane* talilne faze je segretek površine talilnega vložka presegel temperaturo, ki jo zdrži sam in deli v njegovi bližini. Napetost prekinitve toka, dobljena v *normalni* talilni fazi, je bila okrog 280 mV, v *podaljšani* talilni fazi pa okrog 1100 mV. Ta vrednost je bila dosežena sočasno s stalitvijo prekinitvenih mest na pospajkanem segmentu talilnega elementa in je odvisna od tališča zlitine na teh mestih. Napetost prekinitve toka kaže na kemijsko sestavo taline na prekinitvenem mestu, iz česar je bilo mogoče sklepati na verjeten potek podaljšanja talilne faze. Presoja ugotovitev je bila opravljena tudi s pomočjo rezultatov, ki so bili pridobljeni z drugimi preiskovalnimi metodami.

Breaking of Fuse Element by Occasionally Prolonged Melting Phase

Key words: fuse link, fuse element, breaking characteristics, voltage drop, M-effect, melting voltage, rupture of fuse element

Abstract: In the case of fault in power circuit protected by melting fuse the overcurrent should be interrupted by blow of fuse link in the predetermined breaking time. Depending on the magnitude of overcurrent the breaking time can be either as long as up to one hour or of the order of few milliseconds. The relationship breaking time - overcurrent is defined by the breaking characteristics of fuse link of given type and rated current, which follows the rule of inverse proportionality. It should comply to the relevant standards. The breaking process at moderate overcurrent values, which lasts up to approximately 1000 s is dealt with in this paper.

Samples of fuse links of rated current 63 A were tested in laboratory conditions in order to investigate breaking phenomena at constant test current 100 A. Although the test samples passed the requirements of relevant standard a serious overheating was indicated on some of them during break, which was considered hazardous for the device comprising such fuse links due to thermal damage of plastic parts in their vicinity. The majority of tested fuse links exhibited quite moderate temperature rise during break test.

Due to the risk of thermal damage even at the proper breaking function of fuse link the phenomena of the excessive overheating was investigated in particular. At first the statistics of measured results obtained on tested samples was accomplished. Although considerably high scattering of breaking time was indicated, no correlation between the duration of breaking process and the excessive temperature rise was found. In order to find out physical background of these phenomena the voltage drop on the fuse link terminals at the constant test current (of 100 A) was recorded by time for each of test samples. Regarding the obtained results the voltage drop increased at the beginning as semi exponential function according to the physical rules determining the heating of the electric conductor, which was in this case the Cu strip of fuse element. At the end of this phase, which is called "heating-up period", a significantly steeper increase of voltage drop was observed, which indicates changes of material structure on particular interrupting sites, notches, of fuse element, therefore it is called "melting period". The results of the voltage drop obtained on various test samples of the same type, recorded at the moment of break, were concentrated around two significantly different values. As the voltage drop achieved one of these values by levelling it could be attributed to "melting" voltage drop of notches. In the attempt to extract the voltage drop on notches from the voltage drop between fuse terminals, which was the sum of the rest of fuse link drop in series, the best fit approximation of voltage drop in heating-up period was done in order to extrapolate the contribution of the remainder in series and to subtract it from the measured voltage drop in melting period. By this way an improved course of voltage drop on interrupting sites of fuse element was attained. From these results the melting periods of shorter duration and lower level voltage could be clearly distinguished from the longer ones which have significantly higher level voltage. As the course of voltage drop in both cases followed a similar trace, the physical process in melting phase was probably the same regardless to its duration. So the break event can be principally divided into the interruption by *regular* and *prolonged* melting phase. Due to significantly higher value of Joule integral for fuse element in *prolonged* melting phase compared to *regular* one the excessive temperature rise of fuse link can be explained by higher power dissipation at melting of fuse element.

In order to reveal possible phenomena which probably led to the *prolonged* melting period, an examinations of the state of fuse elements after break test were conducted. Each tested fuse link was carefully opened and its blown fuse element withdrawn out of ceramic body and prepared for optical and SE microscopy. Particular attention was pay to the interrupting sites on the fuse element strip. These sites normally appear close to the strip of solder cover, which is placed near a row of notches in order to initiate the interruption of conducting path by dissolving the base strip in melted solder. The major difference between samples, accomplishing break of overcurrent by *regular* and *prolonged* melting phase, was reflected in the area of base Cu strip subjected to the dissolution process. Observations by optical stereomicroscopy clearly showed that at the *regular* type only the row of notches adjacent to the solder strip was exposed to dissolution in melted solder, while at the *prolonged* type the area of dissolution covered significantly larger area along the perforation of fuse element.

The metallographic structure of material on the interrupting sites was characteristic for the CuSn alloy, which was formed in Cu fuse element strip by alloying primarily with Sn of solder cover. Various CuSn metallurgical phases were found by optical metallography, the chemical composition of each particular one was determined by microprobe EDX analyses. With the help of CuSn binary diagram of states the most probable evolution of interruption process during break of overcurrent was assessed. On the other hand it is well known from the physics of electric contacts that the voltage drop on the

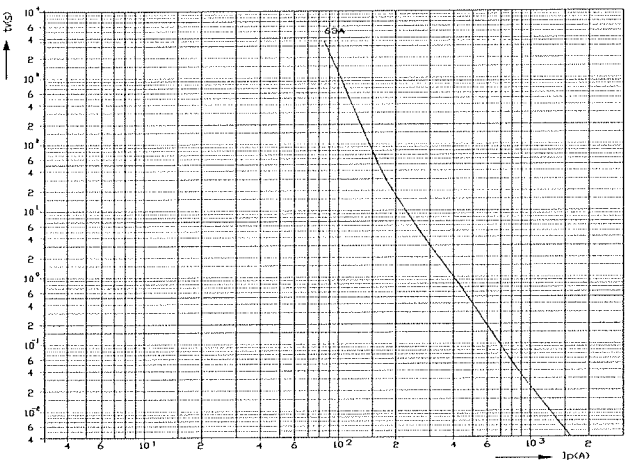
site of current constriction is proportional to the temperature rise on this site. This quantity was also applied as the qualitative indication of temperature on the interrupting sites at moment of break.

In the conclusion, the *regular* melting phase took place at much lower temperature of melting the interruption sites due to the formation of the CuSn alloy with higher content of Sn. At the *prolonged* melting phase the CuSn alloy with higher content of Cu was formed through the dissolution of Cu in solder, which led to higher melting temperature of interrupting sites. The decisive process which led toward the *prolonged* melting phase were proper conditions which enable spreading of liquid solder along the fuse element during interrupting process. If the solder melt remained confined in the narrow area of notches it resulted in the formation of low temperature melting alloy. But when due to certain reasons melted solder was able to spread along the larger area of fuse element, a solid solution with higher melting temperature was formed with regard to the reduced content of Sn in CuSn alloy.

The problem of the detrimental temperature rise of fuse link at break is related to the conditions of the dissolution of Cu fuse strip in solder. In order to avoid the risk of excessive heating of fuse link at break, the propagation of liquid solder over the surface of fuse element strip should be under control. This can be presumably performed by the proper design of fuse element and by select conditions of its manufacturing.

1 Uvod

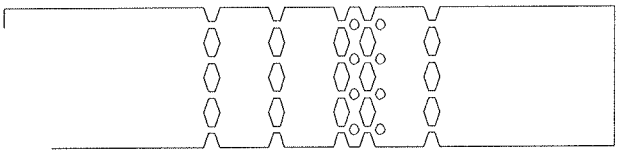
Talilna varovalka varuje tokokrog pred nadtoka, to je toka, ki je večji od nazivnega toka varovalke. Varovalka ima vgrajen talilni vložek s talilnim elementom, ki pri nadtoku prekine tok v času, ki je odvisen od njegove vrednosti. Izklopni čas, ki podaja hitrost reagiranja varovalke pri nadtoku dane vrednosti, je določen z izklopno karakteristiko, ki jo predpisuje standard IEC 60269 /1/. Izklopna karakteristika je funkcijska zveza med izklopnim časom in nadtokom, ki je prikazana grafično je na Sl. 1. Iz nje razberemo, da je izklopni čas varovalke pri večjem nadtoku krajši. Izklopni časi za nadtoke, ki so malo večji od nazivnega toka, so reda velikosti 1000 s, za nadtoke, ki so za faktor 10 ali več večji od te vrednosti, pa so ti časi reda velikosti 10^{-3} s.



Sl. 1: Izklopna karakteristika varovalke D02 gL 63 A

Vsaka varovalka podanega tipa ima za dan nazivni tok svojo standardizirano izklopno karakteristiko. Tej je prilagojena dimenzija in oblika talilnega elementa /2/. Talilni element je funkcionalni del talilne varovalke. Izdelan je iz tankega bakrenega (ali srebrnega) traku, ki je vstavljen v keramično ohišje s priključki in obdan s kremenčevim peskom. Trak je perforiran v obliki segmentov z zoženimi deli (Sl. 2), skozi katere teče električni tok in jih segreva. Na zožitvah enega segmenta je debel nanos spajke, ki je pri nadtoku staljena in raztaplja trak talilnega elementa na prekinitvenih mestih, tako da v določenem času prekine zožitve med segmentoma in s tem tudi električni tok. Pojav je poznan kot pregoretje varovalke.

Proces izklopa toka s talilnim elementom poteka v splošnem po sledečem zaporedju: segrevanje talilnega elementa na prekinitvenih mestih oz. zožitvah perforacije (notches) od delovne temperature do tališča (faza segrevanja) → taljenje prekinitvenih mest (talilna faza) → prekinitev kovinske zveze in nastanek izklopnega električnega obloka → gašenje obloka in izklop toka (obločna faza). Pri manjših nadtokih in dolgih izklopnih časih reda velikosti 100 s do 1000 s prevladuje izklop s predhodno fazo segrevanja, ki ji sledi prekinitev toka v talilni fazi. Pri velikih nadtokih reda velikosti kratkostičnih tokov pa sta fazi segrevanja in taljenja zelo kratki in jo obravnavamo kot predobločno fazo, ki ji sledi eksplozijska prekinitev zoženih mest, kjer se vzpostavi električni oblok s trajanjem reda velikosti 1 ms, dokler ne ugasne.



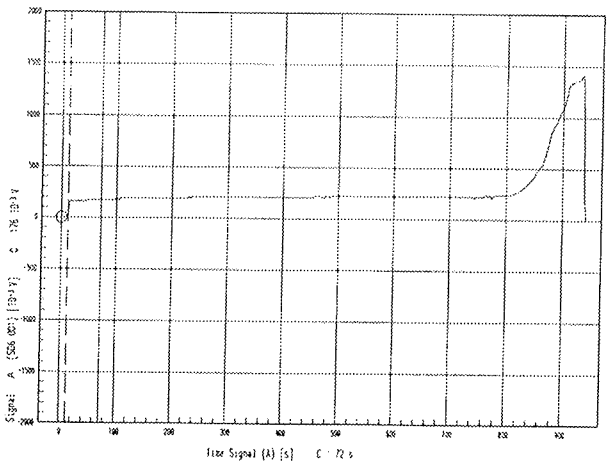
Sl. 2: Talilni element varovalke D02 gL 63 A

Med izklopom se v talilnem elementu razvija toplota, ki jo absorbira kremenčev pesek, od tu pa dalje v keramično telo talilnega vložka ter naprej v ohišje varovalke in v okolico. Pri kratkih izklopnih časih poteka segrevanje skoraj adiabatno, pri dolgih pa je proces kvazistacionaren, kjer se toplota prevaja od talilnega elementa skozi kremenčev pesek na telo talilnega vložka, z njegove površine pa v okolico. Keramično telo vložka se lahko včasih segreje celo do rdečega žara. Zaradi vse pogostejše uporabe plastičnih materialov za ohišja naprav z varovalkami, ki nadomeščajo toplotno odpornejšo keramiko, je priporočljivo privzeti mere, s katerimi omejimo segretek talilnega vložka, ker sicer lahko poškoduje okoliške dele iz plastike.

Pri izklopu s prevladujočo talilno fazo plast posebne spajke, ki je nanesena na izbrani predel perforacije talilnega elementa, omogoča prekinitev s stalitvijo kovinske zveze pri temperaturi, ki je manjša od tališča osnovnega materiala traku, iz katerega je izdelan talilni element. Pri nadtoku se na prekinitvenih mestih spajka segreje do njenega tališča, zaradi česar začne raztapljati material talilnega elementa na prekinitvenih mestih, dokler ne nastanejo tekoči mostiči, na katerih se tokokrog prekine. Proces raztapljanja določa trajanje izklopa.

2. Raziskava časovnega poteka izklopa varovalke v talilni fazi

Izklop varovalke s talilno fazo je bil raziskan v talilnem vložku tipa gL za nazivni tok 63 A pri konstantnem enosmernem preskusnem toku 100 A, kar znaša 1,6-kratnik nazivnega toka. V skladu z predpisano izklopno karakteristiko mora biti izklop toka opravljen manj kot 3600 s. Testni talilni vložek je bil za preskus vstavljen v standardno podnožje varovalke gL 63 A s priključki. Računalniško krmiljen laboratorijski vir preskusnega toka z napetostjo med odprtima sponkama 30 V /3/ je dovajal stabiliziran enosmerni tok na sponke testnega podnožja. Meritev in zapis časovnega poteka napetosti med priključki varovalčnega podnožja je opravljal merilni modul, ki je sestavni del računalniškega krmilnega modula laboratorijskega tokovnega vira. Vgrajena zaščita merilnega modula pred prenapetostjo odprtih sponk zagotavljajo hitri izklop toka pri prekinitvi talilnega elementa in drugih nenadnih spremembah napetosti na merilniku. Merilni rezultati so bili shranjeni v pomnilniku krmilnega računalnika za nadaljnjo analizo kot množica parov (napetostna razlika na priključkih, čas tokovne obremenitve) s širimo razreda za napetosti 16 mV in s širino razreda za čas 1 s. Mogoč je tudi neposreden grafični prikaz izmerjenega poteka izklopa, ki je kot primer izklopa s talilnim vložkom prikazan na Sl. 3.



Sl. 3: Časovni potek izklopa talilnega vložka s talilno fazo pri izklopnem preskusu

Kot je videti iz poteka izklopa na Sl. 3, napetost med priključki po vključitvi preskusnega toka narašča semi-eksponentno zaradi segrevanja talilnega elementa. Njegova električna upornost narašča zaradi temperaturnega koeficienta električne upornosti materiala, uporabljenega za talilni element, sorazmerno temperaturi. Na koncu faze segrevanja začne napetost strmo naraščati in bi pri prekinitvi toka skokovito narasla na napetost odprtih sponk tokovnega vira, če ne bi zaščitni modul tokovnega vira sprožil hitri izklop toka in zmanjšal napetost vira na ničlo.

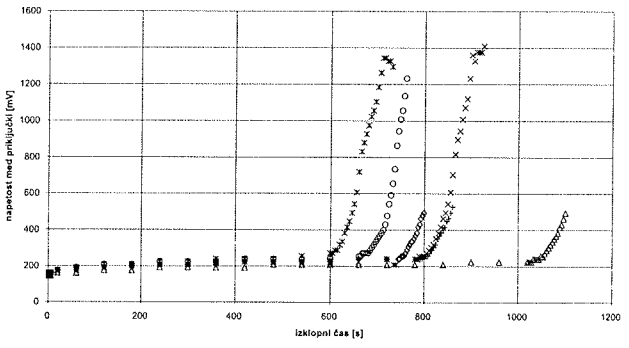
Pri raziskavi talilne faze je bil na vsakem preskusnem talilnem vložku izmerjen celoten časovni potek napetosti med priključki. Za nadaljnjo analizo je bil uporabljen fizikalni model segrevanja talilnega vložka in s tem je bila po metodi najboljšega približka opravljena ocena vplivnih parametrov in korelacije z časom trajanja talilne faze.

Električnim meritvam je sledila preiskava stanja talilnega elementa po opravljenem izklopu. Vsak talilni vložek je bil po preskusu previdno odprt, tako da smo iztresli iz keramičnega telesa najprej kremenčev pesek, potem pa izvlekli iz njega še talilni element, ki zaradi tega posega ni smel biti poškodovan. Na prekinitvenih mestih so bila na traku vtaljena zrna kremenčevega peska, katera so bila deloma odstranjena, da je bilo mogoče izvesti mikroskopsko preiskavo prekinitev na perforaciji traku zaradi izklopa toka.

Rezultati električnih meritev in mikroskopskih preiskav pregorelega talilnega elementa naj bi omogočili razpoznati procese v talilnem elementu med potekom izklopa in pojasniti vzrok za podaljšanje talilne faze.

3. Analiza merskih rezultatov izklopa s talilno fazo in diskusija

Rezultati preskusov izklopa varovalke so za nadaljnjo računalniško analizo prirejeni v obliki, kot jo prikazuje graf na Sl. 4. Na njem so zbrani rezultati časovnega poteka napetosti na priključkih varovalke med izklopom za več talilnih vložkov istega tipa. V trenutku $t = 0$ je bil vključen preskusni tok, izmerjena napetost približno 150 mV pa je razlika med priključki v začetnem hladnem stanju. S trajanjem tokovne obremenitve napetost zaradi segrevanja talilnega elementa v vložku počasi narašča, dokler po kakih 500 do 1000 s hitro naraste od kakih 300 mV preko 500 mV celo do 1400 mV. Tam napetost v hipu pade pod 100 mV zaradi delovanja zaščite merilnega modula, dejansko pa skoči na napetost vira 30 V, ker pride do prekinitve preskusnega toka oz. izklopa. Izklopni čas pri nastavljenem preskusnem toku je trajanje tokovne obremenitve od vključitve toka do prekinitve in je za preskušance, katerih rezultate kaže graf na Sl. 4, v razponu od 700 s do 1100 s, kar ustreza zahtevam glede na izklopno karaktertistiko.



Sl. 4: Prikaz rezultatov izklopa za preskusne talilne vložke, uporabljenih za nadaljnjo analizo

Začetni dolgotrajnejši časovni interval s počasnim naraščanjem napetosti med priključki $U(t)$ pripada *fazi segrevanja*,

kateri sledi precej krajša *talilna faza* s hitrim naraščanjem $U(t)$ preko 400 mV, ki se konča s prekinitvijo toka pri $t = t_{izk}$ pri neki največji izmerjeni napetosti $U(t_{izk}) = U_p$. Na prikazanem grafu lahko vidimo dva načina izklopa: *prvega* s kratkotrajno talilno fazo in majhno napetostjo prekinitve toka $U_p \approx 500$ mV, *drugega* pa z daljšo talilno fazo in precej večjo napetostjo prekinitve toka $U_p \approx 1400$ mV. V *prvem* načinu izklopa je talilna faza *normalna*, talilni vložek segreje do največ 300°C, kar pomeni nevarnosti za poškodbe sosednjih delov tokokroga, zato je tak časovni potek izklopa varovalnega elementa tokokroga *normalen*. Včasih se zgodi tudi *drugi* način izklopa s *podaljšano* talilno fazo. Z izračunom joulskega integrala izklopa je bilo ocenjeno, da se v podaljšani talilni fazi sprosti vsaj petkrat več toplote, kot v normalni. Ker se toliko toplote razvije v času, krajšem od 0,2 τ, kjer je τ termična časovna konstanta varovalke se keramično telo vložka na površini segreje lahko za več kot 600°C (temperatura rdečega žara!). Zato naj bi analiza rezultatov preskusov predvsem omogočila dognati, kaj je odločujoče za nastop podaljšane talilne faze in v skladu s tem, če ima vpliv na trajanje talilne faze tudi predhodna faza segrevanja.

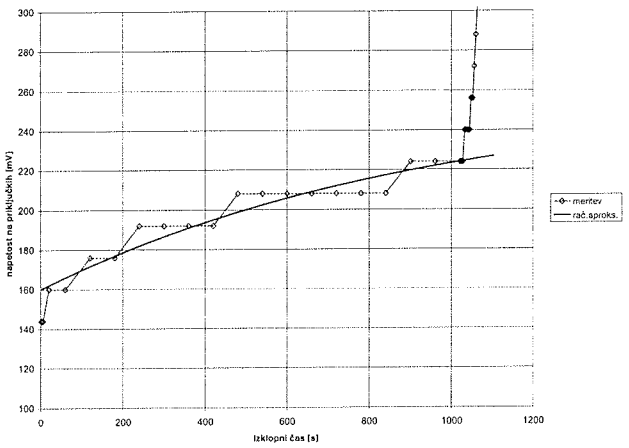
3.1. Analiza poteka faze segrevanja

Potek $U(t)$ v fazi segrevanja je značilen za procese segrevanja vodnika s tokom. S fizikalnim modelom vodnika enostavne geometrije dobimo aproksimativen časovni potek $U_{apr}(t)$ za fazo segrevanja v obliki:

$$U_{apr}(t) = U_0 + B\left[1 - \exp\left(-\frac{t}{\tau}\right)\right] \qquad (1)$$

s katerim po metodi najboljšega približka iz merskih rezultatov $U(t)$ določimo koeficiente U_0, B in τ kot parametre, ki so značilni za vsak posamezen preskušanec. V začetku tokovne obremenitve ($t < 10$ s) se $U_{apr}(t)$ ne prilega najbolje $U(t)$, ker v aproksimativnem modelu obstaja le termična časovna konstanta celotne varovalke, pojavi hitrejšega segrevanja delov perforacije pa so zanemarjeni. Zato se aproksimacija prilega meritvi šele po času, primerljivem s časovno konstanto τ varovalke. Prileganje je prikazano na grafu Sl. 5, iz katerega lahko ocenimo tudi natančnost merskih rezultatov.

Natančnost ocene parametrov U_0, B in τ očitno ni odvisna le od natančnosti aproksimacijskega modela, ampak v fazi segrevanja tudi od širine razreda, s katero pomnilnik merilnega modula zajema merske rezultate. Ta je v redu velikosti spremembe merjene količine, zato je ocena koeficientov aproksimacijske funkcije premalo natančna, da bi lahko iz njihove deviacije za vsak posamezen preskušanec lahko sklepali na vplive, ki vodijo v podaljšano talilno fazo. Dovolj zanesljiva je le korelacija med koeficientom U_0 in časom trajanja faze segrevanja, oz. časom prehoda v talilno fazo. Hitrost segrevanja talilnega elementa je odvisna od njegove začetne upornosti, od tega pa je odvisen tudi čas, ko se prekinitvena mesta na njem segrejejo do tališča spajke. Iz rezultatov grafa Sl. 4 pa je mogoče sklepati, da trajanje



Sl. 5: Funkcija najboljšega prileganja za rezultate faze segrevanja

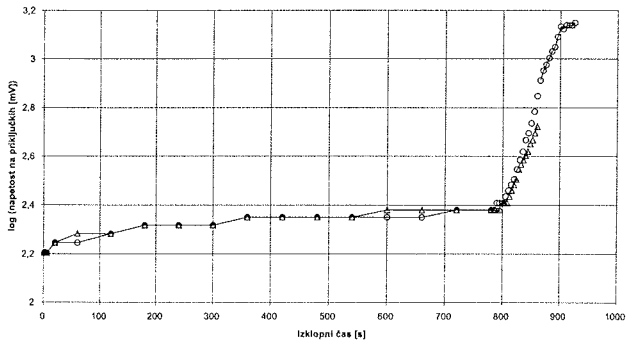
faze segrevanja ni v povezavi z nastopom podaljšane talilne faze.

Numerična simulacija segrevanja talilnega elementa je pokazala, da upornost priključnih elementov med dovodnim vodnikom in samim talilnim elementom bistveno vpliva na merilne rezultate in da je deviacija te vrednosti po posameznih preskušancih nezanemarljiva, zato ta lahko prekriva vpliv osnovnih parametrov talilnega elementa na potek izklopa v fazi segrevanja. Zato se vpliva faze segrevanja na nastop podaljšane talilne faze ne da ugotoviti dovolj zanesljivo.

3.2. Analiza poteka talilne faze

Časovni potek napetosti med priključki $U(t)$ preskušanega talilnega vložka se v talilni fazi ne da simulirati na osnovi enostavnega fizikalnega modela, kot v fazi segrevanja, ker v njej potekajo kompleksni procesi topljenja perforiranih delov talilnega elementa v spajki in spremembe kemijske sestave teh segmentov. Potek teh procesov lahko vodi v *normalno* ali v *podaljšano* talilno fazo. Nastanek ene ali druge vrste talilne faze ne korelira niti z začetno upornostjo talilnega elementa, niti s trajanjem faze, kot sledi kvalitativno tudi iz medsebojne primerjave grafov $U(t)$ za posamezne talilne vložke na Sl. 4. Na grafu Sl. 6 lahko primerjamo rezultate dveh preskušancev, katerih izmerjeni potek $U(t)$ je skoraj identičen v času trajanja tokovne obremenitve skoraj do izklopa, vendar je eden izklopil z *normalno* in drugi s *podaljšano* talilno fazo.

Zaradi raziskave poteka procesa raztapljanja perforiranega segmenta talilnega elementa v spajki bi bil v talilni fazi relevanten podatek napetost na perforiranem segmentu s plastjo staljene spajke. Ker neposredni dostop z merilnimi sondami do tega mesta ni mogoč, je bil ta podatek pridobljen iz meritev posredno: če od izmerjene $U(t)$ odštejemo v skladu z (2) aproksimacijsko funkcijo faze segrevanja $U_{apr}(t)$ po vsem časovnem intervalu tokovne obremenitve, s tem računsko eliminiramo prispevek ostalih delov talilnega elementa in dobimo z dosegljivo natančnosti potek napetosti na talečih se segmentih, $U_M(t)$:



Sl. 6: Primer skoraj identičnega poteka izklopa za preskušanca z normalno in podaljšano talilno fazo

$$U_M(t) = U(t) - U_{apr}(t) \quad (2)$$

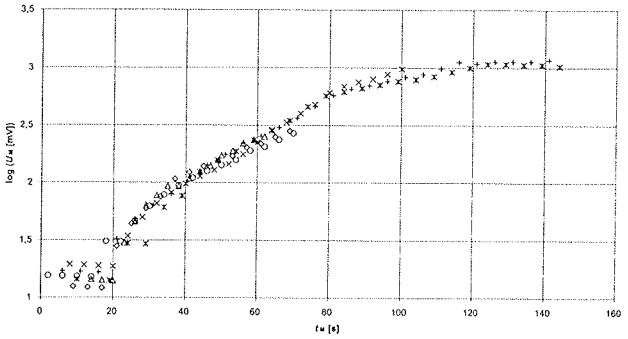
S prej utemeljenim privzetkom, da procesi *faze segrevanja* ne določajo poteka procesov *talilne faze*, lahko merske rezultate $U_M(t)$ za talilno fazo vsakega posameznega preskušanca premaknemo v isto časovno izhodišče in jih na tak način napravimo medsebojno primerljive. Tako je talilna faza opisana s potekom napetosti $U_M(t_M)$ za $t_M < 0$ od trenutka nastopa te faze pri $t_M = 0$, kjer je $U_M(t_M) = 0$. Tako obdelani merski rezultati za posamezne preskušance so skupaj grafično prikazani na Sl. 7, rezultati preskušancev z *normalno* talilno fazo so predstavljeni s simboli ◊, Δ in O, rezultati za *podaljšano* talilno fazo pa s simboli +, x in *.

Kot je razvidno iz grafa, se rezultati $U_M(t_M)$ za preskušance z istovrstno talilno fazo precej dobro skladajo med seboj. Rezultati za normalno talilno fazo so nekoliko manjši od rezultatov za podaljšano talilno fazo, vendar so tudi rezultati obeh vrst talilne faze kvalitativno toliko skladni med seboj, da lahko predpostavljamo enak potek fizikalnih pojavov v talilni fazi v večjem delu časovnega intervala, ki je skupen obema vrstama talilne faze. Šele malo pred izklopom z *normalno* talilno fazo se pojavi opaznejše razhajanje rezultatov, ki spelje dogajanje v *podaljšano* talilno fazo. Potek talilne faze torej lahko preusmerijo že sorazmerno majhni učinki, ki jih je težko nadzorovati.

3.3. Talilna napetost

Električna napetost med dvema točkama na vodniku s tokom, ki odvaja toploto le s prevajanjem po njem, je odvisna le od temperature na odseku med tema dvema točkama in od snovnih parametrov vodnika /4/. V fazi segrevanja je potek $U(t)$ odražal le spremembo temperature talilnega elementa, na potek $U_M(t_M)$ v talilni fazi pa dodatno vplivajo še snovne spremembe na prekinitvenem segmentu. Zato $U_M(t_M)$ ne moremo matematično podati z enostavno analitično funkcijo. Vendar ima v trenutku izklopa ($t_M = t_{M,izk}$) za posamezne preskušance $U_M(t_{M,izk})$ neko značilno vrednost, ki je odvisna od vrste talilne faze, za *normalno* 250 ÷ 280 mV in za *podaljšano* 1000 ÷ 1100 mV. Kvalitativno ima ta količina značilnosti spremembe agregatnega stanja. Za nekatere kovine je podana v [V] talilno napetost /4/, ki ustreza temperaturi tališča.

Vrednost 250 ÷ 280 mV oz. 1000 ÷ 1100 mV lahko privzamemo kot napetost stalitve perforiranih mest na talilnem elementu, ki prekine električni tok. Nobeno od teh vrednosti se kvantitativno ne da uskladiti s podatki v literaturi, npr. v /4/, ker navedene vrednosti veljajo za segrevanje stičnega mesta na električnih kontaktih, kjer dejanske razmere ustrezajo teoretičnim izhodiščem. V literaturi so podane predvsem vrednosti za čiste kovine in enostavne zlitine, za zlitine v splošnem pa so podatki nezanesljivi. Na grafu Sl. 7 lahko zlasti za časovni potek $U_M(t_M)$ na preskušancih s *podaljšano* talilno fazo tik pred izklopom, to je pri $t_M \rightarrow t_{M,izk}$, opazimo približevanje konstantni vrednosti, kar res ustreza poteku temperature pri spremembi agregatnega stanja, kot je npr. prehod trdno → tekoče. Pri *normalni* talilni fazi se na samem poteku $U_M(t_M)$ tega prehoda ne opazi, ker se prekinitveni segment pri izklopu na zelo omejenem območju. Iz prekinitvene napetosti $U_M(t_{M,izk})$ za obe vrsti talilne faze se da vsaj kvantitativno zanesljivo oceniti, da je temperatura prekinitvenega mesta v trenutku izklopa pri *normalni* talilni fazi precej manjša, kot pri *podaljšani*. V *podaljšani* talilni fazi je temperatura precej večja, taljenje zajame precejšen del talilnega elementa, kar nakazuje tudi potek $U_M(t_M)$ pred prekinitvijo toka.



Sl. 7: Potek talilne napetosti $U_M(t_M)$ za preskušance z normalno in podaljšano talilno fazo

Metalografska analiza, ki je bila opravljena na vseh preskušancih po opravljenem preskusu izklopa toka, je pokazala značilne razlike med predstavniki *normalne* in *podaljšane* talilne faze v končnem stanju talilnega elementa. Na talilnem elementu z *normalno* talilno fazo lahko razločimo nastanek zlitine med spajko in osnovnim materialom, ki je ostro omejena okoli prekinitvenih mest le na enem segmentu perforacije, ter vsebuje metalurške faze z *nižjim tališčem* /5/. Na talilnem elementu s *podaljšano* talilno fazo je nastala zlitina na prekinitvenih mestih preko več zaporednih segmentov talilnega elementa. Tališče metalurških faz te zlitine je veliko in je blizu tališča osnovnega materiala talilnega elementa /5/.

4. Sklep

Pri izklopu toka s talilno varovalko se v *podaljšani* talilni fazi izklopa sprosti do petkrat več toplote, kot v *normalni*. Večina sproščene toplote se porabi za segrevanje talilnega vložka, zato se ta segreje do temperature, ki lahko preseže

termično odpornost varovalke in materialov v njeni okolici, zato je treba odpraviti vzroke za pojav podaljšanja talilne faze. Trajanje talilne faze je odvisno od poteka raztapljanja talilnega elementa v spajki na prekinitvenih mestih. Če na teh mestih nastane zlitina z veliko vsebnostjo komponent iz spajke, ki ima majhno temperaturo tališča, je talilna faza krajša oz. *normalna*. Taljenje zajame ozko omejeno območje na prekinitvenih mestih enega segmenta. Talilna faza se podaljša, kadar se spajka razlije po širšem področju talilnega elementa, ne da bi se izvedla prekinitev. Koncentracija komponent iz spajke je v nastali zlitini z osnovnim materialom talilnega elementa manjša in temperatura tališča večja. To temperaturo doseže talilni element kasneje, zato se talilna faza podaljša.

Iz predhodnih rezultatov meritev električnih količin ni mogoče vnaprej razpoznati in izločiti talilnih vložkov, ki bodo izklopili s podaljšano talilno fazo. Trajanje talilne faze je odvisno od lege nanosa spajke na perforiranih segmentih in od količine nanesene spajke. Verjetnost, da bo talilna faza podaljšana, je odvisna od tehnološko dosegljivih toleranc pri nanašanju plasti spajke v proizvodnji varovalk.

Rezultati raziskave so podali izhodišče za konstrukcijo talilnega elementa varovalk iz serijske proizvodnje, za katere je verjetnost podaljšanja talilne faze zanemarljiva.

5. Literatura

/1/ IEC 60269-1 (1998-12), IEC 60269-3-1 (2001-06)

/2/ H. Johann: Elektrische Schmelzsicherungen fur Niederspannung, Springer-Verlag, Berlin Heidelberg New York, 1982

/3/ Navodilo za uporabo tokovnega generatorja v Laboratoriju ETI

/4/ P. G. Slade: Electrical Contacts, Marcel Dekker, 1999

/5/ C. J. Smithells: Metals Reference Book, Volume II, Butherworths, 1967

Dr. Martin Bizjak, univ. dipl. ing. fiz
Sv. Duh 275, 4220 Škofja Loka
E-pošta: bizjakm@email.si
tel.: +386 4 5131 521

Mitja Koprivšek, univ. dipl. ing. el
ETI d. d. Izlake, Obrezija 5, 1411 Izlake
E-pošta: mitja.koprivsek@eti.si
tel: +386 3 5657 450, fax: +386 3 5674 077

Franc Pikl, ing. el
ETI d. d. Izlake, Obrezija 5, 1411 Izlake
E-pošta: franc.pikl@eti.si
tel: +386 3 5657 460, fax: +368 3 5674 077

Prispelo (Arrived): 08.01.2003 *Sprejeto (Accepted): 26.08.2003*

Zavod TC SEMTO

Tehnološki center za sklope, elemente, materiale, tehnologije in opremo za elektrotehniko

Smo neprofitni zavod, ustanovljen v letu 2000 od zainteresiranih proizvodnih podjetij in raziskovalnih institucij v Sloveniji v skladu z zakonom in priporočili Ministrstva za znanost in tehnologijo z namenom, da se vzpostavi boljša povezava med industrijo ter znanostjo in podjetji in da se bolje izkoristi obstoječe znanje, oprema, priložnosti, ...

Znanje in izkušnje je treba čim večkrat uporabiti in prodati!

USTANOVITELJI:

INDUSTRIJA:

Iskra TELA, Iskra Feriti, Iskra SEM, Iskra Kondenzatorji, TEM Čatež, Varsi, Iskra Zaščite, Magneti, Kekon, Kolektor Idrija, RLS, Iskraemeco, Noviklik

UNIVERZE IN INŠTITUTI:

Institut Jožef Stefan, Fakulteta za elektrotehniko in Fakulteta za strojništvo Univerze v Ljubljani, Fakulteta za elektrotehniko, računalništvo in informatiko Univerze v Mariboru, Inštitut za kovinske materiale in tehnologije, Slovenski inštitut za kakovost

TC SEMTO je odprt še za nove interesente. Včlani se !!

Cilji Zavoda TC SEMTO:

- Glavna naloga TC SEMTO je zbirati in posredovati informacije o razpoložljivem in potrebnem znanju ter pospeševati skupne akcije.
- Poglobiti medsebojno poznanje in bolje izkoristiti obstoječe znanje in opremo ter ustvarjati novo znanje v vertikalnem in horizontalnem sodelovanju na področju elektromehanskih in elektronskih elementov (sklopi, izdelki, materiali, tehnologije in oprema)
- Izmenjati izkušnje in organizirati skupno izobraževanje
- Zagotoviti udeležencem konkurenčne prednosti pri skupnem in koordiniranem delu
- Poglobiti sodelovanje z vladnimi organizacijami in drugini tehnološkimi centri

Sodelovanje med udeleženci ni novo. Začelo se je pred desetletji, ko so bili nekateri udeleženci še v velikih sistemih. Sedaj so ponovno spoznali, da je pretok informacij o možnostih sodelovanja prešibek in ga žele ojačiti. Sodelovanje temelji na podjetniških odnosih v korist vseh sodelujočih.

Rezultati sodelovanja se zrcalijo v poslovnih rezultatih udeležencev.

Raziskovalci svoje rezultate tudi objavljajo v strokovnih publikacijah ter predstavljajo na strokovnih konferencah.

Večina razvojnikov je članov v profesionalnem društvu za mikroelektroniko, elektronske komponente in materiale - MIDEM in v njegovi publikaciji Informacije MIDEM in drugih strokovnih publikacijah (Materiali in tehnologije, ...) objavlja svoje prispevke.

Raziskovalna oprema

V TC SEMTO so se vključile raziskovalne organizacije in podjetja, ki imajo lastne raziskovalno-razvojne oddelke. Vsak od njih ima tudi veliko specifične raziskovalne opreme in specifično znanje in izkušnje. Vsi skupaj lahko zagotovijo potrebne kapacitete za realizacijo skoraj vseh raziskovalnih in razvojnih projektov s področja delovanja TC SEMTO. **Vsi člani so se zavezali dati za izvajanje projektov TC SEMTO na razpolago svoje laboratorije in opremo ter vključiti svoje kadre**. Raziskovalci in razvojniki podjetij se lahko za čas trajanja projektov redno ali dopolnilno zaposlijo v TC SEMTO in delajo na najprimernejši lokaciji. Zavod TC SEMTO zato ne odpira dodatnih raziskovalnih prostorov in opreme ampak skrbi, da je obstoječa draga raziskovalna oprema optimalno izkoriščena. Več informacij na strani člani TC in partnerji ter na spletnih straneh članov.

Način delovanja:

- Zavod TC SEMTO je evidentiran pri MŠZŠ kot raziskovalna organizacija (evid.:B) pod št. 1689
- TC Semto je registriral "Raziskovalno skupino za materiale ter elektronske komponente in tehnologije" (1689-001). V njej lahko pri skupnih projektih sodelujejo raziskovalci iz raznih podjetij in raziskovalnih institucij.
- TC SEMTO lahko nastopa kot izvajalec raziskovalnih ali razvojnih projektov. Raziskovalci in razvojniki inštitucij in podjetij se lahko za čas trajanja projektov redno ali dopolnilno zaposlijo v TC SEMTO. (okrog 20 vsako leto)
- Člani TC SEMTO izvajajo projekte v lastnih laboratorijih ali v mešanih skupinah v okviru TC SEMTO na najprimerneje opremljeni lokaciji. Vsi člani so se zavezali dati za izvajanje projektov TC SEMTO na razpo-

lago svoje laboratorije in opremo ter vključiti svoje kadre.
- Člani namenjajo v okviru svojih strateških načrtov precejšnja sredstva za razvoj in prijavljajo svoje projekte na razpise za subvencije MG in MŠZŠ.
- Nekateri člani imajo evidentirane lastne razvojne enote in lahko nastopajo kot izvajalci raziskav v subvencioniranih projektih.
- Delovanje nadzira skupščina in nadzorni odbor.
- Informiranje o potrebah in možnostih izmenjave znanja in izkušenj poteka na srečanjih, posvetih in v neposrednih stikih. Pobude za sodelovanje se lahko dajejo tudi preko te spletne strani.

Izmenjava informacij:

- Izdajamo *Obvestila S-xy*. Vsi člani jih prejmejo po elektronski pošti. Po Obvestilih tudi sklicujemo srečanja. Člane obveščamo o novih razpisih, ...
- Srečanja so vedno pri enem od ustanoviteljev, ko ta predstavi svojo dejavnost, potrebe in ponudbo znanja. Na srečanjih obravnavamo tudi teme skupnega interesa in izmenjamo mnenja.
- Kadar je potrebno, delujejo srečanja kot "skupščina TC SEMTO"
- Člani so pozvani, da dajo pobude in sodelovanje za delovanje TC SEMTO.

Kako se TC financira?

- Ustanovitelji so plačali manjši prispevek k stroškom začetnega delovanja.
- Ustanovitelji so dolžni prispevati sredstva za delovanje TC. O višini in načinu se sklepa na vsakoletni skupščini. V prvih letih so morali prejemniki odstopiti TC določen procent sredstev, ki so jih prejeli kot subvencijo od ministrstev za projekte, ki so bili na razpisih uvrščeni v subvencioniranje. V letu 2002 je bilo dogovorjeno, da namesto tega člani plačajo članarino.
- TC SEMTO kandidira tudi za subvencijo MG za tehnološke centre.

Kaj od članov pričakujemo?

- Da sledijo smerem razvoja svoje panoge in trga in naredijo strateške razvojne načrte na osnovi analize SWOT svojih programov ter predvidijo možne rešitve za doseganje konkurenčnih prednosti in iz okolice pritegnejo manjkajoča znanja.
- Da si vzamejo čas, da drugim predstavijo svoje možnosti in probleme.
- Da pripravijo dobre predloge projektov za domače in mednarodne razpise.
- Da predlagajo zanimive teme za skupno obravnavo (pobude in sodelovanje).
- Da se kadrovsko krepijo.
- Ustanovitelji TC SEMTA so opozorjeni, da bodo imeli od tega centra toliko, kolikor bodo pri njegovem delu aktivno sodelovali.
- TC lahko predlaga vrsto akcij in rešitev, ki pa za člane niso obvezne. Od njihovega zanimanja in sodelovanja je odvisna uspešnost TC SEMTA.
- Razni tehnološki centri (TC) so po namenu in organizaciji zelo različni. Imajo pa vrsto sorodnih problemov in v sodelovanju jih je možno rešiti, če so za posamezen TC prezahtevni (izobraževanje, ...). TC SEMTO sodeluje s tehnološkimi centri TECES, Maribor, ARI, Ljubljana, Iskra TECHNO, Vakuum-TC, ...

Igor Pompe
TC SEMTO
Stegne 25
1521 Ljubljana
telefon +386 01 5191281
FAX +386 01 5111295
semto@guest.arnes.si

Informacije MIDEM
Strokovna revija za mikroelektroniko, elektronske sestavine dele in materiale

NAVODILA AVTORJEM

Informacije MIDEM je znanstveno-strokovno-društvena publikacija Strokovnega društva za mikroelektroniko, elektronske sestavne dele in materiale - MIDEM. Revija objavlja prispevke s področja mikroelektronike, elektronskih sestavnih delov in materialov. Ob oddaji člankov morajo avtorji predlagati uredništvu razvrstitev dela v skladu s tipologijo za vodenje bibliografij v okviru sistema COBISS.
Znanstveni in strokovni prispevki bodo recenzirani.

Znanstveno-strokovni prispevki morajo biti pripravljeni na naslednji način:

1. Naslov dela, imena in priimki avtorjev brez titul, imena institucij in firm
2. Ključne besede in povzetek (največ 250 besed).
3. Naslov dela v angleščini.
4. Ključne besede v angleščini (Key words) in podaljšani povzetek (Extended Abstract) v angleščini, če je članek napisan v slovenščini
5. Uvod, glavni del, zaključek, zahvale, dodatki in literatura v skladu z IMRAD shemo (Introduction, Methods, Results And Discsussion).
6. Polna imena in priimki avtorjev s titulami, naslovi institucij in firm, v katerih so zaposleni ter tel./Fax/Email podatki.
7. Prispevki naj bodo oblikovani enostransko na A4 straneh v enem stolpcu z dvojnim razmikom, velikost črk namanj 12pt. Priporočena dolžina članka je 12-15 strani brez slik.

Ostali prispevki, kot so poljudni cčlanki, aplikacijski članki, novice iz stroke, vesti iz delovnih organizacij, inštitutov in fakultet, obvestila o akcijah društva MIDEM in njegovih članov ter drugi prispevki so dobrodošli.

Ostala splošna navodila

1. V članku je potrebno uporabljati SI sistem enot oz. v oklepaju navesti alternativne enote.
2. Risbe je potrebno izdelati ali iztiskati na belem papirju. Širina risb naj bo do 7.5 oz.15 cm. Vsaka risba, tabela ali fotografija naj ima številko in podnapis, ki označuje njeno vsebino. Risb, tabel in fotografij ni potrebno lepiti med tekst, ampak jih je potrebno ločeno priložiti članku. V tekstu je treba označiti mesto, kjer jih je potrebno vstaviti.
3. Delo je lahko napisano in bo objavljeno v slovenščini ali v angleščini.
4. Uredniški odbor ne bo sprejel strokovnih prispevkov, ki ne bodo poslani v dveh izvodih skupaj z elektronsko verzijo prispevka na disketi ali zgoščenki v formatih ASCII ali Word for Windows. Grafične datoteke naj bodo priložene ločeno in so lahko v formatu TIFF, EPS, JPEG, VMF ali GIF.
5. Avtorji so v celoti odgovorni za vsebino objavljenega sestavka.

Rokopisov ne vračamo. Rokopise pošljite na spodnji naslov.

Uredništvo Informacije MIDEM
MIDEM pri MIKROIKS
Stegne 11, 1521 Ljubljana, Slovenia
Email: Iztok.Sorli@guest.arnes.si
tel. (01) 5133 768, fax. (01) 5133 771

Informacije MIDEM
Journal of Microelectronics, Electronic Components and Materials

INSTRUCTIONS FOR AUTHORS

Informacije MIDEM is a scientific-professional-social publication of Professional Society for Microelectronics, Electronic Components and Materials - MIDEM. In the Journal, scientific and professional contributions are published covering the field of microelectronics, electronic components and materials.
Authors should suggest to the Editorial board the classification of their contribution such as : original scientific paper, review scientific paper, professional paper...
Scientific and professional papers are subject to review.

Each scientific contribution should include the following:

1. Title of the paper, authors' names, name of the institution/company.
2. Key Words (5-10 words) and Abstract (200-250 words), stating how the work advances state of the art in the field.
3. Introduction, main text, conclusion, acknowledgements, appendix and references following the IMRAD scheme (Introduction, Methods, Results And Discsussion).
4. Full authors' names, titles and complete company/institution address, including Tel./Fax/Email.
5. Manuscripts should be typed double-spaced on one side of A4 page format in font size 12pt. Recommended length of manuscript (figures not included) is 12-15 pages
6. Slovene authors writing in English language must submit title, key words and abstract also in Slovene language.
7. Authors writing in Slovene language must submit title, key words and extended abstract (500-700 words) also in English language.

Other types of contributions such as popular papers, application papers, scientific news, news from companies, institutes and universities, reports on actions of MIDEM Society and its members as well as other relevant contributions, of appropriate length , are also welcome.

General informations

1. Authors should use SI units and provide alternative units in parentheses wherever necessary.
2. Illustrations should be in black on white paper. Their width should be up to 7.5 or 15 cm. Each illustration, table or photograph should be numbered and with legend added. Illustrations, tables and photographs must not be included in the text but added separately. However, their position in the text should be clearly marked.
3. Contributions may be written and will be published in Slovene or English language.
4. Authors must send two hard copies of the complete contributon, together with all files on diskette or CD, in ASCII or Word for Windows format. Graphic files must be added separately and may be in TIFF, EPS, JPEG, VMF or GIF format.
5. Authors are fully responsible for the content of the paper.

Contributions are to be sent to the address below.

Uredništvo Informacije MIDEM
MIDEM pri MIKROIKS
Stegne 11, 1521 Ljubljana, Slovenia
Email: Iztok.Sorli@guest.arnes.si
tel.+386 1 5133 768, fax.+386 1 5133 771