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

INFORMACIJE MIDEM 1·2003

Strokovno društvo za mikroelektroniko elektronske sestavne dele in materiale

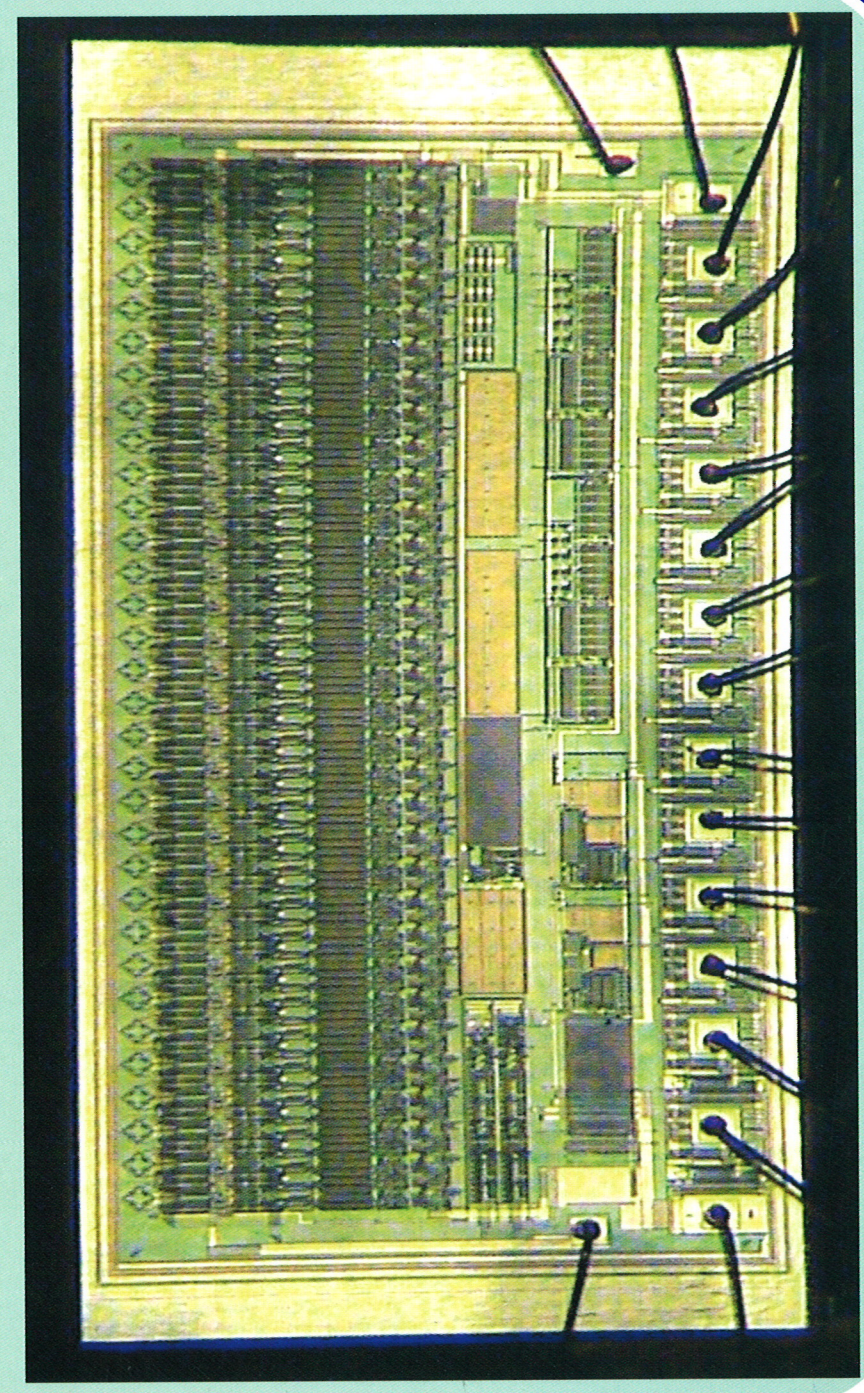
Strokovna revija za mikroelektroniko, elektronske sestavne dele in materiale
Journal of Microelectronics, Electronic Components and Materials

INFORMACIJE MIDEM, LETNIK 33, ŠT. 1(105), LJUBLJANA, marec 2003

UNIVERZA V LJUBLJANI

LMFE

FAKULTETA ZA ELEKTROTEHNIKO



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

INFORMACIJE MIDEM 1 ◦ 2003

INFORMACIJE MIDEM LETNIK 33, ŠT. 1(105), LJUBLJANA, MAREC 2003
INFORMACIJE MIDEM VOLUME 33, NO. 1(105), LJUBLJANA, MARCH 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
Elektrotehniška zveza Slovenije
Dunajska 10, 1000 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)1, Ljubljana

ZNANSTVENO STROKOVNI PRISPEVKI		PROFESSIONAL SCIENTIFIC PAPERS
A.Levstek, J.Furlan: Električno polje in potencial v okolici atomov primesi v polprevodnikih	1	A.Levstek, J.Furlan: Electric Field and Potential Around Impurity Atoms in Semiconductors
I.Macarol, R.Osredkar: FOLIS, programsko orodje za simulacijo fotolitografskega procesa	8	I.Macarol, R.Osredkar: FOLIS, a PC Compatible Photolitography Simulation Tool
I.Kramberger, Z.Kačič: Izvedba parametričnega nelinearnega filtra za iskanje kožnih značnic v digitalni sliki z FPSLIC tehnologijo	14	I.Kramberger, Z.Kačič: Implementation Of Parametrical Nonlinear Digital Filter For Skin Features Identification in Digital Image Using FPSLIC Technology
M.Milanovič, R.Kovačič: Malosignalni model pretvornika z resonančnim povezovalnim krogom	24	M.Milanovič, R.Kovačič: Small-signal Model of Resonant Link Converter
G.Hrovat, A.Hamler, M.Trlep, M.Bizjak: Analiza segretja motorskega zaščitnega stikala pri trajni tokovni obremenitvi	32	G.Hrovat, A.Hamler, M.Trlep, M.Bizjak: Analysis of Heating Motor Protection Switch by Permanent Current Load
J.Rozman, M.Bunc: Senzor sile za meritve krčenja drobnih mišic	38	J.Rozman, M.Bunc: Sensor of Forces in Small Volume Contracting Tissues
R.Karba, M.Anatasijević-Kunc, A.Belič: Vloga senzorjev v sistemih vodenja procesov	41	R.Karba, M.Anatasijević-Kunc, A.Belič: The Role of Sensors in Control Applications
D.Belavič, M.Hrovat, M.Pavlin, M.Santo Zarnik: Debeloplastna tehnologija za senzorske aplikacije	45	D.Belavič, M.Hrovat, M.Pavlin, M.Santo Zarnik: Thick-Film Technology for Sensor Applications
M.Pavlin, D.Belavič, M.Možek: Senzorji tlaka za medicinsko in industrijsko uporabo	49	M.Pavlin, D.Belavič, M.Možek: Pressure Sensors For Medical and Industrial Applications
J.Holc, M.Kosec, F.Levassort, L. Pascal Tran-Huu-Hue, M. Lethiecq: Integrirani ultrazvočni piezoelektrični pretvorniki za uporabo v medicini	53	J.Holc, M.Kosec, F.Levassort, L. Pascal Tran-Huu-Hue, M. Lethiecq: Integrated Ultrasonic Piezoelectric Transducers for Medical Applications
A.Vodopivec: Sinteza analognih integriranih vezij	57	A.Vodopivec: Synthesis of Analog Integrated Circuits
S.Starašinič: Sprotno dekodiranje s šumom pomešanih signalov	60	S.Starašinič: Real Time Decoder for Coded Signals Mixed with Noise
A.Pleteršek: Končna ojačevalna stopnja z lastnostmi, nastavljivimi z napajalno napetostjo	63	A.Pleteršek: Ratiometric-to-supply Voltage Output Buffer Design
MIDEM prijavnica	70	MIDEM Registration Form
Slika na naslovnici: Integrirani mikrosistem s poljem 32 magnetnih senzorjev za natančno merjenje pozicije, Laboratorij za mikroelektroniko Fakultete za elektrotehniko iz Ljubljane		Front page: Integrated Microsystem with Array of 32 Magnetic Sensors for Precise Position Measurement, Laboratory for Microelectronics, Faculty of Electrical Engineering, Ljubljana

VSEBINA CONTENT

ELECTRIC FIELD AND POTENTIAL AROUND IMPURITY ATOMS IN SEMICONDUCTORS

Andrej Levstek, Jože Furlan

University of Ljubljana, Faculty of electrical engineering, Ljubljana, Slovenia

Key words: physics, electronics, electrostatics, semiconductor modeling, Coulomb potential, microscopic electric potential, microscopic electric field intensity, Debye-Hückel screening, analytical approximation, numerical solution

Abstract: The paper describes different models for the microscopic electric field intensity and electric potential in the surroundings of ionized impurity atoms in semiconductors. The emphasis is placed on a novel comprehensive model that is an improvement of the Debye-Hückel screening applied to semiconductors. In contrast to other described models, the improved model is featured by respecting all three mechanisms of electric field attenuation in solids: dielectric polarization, free carrier screening, and spatial distribution of impurity atoms. Electric potential and electric field intensity profiles are obtained as numerical solutions of the Poisson equation fully respecting the non-linear space charge dependency. Proposed analytical approximations of numerical results facilitate their further use.

Električno polje in potencial v okolici atomov primesi v polprevodnikih

Ključne besede: fizika, elektronika, elektrostatika, modeliranje polprevodnikov, Coulombov potencial, mikroskopski električni potencial, mikroskopska električna poljska jakost, Debye-Hücklov model zakrivanja, analitična aproksimacija, numerična rešitev

Izvleček: V prispevku so opisani različni modeli mikroskopskega električnega polja in potenciala v okolici ioniziranih atomov primesi v polprevodnikih. Poseben poudarek je posvečen novemu izčrpnemu modelu, ki predstavlja izboljšavo uporabe Debye-Hücklovega modela zakritega polja v polprevodnikih. Za razliko od ostalih opisanih modelov, se izboljšana inačica odlikuje s tem, da so upoštevani vsi trije mehanizmi slabljenja električnega polja v trdnih snoveh: dielektrična polarizacija, zakrivanje polja s prostimi nosilci in vpliv sosednjih, prostorsko razporejenih, ioniziranih atomov primesi. Poteki električnega polja in potenciala so izračunani z numeričnim reševanjem Poissonove enačbe z upoštevanjem nelinearne odvisnosti prostorskega naboja. Dobljeni numerični poteki so aproksimirani z analitičnimi funkcijami, ki olajšajo njihovo nadaljnjo uporabo.

1 Introduction

Properties of various semiconductor devices are usually described by analytical approaches based on the macroscopic model of the semiconductor structure. In such models uniform microscopic structure is presumed, e.g., the edge of the conducting band varies only macroscopically because of the built-in impurity concentration profile. Analytical approaches are usually applied only to one-dimensional semiconductor structures. Two or even three-dimensional analyses are prevalently carried out numerically as in many cases analytical solutions are not feasible.

For analysis, description and understanding of certain phenomena it is inevitably necessary to consider local variations of microscopic space charge, electric field and potential. Scattering and capture of free charge carriers by localized charges consisting of ionized impurities represent an important type of such phenomena. Microscopic gradients of the electric potential play also an important role at the increase of the concentration of thermally emitted free carriers from the energy states within the gap across the lowered potential barrier into the conduction or valence band (Poole-Frenkel effect), tunneling of charge carriers through thin potential barriers, etc.

The microscopic electric potential in the surroundings of an ionized impurity atom is affected by three main factors, namely, the atoms of the base semiconductor, mobile charge carriers, and adjacent ionized impurities. Each of these factors weakens the electric field strength and electric potential in its own way. The most accurate result is obtained when all factors are taken into account. Of course, this is not an easy task especially if analytical expressions for the electric potential are desired. However, it is possible to find approximate analytical expressions if some effects are simplified or neglected.

In order to maintain the main ideas clear we will constrain this discussion to n-type semiconductor that is easier to describe and understand. Majority mobile charge carriers in n-type are electrons hence the physical picture is less demanding, as electrons are real particles. Further, all numerical examples and presented diagrams are calculated for silicon since it is the most important semiconductor material.

From the beginning to the end of this article, we shall increase the degree of complexity involved with the described solutions of the microscopic electric potential.

2 Coulomb potential

Silicon crystalline structure is tetrahedral, i.e. each silicon atom is tied to its four neighbors by covalent bonds consisting of a common pair of valence electrons. In n-type silicon, a small part of Si atoms is replaced by donors, i.e., impurity atoms with five valence electrons. Four valence electrons form covalent bonds to the four neighboring silicon atoms leaving the fifth one loosely tied to the core. The space around impurities is filled with host atoms Si that weaken the electrostatic force in such an extent that at room temperature almost all donors are ionized /1/. In other words, the fifth electron of almost every donor atom gets sufficient kinetic energy to become mobile. Impurity atoms themselves remain firmly bound to the host lattice and can be treated as a uniform spatial distribution fixed-point charges $+q$. However, the spatial pattern of impurities is not a strict arrangement in the sense of a crystal lattice, but it is uniform on the average with minor deviations from its regular positions. The process of ionization does not change the neutrality of the observed volume on the macroscopic level, because the interstitial space is filed with mobile electrons, which contribute the negative charge that exactly compensates the positive charge of fixed ions.

2.1 Electric potential of an isolated ion

The straightforward solution of the electric potential of an ionized impurity atom can be derived from Coulomb's law /2/ since an ionized impurity can be viewed as an isolated point charge $+q$

$$V(r) = \frac{q}{4\pi\varepsilon r}, \qquad (1)$$

where q denotes the elementary charge, r is the distance from the point charge and ε is the dielectric constant (permittivity) of the environment. Eq. is well known as Coulomb potential. The potential at a certain point in space is meaningful only when a reference zero-potential point is specified. In most cases, this point is taken at infinity.

The electrostatic potential of an ionized impurity atom alters the shape of band-edge potentials in the close surroundings. The resulting joint profile is obtained by simple scalar addition of the potentials. This is a significant advantage of using electric potential V rather than electric field intensity **E**. The diagram of band-edge potentials, shown in Fig. 1, illustrates the influence of the Coulomb potential of an ionized donor in silicon.

The diagram in Fig. 1 is only one-dimensional representation of the spherically symmetric spatial field. It is actually the plot of the potential along a straight line, which is laid through the charged center. The valence and conducting band-edge potential are denoted by V_V and V_C, respectively. With increasing distance r, those potentials asymptotically approach their macroscopic values V_{V0} and V_{C0}.

In the described Coulomb potential model, only dielectric attenuation is taken into account. Dielectric polarization of

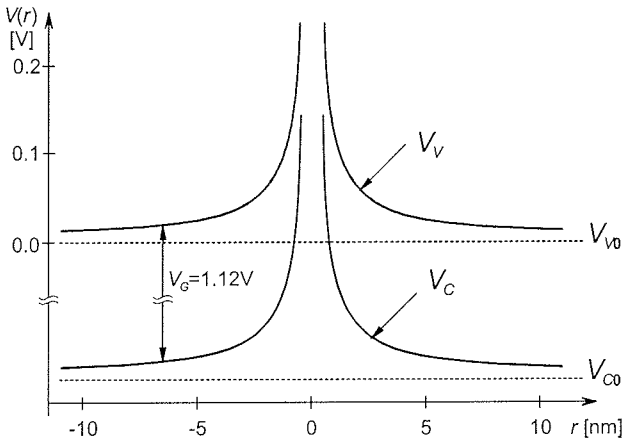


Fig. 1: Band-edge potentials in silicon modified by the Coulomb potential of an isolated ionized donor atom at r = 0.

host semiconductor atoms is responsible for the reduction of the electric field. The application of the macroscopic dielectric constant for microscopic fields /1/, /3/ is justified by the fact that the distance between host atoms is much smaller, e.g., 0.235 nm for silicon, than the distances of interest.

Permittivities of materials are usually expressed as a product of the permittivity of empty space ε_0 and the relative permittivity ε_r of a particular material. For common semiconductor materials ε_r ranges approx. from 9 to 17, in particular for silicon we get ε_r = 11.7 /4/. These relatively high values of ε_r reduce the electric potential and the associated binding energy of donors after being incorporated in host semiconductor.

Finally, the term *isolated impurity* needs some explanation. We use this expression for semiconductor doped with very low concentrations, which result in a sparse spatial distribution of impurities, such that the mutual influences between immediate neighbors are diminished by large average distances to a negligible level. There is no explicit margin of the doping N_D, below which the impurities are considered isolated, but as a rule of thumb $N_D = 10^{14}$ cm^{-3} can be used for silicon.

2.2 Effect of neighbors

Practical doping concentrations are substantially higher than 10^{14} cm^{-3} therefore, influences of the ionized impurity atoms in the neighborhood have to be taken into account. Neighbors modify the potential and electric field distribution of an isolated ion (1), introducing saddles between adjacent charged centers /5/. The modified circumstances can be easily explained by the rough draft of the electrical potential profile along the straight-line laid through two adjoining charged centers shown in Fig. 2.

The joint potential is obtained by adding the Coulomb potentials (1) of each ionized atom, one at the origin and the other at $r = 2R$. For distances $0 < r < 2R$ we get

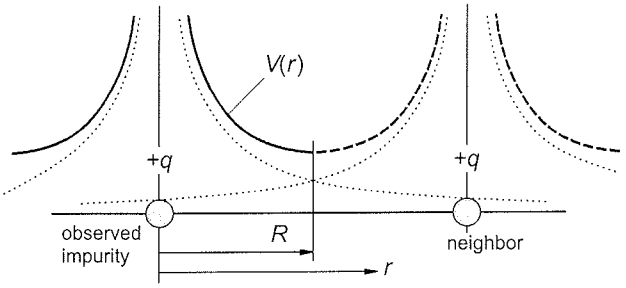


Fig. 2: Schematic representation of the electric potential along the straight-line running through two adjoining ionized impurity atoms in n-type semiconductor.

$$V(r) = \frac{q}{4\pi\varepsilon r} + \frac{q}{4\pi\varepsilon(2R-r)}. \quad (2)$$

Adjacent impurity atoms are brought closer to each other as their concentration N_D increases regardless of their spatial distribution. Consequently, an increased impurity concentration raises the potential saddle at $r = R$. The diagram shown in Fig. 2 is valid also for other directions in space where other neighbors are positioned. If impurities are assumed to have a simple body-centered cubic lattice then each ion is surrounded with six closest neighbors. Though the ideal spherical symmetry of the isolated ion potential is perturbed, high degree of central symmetry remains for radii $r \leq R$ where Eq. (2) may be applied in any direction. It is important to note that the magnitude of the electric field intensity E drops to zero at $r = R$ where the potential has a local minimum. At this stage, we will not discuss the exact diagram of band-edge potentials because in this case the zero-valued reference point at infinity does not exist. The modified band structure will be presented in section 4.1.

3 Debye-Hückel screened potential

Free carriers, which are randomly moving in the space between the ionized impurities, compensate the space charge of fixed ions. The electrostatic field is screened by mobile carriers due to electric forces that concentrate the cloud of charge carriers near ionized impurity atoms. Such phenomena have been noticed and first addressed in the context of electrolytes by Debye and Hückel /6/. This model can be applied to analyze the screened Coulomb potential of one ionized impurity atom. One has to be aware that the Debye-Hückel (DH) theory of screening is based upon simplifications, which are justified for dilute ionic solutions, in order to obtain an analytical expression for the screened potential. Thus, the DH model involves some deficiencies, which become rather important when this model is applied to semiconductors.

The expression for the screened potential is derived by solving Poisson's equation /1/,/7/

$$\nabla^2 V = -\frac{\rho}{\varepsilon}, \quad (3)$$

where e is the macroscopic permittivity of the host semiconductor and ρ is the space charge consisting of rigid ionized impurities and mobile carries. For n-type semiconductor the charges of ionized impurity atoms and free electrons mutually cancel each other. Only the charge of one electron remains, since the positive charge of the observed ion is left out. The solution of (3) is obtained by the linearized Boltzmann distribution that is used for ρ(*V*).

For an n-type semiconductor with impurity concentration N_D the DH screened Coulomb potential around the fixed ionized impurity atom is expressed by /8/

$$V(r) = \frac{q}{4\pi\varepsilon r}\exp\left(-\frac{r}{L_D}\right), \quad (4)$$

where L_D is *Debye length,* given by

$$L_D = \sqrt{\frac{kT\varepsilon}{q^2 N_D}}. \quad (5)$$

Two important distinctions should be made between the circumstances in semiconductors and those in electrolytic solutions, on which the DH model is focused:

i. Ions are rigidly built in the crystal structure of the semiconductor
ii. Mobile carries are unipolar (of only one polarity) as long as the observed semiconductor is not intrinsic, i.e., N_D is at least an order of magnitude bigger than n_i.

The DH screened potential has zero-valued reference point at infinity hence it can be easily compared with the Coulomb potential V_{Cou}. Plots of both potentials vs. distance *r* for silicon ($N_D = 10^{17}$ cm^{-3}) are shown diagram in Fig. 3.

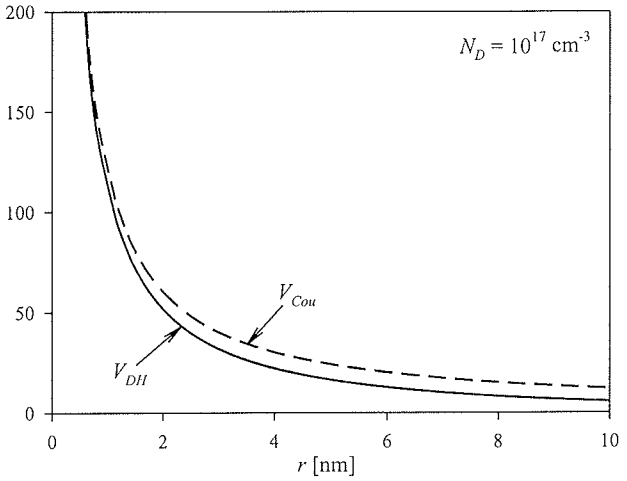


Fig. 3: Microscopic potential against distance r *from ionized impurity in Si.*
(¦¦): screened DH potential Eq. (3)
(—): unscreened Coulomb potential Eq. (1)

The plot of DH screened potential V_{DH} declines steeper and is always lower than the Coulomb potential. The degree of reduction is governed by the exponential factor in Eq. (4) that depends on impurity concentration N_D. The DH approach does not include the influence of neighboring ions that significantly modify the shape of the potential as it is shown in section 3.2.

4 Comprehensive model of the microscopic potential

4.1 Numerical solution

The exact screened potential, which respects the above-mentioned effects at the highest possible extent, can be calculated only numerically. In this section, we will present only the important starting assumptions and the main steps of the method. More details can be found in /9/, /10/.

Unlike the Debye-Hückel approach, which examines the potential of a single ion within an unlimited space, the combined approach is focused on a finite volume confined by the same volumes placed around the neighboring fixed ions. The volume is approximated with a sphere whose diameter is equal to the closest distance between adjoining ions. This approximation substantially simplifies the mathematical complexity of the problem, as the electric potential field possesses spherical symmetry. As it is shown in section 2.2, the potential $V(r)$ becomes flat at midway between adjacent impurities

$$\left.\frac{dV}{dr}\right|_{r=R} = 0 \qquad (6)$$

The microscopic potential is obtained by solving the Poisson differential equation (3) for boundary condition (6). In spherical coordinates, Eq. (3) becomes an ordinary differential equation

$$\frac{d^2V}{dr^2} + \frac{2}{r}\frac{dV}{dr} = -\frac{\rho}{\varepsilon}, \qquad (7)$$

as central symmetry of the potential and space charge is presumed. Space charge density ρ of the electron cloud is expressed by electron concentration $n(r)$ that is determined by the density of states distribution in the conduction band and by Fermi-Dirac occupation probability, defining

$$\rho = -qn = -qN_c F_{1/2}(\eta_C), \qquad (8)$$

where N_c is the effective conduction band density of states and $F_{1/2}(\eta_c)$ is the Fermi-Dirac integral, which is approximated by analytical functions in various regions of normalized potential $\eta_c = q(V_F - V_C)/kT$ /11/.

The screened potential of the localized charge V, which modifies V_C and thus the space charge density ρ (8), has to meet the neutrality condition that can be expressed in integral form

$$\iiint_{V_S} \rho\, dv + q = 0, \qquad (9)$$

where V_S denotes the observed sphere. The a priori unknown difference $V_R = V_C - V_{C0}$ at midway $r = R$, upon which the total charge of the electron cloud depends, is determined by an iterative algorithm. The non-linear Poisson equation, obtained by inserting (8) into (7), is solved by numerical integration, which starts at midway $r = R$, using boundary condition (6) and an initial guess value V_R. After each iteration, V_R is successively adjusted until the neutrality condition (9) is achieved.

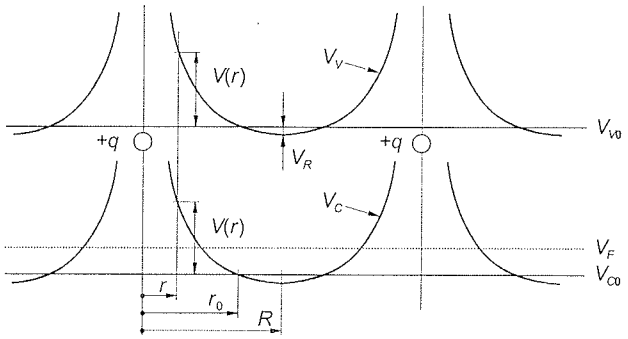


Fig. 4: Schematic diagram of band-edge potentials on a straight line running through two positive ions. Macroscopic band-edge potentials V_{V0} *and* V_{C0} *are modified by the screened potential* V(r).

The band-edge potentials are then given by

$$V_C = V_{C0} + V, \qquad (10)$$

and similarly

$$V_V = V_{V0} + V. \qquad (11)$$

Schematic diagram of screened band-edge potentials is shown in Fig. 4, where some important properties of the final solution can be seen. The polarity of V is both, positive and negative, thus conduction band-edge potential V_C is located partly above and partly below its macroscopic value V_{C0}. This variation of V_C is accompanied with similar but more intense deviations of the electron density n from its macroscopic equilibrium value N_D. In the spherical region close to the ion, the concentration of free electrons n is much above N_D. This increased negative space charge is compensated by $n < N_D$ in the outer shell in order to meet the neutrality condition (9). The screened potential V is always negative for radii r beyond a certain $r > r_0$ (see Fig. 4), hence $V_R = V(R)$ is always slightly negative.

The range of the horizontal axis in Fig. 5 covers all the radii within the observed sphere from 0 to R, because for simple cubic spatial distribution of impurities and $N_D = 10^{17}\text{cm}^{-3}$, we get $R = 10^{-8}$ m.

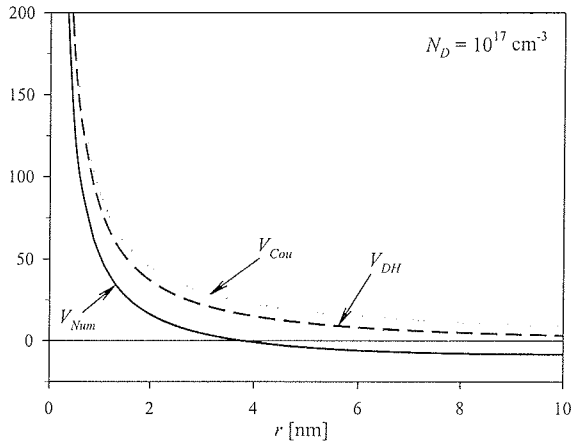


Fig. 5: Numerical and classical potentials against distance r from ionized impurity in n-type Si, $N_D = 10^{17} cm^{-3}$.
(¦¦¦): screened potential - numerical solution of Eq. (7)
(——): screened DH potential Eq. (3)
(·····): unscreened Coulomb potential Eq. (1)

The numerically computed potential, combines all three major mechanisms that attenuate the electric potential as we move from the charged center, i.e., dielectric polarization, space charge screening, and the influence of adjacent ions, thus it exhibits a significant improvement upon the classical expressions, i.e., the Coulomb and DH potential, respectively.

4.2 Analytical approximations

4.2.1 Screened electric field intensity

The main drawbacks of the numerical method are, first, the very nature of numerical results, which are usually obtained in tabular form, and second, the extensive and time-consuming iterative algorithm. In this section, we present analytical approximations proposed in /10/ for the numerical solution of the microscopic electric field and potential, respectively. Throughout this article, the main emphasis is put on the electric potential, owing to its scalar nature and its tight connection to energy bands in semiconductor. Though in some cases, especially those involved with kinematics of charged particles, electric field intensity **E** seems to be more appropriate.

The magnitude of electric field intensity E of a positive charge q given by Coulomb's law

$$E = \frac{q}{4\pi\varepsilon r^2}, \qquad (12)$$

becomes zero only when the distance r gets infinite. The same applies to DH screened electric field, which is obtained by differentiating the potential (4). The boundary condition (6), which is postulated to mimic the effect of neighbors, can be met by an approximate expression for the electric field in the range $0 < r < R$ in the form

$$E = \frac{q}{4\pi\varepsilon r^2}\left[1-\left(\frac{r}{R}\right)^m\right] \qquad (13)$$

in which the screening effect is contained in the factor $[1 - (r/R)^m]$, where m determines the degree of screening. The general Eq. (13) satisfies the boundary condition (6) for any selected value of m. When lowering the exponent m, the electric field decreases more vigorously with an increasing radius r, thus enhancing the screening effect. At low values of r/R the electric field in Eq. (13) approaches the unscreened Coulomb case.

The optimal value m_{opt}, which minimizes the total squared error between the analytic approximation and the numerical solution, is shown /10/ to depend slightly on the impurity concentration. The exponent m_{opt} decreases from 1.8 to 0.9 as the impurity concentration N_D is being increased from 10^{15} up to 10^{19} cm^{-3}, respectively. If simplicity of the analytical expression is desired then integer values 1 and 2 are preferred.

The diagram in Fig. 6 shows different plots of the magnitude of electric field intensity in the surroundings of an ionized impurity atom, namely, E_{Cou} given by (12) (Coulomb's law), E_{Num} is obtained numerically, using the algorithm presented in previous section, and E_{App} according to the analytical approximation with m = 1.5. The plot of the approximate electric field E_{App} exhibits very good matching to the numerical electric field E_{Num}, which is considered accurate. The values of E_{Num} and V_{Num} are computed simultaneously when the Poisson's equation (7) is being solved.

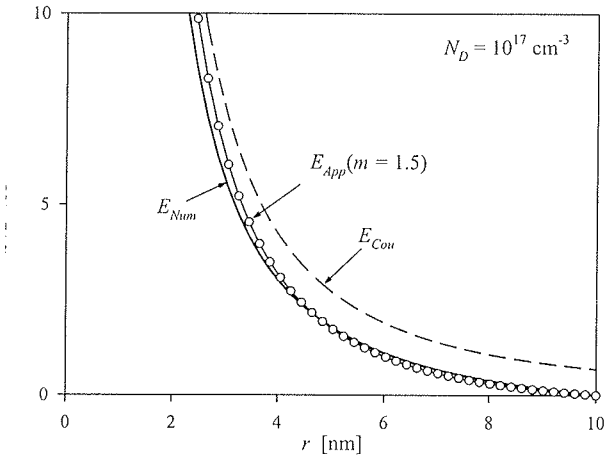


Fig. 6: Magnitude of electric field intensity in the region of an ionized impurity atom in n-type Si, $N_D = 10^{17} cm^{-3}$.
(——): numerical solution
(oooo): analytical approximation Eq. (13) with m = 1.5
(----): unscreened Coulomb electric field Eq. (12)

4.2.2 Screened electric potential

In order to analyze the effect of the proposed approximation (13), an electrostatic potential for arbitrary m is ob-

tained. Integrating Eq. (13) in the range from R to arbitrary r gives the form

$$V - V_R = \frac{q}{4\pi\varepsilon r}\left[1 + \frac{1}{(m-1)}\left(\frac{r}{R}\right)^m - \frac{m}{(m-1)}\frac{r}{R}\right], \quad (4)$$

where V_R is the potential at $r = R$, referenced to the macroscopic band-edge potential V_{C0} (see Fig. 4). The expression is valid even for $m = 1$ since the limit $\lim_{m\to 1}(V - V_R)$ exists. In this special case Eq. (14) changes to

$$V - V_R = \frac{q}{4\pi\varepsilon r}\left(1 - \frac{r}{R} + \frac{r}{R}\ln\frac{r}{R}\right). \quad (15)$$

The exact value of V_R can be obtained numerically from the boundary condition (9) and Eq. (8) for the space charge of the electron cloud. However, it is possible to obtain an approximate value of V_R by expressing the electron concentration with the Boltzmann instead of the Fermi-Dirac distribution. The derivation of V_R /10/ yields two final analytical expressions for the screened potential:

$$V = \frac{q}{4\pi\varepsilon r}\left[1 + \frac{1}{(m-1)}\left(\frac{r}{R}\right)^m - \frac{3}{2}\frac{m(m+1)}{(m-1)(m+2)}\frac{r}{R}\right], \quad (16)$$

for the general value of $m \neq 1$ and

$$V = \frac{q}{4\pi\varepsilon r}\left[1 + \ln\frac{r}{R}\left(\frac{r}{R}\right) - \frac{7}{6}\frac{r}{R}\right], \quad (17)$$

for the special case with $m = 1$.

An evaluation of the approximate screened potential (16) is shown in Fig. 7. The approximation V_{App} with $m = 1.5$ is plotted together with the numerical potential V_{Num} discussed in section 4.1 and Coulomb potential V_{Cou}, which is shown for reference. The value of the exponent m is not exactly the optimal value for $N_D = 10^{17}\text{cm}^{-3}$, but is a reasonable choice for most doping concentrations that appear in electron devices.

The shapes of the potential profiles V_{App} and V_{Num} are in good agreement over the whole range of distances. However, there is small constant difference between the two potentials that are being compared. In contrast with the interweaving plots of electric field E_{App} and E_{Num} in Fig. 6, the potential V_{App} remains slightly above V_{Num} over the entire range of radii. This difference arises from the derivation of V_R, in which linearization of the Boltzmann exponential dependency is used.

5 Conclusion

The intention of this article was to present a review of the various models for the microscopic electric field and potential. The described models were devised or modified

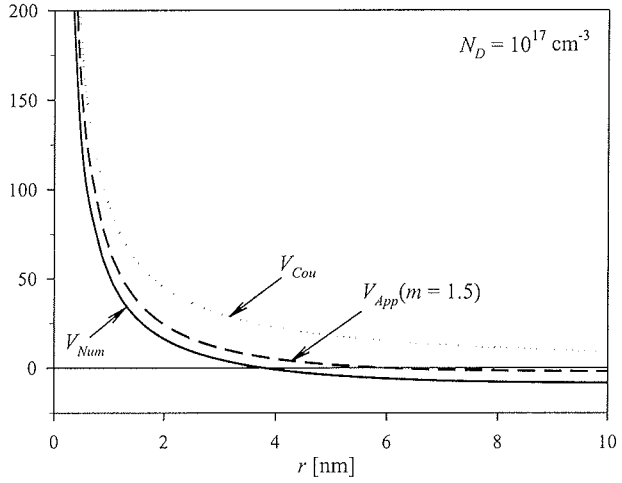


Fig. 7: Approximate and numerical potential versus distance r from ionized impurity in n-type Si, $N_D = 10^{17}cm^{-3}$.
(——): screened potential V_{Num} - numerical solution of Eq. (7)
(- - -): analytical approximation V_{App} Eq. (16) with m = 1.5
(·····): unscreened Coulomb potential Eq. (1)

for the use in semiconductors, however, some of them, e.g. the numerical method, can be applied also in other areas.

In order to maintain informational nature and clearness of the paper many details and results have been omitted. Extensive tests of numerical method for the whole range of interesting doping concentrations have been carried out. The results show god agreement between the approximate and numerical profiles.

Numerically calculated potential and its approximation represent a significant improvement of the DH model, because all three mechanisms of electric field attenuation (dielectric polarization, screening by mobile charge carriers, effects of neighbor impurity atoms) are taken into account. The choice of the appropriate model in a particular case depends on a variety of factors. As general rule, it can be suggested that in cases where higher doping concentrations are concerned a comprehensive model would be more appropriate, since the effect of screening is more intense at high space charge densities.

6 References

/1/ C. T. Sah, *Fundamentals of Solid-State Electronics*, World Scientific Publishing Co., Singapore, 1991

/2/ D. K. Cheng, *Field and Wave Electromagnetics*, Addison-Wesley Publishing Company, Reading 1989

/3/ W. Shockley, *Electrons and Holes in Semiconductors*, D. Van Nostrand Co., New York 1951

/4/ M. E. Levinshtein, S.L. Rumyantsev, *Handbook Series on Semiconductor Parameters*, **vol.1**, M. Levinshtein, S. Rumyantsev and M. Shur, ed., World Scientific, London, 1996

/5/ K. W. Böer, "The Conduction Mechanism of Highly Disordered Semiconductors II. Influence of Charged Defects", *phys. stat. sol.*, vol. 84, pp. 733-740, 1969

/6/ J.O'M. Bockris, A. K. N. Reddy, *Modern Electrochemistry, Vol 1*, Pllenum Press, New York 1970

/7/ W. J. Moore, *Physical Chemistry*, Addison Wesley Longman, London 1972

/8/ K. W. Böer, *Survey of Semiconductor Physics,* Van Nostrand Reinhold, New York 1990

/9/ A. Levstek, J. Furlan: Approaches to the theory of microscopic potential and free carrier distribution in semiconductor, Proc. 36th Int. Conf. on Microelectronics, Devices and Materials and the Workshop on Analytical Methods in Microelectronics and Electronic Materials, October 18 - 20, 2000, Postojna, Slovenia, pp. 329-334.

/10/ A. Levstek, J. Furlan, "Microscopic electric field in the surroundings of ionized impurities in semiconductor", *Journal of Electrostatics,* (in press)

/11/ J. S. Blackmore, Approximations for Fermi-Dirac Integrals. Especially the Function $F_{1/2}$(h) Used to describe Electron Density in a Semiconductor, Solid-St Electron, 25, (1982), 1067-1076.

mag. Andrej Levstek
dr. Jože Furlan
University of Ljubljana
Faculty of Electrical Engineering
Tržaška c. 25, 1000 Ljubljana, Slovenia
Tel.: +386 1 4768 325, Fax: +386 1 4264 630
Email: andrej.levstek@fe.uni-lj.si

Prispelo (Arrived): 29.07.2002 *Sprejeto (Accepted): 25.03.2003*

FOLIS, A PC COMPATIBLE PHOTOLITHOGRAPHY SIMULATION TOOL

Igor Macarol[1], Radko Osredkar[2]

[1] Skupina Atlantis d.o.o., Slovenia, [2] Faculty of Computer Sciences and Faculty of Electrical Eng., University of Ljubljana, Slovenia

Key words: photolithography, projection aligners, process simulation, microelectronics

Abstract: FOLIS, a relatively simple, PC compatible photolithography simulation tool, developed for photolithography process development and demonstration, is presented, and physical models incorporated into software outlined. Examples of simulations by FOLIS of selected photolithography process steps are given.

FOLIS, programsko orodje za simulacijo fotolitografskega procesa

Ključne besede: fotolitografija, projekcijski poravnalniki, simulacija procesov, mikroelektronika

Izvleček: V prispevku je predstavljen FOLIS, relativno preprost in za uporabo na osebnih računalnikih prilagojen programski paket za simulacijo in demonstracijo fotolitografskega procesa. Očrtani so fizikalni modeli, ki so vključeni v program in prikazani primeri simulacij izbranih faz v fotolitografkem postopku.

Introduction

Lithography is the cornerstone of modern IC manufacturing, and lithography tools and process characterization at the core of the lithography process engineering. A great majority of the ICs today are manufactured by optical photolithography. The concept behind it is relatively simple: a light sensitive photoresist is spun onto the wafer forming a thin layer on the surface, which is selectively exposed by light through a mask, containing the pattern of the particular layer to be patterned. The resist is then developed, which transfers the pattern on the mask to the wafer. The remaining resist is finally used as a mask e.g., to etch the underlying layer. As the dimensions of the features to be fabricated on the wafer approach 0.1 μm, the classical limit of the resolution of optical tools used in photolithography, the implementation of the simple concept becomes quite complex. In fact, complex enough for the tradition empirical tools of the IC manufacturing process development to be insufficient for the task at hand. They have to be complemented by different computer simulation tools which make understanding and optimization of the performance of the lithographic process possible. Such tools are based on the Fourier optics to describe the performance of exposure tools, and resist chemistry to describe the formation of the mask pattern in the resist. Photolithography simulation tools generally follow the directions developed by Dill /1, 2/, and a number of commercial products is available for simulation purposes. However, these tools themselves are quite complex and in connection with optimizing performance of the projection Ultrateh Stepper Aligner UTS 1100 in our Microelectronics laboratory, there was a need to develop a relatively simple, PC based simulation tool, which could also serve for demonstration purposes. The result of this efforts is FOLIS (FotoLItografski Simulator), a software simulation tool presented in this contribution.

Optical Considerations

All of the projection exposure systems used in IC manufacturing industry today are diffraction limited optical instruments. Consequently Frauenhofer diffraction theory (3) has to be incorporated in any simulation describing their performance. The basic tenet of the theory is that a point object is imaged by the optical system into a finite patch of light, or blur spot, in the image plain, i. e. the image of the source is not perfect. This obviously has important consequences for imaging the fine, detailed pattern present on the mask on the photoresist covered wafer. In FOLIS the imaging of the mask pattern is modeled by the standard Fourier analysis of the Frauenhofer diffraction. The diffraction theory is used to calculate spatial distribution of the light intensity I(r) in the blur spot, the image of an aperture in the mask. There is a small but significant difference between diffraction patterns of a circular and rectangular aperture. For the former, an axially symmetric spot (Airy disc), the intensity is given by:

$$I(r) = I_o \left(\frac{2J_1(x)}{x} \right)^2$$

where x = (π a sin Θ)/λ, I_o the intensity maximum in the center of the disk, a the diameter of the aperture in the mask, Θ angle between the direction of the intensity maximum and direction of the current position in the image, λ the wavelength of the light, and J_1 the first order Bessel function. The expression for rectangular aperture is analogous and results of a FOLIS simulations for both are shown in Fig. 1. According to the Rayleigh criterion the maximum resolution R of the projection optics can be extracted from the above expression:

$$R = 1.22\,\lambda/\,NA,$$

where NA is the numerical aperture of the projection system. This expression strictly applies to point sources only. In real photolithographic systems, imaging a variety of shapes of finite dimensions on resists of different abilities to distinguish closely spaced features, further details of the imaging, e.g. defocusing of the image plane, have to be considered. Usually all such effects are incorporated into a factor k_{LW} in the corrected expression for R, i.e. the minimal line width that can be imaged:

$$R = k_{LW}\lambda/\,NA$$

In practice k_{LW} is a number between 0.8 and 1.0, resulting in e.g. 0.36 µm resolution for the Hg emission I line in the UTS 1100 aligner. As this projection aligner utilizes I, H, and G emission lines, all 3 are considered in the FOLIS simulations.

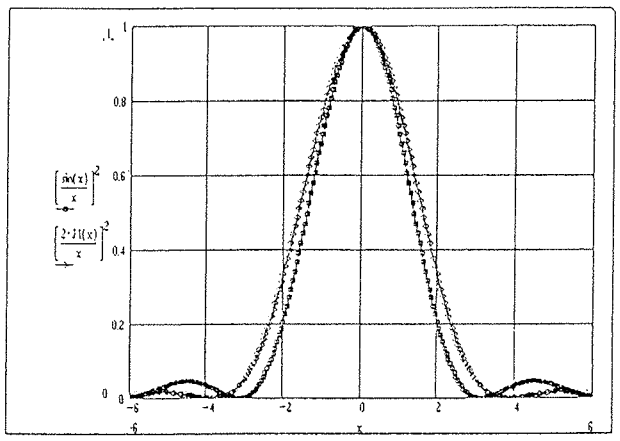


Fig. 1: Airy disk simulaton by FOLIS for circular point source (diamonds) and line source (squares). x–axis divisions are scaled by the objective focal length.

Depth of focus (DOF) of the imaging optics also influences the resolution limit and is considered in FOLIS. The Rayleigh criterium for the DOF is simply translated to the requirement that the lengths of the on-axis and edge of entrance aperture rays not differ by more than $\lambda/4$, and

$$DOF = k_1\,\lambda/(NA)^2$$

where k_1 is a factor analogous to k_{LW}, accounting for the increased DOF for larger features, the dependence of DOF on other factors, e.g. resist process.

Another basic optical concept allowing modeling of the aerial image on the surface of the photoresist film, which is incorporated into the FOLIS, is the modulation transfer function (MTF) /3/. MTF is basically a measure of the contrast in the aerial image produced by the exposure system. It is defined as

$$MTF = \frac{I_{maax} - I_{\min}}{I_{\max} + I_{\min}}$$

where I is the intensity of light at different parts of the image. MTF of a system depends on a variety of factors, including the illumination light wavelength, mask spatial frequency and feature size to be transferred to the photopolymer, the NA of the lens, and spatial coherence of the source. MTF decreases with the mask feature size and is at the resolution limit for an ideal mask only 0.5. Generally, an exposure system needs to achieve a MTF value of at least 0.5 in order for the resist to properly resolve the features incorporated in the mask. In principle the concept of MTF strictly applies only to coherent illumination, however, its approximations for partially coherent radiation are known /4/ and this is incorporated into the FOLIS. The aberrations of the Wynne–Dyson imaging system of the UTS 1100 are neglected in our simulator, and DOF treated as a phase aberration. In Fig. 2 an example of the Bosung plot calculated for a diffraction grating-like mask of 0.6 µm line width and 1.2 µm pitch is shown. It should be noted that in practice actual measurements of such data are quite difficult and therefore an accurate modeling of the image of great value.

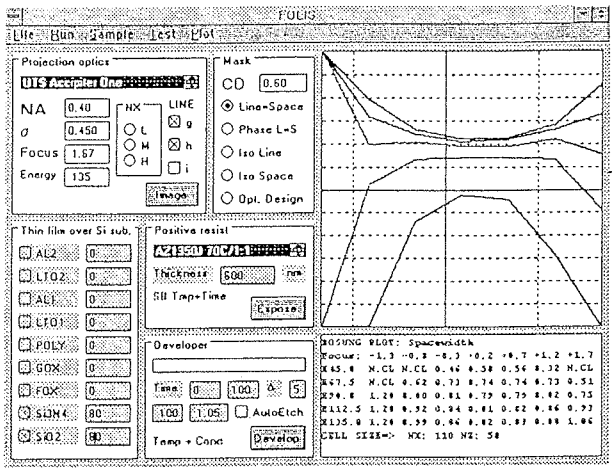


Fig. 2: Bosung plot (UTS 1100 projection stepper and 600 nm AZ1350J photoresist on SN-SO covered wafer) for a 0.6 µm CD line-space mask structure, calculated by FOLIS.

Photoresist Image

Formation of the aerial image, and its quality, by the exposure tool is the first step of the photolithographic process that is considered in modeling. Translation of this image into its 3 dimensional replica in the photoresist is the next

one. Geometric accuracy, exposure speed, and also patterned resist physical and chemical properties have to be considered. While the later aspects of this stage of processing are not modeled by the FOLIS at this time, the first two are.

There are several reasons for the optical intensity pattern of the aerial image to be different from the intensity pattern in the resist layer. One is the reflection of light from underlying structures which results in establishing a 3 dimensional standing wave pattern in the resist, resulting in an exposure pattern that reflects it. Modeling of the standing waves is relatively simple only if it is assumed that the light entering the film is all vertically incident, i.e. parallel rays perpendicular to the wafer plane, and this is the approximation used in the FOLIS. In an exposure systems with a high NA this is clearly not the case. Partial coherency of the light source and the reflected light further complicate the calculations as they involve an integration of effects over the total angle of the incoming light. Non-uniform resist film thickness and similar factors of random nature are not considered.

The calculation of the standing wave pattern in the resist layer is a straightforward application of the electromagnetic theory /3,5/ - electric field at different depths in the film is calculated, taking into account the change in wavelength of the light in the photoresist film due to its refractive index, reflection coefficients at interfaces and absorption coefficient of the film. In FOLIS this is accomplished by the standard technique of introducing the complex refractive index. The light intensity is then simply proportional to the square of the magnitude of the electric field. In Fig 3. a simulation of the standing wave pattern in a uniform photoresist film of 0.65 μm thickness on bare Si substrate, for the 3 wavelengths used in the UTS 1100 projection aligner, is shown. The actual calculation of the standing wave pattern in a more realistic structure, e.g. oxide layer in top of Si, or several films, each with different optical properties, is considerably more complex, but is approached in an analogous manner as the simplest case.

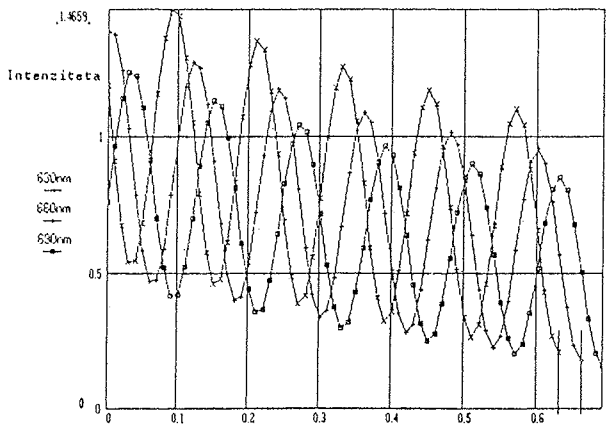


Fig. 3: Standing wave pattern in a 0.65 μm photoresist layer for G, H, and I Hg emission lines, as simulated by FOLIS. Intensity units are arbitrary.

The light intensity in the photoresist film, calculated as described above, is not directly related to the transfer of the aerial image on the film surface into the 3D image structure in the exposed film. During the exposure the optical properties of the resist material change with exposure time (the so called resist bleaching). The effect can be handled by the Dill model of the positive photopolymer /6, for a review see 2/. In this model the film is treated, essentially, as a succession of a number of thinner layers, each thin enough for the bleaching effects to be considered uniform throughout the film subdivision, and certainly thinner than $\lambda/2n$ (where n is the refractive index of the photopolymer material), the spatial frequency of the standing wave pattern in the resist film. The changes in the exposed photoresist are related to the changes of concentrations of one of its components, e.g. the inhibitor, and local bleaching rate calculated from the concentrations and light intensity at a point. The model involves determining 3 parameters (Dill parameters) for the resist and iteratively solving 2 coupled equations, simultaneously with the equation of the standing wave pattern. The results of the procedure in FOLIS, for 1 μm film of the AZ1350J resist on 3 different underlying structures, exposed in UTS1100 projection aligner at 50 mW/cm^2, is shown on Fig. 4. Only G and H Hg emission lines are included in the calculation. As most modern exposure tools use monochromatic light, such effects are even more pronounced, and consequently the control of critical dimensions in the exposed resist film even more difficult.

Photoresist Developing

Photoresist developing is a surface controlled etching process /6/. In modeling it is assumed, that at each point the developer solution etches the surface of the resist isotropically, with the etch rate at each point governed by the local concentration of the resist inhibitor. Of course, it is exactly the local (normalized) concentration of the inhibitor that is calculated by the exposure model described above. The dependence of the etch rate on the inhibitor concentration is nonlinear. In FOLIS, as in most models, terms higher than quadratic are neglected. Time evolution of the developing resist profile is calculated by setting up a 2 dimensional grid, where the inhibitor concentration is defined in each cell, and the developer moving into the resist at local etch rate, thereby evolving the profile. 2 different methods for calculating the evolving pattern are used in FOLIS. One is an interactive cell method that is computationally relatively simple. An example of calculating the developed profile in 0.5 μm AZ1350J resist film by the interactive cell method is shown on Fig. 4. The relatively large size of the cells has been chosen for clarity. The 2 dimensional interactive cell method can easily be extended to 3 dimensions, and an extension of FOLIS in this direction is considered.

The second method for calculating the resist profile in FOLIS is the advancing front method. In this case the lat-

$M=1$

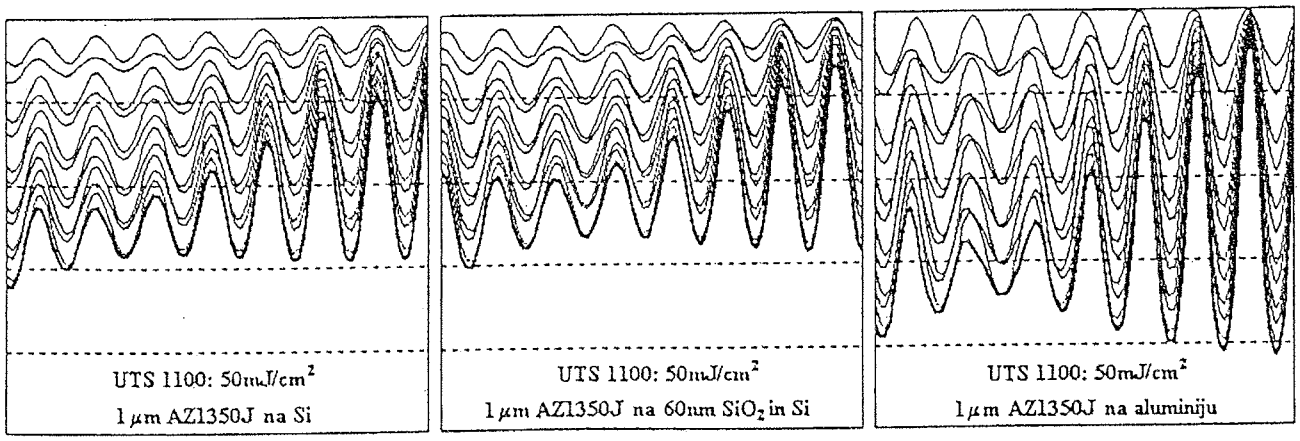


Fig. 4: Time evolution of inhibitor concetration M in photeresist on different substrates (Si, 60 nm SO on Si, and aluminum), as simulated by FOLIS.

tice points define, locally, the interface between the resist and developer and this interface (front) advances into the resist, according to the local etch rate. Such a computation process is easy to relate to the evolving resist profile, however, it is computationally quite demanding and slow. An example of developed 1.5 μm AZ1350J resist film, as calculated by this method, is shown on Fig. 5 and Fig. 6.

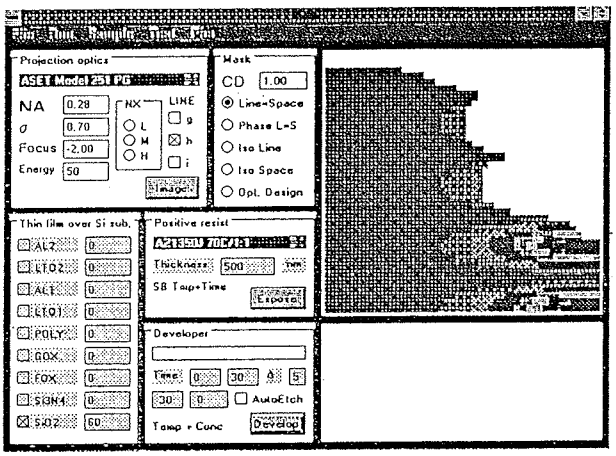


Fig. 5: Photoresist development simulation (interactive cell method) by FOLIS.

Phase Masks

Efforts to improve the resolution of lithography systems are directed not only towards improving the exposure tools and photoresist processing, but also to designing masks with advanced features that allow better transfer of the design pattern to the photoresist. For the projection lithography tools phase masks are of considerable interest and it thus seemed desirable to include, for demonstration purposes, at least some of the features of phase mask operation into the FOLIS. The concept of phase shift masking techniques is not new and has been proposed some time ago /7/, however, its application to the masking process prooved to be quite complex. The concept is based on the elementary fact, that any diffraction pattern observed in the aerial image plane is a result of superposition of electric fields of the light waves scattered by the mask. Diffraction minima in the image are the sites where the fields cancel, and conversely maxima the sites where they add constructively. At an image point the phases of the light waves, arriving from different positions of a standard e.g. chrome "black and white" mask, are determined solely by the different geometrical path lengths from the point to the different positions on the mask. The result of superposition, e.g. for a circular opening in the mask, is the familiar Airy disk pattern of the image. However, the phase of light waves emerging from the mask can also be altered by adding to the mask the so called phase shifters, which are basically transparent films of appropriate thickness and index of refraction. For a 180 deg. phase shift the relation between the thickness d, refractive index n and the wavelength of the light λ is

$$d = \lambda/2(n - 1)$$

By suitable positioning of the phase shift plates on the mask it is possible to arrange for the electric fields at desired sites in the image to be canceled. The photoresist responds to the intensity of the light, which is proportional to the square of the electric field and is thus not effected directly by the phase of the light waves. But by causing the fields to cancel at desired image sites, the resolution and quality of the aerial image can be significantly improved (2,7).

It is not excessively difficult to calculate, by the Fourier methods, the light intensities in the image plane for phase masks with relatively simple geometries. Fig. 7. illustrates a FOLIS simulation for a via mask in which the geometry of the contact via is represented by an opening in the chrome,

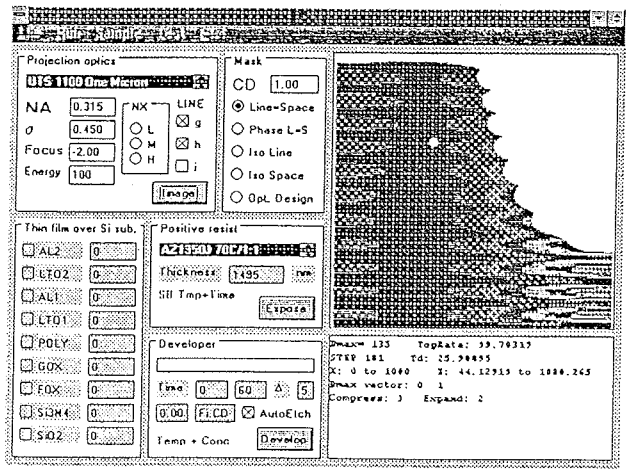


Fig. 6: Photoresist development simulation (advancing front method) by FOLIS.

i.e. dark field, (a), chrome patch (light field, b), 180 deg. phase shift plate with dimensions of the via on light field (c), and 180 deg. phase shift plate behind in an opening (dark field + phase shift, d). Top part of the figure represents the (normalized) electric field intensities immediately behind the mask, and the lower part the light intensities in the image plain. (It should be noted, that the simulation includes a mirroring of the image around the optical axis by the objective lens.) Clearly the contrast afforded by the phase shift mask (b) is largest, exceeding contrasts of other mask types by at least a factor of 2. A simulation of a mask of greater complexity is shown on Fig. 8. As in Fig. 7, on the top the electric field intensity behind the mask is shown, light intensity in the image plain for a phase mask (center), and the light intensity for the same mask geometry without phase shifting (bottom). It can be observed that imaging of the phase mask retains greater detail of the mask geometry than is the case with standard mask, however, the image is distorted. At present FOLIS is incapable of simulating and optimizing image light intensities for a masks with random pattern geometry, which would, presumably, model a real life mask more closely.

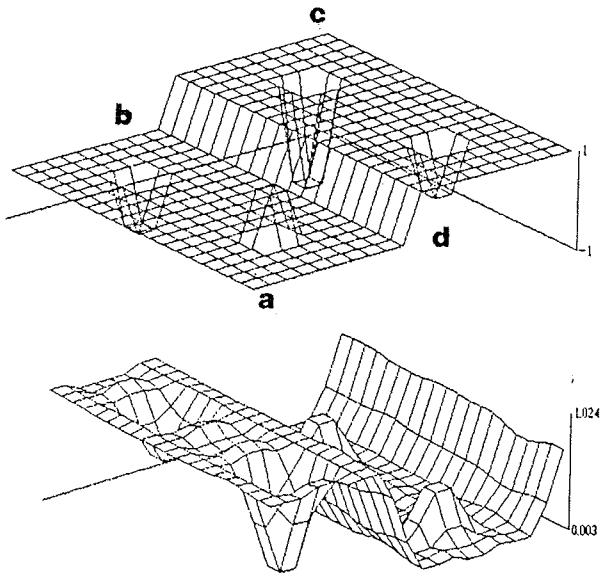


Fig. 7: Via imaging by different masks: (a) dark field, (b) light field, (c) 180 deg. phase shift plate with dimensions of the via on light field, (d) dark field + phase shift, as simulated by FOLIS.

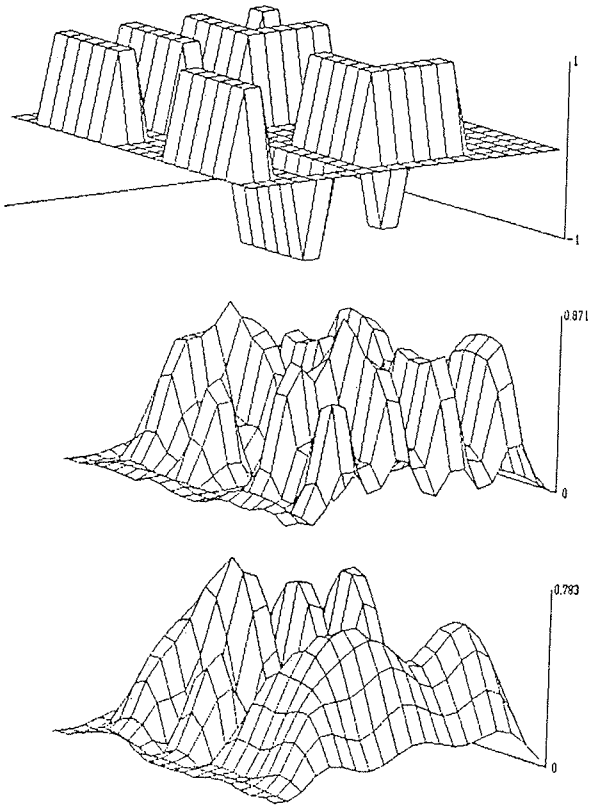


Fig. 8: Via imaging for a »random« phase mask. (a) electric field behind the mask, (b) aerial image at focal plane, (c) image of standard mask of same geometry, as simulated by FOLIS.

Conclusion

The FOLIS simulation tool has been developed to augment the introduction of projection alignment techniques into the Microelectronics Laboratory of the Faculty of Electrical Eng., University of Ljubljana. An important development goal has been to provide a means of demonstrating the relatively complex photolithographic topics and process to the faculty students. At present the FOLIS program package is not being further developed but improvements, indicated in the text, are planned for the future.

Acknowledgements

Part of the FOLIS development work has been performed as partial fulfillment of requirements for an M. Sc. degree of the first author. The work has been supported by a grant

from the Ministry of Education, Science, and Sports of the Republic of Slovenia. The use of IMP (San Jose, Ca., USA) facilities for some of the work presented is gratefully acknowledged.

References:

/1./ F.H. Dill, Optical Lithography, IEEE Trans. Elec. Dev., Vol. ED-22, 1975

/2./ J.D. Plummer et al., Silicon VLSI Technology, Prantice Hall, Upper Saddle River, N.J., USA, 2000

/3./ E.Hecht and A. Zajac, Optics, second printing, Addison-Wesley Publishing Co., Reading, MA. USA, 1977

/4./ A. Offner, Wavelength and Coherence Effects on the Performance of Real Optical Projection Systems, Photographic Sc. Eng., Vol. 23, No. 6, Nov./Dec., 1979

/5./ C.A. Mack, Analytical Expression for the Standing Wave Intensity in Photoresist, Appl. Optics, Vol. 25 1986

/6./ F.H. Dill et al., Modeling Projection Printing if Positive Photoresists, IEEE Trans. Elec. Dev., Vol. ED-26, 1975

/7./ M.D. Levenson et al., Improving Resolution in Photolithography with a Phase Shifting Mask, IEEE Trans. Elec. Dev., Vol. ED-29, 1982

Igor Macarol
Skupina Atlantis d.o.o.,
Parmova 51, 1000 Ljubljana, Slovenia
e-mail: igor.macarol@atlantis.si

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

Prispelo (Arrived): 21.06.2002 *Sprejeto (Accepted): 25.03.2003*

IZVEDBA PARAMETRIČNEGA NELINEARNEGA FILTRA ZA ISKANJE KOŽNIH ZNAČNIC V DIGITALNI SLIKI Z FPSLIC TEHNOLOGIJO

Iztok Kramberger, Zdravko Kačič

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

Ključne besede: digitalni filter, digitalna slika, koža, barva, digitalno procesiranje signalov, FPGA, mikrokrmilnik, FPSLIC, PAL

Izvleček: Članek obravnava razvoj matematičnega modela in strojne izvedbe parametričnega digitalnega filtra za iskanje kožnih značnic v digitalni sliki. Prikazana je implementacija z FPSLIC tehnologijo podjetja Atmel, ki predstavlja učinkovito združenje FPGA programabilnih logičnih celičnih vezij in RISC mikrokrmilnikov. Predstavljen parametrični digitalni filter omogoča filtriranje digitalne slike zapisane v YCbCr barvnem prostoru in na svojem izhodu daje digitalno sliko v obliki binarnih mask. Primarni namen uporabe danega digitalnega filtra je izločanje odtenkov kožne barve v digitalni sliki, kar primarno predstavlja predprocesiranje slike v algoritmih detekcije ali sledenja kožnih formacij. Ob tem je digitalni filter parametrično zasnovan in tako omogoča izločanje barvnih odtenkov v širšem pomenu. Struktura filtra je primerna za izvajanje operacij na nivoju slikovnih točk, ki ne vključujejo časovno povezanih opravil. Predstavljena arhitektura digitalnega filtra s hitrostjo izračunavanja izhodnih vzorcev zadovoljuje potrebe procesiranja digitalne slike v realnem času za standardne PAL slikovne formate v polni ločljivosti. Glede na odprte možnosti perifernih enot FPSLIC integriranih vezij, nudi izvedeni digitalni filter enostavno povezovanje z obstoječimi video dekodirniki in razvojnimi orodji za digitalno procesiranje signalov.

Implementation of Parametrical Nonlinear Digital Filter for Skin Features Identification in Digital Image Using FPSLIC Technology

Key words: digital filter, digital image, skin color, digital signal processing, FPGA, microcontroller, FPSLIC, PAL

Abstract: This article is about development of mathematical model and hardware implementation of the digital filter for skin features identification. With increasing popularity of digital video the computer vision is becoming one of the common media in mainstream electronics. One of the general tasks that are required by consumer computer vision systems is image understanding like detection and tracking of people. Before those tasks can be performed, the images have to be preprocessed. Suchlike tasks can be very cumbersome for general purpose computers. In such cases the skin color can be a very comprehensive feature to achieve the goal. The core of color tracking is color based image segmentation. Introduced parametrical digital filter makes it possible to filter digital images in YCbCr color space and on its output it gives digital images in a binary mask fashion. As the primary purpose of this digital filter is to identify skin color tones in digital images that is the main preprocessing task in detection and tracking of skin color formations, the parametrical design of the filter also offers color detection at comprehensive sense. In contrast to other classical digital filters known for digital audio and video signal processing with convolution equation, this filter features a multitude of threshold functions as transfer function. Appropriated thresholds are defined in HSV color space for the value and saturation components where the hue attribute is reliable. Therefore we use a polyhedron with appropriate threshold values that correspond to the skin-colored clusters with well-defined saturation and value components, based on a large sample set /16/. Each pixel from the digital image that passes through the filter is labeled with appropriate binary information. This binary information creates a binary mask with the same size as the original image where a logical one points to skin-colored pixel. As the applicable level of shown digital filter extends to digital image preprocessing the low level hardware implementation is very suitable. The hardware implementation is shown with FPSLIC technology from Atmel that features high performance combination of FPGA programmable logic circuits and RISC microcontrollers. The given digital filter structure is suitable for low-level non-temporal preprocessing tasks that operate on pixel level. Shown architecture of the digital filter offers sufficient computing power or data bandwidth to satisfy the requirements for full resolution real-time image processing in standard PAL picture formats. For hardware implementation we developed a mathematical model that enables calculation of the digital filter parameters from given boundary values in HSV color space. The digital filter hardware is completely designed in FPGA part of FPSLIC circuit. The embedded FPSLIC RISC microcontroller features parameter computation within the given mathematical model. The appropriate boundary threshold values for skin color identification in HSV color space are passed to the filter through asynchronous serial port and afterwards used for parameter computation. After calculation the computed parameters are directly applied to the digital structure of the filter by the embedded microcontroller. For parameter storage we use 8-bit wide registers that are implemented in the FPGA part of the FPSLIC. Such digital structure enables dynamical changing of the filter parameters as they are computed by the embedded microcontroller and their computation doesn't disturb the digital image filtering which is done by the FPGA logic. As it is shown in the given mathematical model we have expose the polyhedron threshold values to several conditional states that are computed by three logical processing units. We have shown that six conditional states are needed to delineate the given threshold polyhedron in HSV color space. Each of those three processing units is performing two parameters computation so two of conditional states are processed by each unit. Through the parallel structure we get six computed conditional states for each image pixel that are gathered together for the binary pixel label on the filter output by a simply logical function. As FPSLIC integrated circuits utilize many peripheral units with open connection possibilities it is simple to interface the digital filter to other devices like video decoders and development tools for digital signal processing. Through simulation and testing we found that the performance of digital filter mostly depends on lightning conditions of the captured scene. In those cases where lightning conditions are worse the value threshold in HSV color space or the brightness parameters have to be appropriately corrected to increase the performance of the filter in identifying skin-colored features in digital image. In the future we will implement automatic tracing of the lightning conditions in captured scene and re-computation algorithm for specific brightness depending parameters.

1. Uvod

Z naraščajočo popularnostjo digitalnih video kamer postaja elektronski vid eden izmed splošnih medijev potrošniške elektronike. Med splošna opravila, ki jih zahtevajo z elektronskim vidom podprti potrošniški sistemi, spada tudi razumevanje slike, kot na primer detekcija in sledenje ljudi. Preden je možno takšna opravila dejansko izvesti, je potrebno slike predprocesirati. Takšna naloga je lahko izjemo zahtevno opravilo za splošno namenske računalnike. Ne le zaradi računske zapletenosti, temveč tudi zaradi velikih podatkovnih hitrosti.

Barva kože v takšnih primerih nudi primaren način za dosego omenjenega cilja. Jedro barvnega sledenja je na barvi osnovana segmentacija. Glede na predstavitev barvne razdelitve v določenih barvnih modelih, lahko trenutno uporabljene tehnike barvnega sledenja uvrstimo v dva splošna pristopa: parametrični /11, 9/ in ne-parametrični /3, 2/.

Sistem detekcije in sledenja odtenkov kožne barve vključuje modul predprocesiranja, ki omogoča izločanje odtenkov barve kože v živem slikovnem zaporedju. Uporabljena paralelna arhitektura je še posebej primerna za tiste naloge predprocesiranja, ki se izvajajo na nivoju slikovnih točk. Arhitektura je tako uporabna pri elektronskem vidu na nizkem nivoju procesiranja, ki ne vključuje časovno povezanih opravil, saj sistem ne vsebuje slikovnega pomnilnika in je časovno propusten.

Predstavljen digitalni filter za razliko od klasičnih digitalnih filtrov, ki jih poznamo s področja procesiranja zvočnih in slikovnih signalov, kot prevajalno funkcijo /39/ ne uporablja konvolucijske enačbe, temveč množico pragovnih funkcij, ki izhajajo iz opravljene statistične obdelave predhodno zajetih slikovnih vzorcev /16/. Tako gre v našem primeru za označevanje slikovnih točk v obliki binarnih mask, kjer vsaka točka dobljene binarne maske vsebuje informacijo o prisotnosti barve kože. Torej lahko za vsako slikovno točko vhodnega slikovnega toka določimo ustreznost opisnim pogojem odtenkov kožne barve, pri tem se ta označi z logičnim stanjem na izhodu filtra. Ker izhodno slikovno zaporedje v obliki binarnih mask dimenzijsko ustreza vhodnemu slikovnemu zaporedju, lahko govorimo o slikovnem filtru.

Ker aplikativni nivo uporabe digitalnega filtra sega na področje predprocesiranja živega slikovnega zaporedja, ga je smotrno realizirati na logičnem strojnem nivoju. Pri sami izvedbi logičnega stroja smo uporabili *FPSLIC* tehnologijo podjetja *Atmel*. *FPSLIC* tehnologija predstavlja programabilna logična vezja, ki vključujejo vgrajen *RISC* procesor /28/.

2. Identifikacija odtenkov kožne barve

Identifikacijo in sledenje kožno obarvanih regij določamo z uporabo predhodnega poznavanja porazdelitve barve kože v *HSV* barvnem prostoru. Ugotovljeno je bilo, da odtenki kožne barve formirajo precej dobro definirana področja v *HSV* barvnem prostoru /15/. V tem prostoru leži namreč porazdelitev kožnih barvnih tonov dominantno v omejenem območju barvnih odtenkov rdeče in rumene barve ter v določenih primerih v območju škrlatne in rdeče barve za temnejše tipe kože /16/. Komponenta nasičenosti teh formacij nakazuje, da so kožne barve do neke mere nasičene, vendar ne globoko nasičene s spreminjajočim nivojem osvetlitve. Komponenta barvnih odtenkov je najbolj pomembna lastnost za določanje karakteristik kožnih formacij, vendar so lahko barvni odtenki dokaj nezanesljivi v primerih nizke osvetlitve scene ali nizke vrednosti nasičenja na področjih kožnih formacij /16/. Zato je potrebno definirati primerne prage za komponenti vrednosti osvetlitve in nasičenja v primerih, kjer je ustrezen tudi atribut barvnih odtenkov.

V skladu s kožno obarvanimi formacijami iz podatkovne baze zajetih slikovnih vzorcev, je možno zapisati ustrezen polieder s primernimi vrednostmi pragov, ki imajo ustrezno definirani komponenti nasičenja in vrednosti osvetlitve:

$$p(k)=\begin{cases}1, & \text{če je } 340° \geq H(k) \leq 50°, \\ & S(k) \geq 50\%,\ V(k) \geq 35\% \\ 0, & \text{drugje}\end{cases} \qquad (1.1)$$

Polieder velja za poljubno lokacijo točke *p* v sliki, kjer je *k* trenutna koordinata točke.

Kot je iz naštetih pogojev razvidno, je problem dokaj jasno definiran v *HSV* barvnem prostoru, vendar je smotrno problem zasnovati tako, da bo zadoščal uporabi potrošniških slikovnih senzorjev - kamer. Pri večini potrošniških kamer je slikovni tok zapisan v *YCbCr* barvnem prostoru in iz tega razloga so potrebni dodatni računski postopki, ki preslikajo pogoje iz prostora *HSV* v prostor *YCbCr*.

3. Izpeljava matematičnega modela

Kot smo omenili v poglavju 2 je binarna prevajalna funkcija /39/ digitalnega filtra podana v obliki poliedra v *HSV* barvnem prostoru. Zapisan polieder je neparametričnega tipa, saj so pogoji natančno določeni, vendar je v praktični uporabi mnogokrat zahtevano določeno odstopanje, ki ga je nujno potrebno upoštevati na samem začetku snovanja matematičnega modela. Zato smo se odločili za parametrično obliko filtra, ki omogoča spreminjanje pogojev danega poliedra.

Zaradi lažjega prehoda iz *HSV* barvnega prostora v *YCbCr* barvni prostor smo vpeljali še *RGB* barvni prostor, ki zaradi svoje aditivne lastnosti na eni strani omogoča lažje razumevanje pogojev danega poliedra kožnih odtenkov v HSV prostoru in na drugi strani ponuja enostaven prehod v *YCbCr* barvni prostor.

3.1 Predstavitev barvnih prostorov

Človeško oko ima v očesni mrežnici tri tipe celic receptorjev, ki so občutljive na tri različna območja valovnih dolžin elektromagnetnega valovanja. Svetloba povzroča stimulacijo teh celic, ki posledično proizvajajo električne impulze, le-ti pa potujejo do možganov, ki na podlagi jakosti impulzov zaznajo določeno barvo.

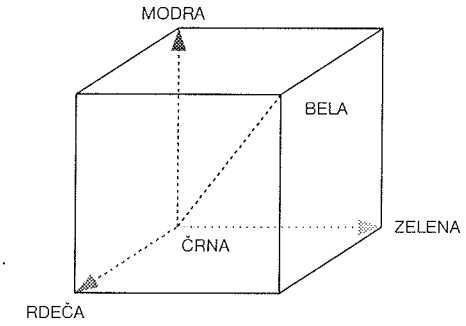
Barvo lahko predstavimo z večimi modeli, ki jim pravimo barvni prostori. Posamezni barvni prostori so prilagojeni različnim tipom aplikacij (računalniška grafika, televizija, video oprema, ...). Zelo pomembno vlogo pri izbiri primernega barvnega prostora igra tip operacij, ki jih moramo izvršiti nad barvno informacijo. Določene operacije so lahko namreč nekemu barvnemu prostoru "naravne" in jih je zato moč lažje in hitreje izvesti. To pomeni, da lahko neko barvno informacijo, ki je zapisana v izvornem barvnem prostoru, pretvorimo v drug prostor, v katerem izvršimo obdelavo in nato pretvorimo informacijo nazaj v prvotni barvni prostor.

3.1.1 Barvni prostor YCbCr

YCbCr barvni prostor je bil razvit za namene prenosa in obdelave video in televizijskega signala. Komponenta *Y* predstavlja intenziteto in nosi vso potrebno informacijo za prikaz slike v črno-beli tehniki. Celotno barvno informacijo je možno interpretirati z uporabo preostalih dveh komponent kromatičnosti *Cb* in *Cr*. Prvotna ideja *YCbCr* barvnega prostora izhaja iz ugotovitve, da so človeške oči občutljivejše na zeleno kot na rdečo ali modro barvo /33/. Glede na razmerja občutljivosti človeškega očesa vsebuje tako kompozitna kot tudi komponentna oblika analognega signala 70 % informacije o zeleni, 20 % informacije o rdeči in 10 % o modri barvi. Glede na dana razmerja so zasnovani tudi slikovni senzorji, ki najpogosteje uporabljajo Bayerjevo razporeditev slikovnih elementov - točk. Bayerjev vzorec uporablja dvakrat več zelenih kot rdečih in modrih elementarnih detektorjev. Iz omenjenih razlogov tako analogni signal kot digitalno slikovno zaporedje vsebuje dvakrat več informacije o intenziteti (pretežno zelena barva), kot o kromatičnosti (rdeča in modra barva). Pri pretvorbi analogne v digitalno obliko živega slikovnega zaporedja se slikovni tok vzorči glede na dana razmerja po vzorčni shemi 4:2:2, kar pomeni, da vsakemu vzorcu intenzitete izmenično pripadata kromatični vzorec modre in rdeče barve. Za nadaljnjo obdelavo slikovnega toka je potrebno vzorčno shemo pretvoriti iz razmerja 4:2:2 v razmerje 4:4:4, kjer vsakemu vzorcu intenzitete slikovnega toka pripadata kromatični vzorec rdeče in modre barve /34/.

3.1.2 Barvni prostor RGB

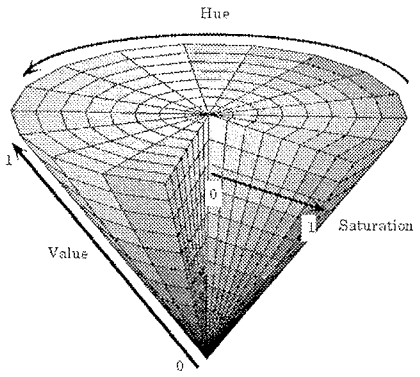
RGB barvni prostor je posebej primeren za uporabo v računalniški grafiki. Barva je določena s tremi komponentami: *R* (rdeča), *G* (zelena) in *B* (modra). Na sliki 1 vidimo, da je RGB barvni prostor predstavljen v obliki tridimenzionalnega kartezijskega koordinatnega sistema. Poljubna barva je določena kot vsota vektorjev osnovnih komponent *RGB*. Za prenos slikovne informacije v *RGB* prostoru je potrebna večja podatkovna širina.



Slika 1: RGB *barvni prostor.*

3.1.3 Barvni prostor HSV

HSV barvni prostor je zelo podoben človekovi percepciji barv. Barvna informacija je v *HSV* prostoru določena s komponentami: *H* - hue (barvni odtenek), *S* - saturation (nasičenje), *V* - value (vrednost osvetlitve). Na sliki 2 je prikazan model barvnega prostora HSV v obliki stožca. Različni barvni toni H so opazni na horizontalnem preseku stožca in sespreminjajo v odvisnosti odkota med izhodiščnim barvnim tonom 0 (rdeča barva) in želenim barvnim tonom. Nasičenje *S* narašča od vrednosti 0 do vrednosti 1 v smeri od središča proti obodu. Osvetljenost *V* pa narašča po vertikalni osi od vrednosti 0 (črna barva) do vrednoti 1 (bela barva). *HSV* barvni prostor je zaradi svoje sorodnosti s človekovo percepcijo barv izjemno primeren za aplikacije elektronskega vida.



Slika 2: HSV *barvni prostor.*

3.2 Matematični model

V uvodnem poglavju smo že omenili, da bo vhodni slikovni tok digitalnega filtra zapisan v *YCbCr* prostoru, med tem ko bo izhodni slikovni tok predstavljen na nivoju binarnih mask. Tako lahko računsko operacijo filtra v splošnem zapišemo z izrazom:

$$M(t) = T(I(t)), \tag{3.1}$$

kjer je *T* prevajalna funkcija filtra, *I* slika vhodnega slikovnega zaporedja in *M* binarna maska izhodnega slikovnega zaporedja v času *t*.

Ker filter izvaja prevajalno funkcijo v obliki pragovnih funkcij na vsaki posamezni slikovni točki vhodnega slikovnega

zaporedja in pri operaciji filtriranja ni medsebojne odvisnosti med posameznimi slikovnimi točkami, lahko matematično formulacijo prenesemo na nivo posamezne slikovne točke v obliki zapisa:

$$m(k,t) = T(i(k,t)), \tag{3.2}$$

kjer je *T* prevajalna funkcija filtra, *i* točka vhodne slike iz slikovnega zaporedja in *m* točka binarne maske izhodnega slikovnega zaporedja s koordinatami *k* v času *t*.

Če upoštevamo zahteve procesiranja v realnem času, lahko ugotovimo, da moramo operacijo prevajalne funkcije izvesti vsaj tolikokrat, kot to pogojujejo dimenzije vhodnega slikovnega toka. Zaradi narave živega slikovnega zaporedja je čas, ki je na voljo za enkraten izračun operacije prevajalne funkcije podan z izrazom:

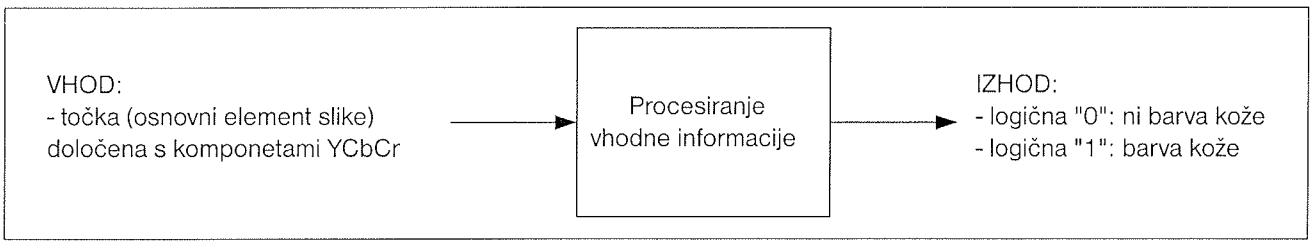
$$t_k = \frac{1}{m \cdot n \cdot f}, \tag{3.3}$$

kjer sta *m* in *n* dimenziji vhodne slike v vodoravni in navpični smeri in *f* frekvenca slik živega slikovnega zaporedja.

Ker moramo pri snovanju filtra upoštevati zahteve procesiranja slikovnega toka v realnem času, je formulacija filtra podana z izrazom:

$$m(k) = T(i(k)).$$

Tako lahko delovanje filtra ponazorimo s sliko 3.



Slika 3: Princip delovanja filtra.

Na sliki 4 je grafično prikazano območje barv v *HSV* barvnem prostoru, ki ustreza barvi kože po pogojih predstavljenih v poliedru prevajalne funkcije filtra, pri čemer so meje pragovnih funkcij parametrično zapisane.

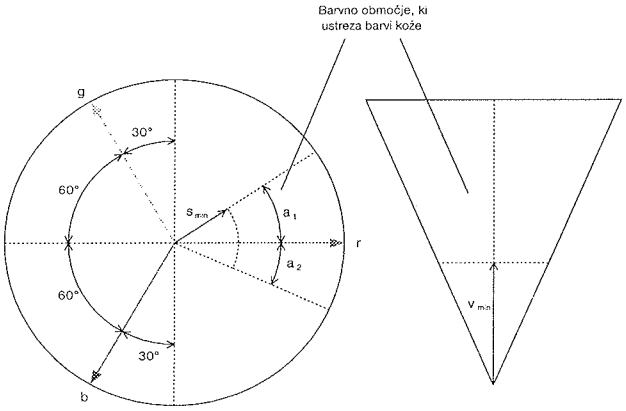
Polieder prevajalne funkcije filtra lahko v skladu s sliko 4 razdelimo na sledeče pragove

$$V \geq v_{\max} \tag{3.4}$$

$$-\alpha_2 \leq H \leq \alpha_1 \tag{3.5}$$

$$S \geq s_{\max} \tag{3.6}$$

pri tem so ti zapisani s parametri v_{max}, α_1, α_2 in s_{max}, ki določajo meje področja, na katerem bo filter na svojem izhodu dajal pozitiven rezultat.



Slika 4: Pogoji v HSV prostoru.

Če barvne vrednosti v *RGB* barvnem prostoru zapišemo kot vektorje v komponentni obliki

$$\vec{r} = (r_x, r_y, r_z), \vec{g} = (g_x, g_y, g_z), \vec{b} = (b_x, b_y, b_z), \tag{3.7}$$

je iz slike 4 razvidno, da lahko z ustrezno postavitvijo *HSV* barvnega prostora v *RGB* barvni prostor, vrednosti posameznih komponent zapišemo kot

$$r_x = R,\ r_y = 0,\ r_z = \frac{1}{\sqrt{3}}R, \tag{3.8}$$

$$g_x = -\frac{1}{2}G,\ g_y = \frac{\sqrt{3}}{2}G,\ g_z = \frac{1}{\sqrt{3}}G, \tag{3.9}$$

$$b_x = -\frac{1}{2}B,\ b_y = -\frac{\sqrt{3}}{2}B,\ b_z = \frac{1}{\sqrt{3}}B. \tag{3.10}$$

Tako lahko s pomočjo barvnih vrednosti *RGB* barvnega prostora zapišemo izraze za nasičenje *S*, barvni odtenek *H* in vrednost osvetlitve *V* v *HSV* barvnem prostoru:

$$S = \sqrt{(r_x + g_x + b_x)^2 + (r_y + g_y + b_y)^2} = \\ = \sqrt{(R-G)^2 + (R-B)(G-B)}, \tag{3.11}$$

$$H = a\tan\frac{r_y + g_y + b_y}{r_x + g_x + b_x} = a\tan\frac{\frac{\sqrt{3}}{2}(G-B)}{R - \frac{1}{2}(B-G)}, \tag{3.12}$$

$$V=\sqrt{r_z^2+g_z^2+b_z^2}=\sqrt{\frac{R^2+G^2+B^2}{3}}, \quad (3.13)$$

Iz izraza (3.11) je razvidno, da vrednost nasičenja *S* limitira proti razliki maksimalne in minimalne komponente *RGB* trojice. Za zadovoljivo oceno vrednosti nasičenja *S*, kjer je vrednost normirana na interval [0-1], lahko izraz (3.11) poenostavimo in zapišemo:

$$S \approx \frac{\max(R,G,B)-\min(R,G,B)}{\max(R,G,B)}. \quad (3.14)$$

Podobno je iz izraza (3.13) razvidno, da vrednost osvetlitve *V* limitira proti maksimalni komponenti *RGB* trojice. Tako lahko izraz (3.13) poenostavimo in oceno vrednosti osvetlitve *V* zapišemo kot

$$V \approx \max(R,G,B). \quad (3.15)$$

Ker aktivno področje filtra leži v okolici rdeče barve, lahko izraz za vrednost osvetlitve *V* dodatno poenostavimo

$$V \approx R. \quad (3.16)$$

Iz tega sledi, da lahko prvi prag (3.4) poliedra prevajalne funkcije zapišemo kot

$$c_1: R \geq v_{\max} \quad (3.17)$$

Iz slike 4 je prav tako razvidno, da ob upoštevanju pogoja c_1 (3.16 - 3.17) velja za barvne odtenke v prvem kvadrantu pogoj

$$c_2: g-b \geq 0 \quad (3.18)$$

in barvne odtenke v četrtem kvadrantu pogoj

$$c_2': b-g \geq 0 \quad (3.19)$$

Tako lahko za območji v prvem in četrtem kvadrantu zapišemo izraz za nasičenje *S* (3.14) z izrazoma

$$S=\frac{R-G}{R} \text{ in } S=\frac{R-B}{R}. \quad (3.20)\,(3.21)$$

Glede na drugi prag poliedra prevajalne funkcije (3.6) lahko zapišemo pogoja nasičenja *S* za prvi in četrti kvadrant kot

$$c_3: R(1-s_{\max})-G \geq 0 \text{ in} \quad (3.22)$$

$$c_4: R(1-s_{\max})-B \geq 0. \quad (3.23)$$

Prav tako lahko s pomočjo pogojev (3.18) in (3.19) preuredimo drugi prag poliedra prevajalne funkcije (3.5) in ga zapišemo s pogojema za barvni odtenek v prvem in četrtem kvadrantu kot

$$c_5: \quad R\tan\alpha_1+G(-\frac{1}{2}\tan\alpha_1-\frac{\sqrt{3}}{2})+B(-\frac{1}{2}\tan\alpha_1+\frac{\sqrt{3}}{2})\geq 0 \quad \text{in} \quad (3.24)$$

$$c_6: \quad R\tan\alpha_2+G(-\frac{1}{2}\tan\alpha_2+\frac{\sqrt{3}}{2})+B(-\frac{1}{2}\tan\alpha_2-\frac{\sqrt{3}}{2})\geq 0 \quad . \quad (3.25)$$

Aktivno področje filtra lahko torej razdelimo na področje v prvem in četrtem kvadrantu. Če je pogoj (18) izpolnjen, določajo aktivno področje pogoji (3.17), (3.23) in (3.24). V obratnem primeru določajo aktivno področje pogoji (3.17), (3.22) in (3.25). Z namenom hitrejšega izvajanja je smiselno preverjati vse pogoje hkrati, pri tem lahko z dodatno logiko ugotavljamo ali smo v aktivnem področju filtra oziroma izven njega. Pogoje lahko zberemo z zapisom

$$n_{i1}\cdot R+n_{i2}\cdot G+n_{i3}\cdot B \geq p_i, \quad (3.26)$$

kjer n_i predstavljajo parametre filtra v *RGB* barvnem prostoru pri pogojnih vrednostih p_i, *i* pa predstavlja indeks posameznega pogoja. Parametre filtra iz *RGB* barvnega prostora pretvorimo v *YCbCr* barvni prostor s pomočjo transformacijske matrike *K* /34/:

$$\begin{bmatrix} m_{i1} \\ m_{i2} \\ m_{i3} \end{bmatrix} = K^T \cdot \begin{bmatrix} n_{i1} \\ n_{i2} \\ n_{i3} \end{bmatrix}, \quad (3.27)$$

pri tem m_i predstavljajo parametre filtra v *YCbCr* barvnem prostoru in *i* indeks posameznega pogoja. Parametre filtra lahko s pomočjo danega matematičnega modela izračunamo iz podanih mejnih vrednosti poliedra prevajalne funkcije. Izhodno funkcijo filtra lahko zapišemo z logičnim izrazom:

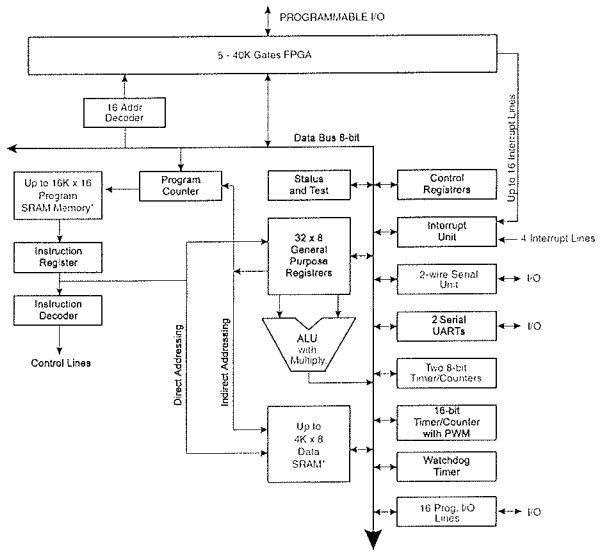
$$o=(c_1\cdot c_2\cdot c_3\cdot c_5)+(c_1\cdot c_2'\cdot c_4\cdot c_6)= =(c_1\cdot c_2\cdot c_3\cdot c_5)+(c_1\cdot \bar{c}_2\cdot c_4\cdot c_6), \quad (3.28)$$

kjer so c_i logični rezultati posameznih pogojev in *o* logična vrednost na izhodu filtra. Ker sta pogoja c_2 (3.18) in c_2' (3.19) komplementarna, je potrebno izračunati le pogoj c_2 in v logičnem izrazu izhodne funkcije namesto c_2' upoštevati njegovo negirano vrednost $\bar{c}_2$.

4. Izvedba digitalnega filtra

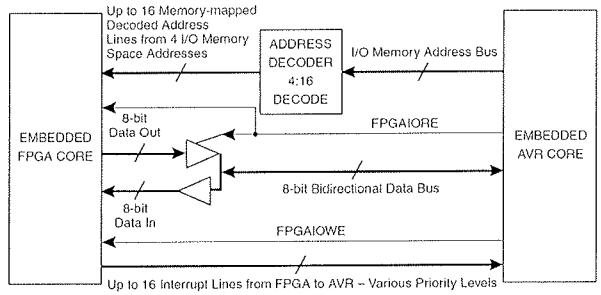
Za izvedbo digitalnega filtra smo izbrali *FPSLIC* tehnologijo podjetja *Atmel*, ki prihaja v obliki integriranih vezij *AT94K*. *FPSLIC* tehnologija predstavlja *FPGA* rekonfigurabilna logič-

na vezja z vgrajenim *RISC AVR* mikrokrmilnikom /28/. Konfiguracija logičnega vezja in vgrajenega mikrokrmilnika je osnovana na statičnem pomnilniku in jo je potrebno naložiti iz zunanjega pomnilnika.



Slika 5: Arhitektura integriranega vezja AT94K.

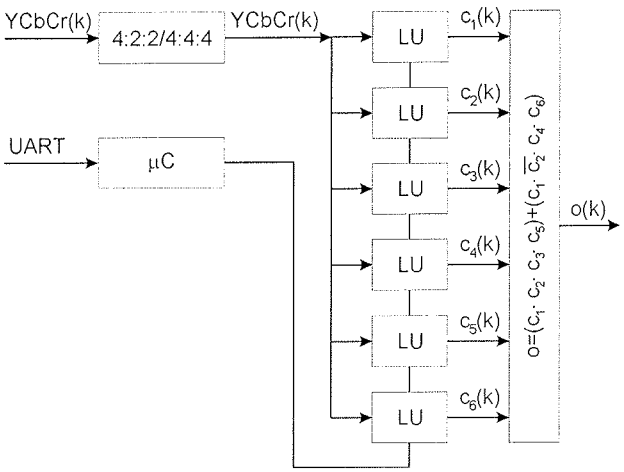
Slika 5 prikazuje notranjo zgradbo *FPSLIC* integriranih vezij, na sliki 6 pa je podrobneje prikazan vmesnik med *FPGA* jedrom in mikrokrmilnikom.



Slika 6: Povezava FPGA *jedra in mikrokrmilnika.*

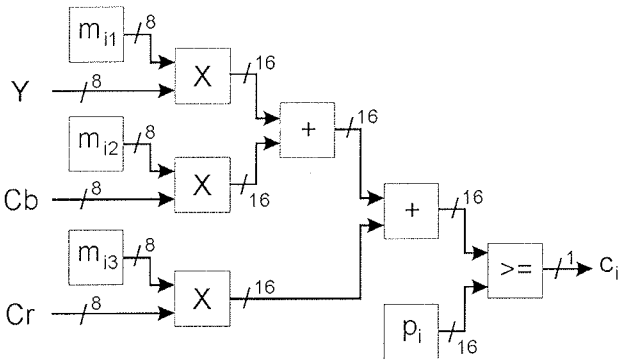
Za samo izvedbo filtra smo uporabili integrirano vezje *AT94K40L*, ki vsebuje *FPGA* polje velikosti 48 x 48 programabilnih logičnih celic. Vgrajen 8-bitni *RISC* mikrokrmilnik lahko konfiguriramo tako, da ima na voljo od 20KB do 32KB programskega pomnilnika in od 4KB do 16KB podatkovnega pomnilnika.Vgrajen mikrokrmilnik ponuja uporabo perifernih enot, kot sta dva zaporedna asinhrona vmesnika *UART*, tri časovnike/števce in uro realnega časa.

Digitalni parametrični filter je v celoti zasnovan v *FPGA* logičnem vezju, pri tem za izračun parametrov skrbi vgrajen mikrokrmilnik. Ta je preko asinhronega zaporednega vmesnika povezan z osebnim računalnikom. Na ta način lahko z osebnim računalnikom spreminjamo mejne pogoje prevajalne funkcije digitalnega filtra, kar je v času preizkušanja zelo koristno. Če se mejni pogoji prevajalne funkcije filtra spremenijo, vgrajen mikrokrmilnik izračuna nove parametre za vgrajene pogoje in jih zapiše v logično strukturo filtra.



Slika 7: Vzporedna zgradba digitalnega filtra (LU – logične enote).

Na sliki 7 je prikazana zgradba digitalnega filtra, iz katere je razvidno, da se po pretvorbi vzorčne sheme 4:2:2 v 4:4:4 za vsako slikovno točko izračunajo pogoji $c_1(k)$-$c_6(k)$ iz katerih logična funkcija poda binarno vrednost $o(k)$ na izhodu filtra. Hkraten izračun pogojev vršijo logične enote *LU*, ki na svojem vhodu prejemajo slikovne podatke v *YCbCr* barvnem prostoru in na svojem izhodu podajajo binarno vrednost glede na izpolnjenost določenega pogoja. Vgrajen mikrokrmilnik na eni strani omogoča sprejemanje mejnih vrednosti prevajalne funkcije in na drugi strani le-te implicira v obliki parametrov v logične enote *LU*.



Slika 8: Splošna zgradba LU *enote.*

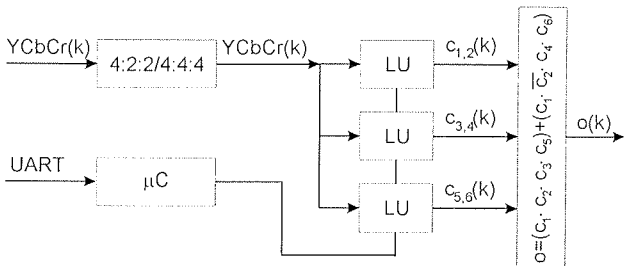
Za izračun pogojev smo zasnovali logično enoto *LU*, ki jo prikazuje slika 8. Na sliki je prikazana splošna logična enota, s katero je možno rešiti poljubnega izmed navedenih pogojev. Matematično funkcijo *LU* lahko zapišemo kot skalarni produkt vektorjev

$$c_i \Rightarrow [m_{i1} \quad m_{i2} \quad m_{i3}] \cdot [Y \quad Cb \quad Cr]^T \geq p_i. \qquad (4.1)$$

LU na svojem vhodu sprejema slikovne podatke v *YCbCr* barvnem prostoru, pri tem je vsaka posamezna komponenta

zapisana z 8-bitno podatkovno širino, kar v skupnem pomeni 24-bitno barvno globino. Komponenta *Y* je podana na intervalu vrednosti [0,255], med tem ko sta komponenti *Cb* in *Cr* podani na intervalu [-128,127]. Vsako komponento vhodnega podatka pomnožimo z ustreznim parametrom m_{ij}, ki je prav tako zapisan z 8-bitno podatkovno širino. Uporabljeni množilniki izvajajo predznačeno celoštevilčno množenje. S pomočjo seštevalnikov tvorimo delno vsoto, ki jo primerjamo s parametri p_i. Primerjanje izvaja primerjalnik, ki na svojem izhodu podaja binarno informacijo o izpolnjenosti pogoja c_i. Parametri m_{ij} in p_i so izvedeni v obliki registrov, do katerih lahko dostopa vgrajen mikrokrmilnik in vanje vpiše ustrezne vrednosti parametrov.

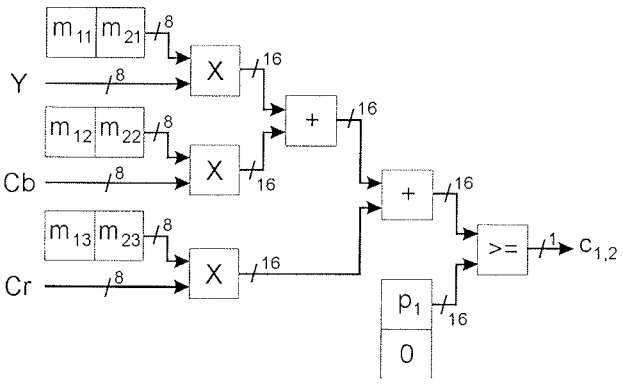
Ker realizacija predstavljene vzporedne zgradbe digitalnega filtra presega število razpoložljivih *FPGA* logičnih celic izbranega *FPSLIC* vezja, smo se odločili za kombinirano vzporedno zaporedno zgradbo digitalnega filtra, kot kaže slika 9.



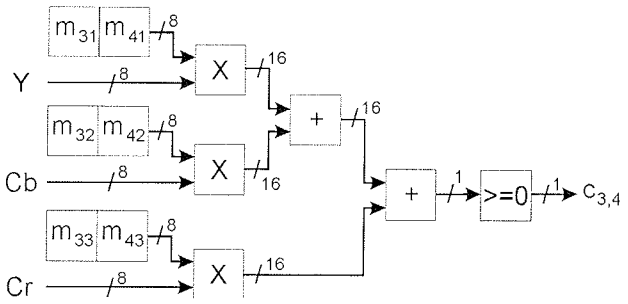
Slika 9: Vzporedno zaporedna zgradba digitalnega filtra.

V zgradbi filtra na sliki 9 vsaka *LU* izmenjujoče izračunava po dva pogoja, pri tem se hkrati izračunavajo trije pogoji. Na ta način upade število potrebnih *FPGA* logičnih celic približno na polovico v primerjavi z zgradbo na sliki 7. Ob tem se poveča zahtevnost po hitrosti *LU*, saj mora ta za vsako slikovno točko izračunati dva pogoja v danem časovnem intervalu.

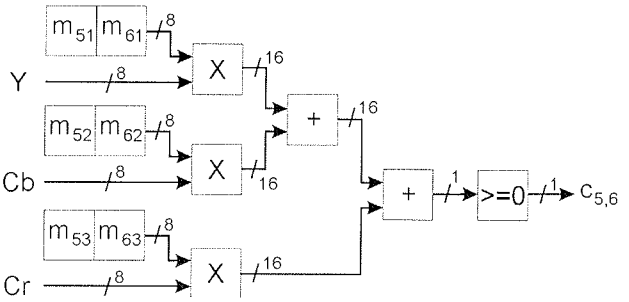
Iz matematičnega modela je razvidno, da med primerjalnimi parametri p_i, le p_1 zavzema pozitivno vrednost, med tem ko so ostali primerjalni parametri $p_2 ... p_6$ enaki 0. Iz tega razloga je možno poenostaviti izhodni primerjalni del *LU* enot, saj je smiselno spremljati le najbolj uteženi bit zadnjega seštevalnika.



Slika 10:Zgradba LU *za pogoja* c_1 *in* c_2.



Slika 11:Zgradba LU *za pogoja* c_3 *in* c_4.

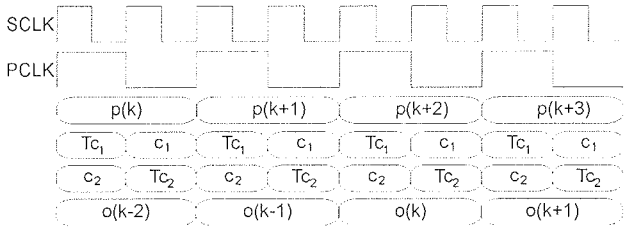


Slika 12:Zgradba LU *za pogoja* c_5 *in* c_6.

4.1 Sinhronizacija

Za delovanje v realnem času moramo digitalni filter sinhronizirati na slikovni izvor. V primeru vzporedno zaporedne zgradbe moramo delovanje filtra uskladiti z dvakrat višjo uro, kot je ura slikovnih točk. V takšnem primeru je vsaka posamezna LU sposobna za eno slikovno točko izračunati dva potrebna pogoja. Tako v prvem ciklu ure za izračun pogoja c_i uporabimo parameter m_i, v drugem ciklu ure pa za izračun pogoja c_{i+1} uporabimo parameter m_{i+1}. Tako na primer znaša v slikovnem formatu *PAL* linijska frekvenca 15.625 kHz in vsebuje skupno 944 slikovnih točk. Od teh je 768 aktivnih točk, ki dejansko nosijo informacijo o sliki. Frekvenca slikovnih točk za dan format znaša 14.75 MHz oziroma njen dvakratnik 29.50 MHz. V slikovnem formatu *PAL ITU-R BT.601* je linijska frekvenca prav tako 15.625 kHz, pri tem ta vsebuje skupno 846 slikovnih točk od katerih je 720 aktivnih. Frekvenca slikovnih točk za ta format znaša 13.50 MHz oziroma njen dvakratnik 27.00 MHz.

Da je digitalni filter sposoben obdelati vse slikovne točke v realnem času, mora hitrost izračunavanja LU enot zadostiti podanim frekvencam. Glede na dane hitrosti izračunavanja smo v enotah za izračunavanje delnih produktov in delnih vsot LU uporabili pristop cevljenja. Tako ima digitalni filter časovno zakasnitev za dve slikovni točki. V trenutku, ko je na vhodu filtra slikovni vzorec *p(k)* je na izhodu filtra vzorec *o(k-2)*. Na sliki 13 je grafično prikazano cevljenje *LU*. Signal *PCLK* predstavlja uro slikovnih točk, signal *SCLK* pa njen dvakratnik. Za izračun enega pogoja so potrebni trije cikli *SCLK* ure. Z uporabo cevljenja se izmenjujoče izračunavata dva pogoja, ki sta med seboj časovno zamaknjena za en cikel ure *SCLK*. Intervali *T* predstavljajo trenutno hranjenje rezultata izračunanega pogoja.



Slika 13: Cevljenje v LU.

Za nemoteno delovanje filtra je smotrno parametre v *LU* enotah spreminjati v času zatemnitve digitalnega video signala.

4.2 Primer izračuna filtra za iskanje kožnih značnic

Predpostavimo, da so želene mejne vrednosti prevajalne funkcije digitalnega filtra podane za področje kožne barve:

$$v_{\max} = 0{,}35,\ s_{\max} = 0{,}5,\ \alpha_1 = 50, \alpha_2 = 20. \quad (4.2)$$

V takšnem primeru bo digitalni filter na svojem izhodu dajal pozitivno vrednost le za točke slike, katerih barva ustreza barvnemu področju kože v *HSV* prostoru, ki ga opisujejo dane mejne vrednosti.

Za izračun parametrov in simulacijo digitalnega filtra smo predlagan matematični model prenesli v programski paket *MATLAB 6*.

Iz mejnih vrednosti po matematičnem modelu izračunamo parametre filtra n_{ij} za *RGB* prostor:

$$n_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, n_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, n_3 = \begin{bmatrix} 0{,}5 \\ 0 \\ -1 \end{bmatrix}, n_4 = \begin{bmatrix} 0{,}5 \\ -1 \\ 0 \end{bmatrix},$$

$$n_5 = \begin{bmatrix} 1{,}1918 \\ -1{,}4619 \\ 0{,}2701 \end{bmatrix}, n_6 = \begin{bmatrix} 0{,}364 \\ 0{,}684 \\ -1{,}048 \end{bmatrix} \quad (4.3)$$

$$p_1 = 0{,}35,\ p_2 \cdots p_6 = 0 \quad (4.4)$$

Tako dobljene parametre filtra pretvorimo s pomočjo transformacijske matrike *K* v parametre m_{ij} za *YCbCr* barvni prostor:

$$m_1 = \begin{bmatrix} 10 \\ 0 \\ 14 \end{bmatrix}, m_2 = \begin{bmatrix} 0 \\ -21 \\ -7 \end{bmatrix}, m_3 = \begin{bmatrix} -5 \\ -17 \\ 7 \end{bmatrix},$$

$$m_4 = \begin{bmatrix} -5 \\ 3 \\ 14 \end{bmatrix}, m_5 = \begin{bmatrix} 0 \\ 10 \\ 27 \end{bmatrix}, m_6 = \begin{bmatrix} 0 \\ -20 \\ 0 \end{bmatrix} \quad (4.5)$$

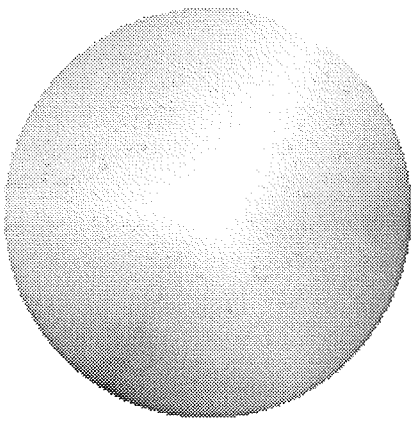
$$p_1 = 893,\ p_2 \cdots p_6 = 0 \quad (4.6)$$

Pri izračunu parametrov filtra se moramo zavedati, da registri za hranjenje parametrov v *LU* lahko sprejmejo le 8-bitne predznačene celoštevilčne vrednosti na intervalu [-128,127]. Pri tem so izjema le parametri p_i, ki so lahko 16-bitne pozitivne celoštevilčne vrednosti. Tako je potrebno vrednosti parametrov po pretvorbi iz *RGB* v *YCbCr* barvni prostor najprej ustrezno skalirati in zaokrožiti. Prikazane vrednosti v enačbah (4.5) in (4.6) so desetkratna zaokrožena števila izračunanih vrednosti.

Na sliki 14 je prikazan barvni krog v *HSV* barvnem prostoru za največjo vrednost osvetlitve *V*. Na sliki 15 je prikazano filtrirano področje slike 14, kjer so v digitalnem filtru uporabljeni predhodno navedeni parametri (4.5)(4.6) oziroma mejni pogoji prevajalne funkcije (4.2).

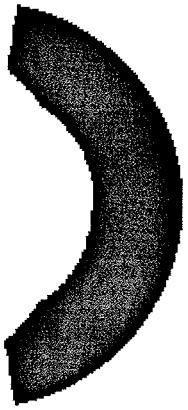
Na sliki 16 je prikazano filtrirano področje, kjer so zahtevani večji mejni pogoji za barvni odtenek *H*. Na sliki so vidne težave pri uspešnosti določanja mejnih vrednosti nasičenja *S* za večje kote barvnega odtenka *H*. Težave nastopajo zaradi kvantizacije parametrov.

Na sliki 17 je prikazana testna slika realnega slikovnega izvora in na sliki 18 njen rezultat filtriranja.



Slika 14:Barvni krog v HSV *barvnem prostoru.*

Slika 15:Filtrirano področje za kožno barvo V≥0.5, S≥0.35, 50°≥H≥340°.



Slika 16:Filtrirano področje za V≥0.5, S≥0.35, 80°≥H≥280°.

Slika 17: Testna slika realnega slikovnega izvora.

Slika 18:Filtrirana testna slika realnega slikovnega izvora.

5. Zaključek

Predstavljen parametrični digitalni filter je uporaben predvsem na področju rdečih barvnih odtenkov, saj je bil specifično zasnovan za iskanje kožnih značnic oziroma odtenkov kožne barve v digitalni sliki. Na digitalnem vhodu je sposoben sprejemati slikovni tok v najbolj razširjenem *YCbCr* barvnem prostoru, pri tem je možno z uporabo različnih transformacijskih matrik za barvne prostore, izračunati parametre tudi za druge barvne prostore, ki so zapisani s tremi komponentami. Mejne vrednosti za dani digitalni filter se podajajo v *HSV* barvnem prostoru, ki je po svoji naravi najbližji človeški percepciji barv. Skupna realizacija filtra omogoča dinamično spreminjanje binarne prevajalne funkcije, saj struktura omogoča sprotno izračunavanje parametrov med samim delovanjem filtra.

Predstavljen digitalni filter bomo uporabljali skupaj z video dekodirnikom *TVP5040* podjetja *Texas Instruments*, ki na svojem vhodu sprejema analogni S-video signal in na digitalnem izhodu podaja slikovni tok v *YCbCr* barvnem prostoru. Slikovni tok iz digitalnega filtra se preko hitrega sinhronega zaporednega vmesnika pošilja v digitalni signalni procesor *TMS320C6711* podjetja *Texas Instruments*.

Pokazalo se je, da je učinkovitost filtra ob iskanju kožnih značnic predvsem odvisna od osvetlitve zajete scene. Z spremembo osvetlitve se spreminja vrednost intenzitete in v primerih slabe osvetlitve scene je potrebno ustrezno znižati prag za vrednost osvetlitve v_{max}. Zato v nadaljnjem razvoju predvidevamo vključitev samodejnega prilagajanja na osvetlitev scene.

Literatura

/1/ Palmer, Steven E. Vision Science. The MIT Press, Cmbridge, MA, 1999.

/2/ 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.

/3/ K. Imagawa, S. Lu, S. Igi. Color-Based hands tracking system for sign language recognition. In Proc. Of Int'l Conf. On face and Gesture Recognition, pages 462-467, 1998.

/4/ S. Ahmad. A usable real-time 3D hand tracker. In proc. IEEE Asilomar Conf., 1994

/5/ 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.

/6/ F. Quek. Unencumbered gesture interaction. IEEE Multimedia, 1997.

/7/ T. Starner. A wearable computer based american sign language recognizer. In Proc. IEEE Int'l Symposium in Wearable Computing, Oct. 1997.

/8/ M. Jones, J. Rehg. Statistical color models with application to skin detection. Technical Report CRL 98/11, Compaq Cambridge Research Lab., 1998.

/9/ C. Wren, A. Azarbayejani, T. Darrel, A. Pentland. Pfinder: Real-time tracking of the human body. In Photonics East, SPIE Proceedings, volume 2615, Bellingham, WA, 1995.

/10/ M. Isard, A. Blake. Contour tracking by stochastic propagation of conditional density. In Proc. Of European Conf. On Computer Vison, pages 343-356, Cambridge, UK, 1996.

/11/ Y. Raja, S. McKenna, S. Gong. Colour model selection and adaptation in dynamic scenes. Ain Proc. European Conf. Computer Vison, pages 460-475, 1998.

/12/ Y. Azoz, L. Devi, R. Sharma. Reliable tracking of human arm dynamics by multiple cue integration and constraint fusion. In Proc. IEEE Conf. Computer Vison and Pattern Recognition, pages 905-910, Santa Barbara, CA, 1998.

/13/ J. Lee, T. Kunii. Model-based analysis of hand posture. IEEE Comput. Graph. Appl., vol. 15, no. 5, pages 77-86, Sept. 1995.

/14/ 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.

/15/ N. Herodotu, A.N. Venetsanopoulos. Image segnemtation for facial image coding of videophone sequences. 13th Internationl Conference on Digital Signal Processing, 1, 223, 1997.

/16/ 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.

/17/ 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

/18/ T. M. Cover, J. A. Thomas. Elements of Information Theory. Wiley, 1991.

/19/ N. Jojic, B. Brumitt, B. Meyers, S. Harris. Detecting and estimating pointing gestures in dense disparity maps. In Proc. IEEE Int'l Conf. On Face and Gesture Recognition, France, 2000.

/20/ R. Kjeldesn, J. Kender. Toward the use of gestures in traditional user interface. In Proc. Of IEEE Automatic Face and Gesture Recognition, pages 151 - 156, 1996.

/21/ Y. Wu, Q. Liu, T. S. Huang. An adaptive self-organizing color segmentation algorithm with application to robust real-time human hand localization. In proc. Of Asian Conference on Computer Vision, Taiwan, 2000.

/22/ I. Kramberger, M. Solar. DSP Acceleration Using a Reconfigurable FPGA. Proc. of IEEE International Symposium on Industrial Electronics - ISIE '99, Bled - Slovenija, 1999.

/23/ New, Bernie. A distributed arithmetic approach to designing scaleable DSP chips. EDN, pp. 107-114, 1995.

/24/ Atmel, Inc.. Recommended Design Methods. Atmel, September, 1997.

/25/ Atmel, Inc.. Implementing Cache Logic with FPGAs. Atmel, September, 1997.

/26/ Atmel, Inc.. Implementing Bit-Serial Digital Filters. Atmel, September, 1997.

/27/ Atmel, Inc.. Implementing FreeRAM inside the FPGA or AT94K Series FPSLIC Using VHDL with IP Core Generator, Atmel, 2001

/28/ Atmel, Inc.. AT94K Series Field Programmable System Level Integrated Circuit. Atmel, 2001.

/29/ MathWorks, Inc.. Image Processing Toolbox User's Guide. Mathworks, 2001.

/30/ MathWorks, Inc.. Signal Processing Toolbox User's Guide. Mathworks, 2001.

/31/ MathWorks, Inc.. Statistics Toolbox User's Guide. Mathworks, 2001.

/32/ B. Furht. Processors Arhitectures for Multimedia. Florida Atlantic University, Department of Computer Science and Engineering, Boca Raton, Florida 1998.

/33/ N. Holzschuch. Color fidelity and Color spaces. University of Cape Town, 1996.

/34/ J. Keith. YCbCr to RGB Considerations, Converting 4:2:2 to 4:4:4 YCbCr. Intersil Application Note March 1997 AN9717.

/35/ J. E. Adams. Design of practical color filter array interpolation algorithms for digital cameras. In Proc. Od SPIE, Bellingham, WA, USA, 1997.

/36/ R. P. Kleihorst et al.. Xetal a low-power high-performance msart camera processor. In Proc. ISCAS2001, Sydney, Australia, 2001.

/37/ J. Yang, A. Waibel. A real-time face tracker. In Proc. IEEE workshop on Applications of Computer Vision, Saratosa, FL, USA, 1996.

/38/ 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.

/39/ N. Pavešić. Razpoznavanje vzorcev: Uvod v analizo in razumevanje vidnih in slušnih signalov. Druga razširjena izdaja. Fakulteta za elektrotehniko, Ljubljana. 2000.

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

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

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

Prispelo (Arrived): 03.06.2002 *Sprejeto (Accepted): 25.03.2003*

SMALL-SIGNAL MODEL OF RESONANT LINK CONVERTER

Miro Milanovič[1] and Robert Kovačič[2]

[1]Faculty of Electrical Engineering and Computer Sciences, University of Maribor, Maribor, Slovenia
[2]IPS d.o.o., Research and Design Department, Ljubljana, Slovenia

Key words: power supplies, high frequency power converters, resonant link, soft switching, modeling

Abstract: The conventional small signal modeling techniques as are state space averaging and injecting-absorbed current method are not appropriate for using in converters based on high frequency resonant link (HFRL). The mentioned methods are appropriate for processes where the switching frequency is constant. Because the operation frequency of the HFRL converters is load dependent the other way of modeling must be used. In this paper a small-signal model of the HFRL converter, operating with a variable resonant link frequency, is developed by using of the estimator for linear system (ELIS) which exists under MATLAB. The high frequency resonant link voltage is modulated by the low frequency signal. The Levenberg-Marguardt aproximation method is used for evaluation of the magnitude and phase of the envelope of the high frequency resonant link voltage.

Malosignalni model pretvornika z resonančnim povezovalnim krogom

Ključne besede: napajanje elektronskih vezij, visokofrekvenčni pretvorniki, resonančni povezovalni krog, mehko preklapljanje, modeliranje

Izvleček: Običajne tehnike za malo-signalno modeliranje, kot so povprečenje v prostoru stanj in metoda iniciranega-absorbiranega toka, niso primerne za uporabo pri pretvornikih, ki so zasnovani na visoko-frekvenčnem povezovalnem krogu. Omenjeni metodi sta primerni v procesih, kjer je frekvenca prožilne enote konstantna. Ker je frekvenca delovanja visoko-frekvenčnega povezovalnega kroga spremenljiva v odvisnosti od bremenske upornosti, moramo uporabiti drugi način modeliranja. V tem članku bomo opisali postopek malo-signalnega modeliranja pretvornika z visoko-frekvenčnim povezovalnim krogom pri spremenljivi resonančni frekvenci. Postopek modeliranja bomo izvedli s pomočjo estimatorja za linearne sisteme (ELIS), ki deluje znotraj programskega paketa MATLAB. V ta namen bo napetost na visoko-frekvenčnem povezovalnem krogu amplitudno modulirana z nizkofrekvenčnim signalom. Za ocenitev vrednosti amplitude in faze ovojnice je uporabljena aproksimacijska Levenberg-Marguardt metoda.

1. Introduction

The term hf-ac or hf-dc resonant link converter usually denotes a circuit, whose main part is the resonant tank circuit. It is also possible to utilize the resonant links to derive energy storage and filtering functions normally obtained by the dc voltage link. Electrolytic capacitors provide low cost, high density energy storage in the dc voltage link of a voltage source inverter. However, the dc link based on electrolytic capacitors has several inherent limitations. One important drawback is the excessive switching loss and device stress which occurs during the switching interval. Introduction of a resonant or quasi-resonant operation principle into the known converter/inverter topology represents a possible solution of this problem. While this principle has been recognized for over ten years, the important advantages in avoiding switching stresses have been appreciated because the conventional hard-switched based converters suffer from the switching losses and hence cannot work at the very high frequency.

Recently, resonant ac or dc links have been studied and suggested as strong candidates for a power conversion link as it was described in /1/, /2/, /6/ and /7/. The using of high frequency ac resonant link principle for energy storage purposes enables to reduce the device losses or device stress by restricting the switching time to the instants of zero current or zero voltage. In general ac or dc high frequency link inverters have been used only for high-power application as are the motor drives, UPS etc. as it is shown in Fig. 1 (a). The converter proposed in /1/, /2/ and [6] are based on bidirectional switches.

In Fig 1 (b) is shown the converter circuit which enables to supply the resonant tank circuit only by using unidirectional switches /3/. This principle of the operation can be used in the low power dc-dc conversion as well, as it is shown in Fig. 2 (a) and (b). The energy storage function is taken over by resonant tank circuit, by using the transformer the energy can be provided to load. The load side of transformer is equiped by rectifier and filter elements. In this paper the structure of dc-dc conversion based on an ac-resonant link is presented. This resonant tank circuit is capable to provide the energy storage function instead of the conventionaly used electrolytic capacitor. The main drawbacks is that the resonant link voltage magnitude and the operation frequency of resonant link circuit is load dependent.

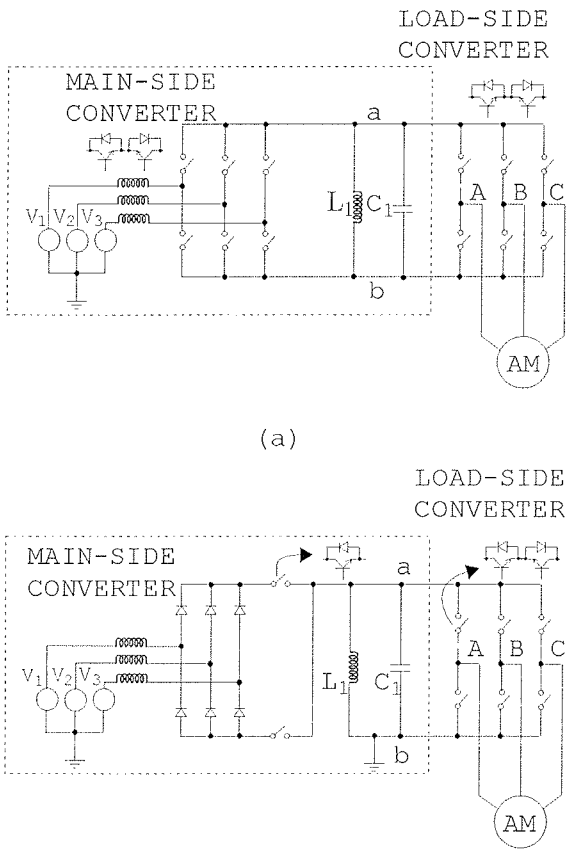


Fig.1: Induction motor drive based on resonant link converter circuit: (a) Main side converter with bi-directional switches; (b) Main side converter with uni-directional switches.

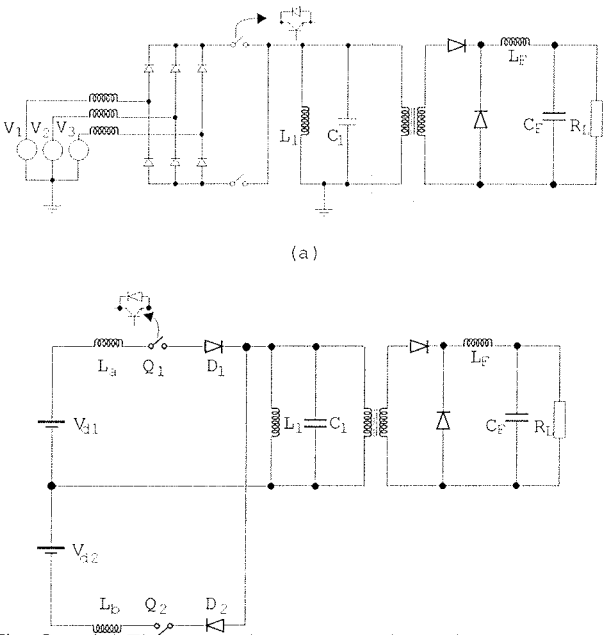


Fig. 2: (a) The ac to dc converter based on resonant link circuit (b) The dc to dc converter based on resonant link circuit.

Hence the conventional methods for the small-signal modelling as are state space averiging, injected-absorbed current method, pwm equivalent circuits /4/ cause the modelling problems because the conventional methods require the constant switching frequency of operation /8/, the method based on the measured data is proposed. The estimator for linear system (ELIS) has been used for small-signal model developing of the HFRL converter, which is operating with a variable resonant link frequency. ELIS is the software available as a toolbox in Matlab-Simulink package. The high frequency resonant link voltage is modulated by the low frequency signal. The Levenberg-Marguardt approximation procedure is used for accurate evaluation of the magnitude and phase of the HFRL voltage envelope from measured data. The control parameter adjustment will be based on developed small signal model.

2. Principle of the operation

The basic scheme for "evolution" of the dc to ac high frequency converter circuits is shown in Fig. 3. In steady state the parallel resonant circuit consists of L_1 and C_1 operates and provides the energy to the resistor R_L (R_L represents the load). Transistor Q_1 could be switched on when the voltage on the parallel resonant link crosses zero as it is shown in Fig. 4 (time instant t_0). Then L_a with the elements of parallel resonant tank circuits L_1 and C_1 establish a series resonant circuit. The current through transistor Q_1 is supposed to be of sinusoidal wave shape (interval A). When the current through Q_1 crosses zero, diode D_1 turns off and transistor Q_1 can be turned off as well (time instant t_1). When the series resonant frequency ω_s is higher than parallel resonant frequency ω_p the soft switch operation of converter has been used. The energy was provided from dc voltage source V_{d1} to parallel tank circuit during the time interval A. Because of similarity with half-wave operation of diode rectifier the circuit can be described as "half-wave" configuration of the dc-ac hf resonant link converter.

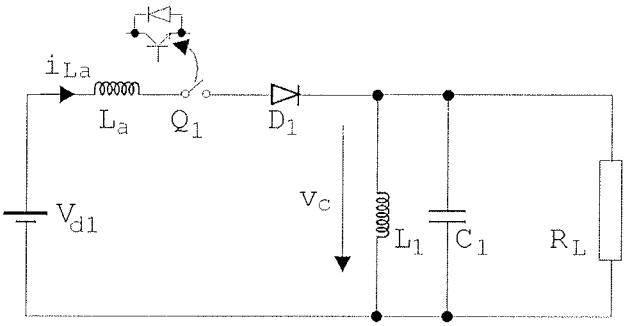


Fig. 3: The "half-wave" topology of dc-ac hf resonant link converter.

Operation of the circuits from Figs. 5 (a) and (b) is similar. The energy is provided from the dc-supply to the parallel resonant tank circuit in both half periods of the output ac high frequency resonant link voltage. Because of that the circuit could be titled a "full-wave" configuration of the ac-

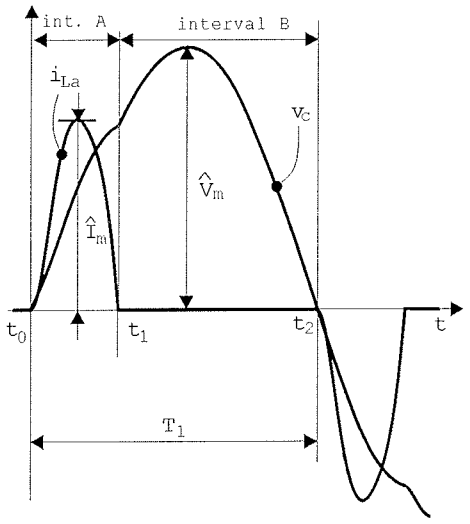


Fig. 4: (a) The current and voltage wave-shape.

hf resonant link converter. The oscilograms measured in the "full-wave" configuration are shown in Fig. 4 (b). In order to provide the energy from main supply to the resonant tank circuit the configurations shown in Figs. 6 (a) and (b) are suggested. The circuit in Fig. 6 (a) is coming as evolution of the circuit shown in Fig. 5 (a). The disadvantages of this circuit are two capacitors, which "simulate" the voltage sources V_{d1} and V_{d2} and there are still two inductors L_a and L_b and two diodes D_1 and D_2.The component minimized converter circuit is shown in Fig. 6 (b). The converter consists of diode bridge, the transistor bridge, inductor L_a and the resonant tank circuit. The function of two series diodes D_1 and D_2 can be taken over by diodes from the rectifier bridge.

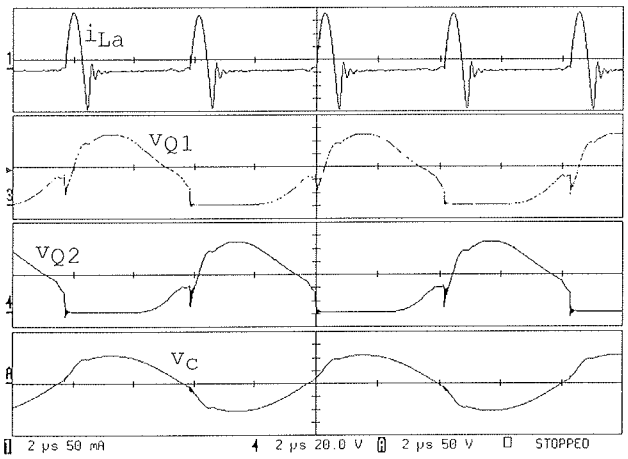


Fig. 4: (b) Experimental results measured in "full-wave" configuration

3. The small signal modeling

The hard switch converters represent the non-linear circuits and because of that it's analysis is so complicated. Because the load, connected at the converter output, requires constant voltage regardless of the current, a voltage controller circuit should be implemented. Many algorithms are developed for the linear time invariant circuits, which can be described in "s" or "z" space. Because of that the linearization method as are the state space averaging, injection-absorbed current method and etc. are widely used. Sometime the modeling process is of great pretension because of non-ideal and parasitic elements. In this case the transfer function, which is necessary for controller parameter design, could be measured.

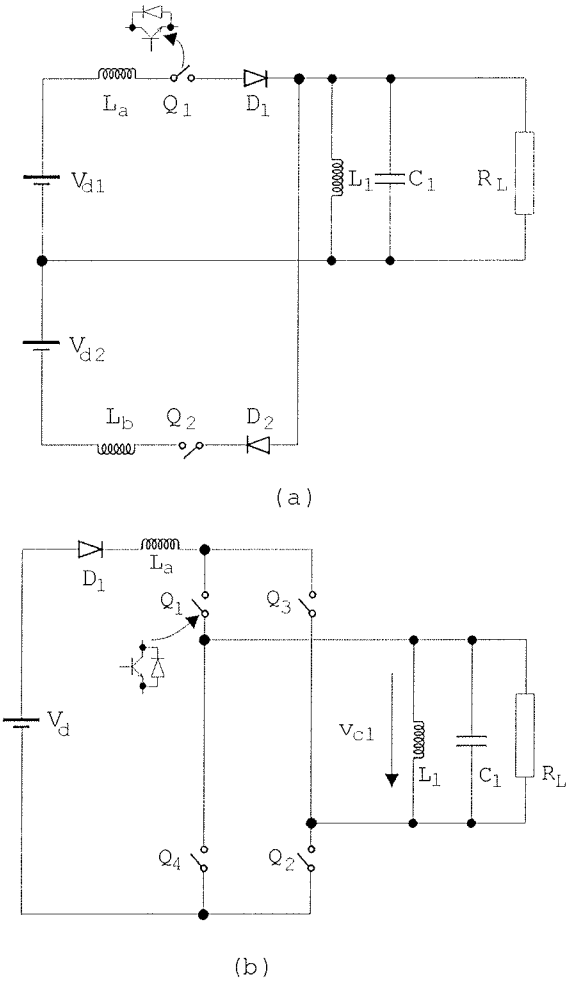


Fig. 5: (a) The "full-wave" topology of dc-ac hf resonant link converter. (b) The "bridge" topology of dc-ac hf resonant link converter.

3.1 The standard procedure for measuring of small signal model

The small-signal model of the converters can be evaluated through the frequency response (FR) of the circuit. The FR of the circuit may be regarded as a complete description of the sinusoidal steady-state behavior of a circuit as a function of the frequency. In Fig. 7 (a) the standard open loop procedure of FR measurement for buck converter is shown. The "small" signal v_{ref} is adding to dc voltage V_{DC}, which defines the steady state of the converter operating points. The large gain of the object causes that the steady

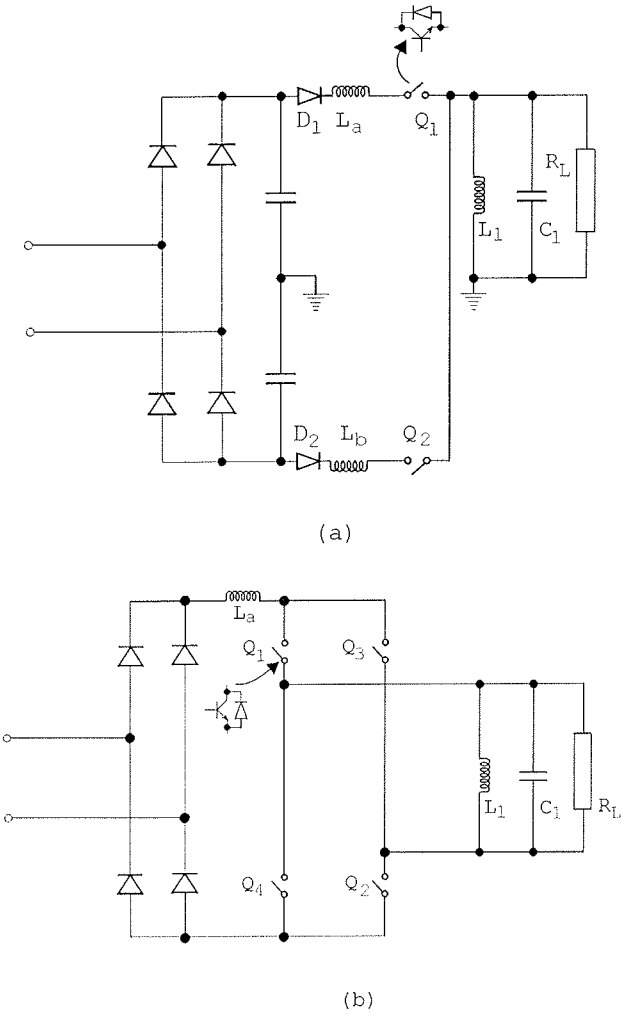


Fig. 6: (a) the "full-wave" topology of dc-ac hf resonant link converter connected with the mains. (b) The "bridge" topology of dc-ac hf resonant link converter connected with the mains.

state is unstable, therefore the small signal perturbation causes the large changing of the output voltage. For example when the gain is 60 *dB* (i. e. 1000) the input small signal perturbation of 1 *mV* will cause the output changing of 1 *V*. Through the measurement process the different operating points of the power stage are exited and such measured frequency response is not accurate. How to avoid this problem is described in /10/. Network analyzers could measure the control objects with large gain where the small signal perturbation has been injected into closed control loop as it is shown in Fig. 7(b). During the measurement process the controller keeps the converter operating point stable. The measurement of the converter FR is very accurate when the network analyzer was used. The ac output has been measured by narrow pass-band filter, which guaranties good disturbance rejection. These instruments also repeat measurement procedure by equal frequency and as result; the arithmetic average value of the FR at the particular frequency is evaluated.

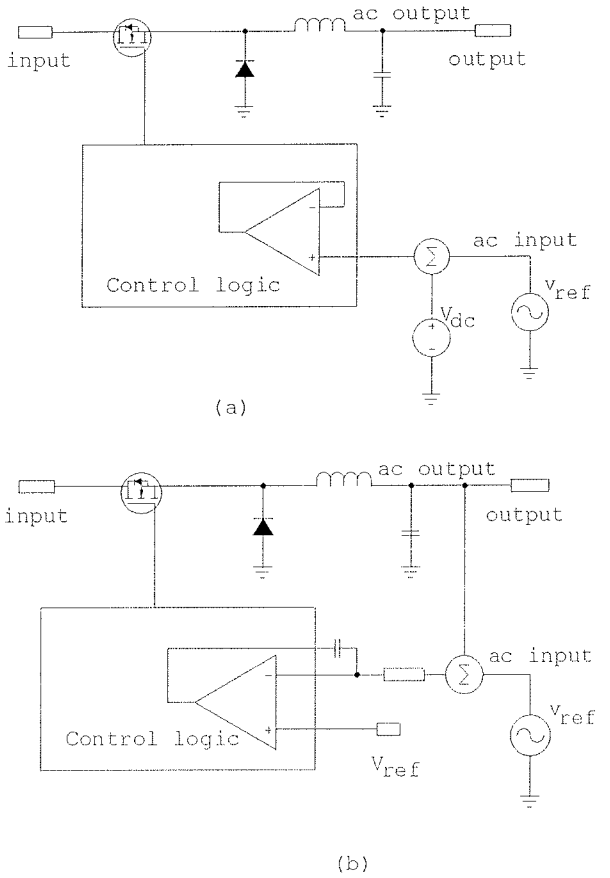


Fig. 7: (a) The buck converter open loop measurement of the frequency response; (b) The buck converter closed loop measurement of the frequency response.

3.2 The measurement procedure for small signal model of HFRL converter

The ac-hf resonant link converter does not have such "nice" properties. The output link voltage is alternative and when small signal perturbation is injected the amplitude modulation will appear on link voltage. The resonant link voltage is alternative but not sinusoidal as it is shown in Fig. 4 (a) and (b). The amplitude changing caused by input small signal perturbation is sinusoidal. The measurement circuit is shown in Fig. 8.

The open-loop principle has been used. This could be used because the open loop gain was only 40 *dB*. The small signal voltage v_{in} and output signal on transistor Q_2, the voltage v_c has been measured by scope *LC334A*. In Fig. 9 the measurement results are shown.

The results shown in Fig. 9 were also available as a file in binary form. This data has been processing with Mathematica, where the Levenberg-Marquardt method for approximation was used to find the unknown parameters of the signal described by (1). Both the input signal v_{in} and the envelope of the output signal v_c could be defined with expression:

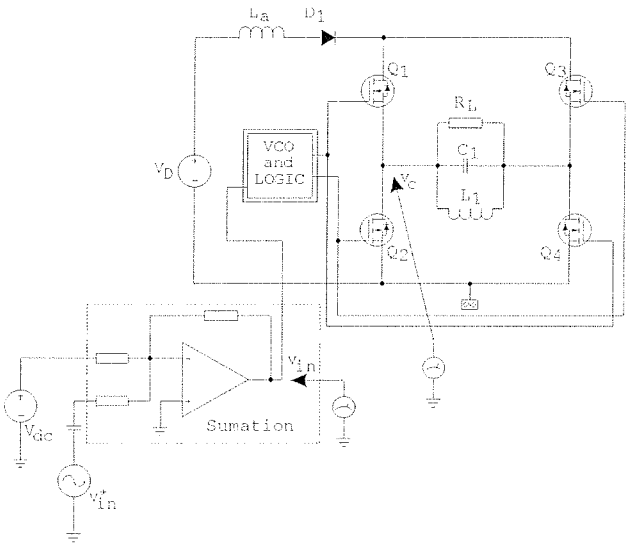


Fig. 8: The open loop circuit for FR measurement of the HFRL converter

$$v_x = V_{DXx} + \hat{V}_x \cos(\omega t + \varphi) \qquad (1)$$

where V_{DXx} is the dc operating points, $\hat{V}_x$ represents the magnitude of the input signal or magnitude of the output signal envelope, ω is the signal frequency and φ is the phase angle regarding scope synchronization point.

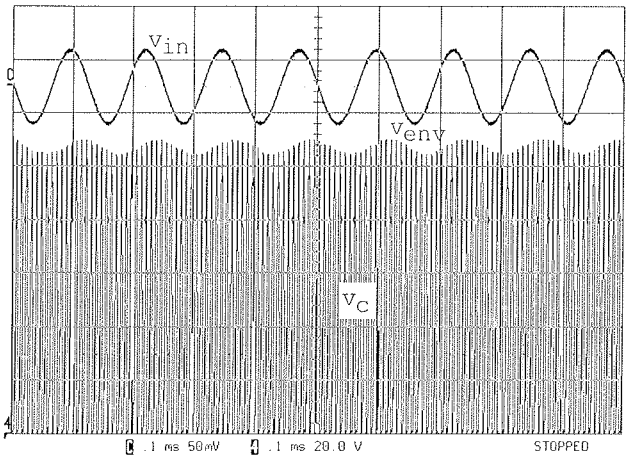


Fig. 9. Measurement results of the input voltage v_{in} and output voltage v_c

From the difference between input phase angle and output phase angle the phase diagram of the HFRL converter could be evaluated as well. Fig. 10 (a) shows the input signal waveform (green) with measurement noise and the solid line is the result of approximation by Levenberg-Marquardt method. The output signal causes some problems because it has been defined as peak values of the resonant link voltage and these points are discrete. The result of this approximation is shown in Fig. 10 (b). From this results all parameters needed in (1) have been defined and the measurement data becomes appropriate for evaluation of the gain and phase waveforms in frequency domain.

By using the estimator for linear system (ELIS) which exists under the Malab the converter model has been derived. The ELIS works better with a set of the measured data, which are measured under the same conditions as shown in Table 1.

To identify the model of the converter a general presumption, i.e. the number of zeros and poles must be defined.

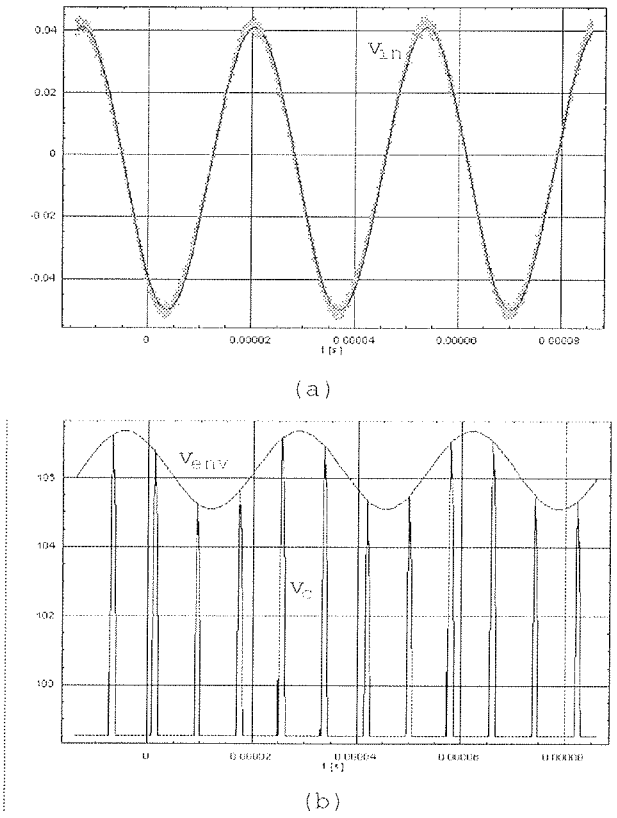


Fig. 10. Measurement results of the input voltage v_{in} and output voltage v_c

In Table 2 the placement of the poles as a result of the evaluation procedure by ELIS is shown. In the first and second columns the evaluation results of the first and third order system are presented. In the third and fourth columns the models of the first order system with the different delays have been shown /5/. When the number of poles has been changed the placement of one of the poles did not change significantly. In the third order model this pole has a natural frequency of 42644 *rd/s*. Two other poles have complex values.

There exists a physical explanation of this natural frequency. The paramount of interest is not the dynamics of the tank circuit itself but the dynamics of the resonant link voltage envelope $\tilde{v}_{env}$ caused by the input voltage $\tilde{v}_{in}(s)$. For $\tilde{v}_{env}$ dynamics could be supposed it is dependent on resistance R_L and capacitance C_P. For chosen R_L and C_P (R_L=23706Ω, C_P=2.014*nF*) the natural frequency of this parallel system is $\omega_{RC} = 1/R_L C_P = 20946$ *rad/s* which is more or less half of the previous derived natural frequency (Table 2).

Table I: Measurement results

	1. experiment		2. experiment		3. experiment	
Freq. *f(Hz)*	Gain *A(dB)*	Phase φ(*deg.*)	Gain *A(dB)*	Phase φ(*deg.*)	Gain *A(dB)*	Phase φ(*deg.*)
200	40.835	-3.0271	40.2397	-2.46356	40.4445	-3.54746
500	41.2899	-4.20871	40.2185	-2.65827	40.1483	-3.10764
800	41.3075	-6.54668	40.1752	-5.17673	40.0429	-4.50148
1000	41.2422	-8.0458	39.9262	-5.29607	40.0305	-3.88876
2000	41.1928	-17.02	39.921	-13.6055	39.9932	-15.9025
3000	40.5761	-25.7001	39.6663	-23.8292	39.6571	-23.6579
5000	39.4706	-40.5851	38.6214	-37.1977	38.691	-38.2392
8000	37.389	-56.2334	36.7191	-53.3234	36.6696	-54.2372
10000	35.842	-64.4906	35.374	-61.4672	35.3644	-61.0895
20000	31.3755	-84.3958	30.8801	-84.5519	30.8145	-82.9609
30000	28.5177	-95.866	28.0225	-98.1729	28.4572	-97.4357

Table II: Pole placement

	1st model	2nd model	3rd model	4th model
Number of zeros	0	0	0	0
Number of poles	1	3	1	1
delay	0	0	1.8μs	1.9μs
The values of poles	-3698	-42644	-44729	-45106
		- 194816 +j445496		
		- 194816 -j445496		

That's why the pole placement could be expressed as:

$$s_1 = -2\omega_{RC} \tag{2}$$

Two complex poles appear in the system because of the discrete nature of the link voltage (only peak of the HFRLV has been observed) and the discrete VCO controlled modulator. The steady state operation frequency ω was $1.6\text{x}10^6$ rd/s (cca. 250*kHz*). The authors in /4/ and /5/ suggest that the discrete structure of such system could be described by time delay as it follows:

$$H(s) = \exp(-sT_d) = \frac{1}{1 + sT_d + \frac{1}{2}sT_d^2 + \ldots} \tag{3}$$

where T_d is half of the period of the resonant link voltage. The gain of the plant can be estimated from the energy stored in the capacitors and dissipated on the load resistor. The magnitude of the current through L_a and the time t_{on} as difference of t_1 and t_0 can be evaluated by formulas:

$$\hat{I}_m = -\frac{i_{L1}(0)L_1}{L_1 + L_a} + \sqrt{\left(\frac{i_{L1}(0)L_1}{L_1 + L_a}\right)^2 + \left(\frac{V_{d1}L_1}{L_1L_a + L_a^2\omega_s}\right)^2} \tag{4}$$

$$t_{on} = \frac{\pi}{\omega_s} - 2\frac{i_{L1}(0)L_1}{L_1 + L_a} \frac{1}{\omega_s\sqrt{\left(\frac{i_{L1}(0)L_1}{L_1 + L_a}\right)^2 + \left(\frac{V_{d1}L_1}{L_1L_a + L_a^2\omega_s}\right)^2}} \tag{5}$$

where:

$$\omega_s = \frac{1}{\sqrt{L_nC_1}}, \qquad L_n = \frac{L_1L_a}{L_1 + L_a}$$

The average value of the current through La in half-period of the resonant link voltage is obtained as

$$I_{La,avg} = \frac{\hat{I}_m t_{on}}{\pi T_1} \tag{6}$$

The energy provided inside whole period is obtained by:

$$W = 4V_{d1}I_{La,avg}T_1 \tag{7}$$

Presume that the resonant link voltage has sinusoidal waveshape, than its rms value is:

$$V_{c,ms} = \sqrt{\frac{WR_L}{2T_1}} \tag{8}$$

and magnitude:

$$\hat{V}_m = \sqrt{2}V_{c,ms} = \sqrt{\frac{WR_L}{T_1}} \tag{9}$$

where ^ represents the magnitude of the current or voltage. The gain consists of the gain of VCO (K_{VCO}) and the converter gain K_Uw which is frequency dependent /3/. By substitution (4)-(8) into (9) the slope could be expressed as:

$$K_{u\omega} = \frac{\partial \hat{V}_c}{\partial \omega_p} = \hat{I}_{La}\sqrt{L_a R_L/\pi}\ \left(2\sqrt{\omega_p}\right)^{-1} \qquad (10)$$

Therefore the whole transfer function is:

$$\frac{V_{env}}{V_{in}} \cong \frac{K_{VCO}K_{u\omega}}{2\omega_{RC}\left(1+\frac{s}{2\omega_{RC}}\right)\left(1+T_d s+\frac{1}{2}T_d^2 s^2\right)} \qquad (11)$$

In Figs. 11 the different frequency responses are shown ("+" is the average value of the measured results from Table 1).

3.3 The control of HFRL converter

In this section the experimental results obtained on a bridge structure of the HFRL converter are shown. The control problem can be defined as a necessity to keep constant magnitude of ac-hf resonant link voltage regardless of load conditions. For small-signal modeling the above-described method has been used. In lab. prototype the measurement is realized by using the appropriate electronics circuit as it is shown in Fig. 12. The control object includes the voltage-controlled oscillator as modulator (VCO), dc-ac high frequency converter and sample and hold circuit (S/H) as a peak voltage detector. Fully analog circuits generate the control signals. For triggering units (instead of pwm) the VCO has been used. The S/H circuit is used for the measurement of the peak of the high frequency voltage. The capacitor current i_{C1} has a delay with respect to voltage v_C of exactly $\pi/2$. If the voltage is sampled in this time instant (when the current i_{C1} crosses zero), the peak value of voltage v_C will be measured. This simple phenomenon eliminates the need for filter in measuring circuit. Based on the developed small signal model the PI controller has been designed. In Fig. 13 the experimental results are shown. In particular time instant the load resistance has been changed which has influence on the high frequency resonant link voltage v_c. From the waveforms it is evident that the control of the high frequency resonant link voltage v_c has been reached.

4. Conclusion

The high frequency resonant link converter has been discussed in this paper. The energy has been provided to the resonant tank circuit through the series resonant process. Because of that the soft switch converter operation has been reached. The main effort has been done at the efficiency study and consequently the efficiency increases up to 92%. The magnitude of the resonant link voltage is always larger than the dc input voltage. This "disadvantage" can be avoided by introducing the transformers and the lower voltage can be reached at the converter dc output. The EMI influences is lower than the hard switch converter usually produces because the current and voltages have sinusoidal wave-shape and because of soft switching operation.

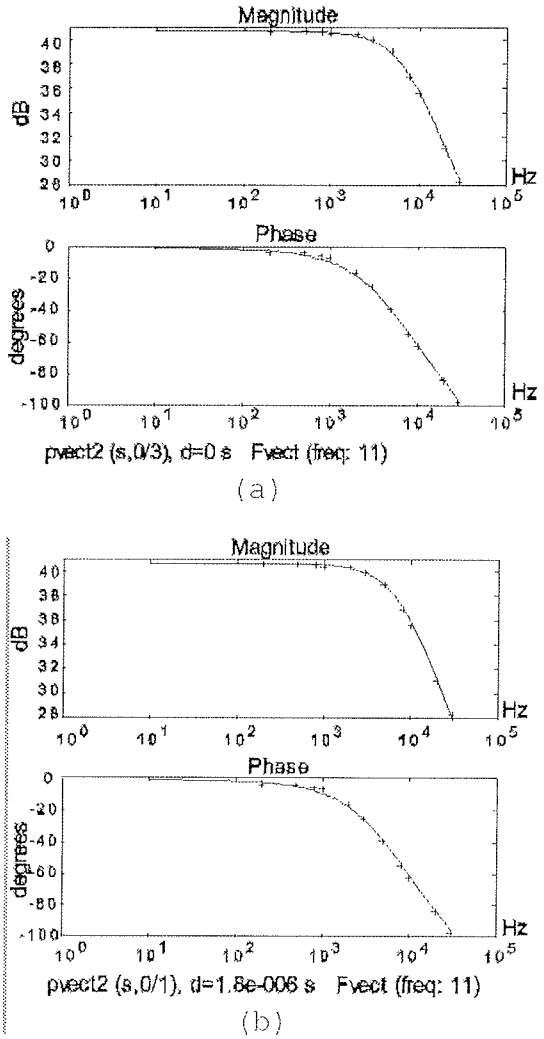


Fig. 11: (i) Frequency response (number of zeros are 0, number of poles are 3, delay is 0μs; (ii) Frequency response (number of zeros are 0, number of poles are 1, delay is 1.8μs)

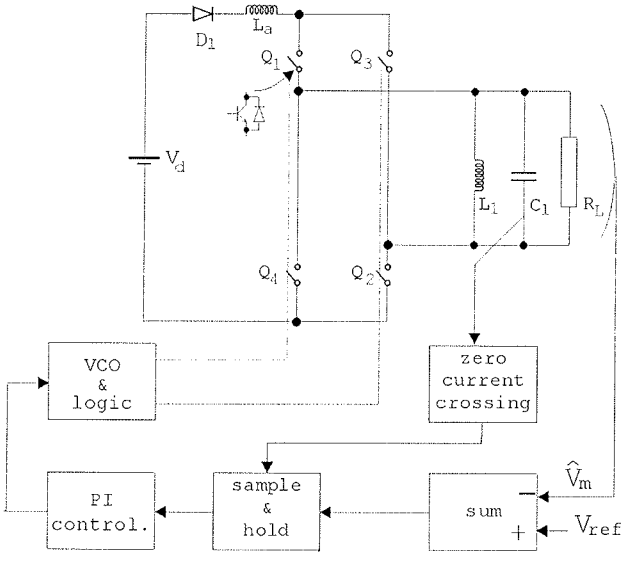


Fig. 12: The control sheme of HFRL converter

For control purposes the method based on measurement data has been used. The MATLAB tool-books ELIS for identification of the small signal model of the ac resonant link converter has been investigated. For accurate identification of resonant link voltage parameters (i. e. magnitude and phase of the voltage envelope) the Levenberg-Marguardt method was used. The experimental results have shown that this procedure is capable of providing a good solution of the frequency response for the non-linear problem. Based on this modeling the control of HFRL voltage has been realized.

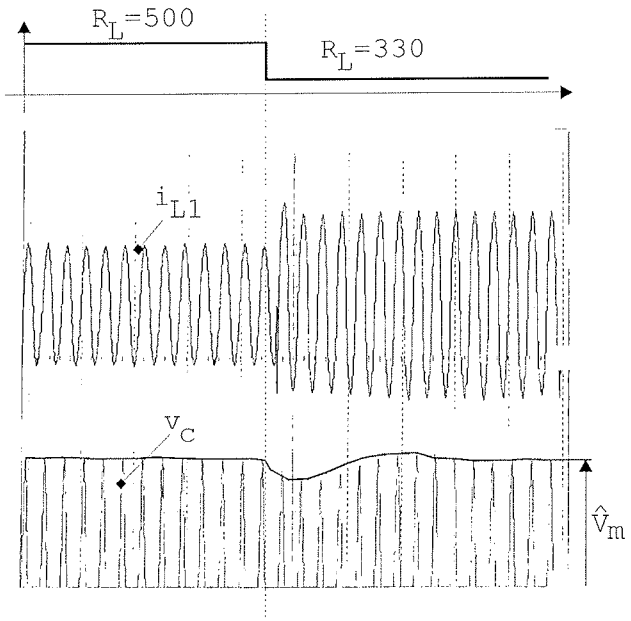


Fig.13: Experimental results when the load resistance has been changed: (uper) the current of parallel resonant link circuit i_{L1} (200 mA/div), (lower) the magnitude of resonant link voltage v_c (20V/div).

References

/1/ SOOD, P. K., LIPO, T. A.,' Power conversion distribution system using a resonant high frequency AC-link', *IEEE Transactions on Industry Applic.* 1988, 24, (2) pp. 586-596.

/2/ DIVAN. D. M., 'The resonant DC-link converter - a new concept in static power conversion', *IEEE Transactions on Industry Applic.* 1989, 25, (2) pp. 317-325.

/3/ MILANOVIC, M., KOVACIC R., MIHALIC F. and BABIC R. 'The control of an AC to HF-AC resonant link converter', *Inf. MIDEM, 1996,* 26, (1), pp. 7-13.

/4/ KISLOVSKI, A. S., REDL, R., SOKAL, N. O., 'Dynamic Analysis of Switching-Mode DC/DC Converters' Van Nostrand Reinhold, New York, 1991.

/5/ TSAKHURUK, T. A., LEHMAN, B., STANKO-VIC, A. M., TADMOR, G., 'Effects of Finite Switching Frequency and Delay on PWM Controlled Systems', *IEEE Transactions on Circuits and System, Fundamental theory and Application,* vol. 47, No.4 pp. 555-567, April 2000.

/6/ S.K.SUL, T.A.LIPO, 'Design and performance of a high frequency link induction motor drive operating at unity power factor', *IEEE IAS Annual Meeting*, Proc., pp. 308-313, Oct. 1988

/7/ Y. MURAI, S. MOCHIZUI, P. CALDERIA, T.A.LIPO, 'Current pulse control of high frequency series resonant dc link power converter', *IEEE IAS Annual Meeting*, Proc., pp. 1023-1030, Oct. 1989.

/8/ J. HOLTZ, 'Pulsewidth modulation - a survey', *IEEE Transaction on Industrial Electronics*, vol.39, No.5, pp, 410-420, October 1992.

/9/ M. MILANOVIC, F. MIHALIC, D. MARKO, K. JEZERNIK, 'The ac to hf/ac resonant link converter' *IEEE PESC Conf. Rec*, pp. II/750-756, 1995.

/10/ R. LENK, 'Practical design of power suppliesr' Van Nostrand Reinhold, New York, 1992.

Dr. Miro Milanovič, univ.dipl.ing.
Univerza v Mariboru
Inštitut za avtomatiko in robotiko
Fakulteta za elektrotehniko,
računalništvo in informatiko
Smetanova 17, Maribor, Slovenija
e-mail: milanovic@uni-mb.si

mag. Robert Kovačič, univ.dipl.ing.
IPS d.o.o., Research
and Design Department
Cesta Ljubljanske brigade 17, Ljubljana, Slovenia
e-mail: kovacic@ips.si

Prispelo (Arrived): 30.05.2002 *Sprejeto (Accepted): 25.03.2003*

ANALIZA SEGRETJA MOTORSKEGA ZAŠČITNEGA STIKALA PRI TRAJNI TOKOVNI OBREMENITVI

Gorazd Hrovat [1], Anton Hamler [1], Mladen Trlep [1], Martin Bizjak [2]

[1] Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko, Maribor, Slovenija
[2] Iskra Stikala d.d., Kranj, Slovenija

Ključne besede: temperaturno polje, MKE, motorsko zaščitno stikalo, ANSYS;

Izvleček: Delo predstavlja analizo segretja motorskega zaščitnega stikala (MZS) pri trajnem toku s programskim paketom ANSYS, ki temelji na metodi končnih elementov (MKE). Ta stikala predstavljajo enostavno obliko motorskega zaganjalnika, v katerem so združene funkcije stikanja, preobremenitvene in kratkostične zaščite. Geometrijski model zaščitnega stikala, ki je bil upoštevan pri analizi, je delno poenostavljen. V analizo je bil vključen le en pol stikala, ki je sicer tripolen. Opravljeni so bili izračuni segrevanja pri dveh različnih koeficientih toplotne prestopnosti na površinah ohišja stikala.

Analysis of Heating of Motor Protection Switch by Permanent Current Load

Key words: temperature field, FEM, motor protection switch, ANSYS;

Abstract: Nowadays, electrical motor is, due to its simplicity, robustness and immediate readiness to operate, the most wide-spread driving machine wherever electrical energy is available. Switching and protection devices should enable undisturbed working, allow complete utilization of motor and interrupt its operation only in cases when the motor is really endangered. One of such devices is the motor protection switch, in which functions of start-up, switch-off, overload and short-circuit protection are combined.

The paper deals with the analysis of heating of motor protection switch by permanent current load of 32 A. Joul losses are caused by that current and all parts of the switch are being heated as a result of heat transition from the site of production to the surrounding parts. Produced heat is partly used to increase the temperature of switch parts and partly conveyed into the surroundings. All three physical heat transition mechanisms perform at the transport of heat (Pic. 1). However, the influence of radiation was negligible because of low temperatures.

Motor protection switch present a very complex 3D geometry, which has been simplyfied during the analysis. Parts of the switch i.e. are control device, mechanism, do not influence the temperature field and for that reason their presence has been neglected. The main source of heat consists of a heating coil with bimetal, so the influence of coil has been neglected, which has slighty changed the current flow of motor protection switch (Pic. 2). Three-pole constructions with partition walls are usually located between the individual poles. Treatment has been reduced to one pole with a suitable boundary condition i.e. heat isolation of partition wall between two poles. The final model of motor protection switch is presented in figure 3. Program package ANSYS, which is based on finite element method (FEM), was used for static calculation of temperature field.

The results of calculations for two different values of convection coefficients on housing of motor protection switch were presented as that boundary condition, which has the most important influence on temperature field distribution. Convection coefficient at the conductor isolation was the same in the both cases. Temperature field distribution with a convection coefficient of 15 W/m^2K is presented in figures 5 and 6. When a convection coefficient increases to 20 W/m^2K temperature field distribution is presented in figures 7 and 8. As seen there, the temperature on housing of motor protection switch never exceeds the maximum heat permitted by a standard. Contact part on the bimetal and heating coil side is more strained by temperature as those two elements present the most important source of heat. Maximum temperature of the conductible part is the middle turns of the heating coil, which is caused by taking away the heat from both sides of the heating coil to the connected parts of the switch.,

1 Uvod

Elektromotor je danes zaradi svoje enostavnosti, robustnosti, ekonomičnosti in takojšnje pripravljenosti na delovanje najbolj razširjen pogonski stroj povsod tam, kjer je na voljo električna energija. V sodobnih tehnoloških procesih pa sta pomembni predvsem zanesljivost in varnost. Stikalne in zaščitne naprave morajo torej omogočati nemoteno obratovanje, dopuščati popolno izkoriščenost motorja in tehnološki proces prekiniti le tedaj, ko je motor resnično ogrožen.

Ena najenostavnejših oblik motorskega zaganjalnika je motorsko zaščitno stikalo (MZS), v katerem so združene funkcije stikanja (vklop - izklop), preobremenitvene in kratkostične zaščite. Njihova stikalna zmogljivost je mnogo večja od zahtevane za samo krmiljenje motorja, zato jih prištevamo kar med odklopnike. Nekateri jih imenujejo tudi »odklopnik za zaščito motorjev«.

Glavni sestavni deli takšnega stikala so:

- mehanizem za ročni vklop in izklop,
- kontaktni sistem s fiksnim in gibljivim kontaktom,
- komore za gašenje obloka in
- nadtokovni sprožnik kot kombinacija bimetalnega in elektromagnetnega (kratkostičnega) sprožnika.

Glede na splošno razvrstitev stikalnih aparatov spada MZS med ročno krmiljene mehanske stikalne aparate in je običajno tripolne izvedbe.

Zgradba in delovanje MZS mora ustrezati zahtevam večih standardov. Glede na njegove lastnosti se morajo pri zasnovi konstrukcije in preizkušanju upoštevati zahteve standardov IEC 60947-1 (splošni standard za nizkonapetostne stikalne aparate), IEC 60947-2 (za odklopnike) in IEC 60947-4-1 (za motorske zaganjalnike) /10/.

Ustreznost delovanja in zmogljivosti MZS se preverja z več standardiziranimi preizkusi. Kosovni preizkusi obsegajo preizkus dielektrične trdnosti stikala, delovanje bimetalnega sprožnika, občasno se opravlja kontrola umerjanja bimetalnega sprožnika. Med tipske preizkuse spada meritev segrevanja priključnih sponk, električna in mehanska trajnost stikala, pa tudi kratkostična izklopna zmogljivost. Vsi testi, ki so zahtevani po standardih tu niso navedeni.

V članku je prikazana analiza segrevanja MZS pri trajnem nazivnem toku 32 A. Analiza temelji na uporabi metode končnih elementov (MKE), ki spada med sodobne numerične metode, saj je njen razvoj šel vzporedno z razvojem računalnikov. Zaradi kompleksnosti geometrije brez poenostavitev le-te ni šlo. Obravnavan je bil le en pol stikala. Rezultati izračunov so predstavljeni za dve različni vrednosti koeficienta toplotne prestopnosti na ohišju stikala.

2. Temperaturno polje

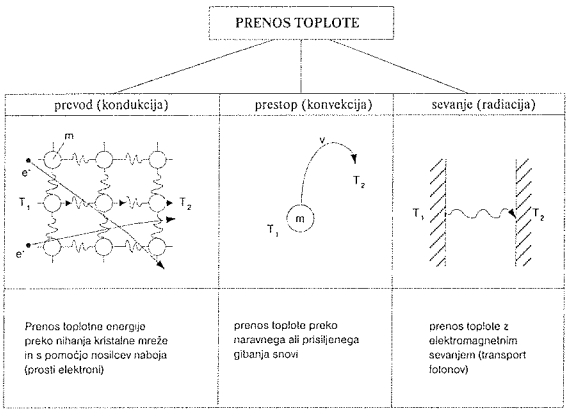
Proces prevajanja toka skozi tokovodeče dele MZS spremljajo joulske izgube, ki segrevajo vse dele stikala, saj prehaja toplota iz mesta nastanka na sosednje dele in okolico. Del nastale toplote se porablja za zvišanje temperature delov stikala, del pa odteka v okolico.

Segrevanje in ohlajanje MZS je splet fizikalnih procesov, ki jih je treba spoznati, da bi lahko razumeli pojav sam. Prenos toplote in mehanika tekočin pa sta tisti veji znanosti, ki nam dajeta na razpolago orodja, s katerimi lahko rešimo zastavljene probleme.

Prenos toplote proučuje transport toplote z enega mesta na drugo, tj. z mesta višje temperature na mesto nižje temperature. Za toplotni tok, ki je posledica krajevno in časovno spreminjajočega temperaturnega polja, veljajo zakoni termodinamike, predvsem drugi, ki govori o smeri, v kateri potekajo termodinamični procesi /5/.

Koliko toplotne energije se prenese iz enega mesta na drugo, lahko določimo le z merjenjem temperature. Torej je poznavanje temperaturnega polja ključnega pomena za izračun toplotnih tokov iz osnovnih zakonov transporta toplote, ki povezujejo temperaturni gradient in toplotni tok.

Pri transportu toplote ločimo tri fizikalno različne mehanizme, ki pa lahko nastopajo istočasno (slika 1).



Slika 1: Mehanizmi prenosa toplote

Pri MZS je temperatura prevodnih elementov majhna, zato so učinki sevanja zanemarljivi. Glavni način odvajanja sproščene toplotne energije je kombinacija prevoda v okoliški zrak in naravne konvekcije. Naravna konvekcija je v notranjosti stikala s površine prevodnih in izolacijskih delov mala, poleg tega pa se v simulaciji ne da zajeti, saj je temperatura v notranjosti stikala neznana.

2.1 Enačba temperaturnega polja

Zakon o ohranitvi energije v telesu prostornine V, ki je omejena s površino A pravi, da je prirastek notranje energije enak netu dotoka toplote skozi površino, in notranjemu izvoru toplote v telesu $\dot{q}$. Matematična oblika zakona se glasi /1/

$$\int_V c\rho \frac{\partial T}{\partial t} dV = -\int \vec{q}\, d\vec{A} \pm \int \dot{q}\, dV. \qquad (1)$$

V gornji enačbi je c specifična toplota, ρ gostota snovi, $\vec{q}$ pa gostota toplotnega toka. Če ploskovni integral transformiramo z Gaussovim diferenčnim stavkom, dobi gornja enačba obliko

$$\int_V \left(c\rho \frac{\partial T}{\partial t} + \operatorname{div} \vec{q} \pm \dot{q} \right) dV = 0. \qquad (2)$$

Ker je integral za poljuben V enak nič, mora biti tudi integrand nič

$$\pm \dot{q} = c\rho \frac{\partial T}{\partial t} + \operatorname{div} \vec{q}. \qquad (3)$$

Gostota toplotnega toka $\vec{q}$ je vektor, ki je normalen na izotermno ploskev, določimo pa ga po (4)

$$\vec{q} = -\lambda \operatorname{grad} T, \qquad (4)$$

kjer je λ toplotna prevodnost. Če v (3) vstavimo (4), dobimo splošno diferencialno enačbo, ki velja za vsa temperaturna polja

$$c\rho \frac{\partial T}{\partial t} = \operatorname{div} (\lambda \operatorname{grad} T) \pm \dot{q}, \qquad (5)$$

ki je v razviti obliki zapisana kot

$$c\rho\frac{\partial T}{\partial t}=\frac{\partial}{\partial x}\left(\lambda_x\frac{\partial T}{\partial x}\right)+\frac{\partial}{\partial y}\left(\lambda_y\frac{\partial T}{\partial y}\right)+\frac{\partial}{\partial z}\left(\lambda_z\frac{\partial T}{\partial z}\right)+\dot{q} \quad (6)$$

Ker imamo ponavadi opravka z notranjimi izvori toplote smo v (6) pred členom $\dot{q}$ pisali pozitivni predznak in ga bomo v nadaljevanju upoštevali.

Če je snov, skozi katero se prevaja toplota, izotropna ($\lambda_x=\lambda_y=\lambda_z$), preide (6) v (7)

$$c\rho\frac{\partial T}{\partial t}=\lambda\nabla^2T+\dot{q}. \quad (7)$$

Če se temperatura s časom ne spreminja torej je $\frac{\partial T}{\partial t}=0$, dobimo Poissonovo diferencialno enačbo:

$$\nabla^2T=-\frac{\dot{q}}{\lambda} \quad (8)$$

Diferencialne enačbe temperaturnih polj lahko rešimo, če poznamo začetne in robne pogoje. Začetni nam podajajo porazdelitev temperature v času t=0, robni pogoji pa določajo toplotne razmere na površini. Ti so lahko:

- Diricletovi - na površini je predpisana temperatura

$$T=T(x,y,z) \quad (9)$$

- Neumanovi - predpisan toplotni tok

$$\lambda\frac{\partial T}{\partial n}=q(x,y,z) \quad (10)$$

- Cauchyjev - predpisan konvektivni toplotni tok

$$\lambda\frac{\partial T}{\partial n}=\alpha(T_s-T_\infty) \quad (11)$$

kjer je α koeficient toplotne prestopnosti, T_S temperatura trde površine, T_∞ povprečna temperatura okoliškega fluida.

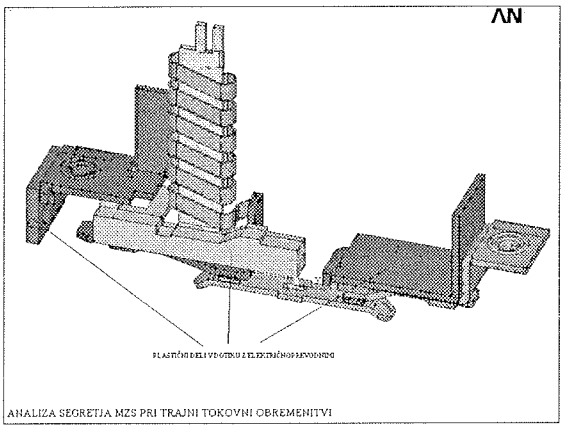
Tako imenovani poseben robni pogoj je, ko je površina toplotno izolirana.

$$\lambda\frac{\partial T}{\partial n}=0 \quad (12)$$

3. Model MZS

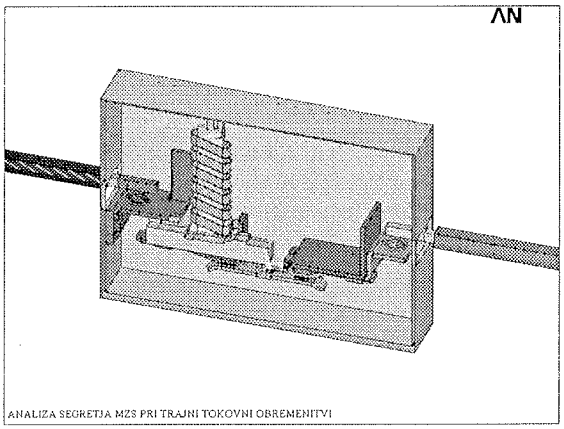
MZS predstavlja zelo kompleksno 3D geometrijo, ki bi jo pri analizi segrevanja težko ali nemogoče zajeli v celoti. Nekateri deli stikala (mehanizem, gasilne komore obloka) na temperaturno polje nimajo skoraj nikakršen vpliv, zato lahko njihovo prisotnost zanemarimo. Tudi obravnavi vseh treh polov se odrečemo, čeprav je že iz samega fizičnega modela razvidno, da je sredinski pol stikala temperaturno bolj obremenjen od zunanjih dveh.

Največji izvor toplote predstavlja grelno navitje z bimetalom, zato lahko vpliv navitja tuljave zanemarimo. S tem se nekoliko spremeni tokovna pot MZS. V stikalu pa je tudi veliko ti. prečnih plastičnih sten (slika 2), ki so v tesnem dotiku tako z električno prevodnimi deli kot tudi z ohišjem stikala. Nastala toplota na prevodnih delih se preko njih odvaja na ohišje in precej vpliva na temperaturno porazdelitev.



Slika 2: Deli tokokroga in prečne plastične stene v notranjosti ohišja MZS

Ohišje takega stikala je prav tako kompleksno, saj vsebuje polno utorov, kanalov, zaskočk, zato tudi tukaj ni šlo brez poenostavitev. Ohišje stikala za en pol smo tako nadomestili s škatlasto obliko, katere dimenzije so približno enake kot pri realnem modelu. Prav tako je pri preskusih segrevanja predpisan presek in dolžina priključnega vodnika. Vodnik mora biti okroglega preseka 6 mm^2 njegova dolžina pa 1 m. Zaradi lažjega modeliranja smo pri simulaciji uporabili vodnik pravokotne oblike, ki ima enak presek kot okrogel vodnik. Končno obliko modela MZS uporabljenega za porazdelitev temperature v stacionarnem stanju nam prikazuje slika 3.



Slika 3: Model MZS uporabljen v simulaciji

4 Robni pogoji

Določitev robnih pogojev predstavlja pri izračunih temperaturnih polj največji problem. Posebej težavno je določevan-

je koeficienta toplotne prestopnosti, ki je odvisen tako od geometrije, lastnosti površine kot tudi od lastnosti medija, ki obkroža geometrijo.

Da skozi stikalo teče tok 32 A, vsilimo na koncih vodnika določen električni potencial. Razlika potencialov med vstopno in izstopno ploskvijo električnega toka vodnika je odvisna od električne upornosti vseh tokovodečih delov stikala. Na teh dveh koncih vodnika smo vsilili tudi temperaturo - Diricletov robni pogoj, ki je enaka temperaturi okoliškega zraka 20 °C.

Med poli tripolnega stikala so vmesne stene, ki pole med seboj ločujejo in predstavljajo toplotno izolacijo, zato velja enačba (11). V nadaljevanju si podrobneje poglejmo načine za določitev koeficienta toplotne prestopnosti in sicer na ohišju stikala (α_{zs}) ter na površini priključnega vodnika (α_v).

4.1 Določanje α_{zs} na površini sten stikala

Kolik je koeficient toplotne prestopnosti α_{zs} s sten na okoliški fluid, je odvisno od več dejavnikov. Najpomembnejši sta oblika stene in vrsta gibanja fluida. Zaradi velikega razpona α je potrebno le tega čimbolj realno določiti.

V tehniki so na voljo trije načini, ki pa terjajo zelo različen trud:

- uporaba približnih izkustvenih vrednosti,
- določanje vrednosti s poskusom in
- izračun vrednosti z empiričnimi formulami.

Pri določanju toplotne prestopnosti na zunanjih stenah stikala smo se poslužili prve variante - izkustvene vrednosti. V literaturah in priročnikih se vrednosti koeficientov toplotne prestopnosti pri naravni konvekciji gibajo med 2 in 30 W/m^2 K. Zaradi velikega razpona smo izračun opravili za dve vrednosti in sicer za α_{zs}=15W/m^2 K in α_{zs}=20W/m^2 K, ki smo jo kot robni pogoj postavili na vse stene razen na tisto, ki meji na sosednji pol stikala. Slednja je namreč toplotno izolirana, torej velja (12).

4.2 Določanje α_v na površini vodnika

Za določitev koeficienta toplotne prestopnosti na površini vodnika (a_v) smo se poslužili empiričnih obrazcev, ki temeljijo na teoriji podobnosti. Uveljavila se je cela vrsta ti. kriterialnih števil, od katerih omenimo le tista, ki so za nas pomembna /4,7/

Grasshofovo število

$$G_r = \frac{g \cdot \beta \cdot \Delta T \cdot h^3}{\nu^2} \qquad (13)$$

Prandtlovo število

$$P_r = \frac{\eta \cdot c_p}{\lambda} \qquad (14)$$

Rayleighevo število

$$R_a = P_r \cdot G_r \qquad (15)$$

Nusseltovo število

$$N_u = \frac{\alpha \cdot L}{\lambda} \qquad (16)$$

V gornjih enačbah je g pospešek prostega pada, β temperaturni razteznostni koeficient fluida, ΔT razlika temperature med fluidom in trdo površino, h višina pri navpični steni ali valju, ν kinematična viskoznost fluida, c_p specifična toplota fluida pri konstantnem tlaku in L karakteristična dolžina.

Koeficient toplotne prestopnosti je najlažje določiti s pomočjo Nusseltovega števila. To število za horizontalni valj določimo po (17)

$$N_u = C \cdot R_a^{\,n}, \qquad (17)$$

kjer sta konstanti C in n odvisni od vrednosti Rayleighevega števila, podani po Michejew-u v tabeli 1 /8/.

Tabela 1: Vrednosti konstant C in n v odvisnosti od R_a

R_a	C	n
$< 10^{-3}$	0,5	0
$10^{-3} \div 5\ 10^{2}$	1,18	0,125
$5\ 10^{2} \div 2\ 10^{7}$	0,54	0,25
$2\ 10^{7} \div 2\ 10^{13}$	0,135	$0,\overline{3}$

Vse snovne vrednosti fluida (zraka) se morajo vzeti pri aritmetični sredini temperatur vodnika in zraka, ki ga obkroža.

Temperatura okoliškega zraka (T_∞) je bila pri izračunu 20°C, povprečno temperaturo izolacije vodnika pa smo ocenili na 25°C. Vse snovne vrednosti za zrak bi torej morali jemati pri temperaturi 22,5°C. Ker v tabelah snovne vrednosti niso podane za to temperaturo smo le-te vzeli pri 20°C /9/.

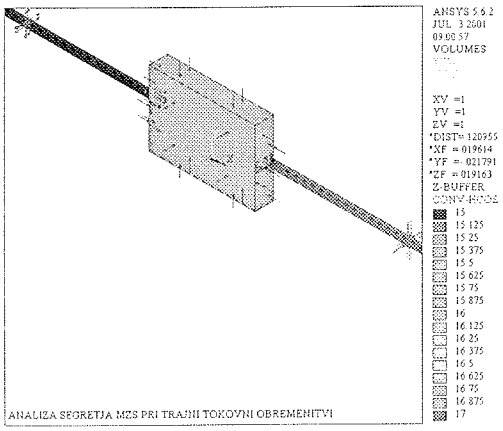
Tabela 2: Snovne vrednosti zraka pri normalnem zračnem tlaku

c_p (kJ/kg K)	η (kg/m s)	ν (m^2/s)	β (1/K)	λ (W/m K)
1,007	$18,24\ 10^{-6}$	$153,5\ 10^{-7}$	$3,421\ 10^{-3}$	0,02569

Po Michejew-u je h v (13) za horizontalen valj pri naravni konvekciji enak

$$h = \pi \cdot r, \qquad (18)$$

kjer je r polmer vodnika z izolacijo in znaša 2,18 mm. Z upoštevanjem enačb od 13 do 18 in tabel 1 in 2 dobimo vrednost koeficienta na površini vodnika 17 W/m^2 K.



Slika 4: Robni pogoji

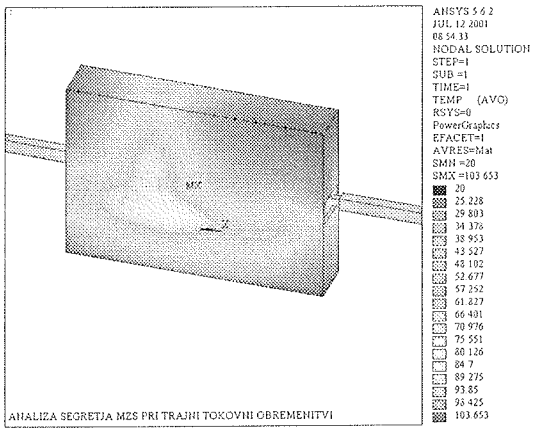
5. Rezultati

Izračuni, ki so bili narejeni s programskim paketom ANSYS, so prikazani za dve različni vrednosti koeficienta toplotne prestopnosti na površinah ohišja stikala (α_{zs}) in pri konstantni vrednosti koeficienta toplotne prestopnosti na površini vodnika (α_v). Temperatura okoliškega zraka je pri obeh izračunih bila 20°C.

Tabela 3: Koncept toplotnega izračuna MZS

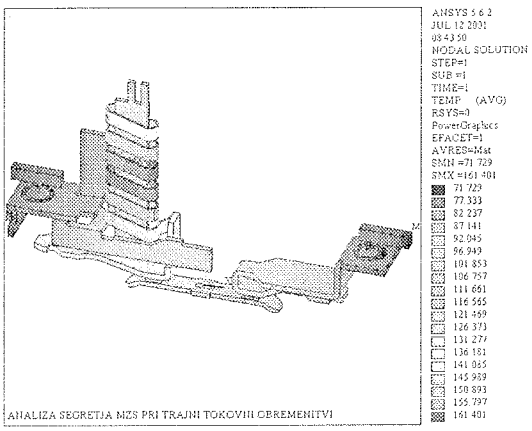
α_{zs} (W/m^2 K)	α_v (W/m^2 K)	T_∞ (°C)
15	17	20
20		

5.1 Analiza temperaturnega polja pri α_{zs}=15 W/m^{2}K



Slika 5: Temperaturno polje celega modela stikala

Sliki 5 in 6 nam podajata porazdelitev temperaturnega polja pri α_{zs}=15 W/m^{2}K na stenah stikala. Iz slike 5 je razvidno, da se temperatura na ohišju giblje med 30°C in 70°C, kar je še v mejah standarda IEC EN 60947, ki pravi, da se lahko zunanja stran plastičnega ohišja segreje za 50°C nad temperaturo okolice, ki je v našem primeru 20°C. Maksi-



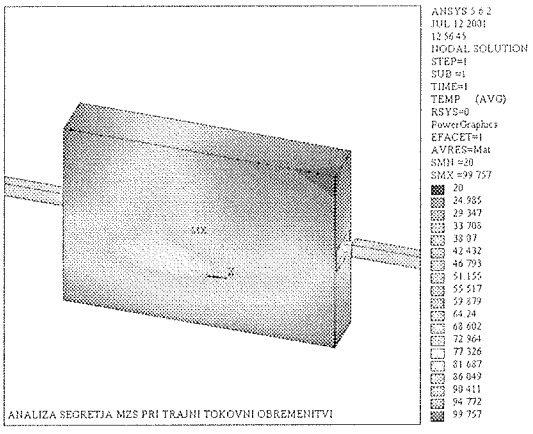
Slika 6: Temperaturno polje električno prevodnega dela

malna temperatura električno prevodnega dela je na srednjih ovojih grelnega navitja in znaša okoli 161°C (slika 6). Da je maksimalna temperatura prav na srednjih ovojih grelnega navitja, ima vzrok v odvajanju toplote z obeh koncev grelnega navitja na z njim spojenima dela stikala.

Iz slike 6 je tudi razvidno, da je kontaktni del na strani bimetala in grelnega navitja temperaturno bolj obremenjen. Prav tako je tudi temperatura priključka na strani bimetala in grelnega navitja za približno 1,5°C višja kot na delu, kjer teh dveh elementov ni in znaša 73°C.

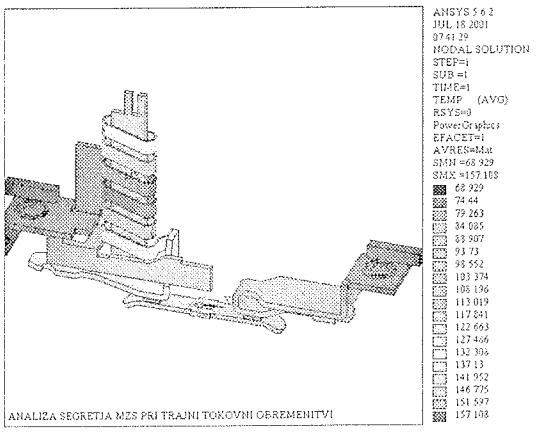
5.2 Analiza temperaturnega polja pri α_{zs}=20 W/m^{2}K

Oblika temperaturnega polja je podobna kot v primeru nižjega koeficienta toplotne prestopnosti na stenah stikala (sliki 5 in 6).



Slika 7: Temperaturno polje celega modela stikala

Višina temperature na posameznih delih stikala se je v povprečju znižala za 2 do 4°C. Tako je maksimalna temperatura na grelnem navitju sedaj okoli 157°C, temperatura priključka na strani bimetala in grelnega navitja pa 70°C.



Slika 8: Temperaturno polje električno prevodnega dela

6. Zaključek

Z uporabo programskega paketa ANSYS (ki temelji na metodi končnih elementov) smo preučevali porazdelitev temperaturnega polja motorskega zaščitnega stikala. Uporabljen 3D model pomeni določeno poenostavitev glede na dejanske razmera v stikalu, kljub temu pa je dovolj podroben in občutljiv, da zajema vpliv vseh tistih pojavov, ki v največji meri vplivajo na temperaturo polje v stikalu.

Iz rezultatov analize je razvidno, da temperatura na priključnih sponkah nikjer ne presega podanega RTI indeksa, ki je za uporabljen izolacijski material 105°C, pri čemer je zunanja temperatura 20°C. Ta temperatura se lahko pri testih segrevanja priključnih sponk po standardu UL 508 in IEC EN 60947 giblje med 10 in 40°C.

Segrevanje MZS je kompleksen proces, ki ga je mogoče raziskati z obsežnimi meritvami in numeričnimi študijami. V prihodnosti se bo potrebno osredotočiti tudi na konvektivni prenos toplote v notranjosti stikala in na določevanje kontaktne upornosti, ki ima lahko velik vpliv na segrevanje. Prav tako se s temperaturo spreminja specifična ohmska upornost električno prevodnih materialov, kar še dodatno povečuje izgube in s tem segrevanje stikala.

7. Literatura

/1/ A. Alujević, P. Škerget, *Prenos toplote*, Fakulteta za strojništvo Maribor, 1990

/2/ C. Groth, G. Müller, *FEM für Praktiker - Die Methode der Finiten Elemente mit dem FE-Programm ANSYS*, 4. Auflage, Grafing, 1997.

/3/ C. Groth, G. Müller, *FEM für Praktiker - Temperaturfelder*, 2 Auflage, Technische Akademie Essligen Weiterbildungszentrum DI Elmar Wippler, 1998.

/4/ D. Pitts, L. Sissom, *Theory and problems of heat transfer*, Second Edition, New York, McGraw-Hill, 1997.

/5/ E. Hering, R. Martin, M. Stohrer, Wärmeübertragung, *Physik für Ingenieure*, 6. Auflage, Springer-Verlag, Berlin, 1997,str. 203-219.

/6/ E. Prelog, *Reševanje enačbe za prenos toplote z metodo končnih elementov*, Strojniški vestnik, Ljubljana, 1973.

/7/ F. P. Incropera, D. P. DeWitt, *Introduction to heat transfer*, third edition, John Wiley & Sons, New York, 1996

/8/ G. Erbe, H. J. Hoffmann, Wärmeübertragung, *Einführunh in die Warmelehre*, 7 Auflage Carl Hanser Verlag, München, 1986, str. 292-309.

/9/ *Berechnungsblätter für den Warmeübertragung*, Verein Deutscher Ingenieure, Auflage 5, VDI Verlag Dusseldorf, 1988.

/10/ J. Pajer, *Nizkonapetostna stikala v praksi*, Kranj, 1997

Gorazd Hrovat, univ. dipl. inž. el.,
e-mail: gorazd.hrovat1@uni-mb.si
izr. prof. dr. Anton Hamler,
e-mail: anton.hamler@uni-mb.si
izr. prof. dr. Mladen Trlep,
e-mail: mladen.trlep@uni-mb.si

Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko
Smetanova ulica 17, 2000 Maribor
tel. (02) 220 70 00

doc. dr. Martin Bizjak
e-mail: martin.bizjak@iskra-stikala.si

Iskra Stikala d.d.
Savska loka 4, 4000 Kranj
tel. (04) 237 22 26

Prispelo (Arrived): 21.05.2002 *Sprejeto (Accepted): 25.03.2003*

SENSOR OF FORCES IN SMALL VOLUME CONTRACTING TISSUES

[1]Matjaž Bunc and Janez Rozman

ITIS d.o.o. Ljubljana, Centre for Implantable Technology and Sensors, Ljubljana, Slovenia

[1]Institute of Pathophysiology, Ljubljana, Slovenia

Key words: sensor of forces, strain gauge, physiology, contracting tissue

Abstract: A single-channel sensor intended for measurement of forces in small volume contracting tissues within the range of mN was designed, developed and experimentally tested. The force sensor was made up of a Wheatstone bridge composed of four semi-conductor strain gauges bonded on a specially designed cantilever with a handle and metallic cover to protect them. The natural frequency of the sensor is 350Hz while the compliance is $5.7X10^{-6}$m/mN. The sensor represents a very linear dependence of the output voltage upon the load. The sensibility of the sensor, at a bridge excitation voltage of 5V, is 0.5mV/mN and the nominal range of the sensor is 0-70mN. Results show that the sensor enables almost isometric measurements of forces in contracting tissues. The results also show that the sensor measures forces with a frequency of up to 300Hz with appropriate accuracy. Finally, the sensor is suitable for isometric measurements of forces in all types of contracting tissues.

Senzor sile za meritve krčenja drobnih mišic

Ključne besede: senzor sile, strain gauge, fiziologija, kontraktilna tkiva

Izvleček: Izdelali in testirali smo enokanalni mehansko-električni pretvornik za meritve moči krčenja drobnih mišic. Senzor sile je izdelan iz štirih polprevodniških strain-gaugov povezanih v Wheatstonov mostiček, pritrjenih na rigidno merilno ročico ter zaščitenih s kovinskim oklepom. Naravna frekvenca senzorja je 350 Hz in podajnost 5.7×10^{-6} m/mN popolnoma ustreza za meritve skoraj izometričnega krčenja drobnih mišičnih tkiv (<5 N, <50 Hz). Občutljivost senzorja z napajanjem 5 V je v razponu 0-70 mN linearna, 0.5 mV/mN. Testiranja so pokazala, da je mogoče s senzorjem zadovoljivo meriti tudi krčenja s frekvenco blizu 300 Hz. Merilec je primeren za meritve sile izometričnega krčenja različnih krčljivih tkiv.

Introduction

In basic research related to the physiology of muscles and other contracting tissues, especially in studies of contraction mechanisms, measurements of elicited forces are very important /1, 2, 3, 4/. Transducers made of strain gauge bridges give the opportunity to develop highly sensitive and reliable force sensors /5, 6, 7/. The development of sensors to measure muscle or vein contractions has a history at least 30 years long and many devices are now commercially available [Axon Instruments (USA), Experimentaria (HUN), 8, 9]. The force sensor should fulfill the following requirements: a) it should be able to evaluate the force elicited by contraction of the muscles and most other contracting tissues, b) electrical response should be as linear as possible in the whole range of expected forces, and c) nevertheless, the sensor should react fast enough to be able to follow a contraction as reliably as possible /10, 11/. The goal of our work was therefore to develop a force sensor that would fit all of the above-mentioned requirements and be easy to manipulate, as a part of the device, especially when testing muscles that are difficult to access.

Material and methods

In the majority of cases of measuring forces of muscle tissues *in vivo* the force sensor should be able to approach the desired contracting tissues at desired angles. Therefore, we designed the cantilever and its holder in such a way that it could move linearly and rotate within a limited space. However, in some cases of measuring forces *in situ* the sensor should be able to be mounted in any position required according to the protocol measurements especially for pharmacological purposes. To obtain adequate characteristics of the force sensor the measurements should be obtained in the way in which the tissues are attached to the force sensor always perpendicularly to the direction of the measured contraction. The cantilever, shown in Fig. 1, was made of highly tempered stainless steel ribbon machined out of the stainless steel bar that acts as the handle of the force sensor at the same time. The dimensions of the section of the cantilever that could be bent upon the applied force were defined according to the request of the gauge's manufacturer (Celesco, USA) and our request to develop the force sensor that would be enough sensitive to measure the elicited forces.

The force sensor itself was made up of a Wheatstone bridge composed of four semi-conductor strain gauges (Celesco, P05-02-500, resistance in ohms: 500.0 ± 0.3%),

bonded on a section of the cantilever where bending is of the highest degree. Strain gauges were actually bonded according to the procedure described by both the manufacturer of the strain gauges and the producer of the adhesive (Micro Measurements, M-Bond 610). The mechanical tension produced by the force within the nominal range elicits elastic deformation of the cantilever, thus resulting as a change of the output voltage that could be amplified and connected to the A/D converter and IBM Compatible PC.

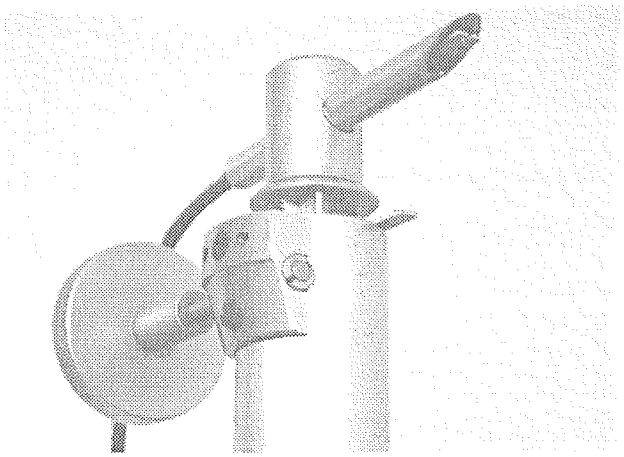


Figure 1. A force sensor mounted within the vice.

The static characteristic of the force sensor was obtained by mechanical connection of the reference force sensor (*ITIS d.o.o., Slo*) to the cantilever of the developed force sensor perpendicularly to the level that is supposed to be the acting point of the measured forces. Then the same manual force was applied to both sensors, exciting them to elicit corresponding voltage signals at the outputs. The output voltage of the developed force sensor was fed to the X input of the X-Y oscilloscope while output voltage of the reference force sensor was fed to the Y input of the oscilloscope. When the applied force reached a nominal value of 70mN, the developed force sensor force was removed. In this way the static charasteristic in both directions was obtained. The compliance of the cantilever was also measured during the recording of the static charasteristics. For this purpose a laser beam was directed to the cantilever and the distance between reflexions when the force sensor was loaded and unloaded was measured. Dynamic behavior of the sensor was defined by eliciting mechanical vibration of the cantilever by striking it with a finger. The output signal of the Wheatstone bridge was amplified at a gain of 100 and fed to a DigiPack 1200 (Axon Instruments) acquisition system and sampled at 2kHz. The recorded data was analyzed using a (*Matlab*) software package enabling Fourier analysis. The dynamic behavior of the force sensor was also tested by a sharp thrust of force applied on the force sensor generated when a weight of 2.75g was dropped on the cantilever of the force sensor from a height of 5cm.

Results

Fig. 2 shows the static characteristics of the force sensor. Considering the data obtained in the aforementioned characteristics and amplification of the output signal, the sensitivity of the force sensor was calculated. Calculations showed that over the nominal range of 0-70mN the sensibility is 0.5mV/mN. The results also showed that the compliance was about 5.7×10^{-6}m/mN. Furthermore, from Fig. 3, showing the dynamic characteristics of the force sensor, the natural frequency and the data describing a behavior of the force sensor below the natural frequency was obtained. The corresponding natural frequency of the force sensor is 350Hz.

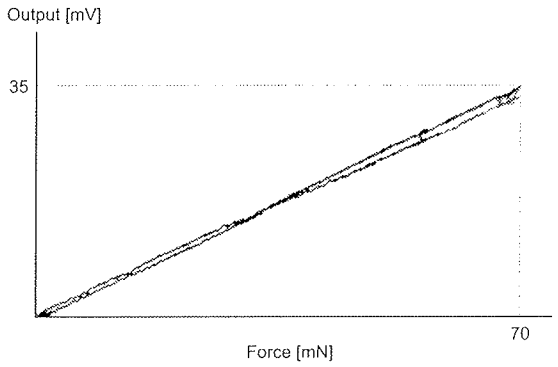


Figure 2. The static characteristic of the force sensor.

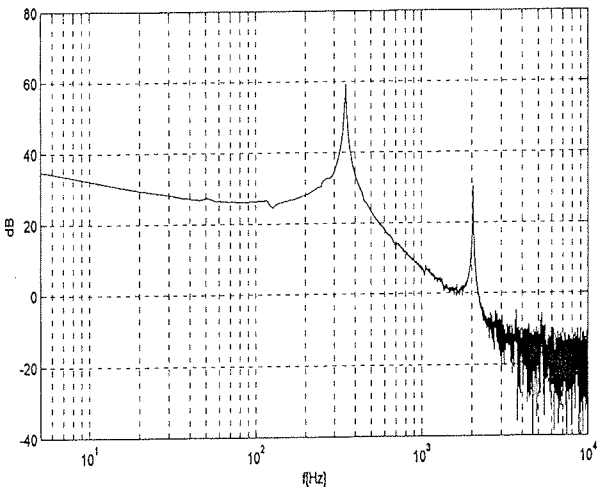


Figure 3. The dynamic characteristic of the force sensor.

Figure 4, however, represents the record of sharp thrust of force applied on the force sensor. It coud be seen that the time from zero to peak force when a weight of 2.75g was dropped on the cantilever of the force sensor from a height of 5cm was 55.23ms.

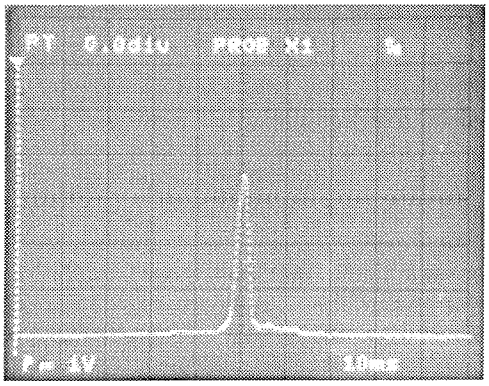


Figure 4. A record of sharp thrust of a force applied on the force sensor.

Discussion

According to the requirements determined in the methods section we can conclude that all of the requirements defined at the start of the development of the force sensor were met. Namely, as can be seen in Fig. 2, the force sensor shows a highly linear dependence of the output voltage upon the load within the whole nominal range of applied forces. However, the different slopes in the trajectories (Fig. 2) arising in opposite directions of the applied force were attributed exclusively to mechanical sliding that occurred at the contact between the two sensors when mounted within the vice. By sliding, the length of the reference sensor was slightly changed. This could be confirmed by the intersection of the two trajectories which is settled exactly at the middle between the points representing zero and maximum load. Moreover, it could be seen in Fig. 3, showing the frequency spectrum contained in the response of the force sensor on the strike applied to the cantilever, that the natural frequency of the sensor is equal to 350Hz. At frequencies below the natural frequency of the force sensor the curve is relatively linear, enabling accurate measurements of forces. Since the natural frequency and sensitivity of the force sensor are relatively high one could conclude that the force sensor is able to record fast changes that could be expected in muscle contraction, also of only a few muscle fibers, without any deformation. In Fig. 4. representing a record of sharp thrust of force applied on the force sensor, one can see that the rise time of 55.23ms corresponds to a frequency of 18Hz, which is far below the natural frequency of the force sensor. Accordingly, the force sensor is also suitable for recording sustained tonic contractions of a muscle. The limitation of the force sensor is that it is designed to measure contractions of 70mN maximum. This sensor with a compliance of 5.7×10^{-6}m/mN enables the recording of forces in a predominantly isometric mode. However, the higher the force, the more isotonic the contraction appears. The important advantage of the force sensor is that it could be orientated in space to reach any point of the hemisphere with a diameter of 15cm from the vertical handle. Because of that the measurements of contracting tissues could always be perpendicular to the axis of the measured muscle contraction. Our force sensor is relatively simple, reliable and is sold at an acceptable price.

Acknowledgement

This work was financed by the following research grants: J2-3415 from the Ministry of Education, Science and Sport, Ljubljana, Republic of Slovenia, and HPRN-CT-2000-00030 from the European Commission.

References

/1./ R. A. Meiss and E. H. Sonnenblick, Controlled shortening in heart muscle: velocity-force and active-state properties, *Am. J. Physiol.*, 222(3) (1972) 630-639.

/2./ P. D. Soden and I. Kershaw, Tensile testing of connective tissues, *Med. Biol. Eng.*, 12(4) (1974) 510-518.

/3./ J. Rozman, B. Zorko, B. and T. Nghiem, Isometric twitch contractions of selectively stimulated muscles in dog´s leg, *Basic and Applied Myology, 4(2)* (1994) 155-163.

/4./ C.S. Fulco, P.B. Rock, S.R. Muza, E. Lammi, A. Cymerman, G. Butterfield, L.G. Moore, B. Braun, S.F. Lewis. Slower fatigue and faster recovery of the adductor pollicis muscle in women matched for strength with men. Acta Physil. Scand., 167(3) (1999) 233-239.

/5./ C.J. De Ruiter, D.A. Jones, A.J. Sargeant, A. De Haan. The measurement of force/velocity relationships of fresh and fatigued human adductor pollicis muscle. Eur. J. Appl .Physiol. Occup. Physiol., 80(4) (1999) 386-393.

/6./ R. A. Meiss, A versatile transducer system for mechanical studies of muscle, *J. Appl. Physiol.*, 37(3) (1974) 459-463.

/7./ R. A. Meiss, An isometric muscle force transducer, *J. Appl. Physiol.*, 30(1) (1971) 158-160.

/8./ J. Rozman, J. Bratanič, B. Sovinec, B. Lenart, A. Jeglič and D. Fefer, Four channel transducer for evaluation of muscle contractions, *FES Conf., Ljubljana, Republic of Slovenia,* (1993) 22-25.

/9./ M. Bunc, J. Rozman and D. Šuput, Measurements of gill movement in fish and water flow through fish mouths using force and pressure transducers, *6th Vienna Workshop on FES, Vienna, Austria, September,* (1998) 22-24.

/10./ A. F. Huxley and R. M. Simmons, A capacitance-gauge tension transducer, *J. Physiol.,* 197(1) (1968) 12P.

/11./ J. S. Petrofsky and C. A. Philips, Determination of the contractile properties of the motor units in skeletal muscle through twitch characteristics, *Med. & Biol. Eng. & Comput., 17* (1979) 525-535.

Corresponding author:
Dr. Janez Rozman
ITIS d. o. o. Ljubljana
Center for Implantable Technology and Sensors
Lepi pot 11, 1001 Ljubljana
Republic of Slovenia

Tel.: ++386 1 470 19 13
Fax.: ++386 1 470 19 39
E-mail: janez.rozman@guest.arnes.si

Prispelo (Arrived): 08.05.2002 Sprejeto (Accepted): 25.03.2003

VLOGA SENZORJEV V SISTEMIH VODENJA PROCESOV

[1]Rihard Karba, [1]Maja Atanasijević-Kunc, [1]Aleš Belič
[2]Juš Kocijan, [2]Janko Petrovčič

[1]Fakulteta za elektrotehniko, Univerza v Ljubljani, Ljubljana, Slovenija
[2]Institut Jožef Stefan,Ljubljana, Slovenija

Ključne besede: sistemi vodenja, merilni sistemi, senzorji, regulacijska zanka, življenjski cikel projektov vodenja.

Izvleček: Pri načrtovanju in vzdrževanju sistemov vodenja v industriji igrajo senzorji odločilno vlogo. Inženirji avtomatike morajo namreč upoštevati dejstvo, da je nemogoče regulirati, če ni na razpolago ustrezno natančnih meritev. Pri tem je pomembna tudi njihova izvedba, saj morajo biti prilagojene okolju, v katerem delujejo, tako po zgradbi kot tudi v smislu dinamičnih lastnosti. Zato je izbira, vgradnja in vzdrževanje merilnih sistemov kompleksen postopek, ki je odločilen za uspešno delovanje obravnavanega procesa.

The Role of Sensors in Control Applications

Key words: control systems, measuring systems, sensors, control projects life-cycle.

Abstract: In the process of control system design and its maintenance sensors play an important role. Control engineers are namely faced with the fact that any control is based on attainability of correspondingly accurate measurements of system outputs. The realisation of measurements must be adapted to the system characteristics and environment. Their dynamic aspects must also be taken into account. Therefore, the choice, mounting and maintenance of measuring systems is complex procedure which is crucial for the successfulness of control systems.

Modern control technology is among most important factors concerning successfulness and progress of state or even world economy. It has infrastructural character which means that its effects can be evaluated through many engineering, economic, social, and other activities. As the requirements of industrial organisations are more and more complex only interdisciplinary groups of experts can tackle such problems efficiently.

Term control represents the procedures which inuence the process behaviour in a way that some previously predicted goal can be achieved. Here open loop (sequential) and closed loop (feedback) control are included.

Control engineers are aware of the fact that only the process variable which can be

accurately measured can also be successfully controlled. Unreliable measurements can be filtered by signal processing only in special cases. Control loop can compensate, to a certain extent, some properties of elements in the loop or disturbances from environment. However, the long term deviation of sensor output, its non-linearity, significant time delay etc. always cause undesired phenomena in the loop behaviour. The comparison between reference (desired) and measured (achieved) value of process variable, which represents the essence of feedback control by generating the error, which is the basis of control action, is not ideal. However, the sensor output is never identical to the value of process variable. If the value of process variable and sensor output are not in good agreement, poor control quality is unavoidable regardless of the rest of the control system. In the control of sequential, batch, semi-batch and continuous process or their combinations, beside simple well known sensors, more and more complex measurement equipment with a lot of knowledge and intelligence incorporated is needed. Measurements of quantitative and qualitative properties of media are very delicate and expensive, therefore, the choice of corresponding sensors is of high importance.

Experts are still arguing, whether everything previously referred to as detector, transducer, transmitter, field device, etc. can be included into the term sensor. However, nowadays it can be concluded that the term sensor represents whole measurement system with all necessary element and signal converters. A very loose definition of sensor could be:A sensor is a black box which converts knowledge of process parameters or outputs as well as about product characteristics into usable information. This description could be

extended in different ways but nevertheless defines the main point: sensors are concerned with the procurement of information. Some conclusions can be summarised as follows:

- significance of sensor technology for international economics is much greater than actual magnitude of sensor production,
- Europe, USA, and Japan divide the world market fairly equally among themselves,
- sensors (only for the area of process automation) contribute approximately one third of the complete investment in process control.
- standard sensors are rather the exception than the rule, custom-made sensors are becoming the usual choice,
- the producers of sensors are huge multinational companies on one side and small producers occupying certain niches in the market on the other,
- trends in sensor production are moving towards compact and miniature devices, enabling better characteristics and lower prices.

1 Uvod

Sodobna tehnologija vodenja spada med t. i. kritične tehnologije, to je tiste, ki so v državnem in celo svetovnem merilu posebej pomembne za uspešnost in napredek gospodarstev. V nasprotju z drugimi tehnologijami, ki dajejo pretežno vidne izdelke in dobrine, pa ima tehnologija vodenja bolj posreden, tj. infrastrukturni pomen. Njeni učinki se namreč prepletajo z mnogimi inženirskimi, ekonomskimi, družbenimi in drugimi dejavnostmi.

Področje tehnologije vodenja je izrazito interdisciplinarnega značaja. Potrebno je povezovanje znanja o sistemih oz. procesih, ki jih želimo voditi, z znanjem iz tehnologije vodenja. Poleg klasičnih zahtev po minimizaciji surovin, energije, časa izdelave in cene so dandanes še zahteve po eksi-

bilnosti in zanesljivosti proizvodnje, po kvaliteti izdelkov, pa tudi po varnosti in humanizaciji delovnih mest ter po varovanju okolja. V teh razmerah so uspešnejše interdisciplinarne skupine strokovnjakov, ki poleg svojega specialnega področja poznajo tudi osnove tehnologije vodenja.

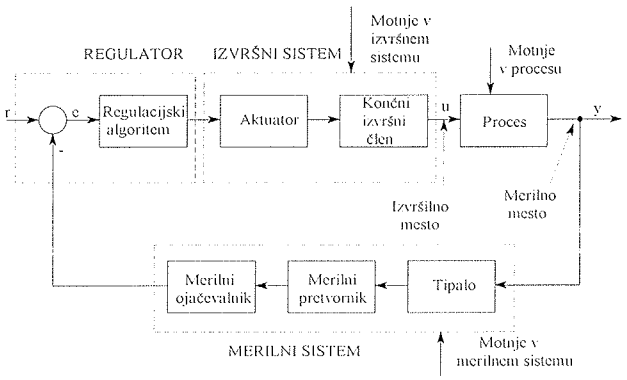
Vodenje je proces, s katerim vplivamo na delovanje sistema z namenom, da dosežemo neki zastavljeni cilj. Gre torej za transformacijo informacij o vodenem procesu in o njegovem okolju v odločitve in ukrepe, ki ob upoštevanju meril in omejitev zagotavljajo želeno vodenje sistema. Pri tem pojem vodenja zajema tako odprtozančno (sekvencno) vodenje - krmiljenje kakor tudi zaprtozančno (povratnozančno) vodenje - regulacijo /1, 2, 3/.

S področjem vodenja sta neločljivo povezana tudi pojma kibernetika, ki se ukvarja s študijem vodenja in komunikacij v živih bitjih in tehničnih sistemih, ter avtomatika, ko se procesi zbiranja informacij o stanju sistema in okolice, priprave ukrepov in odločanje ter ukrepanje opravljajo brez posredovanja človeka. Prav tako pa seveda ne gre brez komunikacijskih, informacijskih in računalniških tehnologij.

V delu predstavljamo pogled inženirja tehnologije vodenja - avtomatika na področje senzorike.

2 Merilni sistemi

Dosegljivost in izbira pravega merilnega sistema je prvi in obenem eden od ključnih korakov pri načrtovanju in realizaciji vodenja nekega procesa. S slike 1 je razvidno, da merilni sistem sestavlja več elementov. Pri delitvi prihaja mnogokrat tudi do terminoloških nesporazumov. V smislu izrazov, ki se pojavljajo v literaturi, bi lahko rekli, da je tipalo ali senzor (prvotno tudi detektor) primarni element merilnega sistema, ki je v neposrednem fizičnem stiku z merjenim medijem. Merilni sistem v tuji literaturi imenujejo pretvornik ali transducer. Sekundarni element merilnega sistema pa je eden ali več merilnih pretvornikov ali prenosnikov (transmitterjev), ki iz izhodne veličine tipala tvorijo uporaben signal, ki ga potrebujemo bodisi za prikazovanje rezultatov meritev, bodisi v regulacijski zanki. Medtem ko imajo primarni elementi raznovrstne konstrukcije in njihovo delovanje temelji na najrazličnejših fizikalnih principih, pa sekundarni elementi težijo k čim več splošnim, skupnim lastnostim in je zato izvedb precej manj. Naloga merilnega sistema je torej meritev neke fizikalne veličine, kar pomeni v bistvu prenos neke informacije obenem s prenosom energije. Zato vse meritve v nekem smislu vplivajo tudi na merjeno veličino, kar kaže na dejstvo, da mora biti proces merjenja zelo skrbno načrtovan /3, 4, 5/. Tipala s pripadajočimi merilnimi pretvorniki so osnova avtomatiziranega vodenja procesov. Procesno veličino, ki jo kvalitetno merimo, lahko v večini primerov tudi uspešno reguliramo. Le v izjemnih primerih lahko slabo in nezanesljivo meritev nadomesti posebna računalniška obdelava signalov. Regulacijska zanka ima sicer čudovito lastnost, da nadomešca oziroma kompenzira neidealnosti elementov zanke in vplive motenj, ne more pa zmanjšati npr. dolgoročnega odmika izhodne vrednosti tipala, njegove nelinearnosti, dolge zakasnitve, neponovljivosti itd. brez bistvenega vpliva na dinamične lastnosti regulacijske zanke. Najbolj kritično mesto vsake regulacijske zanke je namreč mesto primerjave med želeno in dejansko vrednostjo procesne spremenljivke. Regulacijska akcija tako ne temelji neposredno na pogrešku med referencno in regulirano veličino, temveč na pogrešku, ki je odvisen od izhoda merilnega sistema. Če med signalom na mestu primerjave in pripadajočo procesno veličino ni dobre povezave, je proces voden slabo.



Slika 1: Prikaz gradnikov regulacijske zanke, kjer je r referenca, u regulirna, y regulirana veličina, e pa je pogrešek

Tipala lahko razdelimo na proporcionalna in stopenjska. Za regulacijo potrebujemo predvsem proporcionalna. Stopenjska tipala (npr. končna, tlačna in nivojska stikala, fotocelice za detekcijo plamena, pretočna stikala, senzorji bližine itd.) uporabljamo v različne namene, in sicer za:

- alarmiranje
- označevanje začetkov in koncev šarž
- preprečevanje nevarnih situacij v primeru izpada regulacije
- signalizacijo stanja procesa
- sekvencno (logično) vodenje šaržnih in semišaržnih procesov
- v izdelčni proizvodnji.

Proporcionalna tipala pa delimo na osnovna (premik, hitrost, pospešek, sila, nivo, pretok, tlak, temperatura itd.) in na zahtevnejša. Potrebe po objektivnem določanju kvalitativnih in kvantitativnih lastnosti snovi tako v laboratorijskem okolju kot tudi v industrijskih procesih namreč postajajo vse večje. Posebno pri slednjih mora biti merilna oprema prilagojena mnogo težjim razmeram, kot pa so pri laboratorijskih meritvah (korozija, nečistoče, velike spremembe temperature in/ali tlaka itd.). Tako je pravilni izbor merilnika za analizo določene snovi odvisen od fizikalnih in kemičnih lastnosti vzorcev, od značilnosti procesa, od njegove okolice in navsezadnje tudi od poznavanja delovanja merilnikov ter celotnega merilnega sistema. Naštejmo nekaj lastnosti materialov, ki jih pogosto merimo: gostota, viskoznost, vlažnost, toplotna in električna prevodnost, pH-faktor, redoks in vsebnost ter prisotnost ali koncentracija različnih komponent plinov v mediju, kjer gre za analizne postopke.

Merjenje lastnosti snovi in analizne meritve so še bolj kompleksne od merjenj osnovnih procesnih veličin. Meritve so mnogokrat tudi posredne in zahtevajo včasih izredno obsežno in drago opremo (tudi več velikostnih razredov dražjo od merilnikov običajnih procesnih veličin). Zato mora biti taka oprema še posebno skrbno načrtovana, izbrana in vzdrževana. Zanima nas seveda sprotni način merjenja, katerega izhodi bi bili uporabni v regulacijski zanki. Pri tem se moramo zavedati, da je mnogo predvsem analiznih merilnikov šaržnega tipa (analiza poteka na nizu vzorcev), kar daje tudi regulacijski zanki diskretni značaj. Prav tako pa je jasno, da je dinamika tovrstnih merilnikov mnogokrat počasna, pri čemer prihaja tudi do nezanemarljivih časovnih zakasnitev, ki lahko zelo vplivajo na vedenje zanke (stabilnost).

Naštejmo še nekaj najvažnejših dejavnikov, ki jih moramo obenem s specifičnimi zahtevami upoštevati, ko izbiramo merilno opremo za vodenje nekega industrijskega procesa. Pri tem naj opozorimo, da je v nekaterih primerih na razpolago zelo široka izbira komercialno dosegljivih alternativ (npr. meritev temperature, tlaka, pretoka, nivoja, odmika, itd.), medtem ko včasih praktično ni nobene izbire (specialne meritve). Poznavanje možnih alternativ je seveda pri izbiri bistvenega pomena. Če upoštevamo, da je običajno merilni člen vgrajen v regulirani objekt, lahko na izbiro vplivajo predvsem naslednji dejavniki:

- fizična kompatibilnost z reguliranim objektom
- odpornost glede na okolico objekta
- kompatibilnost z regulatorjevimi signali
- zahteve v zvezi z napajanjem in energijo
- razmerje signal - šum in ponovljivost meritev
- dinamične lastnosti (hitrost odziva, linearnost, točnost, merilno območje itd.)
- poreba po vzdrževanju (obstojnost, življenjska doba, zanesljivost, način vgradnje, dosegljivost, potreba po kalibraciji itd.)
- cena.

Pri tem moramo upoštevati tudi koncepte življenjskega cikla sistemov vodenja /2/, ki se začne in konca pri uporabniku. Začetek je identifikacija oz. definicija potreb, cikel pa se nadaljuje s planiranjem, z raziskavami, načrtovanjem, s proizvodnjo ali z gradnjo, vrednotenjem, vgradnjo in uporabo, vzdrževanjem in s podporo pri uporabniku ter konca z "upokojitvijo" sistema ali proizvoda.

3 Sklepi

Vsako obdobje rodi svoje izrazoslovje. Tako so v obdobju računalništva začeli uporabljati izraz senzor v širšem smislu, torej v smislu celotnega merilnega sistema. Osnovno vprašanje, kaj naj bi bil potemtakem senzor, še vedno ni povsem razjasnjeno. Zelo široka in ohlapna definicija bi se glasila:

Senzor je skrinjica, ki pretvori znanje o procesnih parametrih in odzivu sistema ali pa o značaju produktov v uporabno informacijo.

Tudi tej definiciji bi bilo mogoče še marsikaj dodati, vsekakor pa velja, da so senzorji povezani s pridobivanjem in posredovanjem nekih informacij o tehnoloških ali bioloških sistemih.

Tehnološki razvoj senzorjev je izredno raznolik, saj izhaja iz realnih zahtev najrazličnejših okolij. Zato se ob široki izbiri standardiziranih senzorjev vedno bolj razvijajo tudi mnoge specialne izvedbe, bodisi za posameznega uporabnika ali pa za specifična opravila. Zato je jasno, da je pomembnost razvoja tehnologije senzorjev za mednarodno ekonomijo mnogo večja, kot pa je velikost samega področja izdelave in prodaje senzorjev.

Natančna ocena trga senzorjev je praktično nemogoča celo, če bi se omejili le na procesno avtomatizacijo (Pri tem ne upoštevamo tako pomembnih področij in trgov, kot je npr. uporaba senzorjev v različnih vozilih in v gospodinjskih aparatih.). Z dovolj veliko verjetnostjo je možno ugotoviti, da:

- si Evropa, ZDA in Japonska približno enako delijo svetovni trg
- senzorji (v širšem smislu) prispevajo približno eno tretjino celotne investicije v procesno vodenje (ostalo - regulacijski sistem, komunikacije med deli opreme in človekom, dodatki v izvršni sistem itd.)
- so najpomembnejši senzorji za pretok in tlak, nato pa pridejo na vrsto meritve temperature, nivoja in ostalih lastnosti.

Proizvajalci merilne opreme so tudi zelo različni. Od internacionalnih družb z ogromno ponudbo do vedno večjega števila majhnih proizvajalcev, ki se trudijo najti in pokriti posamezne tržne niše ali pa specialne tehnologije.

Glede na razvoj mikroelektronike tudi senzorji težijo k čim večji miniaturizaciji in kompaktnosti elementov. S tem pa seveda pride tudi do boljših lastnosti, boljše kompatibilnosti z mikrovezji in do nižje cene merilnih sistemov, v katere je vgrajeno vedno več znanja in inteligence. Tako se pojavljajo izrazi, kot so: mikromehatronika, tehnologija mikrosistemov, mikroelektromehanski sistemi, mikroinženirstvo itd., ki kažejo, da gre za napredno vejo tehnologije, ki pripomore tako k tehnološkemu napredku kakor tudi v smislu ekonomskih, ekoloških in drugih vidikov.

Literatura

/1/ Kissel, T.E. (2000): Industrial electronics, Prentice Hall, Upper Saddle River, New Jersey, ZDA

/2/ Strmcnik, S. in soavtorji (1998): Celostni pristop k racunalniškem vodenju procesov, Založba FE in FRI, Fakulteta za elektrotehniko, Ljubljana

/3/ Karba, R. (1994): Gradniki sistemov vodenja, Založba FER, Fakulteta za elektrotehniko in racunalništvo, Ljubljana

/4/ Solomon, S. (1999): Sensors handbook, McGraw-Hill, New York, ZDA

/5/ Fraden, J. (1997): Handbook of modern sensors, AIP Press, New York, ZDA

Rihard Karba, Maja Atanasijević-Kunc, Aleš Belič
Fakulteta za elektrotehniko, Tržaška 25,
1000 Ljubljana, Slovenija
e-mail: rihard.karba@fe.unilj.si

Juš Kocijan, Janko Petrovčič
Institut Jožef Stefan, Jamova 39,
1000 Ljubljana, Slovenija

Prispelo (Arrived): 06.06.2002 *Sprejeto (Accepted): 25.03.2003*

DEBELOPLASTNA TEHNOLOGIJA ZA SENZORSKE APLIKACIJE

[1]Darko Belavič, [2]Marko Hrovat, [1]Marko Pavlin, [1]Marina Santo Zarnik

[1]HIPOT-RR, d.o.o., Šentjernej, Slovenija
[2]Institut "Jožef Stefan", Ljubljana, Slovenija

Ključne besede: debeloplastna tehnologija, senzor, senzor sile, senzor tlaka

Izvleček: Kljub pospešenem razvoju mikrosistemske tehnike, ki je integracija številnih področij, kot so: senzorika, aktorika, mikroperiferika, mikromehanika, integrirana optika itd., so prevladujoče tehnologije za izdelavo senzorjev in pripadajoče elektronike še vedno monolitne polprevodniške tehnologije ter tankoplastna in debeloplastna tehnologija. Debeloplastna tehnologija lahko pri izdelavi senzorjev nastopa v dveh vlogah. Prva je izdelava debeloplastnega senzorskega elementa. Druga pa je integriranje senzorskega elementa, ki navadno ni debeloplasten, s kompenzacijskimi elementi in elektroniko za obdelavo električnega signala. V prispevku so prikazane predvsem nekatere aplikacije na področju senzorjev mehanskih veličin (senzorji tlaka in sile), kjer je uporabljena debeloplastna tehnologija v eni ali obeh naštetih vlogah.

Thick-film Technology for Sensor Applications

Key words: thick-film technology, sensor, force sensor, pressure sensor

Abstract: The most important technologies for manufacturing sensors and transducers are semiconductor technology, thin- and thick-film technologies. Thick-film technology is used in two ways, to produce the sensor elements themselves and/or the electronic circuits for signal processing. In both cases are some common particular characteristics and ceramic interconnection technology is one of them. This paper is focused in mechanical sensors (pressure, load, force, etc) where the sensor elements are made with thick-film technology. The first application is thick-film load sensor for kitchen scale. The second application is thick-film sensing element for a low-cost force sensor designed for industrial applications. The third application is thick-film force sensor for belt dynamometer (toco sensor) for medical application in maternity hospital.

1 Uvod

Leta 1990 je bil svetovni trg senzorjev vreden približno 15 milijard EUR, njegova vrednost je narasla do leta 2000 na 30 milijard EUR ter bo leta 2010 predvidoma dosegla 51 milijard EUR z letno rastjo od 5 do 9 %. Pri tem je treba upoštevati, da bo verjetno količinska rast še večja, ker je indeks rasti cen negativen. Regionalno je bil leta 1990 največji trg senzorjev v Evropi (42%), sledila ji je NAFTA (North American free trade Association) s 34 % in Azija s 24 %. Deset let kasneje pa je NAFTA prehitela Evropo in je sedaj na prvem mestu. Glavna področja uporabe senzorjev so industrijski in procesni senzorji, senzorji v avtomobilih ter pri konstrukciji strojev in naprav. Največji indeksi rasti so napovedani za uporabo senzorjev v avtomobilih, na področju informatike in telekomunikacij ter za okolje. Senzorji za okolje naj bi se povzpeli celo na tretje mesto po velikosti trga. Glede na tip senzorjev največji tržni delež pripada senzorjem tlaka, sledijo jim senzorji pretoka, temperature itd.

Prevladujoče tehnologije za izdelavo senzorjev in/ali pripadajoče elektronike so še vedno monolitne polprevodniške tehnologije ter tankoplastna in debeloplastna tehnologija. Monolitne polprevodniške tehnologije imajo izrazito prednost pri miniaturizaciji, integraciji, velikoserijski proizvodnji in cenenosti. Debeloplastna tehnologija nastopa na področju senzorjev in pretvornikov v dveh vlogah. Prva je električna in mehanska integracija senzorskega elementa (ki navadno ni debeloplasten), elektronike za pretvorbo signala in drugih elementov. Druga pa je izdelava samih senzorskih elementov, kot so: senzorji temperature, mehanskih in kemičnih veličin, plinski senzorji, biosenzorji kakor tudi kombinacije le-teh.

V nadaljevanju bo prikazano nekaj rezultatov raziskovalnega dela in nekaj senzorjev mehanskih veličin (senzorji tlaka in sile), kjer je uporabljena debeloplastna tehnologija v eni ali obeh naštetih vlogah.

2 Debeloplastni "strain gauge"

Senzor mehanskih deformacij ("strain gauge") je element, ki pretvarja deformacijo (razteg, skrček, upogib, ...) v električni signal in deluje po principu piezoupornostnega efekta električnih prevodnikov oz. uporov. Piezoupornost (Kelvin, 1857) je lastnost nekaterih materialov, da pri mehanski deformaciji spremenijo električno upornost. Občutljivost materiala za deformacijo pa je podana s faktorjem *GF* (1), ki je razmerje med spremembo upornosti (*dR/R*) in deformacijo (*dl/l*).

$$GF = \frac{\Delta R/R}{\varepsilon} \quad in \quad \varepsilon = \Delta l/l \qquad (1)$$

Debeloplastni upori so narejeni z metodo sitotiska, sušenja (150 °C) in žganja (850 °C) debeloplastnih materialov

na keramični (Al_2O_3) podlagi. Komercialno dostopni debeloplastni uporovni materiali (paste) imajo plastno upornost (R_{PL}) v dekadah od 1 do 10 MΩ. Pri konstantni debelini (t) pa je vrednost upornosti (R) odvisna še od dolžine (l) in širine (w) debeloplastnega upora. Relacije so prikazana v formulah (2):

$$R = \rho \frac{l}{tw} \qquad R_{PL} = \frac{\rho}{t} \qquad R = R_{PL} \frac{l}{w} \tag{2}$$

Piezoupornostni efekt, ki je sicer neželena lastnost pri navadni uporabi, izkazujejo tudi debeloplastni upori. Iz literature /1,2,3,4/ in naših prejšnjih raziskav /5,6,7,8,9,10,11/ izhaja, da je GF odvisen od mikrostrukture materiala, plastne upornosti in delno od geometrije debeloplastnega upora. Glede na orientacijo upora pa poznamo vzdolžni in prečni GF. Pri mehanski deformaciji se upornost debeloplastnega upora spremeni zaradi spremembe plastne upornosti in geometrije. Tako je GF sestavljen iz geometrijskega del (GF=2) in dela zaradi spremembe mikrostrukture materiala (3).

$$\frac{\Delta R}{R} = \left(\frac{\Delta \rho}{\rho}\right) + \left(\frac{\Delta l}{l} - \frac{\Delta w}{w} - \frac{\Delta t}{t}\right) \tag{3}$$

Za uporabo senzorskega elementa so poleg GF pomembne še nekatere druge lastnosti, kot so: temperaturni koeficient upornosti (TKR), temperaturni koeficient GF (TKGF), tokovni šum in dolgoročna stabilnost. Tako je v tabeli 1 prikazana primerjava različnih tehnologij za izdelavo elementov za senzorje mehanskih deformacij.

Tabela 1: Primerjava tehnologij za izdelavo elementov za senzorje mehanskih deformacij

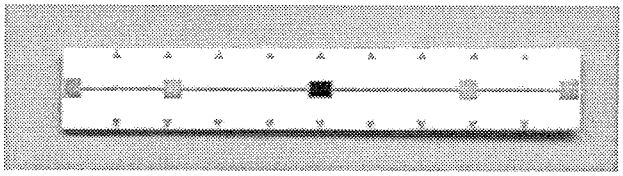
Tehnologija	Metalo-plastna	Debelo-plastna	Polpre-vodniška
GF	2	od 2 do 20	50
TKR (10^{-6}/K)	10	50 (100)	1500
TKGF (10^{-6}/K)	100	300	2000
Dolgoročna stabilnost	Odlična	Zelo dobra	Dobra
Cena	Visoka	Nizka	Nizka*

* Proizvodnja velikega obsega

3 Eksperimentalno delo

Za študij smo izbrali devetnajst debeloplastnih uporovnih materialov različnih proizvajalcev oz. serij (označeni s črkami A, B, ...) in z različnimi plastnimi upornostmi (označeni s številkami 1, 2, ...). Nekateri materiali, označeni z zvezdico (*), so bili namenjeni za tisk na dielektrično podlago.

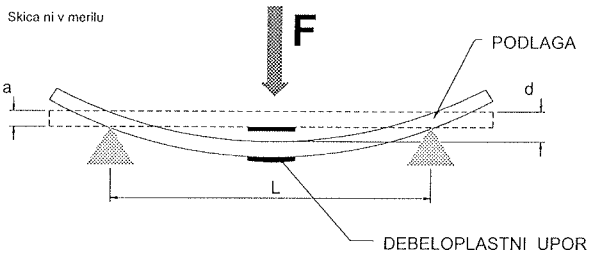
Debeloplastne upore smo tiskali in žgali (850 °C) na keramični podlagi dimenzij 50,8 × 8,0 × 0,64 mm. Dimenzije štirih vzdolžno orientiranih uporov so bile 1×1, 1,6×1,6, 5×1 in 4×4 mm. En upor velikosti 1,6×1,6 mm pa je orientiran tudi prečno. Preskusni debeloplastni upori so bili postavljeni v sredino substrata in so bili s prevodnimi linijami povezani s kontaktnimi blazinicami na robu (slika 1).



Slika 1: Preskusni vzorec za meritev piezoupornostnih lastnosti debeloplastnih uporov

Merjenje piezoupornostnih lastnosti debeloplastnih uporov smo izvedli na način, ki je shematično prikazan na sliki 2. Keramični most smo na sredini upogibali do določenih upogibkov (d) in hkrati merili upornosti. Iz upogibkov smo izračunali deformacijo oz. raztezek debeloplastnega upora (4). Iz tega in spremembe upornosti smo izračunali GF. Rezultati so prikazani v tabeli 2.

$$\varepsilon = \Delta l / l = \frac{6\,d\,a}{L^2}\left(1 - \frac{l}{2\,L}\right) \tag{4}$$



Slika 2: Princip merjenja piezoupornostnih lastnosti debeloplastnih uporov

Poleg občutljivosti je za uporabo teh senzorskih elementov pomembno razmerje signal/šum. Zato smo pri istih preskusnih vzorcih merili tudi tokovni šum. Meritev šuma je bila izvedena po metodi Quan-Tech. Rezultati so podani v μV/V za šum in v dB za indeks šuma.

4 Rezultati in diskusija

Nekateri izmerjeni parametri (plastna upornost, vzdolžni in prečni GF ter indeks šuma) preučevanih uporovnih materialov za upore dimenzij 1,6×1,6 mm so prikazani v tabeli 2. Rezultati kažejo, da so GF večji za višje plastne upornosti. Ravno tako je šum večji pri uporih z višjo upornostjo. Vzdolžni GF so 2- do 4-krat večji od prečnih GF. Zato je potrebna optimalna izbira med dovolj veliko občutljivostjo na eni strani in še dopustnim nivojem tokovnega šuma na

drugi. Pri nekaterih aplikacijah pa je pomembna tudi ohmska upornost in temperaturna odvisnost elementa.

Tabela 2: Plastna upornost, GF in indeks šuma debeloplastnih uporov, izdelanih z različnimi debeloplastnimi uporovnimi materiali

Oznaka	R_{PL} (W)	Vzdolžni GF	Prečni GF	Indeks šuma (dB)
A3	1k	10,2	7,8	-20,9
A4	10k	12,4	10,6	-12,4
A5	100k	14,3	12,9	-1,8
A6	1M	16,3	15,9	>30,0
B3	1k	2,6	2,0	-12,8
B4	10k	3,9	3,4	-5,2
C3	1k	7,5	5,6	-24,5
C4	10k	11,7	9,2	-16,4
C5	100k	13,1	10,6	-5,3
C6	1M	14,7	13,3	5,0
D4	10k	18,0	14,2	-7,2
E4	10k	20,1	12,9	2,0
F3*	1k	4,5	3,5	-18,6
F4*	10k	11,4	9,0	-15,4
F5*	100k	13,7	12,6	-2,5
F6*	1M	14,9	13,5	9,1
G3*	1k	3,8	2,9	-21,4
G4*	10k	10,0	8,0	-16,6
G5*	100k	13,5	12,1	-3,9

5 Aplikacije

Debeloplastni element za senzorje mehanskih deformacij je bil prvič uporabljen že v začetku sedemdesetih let. Najpogosteje se ga uporablja pri senzorjih tlaka in sile. Lahko pa tudi za senzorje pospeškov, vibracij itd. Uporaben je predvsem za aplikacije v zahtevnejših razmerah v okolici. Ravno tako pa je konkurenčen pri maloserijskih naročilih in pri posebnih zahtevah naročnika.

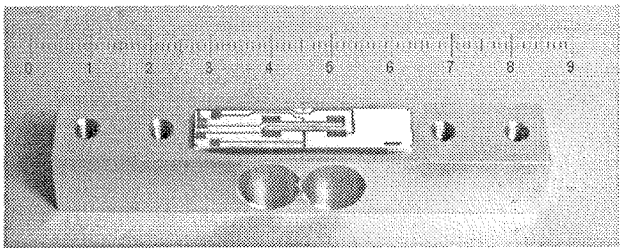
Družba HIPOT-HYB se je že večkrat srečala s tržno zahtevo po debeloplastnih senzorjih sile. Tako je bilo po naročilu razvitih nekaj cenovno konkurenčnih senzorjev sile /12,13,14,15/, ki bodo kratko predstavljeni v nadaljevanju. Vsi omenjeni senzorji uporabljajo debeloplastni senzorski element, izdelan na keramični podlagi s štirimi debeloplastnimi upori, vezanimi v Wheatstonov mostič.

Pri konstruiranju senzorskih elementov pa se s pridom uporablja tudi orodja za numerično modeliranje in računalniško simulacijo mehanskih in elektromehanskih lastnosti /16,17/. Primer simulacije kovinskega nosilca z dvojnim upogibom je prikazan na sliki 3.

5.1 Senzor sile za kuhinjsko tehtnico

Debeloplastni senzorski element je prilepljen na kovinski nosilec z dvojnim upogibom. Senzor je bil razvit za elektronsko tehtnico z merilnim območjem od 10 do 3000 g in s točnostjo ±2% polnega obsega. Senzor je prikazan na sliki 4.

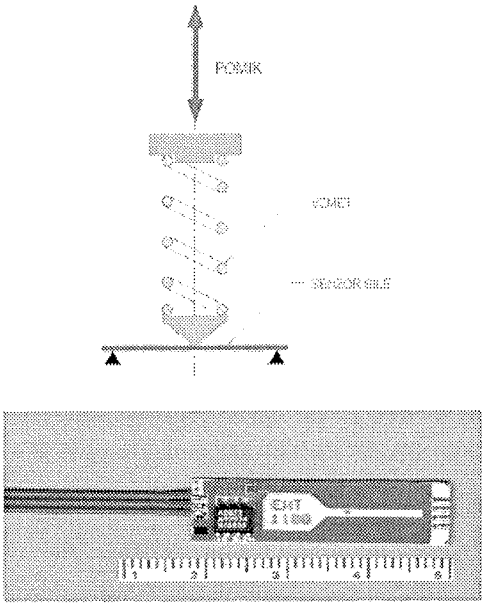
Slika 3: Računalniška simulacija kovinskega nosilca z dvojnim upogibom za senzor sile pri kuhinjski tehtnici



Slika 4: Senzor sile za kuhinjsko tehtnico

5.2 Merilnik pomika

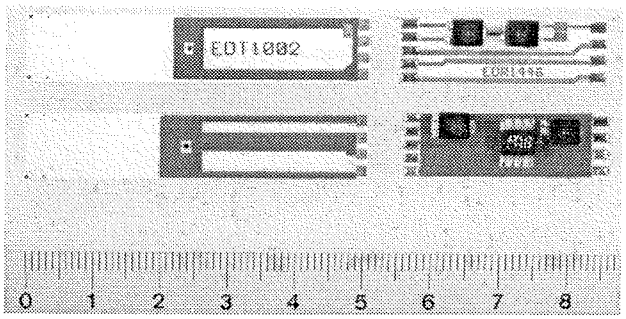
Merilnik pomika ima vgrajen senzor sile za območje do 30 N. Princip delovanja in sam izdelek sta prikazana na sliki 5.



Slika 5: Princip delovanja merilnika pomika (zgoraj) in debeloplasten senzor sile (spodaj)

5.3 Senzor porodnih krčev

Senzor porodnih krčev ima debeloplastni senzorski element, ki je vpet v plastično ohišje in meri sunke porodnih krčev do sile 1,2 N. Prikazan je na sliki 6.



Slika 6: Senzor sile za merjenje sunkov

6 Sklep

Debeloplastna tehnologija je primerna za izdelavo senzorjev zaradi svoje robustnosti (uporaba pri zahtevnejših pogojih okolice), razvojne fleksibilnosti in cenenosti pri manjših serijah. Senzorski elementi se lahko izdelujejo ali z uporabo komercialnih debeloplastnih materialov ali posebej razvitih senzorskih materialov.

Raznovrstnost problematike na področju senzorjev in debeloplastne tehnologije (tudi hibridne tehnologije) pa zahteva interdisciplinarni značaj raziskovalno-razvojne dejavnosti. Pomembna področja so: senzorika, znanost o materialih, keramične tehnologije, različna področja iz fizike in kemije, elektronika, načrtovanje elektronskih vezij in znanja s področja zagotavljanja kvalitete in produktivnosti. Pri tem so potrebne različne kategorije raziskovalnega dela, in sicer od osnovnih raziskav preko aplikativnih do razvojnega in eksperimentalnega dela.

Zahvale

Zahvaljujemo se industrijskemu partnerju, družbi HIPOT-HYB, d.o.o., Šentjernej, ki je dovolila objavo prispevka.

Zahvaljujemo se Ministrstvu za šolstvo, znanost in šport Republike Slovenije za sofinanciranje aplikativnega raziskovalnega projekta L2-3130 z naslovom Debeloplastna tehnologija za senzorske aplikacije.

Avtorji prispevka se zahvaljujemo tudi vsem drugim sodelavcem pri različnih raziskovalnih in razvojnih projektih, ki so bili osnova za nekatere rezultate, navedene v prispevku.

7 Literatura

/1/ B. Puers, W. Sansen, S. Paszczynski, K. U. Leuven, "Miniature highly sensitive pressure - force sensor using hybrid technology", Proc. 6th European Microelectronics Conference, Bournemouth, Anglija, 3.-5.6.1987, 416-420

/2/ S. Chitale, C. Huang, M. Stein, "High gauge factor thick film resistors for strain gauges", Hybrid Circuits Technol., 6(1989)5

/3/ C. Song, D. V. Kerns, Jr., J. L. Davidson, W. Kang, S. Kerns, "Evaluation and design optimization of piezoresistive gauge factor of thick film resistors", IEEE Proc. SoutheastCon 91 Conf., Williamsburg, 2(1991), 1106-1109

/4/ N. M. White, J. D. Turner, "Thick film sensors: past, present and future", Measurements. Sci. Technol., 8(1997)1, 1-20

/5/ M. Hrovat, D. Belavič, S. Šoba, A. Markošek, "Thick-film resistor materials for strain gauges", Proc. 20th Int. Conf. on Microelectronics / 28th Symp. on Devices and Materials MIEL-SD 92, Portorož, 1992, 343-348

/6/ M. Hrovat, G. Dražič, J. Holc, D. Belavič, "Microstructural investigation of thick-film resistors for strain sensor applications by TEM", Proc. 22nd Int. Conf. Microelectronics MIEL-94 / 30th Symp. on Devices and Materials SD-94, Terme Zreče-Rogla, 1994, 207-212

/7/ M. Hrovat, D. Belavič, J. Holc, S. Soba, "An evaluation of some commercial thick film resistors for strain gauges", J. Mater. Sci. Lett., 13(1994), 992-995

/8/ M. Hrovat, G. Dražič, J. Holc, D. Belavič, "Correlation between microstructure and gauge factors of thick film resistors", J. Mater. Sci. Lett., 14(1995)15, 1048-1051.

/9/ M. Hrovat, D. Belavič, G. Dražič, J. Holc, S. Soba, "Preiskave debeloplastnih uporov z visokimi faktorji gauge", Informacije MIDEM, 25(1995)2, 108-114

/10/ M. Hrovat, D. Belavič, M. Jerlah, "Investigation of Some Thick-film Resistor Series for Strain Gauges", Proc. 23rd Int. Spring Seminar on Electronics Technology ISSE 2000, Balatonfüred, 2000, 406-410

/11/ M. Hrovat, Z. Samardžija, J. Holc, D. Belavič, "Microstructural, XRD and electrical characterisantion of some thick film resistors", J. Mater. Sci.; Materials in Electronics, 11(2000)3, 199-208

/12/ D. Belavič, S. Šoba, M. Hodnik, M. Pavlin, S. Gramc, M.Hrovat, "Low cost force sensor for an electronic scale", Informacije MIDEM, 27(1997)3, 172-176

/13/ D. Belavič, S. Šoba, M. Pavlin, S. Gramc, M. Hrovat, "Electronic scale with thick-film resistors strain gauge", Proc. 21st Conference of ISHM Poland, Poljska, Ustron, 5. - 8.10.1997, 9-14

/14/ S. Soba, D. Belavič, M. Pavlin. "Thick-film load sensors array". Proc. 23rd IMAPS Poland'99, Poljska Koszalin Kołobrzeg, 2. - 23.9.1999, 303-308

/15/ D. Belavič, M. Hrovat, M. Pavlin, S. Gramc, "Low-cost thick-film strain-gauge applications". Proc. 13th European Microelectronics and Packaging Conference, Strasbourg, Francija, 30.5. - 1.6.2001, 103-108

/16/ D. Belavič, K. P. Friedel, M. Santo Zarnik, M. Hrovat, "A "design and simulate" approach to improving the characteristics of ceramic pressure sensors". Proc. 37th International Conference on Microelectronics, Devices and Materials and the Workshop on Optoelectronic Devices and Applications, Bohinj, Slovenija, 10. - 12.10.2001, 267-272

/17/ D. Belavič, K. P. Friedel, A. Wymyslowski, M. Santo Zarnik, "Virtual prototyping of the ceramic pressure sensor". Proceedings of the 3rd International Conference on Benefiting from Thermal and Mechanical Simulation in (Micro)- Electronics, ESIME 2002, Pariz, Francija, 15. - 17.4.2002, 38-44

Darko Belavič, univ. dipl. inž. el.
HIPOT-RR, d.o.o.
c/o Institut "Jožef Stefan"
Jamova 39, 1000 Ljubljana, Slovenija

Tel.: +386 1 4773 479, Faks: +386 1 4263 126
E-pošta: darko.Belavič@ijs.si

Prispelo (Arrived): 06.06.2002 Sprejeto (Accepted): 25.03.2003

SENZORJI TLAKA ZA MEDICINSKO IN INDUSTRIJSKO UPORABO

[1]Marko Pavlin, [1]Darko Belavič, [2]Matej Možek

[1]HIPOT-RR, d.o.o., Šentjernej, Slovenija
[2]Fakulteta za elektrotehniko, Ljubljana, Slovenija

Ključne besede: senzor, senzor tlaka, piezouporovni senzorji, pametni senzorji, debeloplastna tehnologija

Izvleček: Družba HIPOT-HYB že vrsto let proizvaja, poleg hibridnih debeloplastnih vezij, tudi senzorje tlaka za medicinsko in industrijsko uporabo. Razvoj tehnologije in zahteve trga usmerjajo tudi razvojno dejavnost. V prispevku bo predstavljena raziskovalno-razvojna dejavnost na področju senzorjev tlaka. Njeni dosežki na tem področju bodo ilustrirani z nekaterimi značilnimi primeri, njene usmeritve pa bodo prikazana na primeru pametnega senzorja tlaka.

Pressure Sensors for Medical and Industrial Applications

Key words: sensor pressure sensor, piezoresistive pressure sensor, smart sensor, thick-film technology

Abstract: Through many years of experience in thick film technology and pressure sensors many medical and industrial pressure sensor applications were developed and produced in the HIPOT-HYB company. For illustration few medical and industrial applications of pressure sensors are described. In all cases the sensor elements were gauge silicon piezoresistive pressure sensors with a periphery made in thick film technology. In this contribution we demonstrated that silicon pressure sensors in chip form in combination with thick film technology could be successfully used in pressure sensor applications. Special attention must be paid to assembling and packaging aspects in the production of pressure sensors. In our case, the silicon pressure sensor die is bonded onto a ceramic substrate with polymeric adhesives and encapsulated with polymeric materials. The mechanical and thermo-mechanical properties of the materials for assembly and housing have a crucial influence on the properties. Another important aspect in the production of pressure sensors is miniaturisation. The goal is achieved by the use of smaller or more complex components, which concentrate more functionality in a smaller volume. An example is shown with the pressure-switch example. The research, development and design activities on the field of the smart sensor are also introduced.

1 Uvod

Po podatkih Intechno Consulting, Basel, Švica je bila vrednost svetovnega trga senzorjev leta 2001 približno 31 milijard EUR z letno rastjo okoli 6 %. Pomembnejša področja uporabe senzorjev so procesna industrija, stroji in naprave, avtomobilska industrija, itd.. V zadnjih letih je avtomobilska industrija z 28-odstotnim deležem postala glavni uporabnik senzorjev in s tem prehitela procesno industrijo, ki je imela pred 10 leti največji, 30-odstotni delež. Delež senzorjev za uporabo v medicini se je tudi povečal od 6 % leta 1991 na 10 % v letu 2001. Glede na tip senzorjev pa največji tržni delež pripada senzorjem tlaka.

Tehnološke poti razvoja na področju senzorjev gredo v dveh smereh. Prva je miniaturizacija v t. i. Mikro-elektromehanske sisteme (MEMS), druga pa je v pametne (smart) senzorje. Mikrosistemske tehnologije za izdelavo MEMS so v največji meri polprevodniške. Uporabljajo pa se tudi keramične (C-MEMS) in hibridne (H-MEMS) tehnologije za izdelavo t.i. mezo- in mikro-elektromehanskih sistemov (M&MEMS). Kombinacija MEMS in modernih elektronskih in informacijskih tehnologij pa je ponudila novo kvaliteto pri uporabi senzorjev. Tako imenovani pametni senzorji so sposobni samodiagnostike in samokalibracije, njihovo delovanje pa se lahko krmili (vklop/izklop, merilno območje, merilna veličina,...). Taki senzorji tudi aktivno sodelujejo pri procesih, so hitro zamenljivi in se jih lahko priredi za komunikacijo v različnih mrežah (CAN, internet, ...).

Splošna usmerjenost na področju senzorjev je tudi zniževanje cen. Najbolj se to pozna pri proizvajalcih senzorskih elementov. Zaradi tega le-ti povečujejo obseg proizvodnje, inovirajo tehnologije in standardizirajo izdelke. V veliki meri cenovna usmerjenost vpliva tudi na naslednjo stopnjo, ki jo imenujemo senzorski modul. tega v minimalnem obsegu sestavljajo: senzorski element, elektronika za procesiranje senzorskega signala in ohišje. Proizvajalci imajo zato večjo fleksibilnost pri izbiri sestavnih delov, razvoju, konstrukciji in sestavi senzorskih modulov. Cenovna usmerjenost na nivoju sistemov pa je zelo odvisna od področja uporabe.

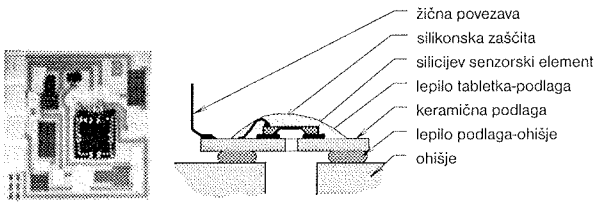
Industrijski partner, HIPOT-HYB, Proizvodnja hibridnih vezij, d.o.o., Šentjernej, je podjetje, ki proizvaja hibridna debeloplastna vezja ter medicinske in industrijske senzorje tlaka. V manjši meri izdeluje tudi senzorje temperature in sile. HIPOT-HYB trži pretežno na razvitih zahodnih trgih, kjer ustvari 75 % prihodkov od prodaje. V programski strukturi je najpomembnejši program medicinskih senzorjev tlaka s 60-odstotnim deležem, sledijo hibridna debeloplastna vezja s 37, 3-odstotni delež pa je bil leta 2001 nov program industrijskih senzorjev tlaka. Podjetje razvija in izdeluje široko paleto različnih senzorjev tlaka: enostavne pasivne senzorje, temperaturno kompenzirane in kalibrirane pasivne sen-

zorje, z dodatkom aktivnih elektronskih komponent pa tudi industrijske pretvornike tlaka in pametne senzorje.

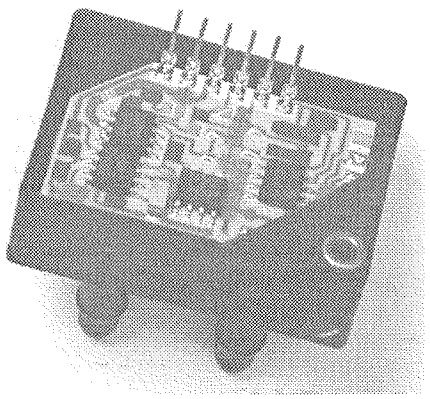
V nadaljevanju bo prikazano nekaj rezultatov raziskovalno-razvojnega dela in nekaj senzorjev tlaka za uporabo v medicini in industriji.

2 Konstrukcija senzorja tlaka

Za izdelavo hibridnih senzorjev tlaka uporabljamo kupljen gol silicijev piezoupornostni senzor tlaka kot senzorski element ter hibridno debeloplastno tehnologijo za montažo senzorskega elementa ter izdelavo kompenzacijskega in umerjevalnega elektronskega vezja (slika 1) /1/. Poleg tega se hibridna debeloplastna tehnologija uporablja tudi za izdelavo elektronskega vezja za procesiranje senzorskega signala (slika 2) /2,3,4,5,6/.



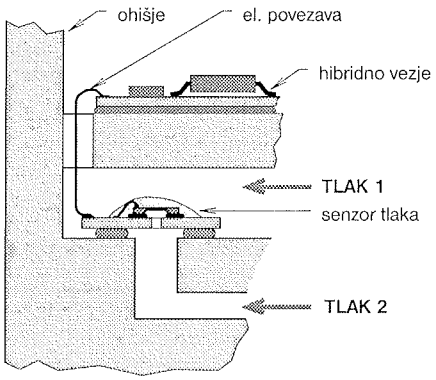
Slika 1: Silicijev piezoupornostni senzor tlaka, integriran v debeloplastno umerjevalno vezje. Na levi je prikazana fotografija izdelka, na desni pa shematsko njegov presek.



Slika 2: Industrijski senzor tlaka. Vidno je elektronsko vezje za procesiranje senzorskega signala, izdelano v hibridni debeloplastni tehnologiji.

Na lastnosti senzorjev tlaka pa pomembno vpliva tudi konstrukcija in izbira materiala za sestavne dele /7,8,9/. Značilna konstrukcija senzorjev (modulov) tlaka je izvedena v dveh delih. Prvi je senzorski element z ustrezno periferijo, drugi pa elektronsko vezje za procesiranje senzorskega signala. Oba dela pa sta potem vgrajena v ohišje, ki je lahko plastično ali kovinsko ter s priključnimi cevkami ali z možnostjo vgradnje v standardna kovinska cevna ohišja. Detajl ene izmed značilnih konstrukcij senzorjev tlaka je prikazan na sliki 3.

Posebno pozornost posvečamo izbiri materialov za ohišja in sestavne dele ter nadzoru proizvodnega procesa. Pri tem je pomembna predvsem skladnost mehanskih, termomehanskih in kemičnih lastnosti sestavnih delov /10,11,12/. V tabeli 1 so podani temperaturni razteznostni koeficienti uporabljenih materialov. Procesi izdelave senzorja pa morajo biti taki, da ne puščajo zaostalih mehanskih napetosti v sestavnih delih.



Slika 3: Detajl ene izmed značilnih konstrukcij senzorjev tlaka

Tabela 1: Primerjava temperaturnih razteznostnih koeficientov materialov, uporabljenih pri konstrukciji senzorja tlaka

Material	**Temperaturni razteznostni koeficient (10^{-6}/K)**
Epoxy	60÷80
Polyester	80÷130
Mehko lepilo (gel)	200÷400
Keramična podlaga	6÷7
Silicijev senzorski element	3÷4

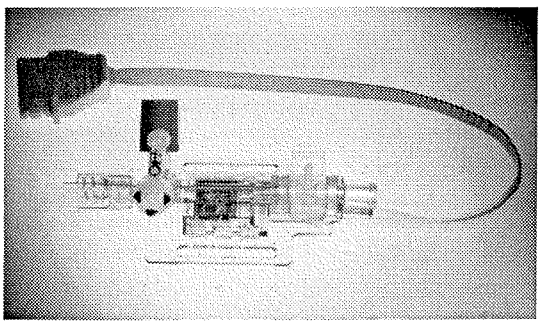
Pri izdelavi senzorjev tlaka praviloma uporabljamo senzorske elemente (silicijeve ali keramične), ki delujejo na osnovi piezoupornostnega efekta. Štirje senzorski upori na membrani so vezani v Wheatstonov mostič, ki je napajan s konstantnim tokom ali napetostjo, izhodna napetost pa je v področju mV. To je osnova, ki jo nadgradimo najprej s temperaturno kompenzacijo, da izničimo vpliv neželenih veličin (temperature). Temu sledi ojačevalnik z diferencialnim vhodom, ki je individualno kalibriran za vsak senzorski element posebej. Izhodna napetost je v standardnem področju od 0,5V do 4,5V. Namesto napetostnega izhoda lahko izdelamo tudi tokovni izhod za tokovne zanke od 4mA do 20mA ali pa digitalni vmesnik. Večjo stopnjo integracije pa dosežemo z uporabo ASIC (Application Specific Integrated Circuit) /13,14/.

Področja tlakov, ki jih obvladujemo z našimi senzorji so od 2 mbar do 10 bar. Temperaturno področje delovanja je od -20 °C do +80 °C. V temperaturnem področju od 0 do 50

°C pa zagotavljamo točnost 1 %. Napajalne napetosti so od 5 V naprej.

3 Senzorji tlaka za uporabo v medicini

V medicini obstajajo vse večje zahteve po točnem spremljanju določenih vitalnih človekovih funkcij pri velikem številu pacientov. Ta smer je še posebej opazna v zelo razvitih državah, kot so: ZDA, Nemčija, Japonska in druge države EZ. Več desetletij je že v uporabi različna nadzorna (monitorska) oprema, ki omogoča v povezavi z ustreznimi senzorji trajno spremljanje človekovih vitalnih funkcij. Tak primer je tudi invazivno merjenje krvnega tlaka pri pacientih v intenzivni negi in merjenje znotrajmaternične tlaka pri nosečnicah. Oboje lahko merimo v principu z istim tipom senzorja, ki pa mora biti prirejen specifični uporabi. Posebnost senzorjev tlaka za medicinsko uporabo je proizvodnja v klinično čistih prostorih in uporaba biokompatibilnih materialov in drugih specifikacijah, ki jih določa standard AAMI. Senzorji, ki so se uporabljali prej, so bili namenjeni za večkratno uporabo. Vzporedno z naraščajočim strahom za krvno okužbo so se v zadnjem desetletju začeli uporabljati senzorji za enkratno uporabo. Tak tip senzorjev je ustvaril novo tržišče, ki dosega samo na področju krvnih senzorjev tlaka že okrog 30 milijonov kosov. Od tega podjetje HIPOT-HYB izdeluje 1,2 milijona kosov senzorjev za invazivno merjenje krvnega tlaka (slika 4) pri pacientih v intenzivni negi.

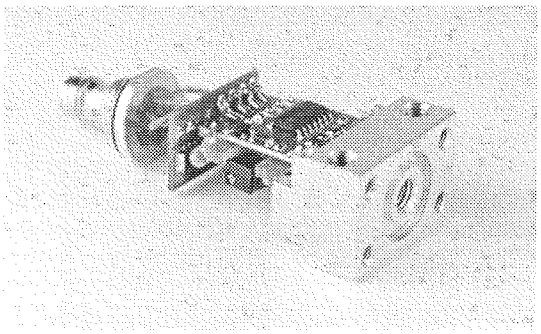


Slika 4: Senzor za invazivno merjenje krvnega tlaka pri pacientih v intenzivni negi

4 Razvojna pot vakuumskega stikala

Vakuumsko ali tlačno stikalo je posebna izvedba senzorja tlaka. Referenčno napetost in napetost ojačenega senzorskega signala, ki pomeni izmerjeni tlak, primerjamo v elektronskem vezju (komparatorju), ki izvede električni preklop. S spreminjanjem referenčne napetost je možno tudi nastavljanje, pri kateri velikosti tlaka bo stikalo preklopilo.

Za znanega kupca smo pred leti po naročilu razvili vakuumsko stikalo (slika 5), ki je bilo takrat najmanjše komercialno dosegljivo vakuumsko stikalo na trgu. Z nadaljnjim razvojno-raziskovalnim delom pa se je evolucija tega izdelka nadaljevala v smeri nadaljnje miniaturizacije (slika 6) kakor tudi v smeri enostavnejših proizvodnih postopkov.

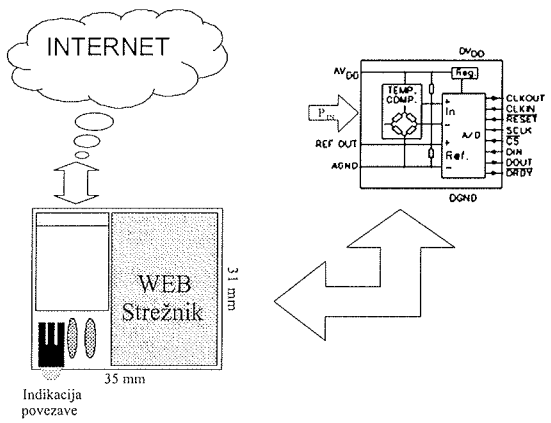


Slika 5: Konvencionalno vakuumsko stikalo

Slika 6: Miniaturno vakuumsko stikalo (levo) in ultraminiaturno vakuumsko stikalo (desno)

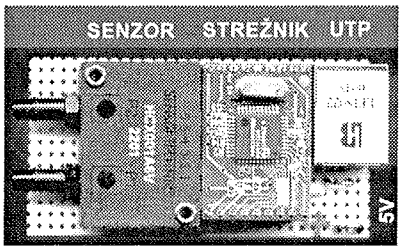
5 Pametni senzor tlaka

Nadgradnja senzorjev tlaka so t. i. pametni senzorji tlaka /15/. Napetost ojačenega senzorskega signala se pretvori v digitalno informacijo, ki se v nadaljevanju procesira z uporabo digitalnih tehnologij, ki omogoča široke možnosti različnih tipov in nivojev "pameti". Taki senzorji imajo poleg digitalnega vmesnika tudi možnost samokalibracije, prilagajanja na proces, konfiguriranja v živo in celo direktne priključitve na omrežje in v internet. Poleg novih lastnosti pa pametni senzorji omogočajo visoko stopnjo avtomatizacije proizvodnje in poenostavljeno izdelavo, ker ni potrebno analogno umerjanje.



Slika 7: Shematski prikaz delovanja tlačnega pretvornika z integriranim strežnikom

Pametni senzorji so sedaj na nivoju prototipov, ki so nastajali v sodelovanju s HIPOT-RR in Fakulteto za elektrotehniko. Do sedaj smo razvili tri tipe pametnih senzorjev: osnovni model s procesorjem MAX1458, kompleksnejši model s procesorjem ADUC816 in tlačni pretvornik z integriranim strežnikom (sliki 7 in 8).



Slika 8: Prototip senzorja tlaka z integriranim strežnikom

6 Sklep

Senzorji tlaka, izdelani v hibridni tehnologiji za medicinsko in industrijsko uporabo, temeljijo na kompatibilnih tehnologijah in materialih, kjer je predvsem pomembna skladnost mehanskih in termo-mehanskih lastnosti silicijevega senzorskega elementa in keramične podlage. Razvoj pa gre v smeri uporabe različnih polprevodniških mikro elektromehanskih sistemov (MEMS), integriranih v različno periferijo (elektronika, aktorika, mikromehanika, mikrofluidika, ...). Za izdelavo omenjene periferije pa je primerna keramična LTCC (Low Temperature Cofiring Ceramics).

Poleg tega nadgradnja senzorjev tlaka v pametne senzorje odpira nove tržne možnosti in vpliva na način proizvodnje.

Zahvale

Zahvaljujemo se industrijskemu partnerju, družbi HIPOT-HYB, d.o.o., Šentjernej, ki je dovolila objavo prispevka.

Zahvaljujemo se Ministrstvu za šolstvo, znanost in šport Republike Slovenije za sofinanciranje aplikativnega raziskovalnega projekta L2-3025 z naslovom Tehnologije silicijevih senzorjev tlaka.

Avtorji prispevka se zahvaljujemo tudi vsem drugim sodelavcem pri različnih raziskovalnih in razvojnih projektih, ki so bili osnova za nekatere rezultate, prikazane v prispevku.

7 Literatura

/1/ D. Belavič, S. Šoba, M. Pavlin, D. Ročak, M. Hrovat, "Silicon pressure sensors with a thick film periphery", Microelectronics International, 15(1998)3, 26-30

/2/ D. Belavič, M. Pavlin, M. Hrovat, "Evaluation of Thick Film Materials for Diffusion Patterning - Preliminary Results", Proc. 34th Int. Conf. Microelectronics, Devices and Materials MIDEM-98, Rogaska Slatina, 1998, 305-310

/3/ D. Belavič, M. Hrovat, M. Pavlin, "Thick film materials for diffusion patterning technology", Proc. The Fifth European Conference on MultiChip Modules, London, Feb. 1-2,1999, 62-72

/4/ M. Hrovat, D. Belavič, M. Pavlin, "Some results obtained with thick film diffusion patterning technology", Proc. 35th International Conference on Microelectronics, Devices and Materials and Workshop on Microsystems, 13. - 15.10.1999, Ljubljana, 163-168

/5/ D.Belavič, M. Pavlin, M. Hrovat, "Cheap MCM-C for sensor applications". Proc. XXIII Conference of the International Microelectronics and Packaging Society, IMAPS POLAND'99, Poland Chapter, Koszalin Kołobrzeg, 21. - 23.9.1999, 155-160.

/6/ D. Belavič, M. Hrovat, M. Pavlin, "Thick-film Resistors and multilayer diffusion patterning technology", Proc. The 6th European Conference on MultiChip Modules, London,. 24. - 25.1.2000, 7-15

/7/ D. Belavič, S. Šoba, M. Pavlin, S. Maček, M. Hrovat, D. Ročak, "Design of thick film hybrid circuits for sensor applications", Proc. 24rd Int. Conf. Microelectronics MIEL-96 / 32st Symp. on Devices and Materials SD-96, Nova Gorica, 1996, 237-242

/8/ D. Belavič, S. Šoba, M. Pavlin, S. Gramc, D. Ročak, "Packaging technologies for thick film sensors", IMAPS/NATO Advanced Research Workshop on Electronic packaging for high reliability, low cost electronics, 10. - 13.5.1997, Bled

/9/ J. S. Bergstrom, W. H. Teat, "Package evaluation for silicon pressure sensors", Proc. Int. Symp. on Microelectronics ISHM'87, Minneapolis, 1987, 89-94

/10/ M. Pavlin, D. Ročak, S. Šoba, S. Amon, "Evaluation of polymer adhesives for use in silicone pressure sensor bonding on ceramic substrates", Proc. 24rd Int. Conf. Microelectronics MIEL-96 / 32st Symp. on Devices and Materials SD-96, Nova Gorica, 1996, 243-248

/11/ S. Šoba, D. Belavič, M. Pavlin, I. Lahne, "Long term stability of pressure transducers", Procedings of 22-nd Conference of the International Microelectrinics and Packaging Society, IMAPS, Poland chapter : Zakopane, 1. - 3.10.1998. Krakov: The International Microelectronics and Packaging Society, 1999, 307-310

/12/ D. Belavič, D. Ročak, J. Fajfar Plut, M. Hrovat, M. Pavlin, "An evaluation of stability of small size untrimmed and laser trimmed thick film resistors", Proc. 23rd Int. Conf. Microelectronics MIEL-95 / 31st Symp. on Devices and Materials SD-95, Terme Čatež,, 1995, 157-16

/13/ M. Pavlin, S. Šoba, D. Belavič, "Cheap ASICs vs. Discrete Electronics in Sensors Applications", Proc. 34th Int. Conf. Microelectronics, Devices and Materials MIDEM-98, Rogaska Slatina, 1998, 257-262

/14/ M. Pavlin, D. Belavič, S. Šoba, "ASICs for sensor applications", Proc. 22nd International Spring Seminar on Electronics Technology, ISSE'99, 18. - 20.5.1999, Dresden, Nemčija, 268-273.

/15/ M. Možek, S. Amon, D. Vrtačnik, D. Resnik, U. Aljančič, "Designing smart pressure sensors", Proc. 37th Int. Conf. Microelectronics, Devices and Materials MIDEM-2001, Bohinj, Slovenija, 10. - 12.10.2001, 179-186

Marko Pavlin, univ. dipl. inž. el.
HIPOT-RR, d.o.o.
Trubarjeva 7, 8310 Šentjernej, Slovenija
Tel.: +386 1 4773 479
Faks: +386 1 4263 126
E-pošta: marko.pavlin@guest.arnes.si

Prispelo (Arrived): 06.06.2002 Sprejeto (Accepted): 25.03.2003

INTEGRIRANI ULTRAZVOČNI PIEZOELEKTRIČNI PRETVORNIKI ZA UPORABO V MEDICINI

Janez Holc[1], Marija Kosec[1], Franck Levassort[2], Louis Pascal Tran-Huu-Hue[2] and Marc Lethiecq[2]

[1] "Inštitut" Jožef Stefan Institute, Ljubljana, Slovenija
[2]LUSSI/GIP Ultrasons, EIVL, François Rabelais University, Blois cedex, Francija

Ključne besede: pretvorniki, ultrazvok, debele plasti, piezoelektriki, integracija

Izvleček: Za izdelavo medicinskih visokofrekvenčnih ultrazvočnih pretvornikov je potrebna zelo tanka piezoelektrična keramika, debeline nekaj 10 μm. Navadno se izdeluje piezoelektrične elemente z rezanjem in tanjšanjem kosa keramike. Zaradi majhne debeline elementa lahko nastane pri lepljenju na podporni dušilec (backing) krušenje in lomljenje. Problem smo rešili tako, da smo na primeren nosilec, ki je imel tudi vlogo dušilca z debeloplastno tehnologijo nanesli piezoelektrično plast. Pretvornik smo izdelali na pozlačenem poroznem nosilcu $Pb(Zr,Ti)O_3$ (PZT). Debelina piezoelektrične PZT-plasti je okoli 40 μm. Zaradi velikega dušenje poroznega PZT- nosilca, njegove primerljive akustične impendance in plasti, ima izdelan modelni pretvornik ustrezno frekvenčno karakteristiko. Zaradi podobne kemijske sestave nosilca in plasti je minimalna tudi kemijska in fizikalna interakcija med plastjo in podlago, kot sta na primer difuzija zaradi podobne sestave in dobro ujemanje termičnih raztezkov.

Integrated Ultrasonic Piezoelectric Transducers for Medical Applications

Key words: Transducers, ultrasound, thick films, piezoelectric, integration

Abstract: Recent development trends in piezoelectric devices are towards smaller size, higher resonant frequencies and a low driven voltage. For high-frequency transducers in medical imaging applications, thin (i.e.<50 μm) piezoceramic elements are necessary. These are usually produced by lapping and machining, however the elements tend to chip and break, and this is a major problem with joining the element to the backing. This problem might be avoided by an integrated device with a thick piezoelectric layer on a suitable substrate that may also serve as a backing. To provide this function, $Pb(Zr,Ti)O_3$ (PZT) ceramics were chosen for the substrate, theoretical density of PZT is close to 8000 kg/m^3.

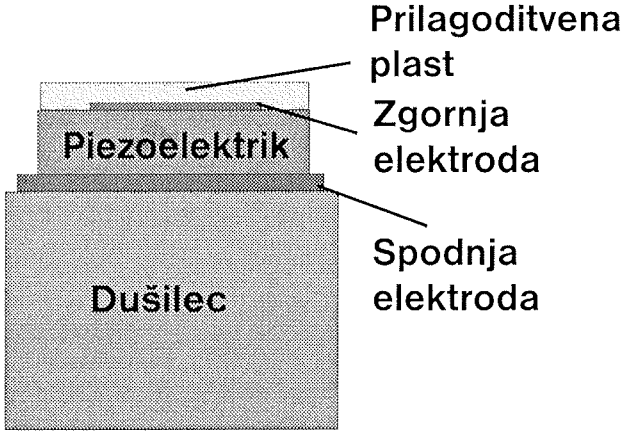
PZT thick films with a thickness around 40μm were succesfully processed on gold-coated alumina and gold-coated PZT substrates. To lower the firing temperature of the PZT layers to 800-900°C the composition was modified with lead germanate (PGO) that forms a low-temperature liquid phase. Films with 80-90% of theoretical density and adequate dielectric properties were prepared by sintering PZT thick films at 800°C. Film thickness and porosity were estimated from SEM images, chemical composition was checked by EDS.

The thick films on PZT substrate appear to be a good solution for transducer fabrication due to their relatively high attenuation (which limits the required backing thickness) and high acoustical impedance, very close to the thick-film acoustical impedance, which increases the transducer bandwidth. Moreover, the high value of thickness coupling factor allows a high transducer sensitivity to be obtained.

1 Uvod

Ultrazvok se v medicini uporablja za slikanje in preiskave tkiv, organov itd. Bistveni del medicinskih ultrazvočnih naprav je pretvornik, ki ultrazvok oddaja in sprejema. To je tanka rezina piezoelektričnega materiala z elektrodama na obeh straneh, polarizirana po debelini in nalepljena na podporni dušilec (backing) (Slika 1). Dober pregled delovanja ultrazvočnih pretvornikov je opisan v referenci /1/.

Debelina piezoelektričnega materiala je enaka polovični valovni dolžini, pretvornik deluje pri resonančni frekvenci. Podporni dušilec ima dvojno vlogo, nosi aktivni element in absorbira večino akustične energije, ki jo element seva nazaj. Akustični impedanci aktivnega elementa in dušilca morata biti primerljivi, da se čim bolj zmanjša sevanje z zadnje strani pretvornika. Tako se zmanjša parazitno sevanje pretvornika. To sicer zmanjša njegovo občutljivost, se pa poveča ločljivost. Ker sta akustični impedanci pretvornika in tkiva zelo različni, se na prednjo stran aktivnega elementa nanese prilagoditveno plast, ki to razliko zmanjša.



Slika 1: Shematski prikaz ultrazvočega pretvornika

Za izdelavo pretvornikov se najpogosteje uporablja keramika na osnovi $Pb(Zr,Ti)O_3$ (PZT). Za izdelavo podpornega dušilca in prilagoditvene plasti se uporablja različne polimere. Za povečanje akustične impendance dušilca se polimeru dodaja kovinske delce, kot na primer volfram, za zmanjšanje pa votle steklene kroglice /2/.

Diagnosticiranje z ultrazvokom gre v smeri večje ločljivosti, kar pomeni uporabo višjih resonančnih frekvenc. Za frekvence okrog 20 MHz se tako uporablja piezoelektrične elemente, debele manj kot 50 μm. Piezoelektrične elemente se navadno izdeluje z rezanjem in tanjšanjem kosa keramike. Zaradi majhne debeline se lahko element pri obdelavi in lepljenju na dušilec poškoduje. Tem težavam se da izogniti z integracijo debele piezoelektrične plasti na primeren nosilec, ki bi imel tudi vlogo dušilca.

V prispevku opisujemo izdelavo in lastnosti integriranega ultrazvočnega pretvornika, ki smo ga izdelali z debeloplastno tehnologijo. Kot nosilec in dušilec smo uporabili porozno PZT in korundno keramiko, na katero smo s sitotiskom nanesli elektrodo in piezoelektrično PZT plast. Z dodatkom svinčevega germanata plasti PZT smo znižali temperaturo sintranja na 800 °C in s tem zmanjšali izgube svinčevega oksida med sintranjem.

2 Eksperimentalno delo

Prah PZT s sestavo $Pb(Zr_{0,53}Ti_{0,47})O_3$ (PZT53/47) smo pripravili z mešanjem oksidov in kalcinacijo pri 900 °C eno uro. Uporabili smo naslednje kemikalije: PbO, ZrO_2 in TiO_2. Po istem postopku smo pripravili tudi svinčev germanat (PGO). Za sintezo smo uporabili PbO in GeO_2, ki smo ju kalcinirali pri 650 °C 2 uri. Po sintezi smo oba prahova mleli v kroglicnem mlinu. Pasto za tiskanje smo pripravili iz PZT-prahu, dodatka 2 mas. % PGO in organskega nosilca z mešanjem v valjčnem mlinu.

Porozne PZT podloge (podporne dušilce) smo pripravili s stiskanjem prahu v jeklenem modelu s pritiskom 100 MPa in sintranjem pri 1100 °C eno uro. Končna geometrijska gostota podlage je bila 80 % teoretične gostote PZT. Korundne podlage smo pripravili z ulivanjem suspenzije Al_2O_3 prahu Alcoa A-16 v modele in s sintranjem pri 1700 °C 4 ure. Kot spodnjo elektrodo smo uporabili zlato pasto Remex 3242. Le-to smo natisnili, posušili in žgali pri 950 °C eno uro. Nato smo natisnili PZT- pasto, jo posušili in žgali pri 800 °C 8 ur v zaprti korundni posodi. Preseke vzorcev smo analizirali z vrstičnim elektronskim mikroskopom. Debeline plasti smo določili iz posnetkov mikrostruktur. Zgornjo zlato elektrodo smo nanesli z naprševanjem. Plasti smo polarizirali pri 150 °C v oljni kopeli pri jakosti električnega polja 12 kV/mm.

Izmerili smo električne, akustične in elektromehanske karakteristike debelinskega nihanja izdelanih pretvornikov /3/. Uporabili smo metodo meritve kompleksne impedance v bližini resonančne frekvence debelinskega nihanja vzorcev. Kompleksno impedanco vzorcev smo merili z impedančnim analizatorjem HP 4195. Za preprost resonator, ki prosto niha, to je v našem primeru piezoelektrična plast, je mogoče admitanco izračunati na osnovi enačbe /4/:

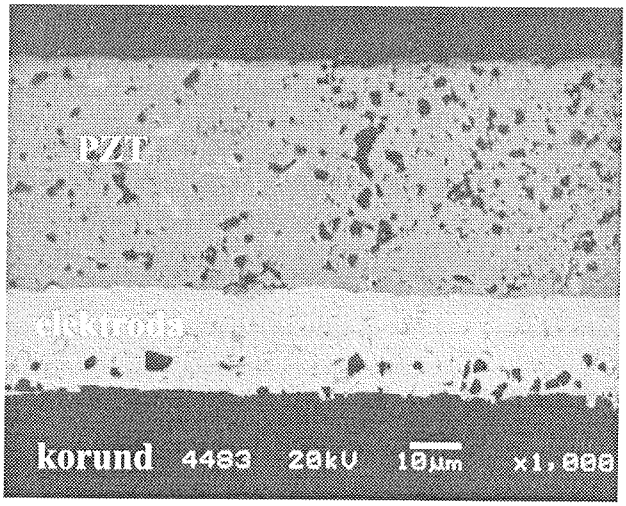
$$Y(\omega)=\left(\frac{i\omega\varepsilon_{33,r}^{S}\varepsilon_0 A}{t}\right)\left[1-k_t^2\frac{\tan\left(\frac{\omega t}{2v_t^D}\right)}{\left(\frac{\omega t}{2v_t^D}\right)}\right]^{-1} \quad (1)$$

kjer je ω kotna frekvenca (rad s^{-1}), $\varepsilon_{33,r}^{S}$ relativna dielektrična konstanta pri konstantni napetosti, ε_0 dielektrična konstanta za vakuum (F m^{-1}), A površina elektrode (m^2), t debelina (m), k_t elektromehanski sklopitveni faktor debelinskega nihanja, v_t^D vzdolžna hitrost nihanja (m s^{-1}), ρ gostota (kg m^{-3}), c_{33}^D prožnostni koeficient (N m^{-2}) in e_{33} piezoelektricni koeficient (C″/m^2). Številka tri pomeni smer polarizacije materiala.

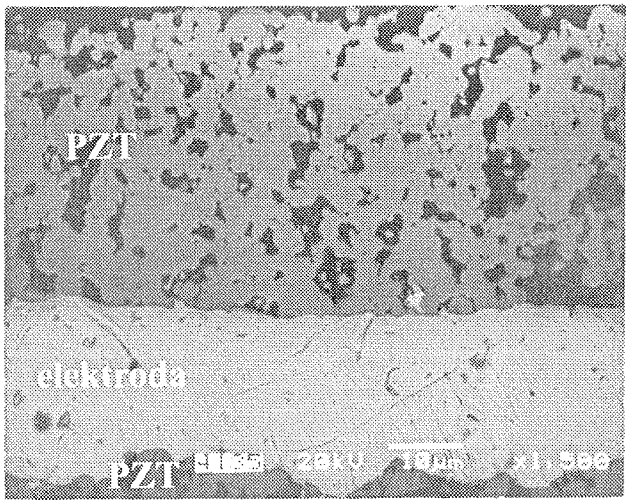
Ker je vzorec kompleksnejši, saj ima v stiku štiri plasti: podlago, spodnjo elektrodo, piezoelektrično plast in zgornjo elektrodo, smo za posnemanje vedenja realnega pretvornika uporabili model K. L. M. /5/. Ta omogoča simulacijo pretvornika v enodimenzionalni obliki kot tudi simulacijo električne impedance. Število spremenljivk v štiriplastni strukturi je zelo veliko. Za dosego zadostne natančnosti pri določanju elektromehanskih karakteristik piezoelektrične plasti je treba zmanjšati število neznanih karakteristik. Določili smo hitrost širjenja valovanja v nosilcu (korund in PZT), znani pa sta bili tudi gostota in debelina posameznih plasti.

3 Rezultati in diskusija

Na slikah 2 in 3 sta prikazana preseka PZT-plasti na korundni in PZT podlagi. Mikrostrukturi sta si podobni, le da ima plast PZT na korundu (slika 2) še svetlejšo fazo, ki vsebuje poleg PZT-komponent še Al in Ge, kar kaže na reakcijo med korundom in PZT-plastjo. Debelina spodnje zlate elektrode je 20 ± 3 μm.



Slika 2: Posnetek preseka PZT-plasti na korundnem nosilcu, prevlečenim z zlatom



Slika 3: Posnetek preseka PZT-plasti na PZT-nosilcu, prevlečenim z zlatom

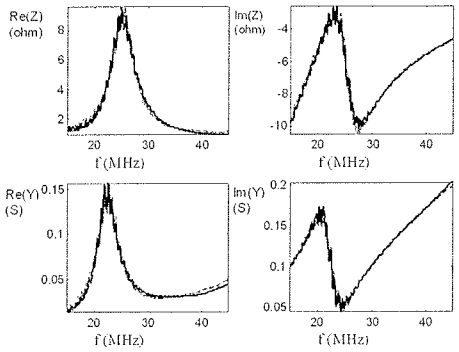
Debelina zgornje zlate napršene elektrode je v primerjavi z debelinami spodnje zlate elektrode in piezoelektrične PZT-plasti zanemarljiva (pod 0,2 μm). Meritve in simulacije vedenja vzorcev z zgornjo elektrodo in brez nje so pokazale, da ima zgornja elektroda zelo majhn vpliv na akustično impedanco pretvornika /3/. Zato se pri nadaljnjih simulacijah zgornja elektroda ni upoštevala.

V tabeli I so izmerjene oz. izračunane količine za dušeno valovanje v PZT in korundnem nosilcu. Ugotavljamo, da je dušenje PZT-nosilca za nekaj redov velikosti večje od korundnega.

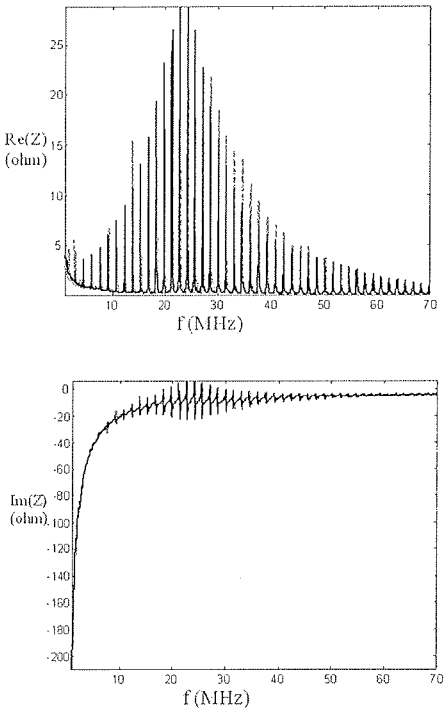
Tabela I: Lastnosti PZT- in korundnega nosilca (v_l vzdolžna hitrost valovanja, e debelina nosilca, α, dušenje valovanja v nosilcu, ρ gostota)

Nosilec	v_l (m s^{-1})	e (mm)	α (dB/mm/MHz)	ρ (kg m^{-3})
PZT	3005	1,94	0,26	6400
korund	1050	3,34	0,006	3900

V tabeli II so zbrane vse spremenljivke in geometrijske karakteristike, ki smo jih uporabili za končno prilagajanje izmerjenih in izračunanih impedanc vzorcev.



Slika 4: Izračunana (črna neprekinjena krivulja) in izmerjena (prekinjena siva krivulja) kompleksne impedance (Z) in admitance (Y) za vzorec 3 (PZT-plast na poroznem PZT) kot funkcija frekvence



Slika 5: Izračunana (črna neprekinjena krivulja) in izmerjena (prekinjena siva krivulja) kompleksne impedance (Z) za vzorec 2 (PZT-plast na korundu) kot funkcija frekvence

Tabela II : Elektromehanske lastnosti merjenih vzorcev (e debelina debelega filma , A površina zgornje elektrode , $\varepsilon_{33}^S/\varepsilon_0$ relativna dielektrična konstanta pri stalni napetosti, v_l vzdolžna hitrost valovanja, k_t debelinski sklopitveni faktor, f_a antiresonančna frekvenca, δ_m mehanske izgube, δ_e dielektrične izgube)

Vzorec (nosilec)	e (μm)	A (mm^2)	$\varepsilon_{33}^S/\varepsilon_0$	v_l (m s^{-1})	k_t (%)	f_a	δ_m (%)	δ_e (%)
1. korund	48	7,96	395	3375	16,5	38,4	2,4	3,6
2. korund	39	7,44	342	3940	39,7	50,5	1,5	2,0
3.PZT	35,5	7,69	347	3345	5,,7	4,0	5,0	10,0
4.PZT	35,5	7,54	334	3238	47	45,8	4,8	4,7

Debela plast PZT na poroznem PZT-nosilcu je zaradi velikega dušenja nosilca (nosilnega dušilca), akustične impedance, primerljive z impedanco debele PZT plasti ter velike občutljivosti pretvornika zaradi visokega sklopitvenega faktorja PZT-plasti na PZT (kt = 51%) uporabna za izdelavo pretvornika.

3.1 Simulacija pretvornika

S simulacijo K. L. M. /5/ smo preverili kvaliteto integrirane PZT-plasti na nosilnem PZT-dušilcu. Lastnosti posameznih elementov pretvornika so podane v tabeli III.

Za podporni dušilec je bila izbrana večja debelina, da bi zmanjšali vpliv odbojev z zadnje strani PZT. Da bi določili optimalno debelino spodnje zlate elektrode, sta bili simulirani dve debelini, in sicer 10 in 20 μm.

Tabela III: Lastnosti posameznih elementov pretvornika, ki smo jih uporabili pri simulaciji (Zakustična impedanca, e debelina elementa)

Element	Z (MRa)	e (mm)
Piezoelektrik	19,4	0,0355
Podporni dušilec	19,2	5,0
Spodnja elektroda	47	0,010 ali 0,020
Prilagoditvena plast	2,8	0,023

Simulirali smo tri različne oblike pretvornika, ki so prikazane v tabeli IV, slika 5 pa prikazuje simuliran elektroakustični odziv pretvornika PZT-plasti na poroznem PZT-nosilcu.

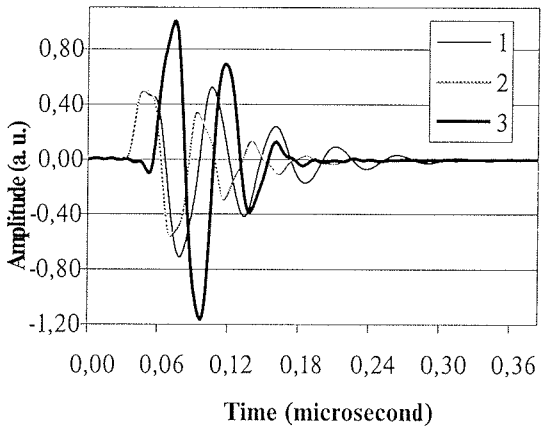
Tabela IV: Karakteristike simuliranih pretvornikov

Oblika	1	2	3
Prilagoditvena plast	ne	ne	da
Debelina spodnje elektrode(μm)	20	10	10
Srednja frekvenca delovanja pretvornika (MHz)	25,0	26,5	23,0
Pasovna širina (-6 dB) (%)	42	55	64

Rezultati simulacije kažejo, da je za paraktično uporabo, večjo občutljivosti in pasovno širino pretvornika potrebna tanjša spodnja elektroda in prilagoditvena plast.

4 Sklepi

Z debeloplastno tehnologijo smo izdelali integrirani ultrazvočni pretvornik in se izognili tehnološkim težavam izdelave pretvornika, to je tanjšanje in lepljenje piezoelektrika ter izdelavo podpornega dušilca. Izdelali smo ga tako, da smo najprej pripravili porozeni PZT-nosilec in nanj nanesli in nato žgali spodnjo elektrodo ter nanjo nanesli piezoelektričen PZT-plast. Po žganju pri 800 °C je bila debelina PZT-plasti okoli 40 μm. Nato smo napršili zgornjo zlato elektrodo. Piezoelektrično plast smo nato polarizirali in karakterizirali. PZT-plast na poroznem PZT ima visok debelinski sklopitveni faktor ter veliko dušenje. PZT-plast na poroznem PZT-nosilcu je zaradi velikega dušenja, akustične impedance, primerljive z impedanco debele PZT-plasti ter velike občutljivosti uporabna za izdelavo ultrazvočnega pretvornika.



Slika 6: Simuliran elektroakustični odziv treh oblik pretvornika (tabela IV)

Zahvala

Delo so podprli: Ministrstvo za šolstvo, znanost in šport RS, projekt EUREKA Pimet EU 1664 in projekt 5.OP Piramid.

5 Literatura

/1/ M. Lethiecq, F. Levassort, G. Feuillard, L. P. Tran-Huu-Hue, Piezoelectric materials for ultrasonic medical diagnostics, Piezoelectric material for end users, Interlaken 2002, to be published

/2/ T. N. Nguyen, M. Lethiecq, F. Levassort, L. Pourcelot, Eksperimental verification of elastic properties using scattering approximation in (0-3) connectivity composite material, IEEE Trans. Ultrason. Ferroelect., Freq. Contr., 43 (1996), 640-645

/3/ M. Kosec, J. Holc, F. Levassort, P. Tran-Huu-Hue, M. Lethiecq, Screen-printed $Pb(Zr,Ti)O_3$ thick films for ultrasonic medical imaging applications, IMAPS 2001; 34th International Symposium on Microelectronics , Baltimore ZDA, 9-11. oktober 2001, 195-200

/4/ D. Royer and E. Dieulesaint, "Ondes élastiques dans les solides : propagation libre et guidée", 1 (1996), Masson, Pariz

/5/ R. Krimholtz, D. A. Leedom and G. L. Mathei, "New equivalent circuit for elementary piezoelectric transducers", *Electron. Lett.*, 38 (1970), 398-399

Janez Holc, Marija Kosec
Jožef Stefan Institute, Jamova 39,
1000 Ljubljana, Slovenia
Franck Levassort,
Louis Pascal Tran-Huu-Hue, Marc Lethiecq
LUSSI/GIP Ultrasons, EIVL, François Rabelais
University, BP 3410, 41034 Blois cedex, Francija

Prispelo (Arrived): 06.06.2002 Sprejeto (Accepted): 25.03.2003

SYNTHESIS OF ANALOG INTEGRATED CIRCUITS

Andrej Vodopivec

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

Key words: analog circuit synthesis, analog circuit, circuit analysis, circuit optimization, topology selection

Abstract: A new approach to synthesis of analog circuits combines circuit rule based topology selection with circuit optimization. With the advent of super fast microprocessors it is possible to choose out the dimensions of devices in an analog circuit which suit best the set of requests using optimization techniques build into modern circuit simulation tools. These tools still basically only analyze a given circuit so an extra layer is needed to select a topology which suits the set of requirements. The selection can be efficiently described for any analog block using programming language OPAN.

Sinteza analognih integriranih vezij

Ključne besede: sinteza analognih integriranih vezij, analogno vezje, analiza vezja, optimizacija vezja, izbira topologije

Izvleček: Nov pristop k sintezi integriranih analognih blokov kombinira na pravilih temelječo izbiro topologije z optimizacijo vezja. Z izboljšavami hitrosti delovanja mikroprocesorjev postaja uporabna izbira dimenzij komponent analognega vezja, ki najbolje izpolnjuje dane zahteve, z uporabo optimizacijskih postopkov vgrajenih v moderna orodja za simulacijo vezij. Ta orodja v osnovi vedno samo analizirajo dano topologijo, zato je za sintezo pomembna izbira ustrezne topologije. Iziro je mogoče učinkovito opisati z jezikom OPAN.

1. Introduction

Accurate and fast circuit simulations of integrated circuits is a critical step in integrated circuit design. Traditional analog circuit simulators like SPICE /1/, RELAX /2/ were slow while new simulation technologies like SPECS /3/ traded accuracy for speed. Both were able to analyze a general analog circuit but were too slow to be used in analog integrated circuit synthesis so device sizing /6/ was used which considerably simplified analysis of specific topologies.

A set of rules /7/ was used to describe changes of properties of the circuit with new sizes of selected devices as well as to guide the selection of new device sizes to ensure requested circuit properties. Circuit modeling /8/ turned out to work fine for operational amplifiers in a tool when used by experienced analog designers. Novice designers were not able to use the synthesis tool because they tend to over-specify the design in so many 'weird' way that the set of rules became unmanageable.

With the advent of super fast microprocessors it is possible to execute hundreds or thousands of circuit analysis required by analog circuit synthesis within a few hours. Building optimization loops in SPICE /4/ turned out very efficient. A number of optimization methods /5/ enable modern implementations of SPICE to choose out optimal device dimensions for a selected set of requirements.

What is left for a synthesis tools (fig. 1) is:

- design of optimization scripts in such a way that a meaningful result is obtained no matter what properties are selected,
- selection of the appropriate topology (e.g. A class) and variation (e.g. micro power), and

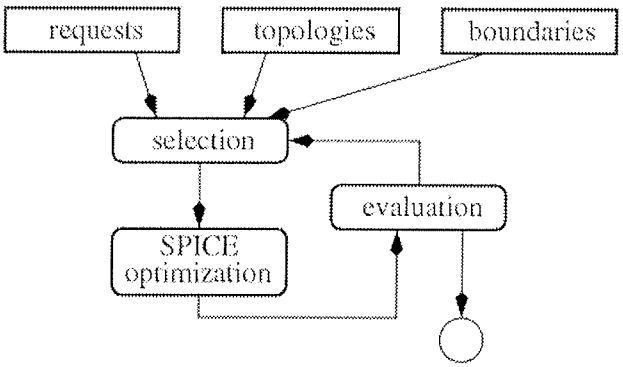


Fig.1: Synthesis of analog circuit

- automation of inter topology boundary selection depending on process parameters used to model devices of the current technology.

2. Describing topology

The topology description consists of:

- circuit netlist
- process parameters
- requested properties
- optimization loop
- corner analysis loop

The description is build modularly so parts of requests and optimization can be omitted to simplify and speed up the optimization process.

The description starts with circuit netlist and process parameters as any other SPICE input file. The netlist is a typical SPICE netlist with the exception of symbolic values

used to denote properties (sizes) of some devices. The sizes of these devices are going to be optimized while others will keep the assigned values. The process parameters are used to pass values to device models of the technology used in the design.

Passing requested properties from the user interface to the optimization script is not straightforward. Since the user is at liberty to select values for the circuit's properties and omit them as well attention must be paid to select default values for certain properties to obtain meaningful result no matter what properties are selected. Some topologies even require properties not available in the user interface to ensure operation.

Since circuit optimization using circuit simulation is a very CPU time consuming operation the optimization loop must be carefully designed. Omission of analysis required by "don't care" properties may result in substantial speed up of the optimization process.

Analysis of the optimized circuit using simulation corners is very important in evaluation of the results of the synthesis. Corner analysis should be placed inside the optimization loop but since this could easily result in prolonging the optimization time by an order of magnitude or more it is more practical to over-specify the circuit's properties a little.

3. Topology selection

Analog block (e.g. operational amplifier) can be build using a databases consisting of tens of topologies (e.g. A class) and variations (e.g. micro power, rail to tail common mode). Optimization of all available topologies for the requested set of block properties and selection of the best final candidate is still a tool CPU time consuming operation. Selection of a small set of topologies to be optimized and evaluated becomes crucial.

Expert knowledge of experienced analog designers is required to set up a good selection procedure. Hard-coding topology selection in C soon turned out to be too complicated for the analog designer so a simple programming language OPAN was used. It consists of assignment statements and control structures. The only control structures needed to describe the topology selection are if-then-else and goto. Despite long formulae the input deck is quite readable since it can be written in TEX.

$$/* \text{ rule } 16 */$$

$$A_0 = A_{0o} * \sqrt{\frac{(\frac{W}{L})_d}{(\frac{W}{L})_{d_o}}} * \sqrt{\frac{I_b}{I_{bo}}} * A_0 t;$$

When a number of topologies or variations meet selection criteria all are evaluated and the final selection is made with a cost function weighting requests and results of all optimized circuits.

4. Adjusting selection to new fab or new minimum feature size

Topology boundaries define the range of requested properties which are optimally synthesized using a given topology. These boundaries vary with the fab or new minimum feature size. Where one topology is best for the given set of request in one process (e.g. AMS 0.8μm) another may be in a different process (e.g. TSMC 0.25μm).

An attempt was made within intensive study of various topologies of operational amplifiers to find general principles which could be used in mapping topology properties from one process to another using process parameters.

As we were not able to find any expressions which could be used in mapping topology boundaries we used optimization capabilities of SPICE building topology boundary tables. A characteristic set of properties was selected for each topology. The topology is then optimized for each property of the set and the optimal values were used in definition of topology boundaries.

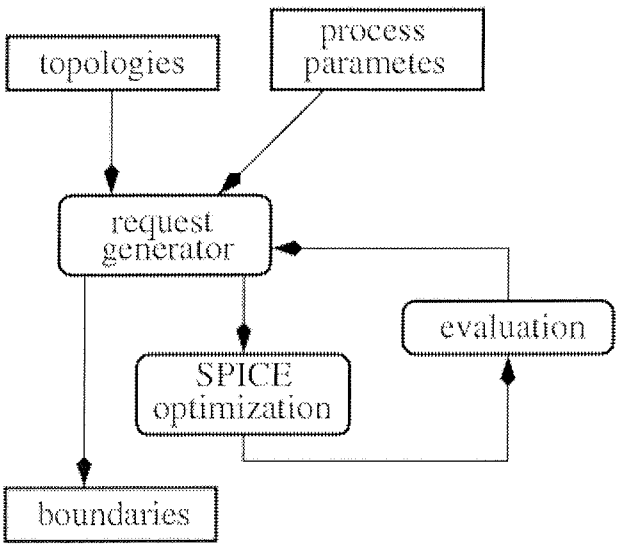


Fig. 2: Topology boundary selection

5. Conclusions

Optimization capabilities build into SPICE make analog synthesis more reliable. What is needed on top of SPICE is good topology selection. This can be accomplished using a rule based programing language OPAN. When applying analog synthesis to real world cases defining topology boundaries becomes a must. SPICE optimization can be used in building topology boundary tables. Computers are still too slow to evaluate corner cases during optimization in real time. As trade offs are always used in design the final request are usually set only when all constraints are evaluated. With this final request one can execute a multi-day corner cases optimization.

References

/1/ L. W. Nagel, "SPICE2: A Computer Program to Simulate Semiconductor Circuits", Elect. Res. Lab. Report ERL-M520, University of California, Berkeley, 1975

/2/ A. R. Newton, A. L. Sangiovanni-Vincentelli, "Relaxation-Based Electrical Simulation", IEEE Trans. Computer-Aided Design, vol. CAD-3, pp 308-330, May 1984

/3/ CC. Visweswariah, R. A. Rohrer, "Piecewise Approximation Circuit Simulation", IEEE Trans. Computer-Aided Design, vol. CAD-10, pp 565-576, May 1991

/4/ J. Puhan, T. Tuma, "Optimization of Analog Circuits with SPICE 3F4", Proc. of the European Conference on Circuit Theory and Design, Budapest, Hungary, pp 177-180, 1997.

/5/ J. Puhan, T. Tuma, I. Fajfar, "Optimization Methods in SPICE, a Comparison", Proc. of the European Conference on Circuit Theory and Design, Stresa, Italy, Vol. 2, pp 1279-1282,1999.

/6/ J. Trontelj, L. Trontelj, A. Pleteršek, A. Vodopivec, G. Shenton, "Synthesis and Layout Compilation Automation of Mixed Analog-Digital ASICs", Proc. Of the IEEE 1990 ISCS, New Orleans pp 816-819, 1990.

/7/ J. Trontelj, L. Trontelj, A. Pleteršek, A. Vodopivec, G. Shenton, "Rule Based CAD Tools for Analog Circuit Synthesis and Layout Compilation", Proc. Of the IEEE 1990 CICC, Boston, pp. 14.8.1-14.8.4, 1990.

/8/ A. Vodopivec, J. Trontelj, "Modeling for Analog Circuit Synthesis", Proc. 1993 MIEL, Bled, pp 61-64, 1993.

Andrej Vodopivec
Faculty of Electrical Engineering, University of Ljubljana, 1000 Ljubljana, Slovenia

Prispelo (Arrived): 06.06.2002 *Sprejeto (Accepted): 25.03.2003*

REAL TIME DECODER FOR CODED SIGNALS MIXED WITH NOISE

Slavko Starašinič

Faculty of Electrical Engineering, University of Ljubljana, Slovenia

Key words: smart card, identification card, contact less integrated circuit, reader, recognition circuit, antenna, bit-stream, data rate, time window.

Abstract: An approach to decoding of serial data bit-stream is described. This approach is based on moving the time window, which makes an average number of received pulses in real time. The average value of the number of pulses is compared on the comparator with a hysteresis and its output shows decoded logic state. Such technique eliminates same spurious and some missing pulses in the incoming data bit-stream and increases the range of communicating devices.

Sprotno dekodiranje s šumom pomešanih signalov

Ključne besede: pametna kartica, identifikacijska kartica, brezkontaktno integrirano vezje, čitalnik, razpoznavno vezje, antena, tok podatkov, podatkovna hitrost, časovno okno.

Izvleček: V članku je opisan način za dekodiranje niza serijskih podatkov. Zasnovan je na premikajočem časovnem oknu, ki povprečuje v realnem času število sprejetih impulzov. Komparator s histerezo primerja povprečno vrednost števila impulzov in na izhodu se pojavi dekodirana logična vrednost. S takšno tehniko uspešno izločamo nekatere lažne in nekatere manjkajoče impulze v podatkovnem nizu, ki prihaja na vhod vezja ter s tem povečujemo razdaljo na kateri lahko uspešno komunicirata identifikacijska kartica in čitalnik.

1. Introduction

There are several methods to decode signals described in ISO standard /1/. Many systems use microprocessors and it is possible to save a part of input data bit-stream and recognize the coded information using this complex hardware. The goal of this work was to develop the minimized hardware, which is able to decode the input signal in real time, because the application requires the integration for high volume production where the reduction of cost is an important parameter.

The developed hardware is located on the chip, which is called reader and receives the signals from the smart card. The distance between smart card and antenna of the reader defines the quality of the received signal. If the smart card is very close to the antenna of the reader, the received signal is very strong and correctly detected in the analog part of the reader chip. This demodulated signal is used as input signal to the described recognition circuit, which decodes incoming bit-stream into the logic data. But increasing distance between smart card and antenna of the reader decreases demodulated signal and pushes it into the noise. The influence of this noise is shown on the demodulated bit-stream as a spurious and missing pulses. The described recognition circuit counts pulses in the moving window. This numbers are compared on the comparator and its output is high, if there is a group of pulses and is low, if there are not pulses or there are only a few spurious pulses. Signal from the comparator is synchronized and formed in the last block. Serial data on the output are decoded and corresponded bit_clock is shifted regarding to the upper mentioned signal.

2. Coded signals

Signal transmitted from the smart card to the reader is coded. This code is defined in ISO standard /1/ and can be used for two data rates. This design is performed for higher data rate (26.48kb/s) and all definitions for logic signals and preambles are valid for this speed of communication. Four times lower data rate has the same definitions. The number of pulses shall be multiplied by four and also all other times will increase by this factor. Increased number of pulses for logic definitions decreases the data rate, but increases the reliability of communication.

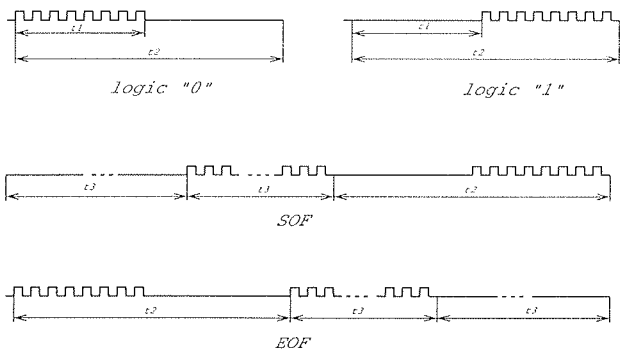


Fig. 1. Definition of coded signals using ISO standard /1/.

Code definitions for high data rate are shown on the figure 1. All definitions use the same sub-carrier frequency. Logic "0" starts with 8 pulses and then follows unmodulated signal with the same duration as is duration of 8 pulses. Unmodulated signal with duration "t1" and additional 8 pulses define logic "1".

The successful protocol with serial coded data requires two additional definitions:

- signal defining beginning of data: start of frame or SOF.
- signal defining end of data: so-called end of frame or EOF.

These signals are also shown on the fig. 1. SOF is defined with unmodulated signal of duration "t3" and then followed by 24 pulses of the same duration "t3" and a coded logic "1". The SOF is required for the synchronization of described recognition circuit. The correct received SOF starts the timing of recognition circuit. After this signal immediately starts decoding of incoming signals and bit clock is also generated. The last important definition is EOF. This signal tells the reader that the transmission of the data is concluded. EOF consists of logic "0" and followed by 24 pulses and an unmodulated signal with the same duration as the duration of 24 pulses. After an EOF is decoded, the receiver is reset and waits the next SOF.

3. Description of the recognition circuit

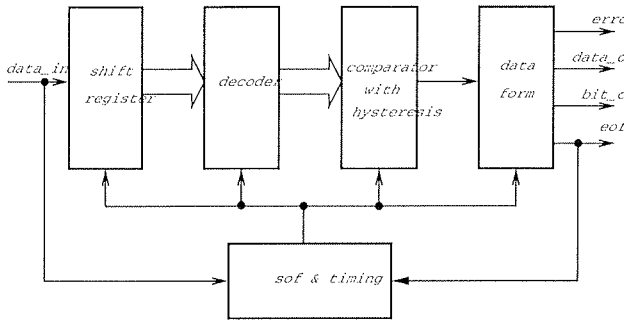


Fig. 2. Block diagram of recognition circuit

Block diagram of the recognition circuit is shown in the fig. 2. Demodulated signal from analog front end is connected to the input of this recognition circuit. Data are checked and after receiving SOF the complete recognition circuit starts with the appropriate timing. Serial data are shifted in the shift register and on its parallel outputs are connected to the inputs of decoder. This block check the numbers of "1" and decodes this number in a binary sequence. These signals are checked on the comparator. Output of comparator goes high, if the number on the inputs is changed from 4 [100] to 5 [101] and goes low, if the number is changed from 4 [100] to 3 [011]. Output from comparator is connected to the data form block where output signals are formed. This block generates output data bit-stream, bit clock and EOF. Bit clock is slightly shifted regarding the data bit-stream. Decoded EOF stops the operation of the complete circuit and the system is now waiting for the next input bit-stream with correct SOF.

3. Principle of operation

The most important signals are shown on the fig. 3., which explain developed algorithm and its realization in the recognition circuit. Upper part of the fig. 3. shows ideal signals. Data bit-stream on the input "data_in" shows "1", "0" and "1" and this shifted and decoded signal is present on the input of comparator as a number of pulses. Shift register creates time window where only eight possible pulses are visible at each time. Hysteresis is built in the comparator. Its output signal is connected to the data form block where this signal is formed and synchronized with the bit clock. Incoming pulses mixed with the noise are shown on the lower part of the fig. 3. There are some missing pulses in the group of pulses and a few additional pulses in unmodulated region of received bit-stream. This inconvenience decrease the span of received numbers at the input

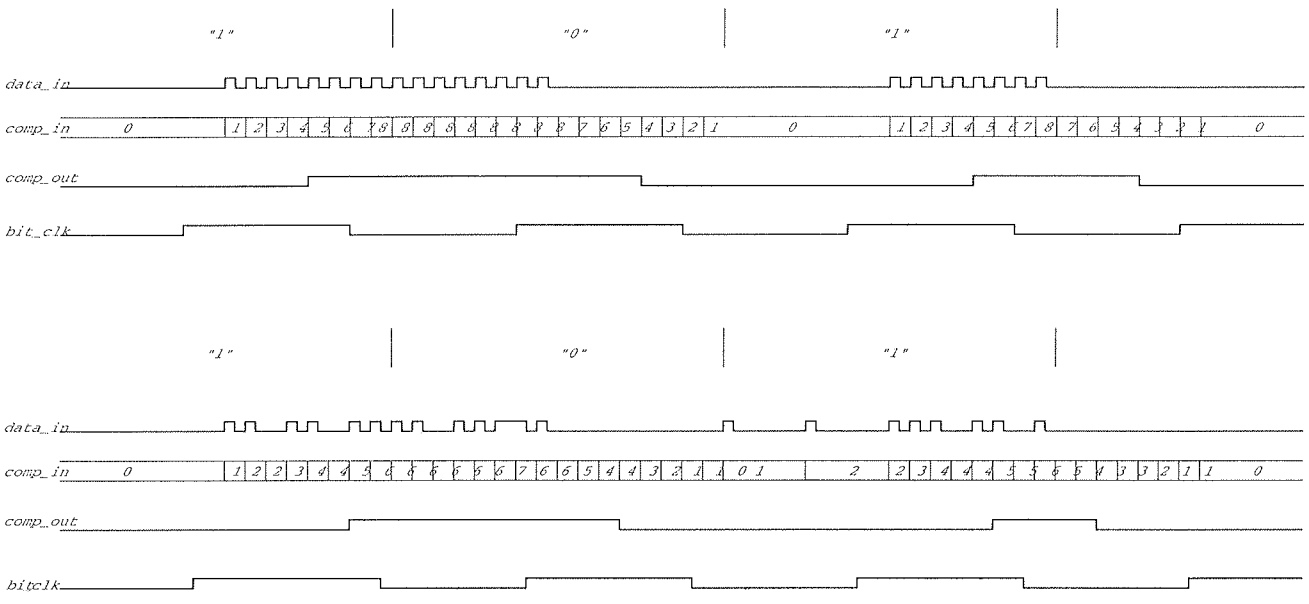


Fig. 3. Main signals of the circuit

Fig. 4. Simulation with correctly decoded output

of the comparator. That means the moving window of selected numbers makes an average of received pulses. The comparator produces logic signal from these numbers, but this signal is now shifted and its duration varies compared to the signal with ideal conditions in the upper part of the figure 3. Shift of the comparator output and its duration are taken into account when bit clock signal is generated. The result of this approach is signals with the jitter. If the noise in the incoming "data_in" signal is to high, this signal is to much corrupted and of course recognition circuit can not decode these signals correctly. In this case the signal "error" goes high.

4. Results

Many simulations were made with ideal and corrupted signals. The fig. 4 shows simulation with corrupted signal and its correct response. Input signal "data_in" mixed with the strong noise is shown on the fig. 5. This signal is too corrupted and comparator can not correctly decode it and "error" goes high.

5. Conclusion

Digital block is designed, which decodes signals from the smart card. The distance from smart card to the antenna of reader defines the quality of received signal. Correct decoding of corrupted incoming data bit-stream increase the reading distance between smart card and reader. The data protocol used corresponds to international standard.

References

/1/ ISO/IEC 15693-2 Identification cards - Contactless integrated circuit(s) cards -Vicinity cards - Part 2: Air interface and initialization, INTERNATIONAL STANDARD.

Slavko Starašinič,
Faculty of Electrical Engineering,
Tržaška 25, 1000 Ljubljana, Slovenia.
E-mail: slavko@kalvarija.fe.uni-lj.si

Prispelo (Arrived): 06.06.2002 *Sprejeto (Accepted): 25.03.2003*

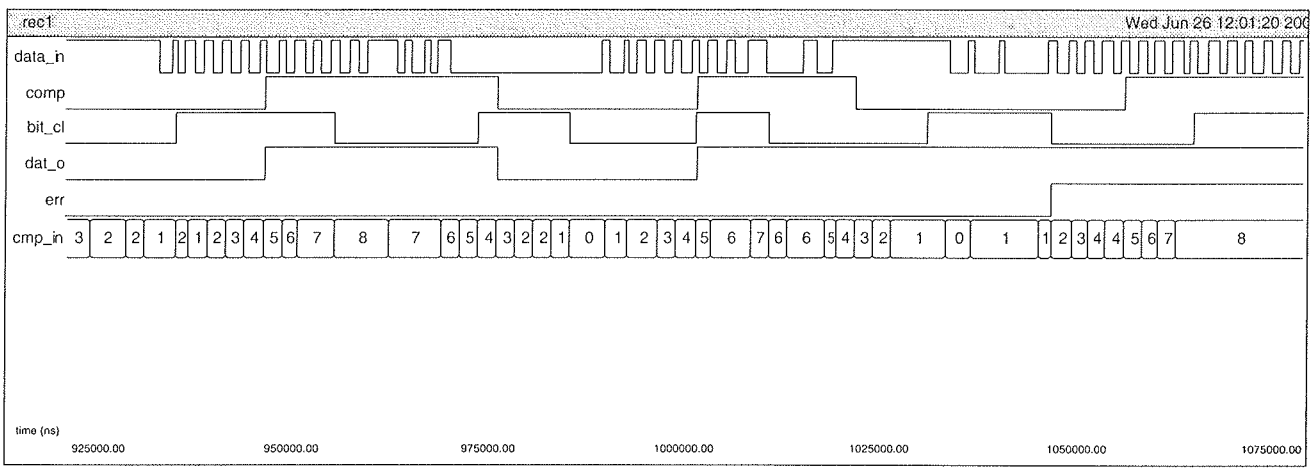


Fig. 5. Noisy signal on the "data_in" and corrupted output

RATIOMETRIC-TO-SUPPLY VOLTAGE OUTPUT BUFFER DESIGN

Anton Pleteršek

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

Key words: ratiometric, transfer function, precision, analog-to-digital, converter, gain control

Abstract: Paper is focused to design the ratiometrical - signal processing gain stage. The stage is a part of complex integrated system and is operating in a narrow frequency band. Its main part is a programmable operational amplifier, having controlled gain, automatic gain tracking with supply voltage and automatic signal ground tracking with supply voltage VDD. The stage cut-off frequency is 100Hz, the control loop -3dB frequency is at 1kHz. Precision of ratiometric control is 0.325%, accuracy is 0.6%. The main topics of present paper is to show ratiometric system behavior by the system description with its transfer function and verification using HSPICE and the MatLab tooling. The circuit is implemented in 0.8um CMOS technology.

Končna ojačevalna stopnja z lastnostmi, nastavljivimi z napajalno napetostjo

Ključne besede: prevajalna funkcija, točnost, AD pretvornik, kontrola ojačanja, sorazmeren napajanju.

Izvleček: V članku je opisana stopnja, katere lastnosti sledijo spremembi pozitivne napajalne napetosti v določenem frakvenčnem pasu. Vezje je zasnovano tako, da sledijo napajalni napetosti - ratiometrično - le lastnosti procesiranega vhodnega signala. Sestavlja ga: DC pretvornik, merilnik napajanja, AD pretvornik, regulacijska zanka za sledenje ojačanja in signalne mase. Cilj članka je opisati vezje s prevajalno funkcijo in jo verificirati s simulacijami (HSPICE), z analizo (MatLab) ter z meritvami na integriranem VLSI vezju. Vezje, kot sestavni del integriranega sistema, je bilo procesirano v 0.8um tehnologiji CMOS.

I. Introduction

The number of different integrated systems have to have output signal proportional to the supply voltage VDD. Such systems require precisely defined power supply rejection outside useful signal pass as well as good definition for signal transfer function. The ratiometric-to-supply feature is a useful function when processed signal besides its main function i.e. follow the supply voltage changes.

The main concept of ratiometry may operate in continuos time or may operate digitally weighted via analog-to-digital converter (ADC) where ratiometric signal ground (RAGND) is generated as continous analog signal with appropriate frequency response. The signal to be processed needs to be converted from system - constant analog ground (AGND) to ratiometric ground (AGND) by processing of the input signal in the level-shift gain stage.

II. Ratiometric signal processing stage

The main system transfer function is expressed as in (1). Rgr is ratiometric-variable resistor with VDD supply. Rg is programmable gain resistor (Fig. 1 in Fig. 2). To complete the function (1) with all the sub-definitions, the transfer function (1) will clearly define the supply voltage dependence, frequency response and temperature stability.

$$T(S) = (1+\frac{R_{gr}}{R_g}) \cdot \frac{1}{1+sR_1C_1} \qquad (1)$$

$$(1+\frac{R_{gr}}{R_g})$$

is the gain of output stage,

and

$$\frac{1}{1+sR_1C_1}$$

is the low-pas filter (100Hz) transfer function.

Gain stage is in non-inverting gain configuration. The gain of the stage is linearly dependent on the VDD supply voltage. Gain is programmable from A=3 to A=12 or 12 in 8 - linear steps by varying the resistor Rg. The resistor Rgr is automatically adjusted with varying of the supply voltage VDD in the range of -5% to +5% from its nominal value (5V) and nominal gain of (3 to 6). Ratiometric-gain-setting precision is 0.325% with worst-case accuracy of 0.6%. Used operational amplifier is class AB stage with open-loop gain of 104dB_min and has input referred noise of 260nV/sqrt(Hz)@ 10Hz to 93nV/sqrt(Hz)@100Hz.

Filter is placed in signal path, having no influence on supply voltage-dependent analog ground (RAGND) and vice-

versa. That principle guarantees good common-mode rejection of the output buffer and therefore no distorts the output signal. Filter in supply voltage (Vd generation) influences therefore only on output stage-gain with its low-pass characteristic. Ratiometric block is shown in Fig. 1, the simplified block diagram is on Fig. 2.

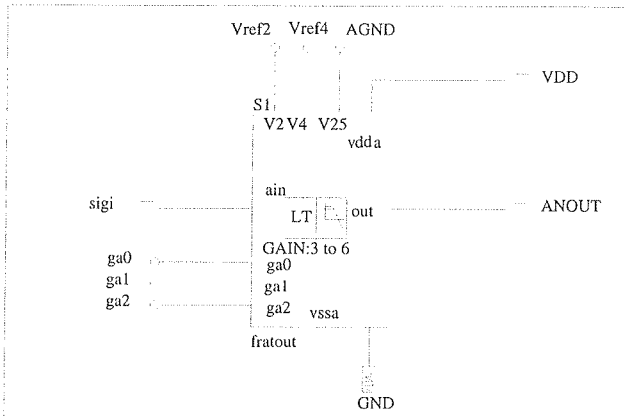


Fig. 1: Ratiometric output gain stage: sigi *is input signal, referred to AGND,* ANOUT *is the gained output signal, referred to ratiometric RAGND.*

A. Ratiometric gain

The feedback network resistor *Rgr* is controlled to follow the supply voltage variations *VDD*, and is included in *Vd*:

$$R_{gr} = R_{nf}[184 + \frac{40V_d}{V_{ref4}} - 4]$$

Rnf is a unit resistor,

$$i = \frac{40V_d}{V_{ref4}} - 4$$

the *i* is a digital code from flash ADC converter, converting the *Vd* voltage which is proportional to *VDD* supply. The accuracy of the code is defined by INL and switching point, which mainly depends on resistors matching and comparators offset voltage.

B. Ratiometric analog voltage Vd

$$V_d = (\lambda \cdot \frac{0.8x}{2x} \cdot VDD - V_{ref2}) \cdot (1 + \frac{R_{df}}{R_{da}}) + V_{ref2} \qquad (3)$$

Vd voltage depends on supply voltage *VDD* and is used for ratiometric gain tracking (1), (2). λ is voltage coefficient of the implanted well resistors. The cutoff frequency of voltage *Vd* is close to 1kHz and is expressed by:

$$Vd(j\omega) = (VDD \cdot \frac{R_3 \,\|\, \frac{1}{j\omega C_2}}{R_0 + \frac{1}{j\omega C_3} \,\|\, (R_2 + R_3 \,\|\, \frac{1}{j\omega C_2})} - V_{ref2}) \cdot \left[1 + \frac{R_{df}}{R_{da}}\right] \qquad (4)$$

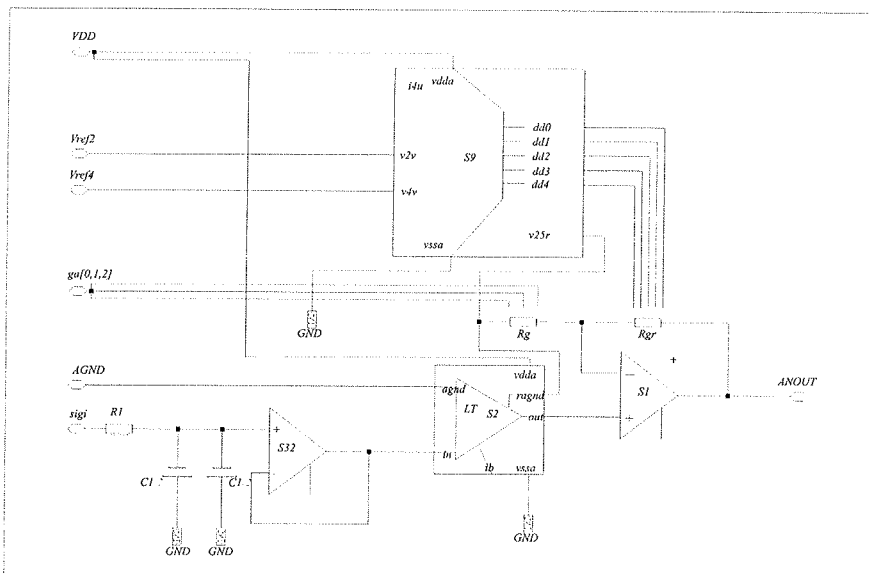


Fig. 2: Ratiometric gain stage, consists of the ratiometric signal ground generation, ratiometric gain control by measuring the supply voltage VDD with gain-control stage via 5 bit AD converter (S9), signal converter (S32), active low-pass filter, and programmable and supply voltage controlled gain - class AB operational amplifier (S1). Signal converter frequency response is also described. It acts as a common mode signal on the input terminals of the output amplifier (s1). The output amplifier has a wide frequency range (BW=9MHz) and has high input common-mode-rejection. RAGND in signal converter is therefore omitted from Transfer function (1), it is included in term Rgr.

C. AD converter

The 5 bit flash AD converter based on resistors string, array of 32 comparators and decoding logic to convert thermometer code to 5 bit binary code. The switching point of the comparator is:

$$V_{sw} = V_{refi} + V_{off}$$

where *Vrefi* is the voltage on the i-th tab of the AD resistor string, supplied by the reference voltage *Vref4* and is:

$$V_{refi} = \frac{V_{ref4}}{40R_n} \cdot (i+4)R_n + \frac{V_{ref4}}{2^5} \cdot \sum_{k=1}^{i} \frac{\Delta R_{nk}}{R_n}$$

$$V_{refi} = V_{ref4} \cdot [\frac{i+4}{40} + \frac{1}{2^5} \cdot \sum_{k=1}^{i} \frac{\Delta R_{nk}}{R_n}]$$

The digital output code is generated when the voltage on i-th tab equals to the analog voltage *Vd*, including comparator offset voltage *Voff*:

$$V_d = V_{sw}$$

From here the digital code i can be found:

$$i = \frac{40V_d}{V_{ref4}} - 4 - \frac{40}{2^5} \cdot \sum_{k=1}^{i} \frac{\Delta R_{ni}}{R_n} \qquad (5)$$

If we assume ideal resistors matching, the last term can be omitted, else the term should be taken too. The maximum accuracy errors may occur in the middle of the resistor string and cause the maximum INL error of:

$$|INL|_{\max} = \frac{V_{ref4}}{2} \cdot \frac{\Delta R_{nk}}{R_n}$$

Rn is resistive unit in the string, *ΔRnk* is the value of the resistance error (difference from ideal *Rn*) due the mismatch. Maximum error therefore may happen around nominal supply voltage.

D. Reference voltage Vref4

Reference *Vref4* is a *G1* times magnified voltage, proportional to the bandgap of silicon:

$$V_{ref4} = G_1 \cdot [V_{be}(T) + A \cdot \frac{k}{q} \cdot T \cdot \ln\left\langle \frac{J2}{J1} \right\rangle] \qquad (6)$$

Vbe is voltage of the forward biased base-emitter junction of the vertical bipolar device, *kT/q* is thermal voltage and is 26mV at 300K, J2/J1 is emitter-current density ratio. *Vref4* is constant with temperature and may vary for 20 ppm/C from its nominal value of 4V.

E. Reference voltage Vref2

Reference *Vref2* is reference voltage based on silicon-bandgap (*Vbgr=1.206V*):

$$V_{ref2} = G_2 \cdot [V_{be}(T) + A \cdot \frac{k}{q} \cdot T \cdot \ln\left\langle \frac{J2}{J1} \right\rangle] \qquad (7)$$

Voltage is stable within 20 ppm/C and has nominal value of 2V.

F. Gain programmability

The gain is adjust via trimming circuit using 3 bits (Fig. 2) which gives eight linear gain steps from nominal gain of 3. Resistors are polysilicon, temperature coefficient is cancelled-out from transfer function (1), also absolute value variations with process is canceled-out due to be fact that amplifier has sufficient open-loop gain and high enough output driving capability:

$$A_{cl} = \frac{A_{ol}}{1 + A_{ol} \cdot \frac{R_g}{R_g + R_{gr}}}$$

Aol is open loop of the class AB differential amplifier and is 104dB.

G. System signal ground

System analog ground voltage (*AGND*) is gained bandgap voltage *Vbgr* of 1.206V:

$$AGND = (1 + G_3) \cdot V_{bgr}$$

and is stable (20 ppm/C) at nominal value of 2,5V.

H. Ratiometric analog ground

The voltage is generated from *VDD* supply as a front-end analog signal to the 5 bit AD converter. The resistor divider network (R0, R2, R3) of (1, 0.2, 0.8 Mohm) and capacitors (C2, C3) of (100pF, 100pF) form a low-pass filter followed by unity gain buffer-driver.

RAGND at ω=0

$$RAGND = VDD \cdot \frac{R_2 + R_3}{R_0 + R_2 + R_3} = \frac{VDD}{2}$$

$$RAGND(j\omega) = VDD \cdot \frac{\frac{1}{j\omega C_3} \| (R_2 + R_3 \| \frac{1}{j\omega C_2})}{R_0 + \frac{1}{j\omega C_3} \| (R_2 + R_3 \| \frac{1}{j\omega C_2})} \qquad (8)$$

and has voltage coefficient due to implanted resistors used.

III. Ratiometric signal tracking with ratiometric analog ground voltage

The signal tracking stage is an inverting gain stage acts as a DC level shift to move the signal-voltage from *AGND* level to ratiometric ground level *RAGND*. Buffer amplifier is a class A stage with open-loop gain (Aol) of 99dB, gain bandwidth of 1.2MHz and input referred noise 88nV/sqrt(Hz)@10Hz to 28nV/sqrt(Hz)@100Hz. Close-loop gain error of the stage with a gain of 0dB is *[Aol/(1+Aol)]* and is negligible. The output voltage of the stage is:

$$V_{rsigi} = RAGND + V_{sigi} = \frac{VDD}{2} + V_{sigi}$$

where is *Vsigi* a signal voltage, relation is valid for ω=0.

IV. Output stage frequency response

The output stage has therefore the frequency response as follows:

$$A[dB] = 10\log[1 + (\frac{\omega_i}{\omega_p})^2]$$

Where ω*i* is a given frequency and ω*p* is the cutoff frequency, defined by low-pas filter constant (1):

$$\omega_p = \frac{1}{R_1 C_1} = \frac{1}{30x\Omega \cdot 2 \cdot 27pF} = 617r/\sec$$

which is close to 100Hz (Fig. 5 and Fig. 6).

All other poles come from level-shift circuit, ratiometric ground, etc, are at much higher frequencies.

Conclusion

The *overall* transfer function which include the frequency behavior, the **VDD** influence, the temperature and the elements matching influence described in the equations (2), (3), (4), (5), (6), ((7), (8)) can be verified by replacing appropriate variables in equation (1). The circuit has been simulated using HSPICE and transfer function verified in MatLab. Results are shown in Fig. 3, Fig. 4, and Fig. 5 -to- Fig. 9. On Fig. 5 is *overall* transfer function (1) included *Rgr(f)*.

References:

/1/ Arthur B. Williams, "ELECTRONIC FILTER DESIGN HANDBOOK", McGraw-Hill Book Company, 1981

Anton Pleteršek
Faculty of Electrical Engineering, Tržaška 25,
1000 Ljubljana, Slovenia
e-mail:anton@kalvarija.fe.uni-lj.si

Prispelo (Arrived): 06.06.2002 *Sprejeto (Accepted): 25.03.2003*

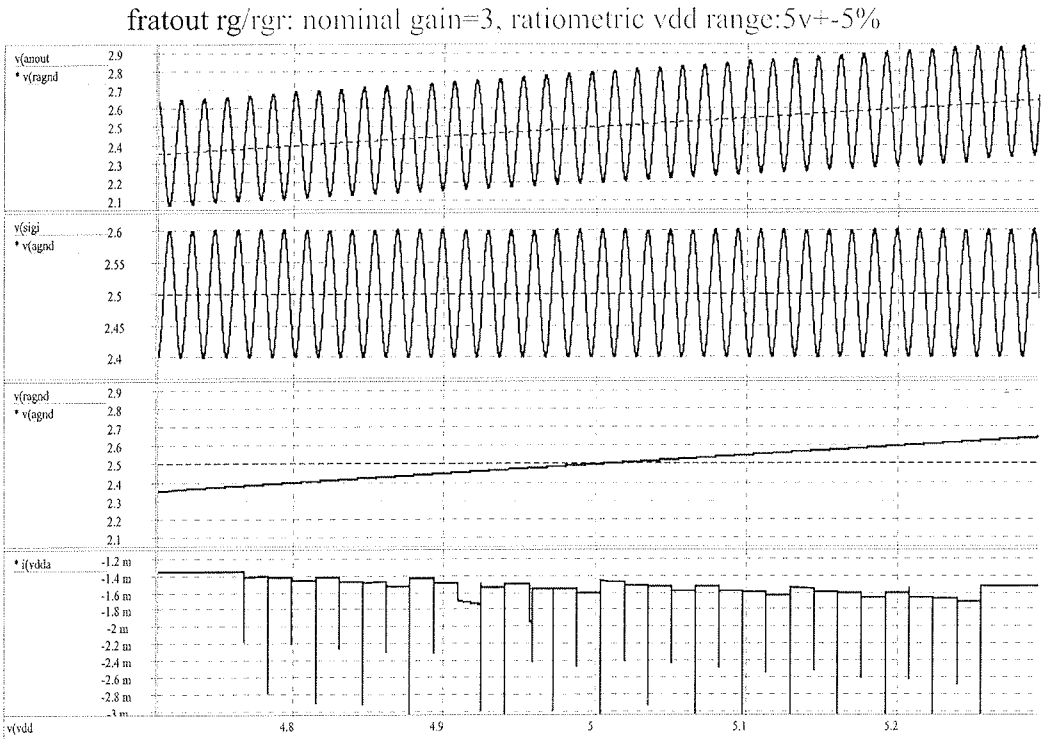


Fig. 3: Ratiometric range is limited to 5V (+-5%). Output DC level and output AC signal amplitude (ANOUT) follow the supply voltage VDD.

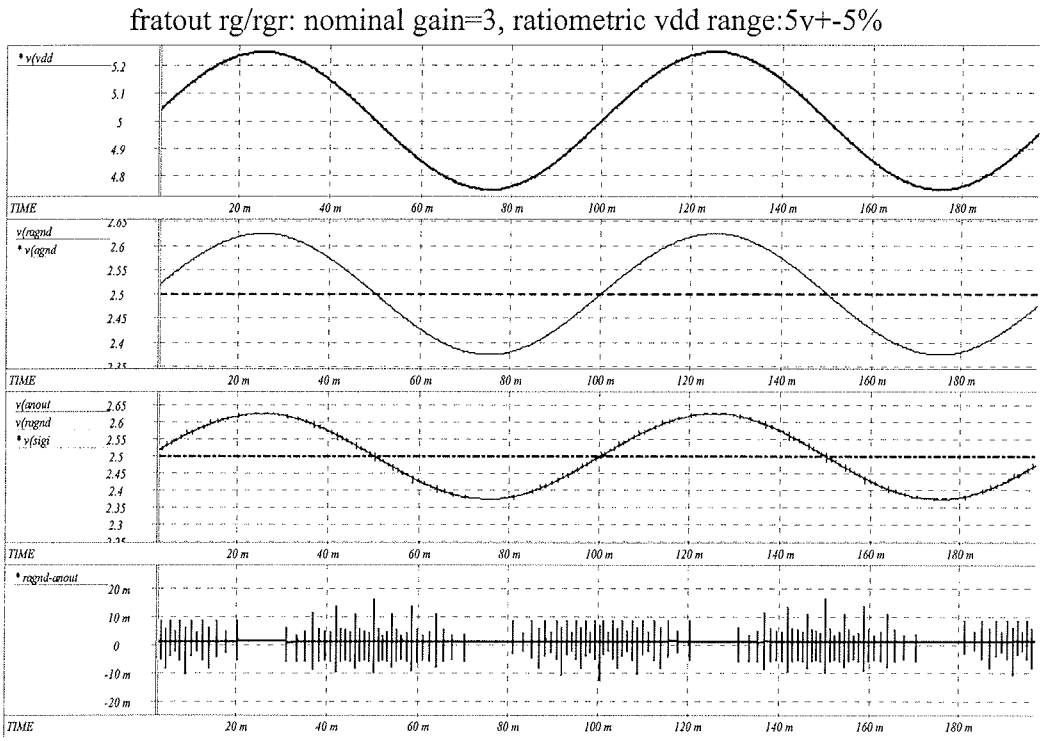


Fig. 4: Ratiometric signal ground (RAGND) follow the supply voltage without significant delay (VDD changes with 10Hz). Output at zero input signal (sigi) follows the ratiometric signal ground, having the same amplitude and frequency response. The low-pass filters have Butterworth filter characteristic, having no overshot and does not have any oscillatory behavior.

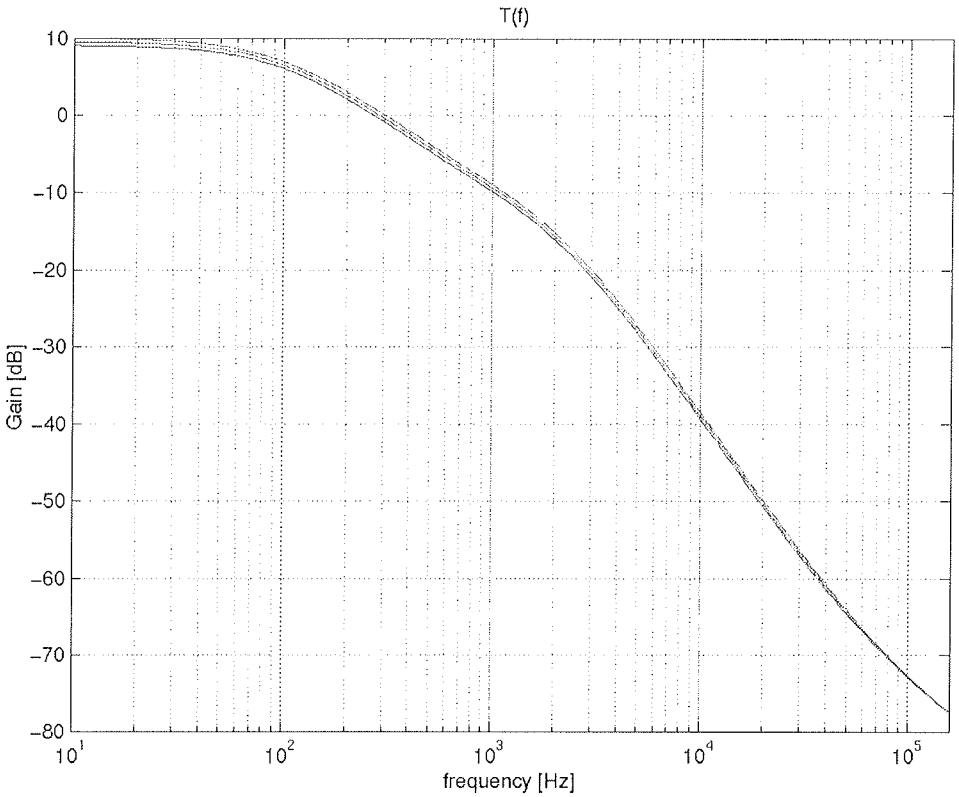


Fig. 5: Gain of "fratout" at nominal setting (g=3) for three different supply voltages (VDD=5V+-5%). Frequency response includes all blocks as is shown from Fig. 2. All equations are combined into equation (1).

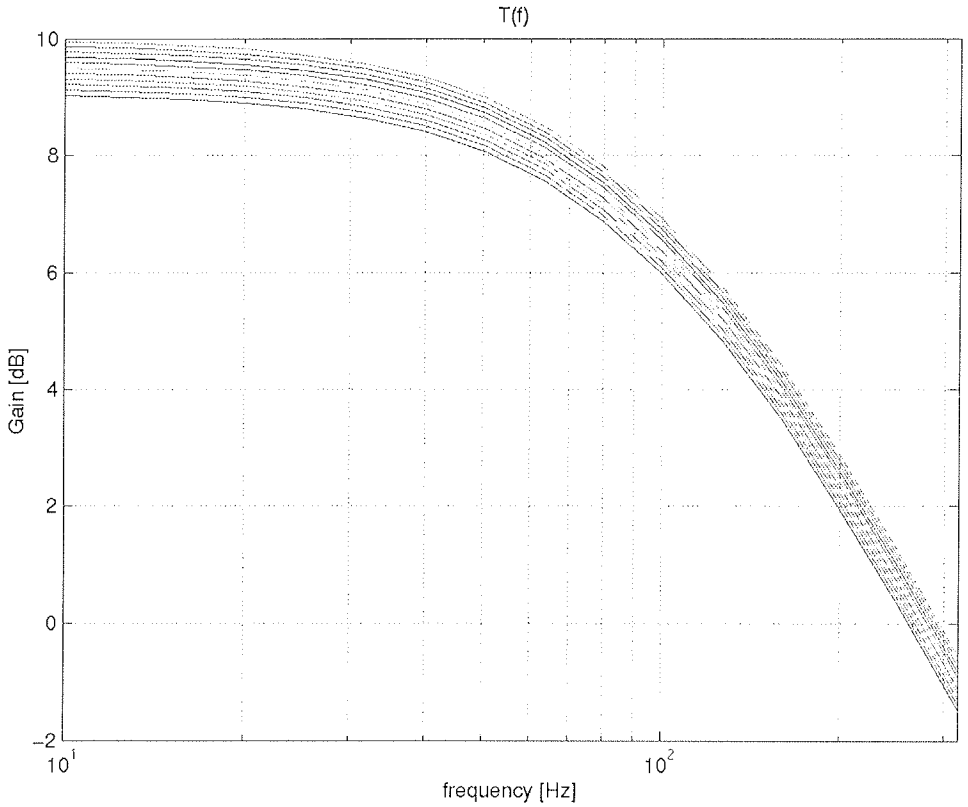


Fig. 6: The "fratout" transfer function with supply voltage VDD from 4.75V to 5.25V.

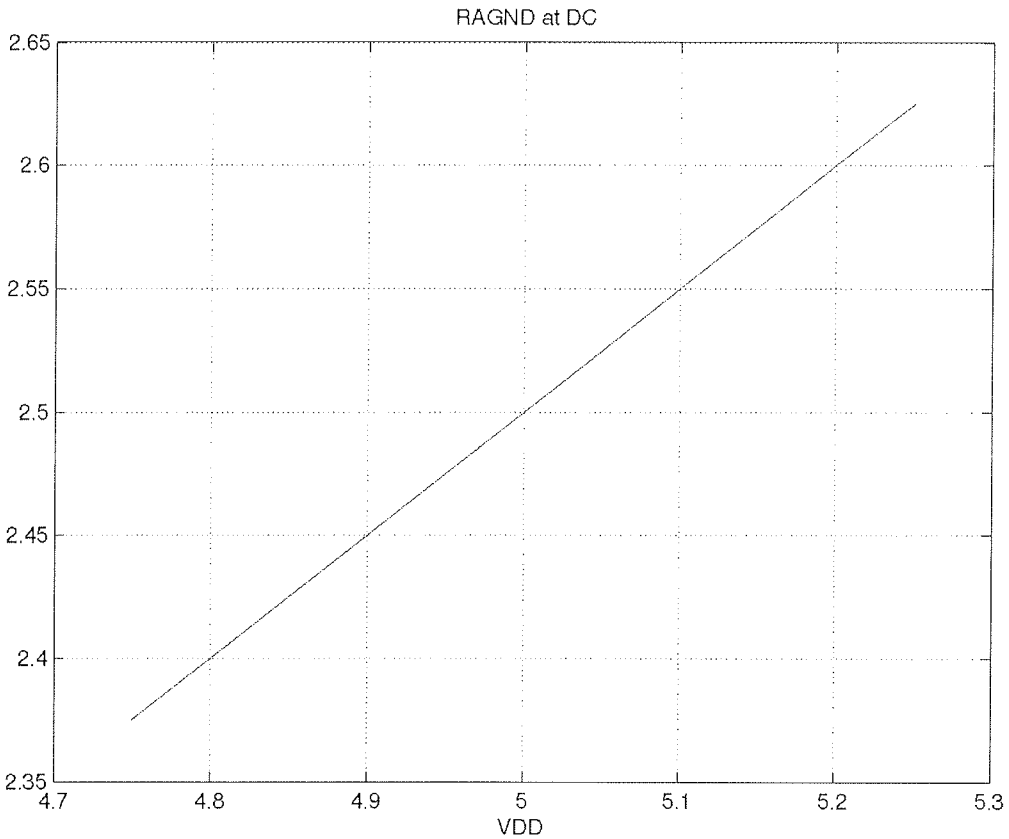


Fig. 7: RAGND variations with VDD. Transfer function is linear.

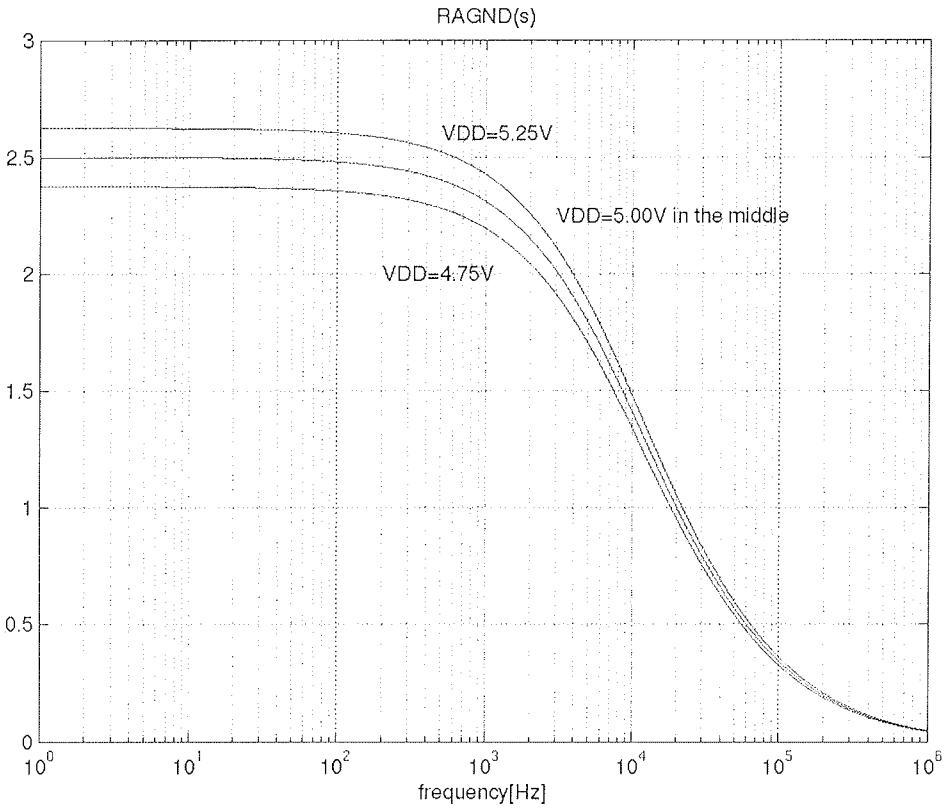


Fig. 8: RAGND frequency response (8) for minimal, nominal and maximal supply voltage.

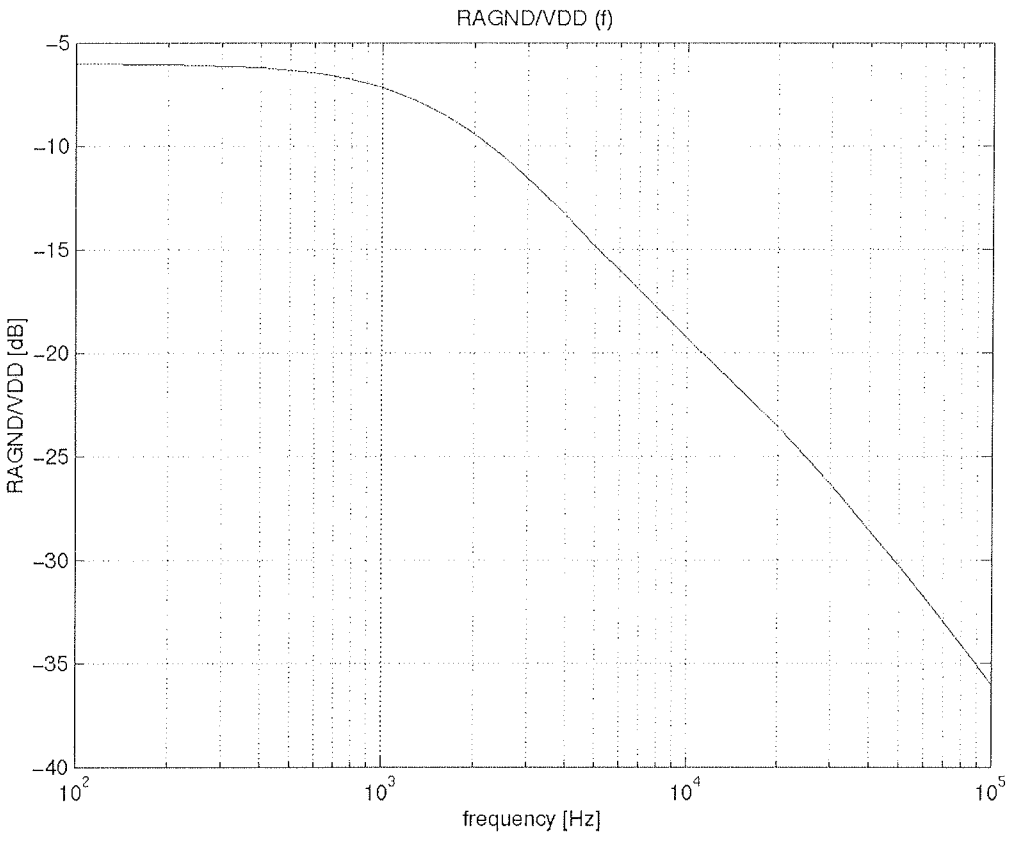


Fig. 9: RAGND/VDD transfer function.

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 domačih in tujih avtorjev s področja mikroelektronike, elektronskih sestavnih delov in materialov, ki so lahko:
izvirni znanstveni članki, pregledni znanstveni članki, predhodne objave, strokovni članki ter predavanja in povzetki s strokovnih posvetovanj.
Strokovni prispevki bodo recenzirani.

Revija objavlja tudi aplikacijske članke, poljudne članke, novi-ce iz stroke, vesti iz delovnih organizacij, inštitutov in fakultet, obvestila o akcijah društva MIDEM in njegovih članov ter druge prispevke.

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 izvleček (največ 250 besed).
3. Naslov dela v angleščini.
4. Ključne besede v angleščini (key words) in podaljšani povzetek (Extended Abstract) v angleščini.
5. Uvod, glavni del, zaključek, zahvale, dodatki in literatura v skladu z IMRAD shemo (Introduction, Methods, Results and Discussion).
6. Polna imena in priimki avtorjev s titulami, naslovi institucij in firm, v katerih so zaposleni ter Tel./Fax/Email podatki.

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 objavljeno v slovenščini ali v angleščini.
4. Uredniški odbor ne bo sprejel strokovnih prispevkov, ki ne bodo poslani v dveh izvodih.
5. Avtorji, ki pripravljajo besedilo v urejevalnikih besedil lahko pošljejo zapis datoteke na disketi (3.5"/1.44 MB) v formatih ASCII ali Word for Windows, ker bo besedilo oblikovano v programu Adobe PageMaker 6,5. Grafične datoteke so lahko v formatu TIFF, EPS, VMF, GIF ali JPEG.

Avtorji so v celoti odgovorni za vsebino objavljenega sestavka. Rokopisov ne vračamo.

Rokopise pošljite na naslov:

Uredništvo Informacije MIDEM
MIDEM pri MIKROIKS
Stegne 11, 1521 Ljubljana
Slovenija
Email: Iztok.Sorli@guest.arnes.si
Tel. 01 513 37 68, fax. 01 513 37 71

Informacije MIDEM
Journal of Microelectronics, Electronic Components and Materials

INFORMATION FOR CONTRIBUTORS

Informacije MIDEM is a professional-scientific-social publication of Professional Society for Microelectronics, Electronic Components and Materials - MIDEM. In the Journal contributions of domestic and foreign authors are published covering the field of microelectronics, electronic components and meterials. These contributions may be:
original scientific papers, review scientific papers, preleminary communications, professional papers, conference papers and abstracts.
All professional contributions are subject to reviews.

Applications articles, scientific news, news from the companies, institutes and universities, reports on actions of MIDEM Society and its members as well as other relevant contributions are also welcome.

Each professional contribution should include the folowing specific components:

1. Title of paper, authors names, name of the institution/ company.
2. Key Words and Abstract (not more than 250 words).
3. Introduction, maintext, conclusion, acknowledgements, appendix and references following the IMRAD scheme (Introduction, Methods, Results and Discussion).
4. Full authors' names, titles and complete company or institution address including Tel./Fax/E-mail.

COMMENT: Slovenian authors who write in English language must submit title, abstract and key words also in Slovene language.

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 are not to be placed into 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 language.
4. Papers will not be accepted unless two copies are recived.
5. Authors may send their files on formated diskettes (3.5"/ 1.44 mb/) in ASCII or Word for Windows format as text will be formatted in Adobe PageMaker 6,5. Graphic files may be in TIFF, EPS, VMF, GIF or JPEG formats.

Authors are fully responsible for the content of the paper. Manuscripts are not refunded.

Contributions are to be sent to the address:

Urednitvo Informacije MIDEM
MIDEM at MIKROIKS
Stegne 11, 1521 Ljubljana
Slovenia
Email: Iztok.Sorli@guest.arnes.si
Tel. +386 1 513 37 68, fax. +386 1 513 37 71