POLSKA AKADEMIA NAUK
KOMITET ELEKTRONIKI I TELEKOMUNIKACJI

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI

ELECTRONICS AND TELECOMMUNICATIONS QUARTERLY

TOM 44 — ZESZYT 3

WYDAWNICTWO NAUKOWE PWN
WARSZAWA 1998

RADA REDAKCYJNA

Przewodniczący

prof. dr hab. inż. ALFRED ŚWIT
członek rzeczywisty PAN

Członkowie

prof. dr hab. inż. DANIEL JÓZEF BEM — czł. koresp. PAN, prof. dr hab. inż. MICHAŁ BIAŁKO — czł. rzecz. PAN, prof. dr hab. inż. STEFAN HAHN — czł. koresp. PAN, prof. dr inż. ANDRZEJ HAŁAS, prof. dr inż. ZDZISŁAW KACHLICKI, prof. dr hab. inż. WŁADYSŁAW MAJEWSKI, prof. dr hab. inż. JÓZEF MODELSKI, prof. dr inż. JERZY OSIOWSKI, prof. dr inż. WITOLD ROSIŃSKI — czł. rzecz. PAN, prof. dr hab. inż. STEFAN WĘGRZYN — czł. rzecz. PAN, prof. dr hab. inż. WIESŁAW WOLIŃSKI — czł. koresp. PAN, prof. dr inż. ANDRZEJ ZIELIŃSKI, prof. dr inż. MARIAN ZIENTALSKI

REDAKCJA

Redaktor Naczelny

prof. dr hab. inż. WIESŁAW WOLIŃSKI

Zastępca Redaktora Naczelnego

doc. dr inż. KRYSTYN PLEWKO

Sekretarz Odpowiedzialny

mgr KRYSTYNA LELAKOWSKA

ADRES REDAKCJI

00-665 Warszawa, ul. Nowowiejska 15/19 Politechnika, pok. 470
Instytut Telekomunikacji, Gmach im. prof. JANUSZA GROSZKOWSKIEGO

Dyżury Redakcji: środy i piątki, godz. 14—16
tel/fax (0-22) 660 77 37, 825 29 18+aut. sekr.

Telefony domowe: Redaktora Naczelnego: 12 17 65
Zastępcy Red. Naczelnego: 826 83 41
Sekretarza Odpowiedzialnego: 825 29 18

WYDAWNICTWO NAUKOWE PWN
Warszawa, ul. Miodowa 10

Ark. wyd. 11,5 Ark. druk. 9,0	Podpisano do druku w grudniu 1998 r.
Papier offsetowy kl. III 70 g. B-1	Druk ukończono w grudniu 1998 r.

Skład: Amel, Warszawa
Druk i oprawa: Drukarnia Braci Grodzickich, Piaseczno, ul. Geodetów 47a

Szanowni Autorzy

„Kwartalnik Elektroniki i Telekomunikacji" – Electronics and Telecommunications Quarterly jest kontynuatorem tradycji powstałego 44 lata temu kwartalnika p.t. „Rozprawy Elektrotechniczne".

Kwartalnik jest czasopismem Komitetu Elektroniki i Telekomunikacji Polskiej Akademii Nauk. Wydawany jest przez Wydawnictwo Naukowe PWN SA w Warszawie. Kwartalnik jest czasopismem naukowym, na którego łamach są publikowane artykuły i komunikaty prezentujące wyniki oryginalnych prac teoretycznych i doświadczalnych, a także przeglądowych. Związane są one z szeroko rozumianymi dziedzinami współczesnej elektroniki, telekomunikacji, mikroelektroniki, optoelektroniki, radiotechniki i elektroniki medycznej.

Autorami publikacji są wybitni naukowcy, znani specjaliści o wieloletnim doświadczeniu, a także młodzi badacze — głównie doktoranci.

Artykuły charakteryzują się oryginalnym ujęciem zagadnienia, interesującymi wynikami badań, krytyczną oceną teorii lub metod, omówieniem aktualnego stanu, lub postępu danej gałęzi techniki oraz omówieniem perspektyw rozwojowych. Sposób pisania matematycznej części artykułów zgodny jest z wytycznymi IEE (International Electronics Commision) oraz ISO (International Organization of Standarization).

Wszystkie publikowane w Kwartalniku artykuły są recenzowane przez znanych krajowych specjalistów, co zapewnia że publikacje te są uznawane jako autorski dorobek naukowy. Opublikowanie w kwartalniku wyników prac naukowych zrealizowanych w ramach „GRANTów" Komitetu Badań Naukowych spełnia więc jeden z wymogów stawianych tym pracom.

Czasopismo dociera do wszystkich zajmujących się elektroniką i telekomunikacją krajowych ośrodków naukowych oraz technicznych a także szeregu instytucji zagranicznych. Jest ponadto prenumerowane przez liczne grono specjalistów i biblioteki.

Każdy Autor otrzymuje bezpłatnie 20 egzemplarzy nadbitek swojego artykułu, co ułatwia przesłanie go do indywidualnie wybranych przez Autora osób i instytucji w kraju, lub za granicą. Ułatwia to dodatkowo fakt, że w Kwartalniku są publikowane artykuły w języku angielskim.

Nadesłane do redakcji artykuły są publikowane w terminie około pół roku, w przypadku sprawnej współpracy Autora z redakcją. Wytyczne dla Autorów dotyczące formy publikacji są zamieszczone w zeszytach Kwartalnika. Można je także otrzymać w siedzibie redakcji.

Artykuły można dostarczać osobiście, lub pocztą pod adresem zamieszczanym na stronie redakcyjnej w każdym zeszycie.

Redakcja

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI, 1998, **44**, z. 3, ss. 307–316

Optymalizacja kompresji informacji w testowaniu układów ze ścieżką brzegową

ANDRZEJ MITAS

Zakład Elektrotechnologii i Elektrotermii Wydział Inżynierii Materiałowej, Metalurgii i Transportu Politechniki Śląskiej

Otrzymano 1998.05.21

Autoryzowano 1998.08.21

W artykule przedstawiono cechy probabilistyczne kompresji informacji wynikające ze stosowania technik nieliniowych i mieszanych. We wstępnej części przypomniano definicje różnych technik kompresji. Zwrócono przede wszystkim uwagę na te techniki, które stosuje się do przetwarzania informacji szeregowej. Techniki te mają interesujące cechy probabilistyczne w sensie małego prawdopodobieństwa maskowania błędów.

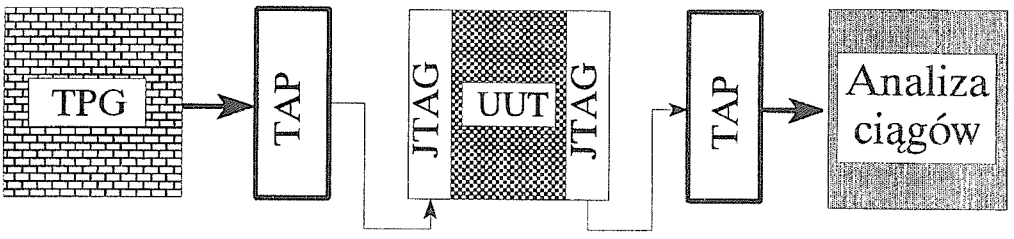
Techniki nieliniowe są mało popularne z powodu ograniczeń w zastosowaniu do kompresji równoległej. Do równoległego przetwarzania wyników, szczególnie ważnego w samotestowaniu wbudowanym nadaje się technika liniowa (równoległy rejestr MISR). Techniki szeregowe można stosować bez ograniczeń w testowaniu układów projektowanych zgodnie z koncepcją JTAG — wyposażonych w brzegową ścieżkę sterująco-obserwacyjną.

W artykule przypomniano autorskie porównanie różnych technik kompresji, w szczególności na gruncie teoretycznego prawdopodobieństwa maskowania. Przedstawiono zależności opisujące prawdopodobieństwo maskowania dla połączeń techniki zliczania jedynek, zliczania przejść i techniki liniowej. Z komentarza do tych wzorów wynika, iż techniki mieszane cechują się „przedziałami interesującymi" cechami probabilistycznymi. Takie techniki kompresji można zastosować w testowaniu układów, w których wprowadzanie pobudzeń testowych odbywa się szeregowo. W artykule nawiązano więc do takich układów (w szczególności z modułami według standardu JTAG), proponując techniki nieliniowe do kompresji wyników testu. Przedstawiono dwie koncepcje rozwiązania problemu kompresji — pierwszą z redundancją sprzętu i drugą z kilkukrotnym powtarzaniem testu, przy każdorazowej rekonfiguracji układu kompresji. W każdym przypadku proponuje się stałą strukturę rejestru liniowego, stosowanego najpierw jako układ kompresji LFSR, a następnie jako rejestr liczący. Przedstawione schematy są punktem wyjścia do programowych symulacji, stanowiących przedmiot dalszych badań.

Słowa kluczowe: BSP, JTAG, kompresja informacji, maskowanie, techniki diagnostyczne, testowanie układów cyfrowych.

WSTĘP

Dynamiczny wzrost stopnia złożoności układów cyfrowych wymusza również stały postęp w dziedzinie testowania. Ze względu na ogromne ilości informacji wynikowej w testowaniu bardzo istotne znaczenie mają techniki kompresji informacji. Powszechnie uznaną i stosowaną techniką jest technika liniowa kompresji informacji ([1], [2], [3]...). Zasadniczą zaletę tej techniki stanowi niezmienne dla wszystkich ciągów prawdopodobieństwo maskowania błędu, również dla rejestrów równoległych MISR, związane z liczbą stopni rejestru n funkcją $P = 2^{-n}$ (w przybliżeniu). Inne znane techniki to techniki zliczeniowe ([4], [5], [6], [7], [8]...), spośród których do najbardziej popularnych należy technika zliczania przejść (TC — *transition counting*) oraz technika syndromowa (SC — *syndrome counting*). W dalszych wywodach stosuje się następujące oznaczenia: m — długość ciągu, t — liczba jedynek w ciągu, k — liczba przejść, czyli zmian wartości w ciągu. Technikę zliczania jedynek w ciągu oznaczono symbolem SC, ponieważ nawiązuje to do techniki syndromowej. Główną zaletą tych technik jest nadzwyczaj mała wartość teoretycznego prawdopodobieństwa maskowania błędów dla ciągów poprawnych, których cechy zliczeniowe istotnie oddalone są od swoich median ([7]). Podstawową wadą (silnie warunkującą niestosowalność w testowaniu wbudowanym) jest natomiast brak możliwości równoległej obserwacji i oceny wyników.



Rys. 1. Schemat blokowy stanowiska diagnostycznego

Techniki szeregowej kompresji informacji mają jednakże szanse implementacji w układach z brzegową ścieżką sterująco-obserwacyjną, szczególnie w kontekście rozpowszechnienia standardu JTAG 1149 ([9], [10]). Przykładowy schemat (z ograniczeniem do jednego układu testowanego) struktury stanowiska diagnostycznego z implementacją koncepcji BSP przedstawia rys. 1 (symbol TPG oznacza generator pobudzeń testowych, UUT — układ testowany, a pośredniczący element TAP to port testowy). Pobudzony układ generuje w trakcie testu odpowiedzi, które następnie są w rejestrze wyjściowym przetwarzane do postaci szeregowej i mogą być poddane dalszej analizie w układzie oceny wyników. Istnieje możliwość nieznacznego wzrostu stopnia złożoności układu weryfikatora w celu istotnego podwyższenia skuteczności (czyli redukcji wartości prawdopodobieństwa maskowania), przy czym istotne jest to, że ten sam weryfikator w koncepcji ścieżki brzegowej może być użyty dla wielu układów testowanych. Celowość działań zmierzających do zmniejszenia

ryzyka błędnej oceny sprawności układu na podstawie porównania jedynie wyniku kompresji z wzorcem jest uzasadniona współczesnym rozwojem jakościowym. W szczególnych przypadkach zastosowań niewykrycie uszkodzenia z powodu realnego maskowania błędu może wiązać się z kosztami, jak i z negatywnym kształtowaniem opinii użytkowników.

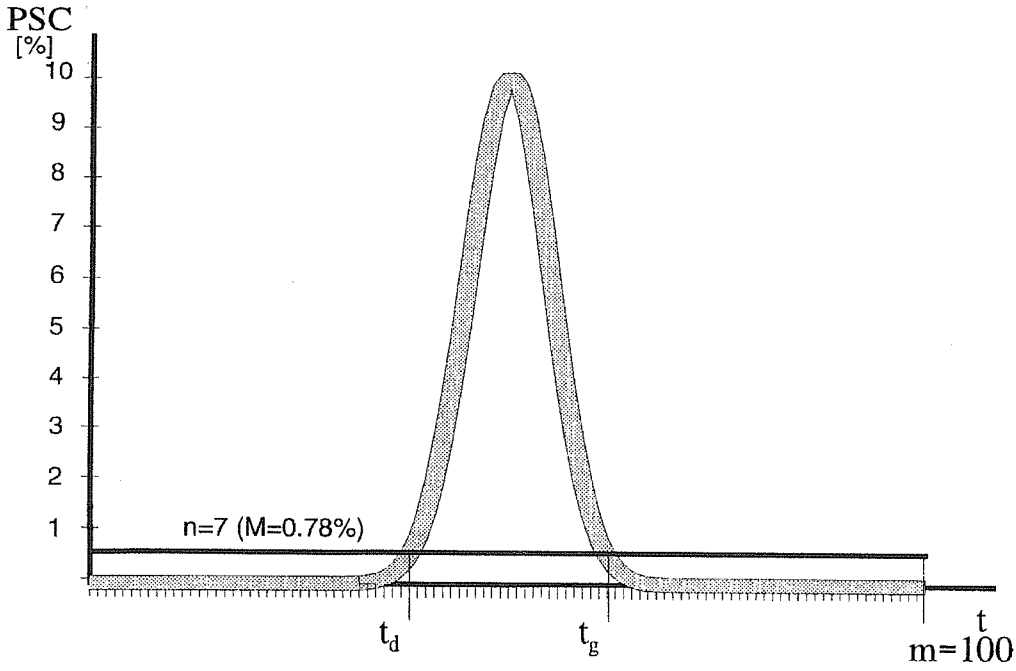
1. MULTIKOMPRESJA — PODSTAWY TEORETYCZNE, MASKOWANIE

Projektowanie procesu diagnostycznego obejmuje dwie fazy:
1) dobór generatora testu,
2) optymalizację etapu weryfikacji odpowiedzi.

W dalszej części skoncentrowano się na drugim zagadnieniu, czyli próbie doboru takiej techniki weryfikacji reakcji układu badanego, by najmniejszym kosztem dokonywać oceny, możliwie najdokładniej. Techniki kompresji informacji wiążą się nieuchronnie z ryzykiem maskowania. Jednym ze sposobów minimalizacji maskowania jest zastosowanie wyselekcjonowanej techniki kompresji lub użycie kilku technik kompresji informacji. Wielokrotna kompresja liniowa odpowiada, jak wiadomo (np. [7]), kompresji liniowej za pomocą rejestru o sumarycznej długości (tzn. równej sumie długości rejestrów składowych). Łączenie takiej samej techniki w różnych odmianach jest uzasadnione skierowaniem właściwości detekcyjnych na przewidywane bądź też dobrze rozpoznane katalogi błędów. Częściej tego rodzaju analizę doboru postaci techniki liniowej (tzn. opisu sprzężenia zwrotnego rejestru) prowadzi się więc w obszarze transmisji danych, gdzie przewidywalność zachowania zakłócającego środowiska jest większa niż w przypadku wtórnego (w wyniku uszkodzenia często po prostu nieznanej struktury) „zakłócania" ciągu danych.

Technika liniowa (AS — *signature analysis*) charakteryzuje się stałą wartością prawdopodobieństwa maskowania w całkowitej przestrzeni ciągów. Pojedynczo stosowane techniki zliczeniowe SC i TC charakteryzują się bardzo małym maskowaniem jeśli odpowiednie cechy zliczeniowe mają wartości oddalone od odpowiednich średnich ($\cong m/2$). Przykładowy przebieg funkcji maskowania jest silnie nieliniowy — rys. 2. Dla wartości cech (np. liczba jedynek) oddalonych od swoich median dostatecznie mocno, tzn. poniżej t_d lub powyżej t_g, maskowanie dla techniki zliczeniowej jest mniejsze niż dla techniki liniowej o „ekwiwalentnej" długości rejestru. Ekwiwalentna długość oznacza taką liczbę stopni (przerzutników), jaka jest niezbędna do zliczenia maksymalnej liczby zdarzeń obserwowanych w badanym ciągu. Jak wynika z dokładnej analizy ([7]) granice, np. dla zliczania jedynek $-t_d/t_g$ dążą szybko do mediany wraz ze wzrostem długości ciągu m (osiągają wartości $\cong m/2$), a z powodu silnej nieliniowości maskowanie szybko maleje wraz z oddalaniem się cech zliczeniowych od swoich wartości granicznych.

Łączne zastosowanie obu tych technik zliczeniowych (multikompresja S&TC — zliczanie jedynek i przejść) stwarza dodatkowe możliwości ograniczenia maskowania. W szczególnych przypadkach może wystąpić sytuacja, że ciąg o średniej liczbie jedynek może mieć niewielką liczbę przejść (lub odwrotnie). Wyznaczenie



Rys. 2. Przykładowy przebieg funkcji maskowania dla techniki zliczania jedynek

teoretycznego maskowania dla techniki równoczesnej kompresji SC i TC w odniesieniu do całkowitej przestrzeni możliwych ciągów (obejmującej 2^m wszystkich m-bitowych ciągów przy założeniu równego prawdopodobieństwa wystąpienia dowolnego ciągu) polega na wyliczeniu w oparciu o rachunek kombinatoryczny liczby ciągów o długości m, w których przy t jedynkach występuje k przejść. Maskowanie dla techniki S&TC określa następujący wzór [7]:

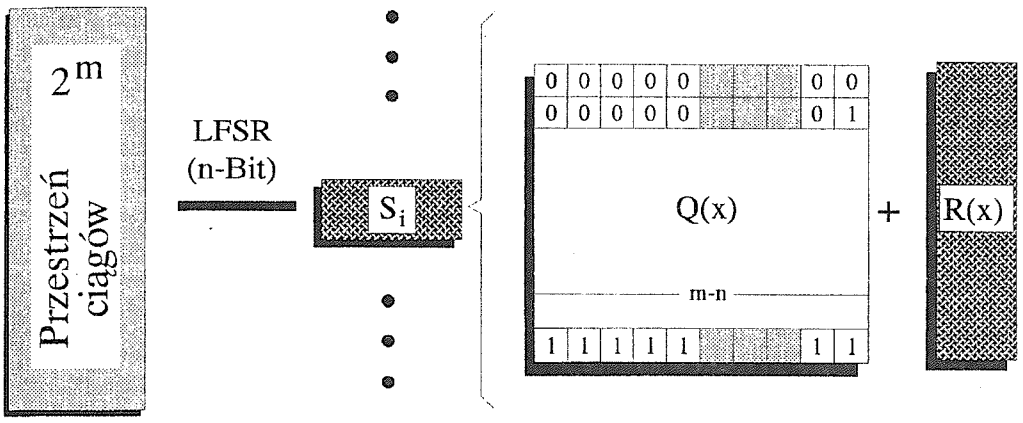
$$P_{S\&TC} = \frac{2\binom{t-1}{t-\left\lfloor\frac{k+1}{2}\right\rfloor}\binom{m-t-1}{m-t-\left\lfloor\frac{k+1}{2}\right\rfloor}\left(\frac{m-k}{k}\right)^{SGN\left(\frac{k+1}{2}-\left\lfloor\frac{k+1}{2}\right\rfloor\right)}}{2^m-1} + \frac{\binom{0}{t}\binom{0}{k}+\binom{0}{m-t}\binom{0}{k}-1}{2^m-1}.$$

Łącznie wystąpienie dwóch symboli Newtona w liczniku powyższego wyrażenia implikuje silne ograniczenie wartości maskowania dla cech zliczeniowych odbiegających od swoich wartości średnich.

2. KONCEPCJA NIELINIOWEJ TECHNIKI SYMULTANICZNEJ — S&TCAS

Podstawą teoretycznych rozważań o technikach kompresji jest zazwyczaj uogólnione maskowanie. Maskowanie to dla techniki liniowej jest stałe w przestrzeni ciągów. Technika S&TC ma bardzo silną nieliniowość, ale dla średnich wartości cech zliczeniowych jest niekorzystna. Technika liniowa ma wówczas lepsze właściwości probabilistyczne. Połączenie technik różniących się rozkładem prawdopodobieństwa maskowania pozwala przedziałami minimalizować (znacznie) ryzyko błędnej oceny.

W celu wyznaczenia maskowania dla techniki symultanicznej S&TCAS (złożenie analizy sygnatury oraz zliczania przejść i jedynek) załóżmy, że ciąg wynikowy (z rejestru ścieżki) poddawany jest wstępnej kompresji zliczeniowej. Następnie wielomianowy iloraz (wielomianowej postaci ciągu wejściowego przez wielomian charakterystyczny rejestru liniowego) podlega współbieżnej kompresji zliczeniowej. W rezultacie dzielenia wielomianu reprezentującego ciąg wejściowy przez wielomian sprzężenia zwrotnego otrzymuje się iloraz $Q(x)$ oraz resztę $R(x)$ — rys. 3.



Rys. 3. Dzielenie wielominaów przez rejestr liniowy — analiza sygnatury

Maskowanie (zgodność sygnatury S_i z wzorcem) wystąpi dla każdego możliwego ciągu o długości $m - n$ (z wyjątkiem ciągu poprawnego). W rezultacie, do badania technikami zliczeniowymi teoretyczną bazę stanowi przestrzeń wszystkich ciągów o długości $m - n$. Taki sposób analizy zawiera wprawdzie pewne „utrudnienie", zapewnia jednak niezależność zdarzeń (dana sygnatura i określone cechy zliczeniowe) w sensie probabilistycznym. (Założenie separacji niewielkiej, n-bitowej, części ciągu można także bez trudu realizować sprzętowo). W oparciu o podstawy rachunku prawdopodobieństwa można wyznaczyć maskowanie dla S&TCAS jako:

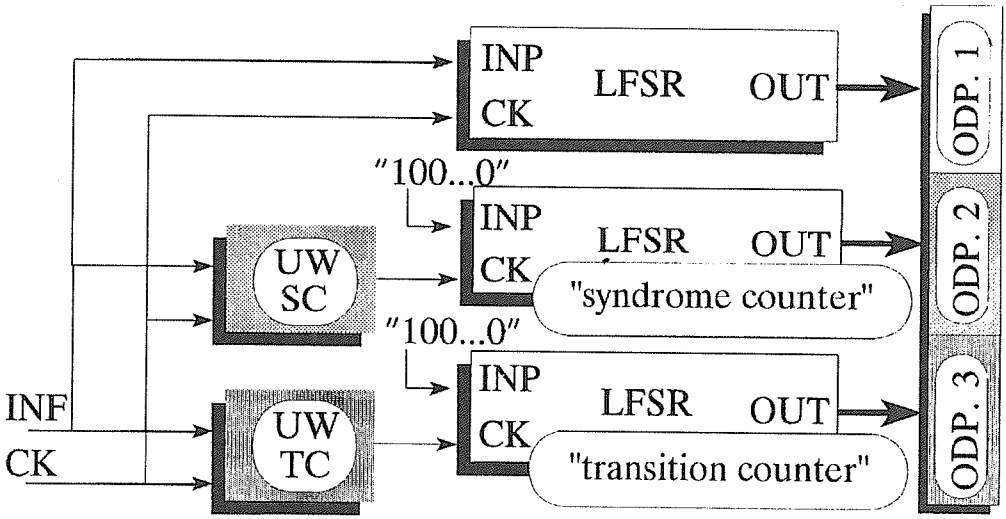
$$P_{\text{S\&TCAS}} = P_{\text{S\&TC}}(m - n) \cdot P_{\text{AS}}(n)$$

gdzie $P_{\text{AS}}(n)$ oznacza maskowanie dla techniki liniowej, przy zastosowaniu rejestru o n stopniach.

Włączając do powyższego równania poprzednie wyrażenie, opisujące prawdopodobieństwo maskowania dla multikompresji nieliniowej (po zastąpieniu parametru długości ciągu m „nową" wartością $m-n$) i po pominięciu przypadków skrajnych wartości cech zliczeniowych ($t=0$, $t=m-n$, $k=0$, $k=m-n-t$), otrzymujemy następujący wzór do probabilistycznego opisu maskowania multikompresji S&TCAS:

$$P_{S\&TCAS} = \frac{2\binom{t-1}{t-\left\lfloor\frac{k+1}{2}\right\rfloor}\binom{m-n-t-1}{m-n-t-\left\lfloor\frac{k+1}{2}\right\rfloor}\left(\frac{m-n-k}{k}\right)^{SGN\left(\frac{k+1}{2}-\left\lfloor\frac{k+1}{2}\right\rfloor\right)}}{2^m-1}.$$

Kompresja wielokrotna może być zrealizowana za pomocą odpowiednich układów, sprzętowo dokonujących wyodrębnienia poszukiwanych cech — rys. 4.

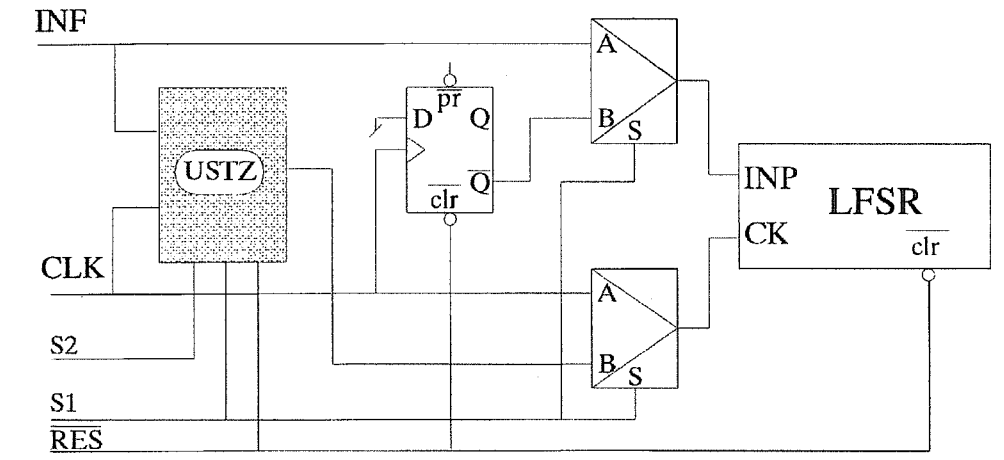


Rys. 4. Kompresja wielokrotna w układzie potrójnym

Do realizacji kompresji liniowej używa się rejestru przesuwającego o liniowym sprzężeniu zwrotnym, czyli LFSR. Wprowadzana szeregowo informacja spowoduje ustalenie na koniec „okna pomiarowego" pewnego stanu rejestru — sygnatury, porównywanej w dalszej kolejności z wzorcem. Do kompresji zliczeniowej należy użyć licznika binarnego, zliczającego „w przód" w chwili występowania określonego zdarzenia w ciągu (np. jedynka na bieżącej pozycji — SC lub zmiana stanu binarnego w ciągu — TC). Rejestr liniowy o sprzężeniu opisanym wielomianem pierwotnym ma cykl pracy o długości $L = 2^n - 1$, gdzie n oznacza długość rejestru. Licznik binarny ma długość cyklu prawie taką samą, więc można do zliczania zdarzeń użyć właśnie rejestru liniowego, funkcjonującego jako rejestr liczący. Zada-

niem użytkownika jest wówczas doprowadzenie do wejścia informacyjnego rejestru takiego ciągu, który jest stosowany do określania długości cyklu (tzn. 100...0), a także odpowiednie selekcjonowanie taktów zegarowych (inicjujących zmianę stanu rejestru jedynie w chwili występowania poszukiwanych zdarzeń binarnych). W efekcie po wykonaniu testu rejestr jest w stanie znamiennym dla określonej liczby jedynek — dla techniki SC lub przejść — dla techniki TC, ponieważ liczby te implikują określoną liczbę taktów zegarowych dla tego rejestru, powodując zatrzymanie się w określonym stanie grafu przejść rejestru.

Z binarną wartością liczby zdarzeń stan rejestru wiąże się iteracyjnym równaniem, opisującym LFSR. Użytkownika interesuje jednakże nie tyle faktyczna postać wyniku, co zgodność z wynikiem wzorcowym, na podstawie czego podejmowana jest decyzja o sprawności badanego układu. Zaletą przedstawionej na rysunku 4 struktury jest jej jednolitość, istotna w realizacji praktycznej. Takie sprzętowe rozwiązanie układu do realizacji multikompresji można zmodyfikować do postaci przedstawionej na rys. 5, odnosząc proponowaną koncepcję multikompresji do testowania układów z brzegową ścieżką obserwacyjną JTAG. W efekcie zmian zwielokrotnienie sprzętowe zastępowane jest redundancją czasu. Badany układ jest mianowicie stymulowany trzykrotnie tą samą sekwencją pobudzeń testowych, a układ kompresji w każdym z obiegów testu pełni inną rolę: analizatora sygnatury, licznika jedynek i licznika przejść.



S1 = 0		$CK_i = CLK_i$	$INP_i = INF_i$
S1 = 1	S2 = 0	$CK_i = CLK_i \wedge INF_i$	$INP_0 = 1$ $INP_i = 0$ dla $i > 0$
USTZ	S2 = 1	$CK_i = CLK_i \wedge (INF_i \vee INF_{i-1})$	$INP_0 = 1$ $INP_i = 0$ dla $i > 0$

Rys. 5. Wielokrotne zastosowanie programowalnego układu kompresji

WNIOSKI

Nieliniowa kompresja wielokrotna może być stosowana w układach z brzegową ścieżką obserwacyjno-sterującą, ponieważ ocenie podlegają wówczas dane przesyłane szeregowo. Zastosowanie technik zliczeniowych ma istotny wpływ redukcyjny na teoretyczne maskowanie, zwłaszcza w przypadku wartości cech zliczeniowych ciągu poprawnego znacznie różniących się od swoich wartości średnich. Technika S&TCAS może być realizowana w systemie diagnostycznym poprzez zwielokrotnienie struktury układu kompresji. W przypadku konieczności minimalizacji nakładów sprzętowych można zastosować ten sam rejestr, przeprogramowując go z funkcji rzeczywistego rejestru liniowego do postaci rejestru liczącego. Występuje wówczas naturalna redundancja czasu testowania.

Taka potrójna technika niesie ze sobą nie tylko statystycznie uzasadnione, obowiązujące dla całej (abstrakcyjnej) przestrzeni ciągów, obniżenie wartości prawdopodobieństwa maskowania, lecz przede wszystkim stanowi uzupełnienie jakościowo odmiennych koncepcji wyodrębniania z ciągu testowego cech rozpoznawczych. Analiza cech probabilistycznych proponowanej techniki multikompresji stanowi odpowiedni punkt wyjścia dalszych badań porównawczych detekcyjnych właściwości w kontekście przykładowych układów ISCAS, pobudzanych np. testową sekwencją pseudolosową.

BIBLIOGRAFIA

1. A. Hławiczkau: *The survey on signature testing methods.* Proc. of 9th Int. FTSD (Fault-Tolerant Systems and Diagnostics) Conf., Brno, 1986
2. T.W. Williamsiin.: *Aliasing errors in signature analysis registers.* IEEE Design & Test, 1987, str. 39–45
3. H.J. Wunderlich: *Hochintegrierte Schaltungen: Prüfgerechter Entwurf und Test.* Springer-Verlag, Berlin Heidelberg, 1991 FTSD
4. J.P. Hayes: *Transition cout testing of combinational logic circuits.* IEEE Trans. on Computers, 1976 C-25 (6), str. 613–620
5. J.P. Robinson iN.R. Saxena: *A unified view of test compression methods.* IEEE Trans. on Computers 1987 C-36 (1), str. 94–99
6. A. Mitas: *How to increase the effectiveness of digital system testing by application of nonlinear compression techniques.* XII Int. FTSD Conf., Praha 1989
7. A. Mitas: *Zwiększanie skuteczności wykrywania uszkodzeń przez zastosowanie różnych (nieliniowych) technik kompresji informacji.* Prace Instytutu Podstaw Informatyki Polskiej Akademii Nauk, zeszyt 659, Warszawa, 1989
8. J. Savir: *Syndrome-testable design of combinational circuits.* IEEE Trans. on Computers, 1980 C-26 (6), str. 442–451
9. P. Fleming iin.: *Standardmaßiges Testen komplexer Systeme.* Elektronik (12) Niemcy, 1989
10. R.P. van Riessen iin.: *Design and implementing, an architecture with boundary scan.* IEEE Design & Test, luty 1990 str. 9–19

A. MITAS

OPTIMISATION OF INFORMATION COMPRESSION IN TESTING OF CIRCUITS WITH BOUNDARY SCAN PATH

Summary

Some special probabilistic features connected with non-linear and mixed information compression techniques have been presented in the paper. At the beginning the definitions of various compression techniques have been recalled. These techniques are pointed out, which are used for the treatment of serial data streams. The techniques have got some interesting probabilistic features — that is very small probability of error aliasing.

Non-linear techiques are not popular because of some limits in application for parallel compression. For this parallel treatment of results, that is very important in built-in self-test, is usually applied the linear technique (parallel MISR-register). The serial techniques can be used without limits in circuits testing, that are designed according to JTAG-conception (they have got the boundary scan-path).

Comparison of various compression techniques, in particular on the base of theoretical probability of error aliasing, have been recalled from earlier author's publications. In the article there are some formulas, that describe the aliasing probability for the joining of the "ones-counting", "transition-counting" and linear techniques. From the commentary for the formulas we can evaluate, that the mixed techniques have got some "interesting intervals" of probabilistic features. The techniques can be applied for testing of circuits, in which the test stimuli are serial introduced. That is why the paper refers to the circuits of this kind (in particular with JTAG-modules). It is suggested to apply the non-linear techniques for test results compression.

Two conceptions of solving the compression problems have been presented. The first — with hardware-redundancy and the other — connected with multiplied test application, accordingly to the reconfiguration of the test unit. In each case there is proposed the identically structure of linear register. It is used as LFSR-compression unit and than as counting register. The presented schemas are some starting points for further researches of software simulations.

Key words: aliasing, BSP, compression methods, digital testing, JTAG, test techniques.

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI, 1998, **44**, z. 3, ss. 317–324

An algorithm of functional decomposition with free set variables coding

DARIUSZ KANIA

Instytut Elektroniki, Politechnika Śląska, Gliwice

Received 1998.04.20

Authorised 1998.06.08

An algorithm of functional decomposition with free set variables coding is presented. The algorithm, based on Curtis' theory, is an enhanced version of the subdecomposition algorithm presented in [7]. Moreover, the use of a transcoder allows us to limit the number of levels and inputs to the subcircuits used to imlement a given logical function. The described algorithm has been implemented in a procedure of prototype decomposition program. Results of such a decomposition for y0 function of rd84 benchmark is presented.

Key words: decomposition techniques, VLSI, FPGA circuits.

1. INTRODUCTION

With the increasing interest in the synthesis of multilevel logic in VLSI design, there is a renewed interest in decomposition. The main topic of this paper is functional decomposition. Decomposition means breaking a large logic blocks into several relatively smaller ones. Functional decomposition is a method to achieve this goal directly by analyzing a function's properties instead of algebraic transformations on its representation.

Over the last few years, several companies have introduced a number of different types of FPGA. In the case of FPGA circuits, decomposition is of primary importance. The devices contain a large number of cells, or of groups of cells, organized into blocks (CLD — Configurable Logic Blocks [13], LAB — Logic Array Blocks [2]). Designing a device based on these blocks requires partitioning the circuit into subcircuits associated with internal blocks described by LUT's (Look-Up Tables) [3, 4, 6, 7, 8, 9, 10, 11, 12]. There are different types of cells, or blocks with different number of inputs and outputs.

2. THEORETICAL BACKGROUND

The latest trend in digital circuit partitioning is the use of functional decomposition as the main (and, often, the only) decomposition technique. Various algorithms using functional decomposition were presented in [4, 6, 7, 8, 9, 10, 11, 12]. They are based on the Ashenhurst decomposition [1] and its enhancement by Curtis [5].

Curtis proved the fundamental theorem:

$$v(X_2/X_1) \leq 2^p \Leftrightarrow f(X_2, X_1) = F[g_1(X_1), g_2(X_1),..., g_p(X_1), X_2] \tag{1}$$

If the column multiplicity $v(X_2/X_1)$ is less or equal 2^p, then the function $f(X_2, X_1)$ can be decomposed into F form as shown in Fig. 1.

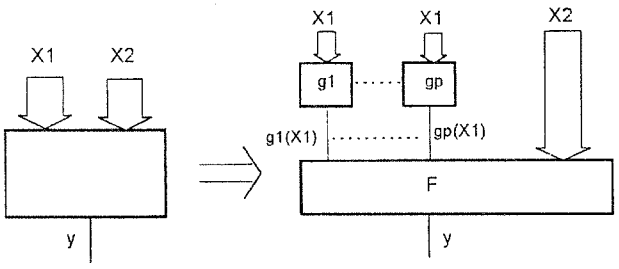


Fig. 1. Curtis decomposition

The algorithm of simple subdecomposition presented in [7] was based on Curtis' theory and allowed a function $y = f(I_{n-1},..., I_1, I_0)$ to be implemented in logical blocks with n_b inputs and m_b outputs.

Sometimes the use of a transcoder additionally allows the number of inputs to subcircuits to be reduced when implementing a given logical function (Fig. 2). This is possible if less than the maximum number of combinations of these variables is needed to define the *on* set of the function to be realized.

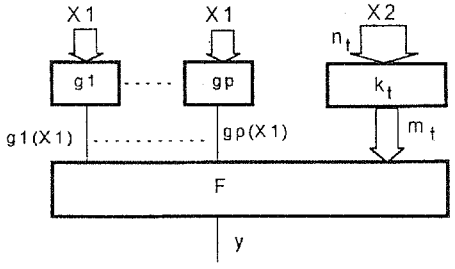


Fig. 2. Functional decomposition with free set variables coding

A transcoder is described by three parameters (Fig. 2):
n_t — the number of inputs, m_t — the number of outputs, and k_t — the number of code words, linked by relationships stated below

$$k_t \leq 2^{m_t}; \quad m_t < n_t \tag{2}$$

Let the function y defined on a set $I = \{I_{n-1},..., I_1, I_0\}$ allow for functional decomposition according to the relationship

$$y = F[g_1(X_1), g_2(X_1),..., g_p(X_1), X_2] \tag{3}$$

$X_1 = \{I_{k-1},..., I_1, I_0\}; \quad X_2 = \{I_{n-1},..., I_{k+1}, I_k\}.$

The X_1 is called bound set and the X_2 is called free set. A direct implementation of functions $g_1(X_1), g_2(X_1),..., g_p(X_1)$ by logical blocks is possible if the number of inputs of these blocks n_b conforms with the relationship $n_b \geq \overline{\overline{X}}_1$. If the number of outputs of logical blocks $m_b > 1$ then it is possible to group the functions g and implement m_b functions simultaneously in one block.

Let the complexity coefficient $w_k(I_{s-1},..., I_{p+1}, I_p)$ will be the number of different words formed by the input variables $I_{s-1},..., I_{p+1}, I_p$. The number of different words formed by the functions $g_1(X_1), g_2(X_1),..., g_p(X_1)$ is $w_k(g_1(X_1), g_2(X_1),..., g_p(X_1))$. If the power of the set X_2 is greater than $[\lg_2 w_k(g_1(X_1), g_2(X_1),..., g_p(X_1))]$, then introducing an external transcoder can limit the number of input. The transcoder would have to possess the following parameters:

$$\begin{gathered} n_t \geq \overline{\overline{X}}_2; \quad k_t > w_k(g_1(X_1), g_2(X_1),..., g_p(X_1)); \\ m_t \geq [\lg_2 w_k(g_1(X_1), g_2(X_1),..., g_p(X_1)] \end{gathered} \tag{4}$$

The last block used to implement the function y is the block implementing the function F. The number of its inputs is equal to

$$p + [\lg_2 w_k(g_1(X_1), g_2(X_1),..., g_p(X_1))].$$

The above considerations allow us to formulate the algorithm of functional decomposition with free set variables coding.

3. THE FUNCTIONAL DECOMPOSITION ALGORITHM WITH FREE SET VARIABLES CODING

Definition 1

By the set of active vectors A_y we shall understand the set containing all the vectors for which the function assumes an active state.

Definition 2

By the full set P_X of the set $X = \{I_{k-1}, I_1, I_0\}$ we shall understand the set 2^k of vectors containing all the combinations L, H of the variables in the set X, where $k = \overline{\overline{X}}$.

Let us denote the vectors in P_X and A_y respectively by

$$P_X = \{P^k_{2-1,..., p_1, p_0}\}_{I_{k-1},..., I_1, I_0}; \quad A_y = \{a_{r-1},..., a_1, a_0\}_{I_{n-1}, ..., I_1, I_0}$$

Definition 3

By the characteristic table of the sets A_y and P_X we shall understand a table $A_y \nabla P_X$ of the size $r \times 2^k$, whose rows correspond to the elements of A_y, columns — to the elements of P_X, and whose elements are determined as follows

$$t_{ij} = \begin{array}{ll} a_i - p_j & \text{if } p_j \subset a_i \\ \varphi & \text{if } p_j \not\subset a_i \end{array}$$

where

$$i = 0, 1, \ldots, r-1; \quad j = 0, 1, \ldots, 2^{k-1} \quad r = \overline{\overline{A}}_y; \quad k = \overline{\overline{X}}$$

Definition 4

By a characteristic subset T_j $(0 \leq j \leq 2^k - 1)$ of the characteristic table $A_y \nabla P_X$ we shall understand the set

$$T_j = \bigcup_{i=0}^{r-1} t_{ij} \quad \text{where} \quad r = \overline{\overline{A}}_y; \quad k = \overline{\overline{X}}; \quad t_{ij} \text{ — element of } A_y \nabla P_X$$

Example 1:

Let

$A_y = \{LLLLL, LLLHH, LHLLH, LHLHL, LHHLL, LHHLH, LHHHL, LHHHH, HHLLL, HHLHH\}_{I3,I2,I1,I0}$

$X = \{I_1, I_0\} \Rightarrow P_X = \{LL, LH, HL, HH\}_{I1,I0}$

The characteristic table $A_y \nabla P_X$ is presented in Table 1. From the columns of the characteristic table it is possible to determine the characteristic subsets T_3, T_2, T_1, T_0 (see Table 1).

Table 1

Characteristic table $A_y \nabla P_x$

I4,I3,I2,I1,I0	LL	LH	HL	HH	I1,I0
LLLLL	LLL	ϕ	ϕ	ϕ	
LLLHH	ϕ	ϕ	ϕ	LLL	
LHLLH	ϕ	LHL	ϕ	ϕ	
LHLHL	ϕ	ϕ	LHL	ϕ	
LHHLL	LHH	ϕ	ϕ	ϕ	
LHHLH	ϕ	LHH	ϕ	ϕ	
LHHHL	ϕ	ϕ	LHH	ϕ	
LHHHH	ϕ	ϕ	ϕ	LHH	
HHLLL	HHL	ϕ	ϕ	ϕ	
HHLHH	ϕ	ϕ	ϕ	HHL	
				I4, I3, I2	

$T_0 = \{LLL, LHH, HHL\}$ $T_1 = \{LHL, LHH\}$ $T_2 = \{LHL, LHH\}$ $T_3 = \{LLL, LHH, HHL\}$

Definitions 1–4 allow us now to present the following algorithm for the implementation of a function $y = f(I_{n-1}, \ldots, I_1, I_0)$ by logical blocks with n_b inputs and m_b outputs each and a transcoder for free set variables coding.

1. Consider partitionings of the set $I = \{I_{n-1},..., I_1, I_0\}$ into disjoint subsets X_1, X_2 such that

$$\overline{\overline{X}}_1 \leq n_b; \quad X_1 \cup X_2 = I; \quad X_1 \cap X_2 = \Phi.$$

2. For each partitioning, determine the characteristic tables $A_y \nabla P_{X_2}$ and the characteristic subsets
3. Verify if the characteristic subsets other than P_{X_1} and Φ can be partitioned into p groups such that an arbitrary characteristic subset belongs to T_i^* or $T_i^{*\prime}$, such that

$$\underset{1 \leq i \leq p}{\forall} \quad T_i^* \cup T_i^{*\prime} = P_{X_1} \qquad T_i^* \cap T_i^{*\prime} = \Phi$$

or, in the particular case if $T_i^{*\prime} = \Phi$ then $T_i^* \subset P_{X_1}$.

4. If it is possible to partition the characteristic subsets then it is also possible to partition the set A_y, into p disjoint parts A_{yi}, each of which is the set of active vectors of a function $g_i(X_1)$, by applying the following principle:
an arbitrary element a_j of the set A_y belongs to the set A_{yi} if there exists a vector $p_k \in P_i$ such that $p_k \subset a_j$

$$T_i^*, T_i^{*\prime} \rightarrow P_i \in P_{X_2} \rightarrow A_{y_2} \rightarrow g_i(X_1).$$

5. Determine $w_k(g_1(X_1), g_2(X_1),..., g_p(X_1))$.
If $[\lg_2 w_k(g_1(X_1), g_2(X_1),..., g_p(X_1))] \leq \overline{\overline{X}}_2$, then it is possible to limit the number of inputs to the block implementing the function $F[g_1(X_1), g_2(X_1),..., g_p(X_1), X_2]$ by using a transcoder with the following parameters:

$$n_t \geq \overline{\overline{X}}_2, \; k_t > w_k(g_1(X_1), g_2(X_1),..., g_p(X_1), \; m_t \geq [\lg_2 w_k(g_1(X_1), g_2(X_1),..., g_p(X_1))].$$

6. Select the best solution (strongest limitation of the number of inputs to the block implementing the function F) by maximizing:

$$\overline{\overline{X}}_1 + \overline{\overline{X}}_2 - p - [\lg_2 w_k(g_1(X_1), g_2(X_1),..., g_p(X_1))].$$

7. If the number of inputs to the block implementing the function F is greater than n_b then seek a functional decomposition of the function F (step 1).

4. EXAMPLE

The functional decomposition algorithm with free set variables coding has been implemented in a procedure of prototype decomposition program (Decomp). The input to the program is a pla file and the output is a set of n_b-input, m_b-output LUTs described by pla files.

Let us try to implement the function $y0$ of benchmark rd84 using Configurable Logic Blocks (CLB) contained in an FPGA structure by XILINX, series 3000. The CLB includes a look-up table that can implement any function of five variables, or any two functions of four variables that use no more than five distinct inputs. The

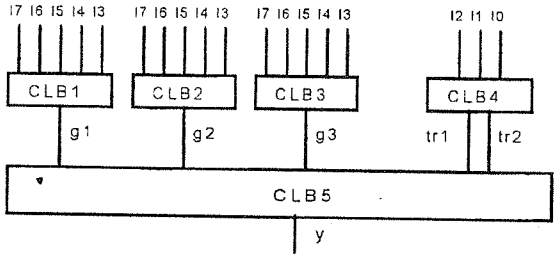
CLB has two outputs. If the CLB is used as a transcoder, it will have the following parameters:

$$n_t = 5; \quad m_t = 1; \quad k_t = 2$$

or

$$n_t = 4; \quad m_t = 2; \quad k_t = 4.$$

Using the functional decomposition with free set variables coding procedure it is possible to implement the function $y0$ of rd84 benchmark in five CLBs, as shown in Fig. 3. The files describing each CLB are presented in Table 2.



$$X_1 = \{I7, I6, I5, I4, I3\}; \qquad X_2 = \{I2, I1, I0\}$$

Fig. 3. Implementing the $y0$ function of benchmark rd84 in CLBs of the XILINX Series 3000

Table 2

Files describing the functions implemented in the individual CLBs

CLB1	CLB2	CLB3	CLB4 (transcoder)	CLB5
.i 5 .o 1 .ilb i7 i6 i5 i4 i3 .ob g1 .p 5 -1111 1 1-111 1 11-11 1 111-1 1 1111- 1 .e	.i 5 .o 1 .ilb i7 i6 i5 i4 i3 .ob g2 .p 20 00011 1 00101 1 01001 1 10001 1 00110 1 01010 1 10010 1 01100 1 10100 1 11000 1 00111 1 01011 1 10011 1 01101 1 10101 1 11001 1 01110 1 10110 1 11010 1 11100 1 .e	.i 5 .o 1 .ilb i7 i6 i5 i4 i3 .ob g3 .p 16 00001 1 00010 1 00100 1 01000 1 10000 1 00111 1 01011 1 10011 1 01101 1 10101 1 11001 1 01110 1 10110 1 11010 1 11100 1 11111 1 .e	.i 3 .o 2 .ilb i2 i1 i0 .ob tr2 tr1 .p 7 001 01 010 01 100 01 011 10 101 10 110 10 111 11 .e	.i 5 .o 1 .ilb g3 g2 g1 tr2 tr1 .ob y p 7 00111 1 0101- 1 01101 1 0111- 1 100-- 1 1010- 1 10110 1 .e

CONCLUSION

The decomposition technique presented in this paper is a specific multiple disjunctive decomposition. The algorithm presented here can be a valuable addition to the algorithms, presented in [7], of partitioning a digital circuit into the individual logical blocks. Experimental results highlight the fact that for those functions which subject decomposition with free set variables coding by introducing external transcoder can be obtain implementation using less number of levels.

REFERENCES

1. R.L. Ashenhurst: *The decomposition of switching functions*, Proceedings of an International Symposium on the Theory of Switching, April 1957
2. Altera: *Data Book*, 1995
3. B. Babba, M. Crastes, G. Saucier: *Input driven synthesis on PLDs and PGAs*. The European Conference on Design Automation, Brussels, Belgium March 16 – 19, 1992
4. S.D. Brown, R.J. Francis, J. Rose, Z.G. Vranesic: *Field Programmable Gate Arrays*. Kluwer Academic Publishers, Boston/London/Dotdrecht, 1993
5. H.A. Curtis: *The Design of switching Circuits*. D. van Nostrand Company, Inc,. Princeton, New Jersey, Toronto, New York, 1962
6. F. Dresig, Ph. Lanches, O. Rettig, U.G. Baitiger: *Functional Decomposition for Universal Logic Cells using Substitution*. The European Conference on Design Automation, Brussels, Belgium March 16 – 19, 1992
7. D. Kania: *Table decomposition methods of digital circuits. Implementation of these circuits in selected PLD structures*. PhD Thesis, Silesian Technical University, Gliwice 1995
8. D. Kania, E. Hrynkiewicz: *Variable partitioning method based on subdecomposition*. Proceedings of the XIXth National Conference Circuit Theory and Electronics Circuits, Kraków-Krynica, 1996
9. T. Łuba, H. Selvaraj, A. Kraśniewski: *A new approach to FPGA-based logic synthesis*. Workshop on Design Methodologies for Microelectronics and Signal Processing, Gliwice-Krakow, 20 – 23 Oct 1993
10. T. Sasao: *Logic Synthesis and Optimization*. Kluwer Academic Publishers, Boston/London/Dotdrecht, 1993
11. H. Selvaraj, T. Łuba, M. Nowicka, B. Bignall: *Multiple-valued decomposition and its applications in data compression and technology mapping*. ICCIMA'97, Chundill, Australia
12. W. Wan, M.A. Perkowski: *A new Approach to the Decomposition of Incompletely Specified Multi-Output Functions Based on Graph Coloring and Local Transformations and Its Applications to FPGA Mapping*, Proceedings EDAC'92, Hamburg, 1992
13. Xilinx: *The Programmable Logic Devices*, 1993

D. KANIA

ALGORYTM DEKOMPOZYCJI FUNKCJONLNEJ Z KODOWANIEM ZMIENNYCH ZBIORU WOLNEGO

Streszczenie

Artykuł przedstawia algorytm dekompozycji funkcjonalnej z kodowaniem zmiennych zbioru wolnego. Zaprezentowany algorytm oparty jest na teorii Curtisa i jest ulepszoną wersją algorytmu subdekompozycji przedstawionego w pracy [7]. Wprowadzenie dodatkowego trankodera pozwala na ograniczenie liczby poziomów i wejść podukładów wykorzystywanych do realizacji funkcji logicznych.

Przedstawiony w pracy algorytm, w postaci oddzielnej procedury, został zaimplementowany w prototypie programu służącego do dekompozycji funkcji logicznych, W niniejszym artykule wynik dekompozycji został zaprezentowany na przykładzie funkcji y_0 zbioru testowego (benchmark) rd 84.

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI, 1998, **44**, z. 3, ss. 325–334

Digital FIR filter design by alternate iterations in the time and frequency domains

EWA HERMANOWICZ, MIROSŁAW ROJEWSKI, MAREK BLOK

Instytut Elektroniki, Telekomunikacji i Informatyki, Politechnika Gdańska

Received 1998.02.23

Authorized 1998.05.29

A technique for designing digital FIR filters aided by a computer program in the MATLAB-WINDOWS environment is presented, based on alternate iterations in the time and frequency domains. The technique is aimed at designing filters with, generally, complex frequency response by forcing the complex approximation error magnitude rather than solely the amplitude response of the filter to the assigned tolerance specifications. The advantages of the proposed approach are illustrated by a design of a Fractional Sample Delay filter. If this filter has to meet demanding specifications then the proposed approach provides a more effective solution than the traditional approach based on windowing technique using classical window functions.

Key words: Digital filters, fractional-sample delay filters, fast Fourier transformation.

1. INTRODUCTION

In this paper a technique for designing finite impulse response (FIR) filters using the fast Fourier transform (FFT) algorithm is presented. The FIR filter design problem is formulated to alternately satisfy the frequency domain constraints on the magnitude of the complex approximation error bounds and the time domain constraints on the impulse response length. The design scheme is based on iterations. Each iteration requires computation of the FFT and its inverse. The technique is simple and versatile. It allows for the approximation of a filter of any arbitrarily chosen complex frequency response characteristics with arbitrary complex approximation error specifications as well. The iterative design process can be readily modified, extended to more than one dimensions and organized in an interactive way giving an insight into the intermediate results.

The problem of designing an FIR filter consists of determining the impulse response $h[n]$ or its frequency response $H(e^{j2\pi f})$, so that given requirements on the

filter response are satisfied. The impulse response here is selected as centred at the origin, i.e. for this response the discrete-time independent variable n for a FIR filter of lenght N takes the values $n = 0, \pm 1, ..., (N-1)/2$, where N is an odd integer. The causal filter $\tilde{h}[n]$, $n = 0, 1, ..., N-1$ can be obtained by delaying $h[n]$ by $(N-1)/2$ sample intervals as given by

$$\tilde{h}[n] = h[n - (N-1)/2], \quad n = 0, 1, ..., N-1. \tag{1}$$

For the causal FIR filter, which is a liner shift-invariant system, the input-output relation is represented by

$$y[n] = \sum_{k=0}^{N-1} \tilde{h}[k]\, x[n-k], \tag{2}$$

where $x[n]$ and $y[n]$ are the input and output sequences, respectively.

The frequency response $H(e^{j2\pi f})$ corresponding to the non-causal impulse response $h[n]$ is expressed as

$$H(e^{j2\pi f}) = \sum_{n=-(N-1)/2}^{(N-1)/2} h[n] \exp(-j2\pi f n), \quad |f| \leq 1/2 \tag{3}$$

and $H(e^{j2\pi f})$ is a periodic function of the frequency f with the period equal to 1. Thus it is sufficient to specify the frequency response of the filter in the region $|f| \leq 1/2$.

In order to meet given specifications, an adequate filter length N needs to be detrmined. Generally, there is a trade-off between the filter length and the frequency domain specifications, which are most frequently formulated in the form of bounds on the filter's amplitude response. The filter length may be determined through an iterative process. If the filter length is fixed, then the tolerance parameters must be properly adjusted to meet the design specifications.

The design technique proposed in this paper is based on specifying bounds on the complex approximation error magnitude rather than on the amplitude response of the filter solely. This allows for an FIR approximation of filters whose impulse response is not necessarily symmetric, as is the case of linear-phase filters. A typical example of a filter whose impulse response is generally non symmetric is the fractional-sample delay (FSD) filter [1], [2] which finds a number of aplications, e.g., in synchronisation and incommensurate resampling. This filter is used further to illustrate the potential of the proposed design technique.

2. THE PROPOSED DESIGN TECHNIQUE

The idea of the proposed technique of designing digital FIR filters by alternate iterations in the time and frequency domains is presented in Fig. 1 in the form of a block scheme of the design algorithm. The iterative method begins with a finite-extent real sequence

$$\{h_{N,0}[n]\}_{-(N-1)/2}^{(N-1)/2} = \{h_{N,i}[n]\}_{-(N-1)/2}^{(N-1)/2}|_{i=0} \equiv h_{N,i}[n]|_{i=0}, \quad n = 0, \pm 1, ..., \pm (N-1)/2 \tag{4}$$

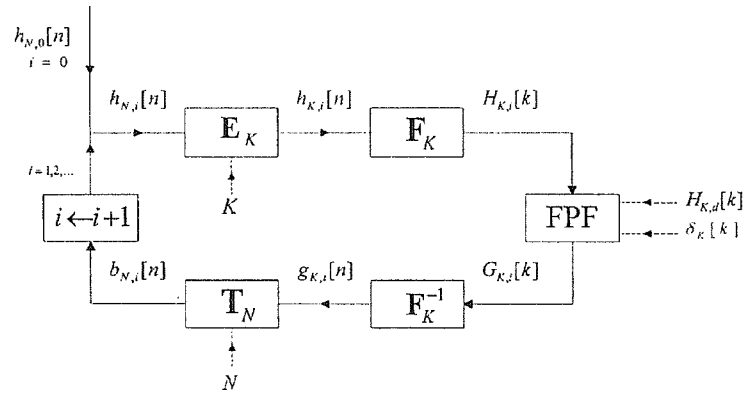


Fig. 1. Block scheme of the proposed algorithm for designing digital FIR filters by alternate iterations in the time and frequency domains. Broken lines with arrows indicate the additional parameters and characteristcs introduced from outside the loop

This sequence is the impulse response of an initial ($i=0$) FIR filter design of length N, where it is assummed that N is an odd integer. A good candidate for the initial estimate can be obtained, e.g., by windowing by means of a classical window [3] the infinite impulse response $h_d[n]$, $n = 0, \pm 1,...$ of an ideal filter whose frequency response is $H_d(e^{j2\pi f})$. This can be achieved by using

$$h_d[n] = \int_{-1/2}^{1/2} H_d(e^{j2\pi f})\, e^{j2\pi f n}\, df \tag{5}$$

with a chosen window function $w_N[n]$, $n = 0, \pm 1,..., \pm(N-2)/2$ to get

$$h_{N,0}[n] = h_d[n]\, w_N[n], \qquad n = 0, \pm 1,..., \pm(N-1)/2. \tag{6}$$

This impulse response is the input to a loop in Fig. 1 where it undergoes processing in successive iterations: $i = 0, 1, 2,....$ In the loop the impulse response (4) is symmetrically elongated according to the following relationship

$$\{h_{K,i}[n]\}_{-(K-1)/2}^{(K-1)/2} \triangleq \mathbf{E}_K\{h_{N,i}[n]\}_{-(N-1)/2}^{(N-1)/2} = \begin{cases} h_{N,i}[n], & |n| \le (N-1)/2 \\ 0, & (N-1)/2 < |n| \le (K-1)/2, \end{cases} \tag{7}$$

where $K > N$ and K is an odd integer. This operation is performed in Fig. 1 by an elongator $\mathbf{E}_K$. Next, the discrete Fourier transform (DFT) of length K of the sequence (7) is computed, as given by

$$\{H_{K,i}[k]\}_{-(K-1)/2}^{(K-1)/2} \triangleq \mathbf{F}_K\{h_{K,i}[n]\}_{-(K-1)/2}^{(K-1)/2} = \sum_{n=-(K-1)/2}^{(K-1)/2} h_{K,i}[n] \exp(-j2\pi nk/K),$$
$$k = 0, \pm 1,..., \pm(K-1)/2 \tag{8}$$

by using the K-point DFT operator $\mathbf{F}_K$.

Of paramount importance in Fig. 1 is a block called FPF, intended for constraining the desired features and parameters of the filter in the frequency domain. This forcing is done here by using the complex approximation error magnitude to impose the desired constraints on the filter frequency response. The complex approximation error is defined as

$$E_{K,i}[k] = H_{K,d}[k] - H_{K,i}[k], \quad k = 0, \pm 1, \ldots, \pm(K-1)/2, \tag{9}$$

where $\{H_{K,d}[k]\}|_{-(K-1)/2}^{(K-1)/2}$ stands for a set of K equidistant frequency points of the ideal frequency resonse $H_d(e^{j2\pi f})$, i.e. $H_{K,d}[k] = H_d(e^{j2\pi f})|_{f=k/K}$ with k as in (9). In the FPF block a new frequency response in the given set of frequency points is created, according to the following relationship

$$G_{K,i}[k] = \begin{cases} H_{K,i}[k] & \text{if} \quad |E_{K,i}[k]| \leq \delta_K[k] \\ H_{K,d}[k] - \delta_K[k]\, E_{K,i}[k]/|E_{K,i}[k]| & \text{if} \quad |E_{K,i}[k]| > \delta_K[k] \end{cases} \tag{10}$$

where $0 \leq \delta_K[k] < 1$, $k = 0, \pm 1, \ldots, \pm(K-1)/2$ is a real-valued deviation characteristics for the complex approximation error magnitude. The imposed constraints are given on the right-hand side of Eq. (10). The core of the technique is outlined below. According to (10), if the complex approximation error magnitude, $|E_{K,i}[k]|$, at a given point k, is not greater than the deviation $\delta_K[k]$, then the frequency response $G_{K,i}[k]$ at this point takes the value $H_{K,i}[k]$. Otherwise, a modified frequency response value is assigned to this point. Namely, the current value $H_{K,i}[k]$ is replaced by the value for which the complex approximation error is $\delta_K[k]\, E_{K,i}[k]/|E_{K,i}[k]| = \delta_K[k]\, e^{j \arg E_{K,i}[k]}$ instead of $E_{K,i}[k]$. Consequently, the error magnitude is replaced by the maximal allowed value $\delta_K[k]$ and the error phase remains unchanged.

The deviation characteristics $\delta_K[k]$ reflects the degree to which it is allowed for the desired filter to deviate from an ideal filter. In the case of selective filters with one passband and one stopband in the nonnegative frequency range $0 \leq f \leq 1/2$, this characteristics assumes the values: δ_p in passband and δ_s in stopband, where δ_p and δ_s stand for the maximum passband and stopband deviation of the complex approximation error magnitude from the ideal zero value. For example, for a zero-phase lowpass filter with passband and stopband cutoff frequencies: $f_p = k_p/K$ and $f_s = k_s/K$ respectively, the ideal frequency response is

$$H_{K,d}[k] = \begin{cases} 1, & k = 0, \pm 1, \ldots, \pm k_p \\ 0, & k = \pm k_s, \pm(k_s+1), \ldots, \pm(K-1)/2 \end{cases} \tag{11}$$

and the deviation characteristics is

$$\delta_K[k] = \begin{cases} \delta_p, & k = 0, \pm 1, \ldots, \pm k_p \\ \delta_s, & k = \pm k_s, \pm(k_s+1), \ldots, \pm(K-1)/2. \end{cases} \tag{12}$$

Normally, the deviation characteristics in the transition band (f_p, f_s) remains unspecified. Then, this band is called the off-care band.

Once the discrete-frequency response $G_{K,i}[k]$ (10) has been obtained, the impulse response

$$\{g_{K,i}[n]\}_{-(K-1)/2}^{(K-1)/2} \triangleq \mathbf{F}_K^{-1}\{G_{K,i}[k]\}_{-(K-1)/2}^{(K-1)/2} = \frac{1}{K}\sum_{k=-(K-1)/2}^{(K-1)/2} G_{K,i}[k]\exp(j2\pi nk/K), \tag{13}$$

$$n = 0, \pm 1, \ldots, \pm(K-1)/2$$

corresponding to (10) is computed by using the inverse K-point DFT operator denoted in Fig. 1 by $\mathbf{F}_K^{-1}$. Next, the time-domain constraint is imposed on the filter's impulse response. This constraint requires that the filter coefficients must be zero outside the region $n = 0, \pm 1, \ldots, \pm(N-1)/2$. According to this, the impulse response $g_{K,i}[n]$ is symmetrically truncated from length K to N, to give the sequence

$$\{b_{K,i}[n]\}_{-(N-1)/2}^{(N-1)/2} \triangleq \mathbf{T}_K\{g_{K,i}[n]\}_{-(K-1)/2}^{(K-1)/2} = g_{K,i}[n], \quad n = 0, \pm 1, \ldots, \pm(N-1)/2. \tag{14}$$

Finally, the sequence $\{b_{K,i}[n]\}_{-(N-1)/2}^{(N-1)/2}$ is treated as the input sequence for the next iteration. Thus at the elongator input we have

$$h_{N,i}[n] = \begin{cases} h_{N,0}[n], & i = 0 \\ b_{N,i-1}[n], & i = 1, 2, \ldots \end{cases} \tag{15}$$

The algorithm is stopped when the magnitude of the complex aproximation error of the discrete-time frequency response $H_{K,i}[k]$ fulfills the condition

$$|E_{K,i}[k]| \leq \delta_K[k] \tag{16}$$

for all k for which the deviation characteristic has been specified. If the algorithm is stopped at the i-th iteration stage, the impulse response of the target filter is

$$h_N[n] = b_{N,i-1}[n], \quad n = 0, \pm 1, \ldots, \pm(N-1)/2. \tag{17}$$

Consequently, each iteration of the algorithm consists of making successive imposition on time and frequency domain constraints onto the current iterate. The i-th iteration consists of the following steps:

- Elongate the i-th iterate from $h_{N,i}[n]$ to $h_{K,i}[n]$ (7).
- Compute the DFT of the i-th iterate $h_{K,i}[n]$ (8).
- Check the condition (16). If this condition is not fulfilled for all k of interest, for which the deviation characteristic has been specified, then impose the constraints given in (10) to obtain $G_{K,i}[k]$. Otherwise exit by taking $b_{N,i-1}[n]$ from the $(i-1)$-st iteration as the target filter impulse response $h_N[n]$ (17).
- Compute the inverse DFT, $g_{K,i}[n]$, of $G_{K,i}[k]$ (13).
- Impose the time-domain constraints by truncating $g_{K,i}[n]$ to obtain $b_{N,i}[n]$ (14).
- Assign $b_{N,i}[n]$ to $h_{N,i+1}[n]$ to be the input for the next iteration stage.

3. DESIGN EXAMPLE AND DESIGN TOOLS

In order to display the performance of the proposed design technique, we present an illustrative numerical example and compare the results with those obtained with the classical windowing technique [3] and with the optimal Chebyshev design [4]. The algorithm was coded in MATLAB [5]. The filter chosen was a fractional-sample delay (FSD) filter [1], [2]. The design here was aimed at reducing the maximal size of the ripples exhibited by the frequency response approximation error magnitude of the FSD filter initially designed using the windowing technique.

The desired frequency response of the FSD filter is defined as given by

$$H_d(e^{j2\pi f}) = \exp(-j2\pi fd), \quad |f| < 1/2, \tag{18}$$

where d stands for the fractional delay which is a noninteger fraction of a sample interwal, and $|d| \leq 0.5$. The amplitude response of this ideal filter is 1 and the group delay response is d. The impulse response corresponding to (18) is

$$h_d[n] = \frac{\sin \pi(n-d)}{\pi(n-d)} = \operatorname{sinc}(n-d), \quad n = 0, \pm 1, \ldots \tag{19}$$

In our example the design specifications for the FIR approximation to this filter in the discrete-frequency domain are as follows (cf. (11)):

$$H_{K,d}[k] = \exp(-j2\pi kd/K), \quad k = 0, \pm 1, \ldots, \pm k_p \tag{20}$$

with $d = 0.3$. The desired bandwidth (DBW) of our FSD filter is specified to be DBW = 0.45. For the convergence of the iterative algorithm in Fig. 1 it is important to have the cutoff frequency represented accurately on the frequency grid of the DFT algorithm. Because of this the passband cutoff frequency for the FSD filter assumed in the example is

$$f_p = \lceil \mathrm{DBW} \cdot K \rceil / K, \tag{21}$$

where $\lceil x \rceil$ stands for the integer part of x not smaller than x, and the number of the frequency point in (20) corresponding to f_p must be $k_p = Kf_p$. Thus the filter has one passband and a transition band. The deviation characteristic for the FSD filter in the passband is specified to be (cf.(12))

$$\delta_K[k] = \delta_p, \quad k = 0, \pm 1, \ldots, \pm k_p \tag{22}$$

with the tolerance parameter $\delta_p = 0.0025$. The discrete-frequency response (20) as well as the deviation characteristic (22) remain unspecified in the transition band. Thus $k = \pm(k_p + 1), \ldots, \pm(K-1)/2$ are the numbers of frequency points belonging to the off-care band. In the example, the approximating filter length is $N = 35$ and the size of the DFT is $K = 11N = 385$. The filter characteristics obtained after 20 iterations are illustrated in the upper part of Fig. 2 in the form of the amplitude response and the group delay response, both zoomed in the desired bandwidth, DBW. The middle part of Fig. 2 shows the performance of the initial filter. This was

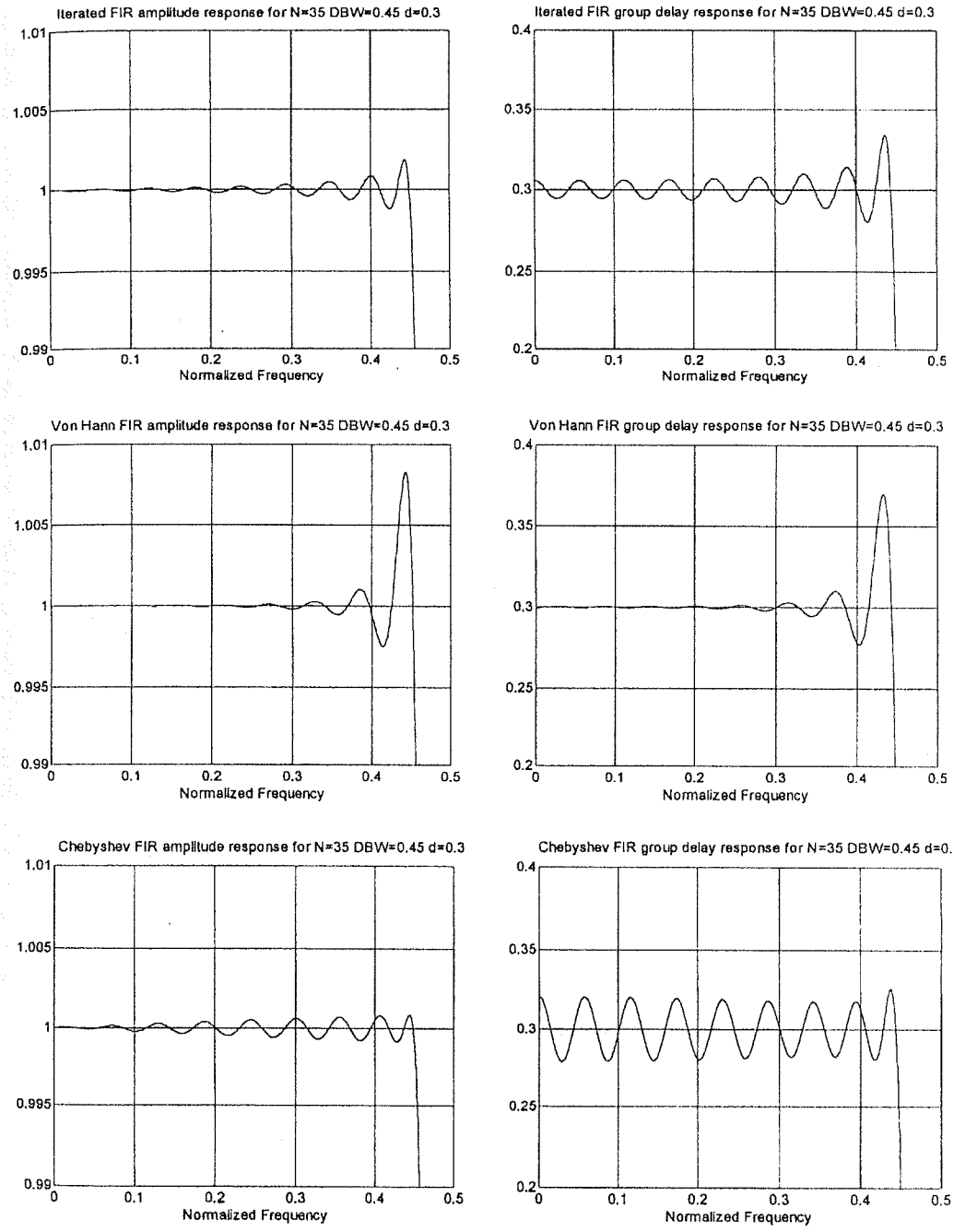


Fig. 2. Amplitude response and group delay response of the FSD filter designed by the proposed iterative method (upper part) and by using the von Hann window (middle part). The lower part shows the reference optimal Chebyshev design

obtained by the classical windowing technique using the offset von Hann window [1] of the same length $N = 35$. The maximal values of passband ripples for the iterated design from Fig. 1 as compared with the von Hann design were significantly reduced from approximately 0.0085 to 0.002 (more than 4 times) on the amplitude response and from 0.07 to 0.035 (twice) on the group delay response. The total peak complex approximation error magnitude measured in the filter passband decreased from the initial value 0.01 (−40dB) to the desired value $\delta_p = 0.0025$ (−52dB). This confirms that the proposed design technique convincingly proves its worth.

If the length of the FFT is very large, the iterations converge to the optimal Chebyshev (minimax) approximation (see the lower part of Fig. 2) which minimizes the Chebyshev norm [4], i.e. the maximal value of the complex approximation error magnitude in the FSD filter passband. However, to achieve the performance of the iterated design close to the optimal one, one must have the possibility to intervene in the design process. this is because ideally the iterative algorithm should be implemented via the discrete-time Fourier transformation with the frequency variable continuous rather than discrete. The convergence of the practical algorithm strongly depends on the FFT length. Moreover, the precise value of the passband edge, the constrained value of the tolerance parameter δ_p and the initial design are also very important for convergence. Experience shows that when the filter has to meet demanding specifications, even very small changes in the constrained parameters may strongly influence the iterative process and speed up the design time. Thus it was indispensable to have a design tool at disposal through which the designer could modify the iterative process in an interactive way having an insight into the intermediate results.

To meet these requirements a graphical user interface in MATLAB-WINDOWS environment was created [5]. Fig. 3 shows the main window for communication with the FIR filter design algorithm from Fig. 1 without programming commands. Shown in this window are buttons for modification of the inital design and for influencing the constrained parameters at every stage of the design process. Also in this window the dynamical changes in the most important filter characteristics can be observed. Actually, Fig. 3 gives a snapshot taken during an interactive process of attempting to achieve an optimal FSD filter design with an equiripple complex frequency response approximation error magnitude (cf. the lower part of Fig. 2). Such a design was reached here after 4172 iteration steps using a FFT of length $K = 8192$.

CONCLUSIONS

In this paper a technique employing alternate iterations in the time and frequency domains has been presented aimed at designing FIR filters with complex frequency response. the iterative algorithm employs the FFT at each iteration. The design process has been organized in an interactive way by means of a graphical user interface in a MATLAB-WINDOWS environment in order to facilitate the conver-

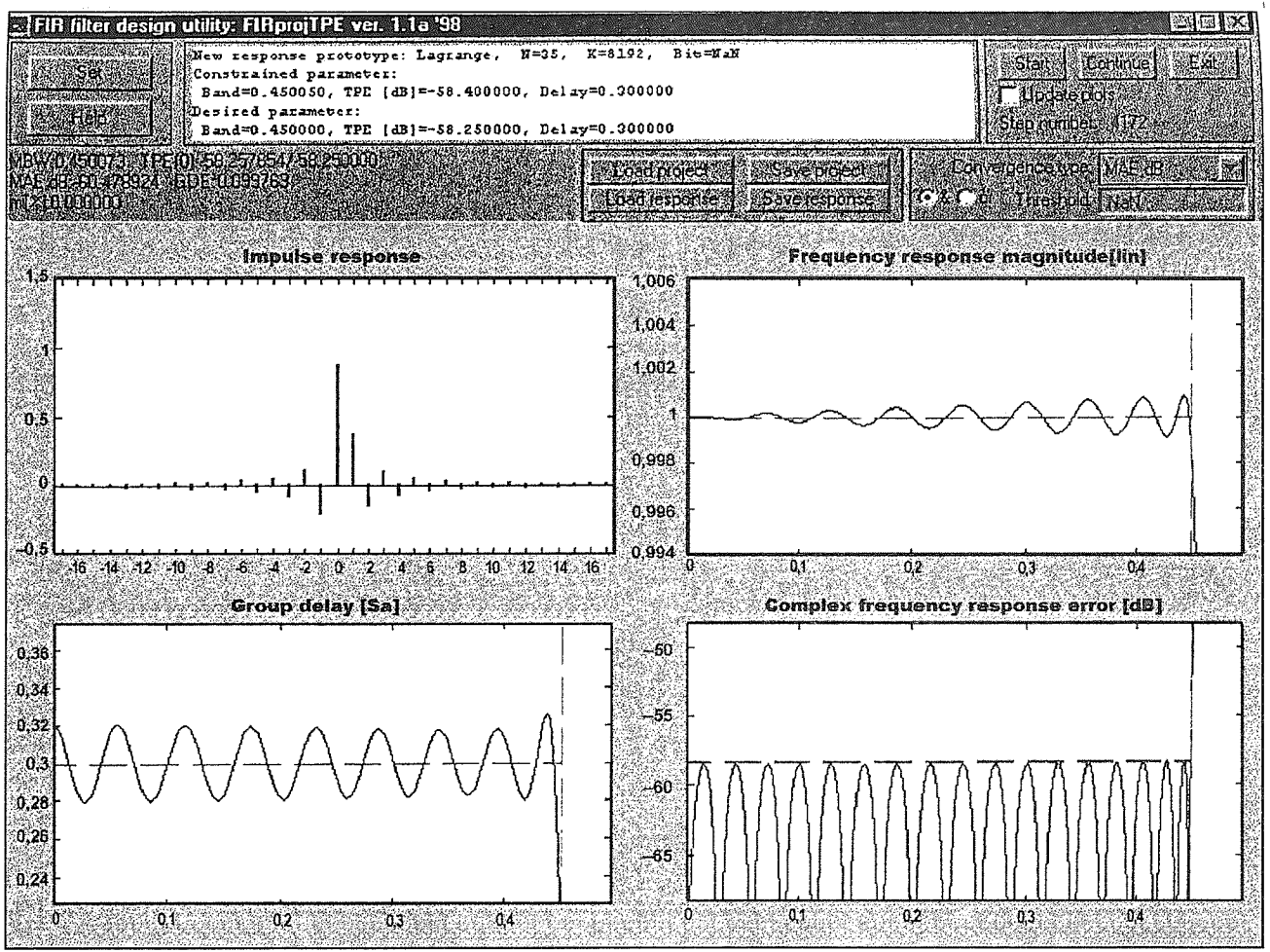


Fig. 3. The graphical user interface for communication with the algorithm for designing FIR filters by the proposed iterative technique

gence of the iterates and to speed up the design time. Experience shows that the algorithm converges to an optimal Chebyshev solution which minimizes the maximal value of the complex frequency response approximation error magnitude.

REFERENCES

1. G.D. Cain, A. Yardim and P. Henry: *Offset windowing for FIR fractional-sample delay.* Proc. Intern. Conference on Acoustics, Speech and Signal Processing, ICASSP'95. Detroit, 8–12 May 1995, vol. 2, pp. 1276–1279
2. E. Hermanowicz, M. Rojewski, G.D. Cain, A. Tarczyński: *Special discrete-time filters having fractional delay. Signal Processing*, 1998, vol. 67, pp. 279–289
3. S.K. Mitra and J.F. Kaiser (eds.): *Handbook for Digital Signal Processing.* J. Wiley 1993
4. L.J. Karam, J.H. McClellan: *Complex Chebyshev approximation for FIR filter design.* IEEE Trans. on Circuits and Systems — III: Analog and Digital Signal Processing March 1995, vol. 2, No 3, pp. 207–216
5. M. Blok, E. Hermanowicz, M. Rojewski: *An algorithm for designing digital FIR filters by alternate iterations in the time and frequency domains.* (Algorytm projektowania filtrw cyfrowych FIR na drodze naprzemiennej iteracji czzsowo-częstotliwościowej). Report No 49/97, Faculty of Electronics, Telecommunications and Informatics, Technical University of Gdańsk (in Polish)

E. HERMANOWICZ, M. ROJEWSKI, M. BLOK

PROJEKTOWANIE FILTRÓW CYFROWYCH FIR NA DRODZE NAPRZEMIENNYCH ITERACJI W DZIEDZINIE CZASU I W DZIEDZINIE CZĘSTOTLIWOŚCI

Streszczenie

W pracy przedstawiono wspomaganą programem komputerowym w środowisku MATLAB-WINDOWS metodę projektowania filtrów cyfrowych FIR na drodze naprzemiennych iteracji w dziedzinie czasu i w dziedzinie częstotliwości. Metoda jest przeznaczona do projektowania filtrów o generalnie, dowolnej zespolonej charakterystyce amplitudowo-fazowej. W ramach specyfikowanego przedziału tolerancji wymuszany jest moduł zespolonego błędu aproksymacji charakterystyki amplitudowo-fazowej, a nie błąd samej tylko charakterystyki amplitudowej. Zalety metody są ilustrowane przykładem projektowania filtru o opóźnieniu ułamkowym. Jeżeli ten filtr ma spełniać ostre wymagania, to proponowana metoda dostarcza bardziej efektywnego rozwiązania niż podejście tradycyjne oparte na metodzie wykorzystującej technikę okien klasycznych.

Słowa kluczowe: filtry cyfrowe, filtry o opóźnieniu ułamkowym, szybkie przekształcenie Fouriera.

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI, 1998, 44, z. 3, ss. 335–344

Wpływ zjawiska samomodulacji fazy na transmisję danych wykorzystującą zjawisko konwersji modulacji częstotliwości na modulację amplitudy

KRZYSZTOF PERLICKI

Instytut Telekomunikacji, Politechnika Warszawska

Otrzymano 1998.04.29

Autoryzowano 1998.06.27

W artykule przedstawiono analizę wpływu zjawiska samomodulacji fazy na parametry systemu transmisji danych wykorzystującego zjawisko konwersji modulacji częstotliwości na modulację amplitudy powodowaną przez dyspersję chromatyczną światłowodu z dwupoziomowym wyjściowym sygnałem optycznym.

Artykuł jest kontynuacją publikacji [1], która zawiera podstawowe informacje dotyczące: zjawiska konwersji modulacji częstotliwości na modulację amplitudy w standardowym światłowodzie jednomodowym z dwupoziomowym optycznym sygnałem wyjściowym i systemu transmisji danych wykorzystującym efekt konwersji. W artykule skoncentrowano się wyłącznie na analizie wpływu nieliniowego zjawiska samomodulacji fazy na parametry omówionego w [1] systemu.

Słowa kluczowe: telekomunikacja światłowodowa, konwersja modulacji częstotliwości na modulację amplitudy, zjawisko samomodulacji fazy.

1. ZJAWISKO SAMOMODULACJI FAZY

Występowanie zjawiska samomodulacji fazy (ang. *Self-Phase Modulation*) w światłowodach spowodowane jest efektem nieliniowego załamania światła polegającym na zależności współczynnika załamania od natężenia światła. Zjawisko to określane jest jako efekt Kerr'a. Współczynnik załamania zależy od częstotliwości i natężenia światła zgodnie z następującą zależnością [2]:

$$n(\omega, |E|^2) = n^2(\omega) + n_2|E|^2, \tag{1}$$

gdzie $n(\omega)$ jest członem liniowym współczynnika załamania, $|E|^2$ — natężenie światła w światłowodzie, a n_2 jest nieliniowym współczynnikiem załamania. Nielinio-

wy współczynnik załamania wynosi około $3{,}2 \cdot 10^{-16}$ cm^2/W [3]. Ze względu na to, że promień modu w światłowodzie jest mały (4–6 μm), a jego tłumienność jest mniejsza od 1 dB/km, to mimo tak małej wartości nieliniowego współczynnika załamania zjawiska nieliniowe można zaobserwować we włóknie optycznym nawet przy niezbyt dużym natężeniu światła. Przesunięcie fazy przez pole elektryczne w światłowodzie po przebyciu pewnej drogi L określa się wzorem:

$$\varphi = knL = k(n + n_2|E|^2)L. \tag{2}$$

Część nieliniowa przesunięcia fazy $\varphi_{NL} = n_2 k|E|^2 L$ powstaje na skutek samomodulacji fazy i jest odpowiedzialna za poszerzenie widma propagujących się w światłowodzie impulsów. Efekty nieliniowe mają wpływ na właściwości propagacyjne światłowodu gdy długość nieliniowa (L_{NL}) jest porównywalna lub mniejsza od długości badanego traktu światłowodowego. Długość nieliniową można określić na podstawie następującej zależności:

$$L_{NL} = \frac{1}{\gamma \cdot P_0} \tag{3}$$

gdzie:
γ — współczynnik nieliniowości;
P_0 — moc szczytowa impulsu na wejściu światłowodu.

2. OPIS BADANEGO SYSTEMU

Analizę transmisji z wykorzystaniem konwersji modulacji częstotliwości na modulację amplitudy przeprowadzono dla systemu przedstawionego na rys. 1 o następujących parametrach:

Nadajnik

- znormalizowana odpowiedź częstotliwościowa lasera wynosi jeden;
- moc wyjściowa z lasera −4 dBm;
- moc wyjściowa ze wzmacniacza końcowego (EDFA1) +13 dBm;

Tor optotelekomunikacyjny

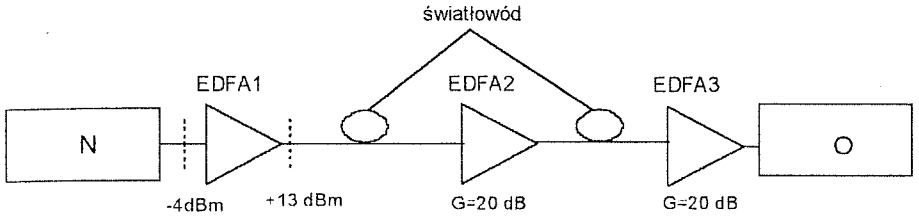
- wzmocnienie wzmacniacza EDFA2 równe 20 dB;
- światłowód jednomodowy standardowy o współczynniku dyspersji równej, dla $\lambda = 1550$ nm, D = 17 ps/nm · km, tłumienności $\alpha = 0{,}2$ dB/km i współczynniku nieliniowości $\gamma = 2{,}17$ W^{-1} km^{-1} [4].

Odbiornik

- wzmocnienie przedwzmacniacza w odbiorniku (EDFA3) równe 20 dB;
- wartość fotoprądu wytworzonego przez fotodiodę wynosi:

$$i(t) = \frac{\eta \lambda e}{hc} P_{wy}(t) + I_d \tag{4}$$

gdzie:
I_d — prąd obejmujący efekty wynikające z obecności szumu,
λ — długość fali światła,
η — sprawność kwantowa detektora,
e — ładunek elektronu,
h — stała Plancka,
c — prędkość światła.



Rys. 1. Schemat badanego systemu N — nadajnik (laser), EDFA1 — wzmacniacz końcowy nadajnika, EDFA2 — wzmacniacz umieszczony na 100 km, EDFA3 — przedwzmacniacz w odbiorniku, O — odbiornik z filtrem

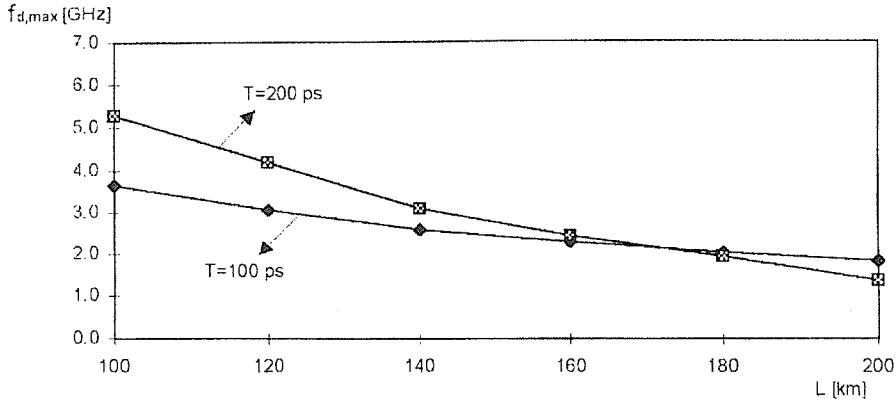
W odbiorniku zastosowano dolnopasmowy filtr Butterwortha pierwszego rzędu o następującej charakterystyce amplitudowej [5]:

$$H_{FR}(j\omega) = \frac{1}{\sqrt{1 + \left(\frac{\omega}{\omega_R}\right)^2}}, \tag{5}$$

gdzie:
ω_R — pulsacja, przy której następuje spadek wartości $H_{FR}(j\omega)$ o 3 dB,
ω — pulsacja.

W celu określenia zachowania się systemu w funkcji długości traktu światłowodowego przyjęto że:
a) dla małych odległości tj. dla L od 0 km do 30 km w systemie występuje przedwzmacniacz EDFA3 i brak jest wzmacniaczy EDFA1 i EDFA2,
b) dla średnich odległości tj. dla L od 40 km do 120 km obok przedwzmacniacza EEFA3 występuje wzmacniacz końcowy EDFA1,
c) dla dużych odległości tj. dla $L \geq 140$ km dodatkowo w celu skompensowania strat na setnym kilometrze umieszczono wzmacniacz przelotowy EDFA2.

Badania zostały przeprowadzone dla traktów optotelekomunikacyjnych długości od 100 km do 200 km. Analizie został poddany system transmisji danych, o przepływności 5 Gbit/s zakodowanych za pomocą kodu bifazowego (Manchester) oparty na konwersji modulacji częstotliwości na modulację amplitudy [1].



Rys. 2. Pary maksymalnej zmiany pulsacji dla stanów o czasie trwania $T = 100$ ps i $T = 200$ ps w funkcji długości toru światłowodowego; $\omega_{d,max} = 2\pi f_{d,max}$

Dewiacja pulsacji optycznej odpowiadająca stanom wysokim i niskim o czasie trwania 100 ps i 200 ps została przedstawiona na rys. 2 [1]. Analiza zachowania się omawianego systemu została przeprowadzona na pseudoprzypadkowej sekwencji stanów niskich i wysokich długości 2^7-1. Do opisu propagacji fali świetlnej w światłowodzie wykorzystano nieliniowe równanie Schrödingera, które zostało rozwiązane metodą dwukrokową (ang. *Split-Step Fourier Method*) [6].

3. ANALIZA WPŁYWU ZJAWISKA SAMOMODULACJI FAZY WYKORZYSTUJĄCĄ KONWERSJĘ MODULACJI CZĘSTOTLIWOŚCI NA MODULACJĘ AMPLITUDY

Jeżeli przyjmiemy, że moc szczytowa impulsu na wejściu światłowodu $P_0 = +13$ dBm i wartość współczynnika nieliniowości $\gamma = 2{,}17$ W^{-1} km^{-1} [4] to z zależności (3) otrzymujemy długość nieliniową równą $L_{NL} = 23$ km. Uwzględniając fakt, że badamy system długości powyżej stu kilometrów należy się spodziewać, że efekt nieliniowy w postaci zjawiska samomodulacji fazy będzie mieć wpływ na właściwości propagacyjne światłowodu, a w konsekwencji na zachowanie się całego systemu. Problemem jest czy to zjawisko nieliniowe zwiększa czy zmniejsza wartości analizowanych przez nas parametrów systemu. Przebadano właściwości systemu dla przypadku, w którym uwzględniono nieliniowości światłowodu i bez jej uwzględniania. Badania przeprowadzono pod kątem analizy dwóch parametrów: parametru Q i stosunku sygnału do szumu (S/N). Parametr Q został zdefiniowany jako stosunek różnicy średniej wartości prądu odpowiadającego stanowi wysokiemu i niskiemu do poziomu szumów [7]:

$$Q = \frac{\bar{I}_1 - \bar{I}_0}{\sigma_1 + \sigma_0}. \tag{6}$$

Wielkości $\bar{I}_1$ i $\bar{I}_0$ odpowiadają średniej wartości prądu związanej odpowiednio ze stanem wysokim i niskim, a σ_1, σ_0 są określone następującymi zależnościami [7]:

$$\sigma_1 = (\sigma^2_{\text{sig-sp},1} + \sigma^2_{\text{sp-sp},1})^{\frac{1}{2}} \tag{7}$$

$$\sigma_0 = (\sigma^2_{\text{sig-sp},0} + \sigma^2_{\text{sp-sp},0})^{\frac{1}{2}}, \tag{8}$$

gdzie:
$\sigma^2_{\text{sig-sp},1}$ i $\sigma^2_{\text{sig-sp},0}$ jest mocą szumu zdudnienia sygnału użytecznego z szumem emisji spontanicznej odpowiednio dla stanu wysokiego i stanu niskiego;
$\sigma^2_{\text{sp-sp},1}$, $\sigma^2_{\text{sp-sp},0}$ jest mocą szumu zdudnienia emisji spontanicznej odpowiednio dla stanu wysokiego i stanu niskiego.

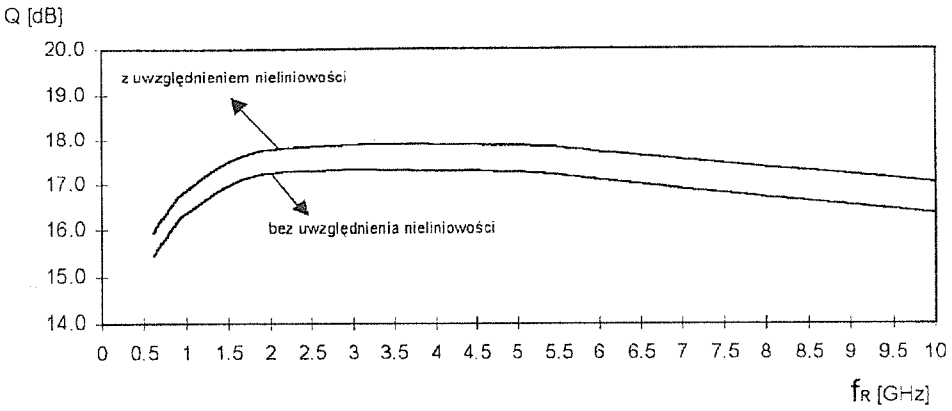
Wartość stosunku sygnału do szumu (S/N), w obecności przedwzmacniacza optycznego typu EDFA określono w następujący sposób [8]:

$$S/N = \frac{\left(\frac{\eta \lambda e}{hc} GP\right)^2}{\sigma^2_{\text{sig-sp}} + \sigma^2_{\text{sp-sp}}}, \tag{9}$$

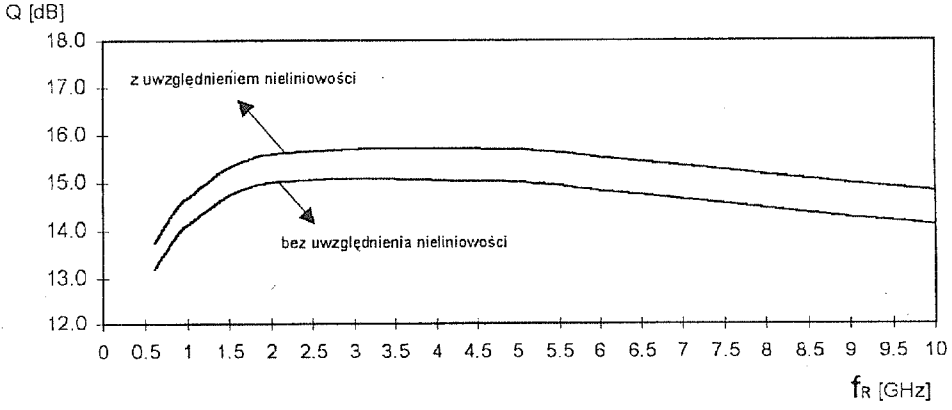
gdzie:
λ, η, e, h, c — jak we wzorze (4);
G — wzmocnienie wzmacniacza;
P — poziom mocy optycznej;
$\sigma^2_{\text{sp-sp}}$ — moc szumu zdudniania emisji spontanicznej;
$\sigma^2_{\text{sig-sp}}$ — moc szumu zdudnienia sygnału użytecznego z szumem emisji spontanicznej.

Na rys. 3, 4, 5, 6, 7 i 8 przedstawiono zmianę wartości parametru Q w funkcji f_R filtru odbiornika dla systemu światłowodowego długości $L = 100$ km, 120 km, 140 km, 160 km, 180 km i 200 km przy uwzględnieniu nieliniowości włókna optycznego i bez jego uwzględnienia. Przeprowadzone symulacje pozwalają stwierdzić, że w wyniku wzajemnego oddziaływania nieliniowego zjawiska samomodulacji fazy z dyspersją chromatyczną światłowodu następuje polepszenie wartości parametru Q w porównaniu z wynikami uzyskanymi dla światłowodu bez uwzględnienia nieliniowości. Wzrost wartości omawianego parametru (ΔQ) przy $f_R = 5$ GHz pokazano w tabeli 1. W tabeli 2 przedstawiono porównanie wartości stosunku sygnału do szumu elektrycznego (S/N) dla poziomu niskiego i wysokiego, zdefiniowanego zależnością (9), dla światłowodu z uwzględnieniem nieliniowości i bez jej uwzględnienia. Przyjęto, że pasmo elektryczne odbiornika $\Delta\nu_{\text{el}} = 5$ GHz.

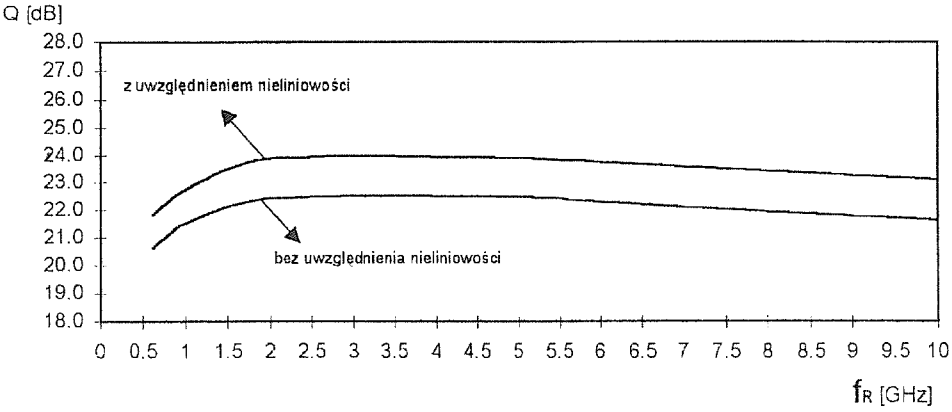
Analizując wyniki prezentowane w tab. 2 można powiedzieć, że dla stanu wysokiego stosunek sygnału do szumu jest większy w przypadku uwzględnienia nieliniowego zjawiska samomodulacji fazy występującego w światłowodzie niż w przypadku jego nie uwzględnienia. Podobne wyniki otrzymano dla transmisji typu „*dimension*" przedstawionej w [9, 10]. Zauważalny wzrost parametru ΔQ i parametru $\Delta S/N_{\text{s.w.}}$, który jest różnicą między $S/N_{\text{s.w.}}$ dla przypadku, w którym uwzględniono



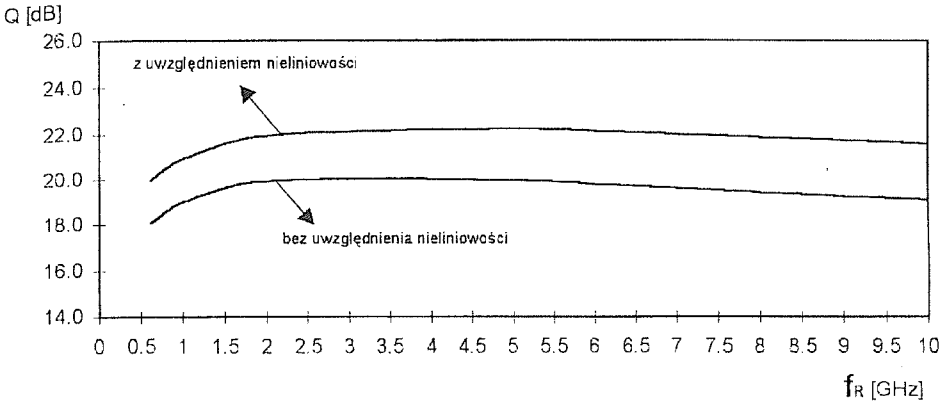
Rys. 3. Porównanie wartości parametru Q dla światłowodu z uwzględnieniem nieliniowości i bez uwzględnienia nieliniowości przy $L = 100$ km



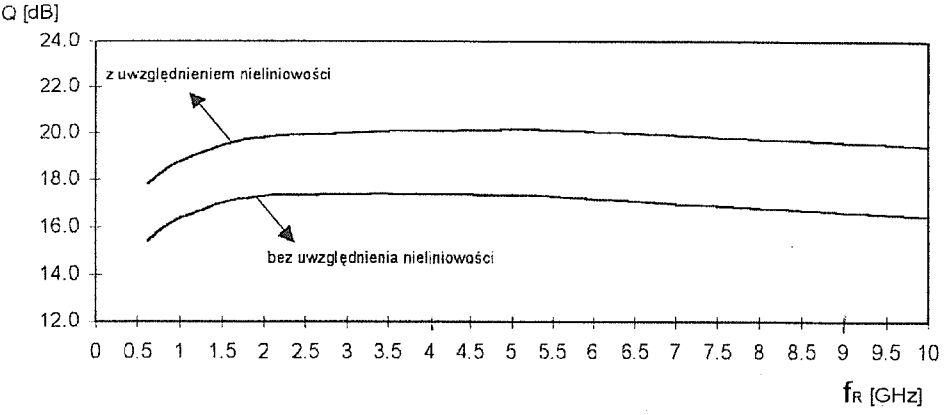
Rys. 4. Porównanie wartości parametru Q dla światłowodu z uwzględnieniem nieliniowości i bez uwzględnienia nieliniowości przy $L = 120$ km



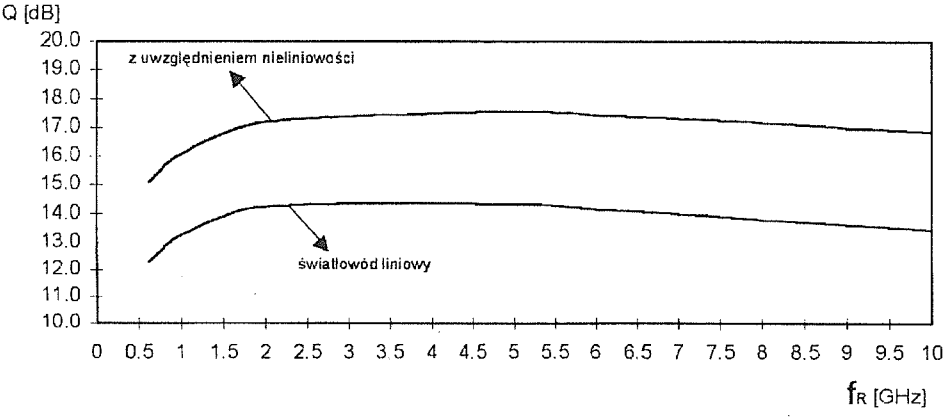
Rys. 5. Porównanie wartości parametru Q dla światłowodu z uwzględnieniem nieliniowości i bez uwzględnienia nieliniowości przy $L = 140$ km



Rys. 6. Porównanie wartości parametru Q dla światłowodu z uwzględnieniem nieliniowości i bez uwzględnienia nieliniowości przy $L = 160$



Rys. 7. Porównanie wartości parametru Q dla światłowodu z uwzględnieniem nieliniowości i bez uwzględnienia nieliniowości przy $L = 180$ km



Rys. 8. Porównanie wartości parametru Q dla światłowodu z uwzględnieniem nieliniowości i bez uwzględnienia nieliniowości przy $L = 200$ km

Tabela 1

Wartość ΔQ przy różnych długościach toru światłowodowego

Długość traktu światłowodowego [km]	ΔQ [dB]
100	0,62
120	0,67
140	0,45
160	2,31
180	2,81
200	3,21

Tabela 2

Wartość stosunku sygnału do szumu dla światłowodów z uwzględnieniem nieliniowości i bez jej uwzględnienia; $S/N_{s.n.}$ — stosunek S/N dla stanu niskiego, $S/N_{s.w.}$ — stosunek S/N dla stanu wysokiego

L [km]	światłowód z uwzględnieniem nieliniowości		światłowód bez uwzględnienia nieliniowości	
	$S/N_{s.n.}$ [dB]	$S/N_{s.w.}$ [dB]	$S/N_{s.n.}$ [dB]	$S/N_{s.w.}$ [dB]
100	46,71	50,55	47,02	50,29
120	42,72	46,46	43,08	46,18
140	58,90	62,74	59,31	62,07
160	54,44	58,98	55,54	57,97
180	50,58	54,91	51,80	53,88
200	46,84	50,78	48,09	49,73

zjawisko samomodulacji faz i bez jego uwzględnienia, wraz ze zwiększeniem długości traktu światłowodowego wynika z faktu, że nieliniowe przesunięcie fazy (φ_{NL}) wzrasta wraz z odległością [3].

Wzrostowi stosunku sygnału do szumu dla stanu wysokiego przy uwzględnieniu nieliniowości światłowodu towarzyszy równoczesny spadek wartości tego stosunku dla stanu niskiego w porównaniu ze światłowodem, w którym nieliniowość nie została uwzględniona. Wskazuje to na wzrost różnicy poziomów mocy optycznej między tymi stanami, czyli na wzrost wartości współczynnika ekstynkcji. **Nieliniowe zjawisko samomodulacji fazy powoduje wyraźny wzrost odstępu między wartościami stosunku sygnału do szumu dla stanu wysokiego i niskiego.**

Poszerzenie widma impulsów na skutek zjawiska samomodulacji fazy związane jest z czasową zależnością φ_{NL}. Zmiana fazy powoduje, że częstotliwość wzdłuż impulsów różni się od częstotliwości środkowej. Różnica ta jest równa [2]:

$$\delta\omega(t) = \frac{\partial \varphi_{NL}}{\partial t} = \frac{\partial |A(0,\ t)|^2}{\partial t} \cdot \frac{L_{sk}}{L_{NL}} \tag{10}$$

gdzie:
$A(0, t)$ — amplituda w $L = 0$ km;
L_{sk} — jest tzw. skuteczną odległością określoną w następujący sposób:

$$L_{sk} = \frac{1}{\alpha}(1 - e^{-\alpha L}) \tag{11}$$

gdzie α jest tu stałą tłumienia.

Powstający na skutek samomodulacji fazy migotanie rośnie wraz ze wzrostem odległości, co powoduje, że nowe składowe spektralne są tworzone w sposób ciągły podczas rozchodzenia się impulsów. Składowe te poszerzają widmo w stosunku do tego, które było na początku [2]. W wyniku wzajemnego oddziaływania poszerzenia widma i zjawiska konwersji modulacji częstotliwości na modulację amplitudy, powodowaną przez dyspersję chromatyczną, następuje wzrost wartości parametru Q i wartości stosunku sygnału do szumu dla stanu wysokiego.

WNIOSKI

W artykule zaprezentowano wyniki analizy numerycznej wpływu zjawiska samomodulacji fazy na wybrane parametry systemu transmisji opartej na zjawisku konwersji modulacji częstotliwości na modulację amplitudy. Na podstawie przeprowadzonych badań stwierdzono, że w wyniku wzajemnego oddziaływania poszerzenia widma sygnału powodowanego przez samomodulację fazy i zjawiska konwersji modulacji częstotliwości na modulację amplitudy następują wyraźne zmiany w wartości parametru Q i stosunku sygnału do szumu. Obecność nieliniowego zjawiska samomodulacji fazy powoduje wzrost wartości parametru Q oraz wartości stosunku sygnału do szumu dla stanu wysokiego przy jednoczesnym spadku wartości stosunku sygnału do szumu dla stanu niskiego; co wskazuje na wzrost wartości współczynnika ekstynkcji. Różnica między wartościami parametru Q i wartościami stosunku sygnału do szumu dla światłowodu dla którego uwzględniono nieliniowość i dla światłowodu bez jej uwzględnienia wzrasta wraz ze wzrostem długości toru optotelekomunikacyjnego.

BIBLIOGRAFIA

1. K. Perlicki: *Zastosowanie konwersji modulacji częstotliwości na modulację amplitudy w światłowodowych systemach transmisji o dużych przepływnościach.* Kwartalnik Elektroniki i Telekomunikacji 1998, t. 44, z. 2, ss. 259
2. G.P. Agrawal: *Nonlinear fiber optics.* Academic Press Inc., New York 1989
3. A. Majewski: *Nieliniowa optyka światłowodowa.* Wydawnictwa Politechniki Warszawskiej, Warszawa 1993
4. C. Kurtzke, A. Gnauck: *Operating principle of in-ine amplified dispersion supported transmission.* Electronics Letters. 1993, vol. 29, nr 22, 1969

5. M.S. Rodan: *Systemy telekomunikacji analogowej i cyfrowej.* WNT, Warszawa 1983
6. G.P. Agrawal, M.J. Potasek: *Nonlinear pulse distorsion in single-mode optical fibres at zero-dispersion wavelength.* Physical Review A. 1986, vol. 33, nr 3, p. 1765
7. G.P. Agrawal: *Fiber-optical communication systems.* John Wiley & Sons, Inc., New York, 1992
8. J. Siuzdak: *Wstęp do współczesnej telekomunikacji światłowodowej.* WKiŁ, Warszawa 1997
9. C. Kurtzke, S. Kindt, K. Petermann: *Impact of residual amplitude modulation on the performance of dispersion supported and dispersion-mediated nonlinearity — enhanced transmission.* Electronics Letters. 1994, vol. 30, nr 12, 988
10. S. Kindt, C. Kurtzke: Design guidelines for high-capacity standard-fiber intercity communication systems based on dispersion-mediated nonlinearity-enhanced transmission. Journal of Optical Communications, 1996, vol. 17, nr 2, p. 65

K. PERLICKI

INPACT OF SELF-PHASE MODULATION ON FM-AM CONVERSION SUPPORTED TRANSMISSION

Summary

The author investigate the impact of self-phse modulation on the FM-AM conversion supported transmission with a binary optical signal at the receiver. The propagation of optical field through a linear and nonlinear fiber is described by nonlinear Schrödinger equation, which is solved numerically by the split-step Fourier method. The computer simulations are caried out for a system with the following characteristics. The transmitted optical signal is modulated by an augmented PRBS of length $2^7 - 1$. The fiber dispersion is 17 ps/nm · km, the loss is 0,2 dB/km and the nonlinear coefficient is 2,17 W^{-1} km^{-1}, the fiber length is from 100 km to 200 km. Optical in-line amplifier is inserted at 100 km. The gain of the amplifier is chosen to compensation of the fiber loss. A first order Butterworth filter at the receiver is used. In particular impact of self-phase modulation on signal to noise ratio is investigated. Results of numerical simulations show that interplay between dispersion of standard fiber and self-phase modulation increases signal to noise ratio.

Key words: optical communication, frequency modulation to amplitude modulation conversion, self-phase modulation.

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI, 1998, 44, z. 3, ss. 345–364

Nieizotermiczne parametry małosygnałowe elementów półprzewodnikowych

WITOLD JERZY STEPOWICZ

Wydział Elektroniki, Telekomunikacji i Informatyki
Politechnika Gdańska

JANUSZ ZARĘBSKI

Wydział Elektryczny, Wyższa Szkoła Morska w Gdyni

Otrzymano 1998.07.02

Autoryzowano 1998.07.21

Zjawiska termiczne wpływają w istotny sposób na własności i parametry elementów półprzewodnikowych oraz układów scalonych. Dotyczy to zarówno ich modeli stałoprądowych jak i modeli dla prądu zmiennego. W sytuacjach, kiedy zjawiska termiczne odgrywają istotną rolę, modele elementów różnią się zasadniczo, zarówno jakościowo jak i ilościowo, od ich odpowiedników znanych z klasycznej teorii nie uwzględniającej tych zjawisk. Z kolei, postać modeli elementów wpływa na parametry robocze, określające własności funkcjonalne, układów elektronicznych skonstruowanych z wykorzystaniem tych elementów.

W pracy przedstawiono metodę wyznaczania nieizotermicznych modeli małosygnałowych elementów półprzewodnikowych i pokazano takie modele dla wybranych elementów, uzyskane na podstawie tej metody. Na kilku przykładach zilustrowano wpływ nieizotermicznej postaci modeli małosygnałowych elementów na własności układów, a także przytoczono inne konsekwencje i korzyści wynikające ze stosowania takich modeli.

Słowa kluczowe: elementy półprzewodnikowe, zjawiska termiczne, modele małosygnałowe.

1. WSTĘP

W klasycznej teorii elementów półprzewodnikowych wpływ temperatury na ich własności i parametry przeważnie sprowadza się do rozważenia wpływu temperatury otoczenia. Jednak w rzeczywistości temperatura wnętrza elementu, nazywana też

często temperaturą złącza, może różnić się znacznie od temperatury otoczenia. Wynika to ze zjawiska samonagrzewania polegającego na zamianie energii elektrycznej wydzielającej się w elemencie na ciepło i na nieidealnych warunkach odprowadzania ciepła do otoczenia. W wyniku tego zjawiska w stanie termicznie ustalonym następuje wzrost temperatury elementu ponad temperaturę otoczenia o składnik zależny od iloczynu mocy elektrycznej rozpraszany w elemencie i ważnego jego parametru termicznego — rezystancji termicznej, świadczącej o efektywności odprowadzania ciepła z elementu do otoczenia. W układach scalonych wskutek umieszczenia elementów w bezpośrednim sąsiedztwie na wspólnym podłożu należy uwzględnić wzajemne oddziaływanie termiczne między nimi. Oczywiście, takie wzajemne oddziaływanie może mieć miejsce także w układach hybrydowych lub układach montowanych klasycznie o dużej gęstości upakowania.

Wpływ zjawisk termicznych na charakterystyki statyczne elementów prowadzi w układach elektronicznych do zmiany współrzędnych punktu pracy, w stosunku do wyznaczonych na podstawie zależności klasycznych, bez uwzględnienia samonagrzewania. Charakterystyki (modele) statyczne elementów uwzględniające zjawisko samonagrzewania są nazywane w niniejszej pracy charakterystykami (modelami) nieizotermicznymi, w przeciwieństwie do zależności klasycznych, które można określić jako izotermiczne.

Przy pobudzeniu elementu przebiegiem zmiennym (typowo jest to prąd lub napięcie) powstaje w nim składowa zmienna temperatury wnętrza. Dotyczy to, oczywiście, zarówno dużego jak i małego sygnału. W niniejszej pracy są rozważane modele małosygnałowe, które analogicznie do sytuacji opisanej powyżej są tu nazywane nieizotermicznymi i to samo określenie dotyczy parametrów w nich występujących. Modele takie są analizowane w literaturze od dawna — należy tu przede wszystkim wymienić pionierską pracę Muellera z 1964 roku [1], ale ostatnio ich znaczenie zdecydowanie zyskuje na ważności, między innymi wskutek zmniejszania rozmiarów elementów w układach scalonych, czy też wskutek rozwoju takich technologii jak SOI (krzem na izolatorze), gdzie zastosowanie podłoża biernego powoduje pogorszenie warunków odprowadzania ciepła z układu do otoczenia, np. [2]. W przypadku nieizotermicznej analizy małosygnałowej parametrem termicznym elementu jest jego przejściowa impedancja termiczna (w dziedzinie czasu) lub zespolona impedancja termiczna (w dziedzinie częstotliwości).

Konsekwencją uwzględnienia w modelach małosygnałowych zjawisk termicznych jest to, że w zakresie małych częstotliwości, gdzie inercja elektryczna elementów półprzewodnikowych ze złączem pn wywołana przez pojemności takie, jak pojemność barierowa i dyfuzyjna, jest do pominięcia, z powodu efektów termicznych nieizotermiczne parametry małosygnałowe elementów stają się wielkościami zespolonymi. Tak więc w rozpatrywanym przypadku inercja termiczna prowadzi do inercji elektrycznej, co ujawnia się poprzez pojawienie się pojemności i indukcyjności w małosygnałowych schematach zastępczych tych elementów.

W niniejszej pracy w rozdziale 2 przedstawiono sposób tworzenia nieizotermicznych małosygnałowych modeli elementów połprzewodnikowych i układów scalo-

nych. W rozdz. 3 pokazano postać takich modeli dla wybranych elementów: diody ze złączem pn (rozdz. 3.1), indywidualnego tranzystora bipolarnego (rozdz. 3.2) oraz scalonego tranzystora bipolarnego (rozdz. 3.3), a także dla indywidualnego tranzystora MOS (rozdz. 3.4). W rozdz. 4 pokazano na wybranych przykładach wpływ postaci nieizotermicznych modeli nieizotermicznych na: szumy typu 1/f diody ze złączem pn (rozdz. 4.1), na wzmocnienie napięciowe układu wzmacniacza z tranzystorem bipolarnym (rozdz. 4.2) oraz na własności prądowego źródła scalonego (rozdz. 4.3). Omówiono też (rozdz. 4.4) sposób wyznaczania rezystancji termicznej tranzystora bipolarnego na podstawie pomierzonej jego nieizotermicznej małosygnałowej rezystancji wyjściowej. Przedstawione w pracy rozważania dotyczą zakresu małych częstotliwości, gdzie w modelach elementów można pominąć inercję elektryczną wnoszoną przez typowe pojemności jak pojemność barierowa, dyfuzyjna oraz pojemność struktury MOS.

Niniejsza publikacja stanowi przede wszystkim podsumowanie prac dotyczących nieizotermicznych parametrów małosygnałowych elementów półprzewodnikowych i układów scalonych, prowadzonych przez autorów w latach 1993 – 1997. W pracy poruszone są także nowe zagadnienia, wcześniej przez autorów nie publikowane. Dlatego też nie zamieszczono tu wielu szczegółów wyprowadzeń poszczególnych podawanych zależności, odsyłając zainteresowanych do wcześniejszych publikacji. Także wykaz bibliografii ograniczono do zasadniczych pozycji, a szerszy przegląd literatury na temat określonych zagadnień można znaleźć w cytowanych pozycjach.

2. METODYKA TWORZENIA NIEIZOTERMICZNYCH MODELI MAŁOSYGNAŁOWYCH ELEMENTÓW PÓŁPRZEWODNIKOWYCH

W celu sformułowania nieizotermicznego małosygnałowego modelu indywidualnego elementu półprzewodnikowego należy [3, 4, 5]:

a) podać jego nieizotermiczne charakterystyki statyczne, tj. zależności, w których uwzględniono wpływ temperatury wnętrza t_J na parametry modelu. Dla dwójnika jest to zależność typu $i(u, t_J)$, dla czwórnika są to dwie zależności, np. w postaci równań admitancyjnych $i_{WE}(u_{WE}, u_{WY}, t_J)$ oraz $i_{WY}(u_{WE}, u_{WY}, t_J)$. Jak widać, temperatura wnętrza staje się jedną ze zmiennych w modelu,

b) podać model termiczny elementu, tj. zależność temperatury wnętrza od mocy elektrycznej wydzielającej się w elemencie,

c) przeprowadzić lineryzację nieizotermicznych charakterystyk statycznych (z punktu a) względem prądu (prądów), napięcia (napięć) i temperatury wnętrza,

d) wyeliminować temperaturę wnętrza z liniowych równań otrzymanych w punkcie c poprzez wykorzystanie modelu termicznego z punktu b.

Dla oddziaływających termicznie na siebie elementów układzie scalonym [4] trzeba w punkcie a) sformułować $2N + M$ równań opisujących ich nieizotermiczne charakterystyki statyczne, gdzie N jest liczbą czwórników, zaś M — dwójników. Trzeba też podać $M + N$ równań opisujących ich modele termiczne. Jeżeli założy się (co uczyniono w niniejszej pracy) pobudzenie elementu małym sygnałem o postaci

harmonicznej, to wówczas należy traktować w równaniach liniowych, otrzymanych według punktu c, „mały sygnał" jako małe amplitudy zespolone. Dotyczy to zarówno prądów i napięć, jak też temperatury wnętrza i mocy elektrycznej. Dla przykładu, T_j oznacza tu amplitudę (zespoloną) temperatury wnętrza, T_J składową stałą temperatury wnętrza w punkcie pracy, zaś t_J wielkość chwilową tej temperatury. Aby uniknąć niejednoznaczności, jako oznaczenie czasu przyjęto tu symbol τ.

W przeprowadzanej analizie model termiczny jest określony przez zespoloną impedancję termiczną, gdyż dotyczy ona dziedziny częstotliwości. Zespolona impedancja termiczna jest transformatą Laplace'a odpowiedzi termicznej elementu na pobudzenie mocą elektryczną w postaci funkcji Diraca (lub transformatę pochodnej odpowiedzi na uskok jednostkowy). Często przyjmuje się, np. [6], odpowiedź termiczną elementu na pobudzenie uskokiem jednostkowym mocy, nazywaną jego przejściową impedancją termiczną $Z_{\text{th}}(\tau)$, w postaci funkcji

$$Z_{\text{th}}(\tau) = R_{\text{th}} \cdot \sum_n a_n \cdot \left[1 - \exp\left(-\frac{\tau}{\tau_n}\right)\right], \tag{1}$$

gdzie R_{th} jest rezystancją termiczną elementu, suma składników a_n jest równa jedności, zaś τ_n są termicznymi stałymi czasowymi. Liczba składników sumy we wzorze (1) odpowiada w przybliżeniu liczbie dających się wyróżnić części w konstrukcji elementu: *chip* półprzewodnikowy, metalowa podstawa, obudowa, radiator itd.

Przy przyjęciu opisu $Z_{\text{th}}(\tau)$ według wzoru (1), zespolona impedancja termiczna $Z_T(j\omega)$ jest dana zależnością

$$Z_T(j\omega) = R_{\text{th}} \sum_n \frac{a_n}{1 + j\omega \cdot \tau_n}. \tag{2}$$

W przypadku wzajemnego oddziaływania termicznego elementów, jak to ma miejsce w układzie scalonym, w dziedzinie czasu o przyroście temperatury wnętrza i-tego elementu wskutek wydzielania się mocy elektrycznej w j-tym elemencie mówi wzajemna przejściowa impedancja termiczna $Z_{\text{thij}}(\tau)$, w odróżnieniu od poprzedniej sytuacji, gdy mowa była o własnej przejściowej impedancji termicznej i-tego elementu $Z_{\text{thii}}(\tau)$. Założono, że postać funkcyjna zależności wzajemnej impedancji termicznej $Z_{\text{thij}}(\tau)$ od czasu jest identyczna jak dla własnej impedancji termicznej $Z_{\text{thii}}(t)$ — wzór (1), zatem postać wzorów na wzajemne zespolone impedancje termiczne $Z_{\text{Tij}}(j\omega)$ jest też identyczna, jak dla impedancji własnej $Z_{\text{Tii}}(j\omega)$ według wzoru (2). Zależność zespolonej amplitudy składowej zmiennej temperatury wnętrza T_j od mocy elektrycznej jest następująca

$$T_j(j\omega) = Z_T(j\omega) \cdot P(j\omega). \tag{3}$$

Składowa zmienna mocy elektrycznej P_d dla diody ma postać

$$P_d(j\omega) = I_D \cdot U_d + I_d \cdot U_D, \tag{4a}$$

gdzie symbole prądu i napięcia I_D, U_D dotyczą współrzędnych punktu pracy, zaś I_d, U_d są amplitudami zespolonymi, zgodnie z uprzednio podanym ustaleniem.

Dla tranzystora bipolarnego przyjęto, że w zakresie aktywnym normalnym składowa zmienna mocy P_b wyraża się wzorem

$$P_b(j\omega) = I_C \cdot U_{ce} + I_c \cdot U_{CE} \tag{4b}$$

zaś dla tranzystora polowego

$$P_p(j\omega) = I_D \cdot U_{ds} + I_d \cdot U_{DS}. \tag{4c}$$

3. NIEIZOTERMICZNE PARAMETRY MAŁOSYGNAŁOWE WYBRANYCH ELEMENTÓW PÓŁPRZEWODNIKOWYCH

W rozdziale tym przedstawiono nieizotermiczne modele małosygnałowe wybranych elementów półprzewodnikowych. Otrzymano je na podstawie procedury opisanej w poprzednim rozdziale, gdzie też podano postać modeli termicznych używanych w analizie dla wszystkich rozpatrywanych elementów. Jak wynika z punktu a) podanej procedury, ostateczna postać modelu małosygnałowego zależy od przyjętego nieizotermicznego modelu statycznego elementu.

3.1. DIODA ZE ZŁĄCZEM p-n

Poniżej pokazano nieizotermiczny model małosygnałowy idealnej diody ze złączem pn [3, 7]. Do rozważań przyjęto diodę idealną, o powszechnie znanej postaci prądowo-napięciowych izotermicznych charakterystyk statycznych

$$i_D(u_D) = I_s \cdot \left(\exp \frac{u_D}{U_T} - 1 \right). \tag{5}$$

Aby otrzymać charakterystyki nieizotermiczne, należy uwzględnić wpływ temperatury na parametry występujące w modelu opisanym równaniem (5): są to potencjał termiczny U_T oraz prąd nasycenia I_s. Jeżeli przyjmie się, że potencjał termiczny zależy liniowo od temperatury, zaś prąd nasycenia wykładniczo, to ostateczna postać nieizotermicznych charakterystyk statycznych diody jest następująca (z pominięciem 1 w nawiasie w zależności (5))

$$i_D(u_D, t_J) = A \cdot \exp \frac{u_D - U_{go}}{\frac{k}{q} \cdot t_J}, \tag{6}$$

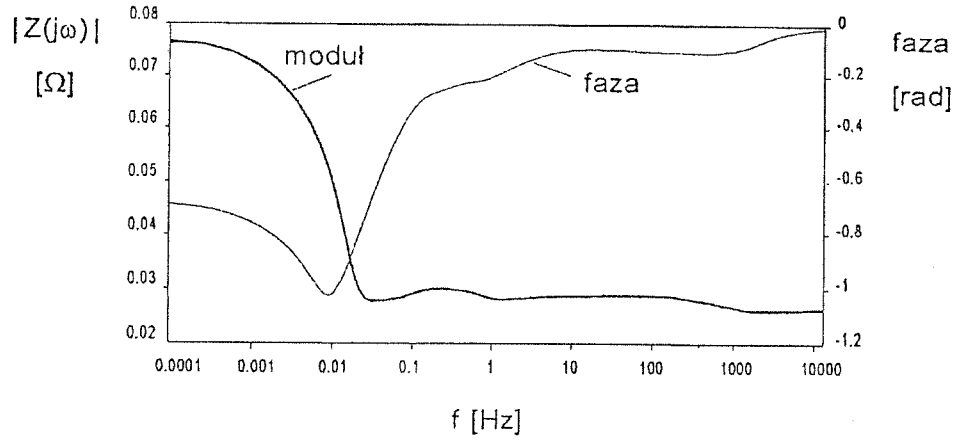
gdzie A oznacza stałą niezależną od temperatury, zaś $U_{go} = 1{,}21$ eV.

Po linearyzacji powyższych charakterystyk oraz po przekształceniach, zgodnie z metodyką opisaną w rozdz. 2, ostatecznie impedancja elektryczna diody jest równa [3, 7]

$$Z(j\omega) = \frac{r_{iz} - \lambda \cdot U_D \cdot Z_T(j\omega)}{1 + \lambda \cdot Z_T(j\omega) \cdot I_D}, \tag{7}$$

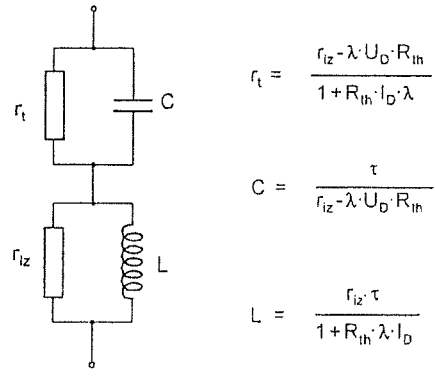
gdzie $\lambda = -\frac{\partial u_D}{\partial t_J} = \frac{U_{go} - u_D}{t_J}$, zaś $r_{iz} = \frac{U_T}{I_D}$.

Na rys. 1 przedstawiono przykładowy przebieg modułu i fazy impedancji opisanej wzorem (7) dla wartości parametrów elektrycznych i termicznych odpowiadających diodzie małej mocy [7]. Jak widać, w rozpatrywanym przypadku inercja termiczna prowadzi do inercji elektrycznej w zakresie od zera do częstotliwości rzędu 0,02 Hz, na co w istotny sposób wpływają przyjęte do obliczeń wartości termicznych stałych czasowych.



Rys. 1. Zależność od częstotliwości modułu oraz kąta fazowego impednacji elektrycznej diody obliczonej wg wzoru (7)

Dla zakresu częstotliwości, w którym zespolona impedancja termiczna diody może być przybliżona jednym składnikiem sumy z równania (2) można podać reprezentację układową impedancji elektrycznej diody. Jest ona pokazana na rys. 2 wraz ze wzorami opisującymi elementy pokazanego tam układu zastępczego diody.



Rys. 2. Małosygnałowy nieizotrmiczny układ zastępczy diody w przypadku jednej dominującej termicznej stałej czasowej τ

3.2. TRANZYSTOR BIPOLARNY

W celu wyprowadzenia nieizotermicznego modelu małosygnałowego tranzystora bipolarnego założono [8] następującą postać jego nieizotermicznych charakterystyk statycznych, słusznych dla zakresu aktywnego normalnego, dla konfiguracji o wspólnym emiterze

$$u_{BE}(i_B, t_J) = U_T(t_J) \cdot \ln \frac{i_B}{I_{BS}} + U_{go} + i_B \cdot r(t_J) \tag{8a}$$

$$i_C(u_{CE}, i_B, t_J) = \beta_0 \cdot [1 + a \cdot (t_J - t_0)] \cdot \left(1 + \frac{u_{CE}}{U_E}\right) \cdot i_B. \tag{8b}$$

W charakterystykach powyższych została uwzględniona wykładnicza zależność prądu nasycenia złącza emiterowego od temperatury (występujący w zależności (8a) parametr I_{BS} nie zależy już od temperatury) oraz liniowa zależność współczynnika wzmocnienia prądowego β od temperatury (β_0 oznacza wartość tego współczynnika w temperaturze doniesienia t_0, zaś „a" to temperaturowy współczynnik zmian β). Występujący w tych charakterystykach potencjał termiczny U_T zależy liniowo od temperatury i taką samą zależność przyjęto dla rezystancji zastępczej r obszarów emitera i bazy. We wzorze (8b) U_E oznacza napięcie Early'ego.

Postępując zgodnie z procedurą opisaną w rozdz. 2 otrzymano następujące wzory na elementy macierzy $[h_e]$ tranzystora [8]

$$h_{11e} = \left(\frac{U_T}{I_B} + r\right) \cdot \left[1 - \frac{I_C \cdot \lambda_t \cdot Z_T \cdot U_{CE} \cdot (1 + a \cdot \Delta T_J)}{(I_B \cdot r + U_T) \cdot M}\right] \tag{9a}$$

$$h_{12e} = -\frac{1}{M} \cdot \lambda_t \cdot Z_T \cdot U_{CE} \cdot (1 + a \cdot \Delta T_J) \cdot \left[\frac{I_C}{U_{CE}} + \frac{I_C}{U_E \cdot \left(1 + \frac{U_{CE}}{U_E}\right)}\right] \tag{9b}$$

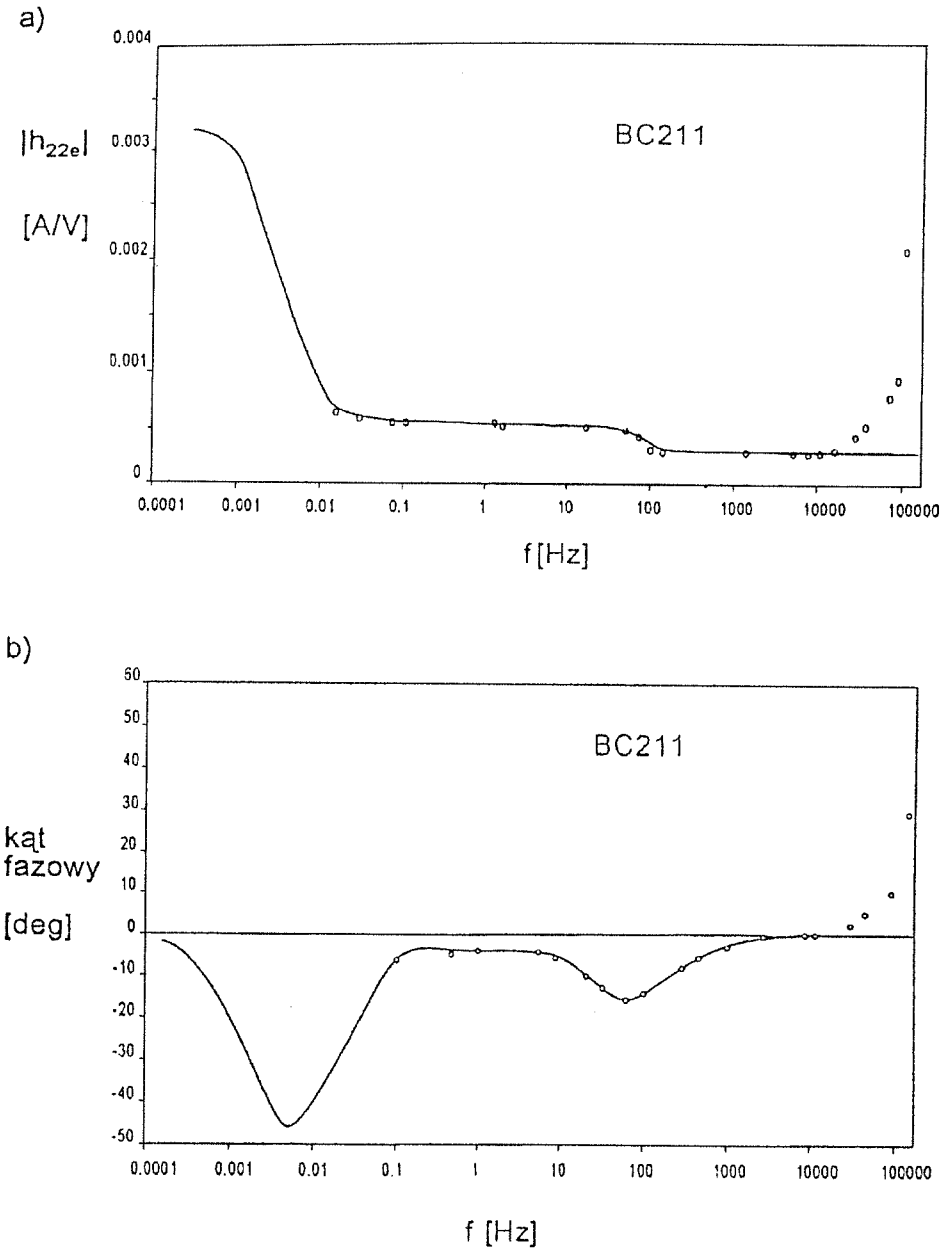
$$h_{21e} = \frac{1}{M} \cdot \beta_0 \cdot (1 + a + \Delta T_J)^2 \cdot \left(1 + \frac{u_{CE}}{U_E}\right) \tag{9c}$$

$$h_{22e} = \frac{1}{M} \cdot \left[I_C^2 \cdot a \cdot Z_T + \frac{I_C \cdot (1 + a \cdot \Delta T_J)}{U_E \cdot \left(1 + \frac{U_{CE}}{U_E}\right)}\right], \tag{9d}$$

gdzie

$$M = 1 + a \cdot \Delta T_J \cdot \left(1 - \frac{Z_T}{R_{th}}\right)$$

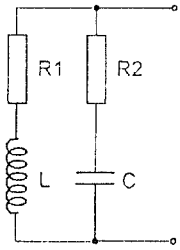
$$\lambda_t = -\frac{\partial u_{BE}}{\partial t_J} = \frac{U_{go} - U_{BE}}{T_J}.$$



Rys. 3. Zależność od częstotliwości modułu (rys. 3a) i kąta fazowego (rys. 3b) nieizotermicznej admitancji wyjściowej tranzystora bipolarnego

W pracy [9] pokazano pomierzone i obliczone zależności od częstotliwości modułu i kąta fazowego elementów macierzy $[h]_c$ dla tranzystora małej mocy BC211. Tu, dla przykładu, na rys. 3 pokazano zależność modułu (rys. 3a) i kąta fazowego (rys. 3b) parametru h_{22c}, tj. nieizotermicznej admitancji wyjściowej [9, 10].

Podobnie jak dla diody, można dla zakresu częstotliwości, w którym dominuje jedna termiczna stała czasowa podać reprezentację układową elementów macierzy $[h]_c$. Na rys. 4 pokazano [9, 10] układ zastępczy odpowiadający admitancji h_{22c}, wraz ze wzorami opisującymi elementy tego układu.



$$R1 = I_C \cdot \left(\frac{1 + a \cdot \Delta T}{U_E \cdot \left(1 + \frac{U_{CE}}{U_E}\right)} + I_C \cdot a \cdot R_{th} \right)$$

$$R2 = \frac{I_C}{U_E \cdot \left(1 + \frac{U_{CE}}{U_E}\right)}$$

$$C = \frac{I_C \cdot (1 + a \cdot \Delta T) \cdot \tau}{U_E \cdot \left(1 + \frac{U_{CE}}{U_E}\right)}$$

$$L = \frac{(1 + a \cdot \Delta T) \cdot \tau}{I_C \cdot \left(\frac{1 + a \cdot \Delta T}{U_E \cdot \left(1 + \frac{U_{CE}}{U_E}\right)} + I_C \cdot a \cdot R_{th} \right)}$$

Rys. 4. Małosygnałowy nieizotermiczny schemat zastępczy odpowiadający admitancji wyjściowej tranzystora bipolarnego w przypadku jednej dominującej stałej czasowej τ

3.3. SCALONY TRANZYSTOR BIPOLARNY

W niniejszym rozdziale przedstawiono nieizotermiczny model małosygnałowy scalonego tranzystora bipolarnego [4, 5]. Dla uproszczenia założono, że oddziałuje on (dalej oznaczany symbolem T1) termicznie tylko z jednym tranzystorem bipolarnym (T2). Założono też, że tranzystory te nie są połączone ze sobą elektrycznie. Wykazano [4, 5], że model małosygnałowy tranzystora T1 jest wówczas dany przez układ równań

$$\begin{bmatrix} U_{be1} \\ I_{c1} \\ U_{be2} \\ I_{c2} \end{bmatrix} = [H]_c \cdot \begin{bmatrix} I_{b1} \\ U_{ce1} \\ I_{b2} \\ U_{ce2} \end{bmatrix} \tag{10}$$

Jak widać z zależności (10), model małosygnałowy tranzystora scalonego zawiera 16 parametrów, których definicje wynikają bezpośrednio z podanego zapisu macierzowego. Dla przykładu, poniżej podano wzór na nieizotermiczną admitancję wyjściową tranzystora T1, która jest opisana następującym wzorem

$$H_{22e} = \left.\frac{I_{c1}}{U_{ce1}}\right|_{I_{b1}, I_{b2}, U_{ce2}=0} =$$

$$= \frac{(G_{E1} + M_1 \cdot a_1 \cdot Z_{11} \cdot I_{C1}^2) \cdot (1 - M_2 P_2) + M_1 \cdot M_2 \cdot a_1 \cdot a_2 \cdot Z_{T12} \cdot Z_{T21} \cdot I_{C1}^2 \cdot I_{C2} \cdot U_{CE2}}{(1 - M_1 \cdot P_1) \cdot (M_2 \cdot P_2) - M_1 \cdot M_2 \cdot a_1 \cdot a_2 \cdot Z_{T12} \cdot Z_{T21} \cdot I_{C1} \cdot I_{C2} \cdot U_{CE1} \cdot U_{CE2}}, \quad (11)$$

gdzie: $P_1 = a_1 \cdot Z_{T11} \cdot I_{C1} \cdot U_{CE1}$,
$P_2 = a_2 \cdot Z_{T22} \cdot I_{C2} \cdot U_{CE2}$,
$M_1 = (1 + a_1 \cdot \Delta T_{J1})^{-1}$,
$M_2 = (1 + a_2 \cdot \Delta T_{J2})^{-1}$,
G_{E1} jest izotermiczną konduktancją wyjściową tranzystora T1, wynikającą z efektu Early'ego o postaci:

$$G_{E1} = \frac{\partial i_{C1}}{\partial u_{CE1}} = \frac{I_{C1}}{\left(1 + \frac{U_{CE1}}{U_{E1}}\right) \cdot U_{E1}}.$$

Pominięcie we wzorze (11) wzajemnego oddziaływania termicznego między tranzystorami prowadzi do wzoru (9d). Porównanie wyników obliczeń i pomiarów omawianego parametru dla tranzystora scalonego podano w pracach [4, 5].

Niektóre z elementów macierzy $[H]_e$ z zależności (10) wynikają jedynie ze wzajemnego oddziaływania termicznego między tranzystorami (przy braku połączenia elektrycznego między nimi). Dla przykładu parametr H_{14e} zdefiniowany następująco

$$H_{14} = \left.\frac{U_{bc1}}{U_{ce2}}\right|_{I_{b1}, I_{b2}, U_{ce1}=0} = Q_2 + \frac{Q_3 \cdot (R_8 + Q_6 \cdot R_7)}{1 - Q_7 \cdot R_7}, \quad (12)$$

gdzie

$$Q_2 = P_7 \cdot Z_{T22} \cdot I_{C2} + \frac{P_7 \cdot Z_{T22} \cdot U_{CE2}}{1 - P_{10} \cdot Z_{T22} \cdot U_{CE2}} \cdot (P_9 + P_{10} \cdot Z_{T22} \cdot I_{C2})$$

$$Q_3 = P_7 \cdot Z_{T21} \cdot U_{CE1} + \frac{P_7 \cdot Z_{T22} \cdot U_{CE2} \cdot P_{10} \cdot Z_{T21} \cdot U_{CE1}}{1 - P_{10} \cdot Z_{T22} \cdot U_{CE2}}$$

$$Q_6 = \frac{P_9 + P_{10} \cdot Z_{T22} \cdot I_{C2}}{1 - P_{10} \cdot Z_{T22} \cdot U_{CE2}}$$

$$Q_7 = \frac{P_{10} \cdot Z_{T21} \cdot U_{CE1}}{1 - P_{10} \cdot Z_{T22} \cdot U_{CE2}}$$

$$R_7 = \frac{P_5 \cdot Z_{T12} \cdot U_{CE2}}{1 - P_5 \cdot Z_{T11} \cdot U_{CE1}}$$

$$R_8 = \frac{P_5 \cdot Z_{T12} \cdot I_{C2}}{1 - P_5 \cdot Z_{T11} \cdot U_{CE1}}$$

$$P_5 = \frac{a_1 \cdot I_{C1}}{1 + a_1 \cdot \Delta T_{J1}}$$

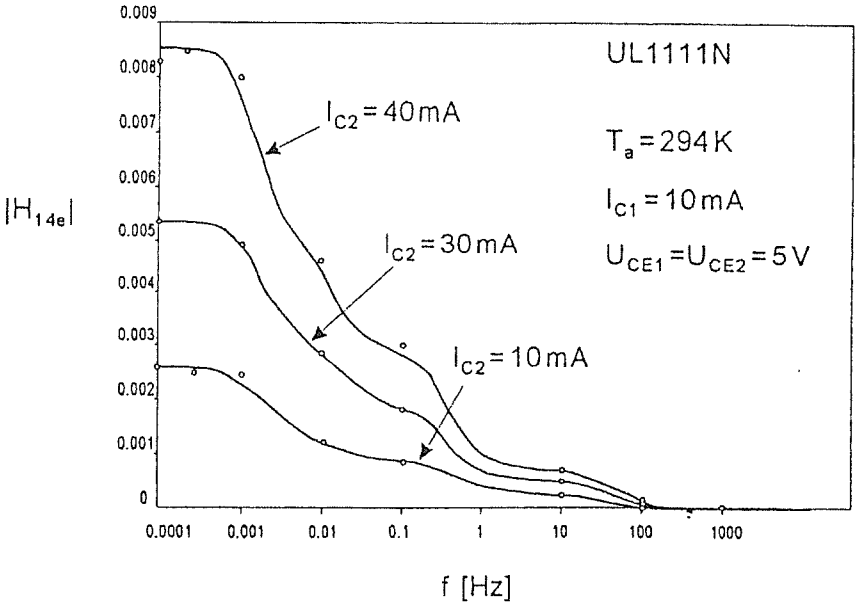
$$P_7 = \frac{U_{BE1} - U_{go}}{T_{J2}}$$

$$P_9 = \frac{1}{U_{E2}} \cdot \frac{I_{C2}}{1 + \frac{U_{CE2}}{U_{E2}}}$$

$$P_{10} = \frac{a_2 \cdot I_{C2}}{1 + a_2 \cdot \Delta T_{J2}}$$

mówi o przenoszeniu się sygnału elektrycznego z wyjścia T2 na wejście T1 wskutek sprzężenia termicznego tych elementów, zatem nie ma on odpowiednika w modelu izotermicznym.

Na rys. 5 pokazano pomierzony i obliczony przebieg modułu parametru H_{14e} dla dwóch tranzystorów z układu scalonego UL1111 [5], dla różnych wartości mocy



Rys. 5. Zależność od częstotliwości modułu parametru H_{14e} tranzystora bipolarnego

wydzielającej się w tranzystorze T2. Jak widać, w rozpatrywanym przypadku wzajemne oddziaływanie termiczne tranzystorów ujawnia się do częstotliwości rzędu kilkudziesięciu herców.

3.4. TRANZYSTOR POLOWY MOS

Dla tranzystora MOS przyjęto następującą postać jego nieizotermicznych charakterystyk statycznych w zakresie nasycenia

$$i_D(u_{GS}, u_{DS}, t_J) = \frac{B}{2}(t_J) \cdot [u_{GS} - U_p(t_J)]^2 \cdot \left(1 + \frac{u_{DS} - U_{DSO}}{U'_E}\right), \tag{13}$$

gdzie napięcie U_{DSO} odpowiada granicy między zakresem nasycenia i nienasycenia. W powyższym modelu uwzględniono wpływ temperatury na napięcie progowe U_p oraz na parametr B. Przyjęto następującą postać tych zależności

$$U_p(t_J) = U_p(t_0) - \lambda_G \cdot (t_J - t_0) \tag{14a}$$

$$B(t_J) = B_0(t_0) \cdot \left(\frac{t_0}{t_J}\right), \tag{14b}$$

gdzie $U_p(t_0)$ to napięcie progowe w temperaturze odniesienia t_0, zaś λ_G jest temperaturowym współczynnikiem zmian napięcia progowego. Parametr U'_E, odpowiednik napięcia Early'ego w tranzystorze bipolarnym, wynika z efektu modulacji napięciowej długości kanału. Oczywiście, dla idealnego tranzystora MOS można przyjąć, że prąd stały bramki nie płynie i dlatego do opisu tego elementu wystarcza tylko jedna zależność statyczna.

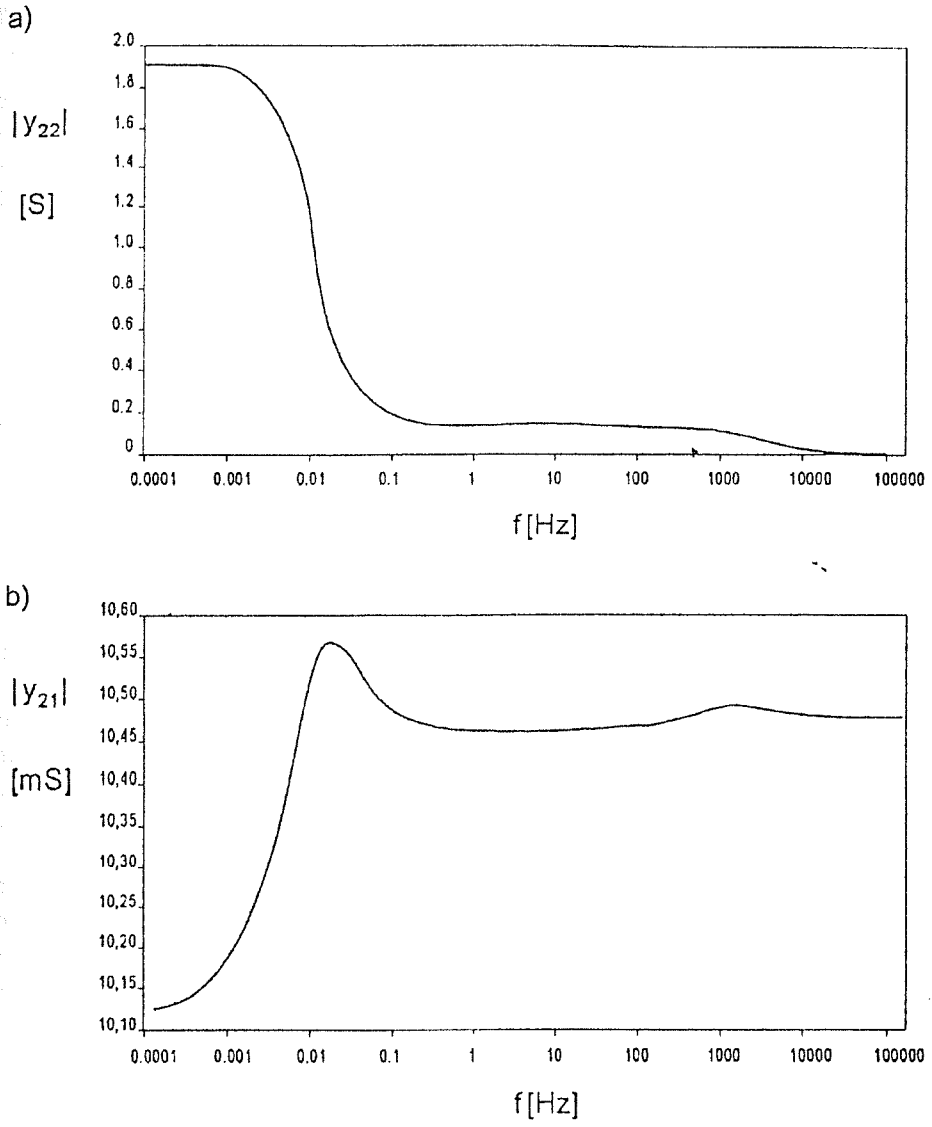
Procedura opisana w rozdziale 2 prowadzi do następującej postaci nieizotermicznych admitancji y_{22} oraz y_{21} tranzystora MOS w punkcie pracy określonym przez współrzędne (I_D, U_{DS}, T_J)

$$y_{22} = I_D \cdot \frac{\dfrac{1}{U'_E} \cdot \dfrac{1}{1 + \dfrac{U_{DS} - U_{DSO}}{U'_E}} + \gamma \cdot Z_T}{1 - Z_T \cdot I_D \cdot U_{DS} \cdot \gamma} \tag{15a}$$

$$y_{21} = \frac{[2I_D \cdot B(T_J)]^{1/2}}{1 - Z_T \cdot I_D \cdot U_{DS} \cdot \gamma} \tag{15b}$$

gdzie $\gamma = \dfrac{2\lambda}{U_{GS} - U_p(T_J)} - \dfrac{1}{T_J}$.

Na rys. 6 pokazano obliczone według powyższych wzorów zależności modułu admitancji wyjściowej y_{22} (rys. 6a) oraz transadmitancji y_{21} (rys. 6b) tranzystora MOS małej mocy. Jak widać, wpływ zjawisk termicznych na wartość tego parametru jest widoczny w zakresie małych częstotliwości do 0,1 Hz, co zależy od przyjętych do



Rys.6 Zależność od częstotliwości modułu nieizotermicznej admitancji wyjściowej y_{22} (rys. 6a) oraz transadmitancji y_{21} (rys. 6b) tranzystora MOS

obliczeń wartości termicznych stałych czasowych modelu termicznego tego elementu. Analogiczne zależności, jak wzory (15), dla nieco zmienionego modelu statycznego tranzystora podano w pracy [3].

4. PRZYKŁADY ZASTOSOWAŃ NIEIZOTERMICZNYCH PARAMETRÓW MAŁOSYGNAŁOWYCH ELEMENTÓW PÓŁPRZEWODNIKOWYCH

Jak wcześniej zaznaczono, postać modeli małosygnałowych wpływa na parametry robocze układów skonstruowanych z użyciem tych elementów, co poniżej zilustrowano na przykładach wzmacniacza napięciowego na tranzystorze bipolarnym (rozdz. 4.2) oraz układu źródła prądowego (rozdz. 4.3). W rozdziale 4.1 pokazano wpływ efektów termicznych na pomiary szumów typu $1/f$ elementów na przykładzie diody pn. W rozdziale 4.4 opisano sposób wyznaczania rezystancji termicznej tranzystora bipolarnego na podstawie pomiaru jego nieizotermicznej rezystancji wyjściowej.

4.1. SZUMY 1/f DIODY ZE ZŁĄCZEM p-n

W zakresie małych częstotliwości dominującym składnikiem szumów wielu elementów półprzewodnikowych są szumy o charakterystyce widmowej typu $1/f^{\alpha}$. Od dawna istnieją liczne kontrowersje na temat pochodzenia tych szumów, a także charakteru ich widma, np. [12]. Według wielu teorii, we wzorze opisującym wydajność prądowego generatora szumów wykładnik α powinien przyjmować wartość równą 1. W wielu przypadkach stwierdzono, na podstawie pomiarów napięcia szumów, zmianę wartości wykładnika α w zakresie małych częstotliwości, co próbowano tłumaczyć w różny sposób. Poniżej wykazano, że w zakresie małych częstotliwości charakterystyka widmowa napięcia szumów diody może wykazywać zmianę wartości wykładnika α w funkcji częstotliwości, mimo ze źródło szumów diody, scharakteryzowane przez generator prądowy, posiada stałą wartość wykładnika α. Tak więc wykazano, że jednym z powodów obserwowanej, na podstawie pomiarów napięcia szumów, zmiany wartości α, charakteryzującej prądowy generator szumów, mogą być zjawiska termiczne.

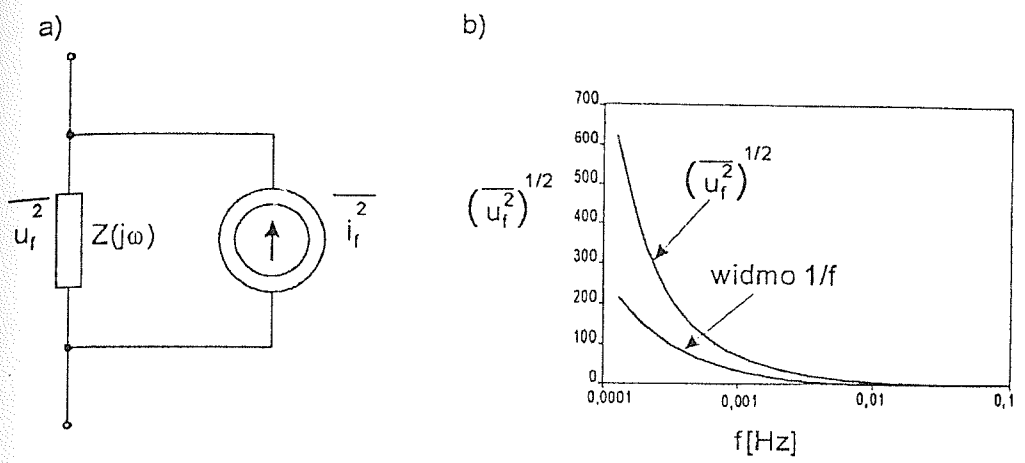
Na rys. 7a pokazano sposób dołączenia źródła szumów prądowych o charaktrystyce widmowej typu 1/f (założono $\alpha = 1$) do diody scharaktryzowanej impedancją elektryczną $Z(j\omega)$, zaś na rys. 7b przedstawiono wyniki obliczeń napięcia szumów (w jednostkach względnych) na zaciskach diody. Jak widać, charakterystyka widmowa napięcia szumów nie może być opisana przebiegiem o stałej wartości α i, dla przykładu, dla bardzo małych częstotliwości α jest rzędu 1,2.

Pomiary szumów typu 1/f często prowadzi się w celach badawczych w zakresie bardzo małych częstotliwości i dlatego pokazany powyżej efekt występujący w tym zakresie może być interesujący.

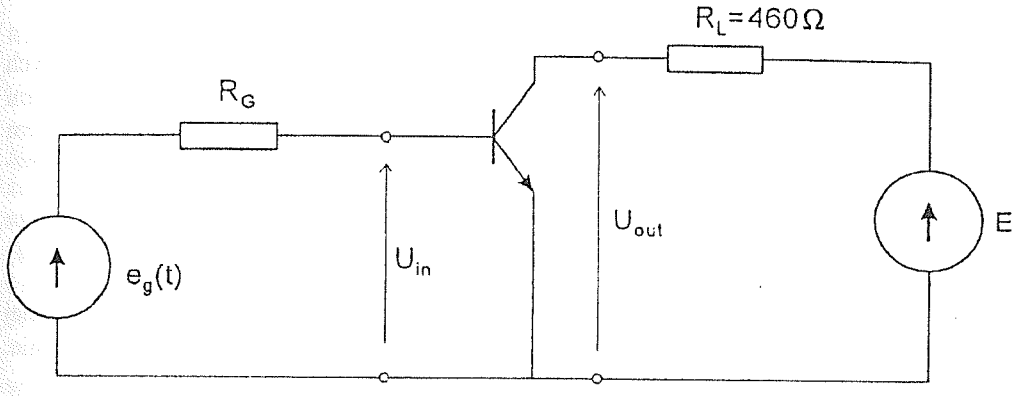
4.2. WZMACNIACZ NAPIĘCIOWY NA TRANZYSTORZE BIPOLARNYM

Na rys. 8 pokazano prosty układ wzmacniacza napięciowego o obciążeniu rezystancyjnym, z wykorzystaniem tranzystora bipolarnego jako elementu aktywnego. Wzmocnienie napięciowe K_u takiego układu jest wyrażone wzorem

$$K_u = \frac{U_{wy}}{U_{we}} = \frac{h_{21e}}{h_{11e} \cdot (h_{22e} + G_L) - h_{12e} \cdot h_{21e}}. \tag{16}$$



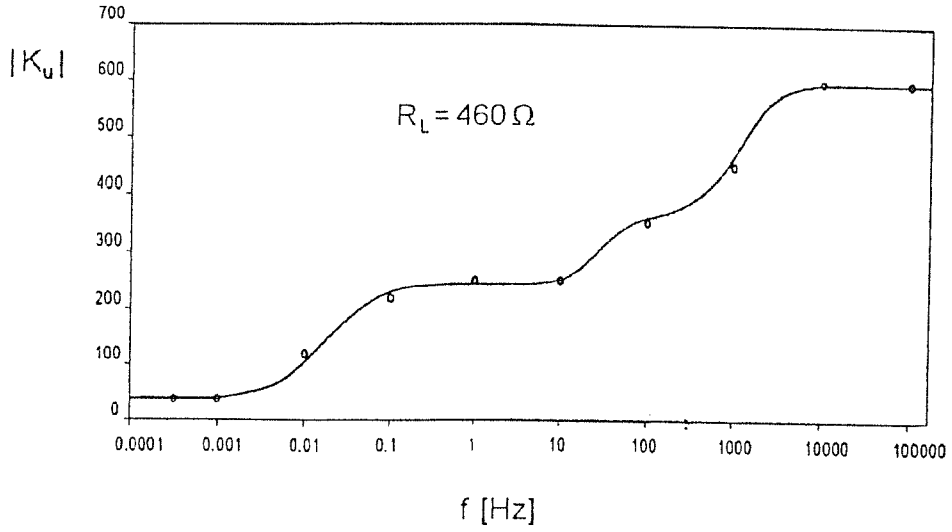
Rys. 7. Małosygnałowy nieizotermiczny schemat zastępczy diody z uwzględnieniem źródła szumow 1/f (rys. 7a) oraz zależność napięcia szumów diody od częstotliwości (rys. 7b)



Rys. 8. Układ wzmacniacza napięciowego

Na rys. 9 pokazano pomierzony i obliczony według wzoru (16) przebieg modułu wzmocnienia napięciowego analizowanego wzmacniacza [5, 9]. Jak widać, znaczne zmiany modułu wzmocnienia w funkcji częstotliwości są w rozpatrywanym przypadku widoczne aż do częstotliwości rzędu 500 Hz. Jak wykazały obliczenia, zakres ten zależy od relacji między rezystancją obciążenia R_L a odwrotnością modułu admitancji wyjściowej. Jeżeli obie te wielkości przybierają wartości współmierne, to zakres częstotliwości, w którym moduł K_u nie jest stały, ulega rozszerzeniu.

W pracy [7] przedyskutowano wpływ oddziaływań termicznych w tranzystorze bipolarnym użytym w torze pomiarowym miernika szumów jako element wzmacniający na wyniki pomiarów szumów elementów. Wykazano, że wyniki te mogą być zniekształcone wskutek opisanego wyżej efektu niestałości modułu wzmocnienia

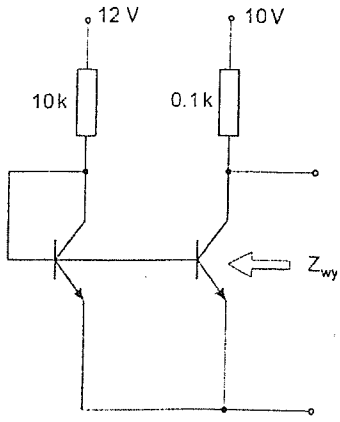


Rys. 9. Zależność od częstotliwości modułu wzmocnienia napięciowego wzmacniacza z rys. 8

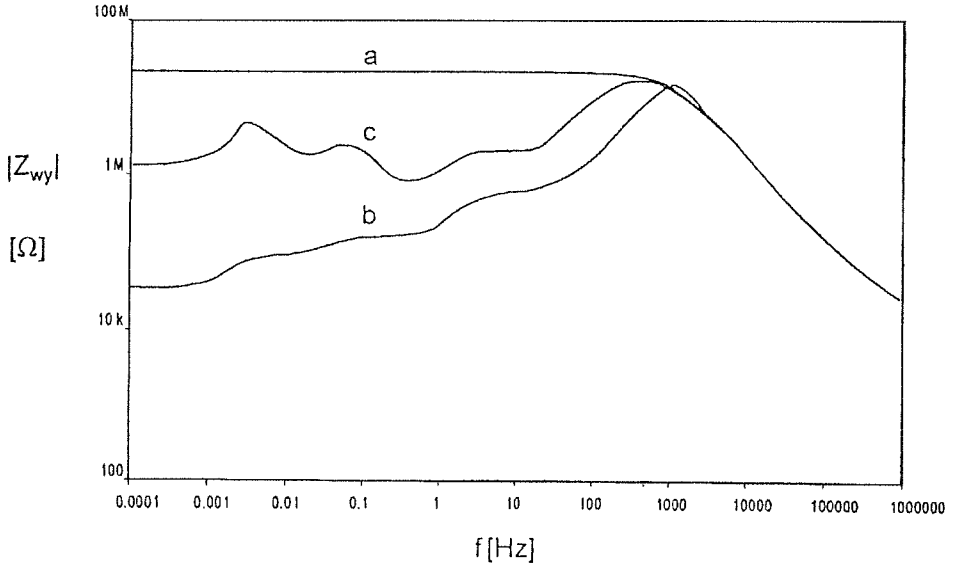
napięciowego układu wzmacniacza w funkcji częstotliwości, widocznego dla małych i bardzo małych częstotliwości. Jest to efekt, wywołany zjawiskami termicznymi związany z własnościami funkcjonalnymi toru pomiarowego, a nie z wpływem zjawisk termicznych na szumy badanego elementu, co poprzednio opisano w rozdziale 4.1 dla diody. Oba te efekty mogą wystąpić niezależnie od siebie.

4.3. SCALONE ŹRÓDŁO PRĄDOWE

W źródle prądowym pokazanym na rys. 10, zrealizowanym w wersji scalonej tranzystory są ze sobą połączone elektrycznie oraz oddziałują na siebie termicznie.



Rys. 10. Scalone źrodło prądowe



Rys. 11. Zależność od częstotliwości modułu impedancji wyjściowej źródła prądowego z rys. 10 (objaśnienie krzywych w tekście)

Na rys. 11 przedstawiono wyniki obliczeń modułu impedancji wyjściowej Z_{wy} tego źródła dla trzech przypadków: analiza izotermiczna (krzywa a), analiza elektrotermiczna bez uwzględnienia wzajemnych sprzężeń termicznych między elementami (krzywa b) oraz analiza elektrotermiczna z uwzględnieniem tych sprzężeń (krzywa c). Przyjęte do obliczeń wartości parametrów elektrycznych i termicznych elementów tworzących źródło prądowe odpowiadają układowi scalonemu UL1111.

Jak widać z rys. 11, wyniki analizy elektrotermicznej różnią się znacznie od wyników analizy izotermicznej w zakresie częstotliwości do około 100 Hz. Wartości modułu impedancji wyjściowej uzyskane z analizy elektrotermicznej są znacznie mniejsze niż z analizy klasycznej, gdy nie uwzględnia się sprzężeń termicznych między elementami. Warto dodać, że o ile do wykonania obliczeń cytowanych w poprzednich rozdziałach użyto prostych, stosownych programów (np. MATHCAD), o tyle przeprowadzenie obliczeń dla nawet stosunkowo prostych układów, w których elementy są sprzężone elektrycznie i termicznie wymagało programów specjalizowanych (np. SPICE). W omawianym przypadku analizy źródła prądowego wykorzystano zmodyfikowaną wersję SPICE'a, przystosowaną do analizy elektrotermicznej [5].

4.4. POMIAR REZYSTANCJI TERMICZNEJ TRANZYSTORA BIPOLARNEGO

W celu określenia nieizotermicznej różniczkowej konduktancji wyjściowej g_d, należy wzór (9d) zapisać dla częstotliwości $f = 0$

$$g_d = a \cdot I_C^2 \cdot R_{th}. \tag{17}$$

Ze wzoru tego widać, że pomiar tej konduktancji pozwala obliczyć rezystancję termiczną tranzystora, przy danym prądzie kolektora, jeżeli dodatkowo wyznaczy się współczynnik „a" zmian temperaturowych parametru β. We wzorze (17) pominięto, nieistotny w tym przypadku, wpływ efektu Early'ego na konduktancję wyjściową. Taką metodę pomiaru można określić jako statyczną.

Przeprowadzone pomiary wykazały [11], że w zakresie słuszności modelu statycznego tranzystora przyjętego do wyprowadzenia zależności (17) uzyskano na podstawie pomiarów g_d wartości rezystancji termicznej tranzystora bliskie (z dokładnością rzędu 10%) wartościom uzyskanym za pomocą metody standardowej. Tak więc ta prosta statyczna metoda, nie wymagająca żadnej aparatury specjalistycznej koniecznej w przypadku wykorzystania metody standardowej pomiaru rezystancji termicznej, może być w niektórych przypadkach z powodzeniem stosowana w praktyce.

ZAKOŃCZENIE

W pracy przedstawiono metodę wyznaczania nieizotermicznych modeli małosygnałowych elementów półprzewodnikowych i pokazano takie modele dla wybranych elementów, uzyskane na podstawie tej metody. Na kilku przykładach zilustrowano wpływ nieizotermicznej postaci modeli małosygnałowych elementów na własności układów, a także przytoczono inne konsekwencje i korzyści wynikające ze stosowania takich modeli.

Warto dodać, że po linearyzacji charakterystyk nieizotermicznych elementu można pozostawić w modelu temperaturę jako jedną ze zmiennych (czyli można pominąć punkt c) metody opisanej w rozdz. 2). Wówczas można zdefiniować parametry elementu mówiące o jego elektrycznej odpowiedzi, w postaci prądu lub napięcia, na sygnał w postaci temperatury. Taką sytuację rozpatrzono w pracy [3].

BIBLIOGRAFIA

1. O. Mueller: *Internal thermal feedback in four-poles especially in transistors.* Proc. IEEE, 1964, vol. 52, str. 924
2. G.O. Workman i in.: *Physical modeling of temperature dependences of* SOI CMOS *devices and circuits including self-heating.* IEEE Trans. on El.Dev., 1998, vol. 45, str. 125
3. W. Janke: *Modele elektrotermiczne elementów półprzewodnikowych.* Zeszyty Naukowe Politechniki Gdańskiej, seria Elektronika, zeszyt LVII, 1984
4. W.J. Stepowicz, J. Zarębski: *Effect of mutual thermal interactions in IC-s on the output admittance of the BJT.* Materiały XVIII Krajowej Konferencji „Teoria Obwodów i Uklady Elektroniczne", Polana Zgorzelisko, 1995, vol. 1, str. 203
5. J. Zarębski: *Modelowanie, symulacja i pomiary przebiegów elektrotermicznych w elementach półprzewodnikowych i układach elektronicznych.* Prace Naukowe Wyższej Szkoły Morskiej w Gdyni, 1996

6. V. Szekely: *A new evaluation method of thermal transient measurements results.* Microelectronics Journal, 1997, nr 28, str. 277
7. W.J. Stepowicz, J. Zarębski: *Wpływ zjawiska samonagrzewania na pomiary szumów 1/f elementów półprzewodnikowych.* Metrologia i Systemy Pomiarowe, 1997, nr 3 – 4, str. 145
8. W.J. Stepowicz, J. Zarębski: *Influence of selfheating on the circuits parameters of the amplifier with the* BJT. Buletin of the Polish Academy of Sciences, Technical Sciences, 1996, vol. 44, str. 283
9. W.J. Stepowicz, J. Zarębski: *Influence of selfheating on the low-frequency voltage gain of the amplifier with the* BJT. 12-th European Conf. on Circuit Theory and Design ECCTD'95, Istanbul, 1995, vol. 2, str. 635
10. W.J. Stepowicz, J. Zarębski: *Influence of selfheating on the output admittance of the* BJT. Materiały XVII Krajowej Konferencji „Teoria Obwodów i Układy Elektroniczne", Polanica Zdrój, 1994, wol. 1, str. 401
11. W.J. Stepowicz, J. Zarębski: *A new method of the thermal resistance measurements for bipolar transistors.* Materiały Krajowego Kongresu Metrologii KKM'98, Gdańsk 1998, vol. 3, pp. 430
12. A. Konczakowska: *Szumy małoczęstotliwościowe a indywidualne prognozowanie niezawodności elementów elektronicznych.* Zeszyty Naukowe Politechniki Gdańskiej, seria Elektronika, zeszyt LXXVII, 1992

W.J. STEPOWICZ, J. ZARĘBSKI

NONISOTHERMAL SMALL-SIGNAL PARAMETERS OF SEMICONDUCTOR DEVICES

Summary

The thermal phenomena influence models of semiconductor devices considerably; it concerns the d.c. characteristics and the small-signal circuits as well. In the well-known models the effect of the ambient temperature is only taken into account, and therefore these models can be denominated as isothermal. In fact, due to the selfheating phenomenon, the junction (inside) temperature differs from the ambient temperature. Under steady-state, the junction temperature rise above the ambient temperature depends on the thermal resistance of the device and on the dissipated power. Under any a.c. current or voltage exitation, the a.c. component of the junction temperature appears. The models including selfheating can be denominated as nonisothermal. In some cases, as in ICs, the mutual thermal interactions between the devices have to be taken into considerations, additionally. In this paper the a.c. nonisothermal small-signal models of some semiconductor devices are discussed. Due to the thermal effects, in the low frequency range the small-signal nonisothermal parameters of the devices become complex. Since in this range the electrical inertia resulting in the diffusion and junction capacitances can be neglected, the thermal intertia leads to the small-signal circuits with the new capacitances and inductances. In the paper the influence of the nonisothermal small-signal parameters on the performance of some circuits: the voltage amplifier and current source is shown. As well, some results concerning the 1/f noise model of the pn diode are given, and a new method of the thermal resistance measurement sof the BJT, based on its nonisothermal output conductance, is presented.

Key words: semiconductor devices, thermal effects, small-signal models.

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI, 1998, **44**, z. 3, ss. 365–372

Thermal breakdown in semiconductor junction devices

WITOLD JERZY STEPOWICZ

Wydział Elektroniki, Telekomunikacji i Informatyki, Politechnika Gdańska

Otrzymano 1998.07.02

Autoryzowano 1998.07.21

Notation

A — constant in eqn (1),
B — constant in eqn (1),
F — function given by eqn (1),
G — function given by eqn (2),
k — Boltzmann constant,
m — exponent in eqn (1),
n — exponent in eqn (1),
q — electron charge,
r_d — nonisothermal differential resistance,
r_i — isothermal differential resistance (at $T_j = T_s$),
R_{th} — thermal resistance,
T_a — ambient temperature,
T_j — junction temperature,
V_{go} — potential corresponding to the bandgap, extrapolated to $T = 0$ K, in silicon $V_{go} = 1{,}21$ V,
Φ_B — height of the potential barrier in the metal-semiconductor (ms) junction.

The breakdown phenomenon is found practically in all semiconductor devices with nonlinear pn or ms junctions. It delimits the safe operation area of these devices. On the other hand, some semiconductor devices, for example the Zener diodes and avalanche photodiodes, operate in the breakdown region. So, understanding the performance of semiconductor devices at breakdown is of great importance.

Key words: breakdown phenomenon in pn and ms junctions, thermal effects.

1. INTRODUCTION

The breakdown region of semiconductor devices with any nonlinear junction (pn or ms) is generally understood as the region in which a small change of the reverse voltage leads to a considerable change in the reverse current. The electrical breakdown can be caused by the impact ionization or/and the tunneling (Zener) effect. At this breakdown the d.c. charcteristics can have the vertical asymptote at the breakdown voltage or, in more sophisticated models, the turnover point, at which the differential resistance changes its sign from positive to negative, occurs, e.g [1].

When the selfheating phenomenon is taken into account, the d.c. characteristics in the breakdown region are determinated by interactions between electrical, mentioned above breakdown effects, and thermal ones, and then such kind of breakdown mode can be called as electrothermal [2]. In most cases the turnover point is present on the d.c. characteristics, which are S-shaped mostly. Then the d.c. characteristics can be called as the nonisothermal ones. In some cases the electrothermal interactions due to selfheating only, without any electrical breakdown included in analysis, lead to the S-shaped d.c. characteristics as well. In such cases this breakdown mode can be called as the thermal breakdown [2].

The aim of this paper is to show that the thermal breakdown is symptomatic for the common, well-known form of the d.c. characteristics $i(v, T_j)$, which describe different junction devices discussed here: the silicon pn diode, the Schottky diode and the bipolar transistor. In Sec. 1 on the basis of the d.c. nonisothermal characteristics the coordinates of the breakdown point (i, v, T_j) are derived. In Sec. 2 some experimental results illustrating the thermal breakdown are shown.

2. THEORY

The d.c. characteristics of some junction semiconductor devices can be presented in the form

$$F = i(v, T_j) = A \cdot v^m \cdot T_j^n \cdot \exp\left(-\frac{B}{T_j}\right). \tag{1}$$

Selfheating causes the junction temperature rise above the ambient temperature due to the electrical power dissipated in the device and nonideal conditions of heat removal to surroundings. The thermal model of the device, i.e. the dependence of the junction (inside) temperature T_j on the disspated electrical power, can be expressed under steady-state condition as

$$G = T_j(i, v) = T_a + i \cdot v \cdot R_{th}(i, v, T_j). \tag{2}$$

In general, the thermal resistance R_{th} can be dependent on the coordinates (i, v) of the operation point and on the temperature (T_j and/or T_a).

The nonisothermal differential resistance r_d of the device, described by the relations (1) and (2), is as follows

$$r_d = \frac{1 - \dfrac{\partial F}{\partial T_j}\dfrac{\partial G}{\partial i}}{\dfrac{\partial F}{\partial v} + \dfrac{\partial F}{\partial T_j}\dfrac{\partial G}{\partial v}}. \tag{3a}$$

For the variables (v, i, T_j) the following identity can be written

$$\frac{\partial i}{\partial v}\frac{\partial v}{\partial T_j}\frac{\partial T_j}{\partial i} = -1. \tag{4}$$

As seen from eqn (4), for the device with the positive derivative $\partial i/\partial v$ the temperature coefficients of current and voltage have to be opposite signs.

Some modifications under assumption of the linear dependence of the junction temperature on electrical power dissipated in the device (ie. with $R_{th} = \text{const}$) lead to

$$r_d = \frac{1 - v \cdot R_{th} \cdot \dfrac{\partial i}{\partial T_j}}{1 - i \cdot R_{th} \cdot \dfrac{\partial v}{\partial T_j}} \cdot \frac{\partial v}{\partial i}, \tag{3b}$$

where $r_i = \dfrac{\partial v}{\partial i}$ can be treated as the isothermal differential resistance, which is positive in the considered case.

It is easy to prove that the derivative $\dfrac{\partial i}{\partial T_j}$ is positive for the device described by eqn (1), and therefore this device has the S-shaped d.c. characteristics. In this case the resistance r_d changes its sign from positive to negative after passing zero value. The change takes place at the thermal breakdown point with the coordinates (T_j, v_t, i_t).

The thermal breakdown temperature T_{jt} is to calculate provided the differential resistance is equal to zero. To simplify the calculations, without losing any important feature of the breakdown, the power „n" in eqn (1) is assumed to be zero. In fact, in eqn (1) the influence of temperature caused by the exponential factor is far stronger than by the factor T_j^n. The above condition is given by

$$\frac{B \cdot (T_{jt} - T_a)}{T_{jt}^2} = 1 \tag{5}$$

and the thermal breakdown temperature T_{jt} calculated from eqn (5) is equal to

$$T_{jt} = \frac{B}{2}\left[1 - \left(1 - \frac{4T_a}{B}\right)^{1/2}\right]. \tag{6}$$

The second solution of eqn (5) gives the value of T_{jt} approximately equal to B, i.e. far beyond the values of the allowable junction temperature. The other coordinates of the breakdown point are as follows

$$v_t = \frac{T_{jt}^2}{R_{th} \cdot A \cdot B} \cdot \exp\left(\frac{B}{T_{jt}}\right) \tag{7}$$

$$i_t = A \cdot \exp\left(-\frac{B}{T_{jt}}\right). \tag{8}$$

To shorten the formulas, the coordinates (i, v_t) given by eqns (7) and (8) are expressed by the coordinate T_{jt} — eqn (6).

As seen, the junction temperature of the thermal breakdown T_{jt} depends on the ambient temperature T_a and the parameter B only. The current and voltage coordinates depend on the parameter A as well, whereas the thermal resistance R_{th} is present in eqn (7) merely. These three coordinates fulfil the following condition

$$\Delta T_{jt} = T_{jt} - T_a = i_t \cdot v_t \cdot R_{th}. \tag{9}$$

To demonstrate the influence of B and T_a on ΔT_{jt}, some calculations results are shown in Table 1.

Table 1

Calculated values of ΔT_{jt}

	T_a [K]	300	350	400	450
ΔT_{jt} [K]	B = 7000 K	14.1	19.5	25.9	33.4
	B = 14000 K	6.7	9.2	12.1	15.5

As seen, the temperature increment ΔT_{jt} increases with the increasing of the ambient temperature at B = const, and it decreases with the increasing of B at T_a = const. The differential coefficients describing these changes of T_{jt} are as follows

$$\frac{\partial T_{jt}}{\partial T_a} = \left(1 - \frac{4T_a}{B}\right) > 0 \tag{10}$$

$$\frac{\partial T_{jt}}{\partial B} = -\frac{\Delta T_{jt}}{B - 2T_{jt}} < 0 \tag{11}$$

The derivaives of i_t, v_t with respect to T_a, B and A can be calculated as well. The derivative $\frac{\partial v_t}{\partial R_{th}}$ is given here only

$$\frac{\partial v_t}{\partial R_{th}} = -\frac{v_t}{R_{th}}. \tag{12}$$

3. EXPERIMENTAL RESULTS

To illustrate the thermal breakdown of some semiconductor junction devices some measurements results are discussed below.

The semiconductor devices mentioned in Sec. 1 are characterizied by different values of the constant B:

a) for the silicon pn diode $B = \frac{q}{2k} \cdot V_{go} = 7000$ K, if the reverse current is determinated by the generation current and $B = 14000$ K, if the reverse current is given by the saturation current,

b) for the Schottky diode $B = \frac{q}{k} \cdot \Phi_B$,

c) for the BJT with the voltage drive $B = \frac{q}{k} \cdot (V_{go} - v_{BE})$.

In the silicon pn diode the condition (9) can be fulfilled, i.e. the thermal breakdown can take place, if the value of the thermal resistance is high or/and the reverse current is high. In practice, this situation is possible at high ambient temperatures. As known, in the high tempratures range the value of the saturation current can be of the same order of magnitude as the generation current.

Measurement scarried out revealed possiblilty of the thermal breakdown in the silicon pn diode at high ambient temperatures [3]. In Fig. 1 the d.c. characteristics of the silicon diode BA 159 measured at $T_a = 486$ K (beyond the allowable temperature for this device) are shown, and the coordinates of the turnover point Q are given. The measured value of B is equal 9400 K, what can evidence, that the both currents — the generation and saturation current form the reverse current at this

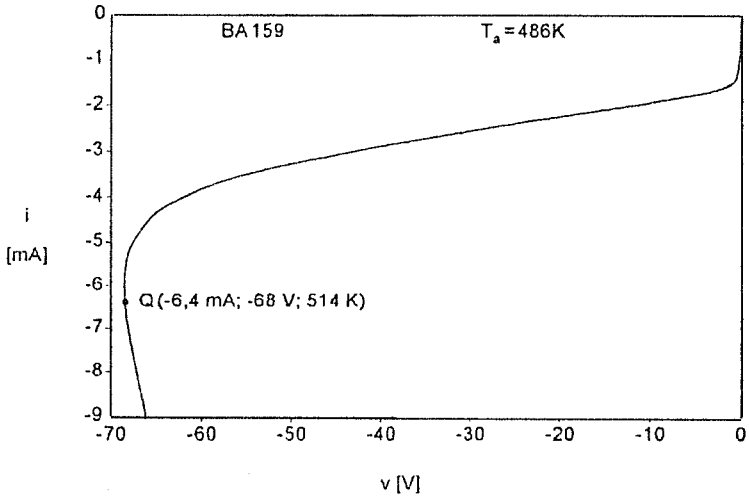


Fig. 1. The d.c. characteristics $i(v)$ of the reverse-biased pn silicon diode; Q denotes the breakdown point

ambient temperature. The measured and calculated (on eqn (6)) values of ΔT_{jt} are practically identical and equal to 28 K. The value of the electrical breakdown voltage of this diode measured at $T_a = 300$ K by a curve tracer was about 1500 V, while at $T_a = 486$ K was about 68 V, as seen from Fig. 1. The observed deterioration of blocking performance of the diode delimits its use in high-temperature electronics.

In Schottky diodes (these considerations concern power diodes first of all) the reverse current is much higher than in silicon pn diodes, and so the condition (9) can be easier fulfilled, even at room ambient temperatures. In Fig. 2 the measured d.c. characteristics of the medium power Schottky diode 1M5832 are shown at the different ambient temperatures [4]. As seen from this figure, the measured values of the junction temperature rises ΔT_{jt} amount about 20 K, and they are in good agreement with the calculated ones. As known, the typical values of the allowable reverse votage do not exceed 200 V for Schottky diodes, and one of the main factors limiting their allowable voltage is the thermal breakdown, at higher ambient temperatures especially.

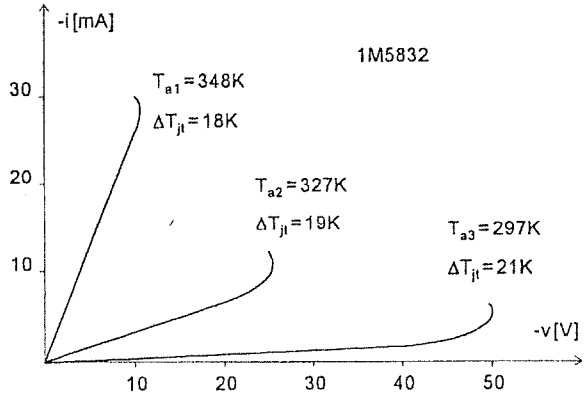


Fig. 2. The d.c. characteristics $i(v)$ of the reverse-biased Schottky diode

In the BJT the thermal breakdown can take place as well. In Fig. 3 the S-shaped input characteristic $i_B(v_{BE})$ of the small-power BJT BC 108 are depicted [5]. The transfer characteristic $i_C(v_{BE})$ have the same S-type shape as $i_B(v_{BE})$, because these both characteristics have the functional form like eqn (1). It is worth mentioning, that in the considered case the output characteristics $i_C(v_{CE})$ with v_{BE} as the parameter are S-shaped as well. The typical value of the junction temperature rise ΔT_{jt} is equal about 15 – 16 K at medium currents and at room temperatures. At high currents the influence of the series resistances of the base and emitter regions on the coordinates of the turnover point becomes important. These resistances can play the stabilizing role in operation of the device [4].

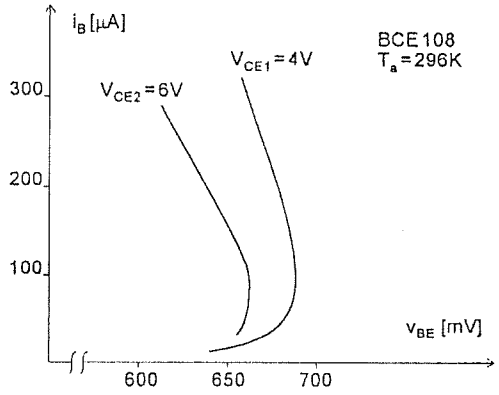


Fig. 3. The d.c. input characteristics $i_B(v_{BE})$ of the BJT

4. CONCLUSIONS

In this paper it is shown that for some semiconductor devices described by the function like eqn (1), the thermal breakdown can affect their d.c. characteristics considerably — it takes place in the reverse-biased silicon pn diode, in the reverse-biased Schottky diode and in the BJT, as it is shown above. The thermal breakdown is understood here as the situation, in which the nonisothermal differential resistance changes its sign, since the considerable changes of current at small changes of voltage, specific for the „classical" breakdown, are not always evident. Moreover, to get the thermal breakdown, the reverse bias of the junction is not necessary — see the d.c. input characteristics of the BJT. Other junction devices described by the formula like (1) and (2) should exhibit the thermal breakdown as well. The S-shaped d.c. characteristics of these devices determine their thermal instability at the supplying from the constant voltage source.

The essential formulas (1), (2) and the other following dependences resulting form the are valid for the thermistor and the intrinsic semiconductors, too. These devices have no nonlinear junction, and therefore it is not admitted to explain the S-type shape of their characteristics by the thermal breakdown. In [6] the thermistor is disccused as an example of „*thermal and electrothermal instabilities*", and the coordinates of the turnover point for this device are given there in the same form as eqns (6)—(8) here. The conditions of the thermal breakdown in semiconductors are discussed in [9], but the scope of this paper is limited to semiconductor crystals, without any pn or ms nonlinear junctions.

The electrothermal interactions, which are responsibile here for the negatively sloped d.c. characteristics, not always lead to the S-type. Unipolar transistors [4], heavily-doped semiconductor resistors [7] and even BJTs at high currents beyond the SOA [8] can have the N-type d.c. characteristics due to thermal effects. For unipolar transistors and semiconductor resistors, as well as for thermistors, using the term

„*breakdown*" is disccusable. For the N-type characteristics the nonisothermal differential conductance changes its sign from positive to negative. In this case the thermal instability can take place, when the device is supplied from the constant current source.

REFERENCES

1. J.C. Biswas et al: *Current-voltage characteristics of bulk semiconductors on impact ionization.* IEEE Trans. Electron Devices, nr 12, 1976, pp. 1353–1356
2. W.J. Stepowicz: *Unified approach to breakdown phenomena in silicon pn junction.* Solid-State Electronics, v. 22, nr 1, 1979, pp. 7–13
3. W.J. Stepowicz, J. Zarębski: *Przebicie termiczne w krzemowych diodach ze złączem pn.* Materiały XVI Sympozjum z Podstaw Elektrotechniki i Teorii Obwodów, v. 1, Gliwice, 1993, pp. 255–260, (in Polish)
4. W.J. Stepowicz: *Nieizotermiczne charakterystyki prądowo-napięciowe wybranych elementów półprzewodnikowych.* Zeszyty Naukowe Politechniki Gdańskiej, seria Elektronika, nr 373, Gdańsk, 1983, pp. 3–84 (in Polish)
5. W. Janke, W.J. Stepowicz: *Temperatura krytyczna tranzystora bipolarnego.* Materiały IV Krajowej Konferencji „Miernictwo Elementów Półprzewodnikowych i Układów Scalonych", Gdańsk, 1980, pp. 94–98 (in Polish)
6. M. Shaw et al: *The physics of instabilities in solid state electron devices.* Plenum Press, 1992, pp. 393–395
7. W.J. Stepowicz: *Relations between thermal stability and nonisothermal d.c. characteristics for semiconductor devices*, Bulletin de l'Academie Polonaise des Sciences, vol. 27, 10/11, 1979, pp. 75–83
8. J. Zarębski: *A new form of the BJT model including electrothermal interactions.* 11th European Conf. on Circuit Theory and Design, vol. 1, Davos, 1993, pp. 431–434
9. K. Somogyi: *Some critical conditions of the thermal breakdown in semiconductors.* 3rd THERMINIC Workshop, Cannes, 1997, pp. 189–191

W.J. STEPOWICZ

PRZEBICIE TERMICZNE W ZŁĄCZOWYCH ELEMENTACH PÓŁPRZEWODNIKOWYCH

Zjawisko przebicia występuje praktycznie we wszystkich elementach półprzewodnikowych ze złączem pn lub mp spolaryzowanych zaporowo. Ogranicza ono obszar bezpiecznej pracy wielu elementow, jednak, z drugiej strony, praca w zakresie przebicia jest typowa dla takich elementów jak, dla przykladu, diody stabilizacyjne i fotodiody lawinowe. Na statycznych charakterystykach $i(u)$ w zakresie przebicia dużym zmianom prądu odpowiadają małe zmiany napięcia. Na charakterystykach tych często występuje punkt, w którym rezystancja różniczkowa elementu zmienia znak z dodatniego na ujemny. Na kształt charakterystyk $i(u)$ w zakresie przebicia mają wpływ zarówno zjawiska elektryczne (jonizacja zderzeniowa i efekt tunelowy), jak i termiczne. W ogólnym przypadku, wskutek sprzężenia zjawisk elektrycznych i termicznych, można mówić o przebiciu elektrotermicznym. W pracy wykazano, że w elementach ze złączami pn lub mp, opisanych dobrze znanymi zależnościami $i(u)$, może wystąpić przebicie o charakterze termicznym, spowodowane jedynie zjawiskiem samonagrzewania złącza oraz wyznaczono współrzędne (i_t, v_t, T_{jt}) punktu przebicia termicznego. Przebicie termiczne zilustrowano na przykładach diody krzemowej ze złączem pn, diody Schottky'ego oraz tranzystora bipolarnego. Oprócz rozważań teoretycznych przedstawiono wyniki stosownych pomiarów przeprowadzonych dla wymienionych elementów.

Słowa kluczowe: zjawisko przebicia w złączach pn i mp, efekty termiczne.

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI, 1998, 44, z. 3, ss. 373–390

Artificial neural network mixed-signal prototype system for model parameter identification

ANDRZEJ MATERKA, PAWEŁ PEŁCZYŃSKI, MICHAŁ STRZELECKI

Instytut Elektroniki, Politechnika Łódzka

Received 1998.02.04

Authorised 1998.05.24

Principles of model parameters identification by means of ANN-like approximators of multivariable mappings are reviewed briefly. The technique offers very high speed and lower noise-induced error compared to traditional methods. However, the ANN training becomes time consuming and somewhat uncontrollable at large range of unknown model parameters and high accuracy level required. It is proposed to use a modular, classifier-approximator architecture to get rid of this deficiency. Results of computer simulation are presented to illustrate the idea. Neural network-based system designed for model parameter identification is proposed. This prototype architecture of the identification system implements both analog and digital network. Classifier network is realised as an analog circuit. FPGA-based digital circuit is used as the classifier control and interface to other subsystems. This modular system has been constructed in the form of two extension cards for PC computers. Learning process is performed by the host PC. Ten chips of Kohonen-type classifier network have been designed and manufactured in 2 μm CMOS technology. Testing of the PC extension cards has been completed. Functional measurements of the whole system have been started.

Key words: parameter identification, neural networks, signal processing, analog classifier.

1. INTRODUCTION

Parametric modelling of dynamic systems is a widely used technique of describing physical systems of interest. Parameterised mathematical equations of a given system provide its consistent model which can be easily used in computer simulation and design processes. Searching for a fast and reliable method of parameter estimation based on system observation is a crucial problem in many applications [1]. It has been proven that artificial neural networks (ANN) that have universal approximation capabilities [5] (such as multilayer perceptron) can be employed for

the dynamic system parameter estimation task [6]. Samples of investigated dynamic system under test (SUT) response are applied to a feedforward ANN. Estimated parameter values are produced at the ANN outputs. Thus the ANN approximates the mapping from the SUT observation space into the SUT parameter space. This mapping is learned by the ANN during a training process. Unfortunately, required size of ANN and the training time increase strongly with the approximated parameter range and assumed accuracy of approximation. Learning process can stick in local minima in case of complex dependencies between parameters and observations. A modular neural network containing observation classifier and parameter approximator was proposed in [16] to overcome these difficulties. The domain of the mapping is partitioned into a number of nonoverlaping regions. Thus the task of parameter estimation can be decomposed into a set of local parameter mapping operations. The classifier makes a decision to which region an observed vector belongs. This information is then used to select an appropriate weight vector of the approximator network. Design of hardware realisation of modular parameter estimation system is described herein.

2. PROBLEM DEFINITION

Consider a physical system represented by its adequate mathematical model. The model can be a set of differential/difference/algebraic equations of constant coefficients (parameters). Assume model structure is known but actual values of the parameters are unknown. To identify the parameters, the system is excited by a predetermined test signal (stimulus), specified either in time or frequency domain depending on the nature and properties of the system. The system response $y = (y_1,..., y_k)'$ to the stimulus is measured, where $(.)'$ denotes the vector transpose. It forms, either directly or after suitable preprocessing, the system observation vector $\varphi = (\varphi_1,..., \varphi_N)'$, $k \geq N$ and $\varphi \in \Phi \subset \Re^N$. The parameter identification problem is defined as follows: given the observation vector φ find estimates $\hat{\theta} = (\hat{\theta}_1,..., \hat{\theta}_p)'$ of the model parameters $\theta = (\theta_1,..., \theta_p)'$, $N \geq p$. The model parameters $\theta \in \Theta \subset \Re^p$ are assumed constant during the observation interval (but the system may be slowly time-varing). For a fixed stimulus, the observation vector is a multivariable function of the model parameters $\{\varphi = f(\theta): \Re^p \to \Re^N\}$ [1].

Some additional measures have to be undertaken to make the problem solvable. Namely, the stimulus signal definition, sampling moments selection (or input frequencies selection in the case of sine-wave testing) and the choice of the preprocessing algorithm have to be decided in a way which ensures that a unique inverse mapping $\{\theta = f^{-1}(\varphi)$: $\Re^N \to \Re^p\}$ from the observation domain Φ to the parameter domain Θ exists. The optimum experiment design theory [2] is useful at this stage. It is assumed in this paper that the inverse mapping of interest exists and is unique.

In most practical cases, the mapping $\varphi = f(\theta)$ is nonlinear and is not analytically invertible. Finding the parameter values in such cases requires iterative computer

```
get y;                          /*measure system response*/
φ = h(y);                       /*compute observation vector*/
θ̂ = θ̂_0;                        /*initialize parameter vector*/
i = 0;

do {
   ε = ‖φ̂_i − φ‖;               /*calculate error in observation*/
i = i + 1;
θ̂_i = θ̂_{i−1} + Δθ̂_i;          /update parameters to minimize observation error*/
   } while ((ε > ε_thresh) and (i < i_max));
```

Fig. 1. Pseudo-code for parameter identification by iterative model tuning

calculations, eg. those necessary to perform a least-square-error (LSE) model tuning [1], see Fig. 1. Limited speed of computers causes that model parameter identification by means of such numerical techniques is a slow process. For some system models, formulae that describe the relationship between the observations φ and the parameters θ can be arranged to take a linear form. Parameter identification is obtained in such cases by solving a set of linear equations for θ [3]. This still takes a substantial amount of computer time which is further increased by the time needed to determine coefficients of the equations, eg. through multiple numerical integration [4]. In conclusion, traditional techniques of model parameter identification are computationally costly and can not be applied to real-time identification for many systems of practical interest.

3. PARAMETER IDENTIFICATION BY MEANS OF ANN APPROXIMATORS

It can be noticed that model parameter identification is the process of evaluating the vector value of the inverse mapping $f^{-1}(\varphi)$ given the observation vector φ. On the other hand, it is well known [5] that artificial neural networks (ANNs) are able to approximate any continuous compact-support function to any degree of accuracy. It was then postulated in [6] that ANNs can be used to approximate the mapping of interest $\hat{\theta} = g(\varphi) \cong f^{-1}(\varphi)$, provided they have been trained properly, see Figs. 2 and 3. The training itself is an iterative process which can take a substantial amount of computer time. However, it is done once only for a given model and its parameter range. After the training, in the recall mode, the ANN produces its response, ie. model parameter values, in a very short time — just in one shot in the case of feedforward neural networks. High speed is therefore one of the advantages of the ANN-based parameter identification technique. This feature makes it suitable for real-time applications.

In the approach of [6], the ANN performs the function of nonlinear associative memory (NAM). Linear associative memory (LAM) was also proposed [7] as

```
get y;                          //measure system response
φ = h(y);                       //compute observation vector
                                //(preprocessing)
θ̂ = g(φ);                       //compute ANN aproximator response
                                //to φ
```

Fig. 2. Pseudo-code for parameter identification with the use of ANN approximator

```
{φ_j, θ_j; j = 0, ..., N};      /*define the set of training examples*/
w = w_0;                        /*initialize weight vector*/
i = 0;                          /*initialize iteration index*/

do {
   ε = 0;                       /*initialize error variable*/
for (j =0; j<N; j++);
   ε = ε + ||θ̂_j − θ_j||;       /*accumulate parameter error*/
   w_i = w_{i−1}+Δw_i;          /*update weight vector to minimise parameter error*/
   } while ((ε>ε_thresh) and (i<i_max));
```

Fig. 3. Pseudo-code for ANN approximator training

a means to obtain preliminary parameter estimates of nonlinear systems. Obviously, LAM-identified parameters are of limited accuracy [7], [8]. Using second-order polynomials for the approximation resulted in accuracy improvement over the LAM [7]. Both LAM and NAM provide higher speed when compared to standard techniques for model parameter identification — by a few orders of magnitude, depending on the model equations [9].

Training the ANN approximator for model parameter identification involves minimising the differences between the actual and target parameters, whereas the model-tunning technique produces parameter values by minimising differences between the model and system observations. Therefore, the ANN trained on noisy observations becomes less sensitive to those input data which would normally affect much the LSE-fitted model parameters. This is another advantage of the ANN-based technique, which is its higher noise immunity when compared to traditional methods. A more detailed analysis of this property can be found in [9].

The novel technique discussed can be applied to any model: linear, nonlinear, continous or discrete, provided that the mapping $\theta = f^{-1}(\varphi)$ exists and there is a network able to approximate it with a tolerable error. The [numerous] applications range from electronic circuit parametric testing [10], [11], through control engineering [12] to medical diagnosis support and biomedical engineering [9] .The technique can also be used for indirect mesurements. Part of the current research work, aimed at futher development of the technique, relates to investigation of the ANN

approximator ability to produce parameter values with an acceptable error at an increased range of model parameters. To illustrate this problem consider an exponential model:

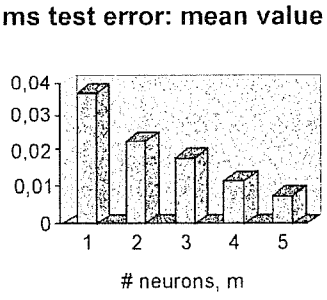
$$y(t) = \theta_3\left[\exp\left(\frac{-t}{10\theta_1}\right) - \exp\left(\frac{-t}{\theta_2}\right)\right], \tag{1}$$

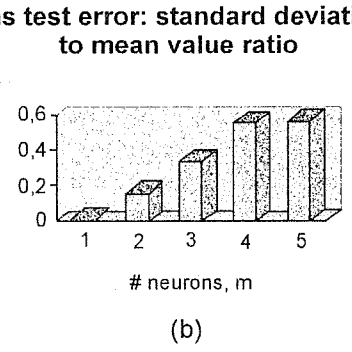
where θ_1, θ_2 and θ_3 are 3 unknown parameters. The response (1) is quite ubiquituous, eg. it may represent biological multicompartmental systems [13] or step response of electronic circuits [10] .It is assumed in this example that the parameters can take their values from a cuboid as follows

$$\{\Theta\colon 0.6 \le \theta_1 \le 1.4,\ 0.2 \le \theta_2 \le 1.0,\ 0.8 \le \theta_3 \le 1.2\}. \tag{2}$$

A well-known single-hidden layer perceptron, with sigmoidal nonlinearity in the hidden layer and linear combiner as the output unit, was trained using a two-stage procedure employing the backpropagation algorithm followed by a direct search, modified Davies-Swann-Campey numerical minimistaion procedure [14]. Three separate networks have been trained, one for each parameter. The inputs to the networks were the model observations, ie. values of $y(t)$ calculated from (1) at 3 distinct time moments: $t_1 = 0.56$, $t_2 = 2.61$ and $t_3 = 12.9$. The training patterns were calculated from (1) on a regular lattice of 8 points per each parameter range, total of 512 points. The respective parameter values were the target quantities to be learned by the ANNs. After the training, the networks were tested on a denser grid of 20 points per parameter, total of 8000 test vectors.

Result of an ANN training depends on the starting point in the weight space. To illustrate this feature, each of the networks considered was trained 30 times, with the initial weights taken at random from the range $[-0.5, 0.5]$. Figs. 4a to 4c show the results of the training as obtained for a number m of hidden-layer sigmoidal units ranging from 1 to 5. The average test error decreases with m as intended, although rather slowly, see Fig. 4a. At the same time, the ratio of the test error standard deviation to the error mean value increases with m, Fig. 4b. Thus by increasing the network complexity in order to achieve higher approximation accuracy, one introduces more uncertainty to the training procedure instead. This can be explained by the fact that with the increased number of weights of the ANN, there appear new local minima of the error function which is minimised during the training. It becomes then more likely that the training algorithm is trapped by a local minimum and is unable to reach the global one. At the same time, an increase of m leads to an exponential increase of the training time, Fig. 4c. Similar results were obtained for identification of the other parameters of the model (1), as well as for a 4-parameter model [10]. In conclusion, this approach to achieving higher accuracy of approximation (through an increase of the number of hidden-layer neurons) does not seem practical. There is a need for developing an ANN architecture whose training will produce more predictable results.





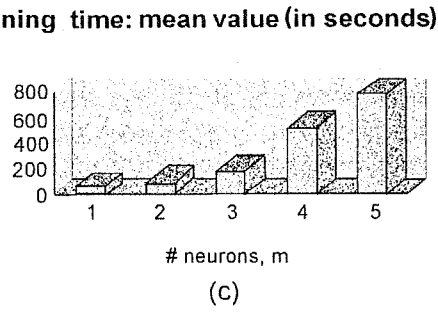


Fig. 4. Identification of parameter θ_1: MLP approximator performance over 30 training sessions

4. MODULAR ANN ARCHITECTURE

It is well known, that in case of polynomial approximation of a smooth single-variable function $v(x)$ on a compact set of real numbers, $x \in <a, b> \subset \Re$, the demand for higher-order polynomial terms, necessary to maintain a given accuracy level, increases with the range $r = |b-a|$ of the independent variable. For a narrow range r, retaining a linear term only may prove satisfactory. Similar property is attributed to single- and multi-variable function expansion using other base functions, not only polynomials. This suggests a solution to the problem addressed in the previous section. Namely, it is proposed to split the observation domain Φ of the mapping into a number M of non-overlapping regions Φ_i, $i = 1,..., M$ such that their union covers the whole domain, $\Phi_1 \cup \Phi_2 \cup ... \Phi_M = \Phi$. Since a unique continuous mapping from the observation space to the parameter space exists, there are M regions Θ_i in the parameter space, such that $\Theta_1 \cup \Theta_2 \cup ... \Theta_M = \Theta$, each corresponding to a respective region Φ_i, $i = 1,..., M$. As M increases, the range of the parameters within each of the regions Θ_i and the range of the observations within each of the regions Φ_i both decrease. The approximation task within each observation region becomes then simpler, since the mapping $\theta = f^{-1}(\varphi)$, $\varphi \in \Phi_i$ is closer to a linear relationship compared to its complexity over the whole domain, $\varphi \in \Phi$, where the nonlinearity is more pronounced. As a consequence, lower complesity approximator network is needed within each of the regions to maintain a given accuracy level. It is necessary, however, to decide which region a given observation vector belongs to.

The idea discussed in the previous paragraph can be realised using a modular ANN architecture which employs a classifier and a bank of q approximators. Each approximator of the bank is preoptimised in terms of its weight vector w to provide lowest possible error wihtin teh respective region. The classifier first allocates an acquired observation vector to one of the q regions, $\varphi \in \Phi_i$. The proper approximator ANN is then invoked, to produce the parameter vector values within the region Θ_i, $\hat{\theta} = g_i(\varphi) \cong f^{-1}(\varphi) \in \Phi_i$. The block diagram of the proposed modular architecture is shown in Fig. 5.

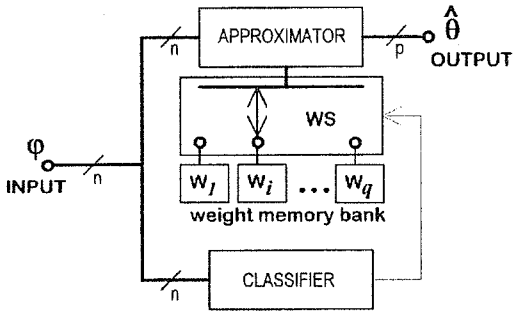


Fig. 5. Proposed modular ANN architecture, WS — weight switching module

5. COMPUTER SIMULATION

Computer simulation was performed to apply the modular ANN to the identification of parameters of model (1). Three independent neural networks were trained, each for an individual parameter θ_1, θ_2 and θ_3. It was found in many of our experiment that such a solution is more efficient than using a single, three output ANN, both in terms of reduced size, reduced training time and increased repeatability of results obtained. A minimum-Euclidean-distance classifier was trained to tesselate the observation domain into $q = 5$, 10, 15 and 20 regions. Two approximator structures were then taken into account, an MLP with 2 sigmoidal neurons (11 weights) and a rational function (RF) network [15] with 13 weights. The q approximators were trained using a numerical minimisation routine and then tested on a dense grid of 8000 points covering the whole parameter domain (2). Some of the

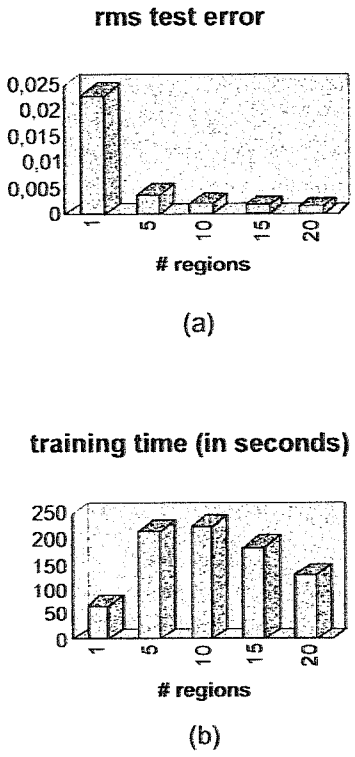


Fig. 6. Identification of parameter θ_1: modular architecture performance for a typical training session using MLP approximators with $m = 2$ sigmoidal neurons

results obtained are shown in Fig. 6, for MLPs approximating parameter θ_1. The standard deviation of the rms error over a number of training sessions tends to decrease with q from already low values, so it is not presented here. With increasing q, the rms test error decreases to much lower values than those obained for the single MLP, compare Fig. 4a nad Fig. 6a. The total time needed to train a lower-error modular ANN structure tends even to decrease with the error value, which is seen from the comparison of Figs. 6a and 6b. On the contrary, the single-approximator architecture needs longer training times to achieve lower errors, cf. Figs 4a and 4c.

These results are encouraging and prove the usefullnes of the modular ANN architecture. In general, for a given q, lower error values and shorter training times were experienced with the RF network which is a promising approximator architecture for this application. It needs futhter investigation to decide which approximator structure, MLP or RF, is more suitable for the task of model parameter identification with high accuracy. Training of the modular architecure is definitely a better controlled and predictable process compared to the training of a single large-size approximator network. The price paid for that is a slightly increased time needed for identification, namely this time is extended by a period dnecessary to classify the observation vector. Parallel classifiers may help reduce this period to acceptable values.

6. HARDWARE ANN SYSTEM FOR MODEL PARAMETER IDENTIFICATION

Functional design of the proposed parameter identification ANN system, as described in Section 4, is illustrated in Fig. 7. Control circuit, programmed by a PC host computer, initialises data sampling and controls data classification stage. First, signal generator excited by the control circuit delivers a test voltage waveform to the system under test (SUT). It is assumed that the observed SUT is characterised by known parametric model and its time response to the test signal depends on

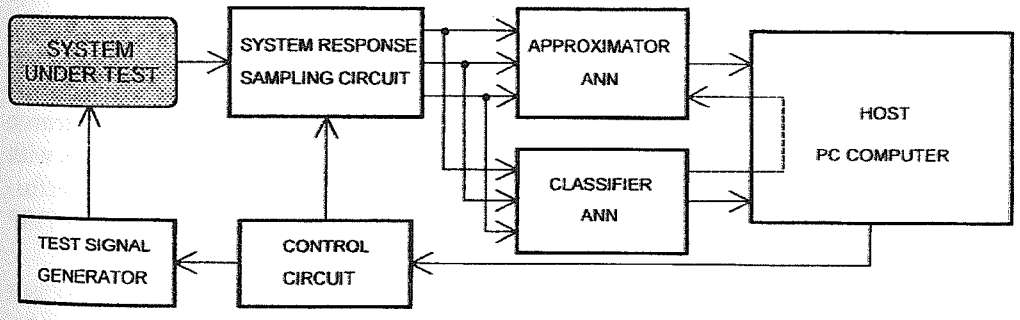


Fig. 7. Block diagram of ANN-based parameter identification system

model parameters. Control circuit determines also time moments of SUT response sampling.

The sampling circuit delivers the vector of SUT response samples to the input of the neural network approximator. Simultaneously, the same signals are fed to the analog classifier, which takes decision as to which weight set should be used for the parametr approximation stage. Classifier output is sampled by the host PC which loads appropriate weight vector to the approximator network. Finally, SUT parameters are estimated by the ANN approximator.

System parameter identification process is controlled and supervised by the host PC computer. The host PC programmes time moments of SUT response sampling, initialises SUT observation and data processing. It reads also A/D converted signal samples, classifier internal voltages and its output. It downloads approximator network architecture and appropriate weight vector to signal processor extension card, transfers SUT observations to the approximator network and reads estimated SUT parameters. Learning process of classifier and approximator ANNs is also controlled by the PC host computer. Two modes of the classifier learning are possible and will be tested in the prototype system. One is off-line training using a computer-simulated classifier and then mapping the weights on the chip. The second is a closed-loop training where the weights are updated on-chip but the weight increments are calculated by the host computer.

7. DIGITAL APPROXIMATOR

Digital realisation of the approximator ANN is preferred to achieve the required high accuracy of dynamic system (SUT) parameter identification. This approach guarantees reproducibility of obtained results. PC extension board containing Analog Devices ADSP21062 digital signal processor emulates the approximator network in the prototype system described. Advanced architecture of the digital signal processor makes it well suited for the task at hand.

Multilayer Perceptron (MLP) network has proven to provide good parameter estimation results in computer simulations [9], [10], [11]. Hardware implementation of such a network is planned. For a given system, separate MLP networks will be used for each parameter estimation of any considered SUT.

Network learning process will be implemented off-line in the host PC computer. Approximator will be trained using the backpropagation algorithm in first experiments. Furthermore, genetic algorithms combined with selected gradient search optimisation method are planned to be tested. Comparison of these approaches with selected numerical optimisation methods of multivariable functions will be performed. The most efficient network training method will be mapped onto the DSP board.

Analog Devices ADSP-21062 SHARC is a very powerful signal processor possessing wide set of features making it well suited for neural network modelling. The 32-bit floating-point computation units: multiplier, ALU and shifter make it

capable of multiplication and accumulation of two wide range arguments in a single machine cycle. This feature is usefull for two vectors dot product calculation which is very common operation in ANN processing. Dual-ported on-chip static RAM can be simultaneously accessed by processor's arithmetic logic unit and its peripheral devices. This solution allows for very fast exchange of modelled neural network weight vectors and other parameters. Host port and multiprocessor interface implemented in ADSP-21062 speeds up information exchange with other devices and allow parallel data processing by several processors. Internal timer can help in time-dependent processes realisation.

As an alternative to MLP, rational function [15] approximator architecture is also considered to be implemented.

8. ANALOG CLASSIFIER

Data sampling and classification subsystems will be placed on a single extension card to simplify control and synchronisation of signal processing stages. Schematic diagram od these systems is presented in Fig. 8. ight-gray colored rectangles represent data acquisition subsystem components. It samples SUT response $y(t)$ to the test signal and stores the quantized samples in the observation digital memory. Analog sample values Y_N are also available on the subsystem outputs and serve as input vector to analog classifier.

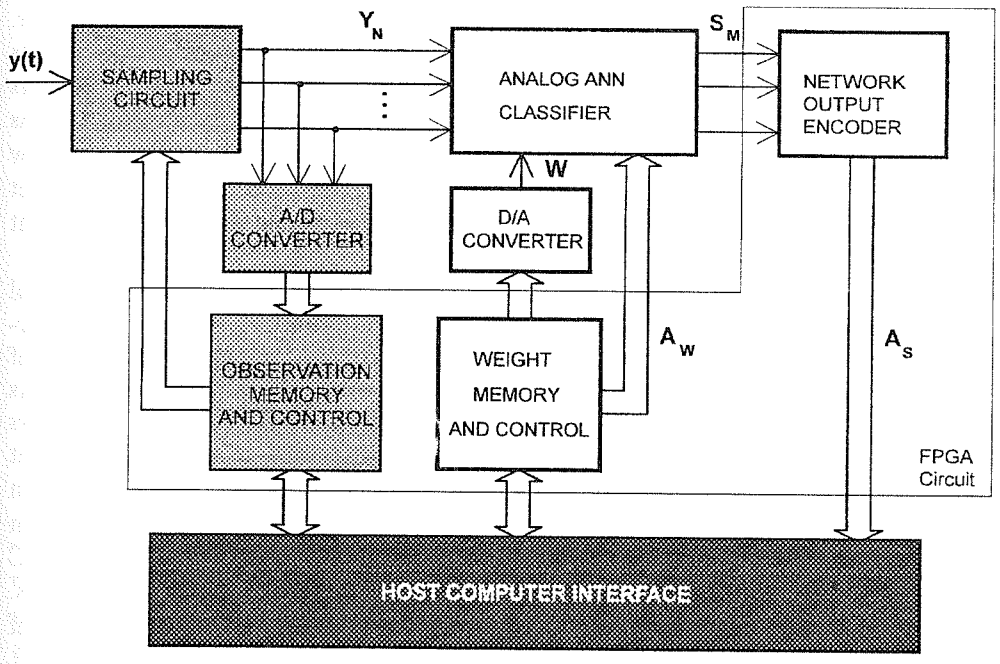


Fig. 8. Data acquisition and classification subsystems

White color blocks in Fig. 8 for the classification subsystem. The ANN classifier realised as analog Kohonen-type network requires weight refreshing circuit which contains sequentially addressed weight memory and D/A converter. Weight value W and its address A_W are simultaneously fed to the ANN classifier. Network output S_M is coded in output encoder into binary word A_S, which can be read by the host computer. The host PC computer interface circuit is represent by a dark gray rectangle in Fig. 8.

The task of the classifier network is to allocate given observation vector to one of the M regions. It selects the appropriate weight vector for the approximator module. A Kohenen-type classifier is proposed as a Learning Vector Quantizer, where splitting the observation damain into regions is performed by the neural network itself. A block diagram of this network, based on circuit described in [17] is shown in Fig. 9. The network input is an N-element vector, connected to the input layer. Each neuron in this layer distributes buffered signal to M computing units in competitive layer through the synapse array. The programmable synapse array is composed of NxM synapse cells. It corresponds to NM-elements array of weight vectors.

To reduce the number of chip inputs, the address decoder for synapse array is implemented providing the number of $\log_2(NM)$ instead of NM inputs. Each of M computing neurons calculates square of Euclidean distance between its weight vector and the input vector. The competitive process of selecting on of the neurons in performed for the whole layer by the winner-takes-all operation. The winner-takes-all cirucit contains M competitive circuits which perform paralleled comparison between M square of Euclidean distance values and then choose a single winner. The wining neuron is determined according to the minimum distance between its weight vector and the input vector. In the learning phase, the synapse weights corresponding to the wining neuron are updated according to the learning rule as specified in [17]. Weight update is performed by an external circuit. Next, the new weight values are stored in the synapse array. During the recall phase, each input vector activates one of M output neurons. This allows to load appropriate weight vector into the approximator network.

Control and weight refreshing of analog classifier will be performed by a digital FPGA circuit. It will implement also observation sampling control circuit, observation memory and a part of the PC host interface. XILINX XC4000 series of FPGA devices are being used. Their specific features make them well suited to this problem. XC4000 series features dual ported memory and reloadable counters implementation. Capability of design decomposition into several chips will make it independent of a specific chip type. Topology of digital circuit implemented in FPGA is stored in its internal RAM memory and must be downloaded each time on power up. Programming of FPGA configuration will be performed by the PC computer. PC interface is being designed using low-scale-integration PLD devices. They will realise serial FPGA programming interface and assure proper location of digital control circuit registers in the PC I/O space.

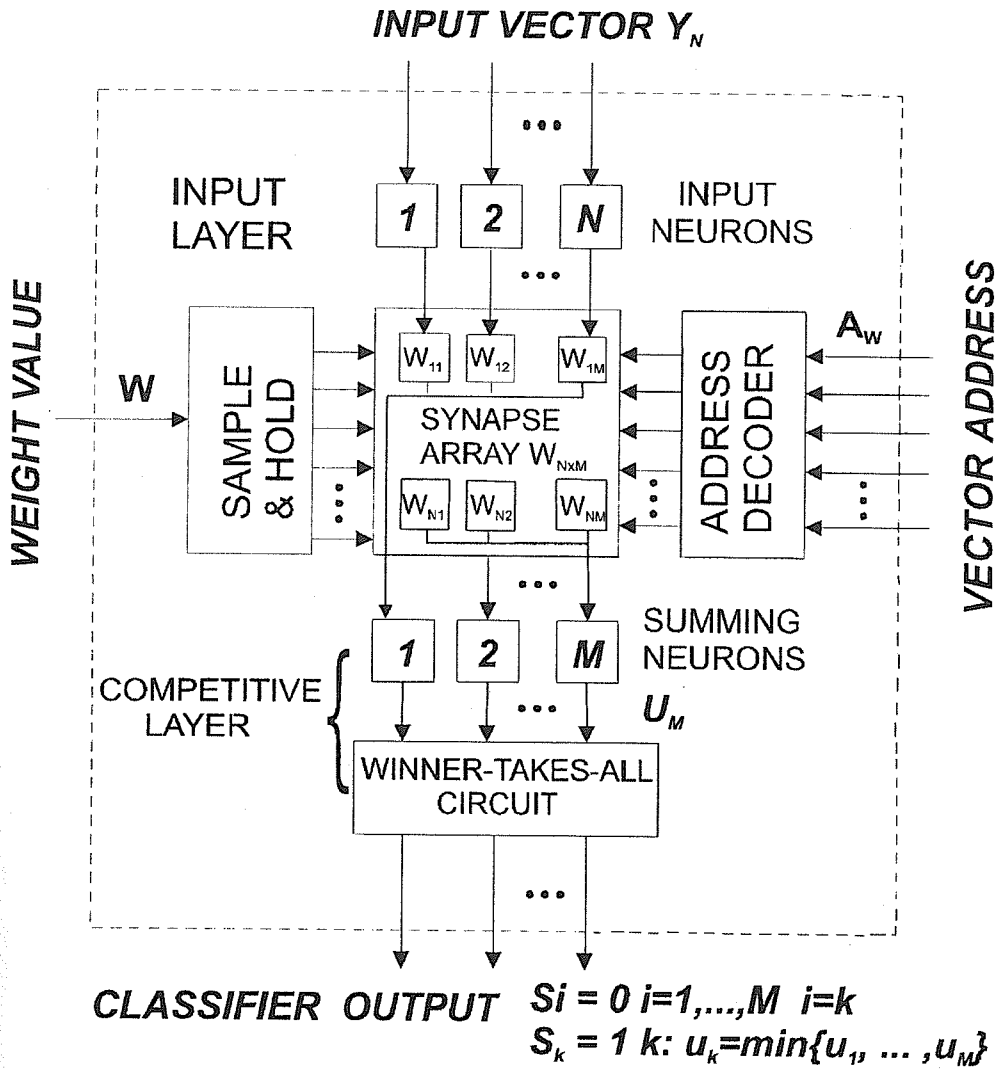


Fig. 9. Block diagram of Kohonen-type neural network analog classifier realisation

9. PRELIMINARY MEASUREMENT OF THE PROTOTYPE SYSTEM

First measurements of the constructed system performance have been carried out. Tests of analog input signal path including A/D conversion showed small noise, not exceeding 2 LSB for 14-bit converter. Fig. 10 illustrates the measured noise signal at the output of A/D converter while the input was shortcircuited to the ground. Subsystem connections and cooperation of different system blocks were also tested out.

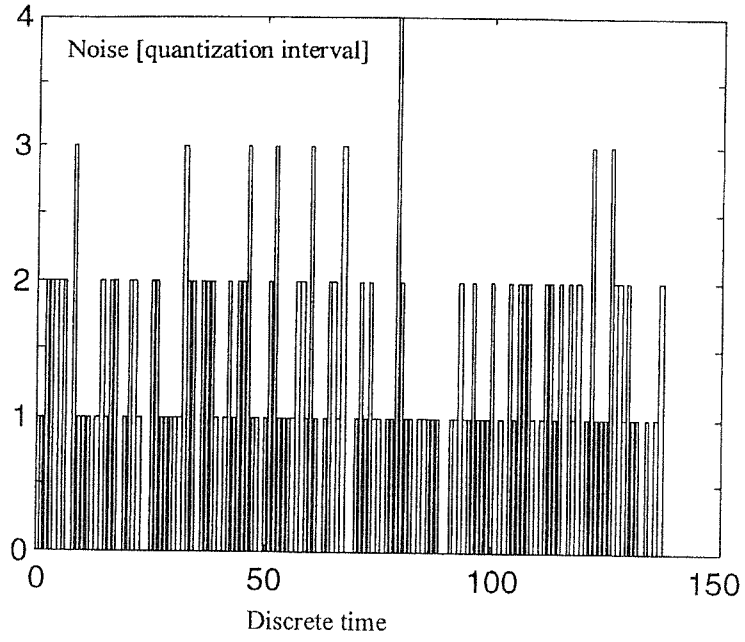


Fig. 10. Measured noise signal at the output of A/D converter

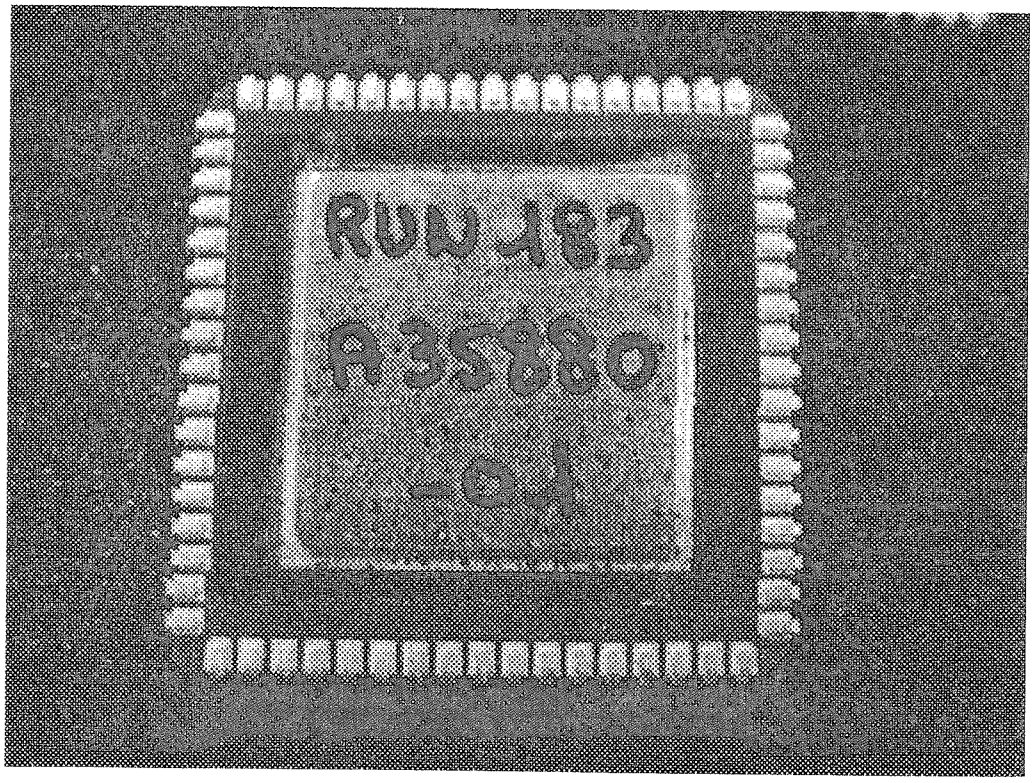


Fig. 11. Three-input, 8-neuron Kohonen neural network chip
(manufactured as a part of KBN grant 8T11F01010)

As an alternative to Kohonen classifier, a discrete analog perceptron network with fixed weight values was built. A comparison of the two networks performance is planned.

Then chips of Kohonen classifier were manufactuired in 2 μm analog CMOS technology. Fig. 11 shows a photo of one of them. As concluded after preliminary testing, 80% of them were fabricated with no detectable technological defects. An example of measured transfer characteristics of an on-chip Gilbert multiplier circuit of a neurone synapse is illustrated in Fig. 12. Functional measurements of the constructed prototype parameter identification system have been already started. The results of them will be published in a separate paper.

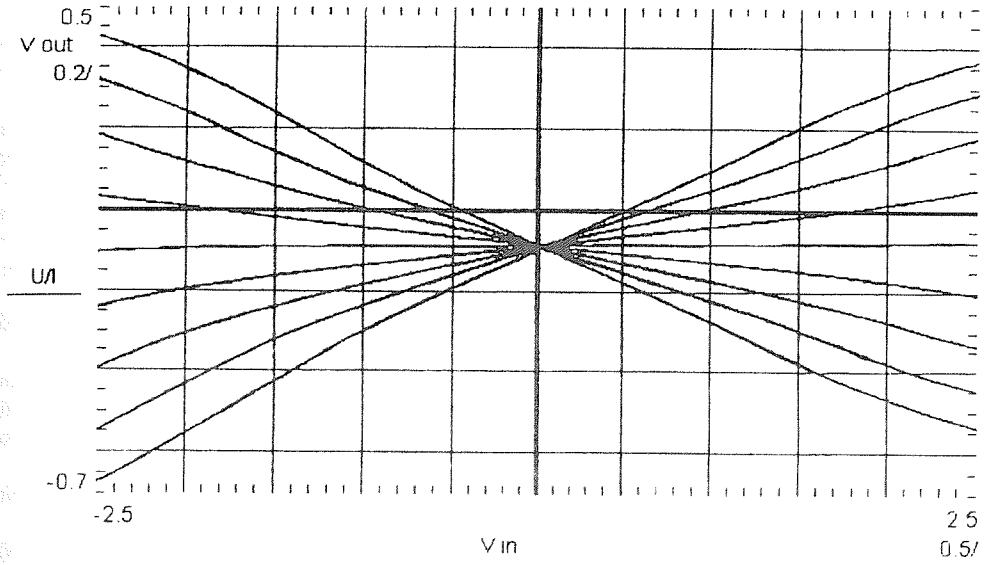


Fig. 12. Measured transfer characteristics of an example Gilbert multiplier circuit

10. DISCUSSION AND FINAL CONCLUSIONS

The prototype design of ANN system for parameter identification was presented in this paper. The role and operation of its functional blocks were discussed. The system is now under CAD design and simulation tests. Once have been built, the prototype ANN mixed-signal system will serve as a tool for experimental verification of the novel, recently proposed technique for fast and robust model parameter identification [6]. Extensive tests are planned using a variety of system under test, including parametric testing of CMOS circuits based on power supply transient responses [11]. The effect of observation noise on the identification process accuracy will also be experimentally investigated to verify results of numerical simulation

[6]–[11], [14], [16]. It is expected that a significant advance to the information processing technology will be made by introducing the proposed ANN-based parameter identification technique into practice.

ACKNOWLEDGMENT

The authors wish to thank Prof. Alexis De Vos for cooperation in designing and manufacturing the ANN chips. Mr Piotr Romaniuk for delivering the artwork of Fig. 10 and Mr Jacek Kowalski for providing data for Fig. 12.

REFERENCES

1. J. Beck, K. Arnold: *Parameter estimation in engineering and science.* J. Wiley, 1977
2. A. Atkinson, A. Donev: *Optimum experiment design.* Oxford Science 1992
3. S. Kay: *Fundamentals of statistical signal processing: estimation theory.* Prentice Hall, 1993
4. D. Feng, S-C. Huang, Z-Z. Wang and D. Ho: *An unbiased parametric imaging algorithm for nonuniformaly sampled biomedical system parameter estimation.* IEEE Trans. Med. Imaging, 1996, vol. 15, no. 4, pp. 512–518
5. K. Hornik: *Some new results on neural network approximation.* Neural Networks, 1993, vol. 6, pp. 1069–1072
6. A. Materka: *Application of neural networks for dynamic system parameter estimation.* 14th Int. Conf. IEEE Eng. Med. Biol. Soc. Paris 1992, pp. 1042–1045
7. B. Kaleba, Z. Lichtenstein, T. Simchony: *Linear and nonlinear associative memories for parameter estimation.* Information Sciences, 1992, vol. 61, no. 1, pp. 45–66
8. B. Tawfik, D. Durrand: *Nonlinear parameter estimation by linear association: application to five-parameter passive neuron model.* IEEE Trans. Biomed. Eng., 1994, vol. 41, no. 5, pp. 461–469
9. A. Materka, S. Mizushina: *Parametric signal restoration using artificial neural networks.* IEEE Trans. Biomed. eng., 1996, vol. 43, no. 4, pp. 357–372
10. A. Materka: *Neural network technique for parametric testing of mixed-signal circuits.* Electronics Letters, 1995, vol. 31, no. 3, pp. 183–184
11. A. Materka, M. Strzelecki: *Parametric testing of mixed-signal circuits by ANN processing of transient responses.* Journal of Electronic Testing: Theory and Applications, 1996, vol. 9, no. 7, pp. 1–16
12. A. Annaswamy, S-H. Yu: *0-adaptive neural networks: a new approach to parameter estimation.* IEEE Trans. Neural Networks, 1996, vol. 7, no. 4, pp. 907–918
13. A. Bahil: *Bio-engineering.* Prentice-Hall, 1981
14. A. Materka: *Application of artificial neural networks to parameter estimation of dynamical systems.* IEEE Conf. Meas. Instrum, Hamamatsu, Japan 1994, pp. 123–126
15. H. Leung, S. Haykin: *Rational function neural network.* Neural Comput., 1993, vol. 5, pp. 928–938
16. A. Materka: *Classifier-approximator modular neural for accurate estimation of dynamic system parameters.* Appl. Math. And Comp. Sci., 1995, vol. 6, no. 3, pp. 447–461
17. W. Fang, B. Sheu, O. Chen, J. Choi: *A VLSI Neural Processor for Image data Compression Using Self-Organization Neural Networks.* IEEE Trans. on Neural Networks, May 1992, vol. 3, no. 3, pp. 506–517

A. MATERKA, P. PEŁCZYŃSKI, M. STRZELECKI

ANALOGOWO-CYFROWY PROTOTYPOWY SYSTEM DO IDENTYFIKACJI PARAMETRÓW MODELI Z WYKORZYSTANIEM SIECI NEURONOWYCH

Streszczenie

W artykule przedstawiono metodę identyfikacji parametrów układu dynamicznego z wykorzystaniem sieci neuronowych. Metoda ta zapewnia dużą szybkość identyfikacji oraz niewielkie wartości błędów związanych z zakłóceniami szumowymi w porównaniu z metodami tradycyjnymi. Z drugiej strony proces uczenia sieci staje się czasochłonny przy rosnącej liczbie estymowanych parametrów i jednoczesnym zachowaniu dużej dokładności estymacji. Dlatego zaproponowano zastosowanie modularnego układu sieci o strukturze klasyfikator-aproksymator. Działanie takiej sieci wyjaśniają przedstawione przykładowe symulacje komputerowe. Zaprezentowano również projekt realizacji sprzętowej prototypowego układu modularnej sieci do identyfikacji parametrów. Układ wykonano w postaci dwóch kart w standardzie ISA współpracujących z komputerem klasy IBM PC. Zawierają one m.in. analogową sieć wykorzystywaną do klasyfikacji. Do sterowania sieci oraz komunikacji z komputerem wykorzystano układy FPGA. Proces uczenia sieci jest realizowany z wykorzystaniem komputera. Zaprojektowano i wykonano 10 układów scalonych w technologii 2 μm CMOS, realizujących funkcję klasyfikatora Kohonena. Przeprowadzono pełne testowanie kart rozszerzających. Rozpoczęto badanie funkcjonalne całego systemu.

Słowa kluczowe: identyfikacja parametrów, sieci neuronowe, przetwarzanie sygnałów, analogowy klasyfikator.

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI, 1998, 44, z. 3, ss. 391–412

Analiza parametrów czasowych architektury potokowej specjalizowanych procesorów sprzętowych

KAZIMIERZ WIATR

Instytut Elektroniki, Akademia Górniczo-Hutnicza w Krakowie

Artykuł dotyczy zagadnień związanych z poszukiwaniem nowych jakościowo metod zwiększania szybkości wstępnego przetwarzania obrazów dla potrzeb szczególnej klasy systemów wizyjnych, będących częścią systemów sensorycznych, stosowanych dla potrzeb automatyki i sterowania w czasie rzeczywistym. Poszukiwanie tych metod polegało na dążeniu do zrównoleglenia obliczeń wstępnego przetwarzania obrazów, poprzez dobór takich elementów obliczeniowych (procesorów) i takiej transmisji pomiędzy nimi (architektura), aby całkowity czas obliczeń był możliwie najmniejszy.

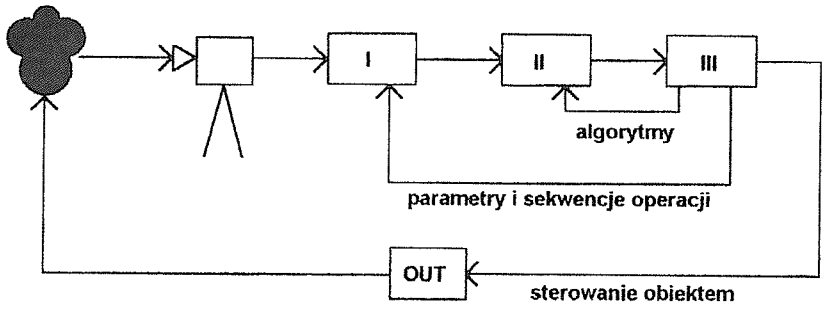
W opracowaniu przeprowadzono rozważania teoretyczne, dotyczące różnych architektur równoległej pracy specjalizowanych procesorów sprzętowych, ze szczególnym uwzględnieniem parametrów czasowych poszczególnych rozwiązań. Uzasadniono wybór architektury potokowej specjalizowanych procesorów sprzętowych. Przedstawione koncpecje zostały zweryfikowane w badaniach eksperymentalnych, ze szczególnym uwzględnieniem parametrów implementacji procesorów w układach FPGA. Ważnym aspektem przedstawionych badań jest fakt, że zostały one ukierunkowane w nurt światowych prac badawczych na polu dedykowanych dla użytkownika struktur obliczeniowych (ang. *Custom Computing Machines*).

W artykule przedstawiono analizę parametrów czasowych dotyczącą wyboru architektury jednostki wieloprocesorowej do szybkiej realizacji zadań wstępnego przetwarzania obrazów wizyjnych. Wykorzystując wcześniejsze doświadczenia autora w zakresie systemów czasu rzeczywistego zaproponowano zaimplementowanie w architekturze potokowej specjalizowanych procesorów sprzętowych zbudowanych w oparciu o układy programowalne FPGA. Przedstawiono także wyniki prac eksperymentalnych ze szczególnym uwzględnieniem czasów opóźnień dla poszczególnych operacji.

Słowa kluczowe: systemy czasu rzeczywistego, architektury systemów wieloprocesorowych, układy programowalne.

1. POZIOMY PRZETWARZANIA I ANALIZY OBRAZÓW

W algorytmach analizy obrazów można wyróżnić kilka poziomów analizy i przetwarzania obrazu. Najczęściej przyjmuje się trzy poziomy analizy obrazów (rys. 1):



Rys. 1. Poziomy analizy obrazów systemu wizyjnego czasu rzeczywistego

I — Najniższy poziom analizy obrazów nazywany jest często wstępnym przetwarzaniem sygnału wizyjnego. Ma on na celu: eliminację zakłóceń, wydobywanie obiektu z tła, detekcję krawędzi, ustalanie poziomów szarości pikseli należących do interesującego nas obiektu na podstawie histogramu, równoważenie histogramu itp.

II — Średni poziom analizy obrazów obejmuje: segmentację obrazu, lokalizację obiektów, rozpoznanie kształtu obiektu i wyróżnienie cech charakterystycznych tego kształtu.

III — Najwyższy poziom przeprowadza analizę złożonej sceny w sensie analizy ruchu obiektu i na bieżąco steruje obiektem oraz pozostałymi poziomami systemu wizyjnego (zadaje algorytmy oraz parametry i sekwencje operacji).

Przedstawiona na rys. 1 struktura systemu wizyjnego ukazuje pnadto występujące powiązania (sprzężenia) pomiedzy różnymi poziomami systemu wizyjnego. Wyniki II i III etapu przetwarzania obrazu mogą mieć wpływ na parametry i kolejność wykonywanych operacji na I najniższym poziomie oraz na algorytmy przetwarzania na II poziomie. Właściwość ta znajdzie zastosowanie w dyskutowanych rozwiązaniach i konkretnych propozycjach autora, w szczególności dotyczących rekonfigurowalnych struktur hardware'owych. Nowe możliwości w tym zakresie, jakie dają układy programowalne wysokiej skali integracji typu FPGA o konfiguracji pamiętanej w wewnętrznej pamięci SRAM, pozwalają na zmianę na bieżąco (ang. *on-line*) zarówno parametrów przetwarzania, jak również algorytmów przetwarzania, co wykorzystano w opisanych dalej oryginalnych rozwiązaniach strukturalnych, zaproponowanych dla systemów wizyjnych w robotyce.

Zadania najniższego poziomu systemu wizyjnego, realizującego wstępne przetwarzanie obrazu, realizowane są jako operacje jednopunktowe (przekodowanie LUT, normalizacje i skalowania, wyrównywanie histogramu, binaryzacja, operacje na dwu obrazach itp). lub jako operacje kontekstowe (konwolucyjne filtry dolnoprzepustowe i górnoprzepustowe, konwolucyjne detektory krawędzi, filtracja medianowa i logiczna, operacje morfologiczne itp.)

2. PRZETWARZANIE OBRAZU W CZASIE RZECZYWISTYM

Istotne znaczenie dla operacji przetwarzania obrazów ma aspekt pojęcia czasu rzeczywistego. Pojęcie to ma bowiem dość zróżnicowane definicje, które ściśle wiążą się z rodzajem obiektu lub procesu, z jakim są związane.

Czas upływa i jest wymiarem, na który ludzka działalność nie ma żadnego wpływu. Dla przedstawionych w tym opracowaniu rozważań ważna jest skala czasu, tzn. jak duże są przedziały czasu ważne z punktu widzenia interesującego nas obiektu lub procesu. W szczególności interesuje nas system mikrokomputerowy wyposażony w kamerę wizyjną i sterujący na podstawie otrzymywanych z niej obrazów obiektem obserwowanym (sterujący system wizyjny). Dla takiego systemu funkcjonowanie w czasie rzeczywistym oznacza, że podejmowane przez system decyzje realizowane są w tempie obsługiwania tego procesu. Można to wyrazić stwierdzeniem, że sterujący system wizyjny działa w czasie rzeczywistym, jeżeli czas reakcji systemu jest niezauważalny przez proces technologiczny. Oznacza to, że opóźnienie pomiędzy zmianami przebiegającymi na obserwowanej scenie a konieczna odpowiedzią systemu sterującego, wynikającą z algorytmu sterowania, jest niewielkie czyli niezauważalne. Często wobec takich systemów używa się określeń: *system o działaniu bezpośrednim* lub *system pracujący na bieżąco (on line)*.

W tym kontekście pojęcie czasu rzeczywistego jest pojęciem względnym. Można to zilustrować przykładami ukazującymi różnorodne systemy sterujące czasu rzeczywistego, tzn. z bardzo różnymi skalami czasu. W szczególności mogą to być systemy stosunkowo wolne (np. sterowanie pojedynczym skrzyżowaniem ulic — dynamika rzędu kilkunastu sekund jest zupełnie wystarczająca), systemy średniej szybkości (np. przemysł maszynowy: obrabiarki, walcownie — wymagane czasy reakcji rzędu 0,1 s) i systemy bardzo szybkie (np. systemy militarne, bardzo szybka robotyka — wymagane czasy liczone w ms). Zatem dla potrzeb niniejszego opracowania możemy przyjąć, że system czasu rzeczywistego przeznaczony do sterowania procesów, to taki system, w którym przetwarzanie zmiennych procesowych będzie realizowane z taką szybkością, aby wynik tego przetwarzania mógł być użyty do skutecznego sterowania tym procesem. Gdyby powyższe czasy reakcji systemu sterującego porównać z czasami obliczeń związanych z analizą obrazów wykonywanych przez klasyczne systemy (np. klasy IBM PC), to okaże się, że opóźnienia wnoszone przez takie systemy są wielokrotnie większe od dopuszczalnych.

Dla ilustracji można przytoczyć kilka pomiarów przeprowadzonych przez autora u początków referowanych w tym opracowaniu badań (rok 1991):

1. obliczenie mediany w oknie 3×3 piksele dla obrazu 512×512 pikseli:
   - IBM-PC 386SX/20 MHz — 407 s,
   - MC68030/20 MHz — 156 s,
2. tworzenie szkieletu dla obrazu 512×512 pikseli:
   - IBM-PC 386SX/20 MHz — 69 s,
   - MC68030/20 MHz — 14 s.

Dlatego uzasadnione jest poszukiwanie skutecznych metod szybkiego przetwarzania i analizy obrazów wizyjnych dla potrzeb systemów czasu rzeczywistego. We wszystkich wspomnianych wcześniej zastosowaniach wymagany czas jest stosunkowo krótki. Dlatego autor przyjął, że minimalny czas, w jakim można dokonywać wstępnego przetwarzania sygnału wizyjnego, związany jest z częstotliwością przetwarzania sygnału z kamery wizyjnej przez przetwornik A/C (wzór 1). Oznacza to podjęcie próby realizacji na bieżąco (*on-line*) operacji wstępnego przetwarzania obrazów wizyjnych (ang. *preprocessing*), w miarę napływania danych z przetwornika A/C (tzn. w czasie odpowiadającym czasowi trwania jednego piksela):

$$t_p = 1/14{,}75 \text{ MHz} = 67{,}8 \text{ ns} \tag{1}$$

Powyższa częstotliwość próbkowania przetwornika A/C wynika z potrzeby stosowania w systemach sensorycznych kwadratowego kształtu piksela dla zachowania parametrów obiektów związanych z ich kształtem [16]. Jeżeli czas ten okazałby się zbyt krótki, należy zrównoleglić obliczenia i/lub wprowadzić linie opóźniające i czas ten wydłużyć o całkowitą wielokrotność tej wartości.

3. PARAMETRY CZASOWE WSTĘPNEGO PRZETWARZANIA OBRAZU

Realizacja algorytmów wstępnego przetwarzania obrazów oznacza w praktyce wykonanie prostych operacji na bardzo dużej ilości danych. Wiąże się to z bardzo dużą liczbą przesłań typu pamięć – procesor – pamięć dla pobrania operandu i zapamiętania wyniku operacji. Szczególnie niekrorzystne w tym zakresie są operacje kontekstowe, dokonujące transformacji punktu wraz z jego otoczeniem w nowy punkt.
Wykonanie samych tylko przesłań wymaga znacznego czasu procesorów i znacznie obciąża (blokuje) magistralę systemu wieloprocesorowego. Ilość wielokrotnych przesłań operandów, dla przykładowych operacji na obrazie o rozdzielczości 512×512 pikseli i 8-bitowej precyzji poziomu szarości, przedstawiono w tabeli 1.

Tabela 1

Liczba przesłań operandów w operacjach wstępnego przetwarzania

Transformacja	Liczba przesłań na 1 piksel	Liczba przesłań na 1 obraz	Liczba przesłań w czasie 1 s
LUT	1	262 144	6 553 600
Mediana 5-punktowa	5	1 310 720	32 768 000
Konwolucja 3×3	9	2 359 296	58 982 400

Ograniczenie liczby przesłań to zadanie, którego rozwiązanie powoduje jakościową zmianę problemu szybkości realizacji tych obliczeń. Podobnie przedstawia się problem czasu realizacji stosunkowo prostych operacji na danych wizyjnych. Zastosowanie mechanizmu eliminacji lub ograniczenia tych problemów postawił sobie za cel autor niniejszego opracowania.

Tabela 2

Liczba prostych operacji we wstępnym przetwarzaniu obrazu

Transformacja	Liczba mnożeń	Liczba dodawań lub porównań	Liczba przesłań operandów i wyniku
LUT	0	6 553 600/s	13 107 200/s
Mediana:			
5-punktowa	0	65 536 000/s	39 321 600/s
9-punktowa	0	294 912 000/s	65 536 000/s
Konwolucja 3 × 3	58 984 400	52 428 800/s	65 536 000/s

W tej sytuacji uzasadniona jest potrzeba dokładnego określenia opóźnienia wnoszonego przez system wizyjny lub jego poszczególne elementy składowe. Zgodnie z rys. 1, odpowiedź systemu wizyjnego na zmiany zachodzące na obiekcie obserwowanym i jednocześnie sterowanym po przetworzeniu wejściowej informacji wizyjnej. Całkowity czas wnoszonego opóźnienia wyniesie:

$$t_{sys} = t_{akw} + t_I + t_{II} + t_{III} + t_{OUT} \tag{2}$$

gdzie: t_{sys} — opóźnienie wnoszone przez system, t_{akw} — opóźnienie wnoszone przez układ wzmacniaczy wejściowych i przetwornika A/C, t_I — opóźnienie wnoszone przez I poziom analizy obrazów (wstępne przetwarzanie), t_{II} — opóźnienie wnoszone przez II poziom analizy obrazów (segmentacja itp), t_{III} — opóźnienie wnoszone przez III poziom analizy obrazów (rozpoznawanie kształtów itp.), t_{OUT} — opóźnienie wnoszone przez układy wyjściowe sterujące obiektem.

Z punktu widzenia niniejszego opracowania ważny jest czas opóźnienia wnoszony przez element t_I ze wzoru (2) czyli wstępne przetwarzanie danych wizyjnych. Analizując dane zawarte w tabelach 1 i 2 łatwo dojść do wniosku, że wielkość tego opóźnienia jest zależna od dwu rodzajów wnoszonych opóźnień: opóźnienia wnoszongo przez element obliczeniowy i opóźnienia związanego z transmisją do/od elementu obliczeniowego danych/wyniku obliczeń, zgodnie ze wzorem (3):

$$t_I = t_{obl\_d} + t_{tr\_d} \tag{3}$$

gdzie: t_{obl_d} — czas związany z obliczeniami, t_{tr_d} — czas związany z transmisją danych.

Biorąc pod uwgę, że operacje obliczeń i transmisji wykonywane są wielokrotnie można zapis wzoru (3) uszczegółowić do postaci wzoru (4):

$$t_I \sum_{i=1}^{M} \sum_{j=1}^{a} \sum_{k=1}^{b} t_{obl\_d_{ijk}} + \sum_{i=1}^{M} \sum_{j=1}^{a} \sum_{l=1}^{c} t_{tr\_d_{ijl}}, \tag{4}$$

gdzie: M — liczba pikseli (np. 512 × 512), a — liczba operacji wstępnego przetwarzania, b — liczba operacji arytmetycznych związanych z obliczaniem nowej wartości piksela dla danej operacji, c — liczba transmisji danych (argumentów i wyniku) związanych z daną operacją dla jednego piksela.

Obliczony we wzorze (4) czas realizacji obliczeń I poziomu analizy obrazów czyli wstępnego przetwarzania danych wizyjnych nie uwzględnia czasu pobrania i dekodowania kodu rozkazów odpowiedzialnych za wykonanie programu związanego z realizacją operacji wstępnego przetwarzania. Należy zatem wzór (3) uzupełnić składnikami odpowiedzialnymi za transmisję i dekodowanie kodu instrukcji programu wstępnego przetwarzania.

$$t'_I = t_{obl\_d} + t_{tr\_d} + t_{tr\_roz} + t_{dek\_roz} \tag{5}$$

gdzie: t_{tr_roz} — czas związany z odczytywaniem kodów rozkazów, t_{dek_roz} — czas związamny z dekodowaniem rozkazów przez procesor.

Uwzględniając elementy składowe poszczególnych opóźnień możemy wzór (5) zapisać w postaci uwzględniającej rozmiary obrazu (liczbę pikseli), liczbę operacji wstępnego przetwarzania oraz elementarne czasy obliczeń i transmisji (wzór 6):

$$t'_I = \sum_{i=1}^{M}\sum_{j=1}^{a}\sum_{k=1}^{b} t_{obl\_d_{ijk}} + \sum_{i=1}^{M}\sum_{j=1}^{a}\sum_{l=1}^{c} t_{tr\_d_{ijl}} + \sum_{i=1}^{M}\sum_{j=1}^{a}\sum_{m=1}^{d} t_{tr\_roz_{ijm}} + \sum_{i=1}^{M}\sum_{j=1}^{a}\sum_{n=1}^{e} t_{dek\_roz_{ijn}} \tag{6}$$

gdzie: d — liczba transmisji związanych z pobieraniem kodów rozkazów i argumentów nie będących danymi wizyjnymi (np. adresy skoków) dla danej operacji, e — liczba rozkazów podlegających dekodowaniu dla danej operacji wstępnego przetwarzania.

Wzór (6) określa czas wykonania kilku operacji (liczba określona parametrem a) wstępnego przetwarzania obrazu. Jednym z celów niniejszego opracowania jest minimalizacja opóźnienia wnoszonego przez I etap przetwarzania obrazu czyli minimalizacja wartości czasu t_I. Można to osiągnąć poprzez minimalizację czasów realizacji obliczeń składowych t_{obl_d} (np. zwiększanie częstotliwości zegara procesora) lub minimalizację czasów transmisji danych i wyników obliczeń t_{tr_d} (bardziej wydajna magistrala) oraz minimalizację czasu transmisji i dekodowania rozkazów (podobnie jak poprzednio). Innym sposobem skrócenia tego czasu jest ograniczenie liczby operacji wstępnego przetwarzania (parametr a) lub rozdzielczości (parametr M).

Wzór (5) należałoby także uzupełnić opóźnieniami wnoszonymi przez układ akwizycji obrazu i układy transmisji przetworzonych danych do dalszego przetwarzania. W niektórych rozwiązaniach czas transmisji przetworzonych danych zawiera się w elemencie t_{tr_d} i w zależności od rozwiązania należy odpowiednio modyfikować wzór (7).

$$t''_I = t_{tr\_we} + t_{obl\_d} + t_{tr\_d} + t_{tr\_roz} + t_{dek\_roz} + t_{tr\_wy} \tag{7}$$

gdzie: t_{tr_we} — czas związany z przekazaniem cyfrowych danych wizyjnych do procesorów I etapu, t_{tr_wy} — czas związany z transmisją danych po przetworzeniu w I etapie do dalszej analizy.

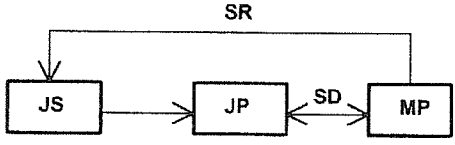
Zamiarem autora jest poszukiwanie optymalnej struktury i architektury logicznej systemu wieloprocesorowego do szybkiego wstępnego przetwarzania obrazu w czasie rzeczywistym, opartej na analizie powyższych parametrów czasowych, tak aby

czas t_I był minimalny, z jednoczesnym zachowaniem dotychczasowych parametrów tych obliczeń (tzn. minimalizacja czasu t_I z zachowaniem wcześniejszej rozdzielczości i liczby operacji wstępnego przetwarzani). W tym celu w dalszych rozważaniach przedstawione zostaną dostępne architektury systemów wieloprocesorowych.

4. ARCHITEKTURY SYSTEMÓW KOMPUTEROWYCH

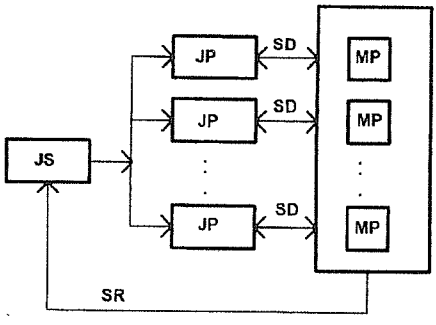
Jedna z pierwszych klasyfikacji architektur systemów komputerowych została przeprowadzona przez M. Flynna [3]. W modelu przyjętym przez Flynna wydzielono w systemie: jednostki przetwarzające JP *Jednostka Procesora), jednostki sterujące JS (Jednostka Sterująca) oraz pamięć MP (Moduł Pamięci). Podstawą tej klasyfikacji jest liczba strumieni danych SD (Strumień Danych) i liczba strumieni rozkazów SR (Strumień Rozkazów). W tym ujęciu wyróżnione są cztery klasy architektur systemów.

Pierwszą z nich jest architektura SISD (ang. *Single Instruction-stream Single Data-stream*) charakteryzująca się pojedynczym strumieniem rozkazów i pojedynczym strumieniem danych. Organizacja SISD charakteryzuje typowe komputery, w których rozkazy są wykonywane sekwencyjnie (rys. 2).

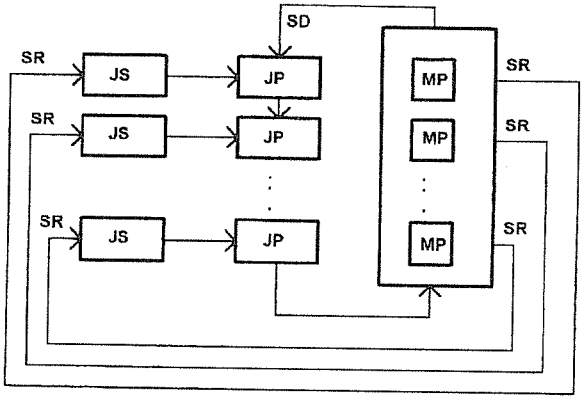


Rys. 2. Architektura systemów typu SISD

Kolejną jest architektura SIMD (ang. *Single Instruction-stream Multiple Data-stream*) charakteryzująca się pojedynczym strumieniem rozkazów i wielokrotnym strumieniem danych. W architekturze tej wiele jednostek przetwarzających JP przetwarza oddzielne strumienie danych. Wszystkie JP są podporządkowane jednej jednostce sterującej JS (jeden strumień rozkazów). Organizacja SIMD (rys. 3) charakteryzuje komputery macierzowe.



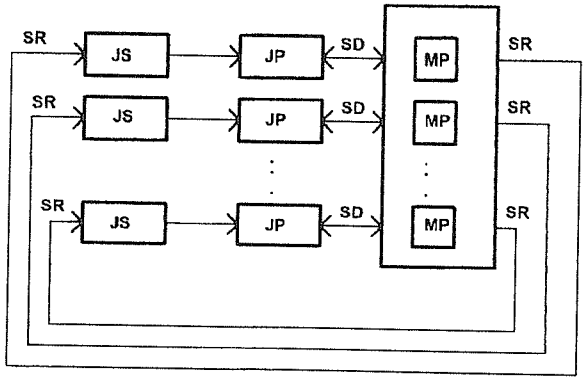
Rys. 3. Architektura systemów typu SIMD



Rys. 4. Architektura systemów typu MISD

Architektura MISD (ang. *Multiple Instruction-stream Single Data-stream*) posiada wielokrotny strumień rozkazów i pojedynczy strumień danych. Według literatury [8] wprowadzenie tej klasy architektur zostało podyktowane głównie kompletnością klasyfikacji, bowiem trudno znaleźć system komputerowy, który jej odpowiada. Według innych autorów architekturze MISD można przyporządkować m.in. komputery systoliczne — w szczególności znany w zastosowaniach obrazowych cytokomputer. W architekturze tej wiele jednostek sterujących JS steruje pracą wielu jednostek przetwarzających JP (wielokrotny strumień rozkazów). Natomiast dane stanowią jeden strumień przekazywany przez jeden procesor JP do następngo (pojedynczy strumień danych). Architektura taka jest przedstawiona na rys. 4.

Architektura typu MIMD (ang. *Multiple Instruction-stream Multiple Data-stream*) charakteryzuje się wielokrotnym strumieniem rozkazów i wielokrotnym strumieniem danych. Taka architektura wykorzystywana jest w dużych systemach — przy czym zdarzają się systemy ściśle i luźno połączone. Schemat blokowy tej architektury przedstawia rys. 5.

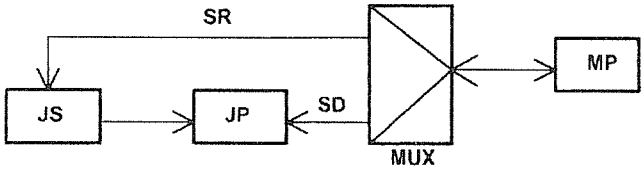


Rys. 5. Architektura systemów typu MIMD

5. WYBÓR ARCHITEKTURY

Analizując przydatność powyżej przedstawionych architektur należy porównać przepływ strumieni danych i rozkazów w systemie wizyjnym i w tych architekturach.

Architektura SISD z rys. 2 w klasycznych rozwiązaniach posiada wspólny dostęp do pamięci danych i rozkazów. Z tego powodu zmodyfikowano rys. 2 wprowadzając multiplekser rozdzielający informacje z modułu pamięci pomiędzy jednostkę sterującą i jednostkę procesora (rys. 6). W rzeczywistych układach multiplekser taki



Rys. 6. Architektura systemów typu SISD z klasycznym modułem pamięci MP

znajduje się w każdej strukturze mikroprocesora. Czas realizacji zadanych do obliczeń algorytmów dla takiego systemu można zapisać odpowiednio modyfikując wzor (6), ponieważ czas $t_{tr\_d} = t_{tr\_roz} = t_{d\_MP}$. Można zatem zapisać:

$$t'_{I_{SISD}} = \sum_{i=1}^{M}\sum_{j=1}^{a}\sum_{k=1}^{b} t_{obl\_d_{ijk}} + \sum_{i=1}^{M}\sum_{j=1}^{a}\sum_{l=1}^{c} t_{d\_MP_{ijl}} \sum_{i=1}^{M}\sum_{j=1}^{a}\sum_{m=1}^{d} t_{d\_MP_{ijm}} + \sum_{i=1}^{M}\sum_{j=1}^{a}\sum_{n=1}^{e} t_{dek\_roz_{ijn}} \quad (8)$$

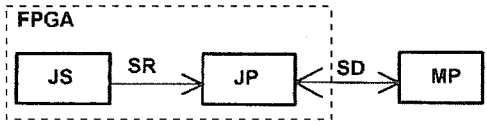
gdzie: t_{d_MP} — jest czasem dostępu do modułu pamięci czyli do zasobów pamiętających dane i rozkazy.

W nowoczesnych mikroprocesorach dekodowanie rozkazów odbywa się równolegle do transmisji następnych bajtów danych i dlatego z dużym przybliżeniem dla współczesnych systemów SISD wzór (8) można zapisać w postaci (9):

$$t''_{I_{SISD}} = \sum_{i=1}^{M}\sum_{j=1}^{a}\sum_{k=1}^{b} t_{obl\_d_{ijk}} + \sum_{i=1}^{M}\sum_{j=1}^{a}\sum_{l=1}^{c} t_{d\_MP_{ijl}} + \sum_{i=1}^{M}\sum_{j=1}^{a}\sum_{m=1}^{d} t_{d\_MP_{ijm}} \quad (9)$$

W systemach wizyjnych czasu rzeczywistego zależy nam, aby czas t_I był jak najmniejszy. Można oczywiście dążyć do minimalizowania poszczególnych składników wzoru (8) lub (9). Prace takie są prowadzone i wynikają z technologicznego rozwoju mikroprocesorów i innych komponentów systemów obliczeniowych, prowadząc jednocześnie także do poprawy parametrów systemów wizyjnych. Zamiarem autora jest jednak poszukiwanie innych jakościowo metod szybkiego przetwarzania wstępnego danych wizyjnych w czasie rzeczywistym, poprzez dobór takich elementów obliczeniowych (procesory) i takiego ich połączenia (architektura), które pozwoliłyby składniki te w istotny sposób ograniczyć lub całkowicie wyeliminować.

W pierwszym rzędzie należałoby ograniczyć lub wyeliminować czas związany z odczytem i dekodowaniem rozkazów programu realizującego proste operacje wstępnego przetwarzania. Są to składniki $t_{tr\_roz_{ijm}}$ i $t_{dek\_roz_{ijn}}$ we wzorze (6). Można to



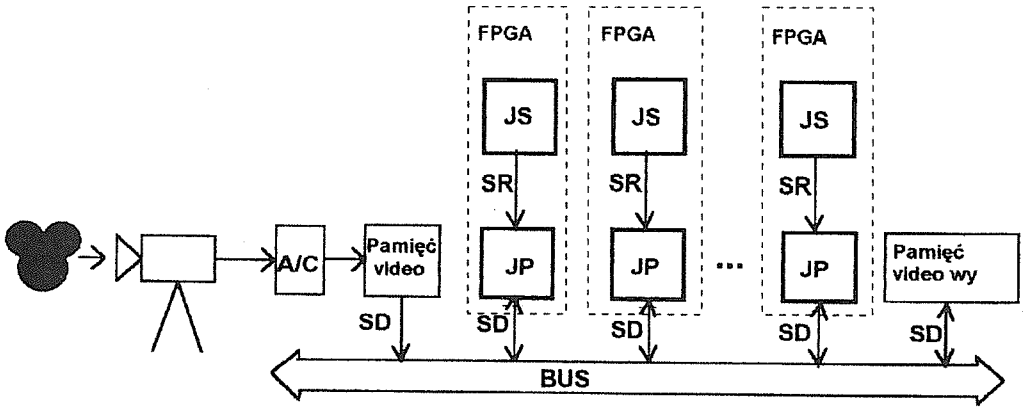
Rys. 7. Architektura systemu SISD pozbawionego transmisji kodów rozkazów

osiągnąć wprowadzając specjalizowane procesory sprzętowe, które w swojej strukturze logicznej będą miały zakodowany program prostych operacji wstępnego przetwarzania. Architektura takiego systemu została przedstawiona na rys. 7, na którym zaznaczono jako jeden moduł strukturę FPGA, w której zawiera się jednostka procesora JP i jednostka sterująca JS. Specjalizowane procesory sprzętowe bardzo korzystnie jest implementować w strukturach FPGA [16, 20].
Wówczas czas wykonania operacji wstępnego przetwarzania danych wizyjnych ograniczy się do jednego składnika ze wzoru (6):

$$t'''_{T_{SISD}} = \sum_{i=1}^{M} \sum_{j=1}^{a} \sum_{k=1}^{b} t_{obl\_d_{ijk}} \tag{10}$$

Połączenie kilku takich procesorów wymaga odpowiedniej ich synchronizacji, aby po wykonaniu pierwszej operacji wstępnego przetwarzania obrazu pierwszy specjalizowany procesor był dezaktywowany i jednocześnie uaktywniał się drugi itd. Wszystkie procesory korzystają bowiem ze wspólnej pamięci. Wprowadzenie kilku zestawów pamięci porządkuje zawarte w nich dane wizyjne, niemniej nie przyspiesza pracy procesorów przyłączonych do wspólnej magistrali. Architekturę takiego systemu przedstawia rys. 8.

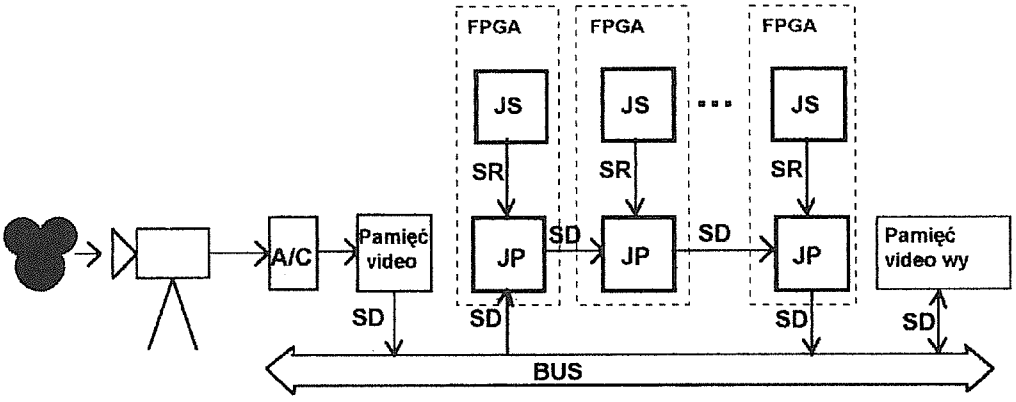
Architektura systemu z rys. 8 bardzo przypomina architekturę systemu MISD z rys. 4. Jest to oczywiście jedynie podobieństwo graficzne. Na rys. 4 poszczególne strumienie danych SD i strumienie rozkazów SR są aktywne jednocześnie, natomiast



Rys. 8. Architektura systemu z procesorami SISD

w systemie z rys. 8 czas magistrali i pamięci są dzielone pomiędzy poszczególne procesory tak, że w danym przedziale czasu aktywny jest tylko jeden procesor (jeden strumień danych SD i jeden strumień rozkazów SR).

Dalsze poszukiwania ograniczenia czasu realizacji I etapu przetwarzania danych wizyjnych prowadzą do próby eliminacji ostatniego dużego składnika we wzorze (6), odpowiedzialnego za transmisję danych wizyjnych pomiędzy procesorami (t_{tr_dijl}). Byłoby to możliwe, jeżeli procesory z rys. 8 pracowałyby synchronicznie, a dane mogłyby być przekazywane z wyjścia pierwszego procesora na wejście drugiego itd. Sytuację taką prezentuje rys. 9. Następuje tu kolizja w dostępie do magistrali systemowej (BUS) przez pierwszy i ostatni procesor (JP). Jednak przy większej liczbie procesorów zyski czasowe są znaczne, a problem dostępu do magistrali można rozwiązać innymi metodami.



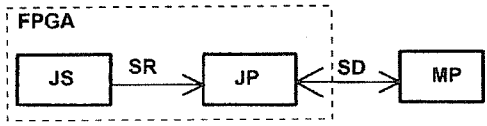
Rys. 9. Architektura MISD specjalizowanych procesorów sprzętowych

Przedstawiona na rys. 9 architektura jest typowym przykładem architektury MISD według klasyfikacji zaproponowanej przez Flynna, bowiem wielokrotny strumień rozkazów SR pracuje jednocześnie (w tym samym przedziale czasu). Architekturę uwidocznioną na rys. 9 można sprowadzić do postaci graficznej bardziej przypominającej to, co zaproponował w swojej klasyfikacji Flynn (rys. 10).

Łatwo można zauważyć, że architektura taka nie wprowadza do czasu obliczeń opóźnienia związanego z odczytywaniem i dekodowaniem kodu rozkazu. W tej sytuacji należy zmodyfikować wzór (6) na czas realizacji obliczeń wstępnego przetwarzania danych wizyjnych dla tej architektury, który przyjmie postać:

$$t'_{I_{MISS}} = \sum_{i=1}^{M} \sum_{j=1}^{a} \sum_{k=1}^{b} t_{obl\_d_{ijk}} + \sum_{i=1}^{M} \sum_{j=1}^{a} \sum_{l=1}^{c} t_{tr\_d_{ijl}} \qquad (11)$$

Należy zatem oczekiwać, że czas obliczeń w architekturze MISD będzie znacznie mniejszy niż w poprzednio prezentowanej architekturze SISD, ze względu na brak składników odpowiedzialnych za transmisję i dekodowanie kodów rozkazów. Po-



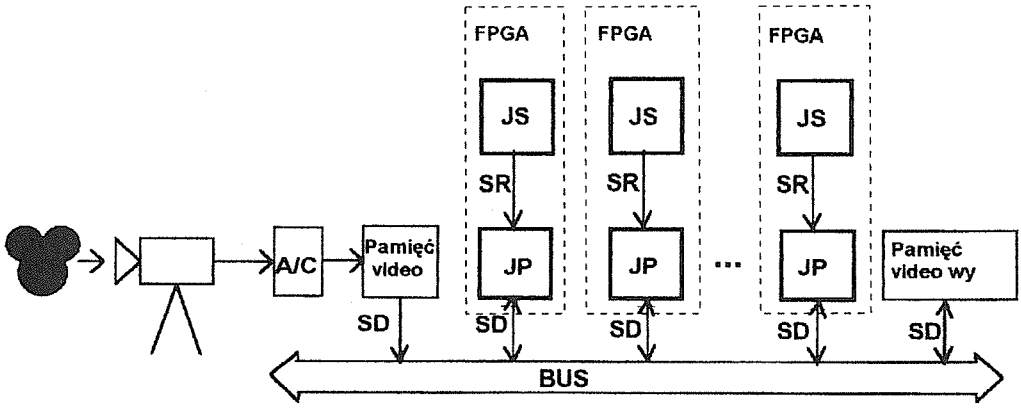
Rys. 7. Architektura systemu SISD pozbawionego transmisji kodów rozkazów

osiągnąć wprowadzając specjalizowane procesory sprzętowe, które w swojej strukturze logicznej będą miały zakodowany program prostych operacji wstępnego przetwarzania. Architektura takiego systemu została przedstawiona na rys. 7, na którym zaznaczono jako jeden moduł strukturę FPGA, w której zawiera się jednostka procesora JP i jednostka sterująca JS. Specjalizowane procesory sprzętowe bardzo korzystnie jest implementować w strukturach FPGA [16, 20].
Wówczas czas wykonania operacji wstępnego przetwarzania danych wizyjnych ograniczy się do jednego składnika ze wzoru (6):

$$t'''_{I_{SISD}} = \sum_{i=1}^{M} \sum_{j=1}^{a} \sum_{k=1}^{b} t_{obl\_d_{ijk}} \tag{10}$$

Połączenie kilku takich procesorów wymaga odpowiedniej ich synchronizacji, aby po wykonaniu pierwszej operacji wstępnego przetwarzania obrazu pierwszy specjalizowany procesor był dezaktywowany i jednocześnie uaktywniał się drugi itd. Wszystkie procesory korzystają bowiem ze wspólnej pamięci. Wprowadzenie kilku zestawów pamięci porządkuje zawarte w nich dane wizyjne, niemniej nie przyspiesza pracy procesorów przyłączonych do wspólnej magistrali. Architekturę takiego systemu przedstawia rys. 8.

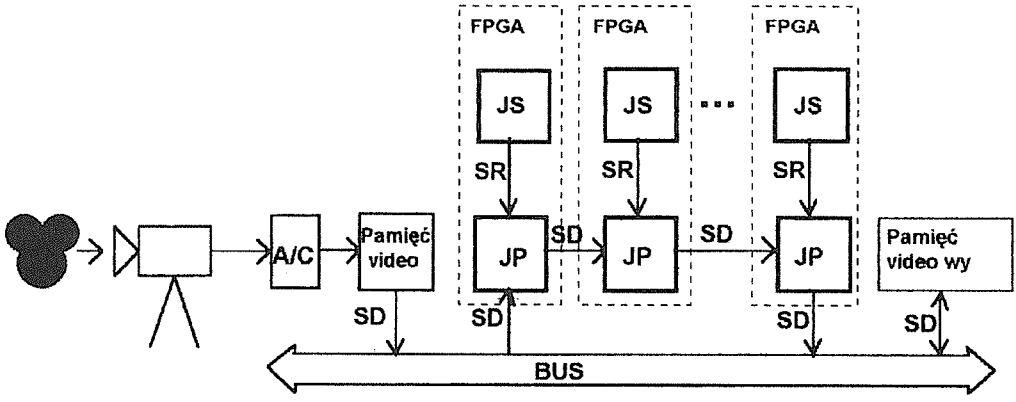
Architektura systemu z rys. 8 bardzo przypomina architekturę systemu MISD z rys. 4. Jest to oczywiście jedynie podobieństwo graficzne. Na rys. 4 poszczególne strumienie danych SD i strumienie rozkazów SR są aktywne jednocześnie, natomiast



Rys. 8. Architektura systemu z procesorami SISD

w systemie z rys. 8 czas magistrali i pamięci są dzielone pomiędzy poszczególne procesory tak, że w danym przedziale czasu aktywny jest tylko jeden procesor (jeden strumień danych SD i jeden strumień rozkazów SR).

Dalsze poszukiwania ograniczenia czasu realizacji I etapu przetwarzania danych wizyjnych prowadzą do próby eliminacji ostatniego dużego składnika we wzorze (6), odpowiedzialnego za transmisję danych wizyjnych pomiędzy procesorami (t_{tr_dijl}). Byłoby to możliwe, jeżeli procesory z rys. 8 pracowałyby synchronicznie, a dane mogłyby być przekazywane z wyjścia pierwszego procesora na wejście drugiego itd. Sytuację taką prezentuje rys. 9. Następuje tu kolizja w dostępie do magistrali systemowej (BUS) przez pierwszy i ostatni procesor (JP). Jednak przy większej liczbie procesorów zyski czasowe są znaczne, a problem dostępu do magistrali można rozwiązać innymi metodami.



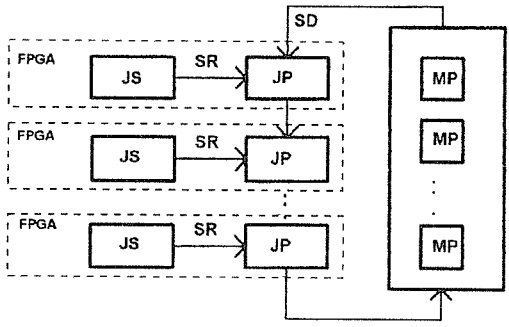
Rys. 9. Architektura MISD specjalizowanych procesorów sprzętowych

Przedstawiona na rys. 9 architektura jest typowym przykładem architektury MISD według klasyfikacji zaproponowanej przez Flynna, bowiem wielokrotny strumień rozkazów SR pracuje jednocześnie (w tym samym przedziale czasu). Architekturę uwidocznioną na rys. 9 można sprowadzić do postaci graficznej bardziej przypominającej to, co zaproponował w swojej klasyfikacji Flynn (rys. 10).

Łatwo można zauważyć, że architektura taka nie wprowadza do czasu obliczeń opóźnienia związanego z odczytywaniem i dekodowaniem kodu rozkazu. W tej sytuacji należy zmodyfikować wzór (6) na czas realizacji obliczeń wstępnego przetwarzania danych wizyjnych dla tej architektury, który przyjmie postać:

$$t'_{r_{MISS}} = \sum_{i=1}^{M}\sum_{j=1}^{a}\sum_{k=1}^{b} t_{obl\_d_{ijk}} + \sum_{i=1}^{M}\sum_{j=1}^{a}\sum_{l=1}^{c} t_{tr\_d_{ijl}} \qquad (11)$$

Należy zatem oczekiwać, że czas obliczeń w architekturze MISD będzie znacznie mniejszy niż w poprzednio prezentowanej architekturze SISD, ze względu na brak składników odpowiedzialnych za transmisję i dekodowanie kodów rozkazów. Po-



Rys. 10. Architektura MISD specjalizowanych procesorów sprzętowych (narysowana zgodnie z konwencją zaproponowaną przez Flynna)

nadto należy zauważyć, że transmisja danych pomiędzy procesorami zachodzi równocześnie. Podobnie obliczenia realizowane są w poszczególnych procesorach równolegle. Dlatego wzór (11) należy odpowiednio zmodyfikować uwzględniając równoczesność tych procesów.

$$t''_{I_{MISD}} = \sum_{j=1}^{a} t_{obl\_d_j} + \sum_{j=1}^{a} t_{tr\_d_j} \tag{12}$$

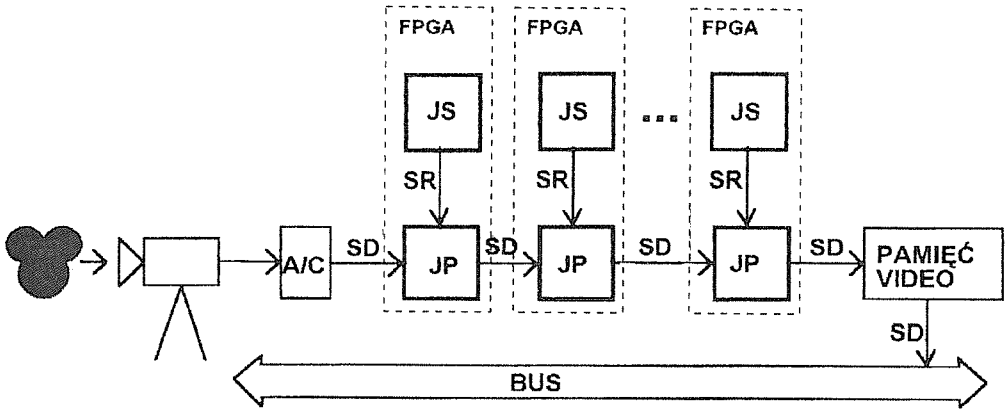
Niestety, zgodnie ze wzorem (7) czas realizacji I etapu przetwarzania obrazu zapisany w (11) należy uzupełnić składnikami odpowiedzialnymi za czas transmisji danych wizyjnych z pamięci wejściowej do pierwszego procesora i z ostatniego procesora do pamięci wyjściowej (wzór 13).

$$t'''_{I_{MISD}} = t_{tr\_we} + \sum_{j=1}^{a} t_{obl\_d_j} + \sum_{j=1}^{a} t_{tr\_d_j} + t_{tr\_wy} \tag{13}$$

W tej sytuacji nietrudno zauważyć, że składniki odpowiedzialne za transmisję danych wizyjnych z pamięci wejściowej do pierwszego procesora i z ostatniego procesora do pamięci wyjściowej można byłoby wyeliminować, gdyby udało się zbudować odpowiednie bloki wieloportowych pamięci wizyjnych, zdolnych w czasie rzeczywistym przekazywać te dane z dużą szybkością przez obie bramy, w sposób synchroniczny do pracy procesorów sprzętowych. Wówczas schemat systemu można byłoby zmodyfikować do postaci z rys. 11, a czas realizacji obliczeń I etapu określlony byłby wzorem (12).

Postać wzoru (12) uwidacznia, jak znacznie opóźnienie wnoszone przez architekturę MISD, złożoną z procesorów dedykowanych, jest mniejsze od analogicznego opóźnienia wnoszonego przez najczęściej używaną architekturę SISD (opisanego wzorem 9 lub 10). Należy się spodziewać, że czas ten jest o kilka rzędów mniejszy, co rzeczywiście odpowiadałoby charakterowi poszukiwań innych jakościowo rozwiązań, prowadzonych przez autora.

Po głębszej analizie przydatności architektury MISD do wstępnego przetwarzania obrazów wydaje się ona w naturalny sposób dostosowana do natury danych



Rys. 11. Architektura MISD specjalizowanych procesorów sprzętowych (z bezpośrednią transmisją danych z/do pamięci)

obrazowych pochodzących z kamery wizyjnej i przetworzonych przez przetwornik analogowo-cyfrowy. Podobne podobieństwo dotyczy strumieni rozkazów. Wymienione w rozdziale 1 operacje wstępnego przetwarzania obrazu są jednakowe dla danej operacji dokonywanej na całym ciągu danych wizyjnych — strumieniu danych (np. operacja LUT lub konwolucja). Następujące po sobie operacje są niezależne, tzn. niezależnie sterowane przez własny strumień rozkazów (każdy procesor w strukturze FPGA posiada własny strumień rozkazów, co daje w efekcie wielokrotny strumień rozkazów). Dzięki temu wnoszone opóźnienie może być tak małe. Architektury SIMD i MIMD nie mają takiego podobieństwa swej struktury do natury sygnału wizyjnego, i w związku z tym ich efektywność, związana głównie ze skomplikowaną ich budową, na tym etapie przetwarzania obrazów nie jest interesująca.

Architektura MISD jest ściśle zsynchronizowana, a jej działanie opiera się na przepływie strumienia danych przez odpowiednio uporządkowany szereg elementów przetwarzających, z których każdy wykonuje przypisaną do siebie operację. W niniejszym opracowaniu określenie architektura potokowa będzie używane dla określenia takiej właśnie architektury (w literaturze pojęcie przetwarzanie potokowe jest także używane do określania sposobu wykonywania następujących po sobie rozkazów). Przetwarzanie danych na strumieniu (potoku) danych wymaga oczywiście zastosowania odpowiedniej architektury jednostek obliczeniowych i odpowiedniego ich przyłączenia do sygnału wejściowego, połączenia pomiędzy sobą oraz wyprowadzenia danych na zewnątrz lub do dalszego przetwarzania. Architektura potokowa pozwala na przetwarzanie sygnału video z kamery wizyjnej jeszcze przed pierwszym zapisaniem go w pamięci komputera. Eliminuje się w ten sposób czasochłonne przesłania pamięć – procesor – pamięć wymagane w innych systemach przetwarzania obrazu (także wieloprocesorowych).

Uzasadniajac dokonany wybór architektury MISD jako optymalnej dla systemu do wstępnego przetwarzania sygnału wizyjnego, należy wymienić szczegółowo jej walory:

1 — **eliminacja bardzo czasochłonnych przesłań pamięć—procesor—pamięć** danych wizyjnych w czasie wstępnego przetwarzania,

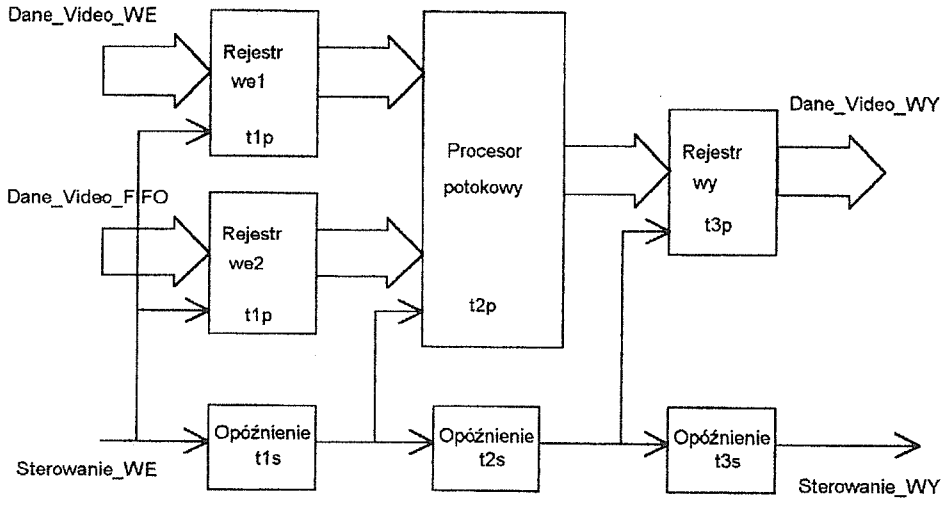
2 — **równoległa realizacja operacji wstępnego przetwarzania** przez dedykowane procesory sprzętowe,

3 — **możliwość przetwarzania danych jeszcze przed zapisaniem ich w pamięci** w przypadku: stosowania kamery z kolejnoliniowym wybieraniem linii, wykorzystywania jednego półobrazu lub stosowania specjalnych translatorów.

Potokowe przetwarzanie sygnału video narzuca duże wymagania wobec procesora potokowego, który ma dokonać pełnego przetworzenia danych punktu (piksela), zanim nadejdzie kolejna porcja informacji z kamery wizyjnej (kolejny piksel). Czas, jakim dysponuje procesor potokowy, jest ściśle związany z częstotliwością próbkowania przetwornika A/C przyłączonego do wyjścia kamery analogowej. Wynika on z czasu przetwarzania jednego obrazu przez kamerę i jego podziału na linie i piksele (por. wzór 1).

6. OPÓŹNIENIA WNOSZONE PRZEZ PROCESORY

W module implementacji projektu możliwa jest interaktywna analiza opóźnień między poszczególnymi punktami układu FPGA, która ma na celu identyfikację problemów związanych z dokładną analizą przebiegów czasowych. Opóźnienia w układach FPGA zależą od sposobu rozmieszczenia logiki w konkretnym układzie scalonym. Przykładowo na rys. 12 przedstawiono model opóźnień w potokowym procesorze logicznym. Opóźnienie dla danych wizyjnych w najprostszych proceso-



Rys. 12. Opóźnienia w układzie potokowego procesora logicznego

rach związane jest z buforowaniem danych wizyjnych na wejściu i wyjściu oraz z realizacją operacji zgodnej z funkcją danego procesora. Sygnały sterujące z jednej strony muszą być opóźnione celem zapewnienia procesorowi odpowiednich sekwencji zboczy sterujących, z drugiej — wnoszone dla sterowania opóźnienie musi być identyczne jak dla danych wizyjnych. W opracowanej architekturze potokowej założono, że opóźnienia te będą wielokrotnością czasu trwania jednego cyklu magistrali potokowej (1/15 MHz = 66 ns), odpowiadającego jednemu pikselowi.

Każde zaprezentowane na rys. 12 opóźnienie składa się z opóźnienia wnoszonego przez logikę kombinacyjną, przerzutniki i połączenia. Można to zapisać wzorami:

$$t_{1si} = t_{i_{BUF\_WE}} + \sum_{j=1}^{m} t_{ij_{LUT}} + \sum_{k=1}^{n} t_{ik_{con}} \tag{14}$$

$$t_{2si} = \sum_{j=1}^{m} t_{ij_{LUT}} + \sum_{k=1}^{n} t_{ik_{con}} \tag{15}$$

$$t_{3si} = \sum_{j=1}^{m} t_{ij_{LUT}} + \sum_{k=1}^{n} t_{ik_{con}} + t_{i_{BUF\_WY}}, \tag{16}$$

gdzie: $t_{i_{BUF\_WE}}$ — opóźnienie wnoszone przez bufor wejściowy, $t_{i_{BUF\_WY}}$ — opóźnienie wnoszone przez bufor wyjściowy, $t_{ij_{LUT}}$ — opóźnienia wnoszone przez tablice LUT bloków CLB użytych do realizacji danej funkcji, $t_{jk_{con}}$ — opóźnienia wnoszone przez połączenia pomiędzy zasobami użytymi do realizacji danej funkcji.

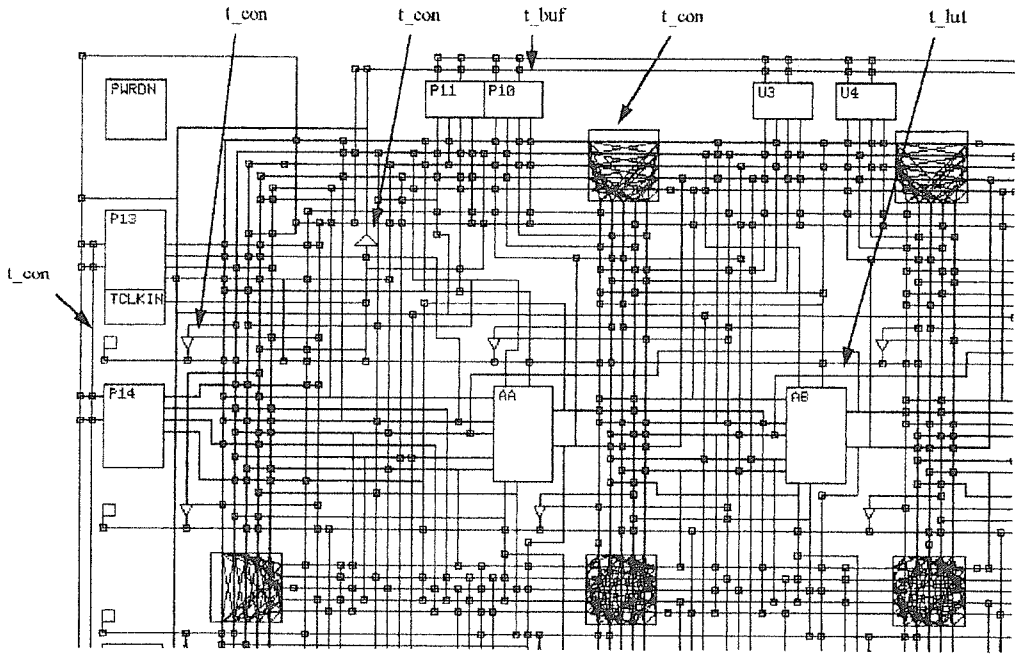
Zaprezentowany we wzorze składnik t_{ikcon} reprezentuje sobą opóźnienia wnoszone przez użyte poszczególne linie sygnałowe (bezpośrednie, długie, specjalne), ich bufory oraz matryce przełączające (rys. 13 oraz tab. 3). Zatem używany we wzorach (12) i (13) czas $t_{\mathrm{obl\_d}}$ oznacza w praktyce sumę czasów t_{ipr}. Opóźnienia wznoszone dla przetwarzanych danych wizyjnych i dla sygnałów sterujących z założenia są takie same (wzór 17).

$$t_{obl\_d_r} = t_{1s_r} + t_{2s_r} + t_{3s_r} = t_{1p_r} + t_{2p_r} + t_{3p_r}, \tag{17}$$

gdzie: t_{ipr} — opóźnienia na drodze danych wizyjnych, t_{isr} — opóźnienia na drodze sygnałów sterujących.

Przykładowa analiza czasów opóźnień dla procesora logicznego:

DEW/S_VID_IN2 to CLB_R7C5.K	
to Clock Input Blk S7	7,1 ns
Blk U240.1 to ODEV/s_pos	9,6 ns
Blk ODEW/S_VID_IN2 to Blk S7	7,1 ns
CKBR2C8 HM1_Q<0>	
SO	17,8 ns
SO to Pad	16,4 ns
	34,2 ns



Rys. 13. Opóźnienia w układzie programowalnym FPGA

Tabela 3

Opóźnienia wnoszone przez niektóre połączenia wewnętrzne układu FPGA

Opóźnienia wnoszone przez połączenia wewnętrzne	XC4003-6	XC4005-6	4010-6
LONGLINES z układu IOB	3,6 ns	3,9 ns	4,5 ns
LONGLINES z logiki wewnętrznej	4,2 ns	4,8 ns	6,0 ns
HALFLINES z układu IOB	3,3 ns	3,9 ns	5,1 ns
HALFLINES z logiki wewnętrznej	4,8 ns	5,4 ns	6,6 ns
Pierwszy *GLOBAL_SIGNAL*	6,1 ns	6,1 ns	6,1 ns
Drugi *GLOBAL_SIGNAL*	6,5 ns	6,5 ns	6,5 ns

Przy założeniach że:

- dane z magistrali potokowej zatrzaskiwane są sygnałem Sterowanie_WE,
- dane z pamięci TRAM (FIFO) wyprowadzane są sygnałem Sterowanie_WE;

dane Dane_Video_WE doznają na drodze wewnątrz układu opóźnienia: Top = 34,2 ns.

Tabela 4

Opóźnienia bloków CLB i I/O

Rodzaj opóźnienia	XC4000-6	XC4000-5
Logika kombinacyjna	8 ns	7 ns
Generowanie przeniesienia	8 ns	6 ns
TBUF z jednym *pull-up* TBUF z dwoma *pull-up*	12,1 ns 5,5 ns	9,9 ns 4,5 ns
Zapis do CLB_RAM	7 ns	5,5 ns
Odczyt z CLB_RAM	10 ns	7,5 ns

Przykładowo dla implementacji potokowego procesora logicznego wynikają z powyższego następujące wnioski:

- maksymalne opóźnienie danych z pamięci FIFO w stosunku do danych przychodzących z magistrali potokowej wynosi 10 ns,
- dane z magistrali potokowej można zatrzaskiwać bezpośrednio sygnałem Sterowanie_WE,
- dane z FIFO spóźnione są o 10 ns w stosunku do sygnału sterującego (należy go opóźnić o 20 ns dla bezpieczeństwa — t_{1s}),
- opóźnienie na odcinku od rejestrów zatrzaskujących dane z magistrali potokowej i dane z FIFO, poprzez sumator, aż do rejestru wyjściowego wynosi 27,6 ns (na tym odcinku należy opóźnić sygnał Sterowanie_WE o 30 ns — t_{2s}),
- rejestr wyjściowy opóźnia dane maksymalnie o 10 ns, dane na wyjściu powinny pojawić się minimum 10 ns przed sygnałem sterującym Sterowanie_WE (na tym odcinku należy opóźnić sygnał Sterowanie_We o dalsze 20 ns — t_{3s}).

Efekty pracy routera są zapisywane w zbiorach komunikatów (*.out, *.rpt) oraz dla udanego przebiegu *routingu* w zbiorze opisującym daną implementację (*.lca). W [16, 19] zaprezentowano przykładowe implementacje dla procesora logicznego (odejmującego dwa obrazy) i medianowego na poziomie schematu LCA oraz wydruk komunikatów (zbiór *.OUT).

W trakcie weryfikacji opracowanych wskazań dla konstrukcji systemów do wstępnego przetwarzania danych wizyjnych dokonano oceny parametrów czasów realizacji algorytmów w opracowanym systemie przetwarzania potokowego. Elementami składowymi opóźnienia wnoszonego przez podsystem wstępnego przetwarzania danych wizyjnych są opóźnienia wnoszone przez poszczególne procesory i komunikacja pomiędzy nimi. Warto w tym miejscu przytoczyć wzór (12):

$$t''_{I_{MISD}} = \sum_{j=1}^{a} t_{obl\_d_j} + \sum_{j=1}^{a} t_{tr\_d_j} \tag{18}$$

gdzie: $t_{obl\_d_j}$ — opóźnienie wnoszone przez dany procesor sprzętowy, a — liczba procesorów sprzętowych w architekturze potokowej.

Wykonywane operacje są ściśle związane ze strukturą danego procesora, a dane wizyjne z jednego procesora automatycznie znajdują się na wejściu następnego. Ewentualny czas transmisji danych pomiędzy procesorami związany jest głównie z buforowaniem danych wizyjnych i dopasowywaniem wnoszonych opóźnień dla poszczególnych sygnałów (danych i sterowania). Odpowiednie bufory znajdują się na module procesora (uzgadnianie opóźnień tych grup sygnałów odbywa się w obrębie danego procesora). Połączenie wyjścia jednego procesora sprzętowego z wejściem następnego wykonane jest jako połączenie galwaniczne o niewielkiej długości. Zatem czas opóźnienia wnoszonego przez procesor $t_{obl\_d_j}$ zawiera w sobie składnik opóźnienia związanego z transmisją danych pomiędzy procesorami potokowymi (składnik ten jest trudny do wydzielenia ze względu na wysoką skalę integracji całości schematu procesora). W tej sytuacji wzór (18) należy zmodyfikować do postaci (19):

$$t'''_{I_{MISD}} = \sum_{j=1}^{a} t_{obl\_d_j}. \tag{19}$$

Opóźnienie wnoszone przez dany procesor $t_{obl\_d_j}$ wynosi od jednego do kilku cykli magistrali potokowej. Jeden cykl magistrali potokowej związany jest z częstotliwością zegara i wynosi 1/15 MHz = 66 ns. Dla konwolucji 3×3 opóźnienie odpowiada czasowi przesłania 3 linii obrazu przez architekturę potokową i wynosi 3×512 cykli, a dla mediany 2 linii obrazu czyli 2×512 cykli. Rzeczywiste czasy wnoszonych przez procesory hardware'owe opóźnień zestawiono w tabeli 5.

Tabela 5

Rodzaj operacji	Wnoszone opóźnienie
Operacja *look-up-table* z pamięcią ROM	66 ns
Operacja *look-up-table* z pamięcią RAM	132 ns
Operacja zliczania histogramu	66 ns
Operacja odejmowania dwu obrazów	132 ns
Operacja mediany 5-punktowej	68 μs
Operacja mediany 9-punktowej	68 μs
Operacja konwolucji 3×3	102 μs
Dwukierunkowy gradient Sobela 3×3	102 μs

Dla porównania przytoczyć należy wartości wnoszonych opóźnień (czasów obliczeń) związanych z realizacją analogicznych operacji przez inne systemy lub moduły. Opóźnienia te podawane w literaturze zebrano w poniższej tabeli porównawczej (tab. 6), przy czym niektóre elementy tabeli są puste ze względu na brak danych.

Tabela 6

System	Liczenie histogramu	Operacje na dwu obrazach	Konwolucja 3×3
Architektura potokowa	66 ns	132 ns	102,0 μs
DT2858 [24]	180 ms	250 ms	1180,0 ms
DT2878 [24]		300 ms	800,0 ms
TMS 320 [97]			2200,0 ms
TMS C40 [1 -June 1995]			141,6 ms
TMS C80 [1 -June 1995]			52,4 ms
FT200 2xi860, bus 200MHz [3]			51,9 ms

7. PODSUMOWANIE

Miarą efektywności przeprowadzonych rozważań jest czas realizacji operacji przetwarzania danych w architekturze potokowej. Wyniki obliczeń zawarto w tabeli 5 i łatwo jest je porównać z dotychczas stosowanymi rozwiązaniami, których parametry zawarte są w tab. 6. Porównanie powyższych czasów może wydawać się na pozór nieuzasadnione. Z jednej strony zestawiono osiągnięte przez autora czasy przetwarzania jednego piksela, z drugiej zaś czasy przetwarzania całego obrazu. Złudzenie takie wynika jednak z zasadniczych różnic w architekturze obu rozwiązań. Przykładowo efektywny czas obliczenia konwolucji przez procesor TMS-C40 wyniesie 141,6 ms od momentu zgromadzenia całego obrazu w pamięci. Procesor potokowy oblicza konwolucję zanim obraz zostanie wpisany do pamięci wnosząc opóźnienie 102 μs przy zapisie każdego piksela, tzn. ostatni piksel zostanie zapisany także 102 μs później niż gdyby nie było procesora konwolucji na drodze sygnału wizyjnego.

Rozważmy następujący przykład: jeżeli algorytm wstępnego przetwarzania danych wizyjnych wymaga zastosowania procesorów potokowych: *look-up-table*, liczenia histogramu, medianowego, procesora konwolucji, dwukierunkowego gradientu Sobela i odejmowania dwu obrazów, to opóźnienie wniesione przez te procesory będzie sumą opóźnień procesorów składowych, co możemy obliczyć powołując się na wzór (19).

$$t''''_{I_{MISD}} = \sum_{j=1}^{6} t_{obl\_dj} = 66\,ns + 66\,ns + 68\,\mu s + 102\,\mu s + 102\,\mu s + 132\,ns = 272{,}3\,\mu s. \quad (20)$$

W tym miejscu warto zebrać zalety opracowanych procesorów potokowych:

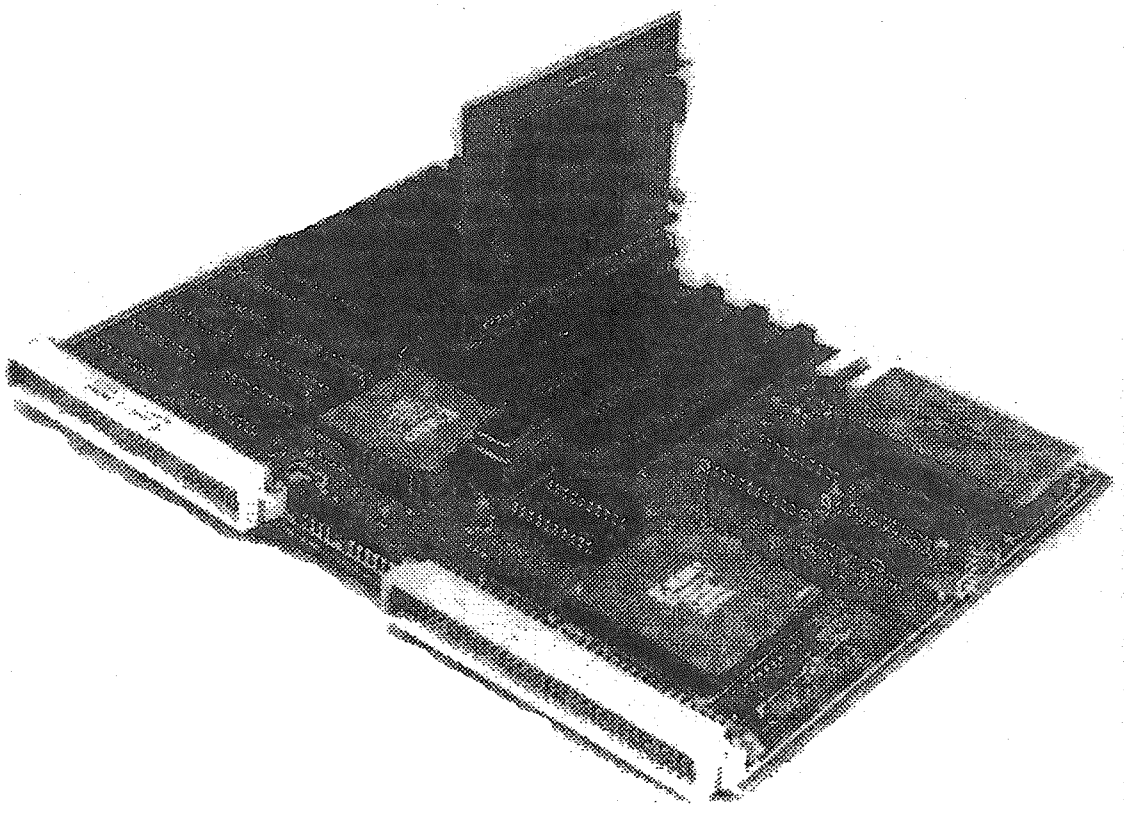
* bardzo krótki czas obliczeń,
* bardzo krótki czas przyjęcia i wysłania danych,
* możliwość dynamicznej zmiany parametrów przetwarzania,
* możliwość dynamicznej rekonfiguracji kolejności operacji,

* procesory są autonomiczne i można je niezależnie projektować.

Wymagania opracowanych procesorów potokowych są następujące:

* konieczność zachowania przez procesory wymogów czasowych:
  - jednakowe opóźnienia dla danych i sygnałów sterujących,
  - wnoszone opóźnienie wielokrotnością czasu przesłania jednego piksela,
* kolejnoliniowy sygnał wizyjny dla procesorów realizujących operacje kontekstowe.

Powyższe prace zostały zweryfikowane eksperymentalnie. Opracowano i wykonano szereg procesorów potokowych z wykorzystaniem układów programowalnych FPGA firmy Xilinx [18]. Wykonano także moduł zapewniający współpracę tych procesorów oraz ich komunikację z kamerą wizyjną i magistralą VME (rys. 14). Prace te były możliwe dzięki finansowemu wsparciu Komitetu Badań Naukowych.



Rys. 14. Moduł architektury potokowej z procesorem LUT

BIBLIOGRAFIA

1. *Advanced Imaging*, PTN Publishing Co., New York, June 1995
2. N. Alexandris: *Design of Microprocessor-Based Systems.* Prentice Hall 1993
3. A.L. Decegama: *The Technology of Parallel Processing – Parallel Processing Architectures and VLSI Hardware.* Prentice Hall, New Jersey 1989
4. M.J. Flynn: *Very High-Speed Computing Systems.* Proc. IEEE, 1966, Vol. 54, Nr12
5. K. Hwang: *Advanced Computer Architecture*, Mc-Graw Hill 1993
6. Y. Iwata, M. Kawamata, T. Higuchi: *Design of fine VLSI array processor for real-time 2-D digital filtering.* IEEE Sympsium on Circuits and Systems, Chicago 1993
7. R.Y. Kainu: *Advanced Computer Architecture.* Prentice Hall 1996
8. S. Kozielski, Z. Szczerbiński: *Komputery równoległe.* WNT, Warszawa 1994
9. Y.H. Lee, C.M. Krishna: *Real-Time Systems* IEEE 1993
10. B. Liebowitz, J. Carson: *Multiple Processor Systems for Real Time Applicationsw.* Prentice Hall – New Jersey 1985
11. T. Łuba, M.A. Markowski, B. Zbierzchowski: *Komputerowe projektowanie układów cyfrowych w strukturach PLD.* WKŁ, Warszawa 1993
12. L.L. Miller, A.R. Hurson, S.H. Pakzad: *Parallel Architecture for Data/Knowledge Based Systems.* IEEE 1995
13. R. Tadeusiewicz: *Systemy wizyjne robotów przemysłowych.* WNT, Warszawa 1992
14. R. Tadeusiewicz: *Komputerowa analiza i przetwarzanie obrazów.* Wydawnictwo Fundacji Postępu Telekomunikacji, Kraków 1997
15. K. Wiatr: *Pipelined Architecture of Reconfigurable Specialised Processors for a Real-Time Image Data Pre-Processing.* Proc. of the IEEE International Conference on Signal Processing – ICSP – 96, Beijing – China 1996, IEEE Press, s. 649 – 652
16. K. Wiatr: *Architektura specjalizowanych procesorów potokowych do obróbki wstępnej danych wizyjnych w czasie rzeczywistym.* Kwartalnik Elektroniki i Telekomunikacji PAN, 1997, Zeszyt 1, s. 121 – 148
17. K. Wiatr: *A Dedicated Hardware Logic Processor for a Real-Time Image Data Pre-Processing.* Proc. of the IASTED & IEEE International Conference: CONTROL'97, Cancun – Mexico, 1997, IASTED Acta Press, Anaheim – Calgary – Zurich, s. 207 – 210
18. K. Wiatr *Specialised Architecture of Dedicated Hardware Processors for a Real-Time Image Data Pre-Processing.* Proc. of the EUROMICRO International Conference: *Real-Time Systems.* Toledo – Spain, 1997, IEEE Computer Press, Washington – Brusssels – Tokyo, s. 180
19. K. Wiatr: *Dedicated Hardware Processors for a Real-Time Image Data Pre-Processing Implemented in FPGA Structure.* 9th International Conference on Image Analysis and Processing – ICIAP'97, Florence – Italy, 1997, Springer – Verlag, Berlin, vol. II, s. 69 – 75
20. XILINX: *Data Book*, San Jose, 1996

K. WIATR

TIME PARAMETERS ANALYSIS FOR PIPELINE ARCHITEKTURE OF SPECIALISED HARDWARE PROCESSOR

Summary

This paper deals with the problems connected with searching for new, faster methods of image pre-processing required in a special class of image processing systems, being part of sensory systems applied in automatic control engineering. This search consisted in striving after parallel arrangement of image pre-processing computations, by selection of such computational elements (processors) and such a transmission between them (architecture), that the total computation time was the least possible.

The paper includes theoretical considerations concerning various architectures of parallel-operating specialised hardware processors, with emphasis on execution times of individual operations. The selection of pipeline architecture of specialised hardware processors has been justified. The concepts presented have been verified by experiments, using processors implemented in FPGA structures. The work presented is a contribution to worldwide intense research on developing user dedicated "Custom Computing Machines".

The paper presents time parameters analysis concern on select multiprocessors architecture for high speed image preprocessing. To verify the concept the following processors have been worked out: look-up-table, median filter, convolver, logic processor, and histogram processor (implemented in FPGA structures). The paper includes experimental delay time for particular operations.

Key words: image processing systems, architekture hardware processors.

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI, 1998, **44**, z. 3, ss. 413–426

Wybrane metody identyfikacji i doboru nastaw stosowane w regulatorach z samonastrajaniem

HENRYK KUNERT

Instytut Cybernetyki Technicznej, Politechnika Wrocławska

Otrzymano 1998.05.05

Autoryzowano 1998.07.10

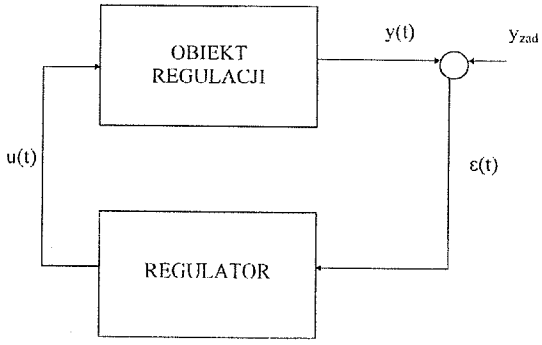
W pracy poddano krytycznej ocenie dwie metody identyfikacyjne najczęściej implementowane w przemysłowych regulatorach mikroprocesorowych PID wraz ze stosowanymi metodami doboru nastaw regulatorów o działaniu ciągłym. Prezentowane metody zaliczane są do dwóch grup metod identyfikacji i doboru nastaw: metody odpowiedzi skokowej, metody oscylacyjne.

Metoda odpowiedzi skokowej zaproponowana przez H.P. Preussa, implementowana jest w serii regulatorów SIPART firmy Siemens. Drugą metodą jest metoda zaproponowana przez J.J. Aströma i T. Hägglunda, wykorzystuje ona zmodyfikowany ekperyment oscylacyjny Zieglera-Nicholsa. Na podstawie przeprowadzonych badań symulacyjnych w niniejszej pracy podano zalety i wady prezentowanych metod.

Słowa kluczowe: metody identyfikacji, regulacja PID, nastawy regulatora PID, symulacja komputerowa.

WSTĘP

Pomimo szybkiego rozwoju algorytmów sterowania, nadal szerokie zastosowanie mają regulatory realizujące algorytm regulacji PID. Rozwój techniki komputerowej sprawił, iż w ostatnich latach coraz szersze zastosowanie znajdują regulatory mikroprocesorowe wyposażone w procedury samoczynnego doboru nastaw tzw. regulatory z samonastrajaniem [3], [5], [6], [8] posiadajace istotną przewagę nad regulatorami klasycznymi. Regulatory z samonastrajaniem stosuje się w układach, w których z upływem czasu zmieniają się parametry procesu regulacji (parametry obiektu, wartość zadana, poziom zakłóceń itp.). W regulatorach tych stosowane są różne rodzaje procedur identyfikacji sterowanego obiektu i doboru nastaw, i jak wykazuje praktyka, nie ma wśród nich metod idealnych.



Rys. 1. Zamknięty układ regulacji

W niniejszej pracy zwrócono uwagę na kilka istotnych zalet i wad cechujących dwie metody samonastrajania. Pierwsza z nich, przedstawiona przez Hansa Petera Preussa [9], dalej nazywana metodą Preussa należy do grupy metod odpowiedzi skokowej. Druga metoda wykorzystuje sposób podany przez J.G. Zieglera i N.B. Nicholsa [10] „zautomatyzowany" przez J. Aströma i T. Hägglunda [1], mianowicie wyznacza się parametry drgań krytycznych zamkniętego układu automatycznej regulacji, w którym regulator o działaniu ciągłym zastąpiono przekaźnikiem dwupołożeniowym. Obie metody dotyczą identyfikacji obiektów sterowania i doboru nastaw regulatorów w zamkniętych układach regulacji (rys. 1).

METODA PREUSSA

Metoda ta zaliczana jest do metod odpowiedzi skokowej, polega na identyfikacji parametrów (k, T, n) tzw. modelu Strejca:

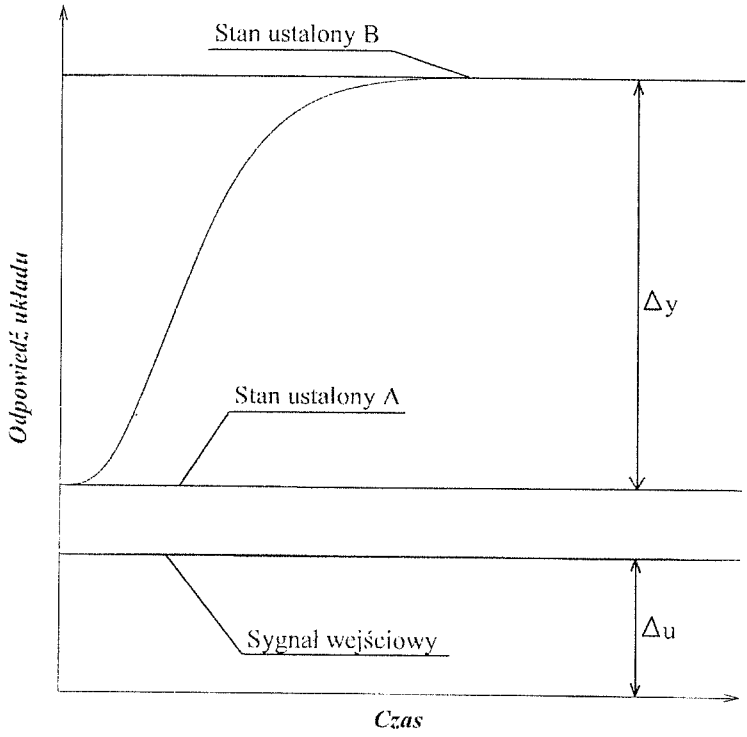
$$G(s) = \frac{k}{(Ts+1)^n}, \tag{1}$$

gdzie: k — wzmocnienie obiektu,
T — stała czasowa obiektu,
n — rząd obiektu.

Metoda ta wymaga otwarcia pętli sprzężenia zwrotnego podczas przeprowadzania eksperymentu identyfikacyjnego. Wzmocnienie k obiektu wyznaczane jest jako iloraz przyrostów sygnałów wyjściowego do wejściowego pomiędzy dwoma stanami ustalonymi (rys. 2) według następującej zależności:

$$k = \frac{\Delta y}{\Delta u}, \tag{2}$$

gdzie Δy jest zmianą wartości sygnału wyjściowego odpowiadająca zmianie sygnału wejściowego o Δu, a Δu — zmianą sygnału wejściowego.



Rys. 2. Identyfikacja parametrów odpowiedzi skokowej

Pozostałe parametry tzn. T oraz n wyznaczane są przy pomocy ogólnie znanej rekurencyjnej metody najmniejszych kwadratów [7]. Zastosowanie ww. metody w wersji rekurencyjnej podyktowane jest brakiem możliwości przechowywania dużej ilości danych pomiarowych w pamięci fizycznego urządzenia identyfikacyjnego np. przemysłowego mikroprocesorowego regulatora PID.

Nastawy regulatora PID o transmitancji:

$$G_R(s) = k_p\left(1 + \frac{1}{T_i s} + T_d s\right)$$

w prezentowanej metodzie wyliczane są w oparciu o następujące kryterium:

$$|G_z(j\omega)| \approx 1 \quad \text{dla} \quad 0 \le \omega \le \omega_g, \tag{3}$$

gdzie: $G_z(j\omega)$ — transmitancja widmowa zamkniętego układu regulacji z regulatorem PI lub PID,

ω_g — maksymalna wartość pulsacji ω, dla której spełniony jest warunek (3)

Optymalne nastawy w sensie kryterium (3) przedstawiono w tabeli 1.

W celu wyznaczenia nastaw dla obiektów rzędu 1-go i 2-go należy skorzystać z innych metod doboru nastaw.

Tabela 1

Optymalne nastawy regulatorów PI oraz PID dla modelu Strejca wg metody Preussa

Typ regulatora	Optymalne nastawy		
	k_p	T_i	T_d
PI	$\frac{1}{4k}\frac{n+2}{n-1}$	$\frac{T}{3}(n+2)$	–
PID	$\frac{1}{16k}\frac{7n+16}{n-2}$	$\frac{T}{15}(7n+16)$	$T\frac{n^2+4n+3}{7n+16}$

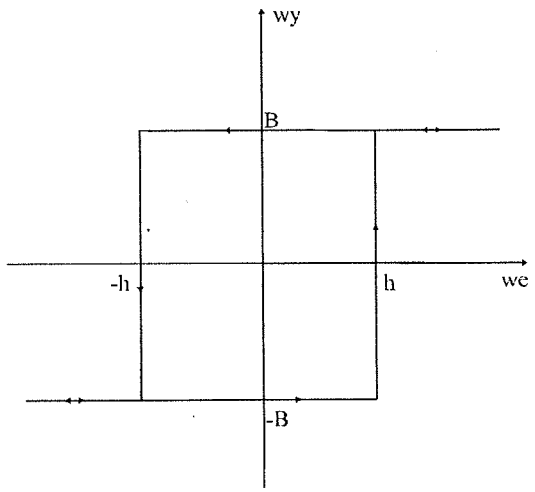
METODA ASTRÖMA–HÄGGLUNDA

W metodzie tej [1] wykorzystano eksperyment Zieglera–Nicholsa polegający na dołączeniu regulatora typu P (proporcjonalnego) do zamkniętego układu regulacji z badanym obiektem o transmitancji $G_0(s)$ i doborze takiego wzmocnienia regulatora k_{kr}, przy którym wystąpią w układzie drgania o stałej amplitudzie A_{kr} i okresie T_{osc}. Powstanie w ukladzie drgań o parametrach A_{kr} i T_{osc} warunkowane jest spełnieniem zależności Nyquista

$$G_0(j\omega)\,G_R(j\omega) = -1, \tag{4}$$

gdzie: $G_0(j\omega)$ — transmitancja widmowa obiektu,
$G_R(j\omega)$ — transmitancja widmowa regulatora.

W metodzie Aströmа–Hägglunda regulator proporcjonalny zastąpiono przekaźnikiem dwupołożeniowym, którego transmitancyjna zastępowana jest tzw. funkcją opisującą (5) tego przekaźnika. Charakterystykę statyczną przekaźnika dwupołożeniowego przedstawiono na rysunku 3. Funkcja opisująca i jej zastosowanie do opisu nieliniowych układów automatyki przedstawiona została w [4].



Rys. 3. Charakterystyka statyczna przekaźnika dwupołożeniowego

$$J(A) = \frac{4B}{\pi A}\left(\sqrt{1-\left(\frac{h}{A}\right)^2} - j\frac{h}{A}\right) \quad \text{dla} \quad A > h \tag{5}$$

gdzie: A — amplituda pierwszej harmonicznej sygnału wejściowego przekaźnika dwupołożeniowego

B — amplituda sygnału wyjściowego przekaźnika dwupołożeniowego,

h — szerokość pętli histerezy przekaźnika dwupołożeniowego.

W powyższej metodzie dokonuje się pomiaru okresu oscylacji krytycznych T_{osc} oraz amplitudy drgań A_{kr}, na podstawie której wylicza się wartość wzmocnienia krytycznego równego co do wartości funkcji opisującej (5) przekaźnika dwupołożeniowego bez histerezy ($h = 0$), mianowicie

$$k_{kr} = \frac{4B}{\pi A_{kr}}. \tag{6}$$

Następnie wyznacza się nastawy regulator a PI bądź PID wg metody Zieglera – – Nicholsa [10] przedstawione w tabeli 2.

Tabela 2

Nastawy regulatora o działaniu ciągłym wg Zieglera – Nicholsa

Typ regulatora	Nastawy regulatora		
	k_p	T_i	T_d
P	$0.5\ k_{kr}$	–	–
PI	$0.45\ k_{kr}$	$0.85\ T_{osc}$	–
PID	$0.6\ k_{kr}$	$0.5\ T_{osc}$	$0.12\ T_{osc}$

PORÓWNANIE DZIAŁANIA PREZENTOWANYCH METOD

Porównanie skuteczności działania przedstwionych metod identyfikacji obiektu i doboru nastaw regulatora ciągłego pracujących w układzie zamkniętym pokazane zostanie dla obiektów wieloinercyjnych współpracujących z regulatorem PI. Dobór nastaw regulatora PI według zaleceń zawartych w metodzie Preussa prowadzi zwykle do otrzymania odpowiedzi aperiodycznej z niewielkim przeregulowaniem, przy skokowym pobudzeniu układu, zarówno dla obiektów wieloinercyjnych o różnych stałych czasowych o transmitancji w postaci:

$$G(s) = \frac{k}{(T_1 s + 1)(T_2 s + 1)...(T_n s + 1)}, \tag{7}$$

jak również dla obiektów opisanych transmitancją w postaci modelu Strejca:

$$G(s) = \frac{k}{(Ts + 1)^n}. \tag{8}$$

Tabela 3

Wartości nastaw regulatora PI w eksperymencie, którego wyniki przedstawione są na rysunkach 4 i 5

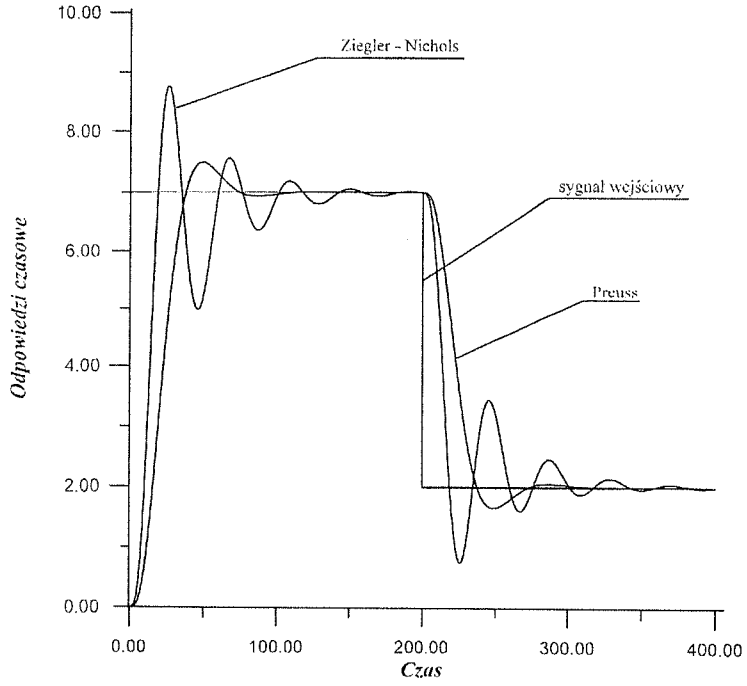
Metoda doboru nastaw	Obiekt 4-go rzędu	Obiekt 20-go rzędu
Zielgera – Nicholsa	$k_p = 0.59$ $T_i = 26.83$	$k_p = 0.20$ $T_i = 167.11$
Preussa	$k_p = 0.17$ $T_i = 10$	$k_p = 0.10$ $T_i = 36.67$

Poza tym, jak się okazuje takie zachowanie układu regulacji jest niezależne od rzędu obiektu (badania obejmowały obiekty od 3-go do 20-go rzędu).
Powyższe wnioski potwierdzają wykresy przedstawione na rysunkach 4 oraz 5.
Obiekty w badanym układzie opisane były transmitancją (8) o następujących parametrach:

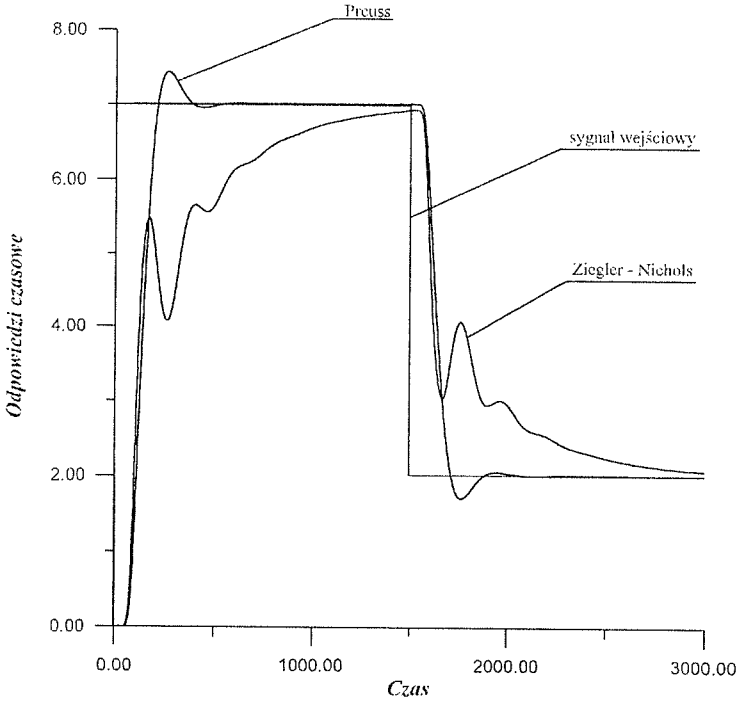
1. $n = 4$, $k = 3$, $T = 5$;
2. $n = 20$, $k = 3$, $T = 5$,

natomiast wartości nastaw regulatora PI przedstawiono w tabeli 3.

Na rysunkach 4 i 5 przedstawiono również wykresy sygnalu wyjściowego dla tego samego układu regulacji, dla którego nastawy regulatora PI dobrane zostały



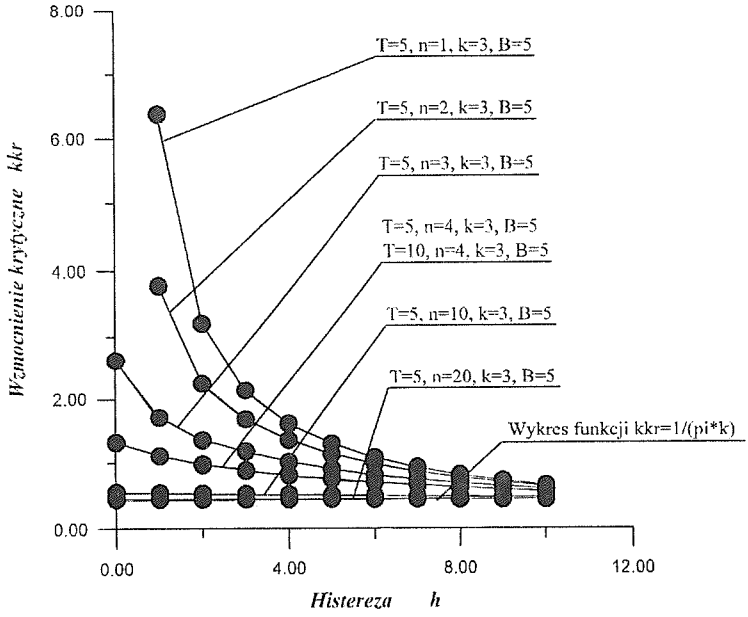
Rys. 4. Przebiegi przejściowe procesu regulacji w układzie zamkniętym dla obiektu inercyjnego 4-go rzędu (nastawy wg Zieglera – Nicholsa oraz Preussa dla regulatora PI)



Rys. 5. Przebiegi przejściowe procesu regulacji w układzie zamkniętym dla obiektu inercyjnego 20-go rzędu (nastawy wg Zieglera – Nicholsa oraz Preussa dla regulatora PI)

przy wykorzystaniu metody Astrőma – Hägglunda i doborze nastaw regulatora wg Zieglera – Nicholsa (wartości nastaw regulatora przedstawiono w tabeli 3).

Można łatwo zauważyć, że metoda Preussa dla obiektów inercyjnych jest bardziej skuteczna w konfrontacji z metodą Astrőma – Hägglunda połączoną z doborem nastaw regulatora wg Zieglera – Nicholsa, lecz nie można na tej podstawie jednoznacznie rozstrzygnąć problemu, która z metod jest lepsza, gdyż metodę Preussa cechuje poważne ograniczenie — można ją stosować wyłącznie wtedy, gdy transmitancję obiektu można przybliżyć transmitancją o postaci (8) zaproponowanej przez Strejca, a do tego potrzebna jest pewna wiedza *a priori* o sterowanym obiekcie, czego nie wymaga metoda Astrőma – Hägglunda. Metoda Astrőma – Hägglunda w połączeniu z doborem nastaw wg Zieglera – Nicholsa posiada również pewne istotne wady. Jedną z nich jest to, że amplituda drgań krytycznych, a w konsekwencji tego rownież wzmocnienie krytyczne k_{kr} wzyznaczane z zależności (6), silnie zależą od histerezy przekaźnika dwupołożeniowego. Metoda ta została opracowana dla przekaźnika bez histerezy ($h = 0$), przy czym wprowadzenie histerezy ($h \neq 0$) powoduje wzrost odporności metody na zakłócenia sygnału wejściowego przekaźnika, który to decyduje o momentach stabilnych przełączeń w procesie identyfikacji. Zależność $k_{kr} = f(h)$ przedstawiona została na rysunku 6.



Rys. 6. Wykres $k_{kr} = f(h)$ dla obiektów o transmitancji $G(s) = \dfrac{k}{(Ts+1)^n}$

Na wspomnianym rysunku widać wyraźnie silną zależność $k_{kr} = f(h)$ dla obiektów niskiego rzędu, dla obiektów o rzędzie inercji $n \geq 10$ zależność ta jest znikoma, stąd wniosek, że dla obiektów wysokiego rzędu histereza przekaźnika w eksperymencie identyfikacyjnym praktycznie nie wpływa na wyznaczenie wzmocnienia krytycznego, a co za tym idzie na wyznaczenie wzmocnienia regulatora ciągłego (patrz tabela 2). Poza tym z rysunku 6 można wywnioskować, że wzmocnienie krytyczne w funkcji histerezy dąży do granicznej wartości:

$$k_{krgr} = \frac{4}{\pi k},$$

gdzie: k — wzmocnienie obiektu.
Ponieważ

$$k_{kr} = \frac{4B}{\pi A_{kr}},$$

oraz przy wzroście histerezy do wartości maksymalnej, przy której nie zostają zerwane drgania w układzie, czyli jest jeszcze spełniony warunek:

$$kB \geq h \tag{9}$$

amplituda drgań w układzie wynosi

$$A_{kr} = kB,$$

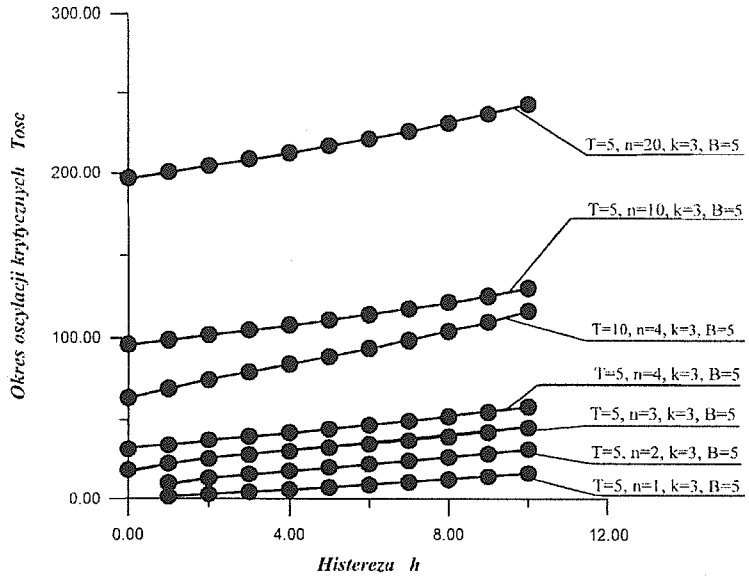
gdzie: B — amplituda sygnału wyjściowego przekaźnika dwupołożeniowego. Stąd

$$k_{krgr} = \frac{4B}{\pi k B},$$

oraz

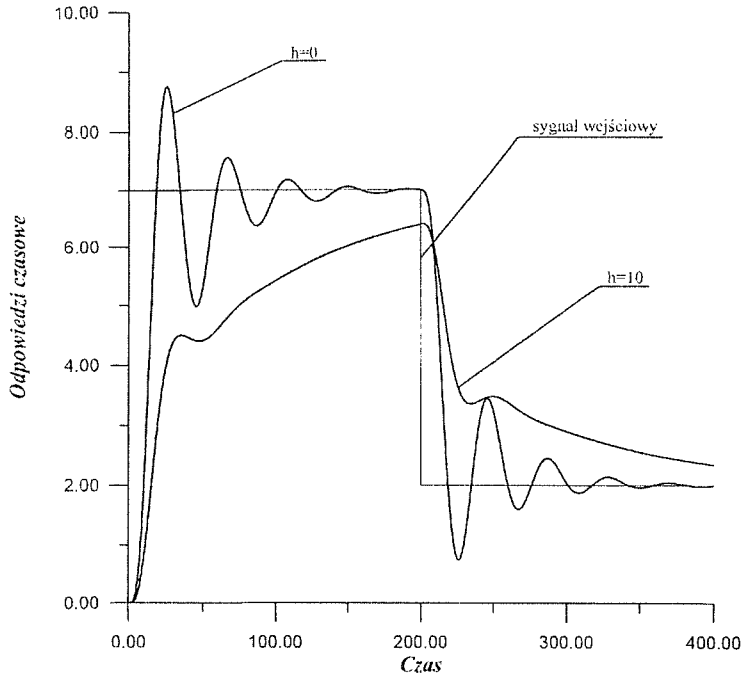
$$k_{krgr} = \frac{4}{\pi k}.$$

Histereza przekaźnika w eksperymencie identyfikacyjnym ma również duży wpływ na wyznaczenie okresu oscylacji krytycznych T_{osc}, który to rośnie prawie liniowo ze wzrostem histerezy (rys. 7), co wraz z niewłaściwie wyznaczonym wzmocnieniem krytycznym k_{kr} przy histerezie $h \neq 0$ prowadzi do poważnych zmian w doborze wartości nastaw regulatora, w następstwie czego następuje znaczne pogorszenie przebiegu procesu regulacji w układzie zamkniętym.



Rys. 7. Wykres $T_{osc} = f(h)$ dla obiektów o transmitancji $G(s) = \frac{k}{(Ts+1)^n}$

Porównanie przebiegów procesu regulacji nadążnej (tu: okresowo stałowartościowej), dla układu, w którym właściwie dobrano nastawy (przy histerezie przekaźnika w eksperymencie identyfikacyjnym $h = 0$) oraz w układzie, w którym dobrano nastawy regulatora na podstawie parametrów drgań krytycznych, uzyskanych w układzie z przekaźnikiem o histerezie $h \neq 0$ przedstawia rysunek 8.

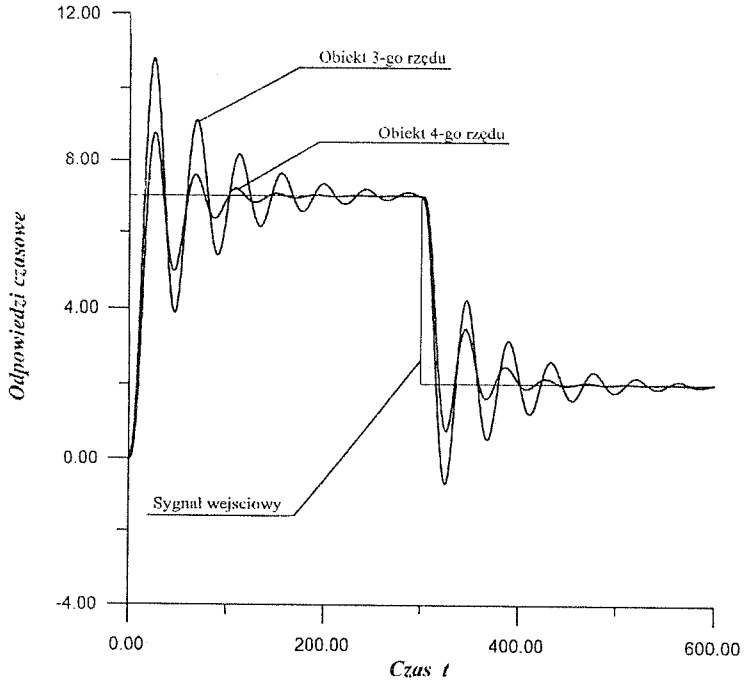


Rys. 8. Przebiegi przejściowe procesu regulacji w układzie zamkniętym dla obiektów wieloinercyjnych — dobór nastaw regulatora z wykorzystaniem metody przekaźnikowej (nastawy wg Zieglera – Nicholsa dla regulatora PI)

Nastawy regulatora PI wynoszą odpowiednio:

- dla $h = 0$:
  $k_p = 0.59$,
  $T_i = 26.83$;
- dla $h = 10$:
  $k_p = 0.25$,
  $T_i = 49.32$.

Wyniki przedstawione na rys. 8 uzyskano dla obiektu o transmitancji, wyrażonej zależnością (8) z następującymi wartościami parametrów: $n = 4$, $k = 3$, $T = 5$. Kolejnym istotnym mankamentem metody Astrőma–Hägglunda „odziedziczonym" po metodzie Zieglera–Nicholsa jest to, iż dane parametry drgań krytycznych A_{kr} (stąd też k_{kr}), T_{osc} charakteryzują nieskończenie wiele różniących się między sobą obiektów regulacji, co prowadzi do doboru jednakowych nastaw (dla danego typu regulatora: P, PI, PID) dla diametralnie różnych obiektów.



Rys. 9. Przebiegi przejściowe procesu regulacji w układzie zamkniętym dla różnych obiektów o takich samych parametrach drgań krytycznych (nastawy wg Zieglera – Nicholsa dla regulatora PI)

Efekt takiego doboru nastaw można zaobserwować na rysunku 9, gdzie dla dwóch różnych obiektów:
3-go rzędu o transmitancji

$$G(s) = \frac{6}{(8.66\,s + 1)^3}$$

oraz 4-go rzędu o transmitancji

$$G(s) = \frac{3}{(5\,s + 1)^4}$$

otrzymuje się te same parametry drgań krytycznych:

$$A_{kr} = 4.78, \quad \text{stąd} \quad k_r = 1.33,$$

$$T_{osc} = 31.42,$$

a co za tym idzie takie same nastawy np. dla regulatora PI:

$$k_p = 0.60,$$

$$T_i = 26.71.$$

Uwaga: Niewielkie różnice wartości nastaw dotyczących układów, których przebiegi procesów regulacyjnych przedstawiono na rys. 8 (dla $h = 0$) oraz rys. 9 wynikają z faktu, iż w pierwszym przypadku nastawy wyliczono w oparciu o eksperymentalne wyznaczenie parametrów drgań krytycznych, natomiast w drugim przypadku parametry te wyznaczono analitycznie.

Stosując eksperyment przekaźnikowy Astrőma–Hägglunda nie można podobnie jak w metodzie Preussa ustalić nastaw dla obiektów 1-go i 2-go rzędu, co wynika z niedokładności metody funkcji opisującej (rozpatruje się tylko 1-szą harmoniczną sygnału wyjściowego regulatora dwupołożeniowego). W myśl tej metody w układzie z obiektem 1-go lub 2-go rzędu i przekaźnikiem bez histerezy ($h = 0$) nie powinny występować drgania swobodne o stałej amplitudzie, chociaż w rzeczywistości drgania takie występują w obu przypadkach. Problemem jest również rozróżnienie czy w układzie znajduje się obiekt 1-go, 2-go czy wyższego rzędu. Jedynie w prosty sposób można odróżnić drgania w układzie z obiektem inercyjnym 1-go rzędu od innych sytuacji poprzez wprowadzenie histerezy do przekaźnika, wtedy amplituda drgań w tym układzie jest równa szerokości histerezy przekaźnika.

PODSUMOWANIE

Reasumując wysunięte wyżej wnioski można wypunktować zalety i wady badanych metod.

Zalety metody Preussa:

1. Identyfikowane modele dobrze odzwierciedlają dynamikę obiektu rzeczywistego dla obiektów intercyjnych.
2. Możliwość implementacji algorytmu w nieskomplikowanych systemach mikroprocsorowych.
3. Dobra dynamika przebiegu procesu regulacji przy wykorzystaniu tej metody do identyfikacji i doboru nastaw.

Wady metody Preussa:

1. Długi czas identyfikacji — zależny od dynamiki obiektu — konieczność wstępnej znajomości dynamiki obiektu.
2. Konieczność przerwania procesu regulacji — pomiarów dokonuje się w układzie otwartym.
3. Występowanie przeregulowania w zamkniętym układzie regulacji przy doborze nastaw regulatora PID wg zaleceń niniejszej metody.
4. Ograniczony zakres stosowania metody do układów inercyjnych oraz inercyjnych z opóźnieniem — niemożność stosowania metody do układów oscylacyjnych (założenie modelu Strejca).

Zalety metody Astrőma – Hägglunda:

1. Automatyzacja procesu doprowadzenia układu do stanu drgań krytycznych.
2. Brak konieczności otwierania pętli sprzężenia zwrotnego w układzie regulacji.
3. Krótki czas ustalania się drgań krytycznych w układzie — krótki czas identyfikacji — krótki czas występowania zakłócenia pomiarowego w układzie regulacji.
4. Wprowadzenie histerezy do przekaźnika dwupołożeniowego zmniejsza wpływ zakłóceń na proces ustalania drgań krytycznych w badanym układzie.
5. Prostota metody umożliwia łatwą jej implemenację w nieskomplikowanych systemach mikroprocesorowych (pomiar T_{osc} przez pomiar przedziału czasu pomiędzy kolejnymi przejściami przez zero sygnału wyjściowego, pomiar amplitudy A_{kr} poprzez wyliczenie powierzchni pod krzywą będącą w przybliżeniu sinusoidą).

Wady metody Astrőma – Hägglunda:

1. Możliwość określenia parametrów drgań krytycznych dla obiektów o transmitancjach rzędu większego niż 2.
2. Warunek występowania drgań w układzie $kB \geq h$.
3. Silna zależność wzmocnienia krytycznego od histerezy dla obiektów niskiego rzędu.
4. Niejednoznaczność identyfikacji — dane parametry drgań krytycznych (A_{kr}, T_{osc}) charakteryzują nieskończenie wiele obiektów regulacji.

BIBLIOGRAFIA

1. J.J. Aström, T. Hägglund: *Automatic tuning of simple regulators with specifications on phase and amplitude margins.* Automatica, 1984, vol. 20, nr 5, s. 645 – 651
2. L. Billmann, J. Rupp: *Adaptive PID-Regler für thermische Prozesse.* Automatisierungstechnische Praxis, 1989, vol. 31, nr 7, s. 327 – 330
3. C.C. Hang, K.K. Sin: *A comparative performance study of PID auto-tuners.* IEEE Control Systems, 1991, vol. 11, s. 41 – 47
4. T. Kaczorek: *Teoria układów regulacji automatycznej.* WNT, Warszawa 1974
5. J. Koj, M. Żelazny: *Badania porównawcze procedur samostrojenia regulatorów mikroprocesorowych.* Pomiary Automatyka Kontrola, 1996, nr 12, s. 325 – 329
6. T.W. Mucai, K. Wojtaś: *Algorytm samostrojenia zastosowany w regulatorze* EFTRONIK XS, typ U 496. Pomiary Automatyka Kontrola, 1997, nr 3, s. 86 – 89
7. A. Niederliński: *Systemy komputerowe automatyki przemysłowej.* t. 2 Zastosowania. WNT, Warszawa 1985
8. Z. Ogonowski: *Samonastrajający algorytm PID w regulatorze* EFTRONIK-M. Pomiary Automatyka Kontrola, 1989, nr 9 – 10, s. 219 – 221
9. H.P. Preuss: *Robuste Adaption in Prozessreglern.* Automatisierungstechnische Praxis, 1991, vol. 33, nr 4, s. 178 – 187
10. J.G. Ziegler, N.B. Nichols: *Optimum settings for automatic controllers.* Trans. ASME, 1943, vol. 65, s. 433 – 444

H. KUNERT

SELECTED IDENTIFICATION AND SETTING METHODS AS USED FOR SELF-TUNING CONTROLLERS

S u m m a r y

Two identification methods most frequently used in industrial microprocessor-based PID controllers including selecting methods for settings of continuous controllers have been subjected to censoriously considerations. The methods presented are classified to two groups of identification and setting selection: step function response methods, oscillation methods.

The step function response method as proposed by H.P. Preuss has been implemented in Siemens SIPART series controllers. The second method set forth by J.J. Aström and T. Hägglund makes use of Ziegler – Nichols oscillation experiment. The paper, basing on the simulation examinations carried out, discusses advantages and drawbacs of these methods.

Key words: identification methods, PID control, PID controller settings, computer simulation.

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI, 1998, 44, z. 3, ss. 427–442

Własności percepcyjne wzroku ludzkiego a charakterystyki monitorów

JERZY SIUZDAK, TOMASZ CZARNECKI

Instytut Telekomunikacji, Politechnika Warszawska

Otrzymano 1998.05.08

Autoryzowano 1998.06.20

Celem badań, których opis zawarto w niniejszej publikacji, było zbadanie zlaeżności między własnościami wzroku ludzkiego a charakterystykami monitorów telewizyjnych dla obrazów monochromatycznych. Omowiono dosyć dobrze znane wasności wzroku ludzkiego w odniesieniu do funkcji wrażliwości na kontrast. Zbadano zależność luminancji ekranu od poziomu sterowania dla rzeczywistych monitorów i mówiono tzw. linearyzację percepcyjną, czyli wzrokową optymalizacją zależności liminancji monitora od poziomu sterowania.

Słowa kluczowe: monitory telewizyjne, wzrok ludzki, luminancja ekranu.

1. WSTĘP

Celem badań, których opis zawarto w nineijszej publikacji, było zbadanie zależności miedzy własnościami wzroku ludzkiego a charakterystykami monitorów telewizyjnych dla obrazów monochromatycznych. W pracy można wydzielić dwie zasadnicze części. W pierwszej omówiono dosyć dobrze znane wlasności wzroku ludzkiego w odniesieniu do funkcji wrażliwości na kontrast. Druga część jest związana z badaniami rzeczywistych monitorów i tzw. linearyzacją percepcyjną, czyli wzrokową optymalizacją zależności luminancji monitora od poziomu sterowania.

Na wstępie, za pracą [9], omówimy bardzo krótko mechanizm widzenia człowieka. Wzrok ludzki składa się z trzech podsystemów [9]: optycznego, detekcji i systemu wstępnego przetwarzania odebranych sygnałow. System optyczny oka [9] składa się z rogówki pełniącej rolę soczewki skupiającej, źrenicy, regulującej ilosć światła wnikając, do oka oraz soczewki mającej zdolności akomodacyjne. Promienie świetlne są odwzorowywane głównie na plamce żółtej będącej fragmentem siatkówki o największej rozdzielczości.

System detekcji [9] bazuje na zjawisku zamiany w fotoreceptorach siatkówki fali elektromagnetycznej na potencjaly elektrochemiczne, które są następnie przenoszone jako impulsy do kory mózgowej. Istnieją dwa rodzaje komórek fotoreceptorowych: pręciki o maksimum absorpcji dla długości fali równej 500 nm oraz trzy rodzaje czopków o maksimach absorpcji na długościach fal 450, 530 i 570 nm. Absorpcja pojedyńczego fotonu [9] zapoczątkowuje reakcje chemiczne powodujące zmianę różnicy potencjałów między wnętrzem komórki fotoreceptora a środowiskiem. Zmiana polaryzacji jednego miejsca komórki nerwowej pociąga za sobą zmianę polaryzacji sąsiednich odcinków komórki, a taka fala zmiany polaryzacji podążająca wzdłuż włókna nosi nazwę impulsu nerwowego. Receptory są zorganizowane w dwa systemy reagowania zależne od natężenia światla [9]:
— widzenie skotopowe, przy małych natężeniach światla, kiedy w procesie widzenia biorą udział pręciki, oko reaguje na energię pojedyńczych fotonów (próg widzenia poniżej 20 fotonów na sekundę), ale rozróżnianie kolorów nie jest możliwe,
— widzenie fotopowe, przy dużych natężeniach światla, w procesie widzenia biorą udział czopki i rozróżnianie kolorów jest możliwe.

W trakcie wstępnego przetwarzania redukowana jest ilość informacji docierająca do mózgu. W oku istnieje około 120 milionów czopków i 6 milionów pręcików, natomiast informacja wzrokowa jst przekazywana w głąb mózgu (nerw wzrokowy) za pomocą około 1 mln komórek zwojowych [9]. Komórka zwojowa łączy się z fotoreceptorami (przez komórkę dwubiegunową) w stosunku 1:1 tylko w obrębie plamki żołtej. W pozostałych obszarach oka jednej komórce zwojowej odpowiada wiele fotoreceptorów zgrupowanych w tzw. pole receptorowe. Podczas przenoszenia bodźca z receptorów do komórek zwojowych również forma bodźca ulega zmianie. Powolne zmiany potencjału receptorowego zostaja zamienione w potencjały iglicowe. Można to porównać do zamiany postaci analogowej na postać binarną.

2. MODELE MATEMATYCZNE WRAŻLIWOŚCI WZROKU LUDZKIEGO

Dla określenia wrażliwości wzroku ludzkiego wykorzystuje się pojęcie kontrastu progowego lub jego odwrotności, a mianowicie tzw. funkcji wrażliwości na kontrast (ang. *contrast sensitivity function*). Kontrast progowy jest to kontrast obiektu, przy którym obserwator zaczyna rozróżniać obiekt od tła. Zgodnie z definicją kontrastu Michelsona, kontrast wyraża się zależnością

$$C = \frac{L_{\max} - L_{av}}{L_{av}}, \tag{1}$$

gdzie $L_{\max}$ jest wartością maksymalną luminancji, zaś L_{av} jest wartością średnią luminancji. Z kolei wrażliwość jest definiowana jako $S = 1/C_T$, gdzie C_T jest kontrastem progowym. Ze względu na różne własności wzroku u różnych osób, powyższe pojęcia uwzględniają uśrednienie badań na dużą grupę obserwatorów. Kontrast progowy lub jego odwrotność jest mierzony zazwyczaj dla bardzo prostych obiektów, takich jak jednorodne kwadraty o różnych wymiarach lub sinusoidalne

siatki o różnej częstości przestrzennej. Jedną z pierwszych prób ilościowego określenia zależności kontrastu progowego została podjęta przez Rogersa i Carela [1]. Na podstawie badań empirycznych i opracowania ich wyników za pomocą analizy regresyjnej drugiego rzędu, znaleźli oni formułę empiryczną uzależniającą kontrast progowy siatki sinusoidalnej od parametrów, takich jak: luminancja siatki, luminancja tła, wymiary kątowe siatki i jej kątowa częstość przestrzenna.

Z nowszych modeli wrażliwości wzroku ludzkiego na kontrast można wspomnieć o modelach Bartena [2] i Daly'ego [3]. Omówimy je nieco dokładniej. Model Bartena [2] uwzględnia następujące zjawiska i parametry:

– wzrost kontrastu progowego przy szumach zewnętrznych (np. szumach kineskopów),

– wymiary przestrzenne rozpatrywanego obrazu,

– szumy wewnętrzne: fluktuacje fotoprądu docierającego do fotoreceptorów w oku, które są zależne od luminancji obrazu i średnicy źrenicy oraz szum neuronowy generowany w systemie nerwowym,

– filtrację górnopasmową na wstępnym etapie obróbki obrazu w systemie nerwowym, spowodowaną przez boczną inhibicję w komórkach zwojów nerwowych; boczna inhibicja polega na odejmowaniu od obrazu jego przefiltrowanej dolnopasmowo wersji,

– funkcję przenoszenia modulacji (*modulation transfer function*) oka ludzkiego, która powoduje filtrację dolnopasmową sygnału przestrzennego, obejmującą również aberracje oraz światło rozproszone.

Uwzględnienie tych zjawisk daje następujące wyrażenie na funkcję wrażliwości na kontrast [1, 2]:

$$S(u) = \frac{1}{k}\sqrt{\frac{T}{2}}\,\frac{M_{\mathrm{opt}}(u)}{\sqrt{\left(\dfrac{1}{\eta h I} + \dfrac{\Phi_0}{[1-F(u)]^2} + \Phi_{\mathrm{ext}}(u)\right)\left(\dfrac{1}{X_0^2} + \dfrac{1}{X_E^2} + \left(\dfrac{u}{N_E}\right)^2\right)}}. \tag{2}$$

W powyższym wzorze u jest częstością przestrzenną w cyklach na stopień kątowy. Zjawiska szumowe są tu uwzględnione w pierwszym nawiasie pod pierwiastkiem: pierwszy składnik jest związany z wewnętrznym szumem fotonowym, drugi — z przefiltrowanym szumem neuronowym, a trzeci $(\Phi_{\mathrm{ext}}(u))$ wyraża gęstość widmową szumów zewnętrznych. Drugi nawias pod pierwiastkiem wynika z ograniczonych zdolności wzroku ludzkiego do uśredniania przestrzennego; istnieje maksymalny wymiar kątowy oraz maksymalna liczba cykli, w których wzrok może uśredniać (całkować) informacje w obecności różnych źródeł szumowych. Drugi człon wyraża to, że kąt uśredniania jest równy najmniejszemu z trzech parametrów: wymiarom kątowym obiektu, maksymalnemu kątowi uśredniania oraz kątowi wizualnemu wyznaczonemu przez liczbę cykli. Optyczna funkcja przenoszenia $M_{opt}(u)$ jest dana zależnością

$$M_{opt}(u) = \exp(-\pi^2\sigma^2u^2), \tag{3}$$

gdzie wartość σ zależy od średnicy źrenicy oka.

We wzorze (2) I jest iluminacją (oświetleniem) oka wyrażoną przez $I = (\pi/4)\,d^2L$ (d — średnica źrenicy, L to iluminancja obiektu w cd/m^2), $k = 3.3$, $\eta = 0.025$, $h = 357.3600$ fotonów/(td sec deg^2), $T = 0.1$ sec. Wariancja kontrastu odpowiadająca szumowi neuronowemu wynosi $\Phi_0 = 3 \cdot 10^{-8}$ sec deg^2. Z kolei funkcja opisująca filtrację szumu neuronowego dana jest przez $F(u) = \exp(-u^2/u_0^2)$, gdzie $u_0 = 8$ cykli/deg. Wartości pozostałych parametrów są następujące: $X_E = 12$ deg, $N_E = 15$ cykli. Zależność (2) dobrze zgadza się z danymi eksperymentalnymi dla $10^{-4} < L < 10^3$ cd/m^2,

$$0.5 < X_0 < 60 \text{ deg}, \quad 0.2 < u < 50 \text{ cykli/deg}.$$

Z kolei model Daly'ego [3] uwzględnia następujące zjawiska:
— własności adaptacyjne siatkówki oka, powodujące nieliniową ależność czułości oka od luminancji,
— zmiany czułości w zależności od częstotliwości przestrzennej i odległości obserwacji spowodowane akomodacją oka, efektywną aperturą fotoreceptorów oraz pasywnymi i aktywnymi połączeniami neuronowymi,
— zmiany wrażliwości oka w zależności od ekscentryczności obiektu (pozycji w polu widzenia) i szumów zewnętrznych.

Natomiast należy zaznaczyć, że nie są w tym modelu uwzględnione zjawiska związane z własnościami detekcyjnymi kory mózgowej, takie jak: hierarchizacja częstości przestrzennych, własności maskujące związane z kontrastem w ssiednich obszarach, funkcje psychometryczne opisujące wzrost prawdopodobieństwa wykrycia przy wzroście kontrastu itd. W tym modelu funkcja wrażliwości na kontrast wyraża się wzorem:

$$S(u, \Theta, L, X_0^2, v_d, E) = S_0 \min\left[S\left(\frac{u}{bw_a \cdot bw_E \cdot bw_\Theta}, L, X_0^2\right), S(u, L, X_0^2)\right], \tag{4}$$

gdzie [3]

$$bw_a = 0.856\, v_d^{0.14},$$

$$bw_E = \frac{1}{1 + kE},$$

$$bw_\Theta = \left(\frac{1 - ob}{2}\right)\cos 4\Theta + \frac{1 + ob}{2}.$$

Z kolei

$$S(u, L, X_0^2) = [(3.23 \cdot (u^2 \cdot X_0^2)^5 + 1]^{-0.2} A \in ue^{-B \in u} \cdot \sqrt{1 + 0.06e^{B \in u}}, \tag{6}$$

gdzie

$$A = 0.8(1 + 0.7/L)^{-0.2}, \quad B = 0.3(1 + 100/L)^{0.15}. \tag{7}$$

S_0 jest maksymalną czułością zależną od wrażliwości osobniczej. Przeciętna wartość wynosi 250. Pozostałe oznaczenia są następujące:
v_d — odległość obserwacji w metrach,
L — luminancja w cd/m^2,
X_0 — wymiary kątowe obiektu w stopniach,
E — ekscentryczność w stopniach, $k = 0.24$,
θ — orientacja obiektu w stopniach, $ob = 0.7$
$\epsilon = 0.9$.

W przypadku obecności szumu zewnętrznego, funkcja wyrażona zależnością (4) musi zostać zastąpiona przez [3]

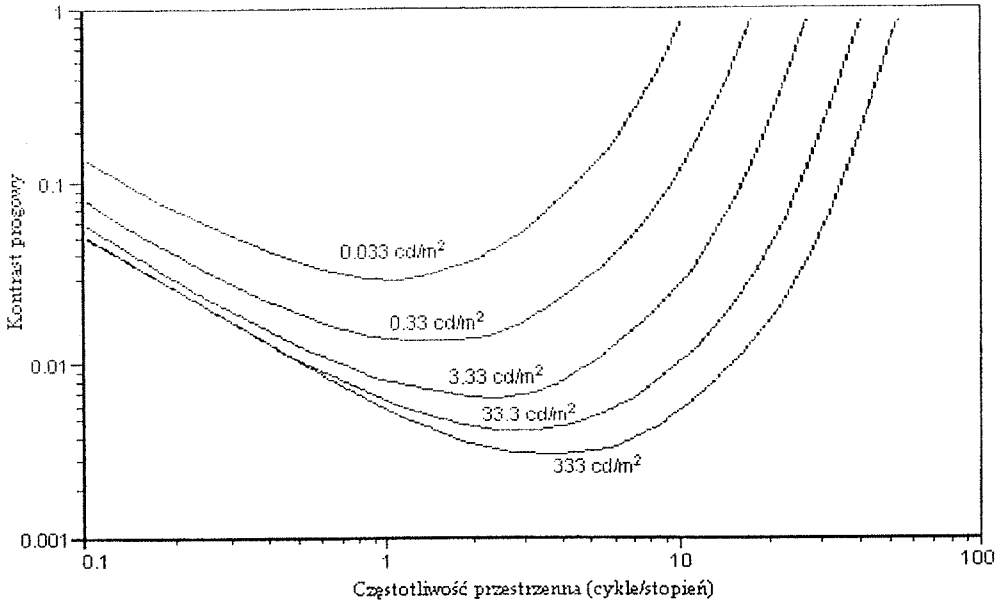
$$S(u, \Theta, L, X_0^2, v_d, E, n_{rms}) = \frac{1}{\sqrt{\frac{1}{S^2(u, \Theta, L, X_0^2, v_d, E) + W_n^2 \cdot n_{rms}^2}}}, \tag{8}$$

gdzie n_{rms} jest wartością średniokwadratowej gęstości przestrzennej szumów na częstotliwość przestrzenną. Model ten daje rezultaty zgodne z doświadczeniem w następującym zakresie parametrów: luminancja $10^{-4} < L < 10^4$ cd/m^2, częstości przestrzenne $0.1 < u < 50$ cykli/deg, orientacja $0 < \theta < 180$ deg, ekscentryczność $0 < E < 20$ deg, odległość obserwacji $0.25 < v_d < 3$ m.

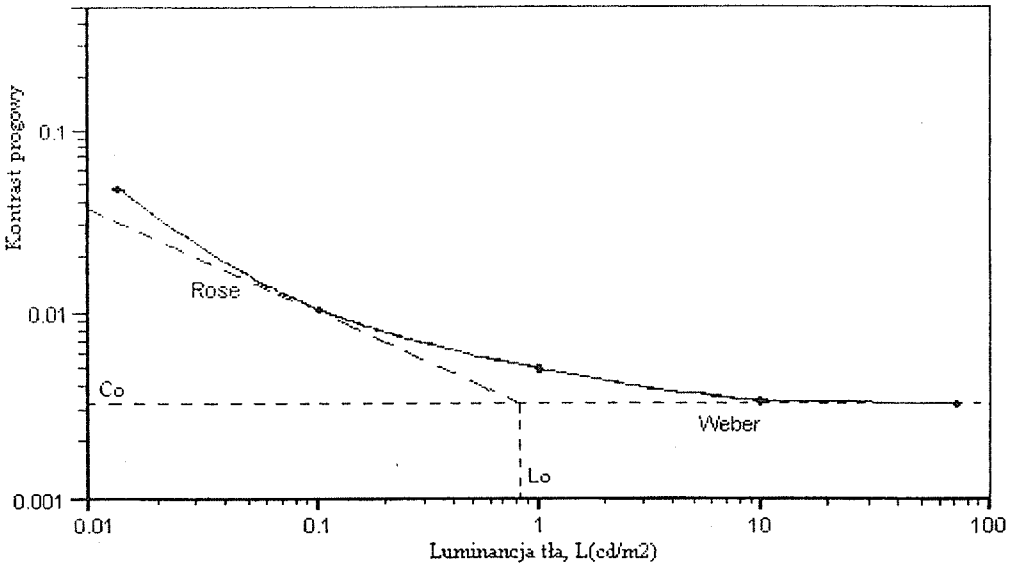
3. NAJWAŻNIEJSZE WŁASNOŚCI WZROKU LUDZKIEGO

Powyższe modele, choć niemal kompletne, są trudne do bezpośredniego zastosowania ze względu na swe skomplikowanie. Stąd w dalszych rozważaniach skoncentrujemy się na pewnych uproszczeniach własności wzroku odzwierciedlanych przez wszystkie modele. Do podstawowych własności wzroku, ujętych przez wszystkie modele i potwierdzonych badaniami doświadczalnymi, należy zależność kontrastu progowego od poziomu luminancji oraz od częstości przestrzennej obserwowanego wzorca. Zależności te pokazano na rys. 1 i rys. 2 [4]. Z obydwu tych rysunków widać, że kontrast progowy maleje wraz ze wzrostem luminancji. Można tu wyróżnić dwa obszary:
— pierwszy tzw. Rose- de Vries, w którym dominuje szum fotonowy o rozkładzie Poissona i w którym kontrast progowy C_T jest odwrotnie proporcjonalny do pierwiastka z luminancji L: $C_T \sim 1/\sqrt{L}$,
— drugi tzw. Webera, w którym kontrast progowy jest niezależny od luminancji: $C_T = \text{const}$.

Należy także zauważyć, że dla każdego poziomu luminancji istnieje pewna częstość przestrzenna, na którą oko ludzkie jest najbardziej wrażliwe. Zmiany tej



Rys. 1. Konrast progowy w funkcji częstości przestrzennej przy różnych poziomach luminancji. Na podstawie [4]



Rys. 2. Typowa zależność kontrastu progowego od luminancji. Na podstawie [4]

częstości dla czterech rzędów wielkości zmian luminancji są stosunkowo niewielkie: dla małych wartości luminancji ta częstość wynosi około 1 cykl/deg, zaś dla dużych poziomów luminancji — około 4 cykl/deg.

4. WŁASNOŚCI SZUMOWE MONITORÓW

Ponieważ o własnościach wzroku ludzkiego mówimy tutaj w kontekście obserwacji obrazu na monitorze trzeba wspomnieć o charakterystycznym w tym przypadku zjawisku, a mianowicie o szumach kineskopów [7, 8]. Można je podzielić na szum czasowy i szum przestrzenny. Szum czasowy jest spowodowany przez szum wzmacniacza sygnału wideo, szum śrutowy wiązki elektronów oraz szum fotonowy. Z kolei szum przestrzenny jest spowodowany przez ziarnistość luminoforu i zmiane jego własności w zależności od położenia. Szum czasowy polega na przypadkowych zmianach w czasie luminancji danego punktu na powierzchni ekranu. Szum przestrzenny jest spowodowany przez przypadkowe zmiany luminancji punktów na powierzchni ekranu w zależności od ich położenia przy sterowaniu mającym zapewnić jednakową luminancję. W odróżnieniu od szumu czasowego, szum przestrzenny nie ulega zmianom w czasie. Jasne jest, że szumy te występują jednocześnie i na przyklad mierząc parametry szumu przestrzennego należy uwzględniać szum czasowy.

Obecność szumu pogarsza zdolność postrzegania. Aby spostrzec różnicę luminancji ΔL musi ona być większa o pewien czynnik od odchylenia średniokwadratowego szumów σ [7, 8]:

$$\Delta L = k_0 \sigma. \tag{9}$$

Tutaj wartość k_0 może być interpretowana jako wartość progowa stosunku sygnału do szumu. Jest ona ściśle związana z prawdopodobieństwem wykrycia (spostrzeżenia). Dla prawdopodobieństwa spostrzeżenia 50% wielkość k_0 powinna być rzędu 3.1, zaś detekcja z prawdopodobieństwem bliskim 100% wymaga wartości k_0 przekraczających 5 [7, 8]. Wartość stosunku sygnału do szumu SNR dla obiektu o małym kontraście, wyświetlanego na ekranie monitora można wyrazić zależnością [7]

$$SNR = \sqrt{\frac{n}{2}} \frac{L_b}{\sigma_p} 2C, \tag{10}$$

gdzie n jest liczbą piksli należących nominalnie do obiektu, L_b — luminancją tła, σ_p — odchyleniem standardowym szumu (czasowego i przestrzennego) kineskopu dla pojedynczego piksla tła, wreszcie C jest kontrastem obiektu zdefiniowanym we wzorze (1). Jeśli przez $SNR_p = L_b/\sigma_p$ oznaczymy stosunek sygnału do szumu dla pojedynczego piksla kineskopu, to wówczas kontrast progowy C_T można wyrazić wzorem [7]

$$C_T = \sqrt{\frac{1}{2n}} \frac{k_0}{SNR_p}. \tag{11}$$

Zależność SNR_p od luminancji dla dwóch różnych typów monitorów pokazano na rys. 4 [7]. Z przedstawionych zależności widać, że dla dużych poziomów luminancji stosunek SNR_p dla pojedynczgo piksla nie zależy od wartości luminancji. Dominującym źródłem szumu jest wtedy szum przestrzenny. Z kolei, dla coraz mniejszych wartości luminancji stosunek sygnału do szumu maleje, a o własnościach szumowychzaczyna decydować szum czasowy. Przy rozpatrywaniu wzoru (11) należy zwrócić uwagę na to, że jest to szczególny przypadek omówionych poprzednio modeli Bartena i Dale'a, w których uwzględniono jedynie wpływ szumów zewnętrznych. Oznacza to w szczególności, że liczba piksli n nie może dowolnie wzrastać i zgodnie z modelem Bartena jest ograniczona maksymalnym kątem, w którym wzrok ludzki może uśredniać obraz ($\sim 12°$). Ponadto wzor (11) nie uwzględnia wpływu częstości przestrzennej, luminancji i innych parametrów mających wpływ na własności wzroku ludzkiego.

5. LINEARYZACJA PERCEPCYJNA

Dla prawidłowego odbioru obrazu istotne jest, aby jednostkowe zmiany poziomu sterowania jaskrawością obrazu robiły na obserwatorze wrażenie jdnakowych niezależnie od poziomu tej jaskrawości (i poziomu sterowania). Linearyzacja percepcyjna polega na właśnie takim przekodowaniu sygnału sterującego, aby równe zmiany w poziomach sterowania (wartościach cyfrowych sygnału wejściowego) wytwarzały zmiany luminancji, które są percepcyjnie równoważne w całym zakresie luminancji.

Dla linearyzacji percepcyjnej wykorzystuje się omówione poprzednio pojęcia kontrastu progowego C_T i funkcji wrażliwości na kontrast. Oznaczymy poziom sterowania (wartości cyfrowe na wejściu) przez G, zaś przez P oznaczymy postrzeganą jaskrawość obrazu. Postrzegana jaskrawość obrazu jest związana z jego luminancją L zależnością [5]

$$P = h(L) \approx \int_{L_{\min}}^{L} \frac{dL}{LC_T(L)}. \tag{12}$$

gdzie $C_T(L)$ jest wartością kontrastu progowego w zależności od luminancji pokazaną na rys. 2. Naszym zadaniem jest znalezienie takiego przekształcenia $L = f(G)$ poziomu sterowania G na luminancję L, która przy zwiększeniu o 1 poziomu sterowania powoduje liniową zmianę w postrzeganej jaskrawości. Łatwo można wykazać [5], że ta funkcja dana jest wzorem

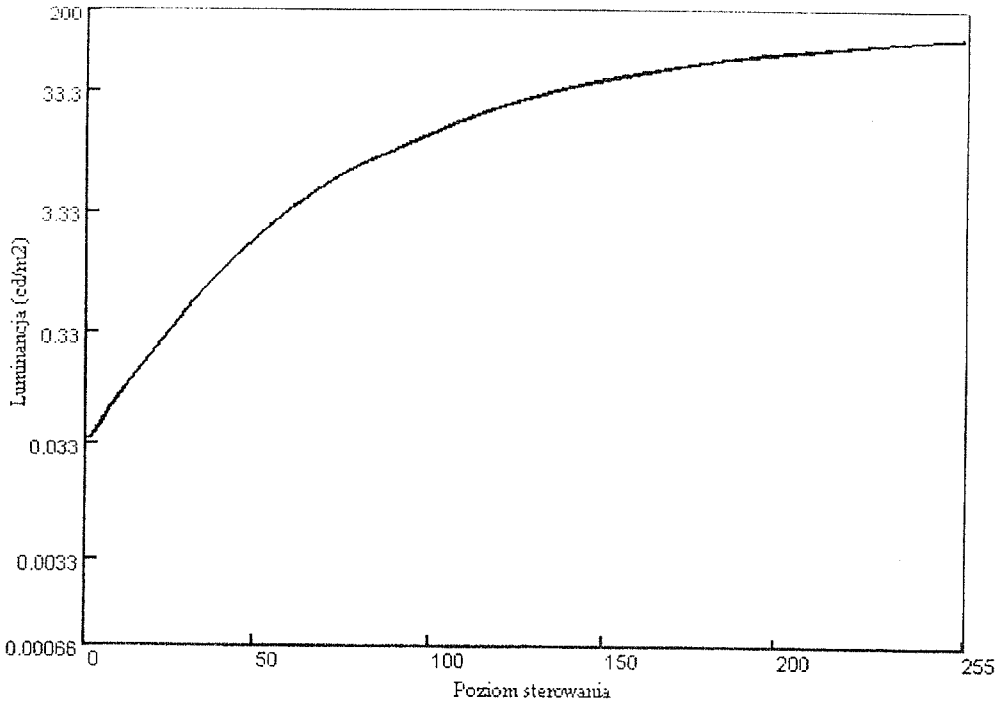
$$L = f(G) = h^{-1}\left(\frac{P_{\max}}{G_{\max}} G\right), \tag{13}$$

gdzie h^{-1} jest funkcją odwrotną do funkcji h wyrażonej przez (12), zaś G_{max} jest maksymalnym poziomem sterowania, a P_{max} jest maksymalną postrzeganą jaskrawością.

Istotnie, mamy

$$P = h[f(G)] = h\left[h^{-1}\left(\frac{P_{max}}{G_{max}}G\right)\right] = \frac{P_{max}}{G_{max}}G \tag{14}$$

Przykład optymalnej funkcji linearyzującej $f(G)$ pokazano na rys. 3.



Rys. 3. Przykład optymalnej funkcji linearyzującej $f(G)$ [6]

Wielkość kontrastu progowego można przybliżyć wzorem [5]

$$C_T(L) = C_0\left[1 + \left(\frac{L_0}{L}\right)^{pm}\right]^{1/m}, \tag{15}$$

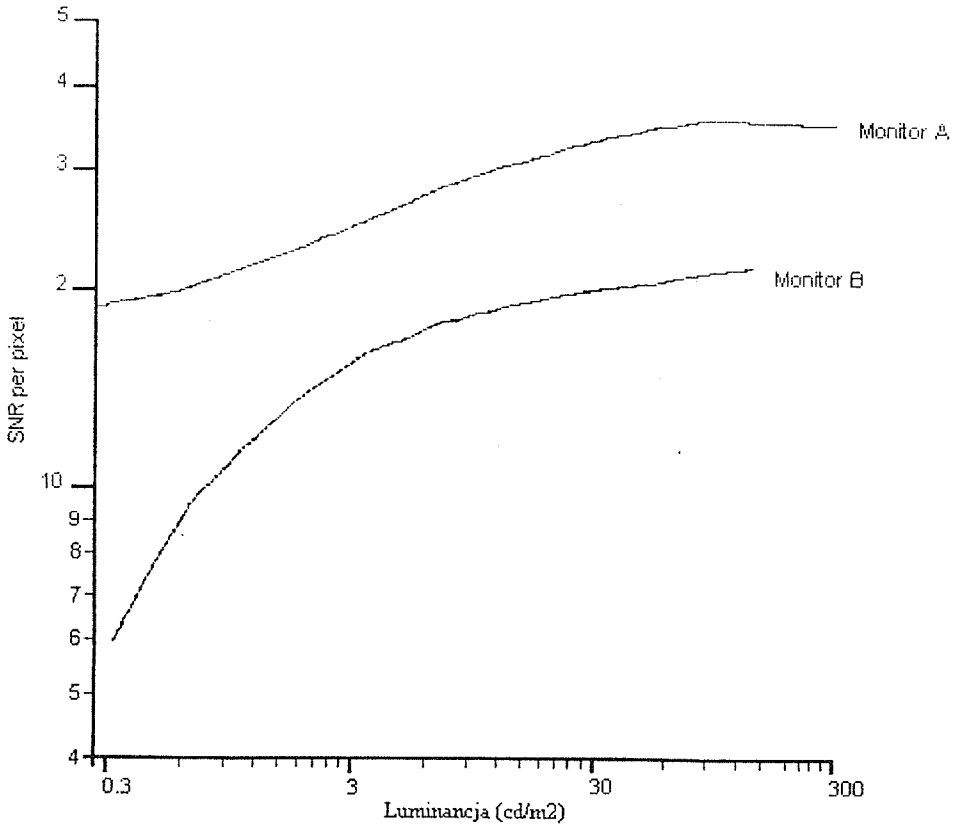
W powyższym wzorze L_0 jest luminancją określającą granicę między regionem Rose-de Vries a regionem Webera, C_0 jest wartością kontrastu progowego, p określa nachylenie charakterystyki w regionie Rose- de Vries, zaś m określa przebieg charakterystyki w rejonie przejściowym. Dla obliczenia funkcji linearyzującej w jawnej postaci skorzystamy z asymptotycznej postaci kontrastu progowego $C_T(L)$, danej w pracy [5]

$$C_T(L) = \begin{cases} C_0 & \text{dla} \quad L \geq L_0 \\ C_0\sqrt{\dfrac{L_0}{L}} & \text{dla} \quad L < L_0 \end{cases}. \tag{16}$$

Wielkości parametrów L_0, C_0 zależą m.in. od rozmiarów, rodzaju i otoczenia obserwowanych obiektów. Podstawiając (16) do (12) otrzymujemy

$$\begin{aligned} h(L) &= \frac{2}{C_0\sqrt{L_0}}(\sqrt{L} - \sqrt{L_{\min}}) \approx \frac{2}{C_0}\sqrt{\frac{L}{L_0}} \quad \text{dla} \quad L \leq L_0, \\ h(L) &= \frac{2}{C_0\sqrt{L_0}}(\sqrt{L_0} - \sqrt{L_{\min}}) + \frac{1}{C_0}\ln\left(\frac{L}{L_0}\right) \approx \frac{2}{C_0} + \frac{1}{C_0}\ln\left(\frac{L}{L_0}\right) \quad \text{dla} \quad L > L_0. \end{aligned} \tag{17}$$

Wreszcie, funkcja linearyzująca percepcyjnie może być otrzymana przez podstawienie wzoru (17) do (13). Dostajemy wtedy



Rys. 4. Zależność SNR_p od luminancji dla dwóch różnych typów monitorów [7]

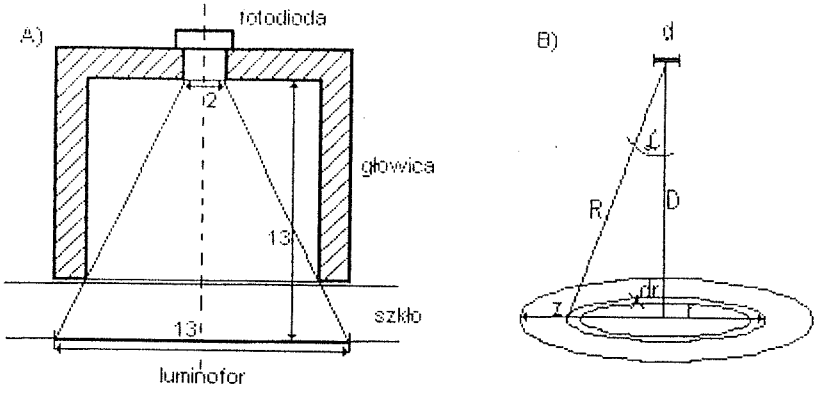
$$L(G)=L_0(1+0.5\ln(L_{max}/L_0))^2\left(\frac{G}{G_{max}}\right)^2 \quad \text{dla} \quad L<L_0, \left(G<\frac{G_{max}}{1+0.5\ln(L_{max}/L_0)}\right),$$

$$L(G)=L_0\exp\left[\left(2+\ln\frac{L_{max}}{L_0}\right)\frac{G}{G_{max}}-2\right] \quad \text{dla} \quad L\geq L_0, \left(G\geq\frac{G_{max}}{1+0.5\ln(L_{max}/L_0)}\right). \tag{18}$$

Jak więc widać, funkcja $L(G)$ nie zależy od wartości C_0 a jedynie od L_{max} oraz L_0. Funkcje linearyzujące percepcyjnie są pokazane na rys. 7 dla 3 różnych wartości parametru L_0.

6. WYNIKI BADAŃ RZECZYWISTYCH MONITORÓW

Interesujące jest zbadanie, jaki przebieg ma funkcja $L(G)$ w rzeczywistych monitorach. Przebadano kilka typów monitorów pod względem zależności luminancji ekranu od poziomu sterowania. W tym celu został napisany krótki program pozwalający nadać całemu ekranowi jednolitą luminancję odpowiadającą wartości sterowania o żądanej wielkości z przedziału 0...255. Monitory były monitorami kolorowymi, ale program pozwalał wyświetlać odpowiednie poziomy szarości. Do szacunkowych pomiarów luminancji energetycznej służyła głowica odbiorcza miernika tłumienności światłowodu, pokazana na rys. 5a). Pomiar polegał na przytknięciu głowicy do powierzchni ekranu i odczytaniu mocy odbieranej przez głowicę odpowiadajacej różnym poziomom sterowania. Następnie pomiary były korygowane poprzez odjęcie mocy światła rozproszonego pochodzącego od źródeł innych niż ekran (mocy światła przy całkowicie wygaszonym ekranie). W ten sposób łatwo jest otrzymać względną zależność luminancji np. w odniesieniu do poziomu maksymalnego, gdyż odbierana przez głowicę moc jest oczywiście do tej luminancji proporcjonalna. Znacznie trudniejszą sprawą jest wyznaczenie luminancji energetycznej, a zwłaszcza luminancji (jasności) fotometrycznej.



Rys. 5. Pomiary luminancji enrgetycznej: a) wygląd głowicy, b) oznaczenia używane przy obliczeniach

Dokładne wyliczenia, oprócz znajomości zależności czułości widmowej fotodetektora (która była znana), wymagają również znajomości charakterystyki widmowej światla emitowanego przez ekran (która nie była znana). Dlatego obliczenia luminancji mają jedynie charakter szacunkowy. Oznaczenia stosowane dla obliczeń strumienia świetlnego pokazano na rys. 5b). Używając oznaczeń z tego rysunku, można wyznaczyć kąt bryłowy, jaki zajmuje apertura fotodetektora widziana z elementu ekranu odległego o r od środka symetrii:

$$\omega_0 = \frac{\frac{\pi}{4} d^2 \cos\alpha}{D^2 + r^2} = \frac{\pi}{4} \frac{Dd^2}{(D^2 + r^2)^{3/2}}. \tag{19}$$

Zmienne α, D, d, r są określone na rys. 5b. Element powierzchni, który widzi aperturę fotodetektora pod takim kątem bryłowym, jest równy

$$\Delta S = 2\pi r\, dr. \tag{20}$$

Z kolei natężenie promieniowania I związane jest z luminancją energetyczną L_e związkiem:

$$I = L_e \cos\alpha\, \Delta S, \tag{21}$$

a z mocą promieniowania Φ zależnością

$$I = \frac{d\Phi}{d\omega}. \tag{22}$$

Zatem łącząc zależności (19) do (22) można obliczyć moc pochodzącą od wszystkich elementów odległych o r od środka obszaru emitującego światło:

$$d\Phi(r) = \omega_0 I = \omega_0 L_e \cos\Delta S = \frac{\pi^2}{2} \frac{d^2 D^2}{(D^2 + r^2)^2} L_e r\,\mathrm{d}r. \tag{23}$$

Całkowita moc wyraża się zależnością

$$\Phi = \int_0^Z d\Phi(r)\,\mathrm{d}r = \frac{\pi^2}{2} L_e d^2 D^2 \int_0^Z \frac{r\,\mathrm{d}r}{(D^2 + r^2)^2}. \tag{24}$$

Zmienna Z jest określona na rys. 5b. Licząc powyższą całkę przez podstawienie $u = D^2 + r^2$, otrzymujemy

$$\Phi = \frac{\pi}{4} L_e d^2 \frac{1}{1 + (D/Z)^2}. \tag{25}$$

Po podstawieniu wartości liczbowych dostajemy ostatecznie

$$\Phi \approx 2 \cdot 10^{-6} L_e \ . \tag{26}$$

Powyższa zależność jest przybliżonym związkiem między luminancją (energetyczną) L_e, a mocą światła padającego na detektor. Jeśli gęstość widmową mocy emitowanej przez źródło światła oznaczyć przez $N(\lambda)$, to wtedy luminancję fotometryczną L można wyrazić wzorem

$$L \approx 0.5 \cdot 10^6 \Phi \cdot K_{\max} \int_0^{+\infty} \frac{R(0.85\ \mu\text{m})}{R(\lambda)} W(\lambda) N(\lambda) \cdot d\lambda. \tag{27}$$

Tutaj Φ jest mocą świetlną wskazywaną przez przyrząd (w watach), $R(.)$ jest czułością fotodetektora na odpowiedniej długości fali, $K_{\max} = 673\,\text{lm/W}$ [9], zaś $W(\lambda)$ jest widzialnością względną dla widzenia dziennego [9].

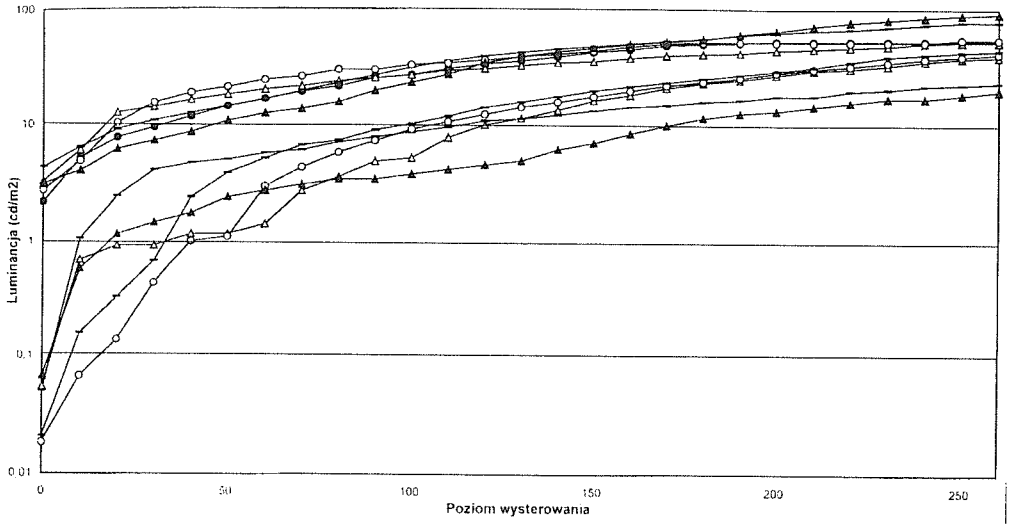
Niestety, jak już wspomniamo, nie była znana charakterystyka widmowa $N(\lambda)$ światła emitowanego przez monitor, co uniemożliwiało dokładne obliczenie, zarówno mocy promieniowania padającego na detektor, jak i luminancji fotometrycznej, gdyż czułość detektora była zależna od długości fali. Dla obliczeń szacunkowych przyjęto, że efektywna długość fali emitowana przez monitor była równa średniej z długości fal dla której poszczególne fotoreceptory w oku mają maksymalną czułość, a więc około 520 nm. Zatem, ostatecznie, wskazanie Φ przyrządu mierzącego moc optyczną można było związać z luminancją fotometryczną za pomocą szacunkowej zależności:

$$L \approx \Phi \cdot \frac{R(0.85\ \mu\text{m})}{R(0.52\ \mu\text{m})} K_{\max} W(\lambda = 0.52\ \mu\text{m}) \cdot 0.5 \cdot 10^6 \approx 0.6 \cdot 10^9 \Phi \frac{cd}{m^2 W}, \tag{28}$$

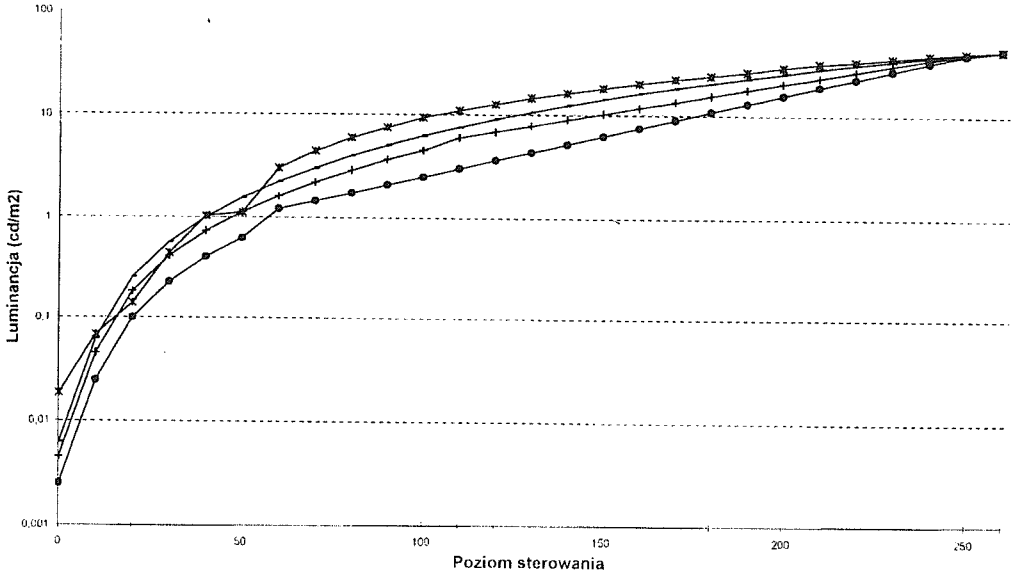
gdzie $W(\lambda = 0.52\ \mu\text{m}) = 0.71$.

Na rys. 6 pokazano wyniki badań zależności luminancji od poziomu sterowania dla różnych monitorów. Widać, że zakres zmian luminancji zwłaszcza dla gorszych typów monitorów jest bardzo niewielki, a charakter tych zmian bardzo odbiega od charakterystyk linearyzowanych percepcyjnie. Natomiast monitory wyższej klasy mają przebieg charakterystyki zbliżony do charakterystyki linearyzowanej percepcyjnie. Należy podkreślić, że bardzo istotny wpływ na przebieg charakterystyk miało ustawienie poziomu kontrastu i jaskrawości monitora. Charakter otrzymanych krzywych jest bardzo podobny do wyników badań monitorów o wysokiej rozdzielczości używanych w radiologii, prezentowanych w pracy [10].

Z badań wynika, że charakterystyki rzeczywistych monitorów należałoby korygować tak, aby były one maksymalnie zbliżone do charakterystyk linearyzowanych percepcyjnie. Można tego dokonać przez odpowiednią zmianę poziomu sterowania monitora. Jeśli linearyzowana zależność luminancji od poziomu sterowania dana jest wzorem (18) $L = L(G)$, zaś rzeczywista zależność dla danego monitora wyraża się



Rys. 6. Zależność luminancji ekranu od poziomu wysterowania dla 10 różnych typów monitorów. Producenci monitorów nie są podani celowo. Zależności otrzymano dla średnich lub zalecanych przez producenta ustawień jaskrawości i kontrastu



Rys. 7. Zależność luminancji ekranu od poziomu wysterowania dla monitora wyższej klasy oraz wzorcowe charakterystyki zlinearyzowane percepcyjnie obliczone zgodnie z zależnością (15) dla $L_0 = 3, 10, 30$ cd/m^2

wzorem $L = L_{rzecz}(G)$, to wtedy dany poziom sterowania G należy zamienić na poziom G', spełniający zależność

$$L = L(G) = L_{rzecz}(G'), \tag{29}$$

stąd szukane odwzorowanie dane jest wzorem

$$G' = L^{-1}_{rzecz}[L(G)], \tag{30}$$

gdzie $L(G)$ dane jest przez (18) zaś przez L^{-1}_{rzecz} oznaczono funkcję odwrotną do $L_{rzecz}(G)$.

Zastosowanie dla rzeczywistych obrazów przekształcenia zgodnego z (30) pozowliło poprawić odbieraną subiektywnie jakość obrazu. Pomimo tego, że obliczenia luminancji są jedynie szacunkowe, a co za tym idzie linearyzacja percepcyjna może być dokonana jedynie w sposób przybliżony, jednoczesna obserwacja obrazu przed i po linearyzacji wykazała wyższość obrazu linearyzowanego percepcyjnie, zwłaszcza jeśli chodzi o widoczność szczegółów obrazu, które na oryginale znajdują się w ciemnych jego częściach. Zwróćmy uwagę raz jeszcze na to, że linearyzacja percepcyjna musi być dokonywana osobno dla każdego monitora, a nawet osobno dla poszczególnych ustawień jaskrawości i kontrastu.

PODSUMOWANIE

Własności percepcyjne wzroku ludzkiego mają zasadniczy wpływ na to, jak odbieramy obraz widoczny na ekranie monitora telewizyjnego. Przeprowadzone badania monitorów wykazały, że własności niektórych typów monitorów, zwłaszcza mniej znanych firm, znacznie odbiegają od optymalnych (linearyzowanych percepcyjnie) charakterystyk. Te wady można częściowo skorygować przez poddanie poziomu sterowania nieliniowemu przekształceniu danemu przez (30). Jak wykazały badania, prowadzi to do poprawy odbieranej subiektywnie jakości obrazu. Niestety, poprawa nie zawsze jest możliwa ze względu na ograniczony zakres zmian luminancji niektórych monitorów. Ponadto należy jeszcze raz podkreślić, że sama linearyzacja była dokonywana w sposób przybliżony, głównie ze względu na szacunkowe obliczenia luminancji monitorów. Poprawa dokładności pomiarów wymaga znajomości widma światła emitowanego z ekranu monitora.

BILBIOGRAFIA

1. H. Blume, S. Daily, E. Muka: *Presentation of medical images on CRT displays.* SPIE Proced., 1993, vol. 1897, str. 215–231
2. P.G.J. Barten: *Physical model for the contrast sensitivity of the human eye.* SPIE Proceed., 1992, vol. 1666, str. 57–72
3. S. Daly: *The visible differences predictor: an algorithm for the assessment of image fidelity.* SPIE Proceed., 1992, vol. 1666, str. 2–15
4. T. Ji, M.K. Sundareshan, H. Roehrig: *Adaptive Imaga Contrast Enhancement Based on Human Visual Properties.* IEEE Trans. Medical Imaging, 1994, vol. 13, nr 4, pp. 573–586

5. T. Ji, H. Roehrig, H. Blume, J. Guillen: *Optimizing the display function of display devices*. SPIE Proceed., 1992, vol. 1653, str. 126–139
6. S.J. Briggs: *Soft copy display of electro-optical imagery*. SPIE Proceed., 1987, vol. 762, str. 153–170
7. T. Ji, H. Roehrig, H. Blume, G. Seeley, M. Browne: *Physical and psychophisical evaluation of CRT noise performance*. SPIE Proceed., 1991, vol. 1444, str. 136–150
8. H. Roehrig, H. Blume, T. Ji, M.K. Sundareshan: *Noise of CRT dispaly systems*. SPIE Proceed., 1993 vol. 1897, str. 232–245
9. M. Ostrowski (koordynator): *Informacja obrazowa*. WNT, Warszawa 1992
10. H. Roehrig, T. Ji, M. Browne, W.J. Dallas, H. Blume: *Signal-to-noise ratio and maximum information content of images displayed by a CRT*. SPIE Proceed 1990, vol. 1232, str. 115–133

J. SIUZDAK, T. CZARNECKI

HUMAN VISION PROPERTIES AND THE CRT MONITORS CHARACTERISTICS

Summary

The dependence between human vision and characteristics of CRT monitors is investigated for monochrome images. The contrast sensitivity function models of Barten and Daly are described. The command level dependence of luminance is measured for various CRT monitors. The concept of perceptual linearization is also discussed.

Key words: image processing, human vision properties, CRT monitors, luminance.

SPIS TREŚCI

A. Mitas: Optymalizacja kompresji informacji w testowaniu układów ze ścieżką brzegową 307
D. Kania: Algorytm funkcjonalnej dekompozycji o swobodnym zestawie zmiennych 317
E. Hermanowicz, M. Rojewski, M. Blok: Projektowanie filtrów cyfrowych FIR na drodze naprzemiennych iteracji w dziedzinie czasu i w dziedzinie częstotliwości 325
K. Perlicki: Wpływ zjawiska samomodulacji fazy na transmisję danych wykorzystującą zjawisko konwersji modulacji częstotliwości na modulację amplitudy 335
W.J. Stepowicz, J. Zarębski: Nieizotermiczne parametry małosygnałowe elementów półprzewodnikowych 345
W.J. Stepowicz: Przebicie termiczne w złączowych elementach półprzewodnikowych 365
A. Materka, P. Pełczyński, M. Strzelecki: Analogowo-cyfrowy prototypowy system do identyfikacji parametrów modeli z wykorzystaniem sieci neuranowych 373
K. Wiatr: Analiza parametrów czasowych architektury potokowej specjalizowanych procesorów sprzętowych 391
H. Kunert: Wybrane metody identyfikacji i doboru nastaw stosowane w regulatorach z samonastrajaniem 413
J. Siuzdak, T. Czarnecki: Własności percepcyjne wzroku ludzkiego a charakterystyki monitorów 427

CONTENTS

A. Mitas: Optimisation of information compression in testing of circuits with boundary scan path 307
D. Kania: An algorithm of functional decomposition with free set variables coding 317
E. Hermanowicz, M. Rojewski, M. Blok: Digital FIR filter design by alternate iterations in the time and frequency domains 325
K. Perlicki: Inpact of self-phase modulation on FM-AM conversion supported transmission .. 335
W.J. Stepowicz, J. Zarębski: Nonisothermal small-signal parameters of semiconductor devices . 345
W.J. Stepowicz: Thermal breakdown in semiconductor junction devices 365
A. Materka, P. Pełczyński, M. Strzelecki: Artificial neural network mixed-signal prototype system for model parameter identification 373
K. Wiatr: Time parameters analysis for pipeline architekture of specialised hardware processor 391
H. Kunert: Selected identification and setting methods as used for self-tuning controllers 413
J. Siuzdak, T. Czarnecki: Human vision properties and the CRT monitors characteristics 427

INFORMACJE DLA AUTORÓW

Redakcja przyjmuje do publikowania prace oryginalne, przeglądowe i monograficzne wchodzące w zakres szeroko pojętej elektroniki. Ponieważ KWARTALNIK ELEKTRONIKI i TELEKOMUNIKACJI jest czasopismem Komitetu Elektroniki i Telekomunikacji Polskiej Akademii Nauk, w związku z tym na jego łamach znajdują się prace naukowe dotyczące podstaw teoretycznych i zastosowań z zakresu elektroniki, telekomunikacji, mikroelektroniki, optoelektroniki, radiotechniki i elektroniki medycznej.

Artykuły powinno charakteryzować oryginalne ujęcie zagadnienia, własna klasyfikacja, krytyczna ocena (teorii lub metod), omówienie aktualnego stanu, lub postępu danej gałęzi techniki oraz omówienie perspektyw rozwojowych.

Artykuły publikowane w innych czasopismach nie mogą być kierowane do druku w Kwartalniku Elektroniki i Telekomunikacji w drugiej kolejności zgłoszenia.

Objętość artykułu nie powinna przekraczać 30 stron po około 1800 znaków na stronie.

Wymagania podstawowe: Artykuły należy nadsyłać w maszynopisie pisanym jednostronnie lub na wyraźnym czarno-białym wydruku komputerowym, w 2 egzemplarzach, w języku polskim lub angielskim wybranym przez autora. Tekst artykułu musi być poprzedzony tytułem pracy, imieniem i nazwiskiem Autora wraz z podaniem miejsca jego pracy. Wszystkie strony muszą mieć numerację ciągłą.

Sposób pisania tekstu: Tekst powinien być pisany bez używania wyróżnień, a w szczególności nie dopuszcza się spacjowania, podkreślania i pisania tekstu dużymi literami z wyjątkiem wyrazów, które umownie pisze się dużymi literami (np. FORTRAN). Proponowane wyróżnienia Autor może zaznaczyć w maszynopisie zwykłym ołówkiem za pomocą przyjętych znaków adiustacyjnych) np. podkreślenie linią przerywaną oznacza spacjowanie (rozstrzelenie), podkreślenie linią ciągłą — pogrubienie, podkreślenie wężykiem — kursywa. Tekst powinien być napisany z podwójnym odstępem między wierszami, tytuły i podtytuły małymi literami. Marginesy z każdej strony powinny mieć około 35 mm. Przy podziale pracy na rozdziały i podrozdziały cyfrowe ich oznaczenia nie powinny być większe niż III stopnia (np. 4.1.1).

Sposób pisania tablic: Tablice powinny być pisane na oddzielnych stronach. Tytuły rubryk pionowych i poziomych powinny być napisane małymi literami z podwójnym odstępem między wierszami. Przypisy (notki) dotyczące tablic należy pisać bezpośrednio pod tablicą. Tablice należy numerować kolejno liczbami arabskimi, u góry każdej tablicy podać tytuł. Tablice umieścić na końcu maszynopisu. Przyjmowane są tablice algorytmów i programy na wydrukach komputerowych. W tym przypadku zachowany jest ich oryginalny układ.

Sposób pisania wzorów matematycznych: Rozmieszczenie znaków, cyfr, liter i odstępów powinno być zbliżone do rozmieszczenia elementów druku. Wskaźniki i wykładniki potęg powinny być napisane wyraźnie i być prawidłowo obniżone lub podwyższone w stosunku do linii wiersza podstawowego. Znaki nad literami i cyframi, całkami i in. symbolami (strzałki, linie, kropki, daszki) powinny być umieszczone dokładnie nad tymi elementami, do których się odnoszą. Numery wzorów cyframi arabskimi powinny być kolejne i umieszczone w nawiasach okrągłych z prawej strony. Nazwy jednostek, symbole literowe i graficzne powinny być zgodne z wytycznymi IEE (nternational Electronical Commision) oraz ISO (International Organization of Standarization).

Powołania: Powołania na publikacje powinny być umieszczone na ostatnich stronach tekstu pod tytułem „Bibliografia", opatrzone numeracją kolejną bez nawiasów. Numeracja ta powinna być zgodna z odnośnikami w treści artykułu. Przykłady opisu publikacji:

- periodycznej: F. Valdoni: A new milimetre wave satelite, E.T.T. 1990, vol. 2, nr 5, p. 553
- nieperiodycznej: K. Andersen: A resource allocation framework. XVI International Symposium, Stockholm (Sweden), may 1991, paper A 2.4
- książki: Y.P. Tvidis: Operation and modelling of the MOS transistors. Mc Graw-Hill, New York 1987, p. 141 – 148

Materiał ilustracyjny: Rysunki powinny być wykonane wyraźnie, na papierze gładkim, lub milimetrowym w formacie nie mniejszym niż 9×12 cm. Mogą być także w postaci wydruku komputerowego. Fotografie lub diapozytywy przyjmowane są raczej czarno-białe w formacie nie przekraczającym 10×15 cm. Na marginesie każdego rysunku i na odwrocie fotografii powinno być napisane ołówkiem imię i nazwisko Autora oraz skrót tytułu artykułu, do którego są przeznaczone. Spis podpisów pod rysunki i fotografie powinien być umieszczony na oddzielnej stronie.

Streszczenia: Do każdego artykułu musi być dołączone obszerne — do 60 wierszy — streszczenie z tytułem artykułu. Streszczenia (Summary) muszą być w języku polskim i angielskim. Streszczenie powinno wyjaśniać główny cel pracy, wskazywać korzyści i ograniczenia, możliwe zastosowania i zalecenia dla dalszego rozwoju danej gałęzi techniki. Pod streszczeniami powinny być podane słowa kluczowe.

Autorowi przysługuje bezpłatnie 20 odbitek artykułu. Dodatkowe egzemplarze odbitek, lub cały zeszyt Autor może zamówić u wydawcy na własny koszt.

Autora obowiązuje korekta autorska, którą powinien wykonać w ciągu 3 dni od daty otrzymania tekstu z Redakcji i zwrócić osobiście, lub listownie pod adresem Redakcji. Korekta powinna być naniesiona na przekazanych Autorowi szpaltach na marginesach ew. na osobnym arkuszu w przypadku uzupełnień tekstu większych niż dwa wiersze. W przypadku nie zwrócenia korekty w terminie, korektę przeprowadza Redakcja Techniczna Wydawcy.

Redakcja prosi Autorów o powiadomienie ją o zmianie miejsca pracy i adresu prywatnego.

INFORMATION FOR THE AUTHORS OF K.E.T.

The KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI (K.E.T.) — ELECTRONICS AND TELECOMMUNICATION QUARTERLY is a journal of the Communittee for Electronics and Telecommunications of the Polish Academy of Sciences and is a direct continuation for 44 years of the quarterly ROZPRAWY ELEKTROTECHNICZE, that has been published also under the auspices of the Polish Academy of Sciences.

Its pages contain scientific articles concerned with theory and applications of electronics, telecommunication, microelectronics, optoelectronics, radioengineerigna and medical electronics.

Accordingly, its pages contain scientific articles, that are original works both contributory and concerned with comparative analyses of methods or systems synthetic reviews of a given field, state-of-art discussions of a given field, and presentations of develeopments in a given technology.

Regular papers are accepted for publication in Polish or Englis, but title, the estensive summary and figure caption must be bilingual. Bilingual summary must shown the significance of the results shown in the paper, emphasizing limitations and advantages, possible applications and recommendations for further works.

K.E.T. is meant for specialists and students with above subjects.

FORMAL PREPARATION OF MANUSCRIPT

Texts

Texts to be submitted to K.E.T. may be typed on one side of the paper only with double spacing between the lines and a 4 cm mrgin or uswing of wordprocessing with text formatter5s based on MS DOS, APPLE/MAC such as Microsoft word. In any case the text formatter must be specified on the diskette. Each text should be preceded by: a sheet with paper title, name (in full) and surname of the Author/s, their Affilltions and relevant addresses.

All text pages should be numbered. Regular papers should have a lenght of no more than 30 pages.

Printing types and symbols

For letteral symbols, units and graphical symbols, the recommendations of IEE (International Electrotecynical Commission) and ISO (International Organisation of Standarisation) must be f0llowed.

- In case of manuscripts conventionally typed (i.e. without wordprocessing), letters to be printed in italic should be underlined; letters to be printed in bold should be double interlined.
- Formulae and footnotes should be numbered in accordance with the quotation order followed in the text.
- The serial number of each formula should be written at right side, between round brackets.
- The serial number of footnotesw should be written between round brackets and upper.
- Mathematical details, anciliary to the main discussion of the paper may be included in one or more appendixes, that are printed after conclusions and before references.

In manuscript containing lot of symbols, a glossary may be included after introduction.

References

References are usually gathered at the end of the text and numbered according to the order of quotation in the text. The serial numbers should be written without square brackets (i.e. 1., 2...)

Illustrations

The originals of illustrations may be either drawings or photographs (black or colour). Each copy of paper must include two complete set of illustrations.

All rights reserved. No part of the publication may be reproduced, stored in a retrival system or transmitted, in any form or by any means: electronic, photocopying, recording or othervise, without the prior permission of the copyright owner — Relaction K.E.T.

Subscription for external subscribers:

The promotional subscription price in 1997 is S 90 including postage for institutions. Subscriptions should be sent to the publisher, Polish Scientific Publishers PWN Ltd, Journal Division, Miodowa 10, 00-250 Warsaw, POLAND, fax (48) (22) 826 09 50, (48) (22) 826 71 63, with a copy by fast to Electronics and Telecommunications Quarterly 00-665 Warsaw, ul. Nowowiejska 15/19, p. 470. Subscription is occupated after showing a checque or transfer documents. Our bank account is as follows: Bank account: PBK VIII O/Warszawa nr 11101037-1052-2700-1-43. At subscripber's request this journal will be air mailed at additional postage to 50% dross price to European countries and 65% overseas.

Subscription orders available through the local press distribution or through the Foreign Trade Enterprise ARS Polona, 00-068 Warszawa, Krakowskie Przedmieście 7, Poland.

Ruch S.A. fulfils foreign customers orders, starting from any issue in the calendar year: tel. (48) (22) 620 120 39; fax (48) (22) 620 17 62.

Cena 9 zł

Kwartalnik Elektroniki i Telekomunikacji

WARUNKI PRENUMERATY

Wpłaty na prenumeratę przyjmują jednostki kolportażowe RUCH S.A., i wszystkie urzędy pocztowe na terenie całego kraju, właściwe dla miejsca zamieszkania lub siedziby prenumeratora oraz doręczyciele w miejscowościach, gdzie dostęp do urzędu jest utrudniony.
Od osób lub instytucji, zamieszkałych lub mieszczących się w miejscowościach, w których nie ma jednostek kolportażowych RUCH ,wpłaty należy wnosić do RUCH S.A. Oddział Krajowej Dystrybucji Prasy, 00–958 Warszawa, ul. Towarowa 28. Konto: PBK S.A. XIII Oddział Warszawa nr 11101053-16551-2700-1-67. RUCH S.A. i POCZTA POLSKA zapewniają dostawę pod wskazanym adresem pocztą zwykłą w ramach opłaconej prenumeraty.

Terminy przyjmowania wpłat na prenumeratę krajową: RUCH S.A. do 5 XII na I kw., do 5 III na II k., do 5 VI na III kw. i do 5 IX na IV kw. POCZTA POLSKA do 25 XI na I kw., do 25 na II kw., do 25 V na III kw. i do 25 na IV kw.

Na zagranicę prenumeratę przyjmuje RUCH S.A. Oddział Krajowej Dystrybucji Prasy, 00–958 Warszawa, u. Towarowa 28. Konto: PBK S.A. XII Oddział Warszawa nr 11101053-16551-2700-1-67. Dostawa odbywa się pocztą zwykłą w ramach opłaconej prenumeraty, z wyjątkiem zlecenia dostawy pocztą lotniczą, której koszt w pełni pokrywa zleceniodawca.
Prenumerata ze zleceniem dostawy za granicę jest o 100% wyższa od krajowej.
Terminy wpłat: do 20 XI na I kw., do 20 II na II kw., do 20 V na III kw. i do 20 VIII na IV kw.
RUCH S.A. fulfils foreign customers' orders, starting from any issue in the calendar year: tel. (48)(22) 620 10 19; fax (48)(22) 620 10 39.

Prenumeratę przyjmuje Zakład Kolportażu Wydawnictwa SIGMA-NOT, 00-716 Warszawa, ul. Bartycka 20, skr. poczt. 1004, termin wpłat do 5 XII roku poprzedzającego.
Konto PBK S.A. III Oddział Warszawa nr 11101024–1573–2720–3–28
Prenumerata ze zleceniem wysyłki za granicę jest dwukrotnie wyższa od ceny krajowej. Zlecający powinien podać dokładny adres odbiorcy. Dodatkowe informacje można otrzymać pod nr telefonu 49-30-86 lub 40-00-21 wew. 249, 295, 299.

Bieżące numery można nabyć w **Księgarni Wydawnictwa Naukowego PWN**, ul. Miodowa 10, 00-251 Warszawa, tel. (22) 695 40 26. Również można je nabyć, a także zamówić (przesyłka za zaliczeniem pocztowym) w Księgarni Ośrodka Rozpowszechniania Wydawnictw Naukowych PAN, ul. Twarda 51/55, 00-818 Warszawa, tel. (22) 697 88 35.

All journals edited by PWN are available through:

Foreign Trade Enterprise
ARS POLONA
Krakowskie Przedmieście 7,
00–068 Warszawa, Poland
fax (48) (22) 826 86 73

or

Polish Scientific Publishers PWN Ltd
10 Miodowa Str.
00–251 Warszawa, Poland
fax (48) (22) 826 09 50
fax (48) (22) 695 42 88

W roku 1997	
prenumerata roczna w kraju	34,00 zł
prenumerata roczna	90 $ USA