POLSKA AKADEMIA NAUK
KOMITET ELEKTRONIKI I TELEKOMUNIKACJI

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI

ELECTRONICS AND TELECOMMUNICATIONS QUARTERLY

TOM 46 — ZESZYT 1

WYDAWNICTWO NAUKOWE PWN

WARSZAWA 2000

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 (022) 660 77 37

Telefony domowe: Redaktora Naczelnego: 812 17 65
Zast. Red. Naczelnego: 826 83 41
Sekretarza Odpowiedzialnego: 633 92 52

WYDAWNICTWO NAUKOWE PWN SA
Warszawa, ul. Miodowa 10

Ark. wyd. 8,5. Ark. druk. 7,75	Podpisano do druku w czerwcu 2000 r.
Papier offset. kl. III 70 g. B-1	Druk ukończono w lipcu 2000 r.

SKŁAD I ŁAMANIE: Agnieszka Chmielewska, Warszawa, ul. Korytnicka 28
DRUK I OPRAWA: Drukarnia Braci Grodzickich, Piaseczno, ul. Geodetów 47a

Szanowni Autorzy

„Kwartalnik Elektroniki i Telekomunikacji" — Electronics and Telecommunications Quartery jest kontynuatorem tradycji powstałego 46 lat temu kwartalnika p.t. „Rozprawy Elektrotechinczne".

Kwartalnik jest czasopismem Komitetu Elektroniki i Telekomunikacji Polskiej Akademii Nauk. Wydawany jest przez Wydawnictwo Naukowe PWN SA w Warszawie. Kwartalinik 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, otoelektroniki, 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 rozowojowych. Sposób pisania matematycznej części artykułów zgodny jest z wytycznymi IEC (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 indywidualnych 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, 2000, 46, z. 1, ss. 5–19

Korekcja wielowymiarowych systemów stacjonarnych przyczynowych*

GRZEGORZ CIESIELSKI

Instytut Elektrotechniki Teoretycznej, Metrologii i Materiałoznawstwa, Politechniki Łódzkiej

Otrzymano 1999.05.10
Autoryzowano 2000.02.24

Artykuł dotyczy korekcji wielowymiarowych systemów stacjonarnych (niekoniecznie liniowych). Rozważane są tu systemy, które dają się opisać za pomocą odwzorowań nazywanych operatorami.

Autor artykułu od szeregu lat pracuje nad jednolitą ogólną teorią operatorową modelowania i korekcji wspomianych tu systemów. Prace te doprowadziły go do uniwersalnego rozwiązania problemu doboru korektora w tych systemach. Okazuje się, że problem korekcji daje się sprowadzić do zadania modelowania operatora wzorcowego, który pojawia się w zadaniu korekcji. Rozwiązaniem pojawiającego się tu problemu jest zatem twierdzenie rozwiązujące zadanie modelowania wielowymiarowych systemów stacjonarnych ciągłych, o których mowa jest w rozprawie habilitacyjnej autora [8], s. 51 oraz w artykułach [10] i [14].

Uzyskane wyniki, podobnie jak w rozprawie habilitacyjnej [8], s. 51 oraz w artykułach o modelowaniu [10] i [14], są na tyle ogólne, że nie są bezpośrednio związane z żadnym ze znanych operatorów (np. Volterry, czy Wienera). Wytyczają one jedynie ścisłe granice, których nie można przekroczyć w każdym procesie korekcji wspomianych wcześniej systemów.

Streszczony tu artykuł jest naturalną kontynuacją prac autora prezentowanych w latach: 1992 i 1995, w Kwartalniku Elektroniki i Telekomunikcaji oraz w latach 1992, 1994 i 1996–1998, na sympozjach dotyczących modelowania i symulacji systemów.

Słowa kluczowe: ogólna funkcjonalna teoria systemów, korekcja, modelowanie, identyfikacja strukturalna, identyfikacja parametryczna, systemy wielowymiarowe, systemy nieliniowe, systemy stacjonarne, systemy przyczynowe, sploty, przestrzeń stacjonarna, przestrzeń splotowa, przestrzeń przyczynowa.

* Pracę wykonano w ramach projektu badawczego nr 8 T10C 031 09 finansowanego przez Komitet Badań Naukowych w latach 1995–1997.

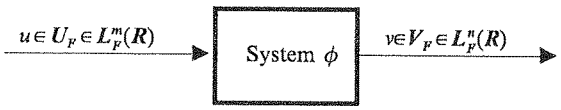
1. WSTĘP

Przedmiotem artykułu jest korekcja wielowymiarowych systemów stacjonarnych (niekoniecznie liniowych). Przypomnijmy, że zadanie koleracji dowolnego systemu, podobnie jak zadanie modelowania systemu, ma na celu wyznaczenie matematycznego modelu korektora tego systemu. Zadanie to składa się z identyfikacji strukturalnej oraz identyfikacji parametrycznej modelu (J. Jaworski, R. Morawski, R. Olędzki [25], s. 65) korektora. W artykule rozważane są systemy, które dają się opisać za pomocą odwzorowań, nazwanych operatorami. Operatory te są odwzorowaniami ciągłymi, dla których jedną z form reprezentacji jest układ równań różniczkowych, być może nieliniowych. Korygowane systemy rozważane w artykule są wielowymiarowe, czyli mają dowolną skończoną liczbę wejść i wyjść. Sygnały na poszczególnych wejściach mogą przyjmować wartości rzeczywiste lub zespolone. Ponadto przyjęto, że zbiór sygnałów wejściowych ma strukturę przestrzeni Hilberta.

W literaturze naukowej nie ma publikacji omawiających problem korekcji nieliniowych systemów stacjonarnych ciągłych, które są wielowymiarowe. Istniejące w literaturze publikacje ograniczają się do zadania korekcji systemów nieliniowych z jednym wejściem i jednym wyjściem. Na szczególną uwagę zasługuje tutaj monografia K. A. Pupkowa, W. I. Kapalina i A. S Juszczenki ([30]), która jest matematycznie dość ścisła. Pewnym problemem jest jednak korygowanie wielowymiarowych systemów stacjonarnych ciągłych. Autor artykułu od szeregu lat pracuje nad jednolitą ogólną teorią operatorową modelowania i korekcji wspomianych tu systemów ([2]–[21]). Prace te doprowadziły go do uniwersalnego rozwiązania problemu doboru korektora w tych systemach. Okazuje się, że tak sformułowany problem korekcji daje się sprowadzić do zadania modelowania operatora wzorcowego, który pojawia się w zadaniu korekcji systemu. Rozwiązaniem pojawiającego się problemu jest zatem twierdzenie rozwiązujące zadanie modelowania wielowymiarowych systemów stacjonarnych ciągłych, o których mowa jest w rozprawie habilitacyjnej autora artykułu [8], s. 51 oraz w artykułach o modelowaniu [10] i [14]. Uzyskane wyniki, podobnie jak w rozprawie habilitacyjnej [8], s. 51 oraz w artykułach o modelowaniu [10] i [14], są na tyle ogólne, że nie są bezpośrednio związane z żadnym ze znanych operatorów (np. Volterry, czy Wienera). Wytyczają one jedynie ścisłe granice, których nie można przekroczyć w każdym procesie korekcji wspomianych wcześniej systemów.

2. PRELIMINARIA MATEMATYCZNE

Dowolny system moża traktować jako odwzorowanie zbioru U sygnałów wejściowych zwanych *wymuszeniami* w zbiór V sygnałów wyjściowych zwanych *odpowiedziami*. Będziemy dalej zakładać, że oba te zbiory wyposażone są w strukturę



Rys. 2.1. Podstawowy model systemu.

przestrzeni liniowej, a zatem określona jest w nich operacja dodawania sygnałów i mnożenia sygnału przez liczbę z ciała liczbowego F[1].

Niech R[2] będzie zbiorem chwil czasowych, dla których określone są sygnały z U oraz V. Załóżmy ponadto, że $m \in N$[3] jest liczbą wejść, a $n \in N$ jest liczbą wyjść systemu. Wartości sygnałów wejściowych należą zatem do F^m[4], a wartości sygnałów wyjściowych do F^n. Wówczas przestrzeń sygnałów wejściowych U_F[5] jest podprzestrzenią liniową przestrzeni $L_F^m(R)$[6]. Podobnie, przestrzeń sygnałów wyjściowych V_F jest podprzestrzenią liniową przestrzeni $L_F^n(R)$, co zostało pokazane na rys. 2.1. Ponieważ U oraz V są przestrzeniami funkcyjnymi, których elementy określone są na tej samej dziedzinie R, a dowolne odwzorowanie ϕ takich przestrzeni nazywane jest operatorem, to każdy system można opisać za pomocą operatora, dla którego jedną z możliwych form reprezentacji jest układ równań różniczkowych. By uprościć dalszą dyskusję, pojęcia: *system opisany za pomocą operatora* ϕ i *operatora* ϕ, będą więc traktowane jako równoważne.

W modelowaniu systemów wykorzystuje się trzy elementarne operatory: przesunięcie, obrót i splotowe przesunięcie funkcyjne (G. Ciesielski [5]-[12] oraz [14]-21]).

Przesunięciem funkcyjnym o $\sigma \in R^k$ nazywamy operator oznaczony przez $\nabla_\sigma: L_F^m(R^k) \mapsto L_F^m(R^k)$, który $\forall u \in L_F^m(R^k)$ i $\forall t \in R^k$ spełnia warunek:

$$[\nabla_\sigma(u)](t) := u(t+\sigma)^7.$$

Funkcję $\nabla_\sigma(u)$ nazywamy *przesunięciem funkcji* u o σ.

Obrotem funkcyjnym nazywamy operator, który oznaczamy przez $\nabla^*: L_F^m(R^k) \to L_F^m(R^k)$ taki, że $\forall u \in L_F^m(R^k)$ i $\forall t \in R^k$

$$[\nabla^*(u)](t) := u(-t).$$

Funkcję $\nabla^*(u)$ nazywamy *obrotem funkcji* u.

[1] W całym artykule przez F oznaczamy ciało liczbowe. W szczególności F oznacza ciało liczb rzeczywistych R lub ciało liczb zespolonych C.

[2] Zbiór X wyposażony w dowolną strukturę algebraiczną i/lub topologiczną będziemy oznaczać literą X, a zatem przez R oznaczamy zbiór liczb rzeczywistych, natomiast przez R ciało liczb rzeczywistych.

[3] Przez N oznaczamy zbiór liczb naturalnych.

[4] Dla dowolnego zbioru X przez X^n oznaczamy n-krotny iloczyn kartezjański (produkt) zbioru X.

[5] Przez X_F oznaczamy przestrzeń liniową nad ciałem skalarów F. Jeżeli F wynika z kontekstu, to jest ono pomijane w jej oznaczeniu.

[6] Przez $L_F^n(X)$ oznaczamy przestrzeń odwzoworwań określonych na X o wartściach z F^n. Oznaczenia $L_F^1(X)$ oraz $L_F(X)$ traktujemy jako równoważne, a gdy zbiór X wynika z kontekstu, to piszemy w skrócie L_F^n lub L_F.

[7] Przez $:=$ oznaczamy równość definicyjną.

Splotowym przesunięciem funkcyjnym o $\sigma \in \boldsymbol{R}^k$ nazywamy operator, który oznaczany jest przez $\nabla_\sigma^*: \boldsymbol{L}_F^m(\boldsymbol{R}^k) \mapsto \boldsymbol{L}_F^m(\boldsymbol{R}^k)$ oraz $\forall u \in \boldsymbol{L}_F^m(\boldsymbol{R}^k)$ i $\forall t \in \boldsymbol{R}^k$

$$[\nabla_\sigma^*(u)](t) := u(\sigma - t).$$

Funkcję $\nabla_\sigma^*(u)$ nazywać będziemy *przesunięciem splotowym funkcji* u o σ.

Zauważmy tu, że przesunięcie ∇_σ i ∇_σ^* są operatorami liniowymi na przestrzeni $\boldsymbol{L}_F^m(\boldsymbol{R}^k)$ i odwzorowują przestrzeń $\boldsymbol{L}_F^m(\boldsymbol{R}^k)$ w nią samą. Ponadto

$$\nabla_\sigma \circ \nabla_{-\sigma} = \mathrm{id}\ ^{8}, \qquad \nabla_\sigma^* \circ \nabla_\sigma^* = \mathrm{id}, \qquad \nabla_\sigma^* = \nabla^* \circ \nabla_\sigma = \nabla_{-\sigma} \circ \nabla^* \quad \text{oraz} \quad \nabla_0^* = \nabla^*.$$

Zauważmy również, że przesunięcie funkcji $\nabla_\sigma(u)$ jest funkcją u przesuniętą w lewo o σ, obrót funkcji $\nabla^*(u)$ jest funkcją u obróconą o π wokół osi rzędnych, a splotowe przesunięcie funkcyjne ∇_σ^* jest złożeniem operatorów ∇^* z ∇_σ, przy czym złożenia tego nie można wykonać w dowolnej kolejności (G. Ciesielski [8], s. 17, oraz [5] i [6].

Przejdźmy teraz do definicji podstawowych właściwości operatorów opisujących systemy.

Operator ϕ na $\boldsymbol{U}_F \subset \boldsymbol{L}_F^m(\boldsymbol{R})$ nazywamy *stacjonarnym*, jeżeli

$$\forall u \in \boldsymbol{U} \forall \sigma \in \boldsymbol{R}\big((\nabla_\sigma(u) \in \boldsymbol{U}) \wedge (\nabla_\sigma(\phi(u)) = \phi(\nabla_\sigma(u)))\big).$$

Innymi słowy, operator jest stacjonarny, gdy odpowiedź operatora na wymuszenie u przesunięte w czasie o σ jest równa przesuniętej w czasie o σ odpowiedzi operatora na wymuszenie u. Zauważmy, że przestrzeń liniowa $\boldsymbol{U}$ musi być zbiorem zamkniętym ze względu na przesunięcia funkcyjne, a zatem musi zawierać wszystkie przesunięcia funkcyjne swoich elementów. Przestrzeń o takiej właściwości będziemy nazywać *przestrzenią stacjonarną*. Kolejną ważną właściwością operatora jest przyczynowość.

Operator ϕ na $\boldsymbol{U}_F \subset \boldsymbol{L}_F^m(\boldsymbol{R})$ nazywamy *przyczynowym*, jeżeli

$$\forall u, v \in \boldsymbol{U} \forall t \in \boldsymbol{R}\big(\forall \tau \in \boldsymbol{R}\big((\tau \leqslant t) \Rightarrow (u(\tau) = v(\tau))\big) \Rightarrow \big([\phi(u)](t) = [\phi(v)](t)\big)\big).$$

Wynika stąd, że operator jest przyczynowy, gdy wartość sygnału wyjściowego w chwili t zależy jedynie od wartości sygnału wejściowego do chwili t.

Wprowadźmy teraz definicje pewnych przestrzeni, które będziemy wykorzystywać dalej. Przestrzenie te będą stanowiły podstawę opisu przestrzeni wymuszeń i odpowiedzi rozważanych systemów.

Załóżmy, że μ jest miarą na σ-ciele w X.

Jeżeli $1 \leqslant p < \infty$, to *przestrzenią* $\boldsymbol{L}_{Fp\mu}^m(X)$ nazywamy zbiór klas równoważności ze względu na relację $\mu - \forall\exists$ [9], funkcji mierzalnych (μ-mierzalnych) $u: X \mapsto \boldsymbol{F}^m$ takich, że

$$\int_X |u|^p d\mu < \infty.$$

[8] Przez id oznaczamy odwzorowanie identycznościowe (identyczność).

[9] Przez $\mu - \forall\exists$ oznaczamy zwrot „μ-prawie wszędzie" lub „dla μ-prawie każdego".

Normą w $L^m_{Fp\mu}(X)$ nazywamy funkcję, którą $\forall u \in L^m_{Fp\mu}(X)$ oznaczamy przez

$$\| u \|_p := \left(\int_X |u|^p d\mu \right)^{1/p}.$$

Funkcje należące do $L^m_{Fl\mu}(X)$ nazywamy *całkowalnymi* (*w sensie Lebesgue'a*).

Przestrzenią $L^m_{F\infty\mu}(X)$ nazywamy zbiór klas równoważności ze względu na relację $\mu - \forall\exists$, funkcji mierzalnych (μ-mierzalnych) $u: X \mapsto F^m$, dla których

$$\operatorname*{supess}_{x \in X} |u(x)| < \infty,$$

gdzie

$$\operatorname*{supess}_{x \in X} |u(x)| := \inf\{M \in \overline{R}_+ : \mu - \forall\exists x \in X(|u(x)| \leqslant M\}\ ^{10},$$

nazywamy *istotnym kresem górnym* funkcji $|u|$. *Normą w* $L^m_{F\infty\mu}(X)$ nazywamy funkcję, którą $\forall u \in L^m_{F\infty\mu}(X)$ oznaczamy przez

$$\|u\|_\infty := \operatorname*{supess}_{x \in X} |u(x)|.$$

Funkcje należące do przestrzeni $L^m_{F\infty\mu}(X)$ nazywamy *istotnie* (μ-*istotnie*) *ograniczonymi* (J. Musielak [28], s. 53 i 59).

Przestrzenią $L^m_{F2\mu}(X)$ nazywamy zbiór klas równoważności ze względu na relację $\mu - \forall\exists$, funkcji mierzalnych (μ-mierzalnych) $u: X \mapsto F^m$, dla której *iloczynem skalarnym* nazywamy funkcję, którą $\forall u, v \in L^m_{F2\mu}(X)$ oznaczamy przez

$$\langle u, v \rangle := \int_X v^* u \, d\mu\ ^{11}.$$

Normą wyznaczoną przez iloczyn skalarny w $L^m_{F2\mu}(X)$ nazywamy funkcję, którą $\forall u \in L^m_{F2\mu}(X)$ oznaczamy przez

$$\|u\|_2 := (\langle u, u \rangle)^{1/2}.$$

Dla dowolnego układu $u := (u_i \in L^m_{F2\mu}(X))_{i=1}^k$, *macierz Grama* oznaczaną przez

$$G(u) := (\langle u_i, u_j \rangle)_{i,j=1}^k.$$

Załóżmy dalej, że $1 \leqslant p \leqslant \infty$.

Przez $C^m_{Fp\mu}(X)$ będziemy oznaczać podprzestrzeń funkcji ciągłych z przestrzeni $L^m_{Fp\mu}(X)$.

Jeżeli zbiór X wynika z kontekstu, to będziemy pisać w skrócie $L^m_{Fp\mu}$ lub $C^m_{Fp\mu}$. Oznaczenia $L^1_{Fp\mu}$ i $C^1_{Fp\mu}$ oraz odpowiednio $L_{Fp\mu}$ i $C_{Fp\mu}$ będziemy traktować jako

[10] Przez $\overline{R}_+$ oznaczamy zbiór liczb rzeczywistych nieujemnych uzupełnionych o liczbę ∞.

[11] Dla dowolnej macierzy A przez A^* oznaczamy jej transpozycję ze sprzężeniem, a więc $A^* := \bar{A}^T$.

równoważne. W przypadku gdy miara μ jest miarą Lebesgue'a na R^k, którą będziemy oznaczać przez λ_k, w skrócie λ, to oznaczenia $L^m_{Fp\lambda}$ i $C^m_{Fp\lambda}$ oraz odpowiednio L^m_{Fp} i C^m_{Fp}, jak tez $\lambda-\forall\exists$ oraz $\forall\exists$, są równoważne.

Przejdźmy teraz do definicji uogólnionego splotu funkcji.

Niech λ_k będzie miarą Lebesgue'a na R^k. *Splotem funkcji* mierzalnych $u, v: R^k \mapsto F^m$ nazywać będziemy funkcję oznaczaną przez $u^*v: R^k \mapsto F$, która o ile istnieje, spełnia warunek:

$$\lambda_k - \forall\exists t \in R^k \left([u^*v](t) := \int_{R^k} u^T(t-\tau)v(\tau)d\tau \right).$$

Zwykle splot definiuje się dla funkcji o wartościach z F, a nie jak tutaj z F^m. Takie rozszerzenie pojęcia splotu wynika z jego związku z iloczynem skalarnym (G. Ciesielski [8], s. 33). Zauważmy również, że całka Lebesgue'a może przyjmować wartości nieskończone, które nie należą do ciała F[12], podczas gdy wartości splotu funkcji należą do F. Z warunku definicyjnego wynika, że splot funkcji istnije, gdy zbiór tych $t \in R^k$, dla których całka przyjmuje wartości spoza F jest miary zero. Główne właściwści tych splotów omówiono w [8], s. 29, oraz w [7] i [11].

Przekonamy się dalej, że przestrzenie stacjonarne, splotowe i przyczynowe odgrywają bardzo ważną rolę w modelowaniu systemów stacjonarnych.

Załóżmy, że μ jest miarą na R^k oraz $1 \leqslant p \leqslant \infty$.

Przestrzenią stacjonarną oznaczaną przez $L^m_{Fp\mu\nabla}(R^k)$, nazywamy podzbiór klas równoważności z przestrzeni $L^m_{Fp\mu}(R^k)$ ze względu na relację $\mu-\forall\exists$ o postaci:

$$L^m_{Fp\mu\nabla}(R^k) := \{u \in L^m_{Fp\mu}(R^k) \forall t \in R^k (\nabla_t(u) \in L^m_{Fp\mu}(R^k))\}.$$

Przestrzenią splotową oznaczaną przez $L^m_{Fp\mu*}(R^k)$, nazywamy podzbiór klas równoważności z przestrzeni $L^m_{Fp\mu}(R^k)$ ze względu na relację $\mu-\forall\exists$ o postaci:

$$L^m_{Fp\mu*}(R^k) := \{u \in L^m_{Fp\mu}(R^k) \forall t \in R^k (\nabla^*_t(u) \in L^m_{Fp\mu}(R^k))\}.$$

Przestrzenią przyczynową oznaczaną przez $L^m_{Fp\mu+}(R^k)$, nazywamy podzbiór klas równoważności z przestrzeni $L^m_{Fp\mu}(R^k)$ ze względu na relację $\mu-\forall\exists$ o postaci:

$$L^m_{Fp\mu+}(R^k) := \{u_{-1}uu \in L^m_{Fp\mu}(R^k)\}$$ [13].

[12] Załóżmy dla uproszczenia, że $F=R$. Mówimy, że *istnieje całka Lebesgue'a* funkcji μ-mierzalnej $f: X \mapsto R$, jeżeli

$$\int_X f d\mu \in \overline{R},$$

gdzie przez $\overline{R}$ oznaczamy zbiór liczb rzeczywistych R uzupełniony o liczby $-\infty$ i $+\infty$. Z istnienia całki Lebesgue'a nie wynika zatem całkowalność funkcji f.

[13] Przez u_{-1} oznaczamy *skok jednostkowy*, który $\forall t := (t_i \in R)^k_1$ zdefiniowany jest przez wzór:

$$u_{-1}(t) := \begin{cases} 1, & \forall i \in \{1, \ldots, k\}(t_i \geqslant 0) \\ 0, & \exists i \in \{1, \ldots, k\}(t_i < 0) \end{cases}.$$

Jeżeli $u: X \mapsto F$ oraz $v: X \to F^n$, to $\forall x \in X$ *iloczyn funkcji* u i v oznaczamy przez uv, zdefiniowany jest za pomocą wzoru:

$$[uv](x) := u(x)v(x).$$

Właściwości tych przestrzeni są omówione w [8], s. 32, oraz w [7] i [11].

Przez $C^m_{Fp\mu\nabla}(X)$, $C^m_{Fp\mu *}(X)$ i $C^m_{Fp\mu +}(X)$ będziemy oznaczać podprzestrzeń funkcji ciągłych odpowiednio z przestrzeni $L^m_{Fp\mu\nabla}(X)$, $L^m_{Fp\mu *}(X)$ i $L^m_{Fp\mu +}(X)$.

Jeżeli zbiór R^k wynika z kontekstu, to będziemy pisać w skrócie $L^m_{Fp\mu\nabla}$, $L^m_{Fp\mu *}$, $L^m_{Fp\mu +}$, $C^m_{Fp\mu\nabla}$, $C^m_{Fp\mu *}$ oraz $C^m_{Fp\mu +}$. Oznaczenia $L^1_{Fp\mu\nabla}$, $L^1_{Fp\mu *}$, $L^1_{Fp\mu +}$, $C^1_{Fp\mu\nabla}$, $C^1_{Fp\mu *}$, $C^1_{Fp\mu +}$ oraz odpowiednio $L_{Fp\mu\nabla}$, $L_{Fp\mu *}$, $L_{Fp\mu +}$, $C_{Fp\mu\nabla}$, $C_{Fp\mu *}$ i $C_{Fp\mu +}$ będziemy traktować jako równoważne. W przypadku gdy miara na R^k jest miarą Lebesgue'a, to jej symbol będzie pomijany w oznaczeniach zdefiniowanych przestrzeni.

Zauważmy, że zgodnie z definicją stacjonarności operatora, sformułowaną w tej części artykułu, przestrzeń wymuszeń operatora stacjonarnego musi zawierać wszystkie przesunięcia funkcyjne swoich elementów. Przestrzeń stacjonarna $L^m_{Fp\mu\nabla}(R)$ ma tą właściwość, co pozwala nam definiować na niej operatory stacjonarne. Zauważmy tu również, że podobnie przestrzeń splotowa $L^m_{Fp\mu *}(R)$ zawiera wszystkie splotowe przesunięcia funkcyjne swoich elementów. Z kolei elementy składające się na przestrzeń przyczynową $L^m_{Fp\mu +}(R)$ są funkcjami, które przyjmują wartość 0 dla ujemnych wartości zmiennej niezależnej. *Sygnały* należące do tej przestrzeni nazywane są *przyczynowymi* (A. Wojnar [39], s. 48).

Można wykazać (G. Ciesielski [8], s. 18, oraz [5] i [6]), że dowolny operator ϕ na $L^m_{F2k\nabla}(R)$ jest stacjonarny, wtedy i tylko wtedy, gdy:

$$\phi = \varphi \circ \nabla .,$$

gdzie przekształcenie

$$\varphi := [\phi(\cdot)](0).$$

Z kolei, $\forall t \in R$ obrazem przestrzeni $L^m_{F2\kappa\nabla}(R)$ w odwzorowaniu ∇_t jest ta sama przestrzeń $L^m_{F2\kappa\nabla}(R)$, a zatem obrazem produktu $L^m_{F2\kappa\nabla}(R) \times R$ w odwzorowaniu ∇ jest również przestrzeń $L^m_{F2\kappa\nabla}(R)$. Jest to bardzo ważny wynik, który pozwala ograniczyć analizę struktury dowolnego stacjonarnego operatora ϕ na $L^m_{F2\kappa\nabla}(R)$, do analizy struktury odpowiadającego mu przekształcenia φ również na $L^m_{F2\kappa\nabla}(R)$ (G. Ciesielski [8], s. 18).

Załóżmy, że

$$L^m_{F2\kappa\nabla}(R) = L^m_{F2\kappa *}(R) = L^m_{F2\kappa}(R)$$

oraz $\varphi \in C^n_{F\infty\mu}(L^n_{F2\kappa}(R))$ jest przekształceniem ograniczonym, dla którego złożenie $\varphi \circ \nabla .$ jest operatorem stacjonarnym przyczynowym. Wynika wówczas, że każdy operator $\varphi \circ \nabla .$ można dowolnie dokładnie przybliżyć za pomocą liniowej kombinacji operatorów o postaci:

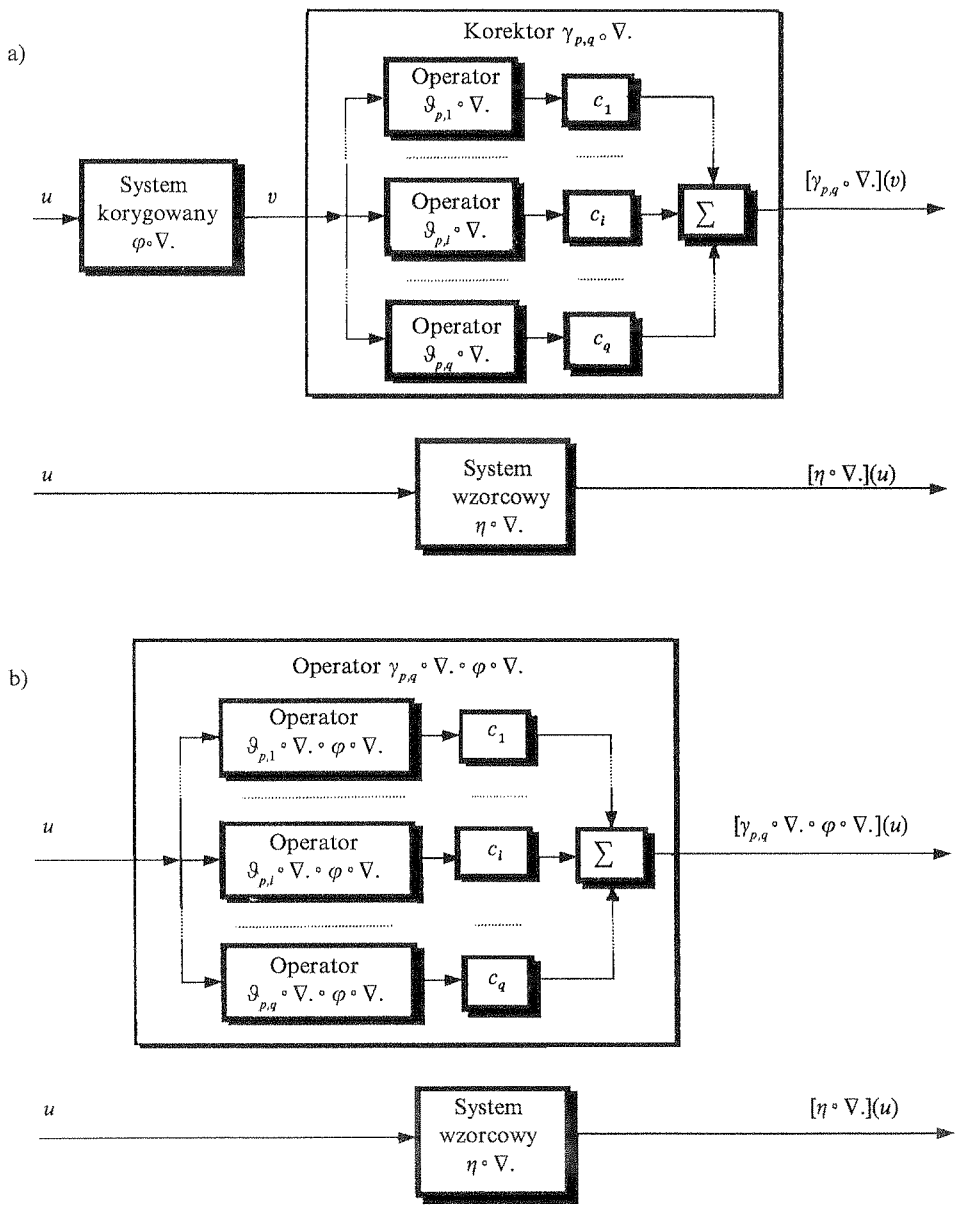
$$[\theta_{k[l]}(u)](t) := ([\theta_{k,f}(u)](t))_1^t := \vartheta_{k[l]}(\nabla_t(u)) = \chi_{[l]}([\overline{wv}_{[k]} * u](t)),$$

gdzie $\chi_{[l]}$ jest układem liniowo niezależnym w $L^n_{F2v_t}(F^k)$, a $v_{[k]}$ jest układem ortonormalnym w $L^m_{F2\kappa +}(R)$ (G. Ciesielski [8], s. 51, oraz [10] i [14]).

3. ZADANIE KOREKCJI SYSTEMÓW

Zadanie korekcji systemu zwykle jest formułowane w następujący sposób. Załóżmy, że κ jest taką miarą na $\boldsymbol{R}$, że

$$L_{F2\kappa\nabla} = L_{F2\kappa}(\boldsymbol{R})$$



Rys. 3.1. Niekonstruktywna (a) i konstruktywna (b) postać zadania korekcji systemu.

oraz $\varphi \in \boldsymbol{L}^m_{F2\mu}(\boldsymbol{L}^l_{F2\kappa}(\boldsymbol{R}))$ jest przekształceniem, dla którego złożenie $\vartheta \circ \nabla$. jest operatorem opisującym korygowany system. Załóżmy też, że

$$\vartheta_{p[q]} := (\vartheta_{p,i})^q_l$$

jest układem przekształceń liniowo niezależnych w $\boldsymbol{L}^n_{F2\nu}$ $(\boldsymbol{L}^m_{F2\kappa}(\boldsymbol{R}))$,

$$\vartheta_{r[s]} := (\vartheta_{r,i})^s_l$$

jest układem przekształceń liniowo niezależnych w $\boldsymbol{L}^n_{F2\kappa}(\boldsymbol{R}))$ oraz spełniony jest warunek, że $\vartheta_{\mathrm{kp}[q]} \circ \varphi \circ \nabla$. jest układem przekształceń należących do $\boldsymbol{L}^n_{F2\mu}(\boldsymbol{L}^l_{F2\kappa}(\boldsymbol{R}))$. Dla danego przekształcenia $\eta \in \boldsymbol{L}^n_{F2\mu}(\boldsymbol{L}^l_{F2\kappa}(\boldsymbol{R}))$, dla którego złożenie $\eta \circ \nabla$. jest operatorem opisującym wzorcowy system, poszukujemy takie przekształcenia $\gamma_{p,q} \in \mathrm{span}\partial_{p[q]}$ opisujące korektor $\gamma_{p,q} \circ \nabla$. systemu $\varphi \circ \nabla$., dla którego $\gamma_{p,q} \circ \varphi \circ \nabla \in \mathrm{span}\vartheta_{p[q]} \circ \varphi \circ \nabla$. jest najlepszym przybliżeniem przekształcenia η w sensie normy w przestrzeni $\boldsymbol{L}^n_{F2\mu}(\boldsymbol{L}^l_{F2\kappa}(\boldsymbol{R}))$.

Tak sformułowane zadanie korekcji zostało schematycznie przedstawione na rys. 3.1.(a). Jest to niekonstruktywna postać tego zadania. Niekonstruktywność jest spowodowana faktem, że z postaci tej nie wynika w sposób oczywisty rozwiązanie. Zauważmy jednak, że korekcja systemu $\varphi \circ \nabla$. jest równoważna wyznaczeniu średniokwadratowego przybliżenia przekształcenia η za pomocą układu przekształceń $\vartheta_{p[q]} \circ \varphi \circ \nabla$., co zostało pokazane na rys. 3.1.(b). Ponieważ w ogólnym przypadku, z liniowej niezależności układu $\vartheta_{\mathrm{kp}[q]}$ nie wynika liniowa niezależność układu $\vartheta_{p[q]} \circ \varphi \circ \nabla$., to rozwiązaniem zadania korekcji jest twierdzenie o aproksymacji średniokwadratowej (G. Ciesielski [8], s. 25), które zostało rozszerzone na przypadek układów liniowo zależnych.

W zadaniu korekcji systemu zakładamy, że elementy układu [przekształceń $\partial_{p[q]} \circ \varphi \circ \nabla$. należą do przestrzeni $\boldsymbol{L}^n_{F2\mu}(\boldsymbol{L}^l_{F2\kappa}(\boldsymbol{R}))$.Warto zatem sprawdzić jakie warunki muszą być spełnione by tak było w istocie.

Załóżmy, że $\varphi \circ \nabla$. jest ograniczonym operatorem mierzalnym, $\vartheta_{p[q]}$ jest układem przekształceń lokalnie ograniczonych z $\boldsymbol{L}^n_{F2\nu}(\boldsymbol{L}^m_{F2\kappa}(\boldsymbol{R}))$, miara ν jest określona na σ-ciele zbiorów borelowskich oraz miara μ jest skończona. Wówczas złożenia $\vartheta_{p[q]} \circ \varphi \circ \nabla$. jest układem przekształceń z $\boldsymbol{L}^n_{F2\mu}(\boldsymbol{L}^l_{F2\kappa}(\boldsymbol{R}))$, gdyż są one mierzalne (P. Billingsley [1], s. 182), a z uwagi na ich ograniczoność oraz skończoność miary μ, są również całkowalne. Pozostaje więc sprawdzić, czy operatory o postaci $\varphi \circ \nabla$. mogą być mierzalne? Odpowiedź pozytywną na to pytanie udziela nam następujące twierdzenie.

TWIERDZENIE. *Załóżmy, że κ jest taką miarą skończoną na σ-ciele zbiorów borelowskich w $\boldsymbol{R}$, która spełnia warunek:*

$$\forall u \in \boldsymbol{L}_{F2\kappa}(\boldsymbol{R}) \forall t \in \boldsymbol{R} \exists M{:}\boldsymbol{R} \mapsto \boldsymbol{R} \ (\| \nabla_t(u) \|_2 \leqslant M(t) \| u \|_2).$$

λ jest miarą Lebesgue'a na $\boldsymbol{R}$ oraz $\kappa \equiv \lambda$ [14]. Jeżeli $\varphi{:}\boldsymbol{L}^m_{R2\kappa}(\boldsymbol{R}) \to \boldsymbol{F}^n$ jest przekształceniem borelowskim takim, że złożenie $\varphi \circ \nabla.{:}\boldsymbol{L}^n_{F2\kappa}(\boldsymbol{R}) \mapsto \boldsymbol{C}^n_{F\infty\kappa}(\boldsymbol{R})$, to $\phi := \varphi \circ \nabla$ jest operatorem borelowskim.

[14] Przez $\mu \equiv \nu$ będziemy oznaczać *miary równoważne*, czyli miary takie, że *miara* μ jest absolutnie ciągła względem miary ν (*miara* ν *dominuje miarę* μ), a zatem dla każdego mierzalnego zbioru A.

$$(\nu(A) = 0) \Rightarrow (\mu(A) = 0),$$

co zapisujemy $\mu < < \nu$, oraz *miara* ν jest *absolutnie ciągła względem miary* μ, czyli $\mu > > \nu$ (P. Billingsley [1], s. 424).

D o w ó d. Z założenia, że $\forall u \in L^m_{F2\kappa}(R)$ oraz $\forall t \in R\ \exists M: R \mapsto R$, dla którego

$$\| \nabla_t(u) \|_2 \leqslant M(t) \| u \|_2$$

wynika, iż $\forall t \in R$ obraz

$$\nabla_t(L^m_{F2\kappa}) = L^m_{F2\kappa},$$

a zatem przestrzeń

$$L^m_{F2\kappa\nabla} = L^m_{F2\kappa}.$$

Ponieważ przekształcenie φ jest określone na $L^m_{F2\kappa}$, to zauważmy, że $\phi := \varphi \circ \nabla$. jest dobrze zdefiniowanym operatorem na $L^m_{F2\kappa}$.

Miara κ jst skończona, a stąd wynika, że przestrzeń $C^n_{F\infty\kappa} \subset L^n_{F\infty\kappa} \subset L^n_{F2\kappa}$ (J. Musielak [28], s. 54. Załóżmy, że τ jest typologią wyznaczoną przez iloczyn skalarny w $L^m_{F2\kappa}$ oraz τ_V jest typologią indukowaną w $C^n_{F\infty\kappa}$ przez topologię wyznaczoną przez iloczyn skalarny w $L^n_{F2\kappa}$. Ponieważ $\phi: L^m_{F2\kappa} \to C^n_{F\infty\kappa}$ to wystarczy wykazać, że $\forall V \in \tau_V$ przeciwobraz $\phi^{-1}(V)$ należy do σ-ciała zbiorów borelowskich $\sigma(\tau_U)$ [15].

W tym celu, zdefiniujmy $\forall t \in R$ przekształcenie $\pi_t: C^n_{F\infty\kappa} \mapsto F^n$ w następujący sposób:

$$\forall v \in C^n_{F\infty\kappa}(\pi_t(v) := v(t)).$$

Wówczas $\forall V \in \tau_V$ przeciwobraz

$$\phi^{-1}(V) := \{u \in L^m_{F2\kappa} : \phi(U) \in V\} = \{u \in L^m_{F2\kappa} : \forall t \in R([\phi(u)](t) \in \pi_t(V))\} =$$

$$= \bigcap_{t \in R} \{u \in L^m_{F2\kappa} : [\phi(u)](t) \in \pi_t(V)\} = \bigcap_{t \in R} \{u \in L^m_{F2\kappa} : \varphi(\nabla_t(u)) \in \pi_t(V)\} =$$

$$= \bigcap_{t \in R} [\varphi \circ \nabla_t]^{-1}(\pi_t(V)) = \bigcap_{t \in R} \nabla_{-t}(\varphi^{-1}(\pi_t(V))).$$

Oznaczamy przez τ_X topologię w F^n. Wykażemy teraz, że $\forall t \in R$ oraz $\forall V \in \tau_V$ zbiór $\pi_t(V) \in \tau_X$, czyli jest otwarty. Niech $\| \cdot \|_\infty$ będzie normą w $L^m_{F\infty\kappa}$. Zgodnie z definicją topologii indukowanej przez metrykę (S. Gładysz [24], s. 48) wystarczy pokazać, że $\forall a \in C^n_{F\infty\kappa}$, $\forall r \in R_+$, oraz $\forall t \in R$ obraz kuli otwartej

$$S_V(a, r) := \{v \in C^n_{F\infty\kappa} : \| a - v \|_\infty < r\}$$

w odwzorowaniu π_t jest zbiorem otwartym.

Ponieważ elementy przestrzeni $C^n_{F\infty\kappa}$ są funkcjami ciągłymi na R, a przestrzeń $C^n_{F\infty\kappa} = C^n_{F\infty\lambda}$, gdyż założyliśmy, że miary κ oraz λ są równoważne, to $\forall v \in C^n_{F\infty\kappa}$ norma

$$\| v \|_\infty := \underset{\sqrt{t \in R}}{\text{supess}} \mid v(t) \mid = \underset{\sqrt{t \in R}}{\sup} \mid v(t) \mid ,$$

[15] Jeżeli $\mathcal{F}$ jest rodziną podzbiorów zbioru X, to przez $\sigma(T)$ będziemy oznaczali σ-ciało generowane w X przez rodzinę $\mathcal{F}$ (P. Billingsley [1], s. 32).

czyli $\forall a \in C^n_{F\infty\kappa}$ oraz $\forall r \in R_+$

$$S_V(a,r) = \left\{ v \in C^n_{F\infty\kappa} : \sup_{t\in R} |a(t) - v(t)| < r \right\}.$$

Zauważmy, że $\forall t \in R$ zbiór

$$\pi_t(S_V(a,r)) = \{v(t) \in F^n : v \in S_V(a,r)\} = \{v(t) \in F^n : \mid a(t) - v(t) \mid < r\} =: S_X(a(t), r).$$

Wynika zatem, że $\forall t \in R$ obrazem dowolnej kuli otwartej S_V z $C^n_{F\infty\kappa}$ w odwzorowaniu π_t jest kula otwarta S_X w F^n. Mamy więc, że $\forall t \in$ R oraz $\forall V \in \tau_V$ zbiór $\pi_t(V) \in \tau_X$, a zatem jest otwarty.

Wykażemy teraz, że $\forall t \in R$ oraz $\forall V \in \tau_V$ zbiór $\pi_t(V)$ jest przeliczalną sumą zbiorów z τ_X. W tym celu oznaczamy przez β_X bazę topologii τ_X w F^n. Można wykazać (S. Gładysz [24], s. 50), że istnieje przeliczalna baza

$$\beta_X := \{A_i \in \tau_X : i \in N\}$$

topologii τ_X. Bazę tą mogą tworzyć wszystkie kule otwarte $S_N(a,r)$, które $\forall a \in F^n$ oraz $r \in R_+$ spełniają warunki:

$$\mathrm{Re}(a) \in Q^n\,[16], \qquad \mathrm{Im}(a) \in Q^n \quad \text{oraz} \quad r \in Q.$$

Ponieważ wykazaliśmy już, że $\forall t \in R$ oraz $\forall V \in \tau_V$ zbiór $\pi_t(V) \in \tau_X$, to istnieje przeliczalny zbiór indeksów $N_{Vt} \subset N$, dla którego rodzina zbiorów

$$\beta_{Vt} := \{A_i \subset \beta_X : i \in N_{Vt} \subset N\}$$

spełnia równość:

$$\pi_t(V) = \bigcup_{i\in N_{Vt}} A_i .$$

Wówczas przeciwobraz

$$\phi^{-1}(V) = \bigcap_{t\in R} \nabla_{-t}\left(\varphi^{-1}(\pi_t(V))\right) = \bigcap_{t\in R}\left(\nabla_{-t}\left(\varphi^{-1}\left(\bigcup_{i\in N_{Vt}} A_i\right)\right)\right) = \bigcap_{t\in R}\left(\bigcup_{i\in N_{Vt}} \nabla_{-t}(\varphi^{-1}(A_i))\right).$$

Z założenia φ jest mierzalne, a zatem $\forall i \in N$ przeciwobraz $\varphi^{-1}(A_i)$ należy do σ-ciała $\sigma(\tau_U)$ czyli jest mierzalny. Pokażemy teraz, że $\forall t \in R$ oraz nawet $\forall A \in \tau_X$, zbiór $\nabla_{-t}(\varphi^{-1}(A))$ też jest mierzalny.

W tym celu załóżmy, że ciąg $(u_i \in L^m_{F2\kappa})_N$ jest zbieżny w sensie normy $\|\cdot\|_2$ w przestrzeni $L^m_{F2\kappa}$, do pewnego $u \in L^m_{F2\kappa}$. Ponieważ zgodnie z założeniem $\forall u \in L^m_{F2\kappa}(R)$ oraz $\forall t \in R$ $\exists M : R \to R$, że

$$\| \nabla_t(u) \|_2 \leq M(t) \| u \|_2 ,$$

[16] Przez Q będziemy oznaczać zbiór liczb wymiernych. Przypomnijmy, że zbiór ten jest przeliczalny (K. Kuratowski [26], s. 60).

to

$$\| \nabla_t(u) - \nabla_t(u_i) \|_2 \leqslant M(t) \| u - u_i \|_2 .$$

Operator ∇_t jest więc ciągły na $L^m_{F2\kappa}$, a zatem jest również mierzalny, gdyż $\forall U \in \tau_U$ zbiór $\nabla_{-t}(U) \in \tau_v \subset \sigma(\tau_v)$. Wynika stąd, że $\forall W \in \sigma(\tau_u)$ zbiór $\nabla_{-t}(W)$ jest mierzalny (P. Billingsley [1], s. 182), czyli $\forall t \in \boldsymbol{R}$ oraz $\forall A \in \tau_X$ zbiór $\nabla_{-t}(\varphi^{-1}(A))$ jest rzeczywiście mierzalny.

Niech $\forall V \in \tau_Y$ oraz $\forall t \in \boldsymbol{R}$ zbiór

$$U_{Vt} := \bigcup_{i \in N_{Vt}} \nabla_{-t}(\varphi^{-1}(A_i)) .$$

Zbiór ten jest oczywiście mierzalny, gdyż każde σ-ciało zawiera wszystkie przeliczalne sumy swoich zbiorów. Pokazaliśmy wcześniej, że $\forall V \in \tau_Y$ przeciwobraz

$$\phi^{-1}(V) = \bigcap_{t \in \boldsymbol{R}} U_{Vt} .$$

Ponieważ $\boldsymbol{R}$ jest zbiorem nieprzeliczalnym, to $\forall V \in \tau_Y$ iloczyn $\bigcap_{t \in \boldsymbol{R}} U_{Vt}$ jest nieprzeliczalny i może nie należeć do $\sigma(\tau_U)$. Przypomnijmy jednak $\forall V \in \tau_Y$ przeciwobraz

$$\phi^{-1}(V) = \bigcap_{t \in \boldsymbol{R}} \left(\bigcup_{i \in N_{Vt}} \nabla_{-t}(\varphi^{-1}(A_i)) \right).$$

Oznaczając

$$N_V := \bigcap_{t \in \boldsymbol{R}} N_{Vt} ,$$

możemy napisać, że

$$\phi^{-1}(V) = \bigcup_{i \in N_V} \nabla_{-t}(\varphi^{-1}(A_i)) .$$

Oczywiście zbiór $N_V \subset \boldsymbol{N}$ i jest tym samym przeliczalny. Wynika stąd, że przeciwobraz $\phi^{-1}(V)$ również należy do σ-ciała $\sigma(\tau_U)$ a zatem ϕ jest operatorem mierzalnym, co kończy dowód. ∎

4. PODSUMOWANIE

Założenie, że odpowiedzi operatora ϕ są sygnałami ciągłymi i ograniczonymi, nie jest zbyt drastyczne. Zauważymy, że każdy fizycznie realizowany system ma ograniczone pasmo przenoszenia, a stąd wynika ciągłość jego odpowiedzi. Z kolei założenie, że odpowiedź operatora ϕ jest ograniczona, można uzasadnić ograniczonymi napięciami i prądami zasilania, ograniczoną liczbą bitów, skończoną skalą pomiarową itp.

Zauważmy jeszcze, że jeżeli κ jest miarą skończoną, to operator ϕ, o którym mowa w ostatnim twierdzeniu, można traktować jako odwzorowanie przestrzeni $L^m_{F2\kappa}$ w $L^n_{F2\kappa}$, albowiem $C^n_{F\infty\kappa} \subset L^n_{F\infty\kappa} \subset L^n_{F2\kappa}$(J. Musielak [28], s. 54). Jest on wówczas również mierzalny, gdyż σ-ciało w $C^n_{F\infty\kappa}$ jest wówczas σ-ciałem indukowanym przez σ-ciało w $L^n_{F2\kappa}$ (P. Billingsley [1], s. 159). Wynika stąd, że elementy składowe układu $\vartheta_{p[q]} \circ \vartheta \circ \nabla$ określonego w zadaniu korekcji systemu, mogą należeć do przestrzeni $L^n_{F2\mu}(L^l_{F2\kappa}(\boldsymbol{R}))$.

BIBLIOGRAFIA

1. P. B i l l i n g s l e y (tłum. z ang. K. Kizeweter, J. E. Rogulski): *Prawdopodobieństwo i miara*. PWN, Warszawa, 1987.
2. G. C i e s i e l s k i: *Aproksymacja średniokwadratowa operatorów nieliniowych opisujących wielowymiarowe systemy pomiarowe, XXIII Międzyuczelniana Konferencja Metrologów*. Politechnika Warszawska, Warszawa, 1991 (referat nie opublikowany).
3. G. C i e s i e l s k i: *Modele Wienera wielowymiarowych systemów pomiarowych opisanych za pomocą operatorów nieliniowych, Modelowanie i symulacja systemów pomiarowych*. Materiały Sympozjum, AGH, Kraków, 1992, ss. 41 – 50.
4. G. C i e s i e l s k i: *Operatorowe modele wielowymiarowych systemów pomiarowych*. Seminarium Naukowe Komisji Kształcenia Komitetu Metrologii i Aparatury Naukowej Polskiej Akademii Nauk. 1994 (referat nie opublikowany).
5. G. C i e s i e l s k i: *Identyfikacja ogólnej struktury wielowymiarowych systemów pomiarowych, Modelowanie i symulacja systemów pomiarowych*. Materiały Sympozjum, AGH, Kraków, 1994, ss. 17 – 23.
6. G. C i e s i e l s k i: *Identyfikacja ogólnej struktury wielowymiarowych systemów stacjonarnych*. Kwartalnik Elektroniki i Telekomunikacji, t. 40, z. 3, 1994, ss. 419 – 428.
7. G. C i e s i e l s k i: *Modelowanie wielowymiarowych liniowych systemów stacjonarnych za pomocą splotów*. Kwartalnik Elektroniki i Telekomunikacji, t. 40, z. 3, 1994, ss. 429 – 448.
8. G. C i e s i e l s k i: *Modelowanie i korekcja wielowymiarowych stacjonarnych systemów pomiarowych za pomocą operatorów nieliniowych*. Rozprawa habilitacyjna, Politechnika Łódzka, zeszyty Naukowe nr 699, Łódź, 1994.
9. G. C i e s i e l s k i: *Identyfikacja szczegółowej struktury wielowymiarowych systemów stacjonarnych*. Kwart. Elektr. i Telekom., t. 41, z. 3, 1995, ss. 305 – 319.
10. G. C i e s i e l s k i: *Modelowanie wielowymiarowych systemów stacjonarnych przyczynowych*. Kwartalnik Elektroniki i Telekomunikacji, t. 41, z. 3, 1995, ss. 321 – 335.
11. G. C i e s i e l s k i: *Modelowanie wielowymiarowych liniowych systemów stacjonarnych w systemach pomiarowych. Modelowanie i symulacja systemów pomiarowych*. Materiały Sympozjum, AGH,Kraków, 1996, ss. 28 – 37.
12. G. C i e s i e l s k i: *Identryfikacja szczegółowej struktury wielowymiarowych systemów stacjonarnych w systemach pomiarowych, Modelowanie i symulacja systemów pomiarowych*. Materiały Sympozjum, AGH, Kraków, 1996, ss. 38 – 42.
13. G. C i e s i e l s k i: *Zastosowanie modeli Wienera w systemach pomiarowych. Systemy pomiarowe w badanich naukowych i w przemyśle*. Materiały konferencyjne, Wyższa Szkoła Inżynierska, Zielona Góra, 1996, ss. 34 – 52.
14. G. C i e s i e l s k i: *Modelowanie wielowymiarowych systemów stacjonarnych w systemach pomiarowych. Modelowanie i symulacja systemów pomiarowych*. Materiały Sympozjum, AGH, Kraków, 1997, ss. 129 – 142.
15. G. C i e s i e l s k i: *Korekcja wielowymiarowych systemów stacjonarnych w systemach pomiarowych. Modelowanie i symulacja systemów pomiarowych*. Materiały Sympozjum, AGH, Kraków, 1997, ss. 143 – 150.
16. G. C i e s i e l s k i: *Pomiar parametrów modeli i korektorów systemów statycznych. Modelowanie i symulacja systemów pomiarowych*. Materiały Sympozjum. AGH, Kraków, 1998, ss. 59 – 66.

17. G. Ciesielski: *Pomiary parametrów modeli liniowych systemów stacjonarnych. Modelowanie i symulacja systemów pomiarowych.* Materiały Sympozjum. AGH, Kraków, 1998, ss. 67–82.
18. G. Ciesielski: *Pomiar parametrów korektorów liniowych systemów stacjonarnych. Modelowanie i symulacja systemów pomiarowych.* Materiały Sympozjum, AGH, Kraków, 1998, ss. 83–90.
19. G. Ciesielski: *Modelowanie liniowych systemów stacjonarnych za pomocą systemów Laguerre'a,* Kwartalnik Elektroniki i Telekomunikacji, (artykuł złożony do druku).
20. G. Ciesielski: *Korekcja liniowych systemów stacjonarnych za pomocą systemów Laguerre'a,* Kwartalnik Elektroniki i Telekomunikacji (artykuł złożony do druku).
21. G. Ciesielski: *General structure of multidimensional time-invariant systems, Communications on Pure and Applied Mathematics.* John Wiley & Sons, Inc. (artykuł w przygotowaniu).
22. R. J. P. de Figueiredo, T. A. W. Dwyer: *A best approximation framework and implementation for simulation of large-scale nonlinear systems,* IEEE Trans. Circuits Syst., vol. CAS-27, no 11, 1980, pp. 1005–1014.
23. R. J. P. de Figueiredo: *Nonlinear circuits and systems applications of functional splines.* Proc. IEEE, 1982, pp. 1245–1247.
24. S. Gładysz: *Wstęp do topologii,* Matematyka dla politechnik. PWN, Warszawa, 1981.
25. J. Jaworski, R. Morawski, J. Olędzki: *Wstęp do metrologii i techniki eksperymentu.* WNT. Warszawa, 1992.
26. K. Kuratowski: *Wstęp do teorii mnogości i topologii.* Bilioteka Matematyczna, t. 9. PWN, Warszawa, 1980.
27. F. Morrison: *Sztuka modelowania układów dynamicznych.* WNT, Warszawa, 1996.
28. J. Musielak: *Wstęp do analizy funkcjonalnej.* PWN, Warszawa, 1989.
29. J. Park, J. W. Sandberg: *Criteria for the approximation of nonlinear systems.* IEEE Trans. Circuits Syst.-I: Fundamental Theory and Applications, vol. 39, no 8, 1992, pp, 673–676.
30. K. A. Pupkow, W. I. Kapalin, A. S. Juszczenko: *Funkcjonalnyje rjady w tjeorii nieliniejnych sistjeme.* Nauka, Moskwa, 1976.
31. I. W. Sandberg: *Criteria for the response of nonlinear systems to be L-asymptotic periodic.* Bell Syst. Tech, J., vol. 60, no 10, 1981, pp. 2359–2371.
32. I. W. Sandberg: *Series expansion for nonlinear systems.* Proc IEEE, 1982, pp. 110–113.
33. I. W. Sandberg: *Expansion for nonlinear systems.* Bell Syst. Tech. J., vol. 61, no 2, 1982, pp. 159–199.
34. I. W. Sandberg: *Volterra expansion for time-varying nonlinear systems.* Bell Syst. Tech. J., vol. 61, no 2, 1982, pp. 201–225.
35. I. W. Sandberg: *Multidimensional nonlinear systems and structure theorems.* J. Circuits, Syst. and Computers, vol. 2, no 4, 1992, pp. 382–388.
36. I. W. Sandberg: *Approximately-finit memory and input-output maps.* IEEE Trans. Circuits Syst.-I: Fundamental Theory and Applications, vol. 39, no 7, 1992, pp. 549–556.
37. M. Schetzen: *The Volterra and Wiener Theories of Nonlinear Systems.* John Wiley & Sons, New York, 1980.
38. T. Söderström, P. Stoisa: *Identyfikacja systemów.* PWN, Warszawa, 1997.
39. A. Wojnar: *Teoria sygnałów.* Podręczniki akademickie EIT. WNT, Warszawa, 1980.
40. R. R. Yager, D. P. Filev: *Podstawy modelowania i sterowania rozmytego.* WNT, Warszawa, 1995.

G. CIESIELSKI

CORRECTION OF MULTIDIMENSIONAL TIME-INVARIANT AND CAUSAL SYSTEMS

Summary

This paper refers to the correction of multidimensional time-invariant systems (not necessarily linear). There are considering the systems, which can be described by maps, which are called operators.

Author of this paper for several years has researched on the uniform general operator theory of modeling and correction of the mentiones systems. This research led him to the versatile resolution of the

corrector choice problem for these systems. It appears that correction problem can be resolve itself into the task of reference operator modeling. This operator appears in the corection task. Resolution of appeared problem is then the theorem, which resolves the task of the multidimensional time-invariant continuos systems modeling. This theorem is presented in the author's past PhD thesis [8], p. 51, and papers: [10] and [11].

The obtained results, similary to the past PhD thesis [8], p. 51, and papers [10] and [11] on modeling are such versatile that it is not directly related to the no one know operators (e.g. Volterra's or Wiener's). This results points out the formal limits, which are not to transgress in any one correction process of systems mentioned before.

The summarized paper is natural continuation of autor's works presented in the years: 1994, and 1995, in Electronic and Telecommunication Quarterly, and in the years: 1992, 1994, and 1996 – 1998, on the symposiums on modeling and simulation of systems.

Key words: universal functional theory of systems, correction, modeling, structure identification, parametric identification, multidimensional systems, nonlinear systems, time-invariant systems, causal systems, operators, convolutions, stationary space, convolution space, causal space.

Porówn

V
nych
Rozw
trzech
zadan
przez
pozos
turze
nienia
efekty
uprze

heury

W kla
każde z z
jaki wyst
spowodo
z szerego
wego po
wymagan
zdefiniow
(ang. Dec
mamy d
W artyku
minimali

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI, 2000, 46, z. 1, ss. 21–34

Porównanie efektywności różnych algorytmów szeregowania zadań wieloprocesowych

MIROSŁAW GAJER

Katedra Automatyki AGH, Kraków

Otrzymano 1999.07.28
Autoryzowano 1999.12.10

W artykule rozważono problematykę szeregowania zadań wieloprocesowych, związanych z dziedziną przetwarzania obrazów, dla wieloprocesowego układu TMS320C80. Rozważony został problem szeregowania zbioru niezależnych zadań wieloprocesowych dla trzech procesorów DSP, wchodzących w skład układu TMS320C80. Celem postawionego zadania było znalezienie takiego planu szeregowania jedno i dwuprocesowych zadań, przeznaczonych dla dedykowania procesorów DSP, aby łączny czas, w którym procesory pozostają w stanie jałowym był minimalny. Dokonano przeglądu proponowanych w literaturze algorytmów oraz zaproponowano nowe oryginalne rozwiązanie rozważonego zagadnienia. Zamieszczono również wyniki eksperymentów, których celem było porównanie efektywności nowo zaproponowanego algorytmu z dwoma algorytmami zaprezentowanymi uprzednio w literaturze.

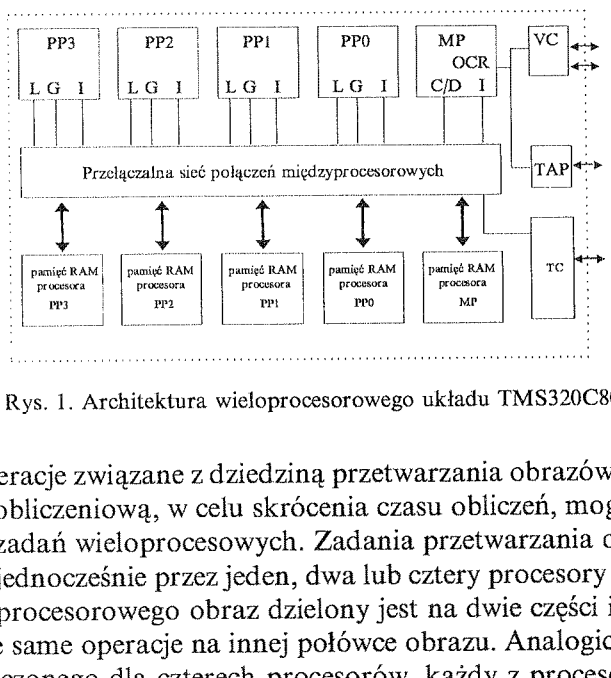
Słowa kluczowe: systemy wieloprocesowe, szeregowanie i alokacja zadań, algorytmy heurystyczne

1. WPROWADZENIE

W klasycznej teorii szeregowania zadań przyjmuje się, że w danej chwili czasu każde z zadań wykonywane może być tylko przez jeden procesor. Jednakże postęp, jaki wystąpił w ostatnich latach w dziedzinie konstrukcji systemów równoległych, spowodował, że w literaturze przedmiotu pojawia się coraz więcej prac związanych z szeregowaniem zadań wieloprocesowych [1]. Do wykonania zadania wieloprocesowego potrzeba, w danej chwili czasu, co najmniej dwóch procesorów. Zbiór wymaganych, do wykonania danego zadania, procesorów może zostać uprzednio zdefiniowany i wówczas rozważany jest przypadek tzw. procesorów dedykowanych (ang. Dedicated processors) lub też wybór procesorów może być arbitralny i wówczas mamy do czynienia z tzw. procesorami równoległymi (ang. Parallel processors). W artykule zostanie rozważony problem poszukiwania optymalnego, ze względu na minimalizację czasu pracy jałowej procesorów, planu szeregowania zbioru niezależ-

nych zadań wieloprocesowych, przeznaczonych dla dedykowania procesorów DSP, wchodzących w skład wieloprocesowego układu TMS320C80. Uwaga zostanie skoncentrowana na tzw. przypadku zadań niewywłaszczalnych (ang. Non-pre-emptive case), gdzie każde z zadań po uruchomieniu musi zakończyć swoje wykonywanie i nie może ulec przerwaniu.

Wieloprocesowy układ tupu TMS320C80 firmy Texas Instruments wyposażony został w cztery, pracujące współbieżnie, procesory DSP (oznaczone jako PP0, PP1, PP2 i PP3), posiadające architekturę zaprojektowaną specjalnie pod kątem przyśpieszenia operacji przetwarzania obrazów cyfrowych. Architektura wieloprocesowego układu TMS320C80 została przedstawiona w sposób schematyczny na rys. 1. Oprócz pięciu procesorów: MP (procesor nadrzędny typu RISC), PP0, PP1, PP2 i PP3 (procesory równoległe), układ TMS320C80 został wyposażony także w dwa kontrolery: VC (sterowanie wymiana danych wizyjnych) oraz TC (nadzorowanie przebiegu komunikacji systemu z urządzeniami zewnętrznymi). Ponadto każdy z procesorów posiada przydzielone 10 KB wewnętrznej pamięci RAM, a komunikacja międzyprocesorowa dokonywana jest poprzez wspólne dla wszystkich procesorów bloki pamięci za pośrednictwem układu przełącznicy (ang. Switching crossbar network), posiadającego zdolność do automatycznej konfiguracji aktualnie potrzebnych połączeń pomiędzy portami procesorów (L - port lokalny, G - port globalny, I - port instrukcji, C/D - port służący do komunikacji z podręczną pamięcią danych). Układ TMS320C80 został wyposażony także w port testowy TAP, służący do nadzorowania stanu pracy procesora z poziomu programu uruchomieniowego.



Rys. 1. Architektura wieloprocesorowego układu TMS320C80

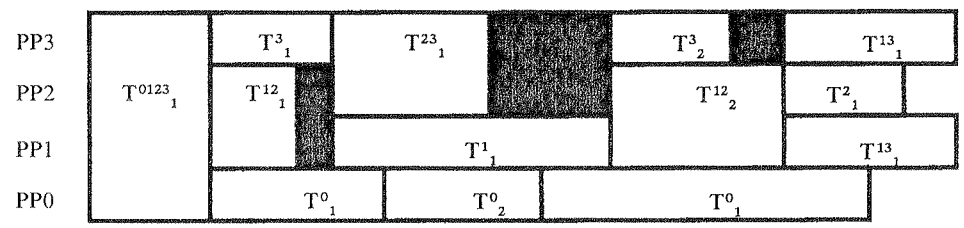
Ponieważ operacje związane z dziedziną przetwarzania obrazów posiadają zwykle dużą złożoność obliczeniową, w celu skrócenia czasu obliczeń, mogą zostać zrealizowane w postaci zadań wieloprocesowych. Zadania przetwarzania obrazów, wykonywane mogą być jednocześnie przez jeden, dwa lub cztery procesory DSP. W przypadku zadania dwuprocesorowego obraz dzielony jest na dwie części i każdy z procesorów wykonuje te same operacje na innej połówce obrazu. Analogicznie w przypadku zadania przeznaczonego dla czterech procesorów, każdy z procesorów wykonuje te

same operacje na przydzielonej mu ćwiartce obrazu. Zatem zrównoleglenie operacji przetwarzania obrazów polega w tym wypadku na implementacji architektur typu SIMD. Ponieważ w strukturę układu TMS320C80 wbudowano przeznaczone dla każdego z procesorów, bloki wewnętrznej statycznej pamięci RAM, przewidzianej między innymi do przechowywania danych, na których wykonywane będą obliczenia. Celowym jest związanie zadań na stałe z dedykowanymi procesorami, które posiadają w swej pamięci wewnętrznej dane potrzebne do wykonywania zamierzonych obliczeń. Przyjęcie takiego rozwiązania przyczynia się niewątpliwie do zmniejszenia narzutów związanych z komunikacją przeprowadzaną, pomiędzy poszczególnymi procesorami, w celu pozyskania danych. Rozwiązanie alternatywne polegałoby na losowym wyborze procesorów do realizacji kolejnych zadań przetwarzania obrazów. W takim jednak przypadku, przed przystąpieniem do wykonywania obliczeń, każdy z procesorów musiałby do związanego z nim bloku wewnętrznej pamięci SRAM załadować dane dotyczące przetwarzanego aktualnie obrazu. W przypadku gdy na danym obrazie ma zostać wykonanych wiele operacji, korzystniej jest rozważać dedykowane procesory, w pamięci, których obraz ten został umieszczony.

2. ZADANIA WIELOPROCESOWE DLA UKŁADU TMS320C80

Niech $P=\{0, 1, 2, 3\}$ będzie zbiorem procesorów DSP układu TMS320C80, odpowiednio PP0, PP1, PP2 i PP3. Z kolei niech $T=\{T_0, T_1 ..., T_n\}$ będzie zbiorem niezależnych zadań wieloprocesowych o przyjętych w sposób arbitralny czasach wykonywania. W celu wykonania, każde z zadań $T_i, 1 \leqslant i \leqslant n$ wykorzystuje przez odcinek czasu $\mathbf{p}(T_i)$ pewien zdefiniowany uprzednio, niepusty zbiór procesorów. W zależności od zapotrzebowania na dostępność procesorów, zbiór zadań T dzielony jest na następujące podzbiory: T^0, T^1, T^2, T^3, T^{12}, T^{13}, T^{23} *i* T^{0123}, gdzie górne indeksy wskazują na numery, potrzebnych do wykonania danego zadania, procesorów. Dla uproszczenia rozważań założono, że procesor PP0 wykorzystywany będzie jedynie do wykonywania zadań jednoprocesowych. Wówczas zagadnienie szeregowania zadań jedno i dwuprocesorowych rozpatrywane jest wyłącznie dla trzech procesorów, dla którego to przypadku zostały podane w literaturze przedmiotu liczne rozwiązania (autorowi nie są znane prace rozważające problem szeregowania zadań jedno i dwuprocesorowych dla czterech dedykowanych procesorów). Na pozostałych procesorach PP1, PP2 i PP3 wykonywane mogą być zarówno zadania jedno, jak i dwuprocesorowe. Problem polega na znalezieniu takiego planu uszeregowania powyższych zadań, który zapewniałby minimalny łączny czas, w którym procesory PP1, PP2 i PP3 znajdować się będą w stanie jałowym (nie będą wykonywały żadnych obliczeń). Ponieważ zadania należące do zbioru T^{0123} wymagają jednoczesnej dostępności wszystkich czterech procesorów, zatem zadania te mogą zostać umieszczone na początku każdego (w tym również optymalnego) planu szeregowania zadań, nie wywierając jakiegokolwiek wpływu na łączną wartość czasu przebywania procesorów systemu

TMS320C80 w stanie jałowym. Zatem bez jakiejkolwiek straty na ogólności rozważań można przyjąć, że zbiór T^{0123} jest pusty. Przykładowy plan szeregowania zbioru zadań T dla wieloprocesowego układu TMS320C80 został zamieszczony na rys. 2. Kolorem czarnym zaznaczono odcinki czasu, w których procesory znajdują się w stanie jałowym.



Rys. 2. Przykładowy plan szeregowania zadań wieloprocesorowych dla dedykowanych procesorów DSP wchodzących w skład układu TMS320C80

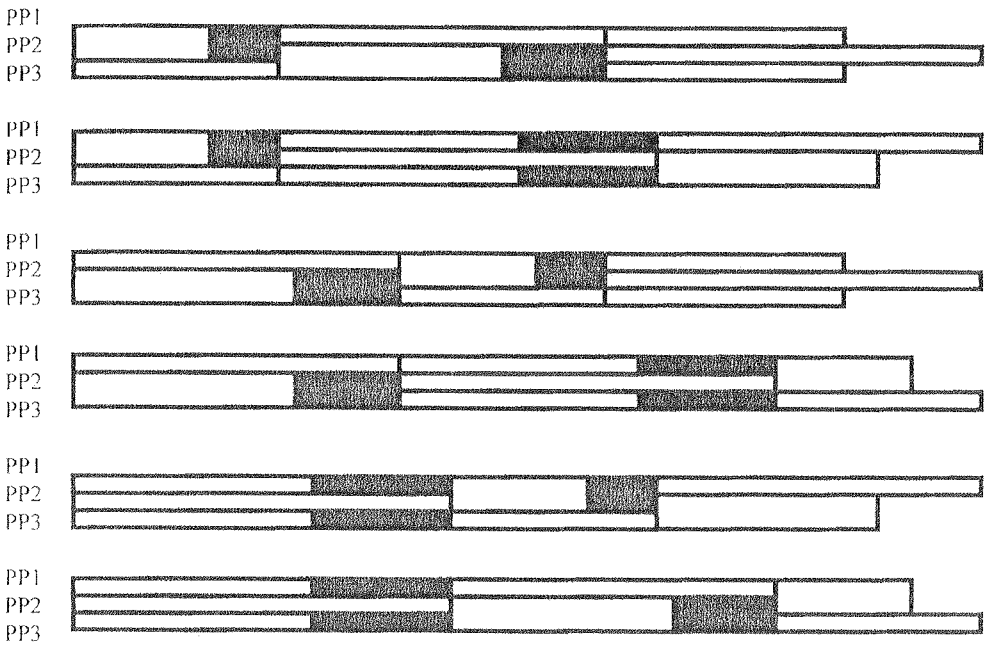
Rozważany problem uległ zatem redukcji do problemu znalezienia możliwego do wykonania planu szeregowania zadań, odznaczającego się minimalnym łącznym czasem przebywania procesorów w stanie jałowym, dla podzbiorów zadań T^i $1 \leqslant i \leqslant 3$, zwanych zadaniami jednoprocesorowymi (ang. Uniorocessor tasks) oraz T^{jk}, $1 \leqslant j \leqslant k \leqslant 3$, zwanych zadaniami dwuprocesorowymi (ang. biprocessor tasks). Ponieważ każdy z procesorów może w danym momencie wykonywać tylko jedno zadanie, dwa różne zadania które wymagają dostępności tego samego procesora nie mogą być oczywiście wykonywane równocześnie. Zadania takie nazywane są niekompatybilnymi (ang. incompatible). Zgodnie z notacją wprowadzona w [2] powyższy problem szeregowania niewywłaszczalnych zadań jedno i dwuprocesorowych dla trzech dedykowanych procesorów jest oznaczany jako $P_3 \mid fix_j \mid C_{\max}$. W pracy [3] wykazano, że problem $P_3 \mid fix_j \mid C_{\max}$ należy do klasy problemów *NP*-trudnych. Z kolei w pracy [4] zaproponowano algorytm heurystyczny (ang. normal schedule heuristic), który może zostać wykorzystany do rozwiązania rozważanego problemu oraz wykazano, że dla algorytmu tego graniczne wykorzystanie zasobów systemu wynosi $\frac{4}{5}$. Natomiast w pracy [5] zaproponowano algorytm dla którego graniczne wykorzystanie zasobów systemu zostało zwiększone do $\frac{6}{7}$. Ponadto w literaturze [6], [7], [8] i [9] podano algorytmy rozwiązania powyższego problemu, wykorzystujące metodę podziału i ograniczeń (ang. branch-and-bound).

W dalszej części wykorzystywana będzie funkcja wagi ω zdefiniowana na podzbiorze zadań T', taka, że dla każdego $T' \subseteq T$, $\omega(T')$ odpowiada łącznemu czasowi wykonywania wszystkich zadań, należących do podzbioru T', zatem

$$\omega(T') = \sum_{T_j \in T'} p(T_j).$$

3. ALGORYTMY HEURYSTYCZNE SZEREGOWANIA ZDAŃ WIELOPROCESOROWYCH

Zaproponowany w pracy [3] algorytm *A1*, szereguje w sposób ciągły wszystkie zadania, które do wykonania wymagają tego samego podzbioru procesorów. Spośród sześciu możliwych planów uszeregowania rozważanych podzbiorów zadań, które zostały przedstawione na rys. 3, wybierany jest ten, dla którego łączny czas pracy jałowej procesorów jest minimalny.



Rys. 3. Przykładowe plany szeregowania zadań, uzyskane w wyniku zastosowania algorytmu A1

Z kolei, zaproponowany w pracy [5] algorytm *A2* składa się z trzech etapów. W kroku pierwszym zbiór zadań jednoprocesorowych Ti dzielony jest na dwie części, a następnie zadania należące do każdego z uzyskanych w ten sposob dziewięciu podzbiorów zadań, szeregowane są w sposób ciągły. To znaczy podczas wykonywania zadania, należące do jednego z podzbiorów zadań, wykonywane są jedno po drugim i pomiędzy takimi zadaniami nie może znaleźć się żadne zadanie należące do innego podzbioru zadań. Spośród wszystkich uzyskanych w ten sposób planów szeregowania zadań, wybierany jest ten, dla którego łączny czas pracy jałowej procesorów jest minimalny. Zbiór zadań jednoprocesorowych T^i, $1 \leqslant i \leqslant 3$ dzielony jest na tzw. część długą T_l^i oraz krótką T_s^i, tak aby był spełniony warunek $\omega(T_l^i) \geqslant (T_s^i)$. Podział zbiorów zadań jednoprocesorowych dokonywany jest w sposób następujący:

- Jeżeli jedno z zadań, należących do zbioru T^iposiada czas wykonania większy bądź równy $\frac{1}{3}\omega(T^i)$, to zadanie takie zostaje wyizolowane w jednym z pod-

zbiorów, natomiast w drugim z podzbiorów zostają umieszczone wszystkie pozostałe zadania.

- W przypadku przeciwnym, zadania zostają kolejno przypisane do pierwszego z podzbiorów, aż do momentu gdy całkowity czas wykonania zadań osiągnie wartość zawartą w przedziale pomiędzy $\frac{1}{3}\omega(T^i)$ i $\frac{2}{3}\omega(T^i)$.

Główna idea zaproponowanego przez autora niniejszego artykułu algorytmu *A3*, polega również na dokonaniu podziału (oczywiście w wypadku gdy jest to możliwe) każdego ze zbiorów zadań jednoprocesorowych T^i, $1 \leqslant i \leqslant 3$ na dwa podzbiory T_1^i i T_2^i. Jednakże w przeciwieństwie do omówionego uprzednio algorytmu *A2*, algorytm *A3*, szeregujący zadania jednoprocesorowe, stara się zbliżyć na tak małą odległość, jak to tylko jest możliwe, do całkowitego czasu wykonania kompatybilnych (dla danych zadań jednoprocesorowych) zadań dwuprocesorowych. Zatem zadania muszą zostać wybrane do zbioru T_l^i w taki sposób, aby różnica $|\omega(T^{jk}) - \omega(T_l^i)|$, $\{j, k\} = \{1, 2, 3\} \setminus \{i\}$ była tak mała, jak to tylko jest możliwe, dla każdego i.

Po dokonaniu podziału zbioru zadań jednoprocesorowych otrzymuje się co najwyżej dziewięć podzbiorów zadań. Powyższe podzbiory zadań zostają uszeregowane z uwzględnieniem faktu, że wszystkie zadania należące do tego samego podzbioru muszą być wykonywane kolejno po sobie w sposób ciągły, to znaczy nie może pomiędzy nimi wystąpić żadne zadanie należące do innego podzbioru. Spośród wszystkich możliwych planów uszeregowania tych zadań wybierany jest ten, dla którego całkowity czas przebywania procesorów w stanie jałowym jest minimalny.

Podstawowy problem, związany z rozważanym algorytmem szeregowania zadań *A3* polega na znalezieniu takiego podzbioru $T_l^i \subseteq T^i$, który pozwoliłby przybliżyć się w stopniu dostatecznym do wartości $\omega(T^{jk})$. W przypadku ogólnym, znalezienie najlepszego podziału $T_l^i \cup T_2^i = T^i$, polega na rozwiązaniu jednego z dwóch problemów poszukiwania maksymalnej wartości sumy podzbioru T^i. Dla pierwszego z tych problemów wartość progowa wynosi $\omega(T^{jk})$, natomiast dla drugiego $\omega(T^i) - \omega(T^{jk})$. Problem poszukiwania maksymalnej wartości sumy podzbioru polega, dla danej wartości $\boldsymbol{b}$ oraz zbioru $A = \{a1, \ldots a_n\}$ dodatnich liczb całkowitych, na znalezieniu takiego podzbioru $A' \subseteq A$, dla którego suma częściowa $s(A') = \sum_{a_i \in A'} a_i$ osiąga maksymalnie dużą wartość, nie przekraczającą jednakże wartości $\boldsymbol{b}$. Ponieważ problem poszukiwania maksymalnej wartości sumy podzbioru jest problemem NP-trudnym, zaproponowano między innymi w [10], [11] i [12] kilka algorytmów pozwalających na przybliżone rozwiazanie tego problemu. Dla przykładu zaproponowany w [12] algorytm pozwala na znalezienie takiego podzbioru A', że (o ile taki zbiór istnieje) zachodzi następujący warunek $b \geqslant s(A') \geqslant (1-\varepsilon) \cdot b$. W celu znalezienia rozwiązania, przy pomocy rozważanego algorytmu, należy wykonać $O\left(\min\left(\frac{n}{\varepsilon}, n + \frac{\log\frac{1}{\varepsilon}}{\varepsilon^2}\right)\right)$ kroków. Jednakże w dalszej części rozważony zostanie jedynie

prosty, szybki zachłanny algorytm (ang. fast greedy algorithm) znajdowania maksymalnej wartości sumy podzbioru zadań jednoprocesorowych, którego działanie polega na tym, że zadania T^i włączane zostają kolejno do zbioru T^i_1, tak długo jak wartość wyrażenia $|\omega(T^{jk}) - \omega(T^i_1)|$ ulega zmniejszeniu. Pozostałe zadania zostają przypisane do zbioru T^i_2.

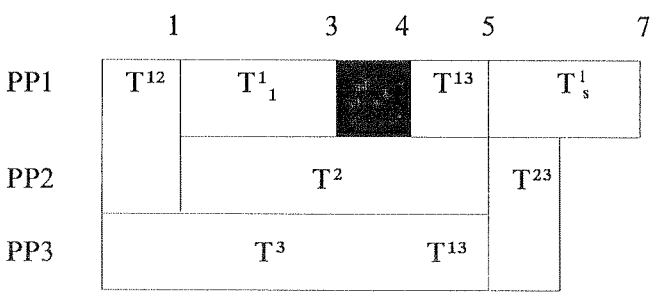
Jak pokazano powyżej, zaproponowany w [5] algorytm *A2* dokonuje również podziału zbioru zadań jednoprocesorowych T^i na dwa podzbiory, jednakże różni się od zaproponowanego przez autora niniejszego artykułu algorytmu *A3*, sposobem w jaki podział ten zostaje przeprowadzony. W celu pokazania różnicy w działaniu algorytmów *A2* i *A3* zostanie rozważony przypadek, w którym każde z zadań posiada jednostkowy czas wykonania. Wartym odnotowania jest fakt, że dla przypadku jednostkowego czasu wykonania każdego z zadań (ang. unit execution time) w [13] został podany algorytm znajdujący optymalne rozwiązanie w czasie wielomianowym. Ponadto dla przypadku jednostkowego czasu wykonania każdego z zadań możliwe jest udowodnienie następującej tezy:

Teza:

Algorytm A3 dla przypadku jednostkowego czasu wykonania każdego z szeregowanych zadań dostarcza rozwiązanie optymalne.

Dowód: Dla każdego ze zbiorów T^i, $1 \leqslant i 3$ zadań jednoprocesorowych, algorytm *A3* będzie wybierał zadania do zbioru T^i do momentu, w którym łączny czas wykonania wybranych zadań będzie dokładnie równy całkowitemu czasowi wykonania odpowiadającego mu kompatybilnego zbioru zadań dwuprocesorowych T^{jk}, przy czym $\{j,k\} = \{1,2,3\}/\{i\}$ (oczywiście jeżeli dokonanie takiego podziału zadań jednoprocesorowych jest możliwe). Optymalny plan szeregowania wszystkich zadań może zostać uzyskany poprzez szeregowania każdego podzbioru zadań T^i równolegle z odpowiadającym mu kompatybilnym zbiorem zadań dwuprocesorowych T^{jk}. Uzyskany plan szeregowania zadań jest optymalny, ponieważ jego czas wykonywania jest równy całkowitemu czasowi pracy co najmniej jednego z procesorów. □

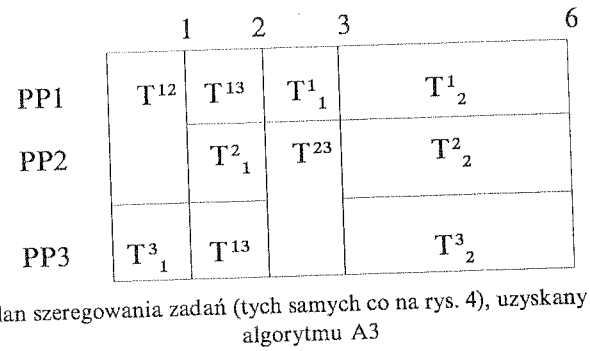
W celu wykazania, że algorytm *A2* nie zawsze jest w stanie dostarczyć optymalnego rozwiązania dla przypadku jednostkowego czasu wykonania każdego z zadań, zostanie rozpatrzony następujący przypadek. Niech dane będą cztery jednoprocesorowe zadania o jednostkowych czasach wykonania dla każdego ze zbiorów T^i,



Rys. 4. Przykładowy plan szeregowania zadań uzyskany w wyniku zastosowania algorytmu A2

$1 \leqslant i \leqslant 3$ oraz dokładnie jedno zadanie dwuprocesorowe o jednostkowym czasie wykonania T^{jk}, przy czym $1 \leqslant j < k \leqslant 3$. Algorytm *A2* dokona podziału każdego ze zbiorów T^i na dwa podzbiory, z których każdy zawierał będzie dokładnie dwa zadania. Dostarczony przez algorytm *A2* plan szeregowania zadań przyjmie postać jak na rys. 4 (dolne indeksy przy podzbiorach zadań: s — ang. short, l — ang. long).

W wyniku zastosowania algorytmu *A2* uzyskano plan szeregowania zadań o długości równej 7 jednostek czasowych, podczas gdy algorytm A3 pozwala na znalezienie optymalnego planu szeregowania rozważanych zadań, posiadającego długość równą 6 jednostek czasowych. Plan taki został przedstawiony na rys. 5.



Rys. 5. Optymalny plan szeregowania zadań (tych samych co na rys. 4), uzyskany w wyniku zastosowania algorytmu A3

4. WYNIKI EKSPERYMENTÓW

W celu porównania efektywności zaproponowanego algorytmu *A3* z algorytmami *A1* i *A2* postanowiono wykonać eksperymenty dla różnych wartości parametrów charakteryzujących zadania, które podlegały procesowi szeregowania. Zastosowanie szybkiego algorytmu zachłannego w celu podziału zbioru zadań jednoprocesorowych pozwoliło na znajdowanie planu szeregowania zadań, podobnie jak w przypadku algorytmów *A1* i *A2*, w czasie wielomianowym $O(n)$. Algorytmy *A1*, *A2* i *A3* zostały przetestowane na zbiorze wygenerowanych w sposób losowy zadań jedno i dwuprocesorowych. Liczba zadań n w kolejnych eksperymentach zwiększona była o 10 i zawierała się w granicach od 10 do 100. Dla Każdej liczby zadań n wygenerowano losowo 1000 różnych przypadków zadań podlegających szeregowaniu. We wszystkich wariantach przeprowadzanych eksperymentów średnie czasy wykonania zadań oraz stosunek średniej częstotliwości występowania zadań wieloprocesorowych do średniej częstotliwości występowania zadań jednoprocesorowych zostały tak dobrane, aby w zdecydowanej większości przypadków był spełniony warunek $\omega(T^i) > \omega(T^{jk})$ dla $1 \leqslant i \leqslant 3$ oraz $\{j,k\} = \{1,2,3\} \setminus \{i\}$, a zatem łączny czas wykonywania zadań jednoprocesorowych danego typu był nie krótszy, niż czas wykonywania odpowiadających im zadań dwuprocesorowych (jest to warunek konieczny do ujawnienia się zalet rozważanych algorytmów). Czas wykonania szeregowanych zadań był zmienną losową posiadającą rozkład normalny. Również dla każdego z procesorów zadania losowane były z jednakowym prawdopodobieństwem. Jako

wielkość służącą porównaniu efektywności badanych algorytmów przyjęto uzyskaną podczas eksperymentów średnią wartość stopnia wykorzystania zasobów systemu. Przy czym stopień wykorzystania zasobów systemu η został zdefiniowany, jako wyrażony w procentach stosunek łącznego czasu T_a aktywnej pracy procesorów PP1, PP2 i PP3 do całkowitego czasu T_c pracy tych procesorów, potrzebnego do wykonania zadań uszeregowanych zgodnie z planem dostarczonym przez testowany algorytm. Czas T_c stanowi sumę łącznego czasu pracy procesorów PP1, PP2 i PP3 i łącznego czasu T_b, w którym procesory te znajdują się w stanie jałowym. Zatem stopień wykorzystania zasobów sytemu może zostać wyrażony wzorem:

$$\eta = \frac{T_a}{T_c} \cdot 100\% = \frac{T_c - T_b}{T_c} \cdot 100\% = \left(1 - \frac{T_b}{T_c}\right) \cdot 100\% \tag{1}$$

Celem przeprowadzonych eksperymentów było dokonanie pomiarów średnich wartości współczynników wykorzystania zasobów systemu η(**A1**), η(**A2**) i η(**A3**), uzyskanych w przypadku zastosowania kolejno algorytmów **A1**, **A2** i **A3**, w zależności od liczby zadań, które podlegały procesowi szeregowania.

W pierwszym z badanych przypadków, prawdopodobieństwo wygenerowania zadania jednoprocesorowego wynosiło $\frac{2}{3}$ i było dwukrotnie większe niż prawdopodobieństwo wygenerowania zadania dwuprocesorowego, które wynosiło $\frac{1}{3}$. Ponadto czasy wykonania zarówno zadań jednoprocesorowych, jak i dwuprocesorowych przybierały z równym prawdopodobieństwem dowolne wartości z przedziału [0...100] jednostek czasowych. Uzyskane podczas przeprowadzonych eksperymentów średnie wartości współczynnika wykorzystania zasobów systemu η zostały zamieszczone w tabl. 1.

Tablica 1

Uzyskane podczas eksperymentów wartości stopnia wykorzystania zasobów systemu (dla dwukrotnie większej częstotliwości występowania zadań jednoprocesorowych niż dwuprocesorowych oraz takich samych średnich czasów wykonywania zadań jednoprocesorowych i dwuprocesorowych)

N	10	20	30	40	50	60	70	80	90	100
η(**A1**)	81,50	83,98	85,21	85,46	85,63	85,60	85,54	85,66	85,56	85,58
η(**A2**)	83,36	85,23	87,25	88,63	89,48	90,13	90,50	90,88	91,18	91.41
η(**A3**)	94,69	95,68	96,72	97,48	97,97	98,34	98,62	98,78	98,96	99,07

W przypadku zmiany prawdopodobieństwa generacji zadania jednoprocesorowego na $\frac{3}{4}$ i dwuprocesorowego na $\frac{1}{4}$ otrzymano wyniki, zamieszczone w tabl. 2.

Następnie prawdopodobieństwo generacji zadania jednoprocesorowego zostało ustalone na $\frac{4}{5}$, a zadania dwuprocesorowego na $\frac{1}{5}$. Uzyskane rezultaty zostały zamieszczone w tabl. 3.

Tablica 2

Uzyskane podczas eksperymentów wartości stopnia wykorzystania zasobów systemu (dla trzykrotnie większej częstotliwości występowania zadań jednoprocesorowych niż dwuprocesorowych oraz takich samych średnich czasów wykonywania zadań jednoprocesorowych i dwuprocesorowych)

N	10	20	30	40	50	60	70	80	90	100
η(A1)	79,04	80,15	80,52	80,29	79,77	79,43	79,12	79,00	78,65	78,60
η(A2)	86,19	87,92	89,91	91,16	92,12	92,77	93,23	93,65	93,96	94,23
η(A3)	96,49	96,81	97,39	97,84	98,26	98,46	98,69	98,87	98,03	99,09

Tablica 3

Uzyskane podczas eksperymentów wartości stopnia wykorzystania zasobów systemu (dla czterokrotnie większej częstotliwości występowania zadań jednoprocesorowych niż dwuprocesorowych oraz takich samych średnich czasów wykonywania zadań jednoprocesorowych i dwuprocesorowych)

N	10	20	30	40	50	60	70	80	90	100
η(A1)	77,37	77,62	77,38	76,68	76,07	75,76	75,31	75,16	74,87	74,73
η(A2)	87,54	88,77	90,51	91,64	92,35	92,83	93,31	93,59	93,87	94,02
η(A3)	97,46	97,40	97,78	98,09	98,38	98,53	98,76	98,88	99,01	99,10

Po zmianie prawdopodobieństwa generacji zadania jednoprocesorowego na $\frac{5}{6}$ i dwuprocesorowego na $\frac{1}{6}$ uzyskano wyniki zamieszczone w tabl. 4.

Tablica 4

Uzyskane podczas eksperymentów wartości stopnia wykorzystania zasobów systemu (dla pięciokrotnie większej częstotliwości występowania zadań jednoprocesorowych niż dwuprocesorowych oraz takich samych średnich czasów wykonywania zadań jednoprocesorowych i dwuprocesorowych)

N	10	20	30	40	50	60	70	80	90	100
η(A1)	76,05	75,72	75,18	74,32	73,68	73,35	72,88	72,69	72,34	72,35
η(A2)	88,49	89,24	90,62	91,52	91,94	92,29	92,59	92,72	92,92	92,93
η(A3)	98,10	97,87	98,04	98,25	98,46	98,57	98,82	98,93	99,04	99,11

W kolejnym etapie eksperymentów ustalono jednakowe prawdopodobieństwa równe 1/2, generacji zarówno zadań jednoprocesorowych, jak i dwuprocesorowych, natomiast zadania jednoprocesorowe posiadały średnio dwukrotnie dłuższy czas wykonywania (czasy wykonywania poszczególnych zadań zawarte były w przedziale [0...200] jednostek czasowych, niż zadania dwuprocesorowe, których czasy wykony-

wania zawierały się w przedziale [0...100] jednostek czasowych. Uzyskane podczas eksperymentów wyniki zamieszczono w tabl. 5.

Tablica 5

Uzyskane podczas eksperymentów wartości stopnia wykorzystania zasobów systemu (dla takiej samej częstotliwości występowania zadań jednoprocesorowych i dwuprocesorowych oraz dwukrotnie dłuższych średnich czasach wykonywania zadań jednoprocesorowych, niż dwuprocesorowych)

N	10	20	30	40	50	60	70	80	90	100
η(A1)	82,00	84,52	85,64	85,90	86,20	86,22	86,20	86,21	86,08	86,09
η(A2)	82,41	84,60	86,73	88,15	89,10	89,64	90,18	90,43	90,85	91,10
η(A3)	82,41	84,60	86,73	88,15	89,10	89,64	90,18	90,43	90,85	91,10

Po zwiększeniu czasów wykonywania zadań jednoprocesorowych, tak aby przyjmowały wartości z przedziału [0...300] jednostek czasowych, przy jednoczesnym pozostawieniu bez zmian wartości czasów wykonywania zadań dwuprocesorowych, uzyskano wyniki zawarte w tabl. 6.

Tablica 6

Uzyskane podczas eksperymentów wartości stopnia wykorzystania zasobów systemu (dla takiej samej częstotliwości występowania zadań jednoprocesorowych i dwuprocesorowych oraz trzykrotnie dłuższych średnich czasach wykonywania zadań jednoprocesorowych, niż dwuprocesorowych)

N	10	20	30	40	50	60	70	80	90	100
η(A1)	80,26	81,44	81,38	81,00	80,59	80,37	80,01	79,82	79,54	79,37
η(A2)	85,34	87,50	89,69	91,35	92,43	92,98	93,53	93,95	94,36	94,68
η(A3)	92,25	93,83	95,48	96,58	97,35	97,71	98,07	98,31	98,44	98,61

Wykonano jeszcze badania dla przypadku, w którym średni czas wykonywania zadania jednoprocesorowego był dwukrotnie dłuższy, niż czas wykonywania zadania dwuprocesorowego oraz zadanie jednoprocesorowe generowane było z dwukrotnie wyższym prawdopodobieństwem, niż zadanie dwuprocesorowe. Uzyskane rezultaty zamieszczono w tabl. 7.

Tablica 7

Uzyskane podczas eksperymentów wartości stopnia wykorzystania zasobów systemu (dla dwukrotnie większej częstotliwości występowania zadań jednoprocesorowych, niż dwuprocesorowych oraz dwukrotnie dłuższych średnich czasach wykonywania zadań jednoprocesorowych, niż dwuprocesorowych)

N	10	20	30	40	50	60	70	80	90	100
η(A1)	78,79	78,35	77,71	77,09	76,50	75,92	75,55	75,39	75,09	74,93
η(A2)	87,89	89,49	91,52	92,67	93,38	93,73	94,16	94,44	94,59	94,76
η(A3)	95,36	95,87	96,77	97,39	97,85	98,17	98,43	98,63	98,77	98,93

Analizując powyższe wyniki widać, że we wszystkich badanych przypadkach najgorsze rezultaty uzyskano stosując najprostszy algorytm *A1*. W rozważanym

przypadku wartość współczynnika wykorzystania zasobów systemu oscyluje pomiędzy 70% a 80% (najlepszy rezultat jaki udało się uzyskać to 86,22%). Dzięki zastosowaniu bardziej rozbudowanego algorytmu *A2* oraz dokonaniu podziału zadań jednoprocesorowych na dwa podzbiory, wartość współczynnika określającego stopień wykorzystania zasobów systemu uległa wyraźnemu zwiększeniu do ponad 90% (najlepszy rezultat, jaki udało się osiągnąć to 94,68%). Zastosowanie nowoopracowanego algorytmu *A3* przyniosło dalszą poprawę stopnia wykorzystania zasobów systemu, gdyż w każdym z badanych przypadków algorytm *A3* okazał się zdecydowanie lepszy od proponowanych w literaturze algorytmów *A1* i *A2*. Najgorszy rezultat uzyskano w przypadku zastosowania algorytmu *A3* to 91,52%, natomiast najlepszy wyniósł aż 99,11%. Zastosowanie w przypadku algorytmu *A3* metody dopasowania długości czasu wykonywania podzbioru zadań jednoprocesorowych do czasu wykonywania odpowiadającego mu zbioru zadań dwuprocesorowych przyniosło zamierzone rezultaty. Ponadto zastosowanie prostego algorytmu, którego działanie polegało na kolejnym doborze zadań jednoprocesorowych, aż do przekroczenia ustalonej wartości, okazało się całkowicie wystarczające. Zatem zastosowanie w tym miejscu bardziej skomplikowanych algorytmów maksymalizacji sumy podzbiorów (posiadających większą złożoność obliczeniową i czas wykonywania) jest nieopłacalne.

Na efektywność rozważanych algorytmów stosunkowo największy wpływ posiada liczba zadań, które podlegały procesowi szeregowania. W przypadku algorytmów *A2* i *A3* można zaobserwować zdecydowany wzrost stopnia wykorzystania zasobów systemu wraz ze zwiększeniem liczby szeregowanych zadań. Na przykład w zamieszczonym w tab. 5 przypadku (dwukrotnie dłuższy średni czas wykonywania zadań jednoprocesorowych, niż dwuprocesorowych, przy identycznych wartościach prawdopodobieństw generacji dwóch typów zadań) dla algorytmu *A2* nastąpił wzrost wartości współczynnika wykorzystania zasobów systemu z 82,41% (dla 10 szeregowanych zadań) do 91,10% (dla 100 zadań). Analogicznie dla algorytmu *A3* zaobserwowano wzrost od wartości 91,52% (dla 10 zadań) do 98,72% (dla 100 zadań). Powodem wystąpienia powyższego zjawiska jest większa możliwość, w przypadku szeregowania większej liczby zadań, dopasowania długości czasów wykonywania podzbiorów zadań jednoprocesorowych do czasów wykonywania odpowiadających im zadań dwuprocesorowych.

W większości badanych przypadków stopień efektywności algorytmu *A1* wykazywał znacznie mniejszą wrażliwość na zmianę liczby szeregowanych zadań, niż miało to miejsce w przypadku algorytmów *A2* i *A3*, przy czym w niektórych przypadkach (tabl. 1, 5) można było zaobserwować wzrost wartości współczynnika wykorzystania zasobów systemu, wraz ze wzrostem liczby szeregowanych zadań. Można było (dla algorytmu *A1*) także zaobserwować przypadki (tabl. 3, 4, 7), w których stopień wykorzystania zasobów systemu wyraźnie malał w przypadku zwiększenia liczby szeregowanych zadań oraz takie (tabl. 2, 6). Dla których stopień wykorzystania zasobów systemu utrzymywał się przez cały czas na mniej więcej stałym poziomie, wykazując jednakże drobne odchyłki o wartościach zarówno dodatnich, jak i ujemnych.

Na zakończenie postanowiono przeprowadzić jeszcze eksperymenty dla przypadku, w którym prawdopodobieństwa generacji zadań zarówno jednoprocesorowych, jak i dwuprocesorowych były identyczne oraz wszystkie zadania posiadały jednakowy średni czas wykonywania. Uzyskane rezultaty zamieszczono w tabl. 8.

Tablica 8

Uzyskane podczas eksperymentów wartości stopnia wykorzystania zasobów systemu (dla identycznej częstotliwości występowania zadań jednoprocesorowych i dwuprocesorowych oraz takich samych średnich czasów wykonywania zadań jednoprocesorowych i dwuprocesorowych)

N	10	20	30	40	50	60	70	80	90	100
η(A1)	83,26	87,11	89,07	90,32	91,25	91,81	92,17	92,85	93,68	93,38
η(A2)	76,88	78,29	79,49	80,22	80,66	80,89	81,10	81,18	81,27	81,35
η(A3)	90,31	92,55	94,22	95,14	95,81	96,17	96,60	96,85	97,06	97,18

Podobnie jak w rozpatrywanych poprzednio przypadkach najlepsze rezultaty przyniosło zastosowanie algorytmu A3, jednakże wyniki najgorsze przyniosło wykorzystanie do szeregowania zadań algorytmu *A2*. Ponieważ, w badanym przypadku, czasy wykonywania zbiorów zadań dwuprocesorowych i odpowiadających im zbiorów zadań jednoprocesorowych były bardzo zbliżone, zastosowana w przypadku algorytmu *A2* metoda podziału zadań jednoprocesorowych na dwa podzbiory okazała się zupełnie nieefektywna i przyniosła rezultaty przeciwne do zamierzonych. Poza tym w przypadku wszystkich algorytmów zaobserwować można wzrost stopnia wykorzystania zasobów systemu wraz ze wzrostem liczby zadań, podlegających procesowi szeregowania.

BIBLIOGRAFIA

1. M. Drozdowski: *Scheduling multiprocessor tasks.* Parallel Computing 25 (1999), pp. 49-61.
2. R. L. Graham, E. L. Lawler: *Optimisation and approximation in deterministic scheduling.* Information Processing Letters 41 (1992), pp. 275-280.
3. J. Blaziewicz, P. Dell' Olmo, M. Drozdowski, M. G. Speranza: *Scheduling multiprocessor tasks on three dedicated processors.* Information Processing Letters 49 (1994), pp. 269-270.
4. P. Dell' Olmo, M. G. Speranza, Z. Tuza: *Efficiency and effectiveness of normal schedules on three dedicated processors,* Discrete Applied Mathematics 61 (1995), pp. 123-133.
5. M. X. Goemans: *An approximation algorithm for scheduling on three dedicated processors.* Discrete Applied Mathematics 61 (1995), pp. 49-59.
6. L. Bianco, P. Dell' Olmo, M. G. Speranza: *Nonpreemptive scheduling of independent tasks with pre-specified processor allocation,* Naval Research Logistics 41 (1994), pp. 959-971.
7. G. Bozoki, J. P. Richard: *A branch-and-bound algorithm for the continuos process task shop scheduling problem.* AIIE Trans. 2 (1970) pp. 246-252.
8. A. K. Amoura: *Problemes d'ordonancement avec communications dans les systemes multiprocesseurs.* Ph. D. Thesis, Orsay, France, Universite de Paris-XI, 1998.
9. A. K. Amoura, E. Bampis, Y. Manoussakis, Z. Tuza: *Scheduling multiprocessir dedicated tasks on three processors.* Technical report 1089, LRI, Orsay, Universite Paris-Sud 1996.

10. D. S. Johnson: *Approximation algorithms for combinatorial problems.* Journal of Computer and System Science 9 (1974), pp. 256-278.
11. T. H. Cormen, C. E. Leiserson, R. L. Rivest: *Introduction to Algorithms.* Combridge, MIT Press, MA, 1990.
12. H. Keller, R. Mansini, U. Pferschy, M. G. Speranza: *An efficient fully polynomial approximation scheme for the subset-sum problem.* University of Brescia, Technical Report, 1997.
13. J. Hoogeveen, S. Van de Velde, B. Veltman: *Complexity of scheduling multiprocessor tasks with pre-specified processor allocation.* Discrete Applied Mathematics 55 (1994), pp. 259-272.

Pracę wykonano w ramach grantu KBN 8T11C 03914 „Analiza i projektowanie systemów czasu rzeczywistego o różnym stopniu rozproszenia".

M. GAJER

THE COMPARISION OF EFFECTIVENESS OF DIFFERENT MULTIPROCESSOR TASKS SCHEDULING ALGORITHMS FOR THREE DEDICATED PROCESSORS

Summary

In the paper the problem of multiprocessor tasks scheduling for the image processing multiprocessor device the TMS320C80 is discussed. The problem of scheduling of independent multiprocessor tasks set for three DSP processors is examinred. The main purpose is to find the optimal schedule of uniprocessor and biprocessor tasks for three dedicated DSP processors so that the total passive time of all the processors should be as short as possible. The proposed in the literature algorithms are discussed and the new solution of the considered problem is presented. The experimental results, the purpose of which is to compare the effectiveness of the proposed tasks scheduling algorithm with two different algorithms known in the literature, are also presented.

Key words: multiprocessor systems, tasks scheduling and allocation, heuristic algorithms.

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI, 2000, 46, z. 1, ss. 35–46

System automatycznej kontroli jakości produkcji tabletek bazujący na wieloprocesorowym układzie wizyjnym

MIROSŁAW GAJER

Katedra Automatyki AGH, Kraków

Otrzymano 1999.12.17
Autoryzowano 2000.03.08

W artykule rozważono możliwości wykorzystania wieloprocesowego układu TMS320C80 do konstrukcji systemu wizyjnego, sprawującego automatyczną kontrolę nad jakością produkcji. Przedstawiono propozycję wykorzystania rozważanego systemu wizyjnego w przemyśle farmaceutycznym do eliminowania z linii produkcyjnej tabletek, posiadających widoczne uszkodzenia mechaniczne. Omówiono szczegółowo wszystkie operacje wykonywane na obrazach badanych tabletek oraz zamieszczono wyniki eksperymentów, mających na celu wykrycie tabletek uszkodzonych. Przeprowadzono także analizę uwarunkowań czasowych, występujących w rozważanym systemie.

Słowa kluczowe: automatyczna kontrola jakości, systemy wizyjne, układy wieloprocesorowe

1. WIZYJNE SYSTEMY AUTOMATYCZNEJ KONTROLI JAKOŚCI

W praktyce przemysłowej często zachodzi konieczność usunięcia z linii produkcyjnej wyrobów pod pewnym względem wadliwych, uszkodzonych, niewłaściwie zmontowanych bądź innych, posiadających pewne niepożądane cechy [1]. We wszystkich przypadkach, w których możliwa jest wzrokowa ocena jakości badanych produktów, przez zatrudnionego na stanowisku kontroli produkcji człowieka, możliwe jest również jego zastąpienie przez automat, wyposażony w system wizyjny, który po wykryciu wady wyrobu, powodował będzie jego usunięcie z linii produkcyjnej, aby produkt ten nie podlegał już dalszym czynnościom, związanym z jego montażem, wykończeniem, paczkowaniem itp. Rozwiązanie takie powoduje eliminację zbędnych czynności (procesów produkcyjnych wykonywanych na wadliwych wyrobach, które i tak w etapie końcowym muszą zostać ostatecznie odrzucone). Fakt

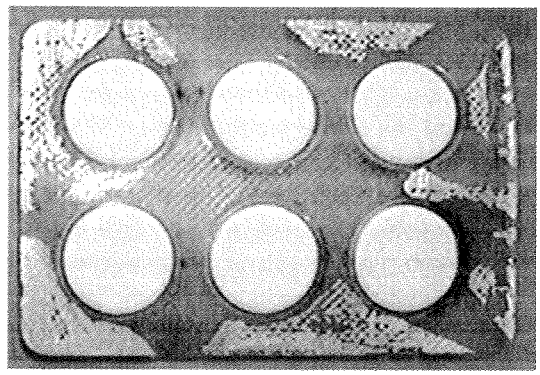
powyższy posiada duże znaczenia, z punktu widzenia rachunku ekonomicznego kosztów produkcji danego wyrobu [2]. Ponadto zastosowanie systemu inspekcyjnego z układem wizyjnym, może przyczynić się także do zmniejszenia liczby, dostarczonych na rynek, wadliwych wyrobów (zwiększenie zaufania klientów do danego produktu, tworzenie odpowiedniego wizerunku firmy). Główną zaletą rozwiązania, polegającego na zastąpieniu człowieka zatrudnionego na stanowisku kontroli jakości produkcji przez komputer wyposażony w system wizyjny, jest to, iż komputer, w przeciwieństwie do człowieka, pracuje zawsze w sposób obiektywny, nie ulega zmęczeniu w trakcie długotrwałego wykonywania cyklicznie powtarzających się, monotonnych czynności. Także decyzje, dotyczące odrzucenia wadliwego wyrobu mogą być podejmowane przez system komputerowy zdecydowanie szybciej, w porównaniu z pracującym na tym stanowisku człowiekiem [3]. Duże tempo produkcji i związana z tym konieczność pracy systemu komputerowego w czasie rzeczywistym, powodują, że istotnym czynnikiem staje się duże wymaganie mocy obliczeniowej, co uzasadnia zastosowanie w rozważanych systemach wizyjnych układów wieloprocesorowych.

W literaturze przedmiotu opisanych zostało wiele systemów automatycznej kontroli jakości produkcji, wykorzystujących układy wizyjne. Dla przykładu można wymienić, opisany w pracy [4] system, nadzorujący proces pakowania elementów elektronicznych. Z kolei praca [5] stanowi przykład systemu, w którym automatyczną kontrolę jakości produkcji połączono z operacją montażu, nadzorowaną przez system wizyjny robota przemysłowego. Znane są także przykłady zastosowań wizyjnych systemów inspekcyjnych w przemyśle tekstylnym [6], maszynowym [7], [8], a nawet w rolnictwie, do sortowania i selekcji sadzonek roślin [9].

W rozdziale niniejszym rozważono możliwości zastosowania wieloprocesorowego układu TMS320C80 firmy Texas Instruments do pracy na stanowisku automatycznej kontroli jakości produkcji tabletek w zakładzie farmaceutycznym. Celem rozważanego systemu wizyjnego jest wykrycie tabletek posiadających widoczne uszkodzenia mechaniczne i ich usunięcia z taśmy produkcyjnej, przed wykonaniem czynności związanych z ich umieszczeniem w plastikowym opakowaniu.

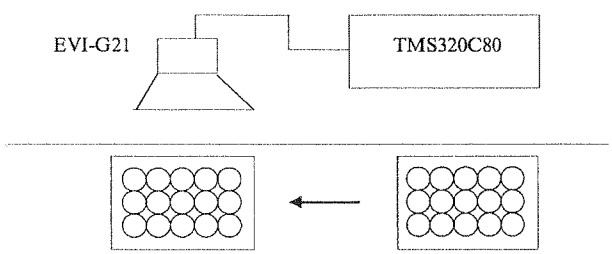
2. ANALIZA OBRAZÓW TABLETEK

W procesie produkcji, tabletki umieszczane są w plastikowym opakowaniu, posiadającym specjalne zagłębienia (służące do umieszczenia w nich tabletek). Następnie na wierzchnią stronę opakowania naklejana jest folia aluminiowa. Zadaniem rozważanego w pracy systemu wizyjnego jest wykrycie tabletek uszkodzonych, w momencie, w którym tabletki zostały już umieszczone w opakowaniu, a na opakowanie nie została naklejona jeszcze folia aluminiowa. W przypadku wykrycia chociaż jednej tabletki uszkodzonej, całe opakowanie musi zostać usunięte z taśmy produkcyjnej, zanim zostanie pokryte folią aluminiową (rys. 1). Przyjęcie takiego rozwiązania zapobiega wykonaniu zbędnych operacji (na uszkodzonych tabletkach) oraz gwarantuje, że wadliwe produkty nie zostaną dostarczone na rynek.



Rys. 1. Opakowanie z tabletkami analizowanymi w systemie wizyjnym

Analiza obrazów tabletek, mająca na celu wykrycie ich uszkodzeń mechanicznych jest procesem wieloetapowym, składającym się z cyklu wykonywanych kolejno po sobie operacji. Na wstępie, pozyskane z kamery obrazy tabletek, poddane muszą zostać procesowi binaryzacji, mającemu na celu ich oddzielenie od tła obrazu. Następnie należy dokonać pomiarów wartości cech, które decydują o stopniu ewentualnego uszkodzenia tabletki. W ostatnim etapie pracy automatycznego systemu kontroli jakości produkcji, na podstawie wykonanych uprzednio pomiarów wartości cech, musi zostać podjęta decyzja o ewentualnym usunięciu z linii produkcyjnej opakowania, zawierającego uszkodzoną tabletkę.



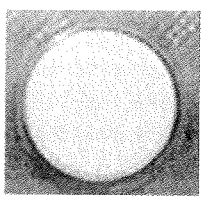
Rys. 2. Stanowisko automatycznej kontroli jakości produkcji z systemem wizyjnym

W dalszych etapach analizy założono, że opakowania z badanymi tabletkami umieszczone zostały na linii produkcyjnej w miejscach posiadających ściśle określone współrzędne przestrzenne. Zatem po umieszczeniu kamery na nieruchomym statywie i ustawieniu jej ogniskowej na stałą wartość, poszczególne tabletki zawsze będą się pojawiać, w z góry zdeterminowanych obszarach pozyskanego z kamery obrazu. Przyjęcie takiego założenia bardzo upraszcza problem segmentacji obrazu, w celu wydzielenia fragmentów reprezentujących pojedyncze tabletki. Sposób pracy rozważanego systemu wizyjnego przedstawiony na rysunku 2.

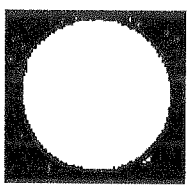
Omawiany system automatycznej kontroli jakości produkcji, składa się z dostarczającej obrazu w standardzie PAL kamery CCD typu EVI-G21 oraz karty

akwizycyjnej Software Development Board z układem wieloprocesorowym TMS320C80 firmy Texas Instruments.

Pozyskane z kamery obrazy tabletek poddawane są na wstępie binaryzacji, której wartość progowa została ustalona eksperymentalnie na 175 poziomie jasności obrazu (obraz posiadał 256 poziomów jasności, a stanowisko było oświetlone źródłem światła o stałym natężeniu). Rozwiązanie takie pozwoliło na wydzielenie fragmentów obrazów, przedstawiających tabletki, które po procesie binaryzacji uzyskały kolor biały (255 poziom jasności obrazu), z ich tła (plastikowe opakowanie), które po binaryzacji uzyskało kolor czarny (zerowy poziom jasności obrazu). Jednakże w procesie produkcyjnym zapewnienie stałego poziomu oświetlenia może sprawiać pewne problemy ze względu na proces starzenia się źródła światła, zabrudzenie oprawy itp. W takim przypadku konieczną może się okazać okresowa kalibracja systemu (wybór wartości progowej), oparta na pomiarze histogramu obrazu. Obrazy tabletki przed i po binaryzacji zostały zamieszczone, odpowiednio na rysunkach 3 i 4.



Rys. 3. Obraz tabletki przed binaryzacją



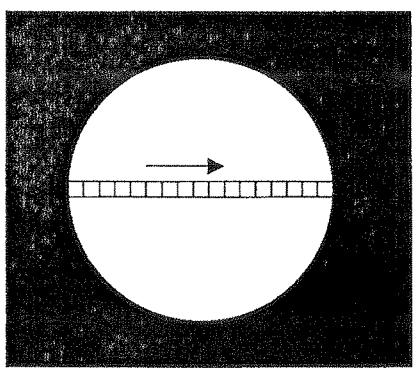
Rys. 4. Obraz po binaryzacji

Uszkodzenia mechaniczne badanych tabletek, polegać mogą albo na braku pewnego fragmentu tabletki, który został od niej odłamany, albo też tabletka może ulec pęknięciu. W celu wykrycia wystąpienia pierwszej z rozważanych możliwości uszkodzenia tabletki, dla każdej z analizowanych tabletek, należy dokonać pomiaru liczby reprezentujących ją (po binaryzacji) białych pikseli. Następnie mierzona w ten sposób liczba białych pikseli musi zostać porównana z wartością progową, reprezentującą minimalną liczbę białych pikseli, charakterystyczną dla nieuszkodzonej tabletki. Jeżeli uzyskana liczba białych pikseli będzie mniejsza, niż przyjęta wartość progowa, wówczas fakt ten będzie świadczył o braku pewnego fragmentu tabletki (jej pole powierzchni będzie pomniejszone,w stosunku do tabletki prawidłowej). Wartość progowa, określająca liczbę białych pikseli dla nieuszkodzonych tabletek została ustalona eksperymentalnie (z pewnym marginesem bezpieczeństwa) na 3450 pikseli.

W celu wykrycia uszkodzeń tabletek (polegających na ich pęknięciu), należy analizować zbinaryzowane obrazy, w celu detekcji czarnych pikseli na białym obszarze reprezentującym tabletkę. Zaproponowano rozwiązanie, polegające na poszukiwaniu czarnych pikseli wzdłuż linii, przebiegających w czterech różnych kierunkach.

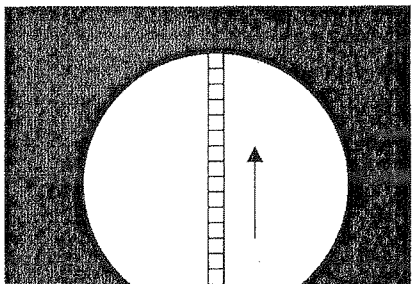
Analizowane są obrazy w formie kwadratu o boku równym 80 pikseli, przedstawiające pojedynczą tabletkę na tle fragmentu jej plastikowego opakowania. Następnie, wybierana jest pozioma linia, dzieląca analizowany obraz na dwie równe

połowy. Wzdłuż tej linii, analizowane są kolejno piksele rozważanego fragmentu obrazu (w tym należy spodziewać się wystąpienia obrazu tabletki), w kierunku od lewej do prawej krawędzi. Wykrycie pierwszego białego piksela powoduje inkrementację specjalnego licznika białych pikseli (licznik przed rozpoczęciem analizy musi zostać wyzerowany). Jeżeli następny (sąsiedni) piksel obrazu posiada kolor biały wartość licznika zostaje również zwiększona o jeden. System pracuje w przedstawiony sposób, aż do momentu napotkania pierwszego czarnego piksela, który świadczy o tym, że albo został osiągnięty przeciwległy brzeg tabletki (pierwszy czarny piksel opakowania), albo też natrafiono na obszar pęknięcia tabletki. Po detekcji pierwszego czarnego piksela (występującego po ciągu białych pikseli), proces analizy obrazu ulega zakończeniu, a aktualna wartość licznika jest porównywana z wartością progową, odpowiadającą typowej średnicy nieuszkodzonej tabletki. Powyższa wartość progowa, została wyznaczona eksperymentalnie, poprzez pomiar średnicy prawidłowych tabletek (Z pewnym marginesem bezpieczeństwa), na 68 pikseli (rozrzut średnic poszczególnych tabletek był nie większy niż 3 piksele). Sposób analizy pikseli zbinaryzowanego obrazu tabletki został przedstawiony na rysunku 5.



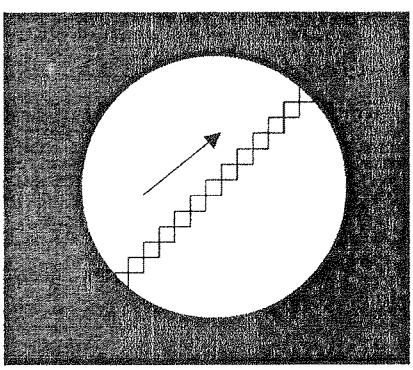
Rys. 5. Analiza pikseli obrazu tabletki wzdłuż linii poziomej

Analogicznie, wybierana jest pionowa linia, dzieląca obraz tabletki na dwie równe połowy. Przy czym założono, że opakowanie z tabletkami znajduje się na linii produkcyjnej zawsze w tym samym miejscu, a ewentualny błąd pozycjonowania jest zaniedbywalny). Sposób analizy, leżących wzdłuż tej pikseli jest identyczny, jak w przypadku linii poziomej. W tym wypadku piksele są analizowane kolejno w kierunku od dolnej do górnej krawędzi obrazu. Uzyskana w wyniku przeprowadzenia pomiaru wartość (inkrementowanego po wykryciu białego piksela licznika), powinna być równa (oczywiście w granicach dopuszczalnej tolerancji) średnicy nieuszkodzonej tabletki (wyrażonej w pikselach). Jeżeli powyższy warunek nie jest spełniony, to fakt ten świadczy o pojawieniu się na białym obrazie tabletki czarnych pikseli, które odpowiadać mogą szczelinie, powstałej w skutek pęknięcia. Sposób analizy obrazu wzdłuż linii pionowej, przedstawiono na rysunku 6.

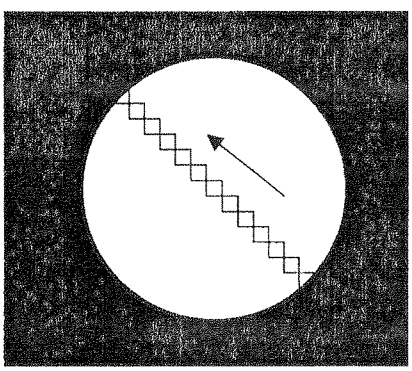


Rys. 6. Analiza pikseli obrazu tabletki wzdłuż linii pionowej

W końcowym etapie analizowane są, w sposób analogiczny, piksele położone wzdłuż dwóch przekątnych fragmentu obrazu, zawierającego pojedynczą tabletkę. Sposób analizy przedstawiono na rysunkach 7 i 8.



Rys. 7. Analiza pikseli obrazu wzdłuż pierwszej przekątnej



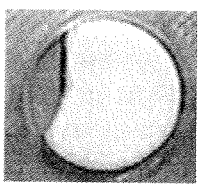
Rys. 8. Analiza pikseli obrazu wzdłuż drugiej przekątnej

W tym przypadku również uzyskanie wyniku pomiaru średnicy tabletki, mniejszego od oprzyjętej uprzednio wartości progowej, świadczy o wystąpieniu na białym obrazie tabletki czarnych pikseli, odpowiadających szczelinie, powstałej w skutek pęknięcia.

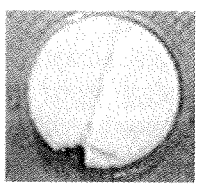
Omówione powyżej operacje, polegające na pomiarze (wyrażonym w pikselach) pola powierzchni tabletki oraz jej średnicy (w czterech różnych kierunkach), muszą zostać powtórzone (z osobna) dla wszystkich tabletek, wchodzących w skład testowanego opakowania. Wykrycie przynajmniej jednej uszkodzonej tabletki, dla której zaproponowane testy nie przebiegły pomyślnie, stanowi podstawę do usunięcia z linii produkcyjnej całego opakowania, przed naklejeniem na niego folii aluminiowej. Rozwiązanie takie gwarantuje, że wadliwy produkt nie trafi do konsumenta.

3. PRZYKŁADY OBRAZÓW

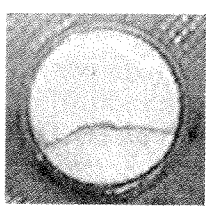
W celu stwierdzenia poprawności zaproponowanej metody analizy obrazów tabletek, przeprowadzono liczne eksperymenty, zarówno z tabletkami normalnymi, jak i uszkodzonymi. Podczas eksperymentów analizie poddano łącznie 50 tabletek, w tym 10 prawidłowych i 40 uszkodzonych(o różnym stopniu i sposobie uszkodzenia). Wszystkie z analizowanych, prawidłowych tabletek przeszły pomyślnie każdy z rodzajów testów. Natomiast wszystkie z uszkodzonych tabletek zostały przez wizyjny system kontroli jakości produkcji odrzucone. Uzyskane wyniki eksperymentów, przemawiają na korzyść (oraz świadczą o dużej skuteczności) zaproponowanej metody wykrywania uszkodzeń tabletek. Poniżej na rysunkach 9÷11 zamieszczono, przykładowe (pozyskane z kamery) obrazy uszkodzonych tabletek. W pierwszym przypadku (rysunek 9) brakuje fragmentu tabletki, natomiast w pozostałych (rysunki 10 i 11) tabletka jest pęknięta.



Rys. 9. Obraz tabletki uszkodzonej (brakujący fragment)



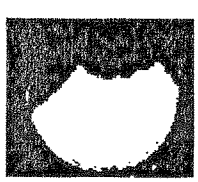
Rys. 10. Obraz tabletki uszkodzonej (pęknięcie)



Rys. 11. Obraz tabletki uszkodzonej (pęknięcie)

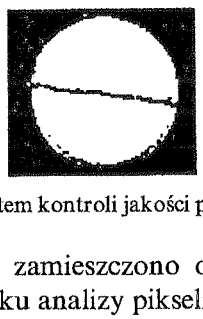
W celu ilustracji sposobu pracy rozważanego systemu wizyjnego, opisano także kilka przypadków wykrycia wadliwej tabletki, dla różnych sposobów jej uszkodzenia.

Na rysunku 12 pokazano tabletkę (obraz po binaryzacji), której fragment uległ odłamaniu. Tabletka ta została odrzucona już po przeprowadzeniu pierwszego tekstu, związanego z pomiarem jej pola powierzchni.



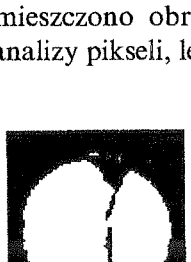
Rys. 12. Tabletka odrzucona przez system kontroli jakości produkcji (pomiar pola powierzchni)

Z kolei na rysunku 13 zamieszczono obraz tabletki (po binaryzacji), która posiadała pęknięcie w kierunku poziomym. W tym wypadku o odrzuceniu tabletki zadecydował negatywny wynik testu, polegającego na analizie kolejnych pikseli linii pionowej, dzielącej obraz na dwie równe części.



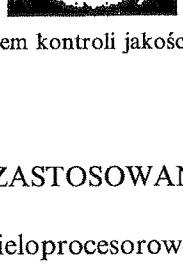
Rys. 13. Tabletka odrzucona przez system kontroli jakości produkcji (detekcja pęknięcia na linii poziomej)

Natomiast na rysunku 14 zamieszczono obraz tabletki pękniętej w kierunku pionowym, odrzuconej w wyniku analizy pikseli, leżących na poziomej linii, dzielącej obraz na dwie połowy.



Rys. 14. Tabletka odrzucona przez system kontroli jakości produkcji (detekcja pęknięcia na linii pionowej)

Z kolei na rysunku 15 zamieszczono zbinaryzowany obraz tabletki z pęknięciem zlokalizowanym w prawym górnym rogu obrazu. W tym przypadku o odrzuceniu tabletki przez system kontroli jakości, zadedydował negatywny wynik testu, podczas analizy kolejnych pikseli, położonych na przekątnej obrazu.



Rys. 15. Tabletka odrzucona przez system kontroli jakości produkcji (detekcja pęknięcia na przekątnej)

4. OPIS ZASTOSOWANEGO SPRZĘTU

Pojedynczy procesor DSP wieloprocesorowego układu TMS320C80 potrzebował na wykonanie pomiaru (wyrażonego w pikselach) pola powierzchni tabletki $2{,}54*10^{-3}$ s. W tym sam proces binaryzacji obrazu zajmował $1{,}28*10^{-3}$ s. Podobnie,

analiza pikseli wzdłuż pionowej lub poziomej linii obrazu wymagała czasu równego $3{,}17*10^{-5}$ s. Wykonanie analogicznej operacji dla przekątnej obrazu wymagało natomiast $4{,}46*10^{-5}$ s. Zatem całkowity czas obliczeń związanych z analizą obrazu pojedynczej tabletki, wynosił $269*10^{-3}$ s.

Wykorzystanie możliwości obliczeniowych wieloprocesorowego układu TMS320C80 polegało na implementacji architektury typu SIMD, gdzie każdy z czterech użytych procesorów DSP (są to procesory równoległe wchodzące w skład systemu TMS320C80), wykonywał operacje związane z analizą obrazu innej tabletki. W rozważanym systemie równoległym wartość współczynnika przyspieszenia obliczeń (ang. speedup) wyniosła 3,16. Jest to wartość znacznie mniejsza od maksymalnej dopuszczalnej teoretycznie wartości tego współczynnika dla architektury tupu SIMD, która w przypadku czterech procesorów wynosi 4,00. (Współczynnik przyspieszenia obliczeń wyznaczony jest jako stosunek czasu wykonywanych przez system jednoprocesorowy do czasu realizacji tych samych obliczeń przez system wieloprocesorowy). Fakt powyższy spowodowany został występowaniem, podczas pracy wieloprocesorowego układu TMS320C80, znacznych narzutów komunikacyjnych, związanych z oczekiwaniem poszczególnych jednostek obliczeniowych na przydział prawa dostępu do portu komunikacyjnego, umożliwiającego wymianę danych (pikseli pozyskanych z kamery obrazów) oraz kodów instrukcji (podczas odświeżania podręcznych pamięci programu) z zewnętrzną pamięcią systemu. Wszystkie procesory równoległe DSP posiadają podręczne pamięci instrukcji o wielkości 2 KB, które są odświeżane (uzupełniany jest brakujący aktualnie blok instrukcji) w sposób całkowicie automatyczny, na którego przebieg programista nie ma żadnego wpływu. Pamięci podręczne są w stanie całkowicie pomieścić jedynie niewielkie kody programów, składające się maksymalnie z 512 słów 32-bitowych, zatem zwartość kodu programu (możliwa do uzyskania poprzez programowanie bezpośrednio w języku asemblera) jest tutaj bardzo istotna.

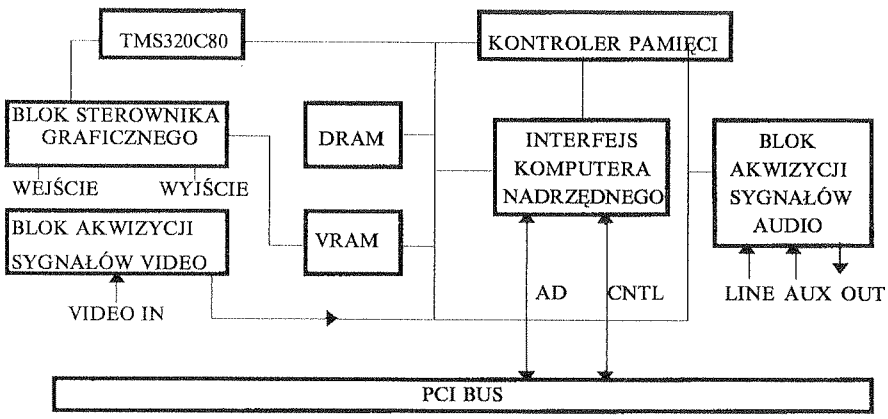
W skład wieloprocesorowego układu TMS320C80 MVP (ang. Multimedia Video Processor) wchodzi 32-bitowy procesor nadrzędny MP, o architekturze typu RISC i mocy obliczeniowej 100 MFLOPS oraz cztery równoległe 32-bitowych, stałoprzecinkowe procesorów DSP (PP0, PP1, PP2 i PP3). Układ TMS320C80 zawiera ponadto kontroler TC, nadzorujący przebieg procesu wymiany danych pomiędzy pamięcią wewnętrzną układu, a urządzeniami peryferyjnymi, kontroler VC sterujący procesem akwizycji i wyświetlania sygnałów wizyjnych, 50 KB wewnętrznej statycznej pamięci RAM oraz układ przełącznicy, pozwalający na automatyczne zestawienie połączeń pomiędzy portami komunikacyjnymi poszczególnych procesorów, a blokami wspólnej (dla wszystkich jednostek układu) wewnętrznej pamięci.

Krzemowa struktura układu TMS320C80 zawiera około czterech milionów tranzystorów CMOS, a moc obliczeniowa, zrealizowanej w oparciu o niego architektury typu MIMD, szacowana jest na ponad dwa miliardy elementarnych operacji typu RISC, wykonywanych w ciągu jednej sekundy. Wieloprocesorowy układ TMS320C80 został zintegrowany z układami akwizycji danych wizyjnych, w ramach współpracującej z komputerem PC karty o nazwie Software Development Board. Karta Software Development Board posiada sześć podstawowych jednostek, stanowiących odrębne bloki funkcjonalne:

- układ wieloprocesorowy TMS320C80 taktowany zegarem 40 MHz

- blok współpracy z komputerem nadrzędnym (ang. host interface block)
- kontroler pamięci MC (ang. memory controller) oraz układy pamięci DRAM/VRAM
- blok akwizycji sygnałów wizyjnych (ang. video capture block)
- blok sterownika graficznego (ang. video display block)
- blok akwizycji i odtwarzania sygnałów audio (ang, audio capture and playback block)

Architektura systemu Software Development Board została schematycznie przedstawiona na rysunku 16.



Rys. 16. Architektura systemu Software Development Board

Ostatecznie czas potrzebny na analizę obrazu jednej tabletki został zredukowany w wyniku paralelizacji obliczeń do $8{,}51*10^{-4}$ s (w porównaniu z czasem analizy obrazu tabletki na pojedynczym procesorze $2{,}69*10^{-3}$), co oznacza, że system jest w stanie w ciągu jednej sekundy przeprowadzić analizę około 1175 tabletek. Uzyskanie takiego wyniku, pozwala na realizację wymienionych powyższych operacji w czasie rzeczywistym, w warunkach pracy na linii produkcji tabletek w zakładzie farmaceutycznym.

Karta Software Development Board posiada możliwość komunikacji z komputerem nadrzędnym typu PC poprzez zgłoszenie przerwania na magistrali PCI. Powyższy sposób komunikacji może zostać zastosowany do zasygnalizowania komputerowi nadrzędnemu konieczności odrzucenia wadliwej tabletki.

5. DYSKUSJA WYNIKÓW

Możliwość przeprowadzenia analizy obrazów tak dużej liczby tabletek w ciągu jednej sekundy, powoduje, że ograniczenia nałożone na pracę systemu wynikają raczej z szybkości procesu akwizycji danych, niż z mocy obliczeniowej zastosowanych jednostek przetwarzających. W przypadku karty Software Development Board, pozyskiwane są obrazy o rozdzielczości 768 na 576 pikseli. Które mogą zawierać

maksymalnie obrazy 63 tabletek (pojedyncza tabletka zajmuje obszar obrazu o wymiarach 80 na 80 pikseli). Kamera dostarcza w ciągu sekundy 25 obrazów, czyli w ciągu jednej sekundy mogą podlegać analizie obrazy najwyżej 1575 tabletek. Dalsze zwiększenie wydajności systemu automatycznej kontroli jakości produkcji można osiągnąć poprzez zwiększenie mocy obliczeniowej systemu analizy obrazów (zastosowanie np. nowszej wersji układu TMS320C80, taktowanej sygnałem o częstotliwości 60 MHz, zamiast 40 MHz – w przypadku starszej wersji tego układu) oraz zwiększenie szybkości akwizycji danych. W tym drugim przypadku jednakże pojawia się konieczność zastosowania bardzo szybkich, specjalizowanych układów akwizycji obrazów (ang. *frame grabber*), dostarczających obrazy o zwiększonej rozdzielczości oraz zapewniających dużą szybkość odświeżania obrazów. Jednakże układy takie wymagają współpracy również ze specjalizowanymi kamerami, dostarczającymi kolejnych kadrów obrazu ze zwiększoną częstotliwością, w stosunku do zastosowanej podczas eksperymentów kamery pracującej w standardzie PAL.

Uzyskana szybkość analizy obrazów tabletek (poniżej tysięcznej części sekundy na jedną tabletkę) jest i tak imponująca w porównaniu z innymi rozwiązaniami tego typu systemów, znanymi z literatury przedmiotu. Dla przykładu w pracy [10] omówiono system wizyjny służący do lokalizacji obiektów. W systemie tym analizowane były obrazy o wymiarach 128 na 128 pikseli, przy czym na lokalizację obiektu potrzeba było około 100 milisekund. Z kolei w pracy [11] podano przykład systemu wizyjnego, którego celem było rozpoznawania montowanych elementów bloków silników spalinowych. Czas potrzebny na rozpoznanie pojedynczego elementu wynosił w tym wypadku aż 13 sekund.

Zatem uzyskany poziom szybkości przetwarzania obrazów tabletek wydaje się być wystarczający dla większości praktycznych zastosowań i użycie droższego sprzętu jest raczej nie uzasadnione ekonomicznie.

Pewien problem w procesie analizy obrazów tabletek mogą stanowić punktowe zakłócenia, objawiające się w postaci pojedynczych białych pikseli umieszczonych na czarnym tle obrazu. W przypadku wystąpienia nieprawidłowości w funkcjonowaniu systemu, należy zastosować dodatkowo operację erozji binarnej, która skutecznie eliminuje zakłócenia mające formę izolowanych pikseli bądź niewielkich zgrupowań takich pikseli. Operacja erozji binarnej przeprowadzona jest w ten sposób, że dla każdego białego piksela sprawdzane jest czy wszystkie piksele sąsiednie (jest czterech najbliższych sąsiadów z góry, z dołu, z prawej i z lewej strony) także mają kolor biały, jeżeli chociaż jeden z tych pikseli ma kolor czarny, wówczas badanemu aktualnie pikselowi również należy przypisać kolor czarny.

BIBLIOGRAFIA

1. L. M a c a i r e, J. G. P o s t a i r e: *Automated visual inspection of galvanised and metallic strips.* 1993 CompEuro Proceeding, Computers in Design, Manufacturing, and production, May 24–27, 1993, Paris-Evry, France, pp. 8–15

2. T. Thomas, M. Cattoen: *System for defects on material.* 1993 CompEuro Proceedings, Computers in Design, Manufacturing, May 24–27, 1993, Paris-Evry, France, pp. 16–22
3. B. Mehenni, M. A. Wahab: APRIS: *Automatic pattern recognition and inspection system.* 1993 CompEuro Proceedings, Computers in Design, Manufacturing, May 24–27, 1993, Paris-Evry, France, pp. 23–28
4. Berger, P. Dunbar, C. Roberts: *Machine vision character recognition in the electronic packaging industry, three case studies.* VISION–85 – Conference proceedings. Machine Vision Association of SME, North Holland, Detroit, Michigen, 1985
5. F. J. M. Van der Heijden: *Assembly of small components by a vision controlled robot using Philips vision system PAPS.* Proceedings of the 5th International Conference on Robot Vision and Sensory Controls. North Holland, Amsterdam, 1985
6. M. Lefley, A. A. Hashim, J. D. Spencer: *Vision for tetile automation.* Proceeding of the 5th International Conference on Robot Vision and Sensory Controls. North Holland, Amsterdam, 1985
7. D. C. Burgess: *Industrial image processing at speed. Proceedings of the 5th International Conference on Robot Vision and Sensory Controls.* North Holland, Amsterdam, 1985
8. E. Schmidberger, R. Ahler: *Quality control with a robot-guided lelectro-optical sensor. Robot Vision and Sensory Controls.* Proceedings of the 4th International Conference on Robot Vision and Sensory Controls, London, 1984, pp. 27–36
9. H. Delepanque, P. Bonnet, J. G. Postaire: *An intelligent robotic system for in vitro plantlet production, Proceedings of the 5th International Conference on Robot Vision and Sensory Controls.* North Holland, Amsterdam, 1985
10. J. Amat, A. Casals, V. Llario: *Location of work-pieces and guidance of industrial robots with a vision systems, Robot Vision and Sensory Controls.* Proceedings of the 4 th International Conference RoViSec, London, 1984, pp. 223–230
11. H. Linderman, H. Hollek: *Supporting sensors for vision guided robots.* Proceedings of the 5th International Conference on Robot Vision and Sensory Controls. North Holland, Amsterdam, 1985

Pracę wykonano w ramach grantu KBN 8T11C 03914 „Analiza i projektowanie systemów czasu rzeczywistego o różnym stopniu rozproszenia".

M. GAJER

THE AUTOMATIC QUALITY CONTROL SYSTEM BASED ON A MULTIPROCESSOR VISION SYSTEM

Summary

In the paper the possibility of the usage of the TMS320C80 multiprocessor system for the construction of a vision system for the automatic production quality control was discussed. The proposed vision system is to be used in the pharmacy factory to eliminate medicines with some mechanical defects. All the image processing operations were thoroughly described and some experimental results were also presented. The time constraints occurring in the system were also discussed.

Key words: automatic quality control, vision systems, multiprocessor systems.

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI, 2000, 46, z. 1, ss. 47–62

Zaawansowane modele probabilistyczne sygnałów w kanałach z zanikami

KRYSTYNA NOGA

Wyższa Szkoła Morska, Katedra Automatyki Okrętowej

Otrzymano 1999.04.27
Autoryzowano 1999.11.10

Podczas transmisji radiowej większość sygnałów odbieranych podlega losowemu uzmiennieniu, spowodowane to jest występowaniem w kanale transmisyjnym zakłóceń multiplikatywnych i addytywnych. Mimo dużej różnorodności emitowanych sygnałów oraz mechanizmów i ośrodków propagacyjnych możliwe jest przedstawienie opisu zjawisk losowych przy pomocy kilku modeli probabilistycznych. Do najczęściej stosowanych modeli należą rozkłady: Rice'a, Rayleigha i Nakagamiego. Często również stosowane są roskłady: jednostronny normalny, Hoyta, trójparametrowy i czteroparametrowy. Modele probabilistyczne opisują powolne fluktuacje obwiedni sygnałów radiowych. W kanałach radiowych, na skutek odbić sygnałów od powierzchni, od warstw zjonizowanych w jonosferze lub od niejednorodności w troposferze, występuje rozpraszanie. Propagacja fal radiowych jest więc zjawiskiem o dużej złożoności. Podstawowy model sygnału radiowego odbieranego z powolnymi zanikami jest dwuparametrowy, np. rozkład Rice'a, Nakagamiego. Dla sygnału całkowicie rozproszonego z zanikami wystarczy rozkład jednoparametrowy, np. Rayleigha. W artykule przedstawiono najczęściej stosowane rozkłady obwiedni sygnału użytecznego oraz powiązania pomiędzy tymi rozkładami.

Słowa kluczowe: kanał radiokomunikacyjny, kanał z zanikami, rozkład prowdopodobieństwa obwiedni.

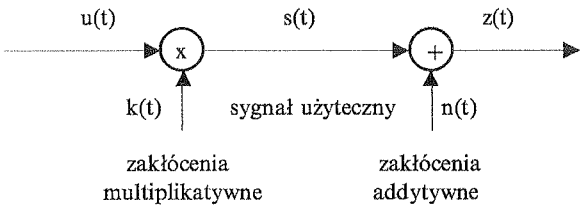
1. WSTĘP

Sygnał nadany u(t) podczas przesyłania przez kanał radiokomunikacyjny podlega losowym zakłóceniom multiplikatywnym k(t) oraz addytywnym n(t). Działanie zakłóceń multiplikatywnych traktujemy jako przypadkowe tłumienie sygnału. Zjawisko to jest określane zazwyczaj mianem zaników. Zaniki są cechą charakterystyczną przede wszystkim radiokomunikacji wykorzystującej propagację jonosferyczna i troposferyczną.

Dla analogowego modelu kanału propagacyjnego (rys. 1) sygnały z(t) odbierane przez odbiornik należy traktować jako sumę sygnału użytecznego s(t) oraz zakłócenia addytywnego n(t), czyli

$$z(t)=s(t)+n(t)=k(t)\cdot u(t)+n(t)=r(t)\cos[\omega_o t+\phi(t)]+n(t) \tag{1}$$

gdzie $r(t)\geqslant 0$ jest obwiednią sygnału użytecznego, $\phi(t)$ jest fazą chwilową sygnału użytecznego. Należy wyraźnie zaznaczyć, że obwiednia sygnału użytecznego zależy wyłącznie od zaników tylko w przypadku sygnałów z modulacją kątową.



Rys. 1. Analogowy model kanału radiokomunikacyjnego

Podstawową cechą sygnałów użytecznych oraz zakłóceń jest losowość ich przebiegów w czasie. Dlatego też często stosowanym modelem sygnałów i zakłóceń jest funkcja losowa. Jej właściwości probabilistyczne określa się na podstawie procesów losowych lub zmiennych losowych. Przy opisie modeli często rezygnuje się z ogólności, np. dzięki założeniu stacjonarności wyznaczanie cech probabilistycznych jest prostsze. Dla analogowego modelu kanału radiokomunikacyjnego sygnały oraz zakłócenia można traktować jako realizacje odpowiednich procesów stochastycznych. W dalszych rozważaniach zakładamy stacjonarność wszystkich sygnałów i zakłóceń występujących w analogowym modelu kanału radiokomunikacyjnego. Jeżeli założymy lokalną stacjonarność sygnałów i zakłóceń oraz powolną zmienność zaników w porównaniu z czasem trwania pojedynczego sygnału elementarnego, to w przekształceniach matematycznych można zamiast procesami stochastycznymi posługiwać się zmiennymi losowymi. Opis zmiennych losowych może być utworzony na podstawie znajomości rozkładów prawdopodobieństw jednej lub dwóch zmiennych. Z analizy literatury wynika, że rozważania bardzo często ograniacza się do rozkładu prawdopodobieństwa rzędu pierwszego i drugiego oraz do momentów do rzędu drugiego włącznie [11].

W dalszej części pracy zostaną przedstawione probabilistyczne modele sygnałów w kanałach z zanikami oraz powiązania pomiędzy tymi modelami. W rozdziale drugim zostanie przedstawiony rozkład Rice'a, który, jak wynika z analizy literatury, jest najczęściej stosowany do modelowania zaników. Dlatego też dla tego rozkładu zostaną podane rozkłady łączne, rozkład obwiedni i fazy oraz moment k-tego rzędu. Również bardzo często do opisu zaników stosuje się rozkłady Rayleigha oraz Nakagamiego [1, 4, 8, 9, 10, 11], które zostaną przedstawione w rozdziałach odpowiednio trzecim i ósmym. W rozdziałach od czwartego do siódmego zostaną omówione rozkłady odpowiednio: jednostronny Gaussa, Hoyta, trójparametrowy i najbardziej uniwersalny — czteroparametrowy, które są również stosowane do modelowania zaników.

2. ROZKŁAD RICE'A

Niech sygnał użyteczny s(t) jest określony zależnością

$$s(t)=c(t)+w_{sz}(t)=[x_c(t)[x_c(t)+x_{sz}(t)]\,x_{sz}(t)]\cos\omega_o t-[y_c(t)+y_{sz}(t)]\sin\omega_o t \tag{2}$$

gdzie: c(t) — zdeterminowany (regularny) sygnał harmoniczny,
$w_{sz}(t)$ — wąskopasmowe zakłócenie (szum) o normalnym rozkładzie wartości chwilowej i zerowej wartości średniej,
ω_o — średnia pulsacja (częstotliwość kątowa),
$x_c(t)$, $y_c(t)$ — składowa synfazowa i kwadraturowa sygnału c(t),
$x_{sz}(t)$, $y_{sz}(t)$ — składowa synfazowa i kwadraturowa zakłócenia $w_{sz}(t)$.

Dla z założenia stacjonarnego szumu $w_{sz}(t)$ składowe $x_{sz}(t)$, $y_{sz}(t)$ są normalne, nieskolerowane i mają zerowe wartości średnie oraz jednakowe wariancje (moce średnie) $\sigma^2_{x_{sz}}=\sigma^2_{y_{sz}}=\sigma^2_{sz}$.

Oznaczmy składową synfazową i kwadraturową wypadkowego sygnału użytecznego jako

$$x(t)=x_c(t)+x_{sz}(t) \tag{3a}$$

$$y(t)=y_c(t)+y_{sz}(t) \tag{b}$$

Wartości średnie tych składowych są równe ich odpowiednim składnikom niezaszumionym

$$E[x(t)]=x_c(t)=m_x(t) \tag{4a}$$

$$E[y(t)]=y_c(t)=m_y(t) \tag{4b}$$

gdzie E jest operatorem uśredniania zmiennych losowych [x(t)] i [y(t)], których realizacjami są x(t) i y(t).

Przebiegi x(t), y(t), $x_c(t)$ i $y_c(t)$ w dowolnej chwili czasu t_0 są liczbami — realizacjami odpowiednich zmiennych losowych. Dlatego też, w celu skrócenia zapisów, w dalej omawianych rozkładach wartości chwilowych argument czasu będzie opuszczany. Wówczas rozkład łączny składowych kwadraturowych x(t), y(t) sygnału użytecznego dla dowolnej chwili t_o wyraża się wzorem

$$p(x,y)=\frac{1}{2\pi\sigma^2_{sz}}\exp\left[-\frac{(x-m_x)^2}{2\sigma^2_{sz}}-\frac{(y-m_y)^2}{2\sigma^2_{sz}}\right] \tag{5}$$

Sygnał użyteczny s(t), określony wzorem (2), można również przedstawić jako

$$s(t)=r(t)\cos[\omega_o t+\phi(t)] \tag{6a}$$

gdzie: $r(t)=\sqrt{x^2(t)+y^2(t)}$ jest obwiednią sygnału użytecznego, (6b)

$\phi(t)=\operatorname{arctg}\dfrac{y(t)}{x(t)}$ jest fazą chwilową sygnału użytecznego. (6c)

Wówczas łączny rozkład prawdopodobieństwa obwiedni r(t) (czyli amplitudy chwilowej) oraz fazy chwilowej $\phi(t)$ przedstawia się następująco

$$p(r,\phi)=\frac{r}{2\pi\sigma_{sz}^2}\exp\left(-\frac{(r\cos\phi-m_x)^2}{2\sigma_{sz}^2}-\frac{(r\sin\phi-m_y)^2}{2\sigma_{sz}^2}\right) \tag{7}$$

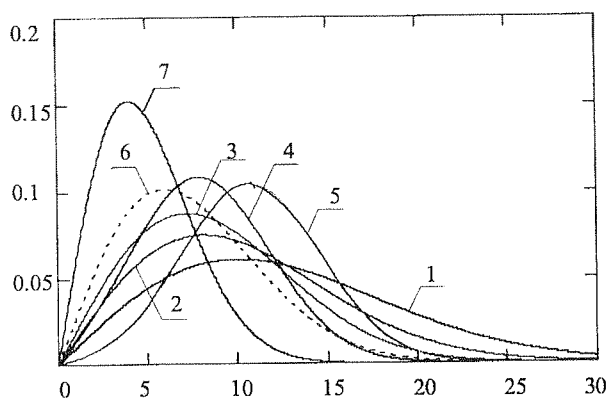
Całkując rozkład (7) względem zmiennej ϕ w przedziale $(0,2\pi)$ otrzymujemy (jako brzegowy) następujący rozkład obwiedni sygnału użytecznego

$$p(r)=\frac{r}{\sigma_{sz}^2}\exp\left(-\frac{r^2+a^2}{2\sigma_{sz}^2}\right)I_o\left(\frac{ar}{\sigma_{sz}^2}\right) \tag{8a}$$

gdzie:

$$I_o(\alpha)=\frac{1}{2\pi}\int_0^{2\pi}\exp(\alpha\cos\phi)d\phi=\sum_{k=0}^{\infty}\frac{(\alpha/2)^{2k}}{(k!)^2} \tag{8b}$$

jest zmodyfikowaną funkcją Bessela pierwszego rodzaju rzędu zerowego. Rozkład (8a) nazywamy rozkładem Rice'a albo uogólnionym rozkładem Rayleigha, czasami rozkładem n. Wykresy gęstości rozkładu Rice'a zostały przedstawione na rysunku 2.



Rys. 2. Wykresy gęstości rozkładu Rice'a i parametry rozkładu wynoszą odpowiednio
1. $a=1$ $\sigma_{sz}=10$ 2. $a=2$ $\sigma_{sz}=8$, 3. $a=5$ $\sigma_{sz}=6$,
4. $a=7$ $\sigma_{sz}=4$ 5. $a=10$ $\sigma_{sz}=4$, 6. $a=0$ $\sigma_{sz}=6$,
7. $a=0$ $\sigma_{sz}=4$

Przy definiowaniu momentów procesu stacjonarnego wystarczy rozpatrzenie zmiennej losowej w dowolnej chwili t_o [11]. Obwiednia r(t) o rozkładzie Rice'a posiada następującą wartość momentu k-tego rzędu [2, 3]

$$E[r^k]=(2\sigma_{sz}^2)^{k/2}\Gamma\left(1+\frac{k}{2}\right){}_1F_1\left(-\frac{k}{2},1,-\frac{a^2}{2\sigma_{sz}^2}\right)\qquad k=0,1,2,\ldots \tag{9}$$

gdzie: $\Gamma(a)=\int_0^\infty t^{a-1}\exp(-t)dt$ jest funkcją gamma, a ${}_1F_1(a, b, x)$ jest funkcją Kummera (zdegenerowaną funkcją hipergeometryczną) zdefiniowaną jako

$$ {}_1F_1(a,\ b,\ x)=\frac{\Gamma(b)}{\Gamma(a)\Gamma(b-a)}\int_0^1 \exp(xt)t^{a-1}(1-t)^{b-a-1}dt; \qquad \mathrm{Re}(b)>\mathrm{Re}(a) $$

lub

$$ {}_1F_1(a,\ b,\ x)=\frac{\Gamma(b)}{\Gamma(a)}\sum_{n=0}^{\infty}\frac{\Gamma(a+n)x^n}{\Gamma(b+n)n!} $$

Przy tym zachodzą następujące zależności:

$$ {}_1F_1(a,\ b,\ 0)=1 $$

$$ {}_1F_1\left(\frac{1}{2},\ 1,\ x\right)=\exp\left(\frac{x}{2}\right)I_0\left(\frac{x}{2}\right) $$

$$ F_1\left(-\frac{1}{2},\ 1,\ -x\right)=\exp\left(-\frac{x}{2}\right)\left[(1+x)I_0\left(\frac{x}{2}\right)+xI_1\left(\frac{x}{2}\right)\right] $$

$$ F_1\left(-\frac{1}{2},\frac{1}{2},\ -\frac{x^2}{2}\right)=\exp\left(-\frac{x^2}{2}\right)+\sqrt{\frac{\pi}{2}}[2F(x)-1]x $$

$$ \Gamma(a+1)=a\Gamma(a);\ \Gamma\left(\frac{1}{2}\right)=\sqrt{\pi} $$

dla $a\in N$, gdzie N jest zbiorem liczb naturalnych, słuszne są następujące zależności:

$$ \Gamma(a)=(a-1)!; \qquad \Gamma\left(a+\frac{1}{2}\right)=\frac{(2a-1)!!}{2^a}\sqrt{\pi} $$

$I_i(\alpha)$ jest zmodyfikowaną funkcją Bessela pierwszego rodzaju rzędu i zdefiniowaną jako

$$ I_i(\alpha)=\sum_{k=0}^{\infty}\frac{(\alpha/2)^{2k+i}}{(k!)\Gamma(i+k+1)} $$

F(x) jest funkcją Laplace'a określoną jako:

$$ F(x)=\frac{1}{2\pi}\int_{-\infty}^{x}\exp\left(-\frac{t^2}{2}\right)dt=\frac{1}{2}+\frac{1}{\sqrt{2\pi}}\int_0^x\exp\left(-\frac{t^2}{2}\right)dt. $$

W analizie często wykorzystuje się funkcję Krampa zdefiniowaną jako

$$\psi(x)=\frac{2}{\sqrt{2\pi}}\int_0^x \exp\left(-\frac{t^2}{2}\right)dt=2F(x)-1$$

Ponadto dla $n\in N$ słuszna jest zależność

$$(2n-1)!!=\frac{(2n)!}{2^n n!}=\frac{(2n-1)!}{2^{n-1}(n)1)!}=\frac{n!}{2^{n-1}}\binom{2n-1}{n}$$

Uwzględniają zależność (9), wartość średnią zmiennej o rozkładzie Rice'a można obliczyć na podstawie wzoru [2, 3]

$$E[r]=\sigma_{sz}\sqrt{\frac{\pi}{2}}F_1\left(-\frac{1}{2},\,1,\,-\frac{a^2}{2\sigma_{sz}^2}\right)=\sigma_{sz}\sqrt{\frac{\pi}{2}}\exp\left(-\frac{a^2}{4\alpha_{sz}^2}\right)\left[\left(1+\frac{a^2}{2\sigma_{sz}^2}\right)I_o\left(\frac{a^2}{4\sigma_{sz}^2}\right)+\frac{a^2}{2\sigma_{sz}^2}I_1\left(\frac{a^2}{4\sigma_{sz}^2}\right)\right] \tag{10}$$

Natomiast wartość średnią kwadratu obwiedni r(t) określa zależność

$$E[r^2]=a^2+2\sigma_{sz}^2 \tag{11}$$

Sygnał opisany zależnością (6a), posiadający obwiednię o rozkładzie Rice'a ma rozkład fazy określony wzorem [2, 3]

$$p(\phi)=\frac{1}{2\pi}\exp\left(-\frac{a^2}{2\sigma_{sz}^2}\right)+\frac{a\cos(\phi-\phi_a)}{\sigma_{sz}^2\sqrt{2\pi}}F\left(\frac{a\cos(\phi-\phi_a)}{\sigma_{sz}^2}\right)\exp\left(\frac{a^2\sin(\phi-\phi_a}{2\sigma_{sz}^2}\right);|\phi-\phi_a|\leqslant\pi \tag{12}$$

gdzie $\phi_a=\operatorname{arctg}\left(\frac{m_y}{m_x}\right)$ jest fazą początkową składowej regularnej sygnału użytecznego. Zależność (12) jest rozkładem brzegowym dla rozkładu określonego wzorem (7).

Sygnał o obwiedni Rice'a występuje na wejściu demodulatora (detektora) przy odbiorze sygnałów telegraficznych z modulacją amplitudy, częstotliwości lub fazy oraz dla radiowej transmisji cyfrowych sygnałów mowy o modulacji PCM w obecności szumu fluktuacyjnego lub częściowego rozpraszania [11]. Ponadto rozkład Rice'a jest często stosowany przy modelowaniu rozproszenia, transmisji wielodrogowej sygnału harmonicznego, gdy na jednej z tras występuje sygnał dominujący bez zaników. Rozkład ten jest również często stosowany przy modelowaniu transmisji w kanale satelitarnym. Przykłady zastosowania rozkładu Rice'a do opisu sygnałów z zanikami oraz do oceny jakości transmisji przedstawiono w [7].

3. ROZKŁAD RAYLEIGHA

Jeżeli sygnał s(t) określony zależnością (6a) nie zawiera składowej regularnej, tj. gdy $a=0$, wówczas rozkład obwiedni (8a) staje się rozkładem Rayleigha o gęstości prawdopodobieństwa

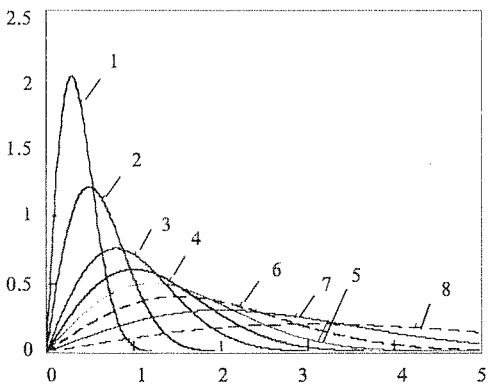
$$p(r)=\frac{r}{\sigma_{sz}^2}\exp\left(-\frac{r^2}{2\sigma_{sz}^2}\right) \tag{13}$$

Wówczas też faza sygnału użytecznego posiada rozkład równomierny

$$p(\phi)=\frac{1}{2\pi};\quad |\phi| \leqslant \pi. \tag{14}$$

Wykresy gęstości rozkładu Rayleigha zostały przedstawione na rysunku 3. Wartość średnia obwiedni o rozkladzie Rayleigha jest proporcjonalna do dyspersji (wartości skutecznej) szumu

$$E[r]=\sqrt{\frac{\pi}{2}}\sigma_{sz} \tag{15}$$



Rys. 3. Wykresy gęstości rozkładu Rayleigha, parametry rozkładu wynoszą odpowiednio: 1. $\sigma_{sz}=0.3$, 2. $\sigma_{sz}=0.5$, 3. $\sigma_{sz}=0.8$, 4. $\sigma_{sz}=1$, 5. $\sigma_{sz}=1.2$, 6. $\sigma_{sz}=1.5$, 7. $\sigma_{sz}=2$, 8. $\sigma_{sz}=3$

W analizie często zakłada się, że zmienna losowa reprezentująca obwiednię sygnału pasmowego o normalnym rozkładzie wartości chwilowej i zerowej wartości średniej oraz sygnału radiowego rozproszonego posiada rozkład Rayleigha [11].

4. JEDNOSTRONNY ROZKŁAD GAUSSA

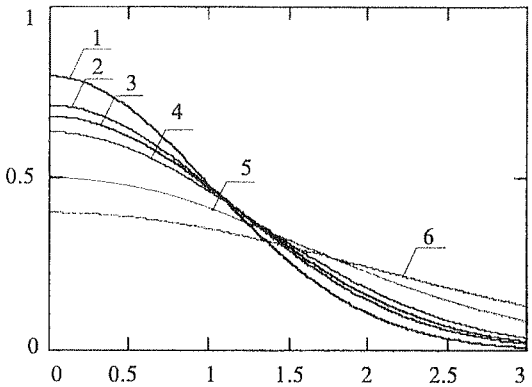
Niech wąskopasmowy sygnał użyteczny s(t) określony jest zależnością

$$s(t)=x(t)\cos\omega_o t-y(t)\sin\omega_o t=r(t)\cos[\omega_o t+\phi(t)] \tag{16}$$

Zakładamy, że składowe ortogonalne tego sygnału mają normalny rozkład o wariancji $\sigma_x^2=0$, $\sigma_y^2>0$ i o wartości średniej $m_x=m_y=0$. Wówczas obwiednia sygnału (16) posiada rozkład jednostronny normalny

$$p(r)=\frac{2}{\sqrt{2q\sigma_y^2}}\exp\left(-\frac{r^2}{2\sigma_y^2}\right) \tag{17}$$

Wykresy gęstości jednostronnego rozkładu Gaussa zostały przedstawione na rysunku 4.



Rys. 4. Wykresy gęstości jednostronnego rozkładu Gaussa, parametry rozkładu wynoszą odpowiednio: 1. σ_y=0dB, 2. σ_y=0.5dB, 3. σ_y=0.7dB, 4. σ_y=1dB, 5. σ_y=2dB, 6. σ_y=3dB

5. ROZKŁAD HOYTA

Zakładamy, że sygnał użyteczny s(t) określony jest zależnością (16) oraz, że składowe ortogonalne tego sygnału mają rozkład normalny o różnej wariancji $\sigma_x^2\neq\sigma_y^2>0$, $\sigma_y^2>\sigma_x^2$ i o wartości średniej $m_x=m_y=0$. Wówczas obwiednia sygnału (16) posiada rozkład Hoyta o gęstości prawdopodobieństwa [3]

$$p(r)=\frac{r}{\sigma_x\sigma_y}\exp\left(-\frac{\sigma_x^2+\sigma_y^2}{4\sigma_x^2\sigma_y^2}r^2\right)I_o\left(\frac{\sigma_y^2-\sigma_x^2}{4\sigma_x^2\sigma_y^2}r^2\right) \tag{18}$$

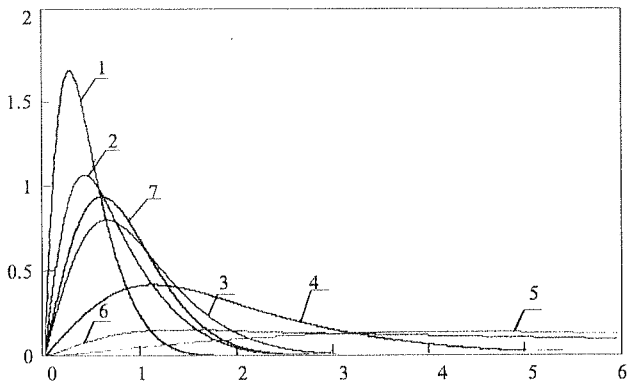
Gęstość prawdopodobieństwa (18) można również przedstawić w następującej postaci [3]

$$p(r)=\frac{r}{\sigma_x\sigma_y}\exp\left(-\frac{r^2}{2\sigma_x^2}\right){}_1F_1\left(\frac{1}{2},\,1,\,\frac{\sigma_y^2-\sigma_x^2}{2\sigma_x^2\sigma_y^2}r^2\right) \tag{19}$$

Rozkład Hoyta nazywany jest też rozkładem pod-Rayleigha (sub-Rayleigha) [3]. Wykresy gęstości rozkładu Hoyta zostały przedstawione na rysunku 5.

Rozkład Hoyta można zastosować do opisu obwiedni sygnału użytecznego w przypadku, gdy składowe ortogonalne tego sygnału posiadają różne moce średnie i gdy w propagacji wielodrogowej nie występuje składowa dominująca. W przypadku różnych mocy średnich składowych ortogonalnych i istnienia składowej dominującej,

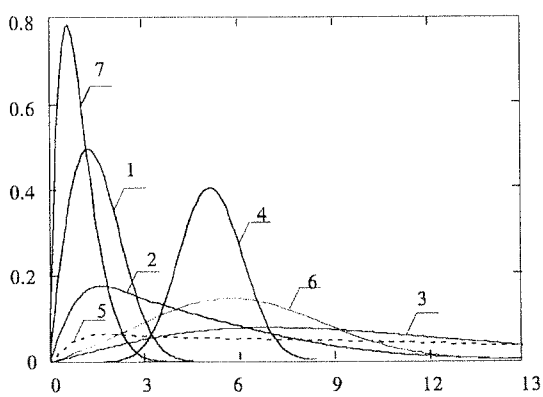
obwiednia sygnału posiada rozkład trójparametrowy, który zostanie omówiony w dalszej części pracy.



Rys. 5. Wykresy gęstości rozkładu Hoyta, parametry rozkładu wynoszą odpowiednio:
1. $\sigma_x=0.2$ $\sigma_y=0.5$ 2. $\sigma_x=0.3$ $\sigma_y=0.8$, 3. $\sigma_x=0.5$ $\sigma_y=1$,
4. $\sigma_x=0.8$ $\sigma_y=2$, 5. $\sigma_x=3$ $\sigma_y=6$, 6. $\sigma_x=0.8$ $\sigma_y=6$
7. $\sigma_x=0.5$ $\sigma_y=0.8$

6. ROZKŁAD TRÓJPARAMETROWY

Zakładamy, że sygnał użyteczny s(t) określony zależnością (16) oraz, że składowe ortogonalne tego sygnału mają rozkład normalny o różnej wariancji $\sigma_x^2 \neq \sigma_y^2 > 0$ i o wartościach średnich $m_x \neq 0$, $m_y = 0$. Wówczas obwiednia sygnału (16) posiada rozkład trójparametrowy o gęstości prawdopodobieństwa [2, 3]



Rys. 6. Wykresy gęstości rozkładu trójparametrowego, parametry rozkładu wynoszą odpowiednio:
1. $m_x=1$, $\sigma_x=1$, $\sigma_y=1$, 2. $m_x=1$, $\sigma_x=5$, $\sigma_y=1$, 3. $m_x=1$, $\sigma_x=5$, $\sigma_y=10$,
4. $m_x=5$, $\sigma_x=1$, $\sigma_y=1$, 5. $m_x=1$, $\sigma_x=14$, $\sigma_y=1$, 6. $m_x=5$, $\sigma_x=3$, $\sigma_y=3$,
7. $m_x=0$, $\sigma_x=0.5$, $\sigma_y=1$

$$p(r)=\frac{r}{2\pi\sigma_x\sigma_y}\int_0^{2\pi}\exp\left[-\frac{(r\cos\phi-m_x)^2}{2\sigma_x^2}-\frac{r^2\sin^2\phi}{2\sigma_y^2}\right]d\phi= \tag{20}$$

$$=\frac{r}{\sigma_x\sigma_y}\exp\left(-\frac{r^2+m_x^2}{2\sigma_x^2}\right)\sum_{i=0}^{\infty}\frac{(2i-1)!!(\sigma_y^2-\sigma_x^2)^i}{i!2^i\sigma_y^{2i}m_x^i}r^iI_i\left(\frac{rm_x}{\sigma_x^2}\right).$$

Wykresy gęstości rozkładu trójparametrowego zostały przedstawione na rysunku 6. Rozkład trójparametrowy dla $\sigma_x^2=\sigma_y^2=\sigma_{sz}^2$ staje się rozkładem Rice'a, natomiast dla $m_x=m_y=0$ i $\sigma_x^2=0$ otrzymujemy jednostronny rozkład normalny. Jeżeli $m_x=m_y=0$ i $\sigma_x^2\neq\sigma_y^2\neq0$ to rozkład trójparametrowy staje się rozkładem Hoyta. Natomiast dla $\sigma_x^2=\sigma_y^2=0$ otrzymujemy rozkład obwiedni sygnału użytecznego w kanale bez zaników.

7. ROZKŁAD CZTEROPARAMETROWY (UOGÓLNIONY ROZKŁAD NORMALNY)

Zakładamy, że sygnał użyteczny s(t) określony jest zależnością (16) oraz, że składowe ortogonalne tego sygnału mają normalny rozkład o różnej wariancji $\sigma_x^2\neq\sigma_y^2>0$ i wartości średniej $m_x\neq m_y$. Obwiednia sygnału s(t) posiada wówczas rozkład czteroparametrowy o gęstości prawdopodobiwństwa [4]

$$p(r)=\sum_{k=0}^{\infty}\frac{G^k}{k!}\sigma^{2G}\frac{\partial^{2G}}{\partial m_I^G\partial m_{II}^G}\left[\frac{r}{\sigma^2}\exp\left(-\frac{r^2+m_I^2+m_{II}^2}{2\sigma^2}\right)I_o\left(\frac{r}{\sigma^2\sqrt{m_I^2+m_{II}^2}}\right)\right] \tag{21}$$

gdzie: $m_I=\frac{m_x+m_y}{\sqrt{2}}$; $m_{II}=\frac{m_x-m_y}{\sqrt{2}}$; $\sigma^2=\frac{\sigma_x^2+\sigma_y^2}{\sqrt{2}}$; $G=\frac{\sigma_y^2-\sigma_x^2}{\sigma_x^2+\sigma_y^2}$.

Wówczas też faza sygnału użytecznego posiada rozkład o gęstości

$$p(\phi)=\frac{\sigma_x\sigma_y}{2\pi(\sigma_y^2\cos^2\phi+\sigma_x^2\sin^2\phi)}\exp\left(-\frac{m_x^2}{\sigma_x^2}-\frac{m_y^2}{\sigma_y^2}\right)\left\{1+K\sqrt{\pi}\exp(K^2)[1+\psi(\sqrt{2}K)]\right\} \tag{22}$$

gdzie: $K=\frac{m_y\sigma_y^2\sin\phi+m_x\sigma_y^2\cos\phi}{\sigma_x\sigma_y\sqrt{\sigma_x^2\sin^2\phi+\sigma_y^2\cos^2\phi}}$.

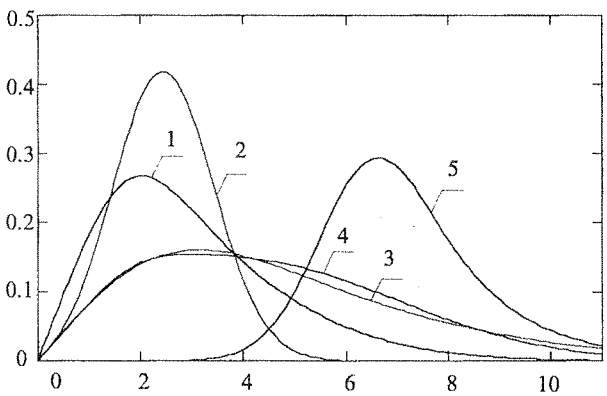
Rozkład prawdopodobieństwa obwiedni określony wzorem (21) można również przedstawić w postaci [2]

$$p(r)=\frac{r}{\sigma_x\sigma_y}\exp\left(-\frac{r^2}{2\sigma_x^2}-\frac{m_x^2\sigma_y^2+m_y^2\sigma_x^2}{2\sigma_x^2\sigma_y^2}\right)\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{(2i+2j)!!(\sigma_y^2-\sigma_x^2)^im_y^{2j}\sigma_x^{2j}}{i!2^i(2j)!\quad\sigma_y^{2i+4j}m_x^{i+j}}r^{i+j}I_{i+j}\left(\frac{rm_x}{\sigma_x^2}\right) \tag{23}$$

albo w postaci [2]

$$p(r)=\frac{r}{\sigma_x\sigma_y}\exp\left(-\frac{r^2}{4}\left(\frac{1}{\sigma_x^2}+\frac{1}{\sigma_y^2}\right)-\frac{1}{2}\left(\frac{m_x^2}{\sigma_x^2}+\frac{m_y^2}{\sigma_y^2}\right)\right)\left\{I_0\left[\frac{r^2}{4}\left(\frac{1}{\sigma_x^2}-\frac{1}{\sigma_y^2}\right)\right]I_0\left[r\sqrt{\frac{m_x^2}{\sigma_x^4}+\frac{m_y^2}{\sigma_y^4}}\right]+\right.$$
$$\left.+2\sum_{k=1}^{\infty}I_k\left[\frac{r^2}{4}\left(\frac{1}{\sigma_x^2}-\frac{1}{\sigma_y^2}\right)\right]I_{2k}\left[r\sqrt{\frac{m_x^2}{\sigma_x^4}+\frac{m_y^2}{\sigma_y^4}}\right]\cos\left[2k\,\mathrm{arctg}\left(\frac{m_y\sigma_x^2}{m_x\sigma_y^2}\right)\right]\right\} \tag{24}$$

Rozkład czteroparametrowy staje się rozkładem:
a) trójparametrowym, gdy $m_x\neq 0$, $m_y=0$,
b)Rice'a, gdy $\sigma_x^2=\sigma_y^2=\sigma_{sz}^2$, $m_x^2+m_y^2\neq 0$,
c) jednostronnym normalnym, gdy $m_x=m_y=0$, $\sigma_x^2=0$,
d) Rayleigha, gdy $m_x=m_y=0$, $\sigma_x^2=\sigma_y^2=\sigma_{sz}^2$,
e) Hoyta, gdy $m_x=m_y=0$, $\sigma_x^2\neq\sigma_y^2\neq 0$.
Natomiast dla $\sigma_x^2=\sigma_y^2=0$ mamy do czynienia z kanałem bez zaników.
Wykresy gęstości rozkładu czteroparametrowego zostały przedstawione na rysunku 7.



Rys. 7. Wykresy gęstości rozkładu czteroparametrowego, parametry rozkładu wynoszą odpowiednio:
1. $m_x=1$, $m_y=1$, $\sigma_x=1$, $\sigma_y=3$, 2. $m_x=1$, $m_y=2$, $\sigma_x=1$, $\sigma_y=1$,
3. $m_x=1$, $m_y=1$, $\sigma_x=2$, $\sigma_y=5$, 4. $m_x=4$, $m_y=1$, $\sigma_x=3$, $\sigma_y=1$,
5. $m_x=1$, $m_y=6$, $\sigma_x=4$, $\sigma_y=1$

Dla rozkładu czteroparametrowego często zamiast parametrów m_x, m_y, σ_x, σ_y stosuje się cztery inne parametry posiadające określony sens fizyczny:

$$q^2=\frac{m_x^2+m_y^2}{\sigma_x^2+\sigma_y^2}=\frac{r_p^2}{\sigma_x^2+\sigma_y^2} \tag{25a}$$

q^2 określa stosunek średniej mocy składowej regularnej sygnału użytecznego do mocy składowej fluktuującej,

$$\beta^2=\frac{\sigma_x^2}{\sigma_y^2} \tag{25b}$$

β^2 określa stosunek wariancji (mocy) składowych ortogonalnych x(t), y(t); parametr ten charakteryzuje więc stopień niesymetrii kanału ze względu na składowe ortogonalne,

$$\phi_a = \operatorname{arctg}\left(\frac{m_y}{m_x}\right) \tag{25c}$$

ϕ_a określa kąt fazowy składowej regularnej sygnału użytecznego w układzie współrzędnych X, Y, przy czym $m_x = a \sin \phi_a$, $m_y = a \sin\phi_a$ w chwili $t = t_0$,

$$r_o^2 = \sigma_x^2 + \sigma_y^2 + m_x^2 + m_y^2 \tag{25d}$$

$r_o^2/2$ określa średnią moc sygnału użytecznego s(t).

Pomiędzy parametrami r_o^2, q^2, β^2, ϕ_a oraz σ_x^2, σ_y^2, m_x^2, m_y^2 istnieją nastepujące zależności:

$$m_x^2 = \frac{q^2 r_o^2 \cos^2\phi_a}{1+q^2} \tag{26a}$$

$$m_y^2 = \frac{q^2 r_o^2 \sin^2\phi_a}{1+q^2} \tag{26b}$$

$$\sigma_x^2 = \frac{\beta^2 r_o^2}{(1+\beta^2)(q+q^2)} \tag{26c}$$

$$\sigma_y^2 = \frac{r_o^2}{(1+\beta^2)(q+q^2)} \tag{26d}$$

Przykłady zastosowania rozkładu czteroparametrowego do opisu sygnałów z zanikami oraz do oceny jakości transmisji przedstawiono w [5, 6].

8. ROZKŁAD NAKAGAMIEGO

Zakładamy, że sygnał użyteczny s(t) określony jest zależnością (16). Do opisu obwiedni sygnału s(t) często stosuje się rozkład Nakagamiego o gęstości prawdopodobieństwa

$$p(r) = \frac{2}{r(m)}\left(\frac{m}{\Omega}\right)^m r^{2m-1}\exp\left(-\frac{m}{\Omega}r^2\right) \tag{27}$$

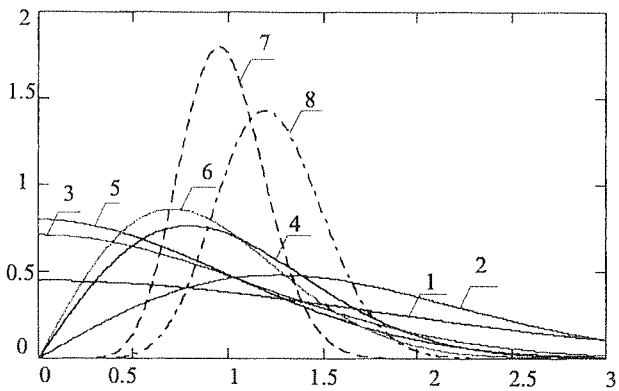
gdzie: $\Omega = E[r^2]$ jest średnią mocą sygnału, $m = \dfrac{\Omega^2}{E[r^2-\Omega]^2} \geqslant \dfrac{1}{2}$ jest odwrotnością unormowanej wariancji kwadratu obwiedni sygnału użytecznego, czyli odwrotnością unormowanej średniej mocy sygnału. Parametr ten charakteryzuje głębokość zaników.

Rozkład Nakagamiego pierwotnie dotyczył geometrycznej sumy obwiedni m niezależnych sygnałów całkowicie rozproszonych. Rozkład ten opisuje wypadkową obwiednię sygnału, gdy jego składowe o rozkładzie Rayleigha mają jednakowe moce średnie. Następnie parametr m rozkładu Nakagamiego został uogólniony, teraz może on przyjmować dowolną wartość rzeczywistą z zakresu $<0.5, \ldots, \infty)$. Rozkład

Nakagamiego w istocie jest rozkładem chi, w którym parametr m może przyjmować także wartości niecałkowite. Rozkład Nakagamiego często jest nazywany rozkładem m. Wykresy gęstości rozkładu Nakagamiego zostały przedstawione na rysunku 8.

Sygnał s(t) o rozkładzie obwiedni Nakagamiego posiada następującą wartość momentu k-tego rzędu

$$E[r^k]=\frac{\Gamma\left(m+\frac{k}{2}\right)}{\Gamma(m)}\sqrt{\left(\frac{\Omega}{m}\right)^k} \tag{28}$$



Rys. 8. Wykresy gęstości rozkładu Nakagamiego, parametry rozkładu wynoszą odpowiednio
1. m=0.5, Ω=5dB, 2. m=1, Ω=5dB, 3. m=0.5, Ω=1dB,
4. m=1, Ω=1dB, 5. m=0.5, Ω=0dB, 6. m=1, Ω=0dB,
7. m=5, Ω=0dB, 8. m=5, Ω=2dB

Rozkład Nakagamiego dobrze aproksymuje inne rozkłady. W szczególności otrzymujemy:

a) jednostronny rozkład normalny, gdy $m=\frac{1}{2}$ (wówczas $m_x=m_y=0$, $\sigma_x^2=0$),

b) rozkład Rayleigha, gdy $m=1$ (wówczas $m_x=m_y=0$, $\sigma_x^2=\sigma_y^2=\sigma_{sz}^2$),

c) rozkład Rice'a, gdy $\Omega=2\sigma^2+a^2$ oraz $m=\frac{\Omega^2}{\Omega^2-a^4}\geqslant 1$

d) rozkład czteroparametrowy, gdy [2]

$$m=\frac{(\sigma_x^2+\sigma_y^2+m_x^2+m_y^2)^2}{2(\sigma_x^4+\sigma_y^4+2\sigma_x^2m_x^2+2\sigma_y^2m_y^2)}=\frac{[(1+\beta^2)(1+q^2)]^2}{2[1+\beta^4+2q^2(1+\beta^2)(\beta^2\cos^2\phi_a+\sin^2\phi_a)]} \tag{29a}$$

$$\Omega=\sigma_x^2+\sigma_y^2+m_x^2+m_y^2. \tag{29b}$$

Natomiast dla $m\to\infty(\sigma_x^2=\sigma_y^2=0)$ mamy kanał bez zaników.

Jeżeli obwiednia sygnału s(t) posiada rozkład Nakagamiego o gęstości prawdopodobieństwa określonej zależnością (27), a faza rozkład równomierny to rozkład wartości chwilowej tego sygnału można przedstawić jako [4]

$$p(s)=\frac{m^{\frac{m}{2}-\frac{1}{4}}\exp\left(-\frac{ms^2}{2\Omega}\right)s^{m-\frac{3}{2}}}{\sqrt{\pi}\,\Gamma(m)\,\Omega^{\frac{m}{2}-\frac{1}{4}}}W_{\lambda\nu}\left(\frac{ms^2}{\Omega}\right) \tag{30}$$

gdzie $\lambda=\nu=\frac{m}{2}-\frac{1}{4}$; $W_{\lambda,\nu}(x)$ jest funkcją Whittakera zdefiniowaną jako

$$W_{\lambda,\nu}(x)=\frac{x^{\nu+\frac{1}{2}}\exp\left(-\frac{x}{2}\right)}{\Gamma\left(\nu-\lambda+\frac{1}{2}\right)}\int_0^\infty \exp(-xt)t^{\nu-\lambda-\frac{1}{2}}(1+t)^{\nu+\lambda-\frac{1}{2}}dt$$

przy czym słuszna jest zależność $W_{\frac{1}{4},\frac{1}{4}}(x)=\exp\left(-\frac{x}{2}\right)\sqrt[4]{x}$.

Przykłady zastosowania rozkładu Nakagamiego do opisu sygnałów z zanikami oraz do oceny jakości transmisji przedstawiono w [8, 9, 10, 11].

Porównując przedstawione probabilistyczne opisy sygnałów z zanikami można stwierdzić, że model Nakagamiego jest znacznie prostszy niż model Rice'a, który w dostępnej literaturze jest znacznie częściej stosowany. Rozkład Nakagamiego uwzględnia szeroką klasę zaników, jednak jeszcze szerszą klasę zaników opisuje rozkład czteroparametrowy. Wyznaczanie funkcjonałów jakości transmisji w kanale Rice'a, w kanale opisanym rozkładem czteroparametrowym, trójparametrowym oraz Hoyta wymaga bardziej skomplikowanych obliczeń niż w kanale Nakagamiego. Dlatego też rozkład Nakagamiego powinien być szerzej rozpowszechniony w badaniach dotyczących jakości transmisji cyfrowej.

9. PODSUMOWANIE

Do opisu elementarnego sygnału całkowicie rozproszonego z zanikami stosuje się rozkłady jednoparameterowe, np. rozkład Rayleigha, w tym przypadku parametr rozkładu charakteryzuje średnią moc sygnału. Przy zaawansowanym modelowaniu obwiedni sygnałów wielodrogowych korzysta się z rozkładów dwuparamerowych, trójparametrowych lub czteroparametrowych. Założenie całkowitego rozproszenia upraszcza model sygnału. Przy rozproszeniu niecałkowitym niezbędne jest zastosowanie modelu zaawansowanego, w którym uwzględnia się stopień rozproszenia oraz głębokość zaników. Możliwe są dwie przyczyny niecałkowitego rozproszenia w ośrodku propagacyjnym: istnienie zdeterminowanej składowej bezpośredniej oraz superpozycja niezależnych sygnałów całkowicie rozproszonych w jednej antenie lub w antenach systemu z odbiorem zbiorczym.

Przedstawione rozkłady można wykorzystać do modelowania sygnałów radiowych odbieranych z powolnymi fluktuacjami, do modelowania kanału z zanikami. Można je wykorzystać do budowy cyfrowych symulatorów nowych systemów radiokomunikacyjnych, zapewniających lepszą jakość transmisji. Przedstawione w literaturze dane doświadczalne potwierdzają słuszność i praktyczność omówionych rozkładów, szczególnie w kanałach wykorzystujących fale decymetrowe lub krótsze. Przedstawione rozkłady umożliwiają ocenę jakości transmisji w kanale z zanikami, umożliwiają określenie dynamicznego prawdopodobieństwa błędu elementowego dla transmisji cyfrowej, prawdopodobieństwa łączności, oraz niezawodności transmisji. Można je również wykorzystać do opisu strumienia błędów. Ponadto można je wykorzystać do modelowania transmisji szerokopasmowej, która posiada właściwości przeciwzanikowe i przeciwzagłuszeniowe.

BIBLIOGRAFIA

1. A. J. Coulson, A. G. Williamson, R. G. Vaughan: *Improved fading distribution for mobile radio*. IEE Proceedings Communications, Vol. 145, No 3, June 1998, pp. 197–202
2. D. D. Kłowskij: *Pieriedacza diskrietnych soobszczenij po radiokanałam*. Swiaź, Moskwa 1969
3. D. D. Kłowskij: *Teorija pieriedaczi signałow*. Swiaź, Moskwa 1973.
4. D. D. Kłowskij, W. J. Kontorowicz, S. M. Szirokow: *Modieli nieprierywnych kanałow swiazi na osnowie stochasticzieskich differiencjalnych urawnienij*. Radio Swiaź, Moskwa 1984
5. K. Noga: *Właściwości strumienia błędów dla transmisji binarnej w kanale z wolnymi zanikami opisanymi rozkładem czteroparametrowym*. Kwartalnik Elektroniki i Telekomunikacji, R. 40, 1994, Z. 3, ss. 357–368
6. K. Noga: *Ocena jakości odbioru transmisji w kanale z zanikami opisanymi rozkładem czteroparametrowym dla odbioru niekoherentnego i wielostanowego kluczowania częstotliwości*. X Krajowe Sympozjum Telekomunikacji 94, Bydgoszcz 94, Referaty Sekcji 1: Problemy podstawowe, 7–9 września 1994, ss. 249–255
7. K. Noga: *Prawodopodobieństwo błędu elementowego oraz właściwości strumienia błędów dla transmisji binarnej w kanale radiokomunikacyjnym z wolnymi zanikami Rice'a i addytywnym szumem gaussowskim*, Rozprawy Elektrotechniczne, 1986, Z. 4, ss. 1237–1251
8. M. K. Simon, M. S. Alouini: *A uniffied approach to the performance analysis of digital communication over generalized fading channels*. Proceedings of the IEEE, Vol. 86, No 9, September 1998
9. M. K. Simon, M. S. Alouini: *A unified approach to the probability of error for noncoherent and differentially coherent modulations over generalized fading channels*. IEEE Transactions on Communications, Vol. 46, No 12, December 1998, pp. 1625–1638
10. M. K. Simon, M. S. Alouini: *A unified performance analysis of digital communication with dual selective combining diversity over correlated Rayleigh and Nakagami – m fading channels*. IEEE Transactions on Communications, Vol. 47, No 1, January 1999, pp. 33–43
11. A. Wojnar: *Teoria sygnałów*. WNT, Warszawa 1988

K. NOGA

GENERALIZED MODEL FOR SIGNALS IN CHANNELS WITH FADING

Summary

In radio communication systems, and especially in mobile systems, the received signal can be affected by fading phenomena due to multipath propagation. In multipath channel, the received signal contains not only a direct ratio wave, but also a large number of reflected and scattered ratio waves. The most general model for the reception of digital signals transmitted through a slowly waves. The most general model for the reception of digital signals transmitted through a slowly verying fading medium is a multipath channel. Multipath fading occurs wherever a signal travels over more than one path between transmitter and receiver, causing fluctuations in the amplitude, phase and time-of arrival of the received signal. Different fading models have been used to describe the fading envelope of the received signal. Under assumptions that the amplitudes of waves have different distributions, the received signal fluctuations are good modeled by the Nakagami distribution. Many authors use the Nakagami distribution for land mobile ratio, for indoor and outdoor propagation. The Nakagami fading model is one of the most versatile. It has a generalized distribution that can model different fading environments. It also has the advantage of including the Rayleigh, Rice and the one-side Gaussian distributions as a special cases. The Rayleigh model is used for the resultant fading assuming a large number of uncorrelated partial waves with identically distributed amplitudes and random phase uniformly distributed on $(0, 2\pi)$.

A various fading channel models and the environments to which they apply are described in this paper. Probability of distribution functions are discussed for characterizing fading processe in moblie radio communications.

Key words: communication channel, fading channels, distribution of envelope, mobile radio.

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI, 2000, 46, z. 1, ss. 63–79

Implementacja sprzętowa skalowanego systemu kodowania obrazów wizyjnych wielkiej rozdzielczości w czasie rzeczywistym

KAZIMIERZ WIATR, PAWEŁ RUSSEK

Katedra Elektroniki, Akademia Górniczo-Hutnicza w Krakowie

Otrzymano 1999.04.14
Autoryzowano 1999.10.20

W pracy przedstawiono problematykę kodowania obrazów wideo wielkiej rozdzielczości SHD (ang. *super high definition*) w oparciu o metodę falkową, nazwaną od nazwiska jej twórcy metodą Shapiro lub metodą EZW (ang. *embedded zero-tree wavelet*). Jednocześnie przedstawiono zagadnienie skalowanego kodowania obrazów i wskazano obszary zastosowania tego rodzaju kodowania. Ukazano także zalety metody Shapiro przy spełnieniu wymogu skalowalności, stawianego kodowaniu.

Na tle powyższych rozważań przedstawiono możliwość realizacji kodera sygnału wideo w oparciu o architekturę przetwarzania danych typu MISD (ang. *Multiple Instruction-stream Single Data-stream*). Przedstawiona architektura kodera cechuje się dużą szybkością relizacji algorytmu kompresji obrazu metodą EZW. Cecha ta kwalifikuje przedstawioną architekturę do kodowania obrazów wielkiej rozdzielczości w trybie czasu rzeczywistego, a obecne prace autorów to implementacja tej architektury w strukturach programowalnych FPGA o wielkich pojemnościach.

Słowa kluczowe: przetwarzanie obrazów, kompresja obrazów, systemy czasu rzeczywistego, układy programowalne.

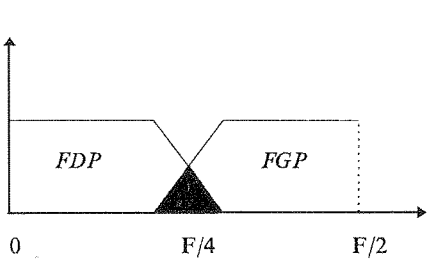
1. WSTĘP

Kompresja i kodowanie obrazów są obecnie ważnym sektorem prowadzonych prac badawczych, zarówno przez ośrodki uniwersyteckie, jak również wiodące firmy. Szczególne miejsce na tym polu zajmuje kompresja obrazów w trybie czasu rzeczywistego [4, 5, 7, 8, 11, 12, 13, 15]. Wzrastające zainteresowanie usługami wideo wielkiej rozdzielczości wpływa na poszukiwanie i rozwój technik kodowania takich obrazów. Za obrazy wielkiej rozdzielczości przyjmuje się obrazy o ziarnistości do 2048×2048 pikseli. Zainteresowanie to wynika nie tylko z rosnących potrzeb użytkowników dotychczas szeroko stosowanych mediów (głównie telewizji), ale także z nowych możliwości usług epoki informacji i telekomunikacji (np. *distance learning*).

Dla potrzeb kodowania obrazów SDH, jedną z grup metod są metody oparte na teorii falek (ang. *wavelet theory*) i wielowymiarowej analizie (ang. *multiresolution analysis*) obrazu. Klasa tych metod nie doczekała się do tej pory szerszego zastosowania praktycznego w przeciwieństwie do metod opartych na transformacji DCT, na której bazują szeroko stosowane standardy JPEG, MPEG, H.261 itd. Wraz z upowszechnianiem się stosowania obrazów SDH można spodziewać się wzrostu znaczenia metod falkowych i pojawienia się opartych na nich nowych standardów kodowania.

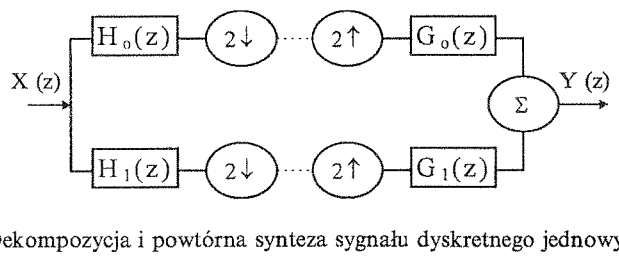
2. DEKOMPOZYCJA FALKOWA OBRAZU

Dekompozycja falkowa obrazu (ang. *subband decomposition*) teorię swoją wywodzi z teorii falek. Praktycznie sprowadza się ona do hiearchicznego filtrowania obrazu przy pomocy filtrów górno i dolnoprzepustowego w dwóch wymiarach. W ten sposób następuje oddzielenie składowych niskoczęstotliwościowych informacji zawartych w obrazie od wysokoczęstotliwościowej. Dobór filtrów nie jest zagadnieniem oczywistym. Należy zwrócić uwagę, że aby nie doprowadzić do utraty części informacji filtry powinny pokrywać cały zakres pasma częstotliwościowego. Rys. 1 pokazuje jak praktycznie budowane filtry mogą zapewnić spełnienie tego warunku.



Rys. 1. Pasma filtrów dolno i górnoprzepustowego dla dekompozycji falkowej

Widać, że częstotliwości w pobliżu F/4 (obszar zaznaczony na czarno) zawierają się po filtracji w częstotliwościowej składowej górnej i dolnej sygnału. Jest to niedogodność, którą należy uwzględnić podczas odtwarzania sygnału. Poprawne odtworzenie obrazu pierwotnego wymaga stosowania filtrów syntezujących — odpowiednio górno i dolnoprzepustowego. Zastosowanie dodatkowego kryterium

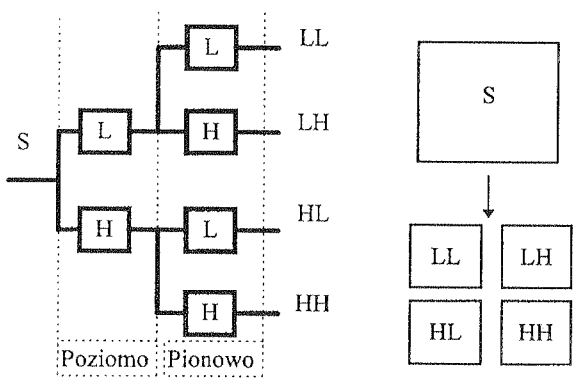


Rys. 2. Dekompozycja i powtórna synteza sygnału dyskretnego jednowymiarowego

umożliwia taki ich dobór, który kompensuje aliasing filtrów analizujących sygnał na wejściu. Cały tor dekompozycji i powtórnej kompozycji dyskretnego sygnału jednowymiarowego przedstawia rysunek (rys. 2).

System taki nosi nazwę systemu QMF (ang. *Quadrature Mirror Filter bank*). Funkcje $H_0(z)$, $H_1(z)$, $G_0(z)$, $G_1(z)$ reprezentują filtry dolno i górnoprzepustowe; H to filtry analizujące, G to filtry syntezujące. Dyskretny sygnał wejściowy reperezentuje funkcja X(z), a dyskretny sygnał wynikowy funkcja Y(z). Po przeprowadzeniu analizy sygnału przy pomocy filtrów H dokonuje się podzielenia częstotliwości próbkowania sygnałów wynikowych przez dwa (ang. *downsampling*). Można to zrobić, bowiem częstotliwości Nyquisa sygnałów składowych są dwukrotnie mniejsze od tejże częstotliwości sygnału wejściowego. Przed odtworzeniem sygnału należy częstotliwość podwoić (praktycznie odbywa się to poprzez wstawienie wartości zerowych pomiędzy istniejące próbki) i odfiltrować za pomocą filtrów syntezujących G. Liczba próbek niesiona przez dwa przedstawione na rysunku tory jest równa liczbie próbek zawartych w sygnałach X(z) i Y(z), a więc wielkość treści nawzajem sobie odpowiada. Następuje jedynie oddzielenie poszczególnych składowych częstotliwościowych. Szczegółowe teoretyczne uzasadnienie poprawności pracy tak zestawionego toru oraz informacje na temat projektowania filtrów analizujących i syntezujących można znaleźć w [16].

System przedstawiony powyżej odnosi się do sygnałów jednowymiarowych. W przypadku obrazów mamy do czynienia z dwoma wymiarami. Metoda stosowana w tym wypadku jest identyczna do przedstawionej powyżej, z tym że stosuje się ją niezależnie do kolumn i do wierszy obrazu. Rys. 3 przedstawia schematycznie sposób dekompozycji obrazu na cztery podstawowe komponenty częstotliwościowe.



L-filtr dolnoprzepustowy
H-filtr górnoprzepustowy

Rys. 3. Dekompozycja obrazu dwuwymiarowego na cztery składowe częstotliwościowe

Obraz wejściowy zamieniany jest na 4 obrazy o dwukrotnie mniejszej rozdzielczości, z których każdy reprezentuje odpowiednią kombinację częstotliwości przestrzennych w pionie i w poziomie.

W przypadku obrazów rzeczywistych większość energii zawarta jest w członach niskoczęstotliwościowych, a w szczególności w członie LL. Dlatego dokonuje się na nim podobnej operacji jak na obrazie pierwotnym dzieląc go na kolejne 4 czterokrotnie mniejsze od obrazu pierwotnego części. Operację tę można powtarzać wielokrotnie, ale dla celów kodowania nie robi się tego więcej niż 3, 4 razy. Wynik dwukrotnego filtrowania przeprowadzonego na obrazie 'Lena' przedstawiono na rys. 4. W celu lepszej rozróżnialności rysunku wartości punktów w elementach wysokoczęstotliwościowych zostały dodatkowo wzmocnione.

LH1		HH1
LH2	HH2	HL1
LL2	HL2	

Rys. 4. Wynik dwukrotnej filtracji obrazu „Lena"

Obszary czarne reprezentują piksele o zerowej wartości. Punkty białe w elementach HH1, HL1, LH1, HH2, LH2, HL2 to punkty, których wartość przekracza 5% maksymalnej wartoscu sygnału.

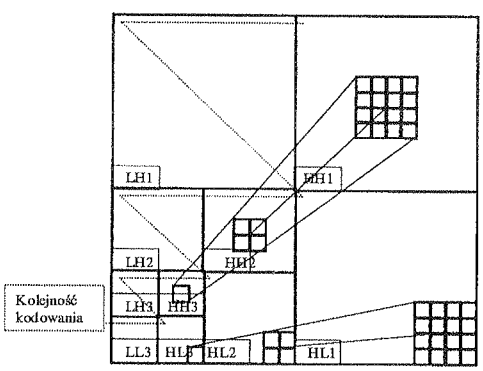
3. ALGORYTM KODOWANIA J. M. SHAPIRO

Poszczególne produkty filtrowania falkowego wykazują różnicę pod względem charakteru i ważności niesionej informacji. Istnieją sposoby kodowania wykorzystujące te różnice. Kodując indywidualnie poszczególne składowe obrazu przy pomocy dedykowanych dla każdej z nich algorytmów, można wydobyć takie ich cechy, aby kodowanie było jak najbardziej efektywne.

Na rys. 4 widać, że częstotliwościowe fragmenty obrazu cechuje również pewne podobieństwo przestrzenne. Można zauważyć, że obszary czarne na elemencie któregoś z poziomów są czarne w odpowiadających elementach na poziomach niższych i wyższych. To samo można powiedzieć o obszarach jasnych. Elementy odpowiadające to (HL1, HL2, HL3, ...); (LH1, LH2, LH3, ...); (HH1, HH2, HH3, ...). Przyjęta notacja odpowiada notacji na rysunku 3 i 4. Pierwsza litera to rodzaj filtrowania w poziomie (L — dolnoprzepustowe, H — górnoprzepustowe), druga litera to filtrowanie w pionie (L — dolnoprzepustowe, H — górnoprzepustowe), a cyfra określa kolejny numer filtrowania składnika LL (w wyniku filtrowania składnika LL n powstają skladniki LL$n+1$, HL$n+1$, LH$n+1$, HH$n+1$).

Ze względu na inną rozdzielczość, każdemu punktowi elementu na poziomie n odpowiadają cztery kunkty na poziomie $n-1$ (rys. 5).

Do kodowania Shapiro używa 4 symboli: plus — P, minus — M, zero zagnieżdżone — EZ (ang. *embedded zero*), zero izolowane — IZ (ang. *isolated zero*) [9]. Występowanie symboli plus i minus wynika z kodowania w procesie kompresji wideo nie kolejnych ramek, ale różnicy dwóch ramek, stąd możliwość występowania zarówno wartości dodatnich jak i ujemnych. Jeżeli dana wartość jest dodatnia kodujemy plus, jeżeli jest ujemna to kodujemy minus. Wartość zerowa jest kodowana przy pomocy dwóch symboli EZ i IZ. Jeżeli dla piksela o wartości zerowej wszystkie leżące na niższych poziomach odpowiadające piksele również przyjmują wartość zero wtedy koduje się symbol zera zagnieżdżonego EZ, w przeciwnym przypadku kodujemy zero izolowane IZ. Proces kodowania przebiega od poziomu najwyższego do poziomu pierwszego (rys. 5). Jeżeli więc dany punkt zostanie zakodowany jako EZ, to w procesie dalszego kodowania nie ma potrzeby kodowania punktów mu odpowiadających i dlatego punkty te są pomijane.



Rys. 5. Kolejne etapy kodowania metodą Shapiro

4. KWANTYZACJA OBRAZU

Przed przystąpieniem do kodowania obrazu metodą Shapiro podlega on procesowi kwantyzacji. Cyfrowa wartość każdego punktu obrazu zostaje podzielona na z góry zdefiniowane części (symbole) odpowiedniejsze do dalszych operacji.

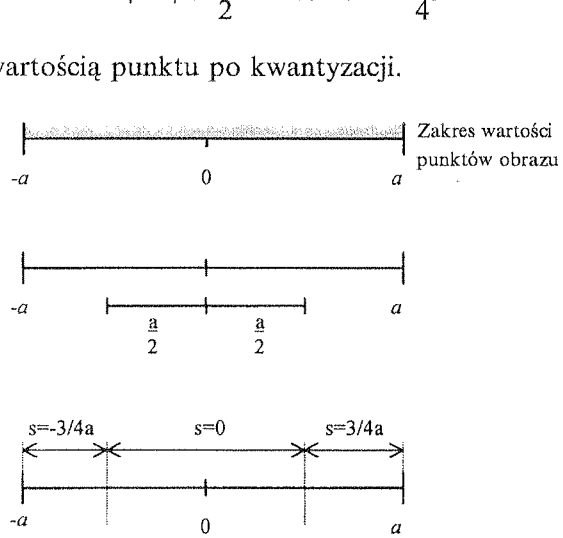
Proces kwantyzacji może przebiegać następująco: dla danego obrazu poszukujemy punktu o największej wartości bezwzględnej wartości sygnału, a następnie wartość tą nadajemy parametrowi a. Zakres zmienności wartości punktów obrazu ustalany jest na $[-a, a]$ (rys. 6). Następnie porównujemy wartość x poszczególnych pikseli z wartością parametru „a" i podstawiamy nowe wartości zgodnie z poniższymi regułami:

$$|x| < \frac{a}{2} \Rightarrow s = 0$$

$$|x| > \frac{a}{2} \wedge x > 0 \Rightarrow s = \frac{3}{4}a$$

$$|x| > \frac{a}{2} \wedge x < 0 \Rightarrow s = -\frac{3}{4}a$$

gdzie s jest nową wartością punktu po kwantyzacji.



Rys. 6. Zakres zmienności dla kwantyzacji

Ponieważ s przyjmuje tylko 3 wartości, to wartości pikseli reprezentowane (są dla potrzeb kodowania Shapiro) także przez 3 symbole nazwane plus, minus i zero. Tak skwantowana ramka nie zapewnia odpowiedniej wierności zakodowanego obrazu z obrazem oryginalnym dlatego kwantyzacja w tym miejscu nie może się jeszcze skończyć.

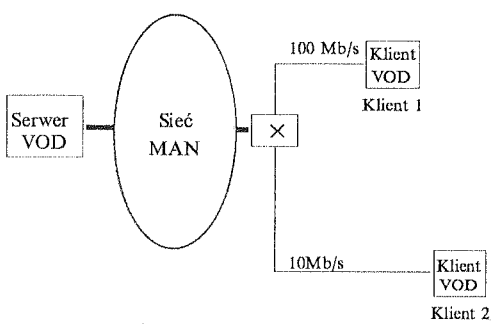
Następnie wartość s odejmuje się od odpowiadających wartości x. W ten sposób powstaje ramka z sygnałem błędu, który poddajemy kwantyzacji analogicznej do metody podanej powyżej, jednak tym razem jako zakres zmienności wartości przyjmujemy przedział $[-a/2, a/2]$. W procesie odtwarzania obrazu należy oczywiście wyniki kolejnych cykli kwantowania dodać do siebie. Operację powtarza się tak długo jak długo nie osiągnie się zadowalająco niskiego pozimu sygnału błędu, za każdym razem skracając przedział zmienności o połowę.

5. SKALOWANIE OBRAZU

Technologia interaktywnego przesyłania obrazów ruchomych na odległość nie jest powszechnie stosowana, ale już istniejące i działające systemy tego typu pozwalają przypuszczać, że w dziedzinie przepływu informacji perspektywa przesyłania obrazu wydaje się jedną z bardziej atrakcyjnych. Interaktywność usług wizyjnych na odległość np. wideokonferencje, wymaga isnienia odrębnego kanału transmisyjnego dla każdego z użytkowników. Jest to obecnie spory problem, bowiem istniejące kanały (telefon, sieć komputerowa) zapewniają przeciętnemu indywidualnem odbiorcy kanały o niewielkiej przepustowości w stosunku do potrzeb stawianych przez systemy wideo. Praktycznie działający standard wideotelefonii H.261, mieści się co prawda ze strumieniem danych w kanale telefonicznym, ale pozwala on na przesłanie obrazu o wielkości i płynności ruchu daleko odbiegających od oczekiwanych rezultatów, nawet w porównaniu ze standardową telewizją.

Oczywiście w miarę upływu czasu przepustowość sieci telekomunikacyjnych będzie coraz większa, ale konieczne wydają się prace nie tylko nad poszerzaniem pasma sieci, ale także nad próbami ograniczenia ilości przesyłanych informacji. Dobrą miarą w tym zakresie byłoby zmieszczenie się z obrazem w tym co jest dostępne. Ponadto uzasadnionym jest przypuszczenie, że nie wszyscy użytkownicy sieci będą dysponowali równie pojemnym kanałem transmisyjnym danych. Kanał taki może ponadto zmieniać swoją przepustowość w zależności od czasu.

Rozważmy przykładowo zdobywającą coraz większą popularność usługę VOD (ang. *Video On Demand*). W tego typu systemie użytkownik siedząc przed ekranem telewizora sam wybiera z biblioteki serwera VOD film lub program, który ma ochotę w danym momencie oglądać (rys. 7).



Rys. 7. Usługa VOD w sieci Klient-Serwer

Klienci 1 i 2 są podłączeni do miejskiej sieci komunikacyjnej. Klient 1 jest połączony z MAN za pomocą łącza o przepustowości 100 Mb/s, a klient 2 za pomocą łącza 10 Mb/s. Naturalnie klient 1 ma możliwość odbierać program o znacznie lepszej jakości niż klient 2. W przypadku, kiedy klient 1 i 2 oglądają ten sam program (np.: wiadomości) nie celowe byłoby przesyłanie poprzez MAN osobnego strumienia danych do klienta 1 i klienta 2 (strumień danych przeznaczonych dla klienta 1 nie mieściłby się w łączu klienta 2). Naturalne w takim przypadku byłoby odrzucenie

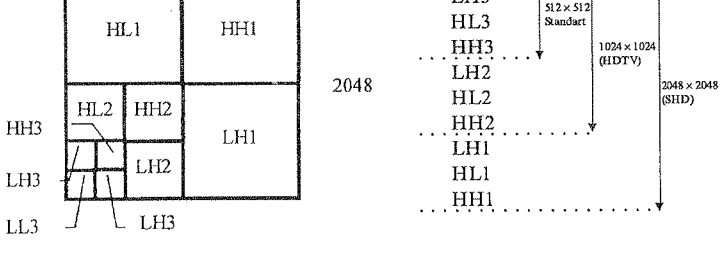
części strumienia przeznaczonego dla klienta 1 na etapie lokalnego koncentratora i skierowanie reszty do klienta 2. Każdy z klientów uzyska jakość obrazu zgodną z zainstalowanym u niego standardem.

Istniej ponadto taka możliwość, że na skutek chwilowego dużego obciążenia w sieci MAN, na etapie któregoś z przełączeń, nie ma wolnego kanału, aby zapewnić przepływ całości strumienia bitów wysyłanego z serwera VOD. Ponieważ informacja opuściła już koder wideo zainstalowany w serwerze, to nie ma już możliwości zmniejszenia ilości bitów składających się na obraz przy pomocy stosowanych w koderze mechanizmów. Sama sieć musi mieć możliwość kontrolowanego obcięcia pewnej ilości danych w taki sposób, aby nie uniemożliwić odtworzenia obrazu u odbiorcy. Obcięcie powinno mieć wpływ tylko i wyłącznie na chwilowe pogorszenie jakości odbieranego obrazu. Ciąg bitów, który umożliwia tego rodzaju operacje jest ciągiem skalowalnego obrazu wideo.

5.1. SKALOWANIE ROZMIARU

Przez pojęcie skalowania rozmiaru (ang. *spatial scaling*) rozumie się możliwość zmniejszenia strumienia bitów poprzez zmniejszenie rozdzielczości obrazu. W sposób naturalny i prosty tego rodzaju skalowalność zapewnia algorytm kodowania Shapiro.

Zwróćmy uwagę, że odtworzenie obrazu małej rozdzielczości jest możliwe już po przesłaniu elementu LL *n*. Przesłanie elementów wysokoczęstotliwościowych umożliwia odtwarzanie obrazów odpowiednio o coraz większej rozdzielczości (rys. 8).



Rys. 8. Zwiększanie rozdzielczości odtwarzania obrazu poprzez dosyłanie kolejnych danych

Jeżeli kolejność kodowania ustalona jest tak jak przedstawiono to na rys. 5 (ang. *zigzag scanning*), to w miarę jak algorytm postępuje obraz rozrasta się. Obraz rozrasta się również w miarę przepływu bitów. Odrzucenie końca takiego strumienia nie eliminuje możliwości odtworzenia obrazu mniejszego niż pierwotny.

5.2. SKALOWANIE JAKOŚCI

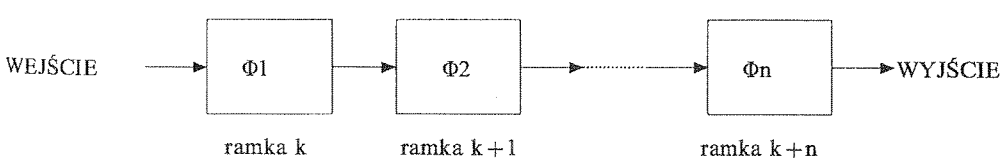
Skalowanie jakości (ang. *SNR scaling*) jest metodą polegającą na obniżaniu pasma potrzebnego do transmisji strumienia wideo poprzez obniżenie współczynnika SNR (ang. *Signal to Noise Ratio*) obrazu.

Tego rodzaju skalowalność można osiągnąć stosując wcześniej przedstawioną metodę kwantowania. Każdy cykl kwantowania to zmniejszenie błędu kwantowania kodowanego obrazu o połowę. Po zdekodowaniu błąd kwantowania wpływa na SNR obrazu wyjściowego względem wejściowego. Jeżeli z jakichś powodów pasmo transmisyjne nie będzie wystarczające do przetransmitowania obrazu o założonej przez koder jakości (ilości cykli kwantowania) obcięcie strumienia danych spowoduje odrzucenie bitów danych związanych z mniejszym przedziałem kwantowania i nie spowoduje zgubienia ramki, a jedynie pogorszenie jakości odbieranej sekwencji.

6. ARCHITEKTURA POTOKOWA MISD

Metoda kodowania Shapiro jest metodą łączącą w sobie dwie podstwowe cechy istotne dla atrakcyjności algorytmu tj. efektywność i prostotę. Obszar jej zastosowania to przede wszystkim skalowalne kodowanie obrazów dużej rozdzielczości. Zastosowanie jej do kodowania obrazów o dużych rozmiarach stwarza konieczność przetwarzania dużej ilości informacji, a to z kolei jest źródłem problemów ze zbudowaniem praktycznego kodeka działającego w trybie czasu rzeczywistego. Współcześnie działające procesory ogólnego zastosowania nie są w stanie podołać zadaniom obliczeniowym stawianym przez algorytmy tej kompresji. Nawet zastosowanie pojedynczych procesorów DSP nie rozwiązuje tego problemu. W przypadku, o którym mowa, mamy bowiem do czynienia z ilościami danych, których transfer wymaga szybkości setek Mbajtów na sekundę. Przy takich wymaganiach samo załadowanie obrazu do systemu kompresji z wykorzystaniem komercyjnie działających magistral zajęłoby czas, w ciągu którego dane powinny być przetworzone.

Rozwiązaniem problemu może być zastosownie struktury obliczeń równoległych w systemie wieloprocesorowym. Przykładem takiej struktury jest architektura potokowa, która zgodnie z klasyfikacją Flynn'a [3] może być zaliczona do architektur typu MISD [17]. W szczególności, dla omawianych w niniejszym artykule potrzeb należy zastosować architekturę dedykowaną, w której proces kodowania obrazu podzielony jest na fazy, przy czym czas przydzielony na obliczenia w każdej z faz odpowiada czasowi trwania jednej ramki obrazu. W celu zapewnienia płynności przetwarzania liczba przetwarzanych równocześnie obrazów powinna być równa ilości faz, na które został podzielony proces kompresji (rys. 9). Efektywny czas kodowania jednego obrazu jest wtedy równy czasowi jego trwania. Każda faza realizowana jest poprzez oddzielny moduł sprzętowy. Autorzy korzystali tu z wcześniejszych doświadczeń przedstawionych w [17, 18, 19, 20, 21]. Technika obliczeniowa

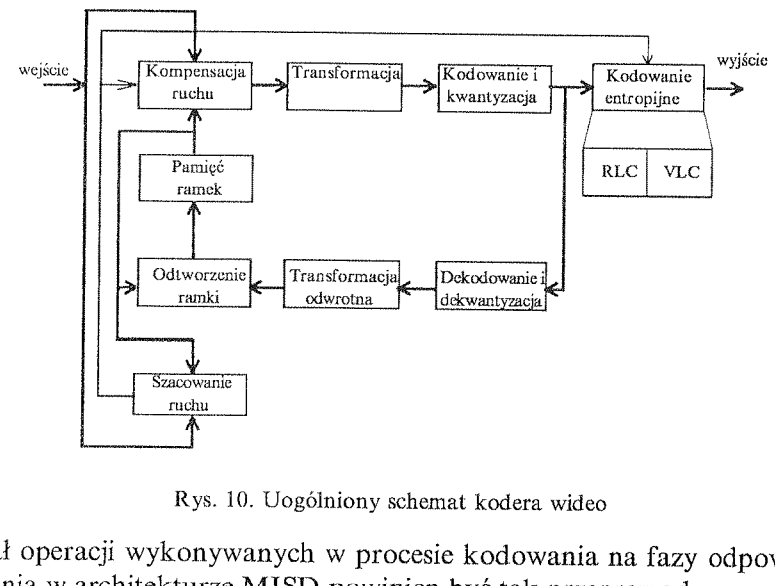


Rys. 9. Architektura przetwarzania potokowego

stosowana przy realizacji każdego modułu jest uzależnona od rodzaju operacji jaka jest w nim wykonywana. Mogą to być procesory RISK ogólnego stosowania, procesory DSP lub dedykowane procesory sprzętowe [22].

7. BUDOWA KODERA (DEKODERA) OPARTEGO NA METODZIE SHAPIRO

Metoda kodowania Shapiro nie definiuje całego procesu kodowania obrazu. Zamiast tego podaje metodę transformacji ramki i sposób kodowania uzyskanych po transformacji współczynników. Pozostałe elementy jak: kompensacja ruchu, kwantyzacja, kodowanie entropijne pozostają wspólne z innymi metodami kodowania i mogą być z uwzględnieniem potrzeb dobrane dowolnie. Na rys. 10 przedstawiony został schemat blokowy uogólnionego kodera wideo. Dla metody Shapiro proces transformacji polega na dekompozycji falkowej, a kodowanie definiuje metoda EZW [9].



Rys. 10. Uogólniony schemat kodera wideo

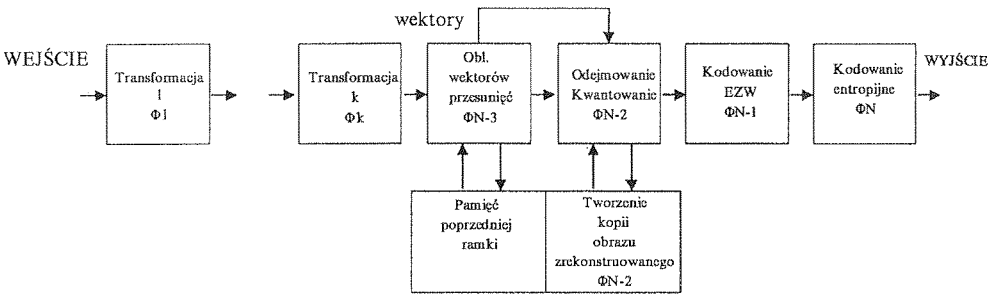
Podział operacji wykonywanych w procesie kodowania na fazy odpowiednie do zastosowania w architekturze MISD powinien być tak przeprowadzony, aby z praktycznego punktu widzenia możliwe było zakończenie obliczeń każdej z faz w czasie jaki trwa pojedyncza ramka obrazu. Zadanie dokonania takiego podziału utrudnione jest poprzez sprzężenie zwrotne niezbędne dla przeprowadzenia komensacji ruchu. Wykonanie operacji kompensacji wymaga bowiem takiego odwzorowania obrazu ramki poprzedniej jaki będzie odtworzony w dekoderze. Obraz taki jest generowany w koderze po przeprowadzeniu szeregu operacji: transformacji, kodowania i kwantyzacji, dekwantyzacji, transformacji odwrotnej i dodawania ramek. Nie wydaje się jednak możliwe, aby wszystkie wymienione operacje można było wykonać w jednej fazie, a to jest konieczne, aby przy takiej kolejności wykonywania operacji jak na rysunku, rozpocząć w odpowiednim czasie kodowanie następnej ramki.

Niezbędne jest dokonanie pewnych zmian upraszczających. Możliwe jest na przykład obliczenie wektorów przesunięć w układzie kompensacji wykorzystując nie zrekonstruowany zakodowany obraz, ale obraz oryginalny. Wynika z tego, że proces kompensacji ruchu przedstawiony w postaci bloku (rys. 10), który obejmuje obliczanie wektorów przesunięć i odejmowanie obrazów, trzeba podzielić na dwie oddzielone od siebie fazy. Jedna faza to ta, w której wykorzystujemy oryginalne obrazy do obliczania przesunięć, a druga to odejmowanie ich od obrazu skwantowanego. Jest to zysk w stosunku do poprzednich założeń zważywszy, że obliczenia wektorów przesunięć to najbardziej pracochłonna operacja w procesie kompresji obrazu.

W metodzie Shapiro wygodne i możliwe jest, jak zostanie pokazane dalej, wykonanie transformacji czyli dekompzycji falkowej przed kompensacją ruchu. Kompensacja ruchu odbywa się na obrazie już przekształconym. W formie już przekształconej wykorzystany jest również poprzedni obraz, który jest odniesieniem przy obliczaniu wektorów. Dzięki temu z pętli sprzężenia zwrotnego można usunąć dwie operacje: transformację i transformację odwrotną. Po takim zabiegu uzyskanie obrazu potrzebnego do odejmowania wymaga już tylko wykonania w pojedynczej fazie odejmowania kodowania i kwantowania. Proces kwantowania ma wpływ na jakość, z jaką kodowany obraz będzie odtworzony. Przy z góry założonym docelowym współczynniku SNR obrazu wyjściowego nie można powiedzieć z jakim skwantowany będzie obraz bez równoczesnego kodowania. Jeżeli jednak krok kwantowania wartości pikseli założyć arbitralnie jako wartość stałą na podstawie maksymalnej wartości współczynników do kodowania, to można w łatwy sposób uzyskać skwantowany obraz przez dokonaniem samego kodowania.

Jeżeli więc wymagania czasowe nie będą mogły być spełnione przy założeniu, że kwantowanie jest nieodłączną częścią kodowania możliwe jest zbudowanie działającego kodera w układzie przetwarzania potokowego, w którym ilość iteracji kwantowania jest stała. Powyższe założenia prowadzą do uproszczeń procesu kodowania dokonywanych w sposób przedstawiony na rys. 11.

Warto w tym miejscu zaznaczyć, że uproszczenie metody kodowania, jakiego dokonano celem sprowadzenia pełnego schematu kodera z rys. 10 do postaci uproszczonej z rys. 11, wiąże się z pogorszeniem parametrów jakości kompresji. Głównym celem autorów było opracowanie metody i układu do kodowania sekwencji obrazów w czasie rzeczywistym. Realizacja tego celu głównego zmusiła autorów do popełnienia szeregu uproszczeń. Jednym z nich jest uproszczenie związane z obliczaniem wektorów ruchu. Ma ono niewątpliwy wpływ na pogorszenie jakości kodowania, mierzonej mniejszym stopniem kompresji sekwencji obrazów. Autorzy nie dokonali pełnej symulacji komputerowej opracowanego kodeka, bowiem niezależnie od stopnia kompresji, układ taki, działający w trybie czasu rzeczywistego, spełnia założone oczekiwania i po jego wykonaniu zostanie zmierzony stopień kompresji. Będzie on i tak zależny od rodzaju kompresowanych obrazów i rodzaju kodowanych sekwencji wideo.



Rys. 11. Koder wideo w układzie przetwarzania potokowego

8. MODUŁY FUNKCJONALNE KODERA

8. 1. MODUŁY TRANSFORMACJI OBRAZU WEJŚCIOWEGO

Moduły transformacji obrazu wejściowego dokonują przekształcenia formy źródłowej obrazu na postać falkową. Proces przekształcania może zostać podzielony na dowolną ilość faz, ale minimalna ich liczba wynika z technicznej konieczności zapewnienia płynności operacji w potoku kodowania tzn. czas wykonywania operacji danej fazy nie może być dłuższy niż czas trwania jednej ramki.

Ponieważ przejście na postać falkową polega na filtrowaniu obrazu za pomocą zestawu filtrów, najefektywniejszą sprzętowo realizacją tego modułu jest sieć filtrów cyfrowych ogólnego zastosowania. Ponadto istniej możliwość wykorzystania układu ADV601 firmy Analog Devices, który wykonuje kompletną dekompozycję falkową lub implementacja tej funkcji w układzie programowalnym typu FPGA.

8. 2. MODUŁ KALKULACJI WEKTORÓW PRZESUNIĘĆ

Moduł ten dokonuje obliczenia przesunięcia wybranych grup punktów dla kolejnych obrazów kodowanej sekwencji. Odpowiednie obliczenia dokonywane są na produkcie LLn transformacji falkowej. W tej sytuacji w naturalny sposób przedstawiona metoda kompresji korzysta z metody kompensacji ruchu, w której dla zmniejszenia ilości koniecznych obliczeń dokonywane są one na obrazie o obniżonej rozdzielczości (ang. *Sub-Pel Motion Estimation*). Dla pozostałych części obrazu (LHk, HLk, HHk; k=0..n) należy dokonać odpowiedniego przeskalowania obliczonych wektorów. Po dokonaniu operacji obraz jest zapamiętywany w celu dokonania obliczeń przesunięć dla ramki następnej.

Istnieje wiele możliwości realizacji sprzętowej tego modułu. Ze względu na ogromną ilość potrzebnych obliczeń, jakie musi wykonać ten układ, aby można spełnić wymagania trybu czasu rzeczywsitego, konieczne jest stosowanie przetwarzania równoległego. Może ono bazować na mocnych obliczeniowo procesorach DSP

(np. TMS3020C80) lub na dedykowanym rozwiązaniu sprzętowym VLSI (np. układ STI3220 firmy SGS THOMSON). Jednak najbardziej uniwersalne jest, wybrane przez autorów, zastosowanie układów programowalnych FPGA.

8. 3. MODUŁ ODEJMOWANIA

Moduł odejmowania przeprowadza operację odjęcia dwóch następujących po sobie ramek obrazu w uwzględnieniem uprzednio skalkulowanych wektorów przesunięć. Wynik odejmowania musi zostać w tym module skwantowany tak jak efektywnie uczyni to moduł kodowania EZW. Aby uzyskać aktualną postać obrazu należy skwantowany wynik odejmowania dodać do poprzedniego obrazu ramki i wynik zachować do kolejnego odejmowania.

8. 4. MODUŁ KODOWANIA EZW I KODOWANIA ENTROPIJNEGO

Na podstawie skompensowanego ruchowo sygnału różnicowego obrazu i algorytmu kodowania Shapiro moduł kodowania EZW generuje strumień symboli kodujących informację obrazową. Moduł kodowania entropijnego przeprowadza kodowanie Huffmana lub arytmetyczne strumienia danych.

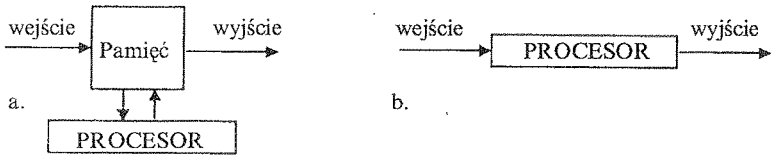
Moduł kodowania entropijnego będzie oparty na procesorach DSP (alternatywą były procesory RISK). Moduły: odejmowania, kodowania EZW są opracowywane jako dedykowane architektury sprzętowe. Wykorzystując wcześniejsze doświadczenia autorów w tym zakresie [19, 21], zostaną one zaimplementowane w układach programowalnych o bardzo dużych pojemnościach typu FPGA firmy Xilinx (seria XC4000) [23]. Przykładowo filtracje do transformacji falkowej realizowane w układach konwolwerów o oknie 3×3, implementowane są w układach XC4013 i wykorzystują 79% zasobów CLB.

9. FORMAT STRUMIENIA DANYCH OBRAZOWYCH

Z punktu widzenia wymagań czasowych omawianej struktury kodera należy zapewnić takie sekwencje i czas napływu danych do procesora, aby nie było konieczności magazynowania w jego strukturze zbyt dużej ilości danych historycznych. Jednocześnie format danych typu linia po linii, czy kolumna po kolumnie, odpowiedni do operacji filtrowania falkowego nie odpowiada operacji obliczania wektorów przesunięć. W tym drugim przypadku bardziej odpowiednie wydaje się wprowadzenie danych blok po bloku.

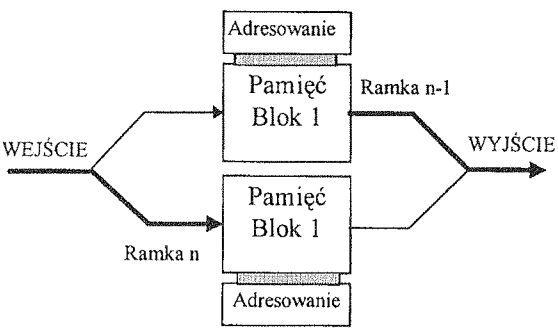
Nie jest możliwe takie zaprojektowanie formatu danych wejściowych, aby były one odpowiednie do operacji wykonywanych przez wszystkie moduły, zachowując prostą strukturę uwidocznioną na rys. 12b. Konieczne jest zatem odpowiednie formatowanie danych przed wejściem do każdego z modułów. Dlatego oprócz modułów wykonawczych wykonujących obliczenia należy zastosować odpowiednie moduły pośredniczące formatujące dane (translatory). Sposób ten, przedstawiony na

rys. 12a, zakłada, że dane wejściowe przeznaczone do obliczeń są najpierw ładowane do pamięci, następnie przetwarzane przez procesor, a na końcu wysyłane do modułu następnego.



Rys. 12. Sposoby wykonywania operacji przez moduły

Na rys. 13 przedstawiono blokowy przykład realizacji takiego modułu z wykorzystaniem pamięci. Bloki pamięci są na przemian czytane i zapisywane danymi z kolejnych ramek obrazu. Kiedy jeden z bloków jest zapisywany, dane z drugiego są czytane. Sposób adresowania różni się podczas czytania i podczas zapisywania w taki sposób, aby odpowiednio zmienić format danych. Zapewniają to odpowiednie zestawy multiplekserów pomiędzy wejściami adresowymi pamięci a układami generującymi adresy. Wielkość pamięci powinna odpowiadać wielkości jednej ramki obrazu.



Rys. 13. Sposób zmiany formatu strumienia danych

10. PODSUMOWANIE

Obecnie trwają prace nad symulacją poszczególnych bloków zaprojektowanej architektury. Platformą technologiczną dla zaprojektowanej implementacji są układy programowalne o bardzo dużych pojemnościach typu FPGA firmy Xilinx. jako narzędzie projektowo-testujące wykorzystywany jest pakiet SYNOPSYS-VHDL (ver. 1998.08) firmy Synopsys zainstalowany na stacji roboczej SUN – Ultra. Jądrem implementacji w strukturach FPGA jest pakiet XACT-ALIANCE STANDARD FOR WS (ALI-STD-WS wer. 1.4) firmy Xilinx.

Przedstawione w niniejszym artykule opracowanie jest dużym zadaniem badawczym i opisanie wszystkich interesujących szczegółów związanych z implementacją

kodeka jest niemożliwe. Aktualnie autorzy opracowują kolejny artykuł, zawierający opis szeregu szczegółowych rozwiązań związanych z implementacją w układach programowalnych FPGA. Należy w tym miejscu zwrócić uwagę, że prezentowane rozwiązanie zawiera nowatorskie podejście do metody kompresji obrazów wysokiej rozdzielczości w trybie czasu rzeczywistego, które pozwoliło zaimplementować układ kompresji obrazów w niestosownej dotychczas wieloprocesorowej architekturze typu MISD.

Przedstawione opracowanie jest ważnym uzupełnieniem do aktualnie stosowanych architektur wieloprocesorowych, ukazującym zastosowanie architektury MISD do kompresji obrazu w trybie czasu rzeczywistego, jako szczególny przykład systemów przetwarzających duże ilości danych. Wartość postawionej tezy, o istotnych walorach stosowania architektury MISD w systemach wizyjnych czasu rzeczywistego, podnosi fakt opracowania szybkich specjalizowanych elementów obliczeniowych (procesorów) i wydajnej dedykowanej struktury ich połączeń (architektury), z możliwością ich dynamicznej rekonfiguracji (w strukturach FPGA) [23]. Jest to istotny wkład autorów do światowych prac badawczych na polu dedykowanych dla użytkownika struktur obliczeniowych *CCM* (ang. *Custom Computing Machines*), których ważną cechą jest rekonfigurowalność architektury systemów *RC* (ang. *Reconfigurable Computers*) i *VC* (ang. *Virtual Computers*).

Pomimo, że kodek zbudowany jest z rekonfigurowalnych układów programowalnych, to sposób łączenia układów (architektura MISD) narzuca pewne ograniczenia związane z głębokością możliwych zmian algorytmów kodowania. Wynikają one ze struktury połączeń pomiędzy poszczególnymi rekonfigurowalnymi układami programowalnymi FPGA, tworzącymi wieloprocesorową architekturę przetwarzania potokowego typu MISD. Trudno dokładnie ocenić, które z algorytmów dadzą się dostosować do postaci, w której mogłyby być implementowane do przetwarzania sprzętowego. Tym bardziej trudność taka związana jest z ograniczeniami, jakie wnosi architektura typu MISD. Rozważanie każdego algorytmu pod tym kątem warte jest każdorazowo odrębnego opracowania i artykułu. Autorzy stwierdzają jedynie, że rekonfigurowalne układy programowalne pozwalają na zmianę parametrów przetwarzania i na zmianę algorytmów w ograniczonym zakresie. Należy jednak w tym miejscu stwierdzić, że opracowana architektura może być zbudowana jako potokowe połączenie kilku rekonfigurowalnych układów programowalnych FPGA lub może być implementowana w jednym układzie FPGA o wielkiej pojemności. Implementacja w jednym układzie programowalnym FPGA całej architektury MISD do kompresji obrazów (procesory + połączenia) jest możliwa dzięki najnowszym układom programowalnym typu Virtex firmy Xilinx, których pojemność przekroczyła już milion bramek. W takim przypadku można już mówić nie tylko o rekonfigurowalnej architekturze MISD do kompresji obrazów, ale o złożonym rekonfigurowalnym procesorze wizyjnym, w którym mogą być implementowane dowolne algorytmy, bez wcześniej wspomianych ograniczeń. Mogą one z powodzeniem pracować jako koprocesor dołączony do procesora ogólnego przeznaczenia typu Pentium lub PowerPC.

Realizacja struktur CCM w układach FPGA pozwoli w przyszłości na wyposażanie komputerów w rekonfigurowalne koprocesory sprzętowe. Dorobek na tym polu stanowi opracowanie i zbadanie architektury potokowej specjalizowanych procesorów sprzętowych, która może być zaimplementowana w dowolnym środowisku. Prezentowane opracowanie, podobnie jak zbadana wcześniej architektura potokowa [17, 21], może być implementowana w uniwersalnej architekturze rekonfigurowalnej jako wysokowydajne narzędzie do przetwarzania danych wizyjnych i innych podobnie zorganizowanych.

Rzetelne porównanie kosztów proponowanych rozwiązań z dotychczas znanymi jest bardzo trudne. Ponadto należy zwrócić uwagę, że autorzy opracowania nie aspirują do stworzenia kompletnego komercyjnego systemu kompresji obrazów w czasie rzeczywistym. Ostatecznym celem autorów jest opracowanie złożonego procesora wizyjnego, który mógłby pracować jako uniwersalny rekonfigurowalny koprocesor, zaimplementowany w układzie FPGA o wielkiej pojemności i współpracujący z procesorami ogólnego przeznaczenia. Obecny etap prac ma zatem charakter przejściowy i dlatego ocena kosztów z jednej strony jest bardzo trudna, z drugiej zaś jest niecelowa ze względu na ostateczny cel prac.

BIBLIOGRAFIA

1. V. Bhaskaran, K. Konstantinides: *Image and Video Compression Standards.* Boston, Dortrecht, Boston, Kluwer Academic Publishers 1997
2. J. Biernat: *Artmetyka komputerów.* Warszawa, PWN, 1996
3. M. J. Flynn: *Very High-Speed Computing Systems,* Proc. IEEE, Vol. 54, 1966, No 12
4. P. H. Nelis: *Data Compression Techniques for Digital Video Recording,* Delft, Technical University of Delft, 1992
5. A. N. Netravali, B. G. Haskell: *Digital Pictures. Representation, Compression and Standards,* New York-London, Plenum Press, 1995
6. B. Pochpień: *Arytmetyka systemów cyforwych,* Gliwice, Wydawnictwo Politechniki Śląskiej, 1997
7. M. Rabbani: *Selected Papers on Image Coding and Compression.* Bellingham, SPIE Optical Engineering Press, 1992
8. M. Rabbani, P. W. Jones: *Digital Image Compression Techniques.* Bellingham, SPIE Optical Engineering Press, 1991
9. J. M. Shapiro: *Embedded Image Coding Using Zerotrees of Wavelet Coefficients.* Transaction on Signal Processing, Vol. 41, No 12, December 1993
10. W. Skarbek: *Metody reprezentacji obrazów cyfrowych.* Warszawa, PLJ, 1993
11. W. Skarbek: *Multimedia — Algorytmy i standardy kompresji.* Warszawa, PLJ, 1998
12. J. A. Stover: *Image and text compressin.* Boston, Kluwer Academic Publisher, 1992
13. J. R. Sullivan, M. Rabbani, B. M. Dawson: *Image Algorithms and Techniques.* Bellingham, SPIE — The International Society for Optical, 1992
14. M. A. Talkap: *Digital Video Processing.* Prentic Hall NJ, 1995
15. B. J. Thomson: *Image Coding and Compression.* Wasington, SPIE Optical Engineering Press, 1992
16. P. P. Vaidyanathan: *Multirate systems and filter banks.* Englewood Cliffs, Prentice Hall, 1993
17. K. Wiatr: *Architektura specjalizowanych procesorów potokowych do obróbki wstępnej danych wizyjnych w czasie rzeczywistym.* Kwartalnik Elektroniki i Telekomunikacji, Warszawa, PWN 1997, z. 1, ss. 121—148

18. K. W i a t r: *Pipelined Architecture of Reconfigurable Specialised Processors for a Real-Time Data Pre-Processing*. Proc. of the IEEE International Conference of Signal Processing — ICSP-96, pp. 649 – 652
19. K. W i a t r: *Dedicate Hardware Processing for a Real-Time Image Data Pre-Processing Omplemented in FPGA Structure*. Proc. of the 9th International Conference on Image Analysis and Processing — ICIAP'97, Springer-Verlag 1997, Vol. II, pp. 69 – 75
20. K. W i a t r: *Pipelined Architecture of Specialised Reconfigurable Processors in FPGA Structures for Real-Time Image Data Pre-Processing*. Proc. of the EUROMICRO International Conference: Digital Systems Desing: Architectures, Methods and Tools IEEE Computer Press 1998, pp. 131 – 138
21. K. W i a t r: *Dedicate System Architecture for Parallel Image Computation used Specialised Hardware Processors Implemented in FPGA Structure*. International Journal of Parallel and Distributed Systems and Networks, Pittsburgh PA 1998, pp. 161 – 168
22. K. W i a t r, P. R u s s e k: *Sprzętowe architektury kompresji obrazu wizyjnego dla potrzeb przetwarzania w czasie rzeczywistym*. Kwartalnik Elektroniki i Telekomunikacji PWN T. 44, z. 2, Warszawa 1998, ss. 245 – 258
23. XILINX: *The Programmable Logic Data Book*. San Jose CA, Xilinx Inc. 1998

K. WIATR, P. RUSSEK

HARWARE IMPLEMENTATION OF SCALING SYSTEM FOR SUPER HIGH DEFINITION IMAGE IN REAL-TIME CODING

Summary

This paper deals withe the problems related to real-time image coding. In particular authors presents scaling system fo SHD (Super High Definition) image coding. The work concerns with Embedded Zero-tree Wavelet (EZW) coding algorithm. EZW method in implemented in multiprocessors MISD (Multiple-stream Single Data-stream) architecture by authors. MISD architecture application guarantee very high speed realisation of image coding algorithm. This aspect recommends MISD architecture for SHD image coding in real-time computation. The subject of coded datastream scalability is also mentioned. Authors implement presented system in the high capacity FPGA structures. The presented work is a contribution to worldwide intense research on developing dedicated „Custom Computing Machines". The work is supported by the Polish Science Committee.

Key words: image processing, image compression, real-time systems, programmable logic device.

Syntez
ty

Stru
zawier
Strukt
poszcz
części,
stawo
zawier
tykule
ciu o
algory
dla st
syntez
uzyska

Słow

Jądro w
MAX9000
firmy Xilin
określoną
poszczegól
tych w teg
syntezy. Is
p-implikant
struktury,
wykorzysta

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI, 2000, 46, z. 1, ss. 81–90

Synteza logiczna wielopoziomowych układów w strukturach typu PAL z trójstanowymi buforami wyjściowymi

DARIUSZ KANIA

Instytut Elektroniki, Politechnika Śląska, Gliwice

Otrzymano 1999.09.01
Autoryzowano 1999.11.15

Struktury CPLD często zawierają bloki logiczne typu PAL. każdy taki blok logiczny zawiera programowalną matrycę AND, która dołączona jest do komórki wyjściowej. Struktura taka posiada ściśle określoną liczbę termów (składników sum) dołączonych do poszczególnych wyjść. Problem podziału całego projektowanego układu na odpowiednie części, które są realizowane na pojedynczych blokach typu PAL, jest jednym z podstawowych problemów syntezy. Bloki logiczne typu PAL występujące w strukturach CPLD zawierają zwykle dodatkowe zasoby logiczne, które mogą ułatwić proces podziału. W artykule przedstawiono prostą metodę syntezy logicznej wielopoziomowych układów w oparciu o bloki logiczne typu PAL zawierające trójstanowe bufory wyjściowe. Opracowany algorytm, zaimplementowany w systemie PALDec (moduł systemu Decomp przeznaczony dla struktur typu PAL) został wykorzystany do podziału układów testowych. Wyniki syntezy logicznej dla układów typu PAL firmy Altera zostały porównane z wynikami uzyskanymi za pomocą systemu MAX+PLUS II.

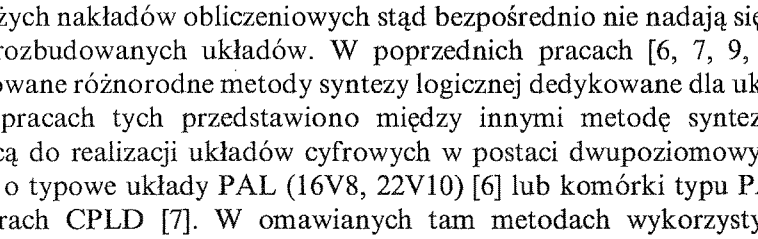
Słowa kluczowe: synteza logiczna, podział, PLD

1. WPROWADZENIE

Jądro większości układów CPLD stanowi struktura PAL (MAX5000, MAX7000, MAX9000 firmy Altera MACH, firmy AMD, pLSI firmy Lattice, XC7000, XC9500 firmy Xilinx, AVT2500, AVT5000 firmy Atmel itp.) [1]. Struktura ta posiada ściśle określoną (najczęściej stałą) liczbę termów (składników sum) dołączonych do poszczególnych komórek wyjściowych. Sposób efektywnego wykorzystania zawartych w tego typu strukturach iloczynów, stanowi jeden z kluczowych problemów syntezy. Istota problemu sprowadza się do realizacji funkcji będącej sumą p-implikantów na blokach logicznych zawierających k-termów, gdy $p>k$. Istnieją struktury, w których zastosowano sprzętowe mechanizmy ułatwiające efektywne wykorzystanie iloczynów, poprzez ich nierównomierny rozdział pomiędzy komórki

wyjściowe. Występują one między innymi w strukturach MACH firmy AMD (logic allocator), w układach firmy Xilinx oraz Altera (MAX 7000, MAX9000). Sprzętowe zasoby układów programowalnych umożliwiające nierównomierny podział termów dołączonych do poszczególnych komórek wyjściowych (logic allocator w układach MACH, mixed mode w układach pLSI itp.) mimo swoich niewątpliwych zalet nie umożliwiają jednak realizacji każdej funkcji w jednym bloku. W tej sytuacji konieczna jest dodatkowa ekspansja liczby termów. Jedna z metod takiej ekspansji może wykorzystywać występujące często w tego typu strukturach bufory trójstanowe [1, 3].

Ekspansja liczby termów łączy się ściśle z problemem dekompozycji, rozumianej jako podział układu na poszczególne bloki, zawierające k-termów. Algorytmy dekompozycji najczęściej dedykowane są jednak dla układów bramkowych [4] lub struktur tablicowych (ang. Look-up Table) [2, 5, 12, 13]. Możliwe jest również ich wykorzystywanie do realizacji układów cyfrowych w strukturach typu PAL [2, 9, 11, 12]. Metodę wykorzystania dekompozycji funkcjonalnej do realizacji układów w strukturach typu PAL przedstawiono w pracy [9]. Metody te wymagają jednak bardzo dużych nakładów obliczeniowych stąd bezpośrednio nie nadają się do syntezy logicznej rozbudowanych układów. W poprzednich pracach [6, 7, 9, 10] zostały zaprezentowane różnorodne metody syntezy logicznej dedykowane dla układów typu PAL. W pracach tych przedstawiono między innymi metodę syntezy logicznej prowadzącą do realizacji układów cyfrowych w postaci dwupoziomowych struktur w oparciu o typowe układy PAL (16V8, 22V10) [6] lub komórki typu PAL zawarte w strukturach CPLD [7]. W omawianych tam metodach wykorzystywano trójstanowe bufory wyjściowe do ekspansji liczby termów.



Rys. 1. (I) Podstawowa struktura bloku logicznego typu PAL,
(II) Symbol bloku logicznego typu PAL zawierającego 3 iloczyny

Celem niniejszego artykułu jest przedstawienie prostej metody syntezy wielopoziomowych struktur w oparciu o bloki logiczne typu PAL zawierające trójstanowe bufory wyjściowe (Rys. 1). Zaletą proponowanej metody jest jej prostota wynikająca z uproszczenia algorytmu wyszukiwania wektorów podziału [7] i ograniczenia ich do pojedynczych zmiennych oraz możliwość jej wykorzystywania w strukturach zawierających bufory trójstanowe tylko w wybranych komórkach. Konsekwencją uproszczenie algorytmu jest konieczność wprowadzenia wewnętrznych sprzężeń zwrotnych prowadzących ostatecznie do uzyskiwania wolniejszych, niż przedstawionych w pracy [6, 7] struktur wielopoziomowych. Ze względu na fakt, że liczba wejść struktur CPLD jest zwykle wystarczająca, proponowany model syntezy logicznej nie obejmuje problemu podziału wejść.

2. PODSTAWY TEORETYCZNE

Niech f będzie n-wyjściową funkcją logiczną odwzorowującą zbiór B^n w zbiór B^1 tzn. $f: B^n \rightarrow B^1$, gdzie $B=\{0, 1\}$. Zminimalizowana funkcja f może być opisana przez zbiór wielowyjściowych implikantów zawierających część wejściową złożoną z elemetów $\{0, 1, -\}$ oraz część wyjściową złożoną z elementów $\{0, 1\}$.

Niech $\mathbf{A}_y^{(0 \text{ lub } 1)}=\{\mathbf{A}_{y(l-1)}, \ldots, \mathbf{A}_{y1}, \mathbf{A}_{y0}\}$ będzie zbiorem wektorów, dla których funkcja $y=f(i_{n-1}, \ldots, i_1, i_0)$ przyjmuje stan aktywny 0 lub 1. Wybrany wektor zbioru $\mathbf{A}_y$ składa się z n-elementów $A_y=(a_{n-1}, \ldots, a_1, a_0)$ odpowiadających zmiennym wejściowym $i_{n-1}, \ldots, i_1, i_0$, przy czym każdy element $a_i \in \{0, 1, -\}$, $(i=0, 1, \ldots, n-1)$. Każdy wektor zbioru $\mathbf{A}_y$ pokrywa 2^r stanów I_j tworzących zbiór stanów $\mathbf{I}$ o mocy 2^r $(\overline{\overline{\mathbf{I}}}=2^r)$, gdzie r jest liczbą elementów $a_i=\{-\}$, przy czym $(i=0, 1, \ldots, n-1)$.

Definicja 1:

Dwa wektory $A_{ys}=(a_{(n-1)s}, \ldots, a_{1s}, a_{0s})$ i $A_{yt}=(a_{(n-1)t}, \ldots, a_{1t}, a_{0t})$, których elementy $a_{is}, a_{it} \in \{0, 1, -\}$ takie, że A_{ys} *pokrywa* 2^{rs} *stanów tworzących zbiór stanów* $\mathbf{I}_s$, a A_{yt} pokrywa 2^{rt} stanów tworzących zbiór stanów $\mathbf{I}_t$, nazywanym wektorami półpokrywającymi wtedy i tylko wtedy, gdy $\overline{\overline{\mathbf{I}_s \cap \mathbf{I}_t}}=2^{rs-1}$ lub 2^{rt-1}

np.: $A_{ys}=(-\ 0\ 0\ -\ 1)_{i_4 i_3 i_2 i_1 i_0}$ i $A_{yt}=(1\ 0\ 0\ 1\ -)_{i_4 i_3 i_2 i_1 i_0}$

Definicja 2:

Wyróżnikiem $\Delta_{i_j}^0$ nazywamy liczbę dziesiętną równą mocy zbioru $\mathbf{A}_y^{i_j=0}$, gdzie $\mathbf{A}_y^{i_j=0}$ jest podzbiorem zbioru $\mathbf{A}_y$, zawierającym wektory, dla których zmienna i_j przyjmuje wartość $\{0\}$ lub $\{-\}$.

Definicja 3:

Wyróżnikiem $\Delta_{i_j}^1$ nazywamy liczbę dziesiętną równą mocy zbioru $\mathbf{A}_y^{i_j=1}$, gdzie $\mathbf{A}_y^{i_j=1}$ jest podzbiorem zbioru $\mathbf{A}_y$, zawierającym wektory, dla których zmienna i_j przyjmuje wartość $\{1\}$ lub $\{-\}$.

Przykład 1:

$$\text{Jeżeli } \mathbf{A}_y=\left\{\begin{matrix} 0 & - & 0 & 1 & 1 \\ - & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & - & 1 \\ - & - & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & - \end{matrix}\right\}_{i_4 i_3 i_2 i_1 i_0}$$

$$\text{to } \Delta_{i_3}^0=\overline{\overline{\mathbf{A}_y^{i_3=0}}}=3 \text{ ponieważ } \mathbf{A}_y^{i_3=0}=\left\{\begin{matrix} 0 & - & 0 & 1 & 1 \\ - & 0 & 0 & 0 & 1 \\ - & - & 0 & 1 & 0 \end{matrix}\right\}_{i_4 i_3 i_2 i_1 i_0}$$

$$\text{natomiast to } \Delta_{i_3}^1=\overline{\overline{\mathbf{A}_y^{i_3=1}}}=4 \text{ ponieważ } \mathbf{A}_y^{i_3=1}=\left\{\begin{matrix} 0 & - & 0 & 1 & 1 \\ 0 & 1 & 0 & - & 1 \\ - & - & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & - \end{matrix}\right\}_{i_4 i_3 i_2 i_1 i_0}$$

Zmienna i_3 dzieli zbiór $\mathbf{A}_y$ na dwa podzbiory $\mathbf{A}_y^{i_3=0}$ i $\mathbf{A}_y^{i_3=1}$. Nazwijmy ją zmienną podziału zbioru $\mathbf{A}_y$.

3. STRATEGIA SYNTEZY UKŁADÓW CYFROWYCH NA BAZIE BLOKÓW LOGICZNYCH TYPU PAL

Ogólna koncepcja systemu syntezy obejmuje następujące etapy:

- Dwupoziomową minimalizację z rozszczepieniem wektorów
- Procedurę wyszukiwania zmiennej podziału i_j, prowadzącą do podziału zbioru, dla których funkcja y przyjmuje stan aktywny, na dwa podzbiory $\mathbf{A}_y^{i_j=0}$ i $\mathbf{A}_y^{i_j=1}$
- Realizację uzyskanych podzbiorów w oparciu o bloki logiczne typu PAL zawierające k termów

3A. DWUPOZIOMOWA MINIMALIZACJA Z ROZSZCZEPIENIEM WEKTORÓW

Celem dwupoziomowej minimalizacji z roszczepieniem wektorów jest uzyskanie zbioru $^*\mathbf{A}_y$ o minimalnej mocy, zawierającego minimalną liczbę elementów $a_i=\{-\}$. Po klasycznej dwupozimowej minimalizacji wykonywana jest procedura rozszczepienia wektorów. Niech przykładowo wektor $A_{ys} \in \mathbf{A}_y$ równy $A_{ys}=(a_{(n-1)s}, \ldots, a_{1s}, a_{0s})$ pokrywa 2^{r_s} stanów tworzących zbiór stanów $\mathbf{I}_s$, natomiast wektor $A_{yt} \in \mathbf{A}_y$ równy $A_{yt}=(a_{(n-1)t}, \ldots, a_{1t}, a_{0t})$ pokrywa 2^{r_t} stanów tworzących zbiór stanów $\mathbf{I}_t$. W przypadku, gdy dwa wektory A_{ys} i A_{yt} są wektorami półpokrywającymi to okazuje się, że istnieje możliwość modyfikacji jednego z wektorów nie zmieniając zbioru $\mathbf{I}_s \cup \mathbf{I}_t$. Wyszukiwanie par wektorów półpokrywających polega na analizie par uporządkowanych $<a_{is}, a_{it}>$, gdzie $i=0, 1, \ldots, n-1$. Jeżeli założymy, że wektor A_{ys} zawiera nie mniejszą liczbę elementów $a_{is}=\{-\}$ niż wektor A_{yt} to wektory A_{ys} i A_{yt} są wektorami półpokrywającymi wtedy, gdy wśród elementów zbioru par uporządkowanych występują pary należące do zbioru $\{<00>, <11>, <-0>, <-1>, <-->\}$ i tylko jedna para $<a_{is}^*, a_{it}^*>$ należąca do zbioru $\{<1->, <0->\}$. Modyfikacja, nie zmieniająca zbioru stanów $\mathbf{I}_s \cup \mathbf{I}_t$ pokrytych przez wektory A_{ys} i A_{yt}, polega na zmianie elementu a_{it}^* zgodnie z zasadą: jeżeli $a_{is}^*=1$ to $a_{it}^*:=0$, natomiast $a_{is}^*=0$ to $a_{it}^*:=1$.

Przedstawiona zasada jest podstawą dwupozimowej minimalizacji z rozszczepieniem wektorów, która polega na poszukiwaniu wszystkich par wektorów półpokrywających i przekształceniu jednego wektora należącego do tej pary. Taka strategia minimalizacji znacząco wpływa na efektywność podziału układu na poszczególne bloki logiczne typu PAL [8].

Przykład 2:

Rozważmy funkcję opisaną zbiorem wektorów $\mathbf{A}_y$, który jest wynikiem klasycznej dwupoziomowej minimalizacji. Po dwupoziomowej minimalizacji z rozszczepieniem wektorów uzyskujemy zbiór $^*\mathbf{A}_y$.

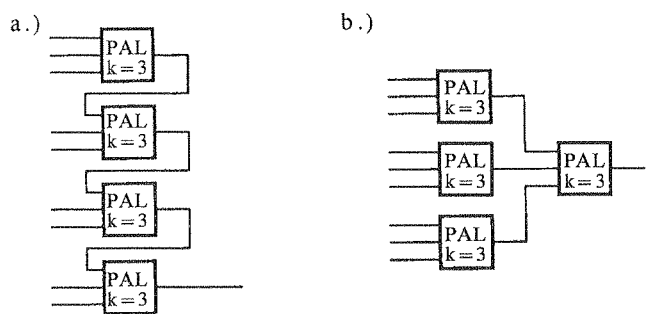
$$\mathbf{A}_y=\begin{Bmatrix} 0&-&0&-&0&1&0\\ 0&-&1&1&1&-&0\\ 0&-&-&1&0&0&0\\ 1&-&1&0&1&0&0\\ 0&-&0&0&1&-&0\\ -&1&-&0&1&-&-\\ 0&-&-&0&-&1&-\\ 0&1&-&-&-&-&- \end{Bmatrix}_{i_6 i_5 i_4 i_3 i_2 i_1 i_0} \qquad ^*\mathbf{A}_y=\begin{Bmatrix} 0&0&0&1&0&1&0\\ 0&0&1&1&1&-&0\\ 0&0&-&1&0&0&0\\ 1&0&1&0&1&0&0\\ 0&0&0&0&1&0&0\\ 1&1&-&0&1&-&-\\ 0&0&-&0&-&1&-\\ 0&1&-&-&-&-&- \end{Bmatrix}_{i_6 i_5 i_4 i_3 i_2 i_1 i_0}$$

3B. PROCEDURA WYSZUKIWANIA ZMIENNEJ PODZIAŁU

Celem procedury wyszukiwania zmiennej podziału wektorów jest podział zbioru $^*\mathbf{A}_y$, dla których funkcja przyjmuje atan aktywny na dwa podzbiory $\mathbf{A}_y^{i_j=0}$ i $\mathbf{A}_y^{i_j=1}$, takie, że realizacja funkcji wykorzystuje minimalną liczbę bloków logicznych typu PAL zawierających k termów. W przypadku większej liczby rozwiązań spełniającej warunek minimalnej liczby wykorzystywanych bloków logicznych, wybierane jest rozwiązanie, które wprowadza mniejszą liczbę poziomów logicznych (AND-OR).

Niech δ_j^0 (δ_j^1) oznacza liczbę bloków logicznych potrzebnych do realizacji wektorów tworzących zbiór $\mathbf{A}_y^{i_j=0}(\mathbf{A}_y^{i_j=1})$. Niech $\lceil x \rceil$ oznacza najmniejszą liczbę całkowitą nie mniejszą od x.

Jeżeli $\Delta_{i_j}^0 > k$ lub $\Delta_{i_j}^1 > k$, gdzie k jest liczbą termów zawartych w pojedynczym bloku logicznym typu PAL, to realizacja wektorów należących do zbiorów $\mathbf{A}_y^{i_j=0}$ lub $\mathbf{A}_y^{i_j=1}$ wymaga użycia więcej niż jednego bloku. Konsekwencją tego jest wprowadzenie wewnętrznych sprzężeń zwrotnych umożliwiających ekspansję liczby termów. Fakt ten sprawia, że mamy do dyspozycji niejako jeden blok zawierający k termów i dowolną liczbę bloków zawierających $(k-1)$ termów. W tej sytuacji realizacja wybranego podzbioru wektorów wymaga użycia odpowiednio $\delta_j^0 = \left\lceil \frac{\Delta_{i_j}^0 - k}{k-1} \right\rceil + 1$ i $\delta_j^1 = \left\lceil \frac{\Delta_{i_j}^1 - k}{k-1} \right\rceil + 1$ bloków logicznych typu PAL zawierających k termów.



Rys. 2. Strategie ekspansji liczby termów wykorzystujące wewnętrzne sprzężenia zwrotne

Istnieją dwa sposoby ekspansji liczby termów umożliwiające ekspansję poprzez wprowadzenie wewnętrznych sprzężeń zwrotnych (Rys. 2). Rysunek przedstawia realizację funkcji, która jest sumą dziewięciu implikantów, w oparciu o bloki logiczne typu PAL zawierające trzy termy. Oczywisty jest fakt, że strategia ekspansji przedstawiona na Rys. 2b, wykorzystuje identyczną liczbę bloków co strategia zobrazowana na Rys. 2a, wprowadzając mniejszą liczbę poziomów logicznych.

W ogólnym przypadku realizacja funkcji będącej sumą $\Delta_{i_j}^0(\Delta_{i_j}^1)$ implikantów za pomocą struktury $\xi_j^0(\xi_j^1)$ poziomowej, gdzie $\xi_j^0(\xi_j^1)$ jest minimalną liczbą naturalną spełniającą nierówność $\Delta_{i_j}^0 \leqslant k^{\xi_j^1}(\Delta_{i_j}^0 \leqslant k^{\xi_j^1})$.

3C. ALGORYTM WYSZUKIWANIA ZMIENNEJ PODZIAŁU

- Dla poszczególnych zmiennych wyznaczane są współczynniki $\Delta_{i_j}^0$ i $\Delta_{i_j}^1$.
- Na podstawie współczynników $\Delta_{i_j}^0$, $\Delta_{i_j}^1$ określane są wartości δ_j^0, δ_j^1 oraz ξ_j^0, ξ_j^1.
- Jako zmienna podziału wybierana jest zmienna i_j, dla której $\delta_j^0+\delta_j^1$ przyjmuje wartość minimalną. W przypadku, gdy istnieje kilka zmiennych, dla których wyrażenie $\delta_j=\delta_j^0+\delta_j^1$ przyjmuje minimum, jako zmienną podziału wybieramy zmienną, dla której $\xi_j=\max(\xi_j^0, \xi_j^1)$ przyjmuje wartość minimalną
- Dokonywany jest podział zbioru $^*\mathbf{A}_y$, na dwa podzbiory $\mathbf{A}_y^{i_j=0}$ i $\mathbf{A}_y^{i_j=1}$, po czym uzyskane podzbiory realizowane są w oparciu o bloki logiczne typu PAL zawierające k termów.

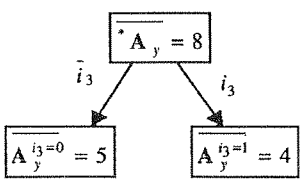
Przykład 3:

Wyznaczmy zgodnie z przedstawionym algorytmen zmienną podziału dla funckji z przykładu 2 zakładając, że chcemy zrealizować funkcję na blokach logicznych typu PAL zawierających 3 termy. Po dwupoziomowej minimalizacji z rozszczepieniem wektorów uzyskujemy zbiór wektorów $^*\mathbf{A}_y$. Wyznaczmy dla poszczególnych zmiennych wartości współczynników $\Delta_{i_j}^0$, $\Delta_{i_j}^1$, δ_j^0, δ_j^1, oraz ξ_j^0, ξ_j^1.

$$^*\mathbf{A}_y = \begin{Bmatrix} 0\,0\,0\,1\,0\,1\,0 \\ 0\,0\,1\,1\,1\,-\,0 \\ 0\,0\,-\,1\,0\,0\,0 \\ 1\,0\,1\,0\,1\,0\,0 \\ 0\,0\,0\,0\,1\,0\,0 \\ 1\,1\,-\,0\,1\,-\,- \\ 0\,0\,-\,0\,-\,1\,- \\ 0\,1\,-\,-\,-\,-\,- \end{Bmatrix}_{i_6 i_5 i_4 i_3 i_2 i_1 i_0}$$

	i_6	i_5	i_4	i_3	i_2	i_1	i_0
	0	0	0	1	0	1	0
	0	0	1	1	1	-	0
	0	0	-	1	0	0	0
	1	0	1	0	1	0	0
	0	0	0	0	1	0	0
	1	1	-	0	1	-	-
	0	0	-	0	-	1	-
	0	1	-	-	-	-	-
$\Delta_{i_j}^0$	6	6	6	5	4	6	8
$\Delta_{i_j}^1$	2	2	6	4	6	5	3
δ_j^0	3	3	3	2	2	3	4
δ_j^1	1	1	3	2	3	2	1
ξ_j^0	2	2	2	2	2	2	2
ξ_j^1	1	1	2	2	2	2	1
δ_j	**4**	**4**	6	**4**	5	5	5
ξ_j	**2**	**2**	2	**2**	2	2	2

Zbiór zmiennych spełniających warunki algorytmu zawiera zmienne i_6, i_5, lub i_3. Jedna z tych zmiennych może zostać wybrana jako zmienna podziału termów. Załóżmy, że zmienną podziału termów będzie zmienna i_3, która dzieli zbiór $^*\mathbf{A}_y$ na dwa podzbiory



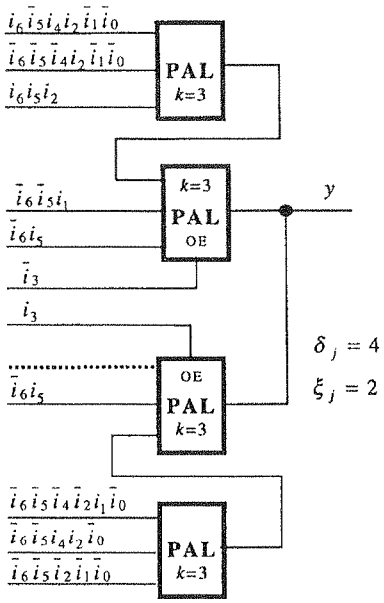
$$\mathbf{A}_y^{i_3=0}=\begin{Bmatrix}1&0&1&0&1&0&0\\0&0&0&0&1&0&0\\1&1&-&0&1&-&-\\0&0&-&0&-&1&-\\0&1&-&-&-&-&-\end{Bmatrix}_{i_6i_5i_4i_3i_2i_1i_0} \qquad \mathbf{A}_y^{i_1=1}=\begin{Bmatrix}0&0&0&1&0&1&0\\0&0&1&1&1&-&0\\0&0&-&1&0&0&0\\0&1&-&-&-&-&-\end{Bmatrix}_{i_6i_5i_4i_3i_2i_1i_0}$$

3D. REALIZACJA PODZBIORÓW $\mathbf{A}_y^{i_j=0}$ I $\mathbf{A}_y^{i_j=1}$ W OPARCIU O BLOKI LOGICZNE TYPU PAL ZAWIERAJĄCE k TERMÓW

Realizacja podukładów opisanych przez podzbiory $\mathbf{A}_y^{i_j=0}$ i $\mathbf{A}_y^{i_j=1}$ oparta jest na dwóch zasadach ekspansji liczby termów. Pierwsza wykorzystuje sprzężenia zwrotne i została przedstawiona w punkcie 3b, natomiast druga wykorzystuje do ekspansji bufory trójstanowe [6, 7].

Przykład 4:

Rysunek 3 przedstawia realizację funkcji z przykladu 2 po wyborze zmiennej podziału przedstawionym w przykładzie 3.



Rys. 3. Realizacja przykładowej funkcji na blokach logicznych typu PAL zawierających 3 termy

4. WYNIKI EKSPERYMENTÓW

Przedstawiony w pracy algorytm został zaimplementowany w systemie PALDec (moduł systemu Decomp przeznaczony do syntezy układów cyfrowych w strukturach programowalnych typu PAL) i stanowi uproszczenie algorytmu przedstawionego

w pracy [7]. W celu zweryfikowania opracowanej metody przeprowadzono syntezę układów testowych (ang. benchmark) dla popularnych struktur typu PAL firmy Altera. Uzyskane wyniki (liczbę niezbędnych do realizacji bloków logicznych o zadanej liczbie termów) porównano z wynikami syntezy, uzyskanymi za pomocą systemu MAX+PLUS II. Podział przeprowadzono dla bloków logicznych typu PAL występujących w strukturach EP1810 oraz układów rodziny MAX5000. Wyniki przedstawione są w tabelach 1 i 2.

Tablica 1

Wyniki syntezy dla struktur EP1810
(liczba wykorzystanych bloków logicznych typu PAL o zadanej liczbie termów ($k=8$), niezbędnych do realizacji odpowiednich układów testowych)

$k=8$	5×p1	bw	Con1	F51m	Inc	Misex1	Misex2	Rd53	Rd53'	Rd73	Rd73'	Root	Squar5	Sqn	Sqr6	Xor5
PALDec	15	28	2	15	11	7	18	5	3	22	12	12	8	8	14	2
MAX+PLUS II	20	28	2	18	16	7	18	7	4	21	13	14	8	11	17	3

Tablica 2

Wyniki syntezy dla struktur MAX5000

$k=4$*	5×p1	bw	Con1	F51m	Inc	Misex1	Misex2	Rd53	Rd53'	Rd73	Rd73'	Root	Squar5	Sqn	Sqr6	Xor5
PALDec*	26	40	3	25	19	13	19	12	6	49	27	21	10	15	21	6
MAX+PLUS II**	22	40	3	20	31	13	19	11	7	39	27	26	11	16	28	4

* blok logiczny zawarty w strukturach MAX5000 realizuje funkcję $out=A_{y1}\oplus(A_{y2}+A_{y3}+A_{y4})$ stąd jeżeli wybierzemy wektor A_{y1}, dla którego spełniony jest warunek $A_{y1}(A_{y2}+A_{y3}+A_{y4})=0$ to w jednym bloku możemy zrealizować 4 wektory. Wykorzystywany sposób minimalizacji tzn. dwupoziomowa minimalizacja z rozszczemianiem wektorów, znacznie upraszcza wybór wektora A_{y1} rozłącznego z wyrażeniem $(A_{y2}+A_{y3}+A_{y4})$.

** strategia minimalizująca liczbę bloków logicznych, nie wykorzystująca ekspanderów NAND

Wyniki eksperymentów jednoznacznie pokazują, że prezentowana metoda jest konkurencyjna w stosunku do syntezy zaimplementowanej w systemie MAX+PLUS II. Należy zwrócić uwagę, że nawet dla struktur MAX5000 uzyskiwane są lepsze rozwiązania. Fakt ten ma duże znaczenie, ponieważ z wielu przeprowadzonych doświadczeń wynika, że mocną stroną systemu MAX+PLUS II jest efektywne wykorzystanie bramki XOR, występującej w bloku logicznym. Realizacja układów testowych Rd53' oraz Rd73', które powstały z układów Rd53 i Rd73 po wyeliminowaniu funkcji XOR, jednoznacznie potwierdza tę tezę.

5. PODSUMOWANIE

Prezentowana metoda syntezy dla programowalnych układów typu PAL zawierających wyjściowe bufory trójstanowe prowadzi do realizacji układów w postaci wielopoziomowych struktur logicznych. Stanowi ona uproszczenie algorytmu prezentowanego w pracy [7]. Niewątpliwie dużą zaletą prezentowanej metody jest jej prostota, wynikająca z ograniczenia rozmiarów wektorów podziału [7] do pojedynczych zmiennych. Konsekwencją tego uproszczenia jest wynik tzn. wielopoziomowa struktura uzyskiwanego układu. Należy podkreślić jeszcze jedną sprawę. Przedstawiona metoda syntezy może być wykorzystywana w strukturach typu PAL, które zawierają trójstanowe bufory wyjściowe tylko w wybranych blokach logicznych.

BIBLIOGRAFIA

1. *AMD, Altera, Atmel, Lattice.* Xilinx Data Book.
2. B. Babba, M. Crastes, G. Saucier: *Input driven synthesis on PLDs and PGAs.* The European Conference on Design Automation, Brussels (Belgium), March 1992
3. M. Bolton: *Digital Systems Design with Programmable Logic.* Addison-Wesley Publishing Company, 1990, pp. 133–140
4. S. D. Brown, R. J. Francis, J. Rose, Z. G. Vernesic: *Field Programmable Gate Arrays, Kluwer Acedemic Publishers.* Boston/London/Dotdrecht 1993, pp. 45–86
5. J. A. Brzozowski, T. Łuba: *Decomposition of Boolean Functions Specified by Cubes.* University of Waterloo Computer Science Departament, CS-97-01, January 1997
6. D. Kania: *Two-level logic synthesis on PALs.* Electronics Letters, 1999, Vol. 35, No. 11, pp. 879–880
7. D. Kania: *Two-level logic synthesis of PAL-based CPLD and FPGA using decompositon.* Euromicro, Milan, 1999, pp. 278–281
8. D. Kania: *Impact of logic minimization on term partitioning effectiveness.* XXII National Conference on Circuits Theory and Electronic Networks, Warszawa – Stare Jabłonki, October, 1999
9. D. Kania: *Synteza logiczna dla układów CPLD typu PAL wykorzystująca dekompozycję.* Kwartalnik Elektroniki i Telekomunikacji, z. 3–4, 1999, ss. 445–454
10. D. Kania: *Efektywna metoda realizacji zespołu funkcji w strukturach typu PAL.* Kwartalnik Elektroniki i Telekomunikacji, z. 3–4, 1999, ss. 433–444
11. T. Łuba: *Multi-level logic synthesis based on decomposition.* Microporocessors and Microsystems, Vol. 18, No 8, October 1994, pp. 429–437
12. G. Saucier, P. Sicard, L. Bouchet: *Multi-level synthesis on PAL's.* European Design Automation Conference, Glasgow, March 1990, pp. 542–546
13. 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

D. KANIA

MULTI-LEVEL LOGIC SYNTHESIS ON PAL-BASED DEVICES WITH THREE STATE OUTPUT BUFFERS

Summary

Complex Programmable Logic Devices (CPLD) often includes the PAL-based logic blocks. Each logic block contains a programmable AND-array that feeds its macrocells. These structures have exactly defined number of terms connected to the individual output. The problem of partition of the whole devices under design into suitable parts, which can be implemented as single PAL-based logic blocks, containing the limited number of terms, is one of basic problems of the synthesis. The logical blocks included into CPLD structures contain usualy additional logic resources that facilitate the partitioning process. The simple method of multi-level logic synthesis on PAL-based logic blocks with three-state output buffers is presented in this paper. Developed algorithms, implemented within the PALDec (PALDecomp) system, have been used for partitioning the benchmark circuits. Results of logic synthesis for Altera PAL-based devices have been compared to the results ibtained with the aid of MAX+PLUS II software.

Key words: logic synthesis, partition, PLD.

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI, 2000, 46, z. 1, ss. 91–90

Działanie i budowa tyrystorów GTO

ZBIGNIEW LISIK

Instytut Elektroniki, Politechniki Łódzkiej

Otrzymano 1999.06.28
Autoryzowano 1999.12.02

Mimo ogromnej ekspansji przyrządów półprzewodnikowych mocy sterowanych polowo, podstawową grupą przyrządów półprzewodnikowych mocy były i jeszcze w dalszym ciągu pozostają przyrządy bipolarne z tyrystorem GTO jako ich głównym przedstawicielem. Na przestrzeni prawie 30 lat, to jest od pojawienia się pierwszego tyrystora GTO, konstrukcja tego przyrządu podlegała szeregu istotnym zmianom mającym na celu polepszenie jego parametrów oraz sprostanie stale rosnącym wymaganiom rynku przyrządów półprzewodnikowych mocy.

W pracy omówiono szczegółowo mechanizm działania tyrystorów GTO ze szczególnym uwzględnieniem zjawisk charakterystycznych dla tych przyrządów, przedstawiono drogi rozwoju ich konstrukcji oraz zasady jej optymalizacji. Szczególnie dużo uwagi poświęcono przedstawieniu nowego rozwiązania tyrystora GTO, tak zwanego tyrystora ze zintegrowaną bramką IGCT (Integrated Gate Commutated Thyristor). Jest to najnowsze osiągnięcie konstruktorów i producentów tyrystorów GTO, będące ich odpowiedzią na konkurencję ze strony tranzystorów IGBT.

Słowa kluczowe: przyrządy półprzewodnikowe mocy, tyrystory GTO, projektowanie przyrządów półprzewodnikowych.

1. WSTĘP

Mimo ogromnej ekspancji przyrządów półprzewodnikowych sterowanych polowo, podstawową grupą przyrządów półprzewodnikowych mocy były i jeszcze w dalszym ciągu pozostają przyrządy bipolarne. Ich głównym przedstawicielem jest tyrystor mocy, który w szeregu zastosowaniach jest niezastąpiony. Na przestrzeni 40 lat, to jest od roku 1958 kiedy firma Westinghouse przedstawiła swój pierwszy tyrystor, konstrukcja tego przyrządu ulegała znacznym modyfikacjom i w efekcie pojawiła się cała rodzina przyrządów określanych mianem tyrystory. Obejmuje ona tyrystory dwukierunkowe TRIAC, tyrystory załączane światłem LTSCR, tyrystory wstecznie przewodzące RCT, tyrystory asymetryczne ASCR, tyrystory o wspomaga-

nym wyłączaniu GATT oraz posiadające obecnie największe znaczenie tyrystory GTO wyłączane prądem bramki.

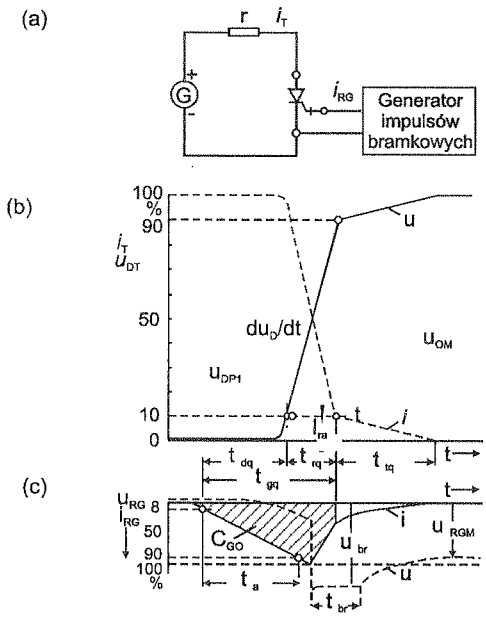
Idea wyłączania tyrystora poprzez doprowadzenie impulsu wstecznego prądu bramki jest prawie tak samo stara jak koncepcja samego tyrystora, tym bardziej, że ta metoda wyłączania była od samego początku możliwa w przypadku przyrządów o bardzo małych prądach przewodzenia. Prace nad wykorzystaniem tego mechanizmu wyłączania w tyrystorach o większych prądach znamionowych sięgają początku lat sześćdziesiątych [1–4]. W roku 1960 van Ligten i Navon [1] stwierdzili, że dla uzyskania tego efektu jest konieczne wprowadzenie rozwiniętej krawędzi złącza emiterowego i redukcja współczynnika wzmocnienia prądowego tranzystora n-p-n lub p-n-p. W 1966 roku Wolley [5] przedstawił pogłębioną teoretyczną analizę procesu wyłączania bramkowego, a na początku lat 70-tych pojawiły się pierwsze publikacje prezentujące wyłączalne tyrystory mocy.

Zo pierwszą taką pracę należy uznać artykuł New'a i in. [6] z 1970 roku. Opisano w nim tyrystor GTO o maksymalnym napięciu blokowania 1000V i jeszcze stosunkowo niewielkim prądzie wyłączalnym 50A, ale już w kolejnej pracy [7] z 1973 roku został przedstawiony tyrystor o prądzie znamionowym 200A. Dalsze prace zmierzały do uzyskania tyrystorów o coraz większych wartościach napięcia blokowania i wyłączalnego prądu przewodzenia. Podstawową trudnością na jaką w nich natrafiono było zwiększenie wartości tego prądu bez zniszczenia przyrządu podczas procesu wyłączania. Było ono wywołane lokalnym wzrostem gęstości prądu, który prowadził do lokalnego wzrostu rozpraszania mocy ponad dopuszczelną wartość. Problem ten został ostatecznie rozwiązany m.in. poprzez wprowadzenie takich modyfikacji konstrukcji tyrystora jak zastosowanie nieskooporowej warstwy p–bazy warstwy n-bazy i złącza bramka katoda o dużym napięciu przebicia [8]. Pozwoliło to na uzyskanie tyrystorów GTO o większych maksymalnych wartościach napięć i prądów, które w obecnie produkowanych przyrządach sięgają, odpowiednio, wartości 5000V i 3000A. W ostatnich latach tyrystory GTO były wypierane z szeregu zastosowań przez burzliwie rozwijające się tranzystory IGBT [9]. Odpowiedzią na to było pojawienie w w drugiej połowie lat 90 [10–12] nowego rozwiązania tyrystora GTO, nazwanego tyrystorem ze zintegrowaną branką IGCT (Integrated Gate Commutated Thyristor), którego parametry pozwalają na skuteczną konkurencję z tranzystorami IGBT.

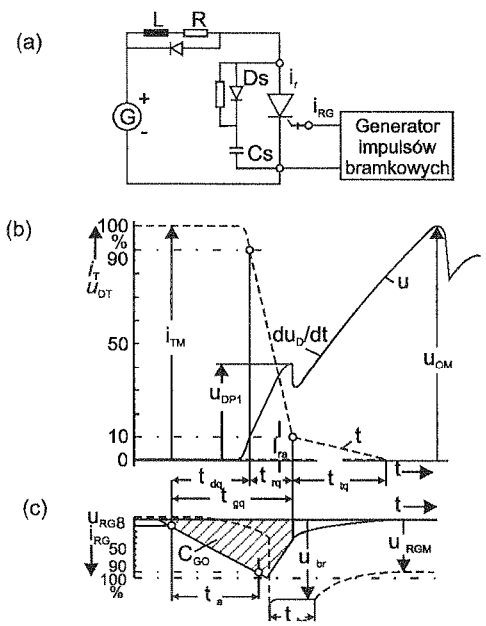
W pracy omówiono szczegółowo mechanizm działania tyrystorów GTO ze szczególnym uwzględnieniem zjawisk charakterystycznych dla tych przyrządów, przedstawiono drogi rozwoju ich konstrukcji oraz zasady jej optymalizacji. W odrębnym rozdziale przedstawiono tyrystor IGCT, najnowsze osiągnięcie konstuktorów i producentów tyrystorów GTO, będące ich odpowiedzią na konkurencję ze strony tranzystorów IGBT.

2. MECHANIZM PROCESU WYŁĄCZANIA BRAMKOWEGO

Proces wyłączania tyrystora GTO jest inicjowany pojawieniem się impulsu wstecznego prądu bramki o odpowiednio dużej amplitudzie I_{RG}, która musi spełniać warunek:



Rys. 1. Obwód elektryczny, w którym jest wyłączany tyrystor GTO (a), zmiany prądu anodowego I_T i napięcia na tyrystorze U_T (b) oraz zmiany prądu bramki I_G i napięcia bramka-katoda U_G (c)

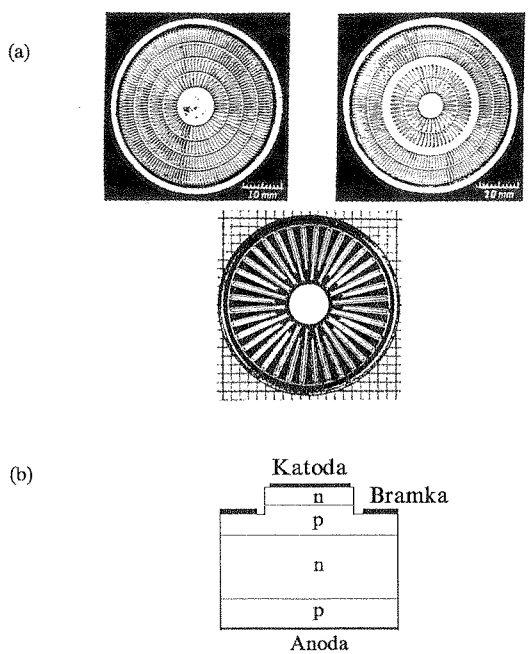


Rys. 2. Obwód elektryczny, w którym jest wyłączany tyrystor GTO wyposażony w tłumik RCD (a), zmiany prądu anodowego I_T i napięcia na tyrystorze U_T (b) oraz zmiany prądu bramki I_G i napięcia bramka-katoda U_G (c)

$$I_{RG} \geqslant \frac{I_T}{g} \tag{1}$$

gdzie I_T — aktualna wartość wyłączanego prądu tyrystora, a g — współczynnik wzmocnienia prądowego przy wyłączaniu, przyjmujący wartości z przedziału $3 \div 10$ w zależności od konstrukcji tyrystora (zwykle ok. 5). Typowe przebiegi zmian prądów i napięć podczas wyłączania samego tyrystora GTO oraz tyrystora wyposażonego w standardowy tłumik RCD przedstawiono, odpowiednio na rys. 1 i 2. Na rysunkach tych zdefiniowano wszystkie parametry używane do opisu procesu wyłączania tyrystorów GTO.

Podstawową różnicą pomiędzy tyrystorami standardowymi a tyrystorami GTO jest występowanie w tych ostatnich bardzo złożonej struktury kontaktu bramki. Jest to podyktowane tym, że prąd sterujący bramki tylko w niewielkim stopniu wnika w obszar pod elektrodą katody, a jego efektywne oddziaływanie na pracę przyrządu zanika praktycznie w odległości rzędu kilkuset mikrometrów od krawędzi kontaktu katody. Aby uzyskać efektywne sterowanie pracą przyrządu na całej powierzchni, obszar kontaktu bramki jest zwykle zlokalizowany nieco poniżej obszaru kontaktu katodowego i obejmuje praktycznie cały obszar pastylki półprzewodnikowej. Obszar kontaktu emitera katodowego jest utworzony przez dużą ilość wąskich długich „palców emiterowych" wystających ponad obszar kontaktu bramki i otoczonych przez ten kontakt, jak to ilustruje rys. 3.



Rys. 3. Widok pastylki półprzewodnikowej tyrystora GTO od strony kontaktu katodowego z „palcami emiterowymi" (a) oraz przekrój poprzeczny tyrystora obejmujący „palec emiterowy" i przyległy obszar kontakru bramki (b)

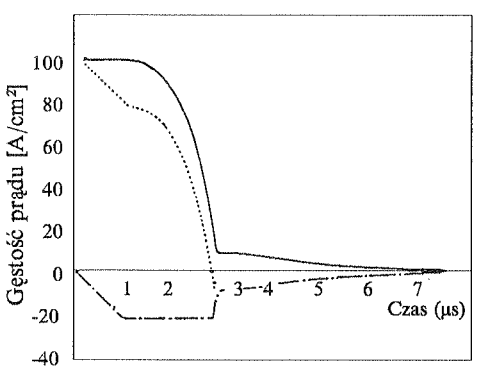
Ta specyfika konstrukcji tyrystora GTO pozwala na potraktowanie go jako równoległego połączenia szeregu pojedynczych komórek obejmujących jeden „palec emiterowy" wraz z otaczającym go kontaktem bramki, które są wprowadzane w stan przewodzenia i blokowania niejako niezależnie, w wyniku przepływu prądu bramki pomiędzy fragmentami elektrody bramki i katody występującymi w danej komórce. W efekcie, przy identyczności komórek, wszelkie procesy zachodzą w nich identycznie, co pozwala ograniczyć analizę pracy tyrystora GTO do obszaru jednej komórki. Wszelkie niejednorodności pomiędzy komórkami prowadzą jednak do pojawienia się dodatkowych efektów poprzecznych związanych np. z ich nierównoczesnym załączeniem lub wyłączeniem.

2.1. PRZEBIEG PROCESU WYŁĄCZANIA POJEDYNCZEJ KOMÓRKI TYRYSTORA GTO

Opisowi procesu wyłączania zachodzącemu w pojedynczej komórce tyrystora GTO jest poświęcona ogromna literatura obejmująca wyniki rozważań teoretycznych, symulacji komputerowych oraz badań eksperymentalnych [5, 13 – 20]. Poniżej, korzystając z zawartych tam wyników, przedstawiono zwarty opis przebiegu procesu wyłączania w typowej komórce standardowego tyrystora GTO przedstawionej na rys. 3b.

2.1.1. Jednowymiarowy opis procesu wyłączania

Wyłączanie bramkowe struktury terystorowej przedstawionej na rys. 3b, jest procesem co najmniej dwuwymiarowym. Podstawowe etapy tego procesu można jednak z powodzeniem opisać jednowymiarowo, śledząc zmiany pola elektrycznego i gęstości nośników wzdłuż osi łączącej anodę z katodą. Zmiany te, wyznaczone przez Naito i inn. [16] dla tyrystora GTO wyłączanego w obwodzie z rys. 1a, są przedstawione na rys. 5, a odpowiadające im zmiany prądów anody, katody i bramiki na rys. 4.

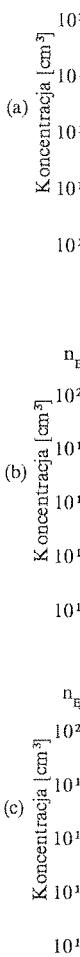


Rys. 4. Zmiany prądu anodowego (linia ciągła), katodowego (linia kropkowana) i bramki (linia przerywana) [16]

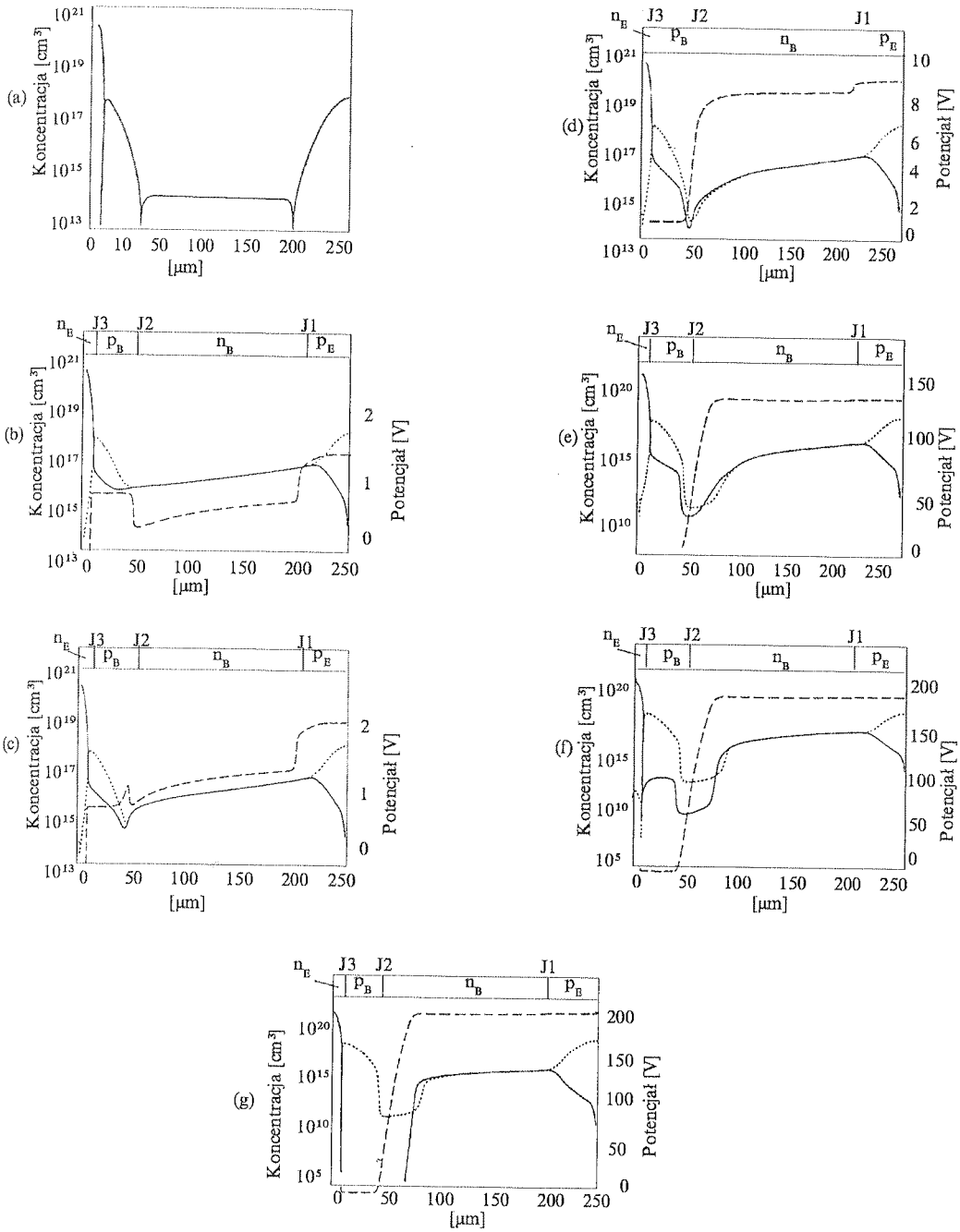
Stan ustalony w tyrystorze GTO (rys. 5b) niczym nie różni się stanu ustalonego w normalnym tyrystorze. Wszystkie złącza są spolaryzowane w kierunku przewodze-

nia, obszar słabo domieszkowany jest zalany przez nośniki nadmiarowe „modulujące" rezystowość tego obszaru, a na tyrystorze występuje napięcie 1.42V przy gęstości prądu 99,3 A/cm^2. Pojawienie się wstecznego prądu bramki powoduje zakłócenie rozkładów w otoczeniu środkowego złącza J2. Rys. 5c i d przedstawiają kolejno stany po 0.925μs i 1.45 μs od pojawienia się wstecznego prądu bramki. Zgodnie z rys. 4, odpowiadają one przedziałowi czasowemu określanemu jako czas opóźnienia, t_{dq} na rys. 1, kiedy nie obserwuje się jeszcze zmian amplitudy prądu anodowego. Na pierwszym z nich wszystkie trzy złącza pozostają jeszcze w stanie polaryzacji w kierunku przewodzenia i obserwujemy jedynie lokalne zmniejszenie koncentracji nośników nadmiarowych po stronie p-bazy złącza J2, któremu towarzyszy niewielki pik w rozkładzie potencjału. Na drugim z nich zmiany koncentracji nośników w otoczeniu złącza J2 są na tyle duże, że doprowadziły już do zmiany polaryzacji złącza J2. Jest ono teraz spolaryzowane zaporowo i zaczyna na nim narastać napięcie w kierunku blokowania, czemu towarzyszy zmniejszanie się prądu tyrystora. Moment zmiany polaryzacji złącza J2 kończy fazę opóźnienia procesu wyłączania i początkuje fazę opadania, w której występuje gwałtowny spadek amplitudy prądu — fazie tej odpowiada przedział t_{rg} na rys. 1. Zmiany zachodzące w tej fazie w strukturze tyrystora pokazują kolejno rys. 5e i f. W otoczeniu złącza J2 koncentracje nośników nadmiarowych gwałtownie maleją umożliwiając tworzenie się w tym obszarze warstwy ładunku przestrzennego, na której odkłada się blokowane napięcie. Istotną różnicą pomiędzy stanami w chwili t=2.65 μs i t=2.98 μs jest fakt, że w pierwszym przypadku oba złącza emiterowe są spolaryzowane przepustowo i wstrzykują nośniki do obszarów baz, natomiast w drugim złącze katodowe przeszło w stan polaryzacji zaporowej i wstrzykiwanie nośników na miejsce od strony emitera anodowego. Na rys. 4 zmianie tej odpowiada zmiana kierunku prądu katody i zmiana charakteru zmian prądu anodowego, który maleje teraz znacznie wolniej. Jest to początek fazy ogona prądowego, której odpowiada na rys. 1 czas ogona t_{tq}. Zmiany występujące w tej fazie ilustrują rys. 5e i g. W wyniku wstrzymania wstrzykiwania elektronów od strony katody ich koncentracja w pobliżu złącza J2 maleje szybciej niż dziur, które są w dalszym ciągu wstrzykiwane od strony anody. Powoduje to podtrzymanie przepływu prądu anodowego. Prąd ten zanika stopniowo wraz ze zmniejszaniem się nadmiarowej koncentracji nośników w obszarze n-bazy tyrystora w wyniku przepływu prądu i rekombinacji.

W procesie wyłączania zachodzącym w obwodzie z rys. 1a wzrost napięcia blokowania jest równie szybki jak spadek wartości prądu anody, a jego wartość w fazie ogona prądowego jest bliska wartości napięcia źródła. Powoduje to, że chwilowe wartości mocy rozpraszanej w tyrystorze są bardzo duże, co może doprowadzić do zniszczenia przyrządu przy dużej wartości czasu ogona prądowego t_{tq}. Aby ograniczyć wielkość energii rozproszonej podczas wyłączania stosuje się tłumiki RCD. Obwód taki, wraz z przebiegiem zmian napięć i prądów podczas wyłączania, przedstawia rys. 2. Dzięki zastosowaniu tłumika napięcie na tyrystoorze jest znacznie mniejsze od napięcia zasilania w całej fazie opadania oraz przez znaczną część fazy ogona prądowego, co istotnie zmniejsza wielkość energii rozproszonej w tyrystorze podczas wyłączania. Ubocznym skutkiem zastosowania tego roz-



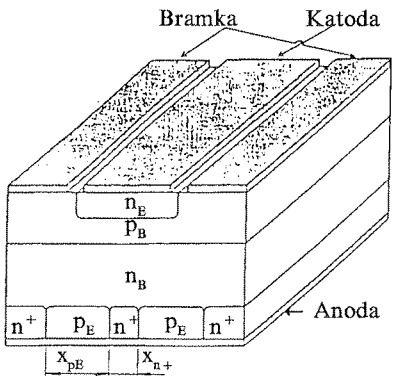
Rys. 5.
elektron
procesu



Rys. 5. Rozkład domieszkowania w analizowanej strukturze tyrystora GTO (a) oraz zmiany rozkładów elektronów (linia ciągła), dziur (linia kropkowana) i potencjału (linia przerywana) dla kolejnych chwil procesu wyłączania przedstawionego na rys. 4: (b) stan ustalony t-0μs; (c) – t = 0.925μs; (d) – t = 1.45μs; (e) – t = 2.65μs; (f) – 2.98μs; (g) – t = 7.08μs [16]

wiązania, wywołanym istnieniem pasożytniczej indukcyjności w gałęzi tłumika, jest wystąpienie skoku napięcia przy przejściu z fazy opadania do fazy ogona prądowego (tzw. „*spike voltage*"). Amplituda tego napięcia, U_{DPI} na rys. 2b, może być na tyle duża, że związany z nią impuls rozpraszanej mocy może mieć istotny wkład do całkowitej energii rozproszonej podczas wyłączania.

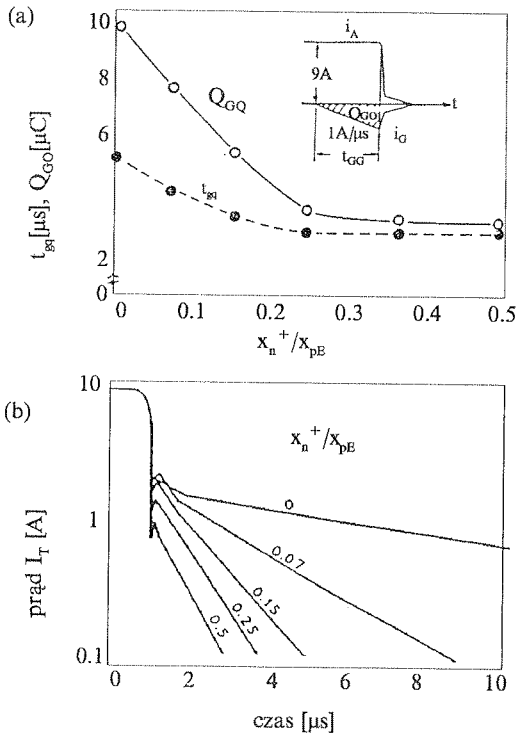
Ograniczenie energii rozpraszanej podczas wyłączania można również osiągnąć na drodze modyfikacji konstrukcji tyrystora prowadzących do skrócenia czasu ogona prądowego t_{tq}. Można to zrobić wprowadzając do obszaru bazy dodatkowe centra rekombinacyjne [20–23] lub stosując zwarcia emiterowe od strony anody [17, 20, 24–26]. W pierwszym przypadku tyrystor zachowuje zdolność blokowania napięcia w obu kierunkach ale wzrasta spadek napięcia w stanie przewodzenia oraz pogarszają się zdolności blokowania napięcia w podwyższonycvh temperaturach. W drugim przypadku tyrystor traci zdolność blokowania napięcia wstecznego i staje się tyrystorem asymetrycznym.



Rys. 6. Typowa komórka tyrystora asymetrycznego GTO ze zwarciami emiterowymi anody

Rys. 6 przedstawia widok typowej komórki asymetrycznego tyrystora GTO, w której występują naprzemiennie od strony kontaktu anody silnie domieszkowane obszary n i p. Parametrem charakteryzującym tą strukturę jest tzw. współczynnik zwarciowości f_s definiowany jako stosunek długości x_n^+ krawędzi obszaru n^+ stanowiącego zwarcie złącza anodowego do długości x_{pE} krawędzi obszaru p tworzącego emiter anodowy. Wartość tego współczynnika decyduje o efektywności wstrzykiwania nośników do bazy od strony anody, zmniejszeniem koncentracji nośników nadmiarowych w bazie w stanie przewodzenia oraz polepszeniem efektywności wyprowadzania nośników z bazy. Wszystkie te czynniki mają wpływ na przebieg procesu wyłączania, a w szczególności na skrócenie czasu ogona prądowego. Ilustrują to przedstawione na rys. 7 wyniki pomiarów przeprowadzonych przez Yatsuo i in. [20]. Rys. 7a pokazuje jak zmieniają się wraz z wartością współczynnika f_s, czas wyłączania t_{gq} i wyłączający ładunek prądu bramki Q_{GQ}, natomiast na rys. 7b pokazano zmiany ogona prądowego wywołane zmianami tego współczynnika.

Wadą wprowadzenia zwarć emiterowych od strony anody jest pogorszenie parametrów załączania tyrystora. Napięcie polaryzujące złącze emiterowe tranzys-

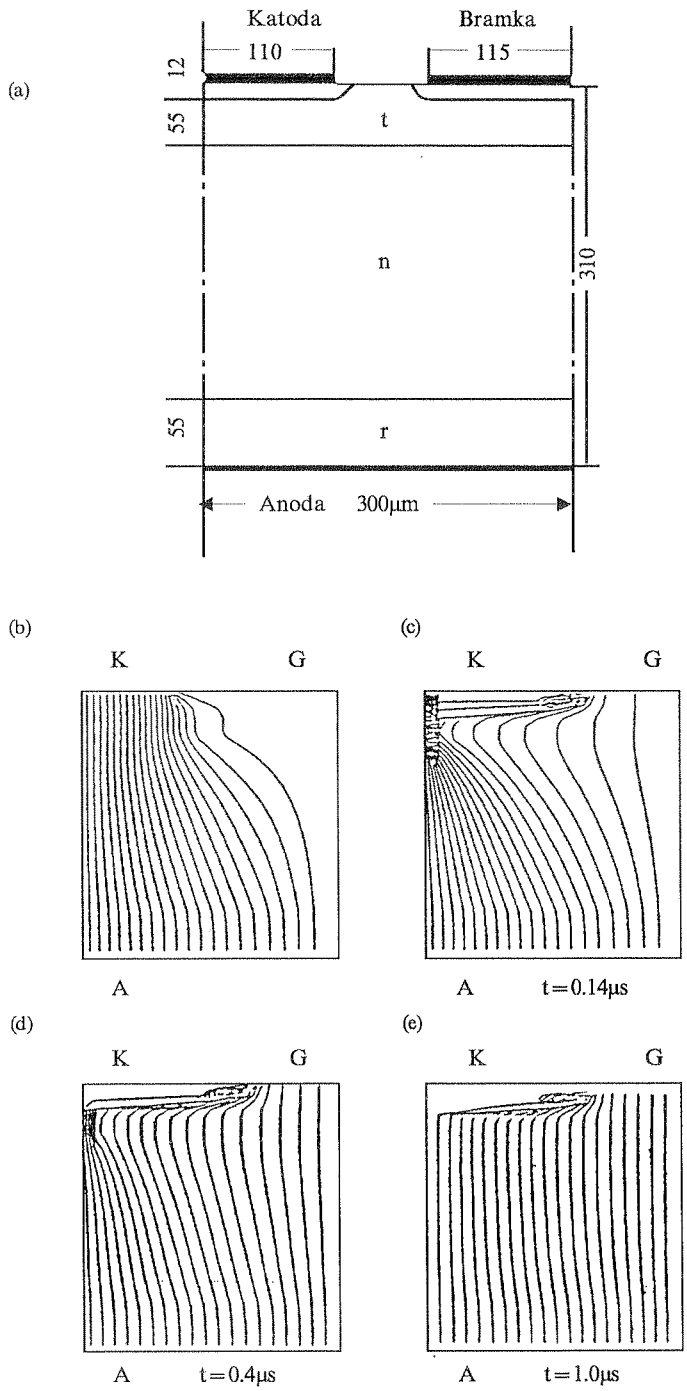


Rys. 7. Wpływ wartości współczynnika zwarciowości F_s na czas wyłączania t_{gq} i wyłączający ładunek prądu bramki Q_{GO} (a) oraz na kształt ogona prądowego (b) [20]

tora pnp jest teraz związane ze spadkiem napięcia wywołanym przepływem prądu w obszarze n^+ zwarcia, który bocznikuje złącze. W efekcie złącze to zacznie wstrzykiwać dziury do n-bazy dopiero po przekroczeniu pewnej wartości prądu anodowego tyrystora, co wymaga odpowiednio dużych wartości prądu bramki podczas załączania tyrystora. Niedogodność tą usunięto wprowadzając w miejsce zwarć tzw. transparentny emiter anodowy [44, 45]. Istotna tego rozwiązania polega na zastosowaniu tak wąskiej warstwy p-emitera, że uzyskany w tej warstwie duży gradient koncentracji nośników nadmiarowych w istotny sposób zwiększa składową prądu elektronowego złącza. Prowadzi to do obniżenia efektywności wstrzykiwania, tak jakby emiter anodowy był „przezroczysty" dla części elektronów docierających do kontaktu anodowego.

2.1.2. Efekty dwuwymiarowe w procesie wyłączania komórki

Jednowymiarowy opis wyłączania tyrystora GTO pozwala na zdefiniowanie podstawowych parametrów charakteryzujących ten proces. Ze swej natury jest to jednak proces dwuwymiarowy, w którym gęstość prądu jest funkcją położenia, a poprzeczne prądy mają decydujące znaczenie dla przebiegu wyłączania bram-

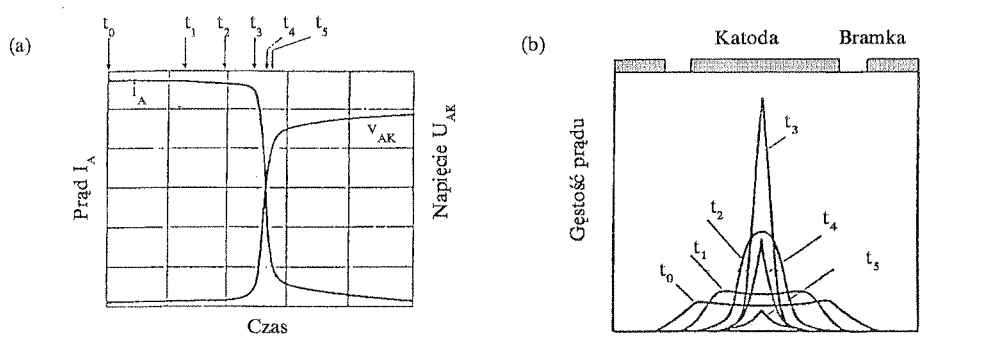


Rys. 8. Zmiany przepływu prądu w komórce tyrystora GTO podczas procesu wyłączania: (a) wymiwary i rozkład domieszkowania, (b–e) linie przepływu prądu w kolejnych fazach procesu [13]

kowego. Ilustruje to rys. 8, na którym przedstawiono za Masadą i inn. [13] zmiany rozkładu gęstości prądu w pojedynczej komórce tyrystora GTO (rys. 8a) w kolejnych fazach procesu wyłączania.

Rys. 8b odpowiada początkowi fazy opóźnienia, w której prąd, podobnie jak w stanie ustalonym, wpływa do komórki całą krawędzią kontaktu anodowego natomiast wypływa poprzez kontakt katodowy. Począwszy od obszaru n-bazy, bliskiemu krawędzi złącza J2, mamy tu do czynienia ze „ściąganiem" prądu do obszaru pod kontaktem katodowym, niemniej gęstość prądu bezpośrednio pod kontaktem jest równomierna. Rys. 8c odpowiada końcowi fazy opóźnienia i początkowi fazy opadania. Obserwujemy tu silne „ściśnięcie" prądu, który tworzy wąski kanał prądowy w obszarze p – bazy ulokowanym pod środkową częścią elektrody katodowej. Większa część tego prądu opuszcza komórkę poprzez centralną część elektrody katody, natomiast pozostała część kieruje się do elektrody bramki tworząc poprzeczny prąd bramiki. Rys. 8d odpowiada fazie opadania, kiedy zgodnie z rys. 4 koncentracja nośników w p – bazie maleje i zaczyna formować się obszar ładunku przestrzennego w otoczeniu złącza J2. Towarzyszy temu stopniowy zanik kanału prądowego z jednoczesnym przejmowaniem coraz większej części prądu przez kontakt bramki. Rys. 8e odpowiada momentowi przejścia od fazy opadania do fazy ogona prądowego, prąd płynie na nim od kontaktu anody do bramki praktycznie równomiernie i tylko na krawędzi bramki od strony katody obserwujemy lokalny wzrost gęstości prądu.

Ściąganie prądu w wąski kanał ulokowany centralnie pod kontaktem katodowym jest charakterystycznym efektem związanym z procesem wyłączania bramkowego. Kanał ten tworzy się w fazie opóźnienia i zanika w fazie opadania, a występujące w nim gęstości prądu mogą osiągnąć wartości rzędu tysięcy A/cm^2. Rozwój i zanik tego kanału ilustruje rys. 9b, na którym pokazano zmiany rozkładu gęstości prądu na środkowym złączu dla kolejnych chwil czasowych procesu wyłączania wyróżnionych na rys. 9a.



Rys. 9. Zmiany kanału prądowego pod kontaktem katody podczas procesu wyłączania tyrystora GTO: zmiany napięcia anoda-katoda V_{AK} i prądu anodu I_A (a) oraz rozkłady gęstości prądu na powierzchni złączu J2 dla wyróżnionych chwil procesu wyłączania (b) [26]

2.2. EFEKTY POPRZECZNE WYWOŁANE NIEJEDNORODNOŚCIĄ TYRYSTORA GTO

Tyrystory GTO są bardzo czułe na wszelkie niejednorodności pastylki półprzewodnikowej, z której są wykonane. Składają się one z ogromnej ilości identycznie zaprojektowanych komórek tworzących elementarne tyrystory GTO połączonych z sobą elektrodami anody, katody i bramki. Warunkiem poprawnej pracy takiego przyrządu jest zapewnienie równomiernego rozkładu prądu przez niewielki obszar przyrządu i w konsekwencji do jego uszkodzenia.

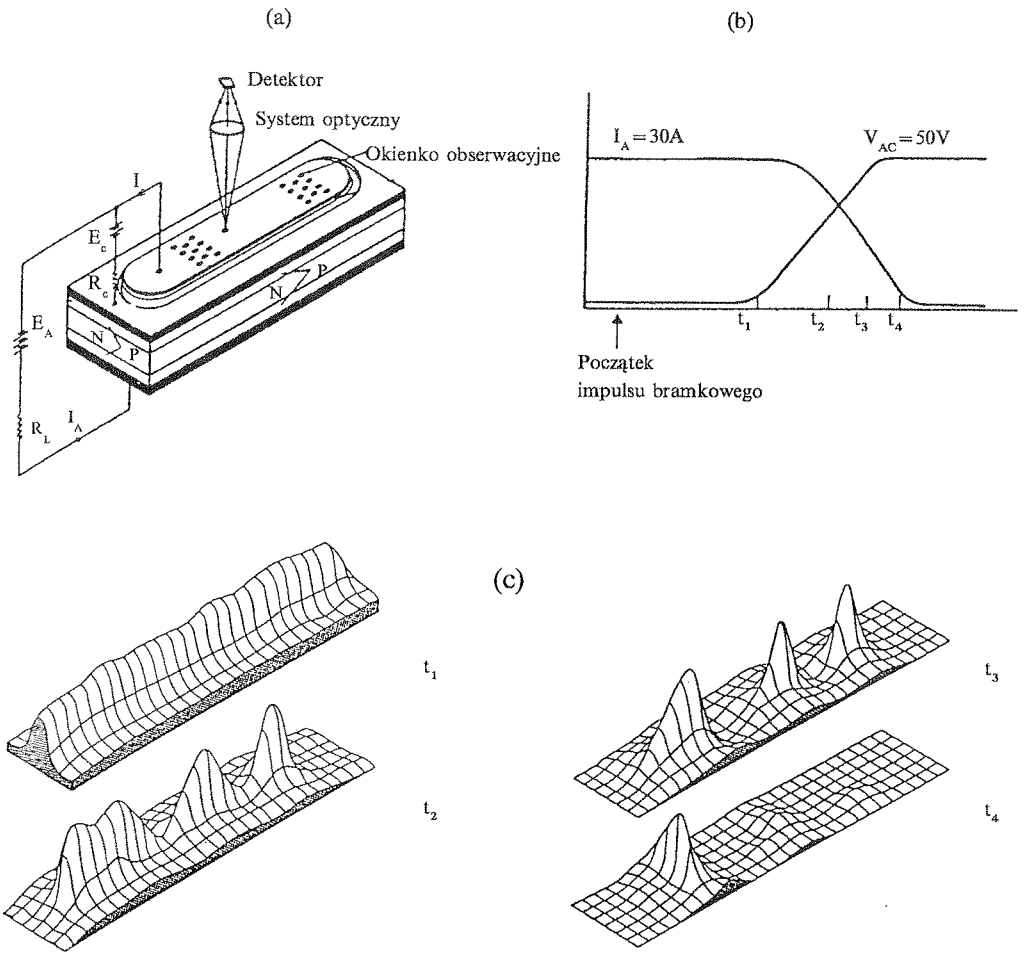
Prawidłowe rozwiązanie tego problemu jest ważne z punktu widzenia niezawodności pracy tyrystora GTO i osiągnięcia przez niego pożądanych parametrów znamionowych. Efekty wywołane fluktacjami koncentracji domieszek (lub czasów życia) są obserwowane zarówno jako efekty lokalne, występujące w obrębie jednej komórki [18, 19, 27, 28], jak i jako efekty globalne, obejmujące swoim oddziaływaniem cały przyrząd [28. 29].

2.2.1. Efekty występujące w obrębie jednej komórki

W komórce rzeczywistego tyrystora GTO, obok opisanych rozdz. 2.1.2 efektów poprzecznych zachodzących w płaszczyźnie prostopadłej do palca emitera katodowego, można także oczekiwać niejednorodności rozkładu gęstości prądu wzdłuż osi prostopadłej do tej powierzchni wywołanej fluktuacjami domieszkowania i czasu życia. Potwierdzają to badania prowadzone w Toshiba Research Center [18, 19], w których korzystając z pomiaru promieniowania rekombinacyjnego w podczerwieni, obserwowano zmiany rozkładu gęstości nośników pod palcem emitera katodowego. Zasada pomiaru wraz z poglądowym widokiem struktury testowej jest przedstawiona na rys. 10a. Pomiar jest wykonywany przez system optyczny o rozdzielczości 20 μm poprzez rozmieszczone równomiernie na powierzchni elektrody emitera katodowego okna pomiarowe (observation windows) o średnicy 30 μm. Pomiary wykonywane dla fazy opóźnienia potwierdziły występowanie ściągania prądu ku środkowi konaktu katody przy zachowaniu względnej jednorodności tego zjawiska wzdłuż całej długości palca. W fazie opadania wystąpiło jednak wyraźne zróżnicowanie gęstości nośników także w tym kierunku, co oznacza że wzdłuż palca wystepuje ściąganie prądu prowadząc do lokalnych maksimów gęstości prądu. Rozkłady uzyskane dla kolejnych chwil czasowych zaznaczonych na rys. 10b, są zebrane na rys. 10c.

2.2.2. Efekty obejmujące obszar pastylki półprzewodnikowej

Jest rzeczą oczywistą, że maksymalna wartość I_{TQRM} wyłączalnego prądu tyrystora GTO jest ograniczona momentem wystąpienia w jednym z segmentów krytycznych warunków przeciążenia prowadzących do uszkodzenia przyrządu w obszarze tego segmentu [14, 18]. Istotną konsekwencją tego faktu jest brak proporcjonalnej zależności pomiędzy przyrostem liczby segmentów w tyrystorach GTO (czyli powie-

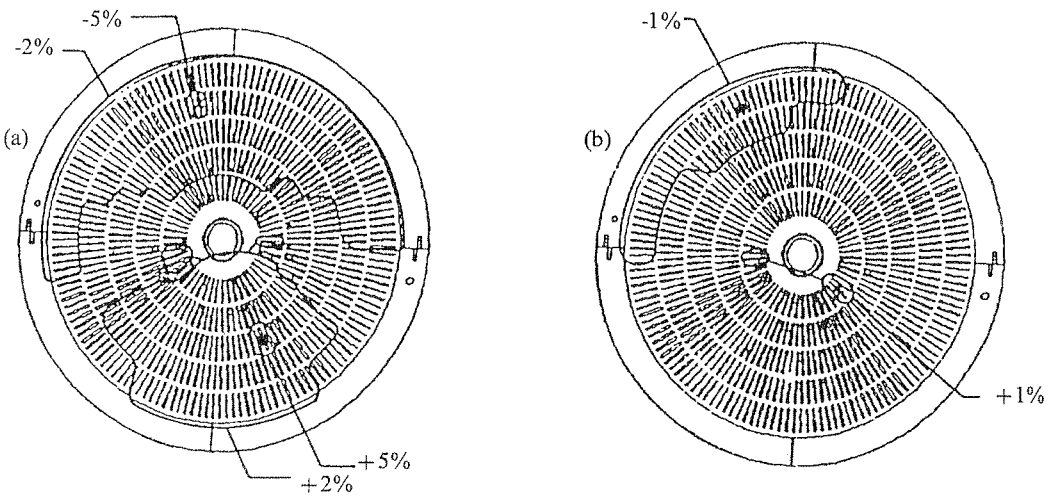


Rys. 10. Pomiar zmian rozkładu koncentracji nośników pod palcem emiterowym w fazie opadania procesu wyłączania: (a) widok systemu pomiarowego; (b) zmiany prądu i napięcia; (c) rozkłady koncentracji nośników dla wyróżnionych chwil czasowych [19]

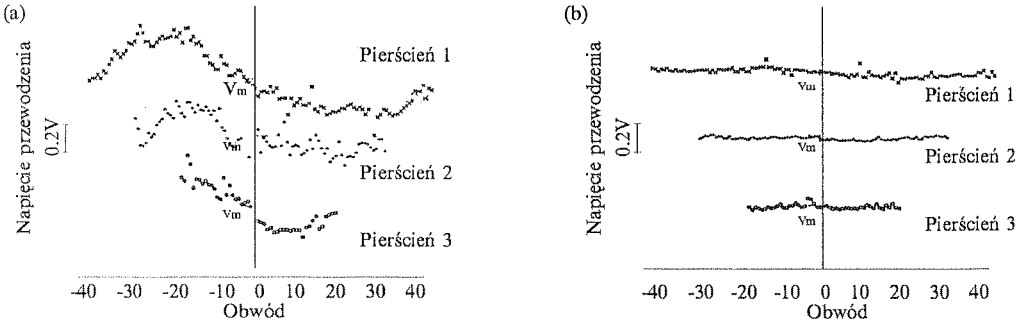
rzchni kontaktu katodowego) a wzrostem wartości prądu I_{TQRM} [26, 28]. Bezpośrednią przyczyną tego jest wystepowanie pomiędzy segmentami dużych tyrystorów elektycznych oddziaływań oraz rozrzutów technologicznych powodujących, że tyrystor GTO nie może być traktowany jako równoległe połączenie identycznych, odseparowanych od siebie segmentów.

Uzyskanie idealnie jednorodnych rezultatów procesów technologicznych na całej powierzchni dużych pastylek półprzewodnikowych jest praktycznie niemożliwe i dlatego wystepowanie fluktuacji parametrów na ich powierzchni jest nieuniknione. Fluktuacje te wystepują już w materiale wyjściowym, a wszelkie procesuy dyfyzyjne, w których np. wprowadza się do struktury domieszki z fazy gazowej zachodzą z lokalnie różnymi prędkościami (typowy rozrzut $\pm 5 \div 10\%$) wprowadzając kolejne niejednorodności. Tak więc, nie mogąc ich całkowicie wyeliminować, należy dążyć do

ich minimalizacji poprzez zastosoanie odpowiedniej technologii. Standardowym posunięciem jest w tym przypadku zastosowanie jako materiału wyjściowego krzemu neutronowego. Uzyskuje się w nim dużą jednorodność rozkładu domieszek poprzez domieszkowanie na drodze przemiany jądrowej atomów krzemu w atomy fosforu w wyniku napromieniowania jednorodną wiązką neutronową (tzw. proces NTD — neutron transmutation doping). W pracy [29] zaproponowano zastąpienie procesami napromieniowania jednorodnymi strumieniami odpowiednich cząsteczek także innych procesów domieszkowania. Przedstawione tam pomiary niejednorodności wprowadzonych przez procesy dyfuzyjne i równoważne im procesy bombardowania strumienia cząsteczek są zebrane na rys. 11 i 12.



Rys. 11. Kontury stałego odchylenia powierzchniowej koncentracji domieszkowania p – bazy określone z pomiaru napięć wstecznych złącz katoda-bramka segmentów, przy założeniu stałej koncentracji powierzchniowej domieszkowania katody; (a) — nakładanie warstwy powierzchniowej p – bazy dyfuzyjne ze źródła BN, (b) — nakładanie warstwy powierzchniowej w procesie implantacji jonów [29]



Rys. 12. Odchylenia spadku napięcia przewodzenia na poszczególnych segmentach tyrystora w stosunku do wartości średniej V_m (każda linia odpowiada jednemu pierścieniowi segmentów): (a) centra rekombinacyjne wprowadzone w procesie dyfuzji złota (820°C, 1 godz.), (b) — centra rekombinacyjne utworzone poprzez bombardowanie elektronami (dawka 1×10^{12} cm^{-2}) [29]

Na rys. 11 porównano oszacowane fluktuacji powierzchniowej koncentracji domieszek boru wprowadzonych do p-bazy tyrystora GTO zawierającego 514 segmentów w procesie dyfuzji z BN (rys. 11a) oraz w procesie implantacji jonowej (rys. 11b), przy niezmienionych pozostałych procesach technologicznych. Zaznaczono na nim punkty o największym odchyleniu od wartości średniej koncentracji oraz obszary o odchyleniu, odpowiednio $\pm 2\%$ i $\pm 5\%$ dla procesu dyfuzji oraz $\pm 1\%$ dla procesu implantacji. W obu przyrządach występują fluktuacje, ich amplituda i zasięg są jednak znacznie mniejsze w przypadku zastosowania implantacji jonów [1].

Na rys. 12 porównano wpływ technologii wprowadzania centr rekombinacyjnych na indywidualne spadki napięcia przewodzenia poszczególnych segmentów tyrystora GTO zawierającego 260 segmentów rozlokowanych w trzech pierścieniach. W pierwszym przypadku źródłem poziomów rekombinacyjnych były atomy złota wprowadzane w procesie dyfizji, a w drugim dyslokacje utworzone w wyniku bombardowania strumieniem elektronów. Wyniki dla każdego pierścienia zostały przedstawione, jako odchylenia od średniej wartości napięcia V_m. Wpływ wybranej metody na jednorodność procesów zachodzących w tyrystorze jest wyraźnie widoczny. W przypadku centr rekombinacyjnych wprowadzanych strumieniem elektronów odchylenia są na poziomie ± 0.07V, a więc rzeczywiście pomijalnie małe.

3. ZASADY OPTYMALIZACJI TYRYSTORÓW GTI

Generalnie, zasada działania tyrystorów GTO i tyrystorów tradycyjnych jest taka sama, przebieg większości zjawisk jest analogiczny, a ich własności opisują te same parametry elektryczne. Także wymagania dotyczące tych parametrów są analogiczne i osiągane w wyniku zastosowania podobnych zasad optymalizacji konstrukcji. Nowe możliwości tyrystorów GTO, związne z możliwością ich kontrolowanego wyłączenia ujemnym prądem bramki, powstały jedynie dzięki zmianom konstrukcyjnym „wzmacniającym" efekty, które były już obserowowane wcześniej. Z uwagi na to bliskie pokrewieństwo nie będą tu omówione te zasady optymalizacji konstrkcji tyrystorów GTO, które są wspólne dla wszystkich rodzajów tyrystorów i z tego względu dostępne w bogatej literaturze dotyczącej tego tematu [31 – 33]. Po ogólnym omówieniu wymagań stawianych podstawowym parametrom tyrystora GTO oraz możliwości ich spełnienia, cała uwaga będzie skupiona na przedstawieniu kryteriów optymalnego doboru parametrów konstrukcyjnych tych fragmentów tyrystora GTO, które stanowią o specyfice tego przyrządu i są istotne dla jego prawidłowej pracy.

[1] Oszacowanie zostało przeprowadzone przy założeniu, że koncentracja powierchniowa domieszek n-emitera nie wykazuje fluktuacji. Ponieważ zostały one wprowadzenie w procesie dyfuzji ze źródła $POLCL_3$, można oczekiwać, że przynajmniej część fluktuacji obserwowanych na rys. 11b jest wywolana fluktuacjami domieszkowania n-emitera.

3.1. WYMAGANIA STAWIANE TYRYSTOROM GTO

• *blokowanie napięcia*

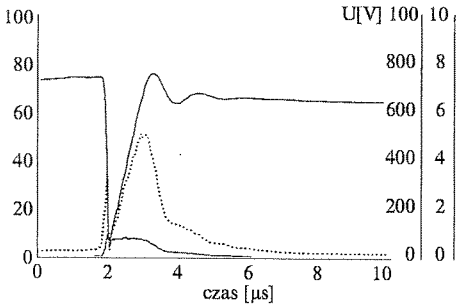
W tyrystorze GTO, analogicznie jak w tyrystorze tradycyjnym, napięcie blokowania odkłada się na ładunku przestrzennym spolaryzowanego wstecznie środkowego złącza tradycyjnie oznaczanym jako złącze J2. Maksymalna wartość tego napięcia zależy od poziomu domieszkowania n-bazy, jej grubości oraz profilu domieszkowania w p-bazie. Parametry te dobiera się analogicznie jak w tyrystorze tradycyjnym i jedynie przy ustalaniu profilu domieszkowania p-bazy może być brana pod uwagę specyfika tyrystora GTO. Na ostateczną wartość maksymalnego napięcia blokwania mają również wpływ efekty powierzchniowe prowadzące do przebicia powierzchniowego. Aby ich uniknąć stosuje się w tradycyjnych przyrządach sfazowanie powierzchni kończącej złącze oraz pokrycia pasywujące. Można tu również zastosować rozwiązanie polegające na wprowadzeniu cienkich warstw pasywujących i pierścieni ograniczających natężenie pola elektrycznego [30].

• *straty w stanie przewodzenia*

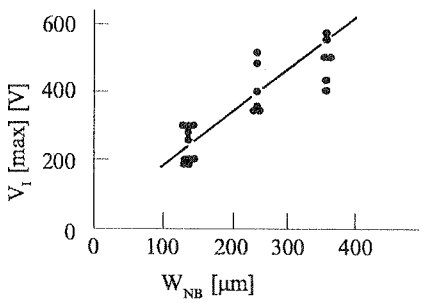
Parametrem decydującym o stratach w stanie przewodzenia jest spadek napięcia na przewodzącym tyrystorze GTO. Podobnie jak tyrystor tradycyjny, zachowuje się on w tym stanie jak przewodząca dioda p-i-n, w której spadek napięcia zależy od efektywności wstrzykiwania p- i n-emitera, szerokości obszaru nisko domieszkowanego oraz szybkości rekombinacji w obu emiterach i obszarze „i". Aby osiągnąć minimalne napięcie przewodzenia szerokość warstwy n-bazy powinna być możliwie mała przy zapewnieniu wymaganego napięcia przebicia, a poziom domieszkowania p-bazy przy złączu J1 branka-katoda możliwie niski w celu zwiększenia efektywności wstrzykiwania n-emitera. W tyrystorze GTO, na te kryteria doboru parametrów n- i p-bazy nakładają się jeszcze inne wynikające ze specyfiki procesu wyłączania bramkowego.

• *straty przy przełączaniu*

W tyrystorach GTO podstawowymi stratami przy przełączaniu są straty związane z wyłączaniem. Zmiany mocy rozpraszanej w tym procesie, na tle odpowiednich zmian napięcia i prądu, pokazuje rys. 13. Proces ten został już opisany w rozdz. 2.1. Powiedziano tam, że większość strat zachodzi w fazie ogona prądowego i ich ograniczenie można osiągnąć poprzez zmniejszenie amplitudy prądu w fazie ogona prądowego oraz czasu trwania tej fazy. Można na nie wpływać m.in. poprzez zmniejszenie efektywności wstrzykiwania p-emitera (złącza anodowe J2), np. w wyniku wprowadzenia zwrać, lub poprzez zmniejszenie czasu życia w n-bazie na drodze wprowadzenia dodatkowych centr rekombinacyjnych. Efekt ten można również osiągnąć poprzez zmniejszenie obszaru zajmowanego przez obszar ładunku przestrzennego w stosunku do calego obszaru n-bazy w wyniku zwiększenia koncentracji domieszek w n-bazie (wpływa na zmniejszenie napięcia blokowania w kierunku przewodzenia). Eksperymentalne prace Ohashi i in. [18], poświęcone badaniu mechanizmów uszkodzenia tyrystora w rezultacie strat podczas wyłączania, po-



Rys. 13. Zmiany napięcia (V), prądu (I) oraz rozpraszanej mocy (P) w tyrystorze GTO (całkowita energia rozproszona podczas wyłączania 9.179 mL) [21]



Rys. 14. Eksperymentalna zależność pomiędzy szerokością n-bazy w_{nb}, a maksymalnym napięciem $V_1(max)$ [18]

zwoliły na wyznaczenie przedstawione na rys. 14 zależnościo pomiędzy szerkością n-bazy w_{nb}, a maksymalnym napięciem $V_1(max)$, które może wystąpić na tyrystorze na początku fazy ogona prądowego nie powodując jeszcze jego uszkodzenia.

- *czułość przy przełączaniu*

Proces załączania tyrystora GTO zachodzi analogicznie jak tyrystora tradycyjnego, wartości prądów wyzwalających bramki są tego samego rzędu, a konstrukcja wielosegmentowa z bardzo rozwiniętą krawędzią złącza emiterowego umożliwia szybkie rozprzestrzenienie się przewodzącego obszaru na całą powierzchnię pastylki półprzewodnikowej.

Proces wyłączania tyrystora GTO jest scharakteryzowany przez dwa podstawowe parametry — maksymalny prąd wyłączalny I_{TQRM} oraz współczynnik wzmocnienia prądowego przy wyłączaniu g. Jest pożądane aby te parametry były maksymalnie duże, co osiąga się odpowiednio modyfikując konstrukcję tyrystora GTO. Generalnie, za podstawę dla tych modyfikacji można przyjąć wzory:

$$g = \frac{\alpha_{npn}}{(\alpha_{npn} + \alpha_{\pi\nu\pi} - 1}, \tag{2}$$

$$I_{TQRM} = 4g\frac{V_{bGK}}{R_B} \tag{3}$$

$$R_B + \frac{\rho_{pb} b_{nc}}{w_{pb} l_{nc}} \tag{4}$$

gdzie: α_{npn} i α_{pnp} — współczynniki wzmocnienia tranzystorów składowych, odpowiednio, npn i pnp, V_{bGK} — napięcie przebicia złącza bramka-katoda, a R_B — rezystancja obszaru p-bazy określona przez rezystywność ρ_{pb} i grubość w_{pb} p—bazy oraz szerokość b_{nc} i długość l_{nc} pojedynczego „palca" emiterowego.

Wzór (2) uzyskano korzystając z jednowymiarowego modelu dwutranzytorowego struktury p-n-p-n i siłą rzeczy nie uwzględnia on efektów dwuwymiarowych. Pokazuje on jednak podstawową prawidłowość — aby współczynnik wzmocnienia g był duży należy dążyć do dużej wartości współczynnika wzmocnienia α_{npn} tranzystora npn, a więc z emiterem kolektorowym, oraz małej wartości współczynnika wzmocnienia α_{pnp} trazystora pnp zawierającego emiter anodowy.

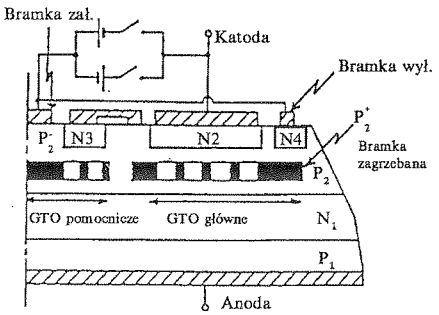
Wzór (3) został podany przez Wolley'a [5]. Zgodnie z nim należy oczekiwać, że maksymalną wartość wyłączalnego prądu uzyskamy podejmując kroki w kierunku zwiększenia napięcia przebicia złącza baza-katoda, co wiąże się z niskim poziomem domieszkowania p-bazy w pobliżu krawędzi złącza, lub zmniejszając rezystancję R_B, co z kolei wymaga silnego domieszkowania p-bazy oraz odpowiednich zmian w wymiarach „palca" emiterowego i grubości p-bazy.

3.2. KRYTERUA DOBORU PARAMETRÓW KONSTRUKCYJNYCH

- *rezystywność p-bazy*

O rezystywności p-bazy decyduje rozkład koncentracji domieszek w obszarze bazy. Występują tu dwa przeciwstawne wymagania — aby uzyskać dużą wartość wyłączalnego prądu tyrystora powinna ona być możliwie mała (duża koncentracja domieszek), gdyż wtedy mała będzie rezystancja R_B, ale zarazem możliwie duża w sąsiedztwie złącza emiterowego (mała koncentracja nośników), gdyż wtedy będzie duże napięcie przebicia V_{bGK}. Dodatkowo, utrzymanie niskiego pozimo domieszkowania przy złączu zwiększa efektywność wstrzykiwania tego złącza, a tym samym współczynnik wzmocnienia α_{npn} tranzystora składowego i w konsekwencji współczynnik wzmocnienia przy wyłączaniu g.

Spełnienie tych przeciwstawnych warunków jest trudne, o ile w ogóle możliwe. Przy dyfuzyjnym wprowadzeniu domieszek akceptorowych stosuje się tu specjalnie dopracowane procesy pozwalające na uzyskanie możliwie niskiej koncentracji domieszxek przy złączu z zachowaniem wysokiej koncentracji w głębi p-bazy. W procesach tych rozkład koncentracji domieszek będzie jednak zawsze określony przez krzywą Gaussa. Dużo większe możliwości wpływania na ostateczny kształt rozkładu domieszek w p-bazie daje domieszkowanie w procesie implantacji jonów lub zastosowanie procesów epitaksji. Pozwala to na takie zaprojektowanie konstrukcji wartstwy p-bazy, aby oba podane wcześniej warunki były w znacznym stopniu spełnione.



Rys. 15. Tyrystor GTO z zagrzebaną bramką

Można to uzyskać wykonując p-bazę jako konstrukcję dwuwarstwową, w której np. stosując technologię epitaksji nakłada się w pierwszej kolejności wysoko domieszkowaną warstwę przylegającą do środkowego złącza J2, a następnie nakłada się nisko domieszkowaną warstwę w której wykonuje się złącze emiterowe. Rozwinięciem tej koncepcji są tzw. tyrystory z bramką zagrzebną [34 – 36] pokazany na rys. 15.

Istotnym ograniczeniem przy wykonywaniu niskoomowej warstwy n-bazy metodami dyfuzyjnymi jest występowanie skokowego wzrostu napięcia na przewodzącym tyrystorze po przekroczeniu pewnej krytycznej wartości I_{crit} przez prąd anodowy [37]. Jest ona wywołana takim zmniejszeniem wartości współczynnika wzmocnienia prądowego α_{npn} w wyniku wzrostu koncentracji C_{FJI} na krawędzi złącza J1, że w wyniku dalszego spadku tej wartości wywołanej wzrostem prądu tyrystora, tranzystor katodowy npn wchodzi w stan pracy aktywnej. Ilościowo, efekt ten opisuje Tablica I, w której zebrano wyniki badań eksperymentalnych przeprowadzonych na

Tablica 1

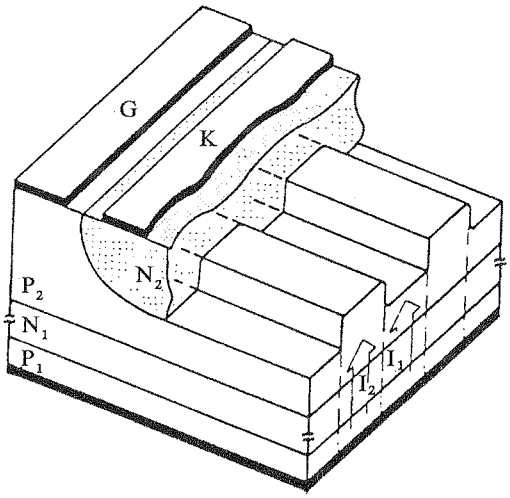
Wymiary i koncentracja domieszek oraz odpowiadające im prądy krytyczne w testowych strukturach tyrystorów GTO [37]

w_{NE}	w_{PB}	w_{NB}	C_{NE}	C_{PB}	C_{NB}	C_{FJI}	I_{crit}
μm	μm	μm	cm^{-3}	cm^{-3}	cm^{-3}	cm^{-3}	A
9.2	43.3	150	$5\ 10^{20}$	$2.8\ 10^{18}$	$2.8\ 10^{14}$	$2.1\ 10^{18}$	20 – 40
11.6	41.9	150	$4\ 10^{20}$	$2.8\ 10^{18}$	$2\ 10^{14}$	$1.8\ 10^{18}$	100 – 300
13.3	40.7	150	$3\ 10^{20}$	$2.8\ 10^{18}$	$2\ 10^{14}$	$1.5\ 10^{18}$	> 1800
15.4	39.8	150	$2.5\ 10^{20}$	$2.8\ 10^{18}$	$2\ 10^{14}$	$1.2\ 10^{18}$	> 1800

różnie domieszkowanych testowych strukturach tyrystorów GTO zawierających 260 segmentów katodowych o łącznej powierzchni katody 3 cm^2.

- *grubość warstwy p-bazy*

Wąska p-baza pozwala na uzyskanie większego współczynnika wzmocnienia α_{npn} tranzystora katodowego, a także poprawia szybkość załączania i zmniejsza straty przy załączaniu. Jest to jednak związane z większą wartością rezystancji poprzecz-



Rys. 16. Struktura tyrystora GTO z bramką podwójnie rozwiniętą [39]

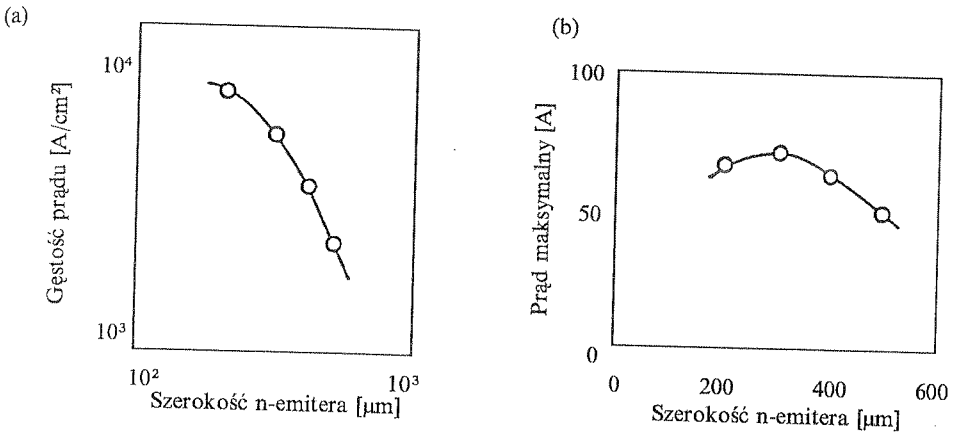
nej warstwy p-bazy, co jst niekorzystne z punktu widzenia maksymalnego prądu wyłączalnego. Wymaga to znalezienia pewnego kompromisu pomiędzy tymi sprzecznymi wymaganiami [38]. Oryginalną próbą uzyskania takiego rozwiązania na drodze zmian konstrukcyjnych jest pokazana na rys. 16 koncepcja tyrystora z bramką podwójnie rozwiniętą [39–41], w której wprowadzono w obrębie jednej sekcji dodatkowy podział na podsekcje o naprzemiennie wąskiej i szerokiej n-bazie.

• *geometria „palaca" emiterowego*

Specyficzną cechą tyrystora GTO jest bardzo wyrafinowany podział powierzchni od strony katody na obszar emitera katodowego i obszar bramki. Zwykle jest to kilkaset krótkich i wąskich „palców" emiterowych rozlokowanych centrycznie w kilku pierścieniach i rozdzielonych obszarem bramki. Rozlokowanie i wymiary tych obszarów są bardzo istotne i dlatego wiele uwagi poświęcono ocenie ich wpływu na poprawną pracę przyrządu [19, 26, 28, 42, 43].

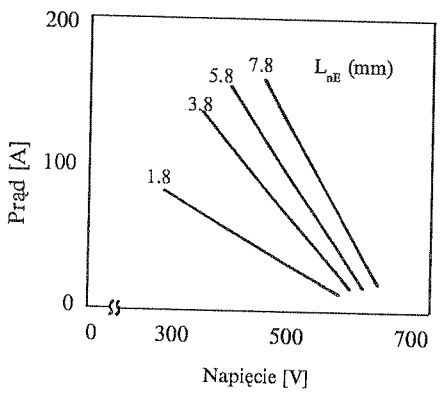
Wielkościami charakterystycznymi dla pojedynczego „palca" emiterowego jest jego szerokość b_{ne} oraz długość l_{ne}. Aby uzyskać jak największą poprzeczną rezystancję n-bazy pod „palcem" emiterowym, powinien on być wąski i długi. W praktyce obie te wielkości nie mogą jednak przyjmować wartości ekstremalnych. W przypadku szerokości b_{ne} istotny wpływ ma tu kwestia wpływu tej wielkości na przebieg ściągania prądu i formowania się kanału prądowego przy przejściu z fazy opóźnienia do fazy opadania. Ilustrują to przedstawione na rys. 17 wyniki pomiarów przeprowadzone dla pojedynczego segmentu tyrystora GTO [26]. Chociaż gęstość wyłączalnego prądu tyrystora rośnie wraz ze zmniejszaniem b_{ne} (rys. 17a), to wartość wyłączalnego prądu segmentu, po osiągnięciu swego maksimum dla ok. 300 μm, zaczyna maleć (rys. 17b). W praktyce wielkość b_{ne} jest utrzymywana w przedziale 250–300μm [20, 24, 26, 38].

Długość l_{ne} i „palca" emiterowego może być znacznie większa, ale jej wpływ na polepszenie parametrów przyrządu nasyca się [19, 26]. Ilustruje to rys. 18, na którym



Rys. 17. Wpływ szerokości „palca" emiterowego na gęstość maksymalnego prądu wyłączanego (a) oraz maksymalny prąd segmentu (b) [26]

zebrano wyniki pomiarów bezpiecznej wartości prądu tyrystora w zależności od blokowanego napięcia, uzyskane dla partii tyrystorów o zmiennej długości „palca" emiterowego i stałej wartości jego szerokości równej 300 μm. W praktyce najczęściej przyjmuje się wartości długości l_{ne} z przedziału 2 – 4 mm.



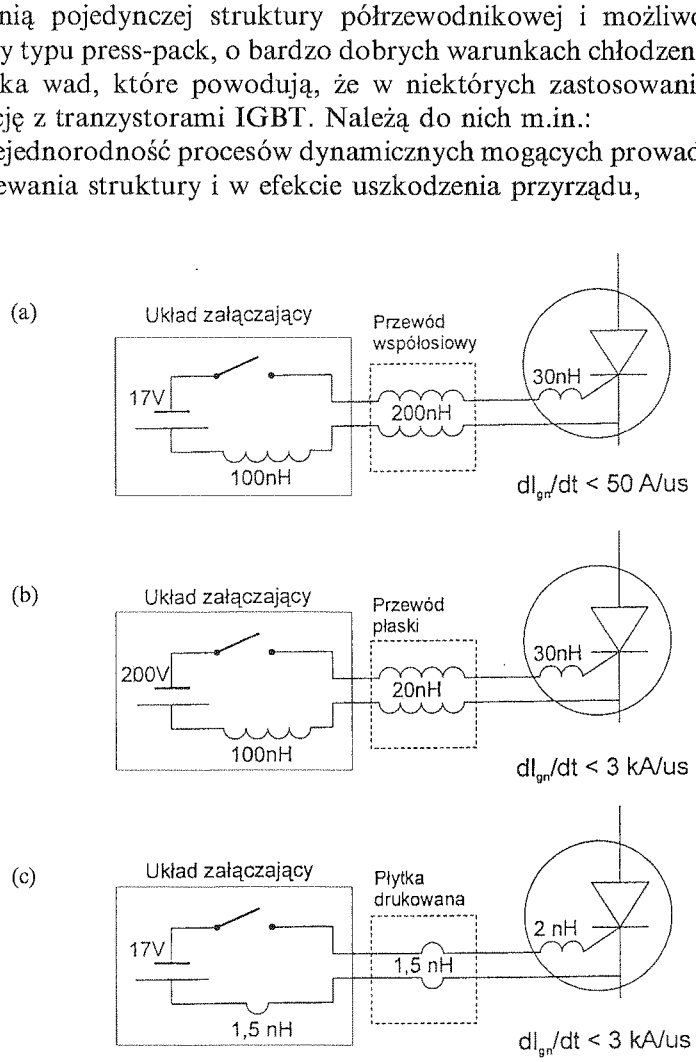
Rys. 18. Wpływ długości „palca" emiterowego na zależność maksymalnej dopuszczalnej wartości prądu od blokowanego napięcia [26]

Ostatnim parametrem geometrycznym, istotnym dla warunków pracy tyrystora GTO jest odstęp pomiędzy sąsiednimi „palcami" oraz grubość warstwy aluminiowego kontaktu bramki. Przy zbyt małych odstępach pojawiają się niepożądane elektryczne oddziaływania międzysegmentowe, a poprzeczne spadki napięcia wzdłuż krawędzi „palca" mogą być na tyle duże, że spowodowane przez nie nierównomierne wysterowanie różnych obszarów segmentu może doprowadzić do uszkodzenia przyrządu w tych obszarach. Problem ten był analizowany w [28], gdzie stwierdzono, że zachowanie odstępów pomiędzy sąsiednimi palcami rzędu 500μm powinno być wystarczające dla zapobieżenia tym niepożądanym efektom.

4. ICGT — TYRYSTOR GTO ZE ZINTEGROWANYM UKŁADEM STEROWANIA

Przedstawiony wyżej opis pracy tyrystora GTO odpowiadał typowym warunkom jego sterowania przedstawionym na rys. 1 i 2, a omówione sposoby polepszenia jego parametrów koncentrowały się na wprowadzeniu odpowiednich zmian w konstrukcji i technologii wykonywania samej struktury półprzewodnikowej. W efekcie tyrystory te posiadają szereg parametrów nieosiągalnych dla większości innych przyrządów półprzewodnikowych mocy. Charakteryzują się one dużą gęstością prądu i małym spadkiem napięcia w stanie przewodzenia, dużym napięciem blokowania oraz dużą powierzchnią pojedynczej struktury półrzewodnikowej i możliwością zamykania w obudowy typu press-pack, o bardzo dobrych warunkach chłodzenia. Posiadają one jednak kilka wad, które powodują, że w niektórych zastosowaniach przegrywają konkurencję z tranzystorami IGBT. Należą do nich m.in.:

- duża niejednorodność procesów dynamicznych mogących prowadzić do lokalnego przegrzewania struktury i w efekcie uszkodzenia przyrządu,

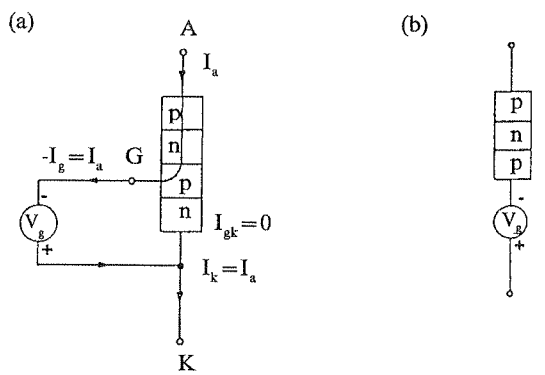


Rys. 19. Kolejne etapy od tyrystora GTO do tyrystora IGCT: (a) standardowy układ z tyrystorem GTO; (b) układ z twardym sterowaniem; (c) tyrystor IGCT [49]

- konieczność stosowania w wielu przypadkach tłumików RCD, zabezpieczających przez nadmiernymi stratami przy wyłączaniu,
- kosztowne i złożone uklady sterowania generujące impulsy bramkowe.

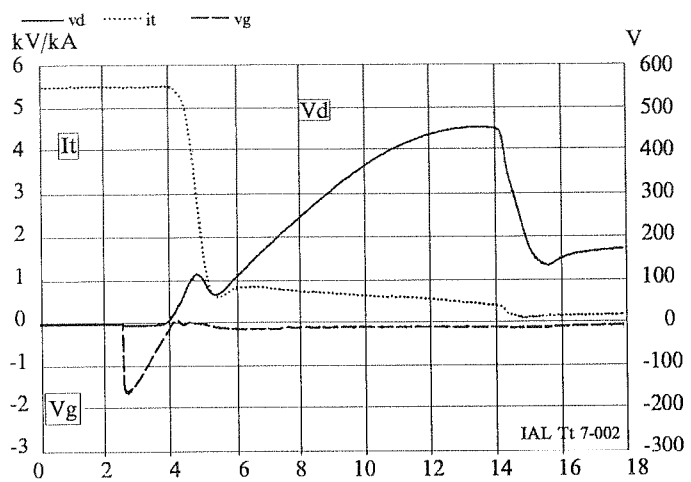
Wady te zostały w latach 90-tych w znacznym stopniu usunięte głównie poprzez zmiany w sposobie sterowania tyrystorem GTO oraz w samym układzie sterowania. Efektem tych zmian jest tyrystor IGCT, mogący skutecznie konkurować z przyrządami mocy sterowanymi napięciowo. Drogę od standardowego tyrystora GTO do tyrystora IGCT ilustruje rys. 19.

Pierwszym krokiem w kierunku polepszenia parametrów użytkowych tyrystorów GTO było wprowadzenie tzw. twardego sterowania (hard drive), polegającego na zastosowaniu krótkich impulsów bramkowych o bardzo dużej amplitudzie. Rozwiązanie to nie jest nowe i było po raz pierwszy zademonstrowane w zastosowaniu do małych tyrystorów już w roku 1986 [46], jednak jego rzeczywiste możliwości zostały odkryte na nowo dopiero w roku 1996 [4, 48]. Opiera się ono na spostrzeżeniu, że po osiągnięciu pewnej szybkości narastania prądu bramki i odpowiednio dużej amplitudzie tego prądu obserwuje się w tyrystorze GTO przejęcie całego prądu anody przez kontakt bramki zanim pojawią się w nim opisane w rozdz. 2.1.2 zjawiska prowadzące do powstania obszarów o dużej gęstości prądu. Osiągnięcie maksymalnej amplitudy prądu bramki, równej co najmniej wyłączanemu prądowi tyrystora (wzmocnienie prądowe $g = 1$), musi nastąpić w początkowej fazie czasu opóźnienia i trwać nie dłużej niż ok. 1 μs, tak, że prąd katody przestaje płynąć zanim nastąpią jakiekolwiek zmiany rozkładów przestrzennych nośników i potencjału w pozostałych obszarach tyrystora. W efekcie proces wyłączania przestaje być procesem wyłączania struktury czterowarstowej i staje się procesem wyłączania struktury tranzystora pnp, jak to ilustruje rys. 20. Proces ten, analogicznie jak np. w tranzystorze IGBT, jest procesem zachodzącym równomiernie w całej objętości przyrządu, bez lokalnych obszarów o dużej gęstości wydzielania ciepła, co pozwala na jego bezpiueczną pracę bez konieczności stosowania tłumików.



Rys. 20. Tyrystor GTO wyłączany metodą „hard drive" (a) oraz obwód zastępczej metody z tranzystorem pnp (b)

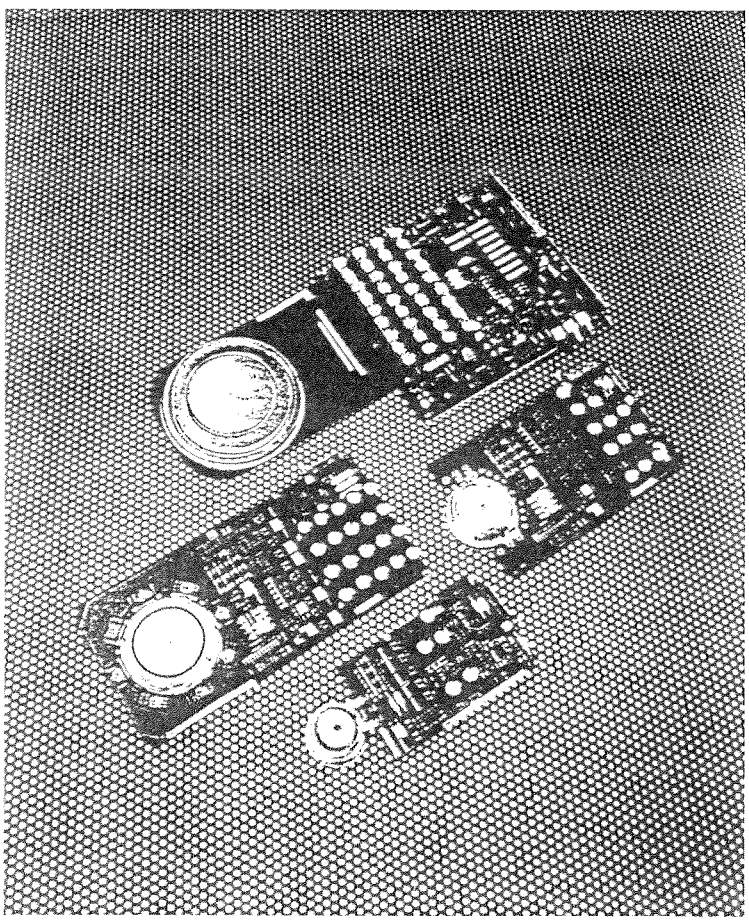
Uzyskanie warunków pozwalających na twarde sterowanie procesem wyłączania wymaga bardzo starannego zaprojektowania układu formowania impulsów bramkowych, co ilustruje rys. 19. Pokazuje on kolejne modyfikacje ukladu sterowania tyrystorem GTO 5SGA 30L4502 firmy ABB, o wartościach znamionowych prądu i blokowanego napięcia, odpowiednio, 3kA i 4.5 kV [48]. Na rys. 19a przedstawiono sytuację przy standardowym sterowaniu tyrystorem GTO. Wewnętrzna indukcyjność układu sterowania jest rzędu 100 nH, indukcyjność kabla koncentrycznego łączącego układ z tyrystorem rzędu 200 nH, co daje łączną indukcyjność rzędu 300 nH. Oznacza to, że przy zastosowaniu standardowego napięcia zasilającego rzędu 17 V można osiągnąć stromość nie większą niż 50 A/μs, a osiągnięcie pożądanej stromości rzędu 3 kA/μs wymagałoby zastosowanie w tym układzie źródła o napięciu co najmniej 900 V i wiązałoby się z energią 1.35 J magazynowaną w indukcyjności podczas impulsu. Są to warunki trudne do zaakceptowania i dlatego opracowano dla celów twardego sterowania nową konstrukcję układu formowania impulsu bramki, której odpowiada rys. 19b. W wyniku zastosowania specjalnego płaskiego kabla łączącego tyrystor z układem oraz niskoindukcyjnego wielowarstwowego obwodu drukowanego udało się zmniejszyć całkowitą indukcyjność do ok. 60 nH. Pozwoliło to na ograniczenie napięcia zasilania do ok. 200 V, a energii magazynowanej w indukcyjności do ok 250 mJ. Proces wyłączania w tych warunkach jest pokazany na rys. 21.



Rys. 21. Oscylogram procesu wyłączania metodą „hard drive" tyrystora GTO typu 5SGA 30L4502 o parametrach 3 kA/4.5 kV [48]

Rozwiązanie z rys. 19b pozwalało już na praktyczne zastosowanie metody twardego sterowania w tyrystorach GTO, które dzięki temu mogły pracować ze znacznie większymi stromościami di/dt (twarde sterowanie przy załączaniu), większymi stromościami du/dt (możliwość rezygnacji z tłumików) oraz większymi prądami wyłączalnymi i mniejszymi czasami wyłączania niż te, deklarowane jako ich

parametry znamionowe. Ich słabością był jednak skomplikowany uklad sterowania oraz konieczność stosowania dużych napięć zasilania w tym układzie. Problem ten został radykalnie rozwiązany przez wprowadzenie zaintegrowanej konstrukcji obejmującej płytkę z układem sterującym oraz tyrystor GTO w nowej zmodyfikowanej tzw. współosiowej obudowie o współosiowym doprowadzeniu sygnału sterującego. W obudowie tej elektrodę bramki stanowi metaliczny pierścień (z otworami na rysunku), do którego prąd bramki jest dostarczany na całym jego obwodzie poprzez szereg równoległych doprowadzeń. Pierścień ten posiada bezpośredni kontakt z pierścieniowym polem kontaktowym na pastylce półprzewodnikowej tyrystora. W efekcie udało się ograniczyć indukcyjność obudowy tyrystora do ok. 2 nH, a łączną indukcyjność układu, pokazanego schematycznie na tys. 19c do ok. 5 nH. Pozwoliło to na zmniejszenie wymagań dotyczących układu sterowania, m.in. umożliwiło ograniczenie napięcia zasilania do wartości poniżej 20 V, oraz jeszcze polepszyło parametry dynamiczne przyrządu.



Rys. 22. Widok rodziny układów IGCT oferowanych przez firmę ABB

Tak zintegrowany układ zawierający tyrystor GTO wbudowany w płytkę drukowaną z układem sterującym generującym impulsy o parametrach wymaganych przy twardym sterowaniu jest oferowany jako IGCT (Integrated Gate-Commutated Thyristore) przez firmę ABB oraz jako GTC (Gate Commutate Turn-off) praz firmę Mitsubishi. Na rys. 22 przedstawiono widok rodziny układów IGCT oferowanych przez firmę ABB. Są to układy zasilane zewnętrznym napięciem na poziomie 19 V, załączenie i wyłączenie tyrystora jest inicjowane sygnałem zewnętrznym dostarczanym poprzez sprzęg światłowodowy, a same tyrystory mogą przełączać prądy do 4 kA przy napięciach blokowania do 4.5 kV.

5. PODSUMOWANIE

Mimo konkurencji ze strony przyrządów półprzewodnikowych sterowanych napięciowo, tyrystory GTO stanowią w dalszym ciągu istotny element rynku przyrządów półprzewodnikowych mocy, a w zastosowaniach do układów wysokonapięciowych oraz wymagających bardzo dużych prądów są w dalszym ciągu bezkonkurencyjne. Są one stosunkowo starym przyrządem, gdyż pierwsze tyrystory GTO pojawiły się ok. 30 lat temu i do niedawna wydawało się, że to co można w nich ulepszyć zostało już zrobione i nie należy oczekiwać rewolucyjnych zmian w ich konstrukcji prowadzących do istotnych zmian ich własności. Dodatkowo, pojawienie się w latach 80 tranzystorów IGBT wydawało się zwiastować zmierzch tyrystorów GTO w szeregu zastosowań nie wymagających przełączania bardzo dużych mocy. Druga połowa lat 90-tych i pojawienie się twardo sterowanych tyrystorów GTO w modułach IGCT pokazała jednak, że ta opinia była zdecydowanie przedwczesna. Wprowadzane obecnie na rynek tyrystory IGCT i GCT mogą z powodzeniem konkurować z tranzystorami IGBT, a pod pewnymi względami, np. przeciążalność prądowa, wielkość pojedynczej pastylki czy napięcie przewodzenia, zdecydowanie nad nimi górują. W Polsce tyrystory GTO nie są obecnie produkowane, aczkolwiek w ramach realizowanego w Politechnice Łódzkiej projektu badawczego KBN nr 8 S501 043 05 [50] zaprojektowano i wykonano w LAMINA Semicoductors International modelowy egzemplarz takiego tyrystora.

W pracy przedstawiono genezę powstania tyrystora GTO szczególny nacisk kładąc na procesy występujące przy jego wyłączaniu bramkowym, które stanowią o jego specyfice w porównaniu ze standardowymi tyrystorami. Omówiono problemy technologiczne związane z realizacją tyrystorów GTO oraz dokonano przeglądu rozwiązań konstrukcyjnych wprowadzanych do jego struktur na przestrzni od powstania koncepcji tyrystora GTO do chwili obecnej. Szczególnie dużo uwagi poświęcono najnowszym pracom nad polepszeniem parametrów eksploatacyjnych tych tyrystorów związanych z koncepcją „twardego sterowania". Jest ona podstawą najnowszej rodziny tyrystorów, tzw. tyrystorów IGCT, które stanowią odpowiedź konstruktorów i producentów bipolarnych przyrządów półprzewodnikowych mocy na obserwowaną ostatnio ekspansję tranzystorów IGBT.

BIBLIOGRAFIA

1. R. H. Van Ligten, D. Navon: *Base turn-off p-n-p-n switches* IRE Wescon Conv. Record. Part 3: Electron Dev., 1960, pp. 49 – 52
2. A. K. Joncher: *Notes of the theory of four-layer semiconductor switches*: Solid-St. Electron., Vol. 2, 1961, pp. 143 – 148
3. M. Klein: *A four-layer p-n-p-n switching device* IRE Trans. Electron Devices, Vol. ED-7, 1960, pp. 314 – 217
4. T. A. Longo, M. Miller, A. L. Derek, J. D. Eknaian: *Planar epitaxial pnpn switch with gate turn-off gate* IRE Wescon Conv Record, Part 3: Electron Dev., 1962, pp.
5. E. D. Wolley: *Gate turn-off in p-n-p-n devices* IEEE Trans. electorn Dev., Vol. ED-13, 1966, pp. 590 – 597
6. T. C. New, W. D. Frobenius, T. J. Desmond, D. R. Hamilton: *High power gate-controlled switch.* IEEE Trans. Electron Dev., Vol. ED-17, 1970, pp. 7-6-710
7. E. D. Wolley, R. Yu, R. Steigerwald, F. M. Matterson: *Conf. Rec. IEEE IAS Meet.*, 1973, pp. 251 – 255
8. M. Azuma, A. Nakagawa, K. Takigani: *High power gate turn-off thyristors*, Jap. J. Appl. Phys., Vol. 17, Sup. 17-1, 1978, pp. 275 – 281
9. Z. Lisik: *Rozwzój konstrukcji tranzystorów IGBT.* Kwart. Elektroniki i Telekom., Vol. 42, 1996, ss. 207 – 225
10. S. Linder, S. Klaka, M. Frecker, E. Carroll, H. Zeller: *A new range of reverse conducting gate-commutated thyristors for high-voltage, medium power applocation.* Proc. EPE'97, Trondheim 1997, pp. 1.117-1.124
11. E. Carroll, S. Klaka, S. Linder: *Integrated gate-commutated thyristors: A new approach to high power elektronics.* Proc. IEMDC'97, Milwaukee 1997
12. K. Satoh, M. Yamamoto, T. Nakagawa, A. Kawakami: *A nwe high power device GCT (Gate Commutated Turn-off) thyristor.* Proc. EPE'97, Trondheim 1997, pp. 2.070 – 2.075
13. E. Masada, M. Tamura, K. Nakajima: *Simulation of switch processes in turn — off thyristors.* IEEE PESC'88, pp. 91 – 98
14. A. Nakagawa, H. Ohashi: *A study on GTO turn-off failure mechanism — A time and temperature-dependent 1-D model analysis.* IEEE Trans. Electron Devices, Vol. D-31, 1984, pp. 273 – 279
15. M. Kurata: *A new CAD-model of a gate turn-off thyristor.* Conf. Rec. IEEE PESC'74, 1974, pp. 125 – 133
16. M. Naito, T. Nagano, H. Fukui. Y. Tarasawa: *One-dimensional analysis of turn-off phenomena for a gate turn-off thyristor.* IEEE Trans Electron Dev., Vol. ED-26, 1979, pp. 226 – 231
17. Y. Shimizu, M. Natio, M. Odamura, Y. Terasawa: *Numerical analysis of turn-off characteristics for a gate turn-off thyristor with a shorted anode emitter.* IEEE Tran, electron Dev., Vol. ED-28, 1981, pp. 1043 – 1047
18. H. Ohashi, A. Nakagawa: *A study on GTO turn-off failure mechanism.* Tech. Dig. IEEE IEDM'81, 1981, pp. 414 – 417.
19. M. Azuma, M. Kurara, H. Ohashi: *Desing consideration for high power GTO's.* Jap. J. Appl. Phys., Vol. 20 – 1, pp. 93 – 98
20. T. Yatsuo. T. Nagano, H. Fukui, M. Okamura, S. i Sakurada: *Ultrahigh-voltage high-current gate turn-off thyristors* IEEE Trans. Electron Dev., Vol. ED-31, 1984, pp. 1681 – 1686
21. E. Huang., J. Barnes: *GTO optimisation*
22. W. J. Findlay, P. D. Taylor, R. TDenter, L. Coulbeck: *Optimised lifetime control in GTO thyristors.* Conf. Rec. IEEE IAS'87, 18 – 23, Oct. Atlanta, 1987, pp. 540 – 543
23. W. J. Findlay, L. Coulbeck, A. D. Milligton: *GTO thyristor optimisation.* IEE Coll. GTO's, Rival Dev. & Appl., 7 March 1988 London
24. H. Bleicher, K. Nordgren, M. Rosling. M. Bakowski, E. Nordllander: *The effect of emoitter shortings on turn-off limitations and device failure in GTO thyristor under snubberless operation.* IEEE Trans. Electon. Dev., Vol. 42, No 1, 1995, pp. 178 – 186

25. T. B Nagano, M. Okamura, T. Ogawa: *A high-power low-forward-drop gate turn-off thyristor*. Conf. Rec. IEEE IAS'78, 1978, pp. 1003
26. T. Nagano, H. Fukui, T. Yasuo, M. Okamura: *A snubber-less GTO*. Conf. Rec. IEEE PESC'82, 1982, pp. 383–367
27. M. Azuma: *GTO thyristors*. Proc. IEEE, Vol. 76, 1988, pp. 419–427
28. M. Bakowski, K. Thorselius, F. Vojdani: *Investogation of the turn-off behaviour for the desing and development of high current GTO thyristors*. 12 Nordic Semicond. Meeting, Jernaker June 1986, pp. 206–215
29. A. Jaecklin: *Gate turn-off thyristors with near perfect technology*. IEEE IEDM, 86, 7–10 Dec. Los Angeles, 1986, pp. 114–117
30. E. Falck, W. Gerlach, W. Reckel: *Calculation of the blocking capability of SOI power devices under the influence of interconnections*. Int. Sem. Power Semicon., 31.09–2.09 1994 Prague, 1994, pp. 37–44
31. A. Blicher: *Thyristor physics*. Springer Verlag, New York Heidelberg Berlin, 1976
32. F. E. Gentry, F. W. Gutzwiller, N. Golonyak, E. E. Von Zastov: *Semiconductor Controlled Rectifiers*. Englewood Cliffs, N. J., Prentice-Hall, 1969
33. Taylor: *Thyristor desing and realization*, John Willey & Sins, Chichester, 1987
34. T. Suzuki, T. Sueoka: *Buried-gate turn-off thyristor (GTO) with amplifying gate structure for power systems*. Meind Rev., Vol. 67, 1983, No 1, pp. 1–4
35. S. Ishibashi. T. Kubo, T. Kawamura, T. Suzuki, T. Sueoka: *High power gate turn-off thyristor with amplifying gate structure*. Conf. Rec. IPEC'83, Tokyo March 1983, pp. 33–41
36. T. Suzuoki, T. Ugazin, M. Kekura, T. Watanable, T. Sueoka: *Switching charakte–ristics of high power buried gate turn-off thyristor*. IEEE IEDM'82, San Francisco 1982, pp. 492–495
37. M. Azuma, K. Takigami: *Anode current limiting effect of high power GTOs*. IEEE Electron Dev. Lett., Vpl. EDL-1, No 10, 1980, pp. 203–205
38. M. Azuma, M. Kurata, K. Takigami: *2500 V 600 A gate turn-off thyristor (GTO)*. IEEE Trans Electron Dev., Vol. Ed-28, No 3, 1981, pp. 270–274
39. A. Silard, S. Rusu, F. Turtudau, B. Kosa: *A double-interdigitated GTO switch*. IEEE Trans. Electron Dev.., Vol. ED-31, 1984, pp. 322–328
40. A. Silard, S. Rusu: *The TIL GTO Thyristor — a power switch with unique turn-off features*. IEEE Electron Dev. Lett., Vol. EDL-4, 1983, pp. 347–349
41. A. Silard. S. Rusu: *A novel gate-cathode configuration with two interdigitation levels for GTO thyristors*. IEEE Electron Dev. Lett., Vol. EDL-4, No 6, 1993, pp. 188–190
42. W. Gerlach: *Qu, FGTO — a fine structured GTO-thyristor with improved switching characteristics*. Arch. für Elektrotechnik, Vol. 74, 1991, pp. 433–444
43. M. Stoisiek, K. G. Oppermann, G. Wachutka: *Turm-off behaviour of GTOs with small emitter elements*. Arch Elektron Ubertrag. Tech., Vol. 43, No 5, 1989, pp. 320–327
44. S. Eicher, F. Bauer, A. Weber, H. R. Zeller, W. Fichtner: *A high-power low-loss GTO with adjustable IGT*. Proc. ISPSD'96, Maui 1996, pp. 262–264
45. S. Eicher: *The transparent anode GTO (TGTO): a new low-loss power switch*. Series in Microelectronics, Vol. 58, Hartunkg-Gorre, Konstanz, 1996
46. F. Wirth: *High speed snubberless operation of GTO's using a new gate drive technique*. Proc. IEEE IAS Conf., 1986, pp. 453–457
47. H. Grüning, E. Zuckerberger: *Hard drive high power GTOs: Better switching capability obtained through improved gate-units*. Proc. IEEE IAS Conf., 1996, pp. 1474–1480
48. H. Grüning, B. Odegard, J. Rees: *High-power hard-drive GTO module for 4.5 kV/3kA snubberles operation*. Proc. PCIM'96. Nürmberg 1996, pp. 169–183
49. H. Grüning, B. Odegard: *High performace low cost MVA inverters realised with integrated gate commutated thyristors (IGCT)*. Proc. EPE'97 Trondheim 1997, pp 2.060–2.065
50. *Raport z realizacji projektu badawczego KBN Nr 8 S501 043 05 pt.: „Termiczna optymalizacja konstrukcji przyrządów półprzewodnikowych mocy z bramką palczastą"*. Łódź 1997, kier. Projektu Z. Lisik

Z. LISIK

CONSTRUCTION AND DESING OF GTO THYRISTORS

Summary

Inspite of a large expansion of field control power semiconductor devices, the bipolar power semiconductor device with their main representative, GTO thyristor, remain still the basic ones for high power applications. Within the compass of almost 30 years, i. e. sine the first GTO thyristor was presented, the construction of GTO thyristors was changed significantly in order to improve their features and to be equal to increasing demands of power semiconductor device mark as well.

In the paper the prnciples of GTO thyristor work and desing have been presented in detail with a special emphasis placed on the phenomena which are characteristic for these devices. A special attention has been also devoted the presentation of a new and very promising solution of GTO, called IGCT (Integrated Gate Commutated Thyristor). It is the newest achievement of the GTO producers and desingers, which is their response on the IGBT competition on the power semiconductor market.

Key words: power semiconductor devices, GTO thyristors, semiconductor devide design

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI, 2000, **46**, z. 1, ss. 121

SPIS TREŚCI

CONTENTS

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.
Objętośc 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 wydrkuku 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 wyróznienia 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 odstepem między wierszami, tytuły i podtytuły małymi literami. Marginesy z każdej strony powinny mięc 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 odnosza. 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 (International 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 tekście 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 modeling of the MOS transistors. Mc Graw-Hill, New York 1987, pp. 141 – 148

Materiały ilustracyjne: 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 Redkacji 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.

Redkacja prosi Autorów o powiadomienie ją o zmianie miejsca pracy i adresu prywatnego.

INFORMATION FOR THE AUTORS OF K.E.T.

The KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI (K.E.T.) — ELECTRONICS AND TELECOMMUNICATION QUARTERLY is a journal of the Committee for Electronics and Telecommunication of the Polish Academy of Sciences and is a direct continuation for 46 years of the quarterly ROZPRAWY ELEKTROTECHNICZNE, that has been published also under the auspices of the Polish Academy of Sciences.
Its pages contain scientific articles concerned with theory and application of electronics, telecommunication, microelectronics, optoelectronics, radioengineering 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 developments ina given technology.
Regular papers are accepted for publication in Polish or English, 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 margin or using of wordprocessing with text formatters based on MS DOS, APPLE/MAC such as Microsoft word. In any case the text formatter must be specified on the Author/s, their Affiliations and relevant addresses.
All text pages should be numbered. Regular papers should have a length of no more than 30 pages.

Printing types and symbols

For letteral symbols, units and graphical symbols, the recommendation of IEE (International Electrotechnical Commission) and ISO (International Orgnisation of Standarisation) must be followed.
- 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 formulae should be written at right side, between round brackets.
- The serial number of footnotes should be written between round brackets and upper.
- Mathematical detials, 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 sets of illustrations.