



ZESZYTY NAUKOWE WYDZIAŁU  
**ELEKTRONIKI  
I INFORMATYKI**  
POLITECHNIKI KOSZALIŃSKIEJ NR 3



POLITECHNIKA KOSZALIŃSKA

**Zeszyty Naukowe  
Wydziału Elektroniki i Informatyki**

**Nr 3**

KOSZALIN 2011

Zeszyty Naukowe Wydziału Elektroniki i Informatyki Nr 3

ISSN 1897-7421

Przewodniczący Uczelnianej Rady Wydawniczej  
*Bronisław Słowiński*

Przewodniczący Komitetu Redakcyjnego  
*Aleksy Patryn*

Komitet Redakcyjny  
*Mirosław Maliński*  
*Dariusz Gretkowski*  
*Krzysztof Bzdyna*

Projekt okładki  
*Tadeusz Walczak*

Skład, łamanie  
*Maciej Bączek*

© Copyright by Wydawnictwo Uczelniane Politechniki Koszalińskiej  
Koszalin 2011

Wydawnictwo Uczelniane Politechniki Koszalińskiej  
75-620 Koszalin, ul. Raławicka 15-17

---

Koszalin 2011, wyd. I, ark. wyd. 7,22, format B-5, nakład 150 egz.  
Druk: ESUS, Poznań

## Spis treści

<i>Mirosław Maliński</i> .....	5
Wspomnienia o profesorze Andrzeju Guzińskim	
<i>Radosław Łuczak</i> .....	9
A gdyby tak posterować prądem ...	
<i>Robert Berezowski</i> .....	29
Minimalizacja opisów funkcji w algebrze bramek prądowych	
<i>Lukasz Bartłomiej Chrobak, Mirosław Andrzej Maliński</i> .....	39
Badania parametrów rekombinacyjnych materiałów krzemowych z wykorzystaniem nieniszczącej techniki MFCA opartej na zjawisku modulacji absorpcji na nośnikach swobodnych	
<i>Dariusz Bernatowicz, Kamil Szymański</i> .....	49
Realizacja fragmentacji mieszanej w natywnych dokumentach XML	
<i>Dariusz Jakóbczak</i> .....	59
Shape Coefficients via Method of Hurwitz-Radon Matrices	
<i>Leszek Bychto</i> .....	73
Układ monitorowania systemu fotowoltaicznego	
<i>Alicja Kilińska – Wiśniewska</i> .....	81
Pomiar czasu przetwarzania rozproszonych zapytań w ewoluujących silnikach serwerów baz danych SQL Microsoft Server	
<i>Wiesław Madej</i> .....	89
Specjalizowany cyfrowy system predykcijny ciągłym pomiarem odległości	
<i>Robert Świta, Zbigniew Suszyński</i> .....	111
Hybrid watershed and region grow segmentation algorithm of color images	
<i>Anna Witenberg, Maciej Walkowiak</i> .....	123
Rozwiązanie równania pola elektrycznego modelującego anteny liniowe w dziedzinie czasu	
<i>Małgorzata Górską, Antoni Drapella</i> .....	133
Do primary event uncertainty distributions impact top event distribution?	





**Prof. dr hab. inż. Andrzej GUZIŃSKI**  
**twórca Wydziału Elektroniki i Informatyki Politechniki Koszalińskiej**

**Mirosław Maliński**

Wydział Elektroniki i Informatyki

Politechnika Koszalińska

## **Wspomnienia o profesorze Andrzeju Guzińskim**

Profesor Andrzej Guziński urodził się w Teczewie 19.06.1939 r. Studia wyższe ukończył w roku 1963 na Wydziale Elektroniki Politechniki Gdańskiej. Na tym samym Wydziale uzyskał w 1970 roku stopień naukowy dr nt. oraz w roku 1975 stopień naukowy dr hab. Tytuł naukowy profesora uzyskał w roku 1987.

Swoją karierę zawodową rozpoczął w ZR Radmor Gdynia, gdzie pracował w latach 1963-1973. Od roku 1971 rozpoczął pracę na Politechnice Gdańskiej w Instytucie Technologii Elektronicznej jako adiunkt, następnie jako docent, a od roku 1992 – profesor. W latach 1983 – 1988 związał swoje losy zawodowe z Wyższą Szkołą Morską w Gdyni.

Od roku 1988 rozpoczął swoją działalność naukową i organizacyjną na Politechnice Koszalińskiej (wówczas Wyższej Szkole Inżynierskiej), gdzie z wielką determinacją i ogromną pasją zaczął tworzyć ELEKTRONIKĘ od podstaw. W 1991 profesor objął funkcję dyrektora Instytutu Elektroniki, a od 1997 do 2001 roku zajmował stanowisko dziekana Wydziału Elektroniki.

Wybitne uzdolnienia, nadzwyczajna pracowitość oraz ogromna intuicja naukowa sprawiły, że jego prace naukowe zawsze należały do dziedziny najbardziej aktualnej w określonym czasie i stanowiły poważny wkład naukowy rozwijający tę dziedzinę.

Dziełem profesora było między innymi:

- opracowanie przemysłowej technologii wytwarzania układów hybrydowych w ZR Radmor (1963-1970),
- opracowanie pierwszych w Polsce układów hybrydowych i pierwszych konstrukcji z ich zastosowaniem w ZR Radmor (1968-1973),
- napisanie 2 monografii książkowych z zakresu technologii i konstrukcji układów hybrydowych stanowiących podsumowanie działalności w latach 1963-1970,
- wniesienie istotnego wkładu do teorii i realizacji technologicznej filtrów aktywnych RC w szczególności filtrów o stałych rozłożonych (1970-1976),
- zainicjowanie i rozwinięcie w Polsce badań z zakresu automatycznego projektowania konstrukcji aparatury elektronicznej i projektowania layoutu układów scalonych VLSI (1978-1981)

- kierowanie zespołem, który opracował 2 systemy automatycznego projektowania obwodów drukowanych dla ZR Radmor (1976-1981) i systemu STOLEM do projektowania układów ASIC dla OBREUS w Toruniu,
- zaproponowanie oryginalnej koncepcji scalenia w technologii CMOS filtrów aktywnych,
- zaproponowanie oryginalnej koncepcji sterowania tranzystora MOSFET z zacisków podłoża i zastosowanie jej do konstrukcji wzmacniaczy transkonduktancyjnych o dużej dynamice,
- zainicjowanie w Polsce badań nad układami pracującymi w trybie prądowym (1991),
- zaproponowanie oryginalnej transformacji układów napięciowych w układy prądowe,
- zaproponowanie nowej klasy filtrów CSC (prądowych filtrów z przełączanymi pojemnościami),
- utworzenie szkoły naukowej w dziedzinie projektowania układów CMOS, w szczególności układów pracujących w trybie prądowym i projektowania Layoutu układów VLSI,
- zaprojektowanie oryginalnej koncepcji prądowych układów cyfrowych.

Prace te zaowocowały ok. 100 publikacjami naukowymi, z których wiele ukazało się w najbardziej prestiżowych periodykach naukowych i na konferencjach o zasięgu światowym. Profesor Guziński był autorem 5 książek. Wypromował 16 doktorów. Tematyka rozpraw doktorskich dotyczyła projektowania układów elektronicznych: filtrów aktywnych, analogowych i cyfrowych, layoutu układów VLSI.

Uznanie Profesora w środowisku naukowym w kraju i zagranicą zaowocowało wieloma prestiżowymi nagrodami i wyróżnieniami w Polsce i za granicą:

- profesor był 6-krotnie wyróżniony nagrodą MNSzWiT oraz Nagrodą Ministra MEN,
- recenzował 19 prac doktorskich, 4 habilitacyjnych, 4 książki, oraz liczne artykuły, opiniował wnioski o nagrodę Ministra, opracowania przemysłowe, wnioski o granty KBN, raporty z zakończonych projektów KBN,
- profesor był członkiem wielu komitetów naukowych, komitetu elektroniki i telekomunikacji PAN,
- członek Sekcji Elektroniki KBN (1992-1993),
- profesor kierował współpracą zespołu z uczelniami zagranicznymi,
- utrzymywał kontakt osobisty z wieloma naukowcami zagranicznymi w USA, Niemczech, Francji, Anglii, Japonii i Korei,
- profesor Guziński pracował także jako Visiting Professor na uczelniach we Francji i Maroku.

Profesor dr hab. inż. Andrzej Guziński zrobił bardzo dużo dla nauki. Należy o tym przypominać szczególnie młodym pracownikom, aby zdawali sobie sprawę, że takim jak Profesor Guziński zawdzięczają renomę nauki polskiej, a także jak wiele jeszcze muszą dokonać, aby pójść w jego ślady. Profesor pozostawił po sobie na Wydziale Elektroniki i Informatyki PK swoich wychowanków doktorantów w osobach: prof. Krzysztofa Wawryna – Prorektora a następnie Rektora PK dwóch kadencji, prof. H. Budzisz – Dziekana Wydziału dwóch kadencji oraz dr. A. Wczgraja obecnie Kanclerza Politechniki Koszalińskiej. Idee naukowe profesora są również kontynuowane na Wydziale w postaci szkoły naukowej układów prądowych.

Profesor Andrzej Guziński zmarł przedwcześnie 18.04.2001 r. w wieku 62 lat.





**Radosław Łuczak**

Katedra Inżynierii Komputerowej  
Wydział Elektroniki i Informatyki  
Politechnika Koszalińska

Koszalin, kwiecień – czerwiec 2011.

## **A gdyby tak posterować prądem ...**

18 kwietnia 2011 r. minęła X rocznica śmierci prof. dr hab. inż. Andrzeja Guzińskiego, pomysłodawcy, założyciela i pierwszego dziekana Wydziału Elektroniki Politechniki Koszalińskiej. Będąc studentem pierwszego naboru elektroniki w naszym Wydziale poznałem prof. A. Guzińskiego, uczestniczyłem w Jego wykładach, a dzisiaj kontynuuję także pracę badawczą zapoczątkowaną przez Niego.

Niech niniejszy tekst będzie sprawozdaniem z działalności naukowej całego zespołu osób, które swoim zaangażowaniem i pracą przyczyniły się do rozwoju idei **bramki prądowej**, wymyślonej przez profesora Andrzeja Guzińskiego.

### **1. Wprowadzenie**

Budowa systemów mieszanych, cyfrowo–analogowych MADS (Mixed Analog–Digital System), w jednej strukturze krzemowej napotyka na wiele trudności związanych z przenoszeniem zakłóceń, powstających w części cyfrowej, na część analogową. Zmiany stanu logicznego w układzie cyfrowym powodują skokowe zmiany prądu pobieranego ze źródła zasilania. Powstają tzw. szpilki prądowe przenoszone po szynie zasilającej i poprzez podłoże, do części analogowej. Widmo powstającej w chwili przełączenia stanu „szpilki” zawiera szereg częstotliwości mieszczących się w paśmie roboczym części analogowej układu. Dochodzi w ten sposób do powstawania zakłóceń, szumów o amplitudzie przekraczającej nawet poziom użytecznego sygnału analogowego [1, 2, 3, 4]. Z tego powodu podejmuje się działania w celu redukcji wpływu zakłóceń powstających w części cyfrowej na część analogową. Są to próby separacji części cyfrowej i analogowej, powodujące wzrost kosztów produkcji układów mieszanych, przy czym, należy nadmienić, że nie są do końca skuteczne. Izolowanie, ekranowanie, separowanie to działania mające na celu redukcję skutków, a nie przyczyny zakłóceń.

Na początku lat 90. ubiegłego wieku, prof. dr hab. inż. Andrzej Guziński, założyciel i pierwszy dziekan Wydziału Elektroniki<sup>1</sup> Politechniki Koszalińskiej zaproponował koncepcję budowy układów cyfrowych pracujących w trybie prądowym (bramki prądowe). Założeniem konstrukcyjnym było zbudowanie układu o stałym poborze mocy ze źródła zasilania. Dzięki temu zmiana stanu układu cyfrowego nie powoduje wystąpienia zakłócenia w postaci szpilki prądowej, a zatem wpływ części cyfrowej na analogową będzie znikomo mały [1, 3].

Przeprowadzono badania, zrealizowano kilka projektów badawczych, przypuszczenia profesora Guzińskiego co do przydatności jego **bramek prądowych** do budowy mieszanych układów cyfrowo–analogowych okazały się słuszne i dość szybko zostały potwierdzone: *Wyniki pomiarów wskazują, że proponowane układy cyfrowe pracujące w trybie prądowym stanowią odpowiedź na potrzebę poszukiwania nowych rozwiązań w zakresie minimalizacji wpływu procesów przejściowych towarzyszących pracy układów cyfrowych na pracę układów analogowych umieszczonych na tym samym podłożu układu scalonego w mieszanych analogowo–cyfrowych układach scalonych*<sup>2</sup>.

Dzisiaj, patrząc z perspektywy kilku minionych lat, poprzez pryzmat efektów uzyskanych przy pracy nad rozwojem technologii bramek prądowych, łatwiej jest nam ocenić istotę wynalazku profesora Guzińskiego, niż Jemu w czasie, kiedy ten powstawał. Wymyślając ideę bramki prądowej, profesorowi Guzińskiemu wydawało się, że otworzył furtkę do niskonakładowych konstrukcji mieszanych układów MADS. Nie mógł wówczas wiedzieć, że w rzeczywistości rozwarł wielkie wrota prowadzące do nowoczesnej elektroniki cyfrowej, do niespotykanej nigdzie prawdziwej logiki wielowartościowej (MVL – Multi Value Logic), do bezpiecznych systemów przetwarzających dane, do systemów kryptograficznych, wrota, o których istnieniu i znaczeniu nikt wówczas nie wiedział ...

## 2. Idea bramki prądowej

Istotą działania bramek prądowych jest wielkość fizyczna niosąca informację o stanie logicznym. W typowych układach cyfrowych ze stanem logicznym związane jest wystąpienie napięcia o określonej wartości, mieszczącej się w granicach wartości dopuszczalnych. W układach prądowych wielkością tą jest prąd wypływający z wyjścia bramki. Jest to cecha, która odróżnia bramki prądowe od innych technologii układów cyfrowych [5].

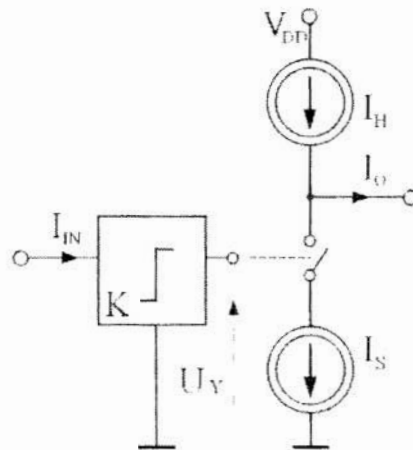
Prąd odpowiadający logicznemu poziomowi „1” ustalany jest w obwodzie polaryzacji (rys. 2). Nazywa się go prądem jedynek i oznacza  $I_1$ . Wartość prądu  $I_1$  zależna jest od technologii wykonania bramki prądowej. Dla CMOS 0,6 [ $\mu\text{m}$ ] wartość tę ustala się na

<sup>1</sup> Dzisiejszy Wydział Elektroniki i Informatyki Politechniki Koszalińskiej, wówczas Instytut Elektroniki przy Wydziale Mechanicznym.

<sup>2</sup> Za prof. dr. hab. inż. Andrzejem Guzińskim.

poziomie kilkunastu do kilkudziesięciu mikroamperów. Ponadto bramka pracuje w reżimie prądowym, tzn. niezależnie od stanu (statycznego czy dynamicznego) pobierany jest praktycznie stały prąd ze źródła zasilania. Prądowy tryb pracy układów oraz wielkość fizyczna niosąca informację o stanie logicznym zdecydowały o nazwie CMCL (Current-Mode Current Logic) dla systemów budowanych w oparciu o bramki prądowe.

Ideę działania bramki prądowej można omówić na przykładzie układu z rys. 1, przedstawiającego koncepcję budowy bramki inwertera prądowego. Zauważa się wyraźny podział na dwa moduły: moduł komparatora **K** i wyjściowy moduł inwertera **I** ( $I_H$ , klucz,  $I_S$ ). Zadaniem modułu komparatora jest reakcja na prąd wejściowy, odpowiadający połowie  $I_1$ , pojawieniem się napięcia sterującego  $U_Y$ . Napięcie to powoduje załączenie klucza i odprowadzenie prądu ze źródła  $I_H$  do masy poprzez źródło  $I_S$  (oba źródła o wydajności  $I_1$ <sup>3</sup>), prąd wyjściowy bramki ( $I_O$ ) jest wówczas bliski 0 [ $\mu$ A]. Jeżeli prąd wpływający do wejścia bramki ma niższą wartość od połowy  $I_1$ , to napięcie  $U_Y$  jest bliskie 0 [V] i klucz jest rozwarty. Prąd wyjściowy  $I_O$  ma wartość  $I_1$ , co wynika z wydajności źródła  $I_H$ . Z przedstawionego opisu wynika, że układ z rys. 1. dokonuje inwersji prądów.

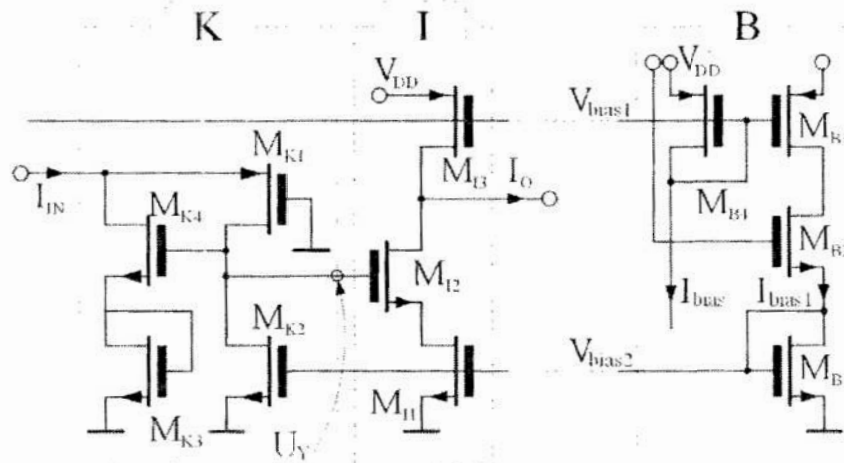


Rys. 1. Idea inwertera prądowego

Na rys. 2 przedstawiono przykład realizacji bramki typu inwerter w technologii CMOS wraz z układem polaryzacji. Jest to już trzecia generacja bramek prądowych. Inwerter wymyślony przez prof. Guzińskiego miał nieco inną, prostszą budowę. Od chwili wynalezienia, bramki prądowe przeszły kilka znaczących modyfikacji w celu poprawy ich właściwości, co pokazuje, że idea jest wciąż żywa i rozwijana.

<sup>3</sup> W bramkach 3. generacji wydajność źródła  $I_S$  jest o ok. 6% większa od  $I_H$ . Stanowi to problem występujący przy próbie budowy bramki MVL, gdyż każdy inwerter w stanie „0” odbiera ok. 6% prądu  $I_1$  od połączonego a nim inwertera w stanie „1”.





**Rys. 2.** Przykład implementacji inwertera prądowego w technologii CMOS, wraz z układem polaryzacji

Modyfikacje bramek prądowych wcześniejszych generacji dotyczyły przede wszystkim modułu komparatora **K**, który na początku realizowany był na pojedynczym tranzystorze w połączeniu diodowym. W bramkach trzeciej generacji komparator **K** zbudowany jest ze źródła prądowego, zrealizowanego na  $M_{K2}$ , o wydajności  $\frac{1}{2} I_1$  oraz układu klucz – obciążenie ( $M_{K4}$ ,  $M_{K3}$ ) odbierającego nadwyżkę prądu. W trzeciej generacji, do syntezy układów cyfrowych wykorzystuje się bramki prądowe zbudowane z udziałem 4 istniejących modułów: **K** – komparator, **I** – moduł inwertera, **AI** – moduł anti-inwertera, **SI** – prosty moduł inwersji napięć.

W tab. 1 podano nazwy i oznaczenia znanych bramek prądowych, ich budowę, realizowane operacje logiczne (różne rodzaje negacji charakterystyczne dla algebry prądowej) oraz tablice prawdy i prądy pobierane ze źródła zasilania.

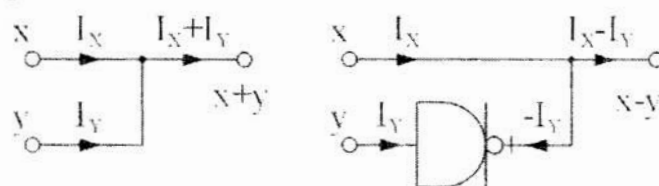
**Tab. 1.** Podstawowe typy bramek CMCL, ich budowa z dostępnych modułów, realizowane funkcje logiczne, tablice prawdy i pobór prądu ze źródła zasilania

Inwerter			podwójny inwerter		
	$y = \bar{x}$	$x \quad y$		$y = \bar{\bar{x}}$	$x \quad y$
	$I = 1 \cdot I_1$	-n 1		$I = 2 \cdot I_1$	-n 0
		0 1			0 0
		n 0			n 1
anti-inwerter			podwójny anti-inwerter		
	$y = \hat{x}$	$x \quad y$		$y = \hat{\hat{x}}$	$x \quad y$
	$I = 0 \cdot I_1$	-n 0		$I = 1 \cdot I_1$	-n -1
		0 0			0 -1
		n -1			n 0

Bramki prądowe mają budowę modułową. Każda posiada jedno wejście. Na wejściu znajduje się komparator **K**. Do komparatora podłącza się do 8 wyjść jednego z 4 typów. Wyjścia realizują 4 różne inwersje prądowe. Realizacja wyjść typu podwójny inwerter oraz podwójny anti-inwerter wymaga zastosowania prostego modułu inwersji (SI) dołączonego do komparatora, co zwiększa liczbę wykorzystanych tranzystorów. Do syntezy układów logicznych bardzo atrakcyjne wydają się wyjścia inwertera i anti-inwertera zbudowane odpowiednio z 3 i 2 tranzystorów ze względu na najmniejszą liczbę tranzystorów występujących w bramce. W układach syntezy z bramek prądowych występują trzy podstawowe stany logiczne: „-1”, „0”, „1”, oraz ich wielokrotności. Z tego względu „logika prądowa” różni się od typowej logiki Boolowskiej.

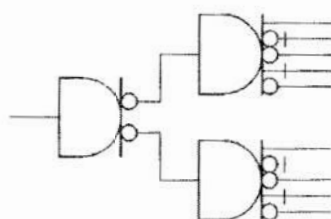
### 2.1. Elementy algebry prądowej

Logika prądowa pozwala na prostą realizację sumy arytmetycznej bez ponoszenia żadnego nakładu sprzętowego. Aby ją otrzymać, wystarczy połączyć ze sobą wyjścia bramek, prąd wypływający z takiego węzła będzie sumą prądów doprowadzonych (rys. 3). Ta cecha układów prądowych wydaje się bardzo atrakcyjna do syntezy układów cyfrowych gdyż realizacja funkcji logicznych odbywa się przy niewielkim nakładzie sprzętowym [5, 6, 7].



Rys. 3. Realizacja operacji dodawania i odejmowania arytmetycznego w układach prądowych

Niestety, aby zrealizować jednocześnie kilka funkcji z udziałem sygnału prądowego  $x$ , należy wcześniej go powielić. Jak już wspomniano, w systemach CMCL wielkością fizyczną niosącą informację o stanie logicznym jest prąd i nie można powielić sygnału cyfrowego poprzez rozgałęzienie ścieżki. Spowodowałoby to rozdzielenie prądu o wartości zdeterminowanej stanem logicznym na części, z których żadna nie niosła by właściwej informacji. Uzyskanie powielenia odbywa się poprzez zastosowanie powielającej bramki wielowyjściowej, co powoduje wzrost nakładu sprzętowego.



Rys. 4. Powielanie sygnału – kaskadowe połączenie bramek prądowych

Opisane powielanie logicznych sygnałów prądowych za pomocą bramek powielających wydają się być głównym powodem dużej złożoności sprzętowej i nadmiernego poboru mocy ze źródła zasilania w systemach CMCL [8, 9].

Należy nadmienić, że na skutek sumowania, oprócz podstawowych stanów logicznych: „-1”, „0”, „1”, w logice prądowej występują także wielokrotności tych stanów. Na wejściu bramki może pojawić się prąd o wartości ze zbioru  $(-n, \dots, 0, \dots, n)$  [10]. Daje to możliwość realizacji systemów pracujących w logice wielowartościowej. Aksjomaty i tożsamości algebry bramek prądowych zostały dokładnie opisane w [6, 7, 11, 12, 13, 14].

## 2.2. Pobór mocy w cyfrowych układach prądowych

Wspomniano, że bramki prądowe pobierają stałą moc ze źródła zasilania. Dzieje się tak dlatego, że prąd jedynki pobierany jest przez pewne obwody w bramce (konkretnie przez prosty układ inwersji **SI** i układ wyjścia typu inwerter **I**) nicustannie, zmienia się jedynie ścieżka przepływu prądu przy przełączaniu klucza sterującego źródłem  $I_s$ . Najmniej korzystną bramką prądową pod kątem konsumpcji mocy jest bramka podwójnego inwertera gdyż zbudowana jest z układu **SI** i modułu wyjściowego **I**, które pobierają prąd odpowiadający logicznej jedynce. Pobór prądu ze źródła zasilania dla wszystkich bramek z jednym wyjściem przedstawiono w tab. 1.

W cyfrowych układach prądowych występuje problem dużego poboru mocy. Całkowity pobór mocy systemu CMCL zależy od sumy układów **SI** i sumy wyjść typu inwerter. Redukcję poboru mocy można uzyskać zmniejszając wartość prądu jedynki, jednakże powoduje to zwiększenie czasu propagacji bramki, który zależy od szybkości przeładowania pasożytniczych pojemności bramki tranzystora (klucza) sterowanego komparatorem. Pojemności te przeładowywane są w czasie przełączania stanu logicznego co najwyżej prądem jedynki. Obniżenie wartości tego prądu redukuje jednocześnie szybkość pracy układu cyfrowego. W praktyce więc redukcji poboru mocy nie da się uzyskać przez zmniejszenie prądu jedynki, należy raczej dążyć do zwiększania tej wartości ze względu na szybkość pracy układu. Znaczące rozwiązanie problemu poboru mocy przedstawione zostanie w kolejnych rozdziałach, przy omawianiu najnowszych wyników związanych z bramkami 4. generacji.

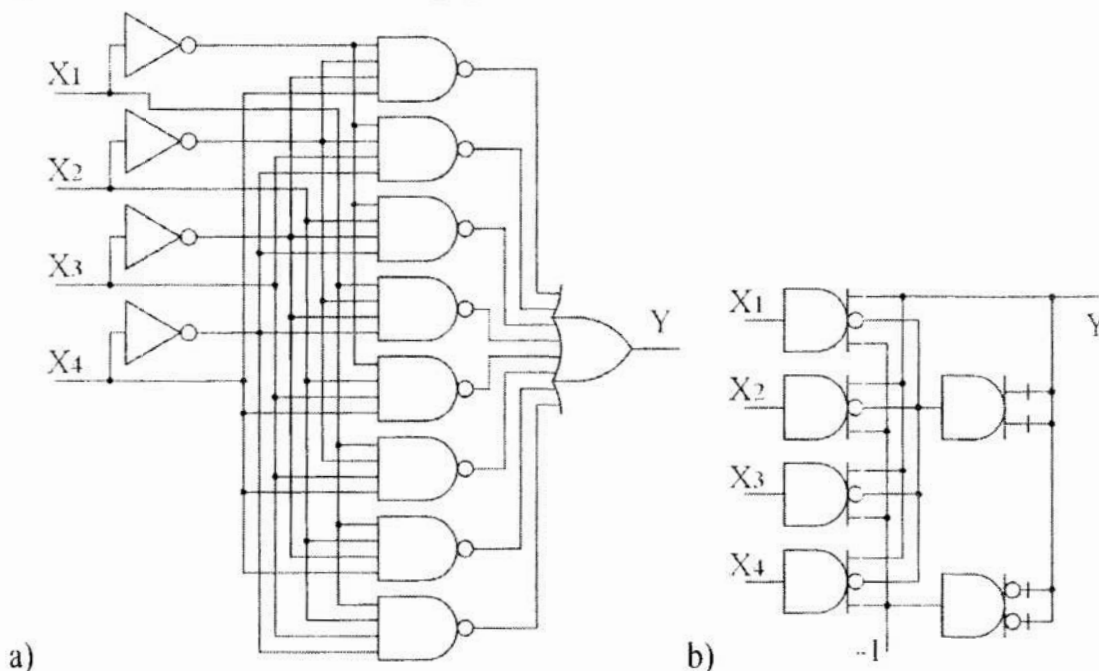
Rozważania na temat mocy pobieranej przez bramki prądowe należy zakończyć przedstawieniem pewnego aspektu związanego ze stałym prądem pobieranym ze źródła. O ile w układzie prądowym nie występują skokowe zmiany mocy przy przełączaniu stanu logicznego, to jednocześnie nie zachodzą znaczące zmiany poziomu pola elektromagnetycznego, generowanego przez układ cyfrowy, co zostanie szerzej przedstawione w rozdziale 3. przy wykorzystaniu CMCL dla celów bezpieczeństwa emisji elektromagnetycznej systemów przetwarzających dane.

### 3. Potencjał bramek prądowych

Bramki prądowe mają dwie charakterystyczne cechy dające potencjalne możliwości wykorzystania ich w takich sytuacjach, gdzie zastosowanie klasycznych bramek stwarza poważne problemy lub jest zwyczajnie niemożliwe. Pierwszą z cech jest praca w trybie prądowym, z czego wynika praktycznie stały pobór mocy, drugą jest reprezentacja stanu logicznego za pomocą prądu. W niniejszym rozdziale przedstawiono potencjalne zastosowania CMCL wynikające z przedstawionych cech tych układów i płynące z tego korzyści [10].

#### 3.1. Prosta realizacja funkcji logicznych

W 2.1. pokazany został sposób realizacji funkcji logicznych w algebrze bramek prądowych. Istotą jest możliwość zrealizowania sumy arytmetycznej prądów poprzez połączenie wyjść bramek, czyli bez udziału funktora logicznego. Wiąże się z tym redukcja nakładu sprzętowego w układzie cyfrowym. Jednocześnie przykład pokazuje, że funkcje logiczne w algebrze bramek prądowych realizowane są zupełnie inaczej niż w przypadku klasycznych bramek cyfrowych. Algebra „prądowa” jest inna od algebry Boolowskiej. Okazało się niezbędne wyprowadzenie podstawowych tożsamości oraz opracowanie metod minimalizacji [12, 13, 14].



**Rys. 5.** Przykład układu kombinacyjnego a) zbudowanego z klasycznych bramek b) realizujący tę samą funkcję logiczną układ z bramek prądowych

Opracowanie odpowiednich, automatycznych technik minimalizacji funkcji logicznych w algebrze bramek prądowych okazało się problemem niezwykle trudnym. Prace badawcze trwały kilka lat. W pierwszym etapie synteza układów prądowych była



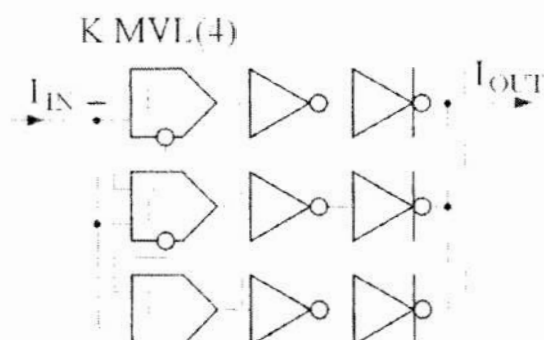
heurystyczna. Struktura systemu cyfrowego zależała od doświadczenia osoby projektującej układ. W efekcie wypracowano jednak metody pozwalające na uzyskanie bardzo prostych (w porównaniu z bramkami napięciowymi) struktur logicznych, realizujących te same funkcje [5, 6, 7, 10, 11]. Przedstawiony na rys. 5 przykład pokazuje, że stosowanie bramek cyfrowych pracujących w trybie prądowym umożliwia skorzystanie z rozszerzonych reguł charakterystycznych wyłącznie dla tych układów. Reguły te wykraczają poza stosowane w klasycznych układach cyfrowych, pracujących według logiki boolowskiej, prawa de'Morgana. Korzyścią jest możliwość wydajnej minimalizacji funkcji logicznych, a w rezultacie: prosta postać układów kombinacyjnych; szybko działające układy; redukcja nakładu sprzętowego oraz radykalne zmniejszenie liczby i całkowitej długości połączeń występujących w układzie cyfrowym.

Największy wkład w opracowanie technik minimalizacji wniósł śp. prof. dr hab. inż. Oleg Maslennikov, który przez wiele lat kierował zespołem badawczym, zajmującym się rozwojem idei prof. A. Guzińskiego. Prace prof. O. Maslennikowa zaowocowały istniejącymi dzisiaj technikami minimalizacji, pozwalającymi na syntezę układów CMCL, charakteryzującymi się minimalną strukturą.

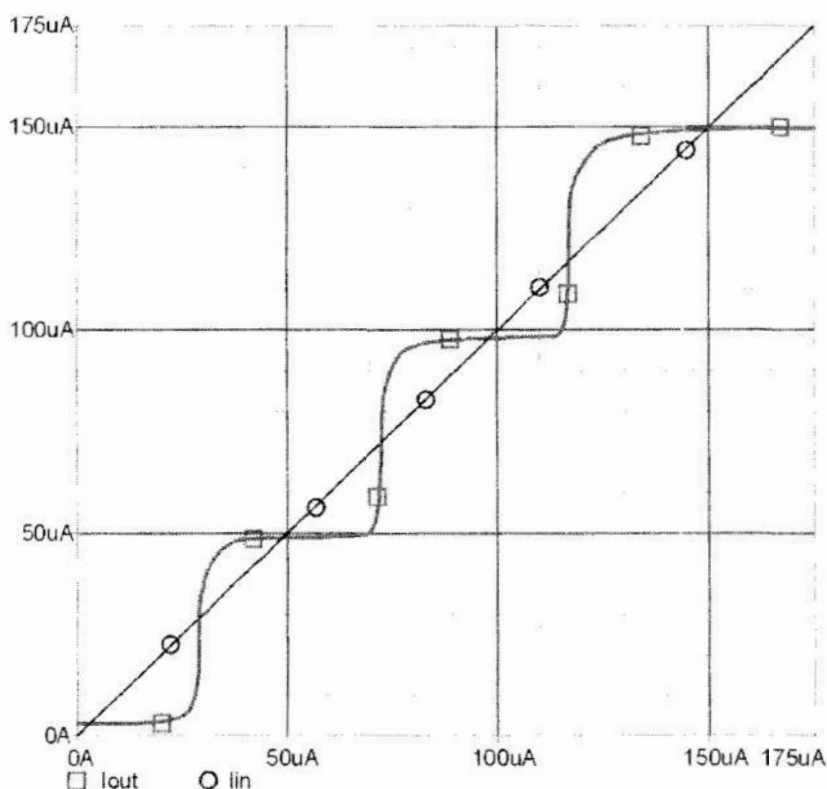
### 3.2. Budowa systemów pracujących w logice wielowartościowej

W praktyce układów cyfrowych spotyka się jedynie systemy binarne. Istnieją próby realizacji układów pracujących w tzw. logice „wielowartościowej”, jednakże występują tam tylko trzy poziomy: niski, wysoki i stan wysokiej impedancji (czyli brak konkretnego poziomu logicznego, stan wysokiej impedancji można traktować jak odłączenie bramki od układu). Wynika to stąd, że w przypadku bramek napięciowych nie można zrealizować prawdziwej logiki wielowartościowej, gdyż zazwyczaj poziom niski reprezentowany jest przez napięcie bliskie  $V_{SS}$ , a wysoki przez napięcie bliskie  $V_{DD}$ . Z tego wynika, że w przedziale  $V_{SS} .. V_{DD}$  (co odpowiada napięciu zasilania) nie ma miejsca na dodatkowe poziomy logiczne. Obniżenie poziomu napięcia związanego ze stanem „1” do poziomu odpowiadającego  $U_1=(V_{DD}-V_{SS})/(N-1)$ , gdzie  $N$  – podstawa systemu, spowodowałaby, po pierwsze, zmniejszenie stosunku sygnał/szum, a ponadto wymusiłaby redukcję tolerancji napięć odpowiadających za stany logiczne i spowodowałaby wzrost nakładu sprzętowego wynikający z konieczności dyskryminacji tych stanów.

W przypadku bramek prądowych implementacja MVL odbywa się w sposób naturalny. Dzięki temu, że poziom logiczny reprezentowany jest przez prąd wypływający z bramki, możliwe jest uzyskanie sumy arytmetycznej kilku prądów  $I_1$ . W CMCL jest to zjawisko powszechne, gdyż w taki właśnie sposób realizowane są funkcje logiczne. Do prawdziwej logiki wielowartościowej konieczne jest jedynie zrealizowanie komparatora MVL(N), który dokona interpretacji prądu wejściowego, reagując na wejściowy prąd o wartości  $n \cdot I_1$  wygenerowaniem  $n$  napięć sterujących  $U_Y$ . Dołączając do komparatora  $N-1$  połączonych modułów wyjściowych jednakowego typu uzyskuje się bramkę prądową MVL(N).



Rys. 6. Bramka prądowa MVL(4) typu podwójny inwerter



Rys. 7. Charakterystyka przejściowa bramki MVL(4) typu podwójny inwerter

W chwili obecnej, powstała czwarta generacja bramek prądowych, która umożliwia budowanie systemów CMCL pracujących w logice wielowartościowej z podstawą do  $N = 8$  włącznic, której poświęcono czwarty rozdział [14].

Zastosowanie MVL w systemach cyfrowych niesie ze sobą kolejną możliwość redukcji całkowitej długości połączeń występujących w układzie scalonym poprzez redukcję szerokości magistrali danych. Przykładowo, dla MVL(4) szerokość magistrali maleje dwukrotnie w stosunku do systemu binarnego. Wiadomo, że szybkość działania współczesnych układów cyfrowych nie zależy tak bardzo od czasu propagacji bramki, jak od opóźnień sygnału cyfrowego w torze logicznym. Decydujący wpływ na te opóźnienia ma długość połączeń występujących między bramkami cyfrowymi. Nie bez

znaczenia jest także powierzchnia zajmowana przez szyny sygnałowe w układzie scalonym. Logika wielowartościowa daje możliwość redukcji powierzchni układu scalonego poprzez redukcję całkowitej długości i ilości połączeń, co wpływa pośrednio na wzrost szybkości przetwarzania układu.

### 3.3. Kryptografia

Obecnie obserwuje się gwałtowny rozwój tzw. inżynierii zabezpieczeń – stosunkowo nowej dyscypliny naukowej, zajmującej się różnymi aspektami ochrony informacji. W ostatnich latach liczba aspektów bezpieczeństwa danych szybko wzrasta. Obecnie dotyczy, oprócz kryptografii i kontroli dostępu do informacji, m.in. problemów odporności sprzętu na penetrowanie fizyczne oraz zapewnienia bezpieczeństwa emisji elektromagnetycznej systemów komputerowych (Emsec – emission security), a w szczególności bezpieczeństwa emisji systemów kryptograficznych. Pojawia się problem ochrony systemów kryptograficznych przed atakami SCA (Side Channel Attack) polegającymi na tzw. „podśluchiwaniu”, tj. na analizie zmian poboru mocy (Power Analysis Attacks – PAA) lub pola elektromagnetycznego (ElectroMagnetic Attack – EMA).

Klasyczne cyfrowe bramki CMOS, które są powszechnie stosowane w systemach komputerowych, w tym kryptograficznych, prawie nie pobierają prądu ze źródła zasilania w stanach ustalonych (stan „0” lub „1”), natomiast pobierają od kilku do kilkunastu tysięcy razy większy prąd w trakcie zmiany stanu. Zmiana stanu logicznego wiąże się ze zmianą polaryzacji sieci podciągającej i ściągającej w bramce. W obszarze przejściowym obie sieci przewodzą, a tranzystory są w nasyceniu. Umożliwia to swobodny przepływ prądu od zasilania do masy, ograniczony jedynie chwilową rezystancją tranzystorów MOS (w nasyceniu). Te względnie duże impulsy prądu (tzw. szpilki prądowe) powodują powstanie charakterystycznych zakłóceń (szumu cyfrowego) na szynach zasilania systemu. Można wykazać, że różne rozkazy mikroprocesora mają różne profile poboru mocy. Atakujący, analizując szum (fluktuacje napięcia i/lub prądu) na szynie zasilania mikroprocesora, może określić, jaka instrukcja jest wykonywana, a nawet na jakich danych! Okazuje się, że mając informacje o tym, jaki mikroprocesor działa w systemie kryptograficznym i jaki algorytm szyfrowania jest zaimplementowany, mierząc prąd pobierany przez mikroprocesor w trakcie szyfrowania, można zrekonstruować klucz szyfrujący (deszyfrujący). Do przeprowadzenia ataku PAA (pomiar, rejestracja i analiza prądu/napięcia zasilania) konieczny jest fizyczny dostęp do systemu, co nie zawsze jest możliwe. Istnieje bardziej wyrafinowany typ ataku – nie wymagający bezpośredniej ingerencji bądź kontaktu z systemem cyfrowym, polegający na analizie zmian emitowanego przez system pola elektromagnetycznego (EMA). W celu pomiaru tych zmian konieczne jest dysponowanie zaawansowanym sprzętem – specjalną anteną i odbiornikiem pomiarowym. Oba rodzaje ataków należą do jednej rodziny SCA, gdyż przyczyna zmian emisji pola jest ta sama – szpilki prądowe powstające w obwodach układu podczas przełączania bramek cyfrowych.

Rozwijane są różne metody skierowane na zapewnienie bezpieczeństwa emisji oraz przeciwdziałania atakom SCA. Metody te działają głównie na poziomie oprogramowania systemu (np. poprzez ujednoczenie czasu trwania operacji; wprowadzenie operacji nadmiarowych; itd.) lub na poziomie sprzętu (np. wprowadzenie generatorów szumu i/lub układów uśredniających pobór mocy; wprowadzenie osobnego, wewnętrznego zegara systemowego zmieniającego swoją częstotliwość, stosowanie układów samo-synchronizujących się, itd.). Innym sposobem jest ekranowanie urządzeń. Wymienione metody (choć niezwykle kosztowne) nie zabezpieczają systemów kryptograficznych przed „podstuchaniem”, a tylko utrudniają jego przeprowadzenie.

Innym kierunkiem zapewnienia bezpieczeństwa emisji oraz przeciwdziałania atakom SCA może być wykorzystanie w systemie bramek CMCL. Ponieważ w/w bramki pobierają prawie stały prąd ze źródła zasilania zarówno w stanach ustalonych, jak i podczas przełączania się, atak polegający na analizie zmian poboru mocy lub pola elektromagnetycznego układu podczas jego działania jest skazany na niepowodzenie w znacznie większym stopniu, niż w przypadku układów i systemów cyfrowych zbudowanych z klasycznych bramek CMOS. Wstępne badania opracowanego układu ASIC, zrealizowanego w technologii 0,35 [ $\mu\text{m}$ ], wykazały, że układ prądowy cechuje się znacznie mniejszym poziomem emisji elektromagnetycznej niż jego odpowiednik klasyczny. Ponadto, zapis dynamiki emisji układu CMCL (tzw. profil emisji) nie zawiera fragmentów charakterystycznych, pozwalających na wykonanie analizy różnicowej. Układy prądowe są odporne na różnicową analizę mocy oraz na analizę emisji elektromagnetycznej (ataki SCA). Pozwoli to na znaczące obniżenie kosztów produkcji systemów kryptograficznych, odpornych na ataki SCA.

### 3.4. Problemy w technologii CMCL

Oprócz wielu zalet i wynikających z nich potencjalnych możliwości układów CMCL istnieją też cechy niepożądane. W układach cyfrowych realizowanych w oparciu o bramki prądowe występują dwa podstawowe problemy: problem dużego poboru mocy i problem dużego czasu propagacji bramki. O poborze mocy pisano już w rozdziale 2.2. W rozdziale 2.1. wspomniano o konieczności powielania sygnałów prądowych, co powoduje wzrost nakładu sprzętowego. Wraz ze wzrostem złożoności układu dochodzi do wzrostu opóźnień w torze sygnału logicznego. Jak się okazuje oba występujące problemy mają wspólne źródło jakim jest złożoność układu cyfrowego. Częściowe rozwiązanie obu występujących problemów zostanie przedstawione w kolejnym rozdziale, który poświęcono kolejnej generacji bramek prądowych, w której uzyskano znaczącą redukcję nakładu sprzętowego, pobieranej mocy i opóźnienia w torze sygnału.

## 4. Bramki 4. generacji

Najnowsze badania związane z rozwojem koncepcji prof. A. Guzińskiego dotyczącej bramek prądowych, przeprowadzone w latach 2009-2011 zaowocowały powstaniem



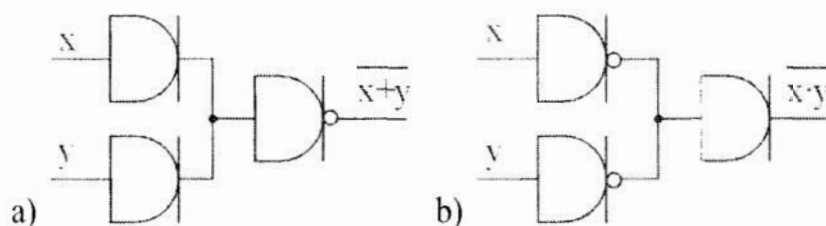
bramek 4. generacji. Wprowadzono dwie znaczące modyfikacje w stosunku do poprzednich realizacji bramek.

Modyfikacja pierwsza – liczba wejść: bramki prądowe do 3. generacji włącznie posiadały jedno wejście, mogły natomiast mieć więcej niż jedno wyjście. W czasie badań nad minimalizacją funkcji logicznych w algebrze prądowej zauważono możliwość redukcji nakładu sprzętowego poprzez zaprojektowanie bramki dwuwejściowej dla realizacji jednej z funkcji charakterystycznych (funkcja wzorcowa typu T), opisujących układ kombinacyjny [9]. Zastosowanie bramek wielowejściowych okazało się wyjątkowo eleganckie. Modyfikacja druga – tolerancja prądów: Ze względu na konieczność ograniczenia prądu wypływającego z bramki w stanie „0” (tzw. prąd szczytkowy  $I_R$ ) w module wyjściowym typu inwerter, w bramkach 3. generacji przyjęto zwiększoną wydajność źródła  $I_S$ , poprzez zwiększenie szerokości kanału tranzystora  $M_{11}$  (rys. 2). Takie działanie pozwoliło na polepszenie parametrów bramek dla systemów binarnych, ale uniemożliwiło zarazem budowanie bramek wielowartościowych. W czasie prac związanych z projektem badawczym, którego celem jest wykorzystanie bramek prądowych w kryptografii, wykorzystanie logiki wielowartościowej okazało się niezwykle przydatne. Dokonano więc zmniejszenia wydajności źródła  $I_S$  do wartości dokładnie  $I_1$ , co poskutkowało wzrostem prądu szczytkowego. Zmodyfikowano zatem moduł wyjściowy typu inwerter poprzez implementację klucza przełączającego n-p w celu redukcji prądu szczytkowego, co zmieniło nieco ideę działania bramki. Uzyskano w ten sposób znaczącą redukcję prądu szczytkowego oraz możliwość budowy układów w logice wielowartościowej [14].

#### 4.1. Bramki wielowejściowe [9]

Przy minimalizacji funkcji logicznych dla CMCL, chętnie wykorzystywaną bramką jest bramka typu anty-inwerter ze względu na najmniejszą liczbę tranzystorów. Zastosowanie anty-inwertera powoduje częste występowanie stanu logicznego „-1” w systemie cyfrowym. Poziom „-1” jest specyficzny, gdyż traktowany jest, z punktu widzenia logiki prądowej, jak stan „0”. W związku z tym operacja logiczna „-1” + „1” powinna mieć wartość „1”, tymczasem suma algebraiczna prądów wynosi 0.

Jeżeli chociaż jeden z sygnałów może przyjmować wartość „-1”, to sumy logicznej nie da się zrealizować poprzez sumowanie prądów w węzle. Konieczne jest zastosowanie wówczas bramek typu podwójny inwerter (lub inwerter) na wejściach, w celu normalizacji sygnałów prądowych i sprowadzenia ich z zakresu  $(-n .. n)$  do jednego ze stanów („0”, „1”). Z tego samego powodu, na wyjściu dołącza się dodatkową bramkę, jak na rys. 8.

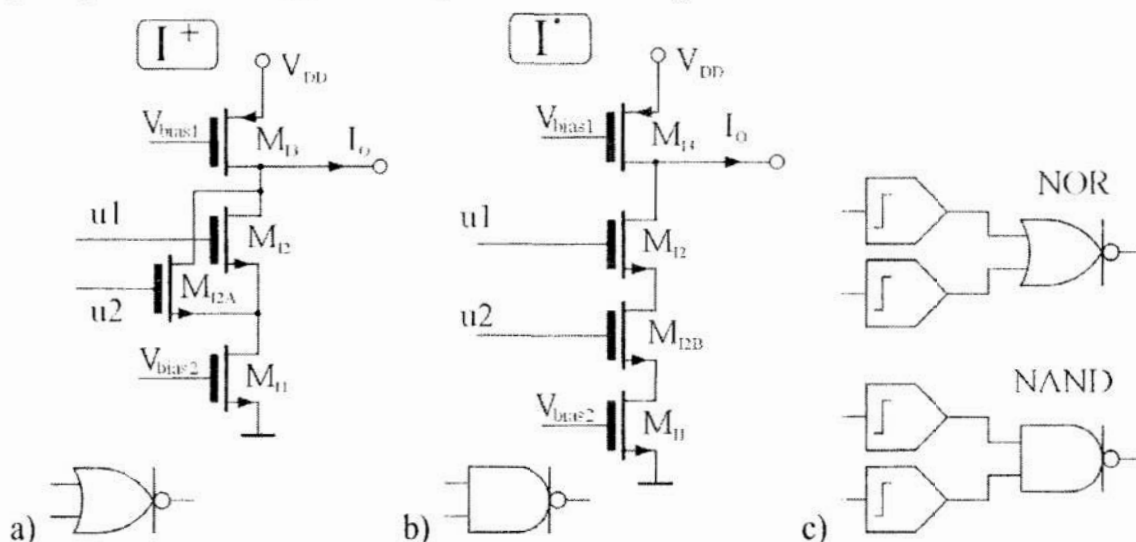


Rys. 8. Realizacja operacji logicznych a) NOR b) NAND na bramkach prądowych

Należy wyjaśnić, że bramki potrzebne są tylko na tych wejściach, na których może wystąpić stan „-n”. Okazuje się, że takie sytuacje występują w praktyce dość często [8, 9].

Struktura z rys. 8.a, realizująca prądową funkcję NOR, zbudowana jest z trzech bramek. Do jej realizacji sprzętowej w technologii CMOS, wykorzystano 27 tranzystorów, a pobór prądu wynosi  $5 \cdot I_1$ . Realizacja funkcji NAND wymaga wykorzystania 24 tranzystorów przy poborze prądu  $4 \cdot I_1$ . Możliwa jest redukcja zarówno ilości tranzystorów, jak i mocy, poprzez zastosowanie dwuwejściowego modułu inwertera.

Do modułu inwertera (I) można dołączyć dodatkowy tranzystor–klucz. Stosuje się równoległe bądź szeregowe połączenie kluczy, oznaczonych na rys. 9 jako  $M_{12}$  i  $M_{12A/B}$ . Uzyskuje się w ten sposób moduł wyjściowy typu inwerter, sterowany sygnałami napięciowymi  $u_1$  i  $u_2$ , który nie realizuje już jedynie negacji (ale sumę lub iloczyn logiczny), gdyż stan wyjścia zależy od stanu obu wejść.



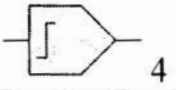
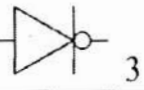
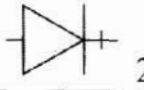
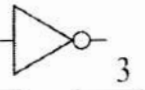
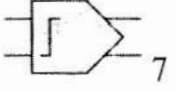
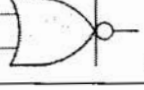
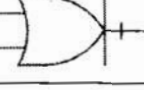
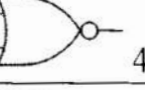
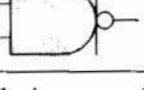
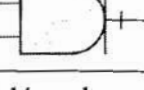
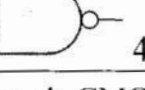
Rys. 9. Dwuwejściowe moduły inwertera i ich symbole a) z kluczami równoległymi b) z kluczami połączonymi szeregowo c) nowa realizacja NOR i NAND

Dołączając komparatory do wejść tak zbudowanych modułów, buduje się dwuwejściowe bramki prądowe realizujące operacje NOR i NAND w algebrze prądowej. Do budowy każdej z nich wykorzystuje się 12 tranzystorów, a prąd pobierany ze źródła jest równy  $I_1$ . W porównaniu z układami z rys. 8, realizującymi te same funkcje logiczne, następuje

**dwukrotna** redukcja złożoności sprzętowej i **czterokrotna** redukcja pobieranego prądu. Wykorzystanie wielowejsściowego modułu inwertera do konstrukcji bramek wielowejsściowych daje znaczący zysk związany z mniejszym nakładem sprzętowym i redukcją mocy pobieranej przez prądowy system cyfrowy.

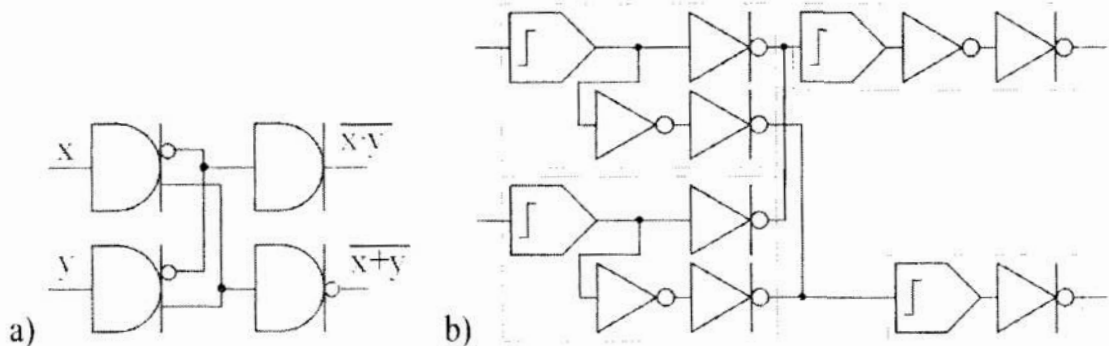
Podobnie można postąpić z modułem anti-inwertera (**AI**) i modułem inwersji napięciowej (**SI**), uzyskując możliwość tworzenia całej gamy nowych, wielowejsściowych bramek prądowych, obok czterech, do tej pory stosowanych bramek realizujących jedynie negacje [9].

**Tab. 2.** Wykorzystywane dotychczas i nowo opracowane moduły do syntezy bramek prądowych

	komparator	inwerter	anti-inwerter	inw. napięcia
znane moduły	 4	 3	 2	 3
nowe moduły	 7	 4	 3	 4
		 4	 3	 4

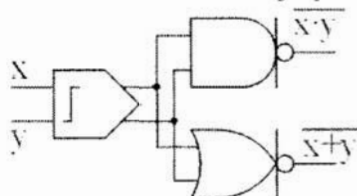
W tab. 2 podano oznaczenia znanych i nowych modułów do syntezy bramek CMCL. Kształt symbolu wiąże się z operacją logiczną, wykonywaną na sygnałach wejściowych, symbol na wyjściu modułów inwertera i anti-inwertera oznacza realizowaną negację prądową. Obok symboli zamieszczono liczby mówiące o ilości tranzystorów.

Jako ilustrację zysku jaki niesie zastosowanie bramek wielowejsściowych zaprezentowany zostanie prosty przykład jednoczesnej realizacji funkcji NOR i NAND na sygnałach  $x$  i  $y$ , pokazany na rys. 10. Należy zwrócić uwagę, że 2 bramki wejściowe służą nie tylko normalizacji wejściowych sygnałów  $x$  i  $y$ , ale także powieleniu tych sygnałów prądowych w celu jednoczesnego wykonywania wielu funkcji logicznych.



**Rys. 10.** Jednoczesna realizacja funkcji logicznych NOR i NAND a) na bramkach prądowych b) z pokazaniem wewnętrznej, modułowej struktury bramek

Do jednoczesnej realizacji funkcji NAND i NOR wykorzystano 4 bramki prądowe. Pobór mocy tego układu jest równy  $9 \cdot I_1$ , a na jego budowę składają się 43 tranzystory. Wykorzystanie dwuwejściowego modułu inwertera i dwuwejściowego komparatora pozwala na zbudowanie struktury pokazanej na rys. 11. Układ składa się z 15 tranzystorów, a prąd pobierany ze źródła wynosi  $2 \cdot I_1$ . Nastąpiła **trzykrotna** redukcja złożoności sprzętowej i ponad **czterokrotna** redukcja pobieranej mocy.



Rys. 11. Dwuwejściowa bramka prądowa NOR/NAND z wykorzystaniem modułów wielowejściowych (budowa modułowa, por. rys. 10.b)

Powstanie wielowejściowych modułów bramki prądowej pozwoliło na utworzenie całego zestawu nowych bramek dwuwejściowych (wielowejściowych), które uzupełniają zbiór znanych bramek prądowych, realizujących jedynie negacje. Daje to większą swobodę przy projektowaniu systemów CMCL. Dzięki nowym modułom możliwa jest realizacja funkcji logicznych OR, NOR, AND ... przy dużo mniejszym nakładzie sprzętowym, dzięki możliwości pominięcia powielania i normalizacji sygnałów prądowych. Uzyskuje się także redukcję mocy pobieranej przez system cyfrowy.

#### 4.2. Zastosowanie klucza n-p w module wyjściowym typu inwerter – koncepcja przełączania prądu w obwodzie bramki prądowej [14]

Wspomniano wcześniej, że bramki prądowe dają potencjalną możliwość realizacji logiki MVL. W praktycznej implementacji układów MVL napotyka się trudności wynikające z tolerancji poziomów prądów odpowiadającym stanom logicznym. W układach binarnych ich wpływ na poprawne działanie układu jest znikomy. Budowa układów MVL wymaga jednak, aby tolerancje te były możliwie małe, gdyż sumowanie prądów łączy się z sumowaniem błędów. Realizacja bramki MVL wymaga połączenia wyjść N-1 modułów inwertera.

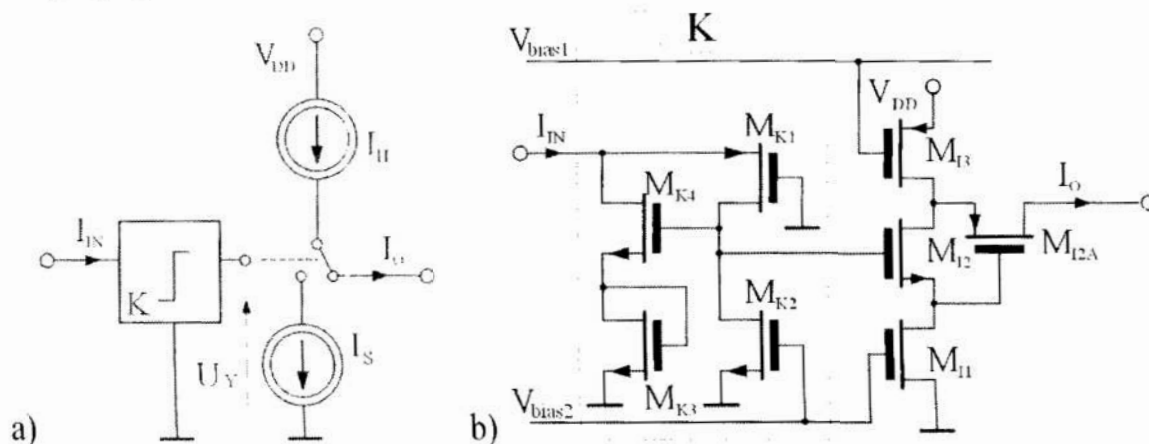
Niestety na wyjściu prądowego inwertera nawet w stanie logicznego zera występuje szczałkowa wartość prądu  $I_R$  (residual current). Cecha ta nie pozwala na łączenie dowolnej ilości wyjść bramek prądowych. Suma tych prądów w węzle może przekroczyć wartość graniczną  $\frac{1}{2}$  prądu jedynek logicznej  $I_1$  i zostanie „zinterpretowana” jako „1” przez układ wejściowy kolejnej bramki (komparator).

W praktycznej realizacji układów binarnych za pomocą bramek 3. generacji, ze względu na konieczność zapewnienia poprawnego działania prądowych układów cyfrowych zwiększono o ok. 6% wydajność źródła  $I_S$  poprzez zwiększenie szerokości kanału tranzystora  $M_{11}$ . Spowodowało to redukcję  $I_R$ , jednakże spowodowało, że każdy inwerter w stanie „0” odbiera  $6\% I_1$  od połączonego z nim inwertera w stanie „1”. Prąd

wypadkowy jest zatem pomniejszony o nadwyżkę wydajności źródła  $I_S$ . Z tego powodu przyjęto ograniczenie w ilości jednocześnie połączonych wyjść bramek do 8. Wprowadzone zmiany i ograniczenia w praktyce uniemożliwiają budowanie układów pracujących w systemach wielowartościowych.

W wyniku badań symulacyjnych określono, że prąd szczytkowy inwertera 3. generacji wynosi ok. 7% prądu  $I_1$ . Dla przyjętego poziomu  $I_1=50[\mu\text{A}]$ , wartość prądu szczytkowego  $I_R$  wyniosła ok.  $3,6[\mu\text{A}]$ . Przeprowadzono także badanie inwertera z ze źródłem  $I_S$  o takiej samej wydajności jak  $I_H$  (nie powiększonej) i zaobserwowano wówczas wzrost prądu  $I_R$  do ok. 11%  $I_1$ .

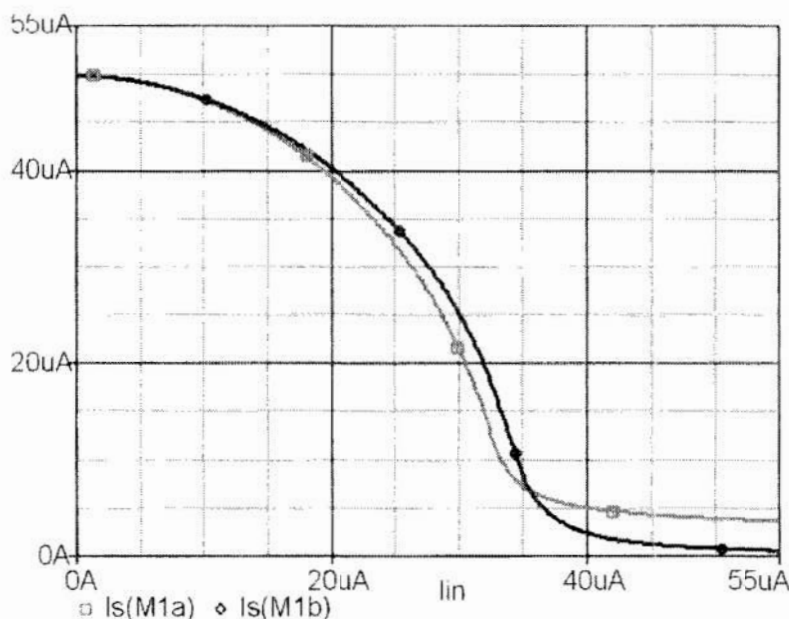
Wyniki przeprowadzonych badań sugerują, że do praktycznej realizacji układów MVL konieczna jest zmiana idei bramki prądowej. Zaproponowano przełączanie prądu ze źródła  $I_H$  między gałęzią zawierającą  $I_S$  a wyjściem bramki, jak na rys. 12.a. za pomocą klucza przełączającego n-p. Spowodowało to redukcję prądu szczytkowego  $I_R$  wpływającego z inwertera w stanie 0 i umożliwiło zastosowanie źródła  $I_S$  o dokładnie takiej wydajności jak  $I_H$  (nie powiększonej).



Rys. 12. a) Nowa koncepcja bramki prądowej – zastosowanie klucza przełączającego; b) jednoweściowa bramka typu inwerter z kluczem n-p (czwartej generacji)

Wykonano symulację porównawczą starej i nowej realizacji inwertera [15]. Rozbieżności kształtu charakterystyki wynikają z różnicy w wydajności źródła  $I_S$ , oraz z wartości prądu szczytkowego  $I_R$ .





**Rys. 13.** Charakterystyka przejściowa bramki typu inwerter trzeciej generacji (IsM1a) i czwartej generacji (IsM1b)

W bramce 4. generacji dokonano redukcji prądu szczytowego do poziomu  $I_R=0,6 [\mu A]$ , co stanowi zaledwie  $0,83\%I_1$ . Pozwala to na połączenie do 30 modułów wyjściowych. Stosując klucz **n-p** w miejsce klucza  $M_{12}$ , kosztem dodania jednego tranzystora, uzyskano ponad czterokrotne zwiększenie dopuszczalnej ilości wyjść bramek łączonych w węzle przy jednakowej wydajności źródeł  $I_S$  i  $I_{II}$ . W praktyce oznacza to, że można łączyć wyjścia nawet 4 bramek MVL o podstawie  $N = 8 \{4 \cdot 7 = 28 < 30\}$ . Dzięki temu umożliwiono praktyczną realizację bramek MVL(4) 4. generacji pokazano wcześniej na rys. 7.

## 5. Podsumowanie

W pracy przedstawiono genezę i rozwój idei budowy cyfrowych systemów prądowych, której autorem był śp. prof. dr hab. inż. Andrzej Guziński. Pokazano złożoność problemów, które napotyka praktyczna implementacja CMCL, jak choćby opracowanie od podstaw praw „algebry prądowej”, czy metod minimalizacji, niezbędnych do realizacji złożonych systemów cyfrowych. Trzeba nadmienić, że przed całym zespołem stoi trudne zadanie opracowania automatycznych metod minimalizacji właściwych dla bramek 4. generacji. Metody te będą inne od metod opracowanych wcześniej, gdyż należy uwzględnić możliwość wykorzystania bramek wielowjściowych. Zadanie dodatkowo utrudnia brak śp. prof. dr. hab. inż. O. Maslennikowa, który w zagadnieniach logiki i minimalizacji był i długo pozostanie niezastąpiony.

W rozdziałach 2.2 i 3.4 pokazano problemy wynikające z konieczności powielania sygnałów prądowych, których skutkiem jest stosunkowo duży pobór mocy i czas



propagacji układów. Przedstawiono częściowe rozwiązania niektórych problemów dzięki nowej koncepcji bramek 4. generacji.

Nie trudności stanowią jednak o potencjale bramek prądowych. To możliwości niespotykane w innych rodzinach układów cyfrowych. Zamysłem autora było przedstawienie atutów układów prądowych. Wynikiem pracy w **trybie prądowym** jest praktycznie stała moc pobierana przez bramki ze źródła zasilania, a co ważniejsze bardzo niski poziom zakłóceń występujących przy przełączaniu stanu logicznego, przenoszonych przez podłoże i szyny zasilające. W wyniku tego bramki CMCL:

- nadają się do budowy systemów cyfrowo–analogowych, w których nie jest konieczne wykonywanie drogich, dodatkowych procesów technologicznych służących ekranowaniu, separacji części analogowej i cyfrowej;
- nadają się do budowy jednostek przetwarzających w systemach przetwarzania danych zapewniając duże bezpieczeństwo emisji elektromagnetycznej (Emsec), wynikiem czego jest odporność na ataki SCA, przy braku konieczności ponoszenia dodatkowych kosztów na zabezpieczenia.

Ponieważ **poziom logiczny** reprezentowany jest poprzez prąd na wyjściu bramki, możliwa jest prosta realizacja funkcji logicznych w algebrze prądowej. Niesie to za sobą możliwość realizacji układów cyfrowych, których budowa jest mniej złożona niż ich odpowiedników wykonanych w innej technologii. Mniejsza liczba bramek skutkuje mniejszą liczbą połączeń oraz mniejszą sumaryczną długością połączeń w układzie scalonym. Bramki prądowe umożliwiają ponadto budowę systemów opartych o logikę wielowartościową o podstawie  $N > 2$ . Także w tym przypadku zysk wynika z redukcji ilości połączeń, co dodatkowo wpływa na całkowitą powierzchnię struktury krzemowej.

Przedstawione aspekty zastosowania bramek prądowych to przysłowiowy czubek góry lodowej. Mówiąc o potencjale CMCL w niniejszej pracy, autorowi udało się ledwie wspomnieć o kilku charakterystycznych cechach układów z tej rodziny. Możliwości ich są z pewnością większe, prawdopodobnie niebawem zaistnieją kolejne potrzeby (aspekt Emsec jeszcze niedawno nie istniał), gdzie możliwe będzie zastosowanie układów CMCL. Istotne jest, że bramki prądowe już dziś stanowią alternatywę dla klasycznych bramek CMOS. Zastosowanie ich w wielu przypadkach daje wymierne korzyści w postaci redukcji kosztów produkcji systemów cyfrowych i analogowo–cyfrowych lub umożliwia budowę systemów, których zrealizowanie z klasycznych bramek cyfrowych jest niemożliwe.

Rozdziały 3. i 4. zawierają wyniki prac wykonanych w ramach Projektu Badawczego **O N515086737**.

## Bibliografia

1. A. Guziński, A. Kielbasiński, *Current-Mode Digital Circuits Operating in Mixed Analog-Digital Systems*, Proc. of the XVIII-th National Conference on Circuit Theory and Electronic Networks, pp. 317-322, Kołobrzeg 1995.
2. G. Blakiewicz, *Porównanie właściwości układów tłumienia zakłóceń podłożowych*, VI Krajowa konferencja Elektroniki: materiały konferencji, Darłówko wschodnie, 11-13 czerwca 2007, T. ½, Darłówko wschodnie, 11-13 czerwca 2007, s. 149-154.
3. Guziński A., Pawłowski P. *Current-mode digital circuits for low-voltage mixed analog-digital systems*, Proc. 6-th Int. Conf. Mixed design of integrated circuits systems, MIXDES'99, Kraków, Poland, 1999, pp. 369-372.
4. P. Pawłowski, *Ocena przydatności bramek cyfrowych pracujących w trybie prądowym w mieszanych systemach analogowo-cyfrowych*, rozprawa doktorska, Politechnika Koszalińska Wydział Elektroniki, Koszalin 2004.
5. A. Guziński, P. Pawłowski, D. Czwyrow, J. Kaniewski, O. Maslennikow, N. Maslennikowa, D. Rataj, *Design of Digital Circuits with Current-Mode Gates*, Bulletin of the Polish Academy of Sciences, Technical Sciences, Vol. 48, No. 1, pp. 73-91, 2000.
6. O. Maslennikow, A. Guziński, J. Kaniewski, R. Berezowski, *Rules of Current-mode Digital Circuit Design and Analysis*, Proc. of the XXII Nat.Conf. on Circuit Theory and Electronic Networks - KKTOiUE, pp. 149-154, Stare Jabłonki 1999.
7. O. Maslennikow, *Minimalizacja funkcji logicznych w algebrze bramek prądowych*, Prace IV Krajowej Konferencji Elektroniki, KKE'2005, pp., Kołobrzeg 2005.
8. O. Maslennikow, N. Maslennikowa, P. Pawłowski, M. Rajewska, R. Berezowski, *Hardware realization of the modular Exponentiation operation in cryptographic systems based on binary and multivalued logic*, 16th Internation Conference Mixdes 2009, pp. 210-214, June 2009, Łódź.
9. Łuczak R., Rajewska M. *Wielowejściowe bramki prądowe*, Elektronika - Konstrukcje, Technologie, Zastosowania 2011/9, Sigma-NOT Warszawa, 2011.
10. M. Białko, O. Maslennikow, N. Maslennikowa, P. Pawłowski, *Układy cyfrowe zbudowane z bramek prądowych: stan obecny, perspektywy rozwoju i zastosowania*, Elektronika - konstrukcje, technologie, zastosowania, nr 12, str. 38-43, 2004 r.
11. O. Maslennikow, *Approaches to Designing and Examples of Digital Circuits Based on the Current-Mode Gates*, Data Recording, Storage & Processing, Vol.3, No.2, pp. 84-98, 2001.
12. O. Maslennikow, M. Rajewska, R. Berezowski, *Hardware Realization of the AES Algorithm S-Block Functions in the Current-Mode Gate Technology*, Proc. 9-th Int. Conf. Experience of Designing and Application of CAD Systems in Microelectronics, CADSM'2007, IEEE Catalog Number 07EX1594, pp. 211-217, Lwów-Polyana, 2007.

13. M. Rajewska, R. Berezowski, O. Maslennikov, *Realizacja S-bloków systemu kryptograficznego DES na bramkach prądowych*, Prace XII Konferencji Krajowej „Komputerowe wspomaganie badań naukowych”, KOWBAN’2005, pp. 121-126, Polanica-Zdrój.
14. R. Łuczak, *An application of a N-P switch for a construction of current-mode gates*, Proc. of 18-th Int. Conf. Mixed Design of Integrated Circuits and Systems, Mixdes 16–18 June 2011, Gliwice, pp. 243-247.
15. Weidong Liu, Xiaodong Jin, James Chen, Min-Chie Jeng, Zhihong Liu, Yuhua Cheng, Kai Chen, Mansun Chan, Kelvin Hui, Jianhui Huang, Robert Tu, Ping K. Ko and Chenming Hu *BSIM3v3.2.2 MOSFET Model, Users’ Manual*, Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94720, 1999.

## Streszczenie

W artykule przedstawiono ideę budowy systemów cyfrowych pracujących w trybie prądowym, które realizowane są w oparciu o bramki prądowe wymyślone przez prof. Andrzeja Guzińskiego. Pokazano ich potencjał i etapy rozwoju oraz dokonano zestawienia potencjalnych możliwości zastosowania bramek prądowych we współczesnych układach elektronicznych.

## Abstract

In present paper the idea of building of digital systems operating in current mode, which are implemented based on current gates invented by Prof. A. Guziński is presented. Their potential and stages of development was shown. A summary of possibilities of the application of current-mode gates in today's electronics was made.

**Robert Berezowski**

Katedra Inżynierii Komputerowej

Wydział Elektroniki i Informatyki

Politechnika Koszalińska

## **Minimalizacja opisów funkcji w algebrze bramek prądowych**

W artykule przedstawiono rozwój metod minimalizacji opisów funkcji w algebrze bramek prądowych jakie opracowano w przeciągu ostatnich dwudziestu lat. Technologia bramek prądowych związana jest z koncepcją bramki prądowej zaproponowanej przez prof. Andrzeja Guzińskiego w latach 90. ubiegłego wieku. Wskazane zostaną również kierunki dalszego rozwoju badań nad optymalizacją układów prądowych.

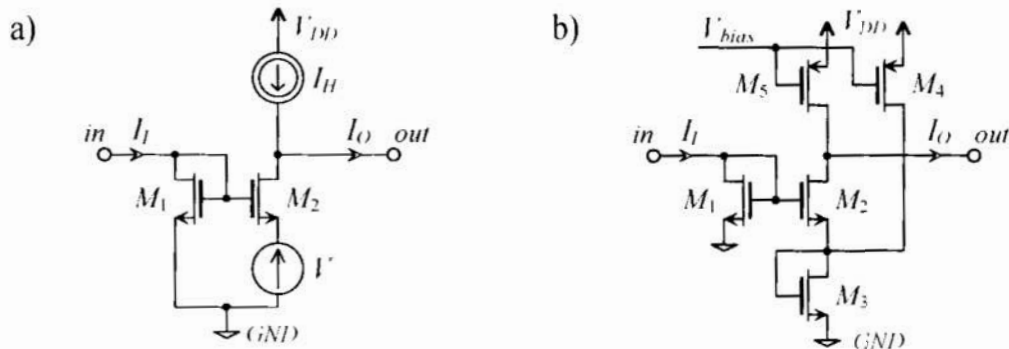
### **1. Wprowadzenie**

Na początku lat 90. ubiegłego wieku profesor Andrzej Guziński opracował koncepcję, a następnie zaimplementował w technologii CMOS pierwszą bramkę pracującą w trybie prądowym [1]. Koncepcja ta stała się pomysłem na nowy kierunek badań w powstającym w tym czasie Instytucie Elektroniki w Koszalinie. Grupa osób zaangażowanych w te badania systematycznie się powiększała, zdobywano nowe doświadczenia, powstawały nowe pomysły w kierunku układów prądowych. Badania prowadzono m.in. w ramach czterech projektów badawczych finansowanych przez Polski Komitet Badań Naukowych. Opracowano w tym czasie podstawy algebry bramek prądowych, metody minimalizacji opisów funkcji w tej algebrze, sposoby projektowania cyfrowych układów prądowych oraz metody ich symulacji i weryfikacji [2-7]. Część opracowanych układów została zrealizowana praktycznie w postaci układów scalonych ASIC [7]. Początkowo bramki prądowe wykorzystywano do implementacji funkcji binarnych (binarnych układów cyfrowych). Kolejne badania wykazały, że bramki prądowe mogą być z powodzeniem wykorzystane do konstruowania układów cyfrowych działających w logice wielowartościowej (MVL, ang. Multiple-Value Logic) z podstawą  $N > 2$  lub w arytmetyce modulo  $N$  [6].

### **2. Koncepcja. Rodzaje oraz cechy bramek prądowych**

Koncepcja opracowanej przez prof. Guzińskiego bramki pracującej w trybie prądowym zakłada, że na wejściu lub wyjściu bramki poziom logicznego 0 odpowiada prądowi o wartości równej zero, a prąd logicznego 1 odpowiada pewnej, ustalonej wartości prądu

większej od zera. Dla bramki z rys. 1, poziomowi logicznemu 1 na wejściu (tj. prądowi wpływającemu  $I_I$ ) odpowiada poziom 0 na wyjściu (prąd wyjściowy  $I_O = 0$ ) bramki i odwrotnie. Z tego powodu bramka ta otrzymała nazwę inwertera prądowego (bramka z wyjściem typu inwerter), a jej parametry elektryczne przedstawione są np. w pracy [1].



Rys. 1. Koncepcja bramki prądowej (a): przykładowa realizacja CMOS inwertera (b)

W kolejnych badaniach opracowano, przetestowano i zakwalifikowano do użycia dodatkowo trzy inne typy bramek: anty-inwerter, podwójny inwerter oraz podwójny anty-inwerter. Bramki z przedrostkiem anty- umożliwiają dodanie trzeciego stanu logicznego -1 (prąd wpływa do bramki). Funkcje jakie realizują przedstawione bramki prądowe różnią się od funkcji jakie realizują bramki klasyczne (napięciowe). Również algebra bramek prądowych nie jest tożsama z algebrą Boole'a. Przede wszystkim na wyjściach bramek prądowych mogą pojawić się 3 stany logiczne: 0, 1 i -1. Bramki prądowe posiadają tylko jedno wejście, mogą posiadać kilka wyjść różnego typu. Najprostszą operacją w algebrze bramek prądowych jest operacja arytmetycznej sumy (połączenie kilku wejść do wspólnego węzła), a to sprawia, że w technologii bramek prądowych łatwiej zrealizować operacje arytmetyczne (dodawanie, odejmowanie) niż logiczne (AND, OR, ...). W algebrze bramek prądowych mogą pojawić się stany logiczne z zakresu  $-n$  do  $n$  (jako wynik sumowania lub odejmowania prądów w węzłach) a nie tylko 0 i 1. Ta ostatnia cecha powoduje, że wykorzystanie narzędzi do minimalizacji, które doskonale sprawdzają się logice binarnej, w algebrze bramek prądowych nie dają najlepszych rezultatów. Zastosowanie algebry bramek prądowych sprawia, że trudno jest znaleźć (przystosować) istniejące narzędzia, które można użyć do symulacji i weryfikacji układów zbudowanych z bramek prądowych.

W dalszej części referatu przedstawione zostaną opracowane metody minimalizacji opisów funkcji w algebrze bramek prądowych na przestrzeni ostatnich 20 lat.

### 3. Metody minimalizacji funkcji w algebrze bramek prądowych

Metody minimalizacji funkcji w algebrze bramek prądowych rozwijały się równolegle z rozwojem technologii bramek prądowych. Pierwsze sposoby opracowane były dla prostych funkcji, których implementacja wymagała zastosowania kilku bramek (2-4



argumentowe funkcje logiczne). Projektowano m.in. sumatory, multipleksery, przerzutniki, rejestry. W dalszych badaniach wskazane było opracowanie metod minimalizacji w postaci graficznej. Forma taka ułatwiała zrozumienie natury algebry bramek prądowych, ich charakterystycznych cech. Realizowano funkcje, których implementacja wymagała zastosowania wielu bramek (różne funkcje wielowyjściowe, układ S-Bloków algorytmu DES), a realizowane funkcje były maksymalnie 6 argumentowe (takie funkcje można przedstawić w postaci diagramów Veitcha-Karnaugh). W kolejnych badaniach założono, że potrzebny jest algorytm minimalizacji, który można wykonać w postaci programu komputerowego. Potrzeba taka wynikała z nowych zastosowań bramek prądowych. Projektowano w tym czasie układy pracujące w logice wielowartościowej, układy kryptograficzne. Funkcje, których opis należało zminimalizować zawierały 6 i więcej argumentów. Nie można było przedstawić ich w postaci graficznej, umożliwiającej projektantowi określenie funkcji bazowych. Potrzebne było narzędzie, które umożliwiłoby minimalizację opisu dowolnej funkcji logicznej w algebrze bramek prądowych.

### 3.1. Konwersja wyrażeń

Pierwszy sposób minimalizacji układów prądowych oparto o wyrażenia konwersji (1) - (4), które szczegółowo opisano w [3]. Wyrażenia te umożliwiają realizację podstawowych funkcji logiki binarnej (AND, OR, ...) w algebrze bramek prądowych.

$$a \cdot b = a + b \quad (1) \qquad a \cdot b = a + b \quad (2)$$

$$a \vee b = a + b \quad (3) \qquad a \vee b = a + b \quad (4)$$

W powyższych wyrażeniach symbolom  $\cdot$ ,  $\vee$  odpowiadają funkcje AND, OR w algebrze Boole'a, zaś symbolom  $+$ ,  $\cdot$  odpowiadają funkcje dodawania arytmetycznego, funkcji inwersji i podwójnej inwersji w algebrze bramek prądowych.

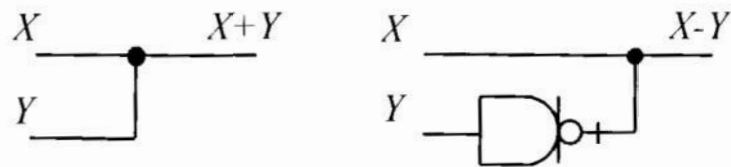
W pierwszym kroku zatem należało przedstawić opis funkcji w postaci wyrażenia boolowskiego. Następnie w tej samej logice należało dokonać minimalizacji opisu przy pomocy dowolnej znanej metody (diagramy Veitcha-Karnaugh, algorytm Quine'a-McCluskey'a, metoda Espresso). Posługując się wzorami konwersji można było otrzymać wynik w algebrze bramek prądowych. W kolejnych badaniach poprawiono tą metodę określając dodatkowe kryteria, które należy przyjąć przy minimalizacji w algebrze Boole'a aby po konwersji otrzymać lepszy rezultat. Sposób ten sprawdzał się przede wszystkim dla średnio skomplikowanych funkcji, znajomość przez projektanta wiedzy o algebrze bramek prądowych nie była wymagana. Nadawał się do minimalizacji dowolnej funkcji binarnej. Liczba bramek prądowych potrzebna do implementacji funkcji była jednakowa jak w przypadku zastosowania klasycznych bramek napięciowych. Należy jednak pamiętać, że bramki prądowe są zbudowane z większej liczby tranzystorów. Złożoność obliczeniowa jest równa złożoności obliczeniowej



wykorzystanej w nim znanej metody minimalizacji. Metoda ta nadawała się do realizacji komputerowej. Jednak nie gwarantowała otrzymania optymalnego rozwiązania.

### 3.2. Funkcje bazowe

Doświadczenie zespołu zdobyte po zaprojektowaniu wielu prostych układów prądowych doprowadziła do opracowania nowego sposobu minimalizacji opisów funkcji logicznych. Sposób ten przeznaczony był do wykorzystania bezpośrednio w algebrze bramek prądowych i oparty był o twierdzenie [3], że dowolna funkcja logiczna może być przedstawiona w algebrze bramek prądowych jako suma algebraiczna kilku innych (prostszych) funkcji logicznych nazywanych funkcjami bazowymi. Metoda ta dawała lepsze rezultaty od poprzedniej, jednak jej skuteczność zależna była od doświadczenia i spostrzegawczości osoby minimalizującej opis funkcji (metoda heurystyczna).



**Rys. 2.** Realizacja operacji arytmetycznych w technologii bramek prądowych

Najważniejszy był bowiem podział całej funkcji na niezależne funkcje bazowe. Metoda ta wykorzystywała operacje arytmetycznego dodawania i odejmowania (najprościej realizowane operacje w algebrze bramek prądowych – rys. 2). Niezależne funkcje bazowe charakteryzowały się brakiem wspólnych jedynek w diagramie Veitcha-Karnaugh. Sposób ten sprawdzał się w przypadku równoczesnej minimalizacji opisów różnych funkcji mających wspólne argumenty. W takim przypadku szukano prostych funkcji, które minimalizowano za stosując wyrażenia konwersji, natomiast pozostałe funkcje wykorzystywały je jako funkcje bazowe. Metoda ta również nie pozwalała uzyskać optymalnych wyników. Metoda ta nadawała się szczególnie do minimalizacji funkcji wielowartościowych. Niestety metoda ta nie nadawała się do realizacji komputerowej.

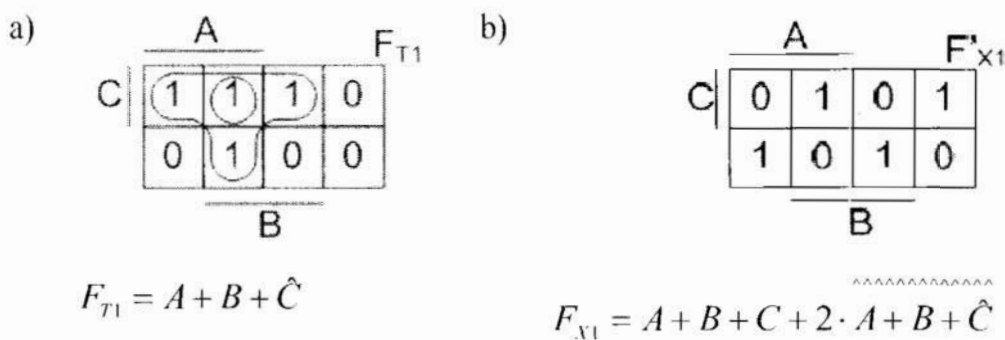
### 3.3. Algorytm ewolucyjny

W metodzie przedstawionej w pracy [11] zastosowano algorytm ewolucyjny z wielowarstwowym chromosomem do optymalizacji kombinacyjnych układów cyfrowych budowanych przy użyciu bramek prądowych. Optymalizacja polegała na minimalizacji liczby tranzystorów wchodzących w skład danego układu. Struktura optymalizowanego układu otrzymywana była podczas projektowania "ręcznego". Metoda ta posiadała jednak ograniczenie związane z liczbą wejść układu. Ograniczenie to jest związane z tzw. eksplozją kombinatoryczną. Przy użyciu metody [11] mogą być

optymalizowane układy o niewielkich rozmiarach (liczba wejść do 10). Jednak wada ta dotyczy wszystkich metod ewolucyjnego projektowania układów cyfrowych.

### 3.4. Funkcje wzorcowe

Kolejne badania (m.in. z użyciem metody z zastosowaniem algorytmu ewolucyjnego) zauważono, że pewne układy jedynek w diagramie Veitcha-Karnaugh, trudne do minimalizacji (prostego opisu) w algebrze Boole'a, w algebrze bramek prądowych pozwalają otrzymać wyrażenia proste (nazwano je funkcjami wzorcowymi). Opisano dwie charakterystyczne funkcje wzorcowe: typu XOR oraz typu T. Dla trzech argumentów funkcje te opisują funkcję sumy (typu XOR) oraz przeniesienia (typu T) w sumatorze jednobitowym (rys. 3).



Rys. 3. Diagram Veitcha i opis funkcji wzorcowej typu a) T b) XOR w algebrze prądowej

Uogólniono opis oraz opracowano algorytmy wyszukiwania tych funkcji wzorcowych (bazowych) w dowolnej minimalizowanej funkcji [9]. Jednocześnie zaproponowano kolejną, heurystyczną, metodę minimalizacji, która pozwala wyszukiwać funkcje bazowe, a dzięki temu znaleźć prostszy opis funkcji. Metoda ta w dużo mniejszym stopniu zależy od doświadczenia osoby, która dokonuje minimalizacji, jednak końcowy wynik nadal zależy od spostrzegawczości projektanta.

Aby przyspieszyć proces minimalizacji w kolejnych etapach określono kolejność wyboru funkcji bazowych spośród wszystkich znalezionych tak by ich liczba była najmniejsza. W pierwszej kolejności wybierano te funkcje bazowe, które zawierały najwięcej niepokrytych do danego momentu implikantów pierwotnych. Opracowano wyrażenia, które obliczały ilość implikantów pierwotnych składających się na określoną funkcję wzorcową. Opracowana metoda nadawała się do realizacji w postaci programu komputerowego.

### 3.5. Zmodyfikowany algorytm Quine'a-McCluskey'a

Opracowanie algorytmów wyszukiwania funkcji bazowych oraz metoda minimalizacji w oparciu o te funkcje zwróciła uwagę zespołu na jej podobieństwo do algorytmu minimalizacji Quine'a-McCluskey'a. Algorytm ten składa się z dwóch etapów: wyszukania implikantów prostych oraz określenia minimalnego pokrycia. Dla funkcji

boolowskiej algorytm ten potrafi znaleźć minimalny opis funkcji. Zaproponowano modyfikację algorytmu Quine'a-McCluskey'a przeznaczonego do minimalizacji opisów funkcji w algebrze bramek prądowych [12]. W pierwszym etapie, oprócz wyszukiwania implikantów prostych, wyszukiwane są funkcje wzorcowe typu T (badania wykazały, że stosowana obecnie realizacja sprzętowa funkcji wzorcowej typu XOR nie pozwala uzyskać opłacalnego zysku, zatem pominięto ją). W drugim etapie wybierane jest najmniejsze pokrycie funkcji spośród wszystkich wyszukanych implikantów prostych i funkcji wzorcowych. Ze względu na różnice w ilości bramek potrzebnych do realizacji implikantu prostego (1 bramka) i funkcji wzorcowej typu T (1 lub 2 bramki w zależności od realizowanej funkcji) etap ten należało odpowiednio zmodyfikować. Tabelę należy utworzyć w taki sposób aby od lewej strony występowały funkcje bazowe zawierające taką samą „wagę” jedynek. Wagę tworzy się obliczając liczbę jedynek, z których składa się funkcja bazowa. Dodatkowo dla funkcji wzorcowych typu T, których realizacja wymaga zastosowania dwóch bramek otrzymana liczba mnożona jest przez 2. Złożoność obliczeniowa proponowanej metody jest zbliżona do metody Quine'a-McCluskey'a. W pierwszym etapie wzrost złożoności polega na dodatkowym utworzeniu (nie jest to wyszukiwanie) funkcji wzorcowych typu T, których maksymalnie może być tyle ile jest implikantów pierwotnych. Drugi etap, utworzenie tabeli (zawierającej implikanty pierwotne oraz funkcje bazowe) oraz wybór najmniejszego pokrycia, z reguły bardziej złożony można zrealizować na dwa sposoby. Umieszczając w nim wszystkie znalezione implikanty proste i funkcje wzorcowe typu T (znaczny wzrost złożoności obliczeniowej) lub upraszczając, można usunąć implikanty proste wchodzące w skład funkcji wzorcowych typu T. W tym drugim sposobie złożoność obliczeniowa będzie podobna do złożoności obliczeniowej algorytmu Quine'a-McCluskey'a. Kolejne badania wykazały, że zastosowana metoda niestety również nie pozwala uzyskać optymalnego opisu funkcji. Wprowadzono trzeci etap, który sprawdza, czy któraś z wybranych funkcji wzorcowych typu T może zostać zredukowana do implikantu prostego (doświadczalnie stwierdzono, że taka możliwość istnieje). Ostatnia opisana metoda wraz z poprawką (trzeci etap) jest aktualnie najlepszą metodą minimalizacji opisu funkcji w algebrze bramek prądowych. Metoda ta nadaje się do realizacji komputerowej i program taki został przez zespół zrealizowany.

#### 4. Podsumowanie

Implementacja prostych układów w technologii bramek prądowych nie wymaga dodatkowych narzędzi ani metod projektowania. Podstawowe układy można narysować oraz przetestować na zwykłej kartce. Kolejnym etapem jest projektowanie większych układów, realizujących bardziej skomplikowane funkcje logiczne. W zespole opracowano m.in. układy mnożące działające w arytmetyce resztowej, konwertery z systemu resztowego do systemu binarnego i odwrotnie, algorytmy kryptograficzne. Projektowanie większych układów wiąże się z opracowaniem dodatkowych narzędzi oraz metod

minimalizacji opisów funkcji. W referacie pokazano krok po kroku rozwój metod minimalizacji opisów funkcji w algebrze bramek prądowych. Począwszy od ręcznych, poprzez automatyzację, algorytmy ewolucyjne, po metody komputerowe.

Stosowane obecnie bramki prądowe posiadają budowę modułową (wyróżnić można 4 różne moduły). Kolejne metody minimalizacji powinny tworzyć prostsze opisy funkcji, operując na funkcjach realizowanych przez te moduły, a nie na bramkach. Jako jeden z etapów rozwoju technologii bramek prądowych prof. A. Guziński zaproponował bramkę wielowejściową [13], której implementacja jest obecnie na ukończeniu. Wprowadzenie nowych bramek oznacza wprowadzenie kolejnych zmian do sposobów minimalizacji opisu funkcji lub wyszukiwania innych funkcji bazowych. Pojawienie się wielu stanów logicznych daje większe możliwości uzyskania prostszych opisów, lecz jednocześnie powoduje, że algorytmy do minimalizacji stają się coraz bardziej skomplikowane.

Wykonano dla Proj. Bad. O N515086737 – Zastosowanie układów cyfrowych pracujących w trybie prądowym w jednostkach przetwarzających systemów kryptograficznych.

## Bibliografia

1. Guziński A., Kielbasinski A., *Current-Mode Digital Circuits Operating in Mixed Analog-Digital Systems* Bulletin of the Polish Academy of Sciences, Technical Sciences, vol. 44, No. 2, pp. 193-198, 1996.
2. Gretkowski D., Guziński A., Kaniewski J., Maslennikow O., *VHDL models of digital combinatorical circuits on the current-mode gates*, Proc. 6-th Int. Conf. MIXDES'99, pp. 253-258, Kraków, Poland, 1999.
3. Maslennikow O., *Approaches to Designing and Examples of Digital Circuits Based on the Current-Mode Gates* Data Recording, Storage & Processing, vol. 3, No. 2, pp. 84-98, 2001.
4. Maslennikow O., Pawlowski P., Sołtan P., Berezowski R., *Current-Mode Digital Gates and Circuits: Conception, Design and Verification* Proc. IEEE Int. Conf. ICECS'2002, pp. 623-626, Dubrovnik 2002.
5. Maslennikow O., Berezowski R., Sołtan P., Rajewska M., *Designing Prototype of the Spartan II FPGA Slice with the Current-Mode Gates* Proc. IEEE Int. Conf. ICCSC'2002, pp. 182-185, St.-Petersburg 2002.
6. Maslennikow O., Gretkowski D., Pawlowski P., *Current-Mode Circuits for Multiple-Valued Logic and Residue Number System Arithmetic* Proc. 10-th Int. Conf. MIXDES'2003, pp. 182-187, 2003.
7. Maslennikow O., *Podstawy teorii zautomatyzowanego projektowania reprogramowalnych równoległych jednostek przetwarzających dla jednoukładowych systemów czasu rzeczywistego*, Koszalin, 2004.
8. Berezowski R., *Jednostki operacyjne zbudowane w oparciu o bramki prądowe dla jednoukładowych systemów VLSI*, Rozprawa doktorska, Koszalin, 2007.



9. Maslennikowa N., *Wiarygodna implementacja algorytmów algebry liniowej w układach FPGA zbudowanych w oparciu o bramki prądowe*, Rozprawa doktorska, Koszalin, 2007.
10. Pawłowski P., *Ocena przydatności bramek cyfrowych pracujących w trybie prądowym w mieszanych systemach analogowo-cyfrowych*, Rozprawa doktorska, Koszalin, 2004.
11. Słowik A., Białko M., O. Maslennikow, *Evolutionary optimization of combinational digital circuits with current-mode gates with respect to transistor count*, Int. Jour. ITIC, V.1, No.1, / 2006 r. / strony: 125-134.
12. Rajewska M., Berezowski R., Maslennikow O., *Wykorzystanie algorytmu Quine'a-McCluskey'a do optymalizacji układów cyfrowych zbudowanych z bramek prądowych*, KOWBAN 2007.
13. Guziński A., Pawłowski P., Kaniewski J., Maslennikow O., Maslennikowa N., Czwyrow D., Rataj D., *Designing digital circuits with the current-mode gates*, bulletin of the polish academy of sciences, 1999.

## Streszczenie

Od początku lat 90. ubiegłego wieku na Politechnice Koszalińskiej prowadzone są badania nad cyfrowymi bramkami prądowymi. W tym czasie opracowano kilka bramek realizujących różne operacje logiczne. Opracowane bramki wymusiły opracowanie innych (różnych od algebry Boole'a) opisów funkcji w algebrze bramek prądowych. Metody minimalizacji funkcji w algebrze bramek prądowych rozwijały się równolegle z rozwojem technologii bramek prądowych. Pierwsze sposoby opracowane były dla prostych funkcji, których implementacja wymagała zastosowania kilku bramek (2-4 argumentowe funkcje logiczne). W dalszych badaniach wskazane było opracowanie metod minimalizacji w postaci graficznej. Forma taka ułatwiała zrozumienie natury algebry bramek prądowych, ich charakterystycznych cech. Realizowano funkcje, których implementacja wymagała zastosowania wielu bramek (różne funkcje wielowyjściowe, układ S-Bloków algorytmu DES), a realizowane funkcje były maksymalnie 6 argumentowe (takie funkcje można przedstawić w postaci diagramów Veitcha-Karnaugh). W kolejnych badaniach założono, że potrzebny jest algorytm minimalizacji, który można wykonać w postaci programu komputerowego. Potrzeba taka wynikała z nowych zastosowań bramek prądowych. Projektowano w tym czasie układy pracujące w logice wielowartościowej, układy kryptograficzne. Funkcje, których opis należało zminimalizować zawierały 6 i więcej argumentów. Nie można było przedstawić ich w postaci graficznej, umożliwiającej projektantowi określenie funkcji bazowych. Potrzebne było narzędzie, które umożliwiłoby minimalizację opisu dowolnej funkcji logicznej w algebrze bramek prądowych. Do optymalizacji zastosowano między innymi algorytm ewolucyjny, funkcje wzorcowe, jak również zaproponowano modyfikację algorytmu Quine'a-McCluskey'a przeznaczonego do minimalizacji opisów funkcji w algebrze bramek prądowych.

**Abstract**

In this paper evolution of methods for minimizing the logic function in the algebra of current mode gates were developed over the last twenty years. Technology of current mode gates is associated with the concept of current gate proposed by prof. Andrzej Guziński in the 90s of the twentieth century. It will also be directions for further development of research on the optimization of current mode systems.





**Łukasz Bartłomiej Chrobak**  
**Mirosław Andrzej Maliński**  
Katedra Podstaw Elektroniki  
Wydział Elektroniki i Informatyki  
Politechnika Koszalińska

## **Badania parametrów rekombinacyjnych materiałów krzemowych z wykorzystaniem nieniszczącej techniki MFCA opartej na zjawisku modulacji absorpcji na nośnikach swobodnych**

### **Wstęp**

Parametry rekombinacyjne materiałów krzemowych są jednymi z ważniejszych, których znajomość jest niezbędna przy projektowaniu urządzeń optoelektronicznych (np. ogniw słonecznych, detektorów, itp.). W ostatnich latach można zaobserwować wzrost zainteresowania nieniszczącymi metodami do wyznaczania parametrów rekombinacyjnych materiałów krzemowych [1-4]. Jedną z nich jest metoda MFCA (ang. Modulated Free Carrier Absorption), która umożliwia wyznaczenie czasu życia nośników, współczynnika dyfuzji nośników, a także prędkości ich rekombinacji powierzchniowej. Jej podstawy teoretyczne, a także przykładowe charakterystyki częstotliwościowe amplitudowe i fazowe uzyskane na płytkach krzemowych przedstawiono w pracy [5]. Modyfikację metody MFCA opartą na przestrzennej separacji wiązki wzbudzającej nośniki i wiązki sondującej przedstawiono w pracach [6, 7]. Wpływ szerokości wiązki sondującej na sygnał MFCA przedstawiono w pracy [8]. Porównanie czułości dwóch odmian metody MFCA opartych na pomiarze sygnału MFCA w funkcji częstotliwości modulacji oraz w funkcji odległości wiązki wzbudzającej i wiązki sondującej przedstawiono w pracy [9]. Rozkład przestrzenny nośników w modelu 3D niezbędny do obliczeń wielkości sygnału MFCA przedstawiono w pracy [10]. Wyniki badań, porównujących metodę MFCA z metodą pomiarową bazującą na wykorzystaniu mikrofal, zweryfikowane pomiarami fotokonduktancji przedstawiono w pracy [11]. Przedstawione w niniejszej pracy zagadnienia dotyczące metody MFCA są ciągle aktualne i znajdują się w spektrum zainteresowań naukowców z międzynarodowych jednostek naukowych [12-14].

## Metodologia pomiarów i rozważania teoretyczne

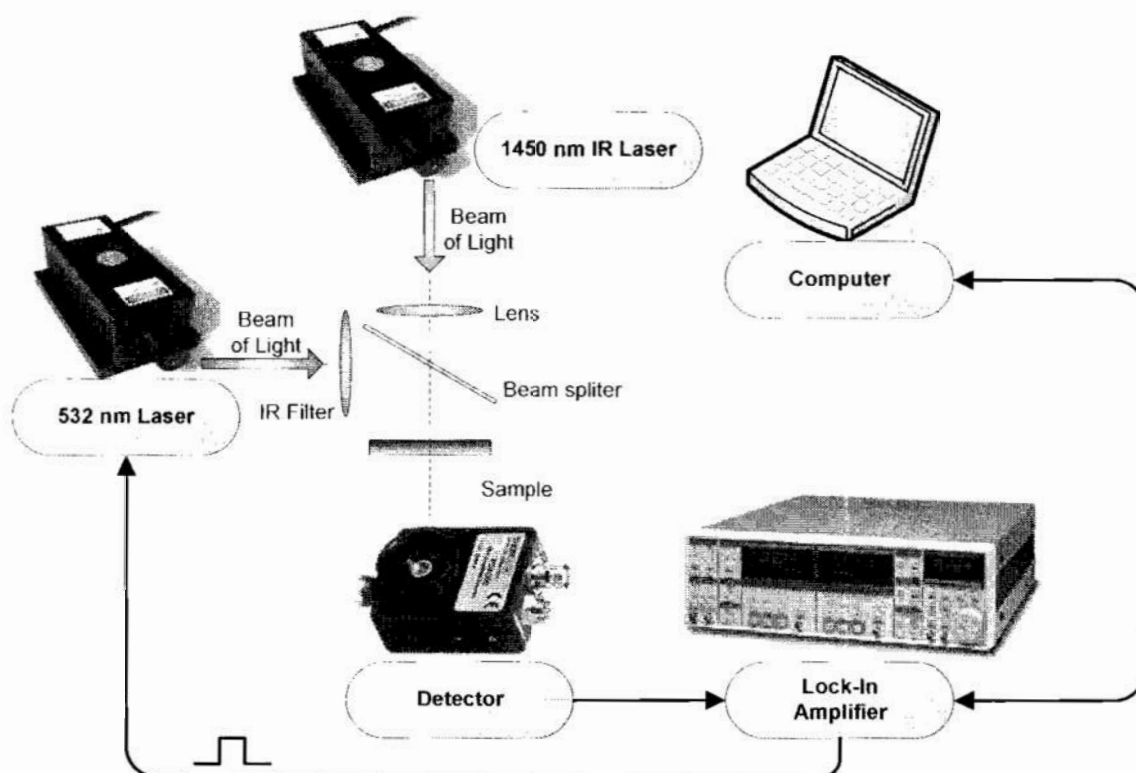
Znajomość parametrów rekombinacyjnych materiałów krzemowych jest wysoce pożądana, zwłaszcza przy projektowaniu i wytwarzaniu ogniw słonecznych. Przykładowe zdjęcie płytki krzemowej wykorzystanej do badań przedstawiono na rys. 1.



**Rys. 1.** Widok płytki krzemowej wykorzystywanej przy produkcji ogniw słonecznych (Dzięki uprzejmości Berlin Solar)

Idea metody MFCA polega na wzbudzeniu nośników z pasma walencyjnego do pasma przewodnictwa w wyniku absorpcji światła lasera wzbudzającego oświetlającego próbkę. W wyniku tego uzyskuje się dyfuzyjne fale plazmowe w próbce, a więc fale gęstości ładunku. Następnie rejestruje się natężenie wiązki sondującej przechodzącej (w bliskiej podczerwieni) przez próbkę. Dla krótszej długości fali niż długość odpowiadająca przerwie energetycznej badanego materiału, prawie cała energia zostanie zaabsorbowana, co powoduje generację nośników swobodnych w paśmie przewodnictwa. Natężenie światła lasera sondującego (pracującego na długości fali większej niż odpowiadająca długość fali przerwy energetycznej) przechodzącego przez próbkę, ze względu na absorpcję na nośnikach swobodnych w paśmie przewodnictwa, będzie mniejsze proporcjonalnie do koncentracji nośników swobodnych wygenerowanych światłem lasera wzbudzającego. Zatem sygnał transmisji światła lasera podczerwonego, przy zmodulowanej natężeniowo wiązce światła lasera wzbudzającego, jest zmodulowany w wyniku periodycznie zmiennej koncentracji nośników, powodującej periodyczną zmianę absorpcji światła na nośnikach swobodnych.

W skład stanowiska pomiarowego wchodziły dwa lasery półprzewodnikowe. Jeden z nich modulowany natężeniowo w dziedzinie częstotliwości i pracujący na długości fali 532 nm pełniący rolę lasera pompującego nośniki z pasma walencyjnego do pasma przewodnictwa oraz laser półprzewodnikowy pracujący w trybie ciągłym na długości fali 1450 nm (bliska podczerwień) pełniący rolę lasera sondującego. Za pomocą precyzyjnej optyki, oba strumienie były skupiane dokładnie w tym samym punkcie na próbce. Natężenie przechodzącego strumienia światła podczerwonego było rejestrowane za pomocą szybkiego detektora firmy Thorlabs® z serii PDA. Zarejestrowany sygnał podawany był na dwukanałowy wzmacniacz fazoczuły (Stanford Research SR 830), pełniący również rolę generatora przebiegu modulującego dla lasera pompującego. Całe stanowisko było sterowane z wykorzystaniem algorytmów sterujących zaimplementowanych w aplikacji napisanej w języku wysokiego poziomu C++. Schemat stanowiska pomiarowego do badań metodą MFCA przedstawiono na rys. 2.



Rys. 2. Schemat stanowiska eksperymentalnego do nieniszczących badań parametrów rekombinacyjnych materiałów optoelektronicznych z wykorzystaniem techniki zmodulowanej absorpcji (MFCA) na swobodnych nośnikach

Zależność na amplitudę i fazę sygnału MFCA w dziedzinie częstotliwości, który jest wprost proporcjonalny do rozkładu przestrzennego ładunku w próbce, obliczane były według wzorów przedstawionych poniżej.

$$\text{Amp}(f, \tau, D, V_1, V_2) = |\theta(f, \tau, D, V_1, V_2)| \quad (1)$$

$$\text{Phase}(f, \tau, D, V_1, V_2) = \frac{180}{\pi} (\theta(f, \tau, D, V_1, V_2)) \quad (2)$$

$$\theta(f, \tau, D, V_1, V_2) = 1 - a \int_0^d \delta n(x) dx \quad (3)$$

$$\delta n(x) = K(f, \tau) \left[ B_1(f, \tau, D, V_1, V_2) e^{\frac{x}{Le(f, \tau)}} + B_2(f, \tau, D, V_1, V_2) e^{-\frac{x}{Le(f, \tau)}} - e^{-\beta x} \right] \quad (4)$$

$$Le(f, \tau) = \left[ \frac{D\tau}{1 + i\omega(f)\tau} \right]^{\frac{1}{2}} \quad (5)$$

$$K(f, \tau) = \frac{I_0 \beta}{2ED \beta^2 - [Le(f, \tau)]^{-2}} \quad (6)$$

$$\varepsilon_b(f, \tau, V_2) = \frac{D}{V_2 Le(f, \tau)} \quad \varepsilon_g(f, \tau, V_1) = \frac{D}{V_1 Le(f, \tau)} \quad (7)$$

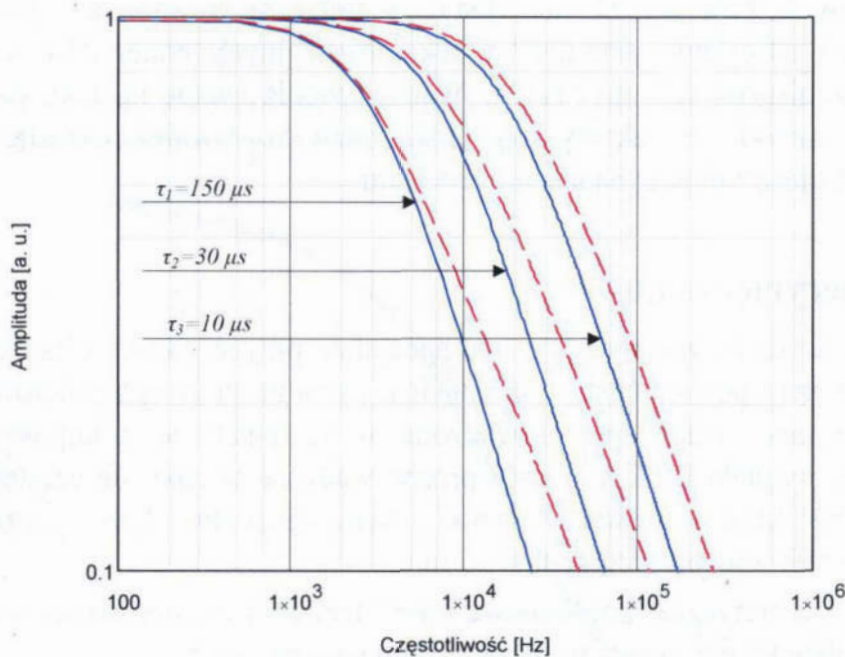
$$\eta_b(f, \tau, V_2) = \varepsilon_b(f, \tau, V_2) \beta Le(f, \tau) \quad \eta_g(f, \tau, V_1) = \varepsilon_g(f, \tau, V_1) \beta Le(f, \tau) \quad (8)$$

$$B_1(f, \tau, D, V_1, V_2) = \frac{[\eta_g(f, \tau, V_1) + 1][\varepsilon_b(f, \tau, V_2) - 1] e^{-\frac{d}{Le(f, \tau)}} - [\eta_b(f, \tau, V_2) - 1][\varepsilon_g(f, \tau, V_1) + 1] e^{-\frac{\beta d}{d}}}{[\varepsilon_g(f, \tau, V_1) + 1][\varepsilon_b(f, \tau, V_2) + 1] e^{\frac{d}{Le(f, \tau)}} - [\varepsilon_g(f, \tau, V_1) - 1][\varepsilon_b(f, \tau, V_2) - 1] e^{\frac{d}{Le(f, \tau)}}} \quad (9)$$

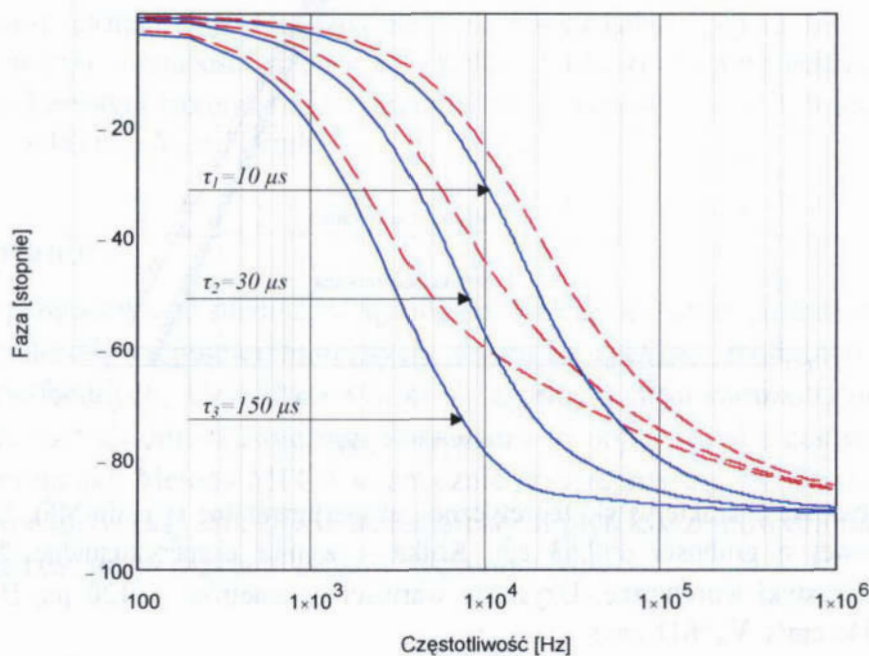
$$B_2(f, \tau, D, V_1, V_2) = \frac{[\eta_g(f, \tau, V_1) + 1][\varepsilon_b(f, \tau, V_2) + 1] e^{\frac{d}{Le(f, \tau)}} - [\eta_b(f, \tau, V_2) - 1][\varepsilon_g(f, \tau, V_1) - 1] e^{-\frac{\beta d}{d}}}{[\varepsilon_g(f, \tau, V_1) + 1][\varepsilon_b(f, \tau, V_2) + 1] e^{\frac{d}{Le(f, \tau)}} - [\varepsilon_g(f, \tau, V_1) - 1][\varepsilon_b(f, \tau, V_2) - 1] e^{\frac{d}{Le(f, \tau)}}} \quad (10)$$

Na dwóch kolejnych rysunkach (rys. 3 oraz rys. 4) przedstawione zostały wyniki symulacji charakterystyk amplitudowych i fazowych sygnału MFCA dla trzech różnych wartości czasów życia nośników. Symulacje przeprowadzone zostały według modelu przedstawionego powyżej. Wartości parametrów przyjęte do symulacji były następujące:  $D = 15 \text{ cm}^2/\text{s}$  (współczynnik dyfuzji nośników dla materiałów krzemowych typu p), prędkości rekombinacji powierzchniowych od strony polerowanej i szlifowanej kolejno  $V_B = 100 \text{ cm/s}$ ,  $V_G = 1500 \text{ cm/s}$  oraz grubość próbki  $d = 0,63 \text{ cm}$ .





**Rys. 3.** Teoretyczne charakterystyki amplitudowe sygnału MFCA dla trzech różnych wartości czasu życia nośników. Linia niebieska odpowiada wzbudzeniu próbki od strony polerowanej, natomiast linia czerwona od strony szlifowanej. Wartości poszczególnych parametrów przyjętych do symulacji:  $d=0,63$  cm,  $D=15$  cm<sup>2</sup>/s,  $V_G=1500$  cm/s,  $V_B=100$  cm/s



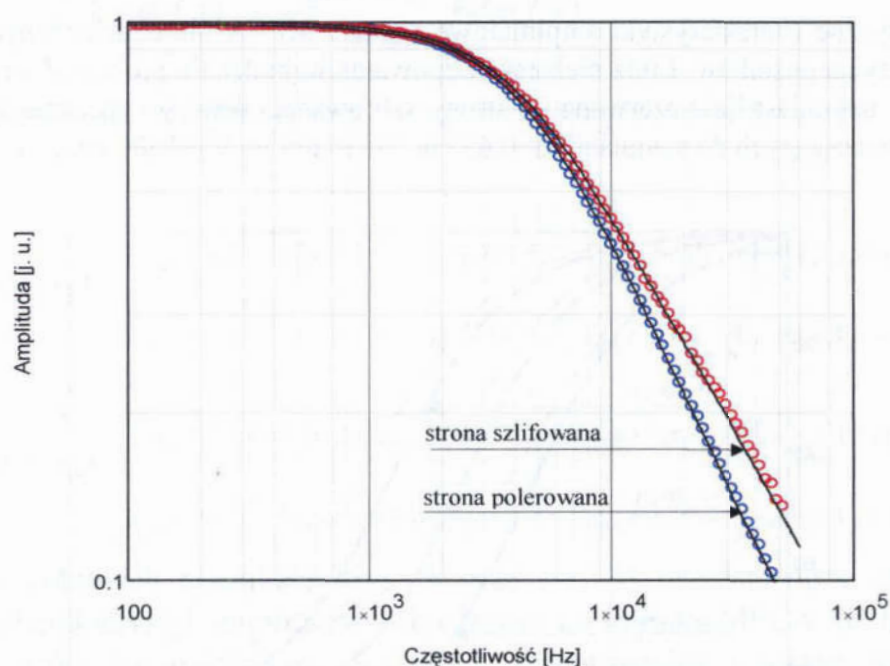
**Rys. 4.** Teoretyczne charakterystyki fazowe sygnału MFCA dla trzech różnych wartości czasu życia nośników. Linia niebieska odpowiada wzbudzeniu próbki od strony polerowanej, natomiast linia czerwona od strony szlifowanej. Wartości poszczególnych parametrów przyjętych do symulacji:  $d=0,63$  cm,  $D=15$  cm<sup>2</sup>/s,  $V_G=1500$  cm/s,  $V_B=100$  cm/s

Z przedstawionych powyżej charakterystyk teoretycznych, można wywnioskować, że możliwe jest określenie parametrów rekombinacyjnych materiałów krzemowych z charakteru przebiegów sygnału MFCA. Warto zwrócić uwagę na fakt, że metoda ta (podobnie jak metoda fotoakustyczna [15]) pozwala również określić prędkości rekombinacji powierzchniowej badanego materiału.

## Wyniki eksperymentalne

Próbki poddane analizie, zostały uprzednio specjalnie przygotowane. Charakteryzowały się dwoma rodzajami powierzchni, a mianowicie jedna ze stron była polerowana, druga szlifowana. Pomiary zostały przeprowadzone w temperaturze pokojowej. Pomiary amplitudy i fazy sygnału MFCA zostały przeprowadzone w zakresie częstotliwości od 100 Hz do 50 kHz. Uzyskane wyniki eksperymentalne były poprawione na charakterystykę referencyjną toru pomiarowego.

Doświadczalne i teoretyczne amplitudowe charakterystyki częstotliwościowe uzyskane na wybranej próbce krzemowej typu 'p' przedstawiono na rys. 5.



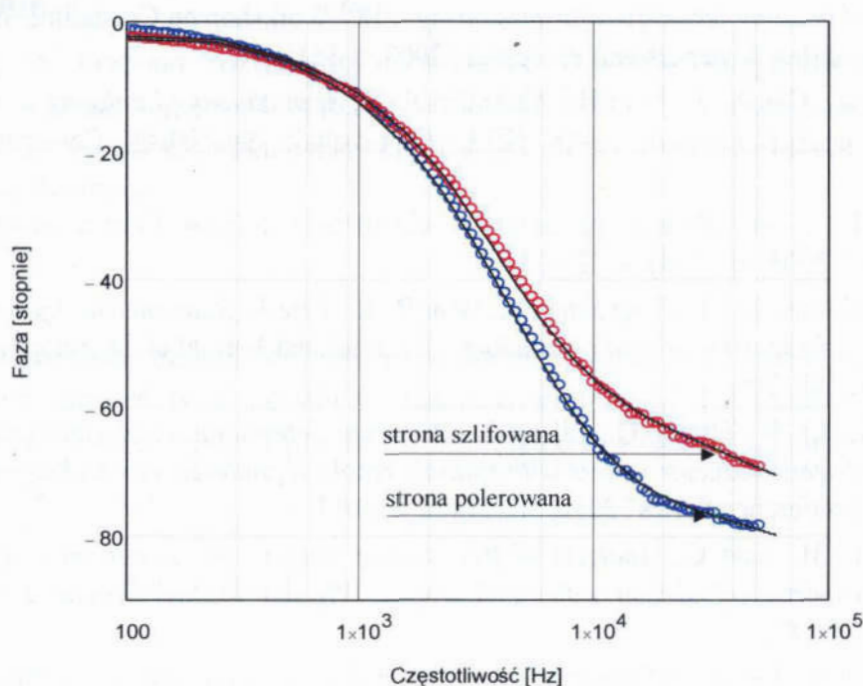
**Rys. 5.** Amplitudowe charakterystyki teoretyczne i eksperymentalne sygnału MFCA dla próbki krzemowej o grubości  $d=0,63$  cm. Kółka – wyniki eksperymentalne, linie ciągłe charakterystyki teoretyczne. Uzyskane wartości parametrów:  $\tau=120$   $\mu\text{s}$ ,  $D=17$   $\text{cm}^2/\text{s}$ ,  $V_G=2046$   $\text{cm/s}$ ,  $V_B=611$   $\text{cm/s}$

Ekstrakcja poszukiwanych wartości parametrów rekombinacyjnych odbyła się przez zastosowanie techniki optymalizacji do dopasowania krzywych teoretycznych do wyników doświadczalnych. Procedura ta realizowana była dwustopniowo, a mianowicie pierwszy uzyskany wynik podstawiany był jako punkt startowy do ponownego procesu



optymalizacji. Uzyskane wartości parametrów były następujące:  $\tau = 120 \mu\text{s}$ ,  $D = 17 \text{cm}^2/\text{s}$ ,  $V_G = 2046 \text{cm/s}$ ,  $V_B = 611 \text{cm/s}$ .

Doświadczalne i teoretyczne fazowe charakterystyki częstotliwościowe uzyskane na tej samej próbce krzemowej typu 'p' przedstawiono na rys. 6.



**Rys. 6.** Fazowe charakterystyki teoretyczne i eksperymentalne sygnału MFCA dla próbki krzemowej o grubości  $d=0,63 \text{ cm}$ . Kółka – wyniki eksperymentalne, linie ciągłe charakterystyki teoretyczne. Uzyskane wartości parametrów:  $\tau=120 \mu\text{s}$ ,  $D=17 \text{cm}^2/\text{s}$ ,  $V_G=2046 \text{cm/s}$ ,  $V_B=611 \text{cm/s}$

## Podsumowanie

W artykule przedstawiono nieniszczącą metodę MFCA do badań parametrów rekombinacyjnych materiałów optoelektronicznych, opartą na zjawisku modulacji absorpcji na nośnikach swobodnych. Uzyskane wyniki są zgodne z wynikami otrzymanymi przy zastosowaniu metody fotoakustycznej. Potwierdza to poprawność i celowość stosowanych obydwu metod. Metoda MFCA w znacznie przystępniejszy sposób (w porównaniu z metodą fotoakustyczną) umożliwia badania całych płytek krzemowych mające na celu sporządzanie tzw. mapy rozkładu czasów życia nośników.

## Podziękowania

Praca naukowa finansowana ze środków budżetowych na naukę w latach 2010-2012 jako projekt badawczy numer N N515604339.

## Bibliografia

1. Warta W., *Defect and impurity diagnostics and process monitoring*, „Solar Energy and Solar Cells” 2002, No 72, s. 389-401.
2. Berman G. M., Call N. J., Johnson S. W., Ahrenkiel R. K., Rotolante R.A., *A comparison of transient and imaging techniques for measuring*, „18<sup>th</sup> Workshop on Crystalline Silicon Solar Cells and Modules: Materials and Processes” 2008, s. 151-154.
3. Johnston S.W., Call N. J., Phan B., Ahrenkiel R. K., *Applications of imaging techniques for solar cell characterization*, „34<sup>th</sup> IEEE Photovoltaic Specialists Conference” 2009, s. 276-281.
4. Schroeder D., *Some Recent Advances in Contactless Silicon Characterization*, “ECS Transactions” 2006, No 3 (4), s. 321-337.
5. Dietzel D., Gibkes J., Chotikaprakhan S., Bein B. K., Pelz J., *Radiometric Analysis of Laser Modulated IR Properties of Semiconductors*, „International Journal of Thermophysics” 2003, No 24 (3), s. 741-755.
6. Li B., Li X., Li W., Huang Q., Zhang X., *Accurate determination of electronic transport properties of semiconductor wafers with spatially resolved photo-carrier techniques*, „Journal of Physics: Conference Series” 2010, No 214, s. 012013.
7. Zhang X., Li B., Gao C., *Analysis of free carrier absorption measurement of electronic transport properties of silicon wafers*, “European Physics Journal Special Topics” 2008, No 153, s. 279-281.
8. Huang Q., Li B., Liu X., *Influence of probe beam size on signal analysis of modulated free carrier absorption technique*, “Journal of Physics: Conference Series” 2010, No 214, s. 012084.
9. Li W., Li B., *Analysis of modulated free-carrier absorption measurement of electronic transport properties of silicon wafers*, “Journal of Physics: Conference Series” 2010, No 214, s. 012116.
10. Mandelis A., Balista J., Shaughnessy D., *Infrared photocarrier radiometry of semiconductors: Physical principles, quantitative depth profilometry, and scanning imaging of deep subsurface electronics defects*, “Physical Review” 2003, No B 67, s. 205208.
11. Schmidt J., *Measurement of differential and actual recombination parameters on crystalline silicon wafers*, “IEEE Transactions on Electron Devices” 1999, No 46 (10), s. 2018-2025.
12. Creazzo T., Redding B., Marchena E., Shi S., Prater D. W., *Free-carrier absorption modulation in silicon nanocrystal slot waveguides*, “Optics Letters” 2010, No 35 (21), s. 3691-3693.
13. Huang Q., Li B., *Self-eliminating instrumental frequency response from free carrier absorption signals for silicon wafer characterization*, “Rev. Sci. Instrum.” 2011, No 82, s. 043104.
14. Huang Q., Li B., *Electronic transport characterization of silicon wafers by combination of modulated free carrier absorption and photocarrier radiometry*, “J. Appl. Phys.” 2011, No 109, s. 023708.

15. Malinski M., Chrobak Ł., Patryn A., *Theoretical and Experimental Studies of a Plasma Wave Contribution to the Photoacoustic Signal for Si Samples*, "Acta Acustica united with Acustica" 2009, No 95, s. 60-64.

## **Streszczenie**

W pracy przedstawiono rozważania teoretyczne i wyniki eksperymentalne badań parametrów rekombinacyjnych materiałów krzemowych z wykorzystaniem techniki MFCA. Wykorzystana technika badawcza bazuje na zjawisku modulacji absorpcji na nośnikach swobodnych.

## **Abstract**

This paper presents theoretical and experimental results of investigations of the recombination parameters of the silicon materials with the MFCA method. The applied technique is based on the phenomena of the modulated free carriers absorption.





**Dariusz Bernatowicz**

**Kamil Szymański**

Katedra Inżynierii Komputerowej

Wydział Elektroniki i Informatyki

Politechnika Koszalińska

## **Realizacja fragmentacji mieszanej w natywnych dokumentach XML**

### **1. Wprowadzenie**

Rosnąca popularność języka XML (ang. Extensible Markup Language) [7] oraz szeroki zakres jego zastosowań stworzyły potrzebę efektywnego składowania i zarządzania danymi XML. Tradycyjnie, dane gromadzone są i zarządzane w relacyjnych bazach danych, które z kolei umożliwiają ich transformację do postaci XML. Jednak wymagania nowoczesnych systemów informacyjnych mają często charakter zbyt złożony, uniemożliwiający ich implementację w relacyjnej postaci. Do wymagań tych zaliczamy głównie systemy sieciowe takie jak: e-commerce, e-banking, geograficzne systemy informacyjne, biologiczne systemy sieciowe itp. oraz aplikacje przetwarzające dużą liczbę dokumentów XML np. dokumentacje techniczne. W celu uniknięcia procesu transformacji danych XML powstały rozwiązania wspierające składowanie dokumentów w systemach natywnych (systemach dedykowanych).

Jednym ze sposobów takiego składowania jest rozproszenie danych realizowane w procesie fragmentacji. Podział dokumentu na określoną liczbę części i rozmieszczenie ich w odpowiednich miejscach pozwala uzyskać przez system informatyczny działający w sieci wiele zalet, a mianowicie: redukcję pobierania i przetwarzania przez aplikacje nadmiernej liczby danych, możliwość równoległego wykonania pojedynczego zapytania poprzez jego rozdzielenie na zbiór zapytań składowych operujących na fragmentach oraz zmniejszenie ilości przesyłanych danych w przypadku wewnętrznej komunikacji.

Hierarchiczna i semistrukturalna właściwość dokumentów XML powoduje podczas ich przetwarzania ogólny wzrost zapotrzebowania na istotne czynniki systemu, tj. wymaganą pamięć i wydajność przetwarzania. Ponieważ większość zapytań na dużych dokumentach XML jest typu selektywnego to korzystanie przez nich z fragmentów dokumentu może przynieść znaczące obniżenie wymagań dotyczących pamięci i wydajności systemu.

W ostatnich latach zagadnienie fragmentacji natywnych dokumentów XML jest tematem szerokich badań lecz mimo to nie udało się osiągnąć jednoznacznego rozwiązania. Dlatego większość opisywanych w literaturze algorytmów ma charakter

heurystyczny i różni się sposobem podejścia oraz ustalonymi kryteriami. W pracy [6] dokonano próby zaadoptowania technik fragmentacji stosowanych w relacyjnych bazach danych dla dokumentów XML. Przedstawiono zarys problematyki procesu fragmentacji oraz zaproponowano metodę podziału poziomego i pionowego. Niestety, wykorzystanie relacyjnej bazy danych umożliwia przechowywanie tylko dokumentów zorientowanych na dane (ang. Data-centric XML documents) i nie pozwala na zastosowanie zaproponowanych rozwiązań w silnie hierarchicznych natywnych dokumentach XML (ang. Document-centric XML documents).

Heurystyczne podejście do minimalizacji kosztów zapytania dla przypadku fragmentacji poziomej przedstawiono w pracy [5]. Wykazano w niej, że minimalny koszt transportu może być osiągnięty lokalnie poprzez przyjęcie fragmentacji pionowej z prostym predykatem charakteryzującym najczęściej występujące zapytania. Niestety wymaga to przeprowadzenia złożonego procesu optymalizacji zapytań w celu określenia wartości predykatu.

Inne podejście przedstawiono w pracy [2], w której szerokość i głębokość drzewa przyjęto jako ograniczenia strukturalne. Aby określić wartości tych ograniczeń i miejsca podziału zastosowano histogramy opisujące strukturę dokumentu. Do głównych wad tej metody możemy zaliczyć brak uwzględnienia logicznych relacji między elementami dokumentu, a także możliwość wygenerowania pustych fragmentów.

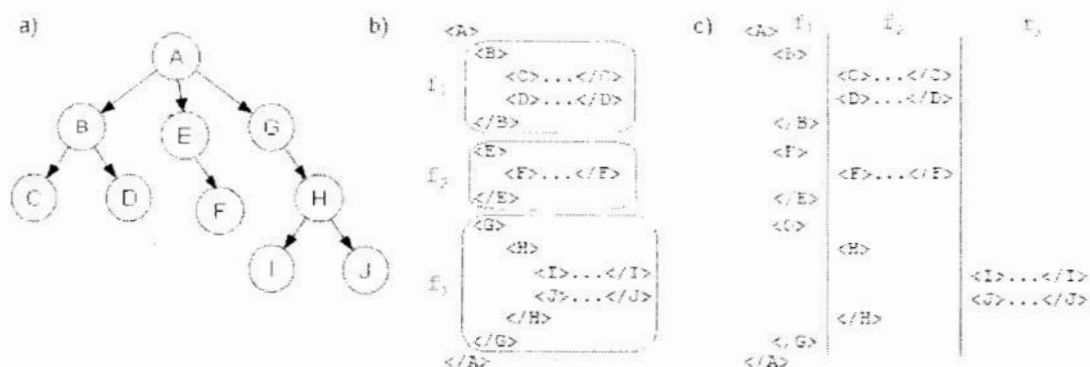
Podejście oparte na statystyce zapytań i graficznej strukturze dokumentu XML przedstawiono w pracy [1]. Zaprezentowano w niej dwa modele: pierwszy to heurystyczny model fragmentacji poziomej oparty na zapytaniach (ang. vertical query based fragmentation model), a drugi to model oparty na rozmiarach struktury (ang. structure-and-size based fragmentation model). Analiza występujących w systemie zapytań wykorzystywana jest w procesie grupowania pokrewnych elementów. Przez pojęcie pokrewne elementy rozumiemy takie elementy dokumentu XML do których zapytania odwołują się jednocześnie. Pozwala to określić elementy najczęściej pobierane przez zapytania z określonych miejsc systemu. Podejście to zostało wykorzystane w artykule, w którym skoncentrowano się głównie na omówieniu procesu fragmentacji mieszanej natywnych dokumentów XML opartych na logicznym modelu danych dedykowanym dla mających drzewiastą strukturę dokumentów np. DOM [3] lub XQuery [8].

## 2. Charakterystyka procesu fragmentacji dokumentów XML

Dokument o strukturze hierarchicznej  $X$  nazywamy pofragmentowanym, jeżeli jest podzielony na minimalny zbiór  $F = \{f_1, f_2, \dots, f_n\}$  wzajemnie rozłącznych części zwanych fragmentami. Fragmenty dokumentu  $f_1, f_2, \dots, f_n$  zawierają dostateczną informację umożliwiającą odtworzenie oryginalnego dokumentu  $X$ . Powszechnie fragmentacja danych realizowana jest na trzy różne sposoby, a mianowicie pionową, poziomą i mieszaną.

## 2.1. Fragmentacja pionowa

Fragmentacja pionowa dokumentu XML (ang. vertical fragmentation of an XML document) określana jest na poziomie ścieżki w której elementy dokumentu X łączone są w grupy, a następnie rzutowane do fragmentów zgodnych z tymi grupami elementów. Najprostszą formą jest dekompozycja dokumentu z umieszczeniem unikalnego identyfikatora w każdym fragmencie. Przykład fragmentacji pionowej przedstawiono na rys. 1b, w której zadany dokument podzielono na trzy części, a rolę identyfikatora spełnia fragment ścieżki określający położenie węzłów w dokumencie. I tak dla pierwszego fragmentu identyfikator określa ścieżkę  $f_1:A/B$ , dla drugiego otrzymujemy  $f_2:A/E$ , a dla trzeciego  $f_3:A/G$ .



**Rys. 1.** Graf A-J ze znakowaniem preorder przykładowego dokumentu XML (a) i jego podział na trzy fragmenty: określone na poziomie ścieżki (b) oraz określone na poziomie warstwy (c)

Odtworzenie pierwotnej postaci dokumentu X uzyskujemy w procesie złączenia wszystkich fragmentów zbioru F przy wykorzystaniu identyfikatorów poszczególnych fragmentów, umożliwiającą zachowanie odpowiedniej kolejności węzłów w dokumencie X. Kwerendę w języku XQuery realizującą złączenie fragmentów dla dokumentu z rysunku 1b przedstawiono na rysunku 2. Składa się ona z klauzuli for, w której tworzymy zmienne odwołujące się do węzłów głównych poszczególnych fragmentów i klauzuli return zwracającej wynik.

```
xquery
for
  $a in doc("fragment1.xml")/A/B,
  $b in doc("fragment2.xml")/A/E,
  $c in doc("fragment3.xml")/A/G
return
  <A>
    { $a }
    { $b }
    { $c }
  </A>
```

**Rys. 2.** Kwerenda języka XQuery realizująca złączenie elementów zawartych w trzech fragmentach dla fragmentacji pionowej

## 2.2. Fragmentacja pozioma

Fragmentacja pozioma dokument XML (ang. horizontal fragmentation of an XML document) określana jest na poziomie warstwy głębokości, w której występuje podział węzłów dokumentu  $X$  na podzbiory. Fragment  $f_i$  jest wynikiem selekcji elementów na dokumencie  $X$  według predykatu  $P$ . Przykład fragmentacji poziomej przedstawiono na rys. 1c, w którym dokument  $X$  podzielono na trzy fragmenty  $f_1, f_2$  i  $f_3$  złożone z elementów umieszczonych w kolejnych warstwach głębokości. Fragment  $f_1$  składa się ze zbioru węzłów  $\{A, B, E, G\}$ , fragment  $f_2$  ze zbioru  $\{C, D, E, H\}$ , a fragment  $f_3$  ze zbioru  $\{I, J\}$ .

```
xquery
for $a in doc("fragment1.xml")/A
return
  <A>
    { $a/B
      { for $b in doc("fragment2.xml")/B
        return
          { $b/C
            { $b/D }
          }
      }
    { $a/E
      { for $c in doc("fragment2.xml")/E
        return { $c/F }
      }
    { $a/G
      { for $d in doc("fragment2.xml")/G
        return
          { $d/H
            { for $e in doc("fragment3.xml")/H
              return
                { $e/I
                  { $e/J }
                }
            }
          }
      }
    }
  </A>
```

**Rys. 3.** Kwerenda języka XQuery realizująca grupowanie fragmentów poprzez wykorzystanie zapytań zagnieżdżonych dla fragmentacji poziomej

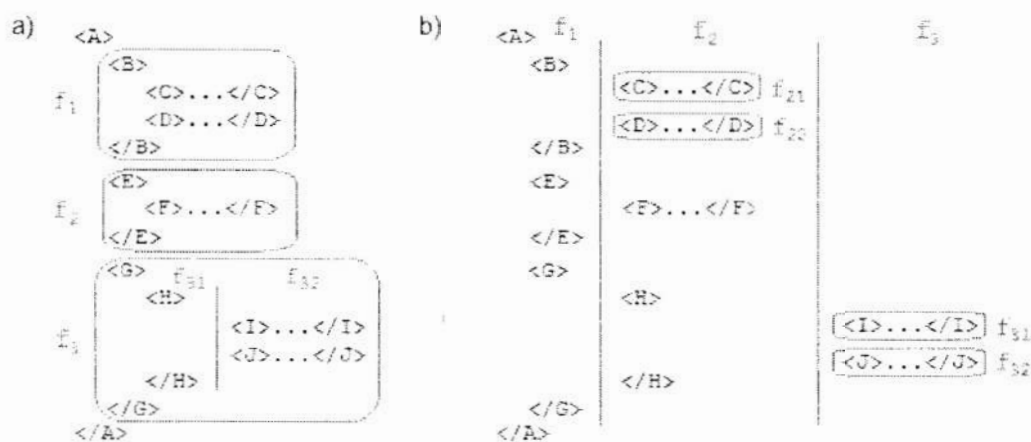
Odtworzenie pierwotnej postaci dokumentu  $X$  uzyskujemy w procesie grupowania poszczególnych fragmentów. W języku XQuery grupowanie możemy zrealizować poprzez odpowiednie zagnieżdżanie wynikowych elementów XML i tak dla przykładu z rysunku 1c odtworzenie dokumentu  $X$  przedstawiono na rysunku 3. Zapytania zagnieżdżone odwołujące się do fragmentu drugiego i trzeciego umieszczono w klauzuli wynikowej return. Odwołanie się do fragmentu występującego w dalszej warstwie zagnieżdżenia odbywa się z warstwy położonej o jeden stopień niżej, a więc rezultat zapytania na fragmencie  $f_3$  zwracany jest najpierw do fragmentu  $f_2$ , a potem dopiero do fragmentu  $f_1$ .

## 2.3. Fragmentacja mieszana

Fragmentacja mieszana dokumentu XML (ang. mixed fragmentation of an XML document) określana jest na poziomie elementu, w której występuje połączenie reguł

fragmentacji pionowej z poziomą. Możemy wyróżnić dwa rodzaje fragmentacji mieszanej, a mianowicie:

- fragmentację poziomą do fragmentów pionowych, w której wybrane fragmenty podzielonego dokumentu X na poziomie ścieżki są dodatkowo dzielone na określoną liczbę części na poziomie warstwy głębokości. Przykład z rysunku 4a przedstawia podział fragmentu  $f_3$  z dokumentu X na podfragmenty  $f_{31}$  i  $f_{32}$  stanowiące dodatkowe warstwy tego fragmentu. Odtworzenie pierwotnej postaci dokumentu X uzyskujemy za pomocą dwóch operacji, a mianowicie sumowania i złączenia. W pierwszym etapie wykonujemy operację sumowania fragmentów podzielonych na poziomie warstwy głębokości, a następnie realizujemy operację złączenia wszystkich fragmentów dokumentu X;
- fragmentację pionową do fragmentów poziomych, w której wybrane fragmenty podzielonego dokumentu X na poziomie warstwy głębokości są dodatkowo dzielone na określoną liczbę części na poziomie ścieżki. Przykład takiego podziału przedstawiono na rysunku 4b, w którym fragmenty  $f_2$  i  $f_3$  podzielono na podfragmenty określone na poziomie ścieżki. Odtworzenie pierwotnej postaci dokumentu X wymaga realizacji najpierw operacji złączenia fragmentów określonych na poziomie ścieżki, a następnie dokonanie operacji sumowania wszystkich fragmentów dokumentu X.



Rys. 4. Fragmentacja mieszana przykładowego dokumentu XML: pozioma do fragmentów pionowych (a) i pionowa do fragmentów poziomych (b)

### 3. Model fragmentacji dokumentu XML

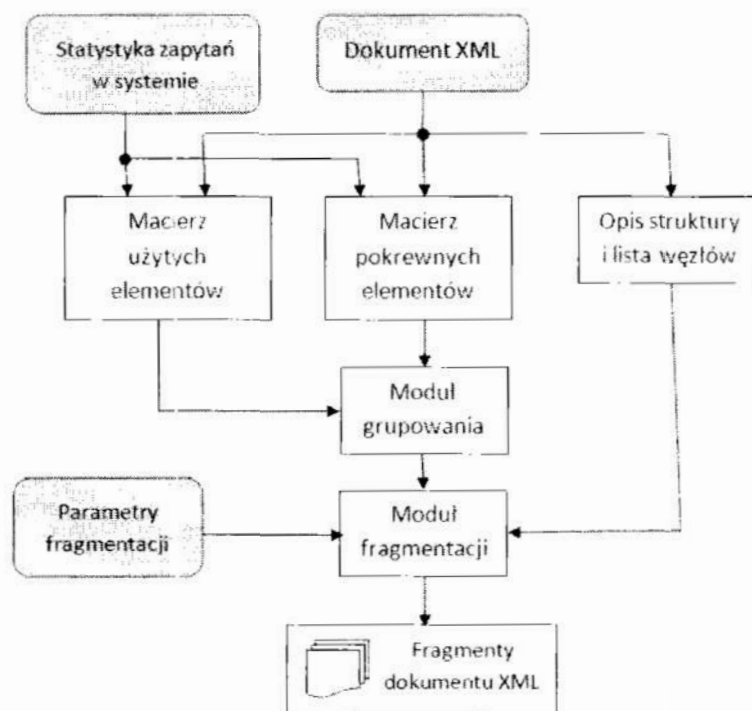
Elementy dokumentu XML w modelu fragmentacji mieszanej łączone są w grupy, a sam dokument rzutowany jest na fragmenty zgodne z uzyskanymi grupami elementów. Na reguły grupowania może mieć wpływ różny poziom szczegółowości analizy wykorzystywanego dokumentu XML oraz dodatkowe informacje charakteryzujące



system rozproszony tj. obciążenie liczbą zapytań poszczególnych miejsc systemu, rodzaj zapytań, charakter pobieranych danych oraz parametry techniczne.

Przez system rozproszony DS rozumiemy system składający się ze zbioru miejsc  $S = \{s_1, s_2, \dots, s_m\}$ , w którym każde miejsce posiada niezależną jednostkę przetwarzania. Do parametrów dodatkowych opisujących system DS. zaliczamy zbiór zapytań  $Q = \{q_1, q_2, \dots, q_k\}$  wywoływanych przez wszystkie miejsca oraz macierz częstotliwości  $F = Q \times S$ , z którą określone zapytanie wykonywane jest w danym miejscu. Przez wyraz macierzy  $f_{ij}$  określamy częstotliwość wykonania zapytania  $q_i$  w miejscu  $s_j$ . Wszystkie elementy występujące w dokumencie  $X$  określono w zbiorze  $E = \{e_1, e_2, \dots, e_p\}$ .

W modelu przedstawionym na rysunku 5 skupiono się głównie na trzech rodzajach danych wejściowych, a mianowicie: dokumencie XML, parametrach fragmentacji i statystyce zapytań występujących w systemie. Pierwotny dokument XML zawierający strukturę dzielonego pliku i listę elementów wykorzystywany jest przy tworzeniu macierzy użytych elementów, macierzy pokrewnych elementów oraz opisie struktury dokumentu XML wykorzystywanego w module fragmentacji. Parametry fragmentacji określają podstawowe właściwości opisujące proces dzielenia dokumentu i zaliczamy do nich rodzaj fragmentacji, liczbę plików na którą ma zostać podzielony dokument główny oraz stopnie głębokości warstwy dla węzłów, w których występuje fragmentacja pozioma. Statystyka zapytań opisuje zachowanie systemu w stosunku do występującego w nim dokumentu. Poprzez określenie teoretyczne lub rzeczywisty pomiar statystyczny określamy macierz częstotliwości wystąpienia zapytań w miejscach systemu oraz zbiór elementów dokumentu do których odwołują się poszczególne zapytania.



Rys. 5. Model fragmentacji dokumentu XML

Do podstawowych komponentów wyróżnionych w modelu fragmentacji dokumentu XML zaliczamy:

- Macierz użytych elementów (MUE), w której wiersze określone są przez zbiór zapytań  $Q$ , a kolumny przez zbiór elementów  $E$  dokumentu  $X$ . Wyrazami macierzy są liczby 1 w przypadku odwołania zapytania  $q_i$  do elementu  $e_j$  lub liczba 0 w przypadku przeciwnym.
- Macierz pokrewnych elementów (MPE) opisująca macierz kwadratową, w której zarówno wiersze jak i kolumny określone są przez zbiór elementów  $E$ , a wyrazami jest liczba odwołań wszystkich zapytań ze zbioru  $Q$  zarówno do elementu  $e_i$  jak i  $e_j$ . Pozwala to na określenie par elementów dokumentu, do których zapytania odwołują się jednocześnie, więc elementy te powinny zostać umieszczone w tym samym fragmencie.
- Moduł grupowania (MG) łączy pokrewne pary elementów w większe bloki i wykorzystując dane z MUE i MPE oblicza miejsce lub miejsca podziału dokumentu. Do obliczenia miejsc podziału wykorzystuje się algorytm BEA (Bond Energy Algorithm)[4].
- Moduł fragmentacji dokonuje podziału dokumentu XML na podstawie opisu struktury dokumentu, parametrów fragmentacji oraz miejsc podziału określonych przez moduł grupowania. Realizuje również proces ekstrakcji wyznaczonych grup elementów dokumentu do utworzonych plików będących fragmentami.
- Fragmenty dokumentu XML określają zbiór plików, z których każdy zawiera określoną część dokumentu  $X$ . Uzyskane fragmenty nie powinny być plikami pustymi, a zawarte w nich zbiory elementów musi cechować nierozłączność.

#### 4. Algorytm fragmentacji mieszanej

W algorytmie podziału dokumentu XML na zadaną liczbę fragmentów stosujemy technikę fragmentacji opartej na stopniu pokrewieństwa między parami elementów dokumentu i próbie połączenia tych par w większe grupy z wykorzystaniem algorytmu BEA. Algorytm ten decyduje o podziale dokumentu na podstawie dostępu aplikacji systemu do jego elementów i może zostać opisany w następujących krokach:

**Krok 1:** Na podstawie statystyki zapytań i struktury dokumentu utworzenie macierzy użycia elementów MUE.

**Krok 2:** Na podstawie statystyki zapytań utworzenie macierz pokrewnych elementów MPE

**Krok 3:** Na podstawie danych z MUE i MPE wygenerowanie przez moduł grupujący macierzy grupowania pokrewnych elementów MGPE, zawierającej poformowane elementy pokrewne.

**Krok 4:** Zastosowanie algorytmu BEA do określenia miejsc podziału poprzez znalezienie zestawów elementów MGPE, które wykazują silne pokrewieństwo.

**Krok 5:** Na podstawie uzyskanych danych dotyczących punktów podziału z kroku 4 i po uwzględnieniu wprowadzonych parametrów fragmentacji do modułu fragmentacji, podzielenie dokumentu X na zadaną liczbę fragmentów. Kolejność realizacji fragmentacji mieszanej zależy od wybranego jej typu. Najpierw wykonywany jest podział wewnętrzny: poziomy lub pionowy, po czym następuje realizacja podziału zewnętrznego.

Umieszczenie zbioru elementów dokumentu X w zadanym fragmencie powinno zostać zrealizowane z uwzględnieniem nieregularnej struktury i zróżnicowanej granulacji charakterystycznej dla natywnych dokumentów XML. Uzyskuje się to poprzez zachowanie odpowiedniej kolejności przy rozmieszczeniu elementów w zadanych fragmentach. Natomiast odpowiednią sekwencję utworzonych fragmentów uzyskujemy poprzez nadanie im unikalnego identyfikatora położenia względem dokumentu X.

## 5. Wnioski

W pracy omówiony został proces realizacji fragmentacji mieszanej natywnych dokumentów XML opartych na logicznym modelu danych dedykowanym dla mających drzewiastą strukturę dokumentów typu DOM lub XQuery. Przedstawione zostały podstawowe rodzaje fragmentacji i sposób odzyskania pierwotnej postaci dokumentu dla każdej z nich. Zaprezentowany został także model blokowy realizacji procesu fragmentacji oraz algorytm podziału i rozmieszczenia poszczególnych elementów dokumentu w zadanej liczbie fragmentów w zależności od rodzaju zapytań i ich częstotliwości występowania w określonych miejscach systemu.

W ramach dalszych rozważań dotyczących powyższego zagadnienia można skupić się również na określeniu kosztów alokacji poszczególnych fragmentów oraz kosztów wymiany komunikatów i danych dla przetwarzania zapytań na otrzymanych fragmentach.

## Bibliografia

1. Birahnu L., Atnafu S., Getahun F., *Native XML Document Fragmentation Model*, Sixth Int. Conf. on Signal-Image Technology and Internet Based Systems, Kuala Lumpur, 2010.
2. Bonifati A., Cuzzocrea A., *Efficient fragmentation of Large XML Documents*, Springer-Verlag, Berlin, Heidelberg, 2007.
3. Dom Working Group: *Document Object Model (DOM) Level 1 Specification, version 1.0*, 1998.
4. Hoffer A., Severance D.: *The use of Cluster Analysis in Physical database design*, First Int. Conf. on very large Database, New York, 1975.

5. Ma H., Schewe D., *Heuristic Horizontal XML fragmentation*, In: proc. OfCAiSE, Brazil, 2005.
6. Ma H., Schewe D., *Fragmentation of XML Documents*. In: proc. OfCAiSE, Brazil, 2005.
7. Navathe S., Ceri S., Wiederhold G., Dov J., *Vertical partitioning algorithm for Database Design*, ACM Trans. on Database Systems, 9, 680-710, 1989.
8. XML Core Working Group: *Extensible Markup Language (XML) 1.0 (Fifth Edition)*, W3C Recommendation 26 Nov 2008, <http://www.w3.org/TR/2008/REC-xml-20081126/>
9. XML Query Working Group and XSL Working Group: *XQuery 1.0 and XPath 2.0 Data Model*, W3C Working Draft 29 Oct 2004, <http://www.w3.org/TR/2004/WD-xpath-datamodel-20041029/>

## Streszczenie

W niniejszej pracy omówiono realizację fragmentacji mieszanej natywnych dokumentów XML, a także scharakteryzowano podstawowe rodzaje fragmentacji i ich podział. Przedstawiony został model blokowy oraz algorytm wykonania podziału i rozmieszczenia poszczególnych elementów dokumentu w zadanej liczbie fragmentów w zależności od rodzaju zapytań i ich częstotliwości występowania w określonych miejscach systemu.

**Słowa kluczowe:** XML, fragmentacja mieszana, bazy danych, algorytm

## Abstract

This study describes an implementation of mixed XML documents fragmentation and characterizes basic types of fragmentation with its division. Box model and an algorithm of the division, that depends on queries type and frequency of occurrence in certain system areas was presented.

**Keywords:** XML, mixed fragmentation, databases, algorithm





**Dariusz Jakóbczak**

Katedra Podstaw Informatyki i Zarządzania

Wydział Elektroniki i Informatyki

Politechnika Koszalińska

## **Shape Coefficients via Method of Hurwitz-Radon Matrices**

### **1. Introduction**

A significant problem in object recognition and computer vision [1] is that of appropriate shape representation and reconstruction. Classical discussion about shape representation is based on the problem: contour versus skeleton. This paper is voting for contour which forms boundary of the object. Contour of the object, represented by contour points, consists of information which allows us to describe many important features of the object as the shape coefficients [2].

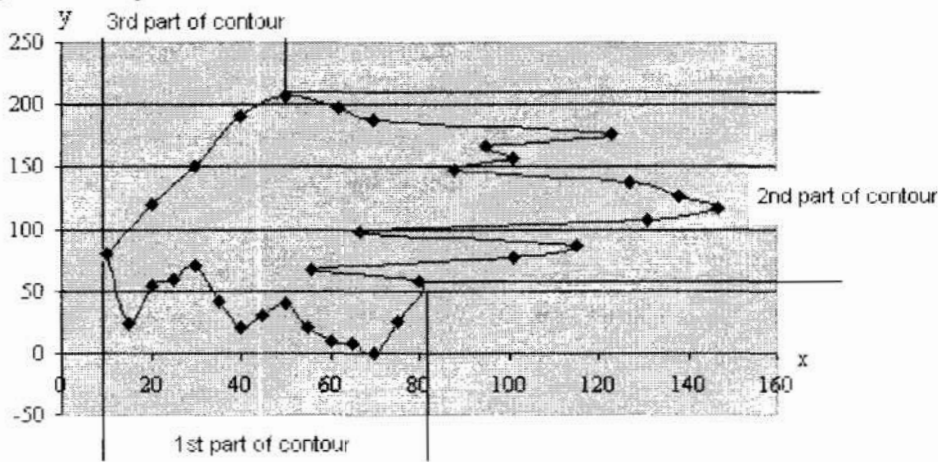
A digital curve (open or closed) may be represented by chain code (Freeman's code). Chain code depends on selection of the started point and transformations of the object. So Freeman's code is one of the method how to describe and to find contour of the object. An analog (continuous) version of Freeman's code is the curve  $\alpha - s$ . Another contour representation and reconstruction is based on the Fourier coefficients calculated in Discrete Fourier Transformation (DFT). These coefficients are used to fix similarity of the contours with different sizes or directions. If we assume that a contour is built from the segments of a line and fragments of circles or ellipses, Hough transformation is applied to detect the contour lines. Also geometrical moments of the object are used during the process of object shape representation [3]. MHR method requires to detect specific points of the object contour, for example in compression and reconstruction of monochromatic medical images [4]. Contour is also applied in the shape decomposition [5]. Many branches of medicine, for example computed tomography [6], need suitable and accurate methods of contour reconstruction [7]. Also industry and manufacturing are looking for the methods connected with geometry of the contour [8]. So suitable shape representation and precise reconstruction or interpolation [9] of the object contour is a key factor in many applications of computer analysis and image processing.

## 2. Contour Representation

The shape can be represented by the object contour, i.e. curves that create each part of the contour. One curve is described by the set of nodes  $(x_i, y_i) \in \mathbf{R}^2$  (contour points) as follows in proposed method:

1. nodes (interpolation points) are settled at local extrema (maximum or minimum) of one of coordinates and at least one point between two successive local extrema;
2. nodes  $(x_i, y_i)$  are monotonic in coordinates  $x_i$  ( $x_i < x_{i+1}$  for all  $i$ ) or  $y_i$  ( $y_i < y_{i+1}$ );
3. one curve (one part of the contour) is represented by at least five nodes.

Condition 1 is done for the most appropriate description of a curve. So we have  $m$  curves  $C_1, C_2, \dots, C_m$  that build whole contour and each curve is represented by the nodes according to assumptions 1-3.



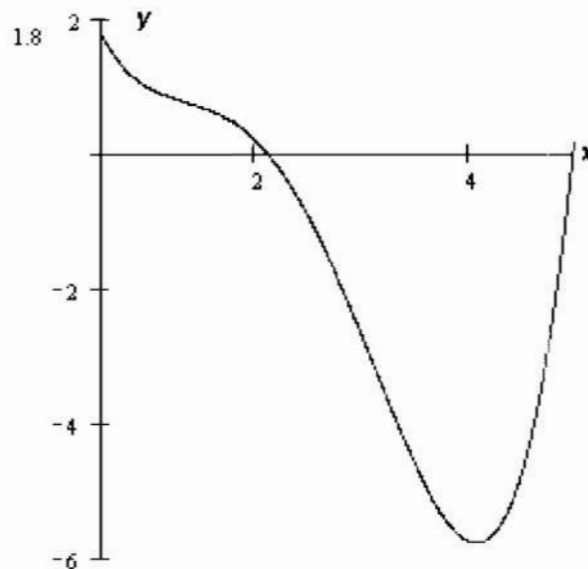
**Fig. 1.** A contour consists of three parts (three curves and their nodes)

Fig.1 is an example for  $m = 3$ : first part of the contour  $C_1$  is represented by the nodes monotonic in coordinates  $x_i$ , second part of the contour  $C_2$  is represented by the nodes monotonic in coordinates  $y_i$  and third part  $C_3$  could be represented by the nodes either monotonic in coordinates  $x_i$  or monotonic in coordinates  $y_i$ . Number of the curves is optional and number of the nodes for each curve is optional too (but at least five nodes for one curve). Representation points are treated as interpolation nodes. How accurate can we reconstruct whole contour using representation points? The contour reconstruction is possible using novel MHR method.

## 3. Contour Reconstruction

The following question is important in mathematics and computer sciences: is it possible to find a method of curve interpolation in the plane without building the interpolation polynomials or other functions? Our paper aims at giving the positive answer to this question. In comparison MHR method with Bézier curves, Hermite curves

and B-curves (*B-splines*) or NURBS one unpleasant feature of these curves must be mentioned: a small change of one characteristic point can make big change of whole reconstructed curve. Such a feature does not appear in MHR method. The methods of curve interpolation based on classical polynomial interpolation: Newton, Lagrange or Hermite polynomials and the spline curves which are piecewise polynomials [10]. Classical methods are useless to interpolate the function that fails to be differentiable at one point, for example the absolute value function  $f(x) = |x|$  at  $x=0$ . If point  $(0;0)$  is one of the interpolation nodes, then precise polynomial interpolation of the absolute value function is impossible. Also when the graph of interpolated function differs from the shape of polynomials considerably, for example  $f(x) = 1/x$ , interpolation is very hard because of existing local extrema of polynomial. Lagrange interpolation polynomial for function  $f(x) = 1/x$  and nodes  $(5;0.2)$ ,  $(5/3;0.6)$ ,  $(1;1)$ ,  $(5/7;1.4)$ ,  $(5/9;1.8)$  has one minimum and two roots.



**Fig. 2.** Lagrange interpolation polynomial for nodes  $(5;0.2)$ ,  $(5/3;0.6)$ ,  $(1;1)$ ,  $(5/7;1.4)$ ,  $(5/9;1.8)$  differs extremely from the shape of function  $f(x) = 1/x$

We cannot forget about the Runge's phenomenon: when the interpolation nodes are equidistance then high-order polynomial oscillates toward the end of the interval, for example close to -1 and 1 with function  $f(x) = 1/(1+25x^2)$  [11]. Method of Hurwitz – Radon Matrices (MHR), described in this paper, is free of these bad features. The curve or function in MHR method is parameterized for value  $\alpha \in [0;1]$  in the range of two successive interpolation nodes.

### 3.1. The Operator of Hurwitz-Radon

Adolf Hurwitz (1859-1919) and Johann Radon (1887-1956) published the papers about specific class of matrices in 1923, working on the problem of quadratic forms. Matrices  $A_i$ ,  $i = 1, 2, \dots, m$  satisfying

$$A_j A_k + A_k A_j = 0, A_j^2 = -I \text{ for } j \neq k; j, k = 1, 2, \dots, m$$

are called a *family of Hurwitz - Radon matrices*. A family of Hurwitz - Radon (HR) matrices has important features [12]: HR matrices are skew-symmetric ( $A_i^T = -A_i$ ) and reverse matrices are easy to find ( $A_i^{-1} = -A_i$ ). Only for dimension  $N = 2, 4$  or  $8$  the family of HR matrices consists of  $N - 1$  matrices. For  $N = 2$  we have one matrix:

$$A_1 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

For  $N = 4$  there are three HR matrices with integer entries:

$$A_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}.$$

For  $N = 8$  we have seven HR matrices with elements  $0, \pm 1$  [4]. So far HR matrices are applied in electronics [13]: in Space-Time Block Coding (STBC) and orthogonal design [14], also in signal processing [15] and Hamiltonian Neural Nets [16].

If one curve is described by a set of representation points  $\{(x_i, y_i), i = 1, 2, \dots, n\}$  monotonic in coordinates  $x_i$ , then HR matrices combined with the identity matrix  $I_N$  are used to build the orthogonal and discrete Hurwitz - Radon Operator (OHR). For nodes  $(x_1, y_1), (x_2, y_2)$  OHR  $M$  of dimension  $N = 2$  is constructed:

$$B = (x_1 \cdot I_2 + x_2 \cdot A_1)(y_1 \cdot I_2 - y_2 \cdot A_1) = \begin{bmatrix} x_1 & x_2 \\ -x_2 & x_1 \end{bmatrix} \begin{bmatrix} y_1 & -y_2 \\ y_2 & y_1 \end{bmatrix},$$

$$M = \frac{1}{x_1^2 + x_2^2} B,$$

$$M = \frac{1}{x_1^2 + x_2^2} \begin{bmatrix} x_1 y_1 + x_2 y_2 & x_2 y_1 - x_1 y_2 \\ x_1 y_2 - x_2 y_1 & x_1 y_1 + x_2 y_2 \end{bmatrix}. \quad (1)$$

Matrix  $M$  in (1) is found as a solution of equation:

$$\begin{bmatrix} a & b \\ -b & a \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}. \quad (2)$$

For nodes  $(x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4)$ , monotonic in  $x_i$ , OHR of dimension  $N = 4$  is constructed:

$$M = \frac{1}{x_1^2 + x_2^2 + x_3^2 + x_4^2} \begin{bmatrix} u_0 & u_1 & u_2 & u_3 \\ -u_1 & u_0 & -u_3 & u_2 \\ -u_2 & u_3 & u_0 & -u_1 \\ -u_3 & -u_2 & u_1 & u_0 \end{bmatrix} \quad (3)$$

where

$$u_0 = x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4, \quad u_1 = -x_1y_2 + x_2y_1 + x_3y_4 - x_4y_3,$$

$$u_2 = -x_1y_3 - x_2y_4 + x_3y_1 + x_4y_2, \quad u_3 = -x_1y_4 + x_2y_3 - x_3y_2 + x_4y_1.$$

Matrix  $M$  in (3) is found as a solution of equation:

$$\begin{bmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}. \quad (4)$$

For nodes  $(x_1, y_1), (x_2, y_2), \dots, (x_8, y_8)$ , monotonic in  $x_i$ , OHR of dimension  $N = 8$  is built [17] similarly as (1) or (3). Note that OHR operators  $M$  (1)-(3) satisfy the condition of interpolation

$$M \cdot \mathbf{x} = \mathbf{y} \quad (5)$$

for  $\mathbf{x} = (x_1, x_2, \dots, x_N)^T \in \mathbf{R}^N$ ,  $\mathbf{x} \neq \mathbf{0}$ ,  $\mathbf{y} = (y_1, y_2, \dots, y_N)^T \in \mathbf{R}^N$ ,  $N = 2, 4$  or  $8$ .

If one curve is described by a set of nodes  $\{(x_i, y_i), i = 1, 2, \dots, n\}$  monotonic in coordinates  $y_i$ , then HR matrices combined with the identity matrix  $I_N$  are used to build the orthogonal and discrete reverse Hurwitz - Radon Operator (reverse OHR)  $M^1$ . If matrix  $M$  in (1)-(3) is described as:

$$M = \frac{1}{\sum_{i=1}^N x_i^2} (u_0 \cdot I_N + D),$$

where  $D$  with elements  $u_1, \dots, u_{N-1}$ , then reverse OHR  $M^1$  is given by:

$$M^1 = \frac{1}{\sum_{i=1}^N y_i^2} (u_0 \cdot I_N - D). \quad (6)$$

Note that reverse OHR operator (6) satisfies the condition of interpolation

$$M^1 \cdot \mathbf{y} = \mathbf{x} \quad (7)$$

for  $\mathbf{x} = (x_1, x_2, \dots, x_N)^T \in \mathbf{R}^N$ ,  $\mathbf{y} = (y_1, y_2, \dots, y_N)^T \in \mathbf{R}^N$ ,  $\mathbf{y} \neq \mathbf{0}$ ,  $N = 2, 4$  or  $8$ .



### 3.2. Method of Hurwitz-Radon Matrices

Key question looks as follows: how can we compute coordinates of points settled between the interpolation nodes? On a segment of a line every number “ $c$ ” situated between “ $a$ ” and “ $b$ ” is described by a linear (convex) combination  $c = \alpha \cdot a + (1 - \alpha) \cdot b$  for

$$\alpha = \frac{b - c}{b - a} \in [0; 1]. \quad (8)$$

When the nodes are monotonic in coordinates  $x_i$ , the average OHR operator  $M_2$  of dimension  $N = 2, 4$  or  $8$  is constructed as follows:

$$M_2 = \alpha \cdot M_0 + (1 - \alpha) \cdot M_1 \quad (9)$$

with the operator  $M_0$  built (1)-(3) by “odd” nodes  $(x_1, y_1), (x_3, y_3), \dots, (x_{2N-1}, y_{2N-1})$  and  $M_1$  built (1)-(3) by “even” nodes  $(x_2, y_2), (x_4, y_4), \dots, (x_{2N}, y_{2N})$ . Having the operator  $M_2$  for coordinates  $x_i < x_{i+1}$  it is possible to reconstruct the second coordinates of points  $(x, y)$  in terms of the vector  $C$  defined with

$$c_i = \alpha \cdot x_{2i-1} + (1 - \alpha) \cdot x_{2i}, \quad i = 1, 2, \dots, N \quad (10)$$

as  $C = [c_1, c_2, \dots, c_N]^T$ . The required formula is similar to (5):

$$Y(C) = M_2 \cdot C \quad (11)$$

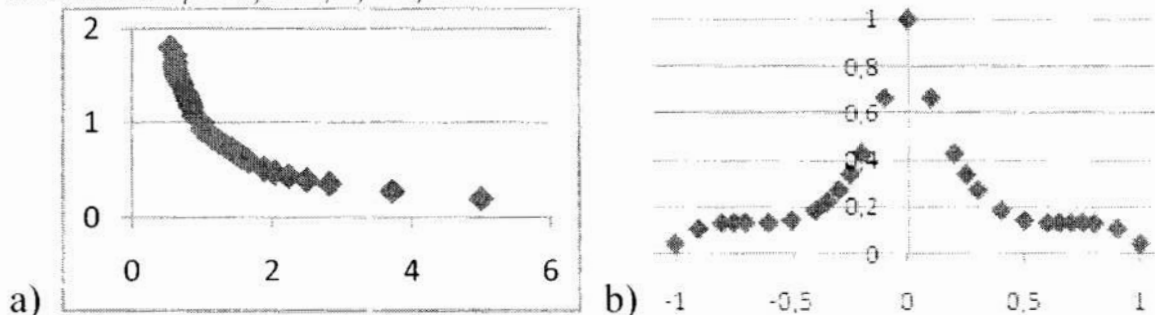
in which components of vector  $Y(C)$  give the second coordinate of the points  $(x, y)$  corresponding to the first coordinate, given in terms of components of the vector  $C$ .

On the other hand, having the operator  $M_2^{-1}$  for coordinates  $y_i < y_{i+1}$  it is possible to reconstruct the first coordinates of points  $(x, y)$ :

$$M_2^{-1} = \alpha \cdot M_0^{-1} + (1 - \alpha) \cdot M_1^{-1}, \quad c_i = \alpha \cdot y_{2i-1} + (1 - \alpha) \cdot y_{2i},$$

$$X(C) = M_2^{-1} \cdot C \quad (12)$$

Contour of the object is constructed with several number of curves. Calculation of unknown coordinates for contour points using (8)-(12) is called by author the method of Hurwitz - Radon Matrices (MHR). Here is the application of MHR method for functions  $f(x) = 1/x$  (nodes as Fig.2) and  $f(x) = 1/(1+25x^2)$  with five nodes equidistance in first coordinate:  $x_i = -1, -0.5, 0, 0.5, 1$ .



**Fig. 3.** Twenty six interpolated points of functions  $f(x)=1/x$  (a) and  $f(x) = 1/(1+25x^2)$  (b) using MHR method with 5 nodes

MHR interpolation for function  $f(x) = 1/x$  gives better result than Lagrange interpolation (Fig.2). The same can be said for function  $f(x) = 1/(1+25x^2)$ .

#### 4. Shape Coefficients

Shape coefficients are the parameters that characterizing and describing the shape of the object. Most of the shape coefficients are calculated using area of the object  $S$  and length of the contour  $L$ . For example [18]:

1. coefficient  $R_S$

$$R_S = \frac{L^2}{4\pi S},$$

2. coefficient of Feret  $R_F$

$$R_F = \frac{L_h}{L_v},$$

3. coefficients  $R_{C1}$  and  $R_{C2}$

$$R_{C1} = 2\sqrt{\frac{S}{\pi}}, \quad R_{C2} = \frac{L}{\pi},$$

4. coefficient of Malinowska  $R_M$

$$R_M = \frac{L}{2\sqrt{\pi S}} - 1,$$

5. coefficient of Blair-Bliss  $R_B$

$$R_B = \frac{S}{\sqrt{2\pi \sum_i r_i^2}},$$

6. coefficient of Danielsson  $R_D$

$$R_D = \frac{S^3}{\left(\sum_i l_i\right)^2},$$

7. coefficient of Haralick  $R_H$

$$R_H = \sqrt{\frac{\left(\sum_i d_i\right)^2}{c \cdot \sum_i d_i^2 - 1}},$$

8. coefficient of compactness  $R_c$

$$R_c = \frac{S}{S_p},$$

9. coefficient  $R_{min}/R_{max}$

$$\sqrt{\frac{R_{min}}{R_{max}}},$$

where:

$R_{min}$  – minimal distance of object contour to center of gravity,

$R_{max}$  – maximal distance of object contour to center of gravity,

$S_p$  – minimal area of the rectangle covering the object,

$L_h$  – maximal horizontal diameter of the object (horizontal Feret's diameter),

$L_v$  – maximal vertical diameter of the object (vertical Feret's diameter),

$r_i$  – distance of object pixel to center of gravity,

$i$  – number of object pixel,

$l_i$  – minimal distance of object pixel to object contour,

$d_i$  – distance of contour pixel to center of gravity,

$c$  – number of contour pixels.

#### 4.1. Length of the contour

The contour is divided into  $m$  curves  $C_1, C_2, \dots, C_m$ . Having nodes  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$  for each  $C_i$  in MHR method [19], it is possible to compute as many curve points as we want for any parameter  $\alpha \in [0; 1]$  (8). Assume that  $k$  is the number of reconstructed points  $p$  together with  $n$  nodes ( $k = n + p$ ). So a curve  $C_i$  consists of  $k$  points that are indexed  $(x_1', y_1'), (x_2', y_2'), \dots, (x_k', y_k')$ , where  $(x_1', y_1') = (x_1, y_1)$  and  $(x_k', y_k') = (x_n, y_n)$ . The length of a curve  $C_i$ , consists of  $k$  points, is estimated:

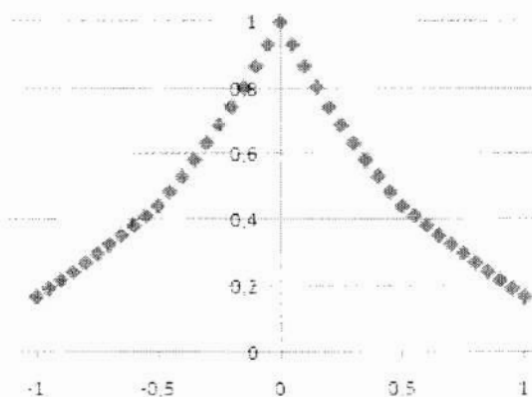
$$d(C_i) = \sum_{i=1}^{k-1} \sqrt{(x_{i+1}' - x_i')^2 + (y_{i+1}' - y_i')^2}. \quad (13)$$

Length of whole contour  $L$  is computed:

$$L = d(C_1) + d(C_2) + \dots + d(C_m). \quad (14)$$

Two examples of estimation a length of the curve via MHR method [20].

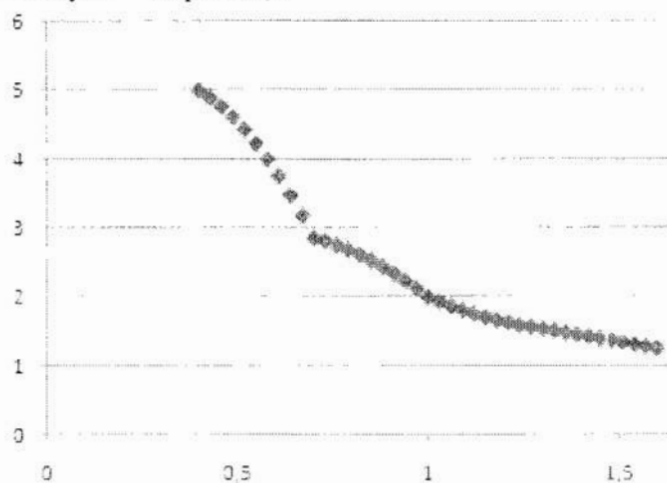
**Example 1** The graph of function  $f(x) = 1/(1+5x^2)$  reconstructed via MHR method (8)-(11) for  $N = 2$  with nodes  $x = -1.0, -0.5, 0, 0.5, 1.0$  ( $n = 5$ ) and calculated points  $p = 36$  looks as follows (the curve is described by  $k = 41$  points):



**Fig. 4.** The curve  $y = 1/(1+5x^2)$  reconstructed by MHR method for five nodes and 36 calculated points

Length of the curve characterized on Fig.4 and estimated by (13) is  $d(C) = 2.643$  whereas precise length is  $d(f) = 2.679$ . There is no Runge phenomenon on Fig.4 and MHR method preserves the symmetry of the curve.

**Example 2** The graph of function  $f(x) = 2/x$  reconstructed via MHR (8)-(11) for  $N = 2$  with nodes  $x = 0.4, 0.7, 1.0, 1.3, 1.6$  ( $n = 5$ ) and calculated points  $p = 36$  looks as follows (the curve is described by  $k = 41$  points):

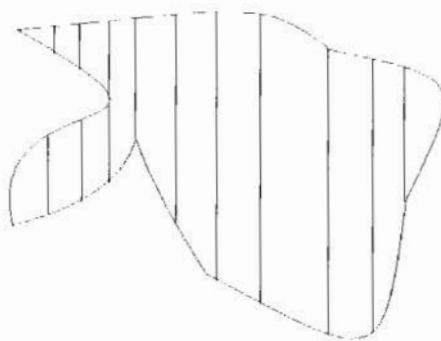


**Fig. 5.** The curve  $y = 2/x$  reconstructed by MHR method for five nodes and 36 calculated points

Length of the curve characterized on Fig.5 and estimated by (13) is  $d(C) = 4.050$  whereas precise length is  $d(f) = 4.045$ .

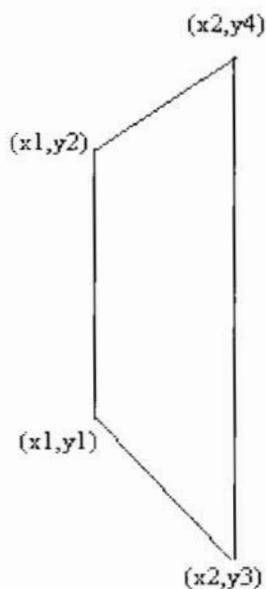
#### 4.2. Area of the object

Area of the object can be divided horizontally or vertically (Fig.6) into the set of / polygons: triangles and quadrangles (squares, rectangles, trapezoids, rhombuses, parallelograms).



**Fig. 6.** The object area consists of polygons

The coordinates of corners for each polygon  $P_i$  are calculated by MHR method [21] and then it is easy to estimate the area of  $P_i$ . For example  $P_1$  as a trapezoid with the corners  $(x_1, y_1)$ ,  $(x_1, y_2)$ ,  $(x_2, y_3)$ ,  $(x_2, y_4)$ :



**Fig. 7.** Trapezoid as a part of the object

Area of a trapezoid  $P_1$  is computed:

$$s(P_1) = \frac{1}{2} |x_2 - x_1| \cdot (|y_2 - y_1| + |y_4 - y_3|).$$

It is easy to compute the area of other polygons with given corners. For example a triangle  $P_2$  with sides  $a$ ,  $b$ ,  $c$  and  $p = (a+b+c)/2$ :

$$s(P_2) = \sqrt{p(p-a)(p-b)(p-c)}$$

and a rhombus  $P_3$  with diagonals  $d_1$  and  $d_2$ :

$$s(P_3) = \frac{1}{2} d_1 \cdot d_2.$$

Estimation of the object area  $S$  is given by a formula:

$$S = \sum_{i=1}^l s(P_i). \quad (15)$$

Feret's diameters (horizontal  $L_h$  and vertical  $L_v$ ) are also possible to calculate having a contour of the object. Contour points, computed by MHR method [22], are applied in shape coefficients.

## 5. Conclusions

The method of Hurwitz-Radon Matrices leads to contour interpolation and shape reconstruction depending on the number and location of contour points. No characteristic features of the curve, significant for classical polynomial interpolations or Bezier curves and NURBS, are important in MHR method. MHR gives the possibility of reconstruction a curve consists of several parts, for example closed curve (contour). The only condition is to have a set of nodes for each part of a curve or contour according to assumptions in MHR method. Any point of the contour can be calculated by MHR method and then parameters of the object used in shape coefficients are computed. Contour representation and curve reconstruction by MHR method is connected with possibility of changing the nodes coordinates and reconstruction of new curve or contour for new set of nodes, no matter what shape of curve or contour is to be reconstructed. Main features of MHR method are: accuracy of shape reconstruction depending on number of nodes and method of choosing nodes; reconstruction of curve consists of  $L$  points is connected with the computational cost of rank  $O(L)$  [19]; MHR method is dealing with local operators: average OHR operators are built by successive 4, 8 or 16 nodes, what is connected with smaller computational costs then using all nodes; MHR is not an affine interpolation [23].

Future works are connected with: geometrical transformations of contour (translations, rotations, scaling)- only nodes are transformed and new curve (for example contour of the object) for new nodes is reconstructed, possibility to apply MHR method to three-dimensional curves and connection MHR method with object recognition [24].

## References

1. Ballard, D.H.: *Computer Vision*. Prentice Hall, New York (1982).
2. Tadeusiewicz, R., Flasiński, M.: *Image Recognition*. PWN, Warsaw (1991).
3. Choraś, R.S.: *Computer Vision*. Exit, Warsaw (2005).
4. Jakóbczak, D., Kosiński, W.: *Application of Hurwitz - Radon Matrices in Monochromatic Medical Images Decompression*. In: Kowalczyk, Z., Wiszniewski, B. (eds.) *Intelligent Data Mining in Diagnostic Purposes: Automatics and Informatics*, pp. 389-398, PWNT, Gdansk (2007).



5. Latecki, L.J., Lakaemper, R.: *Convexity Rule for Shape Decomposition Based on Discrete Contour Evolution*. Computer Vision and Image Understanding 3(73), 441-454 (1999).
6. Cierniak, R.: *Computed Tomography*. Exit, Warsaw (2005).
7. Soussen, C., Mohammad-Djafari, A.: *Polygonal and Polyhedral Contour Reconstruction in Computed Tomography*. IEEE Transactions on Image Processing 11(13), 1507-1523 (2004).
8. Tang, K.: *Geometric Optimization Algorithms in Manufacturing*. Computer – Aided Design & Applications 2(6), 747-757 (2005).
9. Kozera, R.: *Curve Modeling via Interpolation Based on Multidimensional Reduced Data*. Silesian University of Technology Press, Gliwice (2004).
10. Dahlquist, G., Björck, A.: *Numerical Methods*. Prentice Hall, New York (1974).
11. Ralston, A.: *A First Course in Numerical Analysis*. McGraw-Hill Book Company, New York (1965).
12. Eckmann, B.: *Topology, Algebra, Analysis- Relations and Missing Links*. Notices of the American Mathematical Society 5(46), 520-527 (1999).
13. Citko, W., Jakóbczak, D., Sićńko, W.: *On Hurwitz - Radon Matrices Based Signal Processing*. Workshop Signal Processing at Poznan University of Technology (2005).
14. Tarokh, V., Jafarkhani, H., Calderbank, R.: *Space-Time Block Codes from Orthogonal Designs*. IEEE Transactions on Information Theory 5(45), 1456-1467 (1999).
15. Sićńko, W., Citko, W., Wilamowski, B.: *Hamiltonian Neural Nets as a Universal Signal Processor*. 28<sup>th</sup> Annual Conference of the IEEE Industrial Electronics Society IECON (2002).
16. Sićńko, W., Citko, W.: *Hamiltonian Neural Net Based Signal Processing*. The International Conference on Signal and Electronic System ICSES (2002).
17. Jakóbczak, D.: *2D and 3D Image Modeling Using Hurwitz-Radon Matrices*. Polish Journal of Environmental Studies 4A(16), 104-107 (2007).
18. Tadeusiewicz, R., Korohoda, P.: *Computer Analysis and Image Processing*. FPT, Cracow (1997).
19. Jakóbczak, D.: *Curve Interpolation Using Hurwitz-Radon Matrices*. Polish Journal of Environmental Studies 3B(18), 126-130 (2009).
20. Jakóbczak, D.: *Shape Representation and Shape Coefficients via Method of Hurwitz-Radon Matrices*. Lecture Notes in Computer Science 6374 (Computer Vision and Graphics: Proc. ICCVG 2010, Part I), Springer-Verlag Berlin Heidelberg, 411-419 (2010).
21. Jakóbczak, D.: *Application of Hurwitz-Radon Matrices in Shape Representation*. In: Banaszak, Z., Świć, A. (eds.) Applied Computer Science: Modeling of Production Processes 1(6), pp. 63-74. Lublin University of Technology Press, Lublin (2010).

22. Jakóbczak, D.: *Object Modeling Using Method of Hurwitz-Radon Matrices of Rank  $k$* . In: Wolski, W., Borawski, M. (eds.) *Computer Graphics: Selected Issues*, pp. 79-90. University of Szczecin Press, Szczecin (2010).
23. Jakóbczak, D.: *Implementation of Hurwitz-Radon Matrices in Shape Representation*. In: Choraś, R.S. (ed.) *Advances in Intelligent and Soft Computing 84, Image Processing and Communications: Challenges 2*, pp. 39-50. Springer-Verlag, Berlin Heidelberg (2010).
24. Jakóbczak, D.: *Object Recognition via Contour Points Reconstruction Using Hurwitz-Radon Matrices*. In: Józefczyk, J., Orski, D. (eds.) *Knowledge-Based Intelligent System Advancements: Systemic and Cybernetic Approaches*, pp. 87-107. IGI Global, Hershey PA, USA (2011).

## Abstract

Computer vision needs suitable methods of shape representation and contour reconstruction. Method of Hurwitz-Radon Matrices (MHR), invented and described by the author, is applied in reconstruction and interpolation of curves in the plane. Reconstructed curves represent the shape and contour of the object. Any point of the contour can be calculated by MHR method and then parameters of the object, used in shape coefficients, are computed: length of the contour, area of the object, Feret's diameters. Proposed method is based on a family of Hurwitz-Radon (HR) matrices. The matrices are skew-symmetric and possess columns composed of orthogonal vectors. The operator of Hurwitz-Radon (OHR), built from these matrices, is described. The shape is represented by the set of nodes. It is shown how to create the orthogonal and discrete OHR and how to use it in a process of shape representation and reconstruction. MHR method is interpolating the curve point by point without using any formula or function.

## Streszczenie

Komputerowa wizja wymaga odpowiednich metod reprezentacji kształtu obiektu i rekonstrukcji jego konturu. Jedną z takich metod, opracowaną i nazwaną przez autora metodą Macierzy Hurwitza-Radona (MHR), może zostać użyta w interpolacji i rekonstrukcji krzywych płaskich. Odtworzone krzywe przedstawiają kształt i kontur obiektu. Dzięki metodzie MHR możliwe jest wyznaczenie dowolnego punktu konturu i obliczenie parametrów używanych we współczynnikach kształtu: długość konturu, powierzchnia obiektu, średnice Fereta. Metoda ta jest oparta na rodzinie macierzy Hurwitza-Radona (HR). Macierze HR są skośnie-symetryczne i składają się z kolumn tworzących ortogonalne wektory. W pracy pokazano jak konstruować Operator Hurwitza-Radona (OHR) oraz jak wykorzystać go w procesie interpolacji konturu. Kształt obiektu opisany jest za pomocą punktów węzłowych. Metoda MHR rekonstruuje kontur i kształt obiektu punkt po punkcie, bez użycia wzoru opisującego krzywą.



**Leszek Bychto**  
Katedra Podstaw Elektroniki  
Wydział Elektroniki i Informatyki  
Politechnika Koszalińska

## Układ monitorowania systemu fotowoltaicznego

**Słowa kluczowe:** fotowoltaika, monitorowanie, MAX4376, PolySun

### 1. Budowa systemu

Prezentowany system fotowoltaiczny ma strukturę autonomiczną. W skład systemu wchodzi panel fotowoltaiczny zbudowany z dwóch modułów SF50, akumulator kwasowy VARTA Hobby 12V, 90Ah, przetwornica napięcia 12VDC-230AC 150W; regulator systemu RSS-02 oraz wykonany w ramach inżynierskiej pracy dyplomowej układ pomiaru parametrów pracy.

**Panel fotowoltaiczny** – każdy z modułów zawiera 36 ogniw. Całkowita moc modułu wynosi 50 Wp(+/-5%). Napięcie  $V_{oc}=21,1$  V a  $I_{sc}=3,1$  A. Optymalny punkt pracy to 17,1 V i 2,9 A. Podane parametry odnoszą się do warunków oświetlenia AM1,5 i 1000 W/m<sup>2</sup>. Moduły podłączono równolegle. Panel fotowoltaiczny zamocowano na dachu budynku D kampusu Politechniki Koszalińskiej przy ul. Śniadeckich w kierunku wschodnim (takie usytuowanie uwarunkowane było układem mocującym). Kąt między płaszczyzną dachu a panelem wynosi 45°.

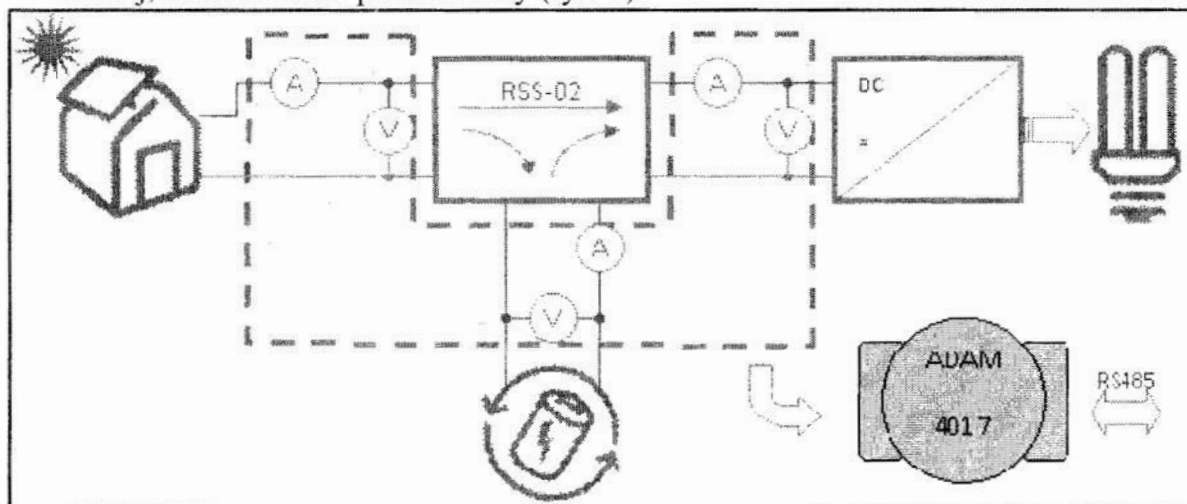
**Regulator systemu** – regulator RSS-02 wykonano w Instytucie Technologii Elektronowej w Krakowie. Do zadań regulatora należy:

- nadzór nad procesem ładowania akumulatora i niedopuszczenie do jego przeładowania;
- odłączenie obciążenia systemu (przetwornicy) z chwilą osiągnięcia minimalnego, dopuszczalnego poziomu naładowania akumulatora. Zapobiega to procesowi pełnego rozładowania i zasiarczenia;
- załączenie obciążenia w momencie osiągnięcia przez akumulator określonego poziomu naładowania.

Regulator może pracować z akumulatorami 12 V lub 24 V. Maksymalny prąd ogniwa i odbiornika wynosi 15 A, maksymalny pobór prądu przez regulator to 5 mA. Regulator dba o dostarczanie energii z baterii słonecznej do obciążenia. Jeżeli bateria jest w stanie dostarczać więcej energii niż potrzebuje obciążenie to nadwyżka gromadzona jest w akumulatorze. Jeżeli napięcie ładowania przekroczy 13,7 V $\pm$ 0,2 V (27,4 V $\pm$ 0,4 V)

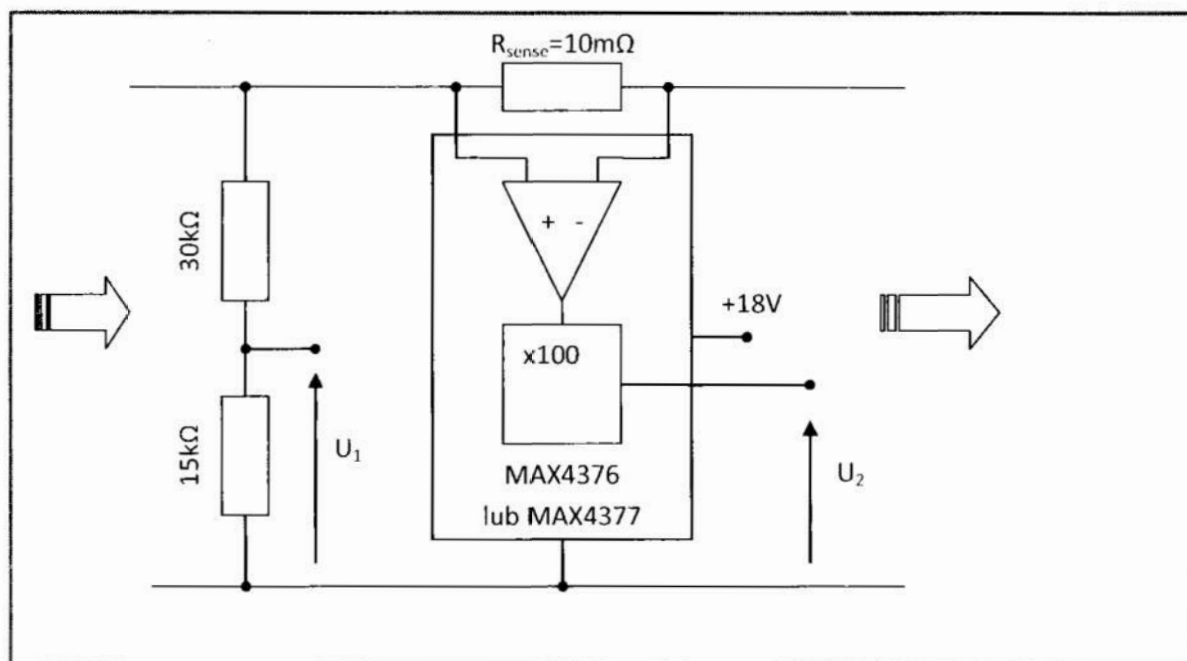
to regulator przerywa proces ładowania akumulatora. Jeżeli warunki słoneczne nie pozwalają na dostarczenie wystarczającej energii do obciążenia to regulator pobiera brakującą część z akumulatora. Pobieranie energii z akumulatora może trwać do chwili osiągnięcia przez akumulator napięcia  $11,1 \text{ V} \pm 0,2 \text{ V}$  ( $22,2 \text{ V} \pm 0,4 \text{ V}$ ). Po osiągnięciu przez akumulator tego poziomu następuje odłączenie obciążenia od systemu. Zapobiega to całkowitemu rozładowaniu akumulatora. Ten stan trwa do momentu osiągnięcia przez akumulator napięcia włączenia obciążenia  $12,6 \text{ V} \pm 0,2 \text{ V}$  ( $25,2 \text{ V} \pm 0,4 \text{ V}$ ). W przypadku długotrwałego wyłączenia obciążenia akumulator jest stale doładowywany tak by utrzymywać poziom naładowania  $13,7 \text{ V} \pm 0,2 \text{ V}$  ( $27,4 \text{ V} \pm 0,4 \text{ V}$ ).

**Układ monitorowania.** Zasadniczo wyróżnia się trzy rodzaje systemów monitorowania: globalny, analityczny i naukowy [3]. W przypadku prezentowanego systemu PV wybrano najprostszy, należący do grupy systemów monitorowania globalnego. W jego skład wchodzi: moduł pomiarowy ADAM4017 firmy Advantech, 3 dzielniki napięcia i 3 układy zamiany prądu na napięcie. Punkty pomiarowe włączono w obwody baterii słonecznej, akumulatora i przetwornicy (rys. 1).



Rys. 1. Schemat blokowy systemu fotowoltaicznego

Dzielniki napięcia służą dopasowaniu napięć do poziomu akceptowalnego przez wejścia karty pomiarowej ( $\pm 10 \text{ V}$ ). Stosunek podziału wynosi 1:3 (rezystory  $30 \text{ k}\Omega$  i  $15 \text{ k}\Omega$ ) (rys. 2). Pomiar prądu zrealizowano poprzez pomiar spadku napięcia na rezystorze pomiarowym  $R_{\text{sense}} = 10 \text{ m}\Omega$ . Spadek napięcia mierzony jest za pomocą wzmacniaczy pomiarowych wbudowanych w układy scalone firmy MAXIM modele MAX4376 i MAX4377 [2]. Są to układy przeznaczone do kontroli prądu w blokach zasilania. Sygnał napięciowy proporcjonalny do mierzonego prądu jest dodatkowo wzmacniany przez te układy 20-krotnie, 50-krotnie i 100-krotnie w zależności od wykonania. W pracy zastosowano układ o wzmacnieniu 100x co oznacza, że maksymalny, możliwy do zmierzenia prąd wynosi  $10 \text{ V} / (100 \times 10 \text{ m}\Omega) = 10 \text{ A}$ . Dość dużą zaletą tych układów jest szeroki zakres napięć zasilających (+3 V do 28 V) co pozwoliło wykorzystać do zasilania zasilacz modułu ADAM4017 (+18 V).



Rys. 2. Schemat jednego z 3 układów pomiaru napięcia i prądu

Niestety układy te są w stanie mierzyć prąd płynący w jednym kierunku dlatego do pomiaru prądu akumulatora wykorzystano podwójny układ MAX4377. Jeden ze wzmacniaczy mierzy prąd ładowania dając napięcie większe od 0 V, gdy płynie prąd rozładowania daje napięcie 0 V. Drugi wzmacniacz pomiarowy włączono tak by mierzył prąd rozładowania.

Ostatecznie zajęto 7 wejść analogowych modułu pomiarowego ADAM4017. Ósme, niewykorzystane wejście podłączono do masy tak by nie generowało sygnałów zakłócających.

Całość wraz z bezpiecznikami zamontowano w szafce elektrycznej. Moduł ADAM4017 podłączono kablem 4-żyłowym z konwerterem RS485<->RS232 (ADAM 4520). Przejście z interfejsu RS485 na RS232 umożliwiło podłączenie systemu pomiarowego do komputera PC. Więcej szczegółów dot. budowy prezentowanego systemu solarnego znajduje się w pracy [1].





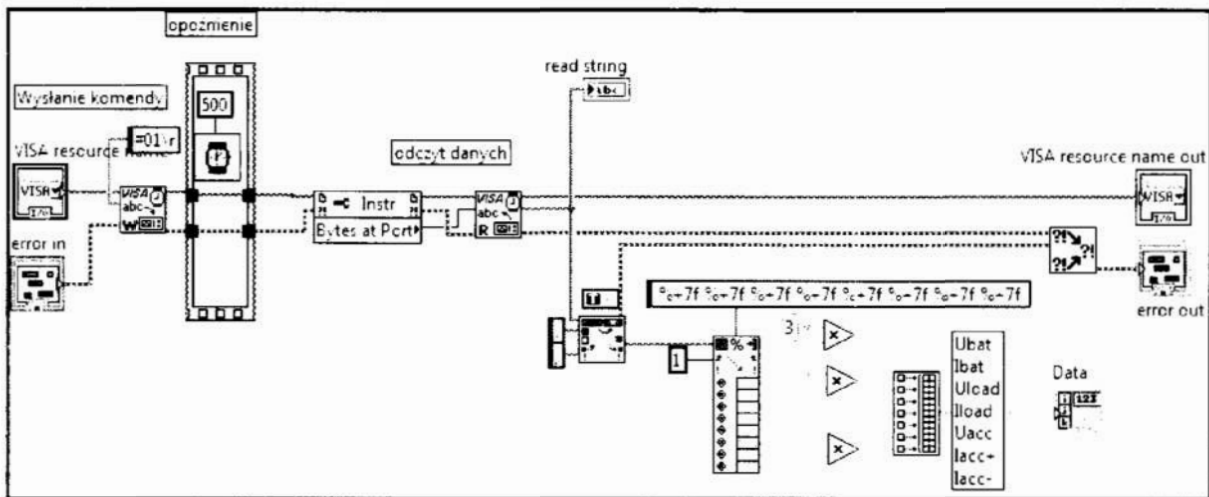
Rys. 3. Widok systemu fotowoltaicznego

W czasie eksploatacji systemu podłączono do przetwornicy żarówkę 50 W, włączaną okresowo przez programator dobowy. Dzięki temu akumulator był w stanie częściowego rozładowania a system mógł, w ciągu dnia, pracować w trybie gromadzenia energii.

## 2. Oprogramowanie

Do komunikacji z systemem pomiarowym napisano aplikację w środowisku NI LabVIEW. Za odczyt danych z modułu pomiarowego i konwersję danych z postaci tekstowej na liczbową odpowiada procedura *'Akwizytor.vi'*. Uruchomienie tej procedury powoduje wysłanie komendy `'#01\r'`. Jest to komenda nakazująca urządzeniu nr 1 podanie napięć występujących na wszystkich wejściach pomiarowych. W odpowiedzi otrzymujemy łańcuch tekstowy zaczynający się znakiem `'>'` i zawierający 8 liczb oddzielonych znakiem `'+'` lub `'-'`.

Na przykład: `„>+12.012+2.800+11.603+2.005+0.005+12.634+0.800+0.001”`.



Rys. 4. Kod procedury realizującej odczyt danych pomiarowych – 'Akwizytor.vi'.

Po odczytaniu łańcucha tekstowego funkcją 'VISA Read' następuje zamiana separatorów liczb dziesiętnych z kropek na przecinki tak aby funkcja 'Scan From String' mogła zamienić 8 liczb występujących w łańcuchu na zmienne numeryczne typu float. Odczytane napięcia są 3 razy mniejsze od rzeczywistych dlatego funkcja dba o przywrócenie ich właściwych wartości. Ostatecznie procedura tworzy wektor zawierający 7 liczb odpowiadających napięciom i prądom w punktach pomiarowych systemu solarnego. Procedura 'Akwizytor.vi' wywoływana jest co 5 sekund a odczytane dane wpadają do tymczasowego bufora. Jeżeli pomiar czasu wykaże przejście do nowej minuty to dane z bufora podlegają uśrednieniu. Następnie do uśrednionych danych dodawany jest znacznik czasu. Wartość minut znacznika czasu odpowiada minucie dla której dokonano pomiarów a wartość sekund ustawiana jest na 30. Utworzony rekord zawiera 8 pól: znacznik czasu (YYYY-MM-DD/hh:mm:30); napięcie fotoogniwa Vbat; prąd fotoogniwa Ibat; napięcie obciążenia Uload; prąd obciążenia Iload; napięcie akumulatora Uacc; prąd ładowania akumulatora Iacc+; prąd rozładowania akumulatora Iacc-. Następnie program wpisuje nowy rekord do pliku. Jeżeli jest to pierwszy rekord danego dnia to zostaje utworzony nowy plik o nazwie zawierającej aktualną datę (YYYY-MM-DD.txt). Pełny plik z danymi zawiera  $60 \times 24 = 1440$  rekordów. Wszystkie pliki zapisywane są do jednego, podanego przez użytkownika, katalogu.

### 3. Analiza wyników

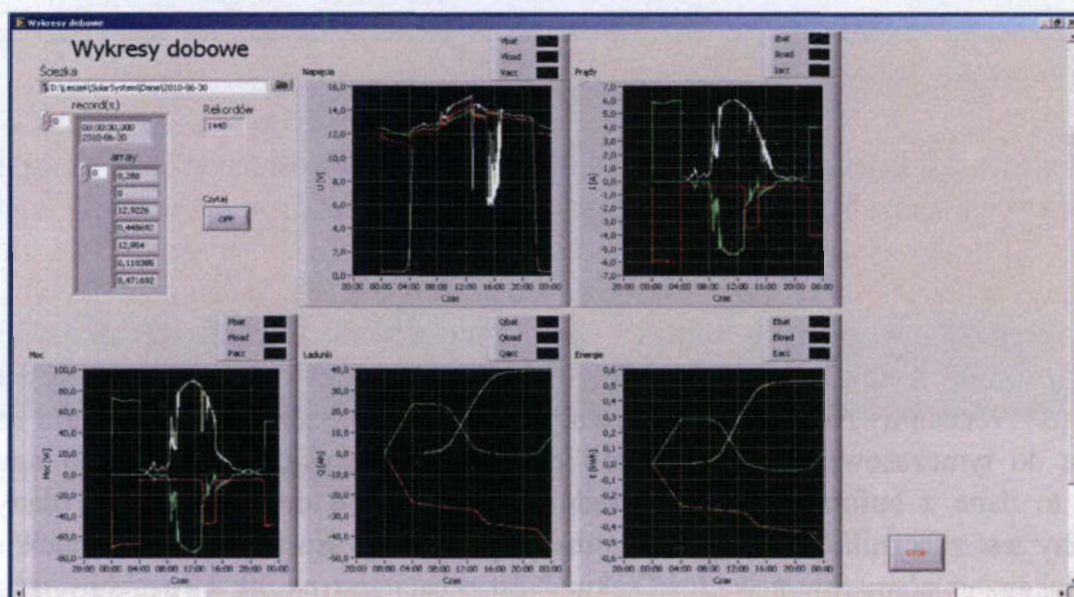
Do analizy danych zawartych w plikach dobowych przygotowano program prezentujący wyniki pomiarów w postaci wykresów. Wyświetlane są wykresy napięć, prądów, mocy, energii i ładunku (rys. 5). Prąd płynący przez obciążenie oraz prąd ładowania akumulatora mają wartości ujemne tak by spełnione było równanie bilansu prądów:

Inne parametry, uzależnione od tych prądów mają również wartości ujemne. I tak, odbiorniki energii posiadają wartości ujemne mocy a źródła energii posiadają moce dodatnie. Napięcia, prądy i moce wyrażone są ich wartościami średnimi za okres 1 minuty natomiast ładunki i energie są przyrostami tych wartości wg. zależności:

$$Q(t)[Ah] = Q(t - 1)[Ah] + I(t)[A]\Delta t[h]$$

$$E(t)[Wh] = E(t - 1)[Wh] + I(t)[A] \cdot U[t][V] \cdot \Delta t[h]$$

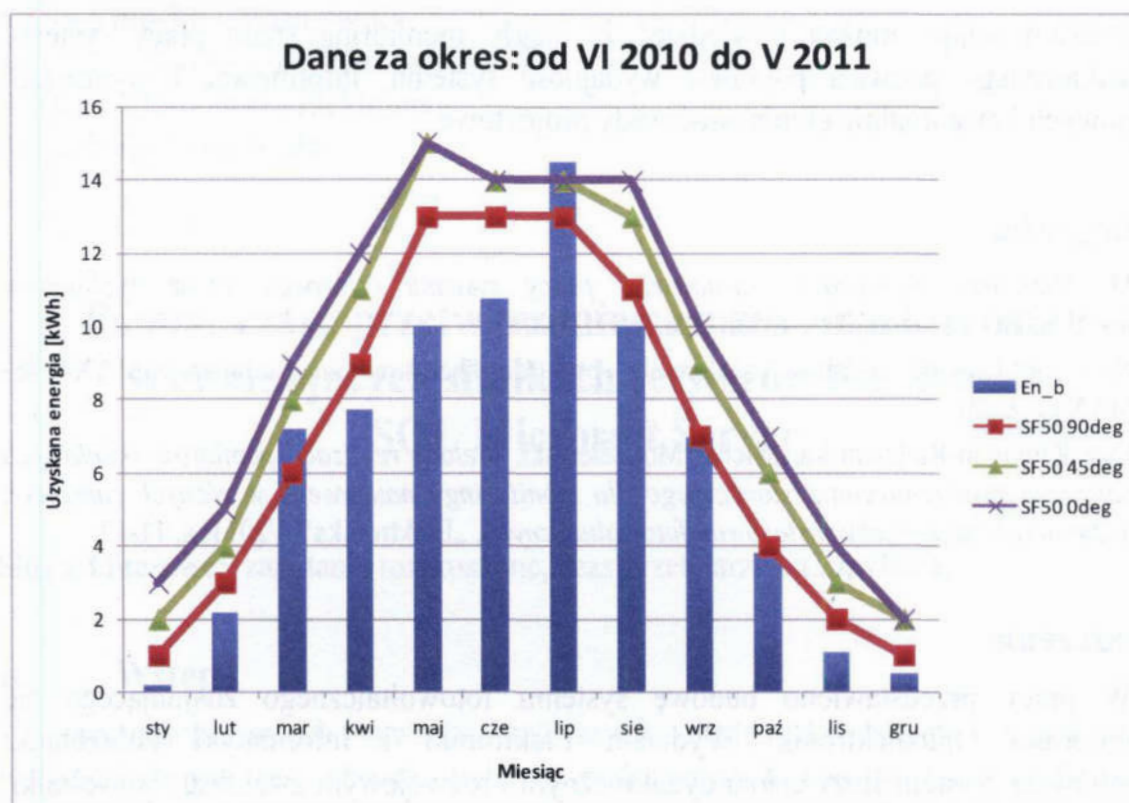
Obliczanie przyrostów ładunków i energii rozpoczyna się od wartości zerowych tzn.  $Q(0) = 0$  i  $E(0) = 0$ . W przyszłości planuje się wykorzystanie sum końcowych z poprzedniego dnia jako wartości startowych.



Rys. 5. Panel użytkownika programu 'Wykresy dobowe'

Na rysunku 6 zaprezentowano miesięczne zestawienie pozyskanej energii z fotowoltaiki za okres od czerwca 2010 do maja 2011 ( $En_b$ ). Na wykres naniesiono również wartość teoretyczną uzyskanej energii otrzymanej w wyniku symulacji modułów SF50 w środowisku 'PolySun'. Symulacji dokonano dla trzech położenia modułu: Wschód ( $90^\circ$ ), Południowy-Wschód ( $45^\circ$ ), Południe ( $0^\circ$ ).





**Rys. 6.** Porównanie rzeczywistej energii uzyskanej z modułu fotowoltaicznego ( $E_{n_b}$ ) z uzyskaną energią otrzymaną w wyniku symulacji programem 'Polysun' dla 3 położenia modułów – Wschód ( $90^\circ$ ), Południowy-Wschód ( $45^\circ$ ), Południe ( $0^\circ$ )

Porównanie wyników rzeczywistych z wynikami symulacji pokazało ich dość dobrą zgodność dla dwóch miesięcy: września i października. W lipcu pozyskano więcej energii niż by to wynikało z symulacji. Wyniki z sierpnia są zaniżone z powodu braku części danych, spowodowanego awariami systemu monitorowania. Energia pozyskana w pozostałych miesiącach dużo niższa niż to przewiduje symulacja.

#### 4. Podsumowanie

W pracy przedstawiono budowę systemu fotowoltaicznego, niewielkiej mocy, przeznaczonego do zadań dydaktycznych i badawczych. Moduł akwizycji danych pomiarowych wzbogacił system o możliwość rejestracji rzeczywistych parametrów pracy systemu i umożliwił uzyskanie bardzo cennych danych na temat monitorowanego systemu.

Otrzymane dane mogą być traktowane jako podstawa do optymalizacji typowego systemu PV pod kątem zwiększenia jego wydajności oraz do oceny zasobów energii słonecznej w obszarze Pomorza Środkowego.

Podsumowując, można powiedzieć iż ciągły monitoring stanu pracy systemu fotowoltaicznego pozwala poprawić wydajność systemu, informować o sytuacjach alarmowych i ewentualnie eliminować wady projektowe.

## Bibliografia

1. M. Marciniak, *Rejestracja parametrów pracy systemu solarnego*, Praca dyplomowa inżynierska - Politechnika Koszalińska. WEiI, 2005.
2. Noty aplikacyjne układów scalonych: <http://datasheets.maxim-ic.com/en/ds/MAX4376-MAX4378.pdf>
3. Ewa Klugman-Radziemska, Michał Modzelewski, *Metody realizacji pomiarów składowych natężenia promieniowania słonecznego dla monitoringu naukowego w różnych zakresach widmowych na potrzeby zastosowań fotowoltaicznych*. „Elektronika” 4/2011, s. 11-17.

## Streszczenie

W pracy przedstawiono budowę systemu fotowoltaicznego znajdującego się w Laboratorium Optoelektroniki Wydziału Elektroniki i Informatyki Politechniki Koszalińskiej. System służy celom dydaktycznym i rozwojowym z zakresu fotowoltaiki. System wzbogacono w układ monitorowania stanu pracy oraz oprogramowanie do gromadzenia danych pomiarowych. W pracy zaprezentowano również przykładowe aplikacje wykorzystane do analizy danych eksperymentalnych. Rzeczywisty zysk energetyczny porównano z zyskiem energetycznym otrzymanym w wyniku symulacji za pomocą programu ‘Polysun’.

## Abstract

The paper presents the photovoltaic system placed in The Laboratory of Optoelectronic, Division of Electronics and Computer Sciences, Technical University of Koszalin. The system is used for didactics and research tasks on the field of photovoltaics. The monitoring circuit and the application for collecting data have been added to the system. The examples of the applications used for analysis of the experimental data are also presented. The real energetic yield was compared with energetic yield obtained from simulation by ‘PolySun’ application.

**Alicja Kilińska – Wiśniewska**  
Katedra Inżynierii Komputerowej  
Wydział Informatyki i Elektroniki  
Politechnika Koszalińska

## **Pomiar czasu przetwarzania rozproszonych zapytań w ewoluujących silnikach serwerów baz danych SQL Microsoft Server**

**Słowa kluczowe:** zapytanie rozproszone, czas przetwarzania zapytania.

### **1. Wstęp**

Organizacja, przechowywanie oraz szybki dostęp do danych umieszczonych w bazach danych jest obecnie jedną z podstawowych idei rozwoju współczesnej informatyki. Wraz ze zwiększonymi wymaganiami systemowymi wzrasta zapotrzebowanie na przechowywanie danych które te systemy generują lub wykorzystują. Dane wzrastają w tempie wykładniczym, co doskonale widać na przykładzie nieustannie rosnących pojemności dysków twardych, które w latach 90. osiągały pojemności rzędu 4,2 Gb, natomiast obecnie dochodzą one do rozmiarów 1 Tb.

Jednocześnie wymaga się coraz szybszego bardziej elastycznego oraz nieprzerwanego dostępu do danych przechowywanych w bazach danych. Dynamiczny rozwój infrastruktury sieciowej, coraz większe możliwości szybkiego przesyłania informacji, współdzielenie i udostępnianie danych spowodowało, że coraz większą rolę zaczęły odgrywać rozproszone bazy danych. Doskonałym przykładem zastosowania rozproszonych baz danych na skalę światową okazała się akcja SETI, która wykorzystywała zasoby komputerów użytkowników rozsianych na całym świecie za pośrednictwem sieci Internet. Była ogromnym sukcesem, aczkolwiek pokazała jak istotny jest czas dostępu do danych rozsianych po całym świecie. Obecnie pracuje się nad przyspieszeniem oraz zabezpieczeniem dostępu do danych przechowywanych w rozproszonych bazach danych.

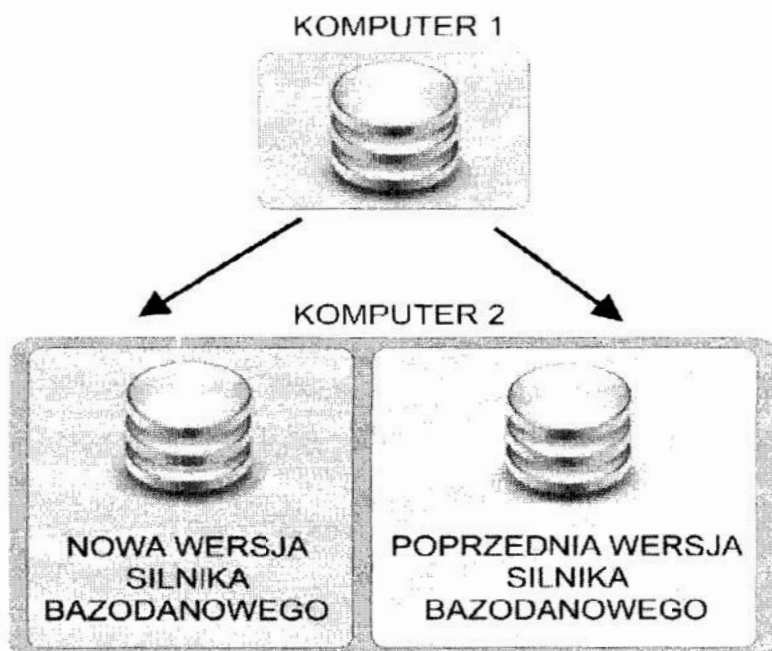
### **2. Motywacja**

Powszechność rozproszonych baz danych sprawia, iż ilość danych przechowywanych w bazach tego typu rośnie bardzo szybko. Wymusiło to na producentach oprogramowania, a dokładniej silników bazodanowych ciągłe modyfikacje istniejących wersji serwerów bazodanowych oraz zbyt częste publikowanie nowych wersji



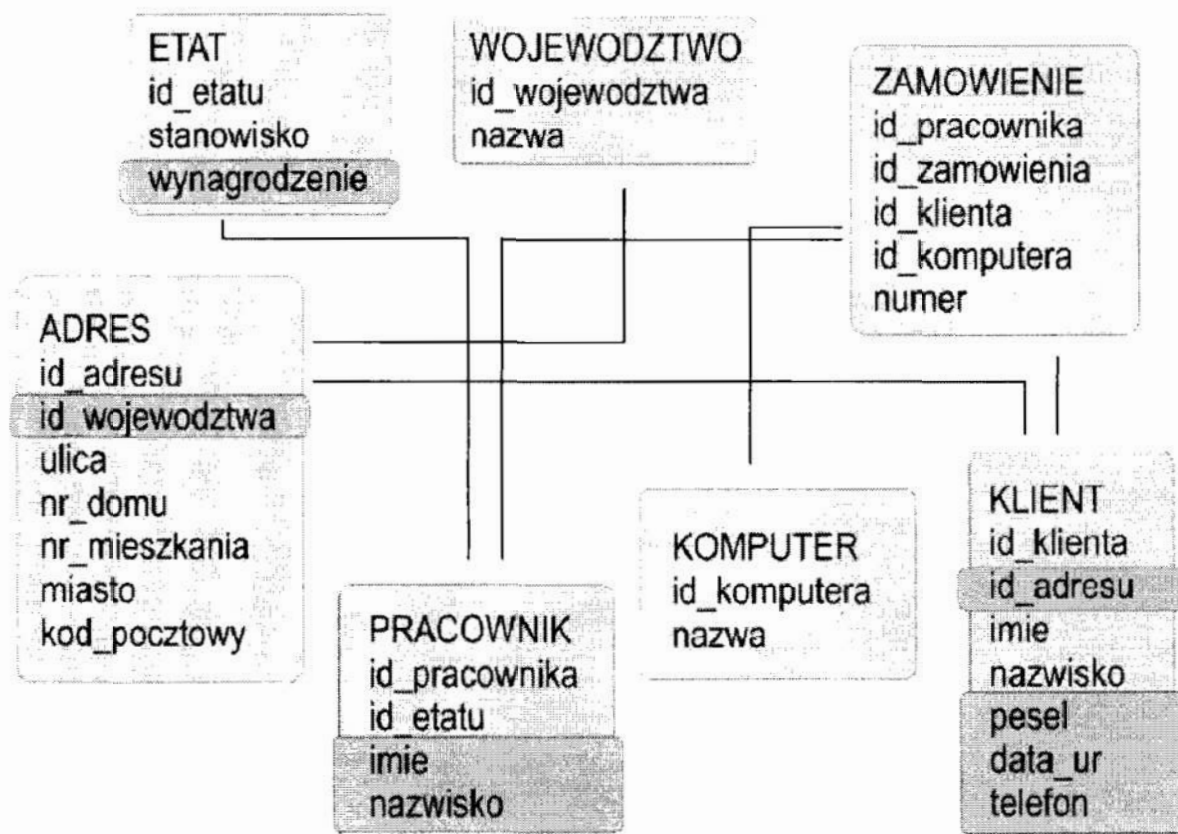
oprogramowania. Korporacje producenckie upatrują w tym łatwy i szybki zysk, gdyż powszechnie przyjętą zasadą jest zakupywanie nowych wersji oprogramowania, które pojawiają się na rynku. Jednakże istnieje możliwość, iż rozproszona baza danych znajdująca się na kilku serwerach różniących się pomiędzy sobą wersjami silnika nie będzie funkcjonowała spójnie, co doprowadzi do spowolnienia działania bazy danych oraz obniży wydajność samej rozproszonej bazy danych.

Rozpatrzmy zatem następujący przykład rozproszonej bazy danych która została zaimplementowana na dwóch silnikach bazodanowych jednego producenta oprogramowania w wersji starszej oraz nowszej. Wspomniane dwie wersje oprogramowania zostały zainstalowane na komputerze położonych w sieci, miało to na celu wykluczenie czynnika różnorodności podzespołów zainstalowanych w samym komputerze, które mogłyby wpłynąć na proces badania. Baza została rozproszona pomiędzy dwoma komputerami co zostało przedstawione na rysunku 1.



**Rys. 1.** Schemat rozproszenia badanej bazy danych

Umieszczony na rysunku 1 schemat przedstawia rozproszenie badanej bazy danych pomiędzy dwoma komputerami umieszczonymi w sieci, które posłużyły do przeprowadzenia badania. Warto podkreślić, że komputery znajdowały się w jednej sieci, która została stworzona specjalnie na potrzeby przeprowadzenia badania, aby wykluczyć czynniki zewnętrzne które mogły być mieć wpływ na wynik przeprowadzonego badania. Baza danych rozstała rozproszona zgodnie ze schematem zamieszczonym na rysunku 2.



Rys. 2. Schemat badanej, rozproszonej bazy danych

Umieszczony na rysunku 2 schemat pokazuje reprezentację głównych elementów opisywanej bazy danych: *etat* – *województwo* – *zamówienie* – *adres* – *pracownik* – *komputer* – *klient*. Ten schemat został przedstawiony jako model fizyczny. Baza została podzielona pomiędzy dwoma silnikami bazodanowymi znajdującymi się na dwóch komputerach. Część bazy danych oznaczona na kolor fioletowy została umieszczona na komputerze numer 1, natomiast część oznaczona na kolor zielony została umieszczona na komputerze numer 2.

Aby uzyskać pełne dane klienta zamawiającego pewien produkt należy odwołać się do tabeli *klient* która została rozproszona.

### 3. Mechanizm przetwarzania zapytań w bazach danych

Istotnym elementem na który należy zwrócić uwagę jest schemat przetwarzania zapytania przez wykorzystane silniki bazodanowe. W tym przypadku zostanie do tego wykorzystana instrukcja SELECT języka SQL. Zapytania pod względem przetwarzania można podzielić na zapytania „ad hoc” oraz składowane, które dodatkowo można podzielić na etapy w których są wykonywane. Zapytanie „ad hoc” składa się z etapów: poprawność syntaktyczna, poprawność semantyczna, optymalizacja zapytania, kompilacja i wykonanie zapytania. Natomiast zapytanie składowe określane jest jako kolejne

wywołanie tego samego zapytania. Tego typu zapytania wykonywane są jedynie w dwóch etapach składających się z pobrania skompilowanej instrukcji oraz jej wykonania.

Instrukcja SELECT została wykorzystana, aby przy jej pomocy dokonać pomiarów czasu jej wykonania na rozproszonej bazie danych. Sama instrukcja służy do pobrania pożądaných danych z bazy. Dodatkowo została wykorzystana również instrukcja FROM, która definiuje z której tabeli lub innego elementu takiego jak widok mają zostać pobrane dane.

#### 4. Pomiar czasu statystycznego

Celem dokonania pomiaru czasu statystycznego było udowodnienie założonej tezy iż starszy silnik bazodanowy znacznie wolniej przetwarza zapytanie rozproszone niż nowszy, co miało ogólny wpływ na szybkość otrzymywania wymaganych wyników. Do tego celu został wykorzystany wyżej opisany model rozproszonej bazy danych rozmieszczonej pomiędzy komputerami w sieci. W ciągu modelowania mierzony był czas statystyczny realizacji podawanych przez serwer zapytań. Pomiar obejmuje czas analizy, kompilacji i realizacji transakcji, które występują po stronie serwera. Otrzymane wartości nie uwzględniają czasu potrzebnego na przesyłanie danych do klienta, co wyklucza błąd pomiaru wynikający z obciążenia sieci komputerowej.

Podczas badania uwzględnione zostały dwie wartości czasowe. Pierwszy wynik dotyczył czasu parsowania w czasie kompilacji zapytania, natomiast drugi wynik był czasem analizy i opracowanie odnoszącym się do czasu opracowania całego zapytania. Obie wartości określają jak długo były wykonywane następujące czynności:

- przetwarzania polecenia SQL, sprawdzanie błędów składni, łamiąc się do polecenia części produkcji wewnętrznej drzewo wykonania,
- kompilacja planu wykonania w pamięci cache od drzewa realizacji, które właśnie zostały wyprodukowane.

Należy jednak nadmienić, że obejmuje to czas, który jest wymagany do przeniesienia danych z tabel powiązanych w pamięci podręcznej, dlatego jeśli wykonane będzie to samo zapytanie dwa razy, a dane z tabeli będą nadal znajdowały się w pamięci podręcznej, to czas parsowania i czas kompilacji będą znacznie niższe podczas drugiego wykonania zapytania.

SQL Server parse and compile time:

CPU time = 0 ms, elapsed time = 0 ms.

Zajętość czasu procesora 0 ms i 0 ms z upływem czasu na zakończenie przetwarzania i kompilacji. Upływający czas obejmuje (CPUTIME, IO Time itp.).

SQL Server Execution Times:

CPU time = 0 ms, elapsed time = 0 ms.

To całkowita ilość czasu ile zajęło wykonanie zapytania. Trwało to 0 ms czasu procesora i 0 ms z upływającego czasu, co oznacza 0 ms na przetwarzanie IO.

## 5. Analiza otrzymanych wyników

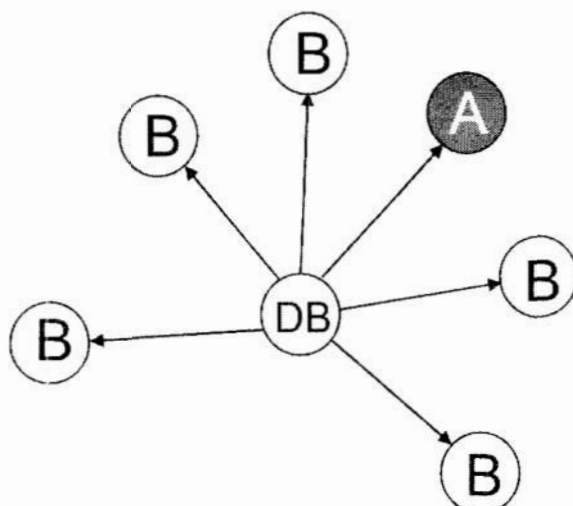
Na podstawie przeprowadzonych badań można jasno poprzeć tezę, iż silniki baz danych różniące się tylko jedną wersją wstecz działają o około 3,75 ms wolniej przy uwzględnieniu wszystkich czasów pomiaru statystycznego.

Średni czas zajętości procesora podczas parsowania zapytania wynosił średnio 0ms dla baz z nowym silnikiem i 3,85 ms dla baz ze starszym silnikiem. Natomiast czas średni czas wykonania IO rozproszonego zapytania wynosił 4,12 ms dla baz z nowym silnikiem i 9,5 ms dla wersji starszej, co za tym idzie można łatwo wysnuć wniosek, że kompilacja przebiega o 5,37 ms szybciej przy użyciu nowego silnika bazodanowego.

Całkowite wykonanie zapytania natomiast zajęło procesorowi średnio 1,87 ms przy nowym silniku, przy starym trwało to średnio 3,87 ms, czyli aż o 2 ms dłużej. Jednakże najistotniejszy był czas wykonania i kompilacji całego zapytania rozproszonego, gdyż tutaj również do porównania został zmierzony czas na lokalnym serwerze. Z pomiarów wynika iż tylko dla kompilacji całego zapytania czas ten jest większy od 0 ms. Pomiar wykazały również opóźnienie względem starszej wersji silnika, gdyż otrzymane wartości wynosiły odpowiednio: dla nowego silnika 23,25 ms, dla wersji poprzedniej 27,12 ms, a dla zapytania wykonanego stacjonarnie 5,12 ms. Różnica pomiędzy wersjami silnika wyniosła aż 3,87 ms, wartość ta oczywiście wraz z rozrostem bazy danych najprawdopodobniej ulegnie zwiększeniu.

Rysunek 3 przedstawia przykładowy graf rozproszenia bazy danych. Wierzchołek DB traktowany jest głównym serwer zarządzający całą bazą, natomiast poszczególne węzły to części rozproszone bazy danych na innych serwerach. Jako zaciemniony wierzchołek na grafie został przedstawiony serwer ze starszym silnikiem bazodanowym oznaczona literą A, natomiast pozostałe silniki są w wersji nowszej oznaczone literą B.

W tym przykładzie uwzględniony zostanie całkowity czas kompilacji zapytania rozproszonego. Zgodnie z przeprowadzonymi badaniami, jeżeli podczas przetwarzania zapytania, które wymaga wykorzystania danych znajdujących się na wszystkich baza danych, średni czas kompilacji zapytania na wierzchołku A wyniesie 27,12 ms, natomiast na pozostałych wierzchołkach średni czas będzie wynosił 23,25 ms. Złożenie zapytania na wierzchołku DB jest maksymalnym czasem wybranym ze wszystkich czasów kompilacji na pozostałych wierzchołkach.



**Rys. 3.** Graf przedstawiający rozproszenie przykładowej bazy danych

Można to zapisać według wzoru:

$$DB = \max(A, B)$$

gdzie:  $A, B$  – czasy przetwarzania zapytania na poszczególnych wierzchołkach.

Zatem otrzymana zostanie wartość najwyższa w tym wypadku 27,12 ms. Przy wykonaniu  $n$  zapytań strata czasu wynikająca z opóźnienia generowanego przez wierzchołek  $A$  będzie systematycznie rosła i może zostać opisana zależnością:

$$OP = \sum_n n \cdot (\max(A, B) - \min(A, B))$$

gdzie:

$OP$  – średni czas opóźnień,

$n$  – ilość wykonanych zapytań.

## 6. Podsumowanie

W pracy przedstawiono wyniki i interpretację badań przeprowadzonych na rozproszonych bazach danych, uzależnionych od wersji silnika bazodanowego. Celem przeprowadzonych badań miało być potwierdzenie lub zaprzeczenie teorii, iż wszystkie używane silniki baz danych w jednej rozproszonej bazie danych muszą być tej samej generacji, jak również potwierdzenie, że ewolucja silników bazodanowych wpływa na czas przetwarzania zapytań w rozproszonych bazach danych. Otrzymane wyniki jednoznacznie potwierdziły postawione tezy. Należy zwrócić również uwagę iż rosnąca popularność baz danych wymusza konieczność stosowania ujednoliconych silników bazodanowych, aby efekty pracy z tymi typami baz danych przynosiły oczekiwane korzyści.



## Bibliografia

Marcin Szeliga, *Transact-SQL. Czarna księga*,  
M. TamerOzs, Patrick Valduricz, *Principles of Distributed Database Systems*

## Streszczenie

Opracowanie zawiera przegląd metod przetwarzania zapytań w rozproszonych relacyjnych systemach zarządzania bazami danych. Omówione zostały podstawowe założenia, pojęcia i metody. Założeniem jest ustosunkowanie się do pomiaru czasu przetwarzania zapytań w rozproszonych bazach danych. Opracowanie omawia także samą koncepcję przetwarzania zapytań w rozproszonych bazach danych opartych na silniku serwerów bazodanowych.

**Słowa kluczowe:** zapytanie rozproszone, czas przetwarzania zapytania.

## Summary

The study contains an overview of methods for query processing in distributed relational database management systems. Discusses the basic assumptions, concepts and methods. The idea is to comment on the timing processing queries in distributed databases on the engine. The on database servers.

**Keywords:** Distributed query, query processing time.



**Wiesław Madej**

Katedra Cyfrowego Przetwarzania Sygnałów

Wydział Elektroniki i Informatyki

Politechnika Koszalińska

## **Specjalizowany cyfrowy system predykcijny ciągłym pomiarem odległości**

**Słowa kluczowe:** moduł obliczeń balistycznych, system czasu rzeczywistego, system predykcijny, przelicznik, warunki strzelania

### **Wstęp**

Zadanie predykcji punktu wyprzedzonego należy do głównych zadań, jakie należy wykonać podczas przygotowywania danych do strzelania w artylerii przeciwlotniczej. Problem ten był dotychczas realizowany przez analogowe systemy predykcyjne, oparte w większości na analogowych lampowych maszynach liczących. Systemy te stały się przestarzałe oraz bardzo zawodne. Dlatego powstała konieczność zaprojektowania nowych systemów predykcyjnych do istniejących zestawów przeciwlotniczych, co stanowi również podstawę do testowania rozwiązań umożliwiających projektowanie systemów predykcyjnych dla nowo powstających zestawów. Zagadnienia związane z projektowaniem artyleryjskich systemów predykcyjnych pojawiają się w literaturze od kilku lat. Różne aspekty związane z projektowaniem specjalizowanych cyfrowych systemów predykcyjnych przedstawione są w pracach [1–8]. W pracy [1] autor przedstawia wpływ braku uwzględnienia poprawek na dokładność pracy specjalizowanego systemu predykcyjnego. Prace [2, 3] dotyczą analizy konieczności zastosowania odpowiednich filtrów niezbędnych do poprawnego określenia prędkości wypracowania prawidłowych wyprzedzeń. W pracach [4, 5] autorzy omawiają wpływ zastosowania algorytmów o zmiennym okresie próbkowania i przyspieszonym algorytmie obliczeń. Jednakże w tych pracach nie przedstawiono konkretnych funkcji, jakie powinny być uwzględnione w algorytmie obliczeń w celu dokładnego odwzorowania toru lotu pocisku do celu, a co za tym idzie poprawnego rozwiązania zagadnienia trafienia. Próbę przedstawienia takich funkcji zawarto w pracy [6], jednak autorzy rozpatrując pełny układ równań ruchu pocisku stwierdzają, że czas potrzebny do obliczeń jest zbyt długi aby można było zrealizować taki algorytm w komputerach przemysłowych lub sterownikach programowalnych. W pracy [8] autor przedstawił algorytm prac specjalizowanego cyfrowego systemu predykcyjnego, jednakże w swoich obliczeniach uwzględnił tylko prostą hipotezę o ruchu celu. W dobie manewrujących obiektów latających hipoteza taka

jest mało użyteczna i praktycznie nieprzydatna. W związku z powyższym powstała konieczność opracowania takiego algorytmu, który uwzględni cele manewrujące. Umożliwi to w sposób dokładny przewidywać tor lotu celu, co jednocześnie wpłynie na poprawność rozwiązania zagadnienia trafienia, a także da możliwość zrealizowania funkcji i algorytmu w komputerze przemysłowym, sterowniku programowalnym PLC lub systemie mikroprocesorowym z procesorem sygnałowym.

## **1. Istota zadania predykcji punktu wyprzedzonego i ogólny algorytm jego rozwiązania**

Strzelanie do celów powietrznych poruszających się w przestrzeni z dużą prędkością i posiadających możliwości manewrowe wymaga uwzględnienia wyprzedzenia odpowiadającego przesunięciu się celu w czasie lotu pocisku do celu. Armatę wcelowuje się nie w punkt  $A_s$  (rys. 1), gdzie cel znajduje się w momencie wystrzału, lecz w pewien punkt  $A_w$  znajdujący się na przyszłej drodze celu, w którym według obliczeń pocisk powinien spotkać się z celem. Punkt ten nazywa się punktem wyprzedzonym. Aby skierować pocisk w punkt wyprzedzony, należy znać jego geometryczne współrzędne, na podstawie których określa się nastawy na działą.

W celu rozwiązania zadania spotkania należy znać bieżące współrzędne celu, wielkość i kierunek wektora prędkości celu, a także charakter ruchu celu w czasie lotu pocisku do punktu wyprzedzonego.

Bieżące współrzędne celu określa się w wyniku śledzenia go przez stację radiolokacyjną lub przyrządy optyczne. Określenie wielkości i kierunku wektora prędkości celu, rozwiązanie zadania spotkania oraz określenie nastaw na działą wykonuje w sposób ciągły przelicznik. Wypracowane nastawy – azymut wyprzedzony i kąt podniesienia są przekazywane w sposób ciągły na działą, przez co skierowuje się je w punkt wyprzedzony.

### **1.1. Hipotezy o ruchu celu**

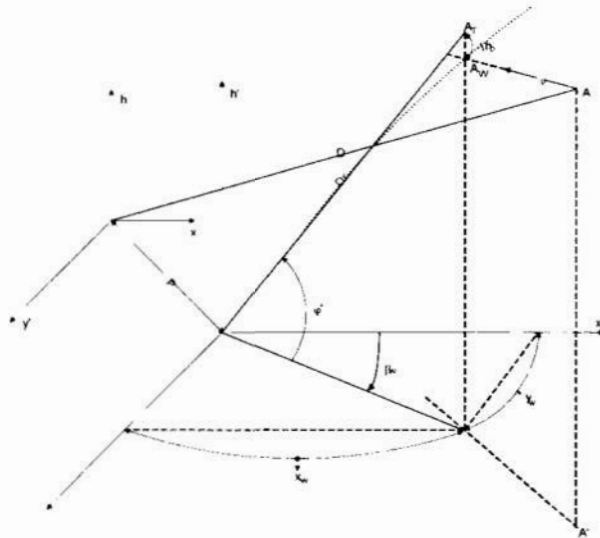
Do rozwiązania zadania spotkania pocisku z celem w określonym punkcie przestrzeni oraz w określonym momencie czasu niezbędna jest znajomość praw ruchu obydwu ciał i możliwość kierowania jednym z nich, w naszym przypadku prawem ruchu pocisku. Prawo ruchu pocisku zostało opracowane przez balistykę i możemy z dowolną dokładnością znać drogę ruchu pocisku, jego prędkość w różnych punktach toru, czas lotu do różnych punktów w przestrzeni. Znany jest nam także wpływ balistycznych i meteorologicznych warunków na lot pocisku i jego położenie w funkcji czasu. Dlatego po odpowiednim wycelowaniu działą można otrzymać pożądaną trajektorię pocisku przechodzącą przez punkt wyprzedzony. Prawo ruchu celu dla danego wystrzału może być ustalone tylko do punktu w przestrzeni, w którym cel znajduje się w momencie oddania wystrzału – punktu  $A_s$ , przy ciągłym określaniu bieżących współrzędnych

i czynników ruchu celu. Przy dalszym ruchu, to znaczy w czasie wyprzedzonym na odcinku  $A_s A_w$  ustalone prawo ruchu celu może być naruszone, czy to z woli pilota (w postaci manewru), czy też z przyczyn od jego woli niezależnych. Możemy jednak zakładać, że cel w czasie wyprzedzenia zachowa poprzedni charakter ruchu. Im czas wyprzedzenia będzie mniejszy, tym bardziej prawdopodobne, że charakter ruchu celu w tym czasie będzie mniej się różnił od charakteru jego ruchu do momentu wystrzału. Z powyższego wynika konieczność zakładania pewnych hipotez ruchu celu w czasie wyprzedzenia.

## 1.2. Istota zadania predykcji punktu wyprzedzonego

Predykcja punktu wyprzedzonego (rozwiązanie zadania spotkania) jest zasadniczym etapem przygotowania wystrzału, w czasie którego określa się geometryczne współrzędne punktu wyprzedzonego  $A_w$ , tj. punktu, w którym według obliczeń powinno nastąpić spotkanie się pocisku z celem. Skierowanie pocisku do punktu  $A_s$  nie ma sensu, ponieważ w czasie lotu pocisku cel przesunie się z tego punktu o odległość  $v_c \cdot \tau$  ( $\tau$  – czas lotu pocisku do punktu wyprzedzonego) i spotkanie pocisku z celem nie nastąpi. Z tego wynika, że aby nastąpiło spotkanie pocisku z celem, należy skierować pocisk do punktu  $A_w$  leżącego na kursie celu i oddalonego od punktu  $A_s$  o wartość  $v_c \cdot \tau$ .

Istota rozwiązania zadania spotkania pocisku z celem polega na uzgodnieniu w czasie dróg celu i pocisku, przebywanych z różnymi prędkościami (stała celu i zmienna pocisku), dających na przecięciu się punkt wyprzedzony  $A_w$ . Samo rozwiązanie polega na określeniu geometrycznych współrzędnych punktu wyprzedzonego w dowolnym układzie współrzędnych.



**Rys. 1.** Istota predykcji punktu wyprzedzonego, gdzie:  $x, y, z$  – układ współrzędnych związany z urządzeniem telemetrycznym,  $x', y', z'$  – układ współrzędnych związany z obiektem sterowania,  $P$  – wektor paralaksy,  $\Delta h_b$  – obniżenie balistyczne,  $A_w$  – predykowany punkt wyprzedzony

### 1.2.1. Predykcja punktu wyprzedzonego

W przypadku prostej hipotezy o ruchu celu, tzn. że cel w czasie wyprzedzenia porusza się prostoliniowo, jednostajnie i w dowolnej płaszczyźnie, gdy znane są jego współrzędne prostokątne  $x$ ,  $y$ ,  $h$  i składowe prędkości  $v_x$ ,  $v_y$ ,  $v_h$  należy określić współrzędne punktu  $A_w(\beta_w, D_w, \varepsilon_w)$ .

Wyznaczanie współrzędnych punktu wyprzedzonego realizuje się według następującego algorytmu:

1. Oblicza się odległość poziomą celu z zależności

$$d_p = \sqrt{x^2 + y^2} \quad (1)$$

2. Oblicza się odległość rzeczywistą celu z zależności

$$D = \sqrt{d_p^2 + h^2} \quad (2)$$

3. Oblicza się czas lotu pocisku do celu według tabel strzelniczych jako funkcję balistyczną

$$\tau = f(D_T, h_w) \quad (3)$$

W pierwszym przybliżeniu jako odległość pozorną  $D_T$  traktuje się odległość rzeczywistą  $D$ , natomiast jako wysokość wyprzedzoną  $h_w$  traktuje się wysokość rzeczywistą  $h$ .

4. Oblicza się współrzędne wyprzedzone według zależności

$$x_w = x + v_x \cdot \tau$$

$$y_w = y + v_y \cdot \tau$$

$$h_w = h + v_h \cdot \tau$$

(4)

5. Oblicza się wysokość pozorną  $h_T$  według zależności

$$h_T = h_w + \Delta h_b \quad (5)$$

gdzie:  $\Delta h_b$  – obniżenie balistyczne.

6. Oblicza się odległość wyprzedzoną poziomą  $d_{pw}$  i pozorną  $D_T$  celu według zależności

$$d_{pw} = \sqrt{x_w^2 + y_w^2}$$

$$D_T = \sqrt{d_{pw}^2 + h_T^2}$$

(6)

7. Oblicza się kąty azymutu  $\beta_w$  i podniesienia  $\varphi$  według zależności

$$\beta_w = \arcsin \frac{x_w}{d_{pw}}$$

$$\varphi = \arcsin \frac{h_T}{D_T}$$

(7)

8. Powraca się do punktu 3.



## 2. Program modułu obliczeń balistycznych

### 2.1. Opis ogólny algorytmu obliczeń

#### 2.1.1. Założenia projektowe

Program obliczeń balistycznych dotyczy armaty przeciwlotniczej 35 mm OERLIKON. Podstawą algorytmu obliczeń są tabele strzelnicze dla amunicji o prędkości początkowej pocisku 1180 m/s.

Uwzględnia się następujące odchyłki od tabelarycznych warunków strzelania:

- odchyłkę prędkości początkowej pocisku  $dV_0$ [m/s];
- odchyłkę temperatury powietrza  $dT$ [deg];
- odchyłkę ciśnienia powietrza  $dP$ [mbar];
- wiatr balistyczny  $W$ [m/s] (jego składową podłużną i boczną);
- zjawisko derywacji pocisku;
- paralaksę strzelania  $P$ [m] (jej składowe prostokątne);

W efekcie do algorytmu obliczeń wprowadza się:

- poprawkę czasu lotu pocisku  $dtV_0$ [s] ze względu na odchyłkę prędkości początkowej pocisku od tabelarycznej;
- poprawkę czasu lotu pocisku  $dtT$ [s] ze względu na odchyłkę temperatury powietrza od tabelarycznej;
- poprawkę czasu lotu pocisku  $dtP$ [s] ze względu na odchyłkę ciśnienia powietrza od tabelarycznego;
- poprawkę czasu lotu pocisku  $dtW$ [s] ze względu na oddziaływanie składowej podłużnej wiatru balistycznego;
- poprawkę przewyższenia balistycznego  $dhW$ [m] ze względu na oddziaływanie składowej podłużnej wiatru balistycznego;
- poprawkę azymutu strzelania  $dbW$ [rad] ze względu na oddziaływanie składowej bocznej wiatru balistycznego;
- poprawkę azymutu strzelania  $dbD$ [rad] ze względu na zjawisko derywacji pocisku;
- poprawki współrzędnych prostokątnych punktu wyprzedzonego  $P_x, P_y, P_z$  [m] ze względu na paralaksę strzelania.

Danymi wejściowymi do obliczeń są:

- współrzędne prostokątne celu – wygładzone;
- składowe wektora prędkości celu w układzie prostokątnym;
- składowe wektora przyśpieszenia celu w układzie prostokątnym;
- kąty odchylenia, pochylenia i przechylenia osi transportera względem osi układu stabilizowanego;

- odchyłki od tabelarycznych warunków meteorologicznych i balistycznych;
- składowe wektora paralaksy strzelania;
- Danymi wyjściowymi modułu obliczeń balistycznych są – opcjonalnie:
- kąty wycelowania armaty w stabilizowanym układzie współrzędnych;
- kąty wycelowania armaty w niestabilizowanym układzie współrzędnych transportera;
- kąty wyprzedzenia i przewyższenia w stabilizowanym układzie współrzędnych;
- kąty wyprzedzenia i przewyższenia w niestabilizowanym układzie współrzędnych transportera;

Przyjęty został lewoskrętny układ współrzędnych stabilizowanych, w którym oś  $x$  skierowana jest na północ, oś  $y$  na wschód natomiast oś  $z$  w górę. W przyjętym układzie współrzędnych kąty poziome rosną w prawo a kąty pionowe w górę.

W układzie współrzędnych związanych z transporterem oś  $x$  skierowana jest wzdłuż transportera do przodu, oś  $y$  w stronę prawej burty natomiast oś  $z$  w górę.

Interpretacja kątów w układzie związanym z transporterem jest następująca:

- kąt odchylenia jest dodatni gdy oś transportera odchylona jest w prawo od kierunku północy;
- kąt pochylenia jest dodatni gdy oś transportera pochylona jest w dół od poziomu;
- kąt przechylenia jest dodatni gdy transporter przechylony jest na lewa burtę.

Kąty: odchylenia, pochylenia i przechylenia zawierają się w przedziałach  $(0, 2\pi)$ .

Do wyrażenia kątów w tysięcznych użyto tysięcznej zdefiniowanej jako  $1/6400$  część kąta pełnego (jak w tabelach strzelniczych).

Przy konstrukcji modelu matematycznego procedury obliczeń czasu wyprzedzonego przyjęto opcjonalnie następujące dwie hipotezy o ruchu celu:

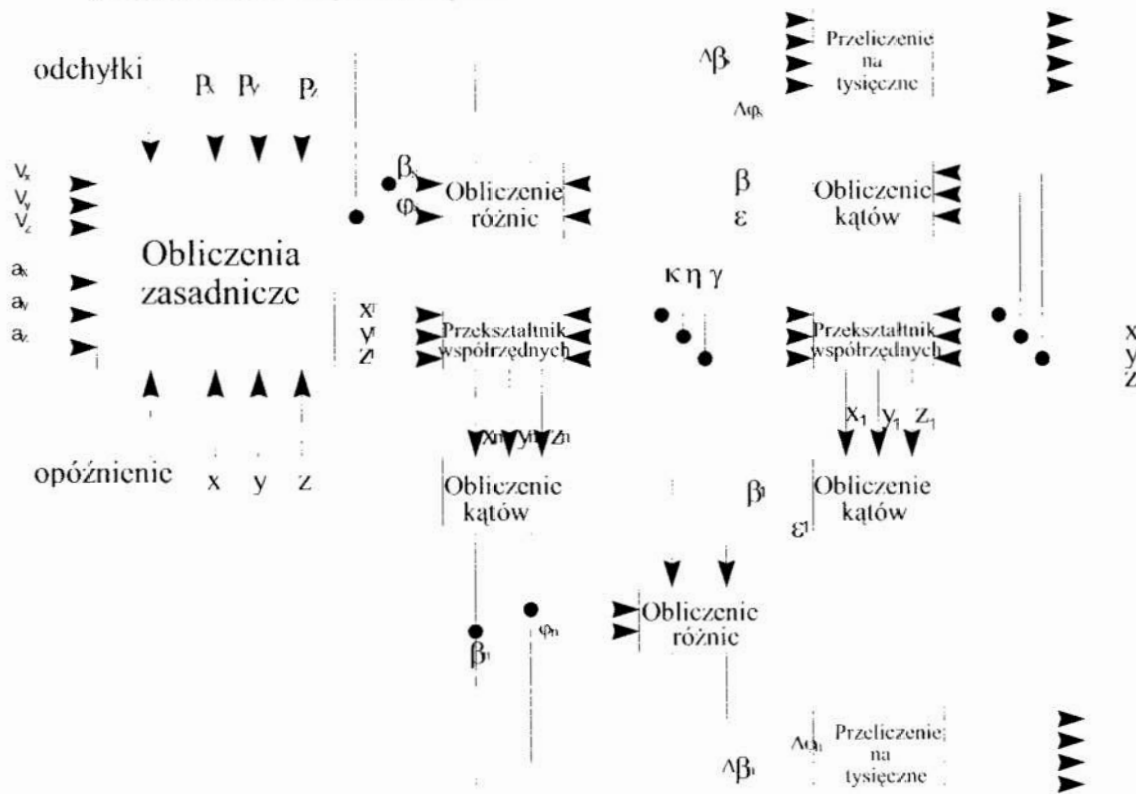
1. *W czasie lotu pocisku do celu cel przemieszcza się ruchem śrubowym po powierzchni walca pionowego ze stałą prędkością kątową oraz stałą prędkością wznoszenia (opadania).*
2. *W czasie lotu pocisku do celu cel porusza się po dowolnej trajektorii zachowując stałe wartości składowych prostokątnych wektora przyspieszenia.*

### 2.1.2. Schemat blokowy modułu balistycznego

Moduł balistyczny składa się z siedmiu różnych bloków wywoływanych w odpowiedniej kolejności przez procedurę obliczeń (rys. 2):

- bloku obliczeń zasadniczych – w pętli;
- bloku obliczania kątów;
- bloku obliczania poprawek donośności i wysokości;
- bloku obliczania poprawek kierunku;
- bloku obliczania różnic kątów;

- bloku przeliczania kątów na tysięczne;
- bloku przekształtnika współrzędnych.



Rys. 2. Schemat blokowy modułu obliczeń balistycznych

Danymi wejściowymi modułu balistycznego są:

- współrzędne prostokątne celu  $x, y, z$ ;
- składowe wektora prędkości celu  $V_x, V_y, V_z$ ;
- składowe wektora przyspieszenia celu  $a_x, a_y, a_z$ ;
- składowe wektora paralaksy  $P_x, P_y, P_z$ ;
- opóźnienie wprowadzane przez filtry wygładzające i różniczkujące;
- kąty przechylenia, pochylenia i odchylenia transportera  $\kappa, \eta, \gamma$ ;
- wektor odchyłek od tabelarycznych warunków strzelania.

Danymi wyjściowymi z modułu balistycznego są – opcjonalnie:

- kąty wycelowania armaty w stabilizowanym układzie współrzędnych  $\beta, \varphi$ ;
- kąty wyprzedzenia i przewyższenia w stabilizowanym układzie współrzędnych  $\delta\beta, \delta\varphi$ ;
- kąty wycelowania armaty w niestabilizowanym układzie współrzędnych  $\beta_a, \varphi_a$ ;
- kąty wyprzedzenia i przewyższenia w niestabilizowanym układzie współrzędnych  $\Delta\beta_a, \Delta\varphi_a$ .

Wszystkie dane wejściowe z wyjątkiem kątów obrotu transportera doprowadzone są do bloku obliczeń zasadniczych. Na podstawie tych wielkości na wyjściu bloku obliczeń zasadniczych uzyskuje się współrzędne prostokątne punktu teoretycznego położonego nad punktem wyprzedzonym o wartość przewyższenia balistycznego  $d_{hb}$  ( $X_t$ ,  $Y_t$ ,  $Z_t$ ) oraz kąty wycelowania lufy ( $\beta$ ,  $\varphi$ ) w stabilizowanym układzie współrzędnych.

Na podstawie bieżących wartości azymutu i kąta położenia celu oraz zapamiętanych wartości azymutu strzelania i kąta podniesienia lufy oblicza się bieżące wartości poziomego kąta wyprzedzenia i pionowego kąta przewyższenia ( $\Delta\beta$ ,  $\Delta\varphi$ ) jako różnice pomiędzy odpowiednimi kątami.

Ostateczne wartości wszystkich kątów mogą być odbierane z modułu balistycznego w radianach lub w tysięcznych artyleryjskich. Przeliczenia kątów z radianów na tysięczne dokonuje się w bloku obliczania tysięcznych.

Współrzędne punktu teoretycznego oraz współrzędne celu przekazane zostają kolejno do przekształtnika współrzędnych.

Wielkościami parametrycznymi tego bloku są kąty odchylenia, pochylenia i przechylenia armaty. Powyższe kąty są skutkiem nierówności podłoża oraz niezgodności kierunku zorientowania i kierunku północy. Wielkościami wyjściowymi bloku przekształtnika są współrzędne prostokątne – punktu teoretycznego ( $X_{ta}$ ,  $Y_{ta}$ ,  $Z_{ta}$ ) i celu ( $X_a$ ,  $Y_a$ ,  $Z_a$ ) w układzie związanym z armatą.

W bloku obliczania kątów dokonuje się zamiany współrzędnych prostokątnych na kąty wycelowania ( $\beta_a$ ,  $\varphi_a$ ) i wizowania ( $\beta_{ca}$ ,  $\varepsilon_{ca}$ ) w niestabilizowanym układzie współrzędnych.

Po obliczeniu różnic tych kątów w bloku różnic otrzymuje się kąty wyprzedzenia i przewyższenia ( $\Delta\beta_a$ ,  $\Delta\varphi_a$ ) dla armaty w jej własnym układzie współrzędnych. Całkowite kąty wycelowania oraz różnice kątów są następnie przeliczane z radianów na tysięczne w takim samym bloku obliczeń dla kanału stabilizowanego i niestabilizowanego.

Blokowa konstrukcja algorytmu modułu balistycznego umożliwia uzyskanie ostatecznych danych do wysterowania armaty w różnorodnej formie. Wybór potrzebnego kompletu danych wyjściowych będzie od przyjętej reguły sterowania armatą.

## **2.2. Uwzględnienie przyspieszeń liniowych celu w algorytmie obliczeń**

### **2.2.1. Procedura pośredniego wyznaczania prędkości kątowej**

W programie wykorzystującym procedurę pośredniego wyznaczania prędkości nie wykorzystuje się danych o przyspieszeniach liniowych. W module balistycznym tego programu wykorzystuje się hipotezę, że cel porusza się po okręgu ze stałą prędkością. Położenie punktu wyprzedzonego wylicza się z zależności:

$$\begin{aligned}
 x_w &= x + v_x \cdot t_x - v_y \cdot t_y \\
 y_w &= y + v_y \cdot t_x + v_x \cdot t_y \\
 z_w &= z + v_x \cdot t
 \end{aligned}
 \tag{8}$$

gdzie:

$$\begin{aligned}
 t_x &= \frac{\sin(\omega \cdot t)}{\omega} \\
 t_y &= \frac{1 - \cos(\omega \cdot t)}{\omega}
 \end{aligned}$$

$t$  – czas lotu pocisku do punktu wyprzedzonego.

Prędkość kątową  $\omega$ , która niezbędna jest do wyznaczania położenia punktu wyprzedzonego, wyznacza się na podstawie zmian kąta kursu celu. Odbywa się to według następującej procedury:

- oblicza się różnicę  $\Delta_q$  pomiędzy aktualnym a poprzednim kątem kursu;
- jeżeli moduł różnicy  $\Delta_q$  jest większy niż  $\pi$ , to modyfikuje się tę różnicę poprzez

$$\text{zależność: } \Delta_q = \Delta_q - \frac{\Delta_q}{|\Delta_q|} \cdot 2 \cdot \pi$$

- następnie oblicza się prędkość kątową  $\omega$  z zależności:

$$\omega = \left[ 2 \cdot \Delta_q + (2 - 1,02 \cdot \text{krok}) \cdot \omega_1 \right] / (2 + 1,02 \cdot \text{krok})$$

gdzie: krok – czas jaki upłynął od poprzedniego obliczania prędkości kątowej;

$\omega_1$  – poprzednia wartość prędkości kątowej.

### 2.2.2. Procedura bezpośredniego wyznaczania prędkości kątowej

Program wykorzystujący tę procedurę różni się od poprzedniego tym, że prędkość kątową niezbędna do prawidłowej pracy modułu balistycznego nie jest wyznaczana w sposób pośredni, lecz bezpośrednio na podstawie przyspieszeń liniowych ( $a_x$ ,  $a_y$ ) i prędkości liniowych ( $v_x$ ,  $v_y$ ) z następującej zależności:

$$\omega = \frac{a_y \cdot v_x - a_x \cdot v_y}{v_x^2 + v_y^2}
 \tag{9}$$

Zależność tę uzyskuje się poprzez zrózniczkowanie względem czasu równania

$$q_c = \arctg \frac{v_y}{v_x},
 \tag{10}$$

które opisuje zależność pomiędzy kątem kursu celu a składowymi prędkości celu.

### 2.2.3. Procedura obliczeń z rozkładu w szereg Taylora

Oprócz dwóch poprzednich przypadków przetestowano również program, w którym zmieniono hipotezę ruchu celu na hipotezę o stałych przyspieszeniach. W hipotezie tej współrzędne punktu wyprzedzonego liczone są z rozwinięcia funkcji w szereg Taylora:

$$\begin{aligned}x_w &= x + v_x \cdot t + \frac{a_x \cdot t^2}{2} \\y_w &= y + v_y \cdot t + \frac{a_y \cdot t^2}{2} \\z_w &= z + v_z \cdot t + \frac{a_z \cdot t^2}{2}\end{aligned}\tag{11}$$

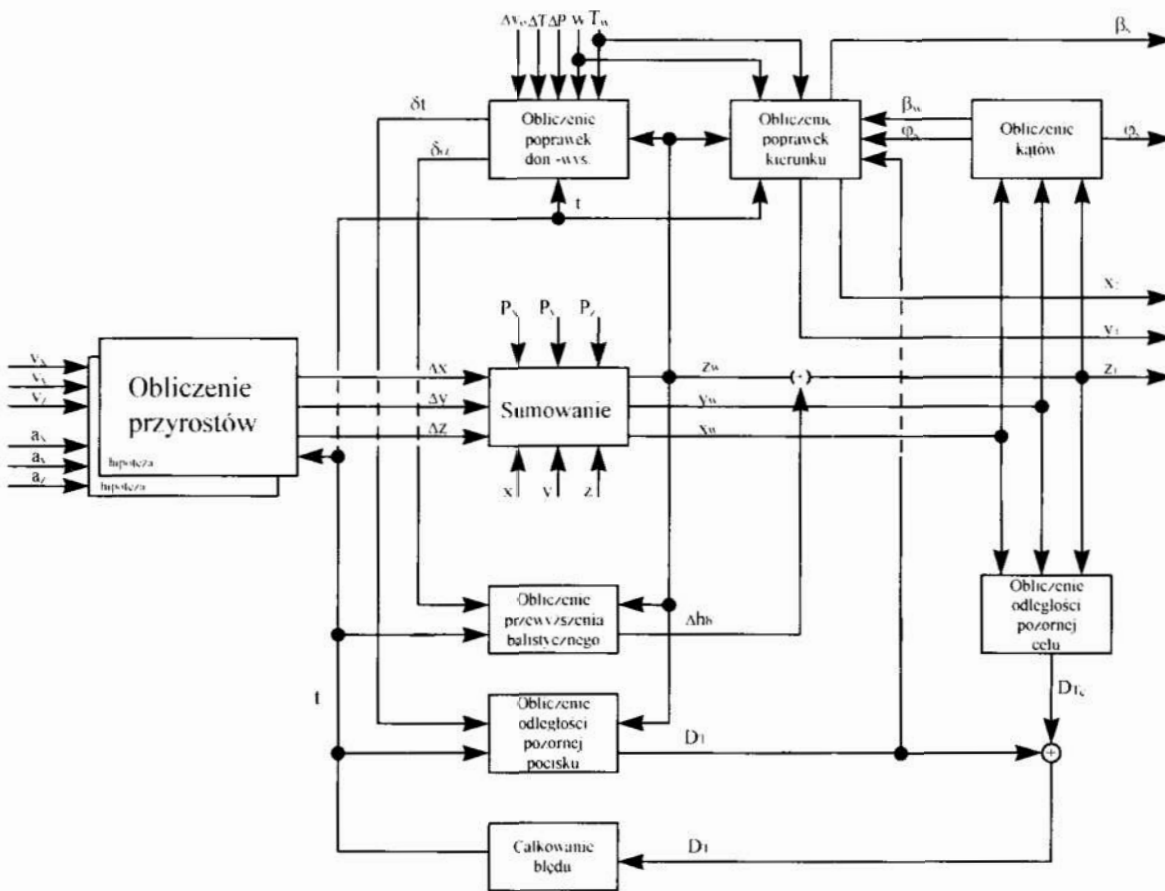
### 2.3. Blok obliczeń zasadniczych z uwzględnieniem poprawek

Ta część obliczeń odbywa się w pętli. Algorytm obliczeń tak długo jest powtarzany z tymi samymi danymi wejściowymi, aż spełniony będzie warunek równości odległości na jakiej znajdzie się punkt pozorny po upływie wyliczonego czasu (DTp) i odległości do punktu leżącego nad celem o wartość przewyższenia balistycznego (DTc) po upływie tego samego czasu od chwili strzału z błędem nie większym niż 1/1000 odległości do punktu pozornego:

$$DTp - DTc \leq 0.001 \cdot DTp$$

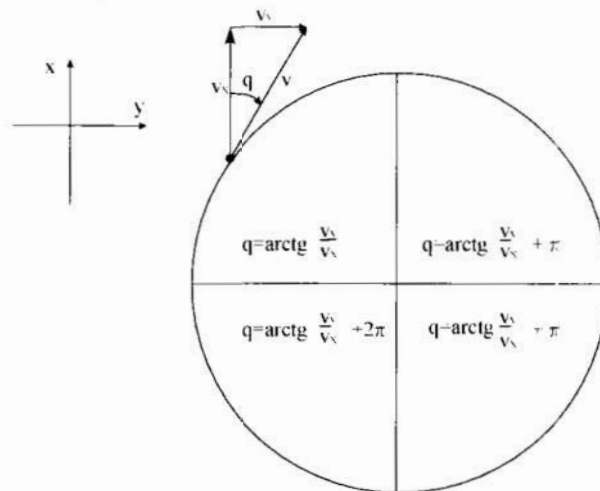
Blokowy schemat algorytmu obliczeń przedstawiony jest na rys. 3.





Rys. 3. Schemat blokowy obliczeń zasadniczych

A) Jeżeli zakłada się pracę według pierwszej hipotezy to na początku oblicza się prędkość kątową celu, bowiem na stałości prędkości kątowej ruchu celu w czasie lotu pocisku opiera się konstrukcja matematyczna modułu obliczania przyrostów współrzędnych. Prędkość kątową celu oblicza się na podstawie składowych wektora prędkości  $v_x, v_y$  jako pochodną kąta drogi.



Rys. 4. Ilustracja do obliczenia wartości kąta drogi

Bez względu na to w której ćwiartce aktualnie znajduje się cel, prędkość kątowna jako pochodna kąta drogi wyraża się zależnością:

$$\omega = \frac{d}{dt} \left[ \arctg \left( \frac{v_y}{v_x} \right) \right]$$

a to w konsekwencji daje prostą funkcję składowych prędkości i przyspieszenia:

$$\omega = \frac{a_y \cdot v_x - a_x \cdot v_y}{v_x^2 + v_y^2}$$

Na podstawie składowych wektora prędkości oraz wyliczonej wcześniej prędkości kątowej obliczone zostają przyrosty współrzędnych  $\Delta x$ ,  $\Delta y$ ,  $\Delta z$  odpowiadające aktualnemu czasowi lotu pocisku  $t$ .

Rzeczywisty czas lotu pocisku musi być skorygowany o wartość opóźnienia wprowadzanego przez filtry współrzędnych celu i filtry składowych wektora prędkości.

W programie zasymulowana została ta korekcja lecz z zerową wartością opóźnienia ( $op = 0$ ).

$$t_1 = t + op$$

Przyrosty współrzędnych w płaszczyźnie poziomej  $\Delta x$ ,  $\Delta y$  są rzutami cięciwy okręgu, po którym według hipotezy porusza się cel na osie układu  $x$ ,  $y$ .

Wyliczane są one według zależności:

$$\begin{aligned} \Delta x &= \frac{v_x \sin(\omega t_1)}{\omega} - \frac{v_x (1 - \cos(\omega t_1))}{\omega} \\ \Delta y &= \frac{v_y \sin(\omega t_1)}{\omega} + \frac{v_y (1 - \cos(\omega t_1))}{\omega} \end{aligned} \quad (12)$$

W przypadku szczególnym, gdy prędkość kątowa jest zerowa to ruch po powierzchni walca przechodzi w ruch w płaszczyźnie pionowej. Przyrosty współrzędnych przyjmują wówczas wartości jakby z wyrażeń:

$$\Delta x = v_x t$$

$$\Delta y = v_y t$$

Gdy prędkość kątowa dąży do nieskończoności to wartości przyrostów współrzędnych dążą do zera co jest teoretycznie zrozumiałe, gdyż oznaczałoby to, że cel zawsze znajduje się w tym samym punkcie. Wobec czego wyprzedzenia nie są potrzebne.

Przyrost trzeciej współrzędnej jest funkcją liniową i liczony jest z zależności:

$$\Delta z = v_z t$$

Gdy cel jest nieruchomy, to składowe jego prędkości są zerowe a zatem i wyprzedzenia zerują się.

B) Jeżeli zakłada się pracę według drugiej hipotezy to przyrosty współrzędnych oblicza się według zależności:

$$\begin{aligned}\Delta x &= v_x t_1 + a_x \frac{t_1^2}{2} \\ \Delta y &= v_y t_1 + a_y \frac{t_1^2}{2} \\ \Delta z &= v_z t_1 + a_z \frac{t_1^2}{2}\end{aligned}\quad (13)$$

Sumy bieżących współrzędnych  $x$ ,  $y$ ,  $z$  oraz odpowiadających im przyrostów i składowych paralaksy strzelania są współrzędnymi wyprzedzonymi punktu trafienia ( $X_w$ ,  $Y_w$ ,  $Z_w$ ).

Wysokość wyprzedzona, aktualny czas lotu pocisku  $t$  oraz poprawki czasu lotu pocisku  $\delta t$  i  $\delta t_w$  są danymi wejściowymi do obliczenia przewyższenia balistycznego  $\Delta hb$  oraz odległości pozornej pocisku  $DTp$ .

Te dwie wielkości jako rezultat aproksymacji zasadniczych tabel strzelniczych i przyjętego modelu oddziaływania odchyłek od warunków tabelarycznych na trajektorię pocisku oblicza się według zależności:

$$\begin{aligned}t_p &= t + \delta t \\ t_b &= t_p - \delta t_w \\ DTp &= \frac{1180t_p}{1 + \frac{1180t_p}{8860 + 0,36z_w} - \frac{t_p^2}{1180}} \\ DTw &= \frac{1180t_b}{1 + \frac{1180t_b}{8860 + 0,36z_w} - \frac{t_b^2}{1180}} \\ dhw &= (DTp - DTw)(\sin \varphi - 0,14 \cos \varphi) \\ \Delta hb &= (2,1 + 0,25 \cdot 10^{-3} z_w) t + \pi^2 + dhw\end{aligned}\quad (14)$$

Czas  $t_p$  jest poprawionym czasem lotu pocisku uwzględniającym wpływ wszystkich czterech czynników. Czas  $t_b$  jest poprawionym czasem lotu pocisku nie uwzględniającym wpływu składowej podłużnej wiatru.

Na podstawie tych czasów liczone są odpowiednio odległości pozorne  $DTp$  oraz  $DTw$ . Odcinek będący różnicą tych odległości nachylony jest do poziomu pod kątem takim jak lufa armaty. Poprawka przewyższenia balistycznego  $dhw$  spowodowana wpływem wiatru jest nieco mniejsza niż składowa pionowa tego odcinka.

Funkcja przewyższenia balistycznego musi być ograniczona od góry ze względu na odporność algorytmu na zakłócenia impulsowe danych wejściowych. W przypadku braku tego ograniczenia w stanach nieustalonych mogłoby się zdarzyć, że przewyższenie balistyczne byłoby większe od odległości pozornej. To z kolei spowodowałoby awarię algorytmu. Opisana sytuacja wynika wyłącznie z charakteru funkcji przyjętych do aproksymacji tabel strzelniczych.

Ograniczenia dokonuje się na poziomie, który w rzeczywistości nigdy nie będzie osiągany według zależności:

jeżeli ( $\Delta hb > 5000$ ) to  $\Delta hb := 5000$ .

Suma przewyższenia i wysokości wyprzedzonej to wysokość pozorna celu  $Z_t$ . Pierwiastek kwadratowy z sumy kwadratów współrzędnych wyprzedzonych poziomych oraz kwadratu wysokości pozornej jest odległością  $DT_c$  do punktu położonego nad celem o wartość  $\Delta hb$ .

$$DT_c = \sqrt{X_w^2 + Y_w^2 + Z_w^2} \quad (15)$$

Różnica  $\Delta Dt$  pomiędzy odległościami  $DT_p$  i  $DT_c$  jest wyznacznikiem błędu czasu lotu pocisku. Dopuszczalny błąd czasu lotu pocisku wynosi 0,01 [s] a to odpowiada w przybliżeniu około 0,001 odległości do celu. Dlatego też przed następnym powtórzeniem algorytmu obliczeń poprawia się czas lotu pocisku o wartość  $\Delta t = 0,001 * \Delta Dt$ .

Proces jest szybko i zawsze zbieżny, pod warunkiem poprawnie dobranego współczynnika w wyrażeniu na poprawkę czasu  $\Delta t$  oraz dozwolonej wartości czasu  $t$  w chwili startu programu. Wspomnianą wartość współczynnika dobrano doświadczalnie, natomiast czas  $t$  w chwili startu programu może być dowolny z przedziału 0 do 30 sekund. Dla prostoty wskazane jest przyjmować czas zerowy.

Blok obliczeń zasadniczych wywołuje blok obliczania kątów, a także procedury obliczania poprawek donośności i wysokości oraz obliczania poprawek kierunku.

Te dwa bloki i dwie procedury tworzą zamkniętą pętlę obliczeń aż do chwili spełnienia założonego warunku wyjścia z pętli.

Współrzędne prostokątne punktu pozornego  $X_w$ ,  $Y_w$ ,  $Z_t$  są podstawą do obliczenia kątów wycelowania armaty. Na ich podstawie otrzymuje się na wyjściu bloku obliczania kątów bieżącą wartość azymutu punktu wyprzedzonego, przed uwzględnieniem poprawki kierunku oraz bieżącą wartość kąta podniesienia lufy w płaszczyźnie pionowej.

### 2.3.1. Procedura obliczania poprawek donośności i wysokości

Procedura uaktywniana jest tylko przy czasach lotu pocisku mieszczących się w granicach przewidzianych w tabelach strzelniczych. Poza przedziałem rzeczywistości np. w stanach przejściowych algorytmu obliczeń, poprawki mogłyby niepotrzebnie wydłużać czas ustalania się procesów.

Najpierw dokonuje się obliczenia składowych wiatru balistycznego.

Rozkład wiatru na składowe wykonuje się według zależności:

$$W_b = W \sin(\beta_w - T_w)$$

$$W_l = W \cos(\beta_w - K_w)$$

gdzie  $K_w$  jest kątem wiatru dla składowej podłużnej obliczanym według zależności:

$$K_w = T_w + \pi \quad \text{dla } T_w \leq \pi$$

$$K_w = T_w - \pi \quad \text{dla } T_w > \pi$$

przy czym  $T_w$  jest topograficznym azymutem wiatru.

Następnie procedura oblicza cztery poprawki czasu lotu pocisku jako funkcje tabelarycznego czasu lotu, wysokości wyprzedzonej oraz odpowiednich odchyłek od warunków tabelarycznych:

$$\delta t V_0 = (4t^2 + 94t + 30 \cdot 10^{-6} z_w^2 - 2,58 \cdot 10^{-3} z_w - 4,5) \cdot 10^{-5} dV_0 [s]$$

$$\delta t T_0 = (-2,2 \cdot 10^{-3} z_w + 34) \cdot 10^{-5} t^2 dT_0 [s]$$

$$\delta t P_0 = -11,5 \cdot 10^{-5} t^2 dP_0 [s]$$

$$\delta t W_1 = (29t^2 - 32t - 64 \cdot 10^{-6} z_w^2 + 102 \cdot 10^{-3} z_w + 63,5) \cdot 10^{-5} W_1 [s]$$

Do bloku obliczeń zasadniczych przekazywana jest sumaryczna poprawka czasu  $\delta t$  oraz dodatkowo poprawka czasu  $\delta t W_l$ .

$$\delta t = \delta t V_0 + \delta t T_0 + \delta t P_0 + \delta t W_1 [s]$$

### 2.3.2. Procedura uwzględnienia poprawek kierunku

W procedurze oblicza się i uwzględnia poprawkę kierunku wywołaną derywacją pocisku oraz poprawkę kierunku wywołaną wiatrem balistycznym. Poprawkę na derywację pocisku oblicza się według formuły uwzględniającej czas lotu pocisku do celu oraz wysokość punktu wyprzedzonego.

$$\delta \beta_{der} = (0,105t^2 + 0,883t - 0,3927 \cdot 10^{-6} z_w^2 + 0,1343 \cdot 10^{-3} z_w - 0,1212) \cdot 10^{-4} [rad]$$

Poprawkę kierunku strzelania ze względu na oddziaływanie składowej bocznej wiatru balistycznego oblicza się według następującej zależności będącej rezultatem aproksymacji tabel poprawek:

$$\delta \beta_w = 0,105 \cdot 10^{-3} t \cdot w_b [rad]$$

przy czym czynnik  $w_b$  jest wartością składowej bocznej wiatru w [m/s] uzyskiwaną z rozkładu wiatru na składowe.

Ostatecznie poprawiony azymut strzelania jest sumą kąta wyliczonego w procedurze obliczeń kąta oraz obydwu poprawek kierunku:

$$\beta = \beta_w + \delta \beta_{wt} + \delta \beta_{der} [rad]$$

Ponieważ w dalszych obliczeniach potrzebne są składowe prostokątne zawierające w sobie obliczone powyżej poprawki kierunku dlatego w tym miejscu dokonuje się ich wyznaczenia według zależności:

$$X_t = DT_p \cos \varphi \cos \beta [m]$$

$$Y_t = DT_p \cos\varphi \sin\beta \text{ [m]}$$

Trzecia współrzędna tj.  $Z_t$  obliczona została już wcześniej.

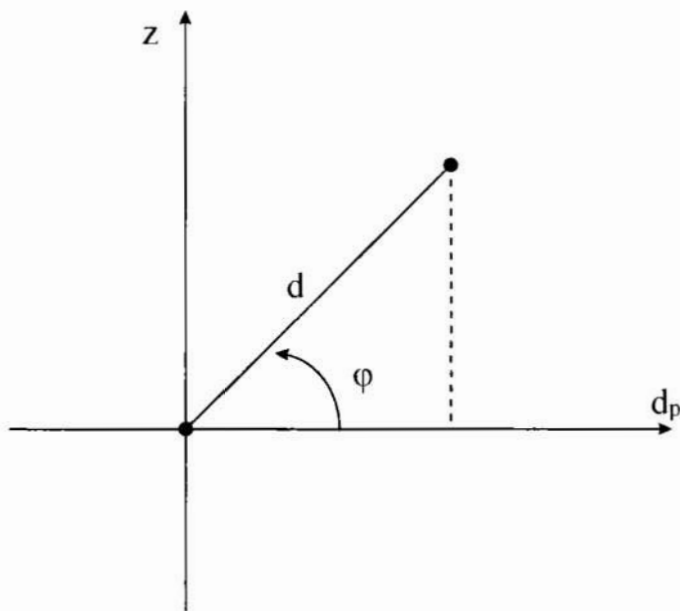
### 2.3.3. Blok obliczenia kątów

W bloku obliczane są dwa kąty:

- kąt w płaszczyźnie pionowej – na podstawie wysokości oraz odległości poziomej;
- azymut w płaszczyźnie poziomej – na podstawie współrzędnych  $x, y$ ;

Kąt w płaszczyźnie pionowej oblicza się dla podstawowego przedziału zmienności funkcji tangens według wyrażenia:

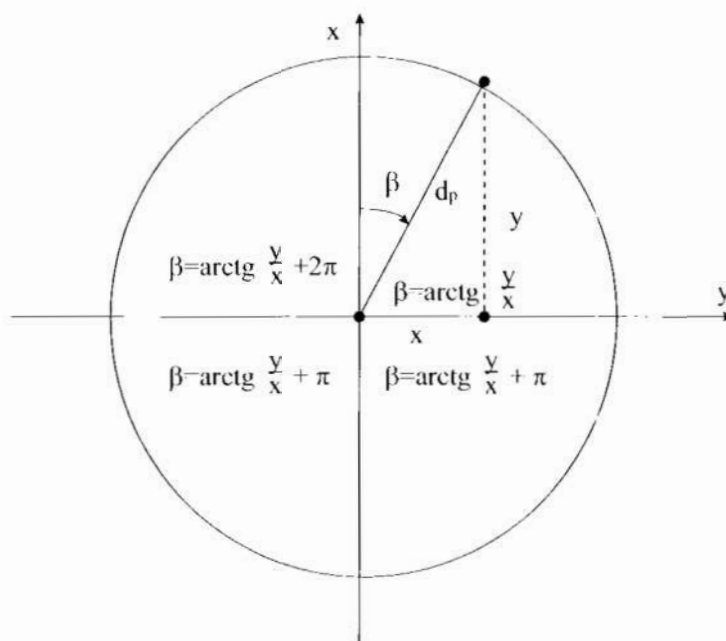
$$\varepsilon = \arctg\left(\frac{z}{\sqrt{x^2 + y^2}}\right)$$



**Rys. 5.** Obliczenie kąta podniesienia lufy

W zależności od wartości i wzajemnych relacji pomiędzy współrzędnymi  $x, y$  azymut oblicza się według wyrażen wpisanych w odpowiednie ćwiartki koła.





Rys. 6. Obliczenie azymutu

Aby uniknąć dwóch różnych wartości azymutu dla tego samego kierunku wprowadzona została instrukcja przypisująca kątowi pełną wartość zerową.

jeżeli kąt =  $2\pi$  to kąt := 0

#### 2.4. Bloki obliczeń pomocniczych

Obliczanie różnic kątów – kąt wyprzedzenia obliczany jest jako różnica pomiędzy azymutem kierunku strzelania i azymutem celu.

$$\Delta\beta = \beta - \beta_c$$

W celu uniknięcia impulsu prostokątnego dużego kąta wyprzedzenia przy przechodzeniu przez zero wprowadzona została instrukcja:

$$\text{jeśli } |\Delta\beta| > \pi \text{ to } \Delta\beta := \Delta\beta - \text{sign}(\Delta\beta) * 2\pi$$

Kąt przewyższenia oblicza się jako różnicę pomiędzy obliczonym kątem podniesienia lufy i kątem położenia celu:

$$\Delta\varphi = \varphi - \varepsilon_c$$

Przeliczanie kątów na tysięczne – całkowite kąty wycelowania oraz kąty wyprzedzenia i przewyższenia przelicza się z radianów na tysięczne poprzez mnożenie przez współczynnik będący ilorazem liczby tysięcznych i liczby radianów w kącie pełnym:

$$\text{mnożnik} = 3200/\pi$$

$$\text{kąt w tysięcznych} = \text{kąt w radianach} * \text{mnożnik};$$

Przekształtnik współrzędnych – to równanie macierzowe o postaci:

$$[w2] = [M][w1],$$

w którym:

- [w1] to trójelementowy wektor współrzędnych prostokątnych przed przekształceniem;
- [w2] to trójelementowy wektor współrzędnych prostokątnych po przekształceniu;
- [M] to dziewięcioelementowa macierz współczynników zależnych od wartości kątów odchylenia, pochyleń i przechyleń armaty.

Elementy macierzy współczynników obliczone zostają przed rozpoczęciem pracy modułu balistycznego jako początkowe warunki startowe programu.

Obliczeń dokonuje się według zależności:

$$m[1,1]=\cos\gamma \cos\kappa;$$

$$m[1,2]=\cos\gamma \sin\kappa;$$

$$m[1,3]=-\sin\gamma;$$

$$m[2,1]=\sin\eta \sin\gamma \cos\kappa - \cos\eta \sin\kappa;$$

$$m[2,2]=\sin\eta \sin\gamma \sin\kappa + \cos\eta \cos\kappa;$$

$$m[2,3]=\sin\eta \cos\gamma;$$

$$m[3,1]=\cos\eta \sin\gamma \cos\kappa + \sin\eta \sin\kappa;$$

$$m[3,2]=\cos\eta \sin\gamma \sin\kappa - \sin\eta \cos\kappa;$$

$$m[3,3]=\cos\eta \cos\gamma;$$

gdzie:

$\kappa$  - kąt odchylenia;

$\gamma$  - kąt pochyleń;

$\eta$  - kąt przechyleń.

### 3. Testowanie systemu

Testowanie systemu odbyło się dwuetapowo. Pierwszy etap polegał na badaniu symulacyjnym w laboratorium.

Na podstawie przeprowadzonych testów laboratoryjnych można wyciągnąć następujące wnioski:

- dla lotu po okręgu najmniejsze błędy uzyskuje się przy bezpośrednim wyznaczaniu prędkości kątowej;
- dla celu manewrującego najmniejsze błędy uzyskuje się przy hipotezie o ruchu celu ze stałymi przyspieszeniami.

Należy zastanowić się nad sposobem wyboru hipotezy o ruchu celu, jaki należy przyjąć w programie obliczeń, gdyż jeżeli założymy, że nie będziemy prowadzić ognia do celu bardzo manewrującego, wystarczy zastosować hipotezę o ruchu po okręgu

i bezpośrednim wyznaczaniu prędkości kątowej, jednakże przy mocno manewrujących celach konieczne jest zastosowanie hipotezy o stałych przyspieszeniach (z rozwinięciem w szereg Taylora).

Drugi etap testów został przeprowadzony na poligonie i polegał na realizacji strzelań do celów imitowanych przez komputer z oceną strzelania przez stację radiolokacyjną sterowaną przez ten sam komputer. Wyniki testów były zadowalające.

#### 4. Podsumowanie

W pracy przedstawiono procedurę postępowania przy projektowaniu specjalizowanego cyfrowego systemu predykcyjnego przeznaczonego dla artylerii przeciwlotniczej. Wyznaczono funkcje balistyczne dla 35-mm armaty Oerlikon oraz przedstawiono algorytm pracy modułu balistycznego. Algorytm został sprawdzony i przetestowany dla różnych przypadków lotu celu. Jego poprawność została zbadana i stwierdzono, że użycie tego algorytmu z wyznaczonymi funkcjami balistycznymi i funkcjami poprawek gwarantuje poprawne rozwiązanie zagadnienia trafienia z dopuszczalnymi błędami i w czasie wystarczającym do realizacji zadania.

#### Bibliografia

1. Borowczyk W., Kaczmarek W., *Badanie wpływu poprawek na dokładność wyliczania nastaw działowych w oparciu o uproszczone tabele balistyczne*, Biuletyn WAT, Vol. LIV, nr 9, 2005, pp. 39-48.
2. Borowczyk W., *Analiza przydatności różnego rodzaju filtrów do wyznaczania wyprzedzeń współrzędnych celu oraz metody ich opracowania*, Biuletyn WAT, Vol. LIV, nr 9, 2005, pp. 23-37.
3. Borowczyk W., *Analiza wpływu rzędu zastosowanych filtrów cyfrowych, przeznaczonych do wyznaczania wyprzedzeń, na dokładność rozwiązania zagadnienia trafienia*, Biuletyn WAT, Vol. LIV, nr 9, 2005, pp. 5-21.
4. Borowczyk W., Kaczmarek W., *Analiza błędów obliczania nastaw działowych przy zastosowaniu algorytmu zmiennego w czasie*, Biuletyn WAT, Vol. LVII, nr 1, 2008, pp. 7-18.
5. Borowczyk W., Kaczmarek W., *Analiza i ocena błędów wyliczania nastaw działowych w przelicznikach przeciwlotniczych zestawów artyleryjskich bliskiego zasięgu*, Biuletyn WAT, Vol. LVII, nr 1, 2008, pp. 19-30.
6. Machowski B., Motyl K., *Analiza możliwości opracowania przelicznika dla systemu artyleryjskiego na podstawie równań ruchu pocisku*, Biuletyn WAT, nr 4 [548], 1998, pp. 109-117.
7. Borowczyk W., Kaczmarek W., *Obliczanie bieżących współrzędnych celu powietrznego w oparciu o współrzędne mierzone przez odległe urządzenia śledzące*, Biuletyn WAT, Vol. LVI, nr 3, 2007, pp. 211-228.

8. Madej W., *Specjalizowany cyfrowy system predykcijny z prostą hipotezą o ruchu celu*, Zeszyty Naukowe WEiI, nr 2, 2010, pp. 15-36.
9. Madej W., Drabarek J., *Realizacja odczytu współrzędnych w specjalizowanym cyfrowym systemie predykcijnym z wykorzystaniem reprogramowalnych układów cyfrowych*, Elektronika, nr 11, 2008, pp. 189-190.
10. Bartkiewicz S., Madej W., *Błędy strzelania jako miara jakości zintegrowanego systemu uzbrojenia*, KKM Warszawa 2001, pp. 577-580.
11. Bartkiewicz S., Ociepa Z., Boroń S., Madej W., *Cyfrowa realizacja łącza selsynowego*, V Krajowa Konferencja „Automatyzacja i eksploatacja systemów sterowania”, Gdynia 1995
12. Boroń S., Madej W., Bartkiewicz S., Ociepa Z., *Artyleryjski przelicznik cyfrowy*, V Krajowa Konferencja „Automatyzacja i eksploatacja systemów sterowania”, Gdynia 1995.
13. Madej W., Boroń S., Bartkiewicz S., Ociepa Z., *Możliwości zwiększenia skuteczności artyleryjskiego zestawu przeciwlotniczego*, V Krajowa Konferencja „Automatyzacja i eksploatacja systemów sterowania”, Gdynia 1995.
14. Ociepa Z., Bartkiewicz S., Boroń S., Madej W., *Cyfrowe sterowanie napędami armaty*, V Krajowa Konferencja „Automatyzacja i eksploatacja systemów sterowania”, Gdynia 1995.

## Streszczenie

Specjalizowane cyfrowe systemy predykcyjne są niezbędnym elementem zautomatyzowanego systemu kierowania ogniem artylerii przeciwlotniczej. Stanowią one podstawowy podsystem, który realizuje zadania modułu balistycznego. Moduł ten odpowiada za wypracowanie odpowiednich nastaw do wykonania zadania ogniowego. W pracy przedstawiono algorytm obliczeń oraz niezbędne funkcje do prawidłowego działania specjalizowanego systemu predykcijnego realizującego predykcję punktu wyprzedzonego dla małokalibrowej artylerii przeciwlotniczej. Algorytm ten może być zaimplementowany w komputerze przemysłowym, systemie mikroprocesorowym z procesorem sygnałowym lub w sterowniku programowalnym PLC, który posiada możliwość programowania w języku C.

## Abstract

Specialized digital prediction systems are necessary elements of the automated system of the fire control of anti-aircraft artillery. They constitute the basic subsystem which realizes tasks of the ballistic module. This module is responsible for working out appropriate settings for performance of the fire task. In the paper the algorithm of computations and necessary functions for the proper operating of the specialized prediction system realizing prediction of the meeting point. i.e. coordinates of the point

in which, according to the computations, the meeting of the target and the missile should take place, for small caliber anti-aircraft artillery are presented. This algorithm can be implemented using industrial computers, microprocessor systems containing a signal processor or a programmable logic controller (PLC) which can be programmed in the C programming language.





**Robert Świta**

**Zbigniew Suszyński**

Katedra Systemów Multimedialnych i Sztucznej Inteligencji

Wydział Elektroniki i Informatyki

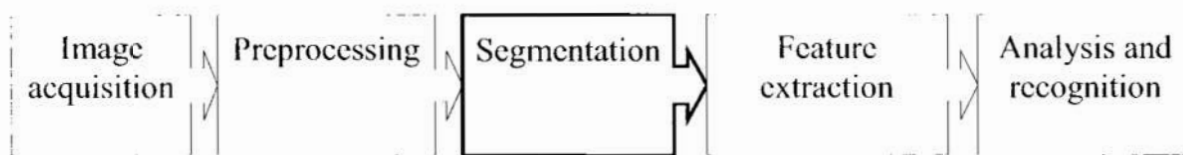
Politechnika Koszalińska

## Hybrid watershed and region grow segmentation algorithm of color images

**Keywords:** segmentation, watershed, region grow, spanning tree, HSV

### Introduction

Image recognition is in general a multi-stage process showed in the diagram



**Fig 1.** Image recognition diagram

Acquisition means sampling and quantization of a video signal in time and magnitude, preprocessing – normalization of image dimensions and noise filtration. The core stage of a recognition is yet segmentation, which results in recognition of objects contained in the image. After segmentation, taking into account purpose of image processing, feature extraction and final classification takes place.

Segmentation is a process of object recognition by means of grouping pixels with common characteristics (brightness, hue, location or texture).

At the end of a process, the image is tiled by objects, which happened to be separated regions build by pixels with common label. Optical Character Recognition System for example segments images and recognize alphanumeric characters. Classic segmentation algorithms work with luminance images, but color images contain additional information, which can be of great use for segmentation purpose.

Stereoscopic systems may use image segmentation for a 3D scene reconstruction.

One can classify segmentation methods as:

- Pixel-oriented, which doesn't take into account pixel's neighborhood (sibling pixels). The main advantage of those algorithms is their speed. It makes possible to optimize segmentation through iteration. The resulting objects are in general not consistent,

build in fact by few separated pixel regions. Typical algorithms of this kind are threshold and cluster methods (K-means) with labeling pixels to a priori known cluster-objects count.

- Area (region and edge) based – pixels are joined upon edge information (gradient) and sibling pixels similarity. Most popular methods are: split & merge, watershed, region growing. Hybrid methods can be very effective [2]
- PDE – Segmentation upon resolving partial differential equations. Most common use this kind of segmentation is in motion pictures, tracing objects and 3D scene reconstruction. The shape of objects is determined by parameters of cubic B-splines minimalizing cost function.
- Model based – main assumption is repeatability of object shapes and using statistical approach. That allows to establish a common pattern and a range of variability. An example of that kind of segmentation is Hough<sup>1</sup> transform, which allows exact recognition of object represented by parametric equations (lines, ellipses, curves, planes, cylinders etc.). In parameter space (Hough space) one can explore similarity of these parameters for defined group of candidate pixels.
- Artificial Neural Networks – special kind of two-dimensional ANN are used for segmentation purposes, which minimize gradient function between sibling pixels. They are called Kohonen<sup>2</sup> maps or Self Organizing Maps (SOM). Pulse Coupled Neural Networks imitate how cats vision cortex works.

## Goals

Hybrid watershed and region grow segmentation algorithm is part of a system which converts visual information into sounds. This system is dedicated and designed for assistance blind people in their selfmobility. It's implementation called SoundVisor is a subject of my Phd work. Segmentation module is one of the core parts of the system and should work very fast, requiring real-time processing. Because segmented images will often be irregular illuminated, the algorithm should be invariant from changes in illumination.

### 1. Gradient function of color images

To compare pixel values in color images defining of cost function is needed. **RGB** values for a color doesn't match human perception of colors. Easily interpretable parameters of colors, such as hue, saturation and brightness are used by **HSV** model. Quite often undesired effect is also change in brightness of pixels as a result of irregular illumination of an object. For these reasons it is better to process image in **HSV** space,

---

<sup>1</sup>Method patented by Paula Hough in 1962

<sup>2</sup>Teuvo Kohonen, professor at the Helsinki University of Technology and the Academy of Finland

ignoring Value component (brightness) for color pixels and Hue component for gray pixels. Color pixels from gray ones are classified based on Saturation component threshold, which defines distance of a color from gray line (HSV cone height). Output gradient depends on gradients of HSV pixels color components.

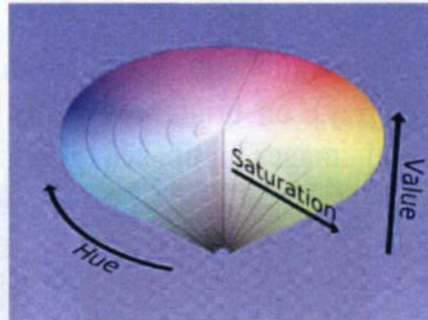


Fig 2. HSV color space

Because Hue component of HSV color model is an angle (normalized here to the range  $<0,1>$ ), gradient Chroma between hues of two pixels P i R is defined by difference of angles or complement of these (the lower one):

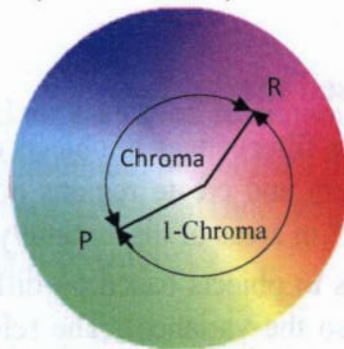


Fig 3. Definition of Hue gradient

Gradient for the pixels P i R could be evaluated as (MATLAB code):

```

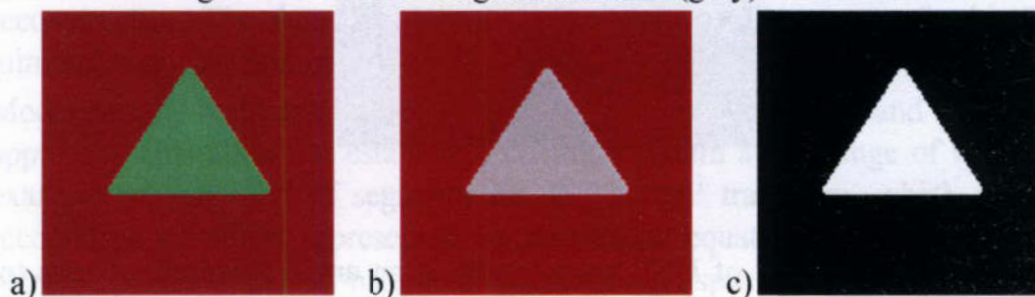
Chroma = abs(P.Hue-R.Hue);
if (Chroma > 0.5) Chroma = 1-Chroma; end; % Chroma = min[Chroma, 1-Chroma]
Chroma = 2*Chroma; % Chroma gradient normalization for range <0,1>
Lumina = abs(P.Val-R.Val);
Satura = abs(P.Sat-R.Sat);
if (P.Sat > GRAY) && (R.Sat > GRAY) % color pixels, gray threshold=GRAY
    Gradient = max([Satura,Chroma]);
else
    Gradient = max([Satura,Lumina]);

```



Colored pixels are combined if the difference in hue and saturation will be small (you can also take into account only the gradient of color Chroma). When one of pixels is considered gray, then saturation and luminance gradients will be compared (gray pixel can be joined to a color pixel if the color is weakly saturated and have similar brightness).

Two objects may vary in hue (although the same luminance and saturation), saturation (one of the objects gray) or brightness. This definition allows correct segmentation of the gradient of both images luminance (gray) and colored.



**Fig 4.** Two objects varying only with: hue (a), saturation (b) or brightness (c)

## 2. Region and edge based methods

Using the edge detectors (high-pass filter) edges of objects can be detected. However, we do need a closed edges to fill them and create regions of pixels. The edge algorithms have so the tendency to "spill over" in areas of discontinuity of edges.

Region methods assign pixels to objects based on difference from the mean, or on standard deviation (sometimes also the variance). The reference value can be changed dynamically when you join the next pixel to the object.

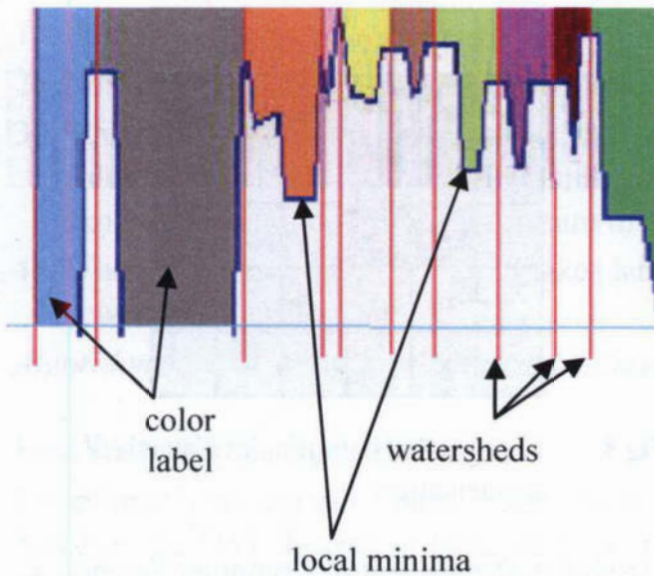
It is possible to combine both techniques. Hybrid methods, as it turns out, can be very effective [2].

## 3. Watershed segmentation

### Flood definition

As an input image serves gradient of the source image. Gradient is treated as topographic map in which pixel value is interpreted as height (the image as a 3D function graph)

Local minima in such image determine the most similar pixels to their neighbors. These minima form pools (objects). The image is flooded, in places where water from different pools joins, the dams are placed (the edges of objects)



**Fig 5.** Image cross-section in watershed segmentation



**Fig 6.** Image flooding in watershed algorithm

### Flood algorithm

- 1) Assign a different label for each local minimum and insert them to set S
- 2) Take out pixel x with a minimum height from the set S. Assign the label of pixel x to all pixels y that do not have labels and are adjacent to x. Put y to S
- 3) Repeat step 2, until S will be empty

*Despite the brevity, the algorithm is very inefficient to implement. It is necessary to fix local minima in advance, and in each iteration of the loop we must find minimum values in the set S*

### Topographic definition and edge weighted graph

The equivalent to the flood definition, topographic definition of watershed segmentation is determined by the principle of a water drop, which flows in the direction of a local minimum. All the pixels, which drops streaming down to the same minimum create a common pool (the object). Number of minima determine the number of pools. Pixels which drops can flow into at least two pools form ridges (edges of objects).

Segmentation problem can be represented in the form of edge-weighted graph, where vertices are the pixels in the image and the edge weights is a function which differs pixels (eg. gradient). Minimum spanning tree (MST) create flowing streams of pixels to a local minimum by selecting the edge with the smallest weight. MST graph spans all vertices (pixels). Each subgraph contains the minimum and is a separate object.



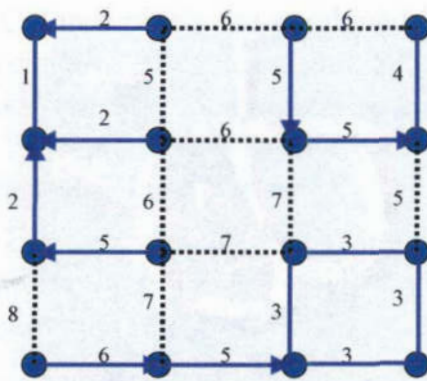


Fig 7. Edge weighted graph

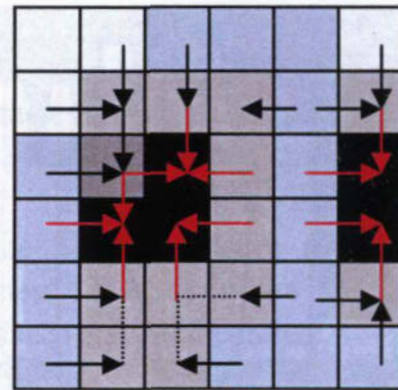


Fig 8. Topographic definition of watershed segmentation

In 2009 a group of French researchers from the Department of Computer Science at the University of Paris in Marne-la-Vallée: J. Cousty, G. Bertrand, L. Najman and M. Couprie [1] published a watershed segmentation algorithm based on the topographic definition, in which, using the weighted graph, optimum spanning tree (OST) is determined.

Optimum Spanning Tree Algorithm (J.Cousty, Bertrand, Najman & Couprie – 2009)

```

foreach  $x \in V$ , dla których  $\text{label}(x) == \text{NO\_LABEL}$  do
   $L \leftarrow \{x\}$ ;  $LW \leftarrow \{x\}$ ;  $\text{lab} = \text{NO\_LABEL}$ ; /*  $LW$  is a  $L$ 's front wave */
  while  $\text{Exists}(y \in LW)$  do
     $LW.\text{Remove}(y)$ ;  $\text{breadth} = \text{true}$ ;
    while  $\text{breadth}$  and  $\text{Exists}\{y, z\} \in E, z \notin L, F(\{y, z\}) == F_{\min}(y)$  do
      if  $\text{label}(z) \neq \text{NO\_LABEL}$  then  $\text{lab} = \text{label}(z)$ ;  $LW = []$ ; break;
      elseif  $F_{\min}(z) < F_{\min}(y)$  then /* depth exploration */
         $L \leftarrow L \cup \{z\}$ ;  $LW \leftarrow \{z\}$ ;  $\text{breadth} = \text{false}$ ;
      else /*  $F_{\min}(z) = F_{\min}(y)$ , breadth exploration */
         $L \leftarrow L \cup \{z\}$ ;  $LW \leftarrow LW \cup \{z\}$ ;
      if  $\text{lab} == \text{NO\_LABEL}$  then  $\text{lab} = \text{no\_labs}++$ ;
      foreach  $y \in L$  do  $\text{label}(y) = \text{lab}$ ;

```

where:

$V$  – set of graph vertices (pixels)

$LW$  –  $L$ 's front wave

$E$  – set of edges

$F()$  – gradient between vertices

$L$  – stream



- 1) Process every pixel in the image.
- 2) Non-labeled pixel becomes stream source.
- 3) Stream flows through graph nodes into local minimum, choosing lowest gradient value. If nodes have the same minimal gradient value, then pool is created (breadth exploration) until drain is found (return to depth exploration).
- 4) When the stream meet another, it takes label of that stream. If it doesn't – it receives new, next numbered label.

*Algorithm computationally efficient, quasi-linear dependence on the size of the image*

### **FastWaterShed Algorithm**

For comparison purposes I have created in MATLAB environment watershed algorithm based on the OST. Instead of building a special wave of expansion, existing stream and pixel expansion counter was used. During construction, the stream gradient is stored in the current stream. Minimum pixel gradient is compared to the current gradient of the stream (if it's equal – we explore in breadth, if it is smaller – we explore in depth). Construction of the stream ends when the expansion count drops to zero.

### **Drawbacks and advantages of watershed algorithm**

Algorithm based on the OST is relatively fast, its computational complexity is  $\sim O(n)$ . It does not create a thick edge between objects, because the streams are "tracked" from a same source - the ridges. The disadvantage is oversegmentation (too many local minima found), which can be limited by reducing the levels of possible values for the gradient (posterization), although it makes easily spill out the algorithm, because pixels are not associated with any object reference value.

## **4. Region grow**

Filling the area starts from a pixel called a seed in the horizontal and vertical directions (4-pixel neighborhood) and also diagonal in 8-pixel. The algorithm is very similar to the Flood Fill algorithm used in computer graphics. Due to the large number of pixels in images however, recursive implementation is avoided.

### **FastSeed algorithm**

This algorithm is based on the principle of spreading, which i've implemented in MATLAB environment. Processed pixel is treated as a seed. If any of the neighbors of the pixel satisfies the condition of belonging to an object, it is attached to an object and becomes the seed. Requirement for membership is determined by the absolute difference between the pixel value and a reference object value (reference gradient). The reference value is changed when adding a pixel to the object so that it is always equal to the average of all pixels.

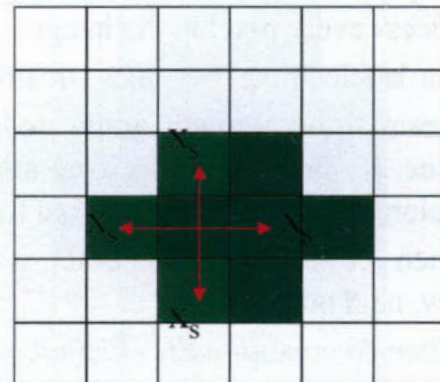
### Region grow algorithm

```

Object.Add(x);
seedCount = 1;
while (seedCount > 0) {
  seedCount--;
  x = Object(Object.Size-1-seedCount);
  foreach (xs: neighbor(x) && Label(xs) == 0)
    if (member_criterion(xs, Object)) {
      Label(xs) = &Object;
      Object.Add(xs);
      seedCount++;
    }
}

```

*Recursion is inefficient and could cause a stack overflow*



**Fig 9.** Region grow in 4-pixel neighborhood

### Drawbacks and advantages of region grow algorithm

The algorithm is very fast, linearly dependent on the size of the image. Objects include only alike pixels. The algorithm does not "spill" because the pixels are related to the reference value of the object. The disadvantage is edge detection. On the edges many small objects are created.

## 5. Hybrid watershed and region grow segmentation

Watershed algorithm based on optimum spanning tree and region grow algorithm are featured by high speed processing and linear computational complexity. The principle of their operation and the result of segmentation, however, are completely different. The watershed algorithm is an edge detection algorithm with a tendency to "spill over" to neighboring objects. Algorithm of region growth has problems with classification of pixels lying between two objects (edges). Therefore, I have made attempt to combine both methods and created a hybrid algorithm, named FastWaterSeed, which combines the advantages of both techniques

### FastWaterSeed algorithm

- 1) Process every pixel in the image.
- 2) Non-labeled pixel becomes stream source (watershed algorithm).
- 3) If the stream leads to local minimum, it becomes a seed and region grow algorithm is engaged. In case when stream meet existing pool, it joins with a pool.

FastWaterSeed algorithm takes advantages of both seeding and watershed algorithms, while eliminating their drawbacks. It flood objects up to the edge with the next object (does not create thick edges). It doesn't spill because the pixels are associated with reference value, which is the only parameter of the algorithm. It is very important that the local minimum determined by the stream becomes the seed of growth, because it connects other, lying nearby, similar local minima and significantly reduces oversegmentation. Therefore, there is no need to posterize image. The algorithm is fully automatic and does not require markers as in the case of semi-automatic watershed methods. Hybrid segmentation retains linear number of operations on the size of the image. In the case of real images, with a lots of objects hybrid algorithm is also faster than both kinds of segmentation.

## **6. Comparison of area segmentation algorithms**

To compare the speed and quality of segmentation FastSeed (FS), FastWaterShed (FWH) and FastWaterSeed (FWS, hybrid) algorithms have been used. All use predetermined gradient of HSV when comparing pixels, and in the case of color pixels gradient is equal to the hue gradient (gradient = Chroma). The neighborhood of pixels was 8-pointed. For each method processing time and below number of identified objects, which determines oversegmentation is given. For the watershed algorithm no posterization is used for gradient values. To make it easier to compare the results of segmentation, the objects were sorted from the largest (containing the most pixels) to the smallest. Then, for all segmentation the same pseudocolor to objects with the same index in the sorted array have been assigned. This means that the largest objects in different segmentation have the same pseudocolor. In the same way one can limit changes of pseudocolors during segmentation of moving pictures (from video or camera).

## **Summary**

Segmentation as a recognition of objects in an image by grouping pixels with similar properties is an essential part of the image recognition process. The article reviews the different methods and classifications. It present latest area segmentation algorithms and a new, universal method formed by merging functionality of region grow and watershed segmentation based on the optimal spanning tree [1]. Hybrid algorithm takes advantages of both methods - bounding pixels with reference value like in region grow (the algorithm does not "spill out" on other objects), and attach the edges to the object - as in the case of watershed segmentation. At the same time drawbacks of these algorithms are eliminated, because disadvantages of one method are advantages in another. It's computational complexity is quasi-linear with respect to image size  $\sim O(n)$ .

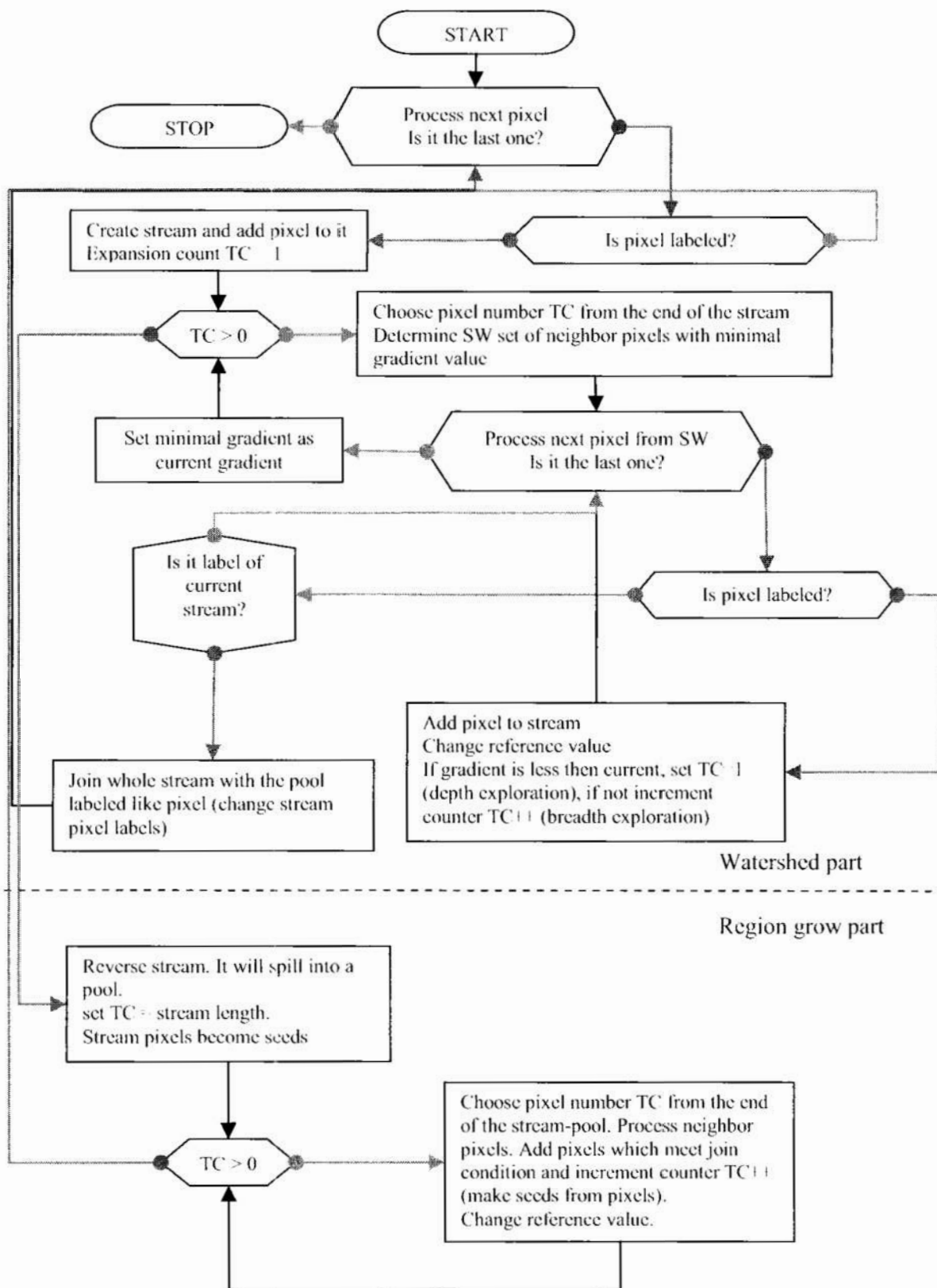
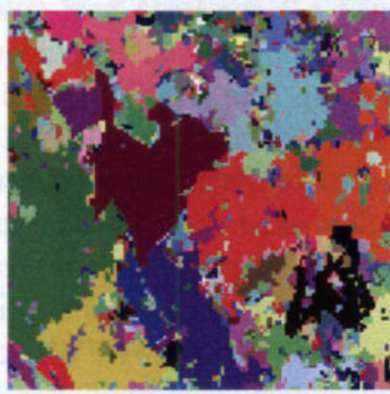


Fig 10. Diagram of hybrid watershed and region grow segmentation algorithm





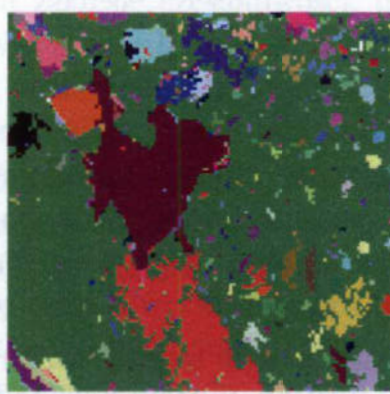
FWH  
0.85s  
1327



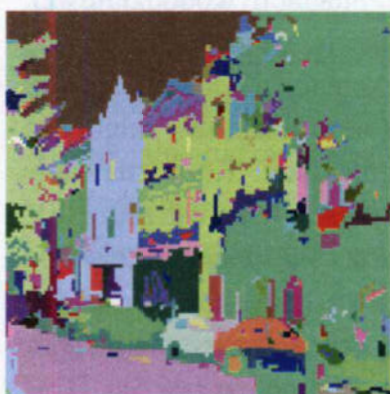
FWH  
0.82s  
1062



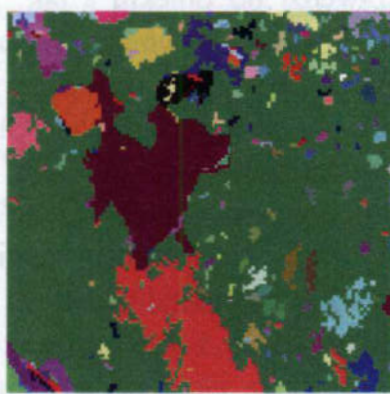
FS  
0.50s  
928



FS  
0.48s  
548



FWS  
0.43s  
413



FWS  
0.34s  
248



## Bibliography

1. J.Cousty, G.Bertrand, L.Najman, M.Couprie. *Watershed Cuts: Minimum Spanning Forests and the Drop of Water Principle*. IEEE Transactions on Pattern Analysis and Machine Intelligence. 31 (8) , pp. 1362–1374. 2009.
2. Skarbek W., Koschan A. *Colour Image Segmentation - A Survey*. Technisches Bericht 94 32 Technische Universitat Berlin, 1994.

## Abstract

In this article I present a new color image segmentation method based on two state-of-the-art region segmentation algorithms – region grow and watershed created by Optimum Spanning Tree [1]. The hybrid algorithm takes advantages of both methods – binding pixels with reference value like in case of region growth, and joining edges to the objects, just like watershed algorithm does. In the same time drawbacks of those methods are eliminated. New method doesn't require posterization of pixel values and does not produce oversegmentation featuring watershed algorithms. It's operations complexity is linear and in case of real images, with a lot of objects, it's faster. Object recognition is made based on HSV (Hue, Saturation, Value) pixel color information, because segmentation algorithm tries to be invariant from changes of objects illumination.

## Streszczenie

W niniejszym artykule przedstawiam nową metodę segmentacji obrazów kolorowych, powstałą w wyniku połączenia dwóch algorytmów state-of-the-art segmentacji obszarowej: algorytmu rozrostu obszaru i segmentacji wododziałowej opartej na optymalnym drzewie spinającym [1]. Algorytm hybrydowy łączy w sobie zalety jednej jak i drugiej metody - związanie pikseli wartością referencyjną jak w przypadku rozrostu obszaru (algorytm nie „rozlewa się” na inne obiekty) oraz dołączanie krawędzi do obiektu - tak, jak w przypadku segmentacji wododziałowej. Metoda hybrydowa nie wymaga posteryzacji wartości pikseli wejściowych i nie wykazuje nadsegmentacji charakterystycznej dla segmentacji wododziałowej. W przypadku obrazów rzeczywistych, o większej liczbie obiektów, algorytm wododziałowo-rozrostowy jest również szybszy od obu segmentacji. Rozpoznawanie obiektów jest dokonywane na podstawie informacji o kolorze pikseli w przestrzeni HSV (barwie, nasyceniu i jasności), ponieważ algorytm stara się uniezależnić od nierównomiernego oświetlenia obiektów.

**Anna Witenberg**  
**Maciej Walkowiak**  
Katedra Telekomunikacji  
Wydział Elektroniki i Informatyki  
Politechnika Koszalińska

## **Rozwiązanie równania pola elektrycznego modelującego anteny liniowe w dziedzinie czasu**

**Słowa kluczowe:** antena liniowa, całkowite równanie pola elektrycznego, metoda kroczenia w czasie (MOT), rozkład prądu w dziedzinie czasu, niestabilności rozwiązania

### **1. Wprowadzenie**

Gdy żądamy dokładnych parametrów badanych anten w wąskim zakresie częstotliwości, to numeryczne modelowanie w dziedzinie częstotliwości jest wówczas efektywne. Jeśli mamy wyznaczyć parametry anten w szerokim zakresie częstotliwości, to obliczenia takie wiążą się z wykonaniem wielu obliczeń w rozmaitych punktach widma. Nic więc dziwnego, że w przypadku szerokopasmowej analizy zjawisk elektromagnetycznych związanych z odpowiedzią prądową anten wykorzystujemy dziedzinę czasu.

Punktem wyjścia badania tych zjawisk są równania Maxwella zapisane w postaci różniczkowej lub całkowitej. Prowadzą one do powszechnie stosowanego całkowitego równania dla pola elektrycznego TD-EFIE (Time Domain Electric Field Integral Equation).

Najpopularniejszym sposobem numerycznego rozwiązywania tego równania jest tzw. metoda kroczenia w czasie MOT (Marching-On In Time), której jawne i uwikłane odmiany przedstawiono w pracach [1, 2]. Klasyczne rozwiązanie równania TD-EFIE rozszerzano i uzupełniano stosując w nim między innymi:

- metodę wykorzystującą uwikłany schemat wykorzystywania kroków czasowych [1],
- specjalne rodzaje czasowych funkcji bazowych [3, 4],
- uśrednianie wartości prądu w czasie i w przestrzeni [6],
- filtrowanie wartości prądu w czasie [7].

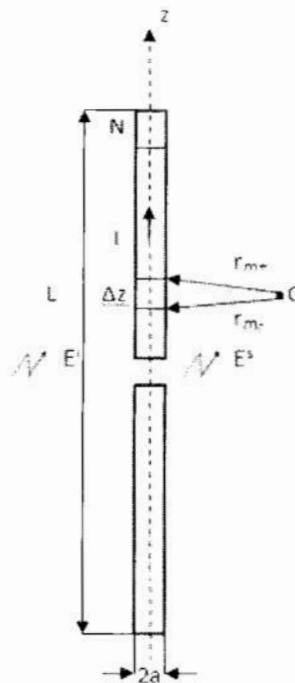
Każde z powyższych rozwiązań (algorytmów) ma wady i zalety, ale żadne z nich nie zapewnia bezwarunkowej stabilności przebiegu prądu w czasie. Żadna odmiana metody MOT nie jest wolna od niestabilności prądu mających charakter narastających oscylacji o wysokiej częstotliwości, które to oscylacje pojawiają się w późnoczasowej części rozwiązania.

Dodajmy, że istnieją bezwarunkowo stabilne algorytmy badania odpowiedzi prądowych, np. FDDM (Finite Difference Delay Modeling) [8] lub MOD (Marching-On in Degree) [5, 12], ale są one nieefektywne (długi czas obliczeń) i wymagają komputerów o bardzo dużych mocach obliczeniowych. Najpoważniejszym jednak zastrzeżeniem do tych algorytmów jest to, że nie wiemy, czy tłumienie oscylacji w późniejszej części rozwiązania nie jest okupione wprowadzaniem do całego rozwiązania błędów systematycznych.

W artykule przedstawiono odmianę metody MOT, wykorzystującą aproksymację skalarnego potencjału elektrycznego „centralnym” ilorazem różnicowym. Podano podstawy teoretyczne metody i wyniki obliczeń numerycznych dla anteny liniowej oświetlonej impulsem Gaussa.

## 2. Równanie TD-EFIE

Rozpatrzmy doskonale przewodzącą antenę liniową o długości  $l$  i promieniu  $a$  ( $a \ll l$ ) umieszczoną w ośrodku o przenikalności magnetycznej  $\mu = \mu_r \mu_0$  i przenikalności elektrycznej  $\varepsilon = \varepsilon_r \varepsilon_0$ . Oś anteny pokrywa się z osią  $z$  układu współrzędnych wykorzystywanego do analizy. Na antenę pada impuls fali elektromagnetycznej w ten sposób, że natężenie  $\mathbf{E}^i$  pola elektrycznego jest równoległe do osi  $z$ . Pole  $\mathbf{E}^i$  indukuje na powierzchni anteny prąd o gęstości liniowej  $\mathbf{J}(\mathbf{r}, t)$  i ładunki elektryczne o gęstości liniowej  $\mathbf{q}(\mathbf{r}, t)$ , które są źródłem promieniowania.



Rys. 1. Antena liniowa oświetlona impulsem Gaussa

Zmiany gęstości prądu  $\mathbf{J}$  wzdłuż osi  $z$  opisuje wektorowy potencjał magnetyczny  $\mathbf{A}(z,t)$ , natomiast zmiany gęstości ładunków  $q$  – skalarny potencjał elektryczny  $\Phi(z,t)$ .

Rozproszone pole elektryczne  $\mathbf{E}^s(\mathbf{r},t)$  z powierzchni anteny opisuje zależność [1]

$$\mathbf{E}^s(\mathbf{r},t) = -\frac{\partial \mathbf{A}(\mathbf{r},t)}{\partial t} - \nabla \Phi(\mathbf{r},t) \quad (1)$$

w którym  $\mathbf{A}(\mathbf{r},t)$  jest wektorowym potencjałem magnetycznym oraz

$$\mathbf{A}(\mathbf{r},t) = \frac{\mu}{4\pi} \int_L \frac{\mathbf{J}(\mathbf{r}',\tau)}{R} dl' \quad (2)$$

natomiast  $\Phi(\mathbf{r},t)$  oznacza skalarny potencjał elektryczny równy

$$\Phi(\mathbf{r},t) = \frac{\mu}{4\pi\epsilon} \int_L \frac{q(\mathbf{r}',\tau)}{R} dl' \quad (3)$$

W równaniach (1) - (3)  $R = \sqrt{|\mathbf{r} - \mathbf{r}'|^2 + a^2}$  jest odległością między dowolnie położonym punktem obserwacji  $\mathbf{r}$  a punktem źródła  $\mathbf{r}'$  na powierzchni anteny,  $\tau = t - \frac{R}{c}$  oznacza czas opóźnienia,  $c = (\mu\epsilon)^{1/2}$  jest prędkością rozchodzenia się fali elektromagnetycznej w ośrodku.

Gęstość prądu  $\mathbf{J}$  i gęstość ładunku  $q$  wiąże ze sobą równanie ciągłości

$$\nabla \mathbf{J}(\mathbf{r},t) = -\frac{\partial q(\mathbf{r},t)}{\partial t} \quad (4)$$

z którego dostajemy

$$q(\mathbf{r},t) = -\int_0^t \nabla \mathbf{J}(\mathbf{r},t') dt' \quad (5)$$

Po uwzględnieniu (5), zależność na skalarny potencjał elektryczny (3) zapisujemy w postaci

$$\Phi(\mathbf{r},t) = -\frac{1}{4\pi\epsilon} \int_L \int_0^t \frac{\nabla' \mathbf{J}(\mathbf{r}',t')}{R} dt' dl' \quad (6)$$

Jeżeli składowe styczne pól elektrycznych: padającego  $\mathbf{E}^i$  i rozproszonego  $\mathbf{E}^s$  na powierzchni anteny spełniają warunek graniczny

$$[\mathbf{E}^i + \mathbf{E}^s]_{\text{tan}} = 0 \quad (7)$$

wówczas całkowite równanie pola elektrycznego w dziedzinie czasu przybiera postać

$$\left[ \frac{\partial \mathbf{A}(\mathbf{r}, t)}{\partial t} + \nabla \Phi(\mathbf{r}, t) \right]_{\tan} = [\mathbf{E}^i(\mathbf{r}, t)]_{\tan} \quad (8)$$

Wskaźnik „tan” w (8) oznacza składową styczną pola elektrycznego na powierzchni anteny.

### 3. Metoda MOT (Marching – On In Time)

Rozwiązanie numeryczne równania TD-EFIE wymaga podzielenia osi czasu na interwały o jednakowej długości  $\Delta t$  (przyjmujemy  $t_k = k\Delta t$ ) i podzieleniu anteny na  $N$  segmentów, każdy o długości  $\Delta z$ .

Pochodne względem czasu w równaniu (8) można aproksymować ilorazami różnicowymi; przyjmujemy  $\mathbf{r} = z$  (punkt obserwacji ma tylko współrzędną  $z$ ) [1,2]

- „do przodu” (forward) przez podstawienie  $t = t_{k-1}$

$$\left[ \frac{A(z, t_k) - A(z, t_{k-1})}{\Delta t} + \Delta\Phi(z, t_{k-1}) \right]_{\tan} = [E^i(z, t_{k-1})]_{\tan} \quad (9)$$

- „wstecz” (backward) przez podstawienie  $t = t_k$

$$\left[ \frac{A(z, t_k) - A(z, t_{k-1})}{\Delta t} + \Delta\Phi(z, t_k) \right]_{\tan} = [E^i(z, t_k)]_{\tan} \quad (10)$$

- „centralnie” (central) przez podstawienie  $t = t_{k-1/2}$

$$\left[ \frac{A(z, t_k) - A(z, t_{k-1})}{\Delta t} + \frac{\Delta\Phi(z, t_k) + \Delta\Phi(z, t_{k-1})}{2} \right]_{\tan} = [E^i(z, t_{k-1/2})]_{\tan} \quad (11)$$

Połączenie równań (9) – (11) w jedną zależność pozwala nam napisać ogólny wzór

$$\left[ \frac{A(z, t_k) - A(z, t_{k-1})}{\Delta t} + (1-p)\Delta\Phi(z, t_k) + p\Delta\Phi(z, t_{k-1}) \right]_{\tan} = [E^i(z, t_{k-p})]_{\tan} \quad (12)$$

w którym dla:

- $p = 1$  dostajemy iloraz różnicowy „do przodu”,
- $p = 2$  dostajemy iloraz różnicowy „wstecz”,
- $p = 1/2$  dostajemy iloraz różnicowy „centralny”.

Aproksymacja pochodnych w równaniu (8) „centralnym” ilorazem różnicowym (równanie (11)) zapewnia największą dokładność i stabilność rozwiązania równania TD-EFIE.

Powyższe rozważania prowadzą do wniosku, że dla wyznaczenia wartości prądu w chwili  $t_k$  muszą być znane jego wartości dla wcześniejszych interwałów czasowych.



Równanie (11) przepisujemy w dogodniejszej postaci

$$2A(z, t_k) + \Delta\Delta\Phi(z, t_k) = 2\Delta E'(z, t_{k-1/2}) + 2A(z, t_{k-1}) - \Delta\Delta\Phi(z, t_{k-1}) \quad (13)$$

i do jego rozwiązania stosujemy metodę momentów.

Istotną cechą tej metody jest zastąpienie (aproksymacja) nieznannej wartości gęstości prądu  $J(z, t)$  zależnością

$$J(z, t) = \sum_{n=1}^N I_n(t_k) f_n(z) \quad (14)$$

w której  $I_n(t_k)$  są nieznanymi współczynnikami wartości natężenia prądu na  $n$ -tym segmencie, a  $f_n(z)$  jest przestrzenną funkcją bazową.

Jako zbiór funkcji bazowych wybrano

$$f_n(z) = \begin{cases} 1 & , \text{ dla } z \in \langle z_n, z_{n+} \rangle \\ 0 & , \text{ pozostałe} \end{cases} \quad (15)$$

Przedstawiona procedura prowadzi do standardowego równania macierzowego [2]

$$[\alpha_{mn}] \cdot [I_{n,k}] = [V_{m,k-1/2}] - [\beta_{m,k-1}] \quad (16)$$

w którym:

$[\alpha_{mn}]$  – macierz niezależnych od czasu współczynników,

$[I_{n,k}]$  – macierz wartości natężenia prądu na  $n$ -tym segmencie w chwili  $k$ ,

$[V_{m,k-1/2}]$  – wektor wartości pola elektrycznego,

$[\beta_{m,k-1}]$  – wektor wartości współczynników natężenia prądu dla czasu  $t \in \langle 0, t_{k-1} \rangle$ .

#### 4. Wyniki obliczeń numerycznych

Antenę liniową o długości  $L = 1$  m i promieniu  $a = 0,002$  m podzielono na 20 segmentów o jednakowej długości i oświetlono w środku impulsem Gaussa modulowanym sygnałem harmonicznym, który opisuje zależność

$$E_z'(z, t) = \frac{4E_0}{cT\sqrt{\pi}} \exp \left[ - \left( \frac{4}{T} \left( t - t_0 - \frac{z}{c} \right) \right)^2 \right] \cdot \cos(2\pi f_0 t) \quad (17)$$

gdzie:

$E_0$  – wektor natężenia pola elektrycznego;  $E_0 = 120\pi$ ,

$T$  – szerokość impulsu,  $T = 6$  ns,

$t_0$  – przesunięcie maksimum impulsu;  $t_0 = 8$  ns,

$f_0$  – częstotliwość, dla której impuls przyjmuje wartość maksymalną w dziedzinie częstotliwości;  $f_0 = 0$  Hz.

W obliczeniach przyjęto długość odpowiedzi prądowej równą 400 ns.

Analizie poddano stabilność odpowiedzi, tzn. wartości prądów indukowanych na antenie w funkcji długości kroków czasowych  $\Delta t$ .

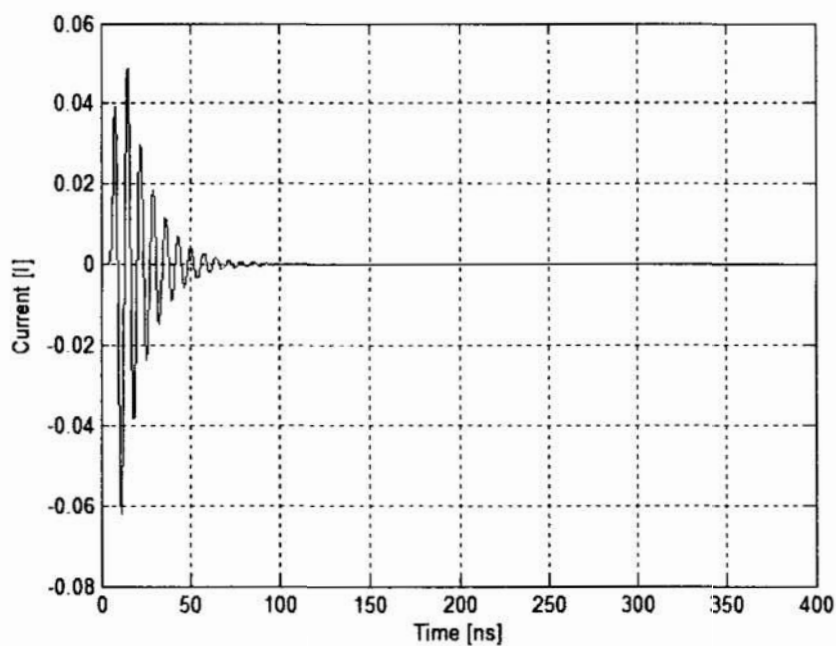
Obliczenia przeprowadzono dla trzech różnych wartości  $\Delta t$ :

$$- \Delta t_1 = 0,2 \cdot \frac{\Delta z}{c} = 0,03334 \text{ ns},$$

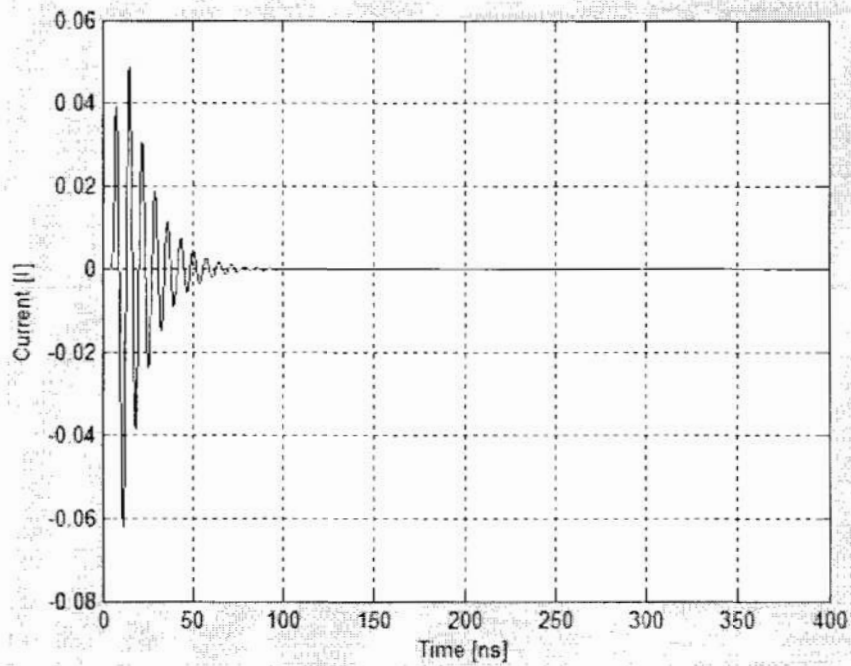
$$- \Delta t_2 = \frac{\Delta z}{c} = 0,1667 \text{ ns},$$

$$- \Delta t_3 = 2 \cdot \frac{\Delta z}{c} = 0,3334 \text{ ns}.$$

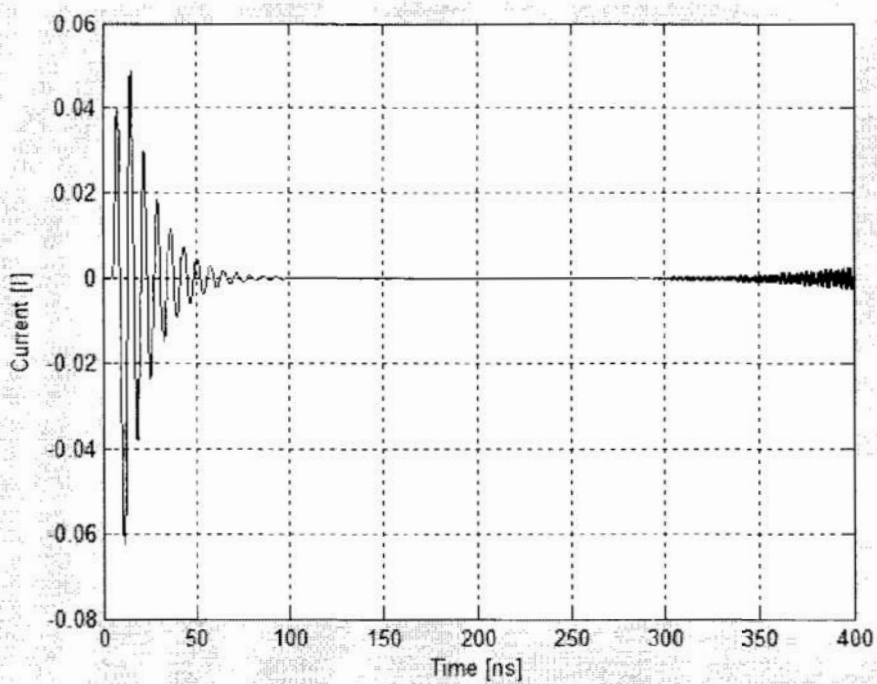
Wyniki symulacji pokazano na poniższych rysunkach.



Rys. 2. Odpowiedź prądowa anteny dla długości kroku  $\Delta t_1$



Rys. 3. Odpowiedź prądowa anteny dla długości kroku  $\Delta V_2$



Rys. 4. Odpowiedź prądowa anteny dla długości kroku  $\Delta V_3$

Ocenę stabilności odpowiedzi prądowej, poza oceną wizualną, dokonano przez podanie wartości zaindukowanego prądu po czasie około 15 ns ( $I_{15}$ ; maksymalna wartość prądu),

100 ns ( $I_{100}$ ), 398 ns ( $I_{398}$ ; maksymalna wartość prądu na końcu odpowiedzi), dla odpowiednich długości kroków czasowych.

Wyniki oceny stabilności odpowiedzi prądowej anteny dla różnych długości kroków czasowych  $\Delta t_i$ :

$$\Delta t_1 = 0,03334 \text{ ns}$$

$$I_{15} = 0,048644 \text{ A}$$

$$I_{100} = 0,00015747 \text{ A}$$

$$I_{398} = 3,2421 \cdot 10^{-13} \text{ A}$$

$$\Delta t_2 = 0,1667 \text{ ns}$$

$$I_{15} = 0,048634 \text{ A}$$

$$I_{100} = 0,00016137 \text{ A}$$

$$I_{398} = 5,7596 \cdot 10^{-13} \text{ A}$$

$$\Delta t_2 = 0,3334 \text{ ns}$$

$$I_{15} = 0,048856 \text{ A}$$

$$I_{100} = 0,00017306 \text{ A}$$

$$I_{398} = 0,0025428 \text{ A}$$

## 5. Podsumowanie

Dokładność i stabilność rozwiązania równania EFIE w dziedzinie czasu silnie zależy od długości (wartości) kroku czasowego przyjętego w procesie obliczeń. Przeprowadzone symulacje z anteną liniową oświetloną impulsem Gaussa prowadzą do wniosków:

- dla długości kroków czasowych  $\Delta t \leq \Delta z/c$  w odpowiedzi prądowej nie obserwujemy narastających z czasem oscylacji (rys. 2 i rys. 3); późnoczesowa odpowiedź prądowa ma postać eksponencjalnie tłumionej sinusoidy,

- dla długości kroków czasowych  $\Delta t > \Delta z/c$  późnoczesowa odpowiedź prądowa zawiera wyraźne i narastające z czasem niestabilności (rys. 4), które na przykład dla wartości  $\Delta t = 2\Delta z/c$  pojawiają się już od około 300 ns; maksymalne wartości prądu na końcu odpowiedzi ( $I_{398}$ ) są około  $10^{13}$  razy większe w porównaniu z maksymalnymi wartościami prądu w odpowiedziach stabilnych.

Długość kroku czasowego  $\Delta t$  ma również wpływ na dokładność i czas trwania obliczeń komputerowych; większe wartości  $\Delta t$  skracają czas obliczeń przy jednoczesnej utracie ich dokładności, szczególnie w późnoczesowej części odpowiedzi.

## Bibliografia

1. Rao S. M., *Time Domain Electromagnetics*, Academic Press, London, 1999.
2. Jung B. H., Sarkar T. K., *Time-domain electric-field integral equation with central finite difference*, *Microwave and Optical Technology Letters*, 2001, vol.31, no.6, pp. 429-434.

3. Manara G., *A space-time discretization criterion for a stable time-marching solution of the electric field integral equation*, IEEE Transactions on Antennas and Propagation, 1997, vol.45, no.3, pp. 527-532.
4. Weile S. D., Pisharody G., Chen N. Y., Shanker B., Michielssen E., *A novel scheme for the solution on the time-domain integral equations of electromagnetics*, IEEE Transactions on Antennas and Propagation, 2004, vol.52, no.1, pp. 283-295.
5. Chung Y. S., *Solution of time domain electric field integral equation using the Laguerre polynomials*, IEEE Transactions on Antennas and Propagation, 2004, vol.52, no.9, pp. 2319-2328.
6. Vechinski A. D., Rao S. M., *A stable procedure to calculate the transient scattering by conducting surfaces of arbitrary shapes*, IEEE Transactions on Antennas and Propagation, 1992, vol.40, no.6, pp. 661-665.
7. Sadigh A., Arvas E., *Treating the instabilities in marching-on-in-time method from a different perspective*, IEEE Transactions on Antennas and Propagation, 1993, vol.41, no.12, pp. 1695-1702.
8. Wang X., Wildman R. A., Weile D. S., Monk P. A., *A finite difference delay modeling approach to the discretization of the time domain integral equations of electromagnetics*, IEEE Transactions on Antennas and Propagation, 2008, vol.56, no.8, pp. 2442-2452.
9. Davis P. J., *On the stability of time-marching schemes for the general surface electric field integral equation*, IEEE Transactions on Antennas and Propagation, 1996, vol.44, no.11, pp. 1467-1473.
10. Makarov S. N., *Antenna and EM Modeling with Matlab*, John Wiley & Sons, New York, 2002.
11. Rao S. M., Wilton D. R., Glisson A. W., *Electromagnetic scattering by surfaces of arbitrary shape*, IEEE Transactions on Antennas and Propagation, 1982, vol.30, no.5, pp. 409-418.
12. Jung B. H., Sarkar T. K., Ji Z., Chung Y., *Time-domain analysis of conducting wire antennas and scatterers*, Microwave and Optical Technology Letters, 2003, vol.38, no.6, pp. 433-436.
13. Rao S. M., Sarkar T. K., *An efficient method to evaluate the time-domain scattering from arbitrarily shaped conducting bodies*, Microwave and Optical Technology Letters, 1998, vol.17, pp.321-325.
14. Liu T. K., Mei K. K., *A time domain integral equation solution for linear antennas and scatterers*, Radio Sci., 1973, vol.8.
15. Smith P. D., *Instabilities in time marching methods for scattering: cause and rectification*, Electromagn., 1990, vol.10, pp. 439-451.
16. Walkowiak M., *Zjawiska przejściowe w antenach liniowych i rozpraszaczach*, WSP, Zielona Góra, 1994.



17. Witenberg A., Walkowiak M., *Rozkład prądu wzdłuż anten liniowych – modele Pocklingtona i Hallena*, Zeszyty Naukowe Wydziału Elektroniki i Informatyki Politechniki Koszalińskiej, 2010, nr 2, s. 37-46.
18. Witenberg A., Walkowiak M., *Wykorzystanie wielomianów Laguerre'a do rozwiązania równania Hallena w dziedzinie czasu*, Metody Informatyki Stosowanej, Szczecin, 2011, nr 1, s. 183-192.

### **Streszczenie**

W artykule przedstawiono jedno z rozwiązań równania pola elektrycznego modelującego anteny liniowe w dziedzinie czasu – metodę postępu w czasie (MOT). Opisano zalety i wady tej metody, ze szczególnym uwzględnieniem wpływu długości kroku czasowego na wyniki uzyskane za pomocą algorytmu MOT – rozkłady prądów zaindukowanych w środku anteny. Podano warunek powstawania późnoczasowych niestabilności w rozkładach prądów i potwierdzono go wynikami obliczeń numerycznych.

### **Abstract**

This paper is focused on time domain modeling of wire antennas. We presented the marching-on time (MOT) method using the time domain electric field integral equation (TD-EFIE) formulation to obtain a transient scattering from antennas – an implicit version of this formulation with mathematical steps. The time derivative in TD-EFIE is approximated by central finite difference. The length of time step influences the stability of the MOT algorithm. This fact is discussed and demonstrated in numerical results.

**Malgorzata Górska**

Akademia Pomorska w Słupsku  
ul. Partyzantów 27, 76-200 Słupsk

**Antoni Drapella**

Akademia Marynarki Wojennej,  
ul. Śmidowicza 69, 81-103 Gdynia

## **Do primary event uncertainty distributions impact top event distribution?**

**Keywords:** fault tree, primary events distributions, top event distribution, uncertainty assessment, mode, uncertainty coefficient, reliability structure function, Monte Carlo simulation.

### **1. Introduction**

Fault trees are of vital importance for complex systems reliability assessment. They are useful for planning a variety of projects and operations at risk. Figure 1 shows an example of the fault tree created by Patrick O'Connor [3]. What is the structure of the fault tree? The tree comprises two kinds of events: primary and intermediate and consists of "AND" or "OR" gates. The top event is connected with intermediate and primary events by means of branches. Probability of the top event is uncertain since particular probabilities of primary events are uncertain as well. In particular, this paper deals with uncertainty of assessment of top event probability. The main aim is to answer the following question: do primary event uncertainty distributions impact the top event distributions in the fault tree? To answer the question four cases for primary event distributions have been employed: the Beta, Lognormal, Johnson  $S_J$  and truncated Normal distribution.

### **2. Calculation environment**

To achieve research aims computer simulation that involve random numbers, commonly known as the Monte Carlo method, was used. Monte Carlo method has been implemented using the Mathcad software package. Mathcad is as versatile and powerful as programming language, yet it is as easy to use as a spreadsheet. This mathematical environment turns out to be an ideal programming and calculation workspace for solving even the complicated problems of reliability engineering. In this paper we present the most important procedures only.

### 3. The function of the fault tree reliability

The „OR” and „AND” gates correspond to specified events and logical operations. In the „OR” gate the output event is takes place if any of the input events take place. From the point of view of the reliability this situation corresponds to the model of series structure [4].

Let  $n$  be the number of inputs in the gate and  $p_i, i=1,2,\dots,n$  are fault probabilities. The reliability structure function for this case is

$$F(p_i) = 1 - \prod_{i=1}^n (1 - p_i). \quad (2.1)$$

In the „AND” gate the output event occurs on condition that all of the input events take place. It corresponds to the model of parallel structure [4]. Appropriately for this case the structure function is

$$F(p_i) = \prod_{i=1}^n p_i \quad (2.2)$$

Figure 1 exemplified the tree of P. O'Connor. The reliability function of this tree is as follows

$$F(p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8) = 1 - \prod_{i=1}^4 p_i \cdot \left[ 1 - \left( 1 - \prod_{i=5}^6 p_i \right) \cdot \left( 1 - \prod_{i=7}^8 p_i \right) \right]. \quad (2.3)$$

The Mathcad procedure used to calculate values of the reliability function according to (2.3) is presented below

```

Proc_03(fT3) :=
  for i ∈ 1..10000
  | for j ∈ 1..8
  |   r_j ← 1 - fT3_{i,j}
  |   R_1 ← ∏_{k=1}^4 r_k
  |   R_2 ← ∏_{k=5}^6 r_k
  |   R_3 ← ∏_{k=7}^8 r_k
  |   R_c ← R_1 · [1 - (1 - R_2) · (1 - R_3)]
  |   fTte_i ← 1 - R_c
  | return fTte

```

This procedure is easily readable and does not require any additional descriptions or comments.

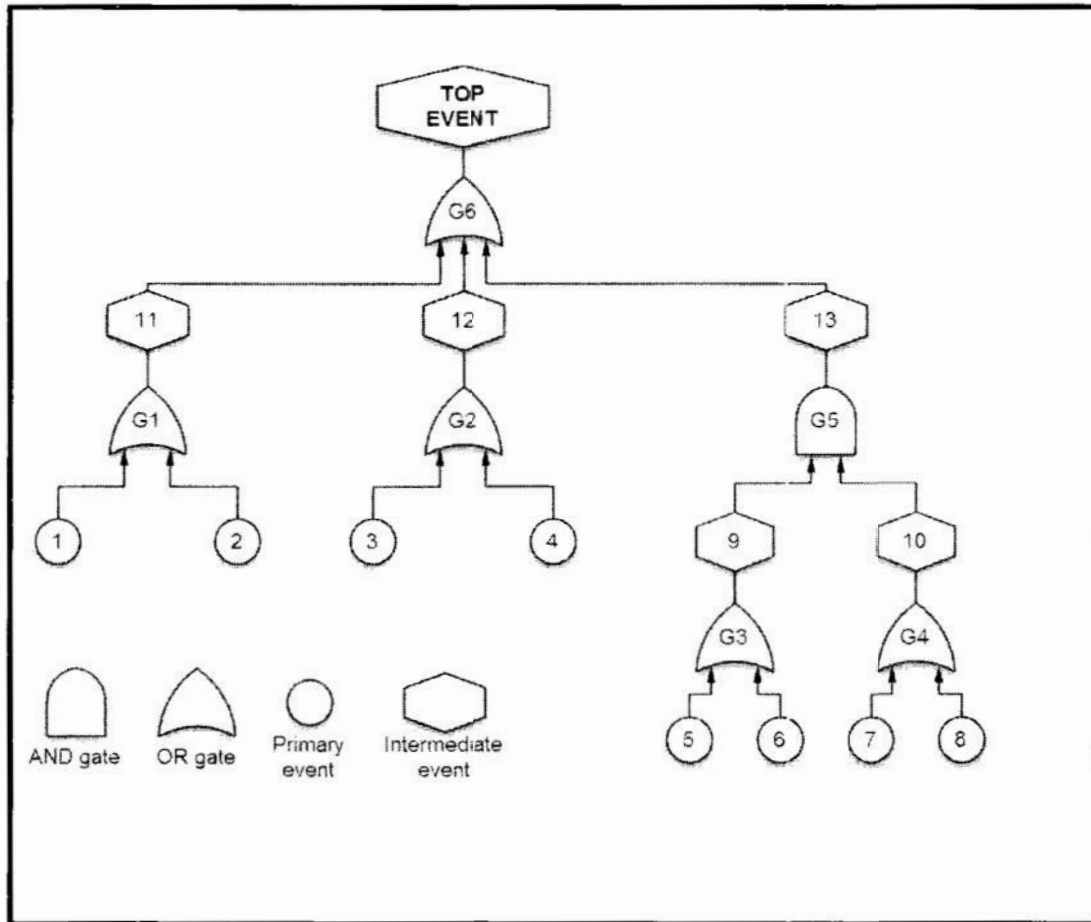


Fig. 1. The fault tree according to O'Connor

#### 4. Uncertainty and randomness

In this paper probabilities of primary events (PEPs) are considered random variables not determined constant. It is because PEPs are results of statistical inference and arbitrary engineering assessments procedures. Let  $n$  be a number of primary events. A set of input data comprises paired values. The first member of each pair is the most probable PEP value (i.e.  $pPEP$ ), the second one is PEP's uncertainty coefficient (PEPu).

PEPu is defined as follows (by Apostolakis [1])

$$PEPu = \frac{x_{0.95}}{x_m} \quad (3.1)$$

where  $x_{0.95}$  is such a value of  $x$ , called 95% quantile, which satisfies the equation  $F(x_{0.95})=0.95$  and  $x_m = pPEP$  is a maximum of density function called the mode.

The mode is understood as the most probable value the r.v. may take. In most cases, just pPEP and PEPu comprise input data.

Underneath denoted symbols were used in [2] for the first time:

**PEPs** - **PRIMARY EVENT PROBABILITIES**,

**pPEP** - the highest **probability of PRIMARY EVENT PROBABILITIES**,

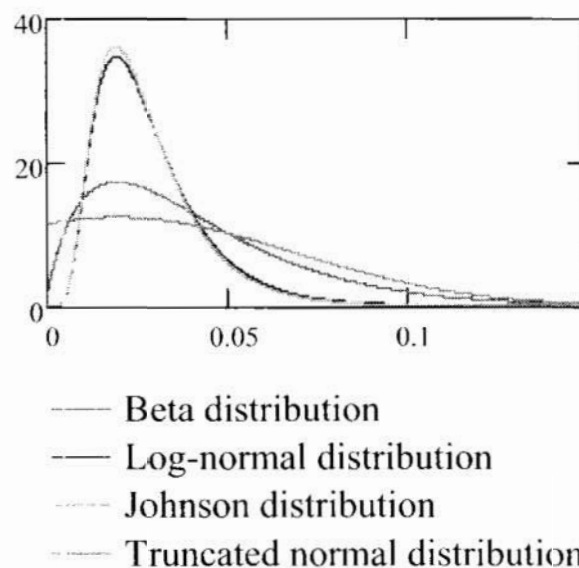
**PEPu** - **PRIMARY EVENT PROBABILITIES uncertainty coefficient**.

## 5. Input data

Table 1 contains all the input data related to primary event tree of Fig. 1. There are preferred values of primary events probabilities and uncertainty coefficients i.e. pPEP values and PEPu values.

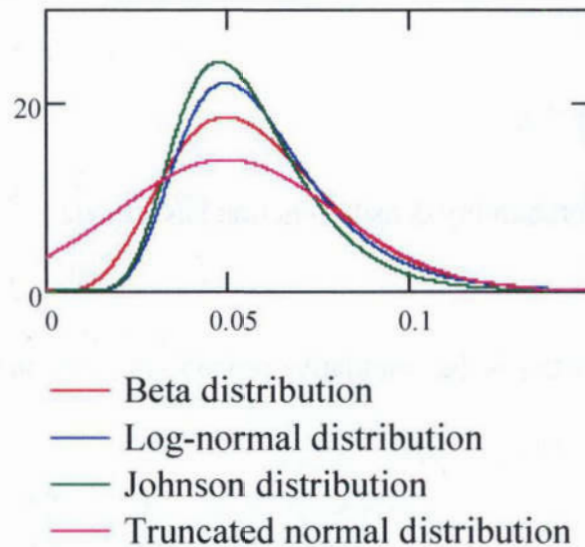
**Tab. 1.** Input data

Primary event	pPEP	PEPu
1	0,01	3
2	0,02	5
3	0,02	5
4	0,02	5
5	0,05	2
6	0,05	2
7	0,05	2
8	0,05	2



**Fig. 2.** Density function of the primary event number 2





**Fig. 3.** Density function of the primary event number 8

All the distributions have the same pPEP and PEPu values but they are diversified in respect of the density function shape. The skewness and the kurtosis are different in this case.

## 6. The scheme of monte carlo simulation

The simulation method was applied in four steps.

*Step 1:* Calculating  $a, b$  parameters values for primary events distribution.

*Step 2:* Generating  $x_{ij}$  input random numbers,  $i=1,2,\dots,10\,000, j=1,2,\dots,8$ .

*Step 3:* Calculating reliability function values (resulting vector comprises 10 000 values).

*Step 4:* Calculating values pPEP, PEPu and  $x_{95}$  quantile for output distribution.

Every calculation was done for four different distributions: beta, lognormal, Johnson and truncated normal.

## 7. The beta distribution employed

The first of two major characteristics of probability distribution is the probability density function denoted  $f(x)$ . In the case of beta distribution we have

$$f(x) = \frac{1}{B(a,b)} x^{a-1} (1-x)^{b-1} \quad (7.1)$$

where  $x$  denotes the random variable;  $a, b$  are beta distribution parameters ( $a > 0, b > 0$ ) and  $B(a,b)$  is the beta function.

By the definition

$$B(a,b) = \int_0^1 t^{a-1} (1-t)^{b-1} dt. \quad (7.2)$$

For  $a > 1$  and  $b > 1$  the probability density function has a mode

$$x_m = \frac{a-1}{a+b-2}. \quad (7.3)$$

The second characteristic is the cumulative probability function defined as follows

$$F(x) = \frac{1}{B(a,b)} \int_0^x t^{a-1} (1-t)^{b-1} dt. \quad (7.4)$$

In particular:

$$\frac{1}{B(a,b)} \int_0^{x_{0,95}} t^{a-1} (1-t)^{b-1} dt = 0,95. \quad (7.5)$$

The parameters  $a$  and  $b$  values we obtain after having solved system of equation (7.3) and (7.5) for values  $x_m$  (pPEP) and  $x_{0,95}$  ( $x_{0,95} = pPEP \cdot PEPu$ ).

The next step is the generation of random variables data and now we can calculate reliability structure function values (Proc\_03).

The final calculation step is to find the values of pPEP,  $x_{0,95}$  and PEPu for output distribution. To achieve this aim the following Mathcad procedure is applied. Particular rows of the procedure are commented in detail. Numbers in brackets are intended as reference numbers.

```

Proc_11(fX) := | nn ← rows(fX) [1]
                | fX ← sort(fX) [2]
                | xd ← fX2000 [3]
                | xg ← fXnn-2000 [4]
                | δ ← 10-6 [5]
                | xs ←  $\frac{x_g + x_d}{2}$  [6]
                | m ← 0 [7]
                | while 2 > 1 [8]
                |   | fd ← Proc_10a(xd, fX) [9]
                |   | fg ← Proc_10a(xg, fX) [10]
                |   | xd ← xs - δ if fg > fd [11]
                |   | xg ← xs + δ if fg < fd [12]
                |   | xs ←  $\frac{x_g + x_d}{2}$  [13]
                |   | m ← m + 1 [14]
                |   | TWPm,1 ← xd [15]
                |   | TWPm,2 ← xs [16]
                |   | TWPm,3 ← xg [17]
                |   | TWPm,4 ←  $\frac{f_g}{f_d} - 1$  [18]
                |   | break if  $\left| \frac{f_g}{f_d} - 1 \right| \leq 0.001 \vee m \geq 100$  [19]
                |   |  $\begin{pmatrix} x_s \\ fX_{9500} \\ fX_{9500} \\ x_s \end{pmatrix}$  [20]

```

The description of Proc\_11 procedure

- [1] - Ascribing the number of fX vector rows to nn variable.
- [2] - Sorting the rows of fX vector in ascending order.
- [3],[4] - Defining of the initial mode interval; ascribe values to  $x_d$  and  $x_g$ .
- [5] - Ascribing the values to  $\delta$  variable.
- [6] - Ascribing the values to  $x_s$  variable.
- [7] - Initialization of m variable and ascribing zero to m.
- [8] - Started the while loop.
- [9], [10] - Ascribing the density function values for  $x_d$  and  $x_g$  arguments to  $f_d$  and  $f_g$  variable, respectively. The Parzen's method (Proc\_10a) was employed to estimate the density function of the top event distribution.
- [11],[12] - Definition of the new limits of the mode interval; conditional ascribing the values to  $x_d$  and  $x_g$  variable.
- [13] - Ascribing the value to  $x_s$  variable.
- [14] - Incrementing the m value.
- [15], [16], [17], [18] - Ascribing the values to columns of TWP matrix.
- [19] - The exit condition of the while loop.
- [20] - Projection the vector of results.

We obtain the following result in the beta case:

pPEP	0,128
$x_{0,95}$	0,228
PEPu	1,776

## 8. The lognormal distribution employed

The lognormal distribution was derived from the well-known normal distribution. It is said that a given random variable follows the lognormal distribution if its natural logarithm follows the normal distribution.

In the case of lognormal distribution we have

$$f(x) = \frac{1}{\sqrt{2\pi} \cdot b \cdot x} \cdot \exp\left[-\frac{1}{2} \cdot \left(\frac{\ln x - a}{b}\right)^2\right], \quad (8.1)$$

where  $x$  denotes the random variable and  $a, b$  denote parameters.

The probability density function has a mode located at

$$x_m = pPEP = \exp(a - b^2) \quad (8.2)$$

The cumulative failure function takes the form

$$F(x) = \frac{1}{\sqrt{2 \cdot \pi} \cdot b} \int_0^x \frac{1}{u} \cdot \exp \left[ -\frac{1}{2} \cdot \left( \frac{\ln u - a}{b} \right)^2 \right] du. \quad (8.3)$$

If  $x$  follows the lognormal distribution then random variable

$$z = \frac{\ln(x) - a}{b} \quad (8.4)$$

follows the normal  $N(0,1)$  distribution. It can be easily read-out from tables of the  $N(0,1)$  distribution that corresponding 95% quantile is equal to 1.64. Consequently, we get

$$\frac{\ln(x_{0.95}) - a}{b} = 1.645 \Rightarrow x_{0.95} = \exp(a + 1.645 \cdot b). \quad (8.5)$$

Substituting (2b) and (5) into (6) we get

$$PEPu = \exp(b^2 + 1.645 \cdot b). \quad (8.7)$$

After transformation of formulas (2b) and (7) we obtain the following equations for  $a$  and  $b$  parameters

$$\begin{aligned} b &= \sqrt{\ln(PEPu) + 0.676} - 0.822 \\ a &= \ln(pPEP) + b^2 \end{aligned} \quad (8.8)$$

The Mathcad procedure named Proc\_11 returns the following result vector:

pPEP	0,101923
$x_{0.95}$	0,159174
PEPu	1,561699

## 9. The Johnson distribution employed

The Johnson distribution of  $S_B$  type has the following probability density function

$$f(x) = \frac{1}{\sqrt{2\pi} \cdot b \cdot x \cdot (1-x)} \cdot \exp \left[ -\frac{1}{2} \cdot \left( \frac{\ln \left( \frac{x}{1-x} \right) - a}{b} \right)^2 \right], \quad (9.1)$$

where  $x$  denote r.v. and  $a, b$  are parameters.

The cumulative distribution function

$$F(x) = \frac{1}{\sqrt{2\pi} \cdot b} \cdot \int_0^x \frac{1}{x \cdot (1-x)} \cdot \exp \left[ -\frac{1}{2} \cdot \left( \frac{\ln \left( \frac{x}{1-x} \right) - a}{b} \right)^2 \right] dx. \quad (9.2)$$



If  $x$  follows the Johnson distribution the random variable

$$z = \frac{\ln\left(\frac{x}{1-x}\right) - a}{b} \quad (9.3)$$

follows the normal  $N(0,1)$  distribution.

In particular

$$\frac{\ln\left(\frac{x_{0,95}}{1-x_{0,95}}\right) - a}{b} = 1,645 \quad (9.4)$$

Consequently

$$b = \frac{1}{1,645} \cdot \left[ \ln\left(\frac{x_{0,95}}{1-x_{0,95}}\right) - a \right] \quad (9.5)$$

For  $x_{0,95}$

$$\frac{df(x)}{dx} = 0 \quad (9.6)$$

Taking (9.5) and (9.6) into consideration we obtain the distribution parameters  $a$  and  $b$ .

The Mathcad procedure named Proc\_11 returns the following result vector.

pPEP	0,096
$x_{0,95}$	0,151
PEPu	1,581

The Beta, Lognormal and Johnson  $S_J$  distributions are a skewed to the right. The right tail is longer than the left side and the bulk of the values lie to the left of the mean. The truncated normal distribution is right-skewed as well; nonetheless the skewness is not such great in this case. This distribution presents more optimistic situation from the reliability point of view.

## 10. The truncated normal distribution

The probability density function of the truncated normal distribution takes the following form

$$f_c(x) = \begin{cases} \frac{1}{c} \cdot f(x) & \text{gdy } x \geq 0 \\ 0 & \text{gdy } x < 0 \end{cases} \quad (10.1)$$

where  $f(x)$  is probability density function of normal distribution

$$f(x) = \frac{1}{\sqrt{2\pi} \cdot b} \cdot \exp\left[-\frac{1}{2} \cdot \left(\frac{x-a}{b}\right)^2\right] \quad (10.2)$$

and

$$c = \int_0^x f(x) dx \quad (c < 1) \quad (10.3)$$

is constant.

The cumulative probability function

$$F_c(x) = \begin{cases} \frac{1}{c} \cdot \int_0^x f(x) g dy & x \geq 0 \\ 0 & \text{gdy } x < 0 \end{cases} \quad (10.4)$$

Left side cut-off normal distribution (10.1) and normal distribution (10.2) have the same values of  $x_m$  and  $x_{0,95}$ .

The appropriately Mathcad procedure gives the following result.

pPEP	0,148
$x_{0,95}$	0,249
PEPu	1,679

## 11. Conclusions

Since Apostolakis [1] has originated uncertainty analysis the lognormal and Johnson  $S_B$  distributions have gained popularity as basic uncertainty distributions. Moreover, one may treat the lognormal distribution as a special case of the Johnson  $S_B$  distribution when  $x \ll 1$ .

In ranges of  $x$  we deal with in uncertainty assessment these distribution are very similar.

In Reality nothing follows the lognormal, Johnson or any other distribution known in probability theory. The distributions are only mathematical models of Reality made for engineering purposes. This is so called "engineering judgment".

It may happen, despite this common believe, that actual uncertainty distribution of primary events will considerably differ from the Lognormal and Johnson distributions. If so, a question is sure to be asked: In what degree does this fact impact uncertainty distribution of the top event?

Table 2 is intended to compare output distributions in terms of pPEP and PEPu.

**Tab. 2.** Mode (pPEP) and uncertainty coefficient (PEPu) of output distributions

	TYPE OF DISTRIBUTION			
	BETA	LOG-NORMAL	JOHNSON	TRUNCATED NORMAL
pPEP	0.128	0,102	0,096	0,148
PEPu	1.776	1.562	1.581	1,679

Taking into account that simulations comprised 10000 repetitions one may conclude that mode and uncertainty coefficient of the top event distribution does not depend on the primary event distributions.

In other words: there is no significant relation between assumptions relating to primary events distribution and uncertainty assessment of output distribution.

## Bibliography

1. Apostolakis G., Lee Y.T., *Methods for the estimation of confidence bounds for the top event unavailability of fault trees*, Nuclear Engineering and Design. 41 (1977), pp 411-419.
2. Drapella A., Górska M., *Uncertainty Propagation in Fault Tree. A Monte Carlo Approach*. Network Centric Warfare, Naval Academy Gdynia 2009.
3. O'Connor P. D.T., *Practical Reliability Engineering*, Wiley 2002.
4. PN-EN 61078 (U), *Techniki Analizy Niezawodności*, KT 009 ds. Niezawodności 2002.

## Abstract

The risk analysis is an essential element of planning, production and operation of technical equipment. This paper deals with the fault tree. The fault tree analysis belongs to the most commonly used risk assessment methods. The main aim of the paper is to ask for the question: does the top event uncertainty assessment have a relationship with adopted assumption of primary events distribution? To achieve this aim a computer simulation that involves random numbers, commonly known as the Monte Carlo method, was used. The research makes use of the Beta, Lognormal, Johnson  $S_J$  and truncated Normal distribution.

**Keywords:** fault tree, primary events distributions, top event distribution, uncertainty assessment, mode, uncertainty coefficient, reliability structure function, Monte Carlo simulation.

## **Streszczenie**

Analiza ryzyka jest nicodzownym elementem w planowaniu, produkcji oraz działaniu sprzętu technicznego. Niniejszy artykuł dotyczy drzewa uszkodzeń. Analiza drzewa uszkodzeń należy do najbardziej powszechnych metod stosowanych w ocenie ryzyka. Głównym celem artykułu jest odpowiedź na pytanie: czy ocena niepewności zdarzenia wierzchołkowego ma związek z przyjętym założeniem, dotyczącym rozkładu zdarzeń pierwotnych? Do osiągnięcia tego celu wykorzystano związaną z liczbami losowymi metodę symulacji komputerowej, powszechnie znaną jako metoda Monte Carlo. Badanie zostało przeprowadzone przy wykorzystaniu rozkładów: beta, log-normalnego,  $S_J$  Johnsona oraz uciętego rozkładu normalnego.

**Słowa klucze:** drzewo uszkodzeń, rozkłady zdarzeń pierwotnych, rozkład zdarzenia wierzchołkowego, ocena niepewności, moda, współczynnik niepewności, funkcja struktury niezawodnościowej, symulacja Monte Carlo.