

Sprawozdanie z realizacji projektu badawczego 7 T11E 039 20:

METODY OPTYMALIZACJI REPREZENTACJI MEDYCZNYCH DANYCH OBRAZOWYCH DO ARCHIWIZACJI I TRANSMISJI TELEMEDYCZNEJ

Zasadniczym celem projektu było **stworzenie efektywnych narzędzi do archiwizacji i szybkiego przeglądania obrazów medycznych, a także przesyłania obrazów w sieciach komputerowych dla potrzeb teledygnicznej**. Planowano stworzenie kompletnego zestawu algorytmów kompresji niezbędnych do medycznych aplikacji, a efektem końcowym miała być **rozprawa habilitacyjna**. W opinii autorów cel ten został zrealizowany.

Przewidywany zakres prac badawczych zrealizowano, przy czym rozszerzono go o zaprojektowanie i wstępną realizację sieci teledygnicznej (intranet) w Zakładzie Diagnostyki Obrazowej Szpitala Wolskiego w Warszawie. Istniejąca infrastruktura została rozbudowana o trzy dodatkowe komputery (przewidywana w projekcie grantu stacja robocza, a także dodatkowy serwer i stacja robocza) i okablowanie łączące szereg pomieszczeń na dwóch kondygnacjach budynku Zakładu. Pozwoliło to zaimplementować opracowane algorytmy kompresji, także znacznie ułatwiło gromadzenie danych, wymianę i przeprowadzenie testów oceny subiektywnej. Ze względu na ograniczenia czasowe (krótki okres realizacji grantu) oraz finansowe nie udało się zrealizować stałego przyłącza do sieci Internet oraz wydzielonego łącza pomiędzy Szpitalem Wolskim a Politechniką (prace są w toku). Jednak powstała infrastruktura daje możliwość współpracy w kolejnych pracach badawczych, a bagaż wspólnych doświadczeń zaowocował zacieśnieniem współpracy naukowo-badawczej pomiędzy Zakładem Elektroniki Jądrowej i Medycznej PW a Zakładem Diagnostyki Obrazowej Szpitala Wolskiego, co należy zaliczyć do największych osiągnięć tego projektu.

Prowadzone w ramach projektu prace miały charakter teoretyczny i praktyczny, dotyczyły zarówno tworzenia pierwotnych koncepcji, budowania modeli, jak i ich algorytmicznej realizacji. Stworzono oprogramowanie, które było weryfikowane w testach. Opis prowadzonych eksperymentów, podsumowania i konkluzje ujmowane były w formie publikacji przy stałej pracy nad przygotowaniem, wreszcie ostateczną redakcją rozprawy habilitacyjnej.

Badania koncentrowały się na następujących obszarach tematycznych:

- Realizacja kompresji odwracalnej, optymalizowanej do zastosowań medycznych
 - [Pub11, rozdz. 5, szczególnie punkt 5.1]
 - [Pub3][Pub4][Pub5][Pub7][Pub8][Pub9][Pub13]
- Konstrukcja kodera falkowego do zastosowań teletransmisji
 - [Pub11, rozdz. 3, szczególnie punkty 3.5.3 i 3.5.4]
 - [Pub1][Pub2][Pub4][Pub7][Pub8]
- Optymalizacja kodera stratnego do zastosowań archiwizacji obrazów
 - [Pub11, rozdz. 3, szczególnie punkty 3.5.2 i 3.5.5]
 - [Pub4][Pub7][Pub8]
- Modelowanie i optymalizacja całego schematu kompresji falkowej
 - [Pub11, rozdz. 2, rozdz. 3, punkt 3.7]
 - [Pub9]

- Budowa wektorowej miary wiarygodności diagnostycznej obrazów
 - [Pub11, rozdz. 4, punkt 4.3]
 - [Pub12]
- Wyznaczenie wzorca diagnostycznego w testach oceny subiektywnej
 - [Pub6][Pub12][Pub14]
- Projektowanie i realizacja testów detekcji patologii
 - [Pub14]
- Praktyczne wykorzystanie algorytmów kompresji w medycznych systemach informacyjnych
 - [Pub11, rozdz. 5, punkt 5.2]
 - [Pub9][Pub10]

Przygotowana rozprawa habilitacyjna pt. „Falkowe metody kompresji danych obrazowych” ma 138 stron, 6 rozdziałów, 161 pozycji bibliograficznych. Opublikowano także 6 artykułów w czasopismach (w tym 4 międzyn.), kolejnych 4 w materiałach międzynarodowych konferencji i sympozjów. Poza tym zgłoszono jeszcze jeden artykuł do czasopisma międzynarodowego (jest w recenzji), a kolejne dwa są w przygotowaniu.

Do największych osiągnięć naukowych projektu należy zaliczyć obok wspomnianej rozprawy także stworzenie kompleksowego systemu kompresji obrazów medycznych do zastosowań we współczesnych szpitalnych systemach informacyjnych, wyposażanych w bogate narzędzia archiwizacyjne i telemedyczne. Zaproponowano także narzędzie obliczeniowej oceny wiarygodności diagnostycznej obrazów, oparte na własnej procedurze ustalenia wzorca diagnostycznego. Przeprowadzono subiektywne testy oceny diagnostycznej obrazów, w których 7 radiologów z trzech ośrodków medycznych oceniało ponad 100 obrazów mammograficznych. W przygotowaniu testu brało udział dwóch doświadczonych radiologów dokonujących selekcji obrazów testowych z ponad 200 zarejestrowanych w postaci cyfrowej badań. Analiza wyników testów doprowadziła do wniosku, iż graniczną wartością dopuszczalnej średniej bitowej falkowej reprezentacji obrazów mammograficznych jest 0.1 bpp.

W sprawozdaniu pokrótce omówione są poszczególne zadania badawcze, scharakteryzowane szerzej w załączonej rozprawie oraz publikacjach. Bardziej szczegółowo przedstawione są zagadnienia dotyczące testów subiektywnej oceny wiarygodności diagnostycznej obrazów, realizowane w ostatnim etapie prac, a także opis przykładowych systemów telemedycznych realizowanych pod kątem wykorzystania w tworzonej sieci wewnętrznej w Szpitalu Wolskim.

Wykaz prac przyjętych do druku lub opublikowanych

Artykuły w czasopismach:

Pub1. Przelaskowski A., "Elastyczność koderów falkowych w systemach archiwizacji i transmisji medycznych danych obrazowych", Prace Naukowe Politechniki Warszawskiej - Elektronika, z. 130, str. 105-122, Oficyna Wydawnicza PW, 2001.

Pub2. Przelaskowski A., "Techniki optymalizacji falkowej reprezentacji obrazów medycznych", Prace Naukowe Politechniki Warszawskiej - Elektronika, z. 130, str. 123-142, Oficyna Wydawnicza PW, 2001.

- Pub3. Przelaskowski A., "Lossless encoding of medical images: hybrid modification of statistical modelling-based conception" *Journal of Electronic Imaging*, vol. 10, issue 4, pp. 966-976, October 2001.
- Pub4. Przelaskowski A., "Lifting-based reversible transforms for lossy-to-lossless wavelet codecs", *Lecture Notes on Computer Science*, Springer Verlag, vol. 2124, pp. 61-70, 2001.
- Pub5. Przelaskowski A., "Hybrid lossless coder of medical images with statistical data modelling", *Lecture Notes on Computer Science*, Springer Verlag, vol. 2124, pp. 92-101, 2001.
- Pub6. Przelaskowski A., " Today's and tomorrow's Medical Imaging ", *Lecture Notes on Computer Science*, Springer Verlag, vol. 2124, pp. 236-237, 2001.

Artykuły opublikowane w materiałach konferencyjnych:

- Pub7. Przelaskowski A., "Performance evaluation of jpeg2000-like data decomposition schemes in wavelet codec", *Proceedings of IEEE International Conference on Image Processing*, pp. 788-791, 2001.
- Pub8. Przelaskowski A., "Effective integer-to-integer transforms for JPEG2000 coder ", *SPIE Conference: Wavelets: Applications in Signal and Image Processing IX*, *Proceedings of SPIE Vol. 4478*, pp. 299-310, 2001.
- Pub9. Przelaskowski A., "Falkowe metody kompresji danych obrazowych jako narzędzia kształtowania optymalnej reprezentacji strumienia danych", *Materiały II Seminarium Radiokomunikacja i Techniki Multimedialne*, str. 23-30, 2001.
- Pub10. Przelaskowski A., Kawczyński M., Maksymiuk K., Skaliński M.: Effective image compression tools and telemedicine systems. *Materiały 1-st Polish-Norwegian Seminar: Selected Research Issues at the Polish and Norwegian Universities, Working group "Medical Technology"*, str. 5-6, Politechnika Warszawska, Październik 2001.

Wykaz prac zgłoszonych do druku (w recenzji)

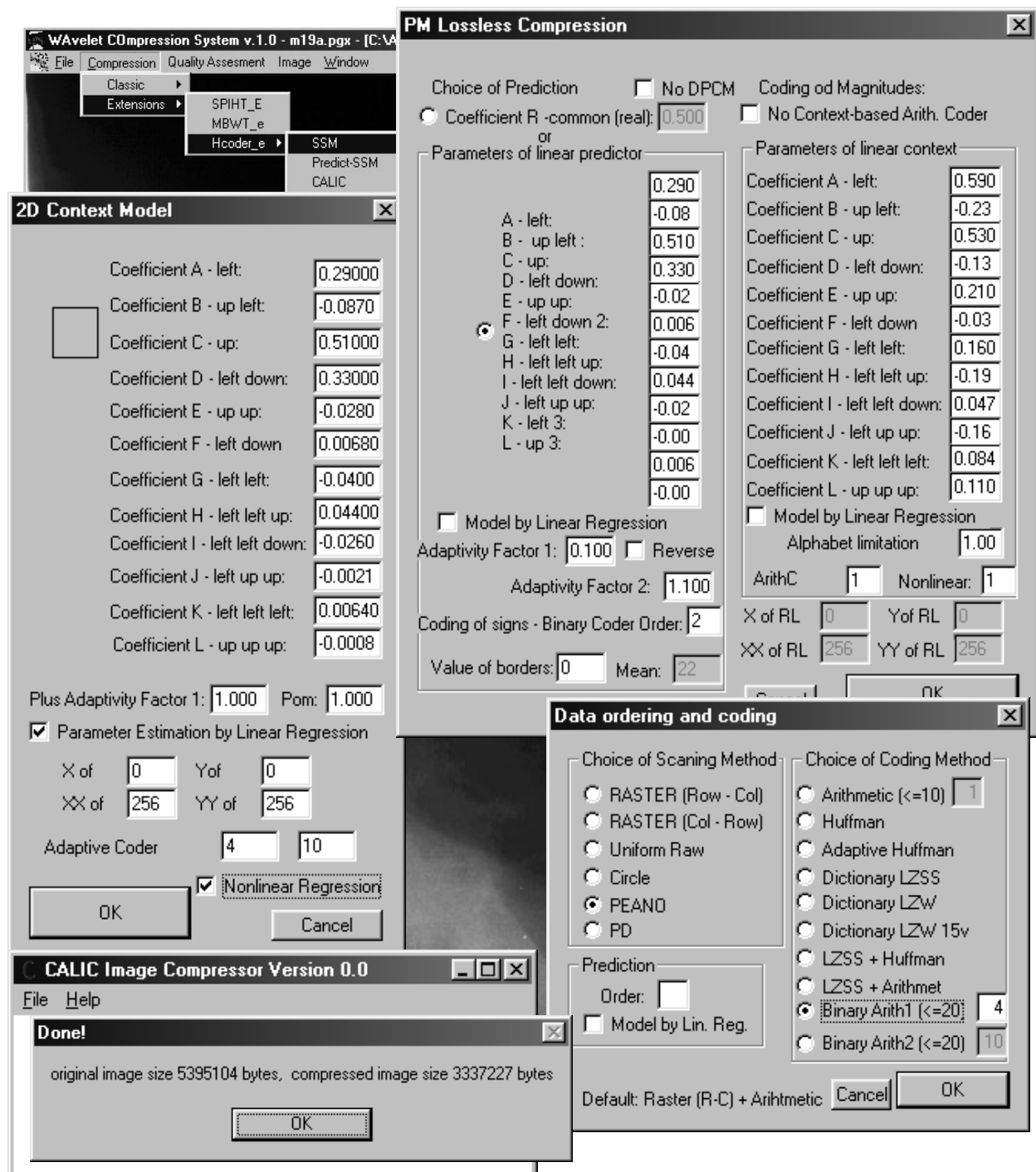
- Pub11. Przelaskowski A., "Falkowe metody kompresji danych obrazowych", rozprawa habilitacyjna złożona do druku w Oficynie Wydawniczej Politechniki Warszawskiej (złożona 4.01.2002 na ręce prof. Z. Pawłowskiego, przesłana, po zapoznaniu, do Redaktora Wydawnictw PW w dn. 4.02.2002 r. z propozycją recenzentów).
- Pub12. Przelaskowski A., "Vector measure with scalar equivalent for quality estimation of compressed medical images", zgłoszony do *Journal of Electronic Imaging*.

Wykaz prac przygotowywanych do zgłoszenia

- Pub13. Przelaskowski A., "Lossless encoding of medical images: wavelet-based conception with statistical modelling", do *Computerized Medical Imaging And Graphics*
- Pub14. Przelaskowski A., "Diagnostic pattern estimation for lossy compression evaluation", do *Computerized Medical Imaging And Graphics*.

1. REALIZACJA KOMPRESJI ODWRACALNEJ, OPTYMALIZOWANEJ DO ZASTOSOWAŃ MEDYCZNYCH

Opracowano hybrydowy koder odwracalny, *de facto* koncepcję całego systemu kompresji, który pozwala uzyskać dużą efektywność przy archiwizacji zróżnicowanej klasy obrazów medycznych. Koncepcja ta została przedstawiona w publikacjach. Poprzez włączenie kodera falkowego system ten umożliwi progresywną transmisję do zastosowań telemedycznych. Wygląd interfejsu użytkownika w zrealizowanym systemie kompresji oraz wybrane opcje optymalizacyjne przedstawiono na rys. 1.



Rys. 1. Hybrydowy system kompresji obrazów medycznych, wykorzystujący metody falkowe, uzupełnione technikami predykcji i statystycznego modelowania kontekstu w

koderach arytmetycznych. Widoczny interfejs użytkownika pozwala optymalizować w opcji Hcoder_e różne metody predykcyjne, w tym algorytm CALIC, a także technikę statystycznego modelowania kontekstu SSM.

2. KONSTRUKCJA KODERA FALKOWEGO DO ZASTOSOWAŃ TELETRANSMISJI

Ważną cechą koder falkowego jest możliwość ustalania kolejności transmisji informacji. Najbogatszy zbiór dostępnych opcji zawarto w standardzie JPEG2000. Testowano trzy aplikacje standardu (interfejsy do optymalizacji przedstawiono na rys. 2):

- model weryfikacyjny (VM v.8.6) (realizacja w C),
- Jasper (ogólnodostępna realizacja w C),
- JJ2000 (ogólnodostępna realizacja w Javie).

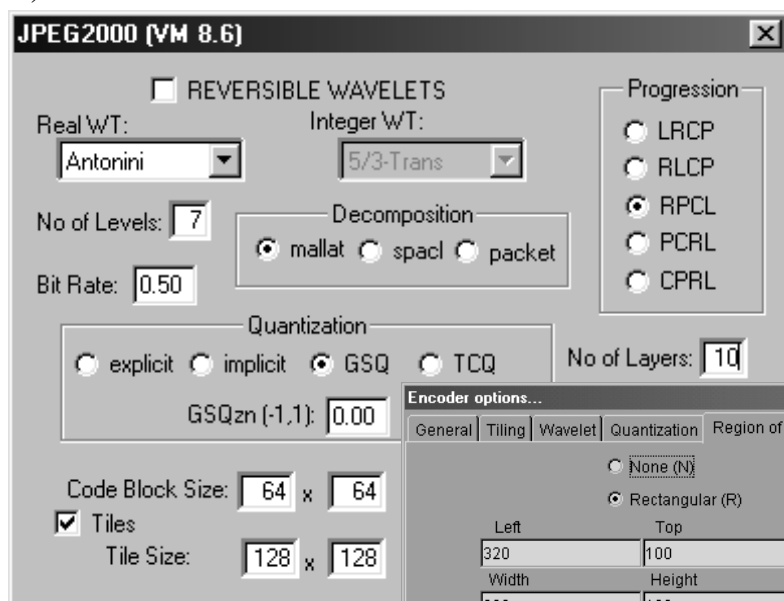
Dobierano postać transformacji falkowej, szczególnie całkowitoliczbowej, pozwalającej zwiększyć efektywność kompresji obrazów testowych (głównie mammograficznych). Testowano także możliwość podziału na części (ang. tile), badano wpływ różnych metod kwantyzacji (schematy explicit, implicit, TCQ, GSQ z doбором przedziału zerowego) i kodowania (wielkość bloków kodowych, zerowanie modelu statystycznego, wprowadzanie znaczników zwiększających odporność na zniekształcenia), parametrów optymalizacji R-D (stopnia zniekształceń źródeł informacji), rodzajów progresji (zorientowanej na jakość, rozdzielczość, komponent, obszar zainteresowań, fragment podpasm), liczby warstw, itd. Obserwowano także wpływ selekcji wybranych obszarów zainteresowań (ROI) na szybkość i efektywność transmisji. Rozważano także możliwości transmisji interaktywnej z aplikacją transkodera zmieniającego progresję jakości na progresję ROI.

Optymalizowano także koder falkowy SPIHT [27], dobierając postać falkowej transformacji i modyfikując schemat sukcesywnej aproksymacji poprzez dobór początkowej wartości proggu oraz progową selekcję, zgodnie z techniką TSUQ. Opracowano także wersję progresywną i osadzającą wersję algorytmu MBWT [21]. Dużym problemem okazała się optymalizacja w sensie R-D osadzonego strumienia tak, aby uzyskać efektywną jego postać nie tylko dla zadanej wartości średniej bitowej BR_t , ale także dla wartości pośrednich (od pierwszych bajtów do BR_t) [16]. Najbardziej użyteczne i elastyczne narzędzie optymalizacji zaproponowano w JPEG2000. Interfejsy stworzonego oprogramowania pokazano na rys. 3, a wybrane wyniki testów - w tabeli 1.

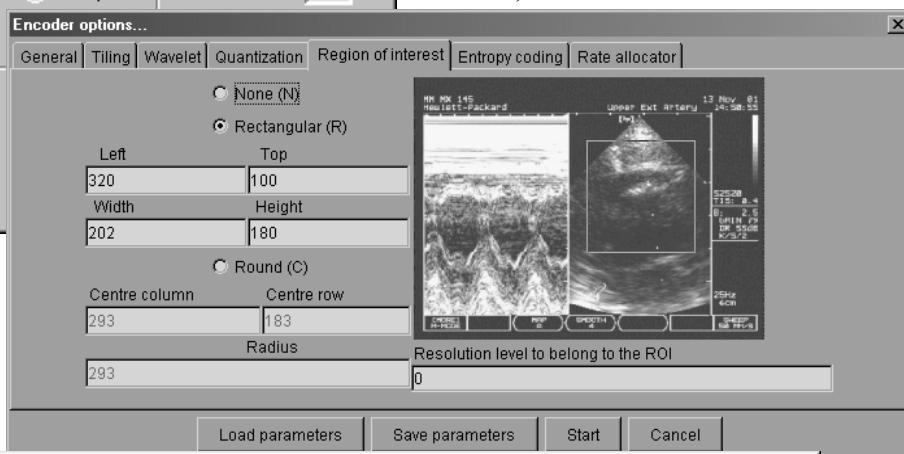
Tabela 1. Porównanie efektywności kodowania koderów falkowych przy pomocy miar obliczeniowych: SPIHT, zoptymalizowany SPIHT (M-SPIHT) [23], MBWT z progresją jakości (wersja z [23] plus modyfikacje transformacji falkowej z [13][19]) oraz JPEG2000. We wstępnej analizie wykorzystano naturalne obrazy testowe, a w tabeli umieszczono wartości PSNR.

Technika kompresji	Lena		Barbara		Goldhill	
	0.25	0.5	0.25	0.5	0.25	0.5
SPIHT	34.11	37.21	27.58	31.39	30.56	33.13
M-SPIHT	34.35	37.39	29.36	33.07	30.64	33.27
MBWT progress	34.49	37.53	29.58	33.38	30.77	33.42
JPEG2000 (VM8.6)	34.23	37.32	29.35	33.11	30.70	33.28

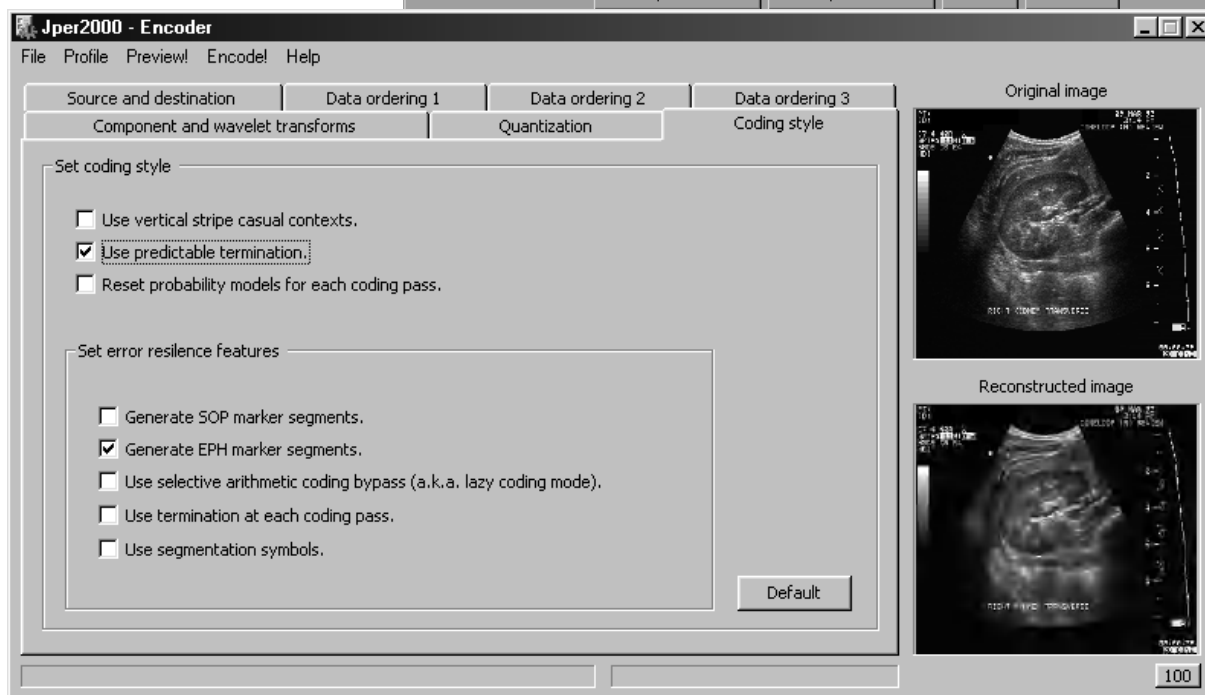
a)



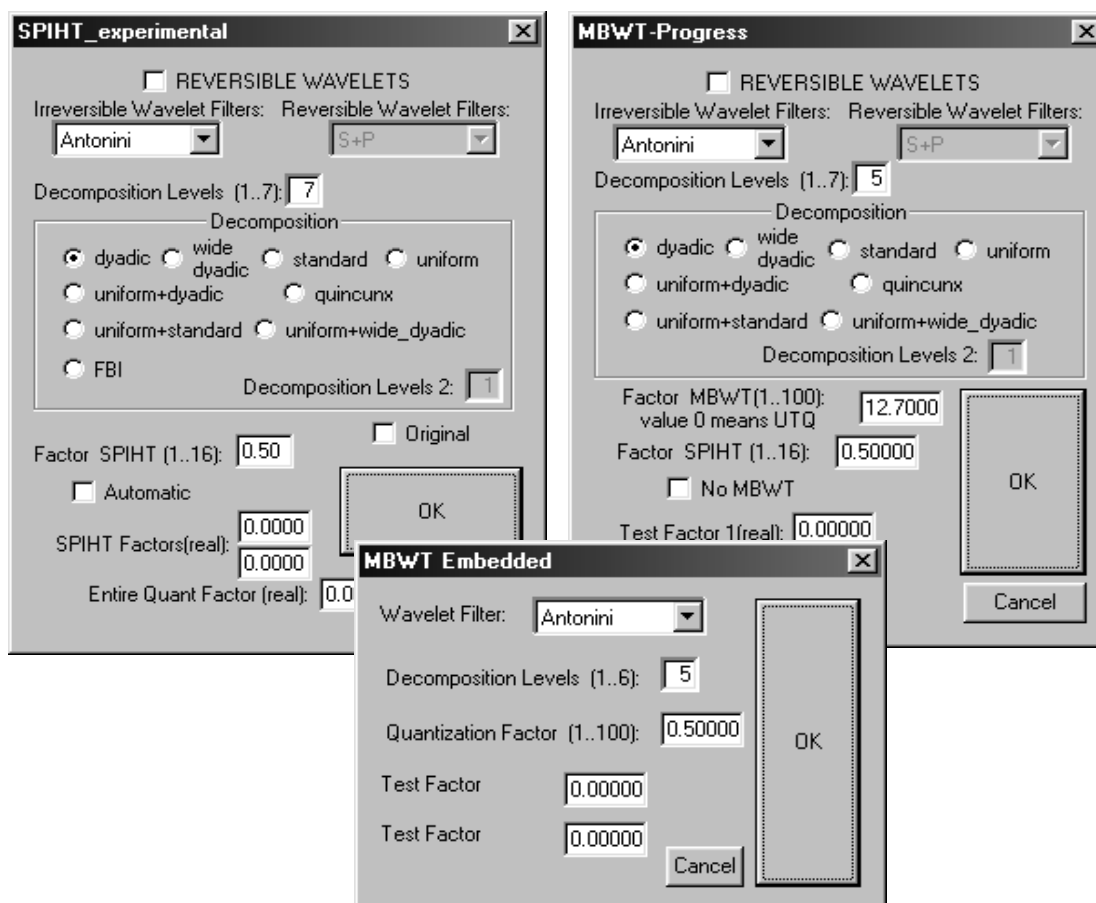
b)



c)



Rys. 2. Interfejsy programów optymalizacji koderów JPEG2000 do zastosowań w systemie teledyecznej: a) VM 8.6, b) JJ2000, c) Jasper.



Rys. 3. Optymalizacja kodera falkowego do zastosowań teletransmisji w oparciu o techniki SPIHT oraz MBWT.

Zgodnie z wynikami z tabeli 1, najlepsze rezultaty w sensie minimalizacji obliczeniowej miary podobieństwa z oryginałem pozwala uzyskać technika MBWT progress. Jednak jest to algorytm uzyskujący gorszy poziom optymalności w sensie R-D dla pośrednich stopni kompresji niż M-SPIHT oraz JPEG2000. Wyniki testów porównawczych technik M-SPIHT oraz JPEG2000, przeprowadzonych na klasie obrazów mammograficznych, przedstawiono w tabelach 2 i 3.

Tabela 2. Porównanie efektywności kodowania koderów falkowych przy pomocy miar obliczeniowych: M-SPIHT oraz JPEG2000. Wykorzystano mammograficzne obrazy testowe, a w tabeli umieszczono wartości PSNR.

Technika kompresji	Obraz A				Obraz C				Obraz G			
	1bpp	0.6bpp	0.1bpp	0.04bpp	1bpp	0.6bpp	0.1bpp	0.04bpp	1bpp	0.6bpp	0.1bpp	0.04bpp
M-SPIHT	52.84	49.74	44.74	42.78	49.67	46.74	42.99	41.46	51.34	48.19	43.56	41.47
JPEG2000	52.18	49.39	44.89	43.57	49.21	46.51	43.11	42.12	50.35	47.94	43.72	42.53

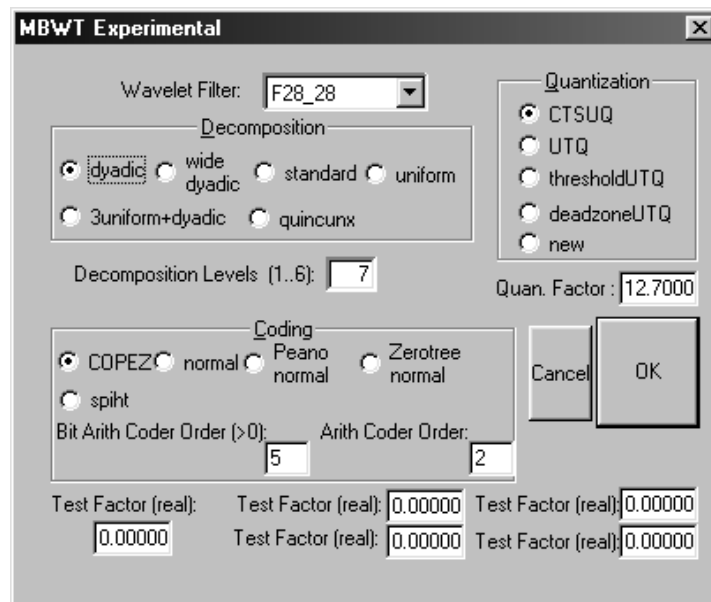
Tabela 3. Porównanie efektywności odwracalnej kompresji koderów falkowych: M-SPIHT oraz JPEG2000. Wykorzystano mammograficzne obrazy testowe, a w tabeli umieszczono wartości średnich bitowych.

	A	B	C	D	E	F	G	H	I	J	K	L
M-SPIHT	8.15	8.65	8.64	8.25	9.36	9.29	8.40	8.92	8.63	8.56	8.46	8.98
JPEG2000	8.30	8.83	8.83	8.45	9.56	9.53	8.58	9.17	8.86	8.73	8.67	9.19

Generalnie, potwierdzona została ogromna przydatność koderów JPEG2000 w teletransmisji. Zaprojektowane postacie transformacji całkowitoliczbowych pozwoliły znacząco zwiększyć efektywność tych koderów, a techniki kwantyzacji poprawiły wydajność koderów w stratnej kompresji obrazów.

3. OPTIMALIZACJA KODERA STRATNEGO DO ZASTOSOWAŃ ARCHIWIZACJI OBRAZÓW

Dokonano optymalizacji algorytmu MBWT (rys. 4), pozwalającego uzyskać bardzo dużą efektywność kompresji w sensie miar obliczeniowych, pod kątem kompresji obrazów medycznych, głównie mammograficznych. Pozwoliło to uzyskać wyraźnie lepszą skuteczność kompresji od kodera JPEG2000, szczególnie jeśli chodzi o jakość rekonstrukcji na poziomie pojedynczego piksela (zachowanie drobnych szczegółów istotnych diagnostycznie - patrz wyniki w tabeli 4). Wykorzystany algorytm kwantyzacji z adaptacyjną selekcją próbek pozwala kształtować postać obrazu rekonstruowanego na podstawie wiedzy dostępnej *a priori*. Także zaprojektowane postacie falkowej transformacji i dekompozycji pasmowej pozwalają lepiej modelować (aproksymować) źródło informacji obrazowej.



Rys. 4. Interfejs użytkownika w oprogramowaniu do optymalizacji techniki MBWT.

Tabela 4. Porównanie efektywności kodowania koderów falkowych przy pomocy miar obliczeniowych: MBWT oraz JPEG2000. Wykorzystano mammograficzne obrazy testowe, a w tabeli umieszczono wartości maksymalnej różnicy w punkcie.

Technika kompresji	Obraz A				Obraz E				Obraz F			
	1bpp	0.6bpp	0.1bpp	0.04bpp	1bpp	0.6bpp	0.1bpp	0.04bpp	1bpp	0.6bpp	0.1bpp	0.04bpp
JPEG2000	361	423	1306	2451	784	1292	2725	5172	688	1320	2695	5523
MBWT	254	359	934	1821	551	885	2017	3166	559	890	2169	3437

4. MODELOWANIE I OPTIMALIZACJA CAŁEGO SCHEMATU KOMPRESJI FALKOWEJ

Stworzenie kompleksowego modelu całego schematu kompresji-dekompresji jest zadaniem trudnym, złożonym i praktycznie nierealizowalnym bez pewnych upraszczających założeń. W [16] przedstawiono szereg rozwiązań wykorzystujących modele 'z pamięcią' źródeł informacji dających wymierne efekty poprawy skuteczności kompresji. Chodzi tu przede wszystkim o estymację całkowitoliczbowej transformacji falkowej, adaptacyjny algorytm kwantyzacji z progową selekcją próbek, a także wspólną optymalizację algorytmów kwantyzacji i binarnego kodowania, opartych na warunkowych modelach kontekstowych.

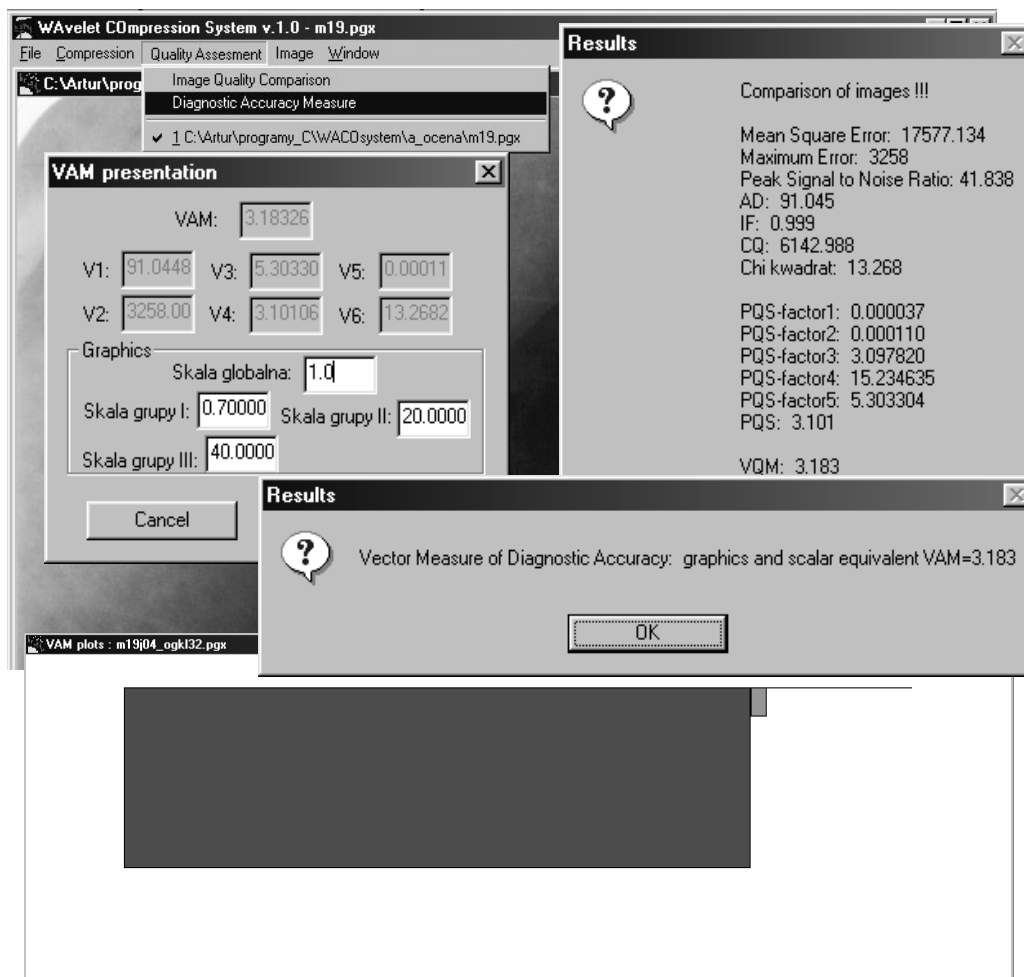
Przedstawiono tam także przykład modelowania całego schematu kompresji w oparciu o uproszczone modele 'bez pamięci' podkreślając ich małą użyteczność. Optymalizacja tak złożonego modelu jest trudna i bardzo złożona, sprowadza się najczęściej do rozwiązania wieloparametrycznego równania Eulera. Koniecznych jest szereg założeń upraszczających, nie sposób włączyć w ten model charakterystyki algorytmu kodowania, optymalizacji operatora transformacji, itp. Wymuszona jest wręcz koncentracja na jednym elemencie schematu, przy założeniu najprostszyc rozwiązań dla pozostałych.

Opracowano w [16] kompleksowy schemat kodera falkowego na poziomie modelu i algorytmu, z opcjonalnym doбором modu (orientacji) transmisji strumienia, globalną optymalizacją efektywności kompresji. Przeprowadzone testy porównawcze z najlepszymi znanymi koderami zweryfikowały słuszność zastosowanych rozwiązań. Dokonano optymalizacji elastycznej reprezentacji kodowego strumienia danych z wykorzystaniem teorii stopnia zniekształceń źródeł informacji (R-D), a także pod kątem wzrostu wiarygodności diagnostycznej kompresowanych stratnie obrazów.

5. BUDOWA WEKTOROWEJ MIARY WIARYGODNOŚCI DIAGNOSTYCZNEJ OBRAZÓW

Wektorowa miara wiarygodności pozwala ustalić w sposób obliczeniowy wiarygodność diagnostyczną kompresowanych stratnie obrazów. Wartość współczynnika korelacji wartości skalarne go ekwiwalentu wiarygodności z wzorcem diagnostycznym wyniosła ponad 0.9, co pozwala dość skutecznie ocenić przydatność koderów falkowych w stratnej kompresji obrazów medycznych.

W systemie wykorzystano szereg obliczeniowych miar jakości pozwalających ocenić kompresowane stratnie obrazy, w tym MSE, PSNR, maksymalny błąd w punkcie, PQS i inne (definicje w [16]) – zobacz rys. 5. Wprowadzono także obliczeniową miarę wiarygodności diagnostycznej (OMW), dającą liczbową reprezentację wiarygodności oraz graficzną prezentację poszczególnych zniekształceń (rys. 5).



Rys. 5. Ocena jakości kompresowanych stratnie obrazów oraz estymacja miary wiarygodności diagnostycznej obrazów mammograficznych. Skrót VAM (ang. Vector Accuracy Measure) oznacza obliczeniową, wektorową miarę wiarygodności.

6. WYZNACZENIE WZORCA DIAGNOSTYCZNEGO W TESTACH OCENY SUBIEKTYWNEJ

Wzorzec diagnostyczny, określający, jako punkt odniesienia, noty wiarygodności dla każdego z badanych obrazów został wykorzystany w optymalizacji OMW. Pozwala on 'naprowadzić' skalarny ekwiwalent wiarygodności miary wektorowej na wyniki możliwie silnie skorelowane z wiarygodnością diagnostyczną kompresowanych stratnie obrazów medycznych. Wyznaczono wzorzec diagnostyczny dla obrazowych badań mammograficznych, przy czym procedura testów przedstawiona w [16] może być zastosowana dla innych badań obrazowych, mniej lub bardziej szczegółowych kategorii diagnostycznych. Trzeba jedynie dostosować kryteria i sposób oceny do specyfiki konkretnych badań obrazowych.

Początkową postać wzorca diagnostycznego, pozwalającego wykonać wstępną selekcję cech wektora OMW oraz uzyskać wyższą od stosowanych dotąd miar korelację z oceną wiarygodności diagnostycznej obrazów, wyznaczono za pomocą testów, w których wzięło udział dwóch radiologów, a także kilkanaście osób z większym lub mniejszym doświadczeniem oceny jakościowej obrazów medycznych (inżynierów i studentów). Dokonano wtedy oceny obrazów testowych CT, MRI i USG [18][26]. Uzyskane wyniki

zweryfikowano następnie w liczniejszych testach przeprowadzonych na bazie ponad 200 badań mammograficznych.

Badania mammograficzne

Podstawową metodą badania sutków u kobiet po 40 roku życia jest mammografia, zarówno w badaniach skriningowych, jak i w diagnostyce stwierdzanych wcześniej patologii. Metoda ta pozwala na wykrycie zmian niemych klinicznie (subklinicznych) i wstępne określenie ich charakteru.

Granica wieku, od której szczególnie zalecane są okresowe badania mammograficzne zależy od rodzaju budowy sutka i wynika ze statystycznie znaczącego wzrostu zachorowań na raka sutka u kobiet w wieku okołomenopauzalnym.

Wartość diagnostyczna mammografii zależy głównie od aparatury (chodzi o odpowiednio wysoki standard określonych parametrów technicznych) oraz jakości pracy techników wykonujących badanie.

Podstawowymi projekcjami w mammografii są projekcje kranio-kaudalna i skośna. Jeżeli na ich podstawie nie można stwierdzić, ew. wykluczyć raka (zmiany), wykonuje się zdjęcia dodatkowe (w innych projekcjach, celowane ze zwiększonym uciskiem, powiększone).

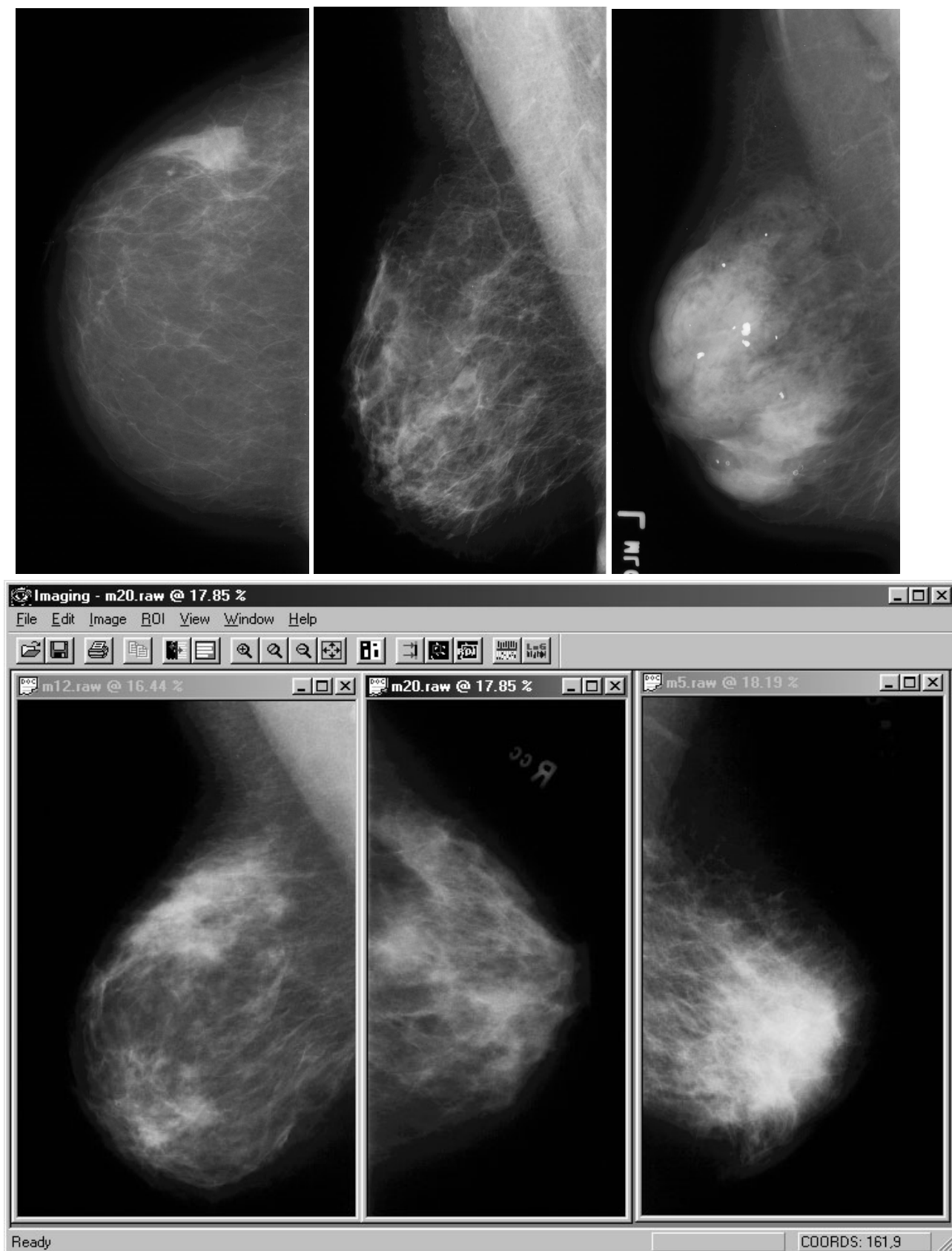
Niezwykle ważne przy ocenie mammografii (MMG) jest doświadczenie radiologa. Do tej pory nie wyznaczano jednoznacznie standardu dla określenia prawidłowego sutka. Odniesieniem w ocenie jest zwykle obraz drugiego sutka, a przy kolejnych badaniach - porównanie ze zdjęciami poprzednimi. W niektórych krajach standardowo wykorzystuje się opinie dwóch niezależnych radiologów, często z różnych ośrodków (otwiera się tutaj możliwość wykorzystania telekonsultacji). W przypadku każdej stwierdzonej w MMG zmiany bardzo ważne jest badanie kliniczne i wywiad, a w wielu wypadkach uzupełniające badania USG; ostateczną odpowiedź daje biopsja aspiracyjna cienkoigłowa i badanie histopatologiczne.

Najwięcej trudności w ocenie sprawiają tzw. „gęste” sutki, czyli sutki z dużą ilością tkanki gruczołowej i włóknistej. W takim przypadku stosuje się ultrasonograficzne badanie uzupełniające. Jednak nawet przy „gęstych” sutkach mammografia jest podstawowym badaniem w ocenie mikrozwapnień. Są one również widoczne w USG (warunkiem jest bardzo dobrej klasy aparat), ale ich ocena ilościowa i jakościowa możliwa jest jedynie w mammografii.

W symptomatologii mammograficznej raka sutka (najogólniej) wyróżnia się:

- objawy bezpośrednie:
 - guz (o różnej morfologii),
 - struktura promienista,
 - zaburzenia architektury linii brzegowej,
 - skupisko mikrozwapnień;
- objawy pośrednie:
 - pogrubienie skóry,
 - wciągnięcie brodawki sutkowej,
 - objawy naczyniowe.

Przykładowe badania mammograficzne wykorzystane w testach przedstawiono na rys. 6.



Rys. 6. Przykładowe testowe obrazy mammograficzne, w dolnym rzędzie prezentowane przy pomocy oprogramowania wykorzystywanego w testach diagnostycznych.

Przebieg testów oceny wiarygodności

Zaprojektowany i zrealizowany test subiektywnej oceny wiarygodności składał się z dwóch części: A (test detekcji) i B (test oceny). Został on przygotowany przez dwóch doświadczonych radiologów oraz inżyniera, tworzących zespół nadzorujący test. Wybrano trudne diagnostycznie, zróżnicowane obrazy z patologiami i bez, ustalono 'złoty standard', sposób i warunki oceny, przygotowano formularze. Jeden z radiologów zespołu nadzorującego brał udział w testach kontrolując ich przebieg i zapewniając jednakowe warunki oceny dla wszystkich uczestników.

W teście A wykorzystano 13 starannie wyselekcjonowanych obrazów z różnymi rodzajami trudno wykrywalnych patologii, w tym skupisk mikrozwapnień, oraz 'bezzmianowe'. Każdy z oryginalnych obrazów cyfrowych (o dynamice 14 bpp, skanowanych z klisz według zasad opisanych w [3]) został poddany kompresji metodą zgodną ze standardem JPEG2000 (11 obrazów) lub techniką MBWT (dwa obrazy) do dwóch wartości średniej bitowej: 0.1 bpp oraz 0.04 bpp, co daje liczbę 39 obrazów testowych. Zespół radiologów biorących udział w testach składał się z 7 osób: 3 z Zakładu Diagnostyki Obrazowej Szpitala Wolskiego w Warszawie, 2 z Pracowni Mammografii Centrum Onkologii w Warszawie i 2 z Zakładu Radiologii Szpitala Grochowskiego w Warszawie. Obrazy oglądano pojedynczo podejmując decyzję dotyczącą faktu wystąpienia patologii oraz jej lokalizacji. Ponadto, w uwagach można było dookreślić własności patologii. Nie było także ograniczeń, co do liczby zmian patologicznych w pojedynczym obrazie, można było więc odnotować nadrozpoznanie, czyli wskazanie zmian patologicznych w miejscach, gdzie według 'złotego standardu' one nie występują. Warunki prezentacji obrazów były w razie potrzeby modyfikowane (w ustalonych granicach) przez uczestnika testu. Pracowano zwykle w dwóch trybach prezentacji: pomniejszonej wersji całego obrazu oraz przy rozmiarach rzeczywistych, z możliwością obserwacji jedynie części obrazu. Obrazy były prezentowane w trzech sesjach, najpierw obrazy potencjalnie najgorszej jakości (0.04 bpp), potem rekonstruowane ze średniej 0.1 bpp, a na końcu obrazy oryginalne. Starano się rozdzielić te sesje jak najdłuższą przerwą, jednak, ze względu na trudności organizacyjne, w praktyce wynosiła ona około pół godziny. Ponadto, w przerwie była wykonywana część testu B na grupie 9 innych badań mammograficznych.

W teście B liczba ocenianych obrazów mammograficznych wyniosła 75. Złożyło się na nią na nią dziewięć obrazów oryginalnych oraz szereg obrazów rekonstruowanych po kompresji do następujących wartości średnich bitowych: 1.0 bpp, 0.6 bpp, 0.1 bpp oraz 0.04 bpp. Wykorzystano te same kodery jak w części A, przy czym sześć obrazów było kompresowanych obydwojema koderami, a pozostałe trzy tylko koderem JPEG2000. Jednocześnie prezentowano więc obraz oryginalny oraz cztery jego wersje po kompresji w różnym stopniu. W przypadku obrazów kompresowanych dwoma koderami, zestawiano wymieszane wersje obrazów (np. 1.0 bpp i 0.04 bpp po kompresji JPEG2000, a 0.6 bpp i 0.1 bpp po kompresji MBWT). Ten sam obraz oryginalny wyświetlany był w dwóch zestawach, można było więc odnotować różnice w ocenie tego samego obrazu, czyli błąd metody oceny subiektywnej. Test B odbywał się w dwóch częściach, na przemian z trzema częściami testu A. Uczestnicy oceniali cztery cechy zmian patologicznych w obrazie (wskazanych w razie ewentualnych wątpliwości): kontrast, ostrość, zarysy oraz kształt w skali od 1 (słabe, niewyraźne, ledwo dostrzegalne, zniekształcone) do 3 (wyraźne, dobrze rozróżnialne, regularne, nie budzące wątpliwości). Według większości radiologów taka skala ocen była wystarczająca.

Przykładowe formularze wykorzystane w testach pokazano na rys. 7, a interfejsy oprogramowania do prezentacji i analizy badań mammograficznych, optymalizowanego pod kątem użytkowników- radiologów oraz planowanych testów przedstawiono na rys. 8.

a)

Test DETEKCJI	Obraz	Patologia (tak/nie)	Lokalizacja	Uwagi
Część pierwsza	da9			
	da12			
	da15			
	da18			
	...			
	...			
	...			
	...			
	da27			
	da224			
	da30			
Uwagi ogólne				

b)

Test OCENY	Obraz	Kontrast 1-3 (uwagi)	Ostrość 1-3 (uwagi)	Zarysy 1-3 (uwagi)	Kształt 1-3 (uwagi)
Część druga	oam				
	oa2k				
	oaoz				
	...				
	...				
	...				
	...				
	...				
	oocvbvdf				
	oogfj7				
	op4mjd8v				
opfk8vk9					
Uwagi ogólne					

Rys. 7. Przykładowe formularze: a) z testu A, b) z testu B.



Rys. 8. Przykładowe interfejsy oprogramowania wykorzystanego w testach diagnostycznych. Widoczne są prezentacje obrazu o rozmiarach rzeczywistych (we fragmencie) oraz pomniejszonych tak, aby widoczny był cały obraz.

Komentarze radiologów

Oceniający w trakcie testów zgłaszali dodatkowe uwagi, które można podzielić na 3 grupy:

1. Uwagi dotyczące samego sposobu oceniania obrazów:

- ★ uzasadnione jest, żeby zdjęcia oglądała jedna osoba wielokrotnie i oceniała je wielokrotnie (zmęczenie oczu) - wiarygodność testu byłaby większa (wady wzroku, zmęczenie oczu);
- ★ za mała skala ocen;
- ★ źle, że nie ma [w teście detekcji] do porównania innej projekcji i obrazu drugiego sutka; niemiarodajne – tylko jedna projekcja; brak porównawczych zdjęć drugiego sutka; brak drugiej projekcji.

2. Komentarze odnoszące się do sposobu prezentowania obrazów w kontekście praktycznej diagnostyki mammograficznej:

- ★ zdjęcie całego sutka [zmniejszone] jest zbyt małe do oceny guzka;
- ★ brak powiększonego [tj. o rozmiarach rzeczywistych, nie zmniejszonego] obrazu całości sutka [test detekcji];
- ★ źle się ocenia fragmenty obrazu [test detekcji];
- ★ oglądanie fragmentami – stwarza złudne obrazy, niebezpieczeństwo pominięcia patologii typu „zaburzenia architektury”.

3. Uwagi dotyczące oceny diagnostycznej wartości wybranych obrazów:

- ★ jakość zdjęć dobra;
- ★ obraz mało czytelny, artefakty ???, [dotyczy pierwszego zestawu obrazów w teście detekcji];
- ★ bardziej czytelny, możliwość interpretacji zdjęć – dobra [dotyczy drugiego i trzeciego zestawu obrazów w teście detekcji];
- ★ w tej serii zdjęć przewaga na plus zarysów i kształtów, kontrast, ostrość dobra [test oceny];
- ★ słabo widzę mikrozwapnienia [j.w.];
- ★ trudno wykryć mikrozwapnienia w zmianach guzkowych;
- ★ mikrozwapnienia są nie do zróżnicowania na zdjęciach całego sutka [pomniejszonych] i powiększonych [o rozmiarach rzeczywistych];
- ★ tylko na jednym zdjęciu widoczne są mikrozwapnienia, na innych niewidoczne, można przeoczyć chorobę.

Wielokrotne obserwowanie obrazów, najlepiej z jedno-dwutygodniowymi przerwami, jak również inne jeszcze zabiegi przybliżające testy do realiów praktyki diagnostycznej znacznie zwiększają złożoność i czasochłonność testów, a także ich koszty. Środki i możliwości w ramach tego projektu były ograniczone, a poza tym główna teza prezentowana w [16] dotyczy propozycji rozwiązań alternatywnych, o znacznie mniejszej złożoności. Stąd procedura testów była optymalizowana głównie pod kątem skuteczności części B, pozwalającej szybko i możliwie jednoznacznie wyznaczyć wzorzec diagnostyczny dla wektorowej miary wiarygodności, przy niewielkich kosztach czasowo-organizacyjnych. Chodziło bowiem o rozwiązanie praktyczne, możliwe do wykorzystania w praktyce klinicznej. Stąd warunki przeprowadzenia klasycznego testu detekcji patologii były ograniczone, dalekie od optymalnych, znanych np. z projektu [1] o bardzo dużych nakładach ludzkich i finansowych.

Starano się jednak uczynić test detekcji możliwie wiarygodnym w tych ograniczonych warunkach, a wykonano go w celu zwiększenia spektrum analizy sposobów oceny oraz w celach porównawczych.

Uwaga o zbyt małej skali ocen była odosobniona, natomiast korzystanie z innych projekcji oraz porównawczych zdjęć drugiego sutka jest jak najbardziej zasadne z punktu widzenia prawidłowości procesu diagnozy. Zasadniczym powodem wykorzystania w testach tylko jednego zdjęcia danego sutka było stworzenie utrudnionych warunków bezwzględnej diagnozy, bez badań dodatkowych, bez porównań. Zaburza to prawidłową procedurę diagnostyczną, w pewnym sensie podważa zasadność takiej oceny zwiększając liczbę błędów, pozwala jednak zarejestrować sugestie radiologów odnośnie miejsc podejrzanych o zmiany patologiczne, wychwycić każdą nieprawidłowość tkanki, która musi zostać zakwalifikowana jako normalna lub potencjalnie chorobowa. Proces detekcji odbywa się więc na niższym poziomie podejrzeń zmian patologicznych, które byłyby zweryfikowane w innych badaniach (przy pomocy dodatkowych zdjęć, projekcji, itp.). W decyzjach radiologów patologia oznacza więc *de facto* jedynie podejrzenie patologii, co daje większą czułość procedury detekcji. Uzasadnia to także binarną skalę decyzji, z wykorzystaniem nadrozpoznań jako dodatku kowego elementu poprawności sugestii diagnostycznych. Metoda przetwarzania danych oryginalnych (kompresja) może wprowadzać dodatkowe zmiany w obrazie sugerujące patologie lub też eliminować takie sugestie, wynikające z obserwacji oryginału.

Zespół nadzorujący, dysponujący pełną bazą badań, nie tylko mammograficznych, w kilku przypadkach przyznał, że właściwie nie sposób na podstawie jednego zdjęcia stwierdzić patologii, można jednak zauważyć pewne hipotetyczne zmiany. Dość duża liczba nieprawidłowych kwalifikacji w decyzjach radiologów z testów detekcji wzięła się więc ze świadomego zwiększenia stopnia trudności testów w stosunku do praktyki klinicznej. Pozwala to z jednej strony przeprowadzić bardziej wiarygodne badania w ograniczonych warunkach, z drugiej - łatwiej jest zauważyć 'pozytywne' zmiany w obrazach rekonstruowanych po stratnej kompresji. Wydobycie pewnych cech poprawiających widoczność zmian zwiększa prawdopodobieństwo poprawnych decyzji lekarzy, co można było zaobserwować w pojedynczych przypadkach testu detekcji, jak również w teście oceny wiarygodności.

Problemy związane ze sposobem prezentacji obrazów wynikają z ograniczeń technicznych. Wykorzystany w badaniach 19" monitor (SONY G420 trynitron) o rozmiarach plamki 0.24 mm i maksymalnej rozdzielczości poziomej 1920 punktów oraz pionowej 1440 linii (przy rozdzielczości zalecanej 1280 punktów na 1024 linii) nie pozwalał na wyświetlenie większości obrazów testowych w całości (rozmiary rzędu 2500×1500). Stąd konieczność pomniejszania obrazów przy prezentacji całego obszaru zdjęcia. Wygodna selekcja obszaru zainteresowań z podglądem całości ułatwiała pracę radiologom. Jednak w opinii dwóch radiologów te ograniczenia utrudniały proces diagnozy. Zakup profesjonalnego monitora do radiologii przekraczał jednak możliwości finansowe tego projektu, wątpliwość budziła także jego przydatność w Szpitalu Wolskim po zakończeniu projektu wobec powszechnego wykorzystywania klasycznej praktyki oceny zdjęć analogowych.

Wyniki testów części B

Błąd metody oceny subiektywnej widać na przykładzie obrazów oryginalnych, ocenianych dwukrotnie w zestawach z obrazami rekonstruowanymi po kompresji dwoma koderami. Tabela 5 pokazuje zróżnicowanie ocen tych obrazów.

Tabela 5. Różnice w ocenie wiarygodności diagnostycznej oryginalnych obrazów testowych.

	Obrazy					
	A	B	C	E	F	I
Ocena 1	10,43	10,14	9,00	9,43	9,43	8,14
Ocena 2	9,00	11,29	10,71	10,57	9,29	10,00
Średnia	9,72	10,72	9,86	10,00	9,36	9,07
Maks. różn. [%]	14,8	10,8	17,3	11,4	1,5	20,5

Zróznicowanie ocen może sięgać nawet 20% estymowanej średniej wartości wiarygodności, co pokazuje niejednoznaczność ocen i duży błąd wyznaczenia wzorca diagnostycznego. Aby uzyskać bardziej wiarygodny wzorzec zrezygnowano z obrazu I, jak również uwzględniono w analizie niejednorodny charakter obrazów testowych. Ocena cech zmian patologicznych jest bardziej jednoznaczna w obrazach F, B i E, natomiast nieco mniej w obrazach A i C.

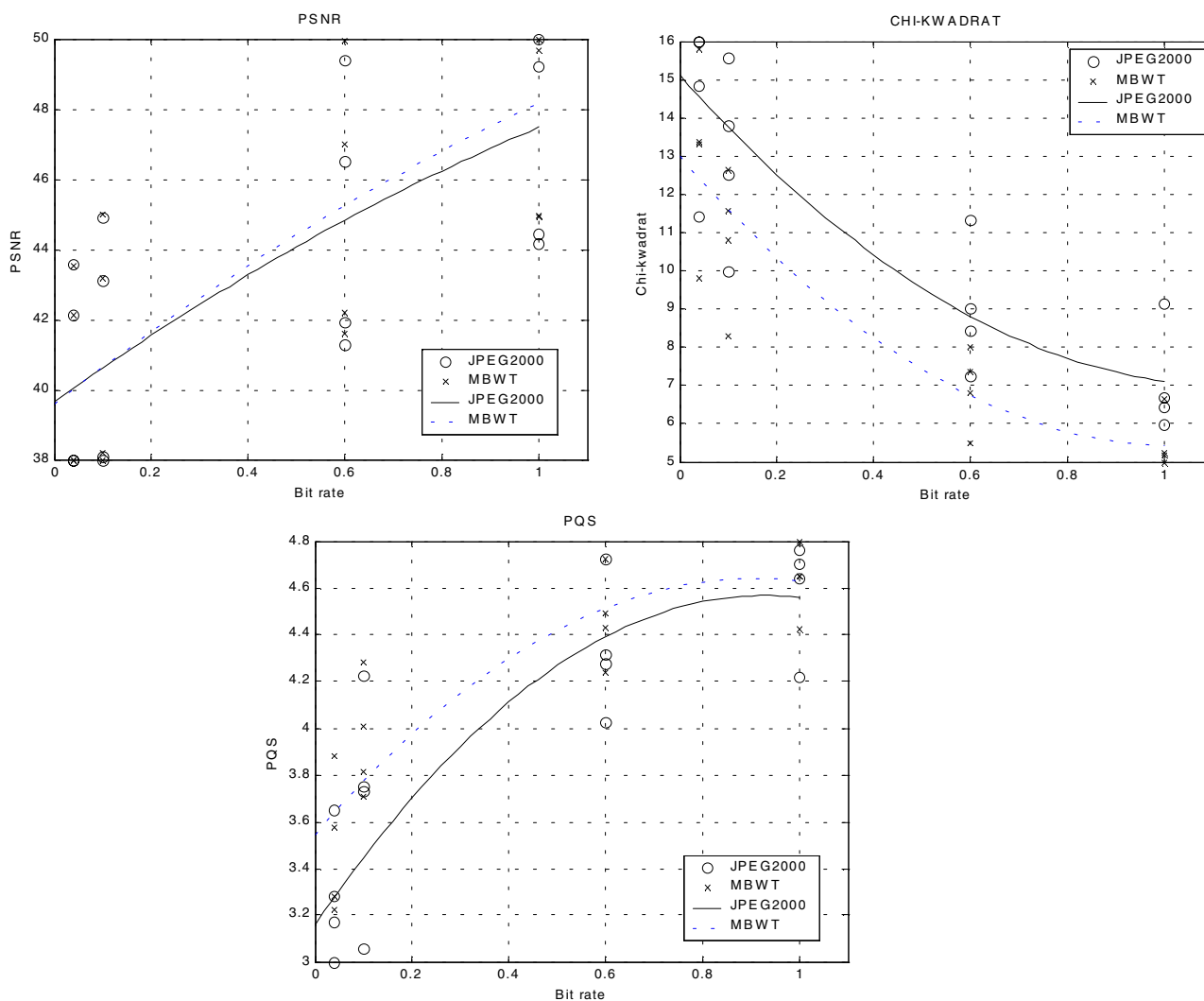
Najwyższe oceny wiarygodności dla poszczególnych obrazów tylko w czterech przypadkach uzyskały oryginały, natomiast w pięciu były to obrazy rekonstruowane (1 bpp i 0.6 bpp). W tabeli 6 pokazano kolejność wersji poszczególnych obrazów wynikającą z przydzielonych im ocen.

Tabela 6. Kolejność ocen wiarygodności diagnostycznej obrazów. Oznaczenia: O – oryginał, 10 – 1 bpp, 6 – 0.6 bpp, 1 – 0.1 bpp, 4 – 0.04 bpp.

Koder	Obrazy								
	A	B	C	D	E	F	G	H	I
JPEG2000	O,10,1,6,4	O,10,6,1,4	10,O,6,1,4	O,10,1,6,4	10,O,6,1,4	10,O,6,1,4	O,6,10,1,4	6,10,O,1,4	10,6,1,O,4
MBWT	O,10,1,6,4	10,O,6,1,4	10,O,6,1,4	-	6,10,O,1,4	6,10,O,1,4	-	-	-

Obraz rekonstruowany ze średniej 0.04 bpp otrzymał najniższą ocenę w każdym przypadku, przy czym była ona zwykle prawie dwukrotnie niższa od ocen pozostałych wersji. Świadczy to o wyraźnie gorszej jakości tych obrazów, a więc niedopuszczalności tak wysokiego stopnia kompresji ze względu na wyraźne pogorszenie jakości rekonstrukcji cech obrazu istotnych diagnostycznie. Pozostałe cztery wersje występują zamiennie, przy czym oryginał, wersja 1 bpp oraz 0.6 bpp wydają się być tej samej, mieszczącej się w granicach błędu metody, wartości diagnostycznej. Wersja 1 bpp występuje częściej na pierwszej pozycji niż oryginał, co może świadczyć nawet o pewnej poprawie wartości diagnostycznej względem oryginału, przynajmniej w niektórych przypadkach. Dokładnie taka sama kolejność w przypadku obrazów E i F oraz koder MBWT, według której wersja 0.6 bpp wyprzedza wersję 1 bpp oraz oryginał, może sugerować poprawę wartości diagnostycznej obrazów w wyniku stratnej kompresji MBWT do poziomu 0.6 bpp. Wersja 0.1 bpp występuje na drugiej bądź trzeciej pozycji od końca, w niektórych przypadkach zachowuje wartość diagnostyczną oryginału, trudno jednak na podstawie uzyskanych rezultatów definitywnie stwierdzić, czy kompresja do średniej 0.1 bpp jest dopuszczalna i nie powoduje utraty wiarygodności diagnostycznej.

Jednoznaczne stwierdzenie, która z metod kompresji daje lepsze rezultaty nie jest możliwe. Uzyskane różnice ocen subiektywnych mieszczą się w granicach błędu metody oceny, trudno też zauważyć jakieś zdecydowane tendencje. Można jednak sformułować kilka przesłanek charakteryzujących właściwości rekonstrukcji obu koderów. Porównanie efektywności tych koderów za pomocą obiektywnych miar obliczeniowych przedstawiono na rys. 9, natomiast przy użyciu ocen subiektywnych – na rys. 10. Bardziej szczegółowe zestawienie wyników ocen kompresowanych obrazów przedstawiono w tabeli 7.

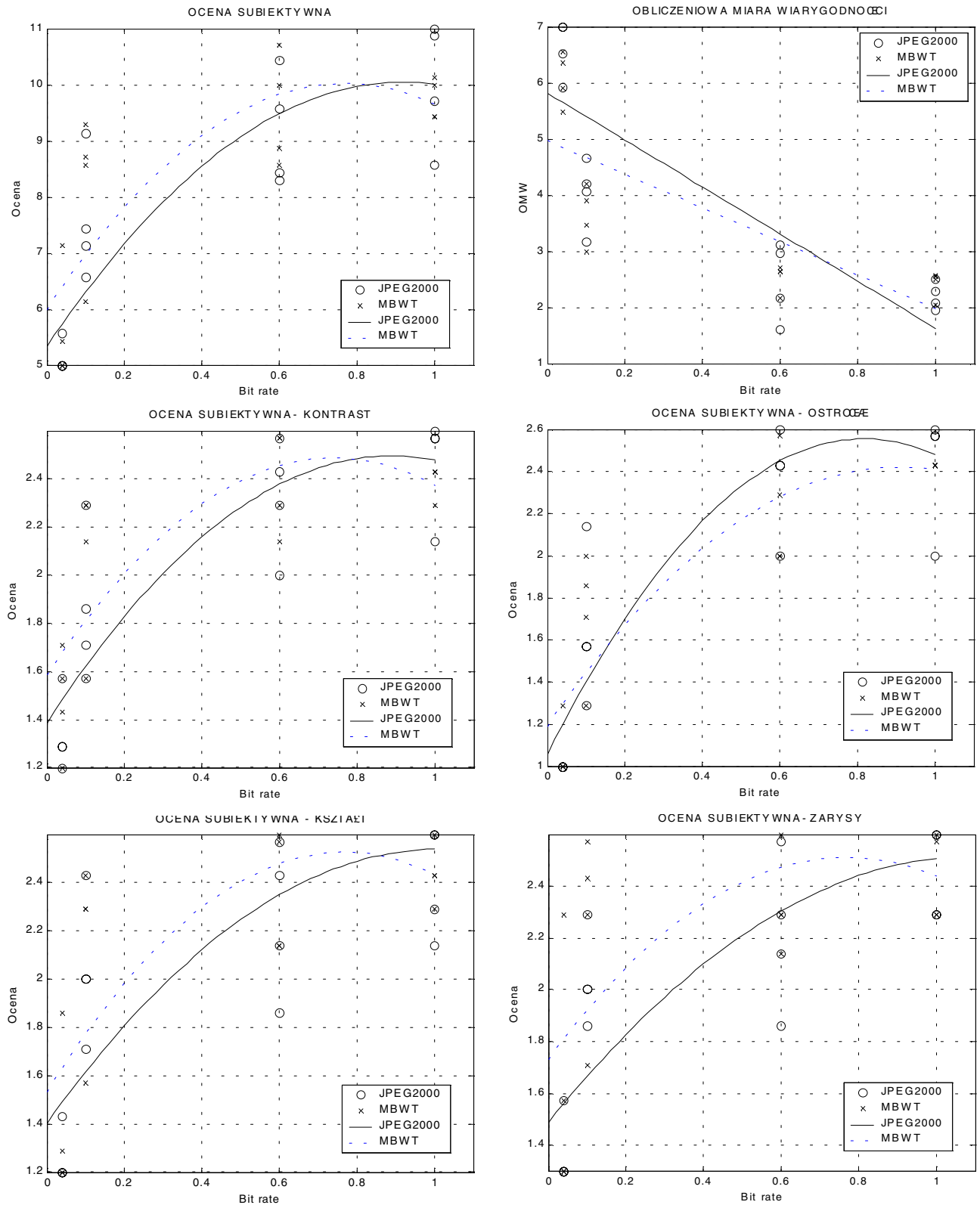


Rys. 9. Porównanie efektywności koderów falkowych: JPEG2000 i MBWT za pomocą obliczeniowych miar jakości: PSNR, chi-kwadrat i PQS.

Według obliczeniowych miar jakości, zarówno lokalnych jak i globalnych, skalarnych jak i wektorowych, lepszym algorytmem kompresji jest MBWT w całym zakresie testowanych średnich bitowych. Jedynie w pojedynczych wypadkach ocena jakości jest porównywalna. Według oceny subiektywnej metoda MBWT jest dominująca w zakresie mniejszych średnich bitowych, tj. 0.04 – 0.6 bpp. Dla średniej bitowej 1 bpp zarówno globalna ocena jakości, jak i jej elementy składowe (kontrast, ostrość zarysy, kształt) wykazują nieco wyższą jakość obrazów kompresowanych metodą JPEG2000, choć jest to właściwie efekt niemal równoważnych ocen dla średnich 0.6 bpp i 1 bpp.

Takie rezultaty można tłumaczyć właściwościami obu algorytmów. Przy większych średnich bitowych zmodyfikowany algorytm kwantyzacji z MBWT odgrywa marginalną rolę, a wobec zoptymalizowanego w sensie R-D (stopnia zniekształceń) strumienia JPEG2000 oraz efektywniejszej metody arytmetycznego kodowania skuteczność koder MBWT może być nieznacznie mniejsza, choć przeczą temu wyniki oceny obiektywnej obliczeniowo. Przy mniejszych średnich bitowych zaproponowany mechanizm kwantyzacji zaczyna odgrywać znaczącą rolę, co wyraźnie znajduje odbicie w lepszych ocenach. Tą poprawę wiarygodności widać szczególnie na wykresach ocen zarysów zmian patologicznych. Silniejsza kwantyzacja tekstur w MBWT może prowadzić do nasilenia efektu rozmycia

zobrazowanej substancji tkankowej. Znalazło to potwierdzenie w gorszych ocenach MBWT w kategorii ostrości zmian.



Rys. 10. Porównanie efektywności koderów falkowych: JPEG2000 i MBWT za pomocą subiektywnych ocen diagnostycznej wiarygodności oraz obliczeniowej miary wiarygodności OMW.

Tabela 7. Ocena obrazów rekonstruowanych za pomocą obliczeniowych miar jakości oraz wiarygodności diagnostycznej. Oznaczenia: A,B,C, ... to kolejne obrazy, j – obraz rekonstruowany z użyciem JPEG2000, m - obraz rekonstruowany z użyciem MBWT, 10,6,1,04 to bitowe średnie rekonstrukcji, odpowiednio 1 bpp, 0.6 bpp, 0.1 bpp, 0.04 bpp.

Lp	Obraz	MSE	MD	PSNR	AD	IF	CQ	Chi-kwad.	PQS(1) [*10 ⁻⁴]	PQS(2) [*10 ⁻³]	PQS(3)	PQS(4)	PQS(5)	PQS	OMW
1.	Aj10	1626.6	361	52.18	31.14	1	4869.4	9.12	.03	.027	.8471	8.804	3.164	4.216	2.5000
2.	Am10	1279.5	254	53.22	27.37	1	4868.3	6.67	.01	.020	.3985	7.660	2.721	4.423	2.5041
3.	Aj6	3086.1	423	49.39	40.97	1	4868.2	11.33	.03	.041	.9279	9.571	3.848	4.024	2.1674
4.	Am6	2725.5	359	49.93	37.58	1	4867.6	8.02	.03	.028	.5269	8.562	3.202	4.240	2.6332
5.	Aj1	8704.9	1306	44.89	62.76	0.999	4865.1	15.56	.15	.141	1.783	13.598	7.161	3.053	4.0647
6.	Am1	8499.5	934	44.99	59.78	0.999	4864.3	11.57	.18	.083	1.355	10.864	4.936	3.710	3.4565
7.	Aj04	11811.5	2451	43.57	71.90	0.999	4863.0	17.87	.47	.332	3.723	16.284	10.999	2.200	6.5220
8.	Am04	11906.0	1821	43.53	69.33	0.999	4861.9	13.38	.57	.186	2.838	12.707	6.842	3.224	5.4876
9.	Cj10	3217.7	464	49.21	43.47	1	4900.6	5.98	.02	.009	.2082	6.297	1.546	4.759	1.9627
10.	Cm10	2896.7	343	49.67	40.92	1	4900.0	4.40	.01	.007	.1425	5.765	1.386	4.850	2.0440
11.	Cj6	5989.4	556	46.51	57.15	1	4899.7	7.22	.02	.010	.2209	6.488	1.618	4.724	1.5998
12.	Cm6	5353.0	522	47.00	53.84	1	4899.0	5.51	.02	.010	.1982	6.470	1.636	4.724	2.1641
13.	Cj1	13117.7	1568	43.11	81.25	0.999	4897.1	9.98	.08	.030	.4296	9.087	2.790	4.225	3.1672
14.	Cm1	12886.7	994	43.19	79.50	0.999	4896.5	8.31	.12	.031	.5406	8.725	2.737	4.281	2.9949
15.	Cj04	16475.4	2960	42.12	90.44	0.999	4895.9	11.43	.26	.079	1.261	11.657	4.591	3.651	5.9249
16.	Cm04	16478.9	2922	42.12	89.19	0.999	4895.4	9.81	.33	.071	1.177	10.563	3.940	3.883	5.9038
17.	Daj10	2267.3	434	50.73	35.80	1	5475.6	7.95	.03	.022	.5623	8.200	2.625	4.366	2.5506
18.	Daj6	4048.7	475	48.22	46.05	1	5474.0	10.52	.03	.037	.6453	9.037	3.379	4.154	2.1121
19.	Daj1	10949.1	1293	43.89	69.36	0.999	5469.8	15.13	.14	.108	1.479	12.841	5.826	3.330	3.6743
20.	Daj04	14342.4	2774	42.72	78.14	0.998	5467.1	17.35	.43	.279	3.403	15.643	9.294	2.512	6.5510
21.	Ej10	10270.7	784	44.17	75.37	1	6781.1	6.69	.02	.008	.3199	7.245	1.498	4.642	2.2887
22.	Em10	8544.2	551	44.97	68.45	1	6780.6	5.19	.02	.009	.3053	7.131	1.528	4.652	2.5561
23.	Ej6	19975.2	1292	41.28	101.62	0.999	6779.2	9.02	.05	.020	.5200	9.165	2.342	4.276	2.9731
24.	Em6	18626.0	885	41.59	98.20	0.999	6778.3	7.35	.05	.016	.5071	8.351	1.988	4.431	2.6439
25.	Ej1	47877.4	2725	37.49	151.63	0.998	6772.2	13.79	.19	.051	1.452	11.857	3.673	3.750	4.2010
26.	Em1	48071.5	2017	37.47	151.31	0.998	6771.2	12.65	.30	.057	1.570	11.618	3.468	3.811	3.9119
27.	Ej04	60206.4	5172	36.49	169.49	0.998	6768.9	16.74	.60	.116	3.483	14.782	5.266	3.170	8.1224
28.	Em04	61776.0	3166	36.38	170.92	0.998	6768.0	15.79	.79	.125	3.486	14.238	4.960	3.283	6.3556
29.	Fj10	9660.9	688	44.44	73.77	1	8252.3	6.42	.01	.006	.5083	6.918	1.381	4.702	2.0753
30.	Fm10	8615.4	559	44.94	69.31	1	8251.6	5.24	.02	.008	.4543	7.179	1.507	4.650	2.5692
31.	Fj6	17302.3	1320	41.91	95.91	1	8250.7	8.43	.03	.017	.8839	8.991	2.267	4.313	3.1221
32.	Fm6	16120.0	890	42.21	92.67	1	8250.2	6.82	.03	.012	.6416	8.110	1.774	4.493	2.7102
33.	Fj1	41611.0	2695	38.10	144.04	0.999	8245.9	12.52	.13	.046	1.978	12.005	3.712	3.731	4.6654
34.	Fm1	40706.5	2169	38.19	141.67	0.999	8245.7	10.79	.17	.030	1.387	10.979	2.611	4.009	4.1934
35.	Fj04	52533.3	5523	37.08	160.75	0.999	8243.5	14.81	.38	.080	3.430	14.401	4.808	3.281	8.6018
36.	Fm04	54437.6	3437	36.93	162.39	0.999	8242.7	13.30	.50	.059	2.585	13.395	3.553	3.575	6.5526
37.	Gj10	2475.7	633	50.35	37.72	1	6993.1	6.94	.02	.016	.4973	8.156	2.317	4.413	3.0406
38.	Gj6	4309.9	515	47.94	48.08	1	6992.2	8.61	.02	.020	.5727	8.713	2.638	4.296	2.5771
39.	Gj1	11410.2	1413	43.72	72.10	0.999	6989.4	11.53	.08	.059	1.166	12.218	4.529	3.584	4.5202
40.	Gj04	14996.9	3026	42.53	81.26	0.999	6987.9	13.54	.26	.153	2.582	14.835	7.268	2.883	7.3083
41.	Hj10	5382.7	563	46.98	55.13	1	5317.9	6.02	.02	.007	.2086	6.302	1.219	4.803	1.9322
42.	Hj6	9043.1	602	44.73	70.57	0.999	5316.7	7.40	.03	.008	.2677	6.692	1.299	4.741	1.5920
43.	Hj1	22114.5	2500	40.84	104.90	0.999	5312.1	11.67	.21	.057	1.021	11.119	3.454	3.873	4.8068
44.	Hj04	29262.2	3184	39.63	118.96	0.998	5309.8	13.31	.48	.080	1.634	13.070	4.066	3.539	6.5323

Wyraźnie wyższe oceny subiektywne dla obrazów 0.1 bpp (rekonstruowanych ze średniej 0.1 bpp) mogą prowadzić do stwierdzenia, iż obrazy te zachowują swą wiarygodność diagnostyczną szczególnie w przypadku kodera MBWT.

Wyznaczenie wzorca na podstawie wyników testów części B

OMW została zweryfikowana poprzez subiektywną ocenę wiarygodności diagnostycznej obrazów mammograficznych według procedury testu B. Wzięło w nich udział 9 lekarzy z trzech ośrodków medycznych, przy czym w samych testach oceniano grupę 9 obrazów z patologiami. Używano przy tym dwóch koderów falkowych: MBWT oraz JPEG2000 w wersjach zoptymalizowanych. Warto nadmienić, iż na etapie przygotowania przejrano ponad 200 różnego typu obrazowych badań mammograficznych, wykonano szereg eksperymentów optymalizacji koderów (w tym także kodera SPIHT). Przytoczone niżej rezultaty pozwalają potwierdzić użyteczność OMW w medycznych systemach informacyjnych, zwiększającej bezpieczeństwo stosowania stratnych wersji rekonstrukcji, a także przydatność tej miary w optymalizacji koderów falkowych. Średnie wartości wszystkich ocen uzyskanych dla poszczególnych obrazów stanowią wzorzec diagnostyczny dla OMW – zamieszczono go w tabeli 8. Stopień korelacji różnych miar obliczeniowych z wzorcem diagnostycznym pokazano w tabeli 9.

Tabela 8. Wzorzec diagnostyczny wyznaczony w testach subiektywnej oceny wiarygodności diagnostycznej obrazów (według testu B). Oznaczenia: A,B,C, ... to kolejne obrazy, j – obraz rekonstruowany z użyciem JPEG2000, m - obraz rekonstruowany z użyciem MBWT, 10,6,1,04 to bitowe średnie rekonstrukcji, odpowiednio 1 bpp, 0.6 bpp, 0.1 bpp, 0.04 bpp.

Lp.	Obraz	Średnia ocena					Lp.	Obraz	Średnia ocena				
		kontrast	ostrość	kształt	zarysy	suma			kontrast	ostrość	kształt	zarysy	suma
1.	A	2,22	2,36	2,43	2,71	9,72	31.	Daj1	2,29	2,29	2,43	2,43	9,43
2.	Aj10	2,14	2,00	2,14	2,29	8,57	32.	Daj04	1,57	1,14	1,43	1,57	5,71
3.	Am10	2,43	2,43	2,71	2,57	10,14	33.	E	2,36	2,50	2,64	2,50	10,00
4.	Aj6	2,00	2,00	2,14	2,14	8,29	34.	Ej10	2,71	2,57	2,86	2,71	10,86
5.	Am6	2,29	2,00	2,14	2,14	8,57	35.	Em10	2,43	2,43	2,43	2,71	10,00
6.	Aj1	2,29	2,14	2,43	2,29	9,14	36.	Ej6	2,43	2,43	2,43	2,29	9,57
7.	Am1	2,29	2,00	2,43	2,57	9,29	37.	Em6	2,57	2,57	2,71	2,86	10,71
8.	Aj04	1,29	1,00	1,00	1,14	4,43	38.	Ej1	1,71	1,29	1,71	1,86	6,57
9.	Am04	1,71	1,29	1,86	2,29	7,14	39.	Em1	2,14	1,86	2,29	2,29	8,57
10.	B	2,65	2,79	2,64	2,64	10,72	40.	Ej04	1,29	1,00	1,14	1,29	4,71
11.	Bj10	2,57	2,43	2,71	2,71	10,43	41.	Em04	1,57	1,00	1,29	1,57	5,43
12.	Bm10	2,57	2,86	2,43	2,57	10,43	42.	F	2,43	2,29	2,36	2,29	9,36
13.	Bj6	2,29	2,29	2,43	2,57	9,57	43.	Fj10	2,57	2,57	2,29	2,29	9,71
14.	Bm6	1,71	1,57	2,00	2,00	7,29	44.	Fm10	2,29	2,43	2,43	2,29	9,43
15.	Bj1	1,86	1,29	1,71	2,00	6,86	45.	Fj6	2,29	2,43	1,86	1,86	8,43
16.	Bm1	1,86	1,57	1,57	1,43	6,43	46.	Fm6	2,57	2,00	2,57	2,29	10,00
17.	Bj04	1,14	1,00	1,00	1,00	4,14	47.	Fj1	1,57	1,57	2,00	2,00	7,14
18.	Bm04	1,29	1,00	1,00	1,00	4,29	48.	Fm1	1,57	1,29	1,57	1,71	6,14
19.	C	2,22	2,43	2,58	2,65	9,86	49.	Fj04	1,00	1,00	1,00	1,00	4,00
20.	Cj10	2,57	2,86	3,00	2,86	11,29	50.	Fm04	1,14	1,00	1,14	1,29	4,57
21.	Cm10	2,43	2,43	2,29	2,29	9,43	51.	G	2,43	2,43	2,29	2,43	9,57
22.	Cj6	2,57	2,71	2,57	2,57	10,43	52.	Gj10	2,14	2,00	2,14	2,14	8,43
23.	Cm6	2,14	2,29	2,14	2,29	8,86	53.	Gj6	2,14	2,29	2,29	2,14	8,86
24.	Cj1	1,86	1,57	2,00	2,00	7,43	54.	Gj1	2,00	2,00	2,00	2,00	8,00
25.	Cm1	2,29	1,71	2,29	2,43	8,71	55.	Gj04	1,14	1,29	1,14	1,14	4,71
26.	Cj04	1,57	1,00	1,43	1,57	5,57	56.	H	2,29	2,43	2,71	2,71	10,14
27.	Cm04	1,43	1,00	1,00	1,00	4,43	57.	Hj10	2,43	2,57	2,43	2,71	10,14
28.	Da	2,43	2,57	2,57	2,71	10,29	58.	Hj6	2,71	2,43	2,71	2,86	10,71
29.	Daj10	2,71	2,43	2,29	2,43	9,86	59.	Hj1	2,29	2,00	2,57	2,57	9,43
30.	Daj6	2,14	2,29	2,43	2,43	9,29	60.	Hj04	1,43	1,29	1,43	1,57	5,71

Uzyskany współczynnik korelacji wartości ekwiwalentu wiarygodności OMW z wzorcem diagnostycznym potwierdza przydatność tej miary do oceny wiarygodności kompresowanych

stratnie obrazów mammograficznych. Wobec wyników testów oceny wstępnej wydaje się, iż OMW może znaleźć zastosowanie w przypadku analogicznych testów oceny także innych medycznych badań obrazowych.

Tabela 9. Korelacja pomiędzy wzorcem diagnostycznym a obliczeniowymi miarami skalarnymi (przedstawione są wartości współczynników korelacji). $PQS(i)$ oznacza i -ty współczynnik miary PQS, a WD –wzorec diagnostyczny. Definicje poszczególnych miar z tabeli 9 podano w [16].

Miary skalarne	Korelacja z WD	Miary skalarne	Korelacja z WD
$PQS(1)$	0.7815	MD	0.8543
$PQS(2)$	0.6115	MSE	0.6162
$PQS(3)$	0.8112	AD	0.5903
$PQS(4)$	0.8060	CQ	0.1644
$PQS(5)$	0.6374	IF	0.6079
PQS	0.7537	PQS (optymalizowany)	0.8459
Chi-kwadrat	0.7266	$AD+MD+Chi-kwadrat$	0.8625
$PSNR$	0.5825	OMW	0.9028

Na podstawie przytoczonych rezultatów widać wyraźnie, że poziom korelacji pomiędzy poszczególnymi miarami skalarnymi a wzorcem diagnostycznym jest silnie zróżnicowany, a w kilku przypadkach zadawalający. Stosowane najczęściej do oceny skuteczności różnych technik kompresji MSE i PSNR dają niezbyt satysfakcjonującą wartość współczynnika korelacji na poziomie 0.6. Podobnie zachowują się dwie inne miary: AD i IF, a CQ koreluje z wartościami oceny subiektywnej jeszcze gorzej. Spośród miar skalarnych wyraźnie najwyższe współczynniki korelacji uzyskano dla MD, co podkreśla duże znaczenie miar lokalnych. Z miar globalnych najbardziej użyteczna okazała się miara chi-kwadrat (χ^2). Wyższe wartości współczynników korelacji niż miara PQS mają trzy jej współczynniki: pierwszy, trzeci i czwarty. Przeprowadzono optymalizację miary PQS, zmieniając nieco wagi operatora liniowego tak, aby uzyskać lepszą korelację z wzorcem diagnostycznym. Dodatkowo, skonstruowano miarę wektorową z AD, MD i chi-kwadrat, co dało większą wartość współczynnika korelacji niż zoptymalizowana PQS. Wreszcie zaproponowana w OMW kombinacja współczynników dała najwyższy poziom korelacji z wzorcem diagnostycznym, powyżej 0.9. Wartość współczynnika korelacji skalarnej miary OMW z poszczególnymi elementami oceny wiarygodności (kontrast, ostrość, itd.) jest bliska 0.8.

7. PROJEKTOWANIE I REALIZACJA TESTÓW DETEKЦИИ PATOLOGII

W teście detekcji wykorzystano 13 mammogramów, wśród których 4 były bez zmian patologicznych. Patologie widoczne na pozostałych to zmiany nowotworowe (potwierdzone histopatologicznie) w postaci guzków (dobrze odgraniczonych i spikularnych), guzkowych zagęszczeń o promienistym charakterze, zaburzeń architektury z mikrozwapnieniami oraz skupiska mikrozwapnień bez zmiany guzkowej. Obrazy prezentowane były pojedynczo, tylko w jednej projekcji (przy możliwości 'zoomigu'). Wybrano dwie wartości średnich bitowych kompresowanych wersji obrazów 0.1 bpp oraz 0.04 bpp, aby stwierdzić możliwość przesunięcia, za pomocą koderów falkowych JPEG2000 oraz MBWT, granicy dopuszczalnej długości reprezentacji obrazów mammograficznych poniżej wartości 0.15 bpp sugerowanej w pracy [1].

Wyniki testów części A

Rezultaty testów zamieszczono w tabeli 10. Zawarto tam oceny stwierdzające (sugerujące) obecność zmian patologicznych we wskazanych miejscach analizowanych obrazów lub też ich brak, zweryfikowane przez zespół nadzorujący w kategoriach poprawnego lub niepoprawnego wskazania rodzaju i lokalizacji zmiany oraz nadrozpoznanie, czyli wskazania zmian patologicznych w obrazach bez patologii lub w miejscach różniących się ze 'złotym standardem' w obrazach z patologiami.

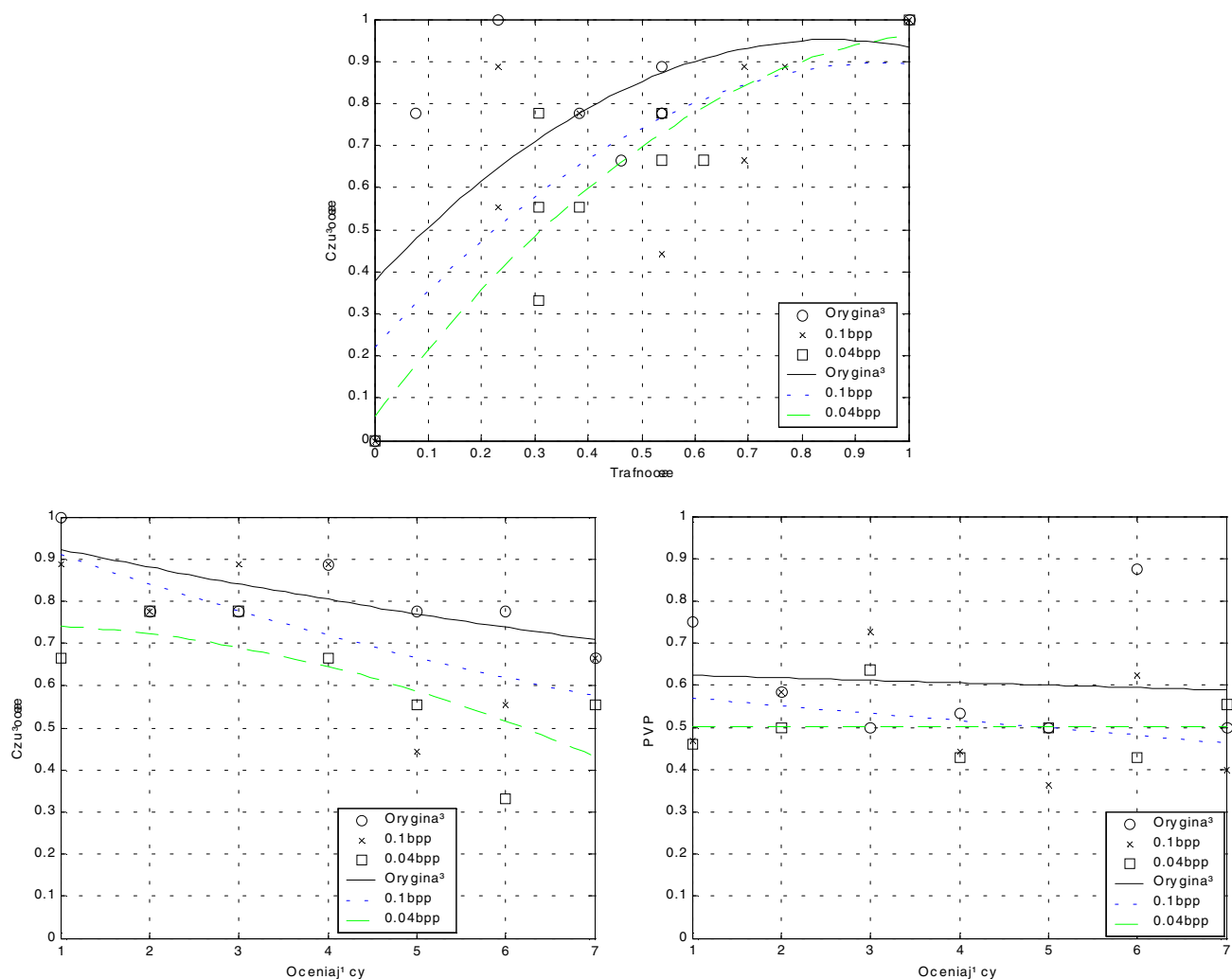
Tabela 10. Dane z testu detekcji patologii. Oznaczenia: **tak** to stwierdzenie obecności patologii, **nie** – brak patologii; **1** w kolumnach weryfikacji oznacza poprawną decyzję diagnostyczną, **0** – niezgodność ze 'złotym standardem'; **1** w kolumnach nadrozpoznanie oznacza zweryfikowany fakt nadrozpoznanie, a **0** to brak – niezgodność ze 'złotym standardem'.

Obraz	Ocena obecności patologii							Weryfikacja ocen							Nadrozpoznanie w ocenach						
	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
1	tak	tak	tak	tak	tak	tak	tak	1	1	1	1	1	1	1	0	1	1	1	0	0	0
1j1	tak	tak	tak	tak	tak	tak	tak	1	0	1	1	1	0	1	1	0	0	1	0	0	1
1j04	tak	tak	tak	tak	tak	nie	tak	1	1	0	1	1	0	1	0	1	0	0	0	0	0
2a	tak	tak	tak	tak	tak	tak	tak	1	1	1	1	1	1	1	0	0	0	1	0	0	0
2aj1	tak	tak	tak	tak	tak	tak	tak	1	1	1	1	1	1	1	1	0	0	1	0	0	1
2aj04	tak	tak	tak	tak	tak	tak	tak	1	1	1	1	1	0	1	1	0	0	1	0	1	0
3	tak	tak	tak	tak	tak	tak	nie	0	0	0	0	0	0	1	1	1	1	1	0	0	0
3j1	tak	tak	tak	tak	tak	nie	tak	0	0	0	0	0	1	0	1	1	1	1	0	1	1
3j04	tak	tak	tak	tak	tak	tak	tak	0	0	0	0	0	0	0	1	1	1	0	1	1	1
4	nie	tak	tak	nie	tak	tak	tak	1	0	0	1	0	0	0	0	1	1	0	1	1	1
4j1	tak	tak	nie	tak	tak	tak	tak	0	0	1	0	0	0	0	1	1	1	1	1	1	1
4j04	nie	tak	nie	tak	tak	nie	tak	1	0	1	0	0	1	0	0	1	0	0	1	0	1
5	tak	tak	tak	tak	tak	tak	tak	1	1	1	1	1	1	1	0	0	1	0	1	0	1
5j1	tak	tak	tak	tak	tak	tak	tak	1	1	1	1	1	1	1	0	0	0	0	0	0	1
5j04	tak	tak	tak	tak	tak	tak	tak	1	1	1	1	1	1	1	0	0	0	0	0	0	0
6	tak	tak	nie	nie	tak	tak	tak	1	1	1	1	0	1	1	0	0	0	0	1	0	0
6j1	tak	tak	nie	tak	tak	tak	tak	1	1	1	1	0	1	1	0	0	0	0	1	0	1
6j04	tak	tak	tak	tak	tak	nie	nie	0	0	0	0	1	1	1	1	1	0	0	0	0	0
7	nie	tak	tak	tak	tak	nie	tak	1	0	0	0	0	1	0	0	1	1	1	0	1	1
7j1	tak	tak	tak	tak	tak	nie	tak	0	0	0	0	0	1	0	1	1	1	1	0	1	1
7j04	tak	tak	tak	tak	tak	nie	tak	0	0	0	0	0	1	0	1	1	1	1	0	1	1
8	tak	tak	tak	tak	tak	tak	tak	1	0	1	1	1	0	0	0	1	0	0	0	0	1
8j1	tak	tak	tak	tak	nie	nie	nie	1	1	1	1	0	0	0	1	0	0	0	0	0	0
8j04	tak	tak	tak	tak	tak	nie	nie	1	1	1	1	1	0	0	0	0	0	0	0	0	0
9	tak	tak	tak	nie	tak	tak	tak	0	0	0	1	0	0	0	1	0	1	0	1	0	1
9m1	tak	tak	nie	tak	tak	tak	tak	0	0	1	0	0	0	0	1	1	0	1	1	1	0
9m04	tak	tak	tak	tak	tak	nie	tak	0	0	0	0	0	1	0	1	1	1	1	0	1	1
10	tak	tak	tak	tak	tak	tak	tak	1	1	1	1	1	1	1	0	0	0	1	0	0	0
10j1	tak	tak	tak	tak	nie	tak	tak	1	1	1	1	0	1	1	0	0	0	1	0	0	1
10j04	tak	tak	tak	tak	nie	tak	nie	0	1	1	1	0	0	0	1	0	0	1	0	1	0
11	tak	tak	tak	tak	tak	tak	tak	1	1	1	1	1	1	1	0	0	0	1	0	0	0
11j1	tak	tak	tak	tak	tak	tak	tak	1	1	1	1	1	1	1	1	0	0	1	0	0	0
11j04	tak	tak	tak	tak	tak	tak	tak	1	1	1	1	1	1	1	0	0	0	1	0	0	0
12	tak	tak	tak	tak	tak	tak	tak	1	1	0	0	0	1	0	1	0	1	1	0	1	1
12m1	tak	tak	tak	tak	tak	tak	tak	1	1	1	1	0	0	0	1	0	0	1	1	0	1
12m04	tak	tak	tak	tak	tak	tak	tak	0	1	1	0	0	0	1	1	0	1	1	1	1	0
13	tak	tak	tak	tak	tak	tak	tak	1	1	1	1	1	1	1	0	0	0	0	0	0	0
13j1	tak	tak	nie	tak	tak	tak	tak	1	1	0	1	0	1	1	0	1	0	1	1	1	0
13j04	Nie	tak	nie	tak	nie	nie	nie	0	0	0	0	0	0	0	0	1	0	1	0	0	0

Analiza wyników

A. Wyznaczenie krzywych ROC poprzez aproksymacje danych funkcjami wielomianowymi oraz jednowymiarowe testy statystyczne

W tabeli 11 przedstawiono dane, na podstawie których wyznaczono krzywą ROC z rys. 11. Obliczono liczbę decyzji prawdziwie pozytywnych (NPP) oraz fałszywie pozytywnych (NFP) wykorzystując, obok decyzji dotyczących patologii w obrazach z patologiami i bez, także nadrozpoznanie. Rozszerzono w ten sposób spektrum decyzyjne. Pozwoliło to wierniej oddać specyfikę pracy radiologa w testach.



Rys. 11. Krzywa ROC oraz wykresy czułości i PVP w funkcji ocen zespołu radiologów. Punkty danych aproksymowano wielomianami stopnia 2.

Krzywa ROC z rys. 11 pokazuje większą wartość diagnostyczną oryginalnych wersji obrazów. W stosunku do ocen obrazów z dwóch pozostałych klas (0.1 bpp i 0.04 bpp) zanotowano więcej poprawnych detekcji patologii przy zbliżonym poziomie pomyłek (nietrafności decyzji). Widać to wyraźnie na wykresie wartości czułości ocen poszczególnych radiologów – wyraźnie najgorszą czułość miały oceny obrazów 0.04 bpp. Poziom wartości PVP jest już mniej zróżnicowany – brak jest wyraźnych tendencji, pozwalających jednoznacznie stwierdzić pogorszenie wiarygodności diagnostycznej obrazów

rekonstruowanych. Wnioski z wykresów zostały potwierdzone przez wyniki testów statystycznych zamieszczone w tabeli 12.

Tabela 11. Dane przygotowane w celu wyznaczenia krzywych ROC i PVP; **1** oznacza obecność patologii, **0** – nie ma patologii, przy czym NFP = NNA(pat) + NNA(bezpat) – podkreślają to widoczne strzałki.

Obraz		Oceny obecności patologii								Nadrozpoznanie w ocenach							
		1	2	3	4	5	6	7	NPPO	1	2	3	4	5	6	7	NNA
Z patologią	1	1	1	1	1	1	1	1	7	0	1	1	1	0	0	0	3
	2	1	1	1	1	1	1	1	7	0	0	0	1	0	0	0	1
	4	1	0	0	1	0	0	0	2	0	1	1	0	1	1	1	5
	5	1	1	1	1	1	1	1	7	0	0	1	0	1	0	1	3
	8	1	0	1	1	1	0	0	4	0	1	0	0	0	0	1	2
	10	1	1	1	1	1	1	1	7	0	0	0	1	0	0	0	1
	11	1	1	1	1	1	1	1	7	0	0	0	1	0	0	0	1
	12	1	1	0	0	0	1	0	3	1	0	1	1	1	0	1	5
	13	1	1	1	1	1	1	1	7	0	0	0	0	0	0	0	0
NPP		9	7	7	8	7	7	6	NFPO	1	3	4	5	3	1	4	
Bez patologii	3	1	1	1	1	1	1	0	6	1	1	1	1	1	0	0	5
	6	0	0	0	0	1	0	0	1	0	0	0	0	1	0	0	1
	7	0	1	1	1	1	0	1	5	0	1	1	1	1	0	1	5
	9	1	1	1	0	1	1	1	6	1	0	1	0	1	0	1	4
NFP		3	5	7	7	7	1	6	NPPO	2	2	3	2	4	0	2	
Z patologią	1j1	1	0	1	1	1	0	1	5	1	0	0	1	0	0	1	3
	2j1	1	1	1	1	1	1	1	7	1	0	0	1	0	0	1	3
	4j1	0	0	1	0	0	0	0	1	1	1	1	1	1	1	1	7
	5j1	1	1	1	1	1	1	1	7	0	0	0	0	0	0	1	1
	8j1	1	1	1	1	0	0	0	4	1	0	0	0	0	0	0	1
	10j1	1	1	1	1	0	1	1	6	0	0	0	1	0	0	1	2
	11j1	1	1	1	1	1	1	1	7	1	0	0	1	0	0	0	2
	12m1	1	1	1	1	0	0	0	4	1	0	0	1	1	0	1	4
	13j1	1	1	0	1	0	1	1	5	0	1	0	1	1	1	0	4
NPP		8	7	8	8	4	5	6	NFPO	6	2	1	7	3	2	6	
Bez patologii	3j1	1	1	1	1	1	0	1	6	1	1	1	1	1	0	1	6
	6j1	0	0	0	0	1	0	0	1	0	0	0	0	1	0	1	2
	7j1	1	1	1	1	1	0	1	6	1	1	1	1	1	0	1	6
	9m1	1	1	0	1	1	1	1	6	1	1	0	1	1	1	0	5
NFP		9	5	3	10	7	3	9	NPPO	3	3	2	3	4	1	3	
Z patologią	1j04	1	1	0	1	1	0	1	5	0	1	0	1	0	0	0	2
	2j04	1	1	1	1	1	0	1	6	1	0	0	1	0	1	0	3
	4j04	1	0	1	0	0	1	0	3	0	1	0	0	1	0	1	3
	5j04	1	1	1	1	1	1	1	7	0	0	0	0	0	0	0	0
	8j04	1	1	1	1	1	0	0	5	0	0	0	0	0	0	0	0
	10j04	0	1	1	1	0	0	0	3	1	0	0	1	0	1	0	3
	11j04	1	1	1	1	1	1	1	7	0	0	0	1	0	0	0	1
	12m04	0	1	1	0	0	0	1	3	1	0	1	1	1	1	0	5
	13j04	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	2
NPP		6	7	7	6	5	3	5	NFPO	3	3	1	6	2	3	1	
Bez patologii	3j04	1	1	1	1	1	1	1	7	1	1	1	0	1	1	1	6
	6j04	1	1	1	1	0	0	0	4	1	1	0	0	0	0	0	2
	7j04	1	1	1	1	1	0	1	6	1	1	1	1	1	0	1	6
	9m04	1	1	1	1	1	0	1	6	1	1	1	1	1	0	1	6
NFP		7	7	4	8	5	4	4	NPPO	4	4	3	2	3	1	3	

Tabela 12. Analiza statystyczna krzywych z rys. 11. Pogrubiono wyniki sugerujące odrzucenie odpowiednich hipotez zerowych (przy poziomie istotności $p = 0.05$).

Wartość	Czułość			Nietrafność			PVP		
	Oryginal	0.1bpp	0.04bpp	Oryginal	0.1bpp	0.04bpp	Oryginal	0.1bpp	0.04bpp
Średnia	0.8095	0.7302	0.6190	0.3956	0.5055	0.4286	0.6060	0.5163	0.5015
Wariancja	0.0112	0.0323	0.0241	0.0324	0.0510	0.0175	0.0221	0.0175	0.0056
Statystyka U	→	1.0067	2.6833	→	-1.0067	-0.3906	→	1.1919	1.6615
Statystyka t		0.9321	2.4842		-0.9321	-0.3616		1.1035	1.5382

Przy poziomie istotności $p = 0.05$ należy odrzucić hipotezę zerową w obu testach wartości czułości przy porównaniu danych oryginalnych i kompresowanych ze średnią 0.04 bpp. Stąd generalna konkluzja skłania do odrzucenia obrazów kompresowanych w tym stopniu jako dopuszczalnych w diagnostyce. Potwierdzają to także komentarze zarejestrowane w czasie przeprowadzania testu.

Warto zauważyć pojedyncze przypadki, kiedy wskutek kompresji poprawiła się zdolność detekcji patologii z obrazów rekonstruowanych w stosunku do ich wersji oryginalnych. Przykładem jest obraz 12, dla którego reprezentacji 0.1 bpp zanotowano więcej poprawnych decyzji diagnostycznych i mniej nadrozpoznań, czy też obraz 5, gdzie systematycznie maleje liczba nadrozpoznań przy zmniejszaniu średniej bitowej reprezentacji.

B. Wyznaczenie binormalnych krzywych ROC oraz bardziej złożone testy jedno- i dwuwymiarowe

Do oceny wyników testu detekcji wykorzystano także narzędzia stworzone na Wydziale Radiologii Uniwersytetu w Chicago przez zespół prof. Ch. E. Metz [8][9][10], w tym oprogramowanie ROCKIT v. 0.9B [2]. Zastosowano:

- o jednowymiarowy test TFP różnic pomiędzy czułością dwóch krzywych ROC (dla ustalonych wartości nietrafności) ze statystyką z według [6]. Hipoteza zerowa jest następująca: zbiory danych pochodzące z binormalnych krzywych ROC mają te same wartości czułości przy ustalonej nietrafności;
- o jednowymiarowy test różnic pomiędzy powierzchnią pod dwoma krzywymi ROC ze statystyką z i hipotezą zerową taką, że zbiory danych pochodzą z binormalnych krzywych ROC o tej samej powierzchni poniżej [9];
- o dwuwymiarowy test chi-kwadrat równoczesnych różnic pomiędzy czułością oraz nietrafnością dwóch krzywych ROC. Hipoteza zerowa jest wtedy następująca: zbiory danych pochodzą z tej samej binormalnej krzywej ROC [9][8].

Estymowano przy tym binormalną krzywą ROC [9], pozwalającą wykreślić zależności pomiędzy czułością i nietrafnością decyzji radiologów opisanych zbiorem odpowiednio przygotowanych danych wejściowych. Ponieważ w testach tych wymagane są decyzje w szerszej niż binarna skali ocen, na podstawie decyzji radiologów z tabeli 10 przygotowano kilka wariantów takich skal obserwując ich wpływ na wynik testów statystycznych. Najbardziej naturalną wydaje się 4 stopniowa skala ocen indywidualnych decyzji lekarskich, powstała poprzez uwzględnienie zweryfikowanych ocen obecności patologii oraz nadrozpoznań dotyczących dwóch grup obrazów: z patologią i bez. Skalę tą pokazano w tabeli 13.

Tabela 13. 4 stopniowa skala ocen indywidualnych decyzji radiologów.

Skala	Obraz z patologią		Obraz bez patologii	
	Weryfikacja	Nadrozpoznanie	Weryfikacja	Nadrozpoznanie
4	1	1	1	0
3	1	0	1	1
2	0	1	0	1
1	0	0	0	0

Przygotowano także trzy inne skale, tzw. sumaryczne, przypisujące ocenę nie pojedynczym decyzjom, ale każdemu obrazowi testowemu na podstawie decyzji siedmiu radiologów (weryfikacja poprawności, nadrozpoznanie) dotyczących tego obrazu. Uzyskane dane wejściowe do testów statystycznych przedstawiono w tabeli 14. Wyniki wspomnianych trzech testów pokazano w tabeli 15, a estymowane krzywe ROC na rys. 12.

Rezultaty testów z tabeli 15 silnie zależą od przyjętej skali ocen. Według skali ocen indywidualnych brak jest podstaw do odrzucenia hipotez zerowych każdego z testów. Uzyskane wyniki statystyk są niskie i przy typowej wartości poziomu istotności $p = 0.05$ decyzje diagnostyczne radiologów podjęte na podstawie dwóch grup obrazów 0.1 bpp oraz 0.04 bpp nie różnią się znacząco od decyzji sformułowanych przy obserwacji wersji oryginalnych tych obrazów. Jednak w dość istotnym teście różnic powierzchni obrazu klasy 0.04 bpp znajdują się bardzo blisko granicy odrzucenia. Test dwuwymiarowy pokazuje wyraźnie większą wartość statystyki dla obrazów 0.04 bpp, jednak jeszcze daleką od granicy odrzucenia. Wydaje się więc bezpieczną opinią o zbyt małej różnicy w charakterze decyzji diagnostycznych oryginałów i obrazów 0.1 bpp, aby stwierdzić istotną różnicę wartości diagnostycznej obrazów tych dwóch klas. W przypadku obrazów 0.04 bpp ostrożne i rozsądne wydaje się stwierdzenie o wystąpieniu istotnego poziomu różnic wartości diagnostycznej tych obrazów w stosunku do oryginałów.

Skale sumaryczne SB i SC uwzględniające nadrozpoznanie zdają się lepiej oddawać charakter podejmowanych decyzji diagnostycznych niż skala SA. Wszystkie one mają charakter uśredniający, mogą więc wprowadzać uproszczenia w opisie procesu decyzyjnego. Według skali ocen SA należałoby zdecydowanie odrzucić wszystkie testowane hipotezy. Mniej jednoznaczne wnioski można wyciągnąć w przypadku dwóch pozostałych skal. Przyjęcie skali SB w teście dwuwymiarowym prowadzi do odrzucenia hipotezy, iż oceny z obrazów oryginalnych i 0.04 bpp pochodzą z tej samej binormalnej krzywej ROC. Brak jest podstaw do odrzucenia takiej hipotezy w przypadku oryginałów i obrazów 0.1 bpp, jakkolwiek wartość statystyki testu znajduje się blisko ustalonej granicy. Wyraźne odrzucenie hipotez w testach powierzchni i TPF dla obu grup obrazów, z zachowaniem jednak hierarchii bardziej zgodnych z oryginałem decyzji z obrazów 0.1 bpp sugeruje zakwalifikowanie obrazów 0.04 bpp jako gorszych pod względem wartości diagnostycznej, natomiast dość niejednoznaczną akceptację wiarygodności diagnostycznej obrazów 0.1 bpp, głównie na podstawie wyniku z najbardziej kompleksowego testu dwuwymiarowego. Wykorzystanie skali SC pozwala przy pomocy tego samego testu dość jednoznacznie potwierdzić zbliżoną do oryginalnych wersji wartość diagnostyczną obrazów 0.1 bpp oraz 0.04 bpp, co jest szczególnie cenne gdyż jako jedyna ta skala sumacyjna uwzględnia wszystkie nadrozpoznanie. Wyraźnie lepiej wypadła także ocena obrazów obu tych grup w testach powierzchni i TPF w porównaniu ze skalą SB.

Tabela 14. Dane przygotowane w celu przeprowadzenia testów statystycznych ze skalą detekcji patologii dla ocen od indywidualnych obserwatorów, a także z, na różny sposób komponowaną, sumaryczną skalą ocen obrazu: SA: S(pat) = NPPO, S(bezpat) = NFPO; SB: S(pat) = NPPO, S(bezpat) = NNA; SC: S(pat) = NPPO-NNA, S(bezpat) = NNA.

Obrazy	Ocena indywidualna (I)							Ocena sumaryczna		
	Obserwator									
	1	2	3	4	5	6	7	SA	SB	SC
ORY - pat										
1	4	3	3	3	4	4	4	7	7	4
2	4	4	4	3	4	4	4	7	7	6
4	4	2	2	4	2	2	2	2	2	0
5	4	4	3	4	3	4	3	7	7	4
8	4	2	4	4	4	1	2	4	4	2
10	4	4	4	3	4	4	4	7	7	6
11	4	4	4	3	4	4	4	7	7	6
12	3	4	2	3	2	4	2	3	3	0
13	4	4	4	4	4	4	4	7	7	7
ORY - bezpat										
3	4	4	4	4	4	3	1	6	5	5
6	1	1	1	1	4	1	1	1	1	1
7	1	4	4	4	4	1	4	5	5	5
9	4	3	4	1	4	3	4	6	4	4
01 - pat										
1j1	3	1	4	3	4	1	3	5	5	2
2j1	3	4	4	4	4	4	3	7	7	4
4j1	2	2	3	2	2	2	2	1	1	0
5j1	4	4	4	4	4	4	3	7	7	6
8j1	3	4	4	4	1	1	1	4	4	3
10j1	4	4	4	3	1	4	3	6	6	4
11j1	3	4	4	3	4	4	4	7	7	5
12m1	3	4	4	3	2	1	2	4	4	0
13j1	4	3	1	3	2	3	4	5	5	1
01 - bezpat										
3j1	4	4	4	4	4	1	4	6	6	6
6j1	1	1	1	1	4	1	2	1	2	2
7j1	4	4	4	4	4	1	4	6	6	6
9m1	4	4	1	4	4	4	3	6	5	5
004 - pat										
1j04	4	3	1	3	4	1	4	5	5	3
2j04	3	4	4	3	4	2	4	6	6	3
4j04	4	2	4	1	2	4	2	3	3	0
5j04	4	4	4	4	4	4	4	7	7	7
8j04	4	4	4	4	4	1	1	5	5	5
10j04	2	4	4	3	1	2	1	3	3	0
11j04	4	4	4	3	4	4	4	7	7	6
12m04	2	4	3	2	2	2	4	3	3	0
13j04	1	2	1	2	1	1	1	0	0	0
004 - bezpat										
3j04	4	4	4	3	4	4	4	7	6	6
6j04	4	4	3	3	1	1	1	4	2	2
7j04	4	4	4	4	4	1	4	6	6	6
9m04	4	4	4	4	4	1	4	6	6	6

Tabela 15. Analiza statystyczna z wykorzystaniem trzech testów na bazie estymacji binormalnych krzywych ROC jak na rys. 12. Pogrubiono wyniki sugerujące odrzucenie odpowiednich hipotez zerowych (przy poziomie istotności $p = 0.05$). Dla testu TPF przytoczono reprezentatywne wyniki przy wartości FPF=0.5.

Testy	Dane wejściowe	ORY - 0.1bpp		ORY - 0.04bpp	
		statystyka	wartość p	statystyka	wartość p
Dwuwymiarowy chi-kwadrat	I	1.9415	0.3788	3.1734	0.2046
	SA	6.4867	0.0390	7.7516	0.0207
	SB	5.8297	0.0542	6.1962	0.0451
	SC	3.9242	0.1406	2.8190	0.2443
Powierzchni	I	1.3499	0.0885	1.6151	0.0531
	SA	2.6661	0.0038	3.2383	0.0006
	SB	2.4498	0.0071	2.6038	0.0046
	SC	2.3071	0.0105	1.7564	0.0395
TPF (FPF=0.5)	I	-0.8714	0.1918	-0.6135	0.2698
	SA	-2.5224	0.0058	-2.7644	0.0028
	SB	-2.3402	0.0096	-2.3971	0.0082
	SC	-1.9780	0.0240	-1.6509	0.0494

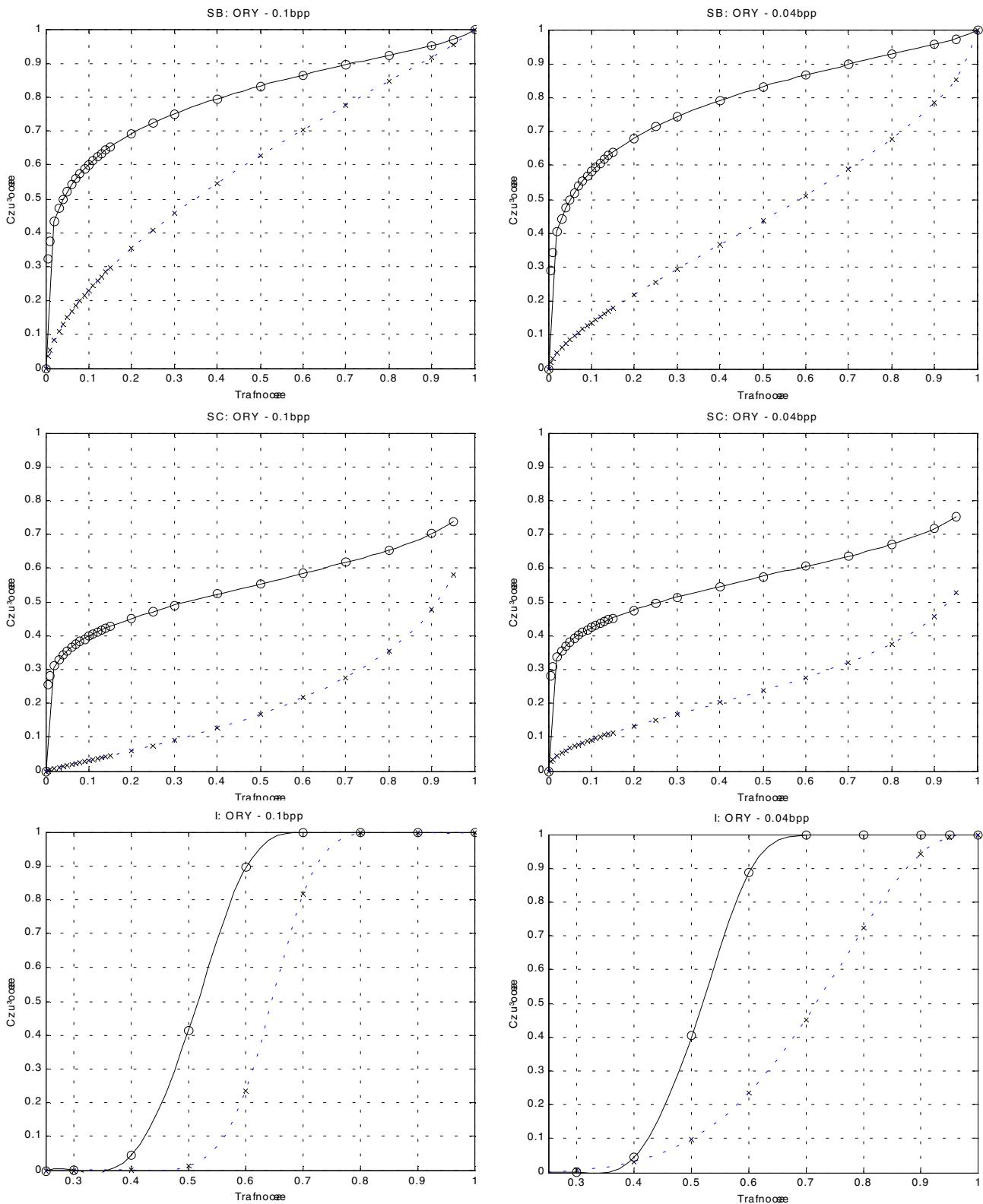
Powyższa analiza pozwala sformułować wiele różnorodnych wniosków w zależności od przyjętych kryteriów istotności poszczególnych decyzji. Jeśli przykładowo nie są ważne nadrozpoznanie, a chodzi tylko, aby radiolog nie 'zgubił' przypadków patologii w klasycznym modelu binarnej krzywej ROC, wówczas zastosowanie skali SA prowadzi do wniosku, iż kompresja do wartości średnich 0.1 bpp oraz 0.4 bpp jest niedopuszczalna.

Generalnie, na podstawie wyników testu dwuwymiarowego można potwierdzić wiarygodność diagnostyczną obrazów rekonstruowanych ze średniej bitowej 0.1 bpp.

Wnioski, zebrane opinie

Według większości radiologów biorących udział w teście, ostateczne wykrycie patologii i jej charakterystyka (niezależnie od wersji obrazu) w większości przypadków są niemiernodajne, ponieważ ocena nie spełnia standardów oceny mammografii (brak możliwości porównania z obrazem drugiego sutka, brak innych projekcji, badania klinicznego i USG). Jedynie ewidentne zmiany, guzki dobrze odgraniczone i spikularne były jednoznacznie definiowane i rozpoznawane. W pozostałych przypadkach właściwa patologia była często pomijana, natomiast podejrzenia budziły miejsca 'bezmianowe', np. zagęszczenia gruczołowe. Był to wynik oglądania obrazu pomniejszonego oraz fragmentarycznie, bez drugiej projekcji i możliwości porównania ze zdjęciami drugiego sutka.

Pomijane były również słabo wysyczone mikrozwapnienia, natomiast rozpoznawano je w sutkach ze zmianami włóknistymi (bez mikrozwapnień). W przypadku wykrycia mikrozwapnień, ocena ich morfologii była często niemożliwa. Również określenie ich ilości w niektórych obrazach przetworzonych nie było miarodajne



Rys. 12. Estymacja binormalnych krzywych ROC przy dwóch sumarycznych oraz indywidualnej skali ocen. Linią ciągłą zaznaczono krzywe dla obrazów oryginalnych.

Bardziej wiarygodnym wydaje się test oceny jakości cech istotnych diagnostycznie konkretnej zmiany. Również w tym teście (tak jak w teście detekcji) problem stwarzały mikrozwapnienia, będące ważnym elementem diagnostycznych w ocenie patologii sutka, a często jej jedynym objawem. W obrazach rekonstruowanych możliwe było ich stwierdzenie,

natomiast nie próbowano określać ich morfologii. Wydaje się istotnym przeprowadzenie w przyszłości testu optymalizowanego pod kątem detekcji i charakterystyki mikrozwapnień.

Zbieżność ocen obrazów oraz opinii radiologów wyrażonych w teście i na temat testu pozwala na wysunięcie wstępnych wniosków, że stosowanie falowych metod kompresji w sugerowanych granicach wartości średnich bitowych nie zmniejsza wartości diagnostycznej obrazów. Porównanie obrazów oryginalnych (skanowanych) i rekonstruowanych nie wykazuje różnic istotnych diagnostycznie. Według zebranych opinii radiologów należałoby przeprowadzić dodatkowy test wykorzystując także oryginalną wersję analogową mammogramu (kliszę). Jeżeli potwierdziłby on zachowanie wartości diagnostycznej zarówno oryginalnej wersji cyfrowej (skanowanej), jak i rekonstruowanej, wówczas wnioski końcowe byłyby bardziej wiarygodne.

Nie sposób było przeprowadzić testu z wykorzystaniem obrazów analogowych w sposób obiektywny i wiarygodny ze względu na obowiązującą regułę wydawania badań pacjentkom. Obrazy wykorzystane w badaniach były gromadzone w postaci cyfrowej przez wiele miesięcy, potem dokonano ich klasyfikacji, selekcji do celów projektowanych testów, itd. Nie można było przechować oryginałów przez tak długi okres czasu. Wykonano jednak pracę, która pośrednio dowodzi równoważności analogowej i cyfrowej postaci wykorzystanych obrazów testowych. Przy dobieraniu parametrów skalowania, optymalizacji rozdzielczości, dynamiki, samego urządzenia skanującego, odwoływano się do opinii radiologów ustalając optymalne warunki, pozwalające zachować pełną wiarygodność wersji cyfrowych. Opisano to w [3].

Na podstawie opinii i uwag lekarzy zebranych w trakcie eksperymentów, wyników statystycznej analizy testów detekcji patologii oraz ocen wiarygodności kompresowanych stratnie obrazów dopuszczalna, graniczna wartość średniej bitowej dla szerokiego spektrum wykorzystanych badań mammograficznych wynosi 0.1 bpp.

Można stwierdzić, że zastosowane metody kompresji obrazów mogą być wykorzystane z powodzeniem w medycznych systemach informacyjnych, a w wybranych przypadkach (typach patologii) okazują się nawet pomocne w ich ocenie. Trzeba jednak pamiętać, że wartości średniej bitowej skompresowanej reprezentacji nie powinna przekroczyć dopuszczalnej wartości granicznej. Gdyby prezentowane tutaj wnioski uznano za w pełni wiarygodne, to w dużym stopniu proponowane metody kompresji rozwiązałyby problem archiwizacji. Jak wspomniano, zdjęcia mammograficzne w większości ośrodków są wydawane pacjentkom. Przy kolejnych badaniach kontrolnych często zdarza się, że pacjentki z różnych powodów nie dostarczają poprzednich wyników badań. Nie sposób więc przeprowadzić tak istotnej analizy porównawczej badań wykonanych w różnym czasie. Stwarza to czasami duże problemy przy interpretacji obrazu, a budzące wątpliwości zmiany okazują się ostatecznie normą dla danej pacjentki.

Innym zagadnieniem są konsultacje zdjęć w przypadkach trudnych, niejednoznacznych w interpretacji zmian. Wydaje się, że bardzo pomocna wtedy telekonsultacja byłaby efektywnie realizowana z wykorzystaniem wskazanych koderów falowych, kształtujących progresywny strumień danych obrazowych.

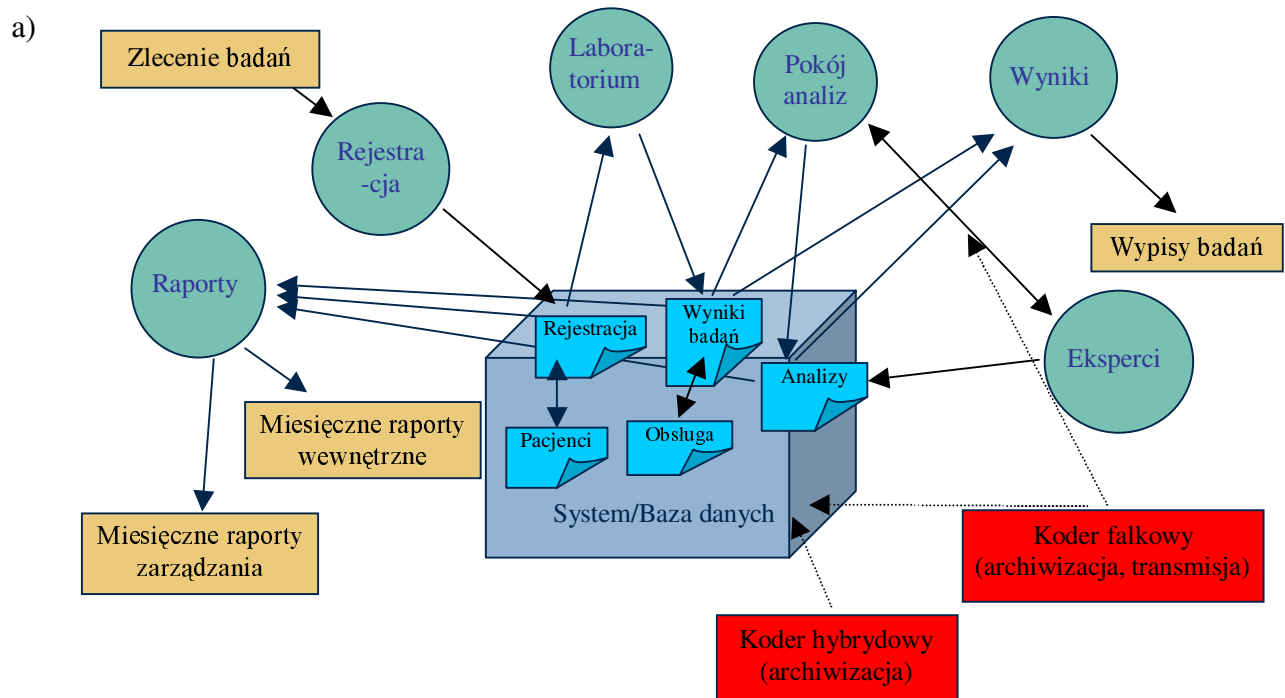
8. PRAKTYCZNE WYKORZYSTANIE ALGORYTMÓW KOMPRESJI W MEDYCZNYCH SYSTEMACH INFORMACYJNYCH

Poniżej przedstawiono kilka własnych koncepcji systemów telemedycznych, w tym przykładowe struktury funkcjonalne, oprogramowanie oraz architektura modułów HIS (ang.

hospital information system). Analizowano, optymalizowano i zrealizowano różne warianty hierarchicznej struktury funkcjonalnej takich systemów, zasad regulujących dostęp do danych chronionych oraz inne rozwiązania związane z restrykcyjnymi prawami korzystania z danych medycznych, zabezpieczenia przed błędami, itp. Wybrane schematy prostych rozwiązań dedykowanych pojedynczemu ośrodkowi medycznemu są realizowane w Szpitalu Wolskim w Warszawie. Zbudowano tam sieć wewnętrzną, która będzie rozbudowywana w przyszłości w miarę możliwości finansowych i organizacyjnych. Fragment projektu tej sieci wraz ze schematem załączono do tego sprawozdania.

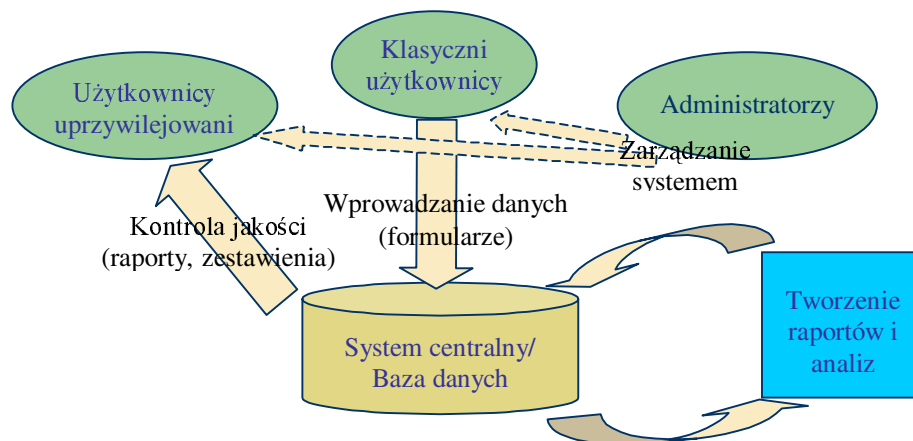
W sieci tej implementowane są różnorodne narzędzia wspomagających decyzje diagnostyczne oraz procedury ułatwiające transmisję i archiwizację danych obrazowych (algorytmy kompresji falkowej MBWT, zgodne z JPEG2000, hybrydowy system kompresji). Wykorzystuje się rozwiązania zapewniające sprawny przepływ informacji pomiędzy różnymi punktami (pokojami, piętrami, rozrzuconymi oddziałami, centrami, zakładami) sieci wewnętrznej (intranet). Dają one możliwość konsultacji specjalistów na odległość. Przykładowe schematy pokazano na rys. 13.

Zestaw zrealizowanych procedur umożliwi telediagnozę w czasie rzeczywistym (z wykorzystaniem ultrasonografu HP Image Point w wersji testowej). Urządzenia digitalizacji analogowych danych z aparatu, procedury przetwarzania obrazów (także tworzenia i analizy obrazów przestrzennych) oraz moduły archiwizacji i komunikacji zapewniają łatwy dostęp do danych obrazowych, zgodnie z żądaniami klienta. Interaktywna konsultacja, ekstrakcja diagnostycznie istotnych cech, wymiana danych obrazowych z bazy serwera, a także obrazów wprowadzanych przez klienta czynią ten system bardziej użytecznym. Inny zestaw algorytmów i procedur realizuje komputerowe wspomaganie diagnozy CAD (ang. Computer-Aided Diagnosis) w badaniach mammograficznych. Zapewniona jest możliwość konsultacji z drugim obserwatorem oraz ekspertem, zaimplementowane są procedury 'inteligentnej prezentacji' uwydatniające cechy obrazu szczególnie ważne w diagnozie, a także automatyczne metody detekcji patologii złośliwych, głównie mikrozwapnień, sugerujące lekarzom obszary godne szczególnej uwagi w procesie diagnozy.

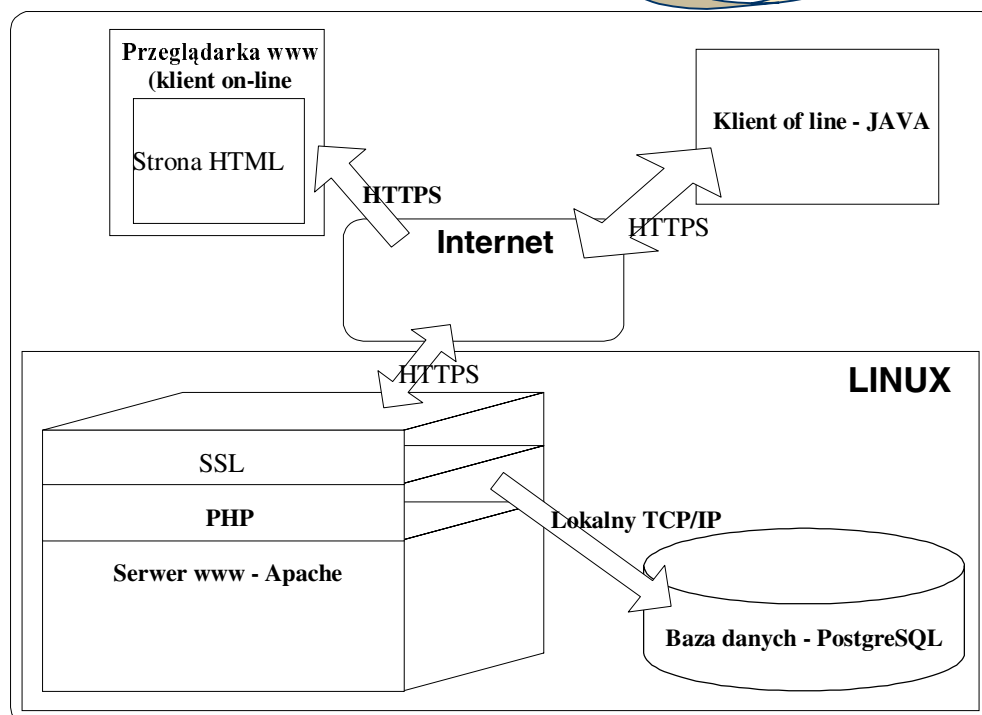


Ponadto, opracowano czteropoziomą hierarchiczną strukturę wspierającą kontrolę jakości badań medycznych (konceptja systemu kontroli jakości powstała we współpracy z Instytutem Matki i Dziecka w Warszawie). Na najniższym poziomie takiej struktury znajduje się zakład czy pracownia, potem pojedynczy szpital czy klinika. Pracę ośrodków medycznych w danym regionie nadzoruje regionalne centrum zarządzania (na poziomie struktury równoległej do Kas Chorych), a pracę regionalnych ośrodków kontroli koordynuje centrum krajowe. Różne rodzaje informacji mogą być wymieniane pomiędzy poszczególnymi poziomami takiej hipotetycznej struktury (system wielodostępny, dostęp nie może być ograniczony przez obszar, wielu użytkowników pracujących jednocześnie). Reguły dostępu muszą być bezpieczne, ale jednocześnie wygodne, ze zróżnicowaną hierarchią użytkowników, często zmienną w czasie i zależną od różnych czynników (aktualna struktura badań, plan pracy poszczególnych lekarzy, tryb kontroli, itd.). Do systemu dostarczana jest duża ilość danych, także obrazowych, która musi być łatwo i szybko dostępna. Archiwum winno mieć więc cechy typowej hurtowni danych z implementacją efektywnych algorytmów kompresji. Koszty systemu powinny być w miarę możliwości minimalne, tak w procesie opracowania i wdrażania, jak i w czasie użytkowania. Należało więc zbudować system na bazie ogólnie dostępnych narzędzi, tanich lub wręcz darmowych. Przykład realizacji tej koncepcji przedstawiono na rys. 14.

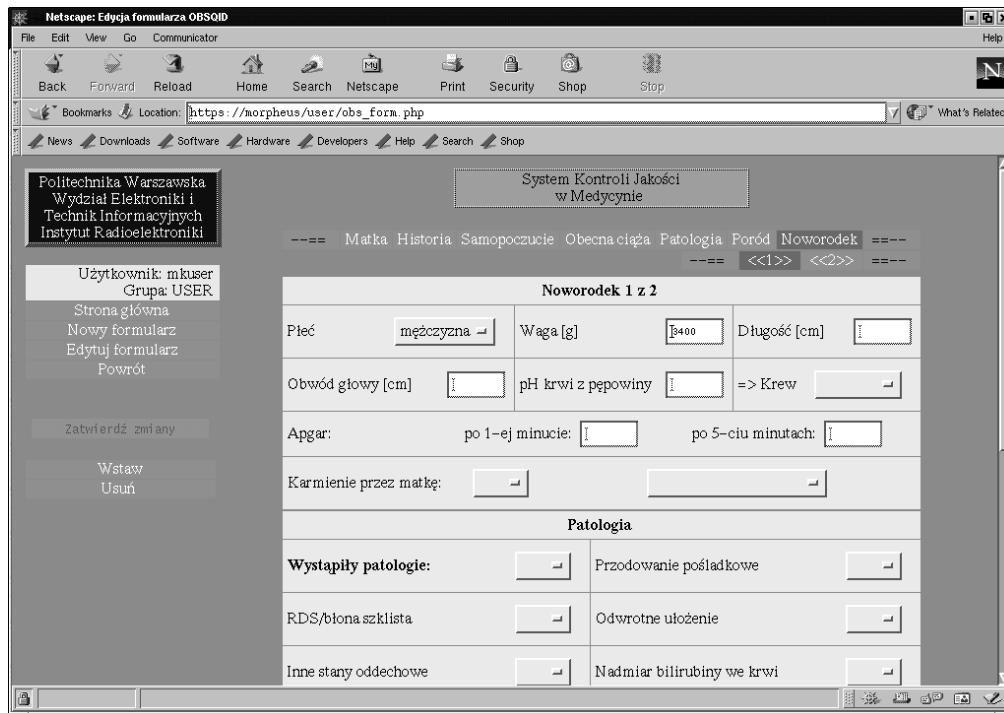
a)



b)



c)

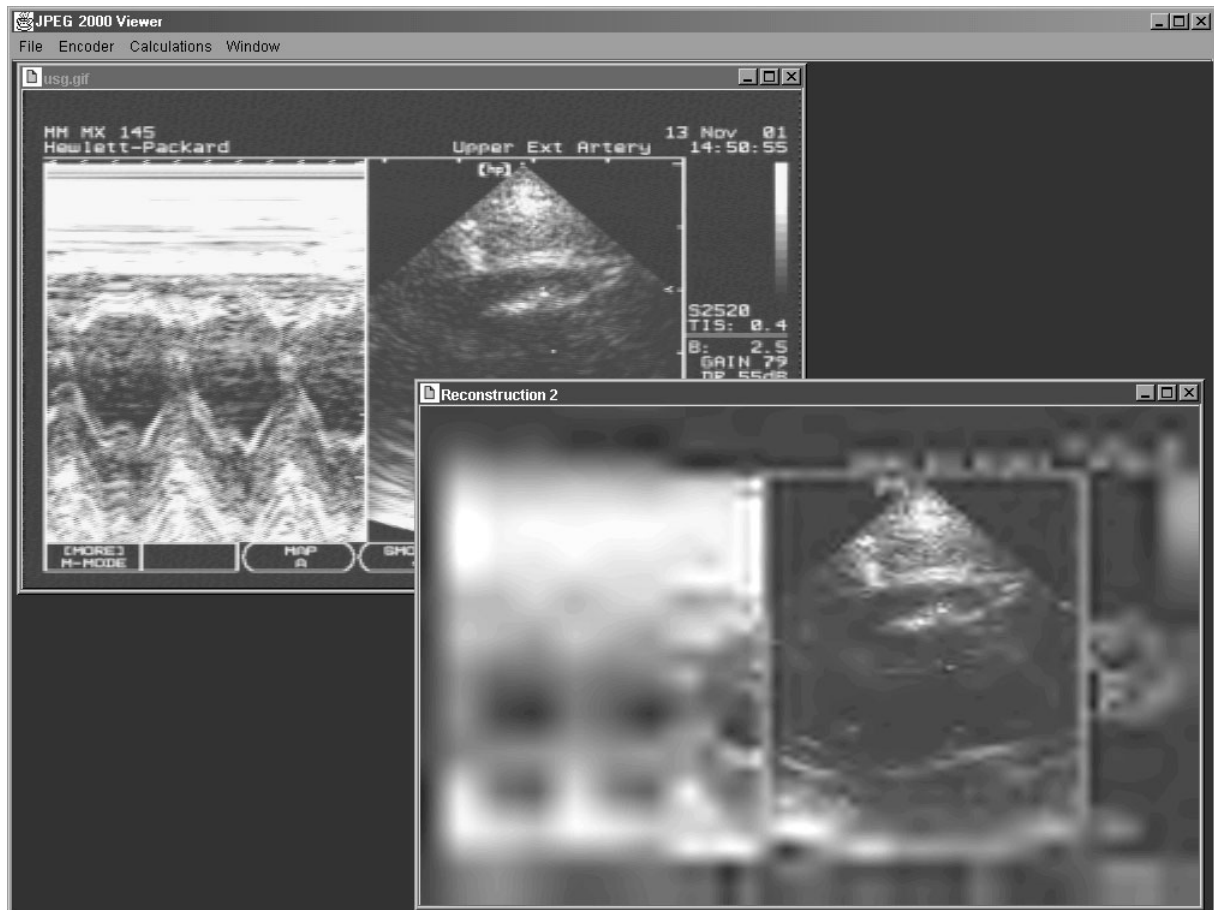


Rys. 14. System kontroli jakości z hurtownią danych: a) schemat funkcjonalny, b) przykładowa architektura, c) wybrany interfejs użytkownika.

Wszystkie te systemy wymagają optymalizacji reprezentacji różnego typu danych, w tym szczególnie ważnej reprezentacji danych obrazowych, głównie ze względu na dużą objętość zbiorów i wysoką wagę w diagnozie, konieczność dodatkowych konsultacji w przypadkach wątpliwych, itp. Istotny jest dobór odpowiedniej techniki kompresji, formatu zapisu danych, protokołu transmisji czy struktury bazy danych, technologii wyszukiwania, itd. Zewnętrzna reprezentacja w standardzie DICOM jest przetwarzana w szereg efektywnych postaci reprezentacji wewnętrznych tak, aby zwiększyć wydajność, szybkość pracy i użyteczność całego systemu telemedycznego. Więcej szczegółów odnośnie przykładowych aplikacji można znaleźć w pracach dyplomowych przygotowywanych przez studentów prowadzonych przez Autora [11][4][7][5].

Aplikacja kodera JPEG2000

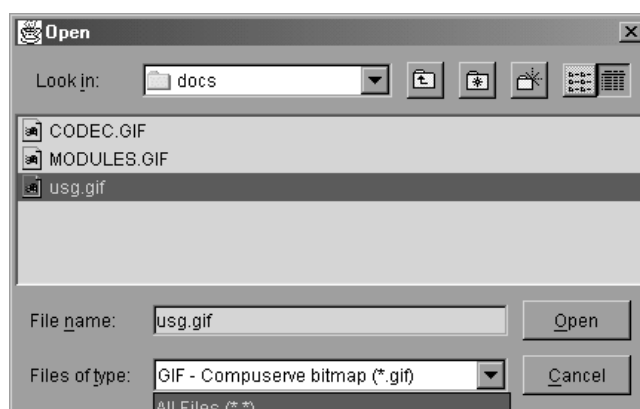
Służy ona do optymalizacji docelowego apletu (Java) realizującego kompresję według JPEG2000 przy zoptymalizowanych parametrach, dobranych dla konkretnej klasy obrazów medycznych. Aplet ten jest włączony w system telemedyczny z przeznaczeniem do progresywnej transmisji obrazów (rys. 15).



Rys. 15. Przykładowy widok aplikacji – przeglądarki z wczytanym obrazem medycznym zapisanym w formacie GIF oraz obraz zrekonstruowany na podstawie strumienia JPEG2000.

Proces wytworzenia strumienia JPEG2000 przebiega w kilku etapach:

- Wczytanie obrazu do programu
Wejściem modułu wczytującego dane jest standardowy interfejs *Image* zdefiniowany w standardowych bibliotekach Javy - czytane są obrazy w formatach: PPM, PGM -8 bitowe-, PGX -16 bitowe- oraz GIF (rys.16).

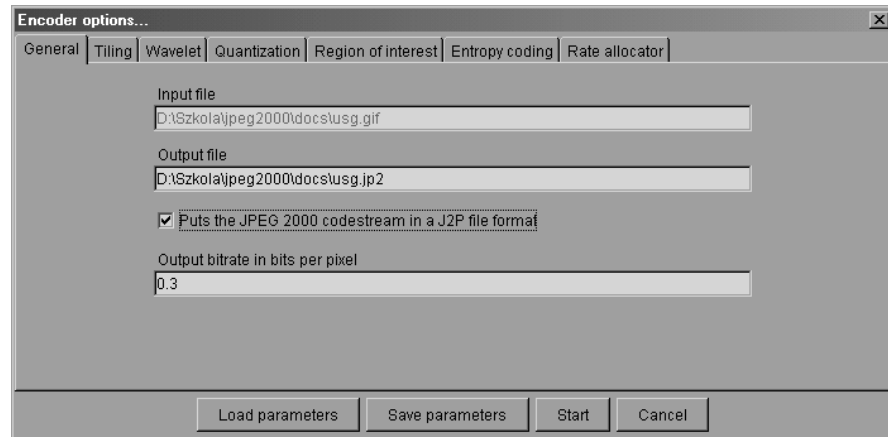


Rys. 16. Funkcja czytania obrazów przygotowanych w odpowiednich formatach.

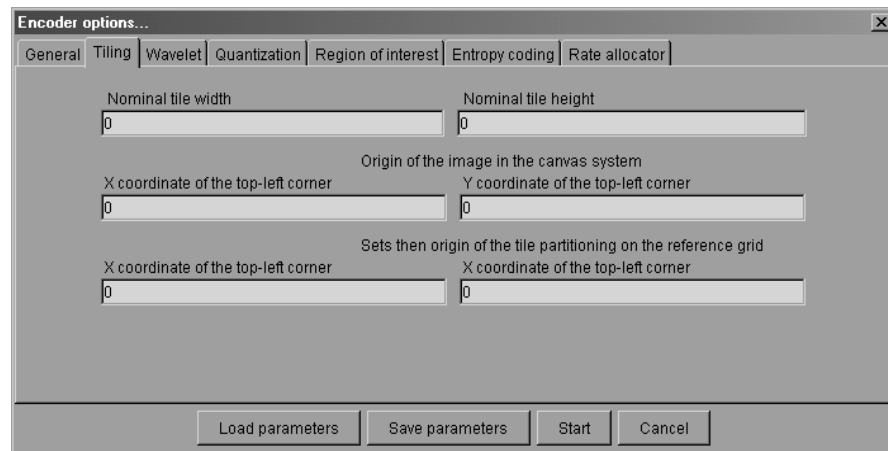
- Określenie parametrów kompresji

Parametry kompresji określone są z okna dialogowego jak na rys. 17:

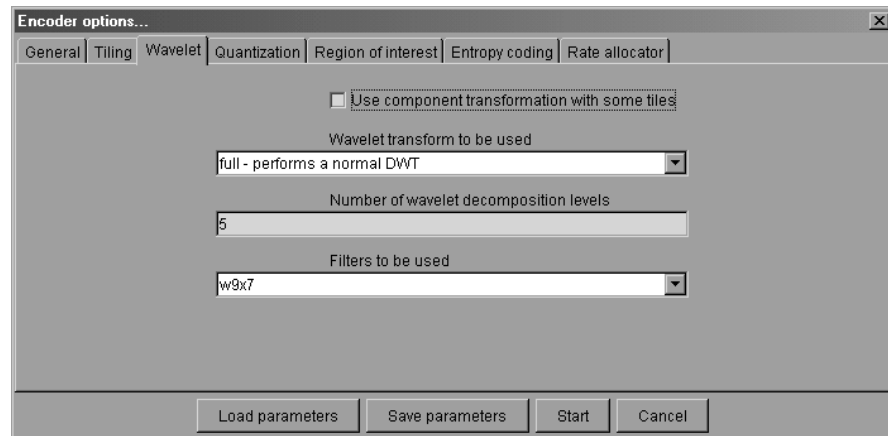
Parametry ogólne



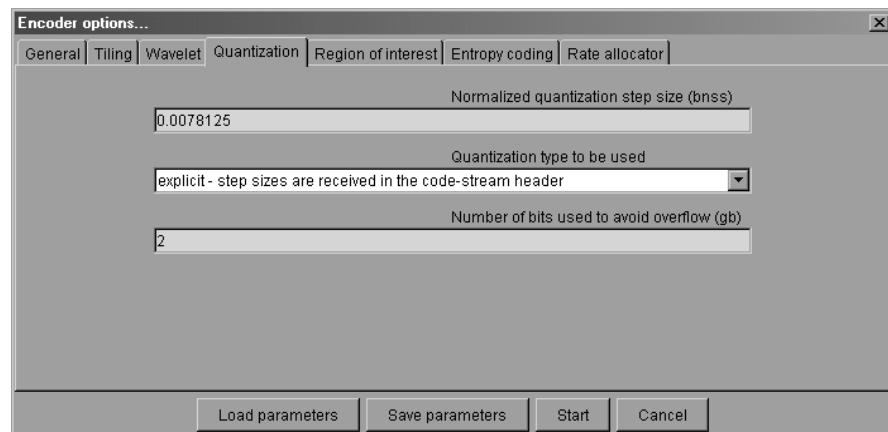
Podział na części
(ang. tiling)



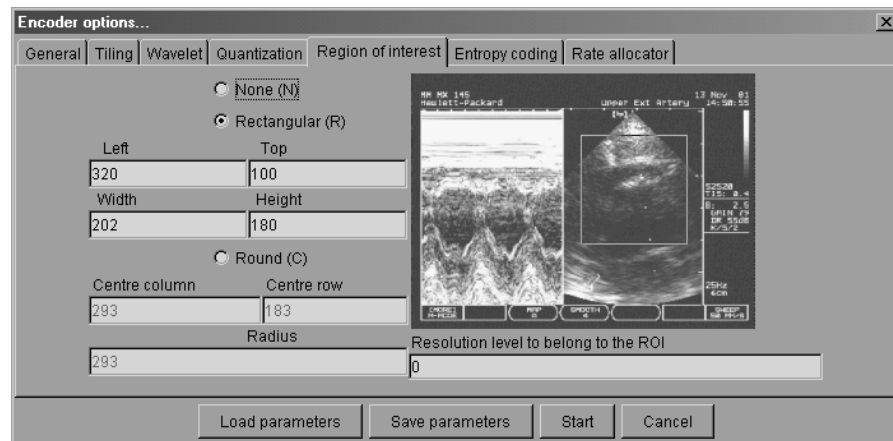
Transformacja
falkowa



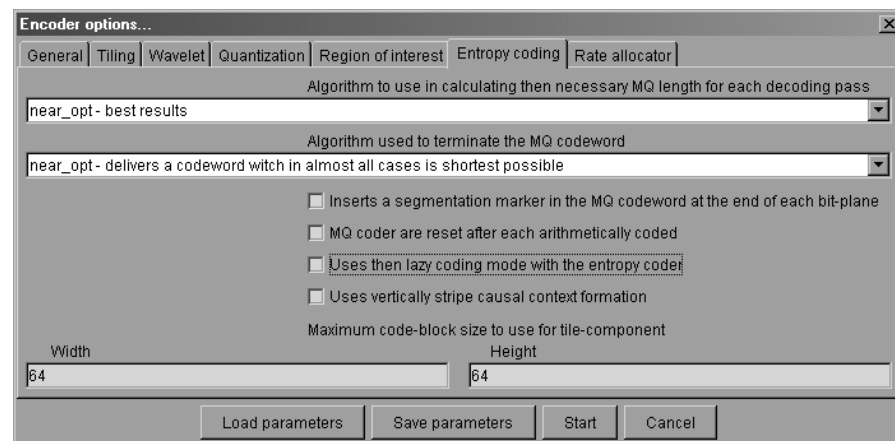
Kwantyzacja



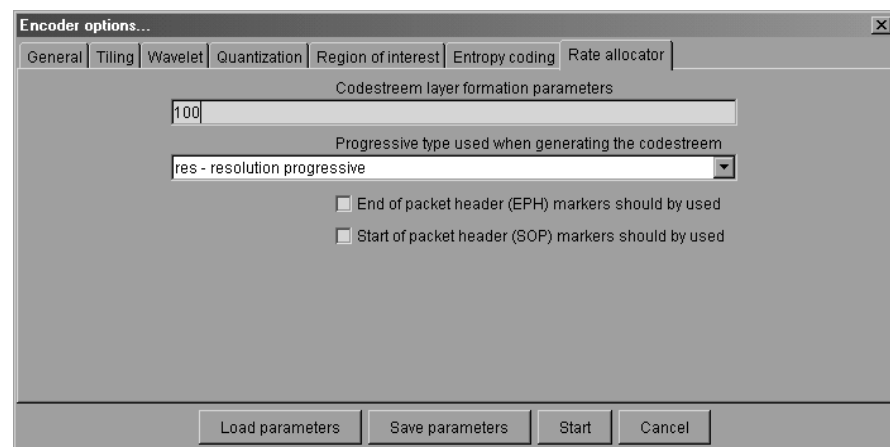
Wybór obszaru zainteresowania (ROI)



Kodowanie entropijne



Progresywna transmisja



Rys. 17. Ustalenie parametrów kompresji w sieciowej wersji kodera JPEG2000.

- Interakcyjna procedura tworzenia strumienia progresywnego

Wykonywana jest kompresja z możliwością przerwania procesu generacji strumienia kodowego w każdej chwili, kontrolą błędów transmisji, możliwością wyboru transmisji dokładniejszej dla wybranego ROI, itd.

Stworzono uniwersalną bibliotekę kodera/dekodera w standardzie JPEG2000 (pakiet jar) umożliwiającą kompresję danych do strumienia jpeg2000 oraz dekompresję strumienia jpeg2000 w różnych schematach: na wejściu strumień bajtowy reprezentowany przez obiekt klasy `DataInputStream` (plik, pamięć, gniazdko sieciowe, ekran i inne) lub wczytana wcześniej bajtowa tablica, na wyjściu zaś strumień bajtowy reprezentowany przez obiekt

klasy `DataOutputStream`. Obecnie biblioteka jest testowana w systemie teledyecznym w Szpitalu Wolskim.

Bibliografia

- [1] Betts B.J., Li J., Cosman P.C., Gray R.M. *et al.*: Image Quality in Digital Mammography. Revision of final report to the Army Medical Research and Materiel Command, *Compression and Classification of Digital Mammograms for Storage, Transmission, and Computer Aided Screening*, <http://www-isl.stanford.edu/~gray/armyfinal.pdf>, September 1998.
- [2] <http://www-radiology.uchicago.edu/krl/toppage11.htm>
- [3] Kawalec T.: System do wspomagania diagnostyki raka sutka. Praca dyplomowa magisterska pod kierunkiem dra A.Przelaskowskiego, Instytut Radioelektroniki PW, 2001.
- [4] Kawczyński M.: Hurtownia danych w teledyecznym systemie kontroli jakości. Praca dyplomowa na ukończeniu, Instytut Radioelektroniki PW.
- [5] M. Skakuj: Kompresja falkowa według standardu JPEG2000. Praca dyplomowa na ukończeniu, Instytut Radioelektroniki PW.
- [6] Ma G, Hall WJ.: Confidence bands for receiver operating characteristic curves. *Med Decis Making* 13: 191 - 197, 1993.
- [7] Maksymiuk K., Skaliński M.: Interaktywne przesyłanie danych w sieci teledyecznej. Praca dyplomowa na ukończeniu, Instytut Radioelektroniki PW.
- [8] Metz C.E, Herman BA, Roe CA.: Statistical comparison of two ROC curve estimates obtained from partially-paired datasets. *Med Decis Making* 18: 110-121, 1998.
- [9] Metz C.E, Wang P-L, Kronman H.B.: A new approach for testing the significance of differences between ROC curves measured from correlated data. *Information Processing in Medical Imaging, Proc. of the 8th Conference*, pp. 432-445, 1984.
- [10] Metz C.E.: Some practical issues of experimental design and data analysis in radiological ROC studies, *Investigative Radiology*, 24(3), pp. 234-245, 1989.
- [11] Pietrasik A.: Komputerowe wspomaganie diagnostyki raka sutka na podstawie badań mammograficznych. Praca dyplomowa na ukończeniu, Instytut Radioelektroniki PW.
- [12] Przelaskowski A., Kawczyński M., Maksymiuk K., Skaliński M.: Effective image compression tools and telemedicine systems. *Materiały 1-st Polish-Norwegian Seminar: Selected Research Issues at the Polish and Norwegian Universities, Working group "Medical Technology"*, str. 5-6, Politechnika Warszawska, Październik 2001.
- [13] Przelaskowski A.: Effective integer-to-integer transforms for JPEG2000 coder. *SPIE Conference: Wavelets: Applications in Signal and Image Processing IX, Proceedings of SPIE*, 4478:299-310, 2001.
- [14] Przelaskowski A.: Elastyczność koderów falkowych w systemach archiwizacji i transmisji medycznych danych obrazowych. *Prace Naukowe Politechniki Warszawskiej - Elektronika*, z. 130, str. 105-122, Oficyna Wydawnicza PW, 2001.
- [15] Przelaskowski A.: Falkowe metody kompresji danych obrazowych jako narzędzia kształtowania optymalnej reprezentacji strumienia danych. *Materiały II Seminarium Radiokomunikacja i Techniki Multimedialne*, str. 23-30, Instytut Radioelektroniki PW, 2001.
- [16] Przelaskowski A.: Falkowe metody kompresji danych obrazowych. *Rozprawa habilitacyjna złożona do druku w Oficynie Wydawniczej Politechniki Warszawskiej*, styczeń 2002.
- [17] Przelaskowski A.: Hybrid lossless coder of medical images with statistical data modelling. *Lecture Notes in Computer Science, Springer Verlag*, 2124: 92-101, 2001.
- [18] Przelaskowski A.: Hybrid Vector Measures of Compressed Medical Images, *SPIE Symposium Medical Imaging : Image Perception and Performance*, <http://sigma.ire.pw.edu.pl/zejim/rois/publikacje/Measure.pdf>, 2000.
- [19] Przelaskowski A.: Lifting-based reversible transforms for lossy-to-lossless wavelet coders. *Lecture Notes in Computer Science, Springer Verlag*, 2124: 61-70, 2001.
- [20] Przelaskowski A.: Lossless encoding of medical images: hybrid modification of statistical modelling-based conception. *Journal of Electronic Imaging*, 10(4): 966-976, October 2001.
- [21] Przelaskowski A.: Modifications of uniform quantization applied in wavelet coder. *Proceedings of IEEE Data Compression Conference, Snowbird, Utah*, str. 293-302, 2000.

- [22] Przelaskowski A.: Performance evaluation of jpeg2000-like data decomposition schemes in wavelet codec. Proceedings of IEEE International Conference on Image Processing, str. 788-791, 2001.
- [23] Przelaskowski A.: Progressive image data compression with adaptive scale-space quantization. Proceedings of SPIE, Internet Imaging, 3964:143-154, 2000.
- [24] Przelaskowski A.: Techniki optymalizacji falkowej reprezentacji obrazów medycznych. Prace Naukowe Politechniki Warszawskiej - Elektronika, z. 130, str. 123-142, Oficyna Wydawnicza PW, 2001.
- [25] Przelaskowski A.: Today's and tomorrow's Medical Imaging. Lecture Notes in Computer Science, Springer Verlag, 2124:236-237, 2001.
- [26] Przelaskowski A.: Vector measure with scalar equivalent for quality estimation of compressed medical images. <http://sigma.ire.pw.edu.pl/zejim/rois/publikacje/HVM.pdf>, zgłoszony do Journal of Electronic Imaging, 2001.
- [27] Said A., Pearlman W.A.: A new fast and efficient image codec based on set partitioning in hierarchical trees. IEEE Trans. Circ. & Syst. Video. Tech., 6:243-250, 1996.