

AUTOMATSKO PREPOZNAVANJE TUMORA DOJKE PRIMENOM MULTIREZOLUCIONOG ALGORITMA

Marina Đoković, *Univerzitet u Kragujevcu, Tehnički fakultet Čačak, marina.m.djokovic@gmail.com*
Aleksandar Peulić, *Univerzitet u Kragujevcu, Tehnički fakultet Čačak, apeulic@tfc.kg.ac.rs*
Nenad Filipović, *Univerzitet u Kragujevcu, Mašinski fakultet, Kragujevac, fica@kg.ac.rs*
Đorđe Damnjanović, *Univerzitet u Kragujevcu, Tehnički fakultet Čačak, fic177@hotmail.com*

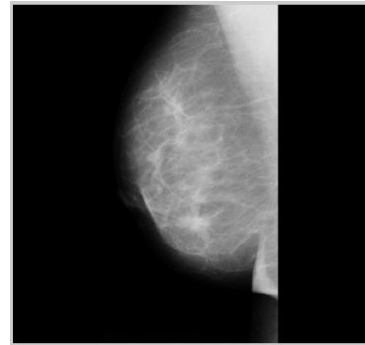
Sadržaj – U ovom radu predstavljen je jedan od načina za detektovanje tumora dojke primenom multirezolucionog algoritma. Primenom multirezolucione analize najpre je uklonjen šum sa mamogramu pa je nakon toga detektovan tumor. Potom je, primenom Embedded Zerotree Wavelet (EZW) algoritma, kompresovana slika sa koje je uklonjen šum. Pokazano je da se primenom algoritma za detektovanje tumora na kompresovanu sliku dobije isti rezultat kao u slučaju nekomresovane slike. Eksperimentalni rezultati dobijeni na osnovu mamografskih slika pacijenata snimljenih u Kliničkom centru u Kragujevcu, pokazuju da je pomoću multirezolucionog algoritma moguće detektovati tumore različitih dimenzija.

1. UVOD

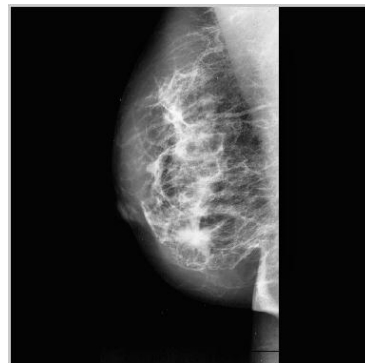
Tumor dojke je najčešći oblik tumora kod žena širom sveta i drugi glavni uzročnik smrti, odmah nakon tumora pluća [1]. Mamografija je danas najbolji način za rano otkrivanje tumora dojke kod žena starijih od 35- 40 godina. Mamografija predstavlja rendgenski snimak dojke i veoma je važna dijagnostička metoda, jer često može da otkrije tumor pre nego što žena oseti bilo kakve promene [2]. Digitalna mamografija donela je veliki napredak u računarskoj obradi slike. Digitalna mamografija koristi kompjutersku sliku dojke, dobijenu nakon konverzije rendgenskog snimka u digitalnu sliku. Pri tome kompjuterska tehnika omogućava proučavanje detalja slike, koji se po potrebi mogu uvećavati ili izdvajati i upoređivati s ranijim nalazima. Digitalna mamografija smanjuje dozu radijacije i omogućava bolje snimanje tkiva dojke, te povećanje određenih mesta na slici radi bolje analize.

Tumor dojke najčešće je zvezdastog oblika ali zvezdasta senka istovremeno je i veoma nesiguran pokazatelj tumora. Centralna masa tumora nije lako uočljiva jer nema jasno definisane ivice. Njihove veličine variraju od nekoliko milimetara do nekoliko centimetara [1]. Iz navedenih razloga tumor nije moguće detektovati na osnovu oblika ili veličine. U ovom radu predložena je metoda detekcije tumora dojke poređenjem parametara ranije pouzdano detektovanog tumora sa parametrima digitalne slike dojke za koju se pouzdano ne zna da li sadrži malignitet. Međutim, uslov koji mora da bude ispunjen da bi rezultati obrade bili korektni jeste upotreba filtriranih slika. U poslednje vreme su postignuti veoma dobri rezultati u eliminaciji šuma sa slike primenom multirezolucionog algoritma odnosno diskretne transformacije talasićima. Pokazano je da se na bazi pogodno izabranog praga može doći do reprezentacije signala koja ima željeni stepen glatкости [3,4].

Često je slika skenirana na takav način da se vrednosti svetlosti ne protežu preko celog dinamičkog opsega. U ovakvim slučajevima slika se koriguje pojačavanjem kontrasta. Pošto se radi o monohromatskoj slici, pojačavanjem kontrasta jasnije će se videti zvezdaste konture na mamogramu. Na slikama 1. i 2. se vidi da se zvezdaste konture jasnije uočavaju na mamogramu sa pojačanim kontrastom.



Slika 1. Originalni mamogram



Slika 2. Mamogram sa povećanim kontrastom

2. MULTIREZOLUCIONA ANALIZA

Kontinualna transformacija talasićima (*Continuous Wavelet Transform – CWT*) je algoritam razvijen kao alternativni pristup, brzom Furijeovoj transformaciji, da bi se prevazišao problem rezolucije. Postoji jedna osnovna razlika između ova dva načina transformacije, a to je da se širina prozorske funkcije menja tokom transformacije signala i to predstavlja najznačajniju karakteristiku transformacije talasićima. Kontinualna transformacija talasićima je definisana na sledeći način:

$$CWT_X^\psi(\tau, s) = \frac{1}{\sqrt{|s|}} \int x(t) \psi\left(\frac{t-\tau}{s}\right) dt \quad (1)$$

Kao što se vidi iz jednačine (1), transformisan signal je funkcija dve promenljive, τ i s , koje predstavljaju translaciju i dilataciju, respektivno, dok je $\psi(t)$ funkcija transformacije.

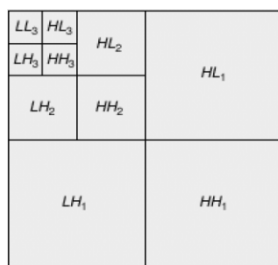
Diskretna transformacija talasićima analizira signal na različitim frekventnim opsezima, različitim rezolucijama, dekomponovanjem signala u grube aproksimacije i detaljne informacije. Originalni signal $x[n]$ je prvo prošao kroz visokofrekventni pojasni filter $g[n]$ i niskofrekventni pojasni filter $h[n]$. Nakon filtriranja polovina uzoraka je eliminisano. Signal je sada pod uzorkovan za faktor 2 jednostavnim odbacivanjem svakog drugog uzorka. Ovo predstavlja jedan nivo dekompozicije i može se matematički izraziti jednačinama:

$$y_{high}[k] = \sum_n x[n] \cdot g[2k - n] \quad (2)$$

$$y_{low}[k] = \sum_n x[n] \cdot h[2k - n]$$

gde $y_{high}[k]$ i $y_{low}[k]$ predstavljaju izlaze visokofrekventnog i niskofrekventnog pojasnog filtra, respektivno, posle pod uzorkovanja za faktor 2. Opisana dekompozicija signala duplo smanjuje vremensku rezoluciju jer polovina uzoraka karakteriše ceo signal. Međutim, istovremeno duplo povećava frekventnu rezoluciju jer frekventni opseg signala obuhvata samo polovinu prethodnog frekventnog opsega čime se neizvesnost u pogledu određivanja tačne frekvencije smanjuje za pola. Ovaj postupak koji je poznat kao pojasno kodiranje može se nastaviti za dalje dekomponovanje signala.

Kada se radi o dvodimenzionalnim signalima kao što je slika, multirezoluciona piramida formira se na način da se na svakom višem nivou čuvaju slika dva puta manje rezolucije nego što je ona na prethodnom nivou i slika detalja neophodnih za rekonstrukciju signala. Multirezoluciona reprezentacija slike se u svakom nivou dekompozicije sastoji od jedne diskretne slike aproksimacije na nižoj rezoluciji i tri slike detalja. Aproksimaciji odgovara deo spektra koji se dobije NF filtriranjem po oba pravca u frekventijskoj ravni. Jedna od slika detalja se dobije horizontalnim NF i vertikalnim VF filtriranjem, druga vertikalnim NF i horizontalnim VF filtriranjem, dok se dijagonalna slika detalja dobije VF filtriranjem po oba pravca. Višestrukim ponavljanjem dolazi se do slika sa sve lošijom rezolucijom, što odgovara piramidalnoj dekompoziciji, slika 3.



Slika 3. Piramidalna dekompozicija slike

2.1. FILTRIRANJE SLIKE PRIMENOM TRANSFORMACIJE TALASIĆIMA

Veoma su raznolike postojeće metodologije uklanjanja šuma. Izbor metoda adekvatnog datoj primeni zavisi kako od konteksta tako i od cilja obrade. Najveći problem koji se javlja prilikom uklanjanja šuma sa slike primenom nekog od digitalnih filtera jeste pojava zamućene filtrirane slike.

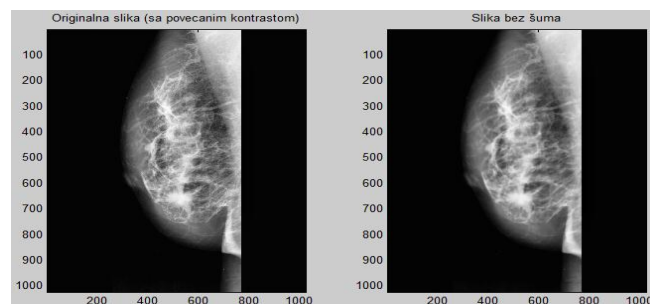
Praktično, filtriranjem signala može se eliminisati samo šum koji se nalazi van frekventijskog opsega tog signala. Dekompozicija signala talasićima pored toga što ukazuje na frekventijski sadržaj signala, ukazuje i na položaj komponenti u vremenu i prostoru. Filtriranjem slike primenom transformacije talasićima prevazilazi se problem pojave zamućenih ivica prilikom uklanjanja šuma [5].

Najčešće se primenjuje metod odsecanja talasića (*wavelet thresholding*) gde se eliminacija šuma vrši prema sledećoj proceduri: 1. Primeni se diskretna wavelet transformacija (DWT) na n odbiraka signala (kod kojih je prisutan šum), čime se dobija n wavelet koeficijenata; 2. Formira se poredak wavelet koeficijenata (najčešće prema njihovoj veličini); 3. Primenom praga bira se m "najznačajnijih" koeficijenata prema utvrđenom poretku iz 2; 4. Estimirani signal (sa uklonjenim šumom) dobija se primenom inverzne wavelet transformacije na izabranim m koeficijenata.

Uklanjanje šuma sa mamograma izvršeno je primenom dvo-nivovske transformacije talasićima, pomoću Haar-ovog talasića i vrednosti praga 44.5. Prag se obično bira na osnovu statističkih pretpostavki o šumu i o samom signalu koji se estimira. Skalirajuća funkcija $\phi(t)$ i funkcija majke talasića $\psi(t)$ Haar-ovog talasića definisane su na sledeći način:

$$\psi(t) = \begin{cases} 1 & 0 \leq t < 1/2 \\ -1 & 1/2 \leq t < 1 \\ 0 & \text{otherwise.} \end{cases} \quad \phi(t) = \begin{cases} 1 & 0 \leq t < 1 \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

Na slici 4. prikazane su originalni mamogram i estimirani mamogram.



Slika 4. Originalni mamogram i estimirani mamogram

2.2. KOMPRESOVANJE SLIKE PRIMENOM TRANSFORMACIJE TALASIĆIMA

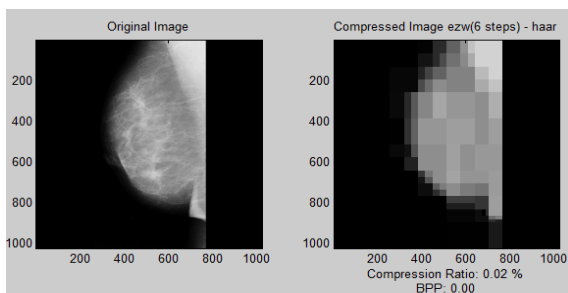
Kompresovanje slike ima za cilj minimizovanje broja bita potrebnog za prikazivanje slike pri čemu će kompresovana slika imati prihvatljiv kvalitet. Talasići predstavljaju optimalno rešenje za ovaj problem. Proces kompresije obuhvata faze kvantizacije, kodovanja i dekodovanja, u kojima se primenjuje obrada talasićima [6].

Mera ostvarene kompresije predstavlja se odnosom kompresije (Compression Ratio- CR) i odnosom bit-piksela (Bit-Per-Pixel- BPP). CR označava da je kompresovana slika sačuvana korišćenjem CR procenata polazne veličine dok BPP je broj bita korišćen za čuvanje jednog piksela slike. Upotrebom metoda za kompresiju pronalazi se kompromis između niskog odnosa kompresije i dobrog kvaliteta slike.

Jedan od jednostavnijih i najčešće korišćenih algoritama za kompresovanje slike je EZW algoritam. Tehnike za kompresovanje slike, posebno nereverzibilne tehnike, su poznate kao tehnike kod kojih raste kompleksnost izračunavanja sa porastom njihove efikasnosti. EZW tehnika prekida međusobnu zavisnost rasta efikasnosti [7]. EZW algoritam se zasniva na četiri ključna koncepta: 1. Diskretna transformacija talasićima; 2. Predviđanje nedostatka važnih informacija korišćenjem bitnih sličnosti kod slika; 3. Entropijski kodovana sukcesivno-aproksimativna kvantizacija; 4. Univerzalno kompresovanje bez gubitka podataka koje se postiže pomoću adaptivnog aritmetičkog kodovanja [8].

Koeficijent talasića x se smatra beznačajnim u odnosu na dati prag T ako je $|x| \leq T$. U suprotnom, koeficijent se zove značajnim. Koeficijent na lošijoj rezoluciji se zove roditelj, a svi koeficijenti koji odgovaraju istoj prostornoj lokaciji na sledećem, finijem nivou rezolucije se nazivaju deca. Koeficijentu x pridružujemo oznaku POS (positive) ako je značajan u odnosu na prag T , a ako je beznačajan pridružujemo mu oznaku NEG (negative). Koeficijentu x pridružujemo oznaku ZTR (zerotree root) ako su on i svi njegovi potomci beznačajni u odnosu na prag T . Koeficijentu x pridružujemo oznaku IZ (isolated zero) ako je on sam beznačajan ali ima značajne potomke. Tokom kodovanja i dekodovanja formiraju se dve liste koeficijenata talasića: dominantna i podređena lista. U bilo kom trenutku u procesu, dominantna lista sadrži koordinate onih koeficijenata koji nisu još uvek označeni kao značajni. Podređena lista sadrži vrednosti onih koeficijenata koji označeni kao značajni. Za svaki prag, svaka lista se skenira jednom. Nakon svakog dominantnog prolaza opseg [prag T ; vrednost najvećeg koeficijenta] se podeli na dva dela, a zatim se koeficijenti koji se nalaze na podređenoj listi koduju sa 1 ako njihova vrednost pripada višem opsegu ili sa 0 ako njihova vrednost pripada nižem opsegu. Postupak se ponavlja u sledećem dominantnom prolazu sa novom vrednošću praga. Koeficijenti prethodno identifikovani kao značajni koeficijenti, sada se smatraju nulama. Obrada se naizmenično nastavlja između dominantnih i sporednih prolaza i može biti prekinuta u bilo kom trenutku [8].

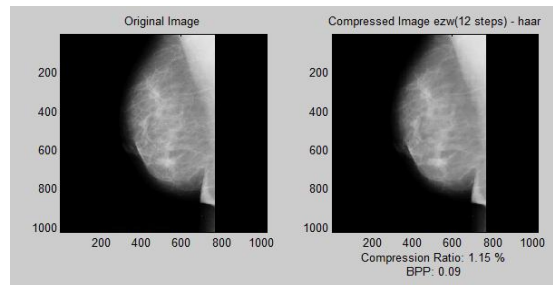
Kod kompresovanja slike pomoću EZW algoritma ključni parametar je broj petlji koji rastući daje kvalitetniju sliku i lošiji odnos kompresije. Na slici 5. mogu se uporediti originalni mamogram i mamogram kompresovan primenom EZW algoritma, 6 petlji i Haar-ovog talasića.



Slika 5. Originalni mamogram i mamogram kompresovan pomoću EZW algoritma (6 petlji) i Haar-ovog talasića

Odnos kompresije i odnos bit-piksel su veoma mali, što je dobro, ali kvalitet slike je veoma loš. Iz tog razloga potrebno je povećati broj petlji. Ako se umesto 6 petlji upotrebi 9 petlji

CR i BPP se malo povećaju u korist kvaliteta kompresovane slike ali kvalitet kompresovane slike i dalje je nezadovoljavajuć. Zadovoljavajući odnos kompresije i odnos bit-piksel i zadovoljavajući kvalitet slike dobije se upotrebom EZW algoritma sa 12 petlji i Haar-ovog talasića (slika 6.).



Slika 6. Originalni mamogram i mamogram kompresovan pomoću EZW algoritma (12 petlji) i Haar-ovog talasića

3. AUTOMATSKO PREPOZNAVANJE TUMORA

Osnovna ideja je da se tumor dojkve zaokruži na osnovu parametara mamograma za koji se pouzdano zna da sadrži tumor. Potrebni podaci o mamogramu koji sadrži tumor, a na osnovu kojeg se vrši detektovanje tumora su koordinate centralnog piksela kao i širina i dužina (izražene preko broja piksela) dela slike koji predstavlja tumor. Sa mamograma na osnovu kojeg vršimo detekciju izdvojili smo tumor koji ima sledeće parametre: vrednost inteziteta piksela je 166.7601 i standardna devijacija je 32.9223. Na osnovu ovih vrednosti tražimo delove mamograma na kojem treba zaokružiti tumor sa približnim vrednostima istih parametara.

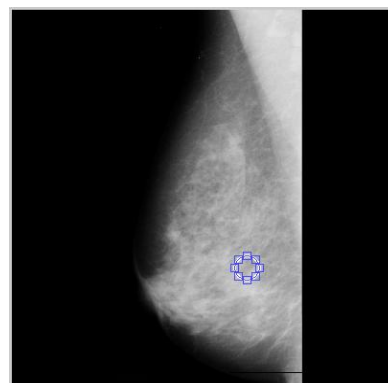
Neka je (i,j) prostorna lokacija na mamogramu u i -toj vrsti i j -oj koloni, a u_{ij} vrednost piksela na lokaciji (i,j) . Za svaku $M \times N$ sliku standardna devijacija je kvadratni koren varijanse i predstavljena je sledećom jednačinom:

$$\sigma = \sqrt{\frac{\sum_{i=1}^M \sum_{j=1}^N (u_{ij} - \mu)^2}{M \times N - 1}} \quad (4)$$

gde μ predstavlja srednju vrednost intenziteta piksela matrice u i predstavljena je sledećom jednačinom:

$$\mu = \frac{\sum_{i=1}^M \sum_{j=1}^N u_{ij}}{M \times N} \quad (5)$$

Nakon uklanjanja šuma sa oba slike, vrši se izdvajanje tumora sa već obrađene slike na osnovu poznatih podataka (koordinata centralnog piksela i veličine tumora) da bi se izračunale vrednosti parametara koje karakterišu tumor. Na slici 7. je prikazana obrađena slika sa zaokruženim tumorom.



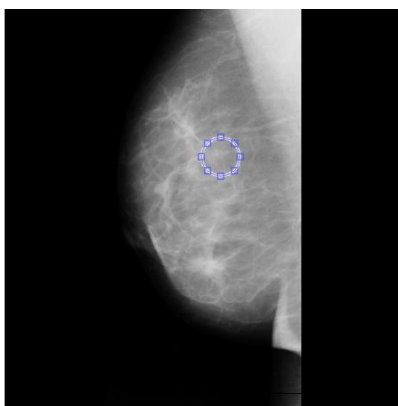
Slika 7. Mamogram sa zaokruženim tumorom na osnovu kojeg se vrši detekcija drugih tumora

Na slici 8. prikazan je tumor koji je zaokružen na slici 7.



Slika 8. Deo već obrađenog mamograma koji predstavlja tumor

Nakon izdvajanja tumora izračunavaju se srednja vrednost inteziteta piksela i standardna devijacija na delu mamograma koji predstavlja tumor. Zadaju se širina i visina dela slike čiji se parametri proveravaju, ovde je to 100 piksela za oboje i dozvoljena odstupanja za parametre slike koji se ispituju. Potom se na mamogramu na kojem treba zaokružiti tumor traži deo veličine 100x100 čija srednja vrednost inteziteta piksela i standardna devijacija su približno jednaki srednjoj vrednosti inteziteta piksela i standardnoj devijaciji tumora kod već obrađenog mamograma, uzimajući u obzir dozvoljena odstupanja. Na slici 9. prikazan je rezultat pretrage tj. originalni mamogram, bez pojačanog kontrasta, na kojem je zaokružen tumor.



Slika 9. Mamogram sa zaokruženim tumorom

Na slici 10. prikazan je deo mamograma, sa pojačanim kontrastom, koji predstavlja tumor detektovan na slici 9.



Slika 10. Tumor detektovan na slici 9.

Rezultati detektovanja tumora na mamogramu, pre i posle kompresovanja mamograma su veoma slični, skoro isti. Razlog je činjenica, da su vrednosti standardne devijacije i srednje vrednosti inteziteta piksela na delu slike koji predstavlja tumor, pre kompresovanja i posle kompresovanja, skoro iste. Kao što smo već rekli, deo mamograma koji predstavlja tumor pre kompresovanja imao je sledeće parametre: srednja vrednost inteziteta piksela je 166.7601 i standardna devijacija je 32.9223. Posle kompresovanja, srednja vrednost inteziteta piksela i standardna devijacija dela mamograma koji predstavlja tumor su 166.6986 i 32.1573, respektivno.

4. ZAKLJUČAK

Pomoću dvodimenzionalne analize talasićima može se efikasno ukloniti šum sa slike i može se kompresovati slika bez žrtvovanja kvaliteta slike. Problem razvijanja algoritama

za detektovanje tumora je nemogućnost kompetentnog testiranja zbog nepostojanja kvalitetne baze digitalnih slika. U ovom radu predstavili smo metodu automatskog detektovanja tumora dojke, poređenjem standardne devijacije i srednje vrednosti inteziteta piksela dva mamograma od kojih se za jedan pouzdano zna da sadrži tumor, a za drugi to nije poznato. Početni rezultati do kojih smo došli pokazuju da je moguće kompresovati mamogram primenom multirezolucione analize i EZW algoritma, a da se rezultat detektovanja tumora ne promeni u odnosu na detektovanje tumora na nekompresovanom mamogramu. U daljem radu nastojaćemo da unapredimo tehniku automatskog detektovanja tumora dodatnim poređenjem vrednosti standardne devijacije histograma i standardne devijacije ugaone orijentacije dva mamograma.

5. ZAHVALNOST Ovaj rad je podržan projektom Ministarstva nauke i tehnološkog razvoja Republike Srbije, broj III41007.

LITERATURA

- [1] S. Liu, C.F. Babbs and E.J. Delp, "Multiresolution detection of speculated lesions in digital mammograms", *IEEE Transactions Image Processing*, vol. 10, pp. 874-884, Jun 2001.
- [2] <http://www.stetoskop.info/Rak-dojke-nocna-mora-zena-1364-c3-content.htm>
- [3] D.L. Donoho and I.M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage", *Biometrika*, vol. 81, no. 3, pp. 425-455, 1994.
- [4] D.L. Donoho, "Denoising by soft thresholding", *IEEE Transactions on Information Theory*, vol. 41, no 3, 613-627, 1995.
- [5] M. Afonso and S. D. Joshi, "A wavelet based scheme for adaptive noise canceling from images", *Proceedings of the 28th General Assembly of the International Union of Radio Science*, 2005.
- [6] M.V. Popović, *Digitalna obrada slike*, Akademska misao, Beograd, 2006.
- [7] A. Said, and W.A. Pearlman, "A New Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, June 1996.
- [8] J.M. Shapiro, "Embedded Image Coding Using Zerotrees of Wavelet Coefficients", *IEEE Transactions on Signal Processing*, vol. 41, no 12, December 1991.

Abstract –In this paper, we present a multiresolution scheme to detect stellate lesions in mammograms. First we removed the noise from mammograms using Multiresolution analysis and then we detected tumors. Then, using the Embedded Zerotree Wavelet (EZW) algorithm, we compressed denoising mammogram and showed that by applying the algorithm to detect tumors in the compressed image we obtained the same results as in the case of non-compressed images. Experimental results obtained from the mammograms of patients recorded in the Clinical Center in Kragujevac, show that using multiresolution algorithm can be detected tumors of different sizes.

AUTOMATIC IDENTIFICATION OF TUMORS USING MULTIREOLUTION ALGORITHM

Marina Đoković, Aleksandar Peulić,
Nenad Filipović, Đorđe Damjanović