

Jakub Smółka*, Maria Skublewska-Paszkowska*

Usuwanie nadmiernej segmentacji w transformacji wododziałowej za pomocą analizy skupień

1. Wprowadzenie

Transformacja wododziałowa jest metodą segmentacji obrazów, która symuluje zalewanie wodą rzeźby terenu utworzonej na podstawie obrazu [7]. Obraz dzielony jest na zlewiska, z których każde jest przypisane do oddzielnego minimum lokalnego. Niestety oznacza to, że liczba zlewisk jest równa liczbie minimów lokalnych w obrazie poddawany transformacji. W praktyce okazuje się, że liczba ta jest znacząco większa niż liczba obiektów przedstawionych na obrazie i nawet dla niezbyt dużego obrazu może wynosić kilkadziesiąt tysięcy. Od pewnego czasu toczą się prace nad rozwiązaniem tego problemu.

Jednym z pierwszych rozwiązań problemu nadmiernej segmentacji jest użycie znaczników [2]. Znaczniki zastępują minima lokalne obrazu w tym sensie, że każde zlewisko tworzone przez transformację wododziałową powiązane jest z innym znacznikiem. Ogranicza to liczbę zlewisk do liczby znaczników, jednak wymaga opracowania sposobu na określanie znaczników.

Innym sposobem na zmniejszenie liczby zlewisk jest wykonanie wielu transformacji wododziałowych [1]. Pierwsza transformacja wykonywana jest na gradientowym obrazie oryginału. Dla każdego z uzyskanych zlewisk obliczana jest średnia wartość punktów (z obrazu oryginalnego) w tym zlewisku. Następnie dla tak uśrednionego obrazu obliczany jest kolejny obraz gradientowy i transformacja wododziałowa wykonywana jest ponownie. Ta sekwencja powtarzana jest do momentu uzyskania satysfakcjonującego wyniku.

Kolejnym sposobem na wyeliminowanie nadsegmentacji jest potraktowanie zbioru zlewisk jako zbioru węzłów grafu nieskierowanego. Krawędzie w tym grafie łączą sąsiadujące ze sobą zlewiska. Po utworzeniu grafu z wybranego węzła uruchamiany jest algorytm rozrostu regionu dzielący obraz na dwie klasy: obiekt i tło [15].

* Politechnika Lubelska w Lublinie

W niniejszym artykule przedstawiono propozycję rozwiązania problemu nadmiernej segmentacji za pomocą analizy skupień. W opisywanej metodzie każde zlewisko traktowane jest jako obiekt opisany wektorem atrybutów ilościowych. Pozwala to, przy wykorzystaniu hierarchicznych metod analizy skupień, na utworzenie hierarchii podobieństwa zlewisk. Podejście to umożliwia bardzo łatwe zmienianie liczby klas uzyskanych w końcowej segmentacji i może być stosowane do różnych typów obrazów, ponieważ stosowane atrybuty mogą być dopasowane do rodzaju segmentowanych obrazów.

2. Transformacja wododziałowa

2.1. Działanie

Transformacja wododziałowa nazywana również segmentacją wododziałową została zaproponowana przez Beuchera i Lantuejoulé'a w 1979 roku [2]. Zasada jej działania jest bardzo prosta i może być wyjaśniona poprzez analogię do deszczu padającego na nierówno ukształtowany teren [7]. Obraz, którego punkty przedstawione są w skali szarości, można zinterpretować jako odwzorowanie rzeźby terenu, gdzie jasne punkty odpowiadają wzgórzom, a ciemne punkty dolinom. Padający deszcz napędzany grawitacją spływa po zboczach pagórków i zbiera się w dolinach. Wszystkie punkty, z których woda zbiera się w okolicy tego samego minimum, nazywane są zlewiskiem. Obszary powstających zbiorników wodnych powiększają się wraz z ilością otrzymanego opadu. W pewnym momencie stają się tak duże, że woda z jednego zbiornika mogłaby się przelać do drugiego. Metoda transformacji wododziałowej nie pozwala na łączenie się takich obszarów poprzez działanie, które można obrazowo nazwać budowaniem zapór. Zapory te można przedstawić na obrazie jako linie działów wodnych rozdzielające poszczególne zlewiska.

Bardziej formalnie wynikiem działania transformacji wododziałowej jest obraz podzielony na zlewiska o ciągłych krawędziach. Ciągłość krawędzi jest istotną zaletą tej metody. Dzięki temu nie jest konieczne stosowanie metod łączących poprzerywane fragmenty granic – tak jak np. w przypadku segmentacji uzyskanej poprzez progowanie obrazu przedstawiającego moduł gradientu oryginalnego obrazu. Wadą opisywanej transformacji jest to, że tworzy ona zlewisko dla każdego minimum lokalnego obrazu, co powoduje zjawisko nadmiernej segmentacji. Obrazem wejściowym dla transformacji wododziałowej zazwyczaj nie jest oryginalny obraz, lecz obraz przedstawiający moduł jego gradientu. Wynika to z faktu, że do poprawnej pracy transformacja wymaga na wejściu obrazu, w którym duże wartości oznaczają występowanie krawędzi, a niskie dolin [7]. Wymagania te spełnia między innymi moduł gradientu obrazu. Rysunek 1 symbolicznie przedstawia proces przygotowania obrazu. Jasny obiekt poddawany jest działaniu filtru gradientowego, a następnie jego wynik transformacji wododziałowej.



Rys. 1. Transformacja wododziałowa obrazu (etapy: poddawanie obrazu działaniu filtra gradientowego, transformacja wododziałowa wykonywana na obrazie gradientowym)

Stosowanie filtra gradientowego niesie jednak ze sobą pewne komplikacje. Pierwszym problemem jest jego wrażliwość na występowanie szumu. Szum to zwiększenie liczby minimów lokalnych w obrazie gradientowym, co pociąga za sobą wzrost liczby zlewisk. Dla typowego niewielkiego obrazu o rozmiarach 256×256 punktów liczba minimów lokalnych może wynosić np. 10 000. Inny problem pojawia się przy próbie zastosowania tego filtra do obrazów kolorowych. Wynika to z faktu, że punkty w takich obrazach opisane są za pomocą wektorów (złożonych zazwyczaj z trzech wartości), natomiast transformacja wododziałowa wymaga obrazu wejściowego, którego punkty są skalarami. Istnieją odpowiednie filtry gradientowe, omówione w dalszej części artykułu, umożliwiające wykonanie takiego przekształcenia.

2.2. Formalny opis transformacji wododziałowej

Istnieje wiele formalnych definicji transformacji wododziałowej oraz wiele różnych algorytmów, które ją realizują. Część z nich jest dokładnie omówiona w pracy [11]. W prezentowanym opracowaniu dla przeprowadzenia transformacji wododziałowej wykorzystano algorytm, który każdemu punktowi obrazu przydziela etykietę (wynik nie zawiera punktów tworzących linie działów wodnych).

W pierwszej fazie działania algorytm ten przydziela wszystkim minimom lokalnym m_k obrazu Φ unikalne etykiety L spełniające warunek:

$$\forall_{i \neq j} L(m_i) \neq L(m_j) \quad (1)$$

Minima mogą być pojedynczymi punktami lub grupami sąsiadujących ze sobą punktów (równinami). Po zakończeniu tego kroku algorytm wszystkim punktom spełniającym warunek:

$$\Gamma(p) \neq \emptyset \quad (2)$$

przypisuje etykietę:

$$L(p) = L(p') \quad (3)$$

przy czym:

$$p' \in \Gamma(p) \quad (4)$$

a $\Gamma(p)$ jest zbiorem najniżej położonych sąsiadów punktu p :

$$p' \in \Gamma(p) \Leftrightarrow \Phi(p') = \min_{r \in N(p)} \Phi(r) \wedge \Phi(p') < \Phi(p) \quad (5)$$

gdzie $N(p)$ – zbiór punktów sąsiadujących z p .

Dzięki takiemu postępowaniu etykieta punktu zależy od tego, do którego minimum lokalnego prowadzi z tego punktu ścieżka najszybszego spadku. Ostatnim etapem działania algorytmu jest nadanie etykiet płaskim obszarom, które nie są minimami lokalnymi. Algorytm przypisuje im etykietę ich najniższego sąsiada:

$$\forall_{q \in P} L(q) = L(r) \quad (6)$$

gdzie:

- P – płaski obszar,
- r – najniższy sąsiad obszaru P .

$$\Phi(r) = \min_{s \in N(P)} \Phi(s) \quad (7)$$

$$N(P) = \{s : s \notin P \wedge \exists_{q \in P} s \in N(q)\} \quad (8)$$

2.3. Moduł gradientu w obrazach czarno-białych i kolorowych

Jak wspomniano powyżej, transformacja wododziałowa wymaga na swoim wejściu obrazu, w którym wysokie wartości punktów odpowiadają krawędziom. Obraz przedstawiający rzeczywiste obiekty można przekształcić do tej postaci za pomocą filtru obliczającego moduł gradientu we wszystkich jego punktach [5, 13].

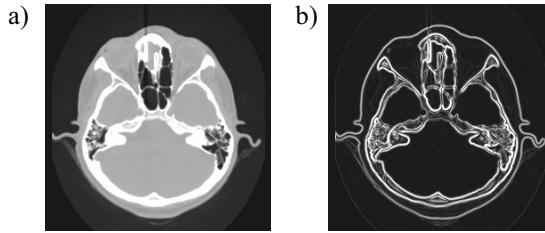
Przypadek obrazu w skali szarości jest dość prosty. Jeżeli przyjąć, że obraz to funkcja postaci:

$$\Phi(x_1, x_2) : \mathfrak{R}^2 \rightarrow \mathfrak{R} \quad (9)$$

wtedy moduł gradientu wynosi:

$$M = \sqrt{\left(\frac{\partial \Phi}{\partial x_1}\right)^2 + \left(\frac{\partial \Phi}{\partial x_2}\right)^2} \quad (10)$$

Rysunek 2 przedstawia efekt działania filtru gradientowego na przykładzie obrazu z tomografu komputerowego.



Rys. 2. Obraz gradientowy obiektu czarno-białego:
a) oryginalny obraz CT (256×256, 12 bity) [6]; b) obraz gradientowy

Obliczanie obrazu gradientowego dla zdjęć kolorowych jest bardziej kłopotliwe. Wynika to z faktu, że ich punkty opisane są za pomocą wektorów. Ponadto, w odróżnieniu od obrazów w skali szarości, nie istnieje jeden najpopularniejszy sposób obliczania obrazu gradientowego.

Kolorowy obraz w przestrzeni barw sRGB może być opisany za pomocą funkcji:

$$\Phi(x_1, x_2): \mathfrak{R}^2 \rightarrow \mathfrak{R}^3 \quad (11)$$

a jego poszczególne składowe:

$$\Phi_j(x_1, x_2): \mathfrak{R}^2 \rightarrow \mathfrak{R}; j = 1..3 \quad (12)$$

W przedstawionych badaniach zastosowano dwa rodzaje filtrów gradientowych przeznaczonych do użycia z obrazami kolorowymi. Pierwszym z nich jest filtr oparty na heurystyce [7], która definiuje moduł gradientu jako sumę kwadratów wszystkich pochodnych cząstkowych wszystkich składowych. Dla obrazu dwuwymiarowego o trzech składowych obliczany jest z wykorzystaniem równania o następującej postaci:

$$M = \sqrt{\sum_{i=0}^2 \sum_{j=0}^3 \left(\frac{\partial \Phi_j}{\partial x_i} \right)^2} \quad (13)$$

W dalszej części artykułu ta metoda obliczania gradientu będzie określana skrótowo *gradient-sk*.

Drugi, z zastosowanych filtrów, oparty jest na analizie głównych składowych. Został zaproponowany przez Sapiro i Ringacha [14]. W tej metodzie pochodne kierunkowe wszystkich składowych obrazu poddawane są analizie głównych składowych. Moduł gradientu obliczany jest jako różnica dwóch największych wartości własnych macierzy zawierającej iloczyn skalarne g_{ij} . Dla dwuwymiarowego obrazu posiadającego trzy kolorowe składowe, jak np. obraz sRGB, obliczany jest za pomocą wzoru [9]:

$$M = \lambda_+ - \lambda_- \quad (14)$$

w którym wartości własne wyrażone są za pomocą wzoru [14]:

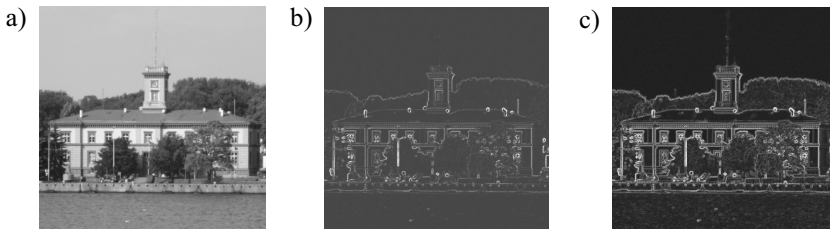
$$\lambda_{\pm} = \frac{g_{11} + g_{22} \pm \sqrt{g_{11} - g_{22}}^2 + 4g_{12}^2}{2} \quad (15)$$

gdzie:

$$g_{ij} = \frac{\partial \Phi}{\partial x_i} \cdot \frac{\partial \Phi}{\partial x_j} \quad (16)$$

Wartości własne λ_+ i λ_- mogą być interpretowane, odpowiednio, jako największa i najmniejsza lokalna szybkość zmian obrazu [14]. Ta metoda w dalszej części artykułu będzie nazywana *gradient-pca*.

Najbardziej widoczną różnicą między opisanymi dwoma sposobami obliczania obrazów gradientowych jest to, że *gradient-pca* wykrywa najbardziej znaczące krawędzie, a ogranicza pojawianie się tych mniej istotnych. *Gradient-sk*, natomiast, jest czuły nawet na bardzo niewielkie zmiany w obrazie, dzięki czemu jest bardziej użyteczny, gdy segmentacja ma zawierać mniej znaczące elementy obrazu. Zazwyczaj *gradient-sk* powoduje większą nadsegmentację. Różnice między filtrami ilustruje rysunek 3. Użycie dwóch różnych metod obliczania obrazu gradientowego ułatwiało dobranie parametrów wstępnego usuwania nadmiernej segmentacji opisanego poniżej.



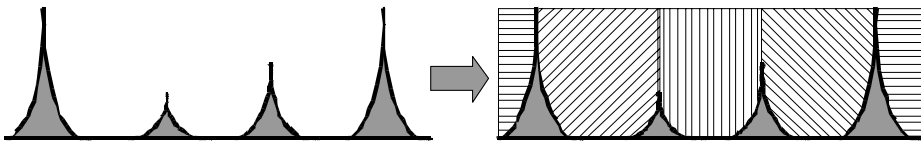
Rys. 3. Rodzaje zastosowanych filtrów gradientowych dla obrazów kolorowych:
a) oryginał (384×384, 24 bity); b) gradient oparty na analizie głównych składowych;
c) gradient oparty na heurystyce

Angulo i Serra [1] opisali inne metody pozwalające na obliczanie gradientu obrazu kolorowego. Do metod tych należą: gradient jasności, kołowy gradient odcienia, ważony nasyceniem gradient koloru, maksymalny gradient koloru, oraz gradienty chromatyczny i percepcyjny.

2.4. Problem nadmiernej segmentacji

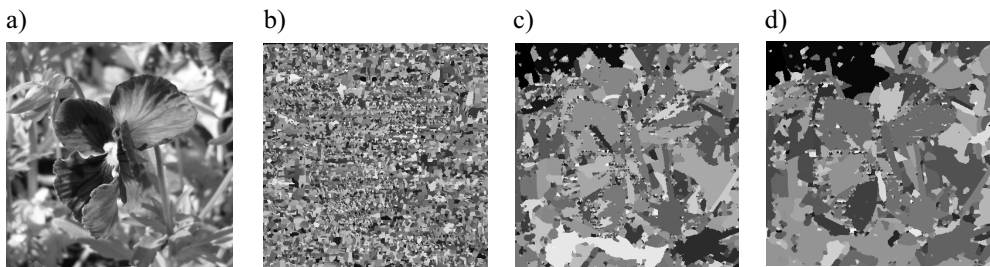
Problem nadmiernej segmentacji wynika z tego, że transformacja wododziałowa tworzy osobne zlewisko dla każdego minimum lokalnego [7], bez względu na to, czy jest ono

otoczone wyraźnymi krawędziami, czy też jest ono wynikiem zaszumienia obrazu. Sytuację pogarsza dodatkowo czułość filtru gradientowego na niewielkie zmiany w obrazie. Rysunek 4 przedstawia sytuację, w której dochodzi do nadmiernej segmentacji. Nawet małe wartości w obrazie gradientowym powodują, że transformacja wododziałowa buduje zapory między zlewiskami i nie pozwala na ich połączenie.



Rys. 4. Ilustracja przyczyn nadmiernej segmentacji – wiele minimów lokalnych

Rysunek 5a przedstawia zdjęcie, które zostało poddane działaniu transformacji wododziałowej. Obraz gradientowy potrzebny do jej działania został obliczony za pomocą *gradientu-sk*. Zdjęcie zostało podzielone na 10 084 zlewiska (rys. 5b).



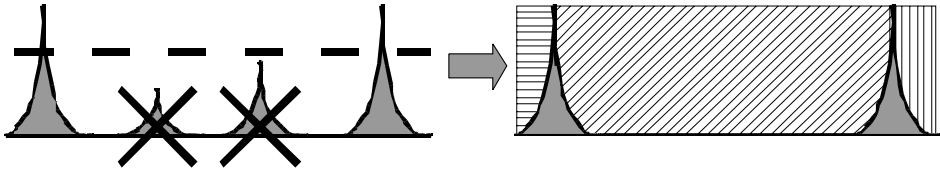
Rys. 5. Wyniki wstępnego usuwania nadmiernej segmentacji: a) oryginalny obraz sRGB; b) próg 0% oraz poziom łączenia 0% (10 084 zlewiska); c) próg 5%, poziom łączenia 0% (3489 zlewisk); d) próg 5%, poziom łączenia 10% (1965 zlewisk)

2.5. Wstępne usuwanie nadmiernej segmentacji

Jak widać na rysunku 5b, transformacja wododziałowa nie daje satysfakcjonujących rezultatów, jeżeli segmentowany obraz nie będzie właściwie przygotowany, a wyniki odpowiednio przetworzone. Opisane metody zmniejszania nadmiernej segmentacji nadają się do użycia zarówno z obrazami w odcieniach szarości, jak i kolorowych.

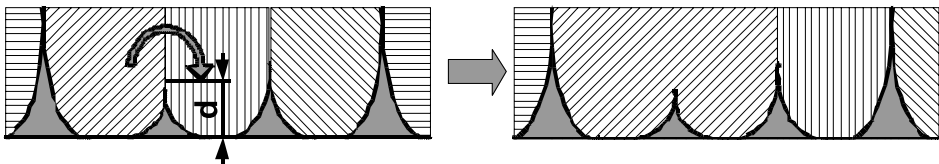
Wszystkie obrazy prezentowane w tym artykule były, przed obliczeniem obrazu gradientowego, poddawane działaniu selektywnego filtru rozmywającego [10]. Filtr taki niweluje niewielkie zmiany wartości w obrazie (zazwyczaj jest to szum lub mało istotne szczegóły), a zachowuje większe zmiany, które odpowiadają zazwyczaj wyraźnym krawędziom.

Mimo zastosowania filtru rozmywającego, obraz gradientowy zawiera wiele minimów lokalnych. Pierwszym sposobem na zmniejszenie tej liczby jest progowanie [7]. Jest to bardzo prosta metoda przedstawiona na rysunku 6. Polega ona na usunięciu z obrazu gradientowego, przed zastosowaniem transformacji wododziałowej, wartości mniejszych niż określony próg. W niniejszym artykule wartość progu wyrażona jest jako procentowa wartość maksymalnej głębokości zlewiska.



Rys. 6. Zmniejszanie nadmiernej segmentacji poprzez progowanie obrazu gradientowego

Drugim sposobem redukowania nadmiernej segmentacji jest łączenie sąsiednich zlewisk na podstawie ich głębokości [7]. Zlewisko może być połączone ze swoim sąsiadem, jeżeli głębokość sąsiada nie przekracza zadanego progu. Proóg ten nazywany jest poziomem łączenia. Tak jak w przypadku progowania, jego wartość jest wyrażona jako procent maksymalnej głębokości zlewisk. Łączenie zlewisk odbywa się po zastosowaniu transformacji wododziałowej do obrazu gradientowego. Ilustrację tego sposobu zmniejszania nadmiernej segmentacji przedstawia rysunek 7.



Rys. 7. Zmniejszanie nadmiernej segmentacji poprzez łączenie sąsiednich zlewisk na podstawie ich głębokości

Rysunek 5 przedstawia efekty stosowania obydwu wspomnianych metod na rzeczywistym obrazie. Niewielkie zmiany wartości w obrazie oryginalnym (rys. 5a) zostały zredukowane za pomocą selektywnego filtru rozmywającego. Do obliczenia obrazu gradientowego zastosowano *gradient-sk*. Następnie przeprowadzono transformację wododziałową. W wyniku otrzymano obraz zawierający 10 084 zlewiska przedstawiony na rysunku 5b. Gdy przed wyznaczeniem transformacji wododziałowej zastosowano progowanie obrazu gradientowego z progiem 5%, otrzymano wynik przedstawiony na rysunku 5c. Liczba zlewisk wyniosła 3489, co oznacza, w przybliżeniu 2,9-krotny spadek. Następnie przeprowadzono łączenie zlewisk z poziomem łączenia równym 10%. Liczba zlewisk wyniosła 1965, czyli nastąpił 5,1-krotny spadek w stosunku do pierwotnej segmentacji. Zazwyczaj obydwie metody stosowane są równocześnie. Należy jednak zaznaczyć, że w większości przypadków nie jest to rozwiązanie wystarczające. Natomiast zbyt wysoko ustawione progi

powodują niedostateczną segmentację – różne obiekty są łączone. Wynika to z faktu, że opisane metody biorą pod uwagę tylko informację z obrazu gradientowego. Informacje o odcieniu lub kolorze są ignorowane.

3. Analiza skupień zastosowana do usuwania nadmiernej segmentacji

Opisane powyżej wstępne metody usuwania nadmiernej segmentacji nie są wystarczające. Do łączenia zlewisk lub zapobiegania ich powstaniu używają tylko informacji dostępnych w obrazie gradientowym oraz wyniku transformacji wododziałowej. Ich stosowanie zazwyczaj nie prowadzi do uzyskania satysfakcjonujących wyników. Dlatego zaproponowano wykorzystanie metod analizy skupień do łączenia zlewisk. Metody te są bardzo elastyczne i pozwalają na uwzględnienie wielu parametrów obrazu lub obliczanych na jego podstawie.

W proponowanym rozwiązaniu zlewiska traktowane są jako obiekty opisane zestawem atrybutów. Przed rozpoczęciem analizy skupień dla każdego zlewiska obliczane są wartości wybranych atrybutów. W przypadku obrazów kolorowych wartość każdego atrybutu jest obliczana dla wszystkich składowych, co oznacza, że dla obrazów w przestrzeni sRGB wektor opisujący zlewisko ma trzy razy więcej elementów niż zastosowany zestaw atrybutów. Po wyznaczeniu wartości atrybutów przeprowadzana jest analiza skupień, a następnie, na podstawie wyników, tworzona jest ostateczna segmentacja obrazu.

3.1. Metody analizy skupień

W badaniach uwzględniono cztery metody analizy skupień: metodę pojedynczych powiązań (*single linkage* – SLINK), metodę pełnego wiązania (*complete linkage* – CLINK), metodę średnich połączeń (*unweighted pair-group method using arithmetic averages* – UPGMA) oraz metodę Warda minimalnej wariancji [12]. Wymienione metody pozwalają na utworzenie hierarchii podobieństwa obiektów opisanych za pomocą wektorów atrybutów. Praktycznie nie ma ograniczeń dotyczących rodzaju, ilości czy typu atrybutów. W przedstawionych badaniach wektory atrybutów zawierały do czterech elementów. Wykorzystano wyłącznie atrybuty ilościowe. Proces analizy skupień jest procesem iteracyjnym. Na początku określone są współczynniki podobieństwa wszystkich możliwych par obiektów. Wynik jest umieszczony w macierzy podobieństwa. Następnie wyszukiwane są dwa najbardziej podobne obiekty i następuje ich łączenie (na początku wszystkie obiekty stanowią jednoelementowe klastry). Uaktualniana jest macierz podobieństwa, tak aby uwzględniała nowo powstały cluster i algorytm wraca do wyszukiwania najbardziej podobnych klastrów.

Trzy pierwsze spośród wymienionych metod są dość podobne [12]. Różnią się tylko sposobem uaktualniania macierzy podobieństwa tzn. sposobem w jaki ustalany jest współczynnik podobieństwa między nowo powstałym klastrem a pozostałymi. W przypadku metody CLINK nowym współczynnikiem jest wartość określająca stopień podobieństwa dwóch najbardziej różnych obiektów wchodzących w skład porównywanych klastrów. Metoda SLINK jest przeciwieństwem CLINK. Tam nowy współczynnik to stopień podobieństwa dwóch najbardziej podobnych obiektów wchodzących w skład porównywanych klastrów. W metodzie UPGMA obliczany jest średni współczynnik podobieństwa wszystkich możliwych par utworzonych z elementów obu klastrów (pary tworzy się wybierając po jed-

nym obiekcie z każdego klastra). Tabela 1 w sposób bardziej formalny przedstawia różnice między opisanymi metodami analizy skupień. Uwzględniono w niej również różne rodzaje współczynników tzn. współczynniki podobieństwa (gdy większa wartość oznacza większe podobieństwo) oraz współczynniki różnicy (gdy mniejsza wartość oznacza większe podobieństwo).

Tabela 1
Różnice między metodami analizy skupień

	Miara różnicy	Miara podobieństwa
SLINK	$d_{C_1C_2} = \min_{\substack{m \in C_1 \\ n \in C_2}} d_{nm}$	$s_{C_1C_2} = \max_{\substack{m \in C_1 \\ n \in C_2}} s_{nm}$
UPGMA	$d_{C_1C_2} = \frac{1}{n_{C_1} \cdot n_{C_2}} \sum_{\substack{m \in C_1 \\ n \in C_2}} d_{nm}$	
CLINK	$d_{C_1C_2} = \max_{\substack{m \in C_1 \\ n \in C_2}} d_{nm}$	$s_{C_1C_2} = \min_{\substack{m \in C_1 \\ n \in C_2}} s_{nm}$

gdzie:

- C_1, C_2 – klastry 1 i 2,
- m, n – obiekty należące odpowiednio do klastrów C_1 oraz C_2 ,
- d_{nm} – miara różnicy obiektów m i n ,
- s_{nm} – miara podobieństwa obiektów m i n ,
- $d_{C_1C_2}$ – miara różnicy klastrów C_1 i C_2 ,
- $s_{C_1C_2}$ – miara podobieństwa klastrów C_1 i C_2 ,
- n_{C_1}, n_{C_2} – liczby obiektów w klastrach C_1 i C_2 .

Metoda minimalnej wariancji Warda różni się w sposób bardziej istotny od metod SLINK, CLINK i UPGMA. Jest to metoda, w której dwa klastry są łączone gdy połączenie to powoduje najmniejszy wzrost sumy kwadratów [4] wyrażonej wzorem (9).

$$E = \sum_{m=1}^c E_m \quad (17)$$

$$E_m = \sum_{l=1}^{t_m} \sum_{i=1}^n (X_{ilm} - \bar{X}_{im})^2 \quad (18)$$

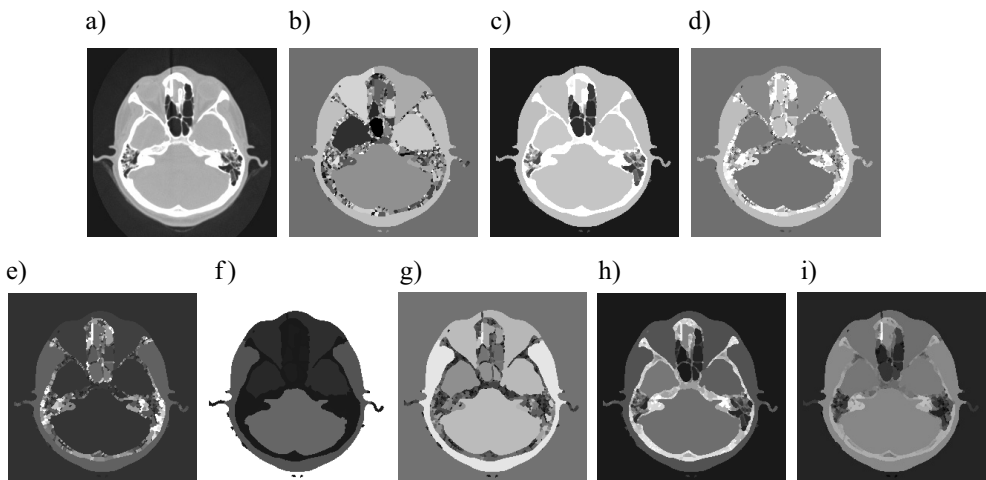
gdzie:

- E – całkowita suma kwadratów błędów,
- E_m – suma kwadratów błędów w m -tym klastrze,
- c – liczba klastrów,
- n – liczba atrybutów,
- t_m – liczba obiektów w klastrze numer m -tym,
- \bar{X}_{im} – średnia wartość i -tego atrybutu w m -tym klastrze.

Algorytm realizujący metodę Warda w każdej iteracji musi sprawdzić wszystkie możliwe połączenia. Ponadto metoda nie wykorzystuje tak jak SLINK, CLINK i UPGMA współczynników podobieństwa opisanych w dalszej części artykułu, lecz korzysta ze swojego „wbudowanego” kryterium wyboru klastrów [12].

3.2. Atrybuty zlewisk

W badaniach nad zastosowaniem analizy skupień do usuwania nadmiernej segmentacji wykorzystano siedem różnych atrybutów [17]. Pierwszym z nich jest rozmiar zlewiska (rys. 8f) wyrażony jako liczba pikseli wchodzących w jego skład. Został włączony ze względu na to, że w niektórych przypadkach tylko niektóre klasy są nadmiernie posegmentowane. Przykładem może być tkanka kostna na rysunku 8b. Drugim rodzajem atrybutu jest średnia wartość punktów w zlewisku. Kolejnymi dwoma atrybutami są wariancja (rys. 8e) i odchylenie standardowe (rys. 8d). Są to atrybuty do pewnego stopnia wrażliwe na teksturę obrazu. Widać to dość dobrze dzięki rysunkom 8c i 8d. Obszary, które mają podobną średnią, dość istotnie różnią się co do wartości odchylenia standardowego. Piątym rodzajem atrybutu jest rozrzut wartości (rys. 8g), a szóstym i siódmym odpowiednio minimalna i maksymalna wartość zlewiska (rys. 8h i 8i).

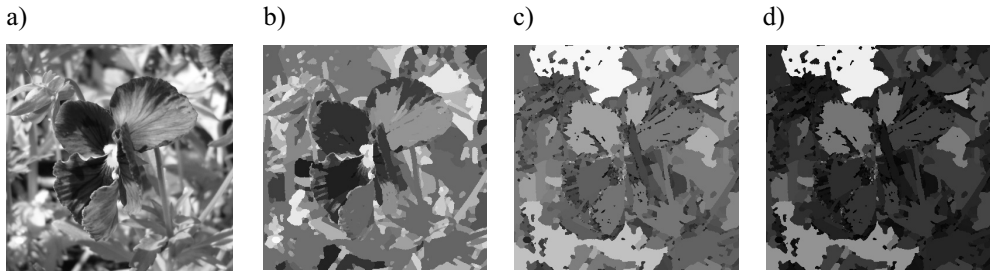


Rys. 8. Atrybuty wykorzystywane w trakcie testów na obrazach czarno-białych:

- a) oryginalny obraz CT (256×256, 12 bitów) [6]; b) zlewiska; c) średnia;
- d) odchylenie standardowe; e) wariancja; f) rozmiar zlewiska; g) rozrzut wartości;
- h) minimalna wartość zlewiska; i) maksymalna wartość zlewiska

W przypadku obrazów w skali szarości atrybuty obliczane są na podstawie obrazu zawierającego etykiety zlewisk oraz obrazu oryginalnego. W przypadku obrazów kolorowych w przestrzeni sRGB wartości atrybutów obliczane są również na podstawie etykiet zlewisk, ale dla każdej składowej barwnej z osobna. W rezultacie otrzymuje się trzy średnie, trzy

wariancje itd. Oczywiście nic nie stoi na przeszkodzie, by obliczać atrybuty tylko dla wybranych składowych. Ta ewentualność zostanie uwzględniona w przyszłych badaniach. Atrybuty zastosowane przy segmentacji obrazów kolorowych przedstawia rysunek 9.



Rys. 9. Atrybuty wykorzystywane w trakcie testów na obrazach kolorowych: a) obraz oryginalny (384×384, 24 bity); b) średnia; c) odchylenie standardowe; d) wariancja

3.3. Metody standaryzacji

Czasami atrybuty obiektów przyjmują wartości z różnych zakresów. Wartości atrybutów mogą się różnić nawet o kilka rzędów wielkości. Aby zapobiec zdominowaniu zestawu atrybutów przez jeden z nich, konieczne staje się stosowanie standaryzacji [12]. Należy jednak zauważyć, co wynika z eksperymentów, że czasami dominujący charakter jednego z atrybutów korzystnie wpływa na wynik segmentacji. W prezentowanych badaniach wykorzystywano cztery różne sposoby standaryzowania atrybutów.

Pierwsze równanie:

$$Z_{ij} = (X_{ij} - \bar{X}_i) / S_i \quad (19)$$

$$\bar{X}_i = \sum_{j=1}^t X_{ij} / t \quad (20)$$

$$S_i = \sqrt{\sum_{j=1}^t (X_{ij} - \bar{X}_i)^2 / (t-1)} \quad (21)$$

gdzie:

j, k – numery obiektów,

t – liczba obiektów,

X_{ij} – wartość i -tego atrybutu w j -tym obiekcie,

powoduje, że każdy z atrybutów ma po standaryzacji wartość średnią $\bar{Z}_i = 0$ oraz odchylenie standardowe $\bar{S}_i = 1$ [12]. Następne dwa równania [12]:

$$Z_{ij} = X_{ij} / X_{\max i} \quad (22)$$

oraz:

$$Z_{ij} = (X_{ij} - X_{\min i}) / (X_{\max i} - X_{\min i}) \quad (23)$$

gdzie $X_{\min i}$, $X_{\max i}$ – minimalna i maksymalna wartość i -tego atrybutu, liniowo odwzorowują przedział wartości atrybutu na przedział $[0, 1]$. W pierwszym przypadku maksymalna wartość w oryginalnych danych odpowiada 1 po standaryzacji. W drugim przypadku, nie tylko maksymalna wartość odpowiada 1, ale także minimalna wartość zawsze odpowiada 0. Ostatnie równanie:

$$Z_{ij} = X_{ij} / \sum_{j=1}^t X_{ij} \quad (24)$$

powoduje, że suma wartości atrybutu po standaryzacji zawsze wynosi 1 [12].

3.4. Współczynniki podobieństwa

Jak wspomniano wcześniej metody SLINK, CLINK i UPGMA do swojego działania potrzebują macierzy podobieństwa obiektów, lecz nie określają sposobu jej obliczania. Musi ona zatem zostać wyznaczona za pomocą dodatkowego współczynnika. W badaniach uwzględniono osiem współczynników przedstawionych poniżej.

Pierwszym współczynnikiem jest odległość euklidesowa [12] opisana wzorem:

$$e_{jk} = \sqrt{\sum_{i=1}^n (X_{ij} - X_{ik})^2} \quad (25)$$

gdzie:

j, k – numery obiektów,

n – liczba atrybutów,

X_{ij}, X_{ik} – wartość i -tego atrybutu w j -tym obiekcie, oraz i -tego w k -tym.

Jest to miara różnicy, co oznacza, że mniejsza wartość jest równoważna z większym podobieństwem.

Drugim współczynnikiem jest średnia odległość euklidesowa [12]:

$$d_{jk} = \sqrt{\sum_{i=1}^n \frac{(X_{ij} - X_{ik})^2}{n}} \quad (26)$$

Różni się ona tym od e_{jk} , że może być obliczona dla niekompletnego wektora atrybutów. Jeżeli dany atrybut nie może być wyznaczony dla pewnego obiektu, jest pomijany a wartość n (oznaczająca liczbę atrybutów) odpowiednio zmniejszana.

Kolejną miarą różnicy między obiektami jest współczynnik różnicy kształtu [12] wyrażony równaniem:

$$z_{jk} = \sqrt{\left[\frac{n}{n-1} \right] (d_{jk}^2 - q_{jk}^2)} \quad (27)$$

gdzie:

$$q_{jk}^2 = \frac{1}{n^2} \left(\sum_{i=1}^n X_{ij} - \sum_{i=1}^n X_{ik} \right)^2 \quad (28)$$

a d_{jk} jest średnią odległością euklidesową. Współczynnik ten porównuje profile dwóch obiektów, nie jest jednak wrażliwy na translacje.

Współczynnik kosinusowy [12]:

$$c_{jk} = \left(\sum_{i=1}^n X_{ij} X_{ik} \right) / \sqrt{\sum_{i=1}^n X_{ij}^2 \sum_{i=1}^n X_{ik}^2} \quad (29)$$

jest miarą podobieństwa, którą można interpretować jako kosinus kąta między dwoma wektorami atrybutów w przestrzeni n -wymiarowej. Ze względu na to, że moduły wektorów nie mają wpływu na kosinus kąta między nimi, współczynnik ten jest niewrażliwy na skalowanie wartości atrybutów (mnożenie wektora przez stałą)

Współczynnik korelacji [12]:

$$r_{jk} = \frac{\sum_{i=1}^n X_{ij} X_{ik} - \frac{1}{n} \left(\sum_{i=1}^n X_{ij} \right) \left(\sum_{i=1}^n X_{ik} \right)}{\sqrt{\left[\sum_{i=1}^n X_{ij}^2 - \frac{1}{n} \left(\sum_{i=1}^n X_{ij} \right)^2 \right] \left[\sum_{i=1}^n X_{ik}^2 - \frac{1}{n} \left(\sum_{i=1}^n X_{ik} \right)^2 \right]}} \quad (30)$$

jest jednym spośród przetestowanych współczynników, który ignoruje zarówno skalowanie jak i translacje.

Współczynnik metryki Canberry [12]:

$$a_{jk} = \frac{1}{n} \sum_{i=1}^n \frac{|X_{ij} - X_{ik}|}{X_{ij} + X_{ik}} \quad (31)$$

wyrównuje wkład każdego z atrybutów do wartości współczynnika różnicy – innymi słowy „nie pozwala”, aby dany atrybut był dominujący.

Ostatnią z zastosowanych miar różnicy jest współczynnik Braya–Curtisa [12]:

$$b_{jk} = \left(\sum_{i=1}^n |X_{ij} - X_{ik}| \right) / \left(\sum_{i=1}^n (X_{ij} + X_{ik}) \right) \quad (32)$$

jest on podobny do metryki Canberra, lecz wpływ poszczególnych atrybutów na wartość współczynnika nie jest wyrównywany.

3.5. Kompletny algorytm segmentacji

Kompletny algorytm segmentacji wykorzystujący transformację wododziałową oraz analizę skupień przebiega w sposób następujący:

1. Obraz jest poddawany działaniu selektywnego filtra rozmywającego.
2. Na podstawie rozmytego obrazu obliczany jest obraz gradientowy.
3. Obraz gradientowy jest progowany w celu usunięcia niewielkich wartości gradientu i zmniejszenia nadsegmentacji.
4. Przeprowadzana jest transformacja wododziałowa.
5. Sąsiadujące ze sobą zlewiska są łączone na podstawie głębokości.
6. Dla każdego z utworzonych zlewisk obliczane są wartości atrybutów.
7. W przypadku użycia metod SLINK, CLINK lub UPGMA obliczana jest macierz podobieństwa/różnicy z wykorzystaniem wybranego współczynnika.
8. Algorytm znajduje dwa najbardziej podobne klastry i łączy je; ponadto uaktualnia drzewo reprezentujące hierarchię podobieństwa.
9. W przypadku użycia metod SLINK, CLINK lub UPGMA macierz podobieństwa/różnicy jest uaktualniana.
10. Jeżeli pozostał więcej niż jeden klastr, algorytm wraca do punktu (8).
11. Na podstawie hierarchii podobieństwa i żądanej liczby klas tworzona jest ostateczna segmentacja.

Tworzenie końcowej segmentacji można obrazowo przedstawić jako przecięcie drzewa reprezentującego hierarchię podobieństwa. Drzewo przecinane jest na poziomie, który zapewnia utworzenie żądanej liczby klas. Dolna część drzewa określa, które obiekty (czyli zlewiska) należy połączyć, natomiast górna jego część jest odrzucana. Im niższy jest poziom, na którym drzewo zostało przecięte, tym więcej klas pojawia się w utworzonej segmentacji.

Poniżej przedstawiono formalny opis algorytmu w postaci pseudokodu:

```
If:=selektywneRozmywanie(I);           //redukcja szumu
Ig:=obliczModulGradientu(I);           //konwersja obrazu na funkcję
                                         // wysokości
It:=progowanie(Ig,t);                   //progowanie (zmniejszenie
                                         //nadsegmentacji)
W:=transformacjaWododzialowa(It);       //transformacja
                                         //wododziałowa
Wm:=lacznZlewiska(W,l);                 //łączenie sąsiednich zlewisk
c:=ileZlewisk(Wm);
A:=obliczAtrybuty(Wm,c);
S:=obliczMacierzPodobienstwa(A,c);     //lub różnicy
```

```

while (c>1)
  //zapisuje wyniki do pierwszy, drugi i odleglosc
  znajdKlastryDoPolaczenia(S,pierwszy,drugi,odleglosc);
  //uaktualnia hierarchie podobienstwa (łączy klastry)
  dodajDoDrzewa(pierwszy,drugi,odleglosc,drzewo);
  //oblicza podobienstwa dla nowego klastra
  uaktualnijMacierzPodobienstwa(S,pierwszy,drugi);
  c:=c-1;
end while;
KoncowaSegmentacja:=przetnijDrzewo(drzewo,liczbaKlasy);

```

gdzie:

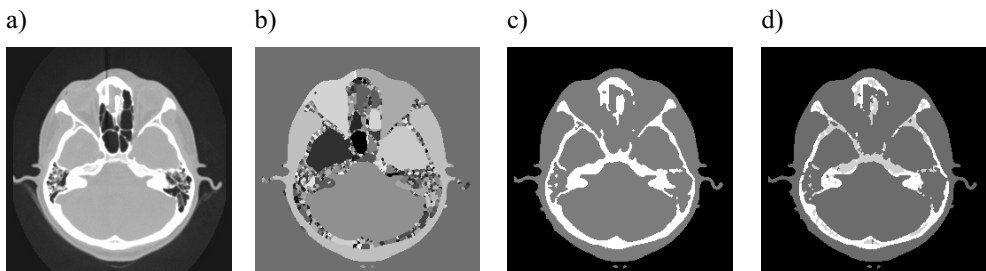
- I – obraz oryginalny,
- I_f – obraz po zastosowaniu selektywnego filtru rozmywającego,
- I_g – obraz gradientowy,
- I_t – obraz gradientowy po progowaniu,
- t – zastosowany próg,
- W – wynik transformacji wododziałowej (obraz z etykietami zlewisk),
- l – poziom łączenia zlewisk,
- W_m – obraz z etykietami po łączeniu zlewisk na podstawie ich głębokości,
- c – bieżąca liczba klastrów (początkowo liczba zlewisk),
- A – tablica przechowująca wartości atrybutów,
- S – macierz podobieństwa/różnicy,
- pierwszy/drugi* – numery klastrów do połączenia,
- odległość* – miara podobieństwa/różnicy klastrów do połączenia,
- drzewo* – drzewo przechowujące hierarchie podobieństwa,
- liczbaKlas* – żądana liczba klas.

4. Wyniki

Wszystkie obrazy, zarówno czarno białe jak i kolorowe, zostały przygotowane tak, aby zmniejszyć stopień nadmiernej segmentacji. Najpierw obrazy były rozmywane za pomocą selektywnego filtru rozmywającego [10, 14]. Dla filtru przyjęto następujące parametry *konduktancja* = 0,3, Δt = 0,12 oraz 5 iteracji. Następnie stosując filtr gradientowy przystosowany do obrazów czarno-białych bądź kolorowych, obliczano obraz gradientowy, który potem progowano w celu usunięcia z niego niewielkich wartości. Po przygotowaniu obrazu przeprowadzano transformację wododziałową. W przypadku obrazów kolorowych wybierano taką kombinację filtru gradientowego (*gradient-pca* lub *gradient-sk*) i progu, aby po transformacji wododziałowej były widoczne wszystkie istotne obiekty. Wartość progu dla obrazów czarno-białych obrazów CT wynosiła 5% maksymalnej głębokości zlewiska, dla obrazów MRI 2%, natomiast dla obrazów kolorowych wartość progu była zróżnicowana. Mimo progowania, w przypadku obrazów kolorowych wynik transformacji wododziałowej był

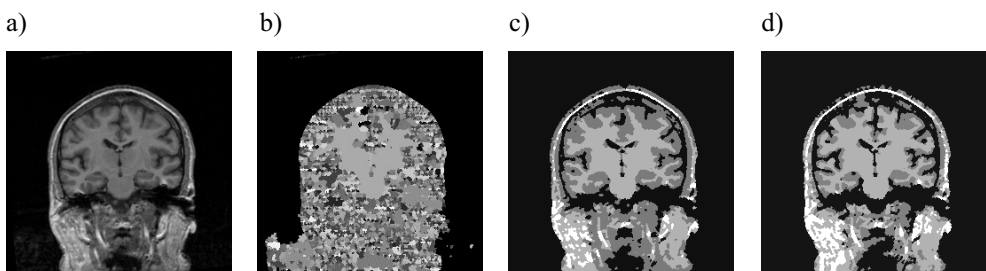
nadal nadmiernie posegmentowany, a liczba zlewisk w obrazie 384×384 często wynosiła około 4000. W celu dalszego zmniejszenia ich liczby zastosowano łączenie zlewisk sąsiednich na podstawie ich głębokości. Wstępne usuwanie nadmiernej segmentacji nie wpływa znacząco na kształt ostatecznej segmentacji, a pozwala na zmniejszenie ilości obliczeń potrzebnych do wykonania analizy skupień.

W testach prezentowanej metody wykorzystano czarno-białe obrazy z tomografu komputerowego (CT) oraz rezonansu magnetycznego (MRI), jak również fragmenty kolorowych zdjęć wykonanych aparatem pracującym w przestrzeni barw sRGB. Artykuł zawiera segmentacje wybranych obrazów.



Rys. 10. Przykładowe segmentacje obrazu CT: a) obraz oryginalny (256×256 , 12 bitów) [6]; b) nadmiernie posegmentowany obraz (885 zlewisk); c) metoda CLINK, odległość euklidesowa, równanie standaryzacyjne (22), 4 klasy, atrybuty: średnia, odchylenie standardowe, rozmiar; d) metoda Warda, równanie standaryzacyjne (19), 4 klasy, atrybuty: średnia, wariancja, rozmiar

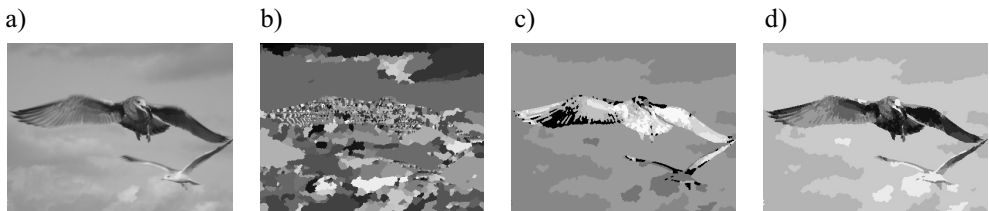
W przypadku obrazów CT i MRI (rys. 10, 11) dobre wyniki najczęściej uzyskiwano, używając metod: Warda, CLINK oraz UPGMA. Metoda SLINK praktycznie nigdy nie dawała satysfakcjonujących rezultatów [16]. W większości udanych segmentacji użyto metody CLINK w połączeniu z odległością euklidesową. Jednakże współczynniki: różnicy kształtu, Braya–Curtisa oraz kosinusowy dawały dobre rezultaty, gdy użyto ich w połączeniu z metodą CLINK lub UPGMA [16]. Metoda Warda nie potrzebuje zewnętrznych miar podobieństwa.



Rys. 11. Przykładowe segmentacje obrazu MRI: a) oryginalny obraz (256×256 , 10 bitów) [8]; b) nadmiernie posegmentowany obraz (2907 zlewisk); c) metoda Warda, brak standaryzacji, 4 klasy, atrybuty: średnia, odchylenie standardowe; d) metoda Warda, brak standaryzacji, 4 klasy, atrybuty: średnia, wartość maksymalna

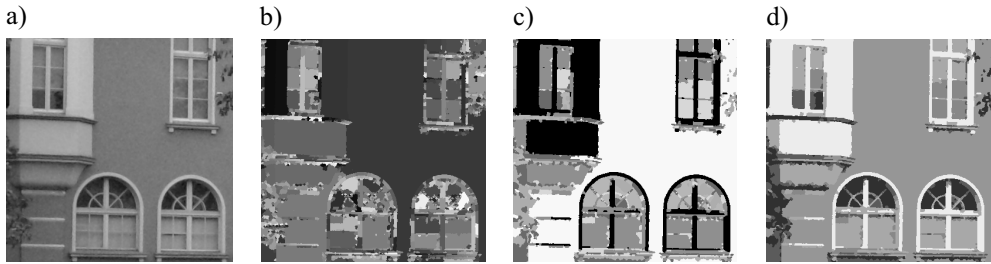
Z przeprowadzonych testów [17] wynika, że to, które zestawy atrybutów dają dobre wyniki, zależy od typu segmentowanego obrazu, jednak wśród najlepszych segmentacji często pojawiały się następujące kombinacje parametrów: średnia i odchylenie standardowe, tylko średnia zlewiska oraz średnia, minimalna i maksymalna wartość. Standaryzacja okazała się nie zawsze potrzebna – dominujący charakter niektórych atrybutów może mieć korzystny wpływ na jakość segmentacji. Należy jednak zauważyć, że w przypadku stosowania standaryzacji najczęściej dobrze sprawdza się standaryzacja liniowa (równania (22), (23)).

Podczas prób wykonywanych na obrazach kolorowych zrezygnowano z użycia metody SLINK, która całkowicie nie sprawdziła się w przypadku obrazów czarno-białych. Ze względu na dużą liczbę możliwych kombinacji parametrów w początkowych badaniach, jako współczynnik różnicy wykorzystano tylko odległość euklidesową oraz ograniczono liczbę parametrów do średniej, wariancji oraz odchylenia standardowego. Dotychczas nie sprawdzono skutków standaryzacji atrybutów, ze względu na to, że zaprezentowane metody traktują każdy atrybut osobno. Oznacza to, że atrybuty obliczone dla różnych składowych byłyby standaryzowane osobno, co prowadziło do zniekształcenia informacji o kolorach. Z przeprowadzonych eksperymentów wynika, że najlepszym zestawem atrybutów była średnia i wariancja zlewiska (co oznacza, że każde zlewisko było scharakteryzowane sześcioma wartościami). Wybranie samej średniej również dawało dobre rezultaty, jednak dodatkowy atrybut (najczęściej wariancja) pozwalał na ograniczenie występowania „zagubionych” klas składających się z niewielu zlewisk i nieodpowiadających rzeczywistym obiektom. Jest to szczególnie widoczne przy niewielkiej liczbie żądanych klas. Dla obrazów kolorowych (rys. 12–15), tak jak w przypadku czarno-białych, najlepsze wyniki uzyskano przy zastosowaniu metody CLINK lub metody Warda, chociaż metoda UPGMA częściej dawała dobre rezultaty.

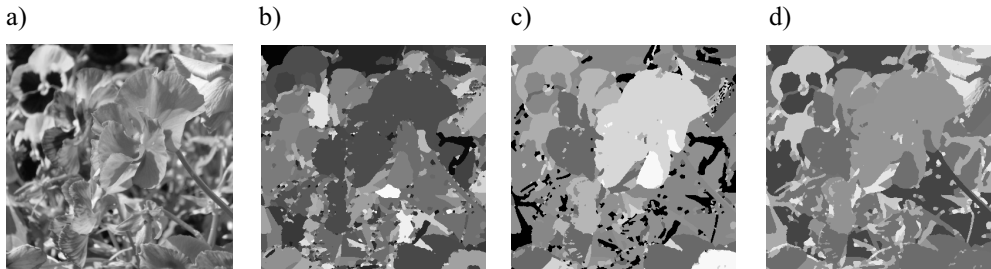


Rys. 12. Przykładowa segmentacja obrazu kolorowego: a) obraz oryginalny (384×288, 24 bity); b) nadmiernie posegmentowany obraz (1967 zlewisk, *gradient-sk* próg 0%, poziom łączenia 7%); c) wynik segmentacji (pokolorowany za pomocą funkcji mieszającej); d) metoda Warda, brak standaryzacji, 9 klas, atrybuty: średnia, odchylenie standardowe

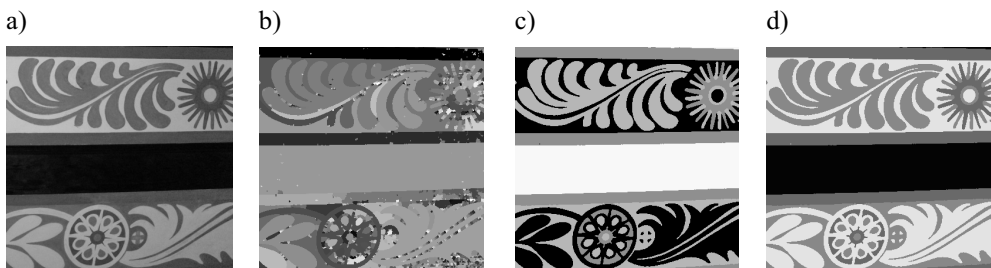
W przedstawionych wynikach liczba klas w ostatecznej segmentacji została wybrana arbitralnie. Celem było uzyskanie wyniku, gdzie wszystkie istotne obiekty są nadal istotne i równocześnie liczba klas jest najmniejsza z możliwych. Gdy parametry zostały odpowiednio dobrane, to zwiększanie liczby klas prowadzi tylko do zwiększania liczby widocznych szczegółów. Niewłaściwe parametry powodują, że istotne obiekty łączone są z innymi, podczas gdy w segmentacji widoczne są nieistotne szczegóły.



Rys. 13. Przykładowa segmentacja obrazu kolorowego: a) obraz oryginalny (384×384, 24 bity); b) nadmiernie posegmentowany obraz (1961 zlewisk, *gradient-pca* próg 0%, poziom łączenia 5%); c) wynik segmentacji (pokolorowany za pomocą funkcji mieszającej); d) metoda Warda, brak standaryzacji, 7 klas, atrybuty: średnia



Rys. 14. Przykładowa segmentacja obrazu kolorowego: a) obraz oryginalny (384×384, 24 bity); b) nadmiernie posegmentowany obraz (1589 zlewisk, *gradient-sk* próg 7%, poziom łączenia 15%); c) wynik segmentacji (pokolorowany za pomocą funkcji mieszającej); d) metoda CLINK, brak standaryzacji, odległość euklidesowa, 13 klas, atrybuty: średnia, odchylenie standardowe



Rys. 15. Przykładowa segmentacja obrazu kolorowego: a) obraz oryginalny (384×384, 24 bity); b) nadmiernie posegmentowany obraz (1618 zlewisk, *gradient-sk* próg 5%, poziom łączenia 10%); c) wynik segmentacji (pokolorowany za pomocą funkcji mieszającej); d) metoda UPGMA, brak standaryzacji, odległość euklidesowa, 4 klasy, atrybuty: średnia, wariancja

5. Wnioski

Segmentacja wododziałowa jest bardzo ciekawą metodą segmentacji. Jej główna wada – czyli nadmierna segmentacja, może być usunięta poprzez odpowiednie przygotowanie obrazu. Istotne jest rozmywanie obrazu wejściowego, zastosowanie odpowiedniego filtra gradientowego oraz dobranie odpowiednich wartości progowania obrazu gradientowego i poziomu łączenia zlewisk. Przygotowanie obrazu i łączenie zlewisk na podstawie ich głębokości nie wystarczają do uzyskania podziału obrazu na niewielką liczbę klas. Wynika stąd konieczność zastosowania dodatkowego etapu przetwarzania wyników transformacji wododziałowej. Jak pokazały próby, do tego zadania nadaje się analiza skupień. Ma ona wiele zalet: wszystkie metody i współczynniki mogą być łatwo wymieniane, można opracować rozmaite zestawy atrybutów dostosowane do konkretnych potrzeb, atrybuty mogą uwzględniać wiele informacji dostępnych w obrazie, np. informację o kolorze.

Literatura

- [1] Angulo J., Serra J., *Color segmentation by ordered mergings*. IEEE Proceedings – International Conference on Image Processing, t. 2, 2003, 125–128.
- [2] Beucher S., Lantuejoul C., *Use of watersheds in contour detection*. International Workshop on Image Processing: Real-time Edge and Motion Detection/Estimation, Rennes, 1979.
- [3] Beucher S., *The watershed transformation applied to image segmentation*. Conference on Signal and Image Processing in Microscopy and Microanalysis, 1991, 299–314.
- [4] Everitt B.S., Landau S., Leese M., *Cluster Analysis*. Arnold a member of the Hodder Headline Group, Londyn, 2001.
- [5] Gonzalez R.C., Woods R.E., *Digital Image Processing*. Addison-Wesley Publishing Company, 1993.
- [6] Kitware 2007: *zestaw danych CT z pakietu vtkdata-5.0.3.tar.gz*. <http://www.vtk.org>.
- [7] Ibanez L., Schroeder W., Ng L., Cates J., *The ITK Software Guide*. Kitware Inc. 2003, <http://www.itk.org>.
- [8] MR 788_6_m 2007: *zestaw danych MR 788_6_m* został udostępniony przez „Center for Morphometric Analysis at Massachusetts General Hospital” i jest dostępny pod adresem <http://www.cma.mgh.harvard.edu/ibsr/>
- [9] NLM Insight 2007: *kod źródłowy biblioteki Insight Toolkit 3.4.0*. <http://www.itk.org>.
- [10] Perona P., Malik J., *Scale-space and edge detection using anisotropic diffusion*. IEEE Transactions on Pattern Analysis and Machine Intelligence, nr 7, t. 12, 1990, 629–639.
- [11] Roerdink J., Meijster A., *The Watershed Transform: Definitions, Algorithms and Parallelization Strategies*. Fundamenta Informaticae, nr 45, 2001, 187–228.
- [12] Romesburg Ch., *Cluster Analysis for Researchers*. Karolina Północna, Lulu Press 2004.
- [13] Russ J.C., *The Image Processing Handbook*. Boca Raton, CRC Press 2002.
- [14] Sapiro G., Ringach D.L., *Anisotropic diffusion of multivalued images with applications to color filtering*. IEEE Transactions on Image Processing, nr 11, t. 5, 1996, 1582–1586.
- [15] Smółka J., *Szybki algorytm rozrostu regionu oparty na zlewiskach*. XIV Krajowa Konferencja Naukowa Biocybernetyka i Inżynieria Biomedyczna, t. 1, 2005, 372–377.
- [16] Smółka J., *Hierarchical cluster analysis methods applied to image segmentation by watershed merging*. Annales UMCS Informatica AI, t. 6, 2007, 73–84.
- [17] Smółka J., Skublewska-Paszowska M., *Comparison of hierarchical cluster analysis methods applied to image segmentation by watershed merging*. Advances in Soft Computing, nr 45, 2007, 84–91.