

Łukasz JAROSIŃSKI, Michał KAWULOK
Politechnika Śląska, Instytut Informatyki

ZASTOSOWANIE ANALIZY WIELOSKALOWEJ DO DETEKCJI LUDZKIEJ SKÓRY W OBRAZACH CYFROWYCH

Streszczenie. Celem prac przedstawionych w niniejszym artykule była poprawa precyzji detekcji ludzkiej skóry w obrazach cyfrowych na podstawie barwy. Wcześniejsze badania autorów wskazywały na możliwość znacznej redukcji błędów segmentacji poprzez zastosowanie analizy obszarów spójnych. W ramach prezentowanych prac zastosowany został mechanizm wieloskalowej analizy ekstremów lokalnych w obrazie transformacji falkowej, pozwalający na wybranie tych obszarów spójnych, które istotnie przedstawiają ludzką skórę. Przedstawione wyniki badań eksperymentalnych potwierdziły skuteczność proponowanych rozwiązań.

Słowa kluczowe: detekcja skóry, analiza wieloskalowa obszarów spójnych

APPLICATION OF MULTISCALE ANALYSIS FOR HUMAN SKIN DETECTION IN DIGITAL IMAGES

Summary. This paper is focused on improving color-based human skin detection in digital images. Our earlier works indicated that the segmentation errors can be significantly reduced using blob detection. The main contribution of this work is the multiscale analysis of local extrema in the wavelet domain, which allows selecting the blobs that indeed represent the human skin. The experimental results have confirmed that the proposed method is effective and reduces the segmentation error.

Keywords: skin detection, multiscale blob analysis

1. Wprowadzenie

Detekcja ludzkiej skóry w obrazach cyfrowych jest istotnym elementem wielu systemów wizji komputerowej oraz znajduje zastosowanie w multimedialnych bazach danych. Jej celem jest wskazanie regionów obrazu przedstawiających ludzką skórę. Opracowanych i opisanych zostało

wiele metod segmentacji skóry na podstawie barwy, które najogólniej można podzielić na parametryczne [4, 8, 9, 13] i statystyczne [5]. Pierwsze z nich funkcjonują na podstawie ściśle określonych reguł, zdefiniowanych w różnych przestrzeniach barwnych, definiujących barwę skóry. Druga grupa metod wymaga statystycznej analizy rozkładu barwy skóry w wielu obrazach stanowiących próbkę treningową, a następnie na podstawie tej analizy przeprowadzana jest detekcja.

Metody te pozwalają na przekształcenie obrazu w mapę prawdopodobieństwa przynależności poszczególnych pikseli do skóry, która następnie poddawana jest progowaniu, w celu wskazania obszarów ludzkiej skóry. Warto zauważyć, że takie postępowanie nie bierze pod uwagę topologii, a każdy piksel rozpatrywany jest niezależnie od swoich sąsiadów. Niestety, analiza spójności obszarów tworzonych przez piksele o wysokim prawdopodobieństwie jest często pomijana w pracach dotyczących tego zagadnienia, co prowadzi do wskazania wielu fałszywych obszarów skóry. Wcześniejsze badania autorów w tej dziedzinie [7] wskazały na możliwość znacznej redukcji błędów poprzez zastosowanie analizy obszarów spójnych i ich odpowiedniej segmentacji.

Badania przedstawione w niniejszym artykule były skoncentrowane na odpowiednim wyborze obszarów spójnych oraz wyznaczeniu ich dokładnych granic. Pierwsze zadanie zostało zrealizowane za pomocą zmodyfikowanego algorytmu detekcji linii maksimów, z wykorzystaniem analizy wieloskalowej. Detekcja ta przeprowadzana jest w obrazie prawdopodobieństwa, wygenerowanym za pomocą istniejących metod, a jej wynikiem jest wskazanie okręgów aproksymujących lokalizację spójnych obszarów obrazu. Dokładne granice obszarów są następnie wyznaczone poprzez algorytmy segmentacji, wykonywane w oryginalnym obrazie barwnym bądź w obrazie prawdopodobieństwa.

Układ artykułu przedstawia się następująco. W rozdziale 2 przedstawione zostały istniejące metody detekcji ludzkiej skóry w obrazach barwnych, które zostały wykorzystane do generowania map prawdopodobieństwa oraz metody segmentacji, pozwalające na wyznaczanie granic tych obszarów. Rozdział 3 zawiera opis algorytmu analizy wieloskalowej z uwzględnieniem jego zastosowania do detekcji obszarów spójnych w obrazach prawdopodobieństwa. Otrzymane wyniki badań eksperymentalnych zostały przedstawione w rozdziale 4. Podsumowanie wraz ze wskazaniem kierunków dalszych badań zostało przedstawione w rozdziale 5.

2. Detekcja ludzkiej skóry na podstawie barwy

2.1. Tworzenie obrazów prawdopodobieństwa występowania ludzkiej skóry

W ramach badań nad detekcją ludzkiej skóry [4, 13] zostało zaobserwowane, że jej barwa ma uniwersalne, dystynktywne właściwości, niezależne od rasy i może ona zostać opisana

w różnych przestrzeniach barwnych za pomocą modeli parametrycznych bądź statystycznych.

Modele parametryczne funkcjonują na podstawie sztywnych reguł decyzyjnych, zdefiniowanych dla różnych przestrzeni barwnych na podstawie analizy rozkładu barwy pikseli, przedstawiających ludzką skórę. Następnie tak stworzone reguły są stosowane dla każdego piksela, w celu rozstrzygnięcia czy przedstawia on fragment ludzkiej skóry.

W literaturze opisane zostały detektory w przestrzeniach RGB [8] i HSV [14], definiujące reguły decyzyjne, na podstawie których kolor piksela może zostać uznany za skórę bądź nie skórę. Według autorów, modele te są dostosowane do zmiennych warunków oświetleniowych, a obrazy poddawane detekcji nie wymagają wcześniejszej normalizacji.

Metoda zaproponowana przez Hsu [4] wykorzystuje wspólne cechy koloru ludzkiej skóry w nieliniowo przekształconej przestrzeni YC_bC_r . Normalizacja sprowadza się do znalezienia 5% najjaśniejszych pikseli w kanale luminancji Y , które służą do zdefiniowania barwy białej. Następnie wszystkie piksele są liniowo skalowane w przestrzeni RGB , tak aby piksele „białe” miały wartość maksymalną we wszystkich trzech kanałach. Po przeprowadzeniu normalizacji stosowany jest eliptyczny model koloru skóry, zdefiniowany w nieliniowo przekształconej przestrzeni YC_bC_r .

Istnieją także techniki operujące w więcej niż jednej przestrzeni, w celu podniesienia dokładności detekcji. Przykładem może być detektor złożony, działający w przestrzeni RGB i YC_bC_r , opracowany przez Kukhareva i Nowosielskiego [9].

Druga grupa metod detekcji skóry na podstawie barwy oparta jest na analizie rozkładu wartości koloru pikseli dla określonego zbioru obrazów treningowych, w których wskazane są obszary ludzkiej skóry [5]. Piksele dzielone są na należące do klasy skóry (C_S) i nie skóry (C_{NS}). Pozwala to na stworzenie modelu statystycznego, umożliwiającego oszacowanie prawdopodobieństwa, że piksel o zadanym kolorze reprezentuje skórę.

Pierwszym krokiem jest zbudowanie histogramów dla obydwu klas pikseli. Następnie, dla każdej wartości piksela (v) obliczane jest prawdopodobieństwo jej wystąpienia w każdej z klas:

$$P(v | C_x) = C_x(v) / N_x, \quad (1)$$

gdzie: $C_x(v)$ jest liczbą pikseli posiadających barwę v w klasie x , a N_x – liczbą wszystkich pikseli w tej klasie.

W celu ograniczenia fałszywych przypadków pozytywnych statystycznego detektora skóry należy nie tylko stwierdzić, że dana wartość występuje często w klasie skóry, ale również konieczne jest zbadanie rozkładu koloru w klasie nie skóry. Prawdopodobieństwo, że dana wartość należy do klasy skóry liczone jest na podstawie prawa Bayesa:

$$P(C_S | v) = \frac{P(v | C_S)P(C_S)}{P(v | C_S)P(C_S) + P(v | C_{NS})P(C_{NS})}, \quad (2)$$

gdzie prawdopodobieństwa *a priori* $P(C_S)$ i $P(C_{NS})$ określone są na podstawie liczby pikseli w każdej z klas. Jeżeli zbiór treningowy jest odpowiednio duży, to prawdopodobieństwo może zostać obliczone dla każdej możliwej wartości koloru, dzięki czemu tworzona jest *mapa prawdopodobieństwa barwy skóry* (P_S). Na podstawie tej mapy, dla piksela o dowolnym kolorze może zostać określone prawdopodobieństwo, że należy on do klasy skóry. Tym samym obraz wejściowy, w którym dokonywana jest detekcja, może zostać przekształcony w *obraz prawdopodobieństwa występowania skóry*.

2.2. Segmentacja obrazów prawdopodobieństwa

Najprostszą metodą segmentacji, przeprowadzaną w celu wskazania obszarów występowania ludzkiej skóry, jest binaryzacja obrazu prawdopodobieństwa na podstawie przyjętego prógu akceptacji (P_{th}). W trakcie prowadzonych badań zostały wykorzystane bardziej zaawansowane metody segmentacji, które są z powodzeniem stosowane w wielu zagadnieniach, związanych z przetwarzaniem obrazu. Wykorzystane metody segmentacji są opisane poniżej.

Algorytm rozrostu ziarna (ang. *flood fill*) [12] to prosty algorytm segmentacji, w którym poszukuje się obszarów zapoczątkowanych przez wybrane punkty (tzw. ziarna). Dla każdego ziarna bada się jego czterech (bądź ośmiu) sąsiadów. Jeżeli sąsiad posiada wartość mieszczącą się w określonym przedziale, wówczas traktowany jest jako część regionu i jego sąsiedzi także są poddawani badaniu. Początek przedziału określa się jako wartość ziarna pomniejszoną o próg dolny, a jego koniec jako wartość ziarna powiększoną o próg górny. W przypadku obrazu o trzech kanałach, należy określić sześć progów. Cała operacja kończy się, gdy nie pozostał żaden sąsiad do przeanalizowania.

Aktywny kontur (ang. *active contour model* lub *snake*) to metoda zaprezentowana w 1988 roku [6]. Dokonuje ona dopasowania konturu określonego za pomocą zbioru punktów na obrazie w skali szarości. Jej celem jest zmniejszenie energii konturu, której wartość wynika z takich czynników, jak: wartość gradientu obrazu, odległości pomiędzy poszczególnymi punktami czy wartością krzywizny.

Działy wodne (ang. *watershed*) [11] bazują na założeniu, że obraz stanowi obszar górzisty (im wyższa wartość, tym wyższa pozycja), a następnie symulują zalewanie go wodą. Ma to na celu znajdowanie zagłębień, w których gromadzi się woda lub całkowite zalanie „terenu” i poszukiwanie jedynie wystających znad wody „wysp”. Zakładając, że kolor czarny to wartość 0, a kolor biały 255 w 8-bitowym obrazie w skali szarości, zagłębienia reprezentować będą obszary spójne ciemniejsze od otoczenia, a wyspy jaśniejsze.

GrabCut [1] został zaprezentowany w 2004 roku jako algorytm przeznaczony do interaktywnej segmentacji obrazów barwnych. Wykorzystuje on wyniki uczenia modelu mieszanin

Gaussowskich (ang. *Gaussian Mixture Model* – GMM), zdefiniowanych osobno dla tła i segmentowanego obiektu. Na wejściu algorytmu należy określić region, wewnątrz którego dokonuje się segmentacji. Zakłada się, że w jego środku znajduje się dokładnie jeden obiekt, a wszystko na zewnątrz jest tłem. Poza regionem można podać punkty wskazujące tło i obiekt, które pomagają zwiększyć precyzję uczenia. Po zakończeniu procesu uczenia tworzony jest graf i dokonywane jest w nim minimalne cięcie, po którym każdy piksel jest sklasyfikowany jako obiekt lub tło.

3. Analiza wieloskalowa w celu detekcji spójnych obszarów skóry

Warto zauważyć, że każda z metod segmentacji zaprezentowanych w podrozdziale 2.2 wymaga wskazania początkowej lokalizacji obszaru, a następnie umożliwia wyznaczenie jego dokładnych granic. W ramach wcześniejszych badań autorów [7] tego rodzaju obszary pierwotne, stanowiące dane wejściowe do segmentacji, były tworzone przez piksele cechujące się bardzo wysokim prawdopodobieństwem. Podejście takie w wielu przypadkach dawało zadowalające rezultaty, mogło jednak prowadzić do wybrania pojedynczych pikseli nienależących do skóry. Z tego względu, w ramach prezentowanych badań, postanowiono sięgnąć po bardziej zaawansowaną metodę detekcji obszarów spójnych działającą w tzw. *przestrzeni skal*. Polega ona na znajdowaniu maksimów w obrazie transformacji falkowej dla różnych skal falek [2]. Maksima te są następnie łączone w linie, na podstawie których wskazywane są obszary spójne.

3.1. Transformacja falkowa w przestrzeni skal

Transformacja falkowa obrazu polega na obliczeniu splotu obrazu $Y(x,y)$ z falką $g(x,y)$ w każdym punkcie (x_0,y_0) :

$$W_g(x_0,y_0) = \sum_i \sum_j g(i,j)Y(x_0+i,y_0+j). \quad (3)$$

Do poszukiwania linii maksimów wykorzystana została falka macierzysta w postaci:

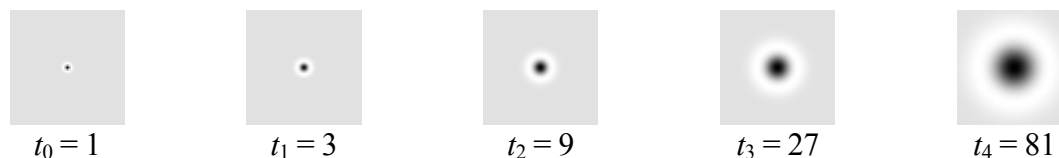
$$g(x,y) = \frac{1}{2\pi} (x^2 + y^2 - 2) \exp\left(-\frac{x^2 + y^2}{2}\right). \quad (4)$$

Znormalizowana falka g_t dla skali t obliczana jest jako:

$$g_t(x,y) = \frac{1}{t} g\left(\frac{x}{\sqrt{t}}, \frac{y}{\sqrt{t}}\right). \quad (5)$$

Transformacja falkowa realizowana jest dla każdej skali t . W proponowanej metodzie zastosowano podejście Ferraza i Binefy [3], zgodnie z którym wartości skal wyznaczane są jako

elementy ciągu: $t_i = \varepsilon t_{i-1}$, $i \in \{1, 2, \dots, n-1\}$. Metoda ma trzy parametry: 1) pierwsza skala t_0 , 2) czynnik ε oraz 3) liczba analizowanych skal n . Na rys. 1 przedstawiono przykładowe obrazy falek, otrzymane dla kilku wybranych skal.



Rys. 1. Przykładowe obrazy falek wygenerowanych dla różnych skal
Fig. 1. Examples of wavelets generated for different scales

3.2. Tworzenie linii ekstremów

Następną czynnością, dokonywaną po wykonaniu transformacji dla zbioru skal, jest poszukiwanie ekstremów lokalnych w uzyskanych obrazach W_g , biorąc pod uwagę sąsiedztwo ośmiopunktowe. Ekstrema liczone są dla każdej skali, a następnie tworzone są linie maksimów. Początkami linii są wszystkie ekstrema znalezione w skali t_0 . Następnie, w skali t_i dla każdej linii maksimów poszukuje się ekstremum o najmniejszej odległości do punktu linii w skali t_{i-1} . Jeżeli jest więcej niż jedno ekstremum w tej samej, minimalnej odległości, uwzględnia się to, którego wartość bezwzględna jest największa. Ekstrema znalezione w skali t_i jest dołączane do linii, jeżeli ekstremum w skali t_{i-1} oraz znalezione ekstremum w skali t_i mają ten sam znak. Jeżeli znaki są różne, ekstremum w skali t_i nie zostaje dołączone, a linia staje się nieaktywna i przestaje być dalej analizowana.

Proponowany algorytm stanowi modyfikację istniejących metod wyznaczania linii maksimów. Zgodnie z metodą Damerval'a [2] poszukuje się najbliższego ekstremum o tym samym znaku, a następnie odrzuca się linię, jeżeli w prostokącie, którego przekątną tworzą dwa ostatnie ekstrema, znajduje się ekstremum o przeciwnym znaku. Jeszcze inne podejście przedstawili Ferraz i Binefa [3], którzy łączenie ekstremów realizują poprzez rozwiązanie problemu optymalizacyjnego za pomocą metody gradientowej. Obydwa alternatywne podejścia analizują wyłącznie maksima, natomiast proponowane podejście znajduje zarówno linie maksimów, jak i minimów.

Gdy nowo dołączone do linii ekstremum ma identyczne współrzędne, jak poprzednie dołączone ekstremum, linia określana jest jako stabilna (nadal jest to linia aktywna), a licznik stabilności ustawiany jest na 1. Jeżeli kolejne ekstremum dodane do linii nie zmienia swojej pozycji względem ostatniego, licznik stabilności jest inkrementowany, natomiast w przypadku zmiany położenia licznik ten jest zerowany.

Gdy licznik stabilności linii przekroczy określoną przez użytkownika liczbę skal, staje się ona nieaktywna. Dla wszystkich aktywnych linii proces zostaje powtórzony. Budowanie linii kończy się, gdy wystąpi przynajmniej jedno z poniższych zdarzeń:

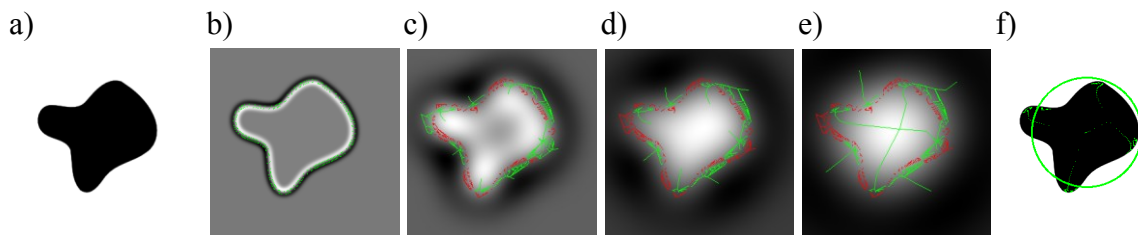
- 1) przeanalizowane zostaną wszystkie skale,

- 2) nie zostanie znalezione ani jedno ekstremum,
- 3) nie pozostała ani jedna aktywna linia.

Po zakończeniu budowania linii, poszukuje się globalnego maksimum osobno dla każdej linii (niezależnie czy jest ona na tym etapie aktywna czy nie), a następnie odrzuca się te linie, których globalne maksimum nie przekroczyło określonej liczby skal (w pierwotnej wersji algorytmu parametr ten wynosi 5 skal).

3.3. Ustalanie obszarów spójnych

Każda linia, która nie jest odrzucona w poprzednim etapie bierze udział w ustalaniu obszarów spójnych. Obszar jest definiowany poprzez okrąg o środku położonym w pozycji globalnego maksimum. Promień okręgu wyznaczany jest jako mediana odległości od tego maksimum do początków wszystkich linii, dla których jest ono globalnym maksimum. Jeżeli dane globalne maksimum są częścią tylko jednej linii, to linia ta jest odrzucana, a obszar spójny nie zostaje wyznaczony. Wynikiem działania algorytmu jest zbiór obszarów spójnych. Każdy z nich jest aproksymowany poprzez okrąg oraz wszystkie ekstrema linii, biorących udział w jego wyznaczeniu.



Rys. 2. Przykład detekcji obszaru spójnego: obraz wejściowy (a), ewolucja linii w kolejnych skalach (b-e), wynik (f)

Fig. 2. Example of blob detection procedure: original image (a), extrema lines in subsequent scales (b-e), the detected blob (f)

Na rys. 2 przedstawiono ewolucję linii ekstremów na przykładzie sztucznie wygenerowanego obrazu (a). Punkty zaznaczone kolorem zielonym (b) oznaczają początki linii ekstremów (po pierwszej skali), które następnie są łączone w linie podczas analizy kolejnych skal: drugiej (c), trzeciej (d) i czwartej (e). Kolorem czerwonym oznaczone są linie nieaktywne. Tło obrazów stanowi obraz po transformacji falkowej w odpowiedniej skali. Analiza prowadzi do wykrycia jednego, spójnego obszaru (f). Przykłady działania algorytmu dla rzeczywistych obrazów są przedstawione w opisie eksperymentów zawartym w rozdziale 4.

4. Walidacja eksperymentalna

Badania eksperymentalne projektu podzielone zostały na dwie części, osobno zbadana została detekcja obszarów spójnych oraz ich segmentacja. Testy przeprowadzono dla 86 obrazów barwnych podzielonych na dwie grupy: 50 obrazów z bazy ECU [10], ukazujących ludzi w codziennych sytuacjach oraz 36 obrazów, przedstawiających gesty języka migowego. Pierwsza grupa to obrazy stosunkowo trudne do analizy, zarejestrowane w niekontrolowanych warunkach. Tło często ma barwę zbliżoną do skóry, co jest przyczyną wielu błędów. Druga grupa obrazów została wykonana w identycznych warunkach oświetleniowych na jednolitym, niebieskim tle. Detekcja skóry może być w nich przeprowadzona z wysoką precyzją za pomocą progowania w obrazie prawdopodobieństwa. Praktycznym celem niniejszych badań było zmniejszenie błędu detekcji dla pierwszej grupy obrazów, przy jednoczesnym utrzymaniu błędu na zbliżonym poziomie wśród obrazów z grupy drugiej. Mapy prawdopodobieństwa zostały wygenerowane z wykorzystaniem modelu statystycznego [5] i były traktowane jako dane wejściowe podczas procesu walidacji.

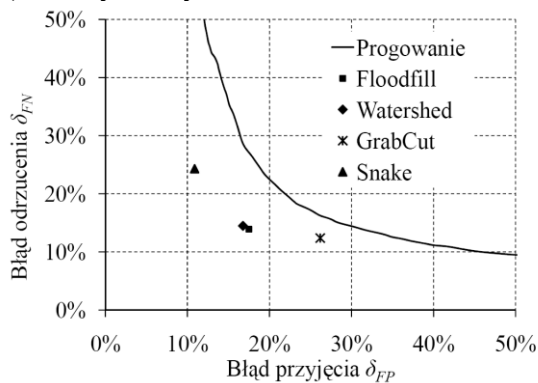
Skuteczność segmentacji była mierzona na podstawie: błędu przyjęcia (δ_{FP}), czyli stosunku liczby pikseli nie skóry błędnie sklasyfikowanych jako będących skórą do liczby wszystkich pikseli skóry i błędu odrzucenia (δ_{FN}), czyli odsetka pikseli skóry, które zostały wskazane jako tło. Powyższe wskaźniki zostały również wykorzystane do określenia skuteczności detekcji obszarów spójnych. W tym przypadku brane pod uwagę były środki okręgów opisujących wykryte obszary spójne oraz fragmenty skóry widoczne w oryginalnym obrazie. Błąd przyjęcia oznaczał odsetek środków wykrytych okręgów leżących poza obszarem skóry. Błąd odrzucenia natomiast oznaczał odsetek faktycznie istniejących, spójnych obszarów skóry, na których nie znajdował się ani jeden środek okręgu.

Wyniki uzyskane dla detekcji obszarów spójnych oraz dla czterech metod segmentacji są przedstawione w tabeli 1. Segmentacja za pomocą aktywnego konturu (*snake*) była przeprowadzana w obrazie prawdopodobieństwa, a w pozostałych przypadkach segmentacja była realizowana w oryginalnym obrazie barwnym. Warto zauważyć, że detekcja spójnych obszarów skóry wykazuje wysoką skuteczność i w przypadku zdecydowanej większości obrazów wszystkie obszary są wykrywane bezbłędnie. Przebadane algorytmy segmentacji pozwoliły na uzyskanie różnych proporcji pomiędzy błędami przyjęcia i odrzucenia. Zdecydowanie najniższy błąd odrzucenia osiągany jest za pomocą algorytmu GrabCut, aczkolwiek odbywa się to kosztem wysokiego błędu przyjęcia, zwłaszcza w przypadku obrazów gestów. Zadowalający balans pomiędzy obydwoma błędami dostarczany jest natomiast przez metodę aktywnego konturu. Wybór najbardziej odpowiedniej metody segmentacji zależy jednak od konkretnego zastosowania i specyficznych dla niego wymagań.

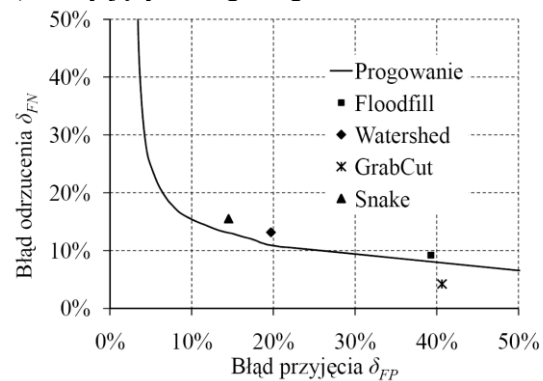
Tabela 1
Błędy detekcji obszarów spójnych oraz segmentacji obszarów ludzkiej skóry

	Gesty języka migowego		Baza ECU	
	δ_{FP}	δ_{FN}	δ_{FP}	δ_{FN}
Detekcja obszarów spójnych	4.4±10.6%	3.7±13.1%	5.1±13.7%	12.1±20.2%
Segmentacja	Flood fill	39.3±13.5%	17.5±21.9%	13.9±11.6%
	Watershed	19.7±8.2%	13.2±7.9%	16.8±17.6%
	Snake	14.5±5.9%	15.5±8.9%	10.9±11.4%
	GrabCut	40.6±8.0%	4.2±6.1%	26.2±18.9%

a) Obrazy z bazy ECU



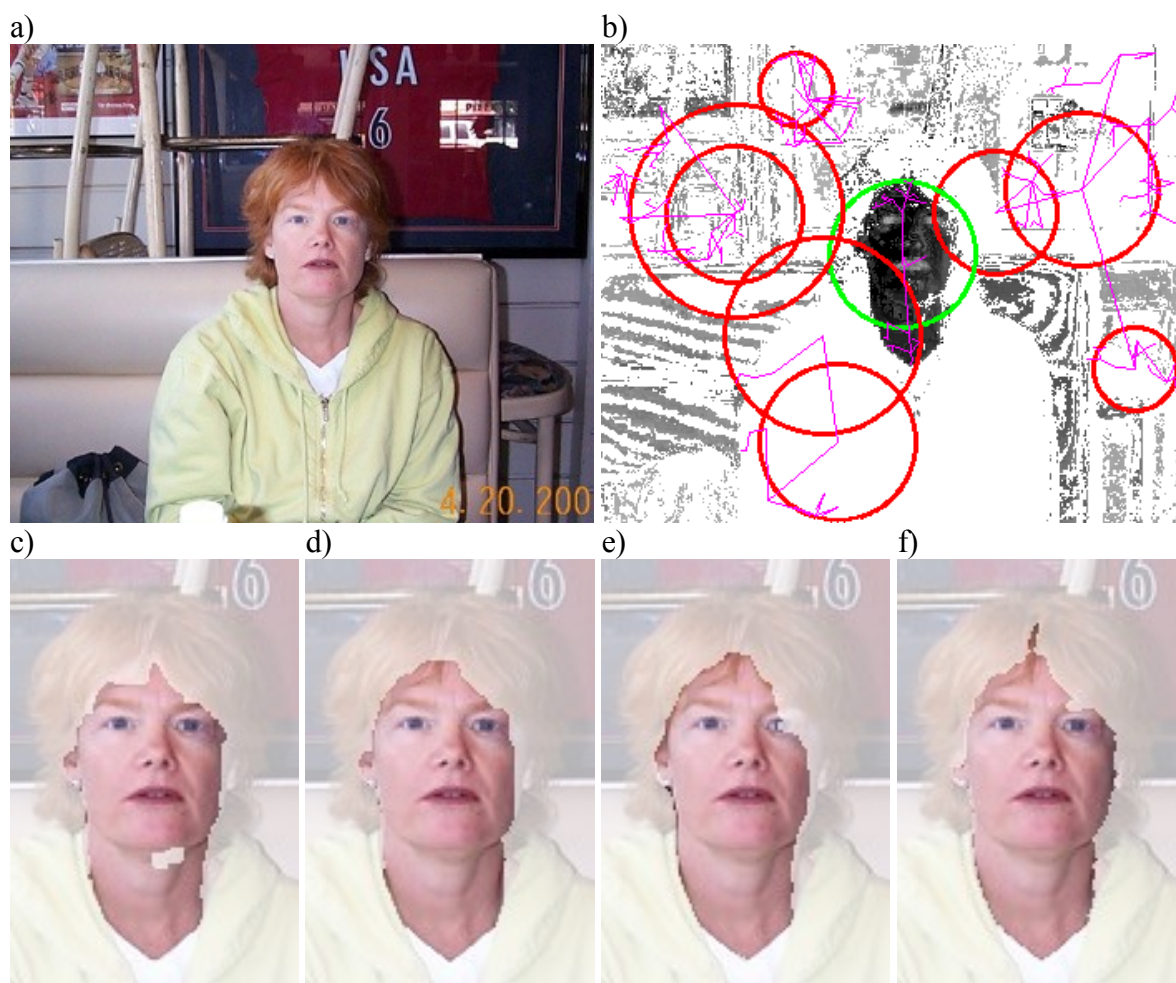
b) Gesty języka migowego



Rys. 3. Krzywe ROC dla progowania obrazów prawdopodobieństwa oraz wyniki uzyskane za pomocą detekcji i czterech metod segmentacji

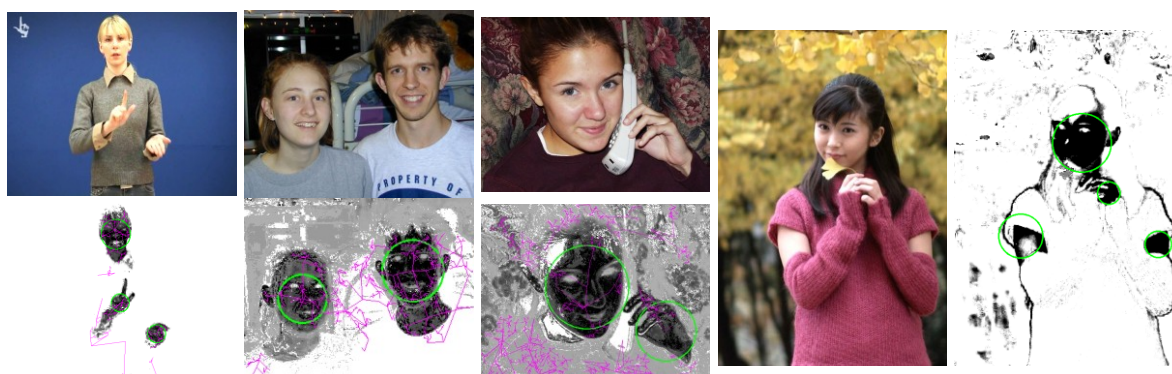
Fig. 3. ROC curves obtained for probability image thresholding and results achieved using blob detection and four segmentation methods

W przypadku progowania opisanego w podrozdziale 2.2 błędy przyjęcia i odrzucenia zależy od wartości progu akceptacji, a wynik działania detektora może być prezentowany w postaci krzywej ROC (ang. *Receiver Operating Characteristic*), ilustrującej zależność błędu odrzucenia od błędu przyjęcia. Zmniejszanie wartości progu pozwala na zmniejszenie błędu odrzucenia kosztem wzrostu błędu przyjęcia. Z kolei zwiększanie wartości progu zawęża obszar klasyfikowany jako skóra, co redukuje błąd przyjęcia, ale powoduje również zwiększenie błędu odrzucenia. Krzywe te są przedstawione na rys. 3 dla obrazów z bazy ECU (a) oraz dla obrazów gestów języka migowego (b). Ponadto, na wykresach są zaznaczone wyniki, uzyskane za pomocą czterech metod segmentacji. Na rys. 3 widać, że stosowanie zaproponowanej detekcji obszarów spójnych pozwala na zdecydowane zmniejszenie błędów w przypadku obrazów z bazy ECU, dla każdej z czterech metod segmentacji (punkty prezentujące uzyskane wyniki znajdują się pod krzywą ROC). W przypadku obrazów przedstawiających gesty języka migowego, zarejestrowanych w kontrolowanych warunkach, zaproponowane rozwiązanie dało wyniki zbliżone do tych, uzyskanych za pomocą progowania. Tym samym, wyniki te potwierdzają, że założony cel został osiągnięty, a proponowana detekcja obszarów spójnych pozwala na zredukowanie błędów segmentacji regionów ludzkiej skóry.



Rys. 4. Obraz wejściowy (a), wykryte obszary spójne (b) oraz segmentacja za pomocą algorytmu rozrostu ziarna (c), GrabCut (d), działów wodnych (e) i aktywnego konturu (f)

Fig. 4. Input image (a), detected blobs (b) and the results achieved using flood filling (c), GrabCut (d), watershed (e) and active contour (f) segmentation algorithms



Rys. 5. Przykłady detekcji obszarów spójnych skóry

Fig. 5. Examples of skin blobs detection

Na rys. 4 przedstawiony jest przykład działania detekcji obszarów spójnych na tle wygenerowanego obrazu prawdopodobieństwa (b). Kolorem zielonym został zaznaczony okrąg reprezentujący obszar wysokiego prawdopodobieństwa, a kolorem czerwonym okręgi wyznaczające tło. Dodatkowo, przedstawione zostały linie ekstremów, na podstawie których

wykonana została detekcja. Pokazane zostały także wyniki segmentacji, uzyskane za pomocą czterech metod (c-f). Przykłady poprawnej detekcji obszarów spójnych są ponadto zaprezentowane na rys. 5.

5. Podsumowanie i kierunki dalszych badań

Zaproponowana metoda detekcji obszarów spójnych za pomocą analizy linii maksimów w obrazie transformacji falkowej pozwala skutecznie wskazać regiony ludzkiej skóry w obrazie prawdopodobieństwa, co zostało potwierdzone przez zaprezentowane wyniki badań eksperymentalnych.

Najsłabszym punktem opracowanej metody jest segmentacja, przeprowadzana po wykryciu obszarów spójnych. Żadna z czterech rozpatrywanych metod nie przyniosła w pełni zadowalających rezultatów i dlatego dalsze prace w prezentowanej dziedzinie zostaną ukierunkowane na poprawę precyzji segmentacji obszarów wskazanych przez opracowany detektor. Ponadto, w szerszym kontekście, planowane jest wykorzystanie informacji o teksturze do generacji bardziej wiarygodnych map prawdopodobieństwa. Ponadto, w trakcie dalszych prac zostanie położony większy nacisk na wykonanie badań porównawczych z najnowszymi metodami detekcji skóry.

BIBLIOGRAFIA

1. Boykov Y. Y., Jolly M.-P.: Interactive graphcuts for optimal boundary & region segmentation of objects in n-d images. *Computer Vision, IEEE International Conference on*, Vol. 1, 2001, s. 105÷112.
2. Damerval C., Meignen S.: Blob Detection With Wavelet Maxima Lines, *Signal Processing Letters*, Vol. 14, No. 1, 2007, s. 39÷42.
3. Ferraz L., Binefa X.: A scale invariant interest point detector for discriminative blob detection, [in:] *IbPRIA'09: Proceedings of the 4th Iberian conference on Pattern Recognition and Image Analysis*, Springer-Verlag, 2009, s. 233÷240.
4. Hsu R.-L., Abdel-Mottaleb M., Jain A. K.: Face detection in color images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*. Vol. 24, No. 5, May 2002, s. 696÷706.
5. Jones M. J., Rehg J. M.: Statistical color models with application to skin detection, *International Journal of Comp. Vision*, Vol. 46, No. 1, 2002, s. 81÷96.
6. Kass M., Witkin A. P., Terzopoulos D.: Snakes: Active contour models. *International Journal of Computer Vision*, Vol. 1, No. 4, 1988, s. 321÷331.

7. Kawulok M.: Energy-based blob analysis for improving precision of skin segmentation. *Multimedia Tools and Applications*, Vol. 49, No. 3, Springer, 2010, s. 463÷481.
8. Kovac J., Peer P., Solina F.: Human Skin Colour Clustering for Face Detection, Eurocon 2003 Ljubljana, Slovenia 2003, s. 144÷148.
9. Kukharev G., Nowosielski A.: Fast and Efficient Algorithm For Face Detection in Colour Images. *Machine Graphics & Vision*, Vol. 13, No. 4, Warsaw 2004, s. 377÷399.
10. Phung S. L.: ECU Face Detection Database, University of Wollongong, 2002.
11. Rother C., Kolmogorov V., Blake A.: Grabcut: interactive foreground extraction using iterated graphcuts. *ACM Trans. Graph.*, Vol. 23, No. 3, 2004, s. 309÷314.
12. Shaw J. R.: QuickFill: An efficient floodfill algorithm. Online article. <http://www.codeproject.com/KB/GDI/QuickFill.aspx>, March 2004, Accessed December 30, 2010.
13. Terrillon J-C., David M., Akamatsu S.: Automatic Detection of Human Faces in Natural Scene Images by Use of a Skin Color Model and of Invariant Moments. In Proc. of the 3rd International Conf. on Automatic Face and Gesture Recognition, Nara, Japan 1998, s. 112÷117.
14. Tsekeridou S., Pitas I.: Facial feature extraction in frontal views using biometric analogies. [in:] Proc. of EUSIPCO '98, Vol.1, Rhodes, Greece 1998, s. 315÷318.

Recenzenci: Prof. dr hab. inż. Ewa Piętka
Dr inż. Łukasz Wyciślik

Wpłynęło do Redakcji 5 stycznia 2011 r.

Abstract

Skin detection is a widely investigated area and many methods for solving this problem have been developed. They work by assigning skin probability to every pixel based on its color. In this way the image is converted into probability image which indicates the areas. Parametric and statistical skin models can be distinguished within the existing approaches to this problem. Some of these methods are quite effective, but in general they produce high false positive errors in case of objects which color is similar to human skin. This is due to the fact that the classification relies exclusively on the pixel color without taking its neighborhood into account, and skin areas are extracted by image binarization with a fixed threshold.

Our earlier works showed that blob analysis performed in the probability image may decrease the false positive error significantly. In the work reported here we propose to apply a blob detection technique relying on the multiscale analysis of wavelet-transformed images. In this way the most relevant blobs are selected in the probability image and their boundaries are determined following one of the well-established segmentation techniques. We have investigated flood fill, GrabCut, watershed and active contour algorithms for this purpose.

The presented experimental results show that the blob detection is very effective and in the vast majority of cases all of the skin blobs were correctly found. After the segmentation the errors decreased for ECU database which contains real-life images. However, a slight increase has been noticed for the gesture images acquired in controlled conditions. In the latter case the simple thresholding delivers sufficiently good results and the proposed technique could not contribute much. It may be concluded from the results that the segmentation aimed at finding the blob boundaries requires further attention and alternative methods will be investigated in the future.

Adresy

Łukasz JAROSIŃSKI: Politechnika Śląska, Instytut Informatyki, ul. Akademicka 16, 44-101 Gliwice, Polska, lukaszj86@gmail.com.

Michał KAWULOK: Politechnika Śląska, Instytut Informatyki, ul. Akademicka 16, 44-101 Gliwice, Polska, mkawulok@polsl.pl.