

Jaromir Przybyło*, Mirosław Jabłoński*, Paweł Wołoszyn*

Detekcja markerów dla celów automatycznej anotacji mimiki twarzy**

1. Wprowadzenie

Lokalizacja cech morfologicznych twarzy jest istotnym etapem w algorytmach rozpoznawania twarzy oraz jej mimiki. Przykładem może być metoda aktywnych modeli kształtu i wyglądu (*active shape and appearance models*) zaproponowana w pracy [2]. Do rozpoznawania i kodowania twarzy wykorzystywane są modele opisujące zarówno przestrzenne rozmieszczenie punktów charakterystycznych, jak i teksturę twarzy. Tworzone są one na podstawie analizy statystycznej zestawu uczącego, zawierającego obrazy twarzy z zaznaczonym (ręcznie) położeniem wszystkich punktów. Również w poprzedniej pracy autorów niniejszego artykułu [4], do analizy i rozpoznawania mimiki twarzy w kolejnych klatkach sekwencji wideo wykorzystano statystyczne modele kształtu.

Istotnym utrudnieniem w wykorzystaniu powyższych metod jest konieczność posiadania bazy obrazów twarzy wraz z anotacją punktów charakterystycznych. Baza taka zawiera z reguły kilkaset do kilku tysięcy obrazów, na których należy ręcznie zaznaczyć położenie od kilku do kilkunastu punktów. Utrudnia to rozwijanie algorytmów rozpoznawania i analizy ze względu na konieczność poświęcenia czasu na ręczną anotację obrazów.

Dlatego też w niniejszej pracy autorzy skupili się na opracowaniu metody automatycznej anotacji obrazów z wykorzystaniem sztucznych markerów. W rozdziale drugim omówiono zagadnienia związane z wyborem lokalizacji markerów. Rozdział trzeci przedstawia opis algorytmu detekcji markerów. Rozdział czwarty objaśnia metodę weryfikacji poprawności detekcji. W piątym rozdziale przedstawiono wyniki przeprowadzonych testów oraz wnioski dotyczące skuteczności opracowanych algorytmów. Artykuł kończy podsumowanie przeprowadzonych prac oraz przedstawienie dalszych planów rozwoju opisanych metod.

2. Wybór lokalizacji markerów

W poprzednich pracach obejmujących zagadnienia rozpoznawania jednostek czynnościowych mimiki twarzy [4] rozważano położenie punktów charakterystycznych wyzna-

* Katedra Automatyki, Akademia Górniczo-Hutnicza w Krakowie

** Praca wykonana w ramach badań własnych – umowa AGH nr 10.10.120.39

czających figury zmieniające kształty podczas wykonywania ruchów mimicznych. Punkty te dobrane zostały według kryteriów zakładających dobrą widoczność, precyzję lokalizacji i międzyosobniczą powtarzalność. Przy wybieraniu położenia punktów kształtu kierowano się cechami wizualnymi poszczególnych lokalizacji oraz ich otoczenia na zarejestrowanym obrazie twarzy. Zastosowanie nanoszonych na skórę twarzy markerów zmusza do rewizji przyjętych wcześniej lokalizacji, ponieważ tym razem w ich ustalaniu należy brać pod uwagę uwarunkowania anatomiczne zamiast wizualnych.

Podstawowym warunkiem, jaki musi spełnić położenie każdego markera, jest możliwość jego umieszczenia na twarzy osoby obserwowanej. Lokalizacje takie, jak na przykład środek źrenicy oka, nie mogą być zastosowane z oczywistych powodów. Z podobnych względów odrzucić trzeba także miejsca położone w okolicach o wyraźnej ruchomości skóry, na przykład kąciaki oczu. Jakkolwiek mogą to być miejsca stosunkowo łatwe do identyfikacji na obrazie, ich anatomiczny kontekst sprawia, że nie jest możliwe oznakowanie tych punktów na etapie wykonywania zdjęć rzeczywistej twarzy, a dopiero w rezultacie późniejszej analizy uzyskanych obrazów.

Ponieważ w dalszej perspektywie prowadzonych prac badawczych leży automatyczna anotacja obrazów twarzy, dlatego też markery nie mogą zmieniać wyglądu bezpośrednich okolic punktów charakterystycznych. W przeciwnym wypadku w wyniku anotacji powstałaby wprowadzicie obraz uzupełniony o współrzędne markerów, ale pozbawiony oryginalnych szczegółów, a zatem w małym stopniu przydatny do dalszego przetwarzania. Aby uniknąć takiej sytuacji, można posłużyć się metodą pośredniego oznaczania punktów charakterystycznych, wykorzystując grupy markerów rozlokowanych w ich pobliżu. Markery same w sobie nie wskazują wówczas bezpośrednio położenia punktu, ale na podstawie ich lokalizacji w grupie można odtworzyć to położenie, uwzględniając geometryczne zależności wiążące ze sobą zajmowane przez markery miejsca (rys. 1).



Rys. 1. Pośrednie wskazanie położenia punktu za pomocą grupy markerów

Zaletą takiego rozwiązania jest zachowanie niezmiennego obrazu w otoczeniu anotowanego punktu, metoda pociąga za sobą jednak kilka konsekwencji. Wymagana jest tutaj większa liczba markerów, które mogą nadal zaburzać obraz twarzy. Aby zredukować owo zaburzenie, można dokonać automatycznego retuszu markerów na obrazie twarzy. Korekta tego rodzaju wymaga umieszczenia markerów w miejscach o jednorodnej kolorystyce i teksturze dającej się łatwo zrekonstruować. Powinny być to również punkty nie odpowiadające charakterystycznym cechom anatomicznym czy wizualnym twarzy. Takie neutralne pod wieloma względami punkty dają się wyznaczyć na gładkiej, nieowłosionej skórze.

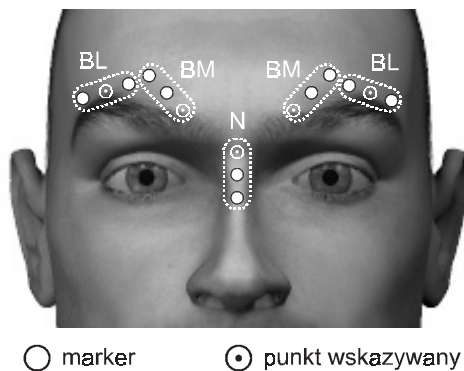
Istotnym problemem przy wykorzystaniu markerów pośrednich jest utrzymanie ich wzajemnych stosunków geometrycznych w poszczególnych grupach podczas ruchów twarzy, zwłaszcza ruchów mimicznych. Przy wykonywaniu przez fotografowaną osobę jednostek czynnościowych mimiki różne obszary skóry przemieszczają się względem siebie

i ulegają deformacji. Aby możliwe było wyznaczenie faktycznej lokalizacji punktu charakterystycznego, niezbędny jest taki dobór grup markerów, by ich położenia względem siebie i względem wskazywanego punktu ulegały możliwie najmniejszym zmianom. W szczególności dużą rolę odgrywa tu przebieg włókien mięśni mimicznych oraz układ struktur szkieletu twarzy.

Tabela 1
Lokalizacja punktów wskazywanych przez markery

Oznaczenie	Lokalizacja
BM	punkt położony na górnym brzegu brwi w tej samej płaszczyźnie strzałkowej, co kąt przyśrodkowy oka
BL	punkt położony na górnym brzegu brwi w tej samej płaszczyźnie strzałkowej, co kąt boczny oka
N	punkt położony w linii pośrodkowej u nasady nosa

Podczas eksperymentów prowadzonych w ramach prac opisywanych w niniejszym artykule posłużono się trzema zestawami markerów wskazującymi niezależne trzy punkty charakterystyczne twarzy, jeden anatomicznie nieparzysty, dwa zaś symetryczne (tab. 1). Każdy z punktów wymagał pośredniego wskazania za pomocą dwóch markerów umieszczonych na skórze twarzy we względnie neutralnym pod względem ruchomości i wyglądu obszarze (rys. 2).

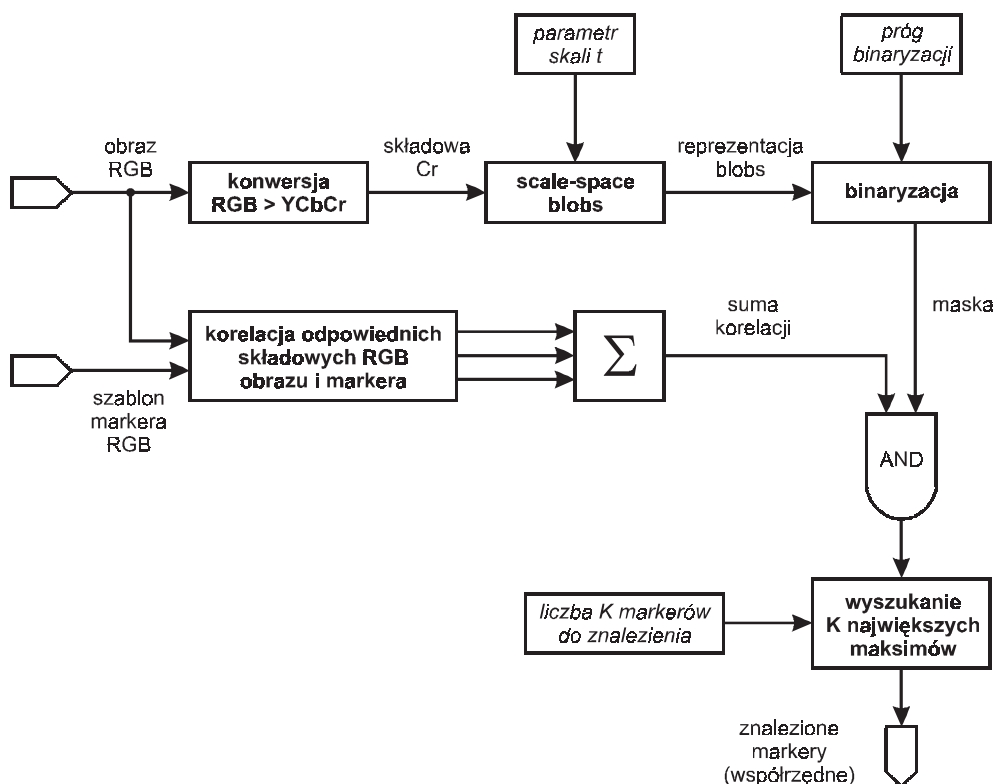


Rys. 2. Położenie grup markerów i wskazywanych punktów twarzy

Istotnym kryterium doboru markerów była również możliwość ich odróżnienia od skóry człowieka. Dlatego zdecydowano się na markery w postaci kół w kolorze niebieskim (niewystępującym w naturalny sposób na twarzy człowieka). W trakcie prac, okazało się jednak, iż tor kamery analogowej wprowadza zniekształcenia barwne – niebieski marker widziany jest jako zbiór pikseli o różnych odcieniach, od białego do niebieskiego. Spowodowało to konieczność rezygnacji z prostej detekcji markerów (binaryzacja) na rzecz algorytmu opisanego w rozdziale trzecim.

3. Algorytm detekcji markerów

Algorytm detekcji markerów został przedstawiony na rysunku 3.



Rys. 3. Schemat blokowy algorytmu detekcji markerów

Wejściowy obraz RGB konwertowany jest do przestrzeni kolorów chrominancji-luminancji YCbCr, z której do dalszej analizy wybierana jest składowa Cr. Wybór składowej Cr podyktowany został występowaniem wyraźnej różnicy między pikselami markera (ciemne, jednolite obszary) a otaczającymi go pikselami twarzy (jasny obszar). Dla innych składowych różnica ta jest dużo mniejsza oraz występują zakłócenia (marker składa się z pikseli o różnych wartościach) – rysunek 4.

W celu wzmocnienia różnicy, dla składowej Cr, wyznaczana jest reprezentacja *scale-space blobs*¹⁾ dla pojedynczej skali dobranej odpowiednio do rozmiarów markerów na obrazie. Reprezentacja *scale-space* jest specjalnym typem reprezentacji wieloskalowej zapro-

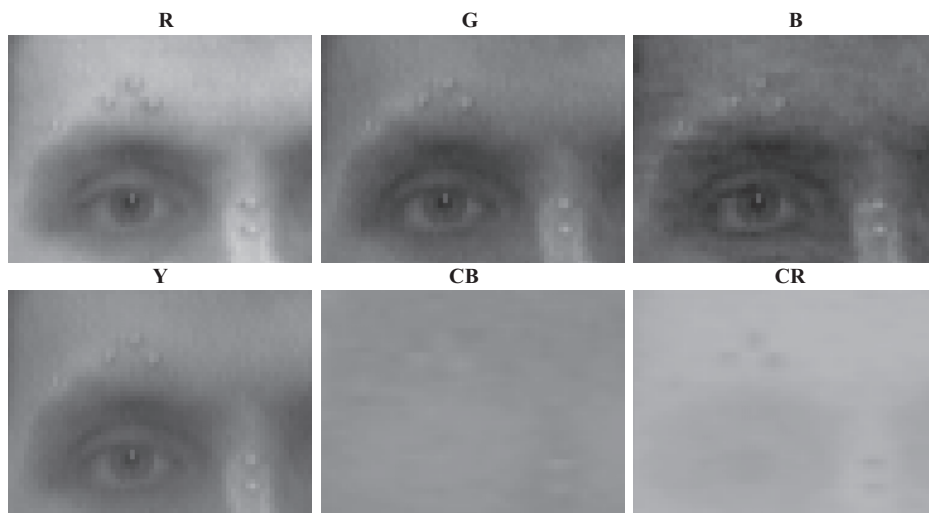
¹⁾ W powszechnie dostępnej literaturze występuje termin anglojęzyczny, dlatego też autorzy tego artykułu używają określenia *scale-space blobs*

ponowanym w [3]. Niech $f : \mathfrak{R}^N \rightarrow \mathfrak{R}$ reprezentuje dowolny dany sygnał. Reprezentacja *scale-space* $L : \mathfrak{R}^N \times \mathfrak{R}_+ \rightarrow \mathfrak{R}$ jest zdefiniowana przez:

$$\begin{aligned} L(;\;0) &= f \\ L(;\;t) &= g(;\;t) * f \end{aligned} \tag{1}$$

gdzie $t \in \mathfrak{R}_+$ jest parametrem skali, a $g : \mathfrak{R}^N \times \mathfrak{R}_+ \setminus \{0\} \rightarrow \mathfrak{R}$ jest N -wymiarowym rozkładem Gaussa:

$$g(x,t) = \frac{1}{(2\pi t)^{\frac{N}{2}}} e^{-\sum_{i=1}^N \frac{x_i^2}{2t}} \tag{2}$$



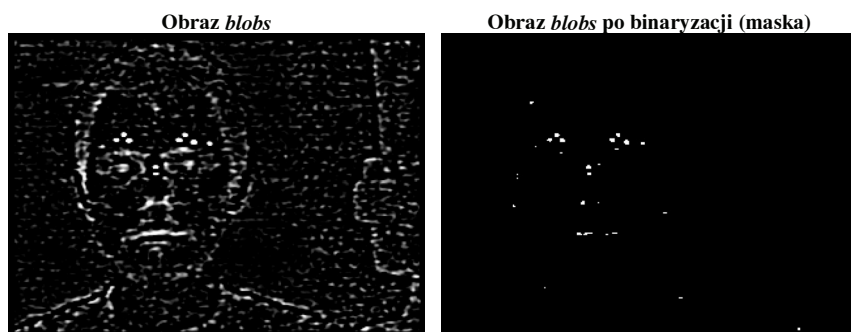
Rys. 4. Składowe wybranego obszaru obrazu dla modeli kolorów RGB i YCbCr

Detekcja cech w pojedynczej skali odbywa się poprzez wyszukiwanie lokalnych przestrzennych maksimów lub miejsc zerowych różniczkowego deskryptora cech – w tym przypadku jako cechy przyjęto plamy (*blobs*). Plama to obszar ciemny otoczony elementami jasnymi lub odwrotnie. Matematycznie plama to ekstremum lokalne laplasjanu

$$\begin{aligned} \nabla^2 L &= L_{xx} + L_{yy} \\ \nabla(\nabla^2 L) &= 0 \end{aligned} \tag{3}$$

Znak laplasjanu mówi nam, czy mamy do czynienia z plamą jasną ($\nabla^2 L < 0$) czy ciemną ($\nabla^2 L > 0$). W wyniku zastosowania powyższego operatora, otrzymujemy obraz

z wyraźnymi maksimumami w miejscach występowania plam o rozmiarze odpowiadającym wybranej skali t . Obraz ten jest następnie binaryzowany z ustalonym progiem, co w rezultacie daje obraz binarny (maskę) zawierający obszary prawdopodobnego występowania markerów (rys. 5).



Rys. 5. Reprezentacja *scale-space blobs* dla przykładowej ramki

Ponieważ obszarów jest znacznie więcej niż markerów (zakłócenia, inne elementy obrazu o podobnej charakterystyce, co markery), dlatego też równocześnie wykonywana jest dodatkowa analiza. Polega ona na wyznaczeniu korelacji pomiędzy poszczególnymi składowymi RGB obrazu, a składowymi RGB szablonu markera. Dwuwymiarowa korelacja jest obliczana według następującego wzoru:

$$C(i, j) = \sum_{m=0}^{(Ma-1)(Na-1)} \sum_{n=0}^{(Na-1)} A(m, n) \cdot \text{conj}(B(m+i, n+j)) \quad (4)$$

$$0 \leq i < Ma + Mb - 1, \quad 0 \leq j < Na + Nb - 1,$$

gdzie:

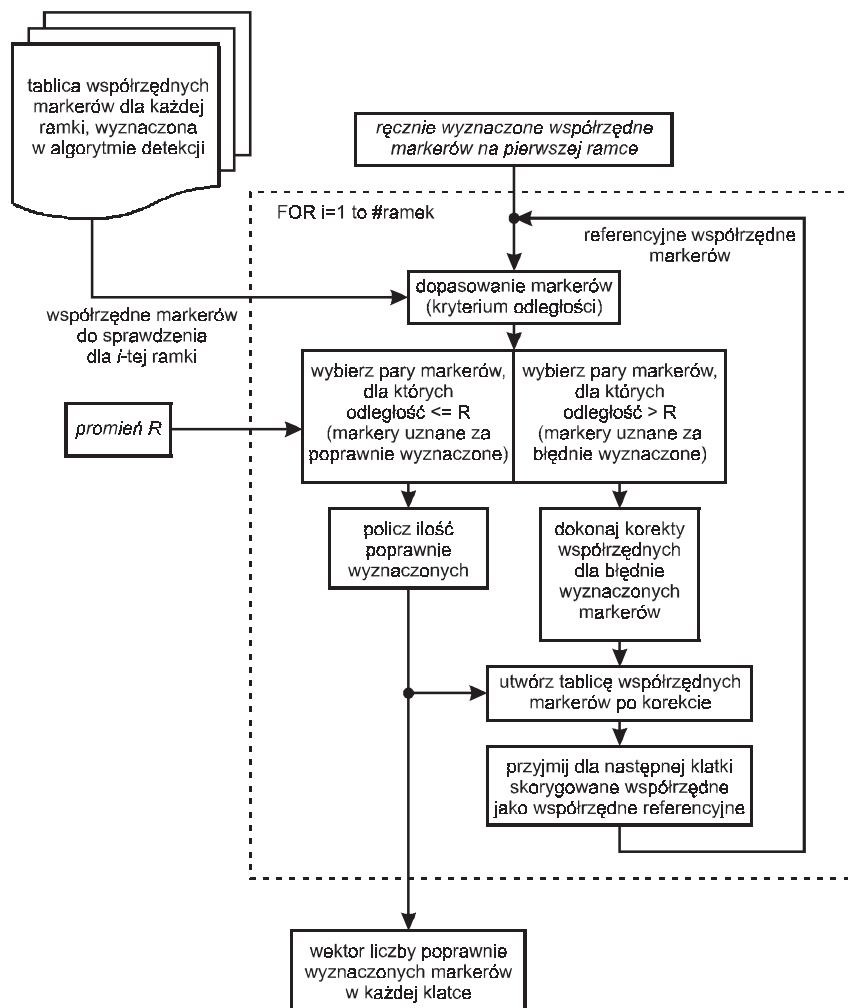
- A – obraz o rozmiarze $[Ma, Na]$,
- B – szablon o rozmiarze $[Mb, Nb]$,
- conj* – operator liczby sprzężonej.

Ponadto wyniki korelacji są normalizowane przez wyrażenie $\sqrt{\sum \text{dot}(I1p, I1p) \cdot \text{dot}(I2, I2)}$, gdzie $I1p$ jest częścią obrazu leżącą pod szablonem markera, *dot* – oznacza mnożenie tablicowe w MATLAB-ie. Szablon markera został wcześniej utworzony poprzez ręczny wybór wszystkich markerów z pierwszej klatki sygnału wideo, a następnie wyznaczenie średniego obrazu markera (dla każdego piksela markera jest to średnia z odpowiadających sobie pikseli kolejnych markerów).

W dalszej części algorytmu z obrazu będącego sumą korelacji dla poszczególnych składowych wybierane są tylko te piksele, dla których wartość maski (wyznaczonej wcześniej z reprezentacji *blobs*) jest równa 1. Spośród tych pikseli znajdowane jest K największych maksimumów, których lokalizacja jest przyjmowana jako współrzędne znalezionych markerów. Markery wykrywane są dla każdej klatki niezależnie – nie są wykorzystywane zależności czasowe, takie jak np. estymacja ruchu markera pomiędzy kolejnymi klatkami.

4. Weryfikacja znalezionych markerów

W celu weryfikacji rezultatów detekcji opracowano algorytm pozwalający na automatyczne określenie ilości poprawnie wyznaczonych markerów na kolejnych klatkach filmu. Schemat blokowy algorytmu przedstawiony został na rysunku 6.



Rys. 6. Schemat blokowy algorytmu walidacji detekcji

Wyznaczone w algorytmie detekcji współrzędne markerów porównywane są z referencyjnymi współrzędnymi markerów. Porównanie odbywa się poprzez obliczenie odległości r_i (według normy euklidesowej), a następnie określenie, które markery spełniają kryterium

$$r_i \leq R, i = 1..K \tag{5}$$

Te, które nie spełniają kryterium, uznawane są za błędnie wyznaczone, natomiast pozostałe za poprawne. Parametr R dobrany został doświadczalnie tak, aby uwzględnić maksymalne przesunięcie markera pomiędzy klatkami, wynikające z ruchu głowy lub ruchu markera wywołanego mimiką twarzy.

Ponieważ znalezione przez algorytm detekcji markery są uszeregowane na kolejnych klatkach w sposób losowy, konieczne stało się ich przyporządkowanie do odpowiadających im markerów referencyjnych. Utrudnieniem dopasowania jest fakt, iż wśród markerów znalezionych mogą znajdować się markery błędnie wyznaczone (np. znajdujące się w innej części obrazu). Problem ten można przedstawić jako liniowe zagadnienie przydziału (*linear assignment problem*).

Niech A oznacza n -elementowy zbiór markerów referencyjnych, B – n -elementowy zbiór markerów znalezionych na aktualnej klatce oraz $C: A \times B \rightarrow \mathbb{R}$ – funkcję kosztu przyporządkowania pomiędzy markerami (w tym przypadku równą odległości między nimi). Naszym celem jest znalezienie funkcji wzajemnie jednoznacznej $f: A \rightarrow B$ tak, aby zminimalizować funkcję celu:

$$\min \sum_{a \in A} C_{a, f(a)} \quad (6)$$

Inaczej mówiąc, poszukujemy najmniej kosztownego przyporządkowania każdemu elementowi ze zbioru A dokładnie jednego elementu ze zbioru B . Do rozwiązania powyższego problemu wykorzystano implementację metody *hungarian algorithm* zaproponowanej w pracy [1]²⁾.

Aby uniknąć konieczności ręcznego wyznaczania położenia markerów na wszystkich klatkach filmu, opracowany został algorytm uaktualniania współrzędnych referencyjnych w kolejnych ramkach. Uaktualnianie odbywa się w sposób następujący:

- dla pierwszej klatki współrzędne referencyjne zostały wyznaczone ręcznie,
- markery uznane za poprawnie wyznaczone przepisywane są bez zmian do następnej ramki jako markery referencyjne,
- dla markerów uznanych za błędnie wyznaczone obliczane są nowe współrzędne zgodnie ze wzorem:

$$\begin{aligned} \text{New}X_i &= \text{Ref}X_i + \frac{1}{k} \sum_{\substack{j \neq i \\ j=1..K}} (\text{Ref}X_j - X_j) \\ \text{New}Y_i &= \text{Ref}Y_i + \frac{1}{k} \sum_{\substack{j \neq i \\ j=1..K}} (\text{Ref}Y_j - Y_j) \end{aligned} \quad (7)$$

²⁾ Wykorzystana implementacja algorytmu w MATLAB-ie: (c) 2002/11/29 Nedialko Krouchev, Krouchen@physio.umontreal.ca

gdzie:

- K – liczba wszystkich markerów,
- k – liczba poprawnie wyznaczonych markerów,
- $NewX_i, NewY_i$ – nowe współrzędne i -tego markera,
- $refX_i, RefY_i$ – współrzędne i -tego markera referencyjnego,
- X_i, Y_i – współrzędne i -tego markera wyznaczonego na aktualnej klatce.

Nowe współrzędne obliczane są poprzez przesunięcie odpowiadających im markerów referencyjnych o średnie przesunięcie pomiędzy poprawnie wyznaczonymi markerami na aktualnej klatce a odpowiednimi markerami referencyjnymi.

5. Wyniki przeprowadzonych testów

Testowanie metody detekcji zostało przeprowadzone na kilku sekwencjach filmowych:

- sekwencja pierwsza – obraz z kamery analogowej, twarz widziana w całości (markery wielkości kilkunastu pikseli), wykonywana jednostka ruchowa AU1+2 (unoszenie brwi), rozdzielczość 320*240;
- sekwencja druga – obraz z kamery analogowej, powiększenie obrazu twarzy (markery składające się z większej ilości pikseli), wykonywana jednostka ruchowa AU1+2 (unoszenie brwi), rozdzielczość 320*240;
- sekwencja trzecia – obraz z kamery cyfrowej, twarz widziana w całości (markery wielkości kilkunastu pikseli), wykonywana jednostka ruchowa AU1+2 (unoszenie brwi), rozdzielczość 512*484.

Powyższe sekwencje charakteryzują się różnym stopniem trudności analizy obrazu. Metoda detekcji markerów była rozwijana przy użyciu sekwencji obrazów uznanej na podstawie wizualnej oceny za najtrudniejszą do analizy – za taką uznano sekwencję pierwszą. Markery na tych obrazach składają się tylko z kilkunastu pikseli, zaś w niektórych przypadkach bardzo trudne jest wizualne odróżnienie markera od otaczającej go skóry. Ponadto kamera analogowa wprowadza znaczne zniekształcenia barwne (w rzeczywistości niebieski marker widziany jest jako zbiór pikseli o różnych odcieniach, od białych do niebieskich).

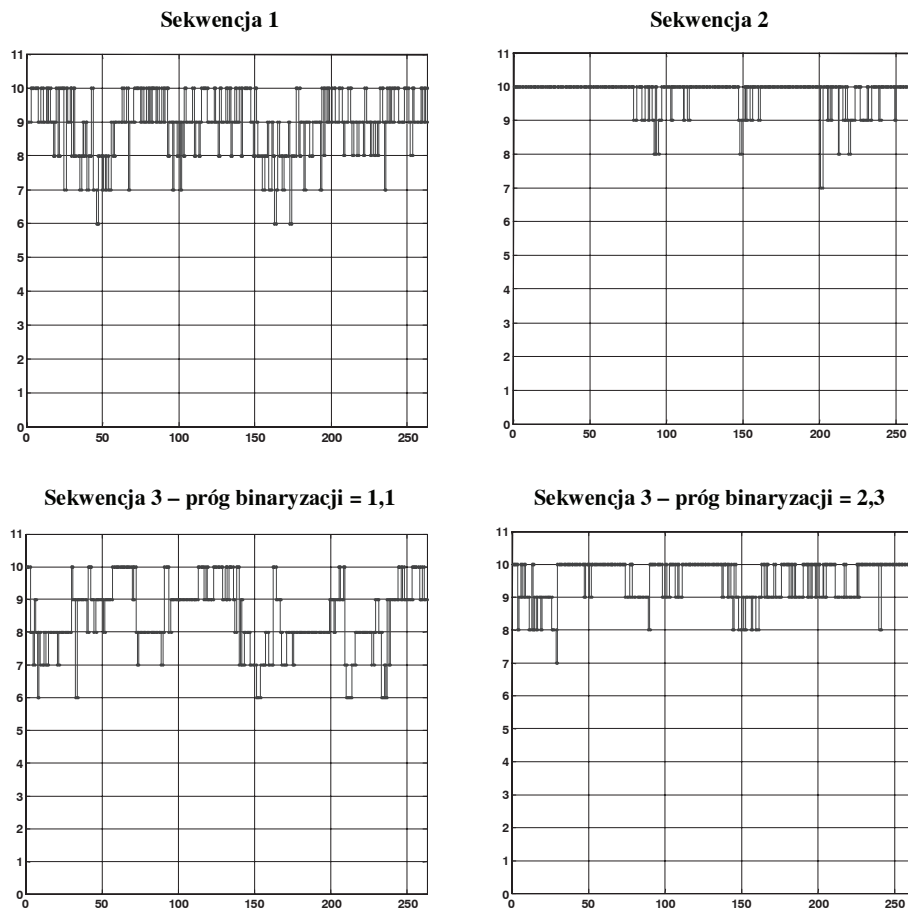
Obraz z sekwencji drugiej, dzięki powiększeniu filmowanego fragmentu twarzy, posiada większe i wyraźniej widoczne markery. Nadal występują jednak zniekształcenia barwne. Natomiast zarejestrowany w sekwencji trzeciej obraz z kamery cyfrowej nie posiada tej wady. Barwy obiektów odwzorowane są bardziej dokładnie, dzięki temu markery są lepiej odróżnialne od otoczenia. Przykładowe ramki dla każdej z sekwencji przedstawione zostały na rysunku 7.

Detekcja dla wszystkich sekwencji przeprowadzona została przy takich samych parametrach algorytmu (parametr skali – 2, próg binaryzacji – 1,1). Dla sekwencji trzeciej wykonano dodatkowo testy przy innej wartości progu binaryzacji obrazu *blobs* (równiej 2,3).

Wyniki przeprowadzonych testów reprezentujące liczbę poprawnie znalezionych markerów przedstawione są na rysunku 8.



Rys. 7. Wybrane ramki z kilku sekwencji filmowych

Rys. 8. Liczba poprawnie znalezionych markerów dla różnych sekwencji filmowych i parametrów (oś x – numer ramki, oś y – liczba poprawnie znalezionych markerów)

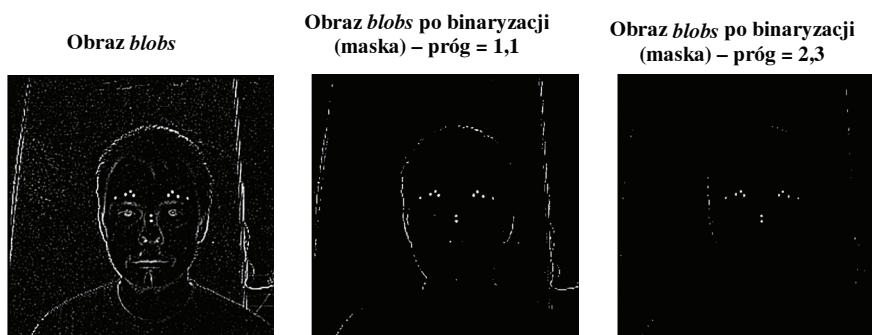
Dla sekwencji pierwszej występują częste i wyraźne zmiany w liczbie poprawnie od-
szukanych markerów. Liczba znalezionych markerów wynosi 6 na 10 (ramka 46) dla naj-
gorszego przypadku. Analizując wyniki pośrednich obliczeń (rys. 9), można zauważyć, iż
reprezentacja *blobs* zawiera wszystkie markery wraz z nadmiarowymi, błędnymi lokaliza-
cjami (jak założono podczas opracowywania algorytmu). Nasuwa się wniosek, iż przyczy-
ną błędnej detekcji jest następny etap detekcji – wyznaczanie korelacji.



Rys. 9. Wyniki pośrednich obliczeń dla sekwencji 1, klatka 46

Skuteczność algorytmu detekcji znacznie poprawia się dla sekwencji drugiej, zawiera-
jącej obraz twarzy w powiększeniu. Na zdecydowanej większości klatek markery zostały
rozpoznane poprawnie. Rezultat ten wynika z faktu, iż markery składają się z większej licz-
by pikseli, a co za tym idzie, korelacja między obrazem a szablonem markera jest większa.

Dla obrazów z kamery cyfrowej obserwujemy tylko nieznaczną poprawę wyników de-
tekcji, pomimo iż obrazy cyfrowe mają lepszą jakość. Potwierdza to tezę, iż przyczyną
błędnych rozpoznań jest słaba korelacja dla niewielkich (składających się z kilkunastu pik-
seli) markerów.



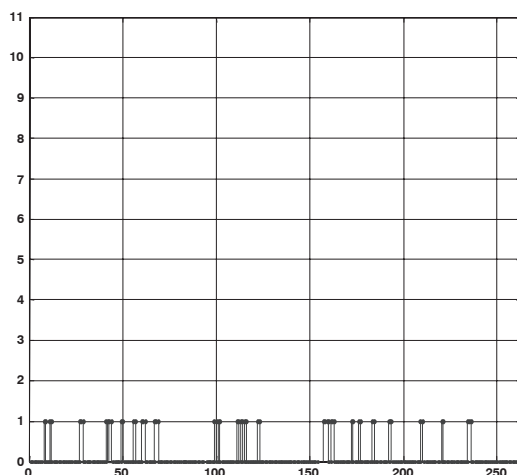
Rys. 10. Reprezentacja *blobs* oraz maska dla różnych progów binaryzacji
(klatka 1, sekwencja trzecia)

Wpływ zmian parametru progu binaryzacji reprezentacji *blobs* na liczbę poprawnych
rozpoznań jest wyraźnie obserwowalny (rys. 8). Odpowiednie dostosowanie progu pozwala
na zmniejszenie ilości potencjalnych markerów, a co za tym idzie, zwiększenie prawdopo-
dobieństwa wykrycia markerów właściwych (rys. 10). Próg ten powinien jednakże być tak do-

brany, aby wszystkie obszary zostały wykryte. W przeciwnym razie, w dalszych etapach algorytmu (korelacja) nie zostaną uwzględnione piksele obrazu należące do właściwych markerów.

Ponieważ kształt opisywany przez markery nie jest figurą sztywną (co oznacza, że w zależności od ruchów mimicznych zmieniają się odległości pomiędzy punktami kształtu), mogą zachodzić sytuacje, dla których algorytm weryfikacji nie będzie działał poprawnie. Metoda korekcji współrzędnych markerów na kolejnych klatkach nie uwzględnia opisanej wyżej sytuacji.

Aby ocenić skuteczność algorytmu weryfikacji, przeprowadzono dodatkowy test, w którym wyniki algorytmu detekcji porównywane są z ręcznie wyznaczonymi współrzędnymi markerów referencyjnych. W tym celu zmodyfikowano algorytm weryfikacji. Zamiast skorygowanych współrzędnych markerów, jako współrzędne referencyjne dla następnej ramki wykorzystywane są położenia markerów wyznaczone ręcznie. Test przeprowadzono dla sekwencji pierwszej. Jego wyniki potwierdzają skuteczność algorytmu weryfikacji. Różnice pojawiają się dla niewielu klatek (31 z 267, co odpowiada około 12%) i są rzędu jednego markera (rys. 11).



Rys. 11. Różnica pomiędzy wynikami weryfikacji z korektą automatyczną a ręcznie wyznaczanymi markerami referencyjnymi (oś x – numer ramki, oś y – liczba poprawnie znalezionych markerów)

6. Podsumowanie i wnioski

W niniejszej pracy zaproponowano algorytm automatycznej detekcji markerów na obrazie twarzy, który może zostać wykorzystany do automatycznej anotacji obrazów. Przedstawiono również metodę weryfikacji skuteczności opracowanego algorytmu oraz zamieszczono analizę wyników eksperymentów wykonanych dla kilku sekwencji wideo różniących się stopniem trudności analizy obrazu.

Przeprowadzone testy sugerują, iż najsłabszym ogniwem algorytmu detekcji jest etap wyznaczania korelacji między obrazem markera a jego szablonem. Dlatego w dalszych pracach powinien on zostać usprawniony, na przykład przez wykorzystanie osobnych szablonów dla każdego z markerów oraz uaktualnianie ich dla kolejnych ramek po wykryciu odpowiednio dużych zmian.

Innym elementem wymagającym przyszłego skorygowania jest sposób określania liczby kandydatów do dalszego przetwarzania jest stała z założenia, co w przypadku przesłonięcia na przykład jednego lub więcej autentycznych markerów spowoduje zaklasyfikowanie innych elementów obrazu jako markerów fałszywych. Wprowadzenie progów detekcji powinno rozwiązać ten problem.

Jednym z usprawnień algorytmu weryfikacji, rozważanym na etapie prac eksperymentalnych, było wykorzystanie zależności przestrzennych pomiędzy markerami, które tworzą charakterystyczny kształt. Powodem rezygnacji z takiej drogi postępowania był fakt, iż łączny kształt tworzony przez markery nie jest sztywny i zmienia się w zależności od mimiki (*non-rigid shape*). Nie wyklucza to tego kierunku badań, wymagać będzie jednak stworzenia bardziej rozbudowanego modelu kształtu poszczególnych grup markerów.

Interesującej obserwacji dostarcza połączenie algorytmu detekcji i weryfikacji w jedną całość, która daje bardziej skuteczny algorytm wykrywania markerów dla sekwencji wideo. Jest to spowodowane uwzględnieniem nie tylko informacji przestrzennych z aktualnej ramki, ale również zależności czasowych w położeniu markerów na kolejnych klatkach. Dalsze prace będą koncentrowały się na udoskonaleniu algorytmów detekcji i weryfikacji oraz dodaniu funkcjonalności automatycznego retuszu markerów.

Literatura

- [1] Ivantchev D., Negler G.: *Network Optimization*. Sofia 1992
- [2] Lanitis A., Taylor C., Cootes T.: *Automatic Interpretation and Coding of Face Images Using Flexible Models*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, no. 7, 1997, 743–756
- [3] Lindeberg T.: *Scale-space theory: A basic tool for analysing structures at different scales*. J. of Applied Statistics, vol. 21(2), 1994, 224–270
- [4] Przybyło J., Wołoszyn P., Jabłoński M.: *Rozpoznawanie jednostek czynnościowych mimiki twarzy dla potrzeb interfejsu człowiek-komputer*. Półrocznik AGH Automatyka, t. 8, z. 3, 2004, 367–379

