

**ZASTOSOWANIE METOD ANALIZY OBRAZÓW
DO AUTOMATYCZNEJ DETEKCJI I POMIARU ŻRENIC OCZU**

**METHODS OF IMAGE ANALYSIS USED FOR AUTOMATIC
DETECTION OF HUMAN EYE PUPILS**

Robert Koprowski¹, Regina Tokarczyk², Zygmunt Wróbel¹

¹Zakład Komputerowych Systemów Biomedycznych, Uniwersytet Śląski

²Zakład Fotogrametrii i Informatyki Teledetekcyjnej,
Akademia Górniczo-Hutnicza w Krakowie

SŁOWA KLUCZOWE: obraz cyfrowy, analiza obrazu, korelacja, sieci neuronowe, analiza morfologiczna

STRESZCZENIE: Potrzeby diagnostyki medycznej wad budowy i postawy człowieka wymagają stworzenia prostej metody jego pozycjonowania w trójwymiarowej przestrzeni. Jedną z metod wyznaczania położenia charakterystycznych punktów ciała jest metoda fotogrametryczna. Pomierzone punkty na zdjęciach służą do zbudowania przestrzennego modelu i za pomocą punktów dostosowania umieszczenia go w układzie odniesienia. Fotogrametryczny system do pomiaru 3D punktów ciała ludzkiego opracowany w Zakładzie Fotogrametrii i Informatyki Teledetekcyjnej WGGiIŚ AGH w Krakowie wymaga pomiaru na zdjęciach cyfrowych wybranych miejsc ciała ludzkiego, sygnalizowanych styropianowymi kulkami-markerami oraz środków źrenic oczu. W ramach automatyzacji pomiaru na zdjęciach prowadzone są badania nad metodami detekcji i pomiaru trzech grup punktów: fotopunktów, markerów na pacjencie oraz źrenic oczu. Niniejsze opracowanie dotyczy metody automatycznego pomiaru trzeciej grupy punktów pomiarowych. Opracowana metoda detekcji źrenic oczu wymaga zrealizowania celów cząstkowych: detekcji markerów na ciele pacjenta, segmentacji obrazu ciała dla wykrycia głowy z rejonem oczu, wykrycia i pomiaru środków źrenic. Pierwszy z celów osiągnięto wykorzystując metodę korelacji krzyżowej. Do wykrycia rejonu oczu wykorzystano analizę skupień zbioru współrzędnych metodą k najbliższych sąsiadów. Dalsze przybliżenie rejonu oczu osiągnięto wykorzystując analizę statystyczną dużej ilości zdjęć pacjentów dla wyznaczenia współczynnika kryterium rejonizacji. Metodyka wyznaczania położenia oczu, centralnych punktów położenia źrenic została oparta na sieciach neuronowych ze wsteczną propagacją błędów. Uzyskana dokładność metody została oceniona na poziomie około 1 piksela, jej skuteczność w znacznej mierze zależy od prawidłowej detekcji markerów na ciele pacjenta.

1. WSTĘP

Potrzeby diagnostyki w rehabilitacji medycznej wymagają stworzenia prostej metody pozycjonowania ciała ludzkiego w przestrzeni trójwymiarowej. Metody tego typu analizy są szeroko rozpowszechnione i nie stwarzają większych problemów przy ich implementacji i codziennym stosowaniu w praktyce medycznej. Jednak bez względu na sposób i metodę

analizy postawy całość zagadnienia sprowadza się do prostej wizualizacji ułożenia poszczególnych punktów charakterystycznych pacjenta w przestrzeni 3D. Ocena, zazwyczaj przemieszczenia względnego oraz asymetrii tych punktów stanowi miarodajną cechę odchylenia postawy ciała od stanu normy.

Jedną z metod wyznaczania położenia punktów charakterystycznych jest metoda fotogrametryczna, wykorzystująca dwa (lub więcej) obrazy pacjenta z naniesionymi na ciało markerami. Pomierzone punkty na zdjęciach służą do zbudowania przestrzennego modelu i za pomocą punktów dostosowania umieszczenia go w układzie odniesienia. Fotogrametryczny system do pomiaru w 3D punktów ciała ludzkiego opracowany w Zakładzie Fotogrametrii i Informatyki Teledetekcyjnej WGGiŚ AGH w Krakowie (Tokarczyk *et al.*, 1999) wymaga pomiaru na zdjęciach cyfrowych wybranych miejsc ciała ludzkiego, sygnalizowanych styropianowymi kulkami-markerami oraz źrenic oczu. Obrazy cyfrowe uzyskuje się za pomocą dwu skalibrowanych kompaktowych aparatów fotograficznych typu *free fokus* o rozdzielczości około 2M. Pomiar ma służyć do wyznaczenia wzajemnego położenia poszczególnych segmentów ciała ludzkiego w różnych pozycjach pionowej postawy, żądania dokładnościowe wyznaczenia współrzędnych mierzonych punktów są rzędu 2-3 mm.

2. CEL PRACY

Celem badań było zaproponowanie i weryfikacja własności oraz cech metrologicznych metody automatycznej analizy i detekcji punktów charakterystycznych ciała ludzkiego na podstawie obrazu w świetle widzialnym. Sformułowano dwa cele cząstkowe:

- wyznaczenie położenia punktów charakterystycznych na ciele pacjenta,
- wyznaczenie położenia oczu.

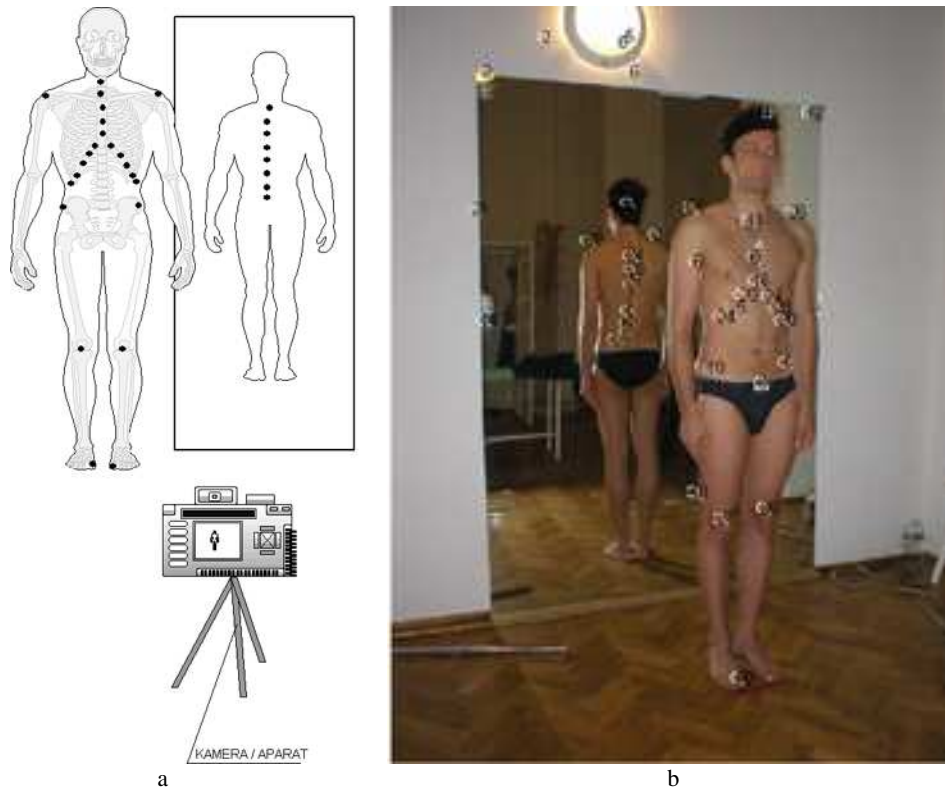
3. METODYKA PRZEPROWADZANIA POMIARÓW

Analiza i algorytm segmentacji punktów charakterystycznych jest ściśle powiązana z akwizycją obrazu. W tym przypadku najskuteczniejszą, w sensie łatwości w jej obróbce, okazała się metoda zaznaczania punktów markerami w postaci kulek styropianowych. Otrzymane przykładowe zdjęcie L_{RGB} pokazano na rysunku 1b.

Zarówno położenie jak też liczba markerów jest uzależniona od stopnia patologii oraz możliwości precyzyjnej oceny charakterystycznego miejsca na ciele pacjenta. Z tego powodu oraz ze względu na konieczność stworzenia algorytmu analizującego obraz w pełni automatycznie niezbędne okazało się przyjęcie następujących założeń:

1. Przeważająca liczba markerów w rzucie na oś Ox , za którą uważa się oś poziomą macierzy obrazu występuje w dwóch miejscach (projekcja przednia oraz odbicie pacjenta w lustrze) w miejscu położenia kręgosłupa, zakłada się jednocześnie, że boczne skrzywienie kręgosłupa nie przekracza IV stopnia.
2. Kąt ustawienia aparatu cyfrowego względem pacjenta umożliwia pozyskanie w tym samym czasie obrazu w projekcji przedniej oraz odbicia tyłu pacjenta w lustrze.

3. Wielkość markerów a tym samym rozdzielczość obrazu jest na tyle duża, że pojedynczy marker ma pole powierzchni nie mniejsze niż 4 piksele i jest koloru białego (lub w innym jednolitym kolorze).
4. Pacjent ma założoną czarną opaskę na czoło, której powierzchnia obejmuje nie mniej niż 10 % całkowitej powierzchni głowy w analizowanej projekcji.



Rys. 1. Akwizycja obrazu: (a) schemat poglądowy przebiegu wykonywania pomiaru, (b) uzyskany obraz L_{RGB} z znacznymi markerami

4. DETEKcja PUNKTÓw CHARAKTERYSTYCZNYCH

4.1. Analiza wstępna

Ze względu na charakter markerów tj. ich niezmienny kształt bez względu na kąt rzutowania (rozmięszczenie kamery) zaproponowano metodę detekcji opartą o korelacje wzajemna $L_{CORR}(m,n,k)$ (gdzie m - wiersz, kolumna, k – składowa R, G lub B), tj.:

$$L_{CORR}(m, n) = \sum_{k=1}^3 \left[\frac{\sum_{m_F=1}^{M_F} \sum_{n_F=1}^{N_F} (L_{F_RGB}(m_F, n_F, k) - \overline{L_{RGB}}) (H1(m_F, n_F) - \overline{H1})}{\sqrt{\left(\sum_{m_F=1}^{M_F} \sum_{n_F=1}^{N_F} (L_{F_RGB}(m_F, n_F, k) - \overline{L_{RGB}})^2 \right) \left(\sum_{m_F=1}^{M_F} \sum_{n_F=1}^{N_F} (H1(m_F, n_F) - \overline{H1})^2 \right)}} \right] \quad (1)$$

gdzie:

$$\overline{L_{F_RGB}} = \frac{1}{M_F \cdot N_F} \sum_{m_F=1}^{M_F} \sum_{n_F=1}^{N_F} L_{F_RGB}(m_F, n_F, k) \quad (2)$$

$$\overline{H1} = \frac{1}{M_F \cdot N_F} \sum_{m_F=1}^{M_F} \sum_{n_F=1}^{N_F} H1(m_F, n_F) \quad (3)$$

M_F, N_F - liczba wierszy i kolumn zarówno maski H1 jak też analizowanego fragmentu obrazu L_{RGB} (z definicji mająca tą samą liczbę wierszy i kolumn),
H1 - maska.

Kolejnym etapem jest progowanie otrzymanego obrazu L_{CORR} a następnie etykietywanie i obliczanie środka ciężkości obiektów. Ze względu jednak na małą rozdzielczość obrazu oraz błędy wprowadzane w czasie akwizycji markery widoczne na obrazie L_{RGB} nie są tej samej jasności, przez co algorytm nie wykrywa w praktyce ok. 20 % markerów. W związku z tym dokonano pierwszej modyfikacji polegającej na zmianie kryterium rozpoznawania markerów.

4.2. Automatyczna lokalizacja markerów

Przynosząca oczekiwane rezultaty (prawidłowe wyznaczenie ok. 99 % markerów) metoda została oparta o elementy analizy morfologicznej (Lee *et al.*, 2002; Min *et al.*, 2004; Park *et al.*, 2000). Obraz wejściowy L_{RGB} został poddany operacji dylatacji elementem kwadratowym strukturalnym o rozmiarze $M_{SE1} \times N_{SE1} = 3 \times 3$ dla każdej składowej RGB, tj. w uproszczeniu:

$$L_{ERGB}(m, n, k) = \min_{m, n \in SE1} (L_{RGB}(m, n, k)) \quad (4)$$

dla $k=1, 2$ i 3 (kolejne składowe RGB)

Następnie przeprowadzono konwersję otrzymanego obrazu L_{ERGB} do obrazu z poziomami szarości L_{GRAY} wykorzystując zależność:

$$L_{GRAY}(m, n) = [0.2989 \quad 0.5870 \quad 0.1140] \cdot \begin{bmatrix} L_{ERGB}(m, n, 1) \\ L_{ERGB}(m, n, 2) \\ L_{ERGB}(m, n, 3) \end{bmatrix} \quad (5)$$

Otrzymany obraz w poziomach szarości został dalej poddany operacji detekcji wzorców wykorzystując operację otwarcia elementem strukturalnym SE₂, tj.:

$$L_{WZRGAY}(m, n, k) = \min \left(L_{GRAY}(m, n) - \max_{m, n \in SE_2} \left(\min_{m, n \in SE_2} (L_{GRAY}(m, n)) \right) \right) \quad (6)$$

gdzie SE₂ – element strukturalny o rozmiarze $M_{SE_2} \times N_{SE_2} = 6 \times 6$ dobrany arbitralnie na podstawie liczby pikseli przedstawiających marker. Ostatnim etapem przed progowaniem wartości jest obliczenie korelacji zgodnie z zależnością (1) dla maski H2 otrzymując obraz L_{CORR} . Uzyskano wyniki w postaci obrazu L_{CORR} oraz obrazu L_{BIN} powstałego w wyniku progowania L_{CORR} . Progowanie obrazu L_{CORR} przeprowadzono zgodnie z typową metodą binaryzacji, tj.:

$$L_{BIN}(m, n) = L_{CORR}(m, n) > p_r \quad (7)$$

gdzie p_r jest progiem binaryzacji ustalonym doświadczalnie na 35. Proóg ten nie może być ustalany automatycznie wykorzystując np. zależności N Otsu czy też inne wskaźniki statystyczne, ponieważ informacje zawarte w obrazie L_{CORR} ściśle są powiązane ze stopniem dopasowania wzorca.

Utworzona macierz zawierająca współrzędne położenia markerów oraz ich identyfikatory zostaje dalej wykorzystywana do zgrubnego i dokładnego umiejscowienia obszaru położenia oczu.

Dokładność wyznaczenia markerów została wyznaczona za pomocą porównania z pomiarem manualnym i wynosi ± 1 piksel, co daje w przestrzeni trójwymiarowej przy rozdzielczości używanych aparatów dokładność wyznaczenia współrzędnych około $\pm 1-2$ mm.

5. DETEKCJA POŁOŻENIA OCZU

Algorytmy detekcji oczu na obrazach pozyskanych w świetle widzialnym są przedmiotem badań specjalistów od wielu lat (Ai *et al.*, 2002; Caetano *et al.*, 2001; Chang *et al.*, 2000; Gandhi, 2004; Gejuš *et al.*, 2003; Hjelmas *et al.*, 2001; Ma *et al.*, 2002; Sahbi *et al.*, 2002; Tian *et al.*, 2001; Tian *et al.*, 2002).

Metody te opierają się w przeważającej części na dwóch algorytmach: wykorzystujących cechy geometryczne położenia oczu względem innych części głowy (Viola *et al.*, 2004) oraz ciała (Viola *et al.*, 2001), oraz metody bazujące na cechach statystycznych (Caetano *et al.*, 2001; Jianfeng, 2004; Tzouveli *et al.*, 2005; Verschae *et al.*, 2003) i sieciach neuronowych (Zhou, 2002). Obie grupy algorytmów przynoszą oczekiwane rezultaty w ściśle określonych warunkach akwizycji obrazu (Hsu *et al.*, 2002; Hsu *et al.*, 2004; Wong *et al.*, 2003; Zhang *et al.*, 2003).

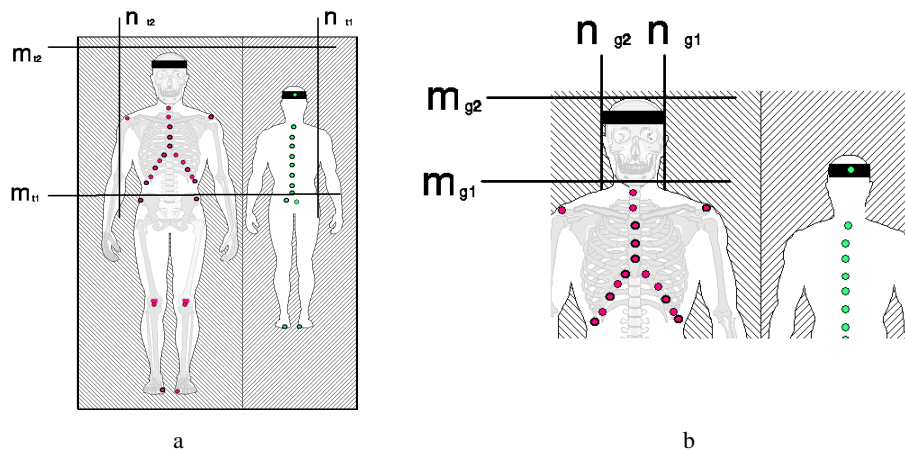
Poniżej opisano praktyczną realizację wybranego podejścia hybrydowego, wykorzystującego zarówno podejście algorytmiczne jak też, w ostatnim etapie, sieci neuronowe.

5.1. Obszar położenia oczu

Ze względu na specyfikę analizowanego obrazu, dla wyznaczenia oczu zaproponowano podejście, które wstępnie ogranicza obszar poszukiwań. W tym celu, na podstawie przedstawionych założeń, dokonano analizy skupień zbioru współrzędnych pa metodą k najbliższych sąsiadów (kNN – k-Nearest Neighbor) (Spath *et al.*, 1985). Otrzymane rezultaty pozwalają na podział na dwa zbiory pa_1 i pa_2 zawierające informacje o położeniu (współrzędnych) markerów przedniej i tylnej części ciała.

Klasyfikator z estymatorem k najbliższych sąsiadów został zastosowany dla jednowymiarowego wektora cech (dla współrzędnych x – drugiej kolumny macierzy pa). Ze względu na postać wektora cech oraz metodykę jego pozyskania przyjęto podział na dwie klasy – 2NN, gdzie $k=2$. Przyjęto metrykę euklidesową jako metodę pomiaru odległości między cechami a obliczonym położeniem środków klas (m_{k1}, n_{k1}) i (m_{k2}, n_{k2}) .

Proces uczenia rozpoczynano wstępnym wyborem położenia środków klas następującymi metodami: losową, wstępnej klasyfikacji pochłaniającej 10 % czasu klasyfikacji całkowitej oraz metodą losowego wstępnego doboru położenia środków klas. Otrzymane wyniki klasyfikacji pokazano na rysunku 2.



Rys. 2. Obraz wejściowy L_{RGB} z zaznaczonym wynikiem klasyfikacji metoda kNN oraz obliczony na tej podstawie obszar zawierający górną część tułowia

Na tej podstawie oraz przyjętych założeń dotyczących konieczności występowania przeważającej części punktów położonych w dwóch obszarach w rzucie na oś Ox (miarodanych w stosunku do położenia kręgosłupa względem tej osi) wyznaczono obszar górnej części tułowia, będący funkcją stałej c_t , przyjętej na podstawie analizy ok. 1000 obrazów na poziomie 0.2.

Kolejnym etapem jest lokalne przybliżenie położenia oczu, które wykorzystuje przedstawione powyżej obliczenia oraz jedno z założeń związane z faktem występowania czarnej opaski.

Opaska została jednoznacznie oddzielona od pozostałej treści obrazu na podstawie kryterium uwzględniającego następujące cechy:

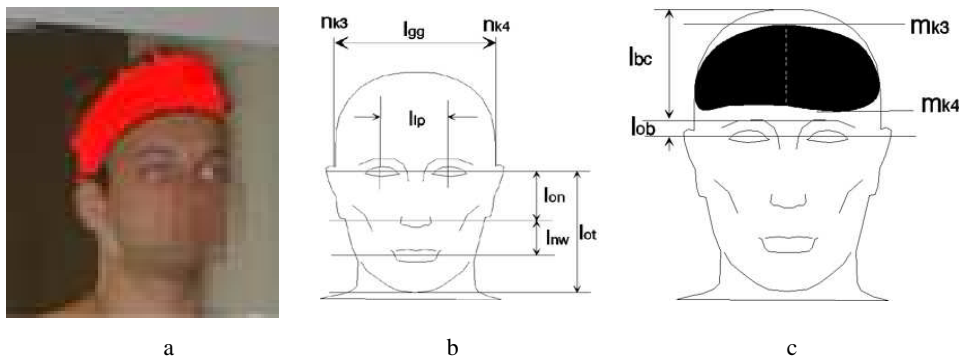
- kolor czarny,
- powierzchnia przeważająca (być może największa na scenie),
- niewielki współczynnik kształtu.

Cechy te umożliwiły konstrukcje kryterium Q w postaci:

$$Q^{(i)} = \frac{\sqrt{P^{(i)}}}{m^{(i)} \cdot (P_k^{(i)})^{\frac{1}{c_q}}} \quad (8)$$

gdzie :

- $P^{(i)}$ - pole powierzchni kolejnych obiektów,
- c_q - stała równa 6,
- $P_k^{(i)}$ - pole powiedzchni konturu .



Rys. 3. Fragment obrazu L_{RGB} obejmujący zakres przybliżenia lokalnego położenia głowy (a) oraz rysunki proporcji między poszczególnymi częściami twarzy (b) i (c)

Maksymalna wartość kryterium Q występuje dla wszystkich 1000 przypadków dla opaski pacjenta widocznej w projekcji przedniej. Przedstawione zatem kryterium jest wystarczające i w wyniku jego stosowania otrzymuje się poprawne wyniki dla wszystkich analizowanych obrazów.

Zgodnie z przyjętą symboliką przedstawiona na rysunku 3b i 3c można przedstawić proporcje głowy w sposób, który prezentuje tabela 1.

Tabela. 1. Średnia proporcja charakterystycznych punktów na głowie

Proporcja	5c	50c	95c	średnia±std
l_{ob}/l_{ip}	0.20	0.21	0.18	0.19 ± 0.02
l_{gg}/l_{ip}	2.54	2.41	2.34	2.43 ± 0.1
l_{ot}/l_{ip}	1.91	1.87	1.90	1.89 ± 0.02
l_{bc}/l_{ip}	1.55	1.51	1.50	1.52 ± 0.03

gdzie: std – odchylenie standardowe średniej (*standard deviation*).

Na podstawie przedstawionych obliczeń i proporcji można oszacować przedział, w którym są położone oczy pacjenta.

Wynik tego etapu przybliżenia położenia oczu na scenie jest wystarczający do zastosowania ostatniego etapu opisywanego algorytmu, wyznaczającego dokładnie położenie oczu.

5.2. Wyznaczanie położenia oczu

Metodyka wyznaczania położenia oczu, centralnych punktów położenia źrenic oczu patrzących na daleki horyzont (lub centralnych punktów gałek ocznych dla pacjentów z zezem), została oparta na sieciach neuronowych ze wsteczną propagacją błędów. Dla obrazu monochromatycznego zawierającego głowę $L_{GRAY}=L_{RGB}(m_{g1}:m_{g2},n_{g1}:n_{g2})$ została przeprowadzona operacja normalizacji obrazu do pełnego przedziału wartości nasycenia.

Następnie został wyznaczony gradient L_{GRADY} w osi Oy obrazu $L_{GRAY} \left(\frac{\partial L_{GRAY}}{\partial y} \right)$. Wybór tej osi jest właściwy ze względu na cechy charakterystyczne konturu oka najlepiej widoczne dla wspomnianej osi Oy .

Ciąg uczący stworzono zaznaczając ręcznie obszary o rozdzielczości $(M_o/4) \times (N_o/4)$ stanowiące stały rozmiar w stosunku do rozmiaru obrazu L_{GRADY} głowy zawierający oko lewe i prawe. Operator wskazując środek położenia oczu jako parę współrzędnych m_l, n_l dla oka lewego i odpowiednio m_p, n_p dla oka prawego tworzył dwa fragmenty $L_{GRADY}^{(l)}$ i $L_{GRADY}^{(p)}$ obrazu L_{GRADY} jako:

$$L_{GRADY}^{(l)} \left(m - m_l + \frac{M_o}{8} + 1, n - n_l + \frac{N_o}{8} + 1 \right) = L_{GRADY}(m, n) \quad \text{dla} \quad \begin{cases} m \in \left(m_l - \frac{M_o}{8}, m_l + \frac{M_o}{8} \right) \\ n \in \left(n_l - \frac{N_o}{8}, n_l + \frac{N_o}{8} \right) \end{cases} \quad (9)$$

oraz analogicznie dla oka prawego:

$$L_{GRADY}^{(p)} \left(m - m_p + \frac{M_o}{8} + 1, n - n_p + \frac{N_o}{8} + 1 \right) = L_{GRADY}(m, n) \quad \text{dla} \quad \begin{cases} m \in \left(m_p - \frac{M_o}{8}, m_p + \frac{M_o}{8} \right) \\ n \in \left(n_p - \frac{N_o}{8}, n_p + \frac{N_o}{8} \right) \end{cases} \quad (10)$$

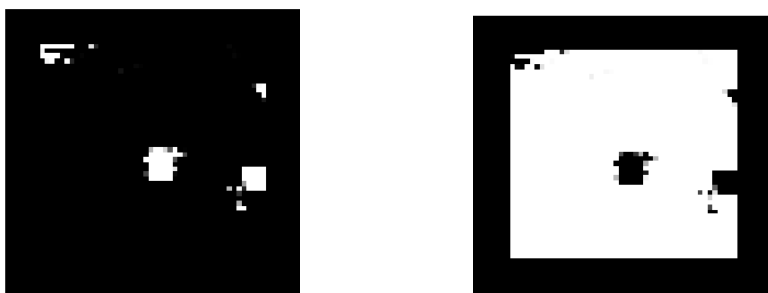
Możliwe zmiany położenia obrazu oka w $L_{GRADY}^{(l)}$ i $L_{GRADY}^{(p)}$ wyeliminowano tworząc dodatkowe obrazy $L_{GRADY}^{(i)}$ w promieniu ± 2 piksele w osi Oy oraz ± 2 piksele w osi Ox eliminując te obszary, które mogłyby wychodzić poza obraz $L_{GRADY}^{(i)}$. Powstałe w ten sposób fragmenty obrazu L_{GRAY} nazwano dalej jako $L_{GRADY}^{(i)}$ gdzie i oznacza numer fragmentu obrazu.

Przyjmując zatem przykładowo, że obraz wejściowy L_{GRADY} jest o rozdzielczości 100×80 pikseli oraz że operator zaznaczył oko prawe i lewe tworząc tym samym parę m_l, n_l i m_p, n_p , otrzymuje się 25 obrazów o rozdzielczości $M_o \times N_o = 25 \times 20$ pikseli.

Na podstawie tego przykładu widać, że rozmiar wektora uczącego zawierającego oko jest znacznie mniejszy (zawiera ok. 7% pikseli obrazu wejściowego) od wektora nie zawierającego oka. Ponieważ najlepsze rezultaty są otrzymywane przy uczeniu sieci ciągiem uczącym zawierającym tyle samo obrazów przedstawiających oko ludzkie, co ich nie przedstawiających, to ciąg uczący zmodyfikowano następująco:

- utworzono zestaw obrazów $L_{GRADY}^{(i)}$ zawierających oko (ok. 25 obrazów dla każdego pacjenta),
- utworzono zestaw obrazów $L_{GRADY}^{(i)}$ nie zawierających oka ludzkiego. Losowo wybrano obrazy nie zawierające oka tak, by ich liczba była zgodna z liczbą obrazów zawierających oko,
- utworzono ciąg wyjściowy dla dwóch wyjść takich, że jedno z wyjść wskazywało na obecność oka (0-fałsz, 1-prawda) natomiast drugie wskazywało na brak oka w obrazie $L_{GRADY}^{(i)}$,
- uczono sieć neuronową losując dla pierwszej iteracji wartości wag,
- przeprowadzono 3000 (lub testowo 5 000) iteracji,
- testując sieć wybrano te przypadki z ciągu uczącego, które sieć rozpoznała nieprawidłowo,
- utworzono zestaw obrazów nie zawierających oka z obrazów nieprawidłowo rozpoznanych, dobierając brakującą resztę z pozostałych obrazów jeszcze nie biorących udziału w uczeniu,
- przeprowadzono taki proces 60-cio krotnie (czyli sumarycznie 3000×60 epok) modyfikując co 3 000 iteracji ciąg uczący.

Przedstawione zależności wykorzystano przy tworzeniu ciągu uczącego dla sieci ze wsteczną propagacją błędów, zawierającą odpowiednio od 10 do 100 neuronów w warstwie ukrytej typu 'logsig' oraz jedną, dwie lub trzy warstwy ukryte.



Rys. 4. Przykładowe obrazy wyjściowe z sieci neuronowych; po prawej wskazujące na położenie oka, po lewej niewskazujące na położenia oka

Korekta położenia oczu obejmowała wykorzystanie operacji indeksacji. Każdy j -ty obiekt (będący wydzielonym obiektem binarnym w wyniku indeksacji) oznaczono dalej jako $L_{BIN}^{(i)}$. Następnie pozostawiono wszystkie te obiekty, które posiadały największą

powierzchnię. Otrzymane wartości j' i j'' pokazano na obrazach L_{GRAY} w formie nałożonych $L_{BIN}^{(j')}$, $L_{BIN}^{(j'')}$ oraz ich środki ciężkości $m_s^{j'}$, $n_s^{j'}$ oraz $m_s^{j''}$, $n_s^{j''}$.



Rys. 5. Przykładowe obrazy będące wynikami omawianych etapów działania algorytmu

Dla sieci neuronowej z 20 neuronami w warstwie ukrytej uzyskano 10 % błędnych rozpoznań dla 20 obrazów L_{GRAY} poddawanych procesowi rozpoznania. By zminimalizować ilość błędnych rozpoznań można zdefiniować kryterium jakości rozpoznania, którego miarą dla danego fragmentu $L_{GRAY}^{(i)}$ jest największa znormalizowana powierzchnia $L_{BIN}^{(i)}$ i największa wartość wyjściowa z sieci. W praktyce dla 20 obrazów L_{GRAY} uzyskano ok. 3 % błędnych rozpoznań. Jakość uzyskanych rozpoznań dla różnych ilości neuronów w warstwie ukrytej nieznacznie wpływa na wynik (rys. 4 i 5).

Z analizy błędu uczenia sieci oraz zgodnie z literaturą (Lee, 2002) wynika, iż najbardziej optymalne rezultaty detekcji uzyskuje się dla $k=2000$ iteracji przy 10 neuronach w warstwie ukrytej.

6. PODSUMOWANIE

Wykorzystane metody przetwarzania i analizy obrazów pozwalają na wykrycie i pomiar źrenic oczu ludzkich. Automatyczna detekcja źrenic oczu w pierwszym etapie bazuje na wykrytych metodą korelacji krzyżowej markerach na ciele pacjenta, które również są punktami pomiarowymi systemu. Służą one do segmentacji obrazu ciała dla znalezienia obszaru głowy, wykorzystując analizę skupień zbioru współrzędnych metodą k najbliższych sąsiadów. Dalsze przybliżenie rejonu oczu osiągnięto wykorzystując analizę statystyczną dużej ilości zdjęć pacjentów dla wyznaczenia współczynnika kryterium rejonizacji. Metodyka wyznaczania położenia oczu, centralnych punktów położenia źrenic została oparta na sieciach neuronowych ze wsteczną propagacją błędów. Warunkami prawidłowo działającej detekcji źrenic są prawidłowo wykryte markery na ciele oraz istnienie czarnej opaski na głowie pacjenta.

Dokładność pomiaru oczu na obrazach jest na poziomie jednego piksela i pozwala na wyznaczenie przestrzennych współrzędnych z błędem około 2-3 mm. Przedstawiona metodyka i tok postępowania znajdzie praktyczne zastosowanie przy przeprowadzanych pomiarach stopnia wady postawy.

7. LITERATURA

- Ai H., Ying L., Xu G., 2002. A Subspace Approach to Face Detection with Support Vector machines. *Proc. IEEE Int'l Conf. on Pattern Recognition*.
- Caetano T. S., Barone D.A.C., 2001. A Probabilistic Model for the Human Skin Color. *Proceedings of ICIAP2001 - IEEE International Conference on Image Analysis and Processing*, Palermo, Italy, s. 279-283.
- Chang H., Robes U., 2000. Face detection.
<http://www-cs-students.stanford.edu/~robles/ee368/main.html>
- Gandhi M., 2004. A Method for Automatic Synthesis of Aged Human Facial Images, Master Thesis. McGill University.
- Gejuš P., Šperka M., 2003. Face tracking in color video sequences. *Proceedings of SCCG 2003*, Budmerice, Slovakia, s. 268-273.
- Hjelmas E., Low B.K., 2001. Face Detection. *A Survey, CVIU*, 83, s. 236-274.
- Hsu R., Abdel-Mottaleb M., Ani1 K., 2002. Face Detection in Color Images. *IEEE transactions on pattern analysis and machine intelligence*, Vol. 24, NOS, s. 696-706.
- Hsu R.L., Jain A.K., 2002. Face Detection in Color Images. *IEEE transactions on Pattern analysis and machine intelligence*, Vol. 24, No.1. s. 34-58 oraz, Vol. 24, No.5, s. 696-718.
- Hsu R.L., Jain A.K., 2004. Image Processing. *ECCV's 2002 Workshop on Biometric Authentication*. s. 597-600.
- Jianfeng R., Lei G., 2004. Applying mixed Gaussian skin models to the automatic face detection. *IEEE Workshop on Machine Learning for Signal Processing*.
- Kadoury S., 2005. Face Detection Using Locally Linear Embedding, a thesis. Department of Electrical and Computer Engineering McGill University, Montreal, Canada.
- Koprowski R., Wróbel Z., 2006. Właściwości wybranych typów sieci neuronowych wykorzystywanych do detekcji oczu pacjenta. *Archiwum Fotogrametrii, Kartografii i Teledetekcji*, Vol. 16 (w druku).

- Lee J.Y., Yoo S.I., 2002. An Elliptical Boundary Model for Skin Color Detection. The 2002 International Conference on Imaging Science, Systems, and Technology, Las Vegas, USA.
- Ma Y., Ding X., 2002. Face Detection Based on Hierarchical Support Vector Machines, Proc. IEEE Int'l Conf. on Pattern Recognition.
- Min J., Flynn P., Bowyer K., 2004. Assessment of time dependency in face recognition. TR-04-12, University of Notre Dame.
- Park M., Park Ch., 2000. Algorithm for detecting human faces based on convex-hull, Yonsei University 134, Shinchon-dong, Seodemun-gu, Seoul, Korea Division of Information Communication Engineering, Paichai University 439-6, Doma Dong Seo-ku, Taejon, 302-735, Korea Recognition, Vol. 33, no 9, s. 1525-1539.
- Sahbi H., Boujemaa, N., 2002. Coarse-to-fine face detection based on skin color adaption. Spatial Histogram Features for Face Detection in Color Images. Hongming Zhang and Debin Zhao Department of Computer Science and Engineering, Harbin Institute of Technology.
- Spath H., 1985. Cluster Dissection and Analysis: Theory, FORTRAN Programs, Examples. translated by J. Goldschmidt, *Halsted Press*, New York, , s. 226.
- Tian Y., Kanade L., Cohn J. F., 2001. Recognizing Facial Actions by Combining Geometric Features and Regional Appearance Patterns. CMU-RI-TR-01-0, Robotics Institute, CMU.
- Tian Y., Kanade T., Cohn J. F., 2002. Evaluation of Gabor-Wavelets Based Facial Action Unit Recognition in Image Sequences of Increasing Complexity, Proc. of the 5th Int. Conf. On Automatic Face and Gesture Recognition, s. 229-234.
- Tokarczyk R., Mikrut S., 1999. Fotogrametryczny system cyfrowy bliskiego zasięgu do pomiarów ciała ludzkiego dla potrzeb rehabilitacji leczniczej. *Archiwum Fotogrametrii, Kartografii i Teledetekcji*, Vol. 9, s. 147-154.
- Tokarczyk R., 2005. Fotogrametria cyfrowa w zastosowaniach medycznych do pomiaru ciała ludzkiego - przegląd i tendencje rozwojowe systemów pomiarowych. *Geodezja, Kartografija i Aerofotogramannia, Mizbidomczij Nauk.-Technicznij Zbornik*, Wipusk 66.
- Tzouveli P., Ntalianis K., Tsapatsoulis N., Kollias S., 2005. Automatic Face Region Watermarking Using Qualified Significant Wavelet Trees. Electrical And Computer Engineering Department, National Technical University Of Athens, Athens, Greece.
- Verschae R. Ruiz-del-Solar J., 2003. A Hybrid Face Detector based on an Asymmetrical Adaboost Cascade Detector and a Wavelet-Bayesian-Detector. Int'l Work-Conf. on Artificial and Natural Neural Networks, s. 742-749.
- Viola P., Jones M., 2001. Rapid Object Detection using a Boosted Cascade of Simple Features. Proc. Of CVPR, Vol. I, s. 511-518.
- Viola P., Jones M., 2004. Robust Real-Time Face Detection. *Int. J. of Computer Vision*, Vol. 57, No. 2, s. 137-154.
- Wong K.-W. , Lam K.-M., Siu W.-C., 2003. A robust scheme for live detection of human faces in colour image. *Signal Process.: Image Commun.*, 18 (2) s. 103-114.
- Zhang L., Samaras D., 2003. Face Recognition under Variable Lighting using Harmonic Image Exemplars, roc. of CVPR.
- Zhou S., Krueger V., Chellappa R., 2002. Face Recognition from Video: A Condensation Approach. Proc. of the 5th IEEE Int. Conf. on Automatic Face and Gesture Recognition, s. 221-226.

Opracowanie wykonano w ramach grantu KBN nr 4 T12E 05227.

METHODS OF IMAGE ANALYSIS USED FOR AUTOMATIC DETECTION OF HUMAN EYE PUPILS

KEY WORDS: digital image, image analysis, correlation, neuron networks, morphological analysis

Summary

The needs of medical rehabilitation requires the creation of a simple method for the positioning of a human body in 3-D space. The method of posture analysis consists in simple visualization of particular characteristic points of the patient's body. Usually, an assessment of the relative displacement and asymmetry of these points are a reliable measure of pathological body posture deviation.

One of the methods of determining the position of characteristic points is the photogrammetric method. The points measured in the images are used to create a spatial model and, based on control points, position it in a reference system. The photogrammetric system for measuring 3D points located on a human body has been developed in the Department of Photogrammetry and Remote Sensing Informatics, Technical University – AGH Kraków. In the system, on digital images (taken with two compact, free-focus-type digital cameras), special markers (light foamed polystyrene balls) located on chosen points of human body and the center of an eye pupil are measured. To make the system more automatic, the research is carried on to automatically detect and measure the three following groups of points: control points, markers located on the human body and eye pupil.

In this paper, the method of automatic measurement of last group of points is considered. The developed method of eye pupil detection requires achieving the following goals: detection of markers located on the human body, segmentation of the body image to detect a head, approximate eye localization and, finally, the measurement of a pupil center. The first goal was achieved by using the cross correlation method. The localization of the eyes was done by analyzing the coordinates' set concentration by k nearest neighbors. The achieved results were divided into two sets pa_1 and pa_2 , which consist of data on the position of markers located on the front and back sides of body. A further approximation of eye location was achieved using statistical analysis of many images to determine the coefficient as a criterion for region membership. The methodology of determination of eye position, and pupil center were based on a neural network with backward error propagation. The achieved accuracy was estimated as 1 pixel, but the efficiency strongly depends on the proper detection of the markers located on the patient's body.

Dr inż. Robert Koprowski
e-mail: koprow@us.edu.pl
tel. +32 2918381 w. 741

Dr inż. Regina Tokarczyk
tokarcz@agh.edu.pl
tel.+12 617 22 88

Prof. Dr hab inż. Zygmunt Wróbel
e-mail: wrobel@us.edu.pl
tel. +32 2918381 w. 736