

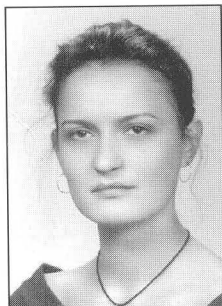
Tomasz KAPUŚCIŃSKI, Joanna MARNIK, Marian WYSOCKI
POLITECHNIKA RZESZOWSKA, KATEDRA INFORMATYKI I AUTOMATYKI

Rozpoznawanie gestów rąk w układzie wizyjnym

dr inż. Joanna MARNIK

W roku 2002 uzyskała doktorat w zakresie Informatyki w Akademii Górniczo-Hutniczej w Krakowie. Interesuje się przetwarzaniem i rozpoznawaniem obrazów cyfrowych ze szczególnym uwzględnieniem metod morfologicznej i sieci neuronowych.

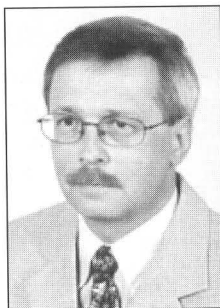
e-mail: jmarnik@prz-rzeszow.pl



dr hab. inż. Marian WYSOCKI

Uzyskał habilitację w zakresie automatyki i robotyki w roku 1990 na Wydziale Elektrotechniki Automatyki i Elektroniki Akademii Górniczo-Hutniczej w Krakowie. Jego zainteresowania dotyczą teorii sterowania i optymalizacji, obliczeń równoległych i wizji komputerowej.

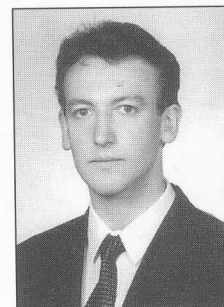
e-mail: mwysocki@prz-rzeszow.pl



mgr inż. Tomasz KAPUŚCIŃSKI

W 1997 ukończył studia na specjalności Automatyka i Informatyka Politechniki Rzeszowskiej. Interesuje się metodami przetwarzania obrazów i zastosowaniami wizji komputerowej.

e-mail: tomekkap@prz-rzeszow.pl

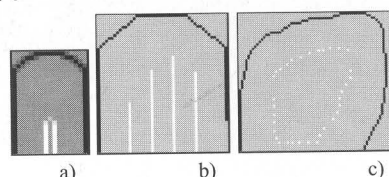


2. Metoda analizy kształtu dłoni

Do rozwiązania zadania rozpoznania kształtu dłoni stosowano w literaturze rozmaite metody, jak np. wykorzystanie cech geometrycznych [20], momentów Zernike [5], składowych głównych [10], elastycznych grafów i filtrów Gabora [18], histogramów orientacji [2], metody Fouriera zastosowanej do brzegu obszaru dłoni [8], modelu dłoni [13].

W niniejszej pracy zastosowano własne rozwiązanie [9], oparte na morfologicznym przekształceniu trafi - nie trafi [17], służącym do wykrywania cech geometrycznych obiektów binarnych. Aby umożliwić analizę kształtów dłoni metodą przekształcenia trafi - nie trafi, najpierw dokonywana jest ekstrakcja dłoni. Autorzy wykorzystywali model barwy skóry w formie dwuwymiarowego histogramu lub jego aproksymacji rozkładem Gaussa w znormalizowanej przestrzeni barw lub też sieć neuronową, na wejście której są podawane wartości składowych R, G, B pikseli obrazu kolorowego. Szczegóły metod i wyniki eksperymentów zostały podane w pracach [6, 9]. Podstawowe znaczenie w analizie kształtu dłoni ma utworzenie bazy elementów strukturalnych. Elementy, które wykorzystywano w eksperymentach omówionych w rozdziale 3. przedstawiono na rys. 1. Kolorem białym zaznaczone są części, które powinny zawierać się w obiekcie (trafi), kolorem czarnym części, które powinny znajdować się w tle (nie trafi), natomiast kolorem szarym obszary, w których wartość pikseli analizowanego obrazu nie ma wpływu na wynik transformacji. Zdefiniowanie takiego obszaru „neutralnego“ jest konieczne ze względu na różnice w proporcjach pomiędzy poszczególnymi fragmentami dłoni u różnych osób, wahania odległości od kamery oraz błędy segmentacji.

Transformacja trafi - nie trafi z zadaniem elementem strukturalnym wykrywa elementy obrazu o postaci zakodowanej w tym elemencie. Do bazy elementów strukturalnych dołączono wszystkie elementy strukturalne z rys. 1. obrócone o odpowiednie kąty [9]. Elementy pogrupowano, przypisując grupom całkowitoliczbowy kod, odzwierciedlający wagę kształtu reprezentowanego przez elementy grupy w zestawie rozpoznawanych znaków. Im bardziej specyficzny jest dany kształt, tym wyższa jest jego waga, a elementy wykrywające podobne kształty mają zbliżone wartości kodów.



Rys. 1. Elementy strukturalne do wykrywania: a) pojedynczego palca, b) czterech złączonych palców, c) zaciśniętej pięści

Fig. 1. Structural elements for detection of: a) single finger, b) four joint fingers, c) clenched fist

Streszczenie

Przedstawiono wyniki prac autorów, dotyczące rozpoznawania kształtu dłoni, wskazań obiektów oraz gestów dynamicznych, przydatnych w robotyce.

Abstract

The paper considers vision-based hand gesture interfaces. Results obtained by the authors, concerning recognition of hand posture, pointing direction, and some dynamic gestures useful in robotics are presented.

Słowa kluczowe: wizja komputerowa, komunikacja człowiek-maszyna, rozpoznawanie gestów rąk.

Keywords: computer vision, human-computer interface, hand gesture recognition.

1. Wstęp

W ostatnich latach duże zainteresowanie wzbudzają komputerowe metody rozpoznawania gestów wykonywanych przez ludzi [3,12]. Rozpoznawanie przez maszyny znaków dawanych za pomocą ręki, ruchów ciała, czy też wyrazu twarzy stanowi atrakcyjne uzupełnienie sposobów komunikacji osób z komputerami lub robotami. Znaki wykonywane rękami należą do podstawowych gestów człowieka. Systemy ich interpretacji mają szereg istotnych potencjalnych zastosowań np.: rehabilitacja i ułatwienie pracy oraz kontaktu z otoczeniem osobom niepełnosprawnym i w podeszłym wieku, zdalne sterowanie maszynami jako uzupełnienie komunikacji werbalnej, poprawa warunków pracy z komputerem, działanie człowieka w środowisku wirtualnej rzeczywistości.

W pracy zostały przedstawione kierunki badań autorów i wybrane, uzyskane przez nich rezultaty, dotyczące rozpoznawania gestów rąk w układzie wizyjnym. W rozdziałach 2. i 3. omówiono kolejno metodykę rozpoznawania kształtu dłoni i wyniki eksperymentów. Rozdział 4. poświęcono gestom dynamicznym. Skupiono się na ujęciu bliższym automatyce i robotyce. W podsumowaniu zasygnalizowano prowadzone równoległe badania związane z rozpoznawaniem języka migowego.

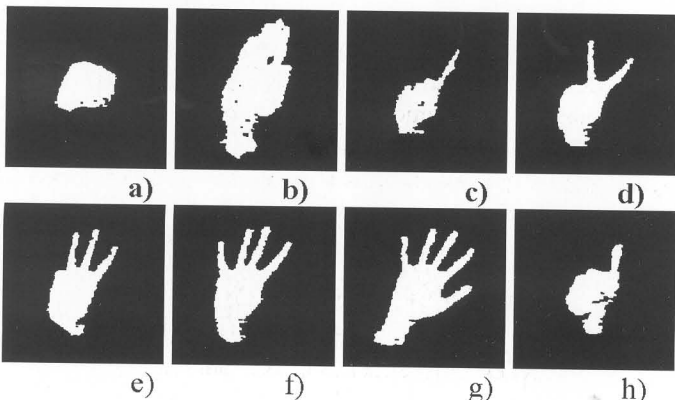
Poza kształtem zakodowanym w elemencie strukturalnym ważny jest także jego rozmiar. Przyjęto więc, że obraz dłoni będzie skalowany. Wartość skali ustala się na podstawie informacji o długości średnicy maksymalnego koła zawartego całkowicie w obiekcie reprezentującym dłoń z wyprostowanymi wszystkimi palcami. Taki kształt dłoni, pokazywany na początku sesji stanowi rodzaj przedstawienia się. Wykorzystując układ stereowizyjny można dokonać przeskalowania na podstawie mapy głębi [1].

Analiza kształtu dłoni polega na wykonaniu transformacji trafi - nie trafi badanego obrazu z użyciem każdego elementu strukturalnego z bazy. Uprzednio z obrazu wycinany jest prostokątny fragment o założonym rozmiarze, zawierający dłoń.

W wektorze cech są zapisywane informacje o fragmentach dłoni wykrytych metodą trafi - nie trafi. Analizie podlegają wszystkie obiekty powstałe w wyniku transformacji. Dla każdego takiego obiektu są wyznaczane następujące parametry: (i) powierzchnia, (ii) kąt nachylenia półprostej, mającej początek w środku ciężkości dłoni i przechodzącej przez środek ciężkości analizowanego obiektu, do poziomu, (iii) odległość środka ciężkości analizowanego obiektu od środka ciężkości dłoni. Powyższe parametry wraz z kodem liczbowym elementu strukturalnego, z użyciem którego powstał obiekt im odpowiadający, stanowią czwórkę liczb opisujących pewien fragment dłoni. Wstępnie przyjęto, że wektor cech składa się z pięciu takich czwórek zapisanych w kolejności malejących kodów. Na podstawie obszernych zbiorów danych, które przygotowano do testowania metody, stwierdzono jednak zasadność redukcji wektora cech. Wskazywała na to analiza macierzy kowariancji i współczynników korelacji wektora cech, a także szczegółowa analiza wrażliwości, którą umożliwiła opcja Sensitivity Analysis pakietu STATISTICA NN [15]. Pominięto elementy wektora wejściowego, których wpływ na skuteczność rozpoznawania przez sieć neuronową okazał się niewielki. Ostatecznie przyjęto do dalszych badań jedenastelementowy wektor cech złożony z: kodów czwórek 1, ..., 5, powierzchni z czwórek 1 i 2, kątów z czwórek 1, 2 i 4 oraz odległości z czwórki 1. Oczywiście nie zawsze zostaną wykryte obiekty ze wszystkich grup. Wówczas odpowiednie elementy wektora cech są wypełniane wartościami -1.

3. Rozpoznawanie kształtu dłoni

Przetestowano dwie metody rozpoznawania gestów opisanych z użyciem morfologicznej transformacji trafi - nie trafi. Pierwsza polegała na skonstruowaniu wielowarstwowej sieci perceptronowej, druga natomiast na zbudowaniu binarnego drzewa klasyfikacyjnego. Badania przeprowadzono dla zestawu ośmiu prostych gestów przedstawionych na rys. 2, przydatnych np. w komunikacji z robotem.



Rys. 2. Rozpoznawane gesty: a) pięść, b) dłoń, c) 1 palec, d) 2 palce, e) 3 palce, f) 4 palce, g) 5 palców, h) kciuk
Fig. 2. Recognized gestures: a) fist, b) palm, c) 1 finger, d) 2 fingers, e) 3 fingers, f) 4 fingers, g) 5 fingers, h) thumb

Wykorzystano 4790 obrazów binarnych, otrzymanych z obrazów kolorowych pozyskanych przy oświetleniu dziennym. 776 obrazów przedstawiało gest kciuk, 541 - 5 palców, 616 - 4 palce, 591 - 3 palce, 526 - 2 palce, 651 - 1 palec, 501 - dłoń i 588 - pięść. Dodatkowo wszystkie obrazy obrócono o losowy kąt z przedziału od -20 do 20 stopni, otrzymując w ten sposób drugą grupę 4790 obrazów. Do eksperymentów wzięto także trzecią grupę 4790 obrazów, otrzymanych z grupy pierwszej poprzez zakłócenie szumem typu sól i pieprz.

Użyto sieć neuronową o logistycznych funkcjach aktywacji. Jej wejścia stanowiło 11 elementów wektora cech. Miała ponadto 25 neuronów w warstwie ukrytej i 8 w wyjściowej. Była uczona na bazie 1696 wektorów cech, po 212 dla każdego gestu, z czego 101 elementów zbioru uczącego pochodziło z pierwszego zestawu testowego, 100 z zestawu drugiego (obrazy obrócone) i 11 z trzeciego (obrazy zakłócone). Ze zbioru uczącego wydzielono 150-elementowy zbiór weryfikujący. Uczenie przebiegało metodą wstecznej propagacji błędów (50 epok) i gradientów sprzężonych (128 epok). W wyniku otrzymano skuteczność rozpoznawania wynoszącą 98 % dla zbioru uczącego, 99 % dla zbioru weryfikującego i 96,4 % dla zbioru testowego, utworzonego z pozostałych obrazów. Bliższe szczegóły dotyczące rozpoznawania gestów ze zbioru testowego podano w tabeli 1.

Drzewo klasyfikacyjne skonstruowano wykorzystując moduł Drzewa klasyfikacyjne pakietu STATISTICA. Zastosowano metodę CART wyczerpującego poszukiwania podziałów jednowymiarowych. Dobroć dopasowania określano za pomocą miary Giniego. Jako regułę zatrzymania podziału wybrano opcję przytnij przy błędzie złej klasyfikacji, przy której dzielenie drzewa jest kontynuowane do momentu, gdy wszystkie węzły są czyste (zawierają obiekty tylko z jednej klasy) lub zawierają nie więcej niż określoną minimalną liczbę obiektów. Liczbę tę ustawiono na 5. Wielkość drzewa była ustalana w oparciu o V-krotny sprawdzian krzyżowy. Przedstawione wyniki otrzymano dla $V = 3$. Zwiększenie liczby podziałów nie dało znaczącej poprawy. Drzewo klasyfikacyjne wygenerowane na bazie tego samego zbioru uczącego jak w przypadku sieci neuronowej ma 18 węzłów końcowych i 17 decyzyjnych. Porównanie wyników klasyfikacji gestów ze zbioru testowego przedstawiono w tabeli 1.

Tab. 1. Wyniki rozpoznawania gestów ze zbioru testowego: SN -sieć neuronowa, DT- drzewo klasyfikacyjne

Tab. 1. Results of recognition of gestures from the test set: SN- neural network, DT - decision tree

Gest	1p	2p	3p	4p	5p	pięść	ręka	kciuk	suma
l. obrazów	1751	1354	1561	1635	1410	1551	1290	2112	12664
+SN [%]	97,2	92,2	92,1	95,8	96,9	99,9	98,4	98,1	96,4
+DK [%]	94,4	95,0	96,2	97,6	98,8	98,3	97,8	96,1	96,7

Oba klasyfikatory dały zbliżone wyniki. Sieć lepiej rozpoznawała gesty 1p, kciuk, i ręka, podczas gdy drzewo klasyfikacyjne okazało się lepsze w przypadku pozostałych gestów.

4. Śledzenie i rozpoznawanie gestów dynamicznych

Gesty wykonywane rękami mają charakter przestrzenno-czasowy. Wykonanie gestu nie jest na ogół perfekcyjne. Ten sam gest jest zwykle przedstawiany inaczej, zarówno w czasie jak i w przestrzeni, nawet przez tę samą osobę. Wykonanie gestu przez człowieka wywołuje dwa różne procesy stochastyczne. Są one związane ze stanami mentalnymi oraz z wynikającymi

akcjami [19]. Stany mentalne nie są obserwowane, akcje natomiast są. W związku z tym wielu badaczy wykorzystuje do rozpoznawania gestów rąk ukryte modele Markowa (HMM) [7, 11, 14]. W niniejszym opracowaniu zastosowano również to podejście.

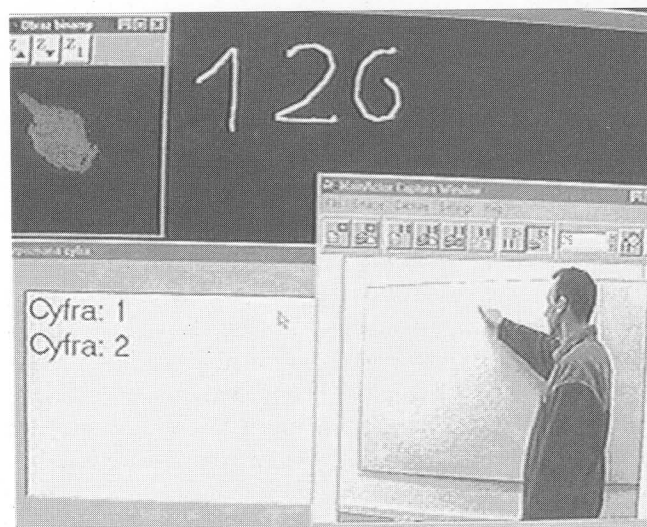
System wykorzystuje komputer PC z kartą pozyskiwania obrazów Matrox Meteor II. Kolorowe obrazy wejściowe o wymiarach 768*576 z kamery Panasonic są rejestrowane z częstotliwością 25 klatek/s. Zastosowano również rzutnik multimedialny wyświetlający na białej tablicy obserwowanej przez kamerę. Rozpoznawane i śledzone gesty wykonywano na tablicy. Eksperymentowi przypisano nazwę „wirtualna tablica”. Oprogramowanie, zawierające moduły: kalibracji kamery, budowania modelu skóry, rozpoznawania skóry, śledzenia dłoni, rozpoznawania kształtu dłoni metodą morfologiczną trafi - nie trafi i rozpoznawania gestów metodą ukrytych modeli Markowa, przygotowano w środowisku Visual C++6.0.

W systemie „wirtualnej tablicy” wykonywane jest śledzenie dłoni tak, aby możliwe było przetwarzanie tylko niewielkiego okna obrazu. Zawężenie przetwarzania do fragmentu obrazu zawierającego dłoń wykonującą gest pozwala na uzyskanie szybkości przetwarzania wystarczającej do rozpoznawania gestów w trybie on-line. Dłoń śledzona jest w oknie o wymiarach 160 x 120. Przyjęto, że znamy początkowe położenie okna. W praktyce oznacza to że na początku użytkownik kładzie dłoń w znanym miejscu tablicy - w jej lewym górnym rogu. Informacja o początkowym położeniu dłoni jest wykorzystywana także do wybrania fragmentu obrazu zawierającego wzorec koloru skóry. Podczas śledzenia przyjęto zasadę, że środek okna ma się pokrywać ze środkiem ciężkości dłoni. Zaprogramowano dwie metody śledzenia. W pierwszej, do bieżących obliczeń wykorzystywane jest okno z poprzedniego kroku. Po wyodrębnieniu dłoni wyznacza się jej środek ciężkości i aktualizuje pozycję okna. Metoda ta działa poprawnie, jeśli rozmiary okna są dostatecznie duże w porównaniu z rozmiarami dłoni w obrazie oraz jeśli ruch dłoni nie odbywa się zbyt szybko. W drugiej metodzie do przewidywania pozycji okna wykorzystano obserwator stanu, zakładając uproszczony model liniowy ruchu ze stałą prędkością. Okazało się, że rozwiązanie bez obserwatora pozwala na ruch dłonią z maksymalną prędkością około 0,68 m/s. Z obserwatorem śledzenie odbywało się poprawnie nawet przy prędkości około 1,15 m/s.

Do wyboru funkcji wykonywanej przez system „wirtualnej tablicy” służą trzy proste gesty. Gest *przesuń* (kształt dłoni jak na rys. 2b) wykorzystywany jest do przesuwania wirtualnych ikon wyświetlanych za pomocą rzutnika multimedialnego w obszarze tablicy, gest *pisz* (jak na rys. 2c) zaś do pisania palcem cyfr na tablicy. Trzeci gest *nic nie rób* polega na przedstawieniu otwartej dłoni z rozłożonymi palcami (rys. 2g). Gesty rozpoznawane są z wykorzystaniem morfologicznej transformacji trafi nie - trafi na podstawie obrazów binarnych otrzymanych w wyniku segmentacji koloru skóry. W celu zwiększenia odporności metody na wahania orientacji dłoni podczas wykonywania gestów dla każdego obrazu binarnego wyznaczane są wartości znormalizowanych momentów centralnych M_{11} , M_{20} i M_{02} [16], które służą do wyznaczenia kąta nachylenia głównej osi obiektu binarnego miał wartość, dla której zaprojektowano elementy strukturalne do operacji trafi - nie trafi. Testy pokazały, że metoda ta sprawdza się nawet przy znacznych wahaniami orientacji dłoni podczas wykonywania gestu.

W eksperymentach dotyczących rozpoznawania gestów dynamicznych wykorzystywano trzy wspomniane kształty dłoni.

Jeśli system rozpozna pierwszy znak *przesuń*, na tablicy jest wyświetlany prostokąt obejmujący dłoń, przesuwany wraz z nią. W drugim przypadku znaku *pisz*, wyświetlana jest na bieżąco linia wykreślana przez wskazujący palec a system analizuje obserwacje używając ukrytych modeli Markowa. By uniknąć niejednoznaczności, użytkownik wykorzystuje gest *nic nie rób* tuż przed i po wykonaniu gestu *pisz*. Jako gesty zdefiniowano przykładowo cyfry 0,1,...,9 i zbudowano HMM dla każdego gestu. Do konstrukcji przygotowano własne oprogramowanie. Dane do uczenia otrzymywane są w systemie wizyjnym. Przygotowano zbiór 600 gestów (60 gestów każdego typu) wykonywanych przez 5 osób. Gesty próbkowane z częstotliwością 25 klatek/s były wykonywane z różnymi prędkościami. Długość kodów charakteryzujących obserwacje w HMM wynosiła od ok. 10 do ok. 30 znaków. Każdy znak kodował aktualny kierunek ruchu, zakwalifikowany jako jeden z ośmiu wariantów. Zbiór uczący 400 gestów wykorzystano do wyznaczenia parametrów HMM metodą Bauma-Welcha. Pozostałe gesty wzięto do testowania. Procent poprawnych odpowiedzi wyniósł 98 w przypadku uczenia i 97,5 podczas testowania. Dwustanowe modele okazały się wystarczające dla cyfr 1, 7, trzystanowe modele dla cyfr 2, 3, 4 i czterostanowe dla cyfr 0, 5, 6, 8, 9. Wspomniane modele są tzw. modelami Bakisa. Rozpoznawanie odbywa się on-line za pomocą algorytmu w przód.



Rys. 3. Rozpoznawanie cyfr: cyfra 6 jest w fazie pisania.
Fig. 3. Recognition of numbers: number 6 is being signed.

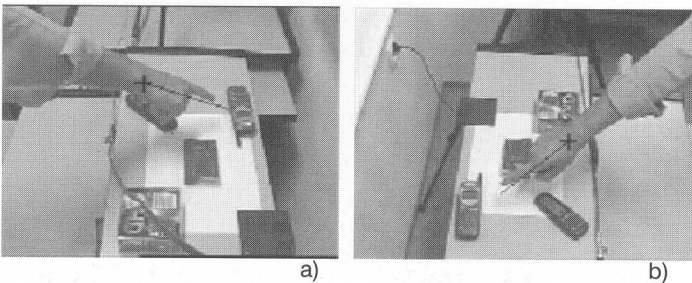
Przetestowano również metodę wskazywania obiektów dłonią w systemie wizyjnym z dwiema kamerami. Podczas uruchamiania systemu wykonywana jest kalibracja kamer oraz tworzone są modele barwy skóry dla obrazów z obu kamer. Modele te wykorzystywane są w trakcie pracy systemu do otrzymywania obrazów binarnych. Następnie wyznaczany jest środek ciężkości oraz nachylenie osi głównej obiektów binarnych odpowiadających dłoni w obrazach z obu kamer. Parametry te wraz ze współczynnikami otrzymanymi na etapie kalibracji kamer służą do przybliżonego określenia wskazywanego kierunku.

Procedurę kalibracji wykonano umieszczając w polu widzenia obu kamer obiekt w kształcie prostopadłościanu o znanych wymiarach. Odpowiedniki wybranych punktów prostopadłościanu w obrazie z kamery znaleziono wskazując je myszką. Znając współrzędne punktów w układzie odniesienia oraz współrzędne punktów w obrazie wyznaczono metodą regresji liniowej parametry macierzy transformacji [1]. Nie uwzględniano zniekształceń optycznych.

Przetwarzanie wykonano na komputerze PC z kartą Matrox Meteor II pozwalającą na dynamiczne przełączanie do 8 kana-

łów wideo. Pomiedzy odpowiadajacymi sobie klatkami z obu kamer wystepuje czasowe przesuniecie rowne 40 ms, za efektywna czestotliwosc przetwarzania z obu kamer to 12,5 Hz.

Wykonano eksperyment polegajacy na wskazywaniu plaskich obiektow lezacych na stole o wymiarach 1 m x 0,5 m. Kamery umieszczono z lewej i prawej strony na wysokosci okolo 0,5 m od blatu i skierowano ku jego srodkowi. Polozenia przedmiotow w ukkladzie odniesienia skojarzonym ze scena byly znane. Ploszczyzna stolu pokrywala sie z ploszczyzna XY sceny. Przyjeta zasade, ze dany obiekt mozna uznac za wskazany, jezeli odleglosc jego srodka od wyznaczonego punktu przeciecia sie ploszczyzny stolu z prostą reprezentujaca wskazywany kierunek jest mniejsza od przytej umownie na podstawie eksperymentow wartosci. Testy wykonano, gdy na stole znajdowalo sie 2, 3, 4 lub 5 obiektow. Za kazdym razem trzy osoby wykonaly po 20 wskazan. Na rys. 4a i 4b przedstawiono obrazy z obu kamer z zaznaczeniem wyznaczonych srodkow ciiezkości oraz kierunkow glownej osi dloni. Praktyka pokazala zasadnosć budowania odrębnego modelu skóry dla kazdej z kamer. Wynikalo to z róznych warunkow oswietlenia.



Rys. 4. Wskazywanie przedmiotow: a), b) obrazy z kamer z zaznaczeniem srodka ciiezkości i glownej osi dloni
Fig. 4. Pointing at objects: a), b) right and left camera pictures with marked centroid and maximum axis of the hand

Przy niewielkiej odleglosci dloni od wskazywanego przedmiotu (ok. 5 cm) wszystkie odpowiedzi systemu byly poprawne. Procent poprawnych wskazan przy odleglosci dwukrotnie wiekszej wynosil odpowiednio 95, 90, 90, 85 dla dwuch, trzech, czterech i pieciu przedmiotow. Jak nalezalo oczekiwac, skutecznosc metody maleje wraz ze wzrostem odleglosci pomiedzy dlonia a wskazywanym obiektem oraz wraz ze wzrostem zagesczenia obiektow na stole. Rozpoznanie wskazania umozliwia systemowi (robotowi) dokonywanie segmentacji tylko w malym fragmencie obrazu wejsciowego, zawierajacym otoczenie wskazywanego palcem celu, a nastepnie dokladne okrelenie pozycji obiektu. Za pomoca umownych gestow statycznych (ksztalt dloni) mozna podpowiedziec sposob uchwycenia przedmiotu. Sladzenie konca palca wskazujacego pozwala natomiast na zgrubne okrelenie sciezki ruchu podczas przenoszenia obiektu z miejsca na miejsce.

5. Podsumowanie

W pracy zostaly przedstawione kierunki badan autorow i wybrane uzyskane przez nich rezultaty rozpoznawania gestow rak w ukkladzie wizyjnym.

Omowiono opracowaną metode rozpoznawania ksztaltu dloni i wyniki eksperymentow dotyczacych zarowno gestow statycznych jak i dynamicznych.

Skupiono sie na zagadnieniach blizszych robotyce. Row nolegle sa prowadzone prace nad rozpoznawaniem Polskiego Języka Miganego [4]. Dotycza one zarowno znakow alfabety palcowego [9], jak i wykonywanych dwiema rekami slow oraz zdań spotykanych w typowych sytuacjach zyciowych [7].

6. Literatura

- [1] B. Cyganek: Komputerowe przetwarzanie obrazow trójwymiarowych. Akad. Ofic. Wyd. EXIT, Warszawa 2002.
- [2] W. T. Freeman, M. Roth: Orientation histograms for hand gesture recognition. Proc. of the IEEE Intl. Workshop on Automatic Face and Gesture Recognition, Zürich, 1995, 296 - 301.
- [3] D. Geer: Will gesture recognition technology point the way? Computer 10, 2004, 20-23.
- [4] J. K. Hendzel: Słownik Polskiego Języka Miganego, Wyd. OFFER, Olsztyn 1995.
- [5] E. Hunter, J. Schlenzig, R. Jain: Posture estimation in reduced model gesture input system. Proc. of the IEEE Intl. Workshop on Automatic Face and Gesture Recognition, Zürich 1995, 290 - 295.
- [6] T. Kapuściński, M. Wysocki: Identyfikacja koloru skóry dloni w róznych przestrzeniach barw. Archiwum Informatyki Teoretycznej i Stosowanej 13(1), 2001, 53-68.
- [7] T. Kapuściński, M. Wysocki: Vision-based recognition of Polish sign language. Proc. of the Symp. on Methods of Artificial Intelligence AIMETH, Gliwice 2003, 145-148.
- [8] A. Licsar, T. Sziranyi: Supervised training based hand gesture recognition system, Proc. of the 16th Int. Conf. On Pattern Recognition ICPR'02, 259-262.
- [9] J. Marnik: Rozpoznawanie znakow Polskiego Alfabetu Palcowego z wykorzystaniem morfologii matematycznej i sieci neuronowych. Rozprawa doktorska, Akademia Górniczo - Hutnicza, Kraków, 2002.
- [10] J. Martin, J. L. Crowley: An appearance-based approach to gesture recognition. Proc of the 9th Intl. Conf. on Image Analysis and Processing, Florence, Italy, 1997.
- [11] J. Martin, J. - B. Durand: Automatic handwriting gestures recognition using hidden Markov models. Proc. of the 4th Int. Conf. On Automatic Face and Gesture Recognition, Grenoble, France, 2000, 403-409.
- [12] V. I. Pavlovic, R. Sharma, T. S. Huang: Visual Interpretation of Hand Gestures for Human - Computer Interaction. IEEE Trans. PAMI, 19, 7, 1997, 677 - 695.
- [13] J. M. Rehg, T. Kanade: Visual tracking of high DOF articulated structures on application to human hand tracking. [w:] Ecklundh J. O., (ed.), Lectures and Notes in Computer Science, vol. 80, Springer, New York, 1994, 34-46.
- [14] T. Starner, J. Weaver, A. Pentland: Real-time American sign language recognition using desk and wearable computer based video. IEEE Trans. PAMI, 20(12):1371-1375, 1998.
- [15] StatSoft Inc.: Statistica Neural Networks. 1998.
- [16] R. Tadeusiewicz: Systemy wizyjne robotow przemysłowych. WNT, Warszawa 1992.
- [17] R. Tadeusiewicz, P. Korohoda: Komputerowa analiza i przetwarzanie obrazow. Wyd. Fundacji Postępu Telekom., Kraków 1997.
- [18] J. Triesch, C. von der Malsburg: A system for person independent hand posture recognition against complex backgrounds. IEEE Trans. PAMI, 23(12), 2001, 1449-1453.
- [19] J. Yang, X. Yangsheng, C. S. Chen: Human action learning via hidden Markov model. IEEE Trans. SMC, Part A, 27(1), 1997, 34-44.
- [20] M. Zhao, F. K. H. Quek, X. Wu: RIEVL: Recursive induction learning in hand gesture recognition. IEEE Trans. PAMI, 20(11), 1998, 1174-1185.

Title: Vision-based hand gesture recognition

Artykul recenzowany