

Analiza wpływu zmian parametrów algorytmu rozpoznawania obiektu na podstawie 2-wymiarowego modelu na poprawność procesu rozpoznania

Piotr KOWALECZKO¹, Przemysław ROKITA²

1 Instytut Matematyki i Kryptologii WAT,
ul. Gen. S. Kaliskiego 2, 00-908 Warszawa;

1 Zakład Integracji Systemów C4ISR, Instytut Techniczny Wojsk Lotniczych,
ul. Księcia Bolesława 6, 01-494 Warszawa
piotr.kowaleczko@o2.pl

2 Instytut Teleinformatyki i Automatyki WAT,
ul. Gen. S. Kaliskiego 2, 00-908 Warszawa
pro@ii.pw.edu.pl

STRESZCZENIE: W artykule przedstawiono wyniki badań wpływu doboru wartości parametrów algorytmu na poprawność i efektywność procesu rozpoznawania obiektu na obrazie z wykorzystaniem metody porównywania z 2-wymiarowym wzorcem. Jako badane parametry przyjęto liczbę iteracji operacji dylatacji wzorcowego konturu, rozmieszczenie punktów próbnych obiektu, liczbę modeli oraz wprowadzanie różnego rodzaju ograniczeń nakładanych podczas klasyfikowania obiektów. Rozpatrywano zarówno analizę obrazów zapisanych w formie plików graficznych, jak i materiału wideo.

SŁOWA KLUCZOWE: rozpoznawanie obrazów, transformata odległościowa, wpływ parametrów.

1. Wstęp

Rozpoznawanie obiektów na cyfrowym obrazie jest procesem skomplikowanym. Trudności pojawiają się w zasadzie na wszystkich etapach procesu. Po dokonaniu akwizycji obrazu (której jakość ma także bardzo duże znaczenie), konieczne jest dokonanie właściwego doboru sposobów wstępnego przetwarzania obrazu (eliminacji zakłóceń, eksponowania jego ważnych cech

itp.). W kolejnym etapie - analizie obrazu, należy wybrać optymalną metodę segmentacji aby zidentyfikować na obrazie obiekty mogące odpowiadać poszukiwanym wzorcom. Najbardziej złożony jest jednak proces określania czy dany obiekt pasuje do posiadanego wzorca, czy też nie. W każdym z etapów możliwe jest wykorzystanie najróżniejszych algorytmów, zaś w każdym z algorytmów konieczny jest właściwy dobór wartości parametrów. Może zdarzyć się, że wykorzystanie nawet najlepszego algorytmu segmentacji dla danego rodzaju zadania okaże się niewystarczające z powodu niewłaściwego doboru wartości parametrów. Analogiczna sytuacja ma miejsce w procesie analizy zaindeksowanego już obrazu. Problem jest tym poważniejszy, że przy analizie obrazu dostępnych jest znacznie więcej algorytmów niż w poprzednich etapach przetwarzania obrazu. Możliwa jest klasyfikacja obiektów na podstawie np. histogramów orientacji [4], współczynników kształtu [1], momentów bezwładności [1] czy też zapisanych w pamięci modeli reprezentowanych przez binarne obrazy [3] lub tzw. „active shape model” [5].

Przy ogromnej liczbie parametrów, dla bardziej złożonych algorytmów, problem właściwego ich doboru staje się kluczowy dla prawidłowego rozpoznania poszukiwanego obiektu na obrazie. W artykule przedstawiony został wpływ zmian wartości parametrów algorytmu rozpoznawania obiektu na poprawność procesu rozpoznawania. Rozpoznawanie wykonywane było na podstawie zapisanego w pamięci 2-wymiarowego wzorca.

2. Opis przyjętej metody

Na potrzeby badań zbudowano aplikację umożliwiającą rozpoznanie otwartej dłoni na wprowadzonym obrazie. Przyjęcie do badań takiego obiektu wynikało z łatwości pozyskania obrazów testowych. Docelowo jednak, planowane jest utworzenie podobnego systemu, rozpoznającego statki powietrzne. Wykorzystany został algorytm porównania znalezionych obiektów z zapisanym w pamięci, dwuwymiarowym modelem (algorytm opisany szerzej w [3]). Model ten, to binarny obraz, w którym piksele aktywne (o wartości 255 – białe) przedstawiają obiekt, zaś tło reprezentowane jest przez piksele o zerowej wartości. Wykorzystany w trakcie badań model przedstawiono na rys. 2 (pierwsze okno).

Zaimplementowany na potrzeby badań algorytm opiera się na przeprowadzeniu dwóch porównań:

- porównanie proporcji znalezionego obiektu z proporcjami modelu;
- porównanie sumy wartości transformowanego metodą odległościową konturowego obrazu modelu (dla wyznaczonych punktów próbnych) z wartością progową.

dokonyje się operacji dylatacji, a następnie transformaty odległościowej, tejże dylatacji. Poszerzając kontur i na tak spreparowanym obrazie dokonując transformaty odległościowej, umożliwia się przypisanie wag (wartości transformaty) modelowemu obiektowi. Im „bliżej” idealnego (oryginalnego, przed dylatacją) kształtu konturu znajdzie się punkt próbny, tym większa wartość zostanie mu przypisana. Granicą, kiedy dany piksel nie będzie w żadnym stopniu uznany za pasujący do modelu jest kontur obrazu poddanego dylatacji. Na rys. 1 zaznaczono także 3 piksele testowe, dla których dokonywano oceny zgodności z modelem opisanym obrazem transformaty. Przyjęto że obraz oryginalny (po lewej) jest fragmentem konturu modelu poddanego operacji dylatacji. Dla piksela oznaczonego zielonym kolorem zgodność wynosi 100% (maksymalna lokalna wartość transformaty), dla piksela żółtego 50%, zaś piksel czerwony został uznany za piksel niepasujący do modelu.

Kolejnym problemem, z którym należało się uporać było ustalenie metody, na podstawie której klasyfikowane były obiekty. Do klasyfikacji wykorzystano opisany wyżej obraz transformaty odległościowej konturu modelu po dokonaniu dylatacji, zwany dalej wzorcem. Punkty próbne wybierane zostają zgodnie z metodą opisaną w dalszej części dokumentu. Ze wzorca odczytywana jest wartość jasności piksela odpowiadającego położeniu punktu testowego. Wartości te są odczytywane dla wszystkich punktów testowych, a następnie zgodnie z prostym wzorem na średnią arytmetyczną (1) wyliczana jest ich średnia wartość.

$$j_{avg} = \frac{j_1 + j_2 + \dots + j_n}{n} \quad (1)$$

gdzie:

j_{avg} – średnia jasność analizowanych punktów konturu,

j_i – jasność i-tego analizowanego punktu konturu,

n – liczba analizowanych punktów konturu.

Proces porównania testowanego obiektu do wzorca polega na porównaniu wartości j_{avg} ze średnią jasnością wszystkich aktywnych punktów wzorca. Wartość ta wyznaczana jest zgodnie ze wzorem (2).

$$m_{avg} = \frac{t_1 + t_2 + \dots + t_n}{n} \quad (2)$$

gdzie:

m_{avg} – średnia jasność wszystkich aktywnych punktów wzorca,

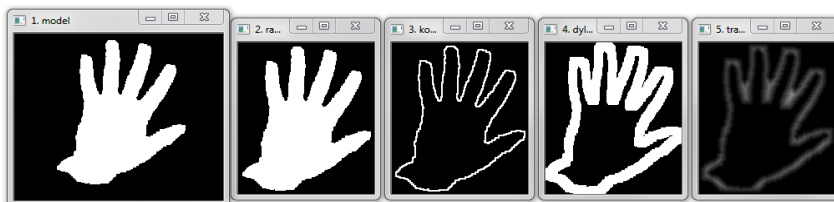
t_i – jasność i-tego aktywnego punktu wzorca,

n – liczba wszystkich aktywnych punktów wzorca.

Dla przykładu, w przypadku rys. 1, wartość $m_{avg} = \frac{246}{115} \approx 2,13$.

Kolejne etapy przetwarzania modelu aż do powstania obrazu wzorcowego

przetawiono na rys. 2.

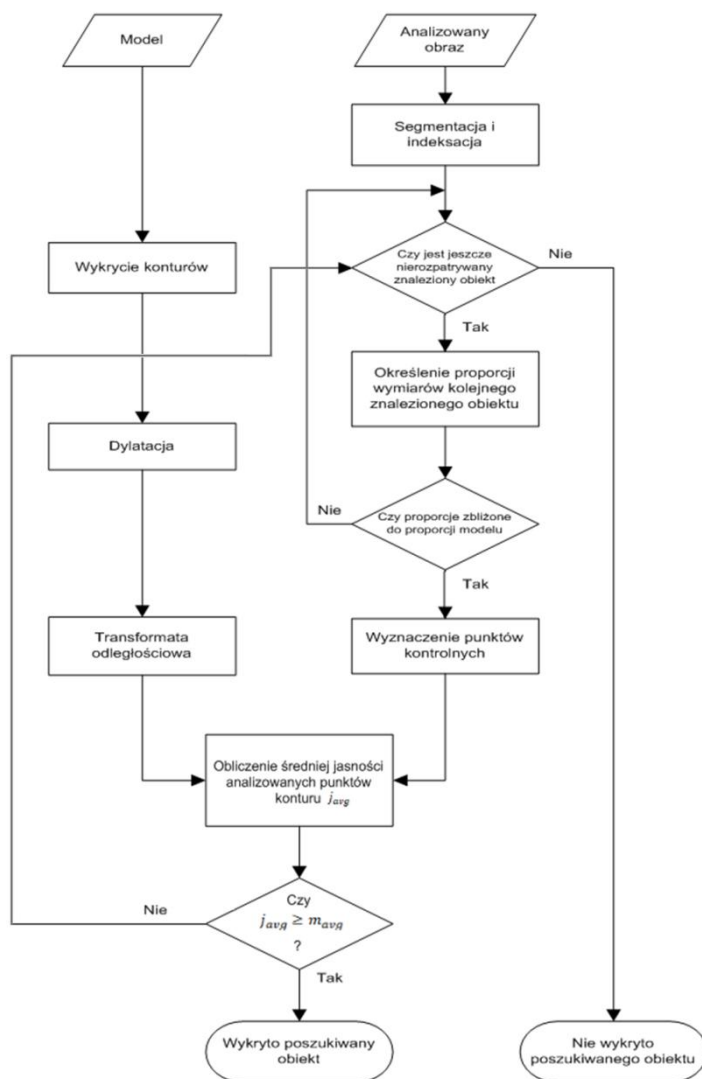


Rys. 2. Kolejne etapy przetwarzania modelu obiektu

Pierwsze z okien przedstawia wczytany plik z modelem, na drugim po indeksacji wycięto tylko sam wykryty obiekt. Dzięki temu możliwe jest określenie rzeczywistych rozmiarów (a więc także proporcji) dłoni. Kolejne obrazy prezentują odpowiednio: obraz konturowy, obraz konturowy po dylatacji oraz obraz konturowy po dylatacji poddany dodatkowo transformacji odległościowej.

Schemat blokowy bazowej wersji zaimplementowanego algorytmu rozpoznawania dłoni przedstawiono na rys. 3.

Przedstawiony, na rys. 3, algorytm bazuje na rozwiązaniach opisanych powyżej i w [3]. Powodzenie procesu rozpoznawania zależy od właściwego doboru wartości niektórych parametrów. Można do nich zaliczyć m.in.: liczbę próbnych punktów konturu analizowanego obiektu, ich rozmieszczenie, liczbę iteracji dylatacji konturu modelu (szerokość konturu przed dokonaniem transformaty odległościowej), czy też sposób wyznaczania i wartość progową przy porównaniu sumy wartości punktów próbnych. Przyjmując, że poszukiwany obiekt może znajdować się w różnych pozycjach (obrót wokół osi przebiegającej przez środek ciężkości modelu), bardzo ważnym parametrem jest także liczba rozpatrywanych przy porównaniach obrotów modelu. Wybierając zbyt duży krok obrotu może bowiem dojść do sytuacji, gdy rozpatrywany obiekt nie będzie pasował do modelu obróconego o np. 30 i 40 stopni, zaś pasowałby do modelu obróconego o 35 stopni. W artykule określony zostanie wpływ zmian wartości wymienionych wyżej parametrów na powodzenie procesu rozpoznawania obiektu. W ramach przeprowadzonych badań, w pierwszej kolejności, analizowano specjalnie spreparowane obrazy, na których obiekty, które miały zostać znalezione umieszczano po dokonaniu obrotu w różnych miejscach obrazu, wśród innych obiektów. W drugiej fazie poszukiwania obiektu opisanego modelem, analizy wykonywano na rzeczywistym obrazie pozyskanym z kamery.

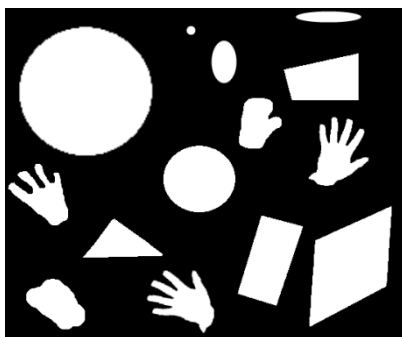


Rys. 3. Algorytm rozpoznawania dłoni na podstawie dwuwymiarowego wzorca przy jednym modelu

3. Wpływ liczby iteracji dylatacji na powodzenie procesu rozpoznawania

W ramach przeprowadzonych badań wykonano 12 prób rozpoznania dwóch modeli ludzkiej dłoni. Obiekty umieszczone zostały na specjalnie

spreparowanym obrazie binarnym o wymiarach 548 x 455 pikseli (rys. 4). Średnia wielkość obiektów (dłoni) wynosi ok. 100 x 100 pikseli. Prócz dwóch obiektów, które miały zostać znalezione (dłoń obrócona o kąt 0 i 80 stopni), na obrazie znajdują się także inne obiekty. Celem było dobranie takiej liczby iteracji dylatacji konturu modelu (algorytm wykorzystuje 24 modele dłoni – obracane co 15 stopni), aby wykryte zostały tylko obiekty – dłonie.



Rys. 4. Obraz poddany operacji rozpoznawania

Tab. 1. Wartości stosunków $\frac{J_{avg}}{m_{avg}}$ i średnie jasności ($\frac{\text{suma jasności aktywnych pikseli}}{\text{suma jasności wszystkich pikseli}}$) dla rozpoznawania obiektów przy zmieniających się liczbach iteracji

Liczba iteracji dylatacji (wzrost szerokości konturu)				Średnia jasność próbek	Średnia jasność modeli	Liczba znalezionych obiektów
	reka_0	reka_75	reka_90			
1 (2x)	0,88	0,23	0,26	0,36	1,98	0
2 (3x)	1,17	0,39	0,37	0,73	2,73	1
3 (4x)	1,35	0,51	0,48	1,26	3,57	1
4 (5x)	1,46	0,82	0,62	1,94	4,4	1
5 (6x)	1,52	1,04	0,74	2,94	5,57	4
6 (7x)	1,55	1,16	0,86	4,59	7,62	6
7 (8x)	1,36	1,25	1,03	6,76	10,45	14
8 (9x)	1,29	1,28	1,09	8,57	12,49	19
9 (10x)	1,33	1,33	1,15	9,7	13,37	23
10 (11x)	1,35	1,38	1,22	10,72	14,06	26
11 (12x)	1,37	1,41	1,27	11,64	14,63	27
12 (13x)	1,41	1,44	1,31	12,52	15,16	32

Zamieszczona powyżej, tab. 1 zawiera wyniki otrzymane przy kolejnych liczbach iteracji operacji dylatacji. Im większa liczba iteracji, tym kontury modelu po wykonaniu transformaty odległościowej stają się „szersze”, zwiększając tym samym prawdopodobieństwo przynależności do nich punktu próbnego konturu testowanego obiektu. Potwierdzają to zamieszczone powyżej

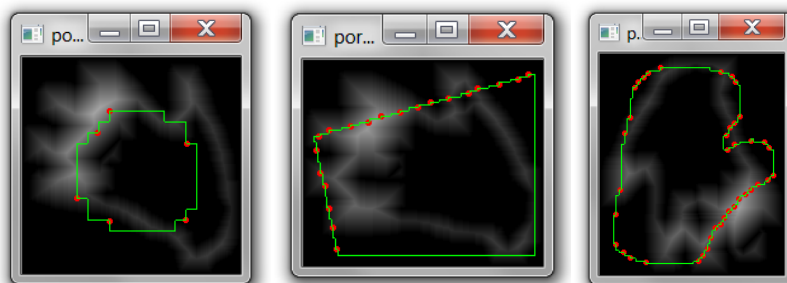
wyniki. Należy dodać, że operacji dylatacji poddawane są jedynie kontury modeli. Przy ich średniej oryginalnej szerokości ok. 2 pikseli przeprowadzenie 1 iteracji powoduje dwukrotny wzrost szerokości konturu (powiększenie linii konturu o 1 piksel z lewej i z prawej strony). Dla każdego analizowanego obiektu wyliczana była średnia wartość jasności punktów próbnych. Jej stosunek do średniej jasności danego modelu zawierają kolumny „reka_0”, „reka_75” i „reka_90”. W przypadku, gdy wynosi on powyżej wartości 1, znaczy to, że średnia wartość jasności punktów próbnych jest większa od wartości średniej dla wzorca, co zostało przyjęte jako powodzenie (rozpoznanie obiektu). Wartości kolumny „średnia jasność modeli” odnoszą się natomiast do wszystkich 24 modeli. Rozpatrywane zostały modele dłoni obrócone o 0, 75 i 90 stopni, ponieważ na obrazie testowym umieszczono obiekty obrócone o kąty 0 i 80 stopni. W przypadku pierwszego modelu wskazanie powinno być jednoznaczne (znaleziono lub nie), w przypadku drugiego natomiast badano dodatkowo czy obrócony o 80 stopni obiekt zostanie zaklasyfikowany jako reprezentant modelu „reka_75” czy „reka_90”. W tab. 1 zamieszczono także średnie wartości jasności punktów transformaty konturu modeli oraz punktów próbnych testowanych obiektów. Dają one pogląd, jak duże są różnice między wartościami dla znalezionych obiektów, a wartościami średnimi. Różnice te można rozumieć jako jakość dokonanego rozpoznania. Analizując wartości zawarte w tab. 1 można zauważyć, że prawdopodobieństwo rozpoznania dłoni przez algorytm jest wprost proporcjonalne do liczby iteracji operacji dylatacji. W przypadku idealnie opisanego modelem obiektu (reka_0) algorytm jest w stanie rozpoznać na obrazie dłoni już po przeprowadzeniu 2 iteracji. Jeżeli obiekt może być jedynie przybliżony jednym z modeli, rozpoznanie zostaje zakończone sukcesem po przeprowadzeniu minimum 5 iteracji. Liczba ta po przełożeniu na wartość procentową (liczba pikseli, o które „powiększył się” obiekt w stosunku do oryginalnej szerokości jego konturu) daje wynik ok. 600%. Niestety wraz ze wzrostem liczby iteracji, prócz obiektów które powinny zostać zaklasyfikowane jako dłoni, rozpoznane zostają także inne, niepożądane obiekty. Po przeprowadzeniu analizy można stwierdzić, że przy przyjętej rozdzielczości najlepiej jest zaimplementować 5 lub 6 iteracji (6- lub 7-krotne „poszerzenie” konturu obiektu). Przy wykorzystaniu dodatkowych metod rozpoznawania obiektu, szansa na poprawne rozpoznanie jest dość duża.

Rys. 5 przedstawia zmiany transformaty poddanego dylatacji obrazu konturowego modelu przy różnych liczbach iteracji operacji dylatacji. Można zauważyć, że wraz ze wzrostem liczby iteracji operacji dylatacji, szerokość konturu w transformacie odległościowej (liczba aktywnych punktów) zwiększa się, co przekłada się na zwiększenie prawdopodobieństwa uznania (z pewną wagą określoną jasnością transformaty w tym pikselu) punktu testowego jako zgodnego ze wzorcem. Można także zauważyć, że w przypadku niewielkiej liczby iteracji dylatacji, obraz transformaty przybiera jeszcze kształt dłoni, podczas gdy po dokonaniu większej liczby dylatacji, obraz oddaje tylko ogólny

kształt dłoni. Właśnie dlatego rozpoznawalność obiektu dłoni wzrasta wraz ze wzrostem liczby iteracji. Zwiększa się wtedy także prawdopodobieństwo zakwalifikowania jako dłoni obiektu nią nie będącego. Wraz z kolejnymi iteracjami powstają też zakłócenia wynikające ze zlewania się konturów sąsiadujących ze sobą palców. Doprowadza to do możliwości powstania dosyć poważnych przekłamań – punkty próbne znajdujące się pomiędzy palcami będą oceniane lepiej niż te, znajdujące się dokładnie na konturze.



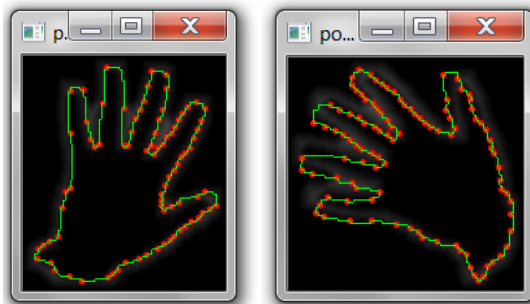
Rys. 5. Zmiany transformaty poddanej dylatacji obrazu konturowego modelu, przy różnych (kolejno: 1 (2x), 4 (3x), 6 (7x), 8 (9x) i 12 (13x)) liczbach iteracji operacji dylatacji (powiększeniach konturu obiektu)



Rys. 6. Niepoprawnie zaklasyfikowane obiekty

Z uwagi na poprawę jakości wyników procesu rozpoznawania wraz ze wzrostem liczby iteracji dylatacji z jednej strony i pojawianiem się zniekształceń modelu

z drugiej, konieczna jest analiza konkretnego problemu rozpoznania i dobranie optymalnej wartości liczby dylatacji dla tego przypadku.



Rys. 7. Poprawnie zaklasyfikowane obiekty

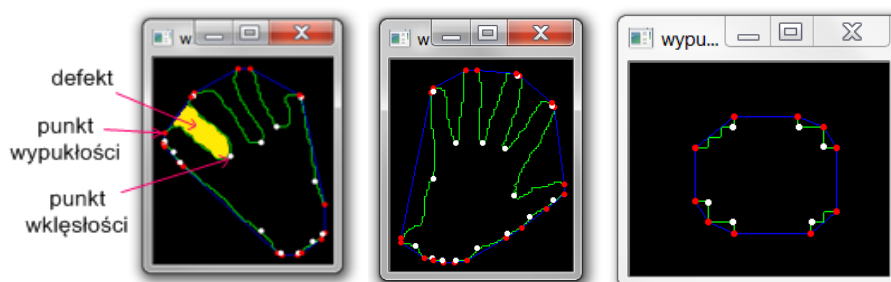
Przykłady błędnie i poprawnie rozpoznanych obiektów (wraz z zaznaczonymi punktami próbnymi ich konturów, dla których wykonywane były obliczenia) przedstawiają odpowiednio rys. 6 i 7.

4. Wpływ rozmieszczenia i wartości jasności punktów próbnych i charakterystycznych konturu na powodzenie procesu rozpoznawania

Pomimo tego, że punkty próbne (dla których wyliczane były wartości sumy) są wybierane równomiernie z uporządkowanego zbioru punktów opisujących kontur, analizując rys. 6 można zauważyć, że ich położenie względem konturu jest nierównomierne. Wynika to ze sposobu opisu kształtu konturowego w bibliotece OpenCV. Kontur opisany jest tam odcinkami, zaś każdy z nich reprezentowany jest przez dwa punkty położone na jego końcach. Stąd w drugim oknie rys. 6, dla analizowanej figury nie był brany pod uwagę żaden punkt leżący na nieukośnych bokach. W przypadku poprawnego, równomiernego rozmieszczenia punktów próbnych dla tego obiektu, nie zostałby on rozpoznany jako kształt dłoni.

Aby zapobiec sytuacji przedstawionych na rys. 6, wprowadzono dodatkowe wymaganie względem analizowanego kształtu: każdy z jego punktów charakterystycznych musi mieć niezerową wartość na obrazie transformaty wykorzystywanej do porównań (rys. 5). Dopiero spełnienie powyższego warunku, wraz ze spełnieniem warunku mówiącego że średnia jasność dla równomiernie rozłożonych punktów próbnych musi być większa lub równa od średniej jasności punktów aktywnych obrazu transformaty, powoduje zaklasyfikowanie analizowanego kształtu jako obiektu opisanego modelem.

Jako punkty charakterystyczne, rozumie się punkty wklęsłości i wypukłości figury [5]. Ich wykrycie na obrazie próbnym realizowane jest przez użycie następujących funkcji biblioteki OpenCV: `cvFindContours` (odnalezienie konturów), `cvConvexHull2` (wyznaczenie najmniejszej figury wypukłej opisanej na konturze) oraz `cvConvexityDefects` (znalezienie defektów – fragmentów konturu niepokrywającego się z wyznaczoną figurą wypukłą). Rys. 8 przedstawia znalezione na konturach (zielona linia) punkty wklęsłości (białe), wypukłości (czerwone), oraz najmniejsze figury wypukłe opisane na konturze (niebieska linia). Zaznaczone punkty wklęsłości to punkty znajdujące się możliwie najdalej (największa głębokość) od linii figury wypukłej, w ramach jednego defektu (pojedynczego uchylenia linii konturu względem linii figury wypukłej – kolor żółty).



Rys. 8. Znalezione na konturze punkty wklęsłości i wypukłości

Aby pozyskać tylko punkty charakterystyczne reprezentujące końcówki palców i zagłębienia między nimi w opracowanym algorytmie należało przyjąć pewien próg głębokości punktu wklęsłości. Przyjęto, że wartość ta musi być większa niż połowa największej wartości punktu wklęsłości dla znalezionych defektów. Przyjęto, że dłoń posiada znaczne różnice między głębokościami punktów wklęsłości i wypukłości w obrębie palców, i nieznaczne w pozostałych fragmentach konturu. Aby zapobiec analizowaniu punktów wklęsłości i wypukłości dla obiektów nieposiadających takich wyraźnych różnic, zastosowano dodatkowe ograniczenie porównujące moduł różnicy między głębokością danego punktu i wyliczoną średnią głębokością punktów. Jeśli wartość ta jest większa od połowy średniej głębokości, uznaje się, że nie ma wyraźnego zarysu palców, więc obiekt nie może być dłonią. Ograniczenia te zapisano poniżej:

$$(d > 0.5 * d_{max}) \quad (3)$$

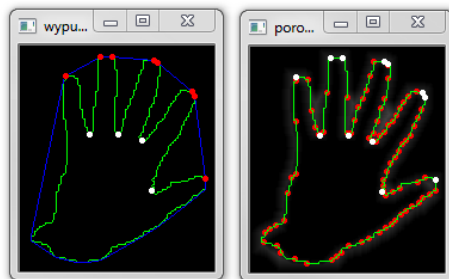
$$(|d - d_{avg}| > 0.5 * d_{avg}) \quad (4)$$

gdzie:

d – głębokość punktu wklęsłości,

d_{max} – największa głębokość dla punktów wklęsłości znalezionych defektów,

d_{avg} – średnia głębokość dla punktów wklęsłości znalezionych defektów.



Rys. 9. Znalezione na konturze punkty wklęsłości i wypukłości palców (lewe okno) oraz wszystkie punkty testowane przy porównaniach (prawe okno)

W wyniku zastosowania w algorytmie opisanego wyżej ograniczenia, w konturach analizowanego obiektu znalezione zostają tylko założone punkty wklęsłości i wypukłości (lewe okno rys. 9). Wszystkie punkty, dla których podczas rozpoznawania obiektu analizowane są punkty transformaty odległościowej pokazane zostały w prawym oknie rys. 9 (białe – punkty wklęsłości i wypukłości, czerwone – równomiernie rozłożone punkty konturu). Warunkiem zaklasyfikowania testowanego kształtu jako zgodnego z modelem jest ostatecznie spełnienie zależności:

$$(j_{avg} \geq m_{avg}) \quad (5)$$

$$\forall c \in C \ j(c) \neq 0 \quad (6)$$

gdzie:

j_{avg} – średnia jasność analizowanych punktów konturu,

m_{avg} – średnia jasność wszystkich aktywnych punktów wzorca (transformaty odległościowej konturu modelu poddanej dylatacji),

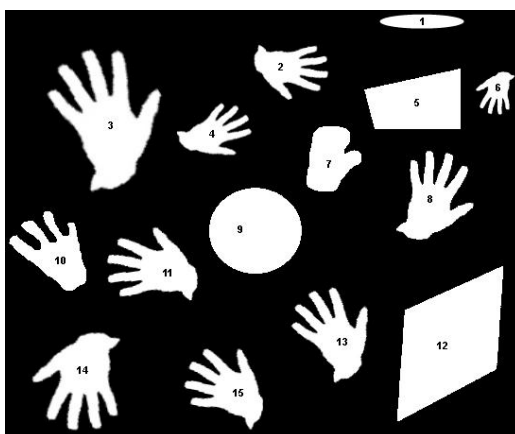
c – punkt charakterystyczny (wypukły i wklęsły) analizowanego konturu,

C – zbiór wszystkich punktów charakterystycznych analizowanego konturu,

$j(c)$ – jasność wzorca dla punktu charakterystycznego c .

Po wprowadzeniu wyżej opisanych modyfikacji, algorytm z dostarczonego mu binarnego obrazu wejściowego poprawnie wykrył dwa znajdujące się na nim obiekty, będące dłońmi. Błędy przedstawione na rys. 6 zostały wyeliminowane. Okazało się jednak, że po wprowadzeniu do obrazu testowego dodatkowego obiektu – ręki, obróconego o niestandardowy kąt,

wymaganie dotyczące aktywności wszystkich punktów charakterystycznych uniemożliwiło poprawne rozpoznanie obrazu. Zwiększenie możliwości rozpoznawania obracanych o różne kąty kształtów może zostać zrealizowane na dwa sposoby: zmniejszając skok obrotu każdego kolejnego modelu (zwiększenie liczby modeli), lub zniesieniem ograniczenia dotyczącego niezerowości wartości transformaty dla punktów charakterystycznych (6). Przy analizie powyższych dwóch przypadków wykorzystano zmodyfikowany obraz wejściowy, przedstawiony na rys. 10 (analizowany obraz pozbawiony był liczb przyporządkowanych obiektom, widocznych na rysunku). Pośród kilku obiektów nie będących dłońmi, zamieszczono tam 9 obiektów - dłoń, obróconych o losowe kąty: 0, 34, 49, 79, 98, 172, 191, 281 i 311 stopni. Dodatkowo, w dwóch z nich (15 i 14) dokonano także minimalnych zmian proporcji.



Rys. 10. Zmodyfikowany obraz poddany operacji rozpoznawania

Aby zwiększyć szanse poprawnego rozpoznania przez algorytm wszystkich obiektów-dłoni, przeprowadzono próby jego modyfikacji: zwiększono dwukrotnie liczbę modeli, a także zmodyfikowano ograniczenie niezerowej wartości wzorca dla każdego punktu charakterystycznego. W pierwszym przypadku usunięto je całkowicie (nie obowiązują wtedy także warunki (3), (4)), w drugim zastąpiono je poniższą, analogiczną do (5) zależnością:

$$(j_{avg} \geq m_{avg}) \quad (7)$$

gdzie:

j_{avg} – średnia jasność wzorca dla analizowanych punktów charakterystycznych konturu,

m_{avg} – średnia jasność wszystkich aktywnych punktów wzorca.

W tab. 2, zamieszczone zostały wyniki działania algorytmu rozpoznawania dla sześciu sytuacji:

- 1) obowiązujący warunek (5), skok obrotu modelu: 15 stopni (24 modele);
- 2) obowiązujący warunek (5), skok obrotu modelu: 7,5 stopnia (48 modeli);
- 3) niezmieniony algorytm: obowiązujące warunki (3), (4), (5) i (6), skok obrotu modelu: 15 stopni (24 modele);
- 4) obowiązujące warunki (3), (4), (5) i (6), skok obrotu modelu: 7,5 stopnia (48 modeli);
- 5) warunek (6) zastąpiony warunkiem (7), skok obrotu modelu: 15 stopni (24 modele);
- 6) warunek (6) zastąpiony warunkiem (7), skok obrotu modelu: 7,5 stopnia (48 modeli).

Tab. 2. Wartości kątów obrotów pasujących do analizowanego obiektu modeli przy zastosowaniu różnych wersji algorytmu rozpoznawania

Kąt obrotu modelu	Numer obiektu	Pasujące modele (kąt obrotu), uwagi					
		Algorytm 1	Algorytm 2	Algorytm 3	Algorytm 4	Algorytm 5	Algorytm 6
0	8	0	0	0	0	0	0
34	3	30, 45	30, 37.5, 45	-	37.5	-	37.5
49	13	45	45, 52.5	45	45, 52.5	45	45, 52.5
79	15	-	-	-	-	-	-
98	11	-	97.5	-	97.5	-	97.5
172	6	-	-	-	-	-	-
191	14	195	187.5, 195	195	187.5, 195	-	187.5
281	2	285	277.5, 285	285	277.5, 285	285.5	277.5, 285
311	4	315	315, 322.5	-	322.5	315	315, 322.5
obiekt 1		-	-	-	-	-	-
obiekt 5		-	-	-	-	-	-
obiekt 7		-	-	-	-	-	-
obiekt 9		-	-	-	-	-	-
obiekt 10		60	52.5, 60	-	-	-	-
obiekt 12		-	-	-	-	-	-

Analizując zamieszczone powyżej wyniki można stwierdzić, że dwukrotne zwiększenie liczby modeli pozwala znacząco zwiększyć szanse poprawnego rozpoznania obiektu. Obiekty, których kąt obrotu wypadł

między kątami kolejnych modeli (np. obiekt 11 lub 3) nie mogły zostać rozpoznane przy przyjętym dla 24 modeli kroku 15 stopni. Należy dodać, że przy dalszych wzrostach liczby modeli, skuteczność rozpoznawania ulegałaby dalszej poprawie. Mógłby się wtedy pojawić problem związany z czasochłonnością pozyskania dużej liczby wzorcowych obrazów. Istnieją jednak metody automatycznej generacji tego typu obrazów, opisane szerzej w [7]. Przy wzroście liczby modeli, zwiększyły się także szanse rozpoznania pozostałych (rozpoznawalnych przy 15 modelach) obiektów, gdyż pasowały one do kilku następujących po sobie modeli, np. obiekt 3 pasował do modeli obróconych względem pierwotnego o 30, 37.5 i 45 stopni. Należy także dodać, że analizując wartości jasności punktów testowych dla kolejnych prób (na ich podstawie dokonywano klasyfikacji), były one tym większe, im bliższa była wartość kąta obrotu modelu względem rzeczywistego kąta obrotu obiektu. Można zatem, w prosty sposób, rozszerzyć algorytm o funkcję wskazywania najbardziej prawdopodobnego kąta obrotu wykrytego i rozpoznanego obiektu względem modelu pierwotnego (0 stopni).

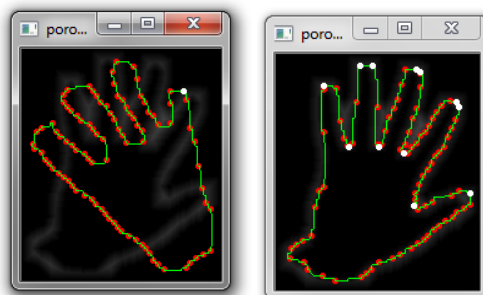
Kolejnym spostrzeżeniem, którego można dokonać analizując zamieszczone wyżej dane jest to, jak duży wpływ na poprawne rozpoznanie obiektu w metodzie dwuwymiarowego wzorca mają zniekształcenia i przeskalowania analizowanych obiektów. Obiekt 6, który został dość istotnie pomniejszony, nie został rozpoznany ani razu. Podczas realizacji pomiarów zauważono, że po dopasowaniu rozmiaru okna z obiektem do rozmiaru okna z modelem (co wiązało się z przeskalowaniem – powiększeniem), kontury zmieniły się dość znacząco, stając się zdecydowanie mniej zaokrąglonymi. Było to spowodowane zastosowaniem interpolacji, która z definicji przy powiększaniu nie pozwala na uszczegółowienie obiektu (żadne nowe informacje nie są wnoszone). Niewielka linia oryginalnego, małego obiektu musiała zostać znacznie powiększona, co spowodowało, że nie dało się idealnie odwzorować jej kształtu w powiększeniu, a to z kolei nie pozwoliło na zaklasyfikowanie konturów obiektu jako zgodnego z modelem. Podobnie sytuacja wyglądała w przypadku obiektu nr 15. Jest to jeden z obiektów, którego proporcje zostały celowo zmienione (został zmniejszony wzdłuż osi X). Spowodowało to, że algorytm odrzucił ten obiekt podczas przeprowadzania analizy zgodności proporcji obiektu z proporcjami modelu. Po przeprowadzeniu próby, w której dopuszczono analizowanie obiektów różniących się znacząco proporcjami względem proporcji modelu, obiekt został poprawnie rozpoznany. Wprowadzenie tego typu ograniczenia pozwala jednak na znaczne przyspieszenie obliczeń. Kolejny zysk w szybkości obliczeń mógłby zostać osiągnięty poprzez ekstrakowanie osi wiodącej rozpatrywanego obiektu, i analizowanie tylko modeli o zbliżonych kątach obrotu do kąta osi wiodącej. Aby umożliwić rozpoznanie obiektów zdeformowanych (które można np. uznać za rzuty obiektów obróconych w przestrzeni trójwymiarowej) należy natomiast

wprowadzić dodatkowe, reprezentujące je modele.

Co ciekawe, uwzględnienie warunków (6) i (7) dało oczekiwany rezultat, jakim było uniemożliwienie zaklasyfikowania jako poszukiwanych obiektów, np. obiektów przedstawionych na rys. 6. Nawet przy istotnym złagodzeniu warunku (5) (pomniejszona o 40% wartość średniej jasności punktów modelu) warunki (6) i (7) gwarantują niezaklasyfikowanie obiektów, o punktach charakterystycznych nie leżących na transformacie konturu modelu.

5. Wpływ liczby wykrytych defektów i ich głębokości na powodzenie procesu rozpoznania

Oprócz opisanych wyżej metod rozpoznania wzorca, można także zastosować algorytm sprawdzający jaka jest liczba defektów w rozpoznanym konturze obiektu (wykorzystany także w [6]). W otwartej dłoni powinny zostać wykryte 4 defekty – przestrzenie między sąsiadującymi palcami. Obiekty o konturach zwartych, nie posiadających tego typu cech, zostaną odrzucone. Metoda ta może stanowić dopełnienie do metody dopasowania punktów próbnych i charakterystycznych do transformaty konturu. Brak możliwości rozpoznania dłoni ze złączonymi dwoma lub więcej palcami nie jest tu wadą, gdyż tego typu obiekt musiałby zostać opisany w wykorzystanym algorytmie odrębnym modelem. Liczba wykrytych defektów jest zliczana na podstawie liczby wykrytych punktów charakterystycznych. Rys. 11 przedstawia rozpoznane na testowym obrazie obiekty z 3 (lewe okno) i 4 (prawe okno) defektami (odpowiednio 9 i 12 punktów charakterystycznych). Należy dodać że rozpoznania dokonane zostały, bez konieczności spełnienia warunków (5), (6) i (7), jedynie na podstawie liczby defektów.



**Rys. 11. Obiekty rozpoznane tylko na podstawie liczby defektów
(3 dla lewego i 4 dla prawego okna)**

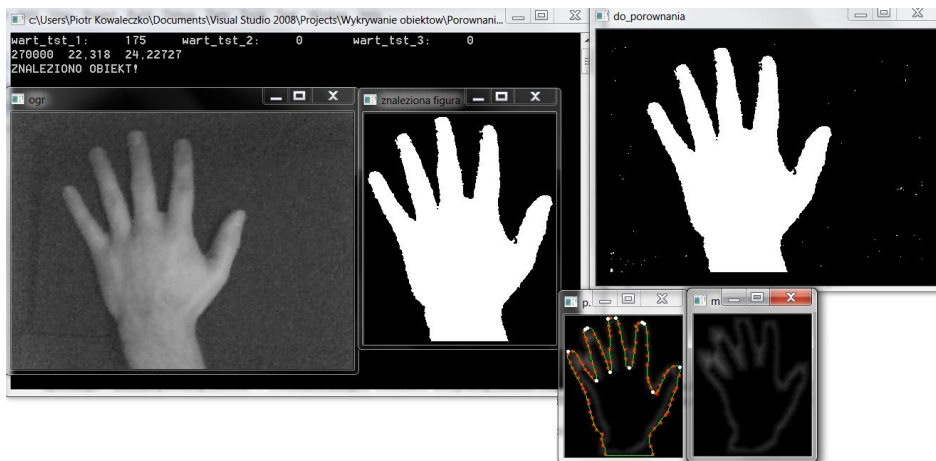
Po połączeniu wyżej opisanej metody z metodą porównania wartości próbnych z punktami transformaty konturu modelu otrzymano dość dobre narzędzie umożliwiające rozpoznanie opisanego modelem kształtu. Podkreślić należy, że metoda ta stanowi jedynie dopełnienie, ponieważ wspomaga metodę rozpoznawania na podstawie transformaty odległościowej modelu. Nie może ona być także wykorzystana w przypadku, gdy konieczne jest rozpoznawanie obiektu o zmieniającej się liczbie defektów (np. dłoń wykonująca gest łączenia się i rozłączania palców).

6. Wpływ zmian wartości parametrów testowych na poprawność rozpoznawania wzorca w obrazie video

Zaimplementowany algorytm dostosowano także do potrzeb przetwarzania obrazów pozyskiwanych z kamery. Wykorzystanie obrazu video umożliwia analizę poprawności działania algorytmu dla znacznie większej liczby danych wejściowych. Pozwala to znacznie zwiększyć prawdopodobieństwo oszacowania przydatności konkretnej metody rozpoznawania danej klasy obiektu. W dalszej części artykułu opisane zostaną wyniki otrzymane dla omawianych wcześniej metod.

W pierwszej fazie badań, obiekty pozyskane z kolejnych klatek video porównywane są z obrazem konturowym tylko jednego modelu. Umożliwiło to określenie wpływu konkretnych metod porównań na wynik rozpoznawania dla konkretnego przypadku. Jako pierwszą metodę klasyfikowania obiektu przyjęto proste sprawdzenie, czy średnia jasność wzorca dla wybranych punktów konturu jest większa od jego średniej jasności (zależność (5)). Daje ona w miarę dobre efekty. Obracając dłoń w kadrze tak, aby odpowiadała ona zarysowi konturowemu modelu, w momencie gdy obrys dłoni w pewnym stopniu pokryje się z transformatą odległościową modelu, obiekt zostaje rozpoznany. Rozpoznanie okazuje się niemożliwe przy obróceniu ręki o zbyt duży kąt, względem modelu (założenie, że wykorzystuje się tylko jeden model) oraz przy nieprawidłowej (niezgodnej z modelem) pozycji dłoni, np. zbyt szerokim lub wąskim rozłożeniem palców. Rys. 12 przedstawia aplikację realizującą omawiany algorytm. Pozyskany z kamery czarno biały obraz (okno „ogr”) poddany zostaje progowaniu (okno „do_porownania”) i indeksowaniu. Następnie każdy znaleziony obiekt poddany zostaje porównaniu z modelem. Lewe dolne okno prezentuje obrys dłoni z wyszczególnionymi punktami, dla których następuje sprawdzenie wartości transformaty modelu (który to też jest widoczny w tle). Zakwalifikowanie obiektu jako dłoni następuje przez wypisanie na konsoli odpowiedniego komunikatu (ZNALEZIONO OBIEKT!) oraz wyświetlenie okna „znaleziona figura”). Należy dodać, że przedmiotem badań była tu poprawność rozpoznawania obiektu przy warunku (5), tak więc

przyjęto założenie, że obiekty mające być poddane porównaniom będzie łatwo wydzielić z tła. W tym celu wykorzystano prosty algorytm progowania, który przy ciemnym tle i dobrze oświetlonej dłoni dał bardzo dobre wyniki.



Rys. 12. Aplikacja rozpoznająca dłonie z obrazu video

Druga faza badań polegała na dodaniu do warunku (5) warunku (7). Jego celem było wyeliminowanie możliwości klasyfikacji jako obiektów pasujących do modelu tych obiektów, których punkty charakterystyczne nie pokrywały się w znacznej mierze (wartość średnia) z transformatą odległościową konturu modelu. W przypadku przetwarzania obrazów video, metoda ta także spełniała swoją rolę. Umożliwiła ona poprawne zaklasyfikowanie dłoni, pomimo że w kadrze umieszczono także inne obiekty. Dało się jednak zauważyć wyraźny spadek liczby dokonanych rozpoznań w stosunku do podjętych prób. Po przeprowadzeniu szeregu prób okazało się, że tylko w około 36% przypadków, gdy na ekranie pojawiła się dłoń pasująca do modelu, została ona rozpoznana. Wyniki wydają się być miarodajne, gdyż sumarycznie przeprowadzono blisko 5200 prób. Przy obowiązywaniu samego warunku (5) wystarczyło jedynie w miarę podobnie (do modelu) ułożyć dłoń, aby została ona poprawnie rozpoznana. Przy uwzględnieniu ograniczenia (7), aby program zaklasyfikował dłoń, należało już dokładnie dopasować kształt jej zarysu do transformaty konturu modelu. Uwzględnienie ograniczenia poprawiającego jakość rozpoznania powoduje więc dość znaczne zmniejszenie wydajności programu. Ostatnim etapem przeprowadzonych prac było zaimplementowanie algorytmu rozpoznawania obiektu jako ludzkiej dłoni na podstawie liczby defektów. W przypadku stwierdzenia, że dany obiekt posiada 4 defekty (znaczące odchylenia konturu od figury na niej opisanej – przestrzenie między palcami), obiekt zostaje zaklasyfikowany jako dłoń. Okazało się, że tego typu ograniczenie pozwala na zaklasyfikowanie obiektu jako zgodnego z opisem aż

w 78% przypadków. Należy jednak dodać, że na analizowanym obrazie nie znajdowały się żadne inne obiekty, tak więc nie dało się zbadać rozwiązania pod względem poprawności rozpoznania dłoni (za dłoń mógł być wzięty dowolny obiekt z 4 dostatecznie dużymi defektami). Tab. 3 zawiera liczbę poprawnie przeprowadzonych rozpoznań w stosunku do podjętych prób dla kolejnych faz badań (przy różnych wprowadzonych ograniczeniach). Należy dodać, że podczas przeprowadzania badań, dłoni była układana przed kamerą w sposób zbliżony do obiektu opisanego modelem tak, aby jak najbardziej „ułatwić” algorytmowi jej poprawne rozpoznanie. Po pierwszym poprawnym rozpoznaniu wprowadzano niewielkie ruchy palców oraz obroty dłoni w różnych płaszczyznach celem zbadania odporności algorytmu na odchylenia od wzorcowego opisu obiektu. Prezentowane w tab. 3 wyniki procentowe pozwalają ocenić w jakim stopniu określone algorytmowi udaje się poprawnie rozpoznać zakłócony i zniekształcony obraz.

Tab. 3. Liczba poprawnie przeprowadzonych rozpoznań w stosunku do podjętych prób przy różnych ograniczeniach

NR PRÓBY	(5)		(5), (7)		Wykrycie 4 defektów	
1	417/546	76%	139/538	26%	415/596	70%
2	348/598	58%	118/514	23%	378/523	72%
3	258/525	49%	94/492	19%	483/517	93%
4	257/523	49%	152/517	29%	495/521	95%
5	383/512	75%	137/507	27%	368/507	73%
6	365/547	67%	129/513	25%	333/543	61%
7	253/510	50%	98/548	18%	411/532	77%
8	228/537	42%	142/523	27%	393/555	71%
9	221/614	36%	127/509	25%	395/500	79%
10	233/556	42%	134/522	26%	364/519	70%
suma	2963/5468	54%	1870/5183	36%	4135/5313	78%

Szybkość działania algorytmu była zadowalająca i zbliżona do czasu rzeczywistego. Wynikała ona ze stosunkowo niewielkiej liczby modeli, które są analizowane podczas porównań. W przypadku wykorzystania wyżej opisanych metod do analizy gestów, czy wszelkiego innego materiału, gdzie kształt rozpoznawanego obiektu może się zmieniać, liczba modeli musiałaby być znacznie większa, a co za tym idzie szybkość działania algorytmu spadłaby dość znacząco. Metody te powinny więc być stosowane wszędzie tam, gdzie szybkość działania nie jest najważniejszym kryterium.

7. Podsumowanie

W wyniku przeprowadzonych prac udało się otrzymać szkielet aplikacji umożliwiającej, przy sprzyjających warunkach (pozwalających na proste wydzielenie obiektów z tła), rozpoznanie obiektu o kształcie ludzkiej dłoni spośród wszystkich znalezionych na obrazie obiektów. Zbadano wpływ zmian parametrów algorytmu na poprawność procesu rozpoznawania i jego efektywność. Przy odpowiednim (ustalonym doświadczalnie) doborze parametrów zaimplementowany algorytm bazujący na 2-wymiarowych modelach daje dość dobre wyniki, zarówno pod względem poprawności rozpoznań, jak i wydajności działania. Wartości parametrów należy jednak dostosować do analizowanego problemu. Jeżeli zmieniłyby się warunki oświetlenia, pojawiłyby się zakłócenia czy też obiekt nie byłby widoczny w całości (np. w przypadku rozpoznawania sylwetki statku powietrznego, część maszyny schowana za chmurą), to konieczne mogłoby okazać się ponowne przeprowadzenie badań mających na celu dobór optymalnych wartości. Żaden z parametrów nie okazał się być także zdecydowanie bardziej istotny od pozostałych. Wystarczy nieodpowiednio dobrana wartość jednego, aby nawet przy optymalnych wartościach pozostałych, rozpoznanie obiektu było niemożliwe.

Literatura

- [1] TADEUSIEWICZ R., KOROHODA P., *Komputerowa analiza i przetwarzanie obrazów*, Wydawnictwo Fundacji Postępu Telekomunikacji, Kraków, 1997.
- [2] TADEUSIEWICZ R., *Rozpoznawanie obrazów*, Państwowe Wydawnictwo Naukowe, Warszawa, 1991.
- [3] PACIURA A., *Obsługa komputera za pomocą kamery internetowej. Interfejs komputerowy wykorzystujący gesty*, Praca magisterska, Politechnika Warszawska, Instytut Informatyki, 2006.
- [4] FREEMAN W., ROTH M., *Orientation Histograms for Hand Gesture Recognition*, Proc. Of the IEEE Intl. Workshop on Automatic Face and Gesture Recognition, Zürich, 1995, pp. 296-301.
- [5] HYLLESTED A. R., NILLSON M. W., *Hand Tracking Using Active Shape Models*, Advanced Image Analysis, 2004,
http://www.itu.dk/stud/projects_f2004/handtracking/report/FullReportDTU.pdf.
- [6] MALIK S., *Real-Time Hand Tracking for Interaction*, 2003
http://www.cs.toronto.edu/~smalik/downloads/2503_project_report.pdf.

- [7] ATHITSOS V., SCLAROFF S., *3d Hand Pose Estimation by Finding Appearance-Based Matches in a Large Database of Training Views*, Technical Report BU-CS-TR-2001-021, shorter version: IEEE Workshop on Cues in Communication, 2001.

Analysis of the influence of parameters changes on accuracy of object recognition process in the case of a two-dimensional model

ABSTRACT: The paper presents results of research on the influence of parameters values on the correctness and efficiency of image object recognition in the case of a two-dimensional pattern. The authors have examined such parameters as: the number of the operation of dilating the model contour, the arrangement of object testing points, the number of models used or imposing various restrictions on the object classifying process. Both graphical files and video data have been analysed.

KEYWORDS: image recognition, distance transform, two-dimensional model, influence of parameters.

Praca wpłynęła do redakcji: 18.07.2012