

Segmentacja i selekcja cech jako jeden z podstawowych problemów w rozpoznawaniu obrazów medycznych

1. Wprowadzenie

Obrazowanie medyczne (ang. *medical imaging*) to metody uzyskiwania i rozpoznawania obrazów, z jakimi mamy do czynienia w badaniach medycznych, podejmowanych zarówno w celach badawczych, jak i diagnostycznych [Chmielewski, Kulikowski, Nowakowski 2000]. Zatem, w dużym uproszczeniu, dwa główne zadania obrazowania biomedycznego to akwizycja obrazu oraz jego rozpoznawanie.

W przypadku obrazowania medycznego, *akwizycja* ma za zadanie pobranie informacji, która dotyczy badanego obiektu. Wykorzystany może być do tego celu cyfrowy rentgen, dzięki któremu otrzymywane są cyfrowe zdjęcia rentgenowskie. Innym przykładem jest tomograf. Kiedy obraz znajdzie się już w pamięci komputera, następuje faza *rozpoznawania*. Faza ta polega na obróbce obrazu, jego wstępnym przetwarzaniu oraz na segmentacji, co w rezultacie prowadzi do otrzymania reprezentacji znajdujących się na nim obiektów. Segmentacja, a tym samym uzyskanie reprezentacji jest krokiem najtrudniejszym z uwagi na to, że obrazy medyczne przedstawiają zazwyczaj wiele obiektów, których położenie, kształt i grubość zależą w dużej mierze od budowy anatomicznej pacjenta. Co więcej, obiekty na obrazie często na siebie nachodzą.

Na podstawie wyeksponowanych cech dokonywana jest analiza, która wykorzystuje teorie z innych, starszych gałęzi informatyki, takich jak teoria języków formalnych [Flasiński 1991; Ogiela 2001] czy statystyka [Stąpor 2005]. Wykorzystanie złożonych teorii czyni ten krok najciekawszym ogniwem rozpoznawania obrazów. Ogniwem, które pozbawione jest jednakże sensu bez poprawnej segmentacji.

Niniejszy artykuł ma na celu przedstawienie trudności związanych z podstawowym problemem rozpoznawania obrazów, jakim jest segmentacja i uzyskiwanie reprezentacji obiektów znajdujących się na obrazie. Proces selekcji cech obiektów na obrazie jest tak skomplikowany, ponieważ obraz zawiera zazwyczaj wiele obiektów o różnorodnym kształcie, które nachodzą na siebie, co często uniemożliwia poprawną segmentację. Za przykład posłużyć mogą kości nadgarstka przedstawione na Rycinie 1, które na siebie nachodzą, przez co ich wyodrębnienie jest niezwykle trudne [Tadeusiewicz, Ogiela, Ogiela 2006].



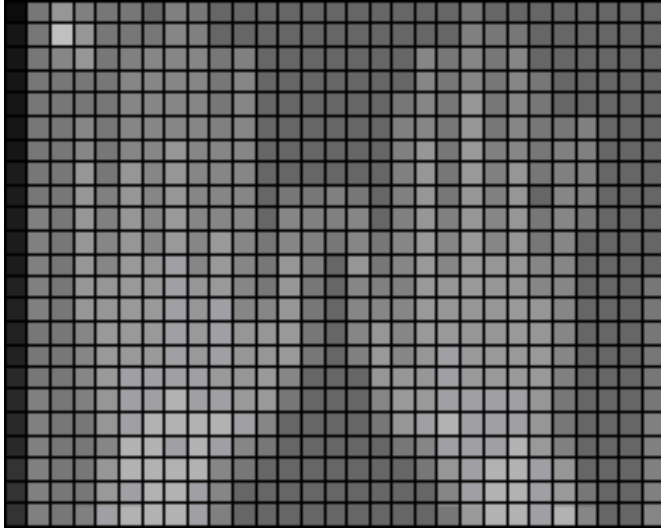
Rycina 1. Zdjęcie rentgenowskie rąk pacjenta zdrowego

W pracy tej zostanie także podjęta próba przedstawienia i porównania dwóch podobnych, a jednocześnie różnych mechanizmów rozpoznawania obrazów – komputerowego i ludzkiego, celem ukazania zachodzących między nimi analogii.

2. Akwizycja obrazu w skrócie

Zgodnie z powyższą definicją obrazowania medycznego, akwizycja obrazów jest zagadnieniem spoza rozpoznawania obrazów, niemniej ma znaczący wpływ na jego efektywność. Efektywność tę bada się za pomocą szeregu cech mających wpływ na jego praktyczne zastosowania, takich jak: jakość techniczna, użyteczność medyczna, dostępność metody zobrazowania [Choraś 2005].

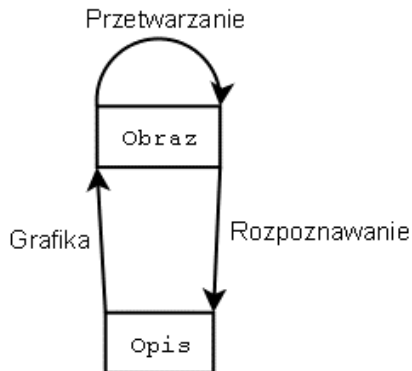
Jako że pozyskiwanie obrazów to zagadnienie bardzo obszerne i niezwiązane ściśle z niniejszą publikacją, przyjęte zostanie, bez wchodzenia w dalsze szczegóły, że obrazy wykorzystywane w rozpoznawaniu obrazów medycznych dostarczane są w postaci pionowo-poziomej kratki z odpowiednio pokolorowanymi pikselami. Taka reprezentacja obrazów nazywana jest grafiką rastrową, a jej przykład przedstawia Rycina 2. Innym rodzajem reprezentacji jest grafika wektorowa, która jednak nie ma zastosowania w obrazowaniu biomedycznym.



Rycina 2. Przykład obrazu o wymiarach 29 x 23 piksele reprezentowanego za pomocą grafiki rastrowej

3. Czym jest przetwarzanie i rozpoznawanie obrazów?

Przetwarzanie obrazów dotyczy tych zagadnień, w których zarówno dane wejściowe, jak i wyjściowe mają postać obrazów. Jednym z jego przykładów jest system przesyłania obrazów, w których istotne są usuwanie zakłóceń oraz kompresja danych. Przetwarzanie obrazów zajmuje się także obrazami prześwietlonymi albo zamazanymi, które są poprawiane dzięki technikom polepszania kontrastu.



Rycina 3. Schemat ilustrujący wzajemne zależności między grafiką, przetwarzaniem obrazów i rozpoznawaniem obrazów

Innym jego przykładem jest redukcja zakresu luminacji i otrzymanie konturu obiektów, kolejnym stworzenie nowego obrazu na podstawie ich zbioru, podobnie jak przy kontrastowaniu obrazów przekroju ludzkiego ciała na podstawie rentgenogramów [Pavlidis 1987].

Rozpoznawanie obrazów to metody tworzenia opisu obrazu wejściowego lub zakwalifikowanie obrazu do jakiejś szczególnej klasy. Proces rozpoznawania zaczyna się od obrazu, który zamienia się w abstrakcyjny opis (zbiór liczb, łańcuch symboli lub graf). Dalsze przetwarzanie takiego opisu prowadzi do zakwalifikowania danego obrazu do jednej z kilku klas. Przykładem może tu być zakwalifikowanie pacjenta, którego zdjęcie widać na Rycinie 1, do klasy pacjentów zdrowych, a pacjenta, którego zdjęcie rentgenowskie znajduje się na Rycinie 4, do klasy pacjentów chorych.



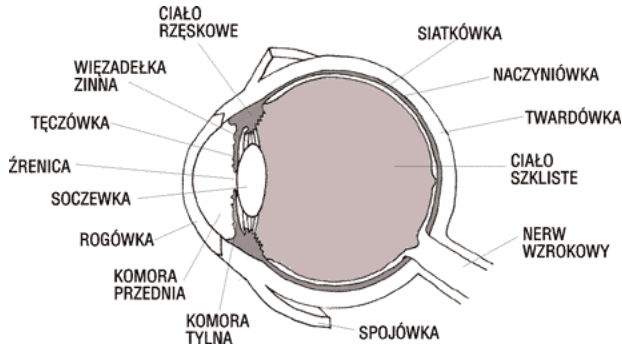
Rycina 4. Zdjęcie rentgenowskie ręki pacjenta chorego

Definiując pojęcie rozpoznawania obrazów, warto nadmienić, że jego nazwa nie-
zbyt precyzyjnie oddaje sens angielskiego wyrażenia *pattern recognition* (rozpozna-
wanie wzorca), od którego się wywodzi. Wzorcem bowiem może być nie tylko rysunek,
ale też litera rękopisu, odcisk palca, stan pacjenta poddawanego diagnostyce, czy
też sygnał mowy [Tadeusiewicz, Flasiński 1991]. W niniejszej publikacji ograniczono
się jednak do rozpatrywania pewnej grupy wzorców, które są obrazami medycznymi.
Mimo iż może się wydawać, że takie postępowanie zawęży obszar rozważań, to
należy zdawać sobie sprawę z tego, że tak zdefiniowany obraz może zawierać bardzo

wiele różnorodnych zdjęć. Mogą to być m.in. zdjęcia tomograficzne, mammograficzne czy rentgenowskie, ale również zdjęcia przedstawiające zmiany skórne, wykonane zwykłym aparatem cyfrowym.

4. Przetwarzanie i rozpoznawanie obrazów przez człowieka

Funkcję analogiczną do funkcji aparatu cyfrowego pełni u człowieka gałka oczna (Rycina 5). Gałka oczna jest narzędziem, które za pomocą nerwów wzrokowych przekazuje wrażenia wzrokowe do mózgu, gdzie są one analizowane i przetwarzane w jeden obraz. Oko ma za zadanie oglądanie, mózg natomiast decyduje o tym, co jest spostrzegane i w jaki sposób rejestrowane [Choraś 2005].



Rycina 5. Budowa gałki ocznej człowieka

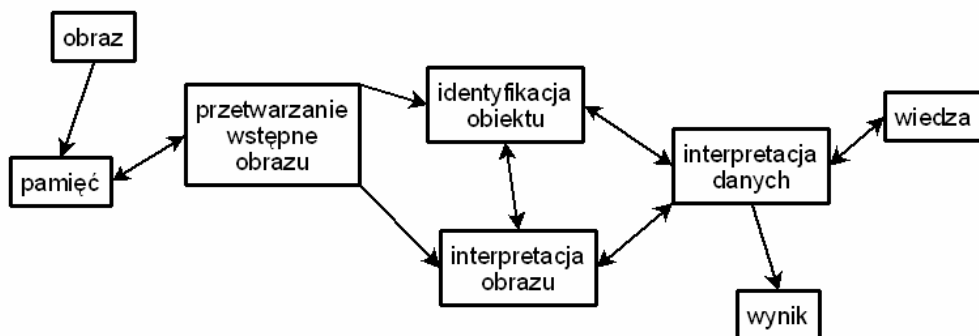
Na samym początku tak otrzymane, bogate w szczegóły odwzorowanie obrazu jest przechowywane w pamięci ultrakrótkotrwałej, która w przypadku przetwarzania bodźców wzrokowych zwana jest pamięcią ikoniczną. Odwzorowanie jest dostępne w pamięci ikonicznej w celu jego identyfikacji. Identyfikacja taka ma charakter symultaniczny, ponieważ w ramach wstępnego przetwarzania zostają wyodrębnione poszczególne elementy obrazu ze względu na ich lokalizację przestrzenną, a także kształt, wielkość i barwę.

Człowiek oglądający obraz nie tylko zapamiętuje odbieraną informację o obrazie, ale wykorzystując odpowiednie obszary mózgu, dokonuje jego interpretacji. Interpretacja dokonywana jest nie tylko na podstawie odebranego sygnału obrazu, ale też w odniesieniu do wcześniej posiadanych wiadomości o tym, co obraz powinien przedstawiać. Porównanie informacji o obrazie z informacją już znajdującą się w pamięci umożliwia interpretację obrazu.

Człowiek przyjmuje do analizy tylko tyle danych z pamięci ultrakrótkotrwałej, ile jest niezbędnych do potwierdzenia hipotezy dotyczącej obrazu i uzupełnia brakującą o nim informację danymi ze swojej pamięci długotrwałej, w której znaj-

duje się model obrazu. Takie uzupełnianie informacji nie jest jednak pełnym i kompletnym przeglądem znajdujących się w pamięci długotrwałej danych dotyczących obrazu, ponieważ byłby to proces zbyt wolny [Choraś 2005].

Opisany powyżej schemat ludzkiego rozpoznawania obrazów przedstawia Rycina 6.



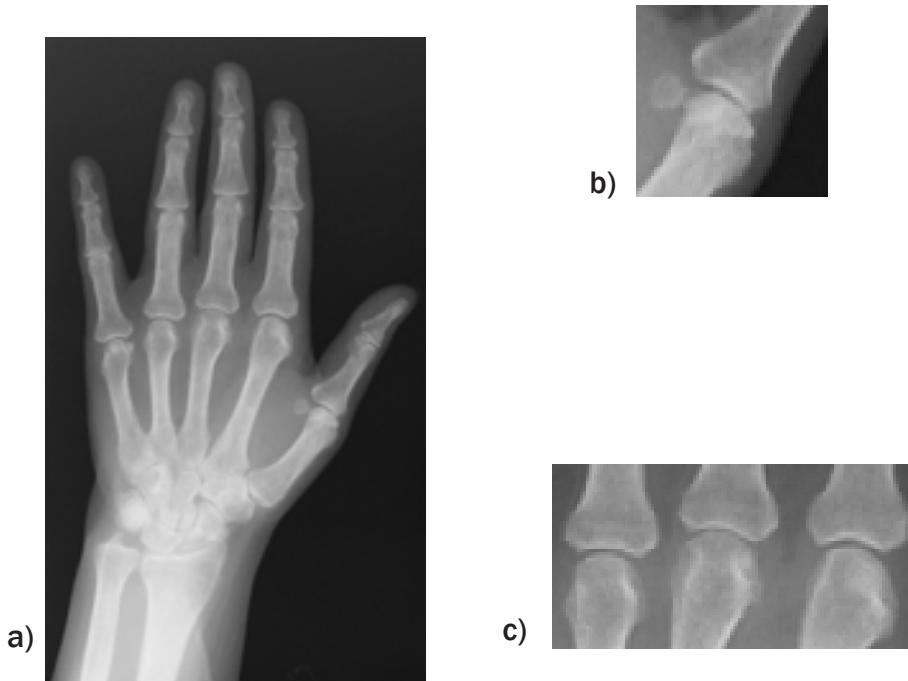
Rycina 6. Schemat rozpoznawania obrazu przez człowieka

Należy także dodać, iż na selekcję i interpretację obrazów u człowieka wpływają jego emocje i motywacje.

Niech za przykład posłuży proces rozpoznawania choroby reumatologicznej na podstawie rentgenowskich zdjęć rąk. Prawdopodobna diagnoza dla pacjenta, którego rentgen przedstawiony jest na Rycinie 2, byłaby optymistyczna, a dla pacjenta, którego rękę przedstawia Rycina 4, pesymistyczna, co byłoby zgodne z prawdą. Można sobie jednak wyobrazić sytuację, że lekarz ma do przeanalizowania wiele zdjęć rentgenowskich i nie ma czasu zagłębiać się w szczegóły każdego z nich. Wtedy pacjent z rentgenem z Ryciny 7a mógłby zostać nieprawidłowo zdiagnozowany, ponieważ na pierwszy rzut oka nie występują tam znaczące patologie.

Skoro patologiczne zmiany nie są widoczne na pierwszy rzut oka, lekarz pobieżnie oglądający zdjęcie stwierdziłby, że wszystko jest w normie i niejako uzupełniłby brakującą, niepokojącą informację o pacjencie, danymi ze swojej pamięci długotrwałej. Taka pomyłka mogłaby okazać się tragiczna w skutkach, ponieważ po dłuższej analizie można zauważyć na zdjęciu z Ryciny 7a pierwsze symptomy reumatoidalnego zapalenia stawów, które objawiają się w tym konkretnym przypadku węższymi szparami śródrečno-palczkowymi kciuka (Rycina 7b) i palca środkowego (Rycina 7c).

Choroba ta jest jednak widoczna dopiero po dokładnym przyjrzeniu się szparom stawowym. Lekarz musiałby w tym wypadku jeszcze raz uważniej spojrzeć na zdjęcie, ponieważ tego typu informacja nie została zapisana w pamięci długotrwałej za pierwszym razem. Przykłady tak otrzymanych zbliżeń, niosących dodatkową informację, przedstawione są na Rycinie 7b i Rycinie 7c.

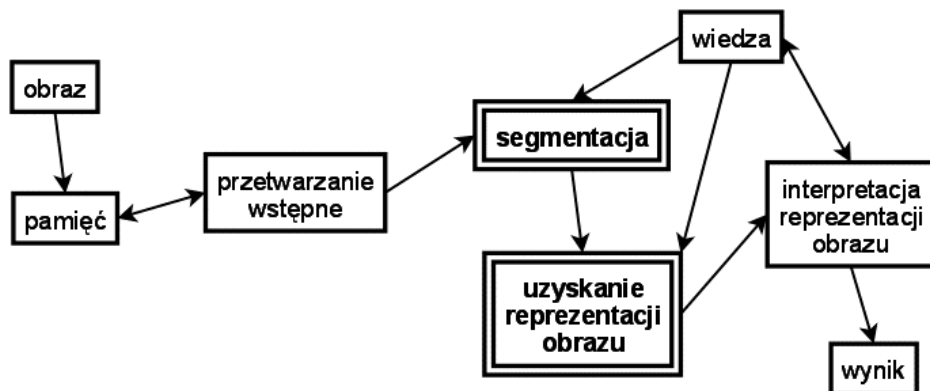


Rycina 7. a) Zdjęcie rentgenowskie ręki pacjenta chorego na reumatoidalne zapalenie stawów, powiększenie b) śródrečno-paliczkowej szpary stawowej kciuka oraz c) palca serdecznego, środkowego i wskazującego

5. Komputerowe przetwarzanie i rozpoznawanie obrazów

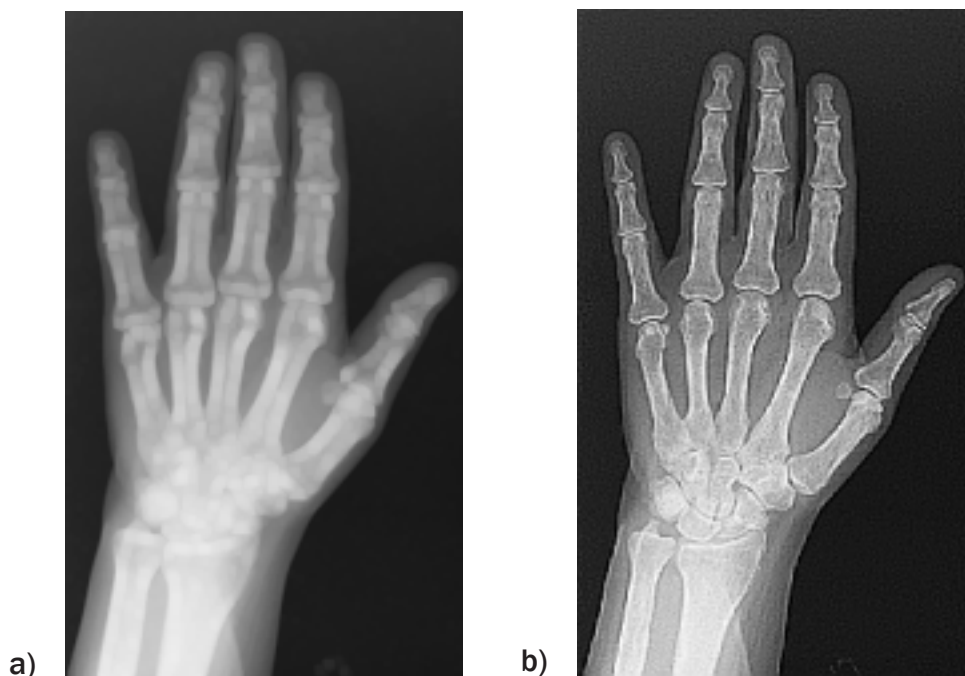
Opisując proces przetwarzania i rozpoznawania informacji przez komputer, należy przede wszystkim powiedzieć, że jest on – w dużym uproszczeniu – analogiczny do procesu zachodzącego u człowieka, który to proces został opisany w poprzedniej części. Jednocześnie system komputerowy posiada zazwyczaj znaczne ograniczenia w porównaniu z człowiekiem.

Rozważmy ogólny schemat rozpoznawania obrazów przez komputer, który przedstawiony jest na Rycinie 8. Mimo że obraz również jest zapisywany do pamięci, co przebiega analogicznie jak u człowieka, to jest on tam zapisywany jako całość. W przeciwieństwie do człowieka komputer nic sobie nie „dopowiada”, co jest oczywiście zaletą, bo nie prowadzi do błędów spowodowanych powstałymi złudzeniami, ale z drugiej strony jest wolniejsze. Co więcej, komputer jest zdolny odnieść się w każdej chwili do wybranego fragmentu obrazu i dokładnie zbadać jego właściwości. U człowieka nastąpić musiałoby powtórne pobranie informacji wzrokowej, która tym razem zawierałaby potrzebną informację.



Rycina 8. Schemat rozpoznawania obrazu przez komputer

Kolejnym krokiem analizy jest *przetwarzanie wstępne*. Podczas tego etapu obraz jest wygładzany, kontrastowany (Rycina 9b) lub poddawany erozji (Rycina 9a). Jak już zostało wspomniane, najważniejszą własnością metod przetwarzania obrazów jest to, że pobierają one do obróbki obraz i obraz zwracają jako wynik przeprowadzonej operacji.



Rycina 9. Obraz z Ryciny 7a poddany a) operacji morfologicznej erozji i b) operacji kontrastowania

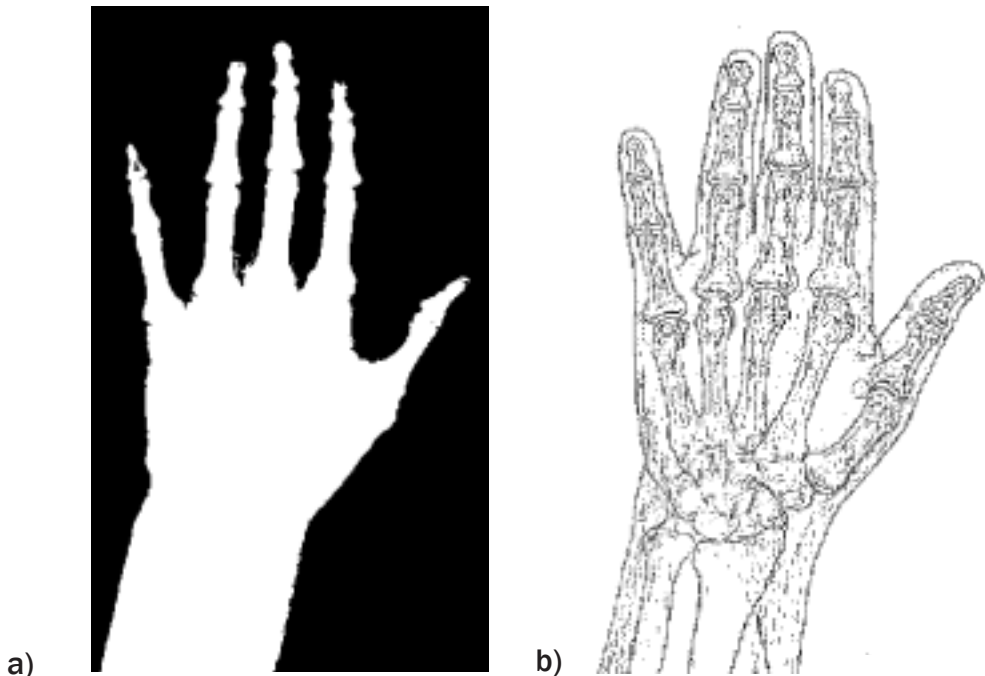
Dzięki odpowiednio przetworzonemu obrazowi możliwa jest z kolei *segmentacja*, która stanowi najtrudniejszą fazę rozpoznawania obrazów. Polega ona na oddzieleniu obiektów od tła i od siebie samych. Dwie najpopularniejsze metody segmentacji to progowanie i wykrywanie krawędzi.

Progowanie (thresholding) polega na porównaniu wartości każdego punktu obrazu z zadaną wartością progową. Odpowiedni wybór wartości progowej umożliwia wyodrębnienie obszarów, na których znajdują się obiekty. Najprostsza metoda stosuje jedną wartość progę i jest realizowana według zależności:

$$J(x,y) = \begin{cases} 1, & \text{gdy } I(x,y) > t \\ 0, & \text{gdy } I(x,y) < t \end{cases} \quad (1)$$

– gdzie t to próg binaryzacji.

Operacja (1) oznacza transformację obrazu źródłowego w odcieniach szarości na obraz binarny. Natomiast punkty, dla których $J(x,y) = 1$, oznaczają zazwyczaj w nowym obrazie punkty obiektu [Malina, Smiatacz 2005]. Prosta operacja progowania jest bardzo nieefektywna, co zauważyć można na Rycinie 10a. Dlatego też zazwyczaj wykorzystuje się ją jedynie do wstępnej segmentacji, która zawiera obiekt i pewne fragmenty tła.

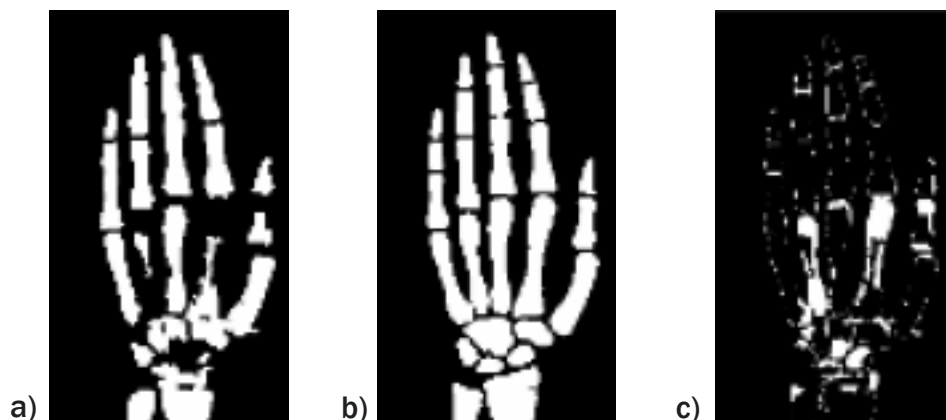


Rycina 10. Obraz z Ryciny 7a poddany a) operacji progowania przy $t = 100$ i b) operacji wykrywania krawędzi za pomocą filtra gradientowego

Wykrywanie krawędzi w obrazach o różnych poziomach kontrastu jest często wykonywane za pomocą filtrów gradientowych. Najogólniej, krawędzie mogą być wykryte poprzez lokalizację lokalnego ekstremum pierwszej pochodnej obrazu I lub wybranych miejsc zerowych drugiej pochodnej. Gradient i laplasjan są podstawowymi funkcjami używanymi do konstrukcji filtrów detekcji krawędzi. Tego typu operacje nie są jednak wykorzystywane bezpośrednio w rozpoznawaniu obrazów medycznych, również ze względu na ich nieefektywność (Rycina 10b).

Ponieważ klasyczne metody segmentacji często zawodzą, kiedy zastosowane do obrazów medycznych tworzone są algorytmy dedykowane konkretnym zastosowaniom. Przykładem takiego algorytmu jest algorytm *bottom-up region merging* [Beier, Lehmann, Seidl, Thies 2005], którego wyniki pokazane są na Rycinie 11. W algorytmie tym podjęta została próba segmentacji z wykorzystaniem skomplikowanych metod morfologicznych oraz wiedzy o budowie anatomicznej ręki. Wyniki nie były jednak zadowalające, ponieważ segmenty odpowiadające poszczególnym paliczkom zlewały się z sobą.

Z uwagi na ogromną trudność otrzymania prawidłowej segmentacji, czasami rezygnuje się z niej na rzecz uzyskania jedynie pewnych cech obiektów znajdujących się na obrazie. Przykładem takiego podejścia jest automatyczne wyznaczenie minimalnej szerokości szpar stawowych zaproponowane przez autora w pracy [Zieliński 2007]. Zrezygowano w niej z karkołomnej operacji segmentacji na rzecz wyszukiwania newralgicznych miejsc obrazu i ich wyłącznej analizy.



Rycina 11. a) Wyniki wykorzystania algorytmu *bottom-up region merging* do segmentacji rentgenowskich zdjęć rąk, b) prawidłowa segmentacja, c) różnica między uzyskanym a prawidłowym wynikiem

Opisana w przytoczonej pracy metoda polegała na uzyskaniu szkieletu ręki, a następnie wyodrębnieniu z niego pięciu gałęzi, które odpowiadają pięciu palcom. Tak uzyskane gałęzie przechodzą przez środek każdego palca (Rycina 12a) i pozwalają uzyskać lokalizację szpar stawowych, które są w tym palcu umiejscowione (Rycina 12b). Następnie tworzone są obrysy górnej i dolnej powierzchni stawowej, które są w tym przypadku *reprezentacją obrazu*.



Rycina 12. Kolejne kroki automatycznego wyznaczania szerokości szpar stawowych: a) stworzenie szkieletu, b) lokalizacja szpar stawowych, c) generowanie obrysu górnej i dolnej powierzchni stawowej

Reprezentacją nie musi być opis wszystkiego, co znajduje się na obrazie. Wystarczy, że znajdują się w niej wyłącznie opisy potrzebne do dalszej analizy.

Ważną cechą współczesnych metod rozpoznawania obrazów jest to, że w kolejnych krokach prowadzących do uzyskania reprezentacji obrazu informacja o strukturze ręki jest cały czas wykorzystywana. Komputer „zdaje sobie sprawę” z tego, że ręka ma pięć palców, a palec ma trzy paliczki (z wyjątkiem kciuka, który ma dwa), bo te informacje są w pewien sposób zaprogramowane. Dzięki temu mogą one być w dalszej kolejności wykorzystywane w czasie segmentacji i selekcji cech. Stąd strzałka od „wiedza” do „segmentacja” i od „wiedza” do „uzyskiwanie reprezentacji obrazu” na Rycinie 8.

6. Podsumowanie

W niniejszej publikacji przedstawiono i porównano dwa złożone systemy rozpoznawania obrazów – komputerowy i ludzki. Okazuje się, że są one bardzo zbliżone, co spowodowane jest tym, że systemy komputerowe od samego początku tworzone były z myślą o tym, by czerpać z wiedzy o systemach ludzkich.

Dlatego też przed rozpoczęciem rozpoznawania obrazu następuje jego przetworzenie, podobnie jak ma to miejsce w gałce ocznej. Z tego samego powodu podczas tworzenia reprezentacji obiektów znajdujących się na obrazie cały czas wykorzystywana jest wiedza o tym, co znajduje się na obrazie – tak zwany model obrazu.

Warto podkreślić także różnice między oboma systemami. System ludzki jest w naturalny sposób dużo szybszy od komputerowego, ale też posiada niedoskonałości występujące podczas akwizycji obrazu. Do powyższego dochodzi też błąd, polegający na „dopowiadaniu” brakujących części obrazu.

Podsumowując, komputerowy system rozpoznawania obrazów imituje system występujący u człowieka, jednocześnie usprawniając proces akwizycji obrazu. Problem pojawia się jednak, gdy dochodzi do fazy segmentacji i selekcji cech. Okazuje się wtedy, że komputer „nie posiada” wystarczającej wiedzy na temat danego zagadnienia, bez względu na to, jak zmyślny algorytm zostanie zastosowany do implementacji jego zachowania. Źródłem tego braku jest fakt, że systemy komputerowego rozpoznawania obrazów wciąż nie są inteligentne.

BIBLIOGRAFIA

- Beier D., Lehmann T., Seidl T., Thies C. 2005, *Segmentation of medical images combining local, regional, global, and hierarchical distances into a bottom-up region merging scheme*, Proceedings of SPIE, nr 5747, s. 546–555.
- Chmielewski L., Kulikowski J.L., Nowakowski A. 2000, *Biocybernetyka i inżynieria biomedyczna, Obrazowanie biomedyczne*, Exit, Warszawa.
- Choraś R. 2005, *Komputerowa wizja – metody interpretacji i identyfikacji obiektów*, Exit, Warszawa.
- Flasiński M. 1991, *Syntaktyczne metody rozpoznawania obrazów*, Skrypty Uczelniane UJ, nr 634, Kraków.
- Malina W., Smiatcz M. 2005, *Metody cyfrowego przetwarzania obrazów*, Exit, Warszawa.
- Ogiela M. 2001, *Syntaktyczne metody rozpoznawania obrazów i ich wykorzystanie w analizie wybranych obrazów medycznych*, Wydawnictwa AGH, Kraków.
- Pavlidis T. 1987, *Grafika i przetwarzanie obrazów*, Wydawnictwa Naukowo-Techniczne, Warszawa.
- Stapor K. 2005, *Automatyczna klasyfikacja obiektów*, Exit, Warszawa.
- Tadeusiewicz R., Flasiński M. 1991, *Rozpoznawanie obrazów*, PWN, Warszawa.
- Tadeusiewicz R., Ogiela L., Ogiela M. 2006, *Image languages in intelligent radiological palm diagnostics*, „Pattern Recognition”, nr 39/11, s. 2157–2165.
- Zieliński B. 2007, *A Fully-Automated algorithm dedicated to computing metacarpophalangeal and interphalangeal joint cavity widths*, „Schedae Informaticae”, nr 16, s. 47–67.