

MICIAK Mirosław, WIATR Roman

METODY PRZETWARZANIA OBRAZU PRZESYŁEK POCZTOWYCH

Streszczenie

W referacie przedstawiono metody przetwarzania obrazów i rozpoznawania znaków w szczególności dla celów pocztowych. W procesie opracowania przesyłek pocztowych w urzędach rozdzielczych szczególną rolę ogrywa odczytanie strony adresowej przesyłki w zwłaszcza opisanej pismem odręcznym. Ze względu na czas trwania procesu przetwarzania obrazu przesyłki, ograniczono się do odczytania pocztowego numeru adresowego (PNA). Dokonano analizy stosowanych rozwiązań w zakresie metod obszarowych i konturowych. Zwrócono szczególną uwagę na metody wykorzystujące sztuczne sieci neuronowe, ukryte modele Markowa, momenty geometryczne, przekształcenie Fouriera. Określono przydatność poszczególnych metod dla specyficznych wymagań przetwarzania obrazu w systemach pocztowych.

WSTĘP

Ze względu na dużą liczbę typów przesyłek w celu usprawnienia łączności pocztowej dąży się do automatyzacji wielu operacji pocztowych; stosuje się automaty do sprzedaży znaczków i kart, maszyny frankujące, kody kreskowe oraz systemy automatycznego sortowania i rozdziału (Rys. 1).



Rys. 1. System sortowania przesyłek pocztowych

Stosowane obecnie rozwiązania w systemach automatycznego opracowywania (sortowania) przesyłek pocztowych wykorzystują mechanizmy OCR (Optical Character Recognition). Moduł OCR jest kluczowym elementem procesu odpowiedzialnym za rozpoznawanie danych z pola adresowego, rozpoznawanie kodów paskowych i sprawdzenie poprawności opłaty pocztowej. Typowy system sortowania stosowany w Centrach Ekspedycyjno - Rozdzielczych (Rys. 2) składa się z modułu akwizycyjnego, modułu wideokodowania oraz części rozpoznającej pismo. Moduł akwizycji wysyła obraz przesyłki

do segmentu OCR. Jeżeli system OCR jest w stanie rozpoznać dane (technologia ta zapewnia około 50% skuteczność dla wszystkich przesyłek), to automatycznie odsyła je do systemu. W przypadku gdy odczyt danych nie jest możliwy, to obraz przesyłki trafia do modułu wideokodowania, gdzie operator odczytuje dane i wpisuje je do systemu [5].

Główne problemy przedstawionego rozwiązania:

- duża liczba przesyłek opracowanych ręcznie 300000-500000 (na 1,5 do 2 mln wszystkich przesyłek opracowywanych dziennie),
- ręczny rozdział przesyłek poleconych. Każda ma numer, musi być policzona i sprawdzona, w ciągu doby opracowanych i liczonych jest od 250 000 do 380 000 listów poleconych,
- sortowane są tylko przesyłki standardowe, typowe, ze znaczkami w prawym górnym rogu, adresem wpisanym w odpowiedniej kolejności: imię, nazwisko, ulica, kod pocztowy i miasto (bez podkreśleń),
- często zdarzają się jednak listy nietypowe (za grube, okrągłe), ale też niestarannie i nieczytelnie zaadresowane, ze złym kodem lub w ogóle bez niego.



Rys. 2. Schemat blokowy typowego systemu sortowania przesyłek pocztowych

Prowadzone są prace w celu podniesienia skuteczności modułu OCR w dziedzinie odczytywania pisma ręcznego. Pomimo, iż uzyskano zadowalające rezultaty z rozpoznawaniem pisma drukowanego, to pismo ręczne jest nadal trudne do interpretacji. Biorąc pod uwagę fakt, że przesyłki opisane ręcznie stanowią 30% całości przesyłek ważne jest, aby zapewnić lepsze wykorzystanie możliwości segmentu rozpoznającego pismo ręczne.

W związku z powyższym dokonano analizy problemów związanych z funkcjonowaniem modułu optycznego rozpoznawania znaków w systemach pocztowych i wskazano szereg problemów, do których można zaliczyć: zniekształcenia procesu akwizycji, niejednorodne tło, zmiany oświetlenia, szumy, obrót, zmiana skali, uszkodzone znaki z brakującymi elementami, przekreślenia itp.

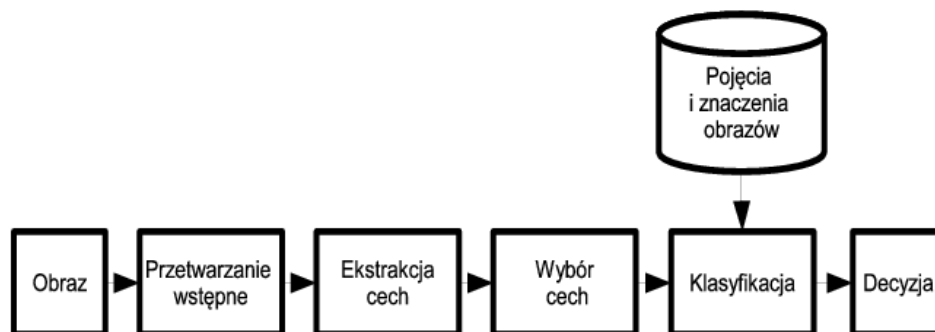
1. LOKALIZACJA DANYCH TELEADRESOWYCH I OPŁATY POCZTOWEJ

Przesyłka pocztowa nim dotrze do adresata, musi zostać opatrzona jego adresem pocztowym. Wzór adresowania zawiera: imię i nazwisko adresata lub nazwę instytucji, miejscowość zamieszkania, położenie w tej miejscowości (ulica, numer budynku, numer mieszkania) oraz kod pocztowy. Większość przesyłek posiada specjalne pole z miejscem na wpisanie adresata. W celu odnalezienia danych teleadresowych na przesyłce pocztowej należy oddzielić elementy zawierające tekst od elementów grafiki, ponieważ na przesyłce bardzo często znajdują się inne obiekty takie jak pieczęci, znaczki, reklamy, logo firmy. Powszechnie w tym celu stosuje się algorytmy morfologiczne [19], algorytmy oparte na izolowaniu

znaków na podstawie cech izotropowych [20], algorytmy bazujące na analizie tła [21], profilu rzutowania [11], czy algorytmy oparte na rozpoznawaniu znaków w izolowanych obiektach [4]. Działanie zaproponowanego algorytmu oparte jest na badaniu kolejnych fragmentów obrazu przesyłki, gdzie uzyskiwane są dwie wielkości: liczba punktów obiektu (1), oraz liczba zmian z 0 na 1 oraz z 1 na 0. Obie wielkości są mnożone, obszar, dla którego pomiar osiągnął wartość maksymalną jest klasyfikowany jako obszar, w którym znajduje się tekst [10]. Dodatkowo stosuje się moduł decyzyjny, który pozwoli określić obszary gdzie znajdują się dane adresata.

2. PROBLEMATYKA PRZETWARZANIA OBRAZÓW

Problematyka przetwarzania obrazów swoim zasięgiem obejmuje zagadnienia związane z dyskretną reprezentacją obrazu, kodowaniem, lokalizacją obiektów, rozpoznawaniem, klasyfikacją czy transmisją w sieci telekomunikacyjnej. Procesy związane z rozpoznawaniem obrazu przedstawiono na rysunku 3.



Rys. 3. Schemat procesu rozpoznawania obrazów

Obraz w postaci cyfrowej poddawany jest przetwarzaniu wstępnemu. Następnie realizowany jest proces ekstrakcji cech charakterystycznych, które są reprezentacją obiektów obrazu. Otrzymane wielkości są selekcjonowane i kodowane dla potrzeb systemu (np. obrazowej bazy danych) i stanowią wektor cech. Zatem w procesie rozpoznawania można wyróżnić etap pozyskiwania i selekcji cech (nazywany też ekstrakcją cech) oraz etap klasyfikacji obiektu do konkretnej klasy. W przypadku obrazu można zdefiniować go, jako zbiór pewnych charakterystycznych właściwości (cech). Cechy można przedstawić zazwyczaj w wielowymiarowym układzie współrzędnych. Obiekty obrazu reprezentowane są przez n cech. Cechy te tworzą n wymiarową przestrzeń cech, w której każda współrzędna reprezentuje odpowiednią cechę. Selekcja cech jest istotnym elementem systemu, stosowana jest w celu znalezienia reprezentacji danych w przestrzeni o mniejszej liczbie wymiarów, w której dane należące do poszczególnych klas będą lepiej separowalne. Niezbędnym elementem tego etapu jest segmentacja obrazu [7] (ang. image segmentation), która umożliwia podział obrazu na regiony homogeniczne pod względem pewnej własności (np. intensywność, kolor, tekstura).

Wybrane metody segmentacji są z powodzeniem stosowane w systemach optycznego rozpoznawania znaków [27]. Istnieje bliski związek pomiędzy omówionym zadaniem rozpoznawania obrazów a zagadnieniem rozpoznawania pisma, w związku, z czym rozwiązania opracowane dla identyfikacji obrazów mogą być przynajmniej w części adaptowane dla celów rozpoznawania pisma ręcznego czy maszynowego. Prace nad problemem rozpoznawania pisma trwają od wielu lat, czego efektem jest wysoka skuteczność systemów odczytujących pismo maszynowe (przekracza poziom 98%). Niemniej jednak w dalszym ciągu prowadzi się prace nad efektywnością rozwiązań przetwarzających pismo, w

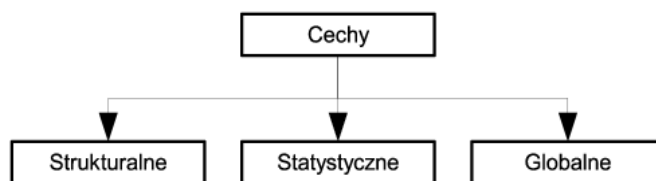
szczegółności uwzględniając jego ręczną odmianę. W aplikacjach pojawiają się trudności związane z technicznymi aspektami procesu akwizycji tekstu. Duże znaczenie mają zakłócenia w postaci różnego rodzaju szumów powstających w samym procesie akwizycji oraz zniekształcenia znajdujące się bezpośrednio na nośniku zawierającym tekst. Zastosowanie odpowiednich filtrów wpływa na podniesienie jakości pracy systemu lecz generuje większe zapotrzebowania na moc obliczeniową danego rozwiązania. Doskonałym przykładem są tutaj specyficzne warunki pracy systemów pocztowych, gdzie ze względu na dużą różnorodność przesyłek i sposobów adresowania pojawiają się wymienione problemy, co w rezultacie prowadzi do zmniejszenia wydajności Centrów Ekspedycyjno Rozdzielczych w zakresie liczby opracowanych przesyłek.

Zasadniczy wpływ na powstawanie błędów ma:

- rodzaj nośnika na którym jest wykonywane pismo (materiał, gramatura, struktura itp.),
- kolorowe tło lub zawierające teksturę,
- rodzaj i kolor środka piśmienniczego (atrament, długopis, druk igłowy itp.),
- nierówność oświetlenia w procesie akwizycji,
- plamy, pożółknięcia papieru, przekreślenia i inne lokalne zaburzenia tła,
- rozmycie krawędzi znaków (np. słaba jakość druku).

Algorytmy segmentacji i rozpoznawania zaimplementowane w systemach OCR nie są w stanie prawidłowo oszacować kształtu znaków, jeśli rozdzielczość obrazu otrzymana w procesie akwizycji nie jest wystarczająca duża. Np. dla dokumentów drukowanych czcionką (10 lub 12 pt.) zwykle minimalna rozdzielczość gwarantująca prawidłową pracę wynosi 300 DPI. Niedostateczna rozdzielczość może wystąpić w przypadku, gdy obraz jest pozyskiwany z kamery analogowej lub aparatu cyfrowego ze znacznej odległości. Na rozdzielczość obrazu ma również wpływ format kodowania obrazu, np. ograniczona głębia koloru lub ograniczona liczba odcieni szarości może spowodować błędy binaryzacji. Zapis obrazu z użyciem kompresji stratnej może powodować błędy w rozpoznawaniu kształtu znaków, ponieważ algorytmy kompresji stratnej powodują np. rozmycie krawędzi.

Architektura systemów rozpoznawania pisma dla celów pocztowych, podobnie jak w przypadku rozpoznawania obrazów oparta jest głównie o klasyfikatory, które umożliwiają określenie przynależności badanego znaku do właściwej klasy. Możliwe jest to w oparciu o zbiór pewnych własności znaków otrzymanych w etapie pozyskiwania i selekcji cech. W idealnym przypadku zbiór parametrów opisujących znak powinien zawierać tylko niezbędne atrybuty umożliwiające poprawne przydzielenie znaku do określonej klasy.



Rys. 4. Podział metod rozpoznawania znaków ze względu na rodzaj wydzielanych cech

Najważniejsze typy cech pisma mogą być zgrupowane następująco [12] (Rys. 4):

- Cechy strukturalne - opisujące geometryczne i topologiczne własności globalnych cech danego obiektu. Do cech strukturalnych wykorzystywanych w metodach rozpoznawania znaków można zaliczyć m.in. :współczynnik proporcji wysokości i szerokości znaku, liczbę punktów przecięcia w znaku, czy też liczbę pętli i rozgałęzień danego znaku. Cechy statystyczne - opierają się głównie na badaniu wielkości związanych z rozmieszczeniem poszczególnych punktów należących do badanego znaku. Do tej grupy cech można zaliczyć dane otrzymane w wyniku analizy statystycznej obrazu.

- Cechy globalne - techniki pozyskiwania cech bazujące na stosowaniu różnego rodzaju transformat pozwalają na redukcję przestrzeni cech i umożliwiają osiągnięcie niezmienności względem przesunięcia, zmiany skali i obrotu. Do grupy przekształceń wykorzystywanych w opracowaniach dotyczących rozpoznawania pisma można zaliczyć: transformatę Fouriera, DWT, Radona, momenty centralne i Zernike, czy też kod łańcuchowy.

Analiza publikacji dotyczących tematyki przetwarzania obrazów i rozpoznawania znaków w szczególności dla celów pocztowych pozwala na stwierdzenie, iż metody rozpoznawania znaków w ogólności bazują na informacji o kształcie. Zatem można je podzielić na:

- Metody konturowe - są łatwe w implementacji, ale bardzo wrażliwe na zakłócenia. W większości bazują na doskonale znanej transformacie Fouriera. Niestety w przypadku niektórych zastosowań w obrazach występują duże zniekształcenia (szумы, bardzo zróżnicowane tło – szare, białe, jasno brązowe, bardzo różny kolor znaków, znaki z dobrze widocznymi fragmentami itp.) dlatego skuteczność tej metody jest niewielka.
- Metody obszarowe wykorzystujące w procesie rozpoznawania znaków najczęściej różnego rodzaju momenty. Wymagają one binaryzacji, normalizacji i dużego nakładu obliczeniowego. Są również bardziej odporne na zakłócenia obrazu.

3. METODY PRZETWARZANIA OBRAZU

Większość przeanalizowanych technik rozpoznawania znaków z wykorzystaniem metod przetwarzania obrazów opiera swoje działanie na porównywaniu wydzielonych cech z badanego obrazu ze wzorcami umieszczonymi w obrazowej bazie danych. Kluczowym zagadnieniem techniki rozpoznawania znaków jest odpowiedni wybór cech, tak aby system mógł działać w czasie rzeczywistym, a wydzielone cechy pozwalały na skuteczną klasyfikację, co jest bardzo istotne w przypadku zastosowań pocztowych. Przeprowadzona analiza powszechnie stosowanych metod rozpoznawania pisma pozwoliła na wskazanie pewnych najliczniej reprezentowanych technik. Poniżej zaprezentowano główne narzędzia wykorzystywane w zadaniu rozpoznawania znaków:

- Metody z wykorzystaniem sztucznych sieci neuronowych - to najczęstsze i najbardziej popularne techniki wykorzystywane w zadaniu rozpoznawania znaków. Pierwsze efektywne rozwiązania systemów automatycznego rozpoznawania zaprezentowano w 1989 roku w pracy [13]. Przedstawiona aplikacja wykorzystywała wielowarstwową sieć neuronową wspieraną przez algorytm uczenia z nauczycielem. Autorzy publikacji [15] zaproponowali system rozpoznawania dla znaków uzyskanych w kolejnych etapach przetwarzania przedstawionych, w [25], które następnie poddano liniowej normalizacji wielkości do wymiarów 16 na 16 punktów, oraz ograniczono wartości jaskrawości do przedziału -1 do 1. Zastosowana tam sieć składała się z czterech ukrytych warstw [14] i praktyce pozwalała na osiągnięcie skuteczności w granicach 93% dla pisma ręcznego oraz niemalże 100% dla znaków drukowanych. Późniejsze rozwiązania bazujące głównie na sieciach ze sprzężeniem feed-forward, składające się z perceptronów uformowanych w architekturę wielowarstwową MLP (ang. multi-layer-perceptron) [3] oraz sieci z funkcjami o symetrii kołowej RBF (ang. radial-basis function) [6] spowodowały, że wielowarstwowe sieci nieliniowe były często stosowane w aplikacjach rozpoznawania pisma ręcznego. Popularność metod opartych o sieci MLP wynika przede wszystkim z stosunkowo łatwego procesu uczenia sieci oraz szybkiego procesu decyzyjnego w trakcie klasyfikacji.

Wadami większości z powyżej przedstawionych metod są przede wszystkim złożone procesy optymalizacji dla poszczególnych klasyfikatorów oraz problem z właściwym określeniem lokalnych obszarów w przestrzeni cech, jak również konieczność przechowywania znacznej ilości danych uczących. W związku z tym już na etapie projektowania i budowy systemów rozpoznawania pisma ręcznego z wykorzystaniem metod

z klasyfikatorami CMC pojawia się dodatkowe utrudnienie w postaci wielkości zestawu uczącego, który może sięgać nawet 1 miliona znaków.

Kontynuacje prac nad zastosowaniem sieci MLP w zadaniu rozpoznawania pisma przyniosły nowe rozwiązania wykorzystujące klasyfikatory zbudowane w oparciu o maszynę wektorów wspierających (ang. Support vector machine, SVM). Klasyfikator zaproponowany w pracy [24] umożliwił takie zdefiniowanie procesu uczenia, że wyznaczane hiperpłaszczyzny rozdzielające elementy należące do dwóch klas, odbywa się z ustaleniem maksymalnego marginesu. Pomimo wielu zalet rozwiązań opartych na sztucznych sieciach neuronowych oraz ich modyfikacjach należy też wspomnieć o wadach powodujących pewne ograniczenia w zastosowaniach aplikacyjnych przedstawionych tutaj metod. Częstym problemem pojawiającym się w większości opracowań to zagadnienie dotyczące właściwej klasyfikacji znaków posiadających pewne zniekształcenia np. uszkodzenia w postaci ubytków czy też całkowitego braku pewnych elementów znaku, przekreślenia, podkreślenia, połączenia znaków, oraz cała gama innych zniekształceń liniowych i nieliniowych, które powodują, że rozpatrywany znak w rzeczywistości nie jest związany w żaden sposób ze zbiorem przykładów uczących.

- Metody wykorzystujące Ukryte Modele Markowa - rozwiązania bazujące na modelach znaków zawierają pewne elementy wspólne i na etapie przechowywania oraz porównywania niektóre informacje na temat danych modelu są powielane. Rozwiązaniem tego problemu okazało się utworzenie modelu, którego podstawowe jednostki składały się z różnych fragmentów znaku (ang. subcharacter). Realizacje metody [2] wskazywały na lepsze wykorzystanie danych uczących ze względu na wprowadzony podział modelu. Trudnością okazuje się jednak sama technika podziału znaku na segmenty. W związku z tym zaproponowano m.in. kryteria podziału uwzględniające występowanie pętli, wypukleń (ang. humps) i wierzchołków (ang. cusps). Jeszcze inne propozycje zakładały dokonywanie segmentacji już na etapie przetwarzania wstępnego. Podział na segmenty generuje trudności w procesie tworzenia modeli, ponieważ na etapie segmentacji mogą być wydzielone pojedyncze znaki. W pracy [7] dokonano dekompozycji znaku ze względu na tzw. grammy tworząc w ten sposób model znaku składający się co najmniej z kilku fragmentów. Taka realizacja pozwoliła na znaczne ułatwienie procesu segmentacji, ponieważ elementem bazowym modelu mógł być dowolny fragment przetwarzanego tekstu. Przykładową aplikacją wykorzystującą tego typu założenia może być system AEGIS (ang. Automatic Evolutional Grammar Interpretation System) zaprezentowany w pracy [7]. Rozwiązanie to dla opracowanego zestawu 93 modeli fragmentów linii osiąga skuteczność rozpoznawania wyrazów na poziomie 94%. Metody rozpoznawania pisma z wykorzystaniem ukrytych modeli Markowa charakteryzuje wysoki współczynnik rozpoznawania. Mogą być stosowane na poziomie zarówno pojedynczych znaków i symboli jak również całych wyrazów. Dobrze sprawdzają się w zadaniu przetwarzania pisma w różnych aplikacjach z interfejsem dotykowym. Aplikacyjne zastosowanie omówionych powyżej metod z wykorzystaniem HMM wskazuje też na dużą złożoność obliczeniową, szczególnie widoczne jest to na etapie segmentacji lub ekstrakcji cech oraz porównywania modeli. Problemem może też być przygotowanie odpowiednich zestawów trenujących szczególnie uwzględniających różne lokacje językowe.
- Metody wykorzystujące momenty geometryczne - to kolejna grupa technik stosowanych w procesie rozpoznawania pisma, które mogą być stosowane nie tylko w celu określenia cech ale również orientacji przetwarzanego znaku. Z tego powodu szybko rozwinęły się techniki umożliwiające rozpoznawanie pisma niezależnie od jego położenia, rozmiaru czy też orientacji poszczególnych wyrazów lub znaków. Przykładem mogą być rozwiązania wykorzystujące w charakterze cech znaków ich momentów geometrycznych [8]. Ponadto do opisu obiektów można też użyć parametrów, które są pewną kombinacją momentów

centralnych oraz są inwariantne względem obrotu i translacji. Wspomniane techniki najczęściej wykorzystują znormalizowane momenty centralne, pozwalające wskazać pewne parametry, które są inwariantne względem np. zmiany skali. W praktycznych zastosowaniach największe znaczenia mają momenty niskich rzędów, dlatego też często bazuje się na momentach przedstawionych przez Hu w roku 1962.

- Metody wykorzystujące transformatę Fouriera (ang. Fourier Transform - FT) - jest narzędziem szeroko wykorzystywanym w przetwarzaniu sygnałów. Umożliwia przejście z funkcji opisanej w dziedzinie czasu do funkcji opisanej w dziedzinie częstotliwości oraz pozwala na operację odwrotną. Cyfrowe przetwarzanie sygnałów przeważnie wykorzystuje dyskretną wersję tego przekształcenia, tzw. dyskretną transformatę Fouriera, która znalazła szerokie zastosowanie w analizie sygnałów cyfrowych jak również w systemach rozpoznawania obrazów. Istnieje możliwość określenia dwuwymiarowej granicy znaku za pomocą jednowymiarowej funkcji odległości wyznaczonej od środka ciężkości x_c , y_c . Dane odwzorowujące np. zewnętrzny kontur znaku mogą być wyznaczone za pomocą funkcji odległości od środka ciężkości (ang. centroid distance function) [27], tworząc reprezentację kształtu inwariantną względem przesunięcia. Praktyczne realizacje systemów przetwarzania pisma na etapie przetwarzania wstępnego dostarczają informacje o znaku w postaci ciągu wartości kolejnych współrzędnych np. zewnętrznego konturu. Tego typu dane mogą być użyte do reprezentacji kształtów za pomocą funkcji współrzędnych zespolonych (ang. complex coordinates function). W ten sposób krzywa cyfrowa reprezentująca kontur znaku podana w postaci parametrycznej może być reprezentowana w postaci zmiennej zespolonej [17,18].

Metody rozpoznawania znaków z wykorzystaniem opisu konturu znaku sprawdzają się w sytuacjach, kiedy rozpoznawane znaki posiadają ciągły kontur. W rzeczywistości jednak obrazy znaków często posiadają uszkodzenia w postaci nieciągłości i szumów, co prowadzi to do wygenerowania kilku konturów w obrębie jednego znaku. Dodatkowo należy uwzględnić wewnętrzne kontury znaków, co znacznie komplikuje cały proces rozpoznawania. Pomimo tego, że niektóre zakłócenia nieciągłości można wyeliminować stosując operacje morfologiczne np. dylatację, to jednak metody konturowe wymagają stosowania szeregu operacji przetwarzania wstępnego.

Należy zauważyć, że określanie cech obrazu niezależnych od obrotu prowadzi do błędnej klasyfikacji niektórych znaków np. cyfr 6 oraz 9 [18].

- Metody wykorzystujące transformacje czasowo-częstotliwościowe, a szczególności transformatę falkową, zdobyły dużą popularność głównie dzięki możliwościom utworzenia reprezentacji sygnału zarówno w czasie i częstotliwości. Dzięki zmiennej rozdzielczości czasowo-częstotliwościowej znalazła bardzo szerokie zastosowanie w wielu dziedzinach związanych z przetwarzaniem obrazu. W praktycznych realizacjach najczęściej spotykane są dwa zastosowania transformaty: ciągła transformata falkowa (CWT) i dyskretna transformata falkowa (DWT). W pracy [26] ciągła odmiana transformaty falkowej została wykorzystana do uzyskania inwariantnych reprezentacji dla zamkniętego konturu znaku, gdzie autorzy uzyskali ponad 95% współczynnik poprawnie sklasyfikowanych ręcznie pisanych znaków (cyfr). Natomiast w pracy [23] zaproponowano zastosowanie odmiany kierunkowej dwuwymiarowej CWT, uzyskując w ten sposób niezmiennie względem przesunięcia współczynniki transformaty reprezentujące cechy przetwarzanego obrazu znaku. Przeprowadzone przez autorów eksperymenty pozwoliły na osiągnięcie ponad 90% skuteczności rozpoznawania dla ręcznie pisanych cyfr.

Warto zaznaczyć, że metody falkowe w ogólności są wrażliwe na rotację przesunięcie czy zmianę skali dlatego też w większości prezentowanych tutaj metod stosowane są operacje przetwarzania wstępnego. Drugim problemem jest odpowiedni wybór cech obrazu znaku ze

względem występowania nadmiarowej reprezentacji obrazu znaku i to zarówno w przypadku ciągłej czy też pakietowej odmiany transformaty falkowej.

- Metody wykorzystujące przekształcenie obrazu w przestrzeń parametryczną, stanowią kolejną grupę przeanalizowanych technik stosowanych w zadaniu rozpoznawania znaków. W większości bazują na przekształceniu obrazu znaku w przestrzeń parametrów (z wykorzystaniem przekształcenia Hougha lub Radona), gdzie można uzyskać informacje o lokalnych właściwościach obrazu znaku (np. linie proste). W pracy [1] autorzy wykorzystali całą reprezentację parametryczną transformaty Radona w charakterze cech obrazu znaku i w przeprowadzonych eksperymentach wykazali skuteczność metody na poziomie 96% (dla obrazów cyfr). Niestety praktyczne wykorzystanie tak zaproponowanego wektora cech może okazać się problematyczne ze względów obliczeniowych jak również z powodu wrażliwości reprezentacji parametrycznej na rotację, zmianę skali i przesunięcia obrazu znaku. Kolejne rozwiązania przedstawione w pracach [16], gdzie autorzy zaproponowali kaskadowe połączenie transformat Radona i Fouriera. Uzyskując w ten sposób niezmiennosc względem przesunięcia obrazu oraz dodatkowo ograniczenie długości wektora cech. Należy jednak zauważyć, że w większości opisywane metody są realizowane na podstawie transformaty ciągłej, która np. wymaga wielu interpolacji wynikających z wyznaczania wartości całek. Dodatkowo w zaproponowanych rozwiązaniach nie uwzględnia się problemów wynikających z przetwarzania obrazów z zakłóceniami czy też z szumami.

PODSUMOWANIE

Przeprowadzona analiza stosowanych rozwiązań jest próbą dokonania podziału technik rozpoznawania znaków wraz z przedstawieniem problemów związanych z realizacją w poszczególnych metodach. Głównym ograniczeniem opisanych metod jest konieczność stosowania różnych operacji przetwarzania wstępnego. W celu osiągnięcia zadowalających rezultatów rozpoznawania wymagana jest m.in. operacja normalizacji znaku do postaci akceptowalnej przez dany system. Do grupy najczęściej stosowanych modyfikacji obrazu w zakresie normalizacji należą operacje: zmiany zakresu jaskrawości, zmiany skali, zmiany orientacji, przesunięcia, szkieletyzacji, ścinienia.

Wskazane niedogodności poszczególnych realizacji wymuszają dalsze prace nad udoskonalaniem znanych metod, a także do implementowania systemów hybrydowych, wykorzystujących jednocześnie kilku rozwiązań. Szczególną uwagę zwrócono na możliwość rozwijania technik bazujących na Transformacie Radona. Przede wszystkim ze względu na możliwość przetwarzania obrazów w skali szarości, wrażliwość reprezentacji parametrycznej na zmianę skali rotacji czy przesunięcia, obserwacji lokalnych cech fragmentów znaku (np. linii prostych) i stosunkowo niewielkiej złożoności obliczeniowej.

BIBLIOGRAFIA

1. Aradhya V. N. M., Kumar G. H., Noushath S., *Robust Unconstrained Handwritten Digit Recognition using Radon Transform*. Signal Processing, Communications and Networking, ICSCN '07, pp. 626-629, 2007 r.
2. Bercu S., Lorette G., *On-line Handwritten Word Recognition: An Approach Based on Hidden Markov Models*. Pre-Proc. IWFHR III, pp.385, USA 1993 r.
3. Bishop C., *Neural Networks for Pattern Recognition*. Oxford University Press, Oxford 1996 r.
4. Bourbakis N., *Methodology for document processing: separating text from images*. Engineering Applications of Artificial Intelligence, Vol. 14, 2001 r.
5. Forella G., *Word perfect. Postal Technology*. 2000 r.

6. Haykin S., *Neural Networks. A Comprehensive Foundation*. Macmillan Publishing Company, UK 1994 r.
7. Hu J., *HMM Based On-Line Handwriting Recognition*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.18, pp.1039--1045, USA 1996 r.
8. Jahne B., *Digital Image Processing*. Springer-Verlag, New York 2002 r.
9. Jain A. K., Murty M. N., Flynn P. J., *Data Clustering: A Review*. ACM Computing Surveys, Vol. 31, 1999 r.
10. Kavallieratou E., Fakotakis N., Kokkinakis G., *An unconstrained handwriting recognition system*. International Journal on Document Analysis and Recognition, vol. 4, nr 4, Springer-Verlag, Heidelberg 2004 r.
11. Khedekar S, Ramanaprasad V., Setlur S., Govindaraju V., *Text - Image Separation in Devanagari Documents*. Document Analysis and Recognition, Edinburgh 2003 r.
12. Khorsheed M. S., *Off-line Arabic character recognition: A review*. Pattern analysis and applications, Vol. 5, pp. 31-45. Springer 2002 r.
13. Le Cun Y., Boser B., Denken J.S., Henderson D., Howard R.E., Hubbard W., Jackel L.D., *Backpropagation Applied to Handwritten Zip Code Recognition*. Neural Computation, vol.1, pp.541--551, 1989 r.
14. Le Cun Y., *Generalization and Network Design Strategies*. Connectionism in Perspective, Elsevier, Switzerland 1989 r.
15. Le Cun Y., Matan O., Boser B., Denken J.S., Henderson D., Howard R.E., Hubbard W., Jackel L.D., *Handwritten Zip Code Recognition with Multilayer Networks*. ICPR, vol.2, pp.35-40, USA 1989 r.
16. Mahmoud S. A., Abu-Amara M. H., *The use of radon transform in handwritten Arabic (Indian) numerals recognition*. WSEAS Transactions on Computers archive, vol. 9, pp. 252--267, USA 2010 r.
17. Maszewski M., Miciak M., *Rozpoznawanie danych teleadresowych z wykorzystaniem współczynników Fouriera i zespolonej dyskretnej transformacji falkowej opartej na projekcji*. Techniki Przetwarzania Obrazu, Serock 2006 r.
18. Miciak M., Marchewka M., *The recognition of Postal Code Using Fourier Transform Method*. XII Konferencja Sieci i Systemy Informatyczne, Łódź 2004 r.
19. Muge F., *Automatic Feature Extraction and Recognition for Digital access of Books of the Renaissance*. Lecture Notes in Computer Science, Vol. 1923, Springer-Verlag, Heidelberg 2000 r.
20. Parodi P., Fontana R., *Efficient and flexible text extraction from document pages*. International Journal on Document Analysis and Recognition, Vol. 2, Springer-Verlag, Heidelberg 1999 r.
21. Parodi P., Piccioli G., *An efficient pre-processing of mixed-content document images for OCR systems*. Proceedings of the 13th International Conference on Pattern Recognition, vol. 3, Wiedeń 1996 r.
22. Pavlidis T., *Grafika i przetwarzanie obrazów*. WNT, Warszawa 1987 r.
23. Romero D. J., Seijas L. M., Ruedin A. M., *Directional Continuous Wavelet Transform Applied to Handwritten Numerals Recognition Using Neural Networks*. JCS, 7, pp. 66--71, 2007 r.
24. Vapnik V.N., *The Nature of Statistical Learning Theory*. Information Science and Statistics, Springer-Verlag, 1995 r.
25. Wang C.H., Srihari S.N., *A framework for object recognition in visually complex environment and its application to locating address blocks on mail pieces*. International Journal of Computer Vision, vol.2, pp.125--151, Springer Netherlands 2004 r.
26. Wunsch P., Laine A. F., *Waveletdescriptors for multiresolution recognition of handprinted characters*. Pattern Recognition,

27. Zhang D., Lu G., *A Comparative Study on Shape Retrieval Using Fourier Descriptors with Different Shape Signatures*. ICIMADE '01, pp. 1--9, USA 2001.

PSYCHOLOGICAL ASPECTS THE METHODS OF POSTAL IMAGE PROCESSING

Abstract

The paper presents a method of image processing and character recognition, in particular for postal applications. In the process of the development of postal offices distribution plays a special role to read the address data in particular described handwriting. Because of the time limitation of the image processing package, the article focused on zip code reading. Particular attention has been paid to the methods based on artificial neural networks, hidden Markov models, the moments of geometric Fourier transform. Determined suitability of particular methods for the specific requirements of image processing postal systems.

Autorzy:

mgr inż. **Mirosław Miciak** – Uniwersytet Technologiczno-Przyrodniczy w Bydgoszcz, Wydział Telekomunikacji, Informatyki i Elektrotechniki, Zakład Systemów Teleinformatycznych, miciak@utp.edu.pl

dr inż. **Roman Wiatr** – Uniwersytet Technologiczno-Przyrodniczy w Bydgoszcz, Wydział Telekomunikacji, Informatyki i Elektrotechniki, Zakład Systemów Teleinformatycznych, rowiat@utp.edu.pl