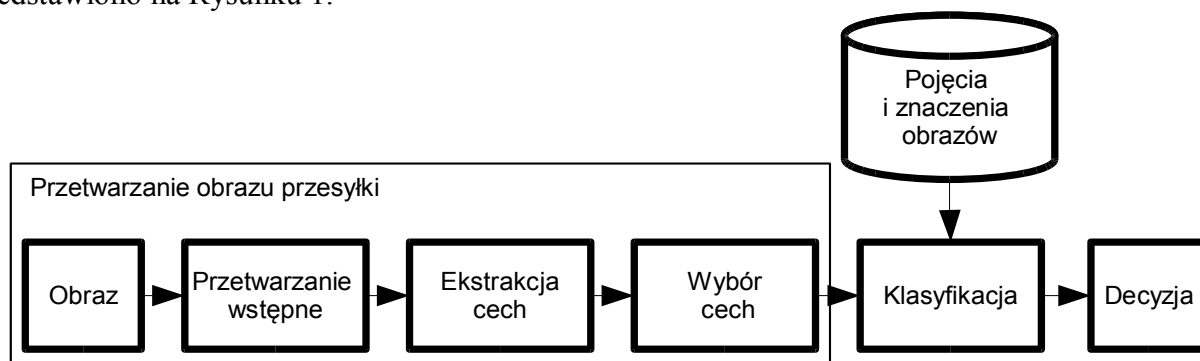


MICIAK Mirosław ¹
 WIATR Roman ¹

Problemy rozpoznawania i klasyfikacji obrazów w procesach logistycznych

WSTĘP

Rozpoznawanie i przetwarzanie obrazów jest powszechnie wykorzystywane do identyfikacji dokumentów. Kierunki rozwoju systemów rozpoznawania uwzględniają możliwości rozpoznawania pisma, jak również ograniczenie ilości przechowywanych danych. Obecnie największymi odbiorcami systemów optycznego rozpoznawania pisma są instytucje pocztowe, w których w sposób automatyczny jest lokalizowany i odczytywany kod z przesyłek. Coraz częściej stosuje się już system automatycznego rozpoznawania znaków do sortowania przesyłek na podstawie adresu bez wpisanego kodu pocztowego. 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 1.



Rys. 1. 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 selekcyonowane 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. Analiza szeregu 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.

1. ANALIZA METOD ROZPOZNAWANIA ZNAKÓW

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

¹ Uniwersytet Technologiczno - Przyrodniczy w Bydgoszczy, Wydział Telekomunikacji, Informatyki i Elektrotechniki, Zakład Systemów Teleinformatycznych; 85-225 Bydgoszcz; ul. Kordeckiego 20. Tel.: + 48 52 340-81-14, 340-81-15, e-mail: [miroslaw.miciak, roman.wiatr]@utp.edu.pl

rozpoznawania pisma pozwoliła na wskazanie pewnych najliczniej reprezentowanych technik. Poniżej zaprezentowano główne narzędzia wykorzystywane w zadaniu rozpoznawania znaków.

1.1. Metody z wykorzystaniem sztucznych sieci neuronowych

Są to najczęstsze i najbardziej popularne techniki wykorzystywane w zadaniu rozpoznawania znaków. Pierwsze efektywne rozwiązania systemów automatycznego rozpoznawania zaprezentowano w 1989 roku w pracy [8]. Przedstawiona aplikacja wykorzystywała wielowarstwową sieć neuronową wspieraną przez algorytm uczenia z nauczycielem. Autorzy publikacji [10] zaproponowali system rozpoznawania dla znaków uzyskanych w kolejnych etapach przetwarzania przedstawionych w [16], 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 [9] 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) [5] 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 omówionych 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 maszyny wektorów wspierających SVM (ang. Support vector machine,). Klasyfikator zaproponowany w pracy [16] umożliwiał takie zdefiniowanie procesu uczenia, że wyznaczone 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.

1.2. Metody wykorzystujące Ukryte Modele Markowa

Są to rozwiązania bazujące na modelach znaków, które 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 [3] 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 [6]. 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 lokalizacje językowe.

1.3. Metody wykorzystujące momenty geometryczne

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 momenty geometryczne [7]. 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.

1.4. Metody wykorzystujące transformatę Fouriera

Metody wykorzystujące transformatę Fouriera FT (ang. Fourier Transform) są narzędziami 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) [20], 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 [12,13]. 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. Należy zauważyć, że określanie cech obrazu niezależnych od obrotu prowadzi do błędnej klasyfikacji niektórych znaków np. cyfr 6 oraz 9 [13].

1.5. Metody wykorzystujące transformacje czasowo-częstotliwościowe

Metody wykorzystujące transformacje czasowo-częstotliwościowe, a szczególnie transformatę falkową, zdobyły dużą popularność głównie dzięki możliwościom utworzenia reprezentacji sygnału zarówno w czasie i częstotliwości. Dzięki zmiennej rozdzielczości metoda 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

[19] 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 [15] 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 względu na występowanie nadmiarowej reprezentacji obrazu znaku i to zarówno w przypadku ciągłej czy też pakietowej odmiany transformaty falkowej.

1.6. Metody wykorzystujące przekształcenie obrazu w przestrzeń parametryczną

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 pracy [11], gdzie autorzy zaproponowali kaskadowe połączenie transformat Radona i Fouriera. Uzyskano 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.

2. METODY KLASYFIKACJI CECH ZNAKÓW

Klasyfikacja cech analizowanego znaku dla metod zaproponowanych w artykule, w ogólności jest realizowana zgodnie ze schematem rozpoznawania obiektów polegającym na wyznaczeniu reguły klasyfikacyjnej w oparciu o pewien zbiór nazywany też zbiorem uczącym. Tworzenie takich reguł jest głównym zadaniem analizy dyskryminacyjnej (ang. discriminant analysis) i polega na rozstrzygnięciu, które zmienne w najlepszy sposób dzielą dany zbiór przypadków na występujące w naturalny sposób grupy. Techniki analizy pozwalają m.in. rozstrzygnąć, czy grupy różnią się ze względu na średnią pewnej zmiennej, czy też wykorzystać zmienną do przewidywania przynależności do danej grupy. Generalnie klasyfikacja polega na przeprowadzeniu dwóch głównych etapów:

- etapu uczenia - czyli budowy modelu, w którym znajdowane są reguły klasyfikacyjne w oparciu o zbiór uczący (próbę statystyczną),
- etapu klasyfikacji - czyli wykorzystania modelu, w którym dokonywana jest klasyfikacja zasadniczego zbioru obiektów, których przynależność jest nieznaną w oparciu o znalezione charakterystyki klas.

Dla potrzeb przeprowadzenia eksperymentów badawczych wybrano trzy typy klasyfikatorów.

2.1. Klasyfikator minimalno odległościowy

Technika ta umożliwia klasyfikację, której kryterium stanowi miara odległości między wektorem badanego znaku a wektorem zbioru uczącego. Miara odległości euklidesowej dana jest następująca zależnością:

$$D(C_i, C_r) = \sum_{j=1}^N [R(j) - A(j)]^2 \quad (1)$$

gdzie: C_i – znak ze zbioru uczącego, C_r – znak analizowany, R – wektor cech znaku aktualnie rozpoznawanego, A – wektor cech znaku ze zbioru uczącego, N – liczba cech.

W przeprowadzonej procedurze rozpoznawania, minimalna odległość między analizowanym wektorem, a przedstawicielami wektorów klasy zbioru uczącego stanowi kryterium dla wyboru klasy znaku.

2.2. Klasyfikator k -najbliższych sąsiadów

Klasyfikator k najbliższych sąsiadów (lub k -nn z ang. k nearest neighbours), został zaproponowany w pracy [4]. Jest algorytmem regresji nieparametrycznej, dla którego dany jest zbiór uczący zawierający obserwacje, z których każda ma przypisany wektor zmiennych objaśniających oraz wartość zmiennej objaśnianej Y . Dana jest również obserwacja C z przypisanym wektorem zmiennych objaśniających dla której chcemy prognozować wartość zmiennej objaśnianej Y . Algorytm polega na:

- krok 1 porównaniu wartości zmiennych objaśniających dla obserwacji C z wartościami tych zmiennych dla każdej obserwacji w zbiorze uczącym.
- krok 2 wyborze k (ustalona z góry liczba) najbliższych do C obserwacji ze zbioru uczącego.
- krok 3 Uśrednieniu wartości zmiennej objaśnianej dla wybranych obserwacji, w wyniku czego uzyskujemy prognozę.

Definicja najbliższych obserwacji w kroku 2 sprowadza się do minimalizacji ustalonej metryki, mierzącej odległość pomiędzy wektorami zmiennych objaśniających dwóch obserwacji.

2.3. Klasyfikator $LTF-C$

$LTF-C$ (ang. Local Transfer Function Classifier) [17,18] to sieć neuronowa do zadań klasyfikacyjnych o architekturze zbliżonej do sieci radialnych (RBF). Składa się z dwóch warstw neuronów. Pierwsza warstwa (tzw. ukryta) zawiera neurony o gaussowskiej funkcji transferu, które wykrywają w danych treningowych skupiska wzorców z tej samej klasy. Każdy neuron tej warstwy ma przypisaną klasę, której skupisko stara się wykryć. Drugą warstwę tworzą neurony liniowe, które segregują odpowiedzi neuronów ukrytych według przypisanych klas i sumują je, formułując ostateczną odpowiedź sieci.

3. PARAMETRYCZNA METODA ROZPOZNAWANIA I KLASYFIKACJI ZNAKÓW

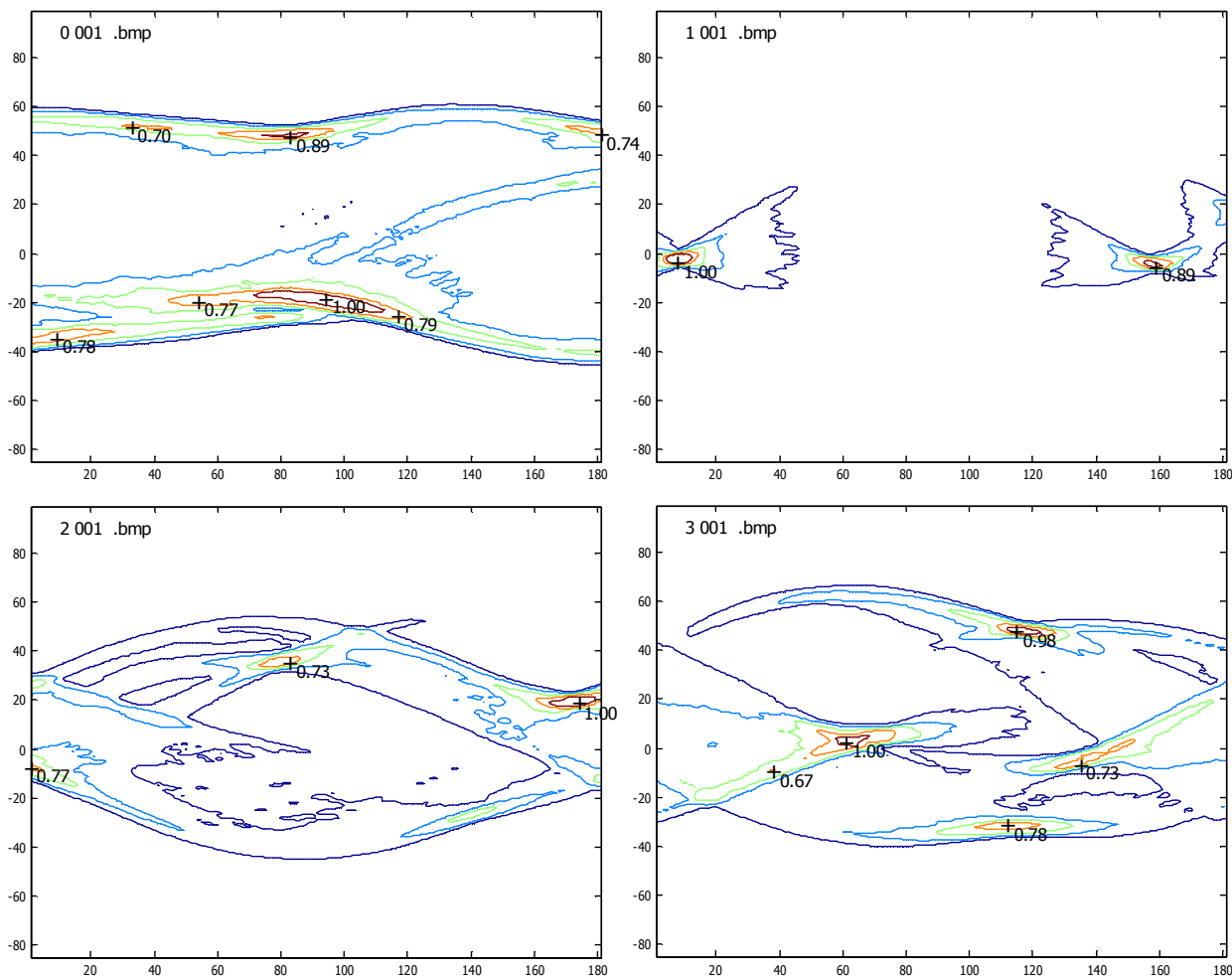
Na podstawie analizy metod wydzielenia, rozpoznawania i klasyfikacji znaków zaproponowano metodę wydzielenia cech, dla której podstawową informacją uzyskiwaną z reprezentacji przestrzeni parametrycznej transformaty Radona [14] są lokalne maksima, które odpowiadają punktom przecięcia linii (lub ich przedłużenia) opisujących kształt badanego znaku. Za pomocą lokalnych maksimów przestrzeni parametrycznej można utworzyć reprezentację obrazu znaku. Dodatkowe informacje to: współrzędne oraz wartość ekstremów w przestrzeni parametrycznej. W rezultacie daje to możliwość sformułowania wektora cech badanego obrazu znaku. Rozmiary reprezentacji parametrycznej transformaty Radona określone są za pomocą parametrów ρ, θ . Ustalając zmiany kąta obrotu co jeden stopień otrzymamy 180 kolumnową reprezentację parametryczną. Natomiast uwzględniając rozmiar obrazu $m \times m$, jednostkowe zmiany długości promienia ρ oraz zależność:

$$\rho = \left\lceil \frac{m}{2} \sqrt{2} \right\rceil \quad (2)$$

to uzyskany zakres zmian długości promienia mieści się w przedziale $\{ -\rho, \dots, 0, \dots, \rho \}$, np. dla obrazu o wymiarach 128×128 uzyskamy 183 wiersze reprezentacji transformaty Radona. Rozpatrując znormalizowaną tablicę akumulatora $A_{cu}(\rho, \theta)$ możliwą do otrzymania na podstawie standardowej realizacji transformaty Radona określone są następujące parametry:

- liczba maksimów tablicy akumulatora - l_{max} ,

- zbiór wartości maksimum tablicy akumulatora - $\{w_{m1}, \dots, w_{m1max}\}$,
- zbiór współrzędnych maksimum akumulatora - $\{(\theta_{l1}, \rho_{l1})w_{m1}, \dots, (\theta_{lmax}, \rho_{lmax})w_{m1max}\}$.



Rys. 2. Reprezentacje parametryczne wybranych obrazów z zaznaczonymi punktami charakterystycznymi

Na podstawie przeprowadzonej analizy przestrzeni parametrycznych bazy obrazów znaków (Tabela 1) można stwierdzić, iż dla różnych klas znaków (jak również ich podzbiorów) liczba maksimum jest na tyle zróżnicowana, że sama nie może stanowić kryterium klasyfikacyjnego

Tab. 1. Przedziały liczby maksimum lokalnych l_{max} akumulatora transformaty Radona A_{cu} dla obrazów z własnej bazy znaków

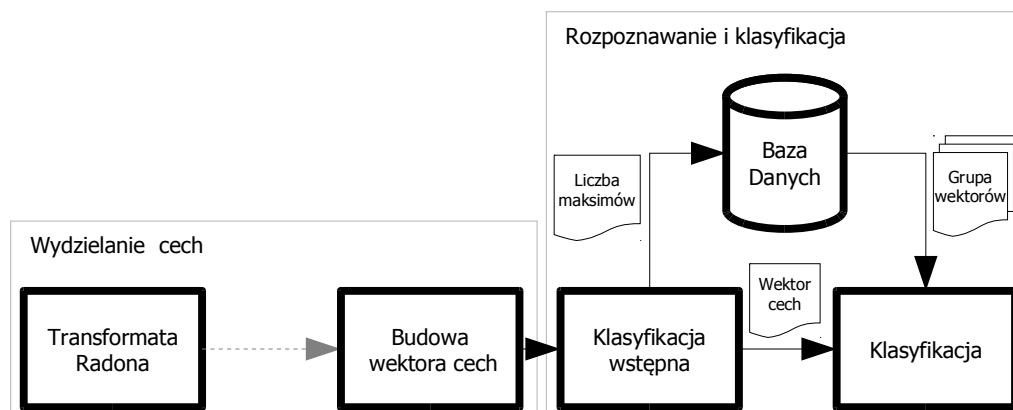
Znaki	0	1	2	3	4	5	6	7	8	9
l_{max}	4-10	1-4	2-5	3-10	3-8	3-9	2-10	2-4	3-10	3-9

W związku z powyższym niezbędne jest wyznaczenie dla każdego maksimum odpowiadającej mu wartości w_m reprezentacji parametrycznej oraz współrzędnych położenia $w_{m\theta}, w_{m\rho}$. Wyniki analizy dla własnej bazy obrazów znaków zostały przedstawione w Tabeli 2. Przykładowe rozmieszczenie wartości maksymalnych dla wybranych znaków ze zbioru ZT1 zostało przedstawione na Rysunku 2. Na podstawie powyższego wektor cech punktów charakterystycznych FV_{CPC} można zapisać w postaci:

$$FV_{CPC} = \{l_{max}, ((w_{m_1}, w_{m\theta_1}, w_{m\rho_1}), (w_{m_{lmax}}, w_{m\theta_{lmax}}, w_{m\rho_{lmax}}))\} \quad (3)$$

Wektor cech dla każdego znaku składa się z dwóch części. Pierwsza zawiera liczbę maksimum l_{max} , druga przechowuje jego wartości w_m oraz współrzędne $w_{m\theta}, w_{m\rho}$. Wartość pierwszego maksimum ze

względem na zastosowanie znormalizowanej tablicy akumulatora będzie równa jedności. Rozdzielenie wektora na części pozwala na przeprowadzenie klasyfikacji wstępnej, bazującej np. na liczbie punktów maksymalnych.



Rys. 3. Schemat klasyfikacji wstępnej

Schemat klasyfikacji wstępnej przedstawiony na Rysunku 3 opiera się na analizie liczby maksimów lokalnych uzyskanych w procesie tworzenia wektora cech. Na podstawie tego parametru określana jest grupa wektorów z bazy danych, które należy dostarczyć do klasyfikatora w celu porównania ich z wektorem znaku badanego. Długość wektora cech (największa liczba punktów maksimów lokalnych) została ustalona na podstawie przeprowadzonych badań z wykorzystaniem własnej bazy znaków.

4. WYNIKI PRZEPROWADZONYCH EKSPERYMENTÓW

W części eksperymentalnej określono skuteczność zaproponowanej metody ze względu na liczbę maksimów lokalnych l_{max} . Zwiększano l_{max} w zakresie od 1 do 10. Okazało się, że liczba poprawnie sklasyfikowanych znaków była największa dla wektora cech o 5 maksimach, natomiast dalsze zwiększanie liczby l_{max} skutkowało spadkiem liczby poprawnie sklasyfikowanych znaków. W tabeli 2 przedstawiono parametry wektora cech FV_{CPC} dla zestawu ZT1.

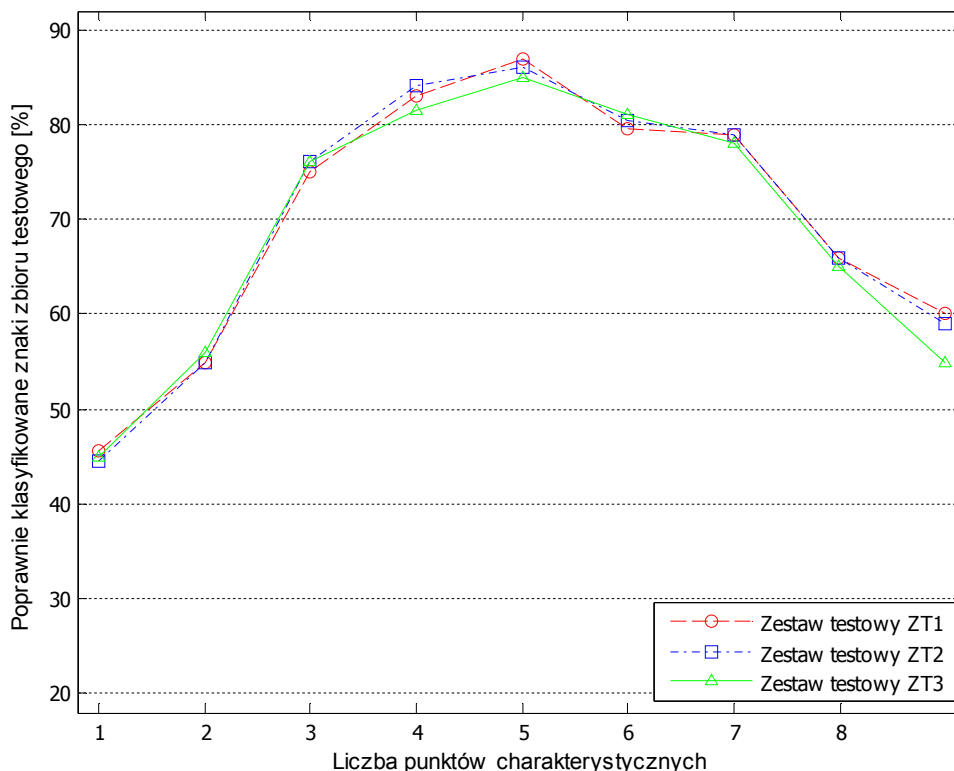
Tab. 2. Parametry wektora cech FV_{CPC} dla wybranych znaków z zestawu ZT1

Cyfra	l_{max}	ω_{m1}	ω_{m01}	ω_{mp1}	ω_{m2}	ω_{m2}	ω_{mp2}	ω_{m3}	ω_{m3}	ω_{mp3}	ω_{m4}	ω_{m4}	ω_{mp4}	ω_{m4}	ω_{m4}	ω_{mp4}
0	5	1	121	68	0,9	110	134	0,8	35	52	0,5	2	137	0,5	16	72
1	2	1	101	83	0,9	71	105									
2	4	1	120	81	0,7	29	64	0,6	128	112	0,6	87	126			
3	4	1	56	97	1,0	110	52	0,8	107	131	0,7	130	106			
4	4	1	30	109	1,0	36	82	0,9	150	104	0,6	53	91			
5	6	1	85	121	0,8	58	84	0,8	91	58	0,7	124	113	0,6	152	110
6	2	1	132	108	0,5	96	94									
7	3	1	130	87	0,8	77	81	0,7	91	117						
8	3	1	65	82	0,9	147	100	0,5	20	97						
9	3	1	59	88	0,7	138	61	0,7	109	142						

Na podstawie powyższego w bloku klasyfikacji wstępnej utworzono pięć grup znaków, gdzie numer grupy odpowiada liczbie maksimów lokalnych. W przypadku znaków o większej liczbie maksimów wybierane są te z największymi wartościami akumulatora A_{cu} .

W części eksperymentalnej zrealizowano badania, na podstawie których określono skuteczność zaproponowanego klasyfikatora w funkcji liczby wydzielonych punktów charakterystycznych. Badania przeprowadzono dla 3 zestawów testowych (ZT1-ZT3) składających się ze znaków kodu pocztowego przesyłek listowych zapisanych ręcznie. Zestaw uczący (baza znaków) zawiera 1440 znaków natomiast zestawy testujące po 200 znaków każdy. Dla wektora cech określonego za pomocą wyrażenia 3 zastosowano miarę minimalno odległościową (1). Na podstawie przeprowadzonych

badania można zauważyć, że najlepsze rezultaty osiągnięto dla 5 punktów charakterystycznych. Okazało się również, że zwiększanie liczby elementów wektora cech skutkowało zmniejszeniem skuteczności rozpoznawania znaków dla wszystkich zestawów testowych. Na Rysunku 4 przedstawiono wykres przedstawiający sprawność klasyfikacji znaków w funkcji liczby punktów charakterystycznych dla zestawów znaków ZT1-ZT3.



Rys. 4. Poprawnie sklasyfikowane znaki w funkcji liczby punktów charakterystycznych dla ZT1-ZT3

Dodatkowo w ramach działań wykonywanych w celu określenia skuteczności zaproponowanych metod przeprowadzono badania rozpoznawania i klasyfikacji kodów pocztowych. Ocenę skuteczności przedstawionych w tym podrozdziale metod przeprowadzono z użyciem klasyfikatorów: minimalno odległościowego, k -najbliższych sąsiadów i LTF-C. Wyniki badań dla poszczególnych klasyfikatorów znaków zaprezentowano w Tabeli 3.

Tab. 3. Sprawność klasyfikatorów

Nazwa klasyfikatora	Skuteczność klasyfikacji dla wektora cech $l_{max}=5$ [%]
Minimalno odległościowy	85,31
k - najbliższych sąsiadów	85,29
LTF-C	88,63

W powyższej tabeli przedstawiono wyniki badań różnych klasyfikatorów dla wektora cech FV_{CPC} z ograniczeniem do $l_{max}=5$. Najlepszą skutecznością charakteryzował się klasyfikator LTF-C, lecz ze względu na uciążliwy proces uczenia zdecydowano się na dwustopniowy klasyfikator minimalno odległościowy.

Podstawową zaletą przedstawionej metody jest zastosowanie dwustopniowej klasyfikacji, możliwość pominięcia niektórych operacji przetwarzania wstępnego. Natomiast zasadniczą wadą może okazać się złożoność obliczeniowa procesu wyznaczania transformaty Radona, szczególnie tej realizowanej na podstawie definicji w postaci ciągłej. Rozwiązaniem tego problemu może być opracowanie realizacji skończonej postaci transformaty dla obrazów dyskretnych.

WNIOSKI

Przedstawione w pracy metody rozpoznawania znaków z wykorzystaniem transformaty Radona pozwalają na eliminację elementów przetwarzania wstępnego pojawiające się we większości spotykanych rozwiązań tego typu i wydają się być skutecznym rozwiązaniem podstawowych problemów aplikacji pocztowych. Dodatkowo uwzględniając możliwość przeprowadzenia procesu normalizacji w dziedzinie parametrycznej uzyskuje się niezmiennosc względem zmiany skali znaku. Zaletą opisanych metod jest zastosowanie dwustopniowego algorytmu klasyfikacji, który umożliwia zmniejszenie liczby porównań wektora cech badanego znaku ze zbiorem wektorów wzorcowej bazy danych. Klasyfikacja wstępna dokonywana jest na podstawie liczby maksimów lokalnych, uzyskanych w procesie tworzenia wektora cech. Na podstawie liczby maksimów określone są grupy znaków z tą samą liczbą maksimów, pozwala to w efekcie na zmniejszenie liczby porównań wektorów bazy znaków wzorcowych, które następnie należy porównać z wektorem cech badanego znaku. Zastosowanie dwu-etapowej klasyfikacji pozwoli w przyszłości na sformułowanie rozwiązania problemu tzw. odrzucenia np. elementów niebędących znakami. Okazuje się, że dla obrazów przedstawiających skomplikowane kształty geometryczne analiza akumulatora wykaże znacznie większą liczbę maksimów lokalnych, niż dla znaków (w tym cyfr). Przekroczenie ustalonej liczby maksimów lokalnych skutkuje odrzuceniem obiektu już na etapie klasyfikacji wstępnej.

Wadą zaproponowanej metody może okazać się mała skuteczność klasyfikacji wstępnej w zakresie odrzucania obiektów nie należących do zdefiniowanych klas. Uzyskana sprawność, która dla przeprowadzonych badań oscyluje w granicach 86% może okazać się niewystarczająca dla wymagań aplikacji rozpoznających pismo. Niemniej jednak dzięki prostej implementacji może okazać się skutecznym narzędziem do rozpoznawania kodów pocztowych. Dużą zaletą przedstawionej metody jest odporność na zakłócenia w obrazie, sprawdza się przy przede wszystkim tam, gdzie nie jest znana cała wiedza o rozpoznawanym znaku, gdzie występują niekompletne dane wejściowe albo dane są silnie zakłócone. Kontynuacja prac będzie realizowana w celu rozwinięcia opracowanych technik w zakresie podniesienia skuteczności systemów pocztowych z wykorzystaniem modyfikacji przekształcenia Radona i eliminacji niektórych procesów przetwarzania wstępnego.

Streszczenie

W pracy przedstawiono analizę metod rozpoznawania znaków dla celów pocztowych. Dokonano prezentacji wybranych klasyfikatorów i możliwości ich zastosowania w systemach rozpoznawania i klasyfikacji przesyłek pocztowych. Dodatkowo zaprezentowano nową metodę rozpoznawania znaków z wykorzystaniem transformaty Radona pozwalającą na eliminację elementów przetwarzania wstępnego pojawiające się we większości spotykanych rozwiązań tego typu. Zaproponowany klasyfikator bazuje na dwuetapowym klasyfikatorze na podstawie liczby maksimów lokalnych, uzyskanych w procesie tworzenia wektora cech. Na podstawie liczby maksimów określone są grupy znaków z tą samą liczbą maksimów, pozwala to w efekcie na zmniejszenie liczby porównań wektorów bazy znaków wzorcowych, które następnie należy porównać z wektorem cech badanego znaku. W końcowej części artykułu przedstawiono wyniki przeprowadzonych eksperymentów dla opracowanej metody.

The problems of recognition and classification of images in logistics tasks

Abstract

In this article we introduced the analysis of character recognition methods used in postal applications. We presented selected classifiers and their possibility to use in the systems of recognition and classification of postmail items. Additionally we presented the new method of character recognition on basis of the Radon transformation. The properties of Radon transformation allow for realisation of the preprocessing stage without few elements of image processing which appear in the standard applications. The proposed classifier basis on two stage processing. The number of feature vector elements describes the groups with the same numbers of characteristic points. This approach allows to decrease amount of comparing vectors from database of known characters with a vector of unknown character. The last part of the article contain experiment results of the proposed method and ratio of various classifiers. We discussed strengths and weaknesses of the method and defined the future directions of development of our method.

BIBLIOGRAFIA

1. Aradhya V. N. M., Kumar G. H., Noushath S., Robust Unconstrained Handwritten Digit Recognition using Radon Transform. ICSCN '07, s. 626-629, 2007.
2. Bercu S., Lorette G., On-line Handwritten Word Recognition: An Approach Based on Hidden Markov Models. Pre-Proc. IWFHR III, s. 385, USA 1993.
3. Bourbakis N., Methodology for document processing: separating text from images. Engineering Applications of Artificial Intelligence, Vol. 14, 2001.
4. Fix E., Hodges J.L., Discriminatory analysis - nonparametric discrimination, Report no. 4, USAF School of Aviation Medicine, s. 261-279, USA, 1951.
5. Haykin S., Neural Networks. A Comprehensive Foundation. MPC, UK 1994.
6. Hu J., HMM Based On-Line Handwriting Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.18, s. 1039-1045, USA 1996.
7. Khedekar S., Ramanaprasad V., Setlur S., Govindaraju V., Text - Image Separation in Devanagari Documents. Document Analysis and Recognition, Edinburgh 2003.
8. Le Cun Y., Boser B., Denken J.S., Henderson D., Howard R.E., Backpropagation Applied to Handwritten Zip Code Recognition. NC, vol.1, s. 541-551, 1989.
9. Le Cun Y., Generalization and Network Design Strategies. Connectionism in Perspective, Elsevier, Switzerland 1989.
10. Le Cun Y., Matan O., Boser B., Denken J.S., Henderson D., Howard R.E., Hubbard W., Handwritten Zip Code Recognition with Multilayer Networks. ICPR, s. 35-40, 1989.
11. Mahmoud S. A., Abu-Amara M. H., The use of radon transform in handwritten Arabic (Indian) numerals recognition. WSEAS TCA, vol. 9, s. 252-267, USA 2010.
12. Maszewski M., Miciak M., Rozpoznawanie danych teleadresowych z wykorzystaniem współczynników Fouriera i zespolonej dyskretnej transformacji falkowej opartej na projekcji. Techniki Przetwarzania Obrazu, Serock 2006.
13. Miciak M., Marchewka M., The recognition of Postal Code Using Fourier Transform Method. XII Konferencja Sieci i Systemy Informatyczne, Łódź 2004.
14. Miciak M., Character Recognition Using Radon Transformation and Principal Component Analysis in Postal Applications, IMCIST 2008, Wisła 2008.
15. Romero D. J., Seijas L. M., Ruedin A. M., Directional Continuous Wavelet Transform Applied to Handwritten Numerals Recognition Using Neural Networks. JCS, 7, 2007.
16. Vapnik V.N., The Nature of Statistical Learning Theory. Information Science and Statistics, Springer-Verlag, 1995.
17. Wojnarski M., LTF-C Architecture, Training Algorithm and Applications of New Neural Classifier, Fundamenta Informaticae, vol.54, s. 89-105, IOS Press, 2003.
18. Wojnarski M., LTF-C - Neural Network for Solving Classification Problems, Parallel Processing and Applied Mathematics, Lecture Notes in Computer Science, vol.2328, s.573-578, 2006.
19. Wunsch P., Laine A. F., Wavelet descriptors for multiresolution recognition of handprinted characters. Pattern Recognition.
20. Zhang D., Lu G., A Comparative Study on Shape Retrieval Using Fourier Descriptors with Different Shape Signatures. ICIMADE '01, s. 1-9, USA 2001.