

Wojciech KACALAK, Maciej MAJEWSKIPOLITECHNIKA KOSZALIŃSKA, WYDZIAŁ MECHANICZNY
ul. Raclawicka 15/17, 75-620 Koszalin**Wybrane problemy efektywnego rozpoznawania pisma odręcznego****Prof. dr hab. inż. Wojciech KACALAK**

Kierownik Katedry Mechaniki Precyzyjnej. Specjalność naukowa: mechatronika, diagnostyka, optymalizacja i automatyzacja procesów mikroobróbki oraz procesów bardzo dokładnej obróbki ścierniej, w tym zwłaszcza materiałów trudno obrabialnych, a także budowa i eksploatacja precyzyjnych urządzeń technologicznych. Dorobek naukowy składa się z ponad 260 publikacji naukowych, w tym wielu zagranicznych, oraz licznych projektów badawczych, monografii i patentów.



e-mail: wojciech.kacalak@tu.koszalin.pl

Dr hab. inż. Maciej MAJEWSKI

Profesor nadzwyczajny Katedry Mechaniki Precyzyjnej. Specjalność naukowa: budowa i sterowanie inteligentnymi systemami interakcji. Stopień doktora n.t. oraz stopień doktora habilitowanego n.t. uzyskał w Politechnice Koszalińskiej. Współautor 53 publikacji naukowych, w tym 39 zagranicznych. Współpracował na zasadzie pobytu w uniwersytetach wschodu i zachodu na świecie przez okres 3 lat.



e-mail: maciej.majewski@tu.koszalin.pl

Streszczenie

W artykule przedstawiono nowatorską metodę efektywnego rozpoznawania pisma odręcznego z zastosowaniem opracowanych sposobów analiz geometrycznych znaków i wybranych metod sztucznej inteligencji. Proponowana metoda analiz geometrycznych znaków oparta na opracowanym sposobie odpowiednich pomiarów odległości wybranych ich punktów pozwala na rozpoznawanie pisma odręcznego niezależnie od stylu i charakteru pisma operatora. W rezultacie zastosowania metody otrzymuje się zakodowaną reprezentację znaku dla efektywnego rozpoznawania przez sztuczne sieci neuronowe. Artykuł również przedstawia system rozpoznawania odręcznego pisma operatora zbudowany z podsystemów wstępnego przetwarzania, analiz geometrycznych, logiki rozmytej, sieci neuronowych oraz ich wyspecjalizowanych modułów. Proponowany inteligentny system może stanowić nowoczesny i efektywny system interakcji urządzeń technicznych i ich operatorów w zadaniach sterowania.

Słowa kluczowe: rozpoznawanie pisma odręcznego, interakcja urządzeń technicznych i ich operatorów, interfejs pisma odręcznego, sztuczna inteligencja, hybrydowe sieci neuronowe.

Selected problems of effective handwriting recognition**Abstract**

In this paper, an innovative method for effective handwriting recognition is presented. It uses the developed methods of geometrical analyses of isolated handwritten characters and selected artificial intelligence methods. The proposed geometrical feature analysis method, based on the developed manner of appropriate measurements of distances of selected character points, allows handwriting recognition independent of different writing and character styles, and writing conditions. As a result of using the method, encoded representations of characters are obtained for effective recognition by artificial neural networks. The paper also presents an operator's handwriting recognition system consisting of the subsystems of preprocessing, geometrical analyses, fuzzy logic, neural networks, and their specialized modules. Handwriting recognition has always been a challenging problem for artificial intelligence researchers, and remains an open issue. It is because of the complexity of the handwriting recognition task. The intelligent handwriting recognition system of the technical device operator's natural writing can be a modern and effective interaction system [3, 5]. In the paper, a review of selected issues is carried out with regards to the handwriting recognition issues, new geometrical analysis method (fig. 1, 2, 3, 4) and concept of a handwriting recognition system (fig. 5, 6). The proposed system is novel in that it integrates efficient geometrical processing with artificial intelligence methods to use neural networks and fuzzy logic for effective handwriting recognition.

Keywords: handwriting recognition, interaction of technical devices and their operators, handwriting interface, artificial intelligence, hybrid neural networks.

1. Wprowadzenie

Pismo odręczne stanowi interesujący sposób komunikacji człowieka z urządzeniami technicznymi mający wiele zalet.

Technologia rozpoznawania pisma odręcznego może wykorzystywać różnorodne mobilne urządzenia komputerowe w istotnych zadaniach inżynierskich. W procesie rozpoznawania pisma naturalnego człowieka obraz zawierający pismo odręczne musi być odpowiednio dostarczony i przetworzony przez odpowiedni system [1]. Zawarty tekst przechodzi przez proces segmentacji i ekstrakcji cech. Niewielkie części tekstu są wynikiem tych procesów, które podlegają rozpoznawaniu przez system [2]. Następnie informacje kontekstowe powinny zostać zastosowane w przetwarzaniu rozpoznanych symboli, w celu weryfikacji wyników. Sztuczne sieci neuronowe zastosowane w rozpoznawaniu pisma naturalnego pozwalają na wysoką zdolność uogólniania oraz nie wymagają szerokiej wiedzy i formalizacji do rozwiązania problemu rozpoznawania pisma odręcznego. Inteligentny system rozpoznawania odręcznego pisma operatora urządzenia technicznego może stanowić nowoczesny i efektywny system interakcji.

2. Metoda efektywnego rozpoznawania pisma

Proponowanym nowatorskim rozwiązaniem w rozpoznawaniu pisma odręcznego, umożliwiającym efektywne jego rozpoznawanie, jest zastosowanie opracowanych metod analiz geometrycznych znaków polegających na odpowiednich pomiarach odległości wybranych ich punktów oraz zastosowaniu metod sztucznej inteligencji do rozpoznawania zakodowanych reprezentacji znaków.

Prezentowane rozwiązanie polega na zastosowaniu podsystemu analiz geometrycznych dokonującego ekstrakcji cech poszczególnych znaków, który składa się z modułu wyznaczania środka ciężkości znaku, modułu detekcji i ekstrakcji cech geometrycznych oraz modułu wyznaczania kodowanej reprezentacji znaku.

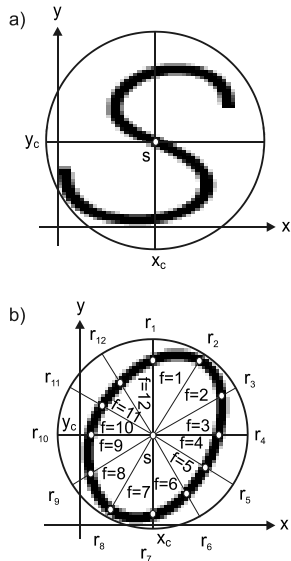
W opracowanej metodzie analiz geometrycznych znaków, proces rozpoznawania znaku rozpoczyna się od znalezienia środka ciężkości znaku z zastosowaniem sieci neuronowych, w pozyskanym jego obrazie. Wyznaczony środek służy do wykreślenia promieni i odpowiednich pomiarów odległości pomiędzy punktami znaku w celu opisu kształtu znaku, oraz utworzenia jego zakodowanej reprezentacji. Za pomocą takiej analizy geometrycznej znaku otrzymywane są charakterystyczne długości odcinków (r) utworzone przez punkt środka ciężkości i punkt należący do znaku na danym promieniu. Analiza geometryczna znaku wyznacza również odcinki w celu dokonania pomiarów charakterystycznych odległości pomiędzy odpowiednimi punktami znaku (l).

Opracowany sposób analiz geometrycznych pozwala na otrzymanie zakodowanych reprezentacji wyizolowanych znaków, które po procesie normalizacji z zastosowaniem reguł logiki rozmytej, są sygnałami wejściowymi sieci neuronowych do rozpoznawania znaków. Rozpoznane znaki są grupowane i dokonywane jest rozpoznawanie ich ciągów. Następnie ciągi znaków podlegają kodowaniu jako obrazy binarne, które stanowią sygnały wejściowe modułu rozpoznawania wyrazów z zastosowaniem wielowarstwowych sieci neuronowych zawierających zestawy uczące wzorców wyrazów. Rozpoznane wyrazy słownikowe, reprezentow-

wane przez neurony wyjściowe sieci, są przetwarzane przez moduł rozpoznawania fraz z zastosowaniem sieci neuronowych.

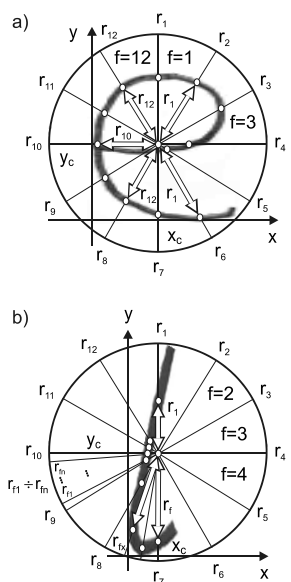
Sieci neuronowe do rozpoznawania fraz zawierają zbiory uczące fraz zbudowane przy pomocy wiedzy kontekstowej z językoznawstwa.

Pomiary długości wyznaczonych odcinków służą do wyodrębnienia cech geometrycznych w postaci zakodowanych reprezentacji znaków dla procesów efektywnego rozpoznawania izolowanych znaków z zastosowaniem sieci neuronowych. Proponowana metoda analiz geometrycznych znaków oparta na opracowanym sposobie odpowiednich pomiarów odległości wybranych ich punktów pozwala na uzyskanie metody rozpoznawania pisma niezależnej od charakteru pisma operatora.



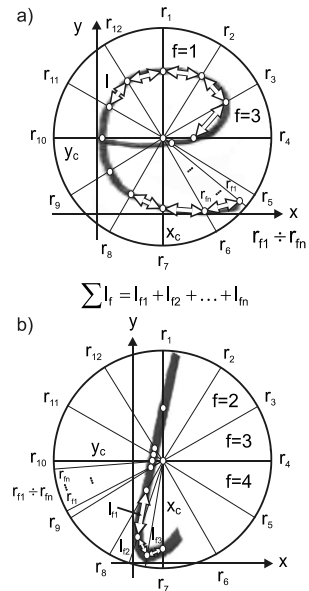
Rys. 1. Analizy geometryczne znaków: a) wyznaczenie środka ciężkości dla przykładowej litery s, b) wyznaczenie punktów przecięcia promieni ze znakiem dla przykładowej litery o

Fig. 1. Geometrical analyses of characters: a) determination of the centre of the mass for exemplary letter s, b) determination of intersection points of the letter and the radiuses for exemplary letter o



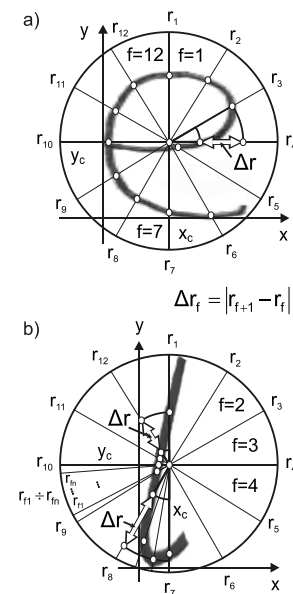
Rys. 2. Analizy geometryczne znaków: pomiary długości odcinków powstałych z każdego promienia na przykładzie: a) litery e, b) litery l

Fig. 2. Geometrical analyses of characters: a) measurement of the length of line segments of each radius for exemplary: a) letter e, b) letter l



Rys. 3. Analizy geometryczne znaków: a) pomiary długości odcinków l powstałych z punktów znaku we fragmentach f dla litery e, b) sumowanie ($\sum l$) pomiarów (a) we fragmentach f , zawierających n promieni dla litery l

Fig. 3. Geometrical analyses of characters: a) measurement of the length of line segments l created by letter points in fragments f for letter e, b) summation ($\sum l$) of measurements (a) in fragments f containing n radiuses for letter l



Rys. 4. Analizy geometryczne znaków: pomiary różnic długości promieni Δr w poszczególnych fragmentach f znaku na przykładzie: a) litery e, b) litery l

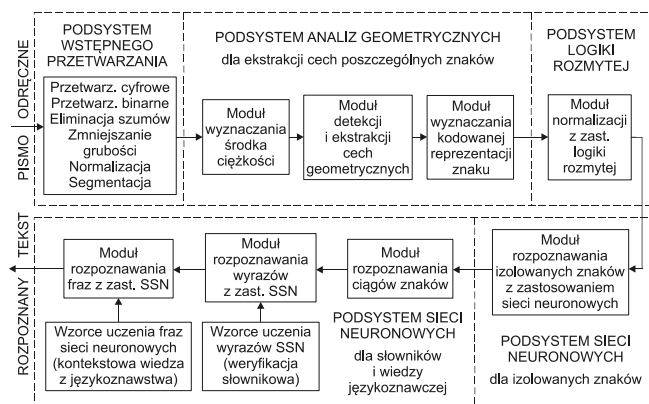
Fig. 4. Geometrical analyses of characters: a) measurements of differences of the radius lengths Δr in each fragment f for exemplary: a) letter e, b) letter l

3. System rozpoznawania pisma odręcznego

W pracach [1, 2] przedstawiono opracowaną metodę rozpoznawania pisma naturalnego (rys. 5). W zaproponowanym rozwiązaniu odręczne pismo operatora podlega procesom przetwarzania wstępnego: przetwarzania cyfrowego i binarnego, eliminowania szumów, zmniejszania grubości znaków, normalizacji oraz segmentacji.

System składa się z podsystemu przetwarzania wstępnego, podsystemu analiz geometrycznych, podsystemu logiki rozmytej, podsystemu sieci neuronowych dla wyizolowanych znaków oraz podsystemu sieci neuronowych dla słowników i wiedzy językoznawczej. Celem przetwarzania wstępnego jest zmniejszenie rozmiarów wejść sieci neuronowych oraz odporność na względn-

ność interpretacji obrazu, ponieważ wszystkie informacje o obrazie są względne do jego środka ciężkości.



Rys. 5. Schemat systemu rozpoznawania odręcznego pisma operatora
Fig. 5. Scheme of an operator's handwriting recognition system

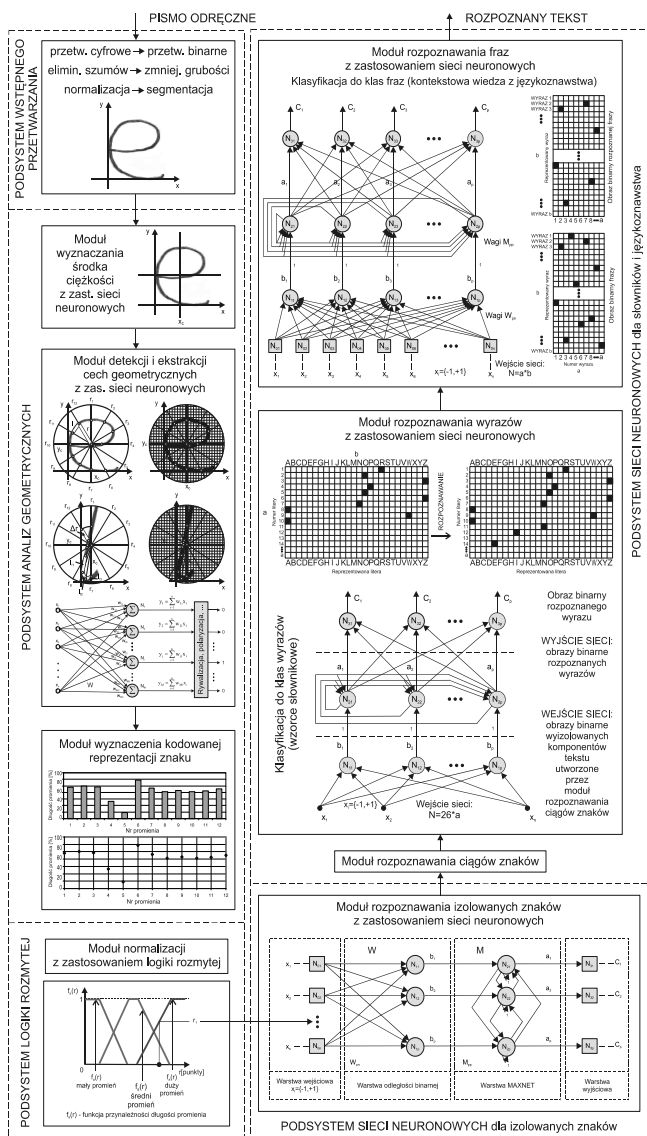
Opracowane analizy geometryczne polegają na przetworzeniu obrazów kształtów znaków do danych je opisujących w postaci zakodowanych reprezentacji znaków. Proces analizy geometrycznej rozpoczyna się od wyznaczenia środka ciężkości znaku (rys. 1a) z zastosowaniem sieci neuronowych, w celu znalezienia punktu początkowego tej analizy. Kolejnym krokiem w algorytmie jest poprowadzenie z punktu środka ciężkości promieni (rys. 1b) o długości odpowiadającej długości odcinka tworzonego przez punkt początkowy oraz punkt, należący do znaku, najbardziej oddalony od punktu początkowego. Powstaje okrąg z wykorzystaniem najbardziej odległego punktu znaku od jego środka ciężkości, który obejmuje cały analizowany znak. Liczba poprowadzonych promieni jest jednym z parametrów dokładności analizy geometrycznej. W miejscach przecięć promieni ze znakiem otrzymywane są punkty, które pozwalają na uzyskanie długości odcinków tworzonych przez punkt początkowy i punkt przecięcia promieni ze znakiem (rys. 2a). Można również wyznaczać punkty przecięć znaków najbardziej odległe od środka ciężkości dla poszczególnych wycinków koła (rys. 2b).

W proponowanej metodzie stosuje się analizy geometryczne polegające na zastosowaniu opracowanych sposobów odpowiednich pomiarów odległości (l) (rys.: 3a, 3b) pomiędzy punktami znaków przez które przebiegają poprowadzone promienie. Analizy geometryczne dotyczą również pomiarów długości odcinków wyznaczonych przez środek ciężkości i punkt znaku leżący na poszczególnych promieniach, najbardziej oddalony od środka znaku. Uzyskuje się w ten sposób różnicę (Δr) (rys.: 4a, 4b) pomiędzy największą długością odcinka a najmniejszą dla każdego z ustanowionych wycinków okręgu. Przy czym efektywność rozpoznawania pisma odręcznego zwiększa się wraz ze wzrostem liczby odpowiednio zdefiniowanych reguł wnioskowania.

Kolejny sposób mógłby uwzględniać liczbę punktów znaku, położonych na poszczególnych promieniach lub w utworzonym fragmencie koła, co może również dostarczać istotną informację dla zakodowanej reprezentacji rozpoznawanego znaku.

W wyniku przeprowadzanej analizy geometrycznej otrzymywane są zakodowane reprezentacje rozpoznawanych znaków na podstawie długości wyznaczonych odcinków (rys. 6). Otrzymane długości odcinków są sygnałami wejściowymi sieci neuronowych Hamminga do rozpoznawania znaków. Sieci Hamminga mają zastosowanie w odpowiednich podsystemach (rys. 6) do rozpoznawania izolowanych znaków, wyrazów oraz fraz.

W sieciach neuronowych Hamminga każdemu wektorowi danych uczących jest przydzielana pojedyncza klasa i podczas fazy rozpoznawania tylko jeden najbliższy wektor do wzorca wejściowego x zostaje znaleziony oraz jego klasa C_i jest zwracana. W tej sieci występują dwie główne fazy działania: uczenia (inicjalizacji) oraz klasyfikacji.



Rys. 6. Algorytm systemu efektywnego rozpoznawania pisma odręcznego
Fig. 6. Algorithm of the system of effective handwriting recognition

Uczenie binarnej sieci neuronowej polega na skopiowaniu wzorców odniesienia do wag macierzy W_{pn} , następująco:

$$w_i = x_i, \quad 1 \leq i \leq p, \quad (1)$$

gdzie p jest liczbą wektorów będących wzorcami wejściowymi x , każdy o tej samej długości n , w_i jest i -tym rzędem macierzy W o wymiarach p rzędów oraz n kolumn. Dla danego n czas obliczeń jest liniowy dla danej liczby wzorców wejściowych p [4].

Zadaniem warstwy rekurencyjnej N_2 jest wybór zwycięskiego neuronu. Cechą charakterystyczną tej grupy neuronów jest połączenie danego neuronu z samym sobą z wagą $m_{ii}=1$ dla wszystkich $1 \leq i \leq p$, podczas gdy wszystkie pozostałe wagi mają ujemne wartości. Inicjalizacja warstwy N_2 polega na przypisaniu ujemnych wartości kwadratowej macierzy M_{pp} z wyjątkiem głównej przekątnej. Pierwotnie Lippmann zaproponował inicjalizację [4]:

$$m_{kl} = -(p-1)^{-1} + \xi_{kl} \quad \text{dla } k \neq l, \quad 1 \quad \text{dla } k = l, \quad (2)$$

$$\text{gdzie } 1 \leq k, l \leq p, p > 1. \quad (3)$$

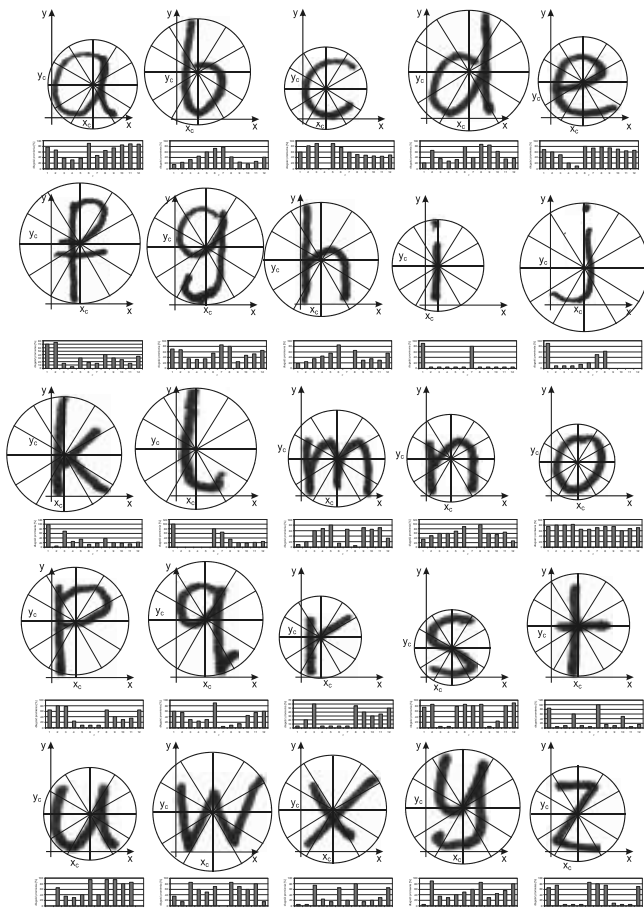
gdzie ξ jest losową wartością dla której $|\xi| \ll (p-1)^{-1}$. Jednak wydaje się, że najbardziej efektywnym i wciąż zbieżnym rozwiązaniem jest ustawienie jednakowych wag dla wszystkich neuronów N_2

które są następnie modyfikowane w każdym kroku podczas fazy klasyfikacji, w następujący sposób:

$$m_{kl} = \varepsilon_k(t) = -(p-t)^{-1} \text{ dla } k \neq l, \quad 1 \text{ dla } k = l, \quad (4)$$

$$\text{gdzie } 1 \leq k, l \leq p, \quad p > 1, \quad (5)$$

gdzie t jest krokiem czasu klasyfikacji. W tym przypadku zbieżność jest osiągnięta w $p-1-r$ kroków, gdzie $r > 1$ oznacza liczbę najbliższych wektorów zapisanych w macierzy W .



Rys. 7. Analiza geometryczna i przykładowe zakodowane reprezentacje liter
Fig. 7. Geometrical analysis and exemplary encoded letter representations

Podczas fazy klasyfikacji, grupa N_1 jest odpowiedzialna za obliczanie odległości binarnej pomiędzy wzorcem wejściowym z oraz wzorcami uczącymi poprzednio zapisanymi w wagach W . Zazwyczaj to jest odległością Hamminga [4]:

$$b_i(z, W) = 1 - n^{-1} D_H(z, w_i), \quad 1 \leq i \leq p, \quad (6)$$

gdzie: $b_i \in [0, 1]$ jest wartością i -tego neuronu w warstwie N_1 ; $D_H(z, w_i) \in \{0, 1, \dots, n\}$ jest odległością Hamminga wzorca wejściowego z oraz i -tego zapisanego wzorca w_i (i -ty rząd W).

W fazie klasyfikacji, warstwa N_2 działa rekurencyjnie, aby wybrać jeden zwycięski neuron. Ten proces jest przeprowadzany przy pomocy następującego równania:

$$a_i[t+1] = \varphi \left(\sum_{j=1}^n m_{ij} a_j[t] \right) = \varphi \left(a_i[t] + \sum_{j=1, j \neq i}^n m_{ij} a_j[t] \right), \quad (7)$$

gdzie $a_i[t]$ jest wyjściem i -tego neuronu warstwy N_2 przy kroku iteracyjnym t , φ jest funkcją progową daną następująco:

$$\varphi(x) = x \text{ dla } x > 0, \quad \varphi(x) = 0 \text{ dla } x \leq 0. \quad (8)$$

W zależności od wybranego schematu (2 i 3)-(4 i 5) wag m_{ij} w (7), otrzymuje się różne dynamiki fazy klasyfikacji. Proces iteracyjny (7) przebiega do punktu, w którym tylko jeden neuron ma wartość inną niż 0, ten neuron jest zwycięzcą.

Moduł rozpoznawania znaków z zastosowaniem sieci neuronowych jest uczony modelami wyizolowanych znaków pisanych odręcznie. Analizy geometryczne znaków oraz ich zakodowane reprezentacje (rys. 7) pozwoliły na wyciągnięcie istotnych wniosków oraz zastosowanie ich w opracowanych algorytmach rozpoznawania pisma odręcznego.

4. Wnioski

Prezentowano opracowaną nowatorską metodę efektywnego rozpoznawania pisma odręcznego z zastosowaniem opracowanych sposobów analiz geometrycznych znaków i wybranych metod sztucznej inteligencji.

Opracowane analizy geometryczne polegają na pomiarach długości odpowiednio wyznaczonych odcinków utworzonych przez wybrane punktu znaku leżące na poszczególnych poprowadzonych promieniach, oraz odcinków pomiędzy środkiem ciężkości znaku i punktem znaku leżącym na poszczególnych promieniach.

Do zalet opracowanej nowej metody efektywnego rozpoznawania pisma odręcznego operatora należy: efektywność i skuteczność metody, analiza geometryczna umożliwiająca rozpoznawanie liter bez względu na krój pisma, styl i charakter pisma, jak również warunki pisania.

W proponowanym rozwiązaniu istnieje możliwość wykorzystania innego typu sieci neuronowych, rozszerzenie zakresu analiz geometrycznych oraz liczne dalsze kierunki rozwoju tej metody.

Proponowany inteligentny system może stanowić nowoczesny i efektywny system interakcji urządzeń technicznych i ich operatorów w zadaniach sterowania.

5. Literatura

- [1] Kacalak W., Majewski M.: A new method for handwriting recognition using artificial neural networks. Artificial Neural Networks in Engineering ANNIE 2006 Conference, Vol. 16, St. Louis, USA, 5-8 November 2006. ASME Press, New York 2006. 459-464.
- [2] Kacalak W., Stuart K., Majewski M.: Selected problems of intelligent handwriting recognition. Analysis and Design of Intelligent Systems using Soft Computing Techniques, Book series of Advances in Soft Computing, vol. 41/2007. Springer 2007. 298-305.
- [3] Kacalak W., Majewski M.: Inteligentny system obustronnej głosowej komunikacji systemu pomiarowego z operatorem dla technologii mobilnych. Pomiary Automatyka Kontrola, Vol. 55, nr 4 (2009). Wydawnictwo PAK 2009. 221-224.
- [4] Lippman R.: An Introduction to Computing with Neural Nets. IEEE Transactions on Acoustic, Speech, and Signal Processing, IEEE Signal Processing Society, Piscataway, 4(3)(1987) 4-22.
- [5] Majewski M.: Podstawy budowy inteligentnych systemów interakcji urządzeń technologicznych i ich operatorów. Monografia nr 172. Wydawnictwo Uczelniane Politechniki Koszalińskiej, Koszalin 2010.