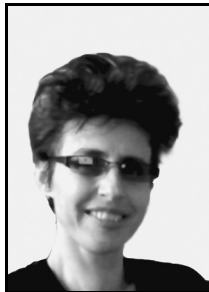


**Teodora DIMITROVA-GREKOW, Agata SWOROWSKA**POLITECHNIKA BIAŁOSTOCKA,  
ul. Wiejska 45A, 15-351 Białystok**Rozpoznawanie wzorców cyfrowych pisma ręcznego z użyciem robota edukacyjnego****Dr inż. Teodora DIMITROVA - GREKOW**

Ukończyła studia na Wydziale Elektroniki Politechniki Sofijskiej, Bułgaria w 1991r., obroniła pracę doktorską na Uniwersytecie Technicznym w Wiedniu w 1997r. Jest adiunktem na Wydziale Informatyki przy Politechnice Białostockiej. Jej zainteresowania naukowe to synteza układów programowalnych, robotyka, mechatronika, analiza i przetwarzanie sygnałów.

e-mail: [t.grekow@pb.edu.pl](mailto:t.grekow@pb.edu.pl)**Inż. Agata SWOROWSKA**

Agata Sworowska ukończyła studia na Wydziale Informatyki, Politechniki Białostockiej w 2013 z tytułem inżyniera na kierunku Informatyka, Politechnika Białostocka, gdzie aktualnie studiuje na kierunku Matematyka. Od 2012 pracuje jako programista w technologii Java. Zainteresowania naukowe to tworzenie aplikacji na urządzenia mobilne oraz analiza i przetwarzanie obrazów.

e-mail: [a.sworowska@gmail.com](mailto:a.sworowska@gmail.com)**Streszczenie**

Niniejsza praca prezentuje zaimplementowanie systemu rozpoznającego ręcznie pisane wzorce cyfrowe z użyciem mobilnego układu edukacyjnego LEGO Mindstorms NXT. Został on wybrany ze względu na prostotę w konstrukcji i równocześnie możliwość złożonego programowania. Zbudowany w ramach projektu robot skanujący znaki pisma ręcznego spełnił założenia początkowe. Wyniki zaimplementowanego algorytmu rozpoznającego również pokryły się z oczekiwaniami - system osiągnął skuteczność na poziomie 100% w warunkach idealnych. We względnie utrudnionych warunkach skuteczność spadła do 91%.

**Słowa kluczowe:** rozpoznawanie wzorców, pismo ręczne, drzewo decyzyjne, graf Hamiltona.

**Handwritten digit pattern recognition based on education robot****Abstract**

Pattern recognition can be classified depending on the data source, the way data is read, processed and on the implementation of the recognition itself [9]. This paper presents a method of pattern recognition identifying handwritten Arabic numbers. The data is collected by a Lego Mindstorms NXT 2.0 mobile robot using a color sensor. Usually, the input data are gathered by high-precision equipment [2,5], and or have an additional multi-sensor subsystem [1]. Very successive recognition approaches are based on neural networks [3, 4,6] additional supported by statistic [8]. Unfortunately, all these methods require powerful calculations. The environment data read by such a simple educational robot contains many drawbacks: noises, relative stabile confidence etc. The solution we propose solves to some extent the problem using a minimal hardware equipment (Fig. 4) and undemanding computation effort. The built recognition system is divided into two parts. The first part presents the data set collection - the handwritten digits scanning (Fig.1) and the data initial processing. The second one consists of primary and secondary classification (Figs. 2 and 3). The algorithm is based on the undirected graph model [10]. The results of the conducted experiments are very interesting (Tabs. 1 and 2). This encourages further exploration of implementation of the well-known and new recognition methods on minimal hardware.

**Keywords:** pattern recognition, handwritten numbers, decision tree, Hamilton graph.

**1. Wstęp**

Rozpoznawanie wzorców PR (ang. *pattern recognition*) jest nauką łączącą obiekty świata naturalnego z maszyną w celu dalszej obróbki tej informacji przez co słusznie zaliczana jest do działu sztucznej inteligencji. Liczba badań w tej dziedzinie nieustannie wzrasta. Rozpoznawanie wzorców można sklasyfikować w zależności od źródła danych, sposobu ich odczytu, zastosowanego aparatu matematycznego a nawet ze względu na cele dalszej obróbki informacji [9]. Obiektem niniejszej pracy jest rozpoznawanie cyfr pisma ręcznego przy użyciu czujnika koloru oraz silników edukacyjnego zestawu firmy Lego Mindstorms NXT 2.0.

Praca każdego kompleksowego systemu wyszukującego i identyfikującego obiekty składa się z szeregu etapów: dokonywania pomiarów, przetwarzania wstępnych danych, segmentacji, normalizacji i w końcu znalezienia najbardziej dopasowanego wzorca (np. numerycznego czy literowego), odpowiadającego badanemu obiektowi. Reprezentacja danych wejściowych zwykle zawiera dużo zakłóceń. Dlatego często pobrane dane pochodzą z wysoko precyzyjnego sprzętu [2, 5], i lub mają dodatkowe wielosensoryczne wejście [1]. Bardzo dobrze sprawdzają się metody rozpoznawania bazujące na sieciach neuronowych [3, 4, 6] dodatkowo wzmacniane zaawansowanymi podejściami statystycznymi [8]. Wszystkie one wymagają jednak zbyt dużo obliczeń, żeby można było je zaimplementować na robocie edukacyjnym.

Proponowane tu rozwiązanie używa minimalną ilość sprzętu oraz wbudowaną wstępną filtrację danych realizowaną podczas ich zbierania. Główny algorytm rozpoznawania został opracowany w oparciu o analizę grafów nieskierowanych [10].

Pomimo uproszczonych warunków przeprowadzone badania pokazały jednak bardzo zachęcające wyniki co pozwala na rozważanie dalszych testów algorytmów rozpoznawania wzorców w poszukiwaniu nieskomplikowanych, a efektywnych rozwiązań.

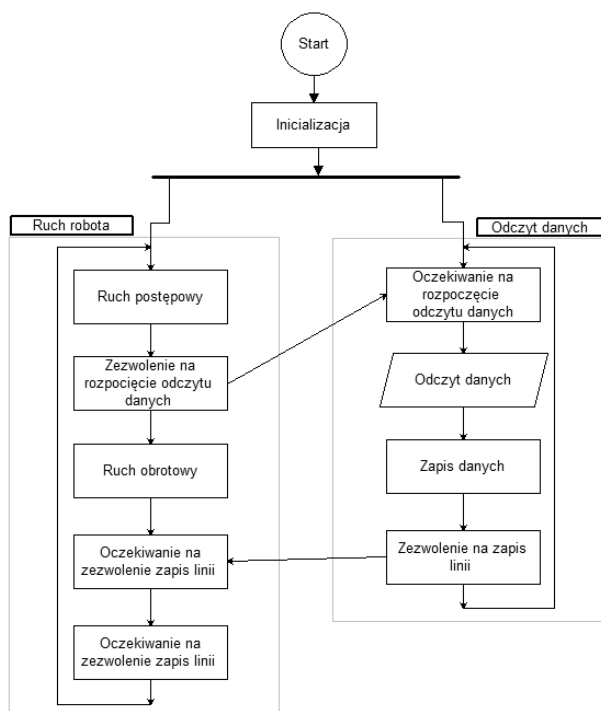
**2. Rozpoznanie wzorów cyfrowych**

Pracę systemu podzielono na dwie główne fazy: tworzenie zbioru danych poprzez skanowanie kartki z ręcznie pisanymi cyframi oraz obróbka informacji w celu rozpoznawania zaskanowanych cyfr. Faza druga ze swojej strony również może zawierać dwa etapy: wstępna i końcowa klasyfikacja. Dalej zostaną omówione poszczególne fazy i etapy.

**2.1. Pobieranie danych**

Zbieranie danych jest oparte na pomiarach natężenia światła za pomocą ruchomego czujnika koloru prowadzonego nad statyczną kartką. Czujnik został zaprogramowany w trybie rozpoznawania kolorów co automatycznie redukuje wiele szumów zewnętrznych. Pomimo, że wybrany sposób wykonania pomiarów zaniża czułość sprzętową, aktualne badania wskazują, że właśnie takie użycie czujnika daje lepsze wyniki w porównaniu do trybu 'surowego' (ang. *raw mode*).

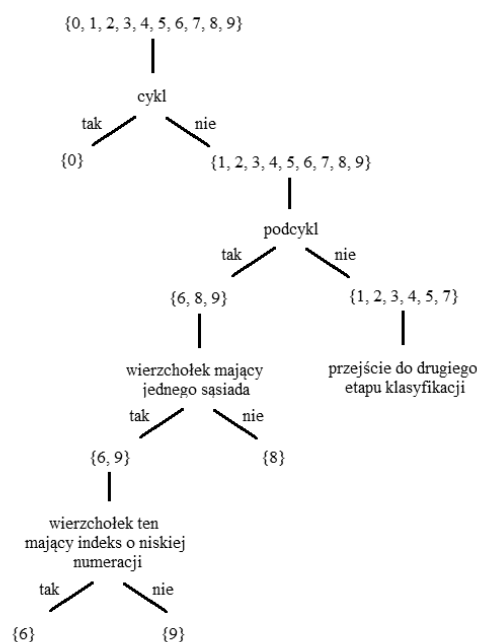
Program zbierający dane ma charakter wielowątkowy. Na schemacie blokowym (rys. 1) widać wątki oraz komunikację między nimi. Całą kinematykę systemu: postępowy ruch robota i obroty ramienia odczytującego wykonuje się w wątku *Ruch\_Robota*. Zbieranie informacji z czujnika umieszczone jest w wątku *Zbieranie\_danych* wykonywanym podczas ruchu obrotowego z pierwszego wątku.



Rys. 1. Algorytm pobierania danych  
Fig. 1. Data collecting algorithm

## 2.2. Klasyfikacja wstępna

Klasyfikacja wstępna polega na zawężeniu ilości możliwych rozwiązań problemu rozpoznawania wzorców cyfrowych. W tym celu został zaimplementowany algorytm odnajdywania cykli w powstałym podczas formatowania danych grafie.



Rys. 2. Drzewo decyzyjne, przedstawiające działanie klasyfikacji wstępnej  
Fig. 2. Decision tree - the initial classification

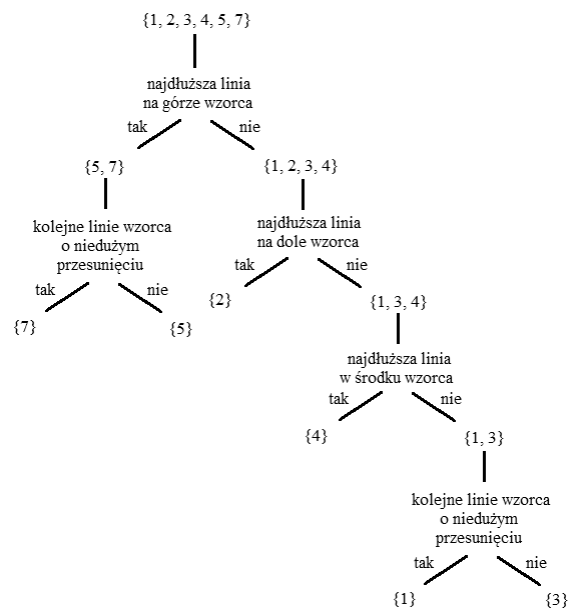
Rys. 2 przedstawia drzewo decyzyjne wykorzystane przy implementacji rozpoznawania wzorców. Jedyną cyfrą będącą grafem hamiltonowskim jest cyfra 0. W przypadku wystąpienia pełnego cyklu zostaje ona rozpoznana. Kolejnym kryterium jest wystąpienie

podcykli. Dzięki temu warunkowi możliwe rozwiązania zostają zawężone do cyfr 6, 8 i 9. Następnie zostaje sprawdzona obecność w grafie wierzchołka o tylko jednym sąsiedzie. Jeżeli jest to spełnione w zbiorze rozwiązań pozostaje tylko 6 i 9. W tym przypadku ważny jest numer wyżej opisanego wierzchołka. Z powodu systemu ich numeracji – od góry do dołu, z lewa na prawo – niski indeks oznacza liczbę 6, wysoki cyfrę 9. W sytuacji, gdy taki wierzchołek nie występuje, rozwiązaniem jest 8. Nerozpoznane cyfry należące do zbioru  $\{1, 2, 3, 4, 5, 7\}$  przechodzą do klasyfikacji końcowej.

## 2.3. Klasyfikacja końcowa

Do tego etapu nie jest potrzebna macierz sąsiedztwa, gdyż ta część algorytmu wykorzystuje oryginalny zestaw danych, na podstawie których była ona stworzona. Rys. 3 pokazuje drzewo decyzyjne wykorzystane przy rozpoznawaniu cyfr należących do zbioru  $\{1, 2, 3, 4, 5, 7\}$ .

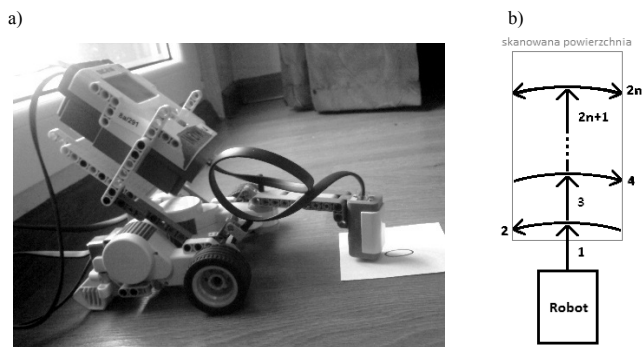
Sprawdzone zostają indeksy najdłuższych poziomych linii w odczytanych danych. Jeżeli taka prosta występuje na górze wzorca rozwiązanie zostaje zawężone do cyfr 5 i 7. Pojawienie się takiej linii w dolnej części wskazuje na znak „2”, natomiast widnienie jej wśród elementów środkowych wyznacza symbol „4”. W przypadku zbiorów  $\{5, 7\}$  i  $\{1, 3\}$  ważne jest zbadanie pojedynczych poziomych linii danych i ich następowanie po sobie. W cyfrach 7 i 1 ich dolne części podzielone na horyzontalne proste charakteryzują nieduże lub żadne przesunięcie względem poprzedniej takiej linii. Dzięki temu te cyfry są odróżniane od znaków „5” i „3”.



Rys. 3. Drzewo decyzyjne, przedstawiające działanie klasyfikacji końcowej  
Fig. 3. Decision tree – the final classification

## 3. Zastosowany sprzęt

Do realizacji zadania wybrany został zestaw LEGO Mindstorms NXT 2.0. Użyta konstrukcję przedstawia rys. 4a. W jej skład wchodzi 3 silniki i sensor koloru. Dwa z silników służy do ruchu postępowego robota. Do trzeciego silnika, zamontowanego pod procesorem, a pomiędzy innymi serwowmotorami, dobudowane jest ramię skanujące. Na końcu tego ramienia znajduje się sensor koloru, odpowiadający za odczytanie wzorca cyfrowego. Ten fragment konstrukcji porusza się po linii łuku. Na rys. 4b są pokazane oba typy ruchów. Najpierw układ mobilny wykonuje ruch postępowy, a następnie ramię z sensorem koloru jest poddawane ruchowi obrotowemu – po linii łuku raz w jedną, raz w drugą stronę.



Rys. 4. a) Robot skanujący; b) kolejność skanowania wzorca  
Fig. 4. a) The scanning robot; b) data scanning sequence

Wzorec odczytywany przez robota za pomocą ramienia skanującego zapisywany jest ręcznie czarnym markerem, na białej kartce.

#### 4. Eksperymenty

Przeprowadzone badania sprawdziły poprawność zaimplementowanego algorytmu oraz możliwość użycia edukacyjnych robotów do zbierania informacji optycznej z realnego środowiska.

Testy były wykonane przy użyciu białych kartek na których czarnym markerem o grubości linii od 1 do 2 milimetrów, zostały ręcznie napisane cyfry. Ich minimalny rozmiar wynosił 15mm na 25mm.

Wszystkie badania zostały wykonane w porze dziennej na stanowisku umieszczonym w pobliżu okna, jednak bez wykorzystania stałej pozycji i oświetlenia.

Od prędkości skanowania zależy liczba próbek na jeden obrót ramienia z sensorem koloru, co zostało dodatkowo wykorzystane do zwiększenia dokładności i skuteczności pracy systemu: skanowano przy użyciu kilkukrotnie większej prędkości w sytuacjach, gdy odczytywane wartości pokazywały niestabilność. Uzyskana w ten sposób rozdzielczość zawiera się w przedziale od 30x12 do 55x20 pikseli.

Pierwszym etapem było zeskanowanie po jednym przykładzie dla każdej cyfry. Test ten bada skuteczność systemu w warunkach prawie idealnych: użyte zostały jedne z najłatwiejszych do rozpoznawania wersji ręcznie napisanych cyfr arabskich. Eksperyment został powtórzony dziesięć razy dla każdej cyfry.

Z tab. 1 widać że dla wszystkich przykładowych wzorców skuteczność algorytmu wynosi 100%. Wyniki te są wstępnym potwierdzeniem założenia, że pomimo ograniczeń układ eksperymentalny jest wystarczająco precyzyjny. Czyli można na nim przeprowadzać wiele interesujących badań z dziedziny obróbki informacji optycznej.

Tab. 1. Skanowanie w warunkach idealnych  
Tab. 1. Scanning under ideal conditions

Cyfra	0	1	2	3	4	5	6	7	8	9
Przykładowy zapis	0	1	2	3	4	5	6	7	8	9
Skuteczność = 100% dla wszystkich testów										

Drugim etapem było przetestowanie rozpoznawalności cyfr o napisanych różnym charakterze pisma. Zapisano każdą cyfrę na różne sposoby i zeskanowano każdej z nich po dziesięciokrotnie.

Tab. 2. Skanowanie w warunkach realnych  
Tab. 2. Scanning under real conditions

Cyfra	0	1	2	3	4	5	6	7	8	9
Skucz, %	90	90	80	100	90	80	100	80	100	100

Tablica 2 przedstawia wyniki drugiego etapu testów. Średnia skuteczność rozpoznawania cyfr jest 91%. Nie dla każdego zapisu ręcznego cyfr można uzyskać stuprocentową rozpoznawalność.

Wadę algorytmu wykazała też cyfra 0 - z powodu niedokładności zapisu klasyfikacja wstępna nie odnalazła pełnego cyklu w grafie i raz została zwrócona cyfra 8.

#### 5. Wnioski

Systemy rozpoznawania pisma zaczęły powstawać już pod koniec lat 80-tych XX wieku. Od tamtej pory ich liczba na rynku zwiększa się. Każdy z nich jest dopracowywany, by zwiększać ich skuteczność, zbliżając ją do 100%.

Wnioskując z przeprowadzonych testów duże znaczenie ma zbliżenie środowiska skanowania do warunków idealnych. Sensor koloru kompatybilny z układami LEGO Mindstorms NXT jest bardzo czuły na zmiany w oświetleniu. Powoduje to błędne odczyty, a co za tym idzie niepoprawne klasyfikowanie wzorców.

Istotne znaczenie ma też sposób zapisu cyfr - ich rozmiar i grubość. Jeżeli jest on nieadekwatny do postawionych warunków w zaimplementowanym algorytmie, skuteczność rozpoznania spada. Jednak zapisując symbol zgodnie z postawionymi założeniami system działa bezbłędnie.

Zbudowany w ramach projektu robot skanujący znaki pisma ręcznego spełnił założenia początkowe. Wyniki zaimplementowanego algorytmu rozpoznającego również pokryły się z oczekiwaniami. System osiągnął skuteczność na poziomie 100% w warunkach idealnych. We względnie utrudnionych warunkach skuteczność spadła do 91%. Możliwe jest poprawienie efektywności poprzez udoskonalenie pracy systemu.

Mimo wykazanych słabości robotów z serii LEGO Mindstorms NXT są one ciekawym narzędziem do eksperymentowania. Pozwalają one na implementacje interesujących i nienadmiernie skomplikowanych algorytmów w praktyce.

#### 6. Literatura

- [1] Pfister M., Behnke S., Rojas M.: Recognition of Handwritten ZIP Codes in a Real-World Non-Standard-Letter Sorting System, Applied Intelligence 12 pp. 95 - 115, Kluwer Acad. Publ. 2000.
- [2] Sesmero M. P., Alonso-Weber J. M., Gutierrez G., Ledezma A., Sanchis A.: A new artificial neural network ensemble based on feature selection and class recoding, Neural Comp&Applic, 21, pp. 771-783, Springer-Verlag London, 2010.
- [3] Hiroki Kurashige, Hideyuki C'ateau: A Method to Construct Visual Recognition Algorithms on the Basis of Neural Activity Data, ICONIP 2011, LNCS 7064, pp. 485-484, Springer-Verlag Berlin Heidelberg 2011.
- [4] Liu Benyong Zhang Jing: An adaptively trained kernel-based nonlinear representation for handwritten digit classification, Journal of Electronics (China), Vol. 23, No. 3, 2006 pp. 379-384.
- [5] Lei Huang, Zhen Li: Feature-based image registration using the shape context, Intern. Journal of Remote Sensing, Vol. 31, No. 8, 2010, pp. 2169- 2177.
- [6] Manabu Kotani, Seiichi Ozaw: Feature Extraction Using Independent Components of Each Category, Neural Processing Letter Vol. 22 2005, pp. 113-124.
- [7] Hairong Lv and Wenyuan Wang: Handwritten Digit Recognition with Kernel-Based LVQ Classifier in Input Space, LNCS 3497, pp. 203-208, 2005, Springer-Verlag Berlin Heidelberg 2005.
- [8] Marisa Morita<sup>1,2</sup>, Robert Sabourin<sup>1,2,3</sup>, Flavio Bortolozzi<sup>3</sup>, Ching Y. Suen: Segmentation and recognition of handwritten dates: an HMM-MLP hybrid approach, IJDAR (2004) 6: pp. 248-262.
- [9] Rahman A. F. R., Fairhurst M. C.: Multiple classifier decision combination strategies for character recognition: A review, IJDAR (2003) 5: pp. 166-194.
- [10] Gajer M.: Systemy optycznego rozpoznawania znaków pisma. Pomiar Automatyka Robotyka 4/2008, s 21-25.
- [11] Steven S. S.: The Algorithm Design Manual. Springer-Verlag, London 2012.