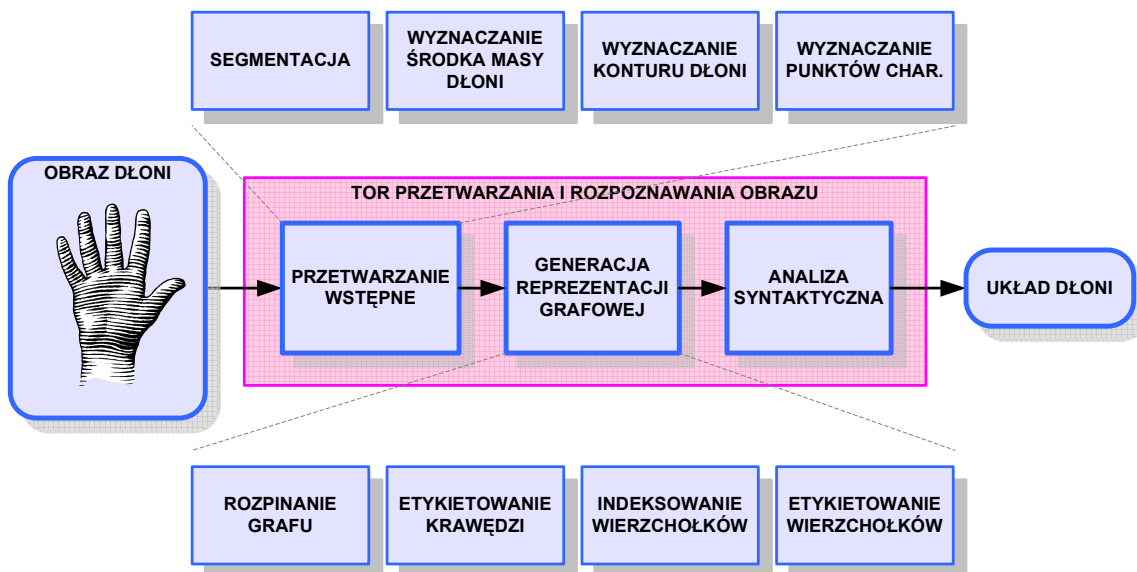


# Rozpoznawanie obrazów dłoni za pomocą gramatyk klasy ETPL(k) w systemach wizyjnych analizy języka migowego.

Autor: mgr inż. Szymon Myśliński  
Afiliacja: Katedra Systemów Informatycznych UJ  
Promotor pracy: Prof. dr hab. Mariusz Flasiński – Uniwersytet Jagielloński  
Recenzenci: Prof. dr hab. inż. Marek Kurzyński – Politechnika Wrocławska  
Prof. dr hab. Marek Ogiela – Akademia Górniczo-Hutnicza

## Cel i zakres pracy

Celem pracy było skonstruowanie metody rozpoznawania układu dłoni (ang. *hand posture*) na obrazie, opartej na modelu parsingu języków grafowych. Metodę przystosowano do obrazów układów dłoni Polskiego Języka Migowego (PJM), który jest podstawowym środkiem komunikacji osób głuchych i niedosłyszących z osobami spoza tego środowiska. Badana metoda klasyfikacji obrazów charakteryzuje się niższą złożonością obliczeniową i większą uniwersalnością od używanego dotychczas podejścia bazującego na reprezentacji grafowej dłoni w procesie rozpoznawaniu obrazów – dopasowania grafów (ang. *graph-matching*). Ta właściwość predestynuje ją do zastosowania w systemach wizyjnych w celu rozpoznawania układów dłoni w czasie rzeczywistym.



Rys. 1 Schemat toru przetwarzania i rozpoznawania obrazu

Rozpoznawanie układu dłoni w prezentowanej metodzie przebiega w następujących etapach (Rys. 1).

1. Obraz wejściowy podlega wstępnemu przetwarzaniu, w wyniku którego wyznaczany jest zbiór punktów charakterystycznych obrazu.
2. Budowana jest grafowa reprezentacja obrazu dłoni, w postaci umożliwiającej zastosowanie efektywnego algorytmu analizy syntaktycznej.

3. W procesie parsingu stworzonej reprezentacji grafowej rozpoznawany jest układ dłoni na obrazie.

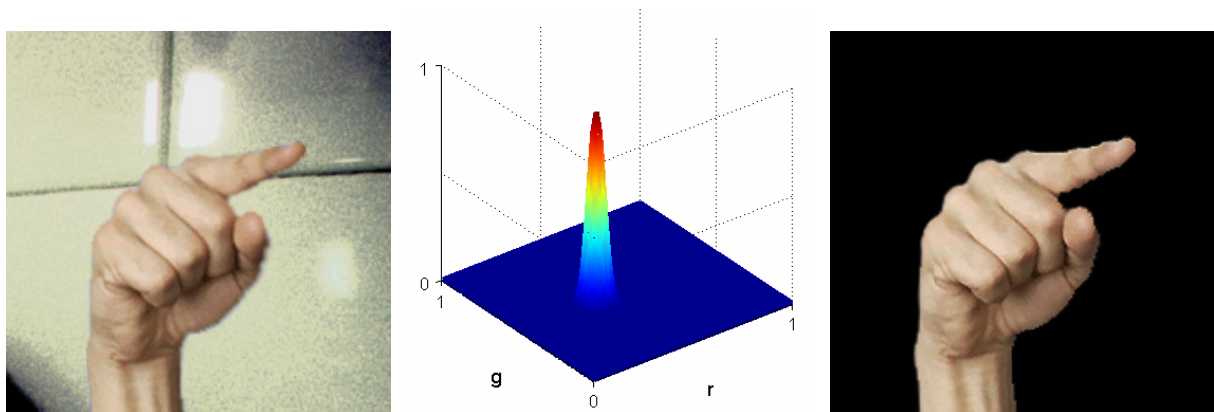
Teza pracy została sformułowana jak następuje. **Syntaktyczne metody rozpoznawania układu dłoni oparte na parsingu języków grafowych pozwalają na uzyskanie najniższej możliwej złożoności obliczeniowej algorytmu rozpoznawania –  $O(n^2)$ , gdzie  $n$  jest liczbą wierzchołków grafu reprezentującego strukturę dłoni.**

W ramach pracy wykonane zostały następujące zadania badawcze.

1. Zidentyfikowanie kroków wstępnego przetwarzania obrazu.
2. Zaprojektowanie algorytmu identyfikacji punktów charakterystycznych na obrazie dłoni.
3. Opracowanie formalnego modelu grafowego reprezentacji obrazów dłoni, dla którego możliwe jest zastosowanie algorytmu parsingu o możliwie najniższej złożoności obliczeniowej.
4. Opracowanie algorytmu rozpinania grafu opisującego przestrzenne relacje pomiędzy wyznaczonymi punktami obrazu dłoni.
5. Zdefiniowanie podklasy TETPL(k) gramatyk grafowych (w ramach klasy ETPL(k)<sup>1</sup>) umożliwiającej generację grafowych reprezentacji obrazów dłoni.
6. Zdefiniowanie algorytmu parsera dla podklasy TETPL(k) gramatyk grafowych.
7. Zaprojektowanie i implementacja parsera podklasy TETPL(k) gramatyk grafowych.
8. Przeprowadzenie eksperymentów weryfikujących opracowaną metodę.

### **Przetwarzanie wstępne**

Pierwszy z etapów przetwarzania – etap przetwarzania wstępnego obrazu – rozpoczyna się od segmentacji obrazu w celu rozdzielenia obszaru dłoni i dopełniającego go obszaru tła. Do tego celu autor wybrał znaną z literatury metodę segmentacji na podstawie modelu barwy skóry ludzkiej (Rys. 2).



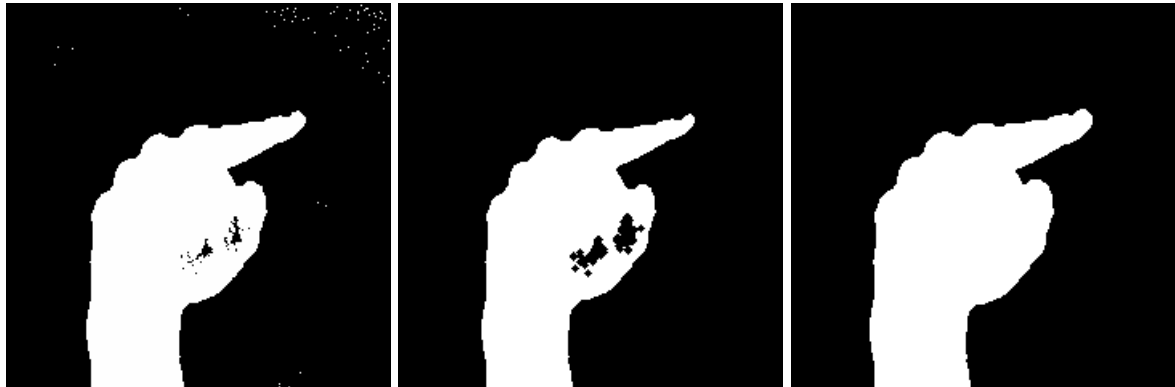
**Rys. 2 Segmentacja obrazu na podstawie modelu barwy skóry ludzkiej**

Wpływ błędów segmentacji, tj. pojedynczych pikseli o wartości 1 na obszarze tła oraz postrzępionych krawędzi obszaru dłoni, na dalsze etapy rozpoznawania obrazu, jest

---

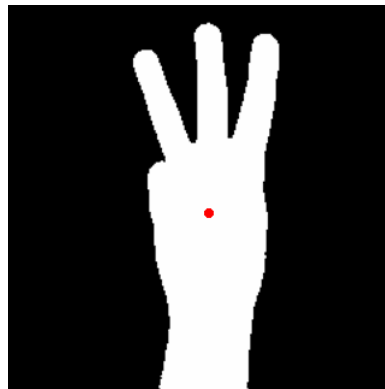
<sup>1</sup> Mariusz Flasiński. *On the parsing of deterministic graph languages for syntactic pattern recognition*. Pattern Recognit., 26:1–16, 1993.

minimalizowany poprzez przeprowadzenie morfologicznej operacji otwarcia przy użyciu kolistego elementu strukturalnego (Rys. 3).



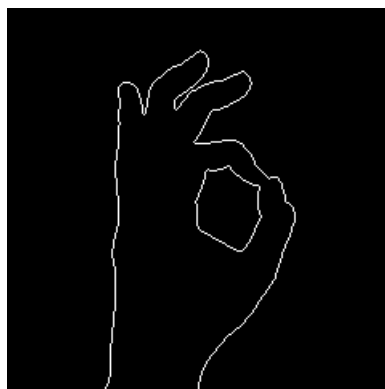
**Rys. 3 Eliminacja błędów segmentacji obrazu metodami morfologii matematycznej**

Dla wyznaczonego w ten sposób obszaru, wyliczane jest położenie punktu środka jego masy. Punkt ten będzie stanowił odniesienie dla pozostałych punktów charakterystycznych wyznaczonych na obszarze dłoni (Rys. 4).



**Rys. 4 Wyznaczanie środka masy obszaru dłoni**

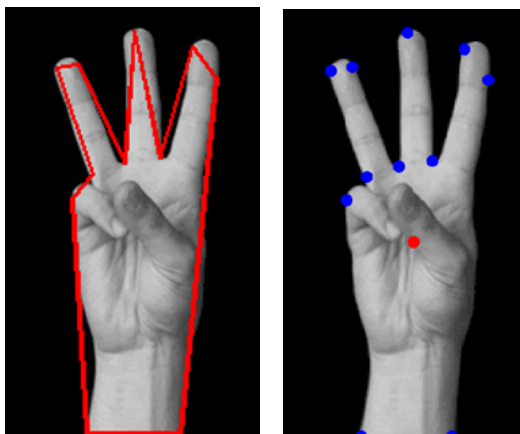
Kolejnym krokiem jest identyfikacja konturów podstawowego i (jeśli taki istnieje) wewnętrznego dłoni (Rys. 5). Ze względu na potrzebę maksymalizacji wydajności procesu, autor zrezygnował ze standardowych metod wykrywania krawędzi na rzecz algorytmu śledzenia konturu obszaru na obrazie binarnym.



**Rys. 5 Wyznaczanie konturu podstawowego i wewnętrznego obszaru dłoni**

Ostatnim krokiem tego etapu było zaproponowanie algorytmu wyznaczającego punkty charakterystyczne obrazu. Autor zaprojektował metodę aproksymacji wyznaczonych

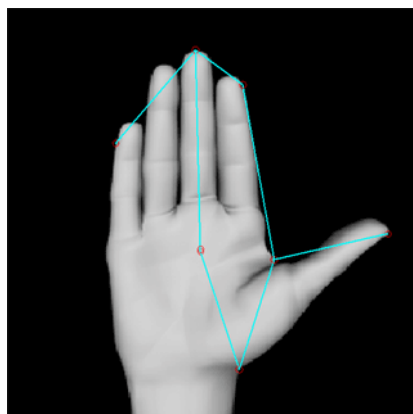
konturów odcinkami. Ich końce, wraz z punktem środka masy obszaru dłoni, tworzą zbiór punktów charakterystycznych obrazu dłoni (Rys. 6).



Rys. 6 Wyznaczanie punktów charakterystycznych obrazu poprzez aproksymację konturu dłoni odcinkami

### ***Generowanie reprezentacji obrazu dłoni***

Kolejny etap przetwarzania polega na budowaniu grafu na wyznaczonych uprzednio punktach. Graf ten stanowi podstawę do klasyfikacji obrazu dłoni na etapie analizy syntaktycznej. Autor skonstruował algorytm rozpinania grafu na punktach charakterystycznych obrazu dłoni. Sposób doboru krawędzi powoduje odzwierciedlenie budowy anatomicznej dłoni w strukturze rozpinanego grafu (Rys. 7).



Rys. 7 Rozpinanie grafu na wyznaczonych punktach charakterystycznych obrazu

W trakcie rozważań, na podstawie przedstawionych przykładów, dowiedziono potrzeby rozszerzenia znanego z literatury formalnego modelu opartego na gramatykach grafowych klasy ETPL(k). Koniecznym okazało się rozwarstwienie zbioru etykiet krawędziowych w celu zapewnienia jednoznaczności reprezentacji grafowej obrazu dłoni. W tym celu autor wprowadził definicję 2DEL-IE grafów (Def. 1).

### Def. 1 2DEL-IE graf

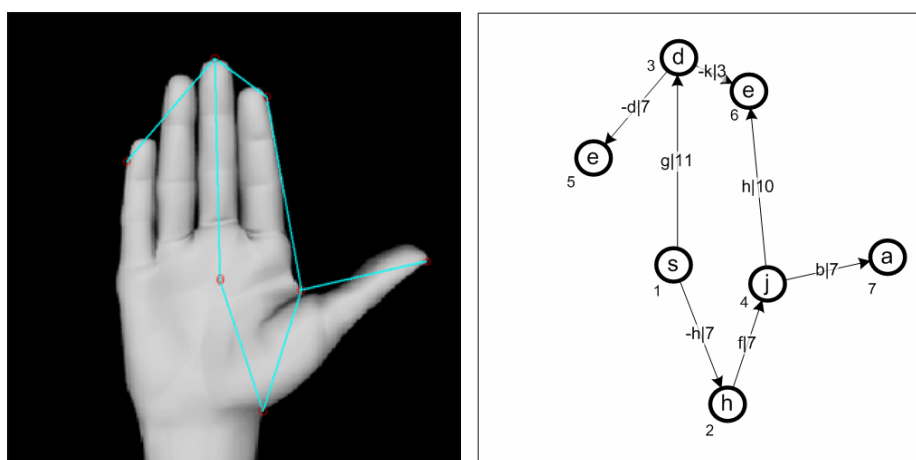
IE grafem o krawędziach z etykietami dwuwymiarowymi, tzw. **2DEL-IE grafem** (ang. Two-Dimensionally Edge Labelled IE Graph), reprezentującym strukturę obrazu nazywamy siódmką:

$$H = (V, E, \Sigma, \Gamma_1, \Gamma_2, \Gamma_3, \varphi)$$

gdzie:

- $V$  jest skończonym, niepustym zbiorem wierzchołków grafu, którym zostały przypisane w sposób jednoznaczny indeksy na bazie obiektów obrazu  $O$ ,
- $\Sigma$  jest skończonym, niepustym zbiorem etykiet wierzchołkowych,
- $\Gamma_1$  jest skończonym, niepustym zbiorem podstawowych etykiet krawędziowych,
- $\Gamma_2$  jest skończonym, niepustym zbiorem pomocniczych etykiet krawędziowych,
- $\Gamma_3$  jest skończonym, niepustym zbiorem nieterminalnych etykiet krawędziowych,  $E$  jest zbiorem krawędzi postaci  $(v, \lambda_1|\lambda_2, w)$ , lub  $(v, \lambda_3, w)$ , gdzie:  $\lambda_1 \in \Gamma_1, \lambda_2 \in \Gamma_2, \lambda_3 \in \Gamma_3, v, w \in V$ , oraz indeks wierzchołka  $v$  jest mniejszy, niż indeks wierzchołka  $w$ ,
- $\varphi : V \rightarrow \Sigma$  jest funkcją etykietowania wierzchołków.

Następnie autor zdefiniował dla nich funkcje etykietujące krawędzie i wierzchołki. Wynikiem tego etapu przetwarzania jest reprezentacja obrazu dłoni w formie 2DEL-IE grafu (Rys. 8), która posłuży do klasyfikacji tego obrazu w dalszej części przetwarzania.



Rys. 8 Reprezentacja obrazu dłoni w formie 2DEL-IE grafu

### ***Analiza syntaktyczna reprezentacji obrazu dłoni***

Etap analizy syntaktycznej zapewnia rozpoznanie układu dłoni na obrazie na podstawie jego reprezentacji grafowej. W wyniku analizy przykładów parsingu stworzonych reprezentacji, autor dowiódł potrzeby dalszej modyfikacji modelu formalnego. Niezbędna okazała się możliwość kasowania krawędzi nieterminalnych w trakcie wyvodu gramatycznego, na co nie pozwalała formalna definicja gramatyk klasy ETPL(k). Autor zdefiniował zatem nową klasę gramatyk TETPL(k), która umożliwia tę operację (Def. 2).

## Def. 2 Gramatyka grafowa klasy TETPL(k)

Gramatykę PL(k) nazywamy **gramatyką grafową klasy TETPL(k)** (ang. Terminal Edge ETPL(k)), jeśli zachodzi następujący warunek. Dla każdej produkcji postaci:

$$(l)A \longrightarrow \begin{array}{cccc} X_1(1) & X_2(2) & \dots & X_m(m) \\ r_1 & r_2 & \dots & r_m \\ E_1 & E_2 & \dots & E_m \\ I_1 & I_2 & \dots & I_m \end{array}, X_e \neq X_f, e, f = 1, \dots, m$$

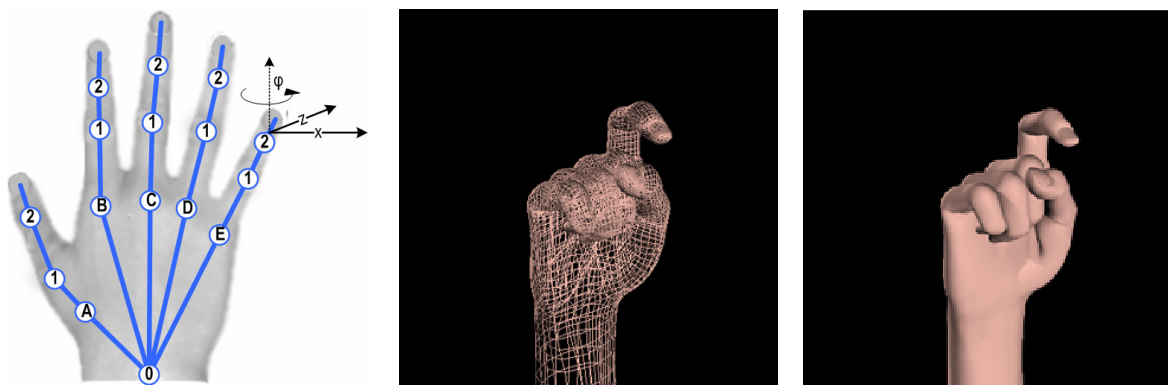
jeśli  $(b, y)$ , gdzie  $y$  jest terminalną etykietą krawędziową, jest potencjalnym kontekstem poprzedzającym dla  $A$ , to istnieje dokładnie jeden element  $(X_i, b, z, in) \in C_l(y, in)$ ,  $i \in \{1, \dots, m\}$ , gdzie  $C_l$  jest transformacją  $l$ -tej produkcji. Jeśli  $i = 1$ , to  $z = y$ , tzn.  $(X_1, b, y, in) \in C_l(y, in)$ .

Wynikiem modyfikacji definicji używanej gramatyki i formalnego modelu grafowego, konieczne stało się ponowne zdefiniowanie związanych z nimi pojęć. Autor wprowadził również algorytm parsera języków TETPL(k) w oparciu o ideę parsera dla języków ETPL(k). Następnie dowiódł, że jego złożoność obliczeniowa to  $O(n^2)$ .

## Eksperymanty

Wszystkie przeanalizowane uprzednio zagadnienia znalazły odzwierciedlenie w projekcie systemu rozpoznawania układu dłoni na obrazie. Dodatkowo, autor przeanalizował wpływ sposobu jego implementacji na wydajność przetwarzania. Zaimplementował on system zgodnie ze stworzonym projektem i przeprowadził pomiary jego wydajności, w celu weryfikacji przyjętych rozwiązań. Ich wyniki potwierdziły słuszność wyboru metody przetwarzania wielowątkowego w celu maksymalizacji wydajności systemu.

Do dalszych eksperymentów niezbędne było zgromadzenie zbioru obrazów testowych. W celu zapewnienia niezależności poprawności rozpoznania od cech anatomicznych dłoni oraz dokładności wykonania jej układów, zdecydowano o generacji obrazów dłoni metodą renderingu. Autor przygotował modele dłoni dla każdego z układów występujących w Polskim Języku Migany (Rys. 9).



Rys. 9 Przygotowanie modeli dłoni i uzyskanie zbioru obrazów metodą renderingu

Następnie, za pomocą serii skryptów, wygenerowano zbiór obrazów referencyjnych oraz lekko zniekształconych. Zniekształcenia te polegały na zmianie proporcji dłoni, dokładności wykonania układu, oraz kąta obserwacji dłoni przez kamerę. Za pomocą skonstruowanego rozwiązania programowego autor przeprowadził serię eksperymentów w celu weryfikacji proponowanej metody.

## **Podsumowanie**

Na podstawie wyników uzyskanych podczas opisanych powyżej badań, można stwierdzić, że postawiony we wstępie cel badawczy został zrealizowany. Co więcej, potwierdzić można tezę stawianą w niniejszej pracy.

Bazując na przebiegu rozważań i wynikach badań przeprowadzonych w niniejszej pracy można sformułować kierunek dalszego rozwoju opisywanej metody. W celu dalszego rozszerzenia jej elastyczności należy wprowadzić mechanizmy analizy syntaktycznej z korekcją błędów dla rozpoznawania obrazów silnie zniekształconych.

## **Publikacje doktoranta**

1. Flasiński M., Jurek J., Myśliński S.: *Parallel computing scheme for graph grammar-based syntactic pattern recognition*, Lecture Notes in Computer Science, Springer, przyjety do druku.
2. Flasiński M., Jurek J., Myśliński S.: *Multi-agent system for recognition of hand postures*, Lecture Notes in Computer Science, 5545, Springer, pp. 815–824, 2009.
3. Myśliński S.: *On the generation of graph representation of hand postures for syntactic pattern recognition*. Journal of Applied Computer Science, 17, 2009.
4. Wołczowski A., Myśliński S.: *Identifying the relation between finger motion and EMG signals for bioprosthesis control*. MMAR06: Proc 12th IEEE int. conf. Methods and Models in Automation and Robotics, pp. 817 - 822, 2006.