

## Katarzyna BARCZEWSKA<sup>1</sup>, Aleksandra DROZD<sup>2</sup>, Łukasz FOLWARCZNY<sup>3</sup>

<sup>1</sup> AGH AKADEMIA GÓRNICZO-HUTNICZA, WEAIIB, KATEDRA AUTOMATYKI I INŻYNIERII BIOMEDYCZNEJ, Al. Mickiewicza 30, Kraków

<sup>2</sup> AGH AKADEMIA GÓRNICZO-HUTNICZA, WEAIIB, KATEDRA METROLOGII I ELEKTRONIKI, Al. Mickiewicza 30, Kraków

<sup>3</sup> SILVERMEDIA Sp. z o. o., ul. Wadowicka 6, Kraków

# Rozpoznawanie gestów z wykorzystaniem czujników inercyjnych o 9 stopniach swobody

### Mgr inż. Katarzyna BARCZEWSKA

Absolwentka Międzywydziałowej Szkoły Inżynierii Biomedycznej, kierunku Inżynieria Biomedyczna, specjalności Elektronika i Informatyka Medyczna na AGH w Krakowie. Aktualnie doktorantka II roku kierunku Biocybernetyka i Inżynieria Biomedyczna w Katedrze Automatyki i Inżynierii Biomedycznej AGH. Zajmuje się biocybernetycznymi systemami analizy ruchu człowieka, rozpoznawaniem gestów, klasyfikacją sygnałów biologicznych.



e-mail: kbarczew@agh.edu.pl

### Mgr inż. Aleksandra DROZD

Absolwentka Matematyki Stosowanej oraz Międzywydziałowej Szkoły Inżynierii Biomedycznej na specjalności Elektronika i Informatyka Medyczna, na Akademii Górniczo-Hutniczej w Krakowie. Jej zainteresowania naukowe obejmują analizę ruchu człowieka, związane z tym systemy kontrolno-pomiarowe, a także modelowanie matematyczne i symulacje komputerowe.



e-mail: drozd@agh.edu.pl

### Mgr inż. Łukasz FOLWARCZNY

W roku 2012 uzyskał tytuł magistra inżyniera na kierunku Informatyka Stosowana wydziału Inżynierii Metali i Informatyki Przemysłowej AGH. Jeszcze w trakcie studiów utworzył i prowadził sekcję „Grupa .net” koła naukowego Metalsoft. Obecnie pracuje w dziale badawczo-rozwojowym firmy Silvermedia, gdzie w współpracy z AGH, zajmuje się algorytmami cyfrowego przetwarzania, analizy i klasyfikacji sygnałów biomedycznych.



e-mail: lukasz.folwarczny@silvermedia.pl

**Keywords:** gesture recognition, inertial sensor, DTW.

## 1. Wstęp

Wraz z rozwojem techniki powstaje coraz więcej alternatywnych sposobów komunikacji człowieka z komputerem, sterowania inteligentnym otoczeniem, czy też grammi komputerowymi. Urządzenia rozpoznające gesty człowieka na potrzeby gier komputerowych mogą być wykorzystane również z myślą o osobach niemówiących w celu rozpoznawania języka gestów, którym posługują się na co dzień, co ułatwi im kontaktowanie się z otoczeniem.

Istnieje zarówno wiele systemów gestów, np. 130 różnych języków migowych [1], systemy gestów, którymi posługują się osoby z zespołem Downa [2], jak i technik akwizycji danych oraz metod ich przetwarzania i analizy. Jednym z pierwszych podejść do tego problemu było skonstruowanie systemu tłumaczącego gesty tajwańskiego języka migowego [3], w którym do akwizycji danych wykorzystano wirtualną rękawiczkę – DataGlove (zawierającą sensory zgięcia i dotyku) i czujnik inercyjny. Jednak wirtualne rękawiczki nie są tak dostępnym i powszechnym produktem, dlatego wykorzystuje się też inne rozwiązania. W związku z tym, że czujniki inercyjne zaczęły być wbudowywane w większość nowoczesnych telefonów komórkowych i stały się znacznie bardziej dostępne dla każdego użytkownika, powstają także systemy rozpoznające proste gesty na podstawie analizy przyspieszenia. Przykładami mogą być systemy opisane w [4-6]. Do akwizycji danych bardzo często wykorzystywane są także nagrania z kamer video [7-11]. Alternatywnym rozwiązaniem dla kamer video, jest urządzenie Kinect, wyprodukowane przez firmę Microsoft do konsoli do gier Xbox 360 [12-14].

Akwizycja sygnałów w trakcie wykonywania gestu jest pierwszym etapem złożonego procesu rozpoznawania. Kolejny etap to albo wskazanie wzorców dla każdego z gestów, albo zbudowanie modelu uwzględniającego charakterystyczne cechy, odróżniające dany gest od innych. W przypadku rozpoznawania gestów przykładowymi cechami charakterystycznymi mogą być: wzajemne położenie rąk, pozycja rąk w przestrzeni, trajektoria ruchu występującego w przekazie znaku, sposób wykonywania ruchu [15]. Do wyznaczenia cech rozróżniających poszczególne gesty między sobą stosuje się m.in. testy statystyczne [11], statystyki niezmiennicze [10], czy też inne metody prowadzące do redukcji informacji. W przypadku wyznaczenia wzorca, stosowany jest np. algorytm *affinity propagation* [6].

Na podstawie zredukowanych cech i modelu lub porównania do wzorca odbywa się klasyfikacja gestów. Wykorzystywane są również różne inne metody: począwszy od sieci neuronowych [10] przez ukryte modele Markova HMM [3, 13], klasyfikatory statystyczne [11], modyfikacje algorytmu DTW [5, 7].

## Streszczenie

Rozpoznawanie gestów za pomocą czujników inercyjnych może być alternatywą dla standardowych interfejsów człowiek-komputer. Do śledzenia gestów wykorzystano czujnik zawierający trójosiowy akcelerometr, magnetometr i żyroskop. W dotychczasowych badaniach bazowano na sygnałach przyspieszenia. Autorzy zaproponowali i porównali rozwiązania wykorzystujące zarówno analizę przyspieszenia, jak i orientacji w przestrzeni, a także umożliwili badanym osobom wykonywanie gestów w sposób naturalny. Wyniki pokazują, że za pomocą algorytmu DTW (Dynamic Time Warping) możliwa jest klasyfikacja indywidualna dla danej osoby (ze skutecznością 92%), a także klasyfikacja uogólniona - na podstawie uniwersalnego wzorca (ze skutecznością 83%).

**Słowa kluczowe:** rozpoznawanie gestów, czujnik inercyjny, DTW.

## Gesture recognition based on 9DOF inertial sensor

### Abstract

Gesture recognition may be applied to control of computer applications and electronic devices as an alternative to standard human-machine interfaces. This paper reports a method of gesture classification based on analysis of data from 9DOF inertial sensor - NEC-TOKIN, Motion Sensor MDP-A3U9S (Fig.1). Nine volunteers were asked to perform 10 different gestures (shown in Fig.2) in a natural way with a sensor attached to their hand. The gesture data base consisting of 2160 files with triaxial acceleration and orientation signals was created. In the first step the data were divided into training and testing sets. The designed system uses the Dynamic Time Warping (DTW) algorithm to calculate similarity of signals (formulas (1)-(3)). Using this method the authors chose representative signals to individual and generalized exemplars data base from the training set. The DTW algorithm was also used in the classification process. Different recognition approaches were tested basing on acceleration-only, orientation-only and acceleration-orientation signals. The results listed in Tab.4 show that the best recognition efficiency of 92% was obtained in the individual recognition (only one person gestures taken into account) for modified exemplars data base. The modification proposed by the authors (Section 3) improved the recognition rate by 10 percentage points. The efficiency rate of 83% (Tab. 5) was reached in the generalized case. The next step of improving the designed recognition system is application of an inertial system with a bluetooth module and real-time gesture classification.

## 2. Materiały i metody

### 2.1. Akwizycja danych

W niniejszej pracy zdecydowano się na wykorzystanie sygnałów z czujnika inercyjnego, zawierającego układy trójosiowe: akcelerometr, magnetometr oraz żyroskop - ze względu na dostępność tego typu czujników i niskie koszty produktu. Przegląd dostępnej literatury pokazuje, że w standardowym podejściu do rozpoznawania gestów wykorzystywane są jedynie dane z czujnika przyspieszenia. Dlatego dodatkowym celem pracy była odpowiedź na pytanie, czy inne dane uzyskane z sensora (np. orientacja w przestrzeni) mogą znacząco poprawić klasyfikację.

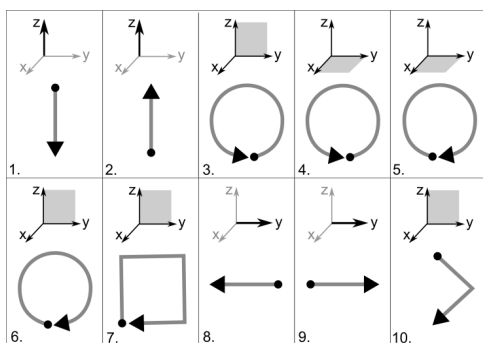
Urządzeniem wykorzystanym do akwizycji danych był czujnik inercyjny o 9 stopniach swobody firmy NEC-TOKIN, Motion Sensor MDP-A3U9S (rys. 1) z trójosiowym: akcelerometrem, magnetometrem oraz żyroskopem.



Rys. 1. Po lewej: umiejscowienie czujnika na dłoni podczas wykonywania gestów wraz z odniesieniem do związanego z nim układu współrzędnych. Po prawej: czujnik NEC-TOKIN, Motion Sensor MDP-A3U9S

Fig. 1. On the left: sensor on the hand during data acquisition. On the right: NECTOKIN, Motion Sensor MDP-A3U9S

Na wyjściu czujnika, oprócz danych dla każdego sensora, otrzymuje się informację na temat orientacji w przestrzeni opisaną za pomocą kątów Eulera. W trakcie akwizycji danych osoba wykonująca gesty miała czujnik przymocowany do palca wskazującego prawej dłoni. Niewielkie rozmiary (20 mm × 20 mm × 15 mm) oraz waga (6 g) modułu pomiarowego pozwalały na swobodne poruszanie dłonią i zginanie palca. Transmisja danych do komputera odbywała się przewodowo przez port USB. Częstotliwość próbkowania wynosiła 25 Hz. Rozpoznaniu poddano 10 prostych gestów, omawianych także w [5, 6, 16]. Schematy rozpoznawanych gestów zestawiono na rys. 2.



Rys. 2. Schematy rozpoznawanych gestów wraz z zaznaczoną płaszczyzną lub kierunkiem ruchu względem osi czujnika.

Fig. 2. Diagrams of the recognized gestures according to the sensor coordinate system.

### 2.2. Baza danych

Zbudowano bazę danych liczącą 2160 plików, z których każdy odpowiadał pojedynczemu gestowi. W 3 sesjach pomiarowych wzięło udział 9 osób, które w każdej pojedynczej sesji wykonywały po 8 powtórzeń dla każdego z gestów. Bazę danych gestów podzielono na 2 zbiory: zbiór treningowy i testowy. Do zbioru treningowego zaliczono po 5 z 8 powtórzeń każdego gestu dla

każdej sesji akwizycji danych, pozostałe 3 powtórzenia zostały włączone do zbioru testowego. Na podstawie analizy zbioru treningowego dla każdej osoby, dla każdego gestu wyznaczono gest wzorcowy, do którego porównywano gesty ze zbioru testowego, sprawdzając skuteczność działania klasyfikatora.

W przeciwieństwie do [5] oraz [6], gdzie założono, że użytkownicy nie zmieniają nachylenia dłoni w trakcie wykonywanego ruchu, pozwolono użytkownikom naturalnie poruszać ręką, przez co, oprócz położenia, zmieniało się także nachylenie dłoni. Dlatego do rozpoznawania gestów wykorzystano zarówno dane z akcelerometru, jak i wartości kątów Eulera w celu wskazania parametrów, które mają większe znaczenie w procesach klasyfikacji.

### 2.3. Algorytm

Podstawowym algorytmem wykorzystanym zarówno przy wyznaczaniu wzorców, jak i klasyfikacji gestów był algorytm DTW (*Dynamic Time Warping*) [17]. Algorytm ten znajduje szerokie zastosowanie w wyznaczaniu podobieństwa między szeregami czasowymi, dla których dopuszcza się transformację czasową. Wybór metody został podyktowany charakterem badanych sygnałów. Zebrane w bazie danych sygnały, odpowiadające tym samym gestom dla różnych osób, charakteryzowały się zmiennością w długości sygnału, rozkładzie czasowym lokalnych ekstremów oraz prędkości wykonywania poszczególnych faz gestu.

Algorytm DTW pozwala na wyznaczenie najmniejszej odległości pomiędzy dwoma sygnałami w następujący sposób. Załóżmy, że mamy dane dwa sygnały, które są opisane za pomocą szeregów czasowych  $X = \{x_1, x_2, \dots, x_n\}$ ,  $Y = \{y_1, y_2, \dots, y_m\}$ . Dla  $X$  i  $Y$  definiujemy macierz odległości  $D$  według następującej reguły:

$$D(i, j) = d(x_i, y_j) \quad (1)$$

gdzie:

$$d(x_i, y_j) = |x_i - y_j| \quad (2)$$

Macierzą kumulacyjną  $P$  nazywamy macierz zdefiniowaną następująco:

$$\begin{aligned} P(1, 1) &= 0 \\ P(i, 1) &= D(i, 1) + P(i-1, 1) \\ P(1, j) &= D(1, j) + P(1, j-1) \end{aligned}$$

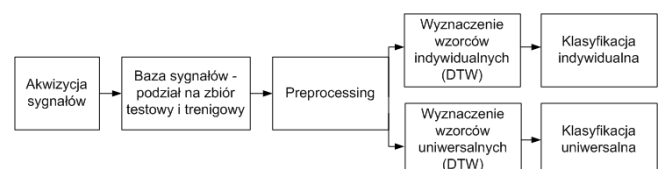
Dla  $i, j > 1$

$$P(i, j) = D(i, j) + \min\{P(i-1, 1), P(i, j-1), P(i-1, j-1)\} \quad (3)$$

Wynikiem działania algorytmu jest zoptymalizowana sumaryczna odległość  $X$  i  $Y$  po dopasowaniu, którą oznaczono  $q_{DTW} = P(n, m)$ .

### 2.4. Klasyfikacja

Klasyfikacja gestów była przeprowadzana dwutorowo (rys. 3).



Rys. 3. Ogólny schemat działania systemu do rozpoznawania gestów  
Fig. 3. General diagram of functioning the gesture recognition system

Pierwszym zadaniem było rozpoznawanie gestów danej osoby na podstawie wzorca wybranego spośród jej gestów (wzorec indywidualny). Uogólnienie tego podejścia doprowadziło do prób klasyfikacji gestów dowolnej osoby na podstawie uniwersalnego wzorca.

Wyznaczenie bazy wzorców indywidualnych odbyło się poprzez porównanie ze sobą wszystkich powtórzeń tego samego gestu ze zbioru treningowego i wyliczenie wartości  $q_{DTW}$  dla wszystkich kombinacji sygnałów dla danej osoby i danego gestu ze zbioru treningowego. Wzorcem zostawał ten, którego suma odległości  $q_{DTW}$  wyliczonych w trakcie porównywania ze wszystkimi pozostałymi gestami była najmniejsza. W podejściu uogólnionym wzorce uniwersalne dla każdego gestu były wybierane z bazy wzorców indywidualnych. Brano pod uwagę sumaryczne odległości każdego z potencjalnych wzorców od sygnałów ze zbiorów treningowych.

Każdorazowo przed zastosowaniem algorytmu DTW, sygnał poddawano wstępnemu przetworzeniu. W celu zminimalizowania różnicy pomiędzy tymi samymi gestami wykonywanymi w różnym czasie, dokonano przekształcenia sygnałów, dzieląc je przez maksymalną bezwzględną amplitudę występującą w sygnale. Sygnały przyspieszeń były dodatkowo poddawane filtracji uśredniającej. Dla każdego zestawu 6 badanych sygnałów dla jednego gestu (przyspieszenia w osi  $x, y, z$  – ozn.  $a$ , oraz zmiany kątowe – ozn.  $\varphi$ ) określano próbkę w czasie, w której rozpoczął się i kończył gest. Do wyznaczenia początku i końca gestu wykorzystano zmiany odchylenia standardowego dla zmian kątowych, które dla tego celu dawały znacznie lepsze rezultaty niż sygnały przyspieszenia.

Rozpoznanie i klasyfikacja gestu polegały na wyliczeniu odległości  $q_{DTW}$  od każdego z 10 wzorców z odpowiedniej bazy wzorców – albo indywidualnej albo uniwersalnej. Najmniejsza odległość  $q_{DTW}$  wskazywała na odpowiadający wzorec. Aby sprawdzić, jaki wpływ na wyniki ma dobór analizowanych parametrów, system testowano uwzględniając sygnały przyspieszeń, sygnały zmian kątowych, a także oba rodzaje sygnałów równocześnie.

### 3. Wyniki

Przetestowanie systemu dla każdego gestu z bazy gestów testowych wymagało przeprowadzenia klasyfikacji indywidualnej i uniwersalnej, a następnie przedstawienia wyników za pomocą macierzy dopasowania. Każdy wiersz macierzy odpowiada jednemu sklasyfikowanemu gestowi. Na przecięciu  $k$ -tego wiersza i  $j$ -tej kolumny znajduje się liczba odpowiadająca procentowej ilości gestu  $k$ , który został rozpoznany jako gest  $j$ . Zatem występowanie wartości zbliżonych do 100% na przekątnej macierzy dopasowania świadczyło o dużej skuteczności rozpoznawania gestów. Wyliczono też całkowitą skuteczność rozpoznania wszystkich gestów (ozn.  $SK$ ) jako stosunek liczby gestów poprawnie rozpoznanych do wszystkich gestów ze zbioru testowego.

Tab. 1. Porównanie skuteczności rozpoznania gestów w metodach z wykorzystaniem przyspieszenia, orientacji w przestrzeni lub obu tych danych

Tab. 1. Comparison of the gesture recognition efficiency for different algorithms

Rodzaj rozpoznawania	SK - $a$ [%]	SK - $\varphi$ [%]	SK - $a, \varphi$ [%]
Indywidualne (bez modyfikacji)	84	69	84
Indywidualne (po modyfikacji)	92	76	92
Uniwersalne (po modyfikacji)	83	55	77

Jak zaznaczono wcześniej, do klasyfikacji użyto sygnałów przyspieszeń w trzech osiach o orientacji w przestrzeni wyrażonej za pomocą kątów Eulera. Tabela 1 pokazuje skuteczność rozpoznania przy uwzględnieniu samych przyspieszeń ( $a$ ), samych zmian kątowych ( $\varphi$ ) oraz obu danych równocześnie, branych pod uwagę z taką samą wagą ( $a, \varphi$ ).

Wyniki pokazują, że skuteczność rozpoznania jest mniejsza w przypadku rozpoznawania na podstawie zmian orientacji w przestrzeni. W rozpoznaniu z wykorzystaniem najprostszego algorytmu polegającego na minimalizacji sumy wszystkich wartości  $q_{DTW}$  dla sygnałów przyspieszenia i orientacji nie uzyskano wyników lepszych niż w przypadku samego przyspieszenia. Może to wynikać z przypisania równych wag obu sygnałom. Ze względu

na większą skuteczność klasyfikacji z wykorzystaniem sygnału przyspieszenia, poniżej zaprezentowano wyniki odnoszące się do tego typu klasyfikacji. Po przeprowadzeniu pierwszych testów, okazało się, że zaproponowany podstawowy algorytm daje bardzo dobre wyniki przy rozpoznawaniu prostych gestów, natomiast występuje duży problem z rozpoznawaniem gestu 10. Gdyby rozważane były gesty z z jego pominięciem, skuteczność rozpoznania indywidualnego wzrosłaby do  $SK_1 = 93\%$ , jednak z jego uwzględnieniem  $SK_1 = 84\%$ . Dla przypadku uniwersalnego skuteczność wyniosła  $SK_U = 73\%$ .

Aby poprawić skuteczność rozpoznawania wszystkich gestów, zaproponowano nową metodę, polegającą na modyfikacji bazy wzorców. Wykonywane gesty są różne dla badanych osób, ponieważ jednak badani mieli wzorować się przy ich wykonywaniu na przedstawionych schematach (jak na rys. 2), można założyć, że gesty powinny być wykonywane w określonych płaszczyznach, a zatem pewne wartości przyspieszeń powinny być bliskie zera.

Zbudowano nową bazę wzorców, w których zastąpiono zarejestrowane wartości pewnych przyspieszeń wartościami zerowymi i przetestowano algorytm z tak zmienioną bazą.

Poniższa tabela pokazuje, które sygnały były zastępowane zerami dla odpowiednich gestów.

Tab. 3. Sygnały, które uległy zmianom przy modyfikacji bazy wzorców  
Tab. 3. Gestures and signals in the modified exemplar database

GEST	SYGNAŁY ZASTĘPOWANE PRZEZ „0”
Gest1 - dół	przyspieszenie w osi $y - ay$
Gest2 - góra	przyspieszenie w osi $y - ay$
Gest4 - koło płaskie lewo	przyspieszenie w osi $z - az$
Gest5 - koło płaskie prawo	przyspieszenie w osi $z - az$
Gest8 - lewo	przyspieszenie w osi $z - az$
Gest9 - prawo	przyspieszenie w osi $z - az$

Dla opisanej modyfikacji bazy wzorców wyliczono skuteczność oraz macierze dopasowania. W macierzach przedstawiono uśrednione procentowe wartości rozpoznania dla wszystkich badanych osób. Wyniki rozpoznania przedstawiono w tab. 4.

Tab. 4. Macierz dopasowania indywidualnego dla klasyfikacji z wzorcami, w których zerowano wartości przyspieszeń (dół, góra, koło płaskie lewo, koło płaskie prawo). Wartości są podane w [%].  $SK_1 = 92\%$

Tab. 4. Confusion matrix for the individual recognition algorithm with exemplar modification. The values are given in [%]. The total efficiency  $SK_U = 92\%$

Gest	1	2	3	4	5	6	7	8	9	10
1	97,5	1,2	0	0	0	0	0	0	0	1,2
2	1,2	95,1	0	0	0	1,2	1,2	0	0	1,2
3	1,2	1,2	86,4	1,2	0	4,9	2,5	1,2	0	1,2
4	0	1,2	8,6	86,4	2,5	0	0	1,2	0	0
5	0	1,2	0	3,7	90,1	3,7	1,2	0	0	0
6	0	0	0	0	0	100	0	0	0	0
7	0	1,2	2,5	0	0	8,6	82,7	0	1,2	3,7
8	0	0	1,2	0	0	0	0	98,8	0	0
9	0	0	0	1,2	0	0	0	2,5	93,8	2,5
10	2,5	0	0	0	2,5	0	1,2	1,2	1,2	91,4

Wyniki pokazują, że w przypadku rozpoznawania trudniejszych gestów, zaproponowana modyfikacja znacząco zwiększa skuteczność klasyfikacji. Stosując zmodyfikowaną bazę wzorców pokazano, że możliwa jest również klasyfikacja uniwersalna, w której baza gestów wzorcowych była wspólna dla wszystkich badanych.

Wyniki te pokazują, że możliwa jest dobra klasyfikacja gestów w przypadku uniwersalnym, a zastosowanie proponowanej modyfikacji zwiększa skuteczność rozpoznania o około 10 punktów procentowych. Porównując klasyfikację uniwersalną i indywidualną należy zwrócić uwagę na fakt, iż w przypadku ogólnym znacząco pogorszyła się skuteczność rozpoznania takich gestów jak koło czy kwadrat. Wiąże się to z faktem, że w przypadku

wielu badanych różne gesty (np. koło i kwadrat) mogą być wykonywane w bardzo podobny sposób.

Tab. 5. Macierz dopasowania uniwersalnego dla klasyfikacji z wzorcami, w których zerowano wartości przyspieszeń (dół, góra, lewo, prawo). Wartości są podane w [%].  $SK_U = 83\%$

Tab. 5. Confusion matrix for the general recognition algorithm with exemplars modification. The values are given in [%]. The total efficiency  $SK_U = 83\%$

Gest	1	2	3	4	5	6	7	8	9	10
1	96,3	0	0	0	0	1,2	0	1,2	1,2	0
2	0	95,1	1,2	0	0	0	0	0	2,5	1,2
3	1,2	0	48,1	7,4	1,2	25,9	2,5	12,3	0	1,2
4	0	0	1,2	88,9	3,7	0	0	3,7	2,5	0
5	0	0	0	0	81,5	12,3	0	0	6,2	0
6	0	0	1,2	1,2	1,2	96,3	0	0	0	0
7	0	1,2	2,5	1,2	2,5	23,5	65,4	0	0	3,7
8	0	2,5	0	0	0	0	0	92,6	0	4,9
9	0	2,5	0	0	3,7	8,4	1,2	1,2	81,5	2,5
10	1,2	0	0	1,2	0	3,7	3,7	0	1,2	88,9

W tabeli 6 zestawiono porównanie skuteczności rozpoznawania podobnych zestawów gestów uzyskane przez różnych autorów. Niższa wartość  $SK_I$  związana jest z faktem, że użytkownicy mogli wykonywać wskazane gesty w naturalny sposób, podczas gdy u pozostałych autorów nie mogli zginać dłoni, ani zmieniać jej nachylenia. Jednak i tak wysokie wartości  $SK_I = 92\%$  oraz  $SK_U = 84\%$  pokazują, że rozpoznawanie naturalnie wykonywanych gestów jest również możliwe.

Tab. 6. Zestawienie wyników rozpoznawania. Oznaczenia: skuteczność rozpoznawania  $SK_I$  - indywidualnego,  $SK_U$  - uniwersalnego, n - ilość gestów

Tab. 6. Summary of the recognition results. Explanation of abbreviations: the efficiency of recognition: individual -  $SK_I$ , universal -  $SK_U$ , n - number of gestures

Metoda	n	$SK_I$ [%]	$SK_U$ [%]	Naturalne gesty
Zaproponowana	10	92,0	84,0	Tak
uWave [16]	8	98,9	75,4	Nie
Continuous HMM [6]	8-18	99,8	75,7	Nie
DTW [6]	8-18	-	69,0	Nie
RP matrix [6]	8-18	99,9	97,6	Nie
Acc. DTW [5]	8	99,2	96,4	Nie

#### 4. Wnioski

Naturalnie wykonywane gesty są odmienne nie tylko dla różnych osób, ale nawet dla tej samej osoby, wykonującej dany ruch w większych odstępach czasowych. Co więcej, czasami zdarza się, że zupełnie różne gesty wykonywane są przez dwie osoby w bardzo zbliżony sposób. Odzwierciedla się to oczywiście w skuteczności ich rozpoznania (tab. 5). Rozwiązaniem tego problemu mogłaby być wcześniejsza demonstracja każdego z gestów, na przykład w postaci pliku video, który stanowiłby instrukcję dla użytkownika, jak wykonywać dany ruch.

W zależności od celu rozpoznawania gestów można dobrać odpowiednią metodę poprawy skuteczności klasyfikacji: od dalszej modyfikacji wzorców, przez zmianę zestawu gestów do rozpoznania (dla zastosowań związanych ze sterowaniem prostymi aplikacjami), po rozbudowanie algorytmu klasyfikacji. W zaproponowanym podejściu nie wykazano wpływu uwzględnienia zmian kątowych na poprawę skuteczności rozpoznania gestów. Zmiany kątów w czasie pozwoliły dokładniej niż w przypadku przyspieszeń wyznaczyć początek i koniec pojedynczego gestu. Ze względu na to, że dla danego gestu niektóre składowe wzorca są bardziej charakterystyczne, w perspektywie dalszych badań rozważa-

ne jest wzbogacenie klasyfikacji opartej na sygnałach przyspieszenia o dane zmian kątowych, uwzględniając je z różną wagą w zależności od gestu. Autorzy rozważają wykorzystanie transmisji bezprzewodowej, a także rozpoznawanie gestów w czasie rzeczywistym.

*Praca była wykonywana w ramach działalności statutowej AGH, grant nr 11.11.120.612*

#### 5. Literatura

- [1] www.ethnologue.com, data odwiedzenia strony: luty 2013 r.
- [2] www.makaton.pl, data odwiedzenia strony: styczeń 2013 r.
- [3] Liang R. H., Ouhyoung M.: A Real-time Continuous Gesture Recognition System for Sign Language, In: Third IEEE International Conference on Automatic Face and Gesture Recognition, Proceedings, Nara, Japan, 1998.
- [4] Xu R., Zhou S., Li W. J.: MEMS Accelerometer Based Nonspecific-User Hand Gesture Recognition, Sensors Journal, IEEE, vol. 12, No. 5, May 2012.
- [5] Hussain S. M. A., Harun-ur Rashid A. B. M.: User Independent Hand Gesture Recognition by Accelerated DTW, IEEE/OSA/IAPR International Conference on Informatics, Electronics & Vision, Proceedings, Dhaka, Bangladesh 2012.
- [6] Akl A., Feng C., Valae S.: A Novel Accelerometer-Based Gesture Recognition System, Transactions on Signal Processing, IEEE, vol. 59, No. 12, December 2011.
- [7] Li H., Greenspan M.: Model-based segmentation and recognition of dynamic gestures in continuous video streams, Pattern Recognition 44, Elsevier, 2011.
- [8] Suk H., Sin B. K., Lee S. W.: Hand gesture recognition based on dynamic Bayesian network framework, Pattern Recognition 43, Elsevier, 2010.
- [9] Rautaray S., Agrawal A.: Design of Gesture Recognition System for dynamic User Interface, IEEE International Conference on Technology Enhanced Education (ICTEE), Proceedings, Amritapuri, India, 2012.
- [10] Paulraj M. P., Yaacob S., Azalan M. S. Z., Palaniappan R.: A Phoneme Based Sign Language Recognition System using 2D Moment Invariant Interleaving feature and Neural Network, IEEE Student Conference on Research and Development, Proceedings, Cyberjaya, Malaysia, 2011.
- [11] Lichtenauer J. F., Hendriks E. A., Reinders M. J.T.: Sign Language Recognition by Combining Statistical DTW and Independent Classification, Transactions on Pattern Analysis and Machine Intelligence, Vol. 30, No. 11, IEEE, 2008.
- [12] www.xbox.com/en-US/kinect, data odwiedzenia strony: styczeń 2013.
- [13] Wang Y., Yang C., Wu X., Xu S., Li H.: Kinect Based Dynamic Hand Gesture Recognition Algorithm System, 4th International Conference on Intelligent Human-Machine Systems and Cybernetics, Proceedings, Nanhang, China, 2012.
- [14] Li Y.: Hand Gesture Recognition Using Kinect. 3rd International Conference on Software Engineering and Service Science (ICSESS), Proceedings, IEEE, Beijing, China, 2012.
- [15] Szczepanowski B.: Podstawy języka migowego, WSIP, Warszawa 1994.
- [16] Liu J., Wang Z., Zhong L., Wickramasuriya J., Vasudevan V.: uWave: Accelerometer-based Personalized Gesture Recognition and Its Applications, Pervasive and Mobile Computing Journal, Vol. 5, Issue 6, Elsevier, Amsterdam, The Netherlands, 2009.
- [17] Müller M.: Information Retrieval for Music and Motion. Chapter 4: Dynamic Time Warping. Springer Verlag 2007.

otrzymano / received: 14.12.2012

przyjęto do druku / accepted: 01.02.2013

artykuł recenzowany / revised paper