

Application of voice communication for creation of flexible contour templates of the objects grasped by a collaborative robot

Zastosowanie komunikacji głosowej do generowania elastycznych wzorców obiektów chwytych przez robota współpracującego



A method for efficiently creating flexible contour templates (FECT) using multimodal interaction (speech, pointing, text) was developed, taking into account the quasi-natural character of the phrases spoken by the user. The use of multimodal commands provides significant time savings in creating a flexible FECT pattern.

W ostatnich latach roboty współpracujące odgrywają coraz istotniejszą rolę, także w przemyśle chemicznym. Ich praca jest często związana z chwytaniem obiektów identyfikowanych za pomocą systemu wizyjnego. Jedną z metod klasyfikacji obiektów na podstawie sygnatur ich zarysów zakłada tworzenie elastycznych wzorców konturów (FECT), umożliwiających automatyczne generowanie danych wejściowych do sieci neuronowej. Jej wadą jest pracochłonność opracowywania FECT przez użytkownika. Przedstawiono metodę efektywnego generowania FECT poprzez wykorzystanie interakcji multimodalnej (mowa, wskazywanie, tekst), uwzględniającej quasi-naturalny charakter fraz wypowiedzianych przez użytkownika.

Keywords: voice communication, collaborative robots

Słowa kluczowe: komunikacja głosowa, roboty współpracujące

W przemyśle chemicznym oraz farmaceutycznym często występuje konieczność pracy z materiałami niebezpiecznymi i szkodliwymi dla człowieka. Zastosowanie robotów, a zwłaszcza robotów współpracujących, przyczynia się znacząco do zmniejszenia związanego z tym ryzyka. Do najbardziej popularnych zadań realizowanych przez roboty zalicza się pakowanie i etykietowanie, jednak prowadzone są też badania dotyczące manipulacji i transportowania takich przedmiotów, jak próbki wirówkowe i próbki na szalkach Petriego¹⁾. Ulbrich i Aggarwal²⁾ przedstawiają perspektywę zastosowania robotów współpracujących do automatyzacji transferu próbek produktów do przyrządów laboratoryjnych. Burger i współpracownicy³⁾ opisują autonomiczną pracę robota mobilnego w laboratorium chemicznym. Istotną

rolę w funkcjonowaniu robotów współpracujących odgrywa zastosowanie systemów analizy obrazu. Dotyczy to również przemysłu chemicznego i farmaceutycznego. Hardner i Schneider⁴⁾ przedstawiają system wizyjny zapewniający dokładne pozycjonowanie systemu zrobotyzowanego służącego do manipulacji próbek umieszczonych na szalkach Petriego. Typowe przedmioty manipulowane przez roboty w przemyśle chemicznym mogą być identyfikowane na podstawie ich zarysów, np. poprzez porównywanie sygnatury tych zarysów z wzorcem^{5,6)}. Obiekty należące do jednej klasy mogą się jednak różnić nie tylko rozmiarami, ale i proporcjami (a zatem i sygnaturą). Dlatego opracowana została metoda bazująca na elastycznych edytowalnych wzorcach konturów FECT (*flexible editable contour templates*), umożliwiająca automatyczne generowanie danych wejściowych do sieci neuronowej realizującej klasyfikację rozpoznawanych obiektów⁷⁾. Prosty przykład opisu wzorca elastycznego dla zarysu w postaci trójkąta równoramiennego o różnych rozmiarach i proporcjach jest przedstawiony poniżej:

$s = 80..120$ \$ zakres wartości dla zmiennej s
 $\alpha = 30^\circ..90^\circ$ \$ zakres wartości dla zmiennej α
 $L = s / \sin(\alpha/2)$ \$ długość ramienia
 $line: a / length: s$ \$ podstawa trójkąta
 $go: left / angle: 90^\circ + \alpha$
 $line: b / length: L$ \$ pierwsze ramię trójkąta
 $go: left / angle: 360^\circ - \alpha$
 $line: c / length: L$ \$ drugie ramię trójkąta



Dr hab. inż. Adam ROGOWSKI (ORCID: 0000-0003-3991-7329) w roku 1983 ukończył studia na Wydziale Mechanicznym Technologicznym Politechniki Warszawskiej. W 1993 r. uzyskał stopień doktora, a w 2015 r. stopień doktora habilitowanego na tym samym wydziale, gdzie obecnie jest adiunktem w Instytucie Techniki Wytwarzania. Specjalność – robotyzacja i automatyzacja.

*** Adres do korespondencji:**

Zakład Automatyzacji i Obróbki Skrawaniem, Instytut Techniki Wytwarzania, Wydział Mechaniczny Technologiczny, Politechnika Warszawska, ul. Narbutta 86, pok. ST405, 02-524 Warszawa, tel.: (22) 849-03-73, e-mail: adam.rogowski@pw.edu.pl

Zdefiniowane na początku zmienne *s* i *alfa* oznaczają odpowiednio długość podstawy trójkąta oraz kąt między jego ramionami. Wymiary te są podawane jako zakresy możliwych wartości. Generowane sygnatury wzorców elastycznych, stanowiące wejście do sieci neuronowej, wykorzystują dowolne kombinacje tych wartości (zmieniających się z pewnym założonym krokiem), co pozwala na uwzględnienie różnych proporcji opisywanego kształtu konturu. Problemem jest jednak pracochłonność tworzenia FECT przez człowieka oraz możliwość popełnienia błędów. Dlatego opracowana została metoda interaktywnego generowania FECT, bazująca na multimodalnych komendach uwzględniających komunikację głosową, wskazywanie, a w uzasadnionych przypadkach ręczne wpisywanie wzorów. Metoda ta umożliwi generowanie opisu obiektów geometrycznych na podstawie relacji zachodzących między nimi (styczność, przecinanie). Ogólna koncepcja jest podobna do rozwiązań powstałych na gruncie głosowo aktywowanych systemów CAD⁸⁾, z dwiema zasadniczymi różnicami. Po pierwsze odległości i kąty nie muszą być podawane jako konkretne wartości, ale także jako zakresy możliwych wartości. Po drugie w przeciwieństwie do wspomnianych rozwiązań, które zakładają koncentrację na komunikacji głosowej i eliminację innych sposobów wprowadzania informacji⁸⁾, multimodalność jest czynnikiem zwiększenia efektywności działania. Zastosowano też mechanizm wyświetlania na bieżąco „podpowiedzi” odnośnie do kolejnych fraz komendy, dzięki czemu nie ma konieczności uczenia się komend.

Język komend głosowych

Za podstawę przyjęto format VCD^{9, 10)}, służący do opisu języka komend głosowych. Aby przedstawić istotę jego struktury, poniżej pokazano fragment opisu komend służących do tworzenia obiektów geometrycznych:

```
#com main
utwórz linię *line
#def line
równoległą do osi iks i przechodzącą przez *point
#def point
punkt bazowy
```

Komenda opisana w instrukcji *#com* odwołuje się do fraz składowych o nazwach zaznaczonych gwiazdką (*), opisanych w osobnych instrukcjach (*#def*) osobnych. Taka struktura zapewnia zwięzły zapis języka *quasi*-naturalnego, uwzględniający różnorodność możliwych sformułowań poszczególnych komend.

Osiągnięcie postawionego celu wymaga jednak, aby opis syntaktyki języka uwzględniał pozostałe metody interakcji. Dlatego rozszerzono format VCD o „frazy specjalne”: *<LINE>* (wskazanie linii), *<CIRCLE>* (wskazanie okręgu) i *<FORMULA>* (wpisanie wyrażenia arytmetycznego). Poniżej przedstawiono fragment opisu języka komend głosowych, obejmujący odwołania do takich fraz:

```
#def circle
```

```
okrąg o promieniu *value styczny do linii <line>
#def value
o wartości <formula>
z zakresu od <formula> do <formula>
```

Analiza semantyczna komend multimodalnych

Celem tej analizy jest utworzenie wyrażen arytmetycznych opisujących wartości (zakresy wartości) zmiennych używanych we wzorcach elastycznych FECT. Opis semantyki ma postać procedur przypisanych każdemu wariantowi poszczególnych fraz komendy. Przykładowo poszczególnym wariantom frazy:

```
#def circle_ tanto_line
styczny do linii <line>, o promieniu *value
styczny do okręgu <circle>, o promieniu *value
odpowiadają procedury oznaczone jako odpowiednio:
#prc circle_ tanto_line /1
#prc circle_ tanto_line /2
```

Każda z tych procedur jest zbiorem instrukcji podstawienia określonych wartości pod zmienne opisujące poszczególne obiekty geometryczne, np. dla okręgu (*circle*) nr 3 są to zmienne opisane symbolami *c3.x*, *c3.y*, *c3.r* (współrzędne środka i promień), a dla linii (*line*) nr 5 są to zmienne opisane symbolami *l5.x*, *l5.y*, *l5.a* (współrzędne punktu należącego do linii oraz kąt, jaki ta linia tworzy z osią *X*). Ponieważ numery są automatycznie przypisywane obiektom geometrycznym dopiero w wyniku realizacji komendy (poprzez wskazania), więc ich zapis w procedurach *#prc* nie może zawierać podawanych *a priori* numerów, tylko odwołuje się do rezultatów tych wskazań. Analogiczna uwaga dotyczy wzorów wprowadzanych ręcznie. Wprowadzone zostały więc następujące oznaczenia:

- <> – numer obiektu tworzonego w wyniku realizacji aktualnej komendy
- <p> – numer obiektu ostatnio wskazanego (p-pointed)
- <p-1>, <p-2> – numery obiektów wskazywanych wcześniej
- <f> – wyrażenie arytmetyczne (f-formuła) ostatnio ręcznie wprowadzone
- <f-1>, <f-2> – wyrażenia arytmetyczne wprowadzone wcześniej
- <n> – numer nowej (n-new) zmiennej pomocniczej
- <c> – numer aktualnie (c-current) utworzonej zmiennej pomocniczej
- <c-1>, <c-2> – numery poprzednio utworzonych zmiennych pomocniczych

Zastosowanie tych oznaczeń można zilustrować przykładem odwołującym się do komendy, w rezultacie której tworzony jest okrąg o promieniu o wartości z podanego zakresu, współśrodkowego z innym wcześniej zdefiniowanym okręgiem:

```
#com main
utwórz okrąg *circle
#def circle
współśrodkowy z okręgiem <circle> oraz *homocentric
```

```
#def homocentric
o promieniu *value.
#def value
o wartości <formula>
z zakresu od <formula> do <formula>
```

Opis semantyki fraz tworzących tę komendę miałyby następującą postać:

```
#pre value /1
v<n> := <f>
#pre value /2
v<n> := <f-1>..<f>
#pre homocentric
c<>.x := c<p>.x
c<>.y := c<p>.y
c<>.r := v<c>
```

Założmy, że podczas realizacji tej komendy multimodalnej istniałoby już pięć zdefiniowanych okręgów (aktualnie tworzony byłby szóstym), zaś wskazanym okręgiem odniesienia byłby okrąg numer 3, a zakres wartości promienia wpisany ręcznie wynosiłby 10–50. W wyniku analizy semantycznej powyższy zapis zostałby przetworzony na następujący ciąg instrukcji przypisania wartości do zmiennych we wzorcu elastycznym FECT:

```
v1 := 10..50
c6.x := c3.c
c6.y := c3.y
c6.r := v1
```

Część doświadczalna

Opracowana metoda została przetestowana w laboratorium, w którym robot Mitsubishi Movemaster-EX miał za zadanie podawanie użytkownikowi różnych obiektów na podstawie obrazu z kamery. W polu widzenia kamery były umieszczone różne obiekty, które miały być rozpoznawane na podstawie wzorców elastycznych FECT (rysunek).

W celu określenia efektywności opracowanej metody, utworzono szereg wzorców elastycznych FECT na dwa sposoby: poprzez manualny zapis FECT oraz poprzez multi-

modalne generowanie FECT. Wzorce elastyczne utworzono dla obiektów o różnych stopniach skomplikowania: szalka Petriego, próbówka wirówkowa, klucz ampulowy i klucz nastawny. Poniżej pokazano utworzony ręcznie opis wzorca obiektu „klucz nastawny”:

```
#cnt klucz nastawny
L:=100..200
S:=30..50
w:=0.4*S..0.6*S
D:=1.8*S..2.4*S
R:=D/2
G:=0.3*D..0.6*D
A:=15°..24°
amax:=arctan(sqrt(D*D-G*G)/D)
alfa=A..amax
r:=(D*cos(alfa)-G)/2/(1-cos(alfa))
beta:=arctan(S/sqrt(4*R*R-S*S))
gamma:=90°+beta
H:=R*cos(beta)
line: a / length: L
arc: b/bend:right; angle: alfa; radius: r
arc: c/bend:left; angle: alfa-A; radius: R
go:left / angle: 90°
line: m / length: w
go:right / angle: 90°
arc: n / bend:left; angle: 90°-beta; radius: R
go:left / angle: gamma
line: d / length: H
go:right / angle: 90°
line: e / length: S-w
go:right / angle: 90°
line: f / length: H
go:left / angle: gamma
arc: g / bend:left; angle: alfa+90°+A-beta; radius: R
arc: h / bend:right; angle: alfa; radius: r
line: i / length: L
arc: j / bend:left; angle: 180°; radius: G/2
```

Czas potrzebny do tworzenia wzorców przedstawiono w tabeli (oczywiście należy uwzględnić fakt, że te wartości mogą się nawet znacząco różnić w zależności od użytkownika, jego zdolności oraz doświadczenia pracy z systemem, tym niemniej różnice w efektywności obu metod są wyraźne).



Figure. Robotic laboratory workstation (left), and vision system interface (right)

Rysunek. Zrobotyzowane stanowisko laboratoryjne (z lewej) i interfejs systemu wizyjnego (z prawej)

Podsumowanie

Jak wynika z przedstawionych badań, zwłaszcza w przypadku wzorców elastycznych opisujących bardziej skomplikowane kontury, których utworzenie wymaga wyprowadzenia złożonych wzorców arytmetycznych, zastosowanie

Table. Time needed to create the FECT depending on the method

Tabela. Czasy tworzenia wzorca elastycznego FECT w zależności od metody

Metoda tworzenia FECT	Obiekt rozpoznawany	Czas tworzenia FECT, min			
		Szalka Petriego	Probówka wirówkowa	Klucz ampulowy	Klucz nastawny
Ręczne tworzenie opisu FECT		2	10	12	38
Multimodalne generowanie FECT		< 1	3	4	8

komend multimodalnych daje znaczące zyski czasowe. Choć nie przeprowadzono badań eksperymentalnych dotyczących częstotliwości popełniania błędów przy ręcznym tworzeniu opisu FECT, to oczywiste jest, że automatyczne generowanie FECT w zdecydowany sposób zmniejsza prawdopodobieństwo wystąpienia błędów w utworzonych wzorcach elastycznych, zwłaszcza w przypadku obiektów o skomplikowanych konturach.

System bazujący na opracowanej metodzie jest obecnie ograniczony do generowania wzorców elastycznych dla obiektów, których kontury składają się wyłącznie z odcinków linii prostej i łuków okręgów. Jak jednak wynika z treści artykułu, bez większych trudności można go rozbudować zarówno o inne obiekty geometryczne (np. krzywe Beziera), jak i wprowadzić nowe relacje geometryczne, na bazie których opisuje się tworzone obiekty. Łatwość ta wynika przede wszystkim z faktu, że w celu tej rozbudowy wystarczy zmodyfikować opis języka komend głosowych

oraz opis semantyki. Tego też dotyczą plany na najbliższą przyszłość związane z przedstawioną tematyką.

Otrzymano: 08-10-2024

Zrecenzowano: 25-11-2024

Zaakceptowano: 28-11-2024

Opublikowano: 20-12-2024

LITERATURA

- [1] R. Mathew, R. McGee, K. Roche, S. Warreth, N. Papakostas, *Appl. Sci.* 2022, **12**, 10895.
- [2] M. Ulbrich, V. Aggarwal, *J. Bus. Chem.* 2019, **16**, nr 2, 76.
- [3] B. Burger, P. Maffettone, V. Gusev, C. Aitchison, Y. Bai, X. Wang, X. Li, B. Alston, B. Li, R. Clowes, N. Rankin, B. Harris, R. Sprick, A. Cooper, *Nature* 2020, **583**, 237.
- [4] M. Hardner, D. Schneider, *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.* 2019, XLII-2/W18, 67.
- [5] H. Elsalamony, *Measurement* 2017, **104**, 50.
- [6] A. Barman, P. Dutta, *Pattern Recognit. Lett.* 2021, **145**, 254.
- [7] A. Rogowski, P. Skrobek, *Sensors* 2020, **20**, 177.
- [8] S. Xue, X. Kou, S. Tan, *Comput. Aided Des. Appl.* 2009, **6**, 125.
- [9] A. Rogowski, K. Bieliszczuk, J. Rapcewicz, *Sensors* 2020, **20**, 7287.
- [10] A. Rogowski, *Sensors* 2022, **22**, 9520.



**Stowarzyszenie Inżynierów
i Techników Przemysłu Chemicznego**
Oddział Gliwice zaprasza



XXXI Ogólnopolska Konferencja Naukowo-Techniczna
ANTYKOROZJA
SYSTEMY - MATERIAŁY - POWŁOKI

31. Ogólnopolska Konferencja Naukowo-Techniczna

ANTYKOROZJA

Systemy - Materiały - Powłoki

19–21 marca 2025 r. Ustroń – hotel JAWOR

- Materiały, powłoki, systemy i technologie,
- Elektrochemiczna ochrona przed korozją,
- Kierunki badań antykorozyjnych,
- Problemy korozyjne w budownictwie,
- Korozja mikrobiologiczna,
- Zabezpieczenie maszyn, urządzeń i środków transportu.

Informacje i kontakt:
tel. 664 421 351, 664 421 349
www.gliwice.sitpchem.org.pl