

Sterowanie głosowe w systemach obróbkowych

dr hab. inż. Adam Rogowski

W artykule omówiono możliwości stosowania sterowania głosowego przy obsłudze gniazd obróbkowych złożonych z obrabiarek CNC. Zagadnienie to ukazano na tle rezultatów dotychczasowych badań dotyczących automatycznego rozpoznawania mowy, a w szczególności jego zastosowania w szeroko rozumianym wytwarzaniu. Omówiony został ogólny schemat algorytmu rozpoznawania mowy oraz specyfika tego algorytmu zależna od stopnia złożoności stosowanych komend. Na bazie tego podane zostały możliwe warianty typów komend służących do obsługi zautomatyzowanych systemów obróbkowych i warianty ich przetwarzania w zależności od zadań, które mają być realizowane w wyniku tych komend. Jako przykład funkcjonującego rozwiązania przedstawiono opracowany na Politechnice Warszawskiej system sterowania głosowego szkoleniowym zrobotyzowanym gniazdem obróbkowym EMCO.

Wprowadzenie

Już od kilku dziesięcioleci komunikacja głosowa człowiek – maszyna jest przedmiotem badań naukowych, zaś w ostatnich latach jesteśmy świadkami bardzo intensywnego rozwoju technik automatycznego rozpoznawania mowy. Zważywszy, że mowa jest najdogodniejszym dla człowieka sposobem komunikowania otoczeniu swoich zamiarów, poleceń, czy zasięgnięcia informacji, nie może dziwić fakt coraz powszechniejszego korzystania z tego środka komunikacji nie tylko w stosunku do innych ludzi, ale również w stosunku do różnego rodzaju maszyn i urządzeń. W pierwszym rzędzie chodzi tu o komputery i inne urządzenia służące do cyfrowego przetwarzania informacji. Coraz częściej wybieramy głosowo numery telefonów (zwłaszcza korzystając z zestawów głośnomówiących podczas prowadzenia samochodu), korzystamy z komputerowych systemów dostępu do informacji itd. Zdobyte w tych dziedzinach doświadczenia są od pewnego czasu prze-

noszone również w dziedzinę sterowania pracą maszyn realizujących konkretne czynności fizyczne. W szczególności dotyczy to sterowania robotami za pomocą komend wypowiedzianych głosem. Podstawowy rodzaj robotów, w stosunku do których zaczęto wykorzystywać tę technikę, to mobilne roboty usługowe [1,4,7,18,23]. Równolegle prowadzone są badania dotyczące zastosowania sterowania głosowego także w stosunku do robotów przemysłowych. Dziedzina ta jest jednakże dość specyficzna z punktu widzenia automatycznego rozpoznawania mowy. Niniejszy artykuł omawia problemy i rozwiązania dotyczące komunikacji głosowej człowiek-maszyna w warunkach przemysłowych, ze szczególnym uwzględnieniem zrobotyzowanych gniazd obróbkowych.

Komunikacja głosowa w systemach wytwarzania

Jak wspomniano, realizacja komunikacji głosowej człowiek-maszyna w warunkach przemysłowych ma swoją specyfikę.

Wynika ona z jednej strony z wysokich wymagań odnośnie niezawodności rozpoznawania komend głosowych, wypowiedzianych na ogół w niesprzyjających warunkach (hałas), a z drugiej z treści tych komend, odnoszącej się do zadań, które mają być realizowane przez maszyny. Może tu wystąpić duże zróżnicowanie: inna będzie na przykład syntaktyka i stopień złożoności komend dotyczących prostych urządzeń dwupołożeniowych jak uchwyty obróbkowe („otwórz” – „zamknij”), a zupełnie inna w przypadku głosowego uczenia robotów, obejmującego podawanie współrzędnych i innych parametrów liczbowych. Należy też zwrócić uwagę na właściwy wybór zadań, które mają być realizowane przez maszyny w wyniku podawanych komend głosowych. Przeprowadzone badania [10] wykazały bowiem, że w wielu sytuacjach bardziej efektywne jest wpływanie na działanie maszyn nie za pomocą głosu lecz np. za pomocą gestów rozpoznawanych przez systemy wizyjne, bądź za pomocą ekranów dotykowych. W szczególności



dotyczy to poleceń powiązanych ze wskazywaniem obiektów. W niektórych przypadkach celowe jest współdziałanie systemu rozpoznawania mowy z innymi metodami wprowadzania danych, dzięki czemu same komendy głosowe mogą być stosunkowo proste, gdyż znaczna część informacji będzie pochodziła z kontekstu.

Jeśli chodzi o zakres potencjalnych zastosowań automatycznego rozpoznawania mowy w wytwarzaniu, to jest on bardzo szeroki. Ze względu na to, że istniejące opracowania i rozwiązania dotyczące sterowania głosowego były dotąd ukierunkowane głównie na roboty (zwłaszcza usługowe), najbardziej podatne na wprowadzanie tej techniki w wytwarzaniu są oczywiście te obszary, w których wykorzystuje się roboty przemysłowe np. spawanie i montaż. W mniejszym stopniu dotyczyłoby to obróbki skrawaniem, jednak zwrócić uwagę, że również w tej dziedzinie roboty mogą być szeroko stosowane do realizacji czynności pomocniczych (manipulacja przedmiotami obrabianymi), a w niektórych przypadkach także do obróbki jako takiej [6].

Oczywiście, sterowanie głosowe nie musi dotyczyć tylko takich zaawansowanych pod względem sterowania maszyn, jak roboty przemysłowe. Można sterować głosowo także urządzeniami prostymi, pod warunkiem jednak, że zostaną one podłączone do komputera (albo i do smartfona) realizującego automatyczne rozpoznawanie mowy. Przykładowo, już w latach osiemdziesiątych ubiegłego wieku opatentowane zostało rozwiązanie dotyczące komunikacji głosowej pomiędzy spawaczem i spawarką [20]. Rozwiązanie to polegało na przetwarzaniu przez komputer głosowych komend spawacza i wysyłaniu stosownych sygnałów do spawarki. W ten sposób użytkownik mógł – nie przerywając wykonywanej pracy – wpływać na sterowanie zasilaniem przez spawarkę, jak również sterować prędkością podawania drutu elektrody.

Wykorzystanie sterowania głosowego w spawalnictwie może być oczywiście dużo szersze. W szczególności dotyczy to komunikacji głosowej z robotami spawalniczymi [12], która może być wykorzystana w różnych sytuacjach np.:

- programowania robota metodą uczenia (całkowite bądź częściowe zastąpienie programatora),

- reagowania głosem w sytuacjach awaryjnych (ekwiwalent stopu awaryjnego),
- zadaniowo zorientowanego sterowania złożonymi funkcjami realizowanymi przez robota (realizacja wypowiedzianej komendy głosowej polega na ładowaniu do układu sterowania robota, a następnie uruchomieniu konkretnego programu).

Potencjalnym obszarem zastosowania głosowej komunikacji człowiek-maszyna w wytwarzaniu są wszystkie te dziedziny, w których można realizować koncepcję bezpośredniej współpracy ludzi i robotów we wspólnej przestrzeni roboczej [9]. Dzięki tej komunikacji można koordynować – poprzez dialog – działania zespołów ludzi i robotów na przykład podczas prowadzenia prac spawalniczych [3]. Koncepcja bezpośredniej współpracy człowiek-robot może być jednak realizowana także podczas różnego rodzaju prac montażowych [22,24].

Jeśli chodzi o zautomatyzowane (zrobotyzowane) systemy obróbkowe, to zastosowanie komunikacji głosowej do ich obsługi może mieć miejsce na wielu różnych poziomach: od formułowania zadaniowo-zorientowanych poleceń dotyczących strategii realizowania procesów technologicznych w gnieździe obróbkowym („proszę rozpocząć obróbkę tulei typu B”), poprzez mające sens tylko w warunkach produkcji jednostkowej wydawanie poleceń obróbki przedmiotu na określony wymiar („toczyć wałek na średnicę trzydziestu milimetrów”), aż do głosowego inicjowania elementarnych czynności realizowanych przez urządzenia proste („zamknij imadło”). Ten ostatni wariant mógłby zachodzić w przypadku gniazd obróbkowych o zróżnicowanym stopniu automatyzacji, w których obok realizacji zadań przez maszyny sterowane numerycznie miałyby miejsce także wykonywanie niektórych czynności ręcznie przez operatora. Operator, wykonując pewne czynności i mając zajęte ręce, mógłby jednocześnie wydawać polecenia „współpracownikom” będącym robotami, obrabiarkami lub uchwytami obróbkowymi [13].

W każdym z wymienionych przypadków przetwarzanie komend głosowych może zachodzić w nieco odmienny sposób. Czasem będzie konieczne uwzględnienie kontekstu na podstawie informacji pochodzących z innych źródeł niż

sama komenda. Również złożoność algorytmu przetwarzania komendy może być bardzo różna: czasem wystarczy proste przypisanie konkretnego ciągu wyrazów do konkretnej instrukcji sterującej maszyną, a czasem może być konieczne przeprowadzenie złożonej analizy syntaktycznej, semantycznej i pragmatycznej wypowiedzianej przez operatora komendy. Zanim więc zostaną omówione różne warianty sterowania głosowego w systemach obróbkowych, trzeba się przyjrzeć ogólnemu algorytmowi przetwarzania komendy i poszczególnym jego etapom.

Algorytm automatycznego rozpoznawania mowy

Niezależnie od wariantu sterowania głosowego, ostatecznym celem automatycznego rozpoznawania mowy w komunikacji człowiek-maszyna jest przetworzenie sygnału mowy na instrukcję, bądź ciąg instrukcji sterujących maszyną. Ogólny schemat tego przetwarzania został przedstawiony na rys. 1 [11]. W wyniku analizy akustycznej i fonetycznej sygnał mowy jest przetwarzany na ciąg fonemów. W wyniku analizy leksykalnej fonemy te łączone są w pojedyncze wyrazy na podstawie słownika opracowanego dla konkretnej aplikacji rozpoznawania mowy. Kolejnym etapem jest analiza syntaktyczna, będąca procesem analizy tekstu, mającym na celu ustalenie jego struktury gramatycznej i zgodności z gramatyką języka. Rezultatem analizy syntaktycznej jest najczęściej utworzenie tzw. drzewa składniowego przedstawiającego strukturę analizowanego zdania. Po niej następuje analiza semantyczna, której istotą jest proces przypisywania znaczenia strukturom syntaktycznym analizowanego tekstu w sposób niezależny od języka, w którym był sformułowany ten tekst. Aby można było na tej podstawie wygenerować stosowne instrukcje sterujące maszyną, konieczne jest jeszcze przeprowadzenie analizy pragmatycznej, która uwzględnia znajomość zasad funkcjonowania sterowanego obiektu (robota przemysłowego, maszyny technologicznej) oraz okoliczności, w jakich to działanie się odbywa.

Oczywiście algorytm przetwarzania komend głosowych został tu przedstawiony w sposób ogólny i uproszczony. Nie wszystkie przedstawione etapy muszą w nim zawsze występować.

Ponadto w ogólnym przypadku przetwarzanie sygnału mowy na instrukcje sterujące dla maszyny nie przebiega według schematu o liniowym charakterze. Realizacja poszczególnych jego etapów może być uzależniona od etapów późniejszych, co oznacza konieczność zastosowania pętli (sprzężeń zwrotnych). Przykładowo analiza fonetyczna może być obciążona błędami. Wynika to z wielu przyczyn m. in. z niedoskonałości współczesnych algorytmów rozpoznawania mowy, z niewyraźnego mówienia, z hałasu. Wpływa to oczywiście na rezultaty analizy leksykalnej. Zwiększenie wiarygodności rozpoznawania ciągów wyrazów może być jednak osiągnięte poprzez zastosowanie sprzężenia zwrotnego od analizy syntaktycznej lub semantycznej. Dzieje się tak dzięki temu, że podczas analizy fonetycznej i leksykalnej brane są pod uwagę tylko takie fonemy i wyrazy, które rzeczywiście mogą wystąpić w danym miejscu prawidłowej pod względem syntaktycznym

i mającej sens komendy głosowej. Hipotezy dotyczące wszelkich innych fonemów i wyrazów są automatycznie odrzucane. Zwróćmy uwagę, że analogiczny mechanizm jest realizowany przez ludzki słuch: jesteśmy w stanie prawidłowo zinterpretować zdania wypowiedziane przez inną osobę nawet wtedy, gdy mówi ona niewyraźnie lub nawet „połyka” niektóre sylaby, gdyż w wielu przypadkach na podstawie dotychczas wypowiedzianych słów „domyślamy się”, jakie powinny być kolejne słowa w wypowiedzianym zdaniu.

Jak wspomniano wcześniej, nie w każdym przypadku muszą być realizowane wszystkie wymienione etapy przetwarzania komendy głosowej. Jeżeli sterowane głosem urządzenie może realizować tylko niewielką liczbę prostych czynności, które mogą być opisane za pomocą niewielkiego zestawu prostych komend, to analiza wypowiedzianej przez operatora komendy sprowadza się do porównania sygnału mowy ze wzorcami możliwych komend. Jest to tzw. metoda dopasowania wzorca. W przypadku jej zastosowania analiza syntaktyczna, semantyczna i pragmatyczna staje się całkowicie zbędna.

Zastosowanie rozpoznawania mowy w zautomatyzowanych systemach obróbkowych

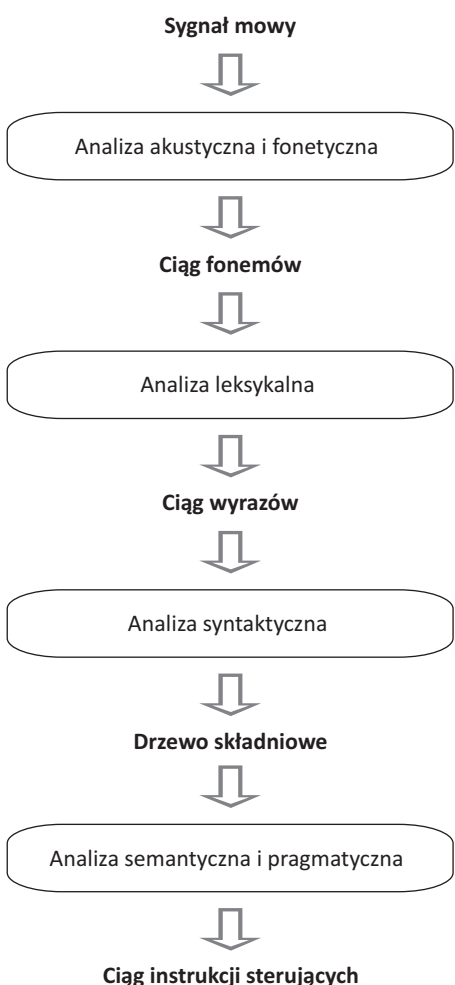
Automatyczne rozpoznawanie mowy może znaleźć bardzo szerokie zastosowanie w zautomatyzowanej obróbce. Warto zaznaczyć, że nie musi to dotyczyć tylko sterowania głosowego maszynami, urządzeniami, czy ich zespołami, ale także technologicznego przygotowania produkcji. Przykładowo, już w latach dziewięćdziesiątych ubiegłego wieku prowadzono badania nad systemem ekspertowym, służącym

do doboru stali szybkoobrotowej dla narzędzi skrawających [2]. Użytkownik miał komunikować się z systemem poprzez zdania w języku quasi-naturalnym wymawiane do mikrofonu.

W niniejszym opracowaniu skupimy się jednak na szeroko rozumianym sterowaniu głosowym, czyli na oddziaływaniu za pomocą mowy na funkcjonowanie maszyn i urządzeń.

Jak wspomniano, schemat przetwarzania komend głosowych może być mniej lub bardziej złożony, może też pomijać niektóre etapy przedstawione na rys. 1. W pierwszym rzędzie zależy to od charakteru procedury egzekucyjnej uruchamianej w rezultacie wypowiedzenia komendy głosowej. Przez procedurę egzekucyjną rozumiemy tu program uruchamiany bądź w układzie sterowania maszyny, bądź na komputerze bezpośrednio sterującym daną maszyną (urządzeniem). W wyniku realizacji procedury egzekucyjnej maszyna (urządzenie) wykonuje jakieś działanie. Procedura egzekucyjna może zawierać pojedynczą instrukcję sterującą urządzeniem, ale może też stanowić złożony program sterujący robotem lub obrabiarką (np. program NC w standardzie ISO), czy też program sterujący całym gniazdem obróbkowym. Procedura ta jest identyfikowana za pomocą nazwy (lub kodu liczbowego), a do tego mogą też ją opisywać pewne parametry liczbowe. W wyniku przetwarzania komendy głosowej, na podstawie treści komendy generowana jest nazwa (kod) procedury egzekucyjnej odpowiadającej tej treści, a dodatkowo mogą być też obliczane wartości parametrów liczbowych procedury egzekucyjnej.

Przykładowo na podstawie komendy „Załaduj wałek numer dwa na tokarkę” (rys. 2) wygenerowana jest nazwa



Rys. 1. Schemat przetwarzania komend głosowych

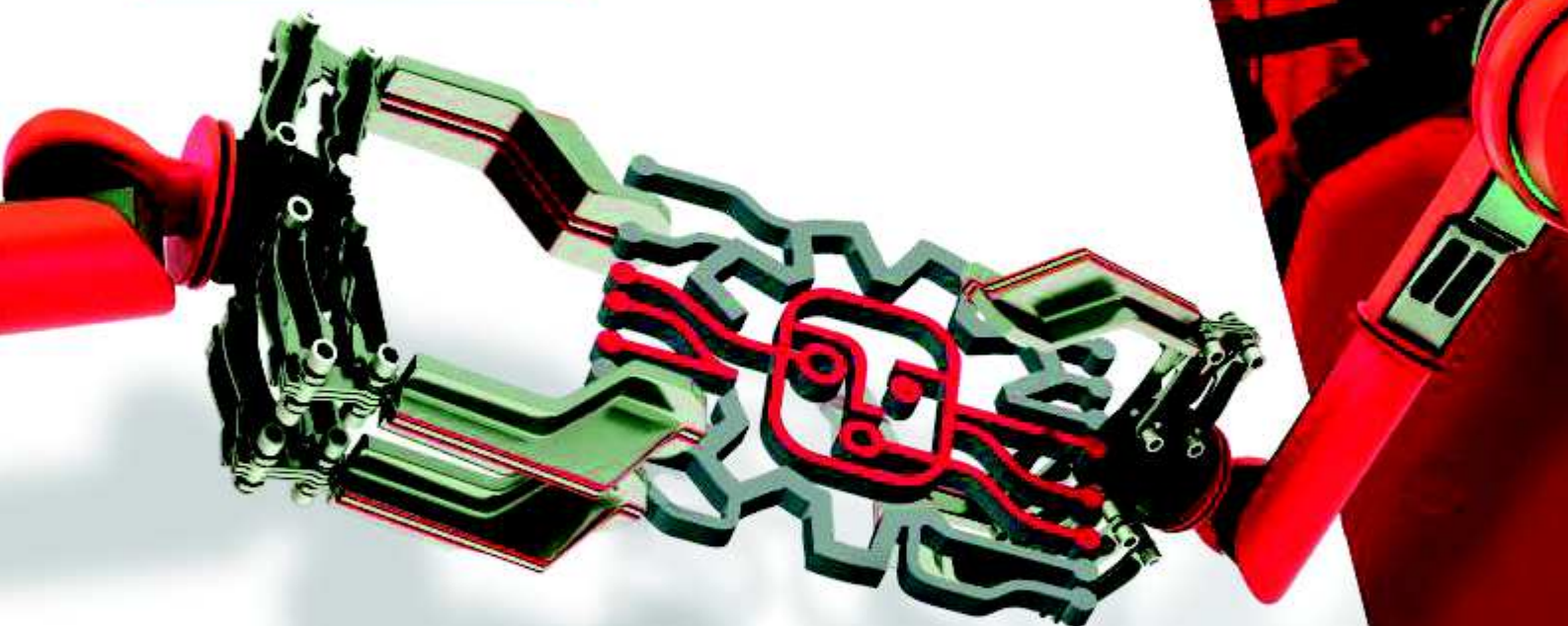


Rys. 2. Przykład przetworzenia komendy głosowej na procedurę egzekucyjną dla robota



 **ITM** INNOWACJE
TECHNOLOGIE
POLSKA MASZyny

5-8 CZERWCA 2018
POZNAŃ



NA DRODZE DO
PRZEMYSŁU 4.0

W TYM SAMYM CZASIE:



MODERNLOG
Targi Logistyki,
Magazynowania
i Transportu



SUBCONTRACTING
TARGI KODERACJI PRZEMYSŁOWEJ



3D SOLUTIONS

(„LOAD”) procedury odpowiedzialnej za realizację programu sterującego robotem ładującym przedmioty na obrabiarki, a oprócz tego obliczone są parametry liczbowe uściślające działanie tego programu: numer pozycji na palecie, z której ma być pobrany przedmiot (parametr liczbowy „2”) oraz numer obrabiarki, na której przedmiot ma być umieszczony (parametr „LATHE”, który odpowiada pewnej liczbie podstawianej pod zmienną w programie sterującym robotem).

W zależności od typu procedur egzekucyjnych stosowanych w danej maszynie (urządzeniu), a także stopnia złożoności komend wykorzystywanych do sterowania głosowego tą maszyną, można wyróżnić – w pewnym uproszczeniu – następujące trzy warianty systemów sterowania głosowego:

1. Procedura egzekucyjna bezparametrowa:

– Typ komendy: ściśle określony ciąg wyrazów np. „Zamknij uchwyt”.

– Zapis procedury egzekucyjnej: pojedyncza nazwa lub kod liczbowy np. CLOSE.

– Algorytm przetwarzania komendy: dopasowanie wzorca.

2. Procedura egzekucyjna z jednym parametrem liczbowym mogącym przyjmować skończoną (i niewielką) liczbę wartości:

– Typ komendy: określony ciąg wyrazów z jednym elementem wymiennym np. „Pobierz przedmiot z pozycji numer <numer>” gdzie <numer> oznacza np. liczbę z zakresu 1–5.

– Zapis procedury egzekucyjnej: pojedyncza nazwa z jednym parametrem liczbowym np. GRASP (4).

– Algorytm przetwarzania komendy: dopasowanie do wielu wzorców, mogące uwzględniać elementy analizy syntaktycznej.

3. Procedura egzekucyjna z wieloma parametrami liczbowymi lub z jednym parametrem liczbowym mogącym przyjmować wiele różnych wartości:

– Typ komendy: ciąg wyrazów będący jednym z bardzo licznych możliwych wariantów (naturalny bądź quasi-naturalny język komend głosowych) np.:

„Toczyć wałek na średnicę <liczba> milimetrów i długość <liczba>” gdzie <liczba> oznacza dowolną dodatnią liczbę całkowitą bądź ułamek dziesiętny.

– Zapis procedury egzekucyjnej: pojedyncza nazwa z jednym bądź wieloma parametrami liczbowymi np. TURN (30.5, 100).

– Algorytm przetwarzania komendy: pełne przetwarzanie, obejmujące analizę syntaktyczną, semantyczną i pragmatyczną.

W wariantcie 1 (bezparametrowa procedura egzekucyjna) komenda jest ściśle określonym ciągiem wyrazów, co jednak nie musi oznaczać zupełnego braku wariantowości wymawianych przez operatora komend. Można bowiem zdefiniować kilka komend, do których będzie przypisana ta sama procedura egzekucyjna np. „Zamknij imadło” i „Proszę zamknąć imadło”. Zapewnia to komfort użytkownikowi, który nie musi się uczyć na pamięć sztywnego zestawu dostępnych komend.

Wariant ten stosowano powszechnie we wczesnych doświadczeniach dotyczących sterowania głosowego maszynami technologicznymi [8]. Jego podstawową wadą jest brak możliwości wykorzystania do sterowania złożonymi czynnościami realizowanymi przez maszynę, a zwłaszcza takimi, które są opisywane za pomocą parametrów liczbowych. Podstawową natomiast jego zaletą jest ograniczona liczba porównywanych przez system wzorców, co znacznie zmniejsza prawdopodobieństwo błędnego zrozumienia komendy przez system. Ponadto przetwarzanie komendy głosowej jest względnie proste.

Jeśli chodzi o zastosowanie tego wariantu w systemach obróbkowych, to głównie dotyczyłby on komunikacji głosowej z prostymi urządzeniami (jak np. uchwyty obróbkowe) w gniazdach obróbkowych o zróżnicowanym stopniu automatyzacji (procedura egzekucyjna obejmowałaby wtedy najczęściej pojedynczą instrukcję sterującą), aczkolwiek w pewnych sytuacjach mógłby on też znaleźć zastosowanie przy wydawaniu zadaniowo-zorientowanych poleceń dotyczących strategii realizowania procesów technologicznych w gnieździe obróbkowym (procedura egzeku-

cyjna mogłaby wówczas realizować ładowanie i uruchamianie nawet bardzo złożonych programów w układach sterowania maszyn).

Wariant 2 (procedura egzekucyjna z jednym parametrem liczbowym mogącym przyjmować skończoną liczbę wartości) jest pewnym rozszerzeniem wariantu pierwszego, umożliwiającym wprowadzenie w ograniczonym zakresie zmiennych elementów do komendy głosowej [21]. Takie same zadania można by wprowadzić realizować przy zastosowaniu wariantu pierwszego, jednak wymagałoby to żmudnego definiowania ogromnej liczby możliwych wzorców komend np. „Pobierz przedmiot z pozycji numer jeden”, „Pobierz przedmiot z pozycji numer dwa” itd. Wariant drugi umożliwia zdefiniowanie tej wielkiej liczby wzorców w stosunkowo zwięzły sposób, tworząc tzw. reguły gramatyki osobno dla wymienionych elementów komendy. Konsekwencją jest występowanie elementów analizy syntaktycznej przy przetwarzaniu takich komend.

Zastosowanie tego wariantu do sterowania głosowego w zautomatyzowanych systemach obróbkowych może dotyczyć głównie robotów przemysłowych realizujących podawanie i odbieranie przedmiotów z obrabiarek.

Wariant 3 (procedura egzekucyjna z wieloma parametrami liczbowymi lub z jednym parametrem liczbowym mogącym przyjmować wiele różnych wartości) przewiduje generowanie quasi-naturalnego języka komend głosowych [5], dzięki czemu za pomocą tych komend można zdefiniować nawet bardzo złożone zadania, w szczególności zadania opisane parametrami liczbowymi mogącymi przyjmować różnorodne wartości. W przeciwieństwie do wariantów 1 i 2, przetworzenie tych komend na procedury egzekucyjne wymaga pełnej analizy syntaktycznej, jak również semantycznej. Również częściej, niż w pozostałych wariantach zachodzić będzie konieczność uwzględnienia w analizie pragmatycznej znajomości kontekstu na podstawie informacji uzyskiwanych np. ze współpracującego systemu wizyjnego, co może wpłynąć na uproszczenie komend głosowych. Przykładowo, jeśli ma być zrealizowane przelotowe toczenie wałka na określoną średnicę, to na pod-

stawie danych z systemu analizy obrazu można automatycznie określić początkową i końcową współrzędną Z noża tokarskiego, dzięki czemu komenda głosowa będzie zawierała tylko informację o średnicy toczenia, a zatem będzie stosunkowo prosta np. „Toczyć wałek na średnicę trzydzieści i pół milimetra”.

Należy zaznaczyć, że przytoczony tu przykład głosowego polecenia obróbki przedmiotu na dany wymiar byłby raczej wyjątkiem w stosowaniu komunikacji głosowej w systemach obróbkowych. Tego typu sytuacja mogłaby zajść raczej tylko w warunkach produkcji jednostkowej. W przypadku produkcji seryjnej generowanie programów NC na podstawie komend głosowych nie byłoby ani opłacalne ani sensowne. Raczej stosowano by wówczas komunikację głosową do wydawania wspomnianych wcześniej zadaniowo-zorientowanych poleceń dotyczących strategii realizowania procesów technologicznych w gnieździe obróbkowym. Ich rezultatem byłoby ładowanie do układów sterowania obrabiarek programów NC utworzonych uprzednio w tradycyjny sposób. Wydaje się natomiast, że wariant 3 znalazłby zastosowanie głównie do sterowania pracą robotów obsługujących obrabiarki.

Jak wspomniano wcześniej, stosowanie wariantu 3 wymaga generowania quasi-naturalnego języka komend głosowych, zawierającego informacje zarówno o syntaktyce, jak i semantyce. Ponieważ możliwa jest wielka różnorodność tych komend, więc należy zapewnić maksymalną zwięzłość opisu języka. Temu celowi służy opracowany format VCD [11,14], wykorzystujący gramatykę formalną, a konkretnie tzw. gramatykę bezkontekstową (context-free grammar – CFG). Ważną cechą formatu VCD jest fakt, że opis języka komend w tym formacie może być wykorzystany na wszystkich etapach przetwarzania komendy: zarówno podczas analizy fonetycznej, jak również syntaktycznej i semantycznej [17]. Tych zalet z reguły nie posiadają rozwiązania bazujące nie na gramatyce formalnej, lecz na przetwarzaniu języka sensu stricte naturalnego. Dlatego też z reguły ograniczają się one do analizy komend wprowadzanych z pomocą klawiatury, a nie głosem [19].

Omówione tu zasady stosowania komunikacji głosowej w wytwarzaniu zostały praktycznie zweryfikowane laboratoryjnie poprzez opracowanie i uruchomienie systemu sterowania głosowego szkoleniowym zrobotyzowanym gniazdem obróbkowym EMCO na Politechnice Warszawskiej (rys. 3).

Gniazdo to składa się z trzech maszyn sterowanych numerycznie:

- robota Movemaster-EX (1),
- tokarki C5CNC (2),
- frezarki F1CNC (3);

oraz z trzech automatycznie sterowanych urządzeń pomocniczych:

- sań pneumatycznych na których umieszczony jest robot (4),
- imadła maszynowego znajdującego się na frezarce,
- konika znajdującego się na tokarce.

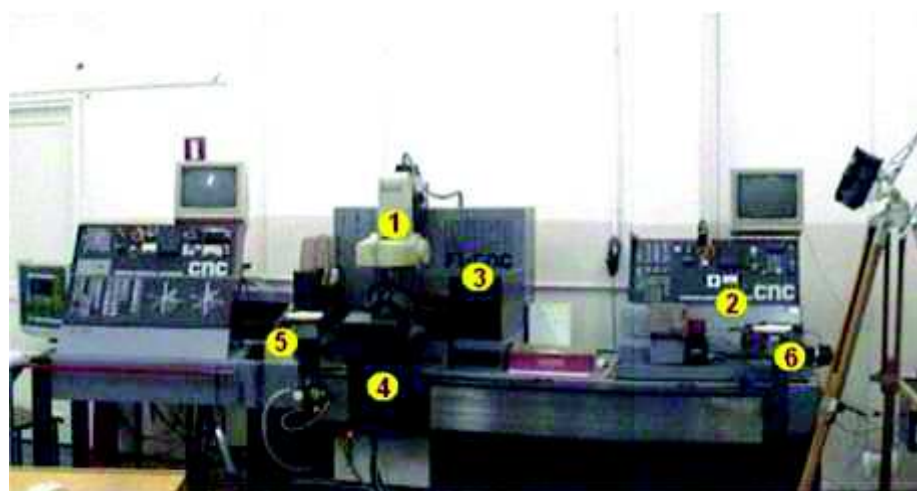
Ponadto przy frezarce i tokarce znajdują się palety (5 i 6) na przedmioty

z systemu rozpoznawania obrazu), jak również wydawanie prostych poleceń urządzeniom pomocniczym.

Na Politechnice Warszawskiej prowadzono również prace nad zdalnym sterowaniem zrobotyzowanym gniazdem obróbkowym (przez Internet) [15,16], co w połączeniu z badaniami nad komunikacją głosową doprowadziło do powstania systemu zdalnego sterowania głosowego [17]. Umożliwia on wydawanie poleceń głosowych maszynom i urządzeniom tego gniazda praktycznie z dowolnego miejsca na kuli ziemskiej.

Podsumowanie

Doświadczenia uzyskane w wyniku badań laboratoryjnych dotyczących komunikacji głosowej w systemach wytwarzania wykazały, że istnieje wiele obszarów potencjalnych jej zastosowań, w tym również w zrobotyzowanych systemach obróbkowych. Komunikacja głosowa może być tu wykorzystana



Rys. 3. Zrobotyzowane gniazdo obróbkowe EMCO

obrabiane typu wałek. Na każdej z palet może się jednocześnie znajdować do pięciu przedmiotów.

Opracowany system sterowania głosowego umożliwia ładowanie programów NC do układów sterowania obu obrabiarek (i ich uruchamianie), wydawanie zadaniowo-zorientowanych poleceń dla robota (dotyczących ładowania i rozładowywania obrabiarek, jak też pobierania i odkładania przedmiotów na palety przy wykorzystaniu informacji

zarówno przy wydawaniu zadaniowo-zorientowanych poleceń dotyczących strategii realizowania procesów technologicznych w gnieździe obróbkowym, w głosowym sterowaniu robotami obsługującymi obrabiarki, w niektórych sytuacjach do generowania programów NC dla obrabiarek w wyniku głosowego wydawania poleceń obróbki przedmiotu na dany wymiar, a wreszcie do sterowania prostymi urządzeniami w gniazdach obróbkowych o zróżnicowanym stopniu automatyzacji.

W zależności od rodzaju maszyny bądź urządzenia, z którym operator będzie się komunikować za pomocą mowy, a także w zależności od rodzaju zadań, które ta maszyna w wyniku podanych komend głosowych ma wykonywać, mogą być realizowane różne warianty sterowania głosowego. Przejawia się to w wykorzystywaniu różnych języków komend, znacząco się od siebie różniących stopniem złożoności, jak również w stosowaniu różniących się od siebie algorytmów przetwarzania wypowiedzianych komend. Jeśli chodzi o złożoność języków komend wykorzystywanych do komunikacji głosowej w zautomatyzowanych gniazdach obróbkowych, to możemy mieć tu do czynienia z szerokim spektrum: od zestawów prostych komend będących ściśle określonymi ciągami wyrazów, poprzez ciągi wyrazów zawierające jeden element wymienny, aż do złożonych języków quasi-naturalnych.

Stosowanie komunikacji głosowej wydaje się mieć szczególnie wiele zalet w systemach obróbkowych o zróżnicowanym stopniu automatyzacji. Szczególnie wyraźną zaletą jest mobilność operatora. Może on wydawać polecenia głosowe dowolnej maszynie lub urządzeniu, bez konieczności podchodzenia do niego, a nawet bez konieczności przebywania w określonym miejscu. Dzięki temu może jednocześnie realizować inne swoje zadania. Wpisuje się to w popularną ostatnio koncepcję bezpośredniej współpracy ludzi i maszyn we wspólnej przestrzeni roboczej.

Efekty, jakie można potencjalnie uzyskać w wyniku dalszych badań nad komunikacją głosową człowiek-maszyna w zautomatyzowanych gniazdach obróbkowych, są z pewnością warte wysiłku.

Literatura

- [1] Bos J., Oka T.: A spoken language interface with a mobile robot. *Artificial Life and Robotics*, Vol. 11(1) (2007), pp 42-47.
- [2] Dobrzański L. i in.: The prototype of an expert system for the selection of high-speed steels for cutting tools. *Journal of Materials Processing Technology*, Vol. 56 (1996), pp 873-881.
- [3] Fong T., Kunz C., Hiatt L., Bugajska M.: The Human-Robot Interaction Operating System. *Proc. of 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, Salt Lake City, Utah, USA, (2006), pp 41-48.
- [4] Kulyukin V.: Human-Robot Interaction Through Gesture-Free Spoken Dialogue. *Autonomous Robots*, Vol. 16 (2004), pp 239-257.
- [5] Kulyukin V.: Talk the Walk: robotic NLP vs. human sublanguage acquisition. *Proc. of AAAI Spring Symposium on Multidisciplinary Collaboration for Socially Assistive Robotics*, Palo Alto, CA, USA, (2007), pp 33-35.
- [6] Mikołajczyk T., Słomion M., Styp-Rekowski M., Matuszewski M.: Wybrane zagadnienia zastosowania robotów w inżynierii produkcji. *Obróbka Metalu*, Vol. 2 (2017), pp 36-41.
- [7] Oka T. i in.: RUNA: a multimodal command language for home robot users. *Artificial Life and Robotics*, Vol. 13(2) (2009), pp 455-459.
- [8] Pires J.: Robot-by-voice: Experiments on commanding an industrial robot using the human voice. *Industrial Robot*, Vol. 32(6) (2005), pp 505-511.
- [9] Pires J.: The industrial robot as a human coworker: the role of the speech interfaces. *Int. Conference on Software Development for Enhancing Accessibility and Fighting Info-exclusion*, Portugal, (2007).
- [10] Profanter S., Perzyło A., Somani N., Rickert M., Knoll A.: Analysis and Semantic Modeling of Modality Preferences in Industrial Human-Robot Interaction. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* Hamburg, Germany, (2015).
- [11] Rogowski A.: Analiza i synteza systemów sterowania głosowego w zrobotyzowanym wytwarzaniu. *Prace Naukowe Politechniki Warszawskiej, Mechanika*, Vol. 244 (2012)
- [12] Rogowski A.: Czy można rozmawiać z robotem spawalniczym? *Przegląd Spawalnictwa*, Vol. 88(1) (2016), pp 5-8.
- [13] Rogowski A.: Głosowa komunikacja człowiek-maszyna w gniazdach obróbkowych o zróżnicowanym stopniu automatyzacji. *Zeszyty Naukowe Politechniki Poznańskiej*, Vol. 15 (2011), pp 131-144.
- [14] Rogowski A.: Industrially oriented voice control system. *Robotics and Computer-Integrated Manufacturing*, Vol. 28(3) (2012), pp 303-315.
- [15] Rogowski A.: Projekt i implementacja systemu programowania i sterowania gniazdem zrobotyzowanym przez Internet. *Prace Naukowe Politechniki Warszawskiej, seria Elektronika*, Vol. 175 (2010), pp 311-320.
- [16] Rogowski A.: Remote programming and control of the flexible machining cell. *International Journal of Computer Integrated Manufacturing*, Vol. 28(6) (2015), pp 650-663.
- [17] Rogowski A.: Web-based remote voice control of robotized cells. *Robotics and Computer-Integrated Manufacturing*, Vol. 29(4) (2013), pp 77-89.
- [18] Savage J. i in.: VIRbot: a system for the operation of mobile robots. *Lecture Notes in Computer Science*, Vol. 5001 (2008), pp 512-519.
- [19] Stenmark M., Nugues P.: Natural Language Programming of Industrial Robots. *Int. Symposium on Robotics (ISR)* Seoul, Korea, (2013).
- [20] Tunnell G., Pomernacki Ch., Gregg J.: Voice controlled welding system. *Patent US 4641292A*, 1987
- [21] Veiga G., Pires J., Nilsson K.: Experiments with service-oriented architectures for industrial robotic cells programming. *Robotics and Computer-Integrated Manufacturing*, Vol. 25(4-5) (2009), pp 746-755.
- [22] Weiss A. i in.: Exploring human-robot cooperation possibilities for semiconductor manufacturing. *Proc. of Int. Conference on Collaboration Technologies and Systems*, Philadelphia, PA, USA, (2011), pp 173-177.
- [23] Yokota M., Sugita K., Oka T.: Natural language understanding based on mental image description language Lmd and its application to language-centered robot manipulation. *Artificial Life and Robotics*, Vol. 13(1) (2008), pp 84-88.
- [24] Zaeh M., Roesel W.: Safety aspects in a human-robot interaction scenario: a human worker is co-operating with an industrial robot. *Progress in Robotics*, Vol. 44(2) (2009), pp 53-62. ■