

Robot społeczny TEPSON

▶ Grzegorz Granosik
Maciej Stanusch
Krzysztof Wójtowicz

W artykule przedstawiono interaktywny system TEPSON umożliwiający multimodalną komunikację z człowiekiem. Zbudowany prototyp stanowi platformę sprzętową i programową, która ma umożliwić testowanie oprogramowania pochodzącego od różnych producentów, zwłaszcza takiego, które może być wykorzystane do budowy robota społecznego.

Roboty z założenia mają zastępować ludzi lub pomagać im w pracach uciążliwych, monotonicznych i niebezpiecznych, tak powstała najlepiej dotąd rozwinięta gałąź – robotyka przemysłowa. Jednak ciągle niezrealizowanym celem jest zbudowanie urządzenia, które może współpracować z człowiekiem, towarzyszyć mu, komunikować się z nim w sposób naturalny.

Od dawna roztaczane są wizje robotów jako autonomicznych i często autonomicznych urządzeń współdzielących przestrzeń działań z człowiekiem, współpracujących z nim lub służących mu. Na drugim biegunie tych wizji znajdziemy roboty w typie terminatora, którymi straszone jest społeczeństwo.

Okazuje się, że rzeczywistość nadąża za tymi wizjami, roboty z grupy edukacji rozrywkowej (*edutainment*) stają się coraz bardziej popularne i dostępne, a to zwiastuje nadejście ery robotów społecznych. Pierwszą jaskółką w tej dziedzinie w Polsce może być prototyp zbudowany na Politechnice Łódzkiej dla Telekomunikacji Polskiej SA. Urządzenie powstało zgodnie z koncepcją stworzoną przez specjalistów polskiego Orange Labs i nazwane zostało TEPSON.

Projekt TEPSON

Badania związane z robotami usługowymi, interakcją człowiek-robot, sztuczną inteligencją, szeroko pojętym zdobywaniem i przetwarzaniem wiedzy przez roboty są coraz częściej podejmowane na świecie, także w ośrodkach europejskich (chociażby projekt CHRIS [1] lub LIREC [3], którego polskim uczestnikiem jest Politechnika Wrocławska). Celem przedstawianego projektu jest budowa urządzenia proaktywnego, zdolnego do multimodalnej komunikacji z człowiekiem i otoczeniem, rozpoznającego rozmówcę i wydawane przez niego polecenia głosowe. Zbudowana platforma sprzętowa i programowa umożliwia testowanie oprogramowania pochodzącego od różnych producentów, poprzez stosunkowo łatwą wymianę modułów oprogramowania. Drugim istotnym elementem jest zbadanie reakcji osób na kontakt z robotem – platformą informacyjną nowe-

go typu – a w szczególności na jego interfejs głosowy i system dialogowy. Ważne jest przetestowanie odbioru społecznego i emocji, jakie wzbudza urządzenie testowane w warunkach naturalnych w Ogrodach Innowacji Orange Labs, miejscu, gdzie można zobaczyć innowacyjne rozwiązania Grupy TP i Grupy FT (France Telecom). To właśnie tam pokazywane jest, jak świat będzie się komunikował w najbliższej przyszłości.

Podstawowe funkcje modelu

1. Wykrycie rozmówcy na podstawie analizy obrazu (ruchu), detekcji głosu lub aktywności klawiatury.
2. Ustawienie klawiatury na wysokości dopasowanej do wzrostu rozmówcy.
3. Identyfikacja rozmówcy, a w przypadku wykrycia kolejnej wizyty, powitanie w odmienny sposób.
4. Nawiązanie dialogu:
 - Komunikacja w kierunku do użytkownika jest prowadzona w formie komunikatów głosowych z wykorzystaniem systemu dialogowego i syntezy mowy
 - Komunikacja w kierunku od użytkownika może być prowadzona w formie zapytań pisanych na klawiaturze (dialog pełny) lub głosowo w obrębie predefiniowanych komend lub fraz.
5. Wyświetlanie animowanej twarzy oddającej emocje towarzyszące rozmowie.
6. Przekazanie informacji z określonej domeny.
7. Możliwość zestawienia połączenia telefonicznego (poprzez głośnik) do predefiniowanych abonentów (np. infolinii), dzięki temu rozmówca robota może rozmawiać ze zdalnym konsultantem.
8. Możliwość sterowania urządzeniami „inteligentnego budynku” podłączonymi do magistrali EIB (punkty świetlne, rolety itp.).
9. W okresie beczynności, gdy nie jest nawiązana interakcja z użytkownikiem, wyświetlanie na ekranie zamiast animowanej twarzy grafik lub materiałów wideo o charakterze informacyjnym.
10. Możliwość rejestrowania strumienia mowy oraz rozpoznanych poleceń, a także zapytań z klawiatury w celu modyfikacji dialogu zgodnie z oczekiwaniami rozmówców.
11. Realizacja scenariuszy modyfikujących zachowanie robota. Aktualnie system realizuje pięć scenariuszy domyślnych. Tworzenie własnych scenariuszy działania możliwe jest z wykorzystaniem modułu /scenari

▶ dr inż. Grzegorz Granosik, Politechnika Łódzka, Instytut Automatyki; mgr Maciej Stanusch, Stanusch Technologies, Sp. z o.o., Ruda Śląska, Górnośląski Inkubator Technologiczny; mgr inż. Krzysztof Wójtowicz, Zakład Rozwoju Prototypów i Usług, Orange Labs, Warszawa



Rys. 1. Wizualizacja 3D platformy TEPSON bez osłon zewnętrznych dolnej części konstrukcji

Podstawowe cechy modelu

1. Stabilna podstawa wyposażona jest w trójdzielną „głowę”, jak pokazano na wizualizacjach 3D (rys. 1). Jedna z „twarzy” ma swobodę ruchu w kierunku pionowym w reakcji na bodźce zewnętrzne, np. zbliżenie się osoby, głos. Pozostałe dwie „twarze” są nieruchomo przymocowane do głowy i po wyposażeniu w dodatkowe monitory mogą pełnić funkcje wyświetlaczy informacyjnych.
2. Wysokość urządzenia ok. 165 cm w górnym położeniu głowy i zakres ruchu „głowy” 35 cm mają umożliwić wygodną rozmowę z osobami o różnym wzroście. Obrys podstawy ok. 70 × 70 cm pozwala na ustawianie urządzenia w typowych ciągach komunikacyjnych spotykanych w biurach i pomieszczeniach wystawowych, wielkość twarzy – owal o średnicach ok. 60 cm i 40 cm; przygotowano dwie wersje graficzne twarzy (rys. 2).



Rys. 2. Dwie wersje graficzne twarzy oraz osłon zewnętrznych dolnej części konstrukcji TEPSON



Rys. 3. Wersja ażurowa z wymiennymi nakładkami magnetycznymi

3. Urządzenie pozwala łatwo zmieniać wersję wyglądu zewnętrznego pomiędzy:
 - ażurową – z widoczną konstrukcją szkieletową środkowej części (szkielet w kolorze srebrzyszarym, kolumny czarne) z możliwością przymocowywania magnetycznych nakładek graficznych na podstawę i kolumny (rys. 3)
 - zamkniętą – z łatwo wymiennymi osłonami bocznymi wykonanymi z lekkiego materiału (PVC) z nadrukiem.

W obu wersjach zmiana owalu „twarzy” następuje przez wymianę osłony wykonanej z plexiglasu.

4. Podstawowe komponenty urządzenia: dwa komputery PC, monitor LCD 19”, napęd umożliwiający ruch monitora, klawiatura, para głośników, mikrofony elektretowe, kamera z napędem w jednej osi, sterowniki elementów wykonawczych i zespół zasilaczy. System ma zasilanie sieciowe 230 V.
5. Komunikacja między podsystemami/modułami platformy wykorzystuje protokół YARP [2], w niższej warstwie połączenie zrealizowane jest poprzez lokalną sieć Ethernet, z bramką umożliwiającą dostęp do Internetu i komunikację bezprzewodową WiFi.
6. System dialogowy korzysta z oprogramowania Wirtualny Doradca, jest wyposażony w bazę wiedzy ogólnej oraz bazę wiedzy specyficznej wskazanej przez TP SA, syntezę dźwięku realizuje oprogramowanie Ivona, a podsystem rozpoznawania mowy oprogramowanie SkryBot, które reaguje na ograniczoną liczbę (500) fraz bądź słów.
7. System wizyjny bazujący na jednej kamerze zawiera aplikację automatycznego wykrywania ruchu (AMR) oraz automatycznego rozpoznawania twarzy (AFR) i tworzenia bazy danych znanych twarzy został także opracowany przez firmę Stanusch Technologies.

Konstrukcja mechaniczna

Urządzenie TEPSON wykonane jest z profili stalowych spawanych o podstawie w kształcie sześciokąta. Dla ułatwienia przestawiania zamocowano 6 kółek (z blokadą) typu kastor o średnicy ok. 8 cm. Dolna część konstrukcji jest zasłonięta blachami stalowymi, przy czym dwie blachy (po jednej z przodu i z tyłu) są mocowane na śruby, aby umożliwić dostęp do komputerów w dolnej części urządzenia.

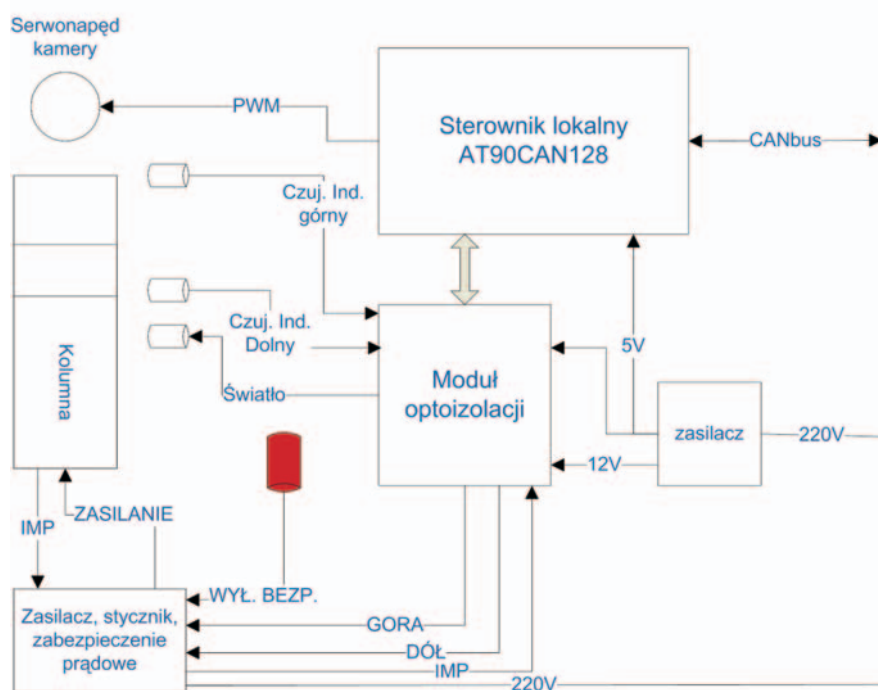


Rys. 4. Monitor z wyposażeniem dodatkowym wewnątrz osłony, zawieszony na kolumnie napędowej

Środkowa część urządzenia zachowuje kształt ostrosłupa sześciobocznego i zawiera 3 kolumny (jedna napędowa i dwie atrapy). U szczytu każdej z nich znajduje się uchwyt do mocowania monitora. Kolumna aktywna posiada ponadto platformę do mocowania klawiatury, głośników, kamery i mikrofonów (rys. 4).

Układ napędowy

Składa się z komercyjnej kolumny TVL2 o udźwigu 80 kg, wysokości minimalnej 570 mm i skoku całkowitym ustawionym na wartość 350 mm (zakres ruchu został dodatkowo zabezpieczony przez zastosowanie czujników indukcyjnych). Ponadto system zawiera serwonapęd kamery HS-475MG. Komponenty napędowe zostały połączone z układem sterowania i zasilania (rys. 5).



Rys. 5. Schemat blokowy połączeń w układzie napędowym

Serwonapęd wraz z kamerą został umieszczony bezpośrednio nad monitorem i umożliwia przemieszczanie kamery od ok. -45° do $+45^\circ$ mierząc od poziomu. Moduł optoizolacji został umieszczony na szczycie konstrukcji w obszarze pomiędzy kolumnami, zaś czujniki indukcyjne górny i dolny umieszczono na wspornikach kolumny napędowej.

Czujniki indukcyjne zabezpieczają ograniczenia dolne i górne ruchu kolumny w zakresie 350 mm. Dodatkowe zabezpieczenie stanowi programowe zliczanie impulsów generowanych przez czujnik Halla umieszczony na silniku kolumny oraz wyłącznik bezpieczeństwa z rygłem umieszczony na górnej obudowie urządzenia, który odcina zasilanie kolumny. Oświetlenie w postaci reflektora diodowego umieszczono pod górnym wspornikiem kolumn. Zapewnia ono dodatkowy efekt wizualny umożliwiając programowalne załączanie i wyłączanie podświetlenia przestrzeni między kolumnami.

Cała przestrzeń między kolumnami znajdująca się za ruchomą twarzą została zasłonięta specjalną metalową obudową z odpowiednimi otworami na poruszającą się kolumnę. Otwory te w miarę przesuwania się twarzy zasłaniane są przez wbudowaną materiałową żaluzję.

Architektura programowa

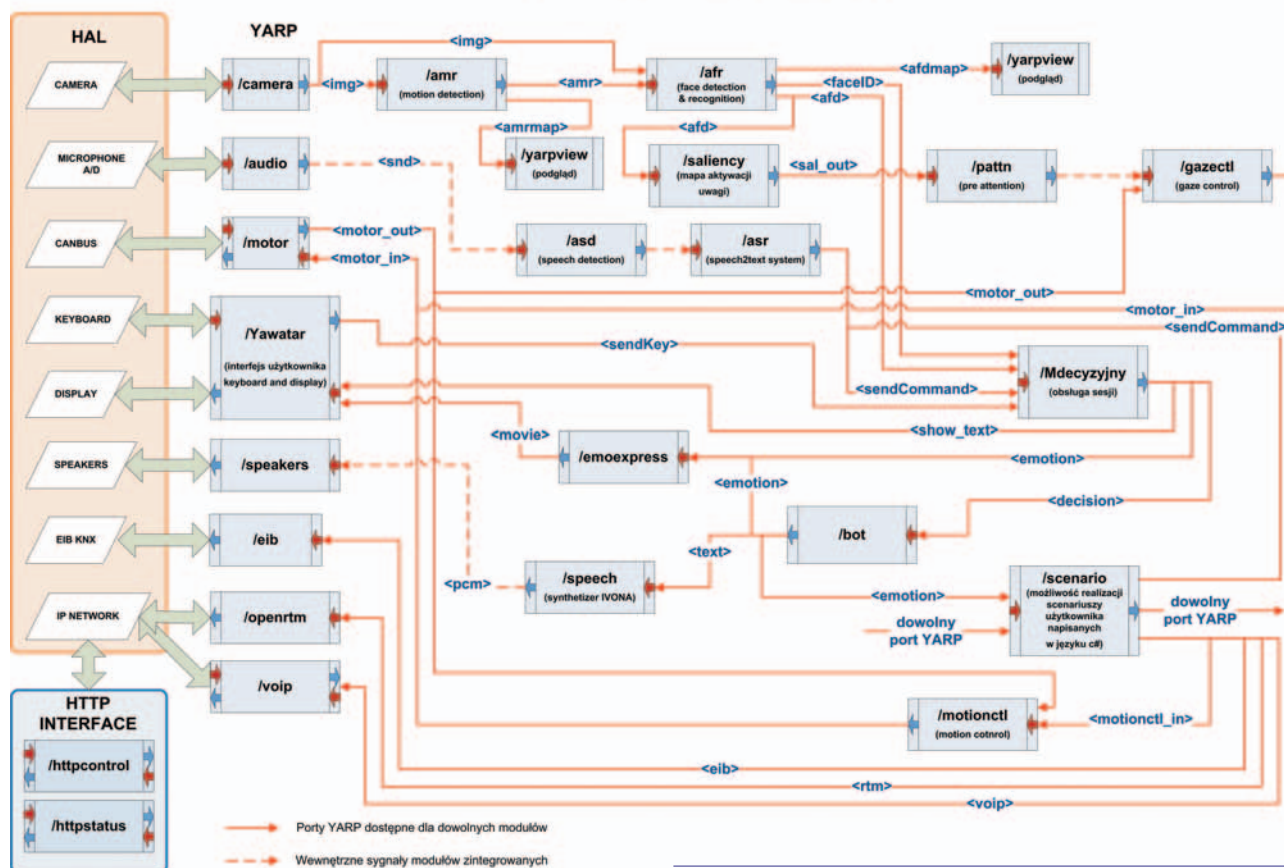
Poszczególne podsystemy urządzenia TEPSON wraz z interfejsami pomiędzy tymi podsystemami oraz system rejestracji strumienia mowy, rozpoznawania mowy i system akwizycji obrazu są zintegrowane przez oprogramowanie pośredniczące (*middleware*) YARP.

YARP (*Yet Another Robot Platform*) jest zestawem bibliotek, protokołów i narzędzi wykorzystywanym w projektach robotycznych, których platforma programowa w założeniach ma być modułowa. Powstał na podstawie doświadczeń przy budowie robotów humanoidalnych. Celem jego twórców było zminimalizowanie czasu poświęconego na tworzenie oprogramowania poprzez wprowadzenie modularności. Powiązanie z tą konkretną gałęzią robotyki nie jest przypadkowe – charakteryzuje się ona częstymi zmianami w części sprzętowej – w takim wypadku kontrola kodu programów jest poważnym wyzwaniem. YARP minimalizuje te problemy oraz porządkuje komunikację między procesami i zapewnia niezależność od platformy programowej – może być z powodzeniem stosowany na platformach Windows, Linux czy QNX. YARP zapewnia wyłącznie komunikację między modułami, nie dostarcza gotowych rozwiązań z zakresu logiki. Zastosowanie platformy YARP daje także możliwość skalowania i rozpraszania systemu, co jest niezwykle istotne przy zmianie aplikacji robota społecznego. Szczegółowy opis działania YARP, dokumentację oraz przykłady można znaleźć na stronach internetowych otwartego i ogólnodostępnego projektu: <http://eris.liralab.it/yarp/>.

Struktura modułów programowych systemu TEPSON (rys. 6) składa się z warstwy sprzętowej – HAL oraz warstwy programowej – YARP. W skład systemu wchodzi następujące moduły programowe:

- akwizycja obrazu, interfejs YARP dla kamery firewire (/camera)
- /amr – moduł wykrywania ruchu
- /afr – moduł wykrywania, rozpoznawania i rejestracji twarzy, zawierający także AdminPanel – moduł administracyjny bazy twarzy
- /saliency – moduły tworzenia mapy uwagi, aktualnie na podstawie informacji wizyjnej; możliwa jest

Architektura modułowa systemu TEPSON



Rys. 6. Struktura oprogramowania TEPSON

dalsza rozbudowa tego modułu na potrzeby fuzji informacji wizyjnej i dźwiękowej oraz określania kierunków największego zainteresowania dla systemu TEPSON

- aktywne ustawianie położenia kolumny oraz kamery na podstawie informacji z systemu uwagi – zintegrowane moduły /pattn i /gazectl
- sterowanie oraz dostęp do statusu silników – moduły /motor i /motionctl – interfejs YARP dla szyny CAN
- akwizycja i obróbka dźwięku: interfejs YARP dla karty dźwiękowej (moduł /audio), zintegrowany z /asd – modułem detekcji i wydzielenia fragmentów mowy oraz /asr – modułem rozpoznawania mowy (interfejs YARP do programu SkryBot)
- moduł /Yawatar – zarządzający standardowym wejściem i wyjściem interfejsu użytkownika: obsługa klawiatury, wyświetlanie wprowadzanego tekstu, wizualizacja twarzy awatara
- /emoexpress – moduł wyrażania emocji
- /Mdecyzyjny – moduł decyzyjny, zarządzający informacjami pochodzącymi z innych modułów a dotyczącymi rozmowy z człowiekiem
- /scenarij – moduł realizujący predefiniowane scenariusze zachowań systemu oraz scenariusze użytkownika
- /bot – moduł dialogowy, silnik sterujący rozmową Tepsiona z człowiekiem
- wyjście na głośniki: interfejs YARP dla karty dźwiękowej (/speakers), zintegrowany z modułem syntezy dźwięku – /speech

- dostęp do urządzeń inteligentnego budynku, interfejs YARP dla sieci EIB/KNX (/eib)
- dostęp do urządzeń zewnętrznych, interfejs YARP dla architektury openRTM (/openrtm)
- dostęp do sieci telefonicznej przez bramkę VoIP (/voip).

Działanie systemu

Aktualnie robot jest szczegółowo testowany w polskim oddziale Orange Labs. Wprowadzane są iteracyjnie poprawki i uzupełnienia niezbędne w obecnej fazie rozwojowej.

Z robotem TEPSON można się przywitać w kilku językach m.in. polskim, angielskim, francuskim i niemieckim oraz nakłonić go, by opowiedział coś o sobie. Oprócz reakcji ogólnych, robot potrafi rozpoznawać zwroty dotyczące usług Ogrodów Innowacji i oferty Telekomunikacji Polskiej. Dzięki temu można w łatwy sposób zapytać go o takie tematy jak chociażby „live-box” czy „kurtka bluetooth”.

System reaguje na ruch. Na podstawie analizy obrazów zarejestrowanych przez kamerę, określa kierunek ruchu i jego dynamikę. Ten bodziec jest wykorzystywany przez inne moduły: rozpoczyna się ruch kamery w celu znalezienia twarzy, rozpoczyna się proces wykrywania twarzy, jej rozpoznania i rejestracji, jeśli twarz zostanie znaleziona ustawienie klawiatury zostanie dopasowane do wzrostu osoby,

My wiemy **kto** wie...



www.automatyka.pl

Serwis branżowy www.automatyka.pl gromadzi informacje o produktach i usługach z branży automatyki przemysłowej. Zasoby Serwisu tworzone są samodzielnie przez zarejestrowane firmy. Każda z nich wprowadza informacje o własnej ofercie, produktach, usługach, wydarzeniach. Dzięki temu Serwis prezentuje żywy, stale aktualny obraz branży. Jest szybkim i skutecznym środkiem komunikacji pomiędzy uczestnikami rynku.

www.automatyka.pl – cała branża w zasięgu ręki

a system dialogowy przywita rozmówcę. Jego zadaniem jest udzielanie informacji klientom nt. oferowanych produktów i usług, specyfiki działalności firmy, czy innych informacji. Program komputerowy umożliwiający rozmowę z użytkownikiem w języku naturalnym wykorzystuje metody i techniki sztucznej inteligencji. Wyposażony jest również w układ syntezy mowy oraz animowaną wirtualną twarz, a także w oprogramowanie rejestrujące strumień mowy i rozpoznające podstawowe polecenia. Gdy system dialogowy wykryje odpowiednie polecenie włącza lub wyłącza oświetlenie, łączy się z określonym numerem telefonu przez bramkę VoIP lub przesuwa klawiaturę.

Podsumowanie

Realizując projekt TEPSON połączyliśmy wiedzę ekspercką z dziedziny robotyki, sztucznej inteligencji, analizy obrazów oraz rozpoznawania mowy tworząc zwarty, ale także elastyczny i skalowalny system. Mamy nadzieję, że powstało coś więcej niż tylko połączenie wymienionych komponentów. Robot potrafi wejść w lokalną interakcję z człowiekiem, ale może także sterować zdalnymi urządzeniami dołączonymi poprzez sieć Internet, EIB lub inne media. Wierzymy, że jest to właściwa droga rozwoju urządzeń klasyfikowanych jako roboty społeczne.

Autorzy dziękują współtwórcom systemu TEPSON: dr. inż. Krzysztofowi Mianowskiemu – konstruktorowi części mechanicznej, mgr. inż. Lechowi Błażejowskiemu, mgr. Jackowi Magnuskiemu, mgr. inż. Markowi Koniewowi, mgr. Michałowi Depcie, mgr. inż. Arturowi Gmerkowi, mgr. inż. Przemysławowi Szewczykowski, mgr. inż. Markowi Gawryszewskiemu – twórcom poszczególnych modułów programowych.

Bibliografia

1. Melhuish C., Hamann, K., Warneken, F., Tomasello, M., *First Year Project Activities Summary* – April 2009 (pdf document), <http://www.chrisfp7.eu/research.html>
2. Welcome to YARP [online], <http://eris.liralab.it/yarpdoc/index.html>
3. <http://lirec.eu/> – strona internetowa projektu

REKLAMA

