

Răzvan Gabriel BOBOC

**INTERACȚIUNE NATURALĂ OM-ROBOT  
PENTRU APLICAȚII DE ROBOTICĂ ASISTIVĂ**



Editura  
Universității  
Transilvania  
din Brașov

2024

## **EDITURA UNIVERSITĂȚII TRANSILVANIA DIN BRAȘOV**

Adresa: Str. Iuliu Maniu nr. 41A  
500091 Brașov  
Tel.: 0268 476 050  
Fax: 0268 476 051  
E-mail: editura@unitbv.ro

**Editură recunoscută CNCSIS, cod 81**

**ISBN 978-606-19-1762-4 (ebook)**

Copyright © Autorul, 2024

Referenți științifici: Prof. dr. ing. Antonya CSABA  
Dr. ing. Adrian-Iulian DUMITRU

# CUPRINS

<b>LISTA DE ABREVIERI</b> .....	7
<b>PREFAȚĂ</b> .....	9
<b>1. INTRODUCERE</b> .....	11
1.1 Aspecte generale.....	13
1.2 Scopul și obiectivele cercetării.....	14
1.3 Structura lucrării.....	16
<b>2. SISTEME DE INTERACȚIUNE NATURALĂ OM – ROBOT ÎN ROBOTICA ASISTIVĂ</b> .....	19
2.1 Introducere.....	20
2.2 Interacțiunea om-robot.....	20
2.2.1 Clasificarea interacțiunilor om-robot.....	23
2.2.2 Interacțiunea naturală.....	27
2.2.3 Concluzii.....	29
2.3 Roboți utilizați în robotica asistivă.....	29
2.3.1 Scurtă istorie a roboților.....	30
2.3.2 Definiții ale roboților.....	32
2.3.3 Clasificarea roboților asistivi.....	34
2.3.4 Exemple de roboți asistivi.....	36
2.3.4.1 Roboți pentru prestări de servicii.....	38
2.3.4.2 Roboți de companie.....	39
2.3.5 Concluzii.....	42
2.4 Percepția asupra mediului înconjurător.....	43
2.4.1 Tipuri de percepție.....	43

2.4.2 Urmărirea mișcărilor corpului.....	46
2.5 Gesturi folosite în interacțiunea om - robot.....	47
2.5.1 Tipuri de gesturi.....	48
2.5.2 Tehnici de recunoaștere a gesturilor.....	49
2.6 Aplicații de interacțiune om-robot în robotica asistivă .....	50
2.6.1 Aplicații bazate pe interacțiune prin gesturi.....	51
2.6.2 Aplicații bazate pe interacțiune prin voce.....	52
2.6.3 Aplicații bazate pe interacțiune multimodală.....	54
2.7 Concluzii.....	56
<b>3. PREMISE HARDWARE ȘI SOFTWARE PENTRU STUDIILE TEORETICE ȘI</b>	
<b>EXPERIMENTALE.....</b>	<b>58</b>
3.1 Dispozitivul Kinect.....	58
3.1.1. Tehnologia Kinect.....	59
3.1.2. Aplicații ale dispozitivului Kinect.....	61
3.1.3. Instrumentul de dezvoltare software.....	62
3.2 Structura fizică a robotului umanoid NAO.....	64
3.2.1 Componentele și senzorii.....	65
3.2.2 Componentele software.....	70
3.2.3 Interfața de comandă a robotului.....	74
3.3 Modelarea robotului și a mediului de lucru.....	76
3.3.1 Supoziții.....	76
3.3.2 Modelarea robotului în mediul virtual.....	77
3.4 Deplasarea robotului.....	78
3.5 Detecția obiectelor în mediu.....	79
3.6 Concluzii.....	81

<b>4. SISTEM DE INTERACȚIUNE NATURALĂ OM-ROBOT</b> .....	82
4.1 Interacțiunea prin gesturi.....	83
4.1.1 Recunoașterea gesturilor umane.....	83
4.1.1.1 Aspecte generale.....	84
4.1.1.2 Structura algoritmului propus.....	85
4.1.1.3 Extragerea caracteristicilor.....	86
4.1.1.4 Metoda DTW (Dynamic Time Warping).....	89
4.1.1.5 Metoda DTW îmbunătățită.....	93
4.1.1.6 Concluzii.....	95
4.1.2 Biblioteca de gesturi.....	96
4.2 Interacțiunea prin voce.....	99
4.2.1 Metoda de recunoaștere vocală.....	99
4.2.2 Comenzi vocale.....	100
4.3 Sistemul de integrare gesturi/voce.....	102
4.4 Paradigma „Point-and-command”.....	105
4.5 Sarcinile robotului.....	111
4.6 Teleoperare bazată pe gesturi.....	114
4.6.1 Aspecte generale.....	114
4.6.2 Imitarea mișcărilor umane.....	116
4.7 Concluzii.....	120
<b>5. SISTEMUL DE LUARE A DECIZIILOR AL APLICAȚIEI DE INTERACȚIUNE NATURALĂ OM-ROBOT</b> .....	121
5.1 Sisteme de suport al deciziei.....	122

5.2 Logica fuzzy.....	124
5.3 Arhitectura sistemului.....	126
5.3.1 Blocul de decizie.....	127
5.3.2 Blocul de comunicare.....	137
5.3.3 Sistemul de control.....	137
5.4 Concluzii.....	138
<b>6. TESTE ȘI EXPERIMENTE CU SISTEMUL DE INTERACȚIUNE NATURALĂ.....</b>	<b>140</b>
6.1 Introducere.....	140
6.2 Definirea mediului de lucru.....	141
6.3 Definirea sarcinilor și operațiilor robotului.....	142
6.4 Experimentul de imitare a mișcărilor utilizatorului.....	143
6.5 Experimentul de testare a interacțiunii multimodale.....	148
6.5.1 Scenariul 1.....	148
6.5.2 Scenariul 2.....	150
6.6 Experimentul de testare a eficienței realizării sarcinilor de către robot.....	153
6.7 Concluzii.....	157
<b>7. CONCLUZII FINALE, CONTRIBUȚII ORIGINALE ȘI DIRECȚII VIITOARE.....</b>	<b>159</b>
7.1 Concluzii finale.....	159
7.2 Contribuții originale.....	161
7.4 Direcții viitoare de cercetare.....	163
<b>BIBLIOGRAFIE.....</b>	<b>165</b>

## ***LISTA DE ABREVIERI***

<b>AR</b>	<b>Assistive Robot</b> - Robot de asistență
<b>3D</b>	<b>Three-Dimensional Space</b> – Spațiu tridimensional
<b>ANN</b>	<b>Artificial Neural Network</b> – Rețea neuronală artificială
<b>API</b>	<b>Application Programing Interface</b> – Interfață de programare a aplicațiilor
<b>BCI</b>	<b>Brain – Computer Interface</b> – Interfață creier - calculator
<b>BN</b>	<b>Bayesian Network</b> – Rețea bayesiană
<b>CCD</b>	<b>Charge-Coupled Device</b> – Dispozitiv cu cuplaj de sarcină
<b>CMOS</b>	<b>Complementary Metal–Oxide–Semiconductor</b> – Semiconductor complementar din oxid metalic
<b>DOF</b>	<b>Degree of Freedom</b> – Grad de libertate
<b>DSS</b>	<b>Decision Support System</b> – Sistem de suport decizional
<b>DTW</b>	<b>Dynamic Time Warping</b> – Aliniere temporală dinamică
<b>EGM</b>	<b>Elastic Graph Matching</b> – tehnică de recunoaștere de obiecte
<b>EOG</b>	<b>ElectroOculoGraphy</b> - Electrooculografie
<b>FL</b>	<b>Fuzzy Logic</b> – Logică fuzzy
<b>FSM</b>	<b>Finite - State Machine</b> – Mașină cu stări finite
<b>GPS</b>	<b>Global Positioning System</b> – Sistem de poziționare globală
<b>GUI</b>	<b>Graphical User Interface</b> – Interfață grafică cu utilizatorul
<b>HCI</b>	<b>Human Computer Interaction</b> – Interacțiune om - calculator
<b>HMM</b>	<b>Hidden Markov Model</b> – Model Markov ascuns
<b>HRI</b>	<b>Human Robot Interaction</b> – Interacțiune om-robot
<b>HSV</b>	<b>Hue, Saturation, Value</b> – Model de culoare Nuanță, Saturație, Valoare
<b>IP</b>	<b>Internet Protocol</b> – Protocol Internet

<b>IR</b>	<b>Infrared Radiation</b> – Radiație infraroșu
<b>MDP</b>	<b>Markov Decision Processes</b> – Proces de decizie Markov
<b>NI</b>	<b>Natural Interaction</b> – Interacțiune naturală
<b>NN</b>	<b>Neural Network</b> – Rețea neuronală
<b>NUI</b>	<b>Natural User Interface</b> – Interfață naturală cu utilizatorul
<b>RF</b>	<b>Radio-Frequency</b> – Radiofrecvență
<b>RFID</b>	<b>Radio-Frequency Identification</b> – Identificare prin radiofrecvență
<b>RGB</b>	<b>Red, Green, Blue</b> – Model de culoare Roșu, Verde, Albastru
<b>SAR</b>	<b>Socially Assistive Robotics</b> – Robotică de asistență socială
<b>SDK</b>	<b>Software Development Kit</b> – Set de unelte software de dezvoltare de aplicații
<b>TCP</b>	<b>Transmission Control Protocol</b> – Protocol de control al transmisiei
<b>UV</b>	<b>UltraViolet</b> – Radiație ultravioletă



## *PREFAȚĂ*

Robotica de asistență socială, un anumit tip de robotică de asistență destinată pentru interacțiunea socială cu oamenii, ar putea juca un rol important în ceea ce privește sănătatea și bunăstarea psihologică a persoanelor în vârstă și nu numai. Obiectivele roboților sociali de asistență sunt considerate a fi utile pentru îngrijirea persoanelor în vârstă din două motive, una funcțională și una afectivă. Astfel de roboți sunt dezvoltati pentru a funcționa ca o interfață cu tehnologia digitală și pentru a contribui la creșterea calității vieții prin furnizarea de companie. Există o atenție tot mai mare pentru aceste dispozitive în literatura de specialitate.

Roboții de servicii sunt dispozitive mecanice mobile care operează în co-prezență cu utilizatorii lor. Aceasta este o provocare pentru proiectarea de interacțiuni om-robot. Interfața robotului trebuie să ajute utilizatorul în înțelegerea stării curente a sistemului și a posibilelor acțiuni următoare. Un aspect în proiectarea unor astfel de interacțiuni este înțelegerea preferințelor și așteptărilor utilizatorilor prin includerea lor în procesul de proiectare.

Interacțiunea om-robot (HRI), definită ca studiul interacțiunilor dintre oameni și roboți, implică mai multe discipline diferite cum ar fi știința calculatoarelor, inginerie, științe sociale și psihologie. Pentru HRI, provocările sunt deosebit de complexe, din cauza necesității de a percepe, înțelege, și a reacționa la activitățile umane în timp real. Principalele aspecte cheie ale percepției sunt multimodalitatea și atenția. Multimodalitatea permite oamenilor să se miște cu ușurință între diferitele moduri de interacțiune, de la vizual la voce și atingere, în funcție de schimbările din context sau de preferință utilizatorului, în timp ce atenția este procesul cognitiv de concentrare selectivă pe un aspect al mediului, în timp ce alte lucruri sunt ignorate. Multimodalitatea permite robotului să interpreteze și să reacționeze la diferiți stimuli ai oamenilor (de exemplu, gest, vorbire, privire).

Tendențele recente în interfețele cu utilizatorul au introdus noi tehnici de interacțiune care fac trecerea de la mouse-ul și tastatura tradiționale,

incluzând interfețe multi-touch și sisteme bazate pe camere video (de exemplu, Microsoft Kinect, Xtion sau Leap Motion Controller). Trăsătura comună a acestor dispozitive este că ele fac interacțiunea om-calculator și om-robot mai naturală și fără efort. Multe dintre ele permit utilizatorilor să folosească mâinile lor și/sau gesturi ale brațelor pentru a trimite comenzi de control și astfel se simt mult mai natural pentru că acesta este modul în care interacționăm cu oamenii zi de zi.

Intenția acestei lucrări este de a prezenta o metodă de interacțiune naturală bazată pe recunoașterea gesturilor umane, inclusiv gesturi de indicare, care sunt fundamentale atunci când se interacționează cu roboți mobili, și un mecanism care permite combinarea comenzilor efectuate prin intermediul brațelor cu comenzi vocale. Scopul este de a oferi un sistem care poate fi utilizat în timp real și poate face față variabilității utilizatorilor în timpul execuției gesturilor. Abordarea propusă se bazează pe tehnica alinierii temporale dinamice (DTW), un algoritm de potrivire a modelelor folosit pentru a recunoaște gesturi dintr-o secvență video.

O paradigmă Point-and-Command este propusă pentru a combina cele două moduri de a trimite comenzi robotului (vocea și gesturile). Astfel, robotul se poate deplasa într-o direcție indicată de utilizator, realizând sarcina cerută de acesta prin comenzi vocale. Dispozitivul utilizat pentru a obține date (atât vizuale, cât și vocale) este un dispozitiv Kinect, în timp ce platforma robotică utilizată este robotul umanoid NAO.

Sistemul dezvoltat se adresează utilizatorilor care sunt profani în domeniul interacțiunilor cu roboți mobili și care doresc să comunice cu robotul personal sau de asistență într-un mod natural, obișnuit, fără multe pregătiri prealabile.

Noiembrie, 2024

Dr. ing. Răzvan Gabriel BOBOC

## **INTRODUCERE**

Din timpuri străvechi omenirea și-a imaginat diferite mașini care imită organisme sau care depășesc oamenii în abilitățile lor și din timp în timp au existat numeroase încercări de construire a unor astfel de mașini. Roboții de astăzi sunt cele mai recente rezultate ale unei evoluții tehnice aflată în curs de desfășurare, care a progresat pe parcursul a câteva secole.

Mulți roboți efectuează operațiuni în fabricile unde se produc bunuri în cantități mari și de calitate, dar în ultimul timp se poate observa apariția unor noi clase de roboți. Acești noi *roboți de servicii* nu sunt concepuți pentru a produce bunuri, ci pentru a oferi servicii utile la locul de muncă sau la domiciliu sau pur și simplu pentru a distra și a oferi confort oamenilor [Graefe, 2009].

Roboții antropomorfi își găsesc locul mai ales în educație și divertisment, cum este cazul robotului umanoid NAO. Roboți precum Asimo (Honda) sau Qrio (Sony) sau Roboray (Samsung) merg pe urmele primilor roboți, iar scopul lor este în cea mai mare parte de a arăta maturitatea tehnică a constructorilor lor. Forma umanoidă poate fi avantajoasă în domeniul în care roboții se află într-un strâns contact cu oamenii sau sunt co-lucrători cu ei (lucrează în același spațiu). Un exemplu este Robonaut, robotul umanoid construit de NASA în colaborare cu General Motors, fiind destinat să lucreze cot la cot cu astronauții umani [Engelberger, 2001].

Domeniul în care robotica se inspiră încă de la oameni și animale este cunoașterea și inteligența. Robotica cognitivă vede cunoașterea umană și cea animală ca un punct de plecare pentru dezvoltarea de procesare a informației robotice. Ținta capacităților cognitive robotice include procesarea percepției, alocarea atenției [Sugiyama, 2007], anticiparea [Koppula, 2013], planificarea, coordonarea motorie complexă, raționamentul despre mediul înconjurător, acțiunile și obiectivele.

Interacțiunea om-robot este un domeniu interdisciplinar în care abordările comportamentale și psihologice față de înțelegerea naturii interacțiunii om-robot completează preocupările din domeniul roboticii și al ingineriei. În acest context, s-a considerat oportună realizarea de cercetări privind conceperea unor sisteme prin care interacțiunea să se facă cât mai natural, la fel cum oamenii comunică între ei în viața de zi cu zi.

Astfel, această cercetare explorează proiectarea și evaluarea unui sistem de interacțiune om-robot pentru roboți de asistență, care folosește limbajul natural și gesturile pentru a interacționa cu aceștia. Cercetarea se concentrează în jurul a trei teme: analiza elementelor care fac posibilă o interacțiune naturală; metode de realizare a unei astfel de interacțiuni; proiectarea interacțiunii om-robot.

Scenariile au fost gândite pentru un robot mobil umanoid destinat serviciilor fetch-and-carry într-un mediu de birou sau unul domestic, în care utilizatorul îi cere robotului obiecte plasate în diferite locații. Folosind aceste scenarii, a fost dezvoltat și analizat un sistem de interacțiune om-robot bazat pe gesturi și voce.

Analiza comportamentelor de comunicare a dus la următoarele observații: utilizatorii trebuie să comunice cu robotul în scopul de a rezolva un obiectiv principal. Pentru a îndeplini acest scop, ei trebuie să înțeleagă faptul că robotul este incapabil de multe ori să se descurce singur și trebuie să îi ofere informațiile necesare pentru a putea îndeplini un obiectiv comun. În timpul interacțiunilor utilizatorii monitorizează în permanență comportamentul robotului, încercând să obțină feedback-ul necesar pentru a înțelege starea sa. În acest sens, feedback-ul obținut de la robot este foarte important pentru utilizator.

Studiile prezentate în această lucrare aduc contribuții la implementarea unor algoritmi de comunicare om-robot și de manipulare a obiectelor, în scopul folosirii roboților mobili ca asistenți în mediile domestice și nu numai.

## 1.1. ASPECTE GENERALE

În prezent, roboții fac trecerea rapidă din laboratoare și fabrici în casele oamenilor, birouri și străzi. Roboții mobili trebuie să exploreze și să navigheze medii întinse din viața de zi cu zi, doar cu ipoteze generale despre proprietățile și structura lor. Pentru a face acest lucru, ei trebuie să simtă mediul, să construiască o reprezentare a mediului, să raționeze despre acea reprezentare și să efectueze acțiuni pentru îndeplinirea scopurilor lor. Ei trebuie, de asemenea, să interacționeze și să comunice cu oamenii într-o formă ușor de înțeles. Cu alte cuvinte, robotul trebuie să posede capacități cognitive cu scopul de a funcționa într-un mediu în care omul își desfășoară activitatea în mod obișnuit în fiecare zi. Câteva dintre aceste capacități cognitive sunt următoarele:

- atitudini de informare (cunoștințe și convingeri);
- atitudini motivaționale (preferințe și scopuri);
- capacități cognitive (raționament, luarea deciziilor, planificare, observare și comunicare);
- capacități fizice (pentru a se mișca în lumea fizică și de a interacționa în condiții de siguranță cu obiecte din acea lume).

Utilizarea roboților mobili pentru execuția de activități în medii domestice este limitată de probleme încă nesoluționate total, precum: percepția mediului, interacțiunea naturală om-robot, învățarea de noi sarcini, planificarea traseelor, evitarea obstacolelor etc. Se impune dezvoltarea de algoritmi de imitare, învățare, recunoaștere a diverselor modalități prin care omul poate interacționa cu roboți mobili.

Atunci când oamenii comunică între ei folosesc anumite comportamente de comunicare cu scopul de a stabili și menține un numitor comun. Pe de altă parte, și comunicarea om-robot are multe caracteristici în comun cu dialogul om-om orientat pe sarcină. Acest lucru se datorează în primul rând faptului că sunt implicați tot oamenii, dar și faptului că robotul poate și trebuie să folosească un limbaj natural pentru schimbul și partajarea de informații despre obiectivele comune.

Câteva dintre caracteristicile care trebuie să fie luate în considerare în proiectarea de sisteme de interacțiune om-robot sunt următoarele:

- comunicarea dintre oameni și roboți trebuie să fie multimodală, încorporând expresii vocale, gesturi, privirea, emoții, posturi;

- mediul în care acționează robotul, spațiul împărțit cu utilizatorul, locul de amplasare și obiectele disponibile formează un scenariu de utilizare complex;
- aspectul robotului și mișcările sale influențează comportamentul și atitudinea utilizatorului;
- feedback-ul de comunicare oferit de către robot influențează calitatea interacțiunii.

Provocările enumerate mai sus pot fi rezumate în următoarele întrebări de cercetare relevante:

*Care sunt modurile de interacțiune adecvate pentru interacțiunea cu roboți mobili de asistență? Care sunt aspectele practice și teoretice relevante?*

*Cum se poate realiza o interacțiune naturală om-robot?*

## **1.2. SCOPUL ȘI OBIECTIVELE CERCETĂRII**

Cercetarea își propune să ofere o soluție pentru interacțiunea dintre operatorul uman și roboții mobili de asistență. Mijloacele de interacțiune sunt foarte importante pentru realizarea unei interacțiuni simple, flexibile, intuitive și în același timp eficiente. Abordarea acestei teme a fost așadar determinată de necesitatea controlului unui robot de asistență (AR) care să execute anumite activități sau operațiuni în medii domestice sau la birouri, astfel încât comunicarea și interacțiunea dintre acesta și operatorul uman să se facă într-un mod cât mai natural și asemănător comportamentului uman.

Înțelegerea interacțiunii naturale a unui utilizator este o provocare care trebuie abordată pentru a permite utilizatorilor începători să folosească roboți într-un mod ușor și intuitiv. Folosirea unui set de comenzi greu codificate pentru a controla un robot este, de obicei, destul de fiabilă și ușor de implementat, însă este supărătoare pentru utilizator, întrucât acest lucru îi cere să învețe și să își amintească comenzi speciale, în scopul de a interacționa cu robotul și nu permite utilizatorului să folosească stilul lui de interacțiune naturală. Înțelegerea limbajului vorbit natural, fără restricții și comportamentul multi-modal al utilizatorilor ar fi de dorit, dar este încă o problemă nerezolvată.

Prin urmare, cercetarea propune o abordare specifică pentru a permite unui utilizator uman să interacționeze într-un mod natural cu un robot mobil în realizarea unor sarcini obișnuite care necesită manipulare de obiecte și deplasare prin mediu. Utilizatorul ajută robotul să ducă la

îndeplinire sarcina, ghidându-l printr-un număr limitat de comenzi exprimate cu ajutorul gesturilor corporale și al expresiilor vocale, arătându-și aprobarea sau dezaprobarea în unele situații.

În scopul de a crea cadrul necesar pentru o interacțiune naturală, au fost create biblioteci de gesturi și de comenzi vocale, iar robotul poate răspunde prin anumite comportamente, astfel încât utilizatorii neexperimentați să se simtă confortabil în preajma unui astfel de robot.

Motivația principală de la care s-a plecat în realizarea acestei cercetări a fost ideea că e nevoie să se construiască și să se programeze roboți care pot înțelege și pot genera comportamente asemănătoare cu cele umane, astfel încât ei să poată coopera cu oameni neexperimentați și să poată învăța de la aceștia, putând fi aplicați cu ușurință într-o gamă largă de contexte. Desigur, limbajul natural joacă un rol foarte important în cadrul acestor comportamente, astfel încât este nevoie de roboți care pot conversa cu ușurință cu oamenii, suportând aspecte importante legate de comunicarea non-verbală, cu scopul de a maximiza eficiența în comunicare și de a permite aplicarea lor rapidă și eficientă.

**Obiectivul principal** al cercetării îl reprezintă așadar studiul, concepția, dezvoltarea, implementarea și testarea unui sistem de interacțiune om-robot caracterizat de o interfață naturală și ușor de utilizat, noi metode de comandă și control realizate prin intermediul gesturilor umane și al comenzilor vocale, precum și combinații ale lor, astfel încât anumite sarcini din mediul casnic să poată fi realizate prin colaborarea dintre om și robot.

Sunt investigate modalităților prin care se poate obține un mod de interacțiune natural, apropiat de comportamentul uman, utilizând mijloace care se folosesc în comunicarea interpersonală și cât mai puține dispozitive care ar putea să împovăreze utilizatorul sau să îl facă să se simtă stingherit. Abordarea este multidisciplinară întrucât implică o serie de cunoștințe din domeniul *roboticii*, al *inteligenței artificiale*, al *ingineriei electrice și electronice*, al *ingineriei mecanice*, dar și al *psihologiei* sau *sociologiei*.

Conform Planului Național de Cercetare, Dezvoltare și Inovare, tema se încadrează în prioritatea principală numărul 1 (Tehnologia informației și comunicații), subprioritatea 1.4 (Inteligență artificială, robotică și sisteme autonome avansate), subpunctul 1.4.6 (Dezvoltarea de sisteme de interacțiune naturală om-calculator minimal dependente de universul discursului) și în prioritatea numărul 1 (Tehnologia informației și

comunicații), subprioritatea 1.3 (Tehnologii, sisteme și infrastructuri de comunicații), subpunctul 1.3.1 (Algoritmi, metode/tehnologii și sisteme de prelucrare în sistemele de comunicații a informației și semnalelor (voce, audio, video/ imagini, date, multimedia) în vederea producerii, prelucrării, transportului la distanță și livrării de conținut informațional).

Câteva dintre **obiectivele specifice**, care au rezultat din obiectivul principal, sunt următoarele:

- analiza bibliografică asupra cercetărilor teoretice și experimentale în domeniul interacțiunilor multimodale om-robot utilizate în medii domestice, precum și a tehnicilor de comunicare și interacțiune dintre oameni și roboți;
- conceperea și dezvoltarea unui sistem de interacțiune prin intermediul gesturilor corporale;
- conceperea și dezvoltarea unui sistem de interacțiune prin intermediul unor comenzi vocale;
- implementarea unui sistem de integrare a comenzilor vocale și gestuale;
- implementarea de algoritmi de detecție și recunoaștere;
- conceperea, dezvoltarea și implementarea unui sistem de control care este capabil să cunoască starea robotului;
- implementarea unei interfețe grafice cu utilizatorul prin intermediul căreia să poată fi folosite funcțiile sistemului propus;
- evaluarea calității și a preciziei algoritmilor dezvoltați și integrarea lor în sistemul global;
- testarea prin experimente a sistemului dezvoltat.

### **1.3. STRUCTURA LUCRĂRII**

Lucrarea este organizată în două părți: prima parte este o introducere a problemei, cu definirea ei, scopul și obiectivele, stadiul actual și premisele hardware și software ale cercetării; a doua parte a acestei cercetări reprezintă o descriere a contribuției cercetării, ilustrând algoritmi și metodele implementate, rezultatele experimentale și concluziile finale.

În *Capitolul 1* sunt prezentate aspecte generale legate de scopul și obiectivele cercetării. Astfel, în prima parte se arată necesitatea proiectării de noi interfețe de interacțiune om – robot care să utilizeze atât modalități



## CAPITOLUL 1

verbale, cât și non-verbale, integrate într-un sistem complex, care să contribuie la sprijinirea persoanelor în activitățile zilnice. Este exprimat astfel în acest capitol obiectivul principal al cercetării, precum și obiectivele secundare, care decurg din acesta, iar în final este redată structura lucrării, prezentându-se succint capitolele care o compun.

*Capitolul 2* prezintă aspecte generale privind sistemele de interacțiune om-robot utilizate în robotica de asistență, cu accent pe interacțiunea naturală. Sunt prezentate câteva moduri de definire a acestui concept de 'interacțiune naturală' în literatura de specialitate și se definesc apoi câteva trăsături pe care trebuie să le posede un sistem de interacțiune naturală. Se face o scurtă trecere în revistă a roboților mobili utilizați în robotica asistivă, cu istoria, definițiile și clasificarea lor. Termenul de *robotică asistivă* a fost folosit în această lucrare cu un sens mai larg, referindu-se în general la toți roboții care dau ajutor sau sprijină utilizatorii umani. Apoi sunt prezentate tipurile de senzori folosiți de către roboți în sarcinile pe care le efectuează, precum și a celor folosiți pentru urmărirea mișcărilor. Se prezintă, de asemenea, gesturile utilizate în general în aplicațiile de interacțiune și tehnicile principale de recunoaștere a lor. Sunt descrise apoi diverse implementări actuale ale sistemelor de interacțiune prin gesturi, voce și interacțiune multimodală legate de robotica asistivă. Capitolul se încheie cu delimitarea problematicilor actuale care necesită rezolvare, sunt trasate concluziile care se desprind de pe urma acestei analize și sunt stabilite obiectivele specifice cercetării.

În *Capitolul 3, Premise hardware și software pentru studiile teoretice și experimentale*, sunt prezentate la început componentele hardware ale sistemului de interacțiune propus, dispozitivul Kinect și robotul umanoid NAO, cu structura lor fizică și senzorii utilizați. Sunt descrise apoi aplicațiile și platformele software folosite pentru implementarea sistemului. Modelarea robotului într-un mediu virtual, deplasarea prin mediu, detecția și identificarea de obiecte sunt, de asemenea, descrise în mod succint în cadrul acestui capitol.

*Capitolul 4, Sistem de interacțiune naturală om - robot*, se concentrează pe descrierea metodelor și algoritmilor folosiți pentru recunoașterea comenzilor trimise robotului prin cele două mijloace de interacțiune verbale și non-verbale. Este prezentată paradigma de interacțiune „Point-and-command”, sarcinile pe care le va îndeplini robotul și răspunsul oferit de acesta în procesul de interacțiune. De asemenea, se oferă detalii despre

modul de funcționare prin imitare, mod prin care robotul reproduce mișcările efectuate de utilizator și poate învăța noi comportamente și abilități prin combinarea de diferite astfel de mișcări.

*Capitolul 5* cuprinde descrierea sistemului de luare a deciziilor al aplicației de interacțiune, care este format din mai multe subsisteme de decizie și de monitorizare a stării sistemului. Este prezentată tot în acest capitol și arhitectura finală a interfeței de interacțiune.

În *Capitolul 6* sunt prezentate experimentele efectuate utilizând interfața de interacțiune, în corespondență cu cercetările teoretice prezentate în capitolele anterioare. Experimentele s-au desfășurat în etape, pe parcursul evoluției sistemului și au fost realizate într-un mediu interior, utilizând platforma robotică dezvoltată de Aldebaran Robotics, NAO. Capitolul se încheie cu ilustrarea concluziilor care se desprind în urma testelor și cercetărilor experimentale, evidențiind și limitările care au fost observate pe parcursul acestor experimente.

*Capitolul 7* cuprinde concluziile, contribuțiile proprii, diseminarea rezultatelor științifice obținute în urma studiului (lista de lucrări publicate și participarea la conferințe internaționale) și direcțiile următoare de cercetare raportate la lucrarea prezentă.

## *SISTEME DE INTERACȚIUNE NATURALĂ OM-ROBOT ÎN ROBOTICA ASISTIVĂ*

În acest capitol se prezintă stadiul actual al interfețelor de interacțiune naturală om-robot utilizate în diferite aplicații. Capitolul descrie câteva din implementările actuale ale sistemelor de interacțiune cu roboți mobili de asistență, identificând anumite limitări ale acestora. Sunt prezentate, de asemenea, și realizări din robotica asistivă și roboți utilizați în acest domeniu.

Roboții de asistență pot fi dispozitive fizice de asistență, dezvoltate de obicei pentru reabilitare și care nu sunt interactive din punct de vedere social. Exemple de astfel de roboți sunt scaune cu roțile robotice inteligente sau membre artificiale și exoscheleți. În acest capitol vor fi abordați roboții sociali, care constituie un alt tip de roboți de asistență. Acești roboți sunt sisteme care pot fi percepute ca entități sociale care pot comunica cu utilizatorul. Există, totuși, și proiecte de roboți sociali care urmează să fie utilizați în scopuri de reabilitare.

Roboții de asistență dezvoltați pentru adulții în vârstă abordează amândouă aspectele menționate, fiind folosiți ca dispozitive independente sau ca o parte a mediilor inteligente. Dispozitive independente răspund nevoilor fizice și cognitive prin sprijinirea mobilității, a activităților de bază, oferind întreținere de uz casnic, monitorizare și menținerea siguranței. Ca parte a unui mediu inteligent, ele sunt de obicei prevăzute să funcționeze ca interfețe, facilitând comunicarea între mediu și utilizatorul său. Concentrându-se pe nevoile sociale, unele studii demonstrează cum roboții pot avea capacitatea de a oferi companie. Aceasta demonstrează cum pot roboții să anticipeze nevoile unui utilizator pentru a construi o relație emoțională cu acesta.

## CAPITOLUL 2

Implementările interfețelor actuale de interacțiune sunt prezentate succint în cadrul acestui capitol. Pe baza acestei analize se formulează la finalul capitolului obiectivele specifice cercetării.

### **2.1. INTRODUCERE**

Cuvântul ‘robot’ sugerează de obicei o mașină autonomă sau un braț robotic care efectuează activități industriale. Cuvântul este adesea asociat cu o întrupare antropomorfică care este capabilă să asiste sau chiar să înlocuiască omul în activitățile zilnice. Un robot inteligent este o mașină capabilă să extragă informații din mediu și să folosească cunoștințele despre lumea din jurul său pentru a se mișca în condiții de siguranță într-un mod autonom și cu un anumit scop.

Cercetările recente în interacțiunea om-robot se concentrează pe crearea de aplicații domestice, odată cu creșterea numărului de roboți personali de servicii, care au început să apară în casele sau birourile oamenilor. Roboții inteligenți oferă suportul lor în multe activități neplăcute, plictisitoare pentru om. Acești roboți trebuie să fie capabili să dobândească suficientă înțelegere a mediului, să fie conștienți de anumite situații, să detecteze și să urmărească oamenii, precum și să stabilească o comunicare reușită cu aceștia pentru a fi capabili să coopereze [Alvarez-Santos, 2013].

În ultimul deceniu cercetarea în domeniul roboticii a trecut de la sisteme robotice staționare la roboți mobili și orientați pe servicii. Pe baza progresului recent în domeniul algoritmilor robotici fundamentali, cum sunt cei de mapare, navigare sau percepție, roboții mobili devin din ce în ce mai pregătiți pentru a fi lansați ca asistenți în medii dificile. O aplicație promițătoare a roboților de servicii se referă la asistenții de zi cu zi din mediile domestice [Breuer, 2012].

### **2.2. INTERACȚIUNEA OM-ROBOT**

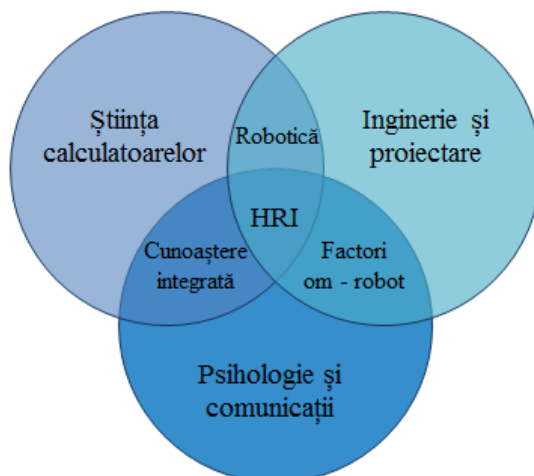
Interacțiunea om-robot este un domeniu multidisciplinar, având contribuții din domeniul interacțiunii om-calculator, inteligență artificială, robotică, înțelegerea limbajului natural și științe sociale.

Roboții sunt proiectați pentru a executa sarcini diverse, de la automatizări industriale la aplicații pentru prestări de servicii sau îngrijire medicală și divertisment. Deși roboții au fost utilizați inițial pentru sarcini repetitive, ei devin implicați în sarcini și activități din ce în ce mai complexe și mai puțin structurate, inclusiv interacțiunea cu oamenii necesară pentru a finaliza aceste sarcini. Această complexitate a determinat apariția unui domeniu de cercetare dedicat interacțiunii om-robot (HRI - Human Robot Interaction), care are ca obiectiv studiul modului în care oamenii interacționează cu roboții și cum este cel mai bine să se proiecteze și să se implementeze sisteme robotice capabile să îndeplinească sarcini interactive. Scopul fundamental al HRI este de a dezvolta principiile și algoritmi pentru sisteme robotice, pentru a le face capabile de interacțiune directă, sigură și eficientă cu oamenii [Feil-Seifer, 2009]. Multe aspecte ale cercetării HRI se referă sau se trag din perspectivele și principiile de psihologie, comunicare, antropologie, filozofie și etică, făcând din HRI un domeniu inerent interdisciplinar.

HRI este văzută în lucrarea [Heckel, 2008] ca un domeniu aflat la intersecția dintre inginerie, știința calculatoarelor și psihologie, iar lucrările de cercetare HRI implică cel puțin două dintre aceste domenii generale.

În figura 2.1 este prezentată diagrama Venn a interacțiunii HRI, așa cum este ea văzută de către autorii lucrării mai sus menționate, unde:

- *cunoașterea integrată* – se referă la sistemele de calcul care imită facilitățile cognitive și afective ale oamenilor;
- *factori om-robot* – se concentrează pe proiectarea sistemelor robotizate și înțelegerea modului în care oamenii răspund la ele;
- *robotica* – domeniu care se axează pe dezvoltarea de noi instrumente pentru platforme robotice, incluzând atât capacități hardware, cât și software.



**Fig. 2.1.** Diagrama Venn a HRI

Interacțiunea om-robot este diversă și ocupă o gamă largă de subiecte. Multe progrese au fost realizate în robotică în ultimul deceniu, iar cercetătorii au depus eforturi pentru a îmbunătăți comunicarea între oameni și roboți (și, de asemenea, între roboți), varietatea și fidelitatea senzorilor aflați pe robot, abilitatea acestuia de a traversa un teren accidentat și nivelul de autonomie pe care roboții sunt capabili să îl atingă [Thavasi, 2012]. Cu toate acestea, au fost realizate relativ puține progrese în optimizarea parteneriatului între oameni și roboți, prin tehnici îmbunătățite de interacțiune.

Interacțiunea om-robot îmbracă în prezent mai multe forme. Activități periculoase, cum ar fi de căutare și salvare sau curățare de materiale periculoase sau deminare [Zhang, 2011], deplasarea pe terenuri accidentate [Loc, 2011] necesită eliminarea operatorului uman din locația fizică a robotului. Roboții care asistă persoanele în vârstă [Wu, 2012], [Mohamed, 2012] sau cu handicap împart același spațiu fizic cu utilizatorii lor, de multe ori transportându-i prin mediul de lucru. Alții, cum ar fi Aibo [Fujita, 1998], oferă divertisment și companie pentru oameni sau sunt folosiți în activități de supraveghere [Lee, 2011], [Surmann, 2008], [Gunn, 2013].

Interacțiunea om-robot este un subdomeniu al interacțiunii om-calculator (HCI – Human-Computer Interaction). HCI a fost definită în multe feluri. Un exemplu este definiția utilizată de către grupul de dezvoltare ACM SIGCHI: *interacțiunea om-calculator este o disciplină care se ocupă cu proiectarea, evaluarea și implementarea unor sisteme de calcul interactive*

pentru uz uman și cu studiul fenomenelor majore care îi înconjoară [Hewett, 1992]. Deoarece roboții sunt sisteme de calcul concepute pentru a beneficia oamenii de către acestea, HRI poate fi susținută de cercetarea în domeniul HCI.

### 2.2.1. CLASIFICAREA INTERACȚIUNII OM-ROBOT

În lucrarea [Scholtz, 2003] se afirmă că interacțiunea om-robot diferă de interacțiunea om-calculator în patru dimensiuni, care pot fi considerate categorii de clasificare a interacțiunii om-robot. Cele patru dimensiuni sunt:

- tipurile de interacțiune (sau rolurile în interacțiune – supervisor, operator, coleg);
- natura fizică a roboților;
- caracterul dinamic al platformelor robotice;
- mediul în care apar interacțiunile.

*Cantitatea de intervenție* necesară pentru controlul unui robot este unul dintre factorii definitorii pentru interacțiunea om-robot. Există un spectru pentru controlul unui robot situat între teleoperare și autonomie deplină, iar nivelul de interacțiune om-robot măsurat prin cantitatea de intervenție necesară variază de-a lungul acestui spectru. Interacțiunea constantă este necesară la nivelul teleoperării, în cazul în care o persoană controlează de la distanță un robot. Cu cât robotul are o autonomie mai scăzută, cu atât este mai necesară intervenția umană. Nivelul de autonomie măsoară procentul de timp în care robotul îndeplinește misiunea sa pe cont propriu. Cantitatea de intervenție necesară măsoară procentul de timp în care un operator uman trebuie să controleze robotul. Roboții teleoperați sunt pe deplin controlați de către un operator aflat de obicei la distanță. La celălalt capăt al spectrului sunt roboții cu autonomie deplină. Exemple de acest tip de control pot fi găsite în multe aplicații, cum sunt roboții care ghidează vizitatorii sau roboții de livrare. Între aceste două puncte este un spectru larg de control al robotului, adesea numit control comun. Cu un control comun, roboții sunt în stare să realizeze o parte a sarcinii, iar operatorul uman trebuie să realizeze cealaltă parte a sarcinii.

Interacțiunea, prin definiție, necesită o comunicare între roboți și oameni. *Comunicarea* între un om și un robot poate lua mai multe forme, dar aceste forme sunt în mare măsură influențate de distanța spațială dintre

om și robot. Astfel, comunicarea și interacțiunea pot fi clasificate în două categorii generale [Goodrich, 2007]:

- interacțiunea de la distanță – operatorul uman și robotul sunt separați spațial sau chiar temporal;
- interacțiunea apropiată – oamenii și roboții sunt co-localizați (de exemplu, roboții pentru prestări de servicii pot fi în aceeași cameră cu oamenii).

În cadrul acestor categorii generale este util să se facă distincția între aplicații care necesită mobilitate, manipulare fizică sau interacțiune socială. Interacțiunea de la distanță cu roboți mobili este adesea menționată ca *teleoperare* sau control de supraveghere, iar interacțiunea de la distanță cu un manipulator fizic este adesea menționată ca *telemanipulare*. Interacțiune apropiată cu roboți mobili poate lua forma unui *robot de asistență* și poate include o interacțiune fizică.

*Interacțiunea socială* include aspecte sociale, emoționale și cognitive. În interacțiunea socială oamenii și roboții interacționează ca și colegi sau însoțitori. De aceea, interacțiunile sociale cu roboți par să fie mai degrabă apropiate decât de la distanță.

Căutarea și salvarea asistată de roboți este un domeniu în care roboții lucrează direct cu utilizatorii finali, care reprezintă în acest caz personalul de salvare special instruit. Roboții sunt de obicei echipați cu o cameră video și senzori de temperatură, eventual chimici și pot fi uneori echipați cu un manipulator. Scopul este de a ancheta rapid o zonă care altfel ar fi nesigură pentru un căutător uman și a aduna informații despre locația victimelor și stabilitatea structurală. Din cauza naturii inerent nestructurate a domeniilor de căutare și salvare, interacțiunile dintre om și robot trebuie să fie foarte bogate [Yanco, 2004].

*Sistemele robotice de asistență* caută să ofere suport fizic, mental sau suport social persoanelor, cum sunt persoanele în vârstă sau cu handicap. *Robotica de asistență* este importantă pentru HRI deoarece pune accentul pe interacțiunea cu oamenii. Provocările HRI din acest domeniu includ furnizarea contactului fizic în condiții de siguranță sau deplasarea în imediata apropiere, suportarea interacțiunilor sociale eficiente prin utilizarea de modele cognitive și emoționale și prin interacțiuni naturale, cum ar fi gesturi sau voce. Tipurile de roboți folosiți în aplicații de asistență variază foarte mult în aspectul lor fizic, incluzând scaune cu rotile, roboți mobili cu manipulator, roboți zoomorfi și umanoizi [Feil-Seifer, 2005].



Un domeniu important de cercetare, în special în țările cu o populație mare a persoanelor în vârstă, este oferirea de sprijin [Pollack, 2005], [Mohamed, 2012], [Leroux, 2013] sau supraveghere [Kaneko, 2014]. Cercetătorii în acest domeniu se concentrează pe asistența în mobilitate [Kubota, 2013], bunăstarea emoțională [Arkin, 2003], [Bemelmans, 2012] și asistență cognitivă [Fridin, 2014]. Multe lucrări în acest domeniu iau în calcul considerațiile etice care apar prin delegarea rolului de *companion* pentru robot [Lin, 2011] și de probleme de încredere care apar în a avea un agent artificial care lucrează cu o persoană cu probleme mentale cum ar fi demența [Mordoch, 2013]. Având în vedere capacitatea de a gestiona de la distanță un robot de asistență personal [van Osch, 2014], acesta ar putea fi un pas în direcția de abordare a acestor probleme etice.

Pentru unele persoane cu probleme fizice și psihice, roboții pot oferi o oportunitate de interacțiune și terapie. O astfel de activitate este în curs de explorare cu copii autiști [Billard, 2006], [Anzalone, 2014]. Mulți dintre acești copii răspund slab sau deloc la replicile sociale, dar răspund bine la dispozitive mecanice. Au fost evaluate și efectele comportamentelor roboților personalizați cu privire la exercitarea și motivația copiilor cu diabet zaharat [Blanson Henkemans, 2013]. Roboții oferă, de asemenea, un posibil rol terapeutic prin utilizarea unui dispozitiv mecanic pentru a îmbunătăți interacțiunile sociale [Kozima, 2007]. Dimensiunea socială a HRI este considerată necesară nu numai în roluri de asistență, dar și în multe domenii de interacțiune apropiată [Breazeal, 2003], [Dautenhahn, 2007], [Feil-Seifer, 2005], [Gockley, 2006], [Kanda, 2004].

Interacțiunea fizică cu roboți nu se limitează la acordarea de asistență pentru cei cu o anumită formă de handicap. În scopul de a oferi asistență pentru persoanele clinic sănătoase, care nu suferă de dizabilități, unii cercetători au investigat modul cum pot fi folosiți roboții pentru a promova educația copiilor, atât la domiciliu, cât și în școli [Keren, 2014]. A fost, de asemenea, investigat modul cum este perceput un robot de către copii [Beran, 2011], [Shahid, 2014]. *Roboții personali pentru prestări de servicii* reprezintă o altă categorie de roboți, utilizați pentru a furniza servicii pentru populația obișnuită.

Roboții Valerie și Tank de la Universitatea Carnegie Mellon au fost concepuți ca recepționeri [Gockley, 2006]. Alți roboți acționează în calitate de ghizi turistici în muzee [Chella, 2009] sau în laboratoare de cercetare [Boboc, 2014a]. Aceste aplicații ar putea fi considerate reprezentative

pentru robotica de asistență, dar sunt încadrate mai degrabă în categoria roboticii de divertisment [Siegwart, 2003a]. Alte studii HRI legate de utilizarea de roboți de divertisment includ povestitori robotizați [Fridin, 2014] sau animale de companie robotizate [Kerepesi, 2006].

Interacțiunea om-robot este centrată pe *interfața om-robot*, care variază foarte mult de la robot la robot [Fong, 2003a]. Având în vedere diversitatea, au fost propuse mai multe clasificări pentru a organiza ceea ce se cunoaște despre interacțiunea om-robot [Bartneck, 2001], [Goodrich, 2007], [Yanco, 2002], [Yanco, 2004]. De exemplu, în lucrarea [Bartneck, 2001] se argumentează că interfețele om-robot ar putea fi clasificate în funcție de patru dimensiuni: (a) jucării/ instrumente; (b) controlate de la distanță/ autonome; (c) interacțiunile inițiate de către utilizator/ interacțiuni inițiate de robot și (d) nivelul de antropomorfism.

De asemenea, în lucrarea [Yanco, 2004] se precizează că interacțiunea om - robot ar putea să fie clasificată pe 11 dimensiuni : (a) sarcina robotului; (b) caracterul critic al sarcinii robotului; (c) morfologia robotului; (d) raportarea oamenilor la roboți; (e) diferite tipuri de roboți în echipe om-robot; (f) structura de comandă a echipelor om-robot; (g) rolul robotului; (h) apropierea omului de robot; (i) informații furnizate de robot pentru a sprijini luarea deciziei de către utilizator; (j) locul unde au loc interacțiunile în termeni de timp și spațiu și (k) nivelul de autonomie al robotului.

Aplicațiile cu roboți de servicii autonomi pot fi împărțite în două grupe principale: în aer liber (de teren) și de interior. Roboții de detectare a minelor și de căutare și salvare, roverele pe Marte, automatele de cărat marfă, vehicule aeriene și subacvatice fără pilot sunt unele aplicații ale roboticii de teren. Termenul *robotică de interior* se aplică de obicei roboților mobili autonomi care se deplasează într-un mediu interior tipic populat. Aspiratoarele robotizate, roboții de divertisment și de companie sau aplicațiile de securitate și de supraveghere sunt, de asemenea, câteva exemple de succes pentru aplicații cu roboți de interior.

Interacțiunea om-robot este un domeniu în creștere de cercetare și de aplicare. Domeniul include multe aspecte și are potențialul de a produce soluții cu impact social pozitiv. Clasficările de mai sus au rolul de a organiza informațiile despre interacțiunea om-robot la un nivel global. În subcapitolul următor va fi introdus conceptul de interacțiune naturală.

### 2.2.2. INTERACȚIUNEA NATURALĂ

În comunicarea interpersonală gesturile, mișcările și expresiile sunt utilizate pentru schimbul de informații. Un aspect important pentru interacțiunea om-robot este de a aplica, de asemenea, acest concept, creând căi de comunicare cu roboții într-un mod natural și eficient. Sistemele robotizate de asistare, ca instrumente care oferă sprijin utilizatorului uman, necesită tehnologii care pot fi controlate prin comunicare naturală pentru asistarea oamenilor în activitățile de zi cu zi [Norman, 2009]. Pentru a facilita o *interacțiune naturală* (NI), roboții ar trebui să fie capabili să perceapă și să înțeleagă modalitățile folosite de oameni în timpul interacțiunii cu aceștia [Tapus, 2007]. O interacțiune naturală este definită ca o interacțiune care poate avea loc în mod inconștient, fără efort, prin care comunicarea în ambele direcții se face atât de natural, încât rezultatul este o fuziune armonioasă a persoanei și mașinii în realizare unei sarcini comune [Stiefelhagen, 2004]. Cu toate acestea există mai multe abordări pentru a defini acest concept, care vor fi menționate în continuare, pentru a găsi unele elemente comune.

Interacțiunea om-robot este un domeniu de studiu dedicat înțelegerii, proiectării, evaluării sistemelor robotizate pentru utilizarea de către/cu oamenii [Goodrich, 2007] și se află la intersecția mai multor domenii de cercetare, cum ar fi psihologia, științele cognitive, științele sociale, inteligența artificială, știința calculatoarelor, robotica, ingineria și interacțiunea om-calculator [Dautenhahn, 2007]. O provocare pentru proiectarea HRI este de a crea interfețe de interacțiune cu roboți de servicii, care funcționează în prezența oamenilor, în casele și birourile lor. Pentru acești roboți este important să fie în măsură să înțeleagă și să interpreteze comenzile umane, iar pentru utilizatorii lor, un lucru de dorit este să interacționeze și să comunice cu roboții într-un mod *natural*.

Ideea de a dezvolta interfețe de interacțiune naturală pentru a comunica cu sisteme inteligente nu este atât de nouă [Bolt, 1980], dar în ultima vreme există o preocupare tot mai mare pentru acest tip de interfețe, datorată în principal progreselor tehnologice, care oferă acum suport pentru interpretarea ușoară a mișcării utilizatorului sau pentru recunoașterea vocii. În acest fel, a apărut ideea de *interfață naturală cu utilizatorul* (NUI – natural user interface). Astăzi, o astfel de interfață este înțeleasă ca o alternativă la o *interfață grafică cu utilizatorul* (GUI – graphical user

interface), ca o tehnologie care recunoaște gesturi umane, voce și alte tipuri de 'intrări' ale omului, dar nu este foarte bine definită încă.

O definiție pentru NUI este dat în [Blake, 2011]: *O interfață cu utilizatorul naturală este o interfață cu utilizatorul proiectată să refolosească abilitățile existente pentru a interacționa în mod direct și plăcut cu tehnologia. În [Internet2.1], o interfață cu utilizatorul naturală este definită ca un sistem de interacțiune om-calculator cu ajutorul căruia utilizatorul operează prin acțiuni intuitive legate de comportamentul uman natural, de fiecare zi.*

A. Valli sugerează că oamenilor ar trebui să li se permită să interacționeze cu tehnologia în același mod în care sunt obișnuiți să interacționeze cu lumea reală în viața de zi cu zi, așa cum i-a învățat evoluția și educația [Valli, 2008]. El identifică două sensuri ale termenului natural: pe de o parte este opusul termenului „abstract” și, pe de altă parte, este sinonim cu „de obicei” - obișnuit, uzual în viața reală [Valli, 2005]. Cuvântul este folosit adesea alternativ cu „intuitiv”, dar acest termen este la fel de ambiguu și nu ajută la înțelegerea conceptelor [George, 2010].

De asemenea, unii autori au arătat că „natural” se referă la un obiectiv în experiența utilizatorului [Internet 2.2]. În acest sens, în lucrarea [Wigdor, 2011] se specifică: *noi percepem termenul <natural> ca referindu-se la modul în care utilizatorii interacționează cu produsul și simt despre produs, sau mai exact, ce anume fac și cum se simt în timp ce îl folosesc.* Astfel, în cazul în care experiența de utilizare a unui sistem se potrivește așteptărilor, astfel încât este întotdeauna clar pentru utilizator cum să procedeze și dacă doar câțiva pași (cu un minim de efort fizic și cognitiv) sunt necesari pentru a finaliza sarcinile comune [Hinckley, 2012], atunci această experiență poate fi numită „naturală”.

Din definițiile de mai sus, se concluzionează că există anumite cerințe pentru ca o interacțiune cu roboți sau cu alte dispozitive să fie naturală, cum sunt următoarele [Boboc, 2015a]:

- trebuie să fie ușor de utilizat, cu comenzi inspirate din lucrurile care sunt învățate în mod natural;
- trebuie să poată fi comparată cu un comportament uman obișnuit;
- trebuie să se adapteze la particularitățile și la abilitățile ființelor umane;
- trebuie să ofere utilizatorilor experiențe plăcute și interactive, cu metode mai bune de răspuns;

- trebuie să ofere utilizatorului posibilitatea de a se descurca fără nici o pregătire sau instruire preliminară complexă.

Acest concept de „interacțiune naturală” va fi exploatat prin dezvoltarea unui sistem de interacțiune cu un robot umanoid folosit pentru aplicații de asistență. Interacțiunile naturale au fost folosite în ultima vreme în multe aplicații, cum ar fi telemedicina [Tavares, 2013], controlul unui braț robotic [Postelnicu, 2011] sau al unui televizor [Boulabiar, 2011]. Roboții umanoizi au atras, de asemenea, atenția cercetătorilor din întreaga lume [Panfir, 2013], [Mu, 2010], deoarece aceștia pot servi ca parteneri sau asistenți pentru a efectua diferite tipuri de sarcini, care sunt neplăcute sau plictisitoare pentru oameni în viața lor de zi cu zi.

### **2.2.3. CONCLUZII**

Interacțiunea om-robot reprezintă așadar studiul interdisciplinar al dinamicii interacțiunii dintre oameni și roboți. Cercetători specializați în HRI provin dintr-o varietate de domenii, inclusiv inginerie, informatică, științe sociale și umaniste.

În acest capitol au fost prezentate câteva aspecte ale interacțiunii om-robot, luând în considerare nivelurile de interacțiune, aplicațiile din acest domeniu și rolurile pe care le pot avea roboții într-o astfel de situație.

### **2.3. ROBOȚI UTILIZAȚI ÎN ROBOTICA ASISTIVĂ**

Termenul de *robotică asistivă* se referea în trecut în special la roboții care asistau persoanele cu dizabilități prin interacțiune fizică. În prezent, termenul este folosit cu un sens mai larg, referindu-se în general la toți roboții care dau ajutor sau sprijină utilizatorii umani.

Roboții sunt obiecte de mare atracție și pot avea un număr mare de modalități de stabilire a unei interacțiuni cu oamenii. Prin urmare, nu este surprinzător faptul că una dintre primele aplicații ale roboților mobili au fost roboții folosiți pe post de ghizi în laboratoare de cercetare și mai târziu, în muzee și expoziții. Shakey, dezvoltat la începutul anilor 1980 este adesea considerat primul robot mobil [Miller, 2008]. Acesta a fost capabil să deplaseze blocuri de lemn în funcție de comenzi date prin intermediul unei tastaturi.

Oamenii au nevoie adesea de asistenți pentru a-i ajuta în activitățile de zi cu zi. Mulțumită roboților personali care au fost dezvoltați în ultimii ani, acest lucru a devenit mai ușor, dar cercetarea în acest domeniu este încă în desfășurare și există multe neajunsuri până când se va ajunge ca un robot să poată ajuta omul într-un mod sigur, eficient și pe deplin autonom. Aceste tipuri de roboți sunt în general numiți *roboți de asistență* și sunt destinați pentru a ajuta o persoană să trăiască mai bine, împărțind mediul cu operatorii lor umani [Ryu, 2006]. După cum se arată în studiul [Dautenhahn, 2005], există mai multe opinii favorabile ideii de robot de companie pentru mediile domestice și cele mai multe dintre ele preferă ca robotul să aibă rolul de *asistent*.

*Robotica de asistență* este un domeniu al roboticii care se ocupă cu utilizarea roboților mai degrabă ca instrumente decât ca sisteme autonome bazate pe sarcini. Roboții de asistență sunt definiți ca *dispozitive care cooperează cu un utilizator prin activitate fizică în mediul utilizatorului* [Meng, 2006]. Au existat numeroase încercări de a dezvolta roboți de asistență de-a lungul anilor, într-o varietate de aplicații, cum ar fi asistarea utilizatorilor într-o bibliotecă [Ramos-Garijo, 2003], asistarea astronauților în expediții [Burrige, 2003], aplicații de transport în spitale [Takahashi, 2010] sau de asistență pentru persoane în vârstă [Quan, 2011]. Conform cu [Van der Loos, 2008], roboții de asistență ar putea avea trei scopuri: manipulare, permiterea/sporirea mobilității oamenilor și scopuri cognitive.

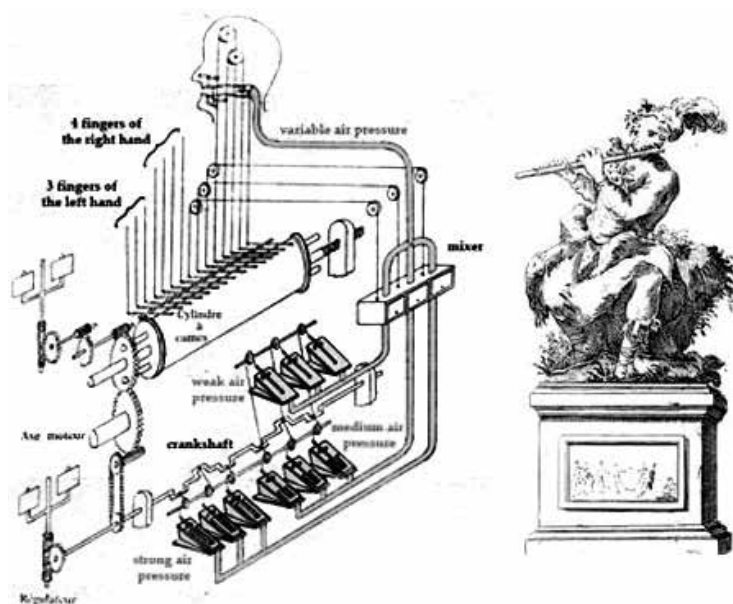
### 2.3.1. SCURTĂ ISTORIE A ROBOȚILOR

Conceptul de *robot* are o istorie foarte lungă, începând din mitologie și folclor, iar primii predecesori mecanici au fost construiți încă din antichitate [Breazeal, 2005], [Mavridis, 2015]. Deși au mai fost unele schițe, precum cele ale lui Leonardo da Vinci din secolul al XV-lea [Rosheim, 2006], primele mașini animate pot fi găsite la începutul secolului al XVIII-lea, printre care faimoasa rață mecanică a inventatorului francez Jacques de Vaucanson, care putea să dea din aripi, să imite jocul de picioare al unei rațe adevărate, să mănânce și chiar să digere cerealele [Iztueta, 2011].

Cuvântul *robot*, așa cum este folosit astăzi, a apărut pentru prima dată în piesa de teatru Rossum's Universal Robots (1921) a scriitorului ceh Karel Čapek pentru a desemna municitori automați de origine organică [Roberts, 2002]. În anul 1941, Isaac Asimov a folosit termenul de *robotică* și a

examinat pentru prima dată conceptele fundamentale ale interacțiunii om-robot, cel mai vizibil în cartea sa *I, Robot* [Feil-Seifer, 2009]. În același an, Grey Walter, considerat un pionier al inteligenței artificiale [Holland, 1997], a publicat o lucrare în care descrie construcția primilor roboți electronici autonomi, descrise ca „broaște țestoase” datorită formei lor și a ritmului lent de mișcare [Internet 2.3].

În perioada modernă tendința multor producători a fost de a umaniza forma roboților, urmând ambiția mai veche a roboticii, aceea de a construi un om artificial, cum este de exemplu cântărețul la flaut construit de Jacques de Vaucanson în 1737 [Boden, 2006]. Acest robot a fost capabil să cânte 12 melodii diferite. Vaucanson a ajuns la aceste sunete prin imitarea fiecărui mușchi implicat în activitatea de cântat. Burduful din interiorul robotului produce un flux variabil de aer. Buzele mecanice se puteau deschide, închide și mișca înainte sau înapoi. O limbă metalică mobilă localizată în interiorul gurii reglementează fluxul de aer și creează pauze (Fig. 2.2).



**Fig. 2.2.** Cântărețul la flaut construit de inventatorul Jacques de Vaucanson [Internet2.4]

După acest moment au fost numeroase încercări. În anul 2001 este produs AIBO, primul robot dotat cu inteligență artificială. În anul 2007 a urmat robotul umanoid NAO produs de Aldebaran Robotics.

## CAPITOLUL 2

Roboții oferă beneficii substanțiale muncitorilor, industriilor și implicit țărilor. În situația folosirii în scopuri pașnice, roboții pot influența pozitiv calitatea vieții oamenilor prin înlocuirea acestora în spații periculoase, cu condiții de mediu dăunătoare omului, cu condiții necunoscute de exploatare etc.

Domeniul roboticii se schimbă într-un ritm rapid. În prezent cei mai mulți roboți operează în medii industriale, unde efectuează sarcini cum ar fi asamblare și transport. Echipați cu sisteme minime de senzori și de calcul, roboții îndeplinesc aceeași sarcină repetitivă de multe ori. În viitor roboții vor oferi servicii direct nouă, oamenilor, la locurile noastre de muncă și în casele noastre.

*Robotica serviciilor* este în cea mai mare parte în fază incipientă, dar domeniul crește într-un ritm mult mai rapid decât robotica industrială. La fel ca și roboții industriali, roboții de servicii profesionale manipulează obiecte și navighează în mediile înconjurătoare. Cu toate acestea, roboții de servicii profesionale ajută oamenii în urmărirea obiectivelor lor profesionale, în mare parte în afara situațiilor industriale.

### 2.3.2. DEFINIȚII ALE ROBOȚILOR

Organizația Internațională de Standardizare (ISO) definește un robot ca *un manipulator reprogramabil controlat automat, multifuncțional, cu trei sau mai multe axe*. Institutul de Robotică din America desemnează un robot ca *un manipulator reprogramabil, multifuncțional proiectat pentru a muta materiale, piese, instrumente sau dispozitive specializate prin diverse mișcări programate pentru a executa o varietate de sarcini*. O definiție mai inspirată este oferită de dicționarul online Merriam Webster, unde se afirmă că un robot este *o mașină care arată ca o ființă umană și realizează diverse acțiuni complexe (ca mersul pe jos sau vorbitul) ale unei ființe umane*.

Se poate spune că nu există nicio definiție precisă, ceea ce l-a determinat și pe Joseph Engelberger, părintele roboților industriali, să afirme: *Nu pot defini un robot, dar când văd unul îl recunosc* [Carlisle, 2000]. Cu toate acestea o idee generală care se poate desprinde este că un robot este o mașină programabilă care imită acțiunile sau aspectul unei creaturi inteligente, de obicei un om.

Pentru a fi calificată drept robot, o mașină trebuie să fie capabilă să primească informații din mediul său înconjurător și să realizeze o activitate



fizică, cum ar fi mutarea sau manipularea obiectelor. Ei pot fi propulsați de roți sau șenile, se pot târî precum șerpii sau pot avea picioare; pot lucra în laboratoare, birouri sau muzee, pot acționa în spațiul cosmic sau înota în adâncul mării. Roboții sunt concepuți pentru a realiza activități murdare, plictisitoare și periculoase, și, mai recent, pentru a distra și a oferi divertisment oamenilor.

Cei mai mulți roboți utilizați în zilele noastre sunt roboții industriali. Ei lucrează în fabricile unde se produc bunuri în cantități mari și au o calitate excelentă la costuri reduse. Ei sunt utili și reprezintă baza pentru o mare parte din prosperitatea de care ne bucurăm în părțile industrializate ale lumii.

Potrivit Asociației Industriilor Robotice, un robot industrial este *un manipulator controlat automat, reprogramabil, multifuncțional programabil, în trei sau mai multe axe care pot fi fie fixe sau mobile, pentru utilizarea în aplicații de automatizări industriale*. Primul robot industrial fabricat a fost numit Unimate și a fost instalat de către General Motors în 1961.

Deși marea majoritate a roboților industriali sunt utilizați pentru fabricarea de produse, progresele tehnologice din ultimii ani permit, de asemenea, roboților să automatizeze multe sarcini în industriile nemanufacturere, cum ar fi agricultura, construcții, îngrijirea sănătății, altele. Acești așa-numiți *roboți pentru prestări de servicii* sau simplu, roboți de servicii, vizează sectorul în creștere rapidă al serviciilor și promet să fie un produs-cheie pentru următoarele decenii. Acești noi roboți de servicii nu sunt concepuți pentru a produce bunuri, ci pentru a oferi servicii utile la locul de muncă sau la domiciliu sau pur și simplu pentru a distra și a oferi confort oamenilor.

În viitor am putea vedea *roboți personali* care există în casele oamenilor sau la locurile de muncă și combină trăsăturile de tovarăși de prietenie, asistenți competenți și funcționari utili. Astfel de roboți vor necesita un nivel de inteligență, robustețe, adaptabilitate și fiabilitate. În timp ce servitorii robotici prezenți există numai în formă de prototipuri în câteva laboratoare de cercetare, se așteaptă ca ei să devină în viitor la fel de omniprezenți cum sunt PC-urile în momentul de față [Graefe, 2009].

Până în prezent sunt cunoscute trei tipuri mai importante de roboți personali. Un tip cuprinde roboții care sunt destinați să facă oamenii să se simtă fericiți, confortabil sau mai puțin singuri sau, în general, să îi afecteze emoțional; acești roboți, de obicei, nu pot și nu este nevoie să facă tot ce este

util într-un sens practic. Ei pot fi considerați animale de companie artificiale sau - în viitor - chiar și tovarăși. Prin urmare, ei mai sunt numiți animale de companie robotice personale sau companioni. Cel mai renumit este AIBO, vândut în număr mare de către compania Sony între 1999-2006. Cântărind aproximativ 2 kg, se aseamănă la aspect și prin unele dintre comportamentele sale cu un câine în miniatură. Un alt tip de roboți personali sunt cei destinați să efectueze sarcini domestice, inclusiv securitatea și supravegherea locuinței. Al treilea tip sunt cei pentru asistența persoanelor în vârstă sau cu handicap.

### 2.3.3. CLASIFICAREA ROBOȚILOR ASISTIVI

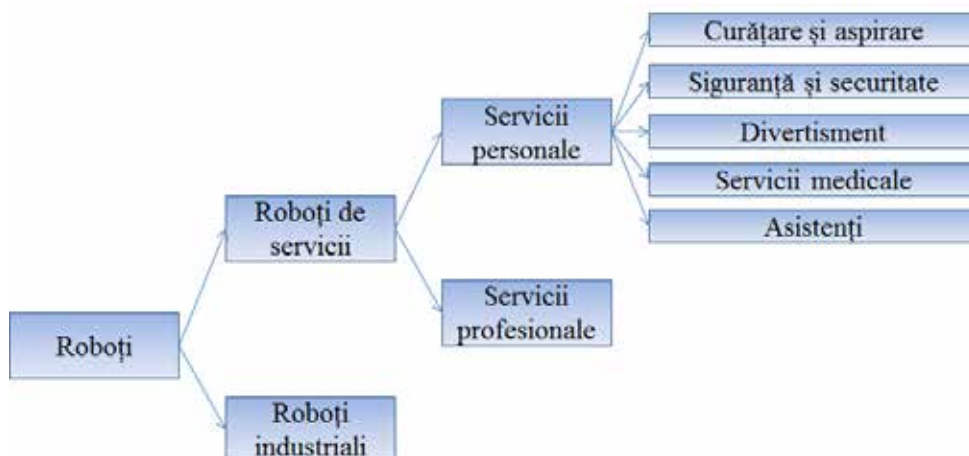
În Fig. 2.3 se prezintă o clasificare generală a roboților. După cum se poate observa și a fost amintit în subcapitolul precedent, ei se împart în două categorii principale: industriali și de servicii [Starețu, 2007].

Federația Internațională a Roboticii definește în felul următor roboții de servicii:

Un *robot de servicii* este un robot care îndeplinește sarcini utile pentru oameni sau echipamente, cu excepția aplicațiilor de automatizări industriale.

Un *robot de servicii personal* sau un robot de servicii pentru uz personal este un robot de servicii folosit pentru o sarcină necomercială, de obicei, de către nespecialiști. Exemple de astfel de roboți sunt asistenții domestici robotizați, scaune cu roțile automatizate, asistenți pentru persoanele cu probleme de mobilitate și animalele de companie robotizate.

Un *robot de servicii profesionale* sau un robot de servicii pentru uz profesional este un robot de servicii folosit pentru o sarcină comercială, fiind operat de obicei de către un operator instruit în mod corespunzător. Exemple de astfel de roboți sunt: roboții de curățare a locurilor publice, roboți de livrare în birouri sau spitale, roboți de stingere a incendiilor, roboți pentru reabilitare și asistenți pentru chirurgie în spitale. În acest context, un operator este o persoană desemnată pentru a începe, a monitoriza și a opri operarea prevăzută a unui robot sau a unui sistem robotic [Internet 2.5].



**Fig. 2.3.** Clasificarea roboților

Până de curând roboții s-au limitat la medii industriale, unde manipolatoare precise au fost dezvoltate pentru a automatiza activitățile plictisitoare, murdare sau periculoase ale oamenilor. „Comunicarea” cu roboții s-a limitat doar la pornirea sau oprirea robotului. Recent acest lucru s-a schimbat. Ultimele două decenii au cunoscut o creștere rapidă în cercetare și utilizarea efectivă a roboților în multe domenii, cum ar fi: asistență la domiciliu și îngrijirea persoanelor în vârstă și a altor categorii de oameni; reabilitare în terapii fizice, cum ar fi accident vascular cerebral, paralizie cerebrală, scleroza multiplă, leziunile măduvei spinării și boala Parkinson; educație pentru dobândirea de cunoștințe generale și dezvoltarea competențelor sociale pentru copiii cu autism; activități de căutare și salvare, cercetare și inovare.

Colaborarea și co-existența între un om și un asistent robotic și problemele inerente în ceea ce privește comunicarea, interacțiunea și colaborarea se pot împărți în două categorii: asistenți în mediile de producție și roboți pentru mediul domestic.

A doua categorie de roboți este reprezentată de roboții pentru menaj și îngrijire a casei, care operează parțial într-un mod autonom și parțial în strânsă colaborare și interacțiune cu omul, în scopul de a îndeplini îndatoririle sale de menaj. Acestea includ sarcini de bază „fetch and carry”, aranjarea mesei, încărcare și descărcare sau sarcini de bază de curățare. O interacțiune naturală, prietenoasă și comunicarea între om și robotul-asistent este de importanță centrală în acest context. Capacitatea de a interacționa cu un utilizator uman oferă, în plus, sistemului robotic

posibilitatea de a face uz de ghidarea umană și sprijinul său pentru extinderea competențelor sale inițiale.

### **2.3.4. EXEMPLE DE ROBOȚI ASISTIVI**

Procesul de îmbătrânire este adesea asociat cu o nevoie crescută pentru asistență fizică și cognitivă: oamenii în vârstă devin cu timpul mai puțin apti fizic și plătânzi, câteodată suferind cu probleme de memorie și percepție. În acest caz robotica avansată este văzută ca o soluție pentru a sprijini traiul independent al oamenilor în vârstă. Progresul recent în câteva domenii științifice și tehnice (știința calculatoarelor, automatică, proiectarea senzorilor și actuatorilor, materiale inteligente) permite translatarea roboților din mediile industriale în mediile domestice, beneficiind de capacitățile și de potențialul ridicat al roboților de a oferi asistență persoanelor. Până acum au existat numeroase proiecte care au demonstrat că sistemele robotice pot acționa autonom în medii cum sunt cele domestice, interacționând verbal și la nivel fizic cu utilizatorii finali. Acești roboți sunt în măsură să ofere sprijin în activitățile obișnuite de zi cu zi, cum ar fi mâncatul, spălatul, îmbrăcatul, folosirea toaletei, pentru mobilitate prin asistență locomotorie și pentru reabilitare.

Cu toate acestea foarte puține soluții robotice au fost cu adevărat acceptate de către cumpărători și pot fi considerate ca un succes comercial [Wu, 2012]. Deși există roboți care pot oferi o gamă largă de funcționalități, nu este destul pentru a face ca sistemele robotice să fie acceptate de către persoanele nespecializate, care caută caracteristici suplimentare pentru acest tip de sisteme. Mulți roboți au fost construiți fără o analiză atentă a aspectelor sociale, estetice, emoționale și culturale, care sunt importante pentru oamenii în vârstă. Pentru o interacțiune socială completă cu utilizatorii, roboții trebuie să fie dotați cu anumite caracteristici suplimentare, cum ar fi înfățișare, emoții, dialog, personalitate, percepție orientată pe om, învățare socială [Breazeal, 2003].

Mulți factori influențează acceptarea și utilizarea unui robot, în primul rând funcția robotului, aspectul și capacitatea socială [Beer, 2011]. În al doilea rând, modul în care ar trebui să interacționeze oamenii cu roboții. Mulți au raportat că doresc să interacționeze cu un robot prin comenzi vocale sau ca acesta să fie preprogramat [Ezer, 2008]. În al treilea rând, un

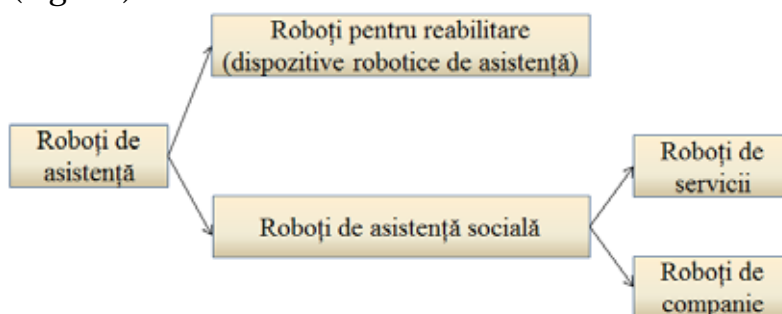
robot trebuie să fie în măsură să ofere sprijin în funcție de capacitățile utilizatorului.

*Robotica de asistență socială (SAR)* a fost prezentată ca răspunsul promițător pentru aceste probleme. SAR este întemeiată pe un punct de vedere orientat pe utilizator, în care aspectele legate de utilitate și acceptare devin predominante.

Robotica de asistență, în general, și robotica de asistență socială, în special, au potențialul de a îmbunătăți calitatea vieții pentru populații largi de utilizatori: persoane în vârstă, persoane cu deficiențe fizice și persoanele aflate în terapie de reabilitare, persoanele cu dizabilități cognitive și de dezvoltare și tulburări sociale [Tapus, 2007].

Persoana în vârstă nu este considerată numai ca un utilizator final, ci și un partener în procesul de proiectare și dezvoltare a sistemului robotic. Proiectele de cercetare includ o fază de specificare a cerințelor, în care utilizatorul final are un rol important când definește aspecte precum interfața cu utilizatorul, caracteristici fizice și funcționalități necesare. De-a lungul ultimului deceniu au fost dezvoltați un număr mare de roboți de servicii, servind pentru diferite scopuri și fiind destinați diferiților utilizatori țintă. Mulți dintre acești roboți sunt încă în stadiul de prototip, în timp ce alții au cucerit deja piața cu diferite niveluri de succes, în funcție de funcționalitatea lor, de cost și de clienții țintă.

Roboții de asistență pot fi clasificați în două categorii: roboți pentru reabilitare și roboți de asistență socială [Broekens, 2009]. Categoria roboților de asistență socială, la rândul ei, cuprinde roboți de servicii și roboți de companie (Fig. 2.4).



**Fig. 2.4.** Clasificarea roboților de asistență

Roboții de servicii au funcții care sprijină activitățile de zi cu zi, în plus față de funcțiile sociale. Aceste funcții sociale au fost proiectate pentru a facilita interacțiunea unei persoane cu robotul. Roboții de companie au fost

creați pentru a spori bunăstarea cognitivă și sănătatea. Ambele tipuri de roboți de asistență socială s-au dovedit a fi benefici pentru adulții în vârstă prin creșterea stării de spirit pozitive, scădere sentimentelor de singurătate, atenuarea stresului și creșterea legăturilor sociale.

Provocările viitoare în proiectarea de roboți asistenți includ abordarea modului în care utilizatorii îi pot învăța noi cunoștințe și sarcini, standarde de siguranță pentru roboți, probleme de confidențialitate, costuri, beneficii de a deține un robot, metode de instruire a adulților în vârstă de a folosi un robot și fezabilitatea pentru un robot de a funcționa în mediul casnic al persoanei (manevrare, perceperea obiectelor într-un mediu aglomerat).

Ideea centrală a multor cercetători este de a dota asistenții robotici inteligenți cu mecanisme puternice și versatile, care permit acestor roboți să comunice, să interacționeze și să colaboreze cu utilizatorii umani într-un mod natural și intuitiv. Aceste mecanisme trebuie să faciliteze predarea intuitivă, programarea și comanda asistenților robotici și să le permită să execute sarcini dificile și complexe sub controlul și în colaborare cu utilizatorul uman. Pentru ca asistenții robotici să acționeze, să se comporte și să comunice într-un mod asemănător oamenilor, aceste mecanisme trebuie să ia în considerare atât forma și mobilitatea corpului uman, precum și performanța și versatilitatea simțurilor umane.

Comunicarea dintre om și robot trebuie să fie prietenoasă și să implice toate simțurile umane și canalele de comunicare, cum ar fi voce, privire și înțelegere a comunicării vizuale (gesturi, mimică), de exemplu, prin gesturi și mimică sau de detectare și înțelegere a forțelor (haptic).

### ***2.3.4.1. Roboți pentru prestări de servicii***

Roboții pentru prestări de servicii apar, de asemenea, în calitate de ghizi interactivi, de exemplu în muzee [Siegwart, 2003b]. Interacțiunea naturală poate avea loc prin voce, gesturi și feedback emoțional (de exemplu, prin expresii faciale și limbajul corpului), care sunt posibile cu ajutorul unui asistent robotic în mod corespunzător. Putem distinge diferite medii inteligente centrate pe om care ar putea beneficia de un asistent personal robotic: mediul domestic, mediul profesional și mediul public.

Robotul din mediul domestic este o personificare a mediului inteligent. El poate încorpora o mulțime de funcționalități: poate fi o interfață pentru

controlul mediului, poate fi conectat la Internet, astfel încât să ofere informații unei persoane aflată la distanță despre ceea ce se petrece în locuința sa, poate fi un tovarăș de joacă interactiv, permițând o interacțiune intuitivă și emoțională.

Roboții din mediul public pot fi utilizați ca ghizi interactivi într-o clădire publică, de exemplu într-un muzeu, într-o sală de expoziții, într-o universitate sau chiar recepționează într-un hotel sau într-o companie.

Pentru mediul profesional, robotul poate avea rolul de asistent inteligent, fiind dotat cu camere video, microfoane, o interfață cu utilizatorul și un ecran, astfel încât să creeze o senzație de teleprezență.

#### *2.3.4.2. Roboți de companie*

Un robot autonom de companie ar putea fi privit ca un tip special de robot de servicii care este conceput special pentru uz personal, la domiciliu. Roboții de companie pot comunica cu utilizatorii neexperimentați într-un mod natural și intuitiv. Roboții concepuți pentru mediul domestic sunt o industrie în creștere atât din perspectiva cercetării, cât și din punct de vedere comercial.

Cercetările în interacțiunea om-robot sunt încă relativ noi în comparație cu robotica de servicii tradițională. Cu toate acestea, din ce în ce mai mulți roboți sunt menși să se angajeze în interacțiunea socială. Roboții de companie în medii casnice ar trebui în mod ideal să ofere posibilitatea de a efectua o gamă largă de activități, inclusiv funcții educaționale, de securitate, divertisment și servicii de livrare a mesajelor, etc. În prezent nu există însă roboți care să fie capabili să efectueze o combinație a acestor sarcini în mod eficient, precis și robust.

În prezent marea majoritate a roboților de servicii este proiectată pentru aplicații industriale sau militare. Dar roboții personali care îndeplinesc sarcini simple încep să apară în viața noastră de zi cu zi. Deși un număr mare de aplicații ale roboticii personale este în faza de identificare sau experimentală (roboți de companie la domiciliu, roboți de predare, etc), mulți oameni cred că următoarea revoluție industrială este robotica personală. Totuși, problemele tehnologice și științifice sunt numeroase, iar cheia pentru robotica personală este în interacțiunea om-robot și capacitatea roboților de a adapta deciziile lor la activitățile umane.

## *Roboți sociali*

O altă categorie demnă de luat în seamă este aceea a așa-numiților roboți sociali. Roboții sociali sunt proiectați pentru a interacționa cu oamenii într-un mod natural, interpersonal - de multe ori pentru a atinge obiectivele social-emoționale în diverse aplicații, cum ar fi educație, sănătate, calitatea vieții, divertisment, comunicare și colaborare. Obiectivul pe termen lung de a crea roboți sociali, care sunt parteneri competenți și capabili pentru persoane este o sarcină provocatoare. Ei vor trebui să fie în măsură să comunice în mod natural cu oamenii folosind atât modalități verbale, cât și non-verbale. Ei vor trebui să se implice nu numai la un nivel cognitiv, dar și la un nivel emoțional și vor avea nevoie de o gamă largă de abilități social-cognitive și o teorie a minții pentru a înțelege comportamentul uman și să fie înțeleși în mod intuitiv de către oameni. O înțelegere profundă a inteligenței umane și a comportamentului pe mai multe dimensiuni (de exemplu, cognitiv, afectiv, fizic, social, etc) este necesară pentru a proiecta roboți care pot juca cu succes un rol benefic în viața de zi cu zi a oamenilor. Acest lucru necesită o abordare multidisciplinară în care proiectarea de tehnologii și metodologii pentru roboții sociali să fie ajutată de robotică, inteligență artificială, neuroștiință, factori umani, antropologie și altele.

Modul în care o persoană interacționează cu un robot social este destul de diferit de interacțiunea cu majoritatea roboților mobili autonomi. Roboții autonomi moderni sunt în general priviți ca instrumente pe care specialiștii le utilizează pentru a efectua sarcini periculoase în medii aflate la distanță (de exemplu, curățarea terenurilor minate, inspectarea puțurilor de petrol, etc). În contrast cu aceștia, roboții sociali sunt concepuți pentru relații interpersonale, de multe ori în calitate de parteneri, în scopul de a atinge obiectivele sociale sau emoționale. Dezvoltarea de roboți inteligenți sociali și abili social conduce cercetarea spre a dezvolta roboți autonomi sau semiautonomi care oferă interfețe naturale și intuitive pentru publicul larg de a interacționa, comunica, lucra ca parteneri și învăța noi capacități.

Roboții sociali exploatează diferite modalități de a comunica și de a exprima comportamentul social-emoțional. Acestea includ mișcarea întregului corp, distanța interpersonală, gesturi, expresii faciale, privirea, orientarea capului, vocalizarea lingvistică sau emotivă, comunicarea prin atingere. Roboții sociali trebuie să fie, de asemenea, capabili să perceapă, să



interpreteze și să răspundă adecvat la replici verbale și non-verbale ale oamenilor.

De-a lungul timpului au fost dezvoltati numeroși roboți umanoizi sociali interactivi, care pot participa la interacțiune socială cu oamenii, cum ar fi dansul [Tanaka, 2006], mersul mână-în-mână [Lim, 2006] sau cântatul unui duet muzical [Chida, 2004]. Brațele și mâinile lor sunt concepute pentru a expune gesturi umane [Miwa, 2004]. Unii dintre ei sunt proiectați cu fețe mecanice pentru a comunica cu oamenii prin expresii faciale și, de asemenea, pot recunoaște expresiile faciale ale oamenilor, în acest scop fiind dezvoltati diferiți algoritmi, cum este cel din lucrarea [Banu, 2012].

Întrucât mulți dintre acești umanoizi au un aspect mecanic, roboții androizi sunt concepuți pentru a avea un aspect foarte asemănător oamenilor, cu piele, dinți, păr și haine. O provocare de proiectare pentru roboții androizi este de a evita abisul straniei (uncanny valley) în care aspectul și mișcarea robotului seamănă mai mult cu un cadavru animat decât cu un om viu. Modele care se încadrează în „valea stranie” provoacă o reacție negativă puternică oamenilor.



**Fig. 2.5.** Roboți sociali și umanoizi: a) NAO; b) ASIMO; c) TORO; d) Leonardo; e) PARO; f) AIBO; g) KISMET; h) DARwIn

Există o serie de mai mulți roboți sociali care și-au luat inspirația lor estetică și comportamentală de la animale. Având în vedere că oamenii îndrăgesc și dezmiardă animalele de companie, comunicarea prin atingere a fost exploatată în mai mulți dintre acești roboți inspirați din lumea animală. Câinele robotic de divertisment dezvoltat de Sony, AIBO [Fujita, 2004] (Fig. 2.5f) este un exemplu comercial bine-cunoscut. Alți roboți din

această categorie au un aspect mai organic, cum ar fi foca robotică de companie terapeutică, Paro [Wada, 2006] (Fig. 2.5e). Cercetătorii au ales, de asemenea, să proiecteze roboți cu un aspect mult mai fantezist, contopind calități antropomorfe cu calități de animale, cum ar fi Leonardo [Breazeal, 2004], [Sidner, 2005] (Fig. 2.5d), sau doar aspecte antropomorfe, cum este capul robotic Kismet [Breazeal, 2003] sau roboții umanoizi (Fig. 2.5a, 2.5b, 2.5c, 2.5h).

Pentru a satisface cerințele utilizatorilor, roboții de servicii personali trebuie să aibă un grad ridicat de autonomie. Prin urmare, ei trebuie să fie capabili să inițieze o interacțiune cu utilizatorii lor mai degrabă decât pur și simplu să aștepte comenzi. De exemplu, un robot care se află în repaus ar putea decide dacă utilizatorul său are nevoie de asistență. Pentru a face acest lucru trebuie să deducă ceea ce utilizatorul intenționează să facă și dacă acesta solicită asistență. Apoi, trebuie să decidă când și cum să interacționeze cu utilizatorul.

### **2.3.5. CONCLUZII**

Industria roboticii de servicii a atins în ultimul timp progrese în mobilitate, percepție și progrese algoritmice care permit roboților să se localizeze într-o hartă bidimensională a lumii și să întocmească harta unui mediu necunoscut cu două dimensiuni. Roboții de curățare folosesc algoritmi foarte simpli pentru a mapa intrări senzoriale și pentru a controla comenzile și acoperă o suprafață bidimensională, evitând în același timp obstacole. Roboții de securitate sunt capabili de a utiliza informații senzoriale pentru a deduce poziția lor într-o lume bidimensională și trimit imagini ale mediului înapoi la un operator uman aflat la distanță. Roboții sunt în măsură să ofere sprijin logistic în medii industriale și la birouri, transportând materiale (pachete, medicamente sau mărfuri) sau conducând vizitatorii pe holuri. Roboții controlați și monitorizați de la distanță sunt, de asemenea, capabili de a intra în medii periculoase sau neplăcute. Exemplele includ vehicule subacvatice comandate de la distanță, roboți de curățare a conductelor și roboți care dezamorsează bombe.

Provocările în robotica de servicii personale include toate provocările prezente în robotica industrială. Mobilitatea este o provocare cheie pentru robotica de servicii. Actuala generație de roboți este capabilă doar să funcționeze în medii bidimensionale, de interior. Deoarece roboții de

servicii trebuie să fie mobili, există provocări pentru proiectarea de roboți care pot fi încărcăți cu propria sursă de alimentare. În cele din urmă, roboții de servicii, în special roboții personali, vor opera în apropierea utilizatorilor umani. Siguranța este extrem de importantă. Pentru că interacțiunea cu utilizatorii umani este foarte importantă în robotica de servicii, este clar că industria are nevoie să depășească provocările semnificative în interfețele om-robot.

## **2.4. PERCEPȚIA ASUPRA MEDIULUI ÎNCONJURĂTOR**

O bună percepție a roboților asupra mediului înconjurător este o parte a ingredientelor critice pentru navigația autonomă în condiții de siguranță. Este foarte important ca robotul să perceapă și să interpreteze mediul său imediat. Acest lucru se realizează prin efectuarea de măsurători folosind diferiți senzori și extrăgând apoi informații semnificative din aceste măsurători.

Pentru a interacționa în mod semnificativ cu oamenii, roboții asistivi trebuie să fie capabili să perceapă mediul înconjurător în același mod ca și oamenii, simțind și interpretând aceleași fenomene pe care le observă oamenii. Acest lucru înseamnă că, în plus față de percepția necesară pentru funcțiile convenționale (localizare, navigație, evitare de obstacole), roboții asistivi trebuie să posede și abilități perceptivă similare cu cele ale omului.

Percepția roboților asistivi trebuie să fie optimizată astfel încât să fie în măsură să urmărească caracteristicile umane (fețe, membre, mâini), să interpreteze vocea umană, inclusiv vorbirea afectivă, comenzile discrete și limbajul natural. De multe ori, ei trebuie să aibă capacitatea de a recunoaște expresiile faciale, gesturile și activitatea umană.

### **2.4.1. TIPURI DE PERCEPȚIE**

Există o mare varietate de senzori utilizați în construcția roboților mobili. Unii senzori sunt utilizați pentru a măsura valori simple, cum ar fi temperatura internă a componentelor electronice ale unui robot sau viteza de rotație a motoarelor. Alți senzori mai sofisticăți pot fi folosiți pentru a obține informații despre mediul robotului sau chiar pentru a măsura direct poziția globală a unui robot [Siegwart, 2011].

Senzorii pot fi clasificați folosind două axe funcționale importante: proprioceptivi/exteroceptivi și pasivi/activi.

*Senzorii proprioceptivi* măsoară valori interne ale sistemului (robot), de exemplu viteza motorului, sarcina roții, unghiuri ale articulației brațului robotului sau tensiunea bateriei.

*Senzorii exteroceptivi* obțin informații din mediul robotului, de exemplu măsurători ale distanței, intensitatea luminii sau amplitudine a sunetului. Prin urmare, măsurătorile senzorilor exteroceptivi sunt interpretate de către robot pentru a extrage caracteristici semnificative de mediu.

*Senzorii pasivi* măsoară energia mediului ambiant care intră în senzor. Exemple de senzori pasivi includ sonde de temperatură, microfoane și camere CCD sau CMOS.

*Senzorii activi* emit energie în mediul înconjurător, apoi măsoară reacția mediului. Deoarece senzorii activi pot gestiona interacțiuni controlate cu mediul, se atinge de multe ori o performanță superioară. Cu toate acestea, senzorul activ introduce mai multe riscuri: energia de ieșire poate afecta chiar caracteristicile pe care senzorul încearcă să le măsoare. Mai mult decât atât, un senzor activ poate suferi de interferență între semnalul său și a celor dincolo de controlul său. De exemplu, semnalele emise de alți roboți din apropiere, sau senzori similari situați pe același robot pot influența măsurătorile rezultate. Exemple de senzori activi includ encodere pentru roți, senzori ultrasonici și telemetre cu laser.

Senzorii proprioceptivi sunt de obicei pasivi și, de obicei, măsoară proprietățile fizice ale robotului, cum ar fi poziția articulațiilor, viteza, sau accelerația, cuplu motor și așa mai departe. Senzorii exteroceptivi, pe de altă parte, pot fi împărțiți în continuare în detectare cu contact și fără contact.

O clasificare a senzorilor tipici folosiți în robotică conform obiectivului măsurat (proprioceptive/exteroceptivi) și al metodei de măsurare (activă/pasivă) este arătată în tabelul următor [Christensen, 2008]:

**Tabelul 2.1.** Clasificarea senzorilor

Clasificare	Tipul senzorului	P/E	A/P
Senzori tactili	comutator / de contact	E	P
	barieră optică	E	A
	de proximitate	E	P/A
Senzori haptici	cu matrice de contact	EC	P

	de forță/de cuplu	P/E	P
	rezistiv	E	P
Senzori de turație	trductoare	P	P
	potențiometre	P	P
	traductori optici	P	A
	traductori magnetici	P	A
	traductori inductivi	P	A
	traductori capacitivi	E	A
Senzori de poziție	compas	E	P
	giroscop	P	P
	înclinometru	E	A/P
Senzori bazați pe unde	GPS	E	A
	optici	E	A
	emițătoare RF	E	A
	emițătoare ultrasunete	E	A
Senzori de proximitate	reflexivi	E	A
	senzori capacitivi	E	P
	senzori magnetici	E	P/A
	cameră video	E	P/A
	ultrasonici	E	A
	laser	E	A
Viteză / mișcare	Radar cu efect Doppler	E	A
	cameră video	E	P
	accelerometru	E	P
Identificare	cameră video	E	P
	RFID	E	A
	laser	E	A
	radar	E	A
	ultrasunete	E	A
	de sunet	E	P

Mai multe aplicații de robotica mobilă, roboți subacvatici și unele în robotica medicală se bazează pe ultrasunete. Clasa generală de senzori este adesea numită SONAR (sound navigation and ranging). Principiul general este că se emite un impuls ultrasonic și se așteaptă revenirea de ecouri de la obiecte din mediul înconjurător. Cunoscând viteza de transmisie în mediu și timpul de revenire, este posibil să se calculeze distanța la obiecte.

Progresele recente, în special cu privire la modelarea mediului și navigație, s-au datorat în multe privințe apariției unor sisteme de înaltă fidelitate cu scanare laser. Scannerul trimite un impuls luminos și măsoară timpul de întoarcere. Un scanner standard permite estimarea de distanțe de până la 80 m la acuratețe de centimetri sau milimetri. Datele senzorului sunt contaminate de zgomot, care trebuie să fie luate în considerare în detectarea de caracteristici sau de integrare a datelor într-o hartă.

Senzorii optici sunt o sursă bogată de informații pentru detectare și de estimare. Senzorii optici vin într-o mare varietate de configurații, variind în funcție de geometria imaginii, rezoluția imaginii, tehnologia senzorilor și gama de benzi spectrale detectate.

În același timp camerele ieftine cu infraroșu (IR) și ultraviolete (UV) au devenit disponibile, permițând dezvoltarea de sisteme avansate de interpretare a imaginilor multispectrale.

### **2.4.2. URMĂRIREA MIȘCĂRII CORPULUI UMAN**

Sistemele de detecție a mișcărilor utilizatorului sunt sisteme indispensabile atunci când este vorba de interacțiunea cu roboți sociali. Sistemele de urmărire a mișcării sunt caracterizate de precizie, rezoluție, răspunsul în timp și ergonomie. Pentru a urmări poziția și mișcarea membrelor sau a întregului corp uman se utilizează diferite tipuri de sisteme de detecție. În funcție de tipul de senzori utilizați, acestea se pot încadra în sisteme: magnetice, mecanice, inerțiale, geo-poziționale, optic-reflexive, sisteme de detecție video și sisteme hibride.

Prin comparație cu alte tehnologii folosite pentru achiziția gesturilor umane, sistemele de detecție video prezintă avantajul că utilizatorul nu este obligat să folosească sau să poarte echipamente suplimentare (spre exemplu mănușile cu senzori). Recunoașterea vizuală a gesturilor apare astfel ca o tehnologie ideală pentru interacțiunea om-mașină care vine să înlăture inconvenientele prezentate de alte metode.

Totuși, trebuie menționate o serie de limitări ale sistemelor de procesare video în ceea ce privește recunoașterea gesturilor, cum ar fi: rezoluția camerelor video nu este suficientă pentru a surprinde atât mișcarea fină a degetelor cât și cadrul mai larg al mișcării efectuate în același timp de braț; 30 cadre pe secundă în cadrul tehnologiei video convenționale sunt deseori insuficiente pentru a captura o mișcare rapidă a mâinii; segmentarea

degetelor poate deveni dificilă datorită ocluziei, degetele putându-se acoperi în mișcarea lor.

## 2.5. GESTURI FOLOSITE ÎN INTERACȚIUNEA OM-ROBOT

Ținând cont de modul în care oamenii interacționează între ei, se urmărește ca această capacitate a lor să fie transmisă și în sfera roboticii, prin construirea de roboți care nu necesită ca oamenii să se adapteze la ei într-un mod special. Acești roboți pot colabora ușor cu oamenii, interacționează cu ei și pot fi învățați într-un mod natural, aproape ca în cazul în care ei înșiși ar fi alți oameni.

Un robot asistent trebuie să fie capabil să interpreteze instrucțiunile verbale și non-verbale date de către om [Richarz, 2007]. În acest context, cercetătorii depun eforturi să găsească noi mijloace de interacțiune mai simple, mai intuitive și asemănătoare cu cele interumane, care să necesite în același timp mai puțină putere de calcul și dispozitive mai puțin sofisticate. Alături de alte abordări mai recente, cum ar fi interfețele bazate pe Electrooculografie (EOG) [Postelnicu, 2012], folosirea gesturilor umane rămâne încă o alternativă naturală și atractivă față de dispozitivele și interfețele împovărătoare ale interacțiunii om-calculator.

Printre alte acțiuni, gestul de indicare (pointing gesture) este natural și, probabil, cea mai intuitivă paradigmă de interacțiune, eficientă chiar și în mediile complexe și utilă pentru comanda sau pentru simpla transmitere a unui mesaj robotului [Park, 2011]. Pe de altă parte, recunoașterea gesturilor umane reprezintă o componentă non-verbală a interacțiunii care poate fi folosită atât în mediul real, cât și în cel virtual [Xu, 2013]. Interacțiunea prin gesturi a fost folosită în multe aplicații, cum ar fi comanda unui robot mobil [Alvarez-Santos, 2013] sau a unui quadrotor [Sanna, 2013], teleghidarea unui manipulator inteligent [Hussain, 2015], controlul unui player media [Suk, 2010], interacțiunea fără atingere cu meniurile dispozitivelor audio [Kajastila, 2013], proiectarea de concepte 3D [Kang, 2013] și așa mai departe.

După cum este menționat în lucrarea [McNeill, 1992], gesturile nu sunt doar mișcări, ci simboluri care au anumite semnificații și ajută la comunicarea unui mesaj. *În plus, gesturile și vorbirea sunt strâns legate în semnificație, funcție și timp; ele au sensuri, roluri și o menire comună.* Prin urmare, gesturile ajută ascultătorii să înțeleagă exprimările verbale atunci

când mesajul este complicat sau neclar [Alibali, 2005]. Nu este de mirare că vorbirea și gesturile de multe ori apar împreună în timpul dezvoltării omului.

Există trei faze asociate cu gesturile umane: pregătirea, care este deplasarea în poziție; efectuarea efectivă a gestului; și retragerea, care reprezintă revenirea la o poziție de repaus [Kendon, 2004]. Gesturile omului au loc în ceea ce se numește spațiu gestual, care este un disc imaginar în fața vorbitorului. Adulții efectuează de obicei gesturile într-un spațiu limitat, în timp ce spațiul gestual pentru copii este mai mare [McNeill, 1992]. Acesta poate fi util pentru gesturile roboților atunci când aceștia interacționează cu un utilizator uman.

### **2.5.1. TIPURI DE GESTURI**

Conform Dicționarului Explicativ al Limbii Române, un gest reprezintă o mișcare a mâinii, a capului etc. care exprimă o idee, un sentiment, o intenție, înlocuind uneori cuvintele sau dând mai multă expresivitate vorbirii [Internet2.6].

Gesturile pot fi clasificate:

- în funcție de forma lor [McNeill, 1992]:
  - § deictice – îndreaptă atenția către un anumit obiect, persoană sau localizare în mediul înconjurător;
  - § convenționale – gesturi care au o anumită semnificație și formă într-o comunitate și sunt, prin urmare, simboluri împărtășite din punct de vedere cultural;
  - § reprezentationale (iconice sau metaforice) – gesturi care fac referire la un obiect, o acțiune sau o relație, creând un aspect al formei sau mișcării referentului; cele iconice reprezintă obiecte fizice sau evenimente, iar cele metaforice reprezintă idei abstracte sau concepte;
  - § bruște – mișcări ale mâinii sau ale capului care evidențiază sau corespund cu prozodia discursului;
- în funcție de relația acestora cu expresia vocală asociată:
  - § pentru consolidarea informațiilor transmise prin vorbire;
  - § pentru a evita ambiguitatea informațiilor transmise prin vorbire;



- § pentru a adăuga informații la informațiile transmise prin vorbire [Cartmill, 2011];
- după funcția îndeplinită în relația om-mediu:
  - § ergotic – acționează asupra mediului;
  - § epistemic – care se referă la aparatul perceptiv al omului (văz, auz, receptori tactili);
  - § semiotic – produc un mesaj informațional pentru mediu.

Gesturile care se folosesc de obicei într-o conversație clasificate conform cu [McNeill, 1992] sunt utilizate în circumstanțe speciale. De exemplu, un gest iconic este strâns legat de conținutul semantic al vorbirii. Un gest metaforic este similar cu gesturile iconice în aceea că ambele reprezintă imagini. Gesturile metaforice sunt mult mai complexe, din cauza naturii lor duale. Un gest deictic este o mișcare de indicare. Aceste gesturi sunt de obicei realizate cu degetul arătător, dar pot fi, de asemenea, făcute cu capul, nasul sau bărbia. În cele din urmă, gesturi bruște nu prezintă nici un sens perceptibil și sunt de obicei bifazice (formate din două mișcări). Este important de remarcat faptul că un singur gest poate avea elemente deictice, de reprezentare și elemente bruște de marcarea a discursului.

### 2.5.2. TEHNICI DE RECUNOAȘTERE A GESTURILOR

Există mai multe tehnici utilizate pentru recunoașterea gesturilor [Pavlovic, 1997], [Mitra, 2007], [Rautaray, 2012]. De obicei aceste tehnici se împart în două categorii principale: bazate pe senzori și bazate pe imagine (vision).

În același timp, gesturile pot fi statice sau dinamice. Pentru recunoașterea gesturilor dinamice în timp real, există probleme în stabilirea punctelor de început și de sfârșit ale unui model gestual semnificativ dintr-un flux continuu [Kang, 2013]. În timp ce recunoașterea gestului static (postură) poate fi realizată în mod obișnuit prin tehnici precum potrivire de șablon (template matching) și recunoaștere a modelelor (pattern recognition), problema recunoașterii gesturilor dinamice implică utilizarea unor tehnici mai avansate [Carmona, 2012]. Unele gesturi au atât elemente statice, cât și dinamice [Mitra, 2007].

Având în vedere aceste observații, cercetătorii au propus diverse soluții pentru a optimiza recunoașterea gesturilor [Kang, 2013], [Linda, 2011], [Li, 2013], [Suk, 2010], [Cheng, 2012], [Park, 2011]. În această lucrare ne vom referi doar la recunoașterea vizuală și, în special, la gesturi ale brațelor. Așa

cum se arată în [Rautaray, 2012], cele mai utilizate tehnici de recunoaștere a mișcărilor corpului sunt Modele Markov ascunse (HMM), aliniere dinamică temporală (DTW), mașini cu număr finit de stări (FSM) și rețele neuronale (NN).

HMM a fost utilizat în [Park, 2011] pentru recunoașterea gesturilor de indicare cu scopul de a controla un robot mobil. Ei au folosit filtre de particule 3D și o cascadă de două HMM pentru a evalua direcția de indicare, ocupându-se atât cu gesturile de indicare mari, cât și mici. În [Suk, 2010] a fost utilizat un model probabilistic, o rețea Bayesiană dinamică (DBN) pentru recunoașterea gesturilor mâinii, care include HMMs și filtre Kalman. De asemenea, rețelele neuronale în combinație cu HMM au fost folosite în [Zhu, 2009] pentru recunoașterea gesturilor mâinii, dar algoritmul presupunea costuri de calcul ridicate.

Tehnica Elastic graph matching (EGM) a fost folosită pentru recunoașterea posturii mâinii, cu îmbunătățirile sale, EGM ponderată [Li, 2013] și EGM ierarhică [Anzalone, 2014]. Alți autori au combinat tehnici bazate pe senzori vizuali și accelerometre pentru o mai bună recunoaștere a gesturilor [Cheng, 2012], [Cheng, 2013].

Alinierea temporală dinamică (DTW) a fost folosită pentru prima dată pentru recunoașterea vorbirii [Sakoe, 1971], dar a fost extins și în alte domenii, inclusiv recunoașterea gesturilor [Linda, 2011]. În [Senin, 2008] este prezentată o analiză a algoritmului DTW.

După cum s-a văzut mai sus, există mai multe tehnici utilizate pentru detectarea și recunoașterea gesturilor umane, dar cele mai populare sunt HMM și DTW. Unele lucrări au demonstrat că rezultate mai bune pot fi obținute cu DTW în loc de HMM, atât în recunoașterea a vocii (de exemplu, sunete emise de animale - [Kogan, 1998]), cât și în recunoașterea gesturilor [Carmona, 2012]. Odată cu apariția senzorului comercial pentru jocuri video (Kinect), recunoașterea gesturilor a fost simplificată și mulți cercetători au folosit acest senzor în lucrările lor [Kajastila, 2013], [Sanna, 2013], [Bautista, 2013], [Susperregi, 2013].

### **2.6. APLICAȚII DE INTERACȚIUNE OM-ROBOT ÎN ROBOTICA ASISTIVĂ**

În ultimele decenii au fost dezvoltati mulți roboți de asistență pentru persoanele în vârstă și pentru persoanele cu handicap. Ei asistă oamenii în

activitățile lor zilnice, cum ar fi igiena personală, educație sau activități de mișcare independente. Cu toate acestea, aplicațiile roboților de asistență sunt mult mai diverse decât furnizarea de ajutor sau îngrijire pentru populația vulnerabilă.

Acest tip de roboți împart o zonă de lucru și interacționează direct cu oamenii. Trăsătura lor distinctivă este capacitatea lor de a percepe mediul și persoanele din jur cu ajutorul unor senzori și algoritmi inteligenți, de a comunica cu oamenii în mod multimodal, de a se deplasa autonom și de a lua decizii în mod independent.

În continuare, vor fi prezentate câteva dintre implementările mai importante legate de interacțiunea multimodală om-robot utilizate în robotica asistivă.

### **2.6.1. APLICAȚII BAZATE PE INTERACȚIUNEA PRIN GESTURI**

Roboții au fost folosiți ca instrumente de cercetare într-o varietate de aplicații [Alvarez-Santos, 2013], [Breuer, 2012], [Park, 2011], [Martin, 2010], [Duguleana, 2012]. Unii dintre ei s-au axat pe modul în care roboții sunt acceptați în societatea actuală [de Graaf, 2013], sugerând că pentru o prezență tot mai mare de roboți inteligenți în viața noastră de zi cu zi, este nevoie de oferirea de interacțiuni naturale. Ultimele cercetări punctează utilizarea gesturilor ca o modalitate de interacțiune om-calculator sau om-robot, ca un mod natural și intuitiv de comunicare sau de selectare a unor opțiuni [Alvarez-Santos, 2013], [Breuer, 2012], [Kajastila, 2013], [Kang, 2013], [Sanna, 2013].

Există, de asemenea, mai multe studii care abordează problema interacțiunii prin gesturi [Jaimes, 2007], [Willems, 2007]. Gesturile sunt mișcări expresive ale corpului care implică mișcări fizice ale degetelor, mâinilor, brațelor, capului, feței, corpului sau cu intenția de a transmite informații semnificative atunci când se interacționează cu mediul înconjurător [Mitra, 2007]. În special gesturile cu mâinile reprezintă o modalitate de interacțiune pentru oameni, mai ales atunci când aceștia încearcă să își exprime și să reprezinte ideile lor în spațiul 3D. Mișcările mâinilor și brațelor sunt, probabil, cele mai studiate gesturi implicate în convorbire [Wagner, 2014]. Pe baza acestei observații și presupunând că interacțiunea om-calculator și mai ales interacțiunea om-robot ar trebui să fie asemănătoare oamenilor (human-like), se poate afirma că gestul asociat

cu vocea este unul dintre cele mai potrivite „modalități de intrare” pentru controlul unui robot.

### **2.6.2. APLICAȚII BAZATE PE INTERACȚIUNEA PRIN VOCE**

Există mai multe studii cu privire la recunoașterea vorbirii cu atenție de a comanda roboți. Scopul principal al lucrărilor în acest domeniu este, în general, de a realiza o comunicare în limbaj natural cu asistentul robotic [Tur, 2011].

În cele mai multe studii vocea umană este folosită ca intrare în sistem, împreună cu alte semnale, într-o interfață de utilizator multimodală. De exemplu, în [Martínez, 2013] comenzile vocale sunt cuplate cu semnale EOG pentru a comanda un robot.

Diversi algoritmi sunt utilizați pentru a realiza recunoașterea vocii. Unul dintre ei este alinierea temporală dinamică, care se bazează pe comparația modelelor, destul de asemănător cu cel folosit în procesarea video [Hong, 2010]. Alte studii folosesc Hidden Markov Models [Caballero Morales, 2013] sau rețele neuronale artificiale (RNA) folosite ca atare sau combinate, de exemplu cu HMM pentru a obține recunoașterea vorbirii [Varchavskaia, 2001].

O altă clasificare a sistemelor de recunoaștere a vocii artificiale (ASR) este prin relația lor cu utilizatorul. Astfel, ele pot fi dependente/independente de vorbitor [Rabiner, 1993]. Sisteme dependente de vorbitor necesită instruirea utilizatorilor, dar au de obicei o acuratețe mai bună, în timp ce sisteme independente de vorbitor sunt mai flexibile, dar cu performanțe de recunoaștere mai mici.

În acest studiu s-a folosit o abordare cu vocabular, uzuală în aplicații de verificare a conceptului [Hong, 2010], [Pourmehr, 2013], [Kollar, 2014], bazată pe biblioteca Microsoft Kinect de recunoaștere a vocii. Algoritmul din spatele bibliotecii este următorul: un flux audio preluat de la senzorul Kinect este analizat și apoi enunțurile vocale sunt interpretate. Dacă motorul recunoaște anumite elemente, acestea sunt trimise la unitatea de procesare. În cazul în care comanda nu este recunoscută, se elimină acea parte din fluxul audio.

Oamenii sunt în general experți la vorbirea cu alte persoane sau la înțelegerea de către alte persoane. Prin urmare, comunicarea verbală reprezintă o interfață om-robot intuitivă [Brueckmann, 2007], [Kollar, 2010].

Recunoașterea vorbirii naturale este dificilă. Limbajul natural este adesea ambiguu și legat de gramatică, care este dependentă de cultură. Prin urmare, în cazul în care comunicarea om-robot se extinde dincolo de un mic set de comenzi stabilite, atunci va fi necesar un sistem de gestionare a dialogului. Sistemele de management al dialogului oferă o structură coerentă pentru interacțiunea om-mașină și permit gestionarea corectă a interacțiunilor mixte [Fong, 2003b].

Pentru a complica lucrurile și mai mult, vorbirea naturală poate fi acompaniată de zgomot de orice fel. Prin urmare, în viitor roboții personali de servicii trebuie să fie în măsură să separe vorbirea utilizatorului de zgomotul de fond. O abordare care ajută la realizarea acestui lucru este localizarea sursei de sunet, care se referă la capacitatea unui ascultător (în acest caz, capacitatea robotului) de a identifica originea unui sunet [Brueckmann, 2007], [Michaud, 2005], [Yamamoto, 2005], [Ikeda, 2007], [Rascón, 2010], [Liu, 2010]. Această capacitate poate completa viziunea și ajută robotul să localizeze o persoană care vorbește într-un mediu dat [Rascón, 2010]. După ce vorbitorul a fost localizat, robotul poate ignora sunetele care provin din alte locații.

Roboții care localizează surse de sunet au folosit între două și opt microfoane [Michaud, 2005], [Takeda, 2006], [Brueckmann, 2007], [Yilu, 2007] [Rascón, 2010]. Cu toate acestea, sistemele de localizare a sursei de sunet care se bazează pe unu sau două microfoane nu sunt adecvate pentru medii zgomotoase deoarece aceste sisteme nu filtrează în mod adecvat zgomotul [Michaud, 2005]. În schimb, sistemele de localizare concepute pentru medii zgomotoase ar trebui să folosească o matrice de microfoane [Ikeda, 2007]. Având în vedere că roboții de servicii personali vor funcționa în medii domestice zgomotoase, ei ar trebui să folosească o matrice de microfoane pentru localizarea sursei de sunet.

După ce robotul personal de servicii a localizat sursa de sunet, el trebuie să recunoască ceea ce se spune. Capacitatea sa de a face acest lucru poate fi îmbunătățită prin adaptarea a ceea ce roboții ascultă la situația actuală. De exemplu [Asoh, 1999] au dezvoltat un sistem de recunoaștere a vorbirii cu un dicționar care a fost adaptat pentru un mediu de birou.

Capacitatea unui robot personal de servicii de a „vorbi” în mod eficient va depinde în mare măsură de sistemul său de sinteză vocală. Calitatea de exprimare sintetică a fost îmbunătățită foarte mult în ultimii ani [Heckel, 2006].

În lucrarea [Fukui, 2005] a fost dezvoltat un robot care vorbește prin organe vocale asemănătoare cu cele umane, cum ar fi plămânii și corzile vocale, și articulatori, cum ar fi limba, buzele, dinții, și cavitatea nazală. Acest robot este interesant pentru că poate oferi perspective noi pentru modul în care vorbirea sintetică poate fi făcută să semene mai mult cu vocea umană.

### **2.6.3. APLICAȚII BAZATE PE INTERACȚIUNE MULTIMODALĂ**

Multimodalitatea este o caracteristică a comunicării umane. Oamenii folosesc adesea mai multe canale de transmitere a informației în interacțiunea lor de zi cu zi și este de dorit ca această caracteristică să fie transferată și la HRI, astfel încât omul să poată comunica într-un mod natural cu roboții. Una dintre primele încercări care abordează această idee este sistemul „Put-That-There”, sistem prezentat în [Bolt, 1980], care este considerat a fi precursorul interfețelor multimodale. Acest sistem combină recunoașterea vocii și a gesturilor pentru a permite utilizatorilor să comande evenimentele de pe un afișaj grafic de format mare. După acel moment au apărut o varietate de noi sisteme bazate pe interacțiunea multimodală cu utilizatorul, iar cele mai multe dintre ele sunt menționate în [Jaimes, 2007] și în [Karray, 2008]. În ciuda numărului mare de sisteme, cu diverse combinații de intrări, diverse metodologii de recunoaștere, diverse domenii de aplicare, există încă multe neajunsuri care trebuie să fie investigate de către cercetători.

În [Böhme, 2003] este propusă o schemă multimodală de interacțiune cu roboți de servicii, concentrându-se pe detectarea mișcării, localizarea sunetului, urmărirea oamenilor, localizarea utilizatorului sau a altei persoane și fuziunea lor. O abordare interesantă este prezentată în [Csapo, 2012a], unde a fost dezvoltat un sistem de interacțiune conversațională multimodal care sporește naturalețea și expresivitatea. O scurtă trecere în revistă a unor aspecte, probleme și provocări cheie în interacțiunea multimodală este oferită în [Turk, 2014]. Există o varietate de moduri de a interacționa cu un robot mobil, dar gestul și comenzile vocale rămân cele mai potrivite.

Oamenii au numeroase modalități de a comunica, multe dintre ele fiind non-verbale. De exemplu, cineva ar putea transmite tristețea prin anumite expresii faciale sau gesturi. Pentru o comunicare om-robot eficientă sunt

necesari roboți care au abilități de comunicare non-verbale [Breazeal, 2003], [Kanda, 2004], [Nieuwenhuisen, 2010]. Așa cum se afirmă în lucrarea [Harris, 2010], *aspectul fizic specific și mișcarea semnificativă sunt tot ceea ce face diferența între roboți și calculatoare.*

Privirea unei persoane este strâns cuplată cu înțelegerea și producerea limbii vorbite [Staudte, 2009]. Privirea, de asemenea, ajută partenerii de conversație să confirme dacă au înțeles conversația și dacă ei au fost înțeleși, fapt care crește performanțele tuturor actorilor [Mutlu, 2009]. Din aceste motive, privirea este considerată o capacitate de conversație umană de bază, împreună cu schimbarea rolului conversațional și pauzele între vorbitori [Trafton, 2008].

Direcția privirii este determinată de o combinație a orientării feței și orientării ochilor [Zhu, 2004]. Oamenii sunt foarte pricepuți în a determina unde se uită alte persoane [Delaunay, 2010]. Astfel, este evident motivul pentru care privirea ar putea fi o formă eficientă de comunicare non-verbală om-robot.

În lucrarea [Staudte, 2009] s-a arătat că privirea robotului ar putea facilita înțelegerea a ceea ce spune acesta. Mai exact, au descoperit că oamenii ar putea înțelege exprimarea robotului mai rapid atunci când aceste exprimări au fost însoțite de privirea corespunzătoare (atunci când robotul se uită la ținta dorită). Similar, în lucrarea [Trafton, 2008] a fost creat un partener de conversație pe care oamenii l-au găsit vigilent și natural prin cuplarea privirii și a vorbirii. Aceste rezultate sugerează că privirea poate fi o parte importantă a comunicării nonverbale om - robot.

De asemenea, este util să se exploreze semnalele fiziologice ca o formă de comunicare non-verbală om-robot. Aceste semnale ar putea permite utilizatorului să comunice cu un robot, fără a trebui să efectueze orice acțiune fizică. Din acest motiv, utilizarea de semnale fiziologice cu un partener robotic este o zonă de interes crescut în interacțiunea om-robot [Millan, 2004].

O astfel de comunicare este de obicei oferită de o interfață creier-calculator (BCI), care convertește semnalele fiziologice în intrări de comandă. Interfețele BCI au fost dezvoltate pentru a exploata o varietate de semnale fiziologice, inclusiv semnalele electroencefalografice [Jones, 2003], [Millan, 2004]; combinații de semnale electrocorticografice, biopotențialele pielii și tensiunile musculare faciale [Saulnier, 2009] și semnale

electrooculografice [Postelnicu, 2012], [Ierache, 2013] sau semnale electroencefalografice și electrooculografice [Postelnicu, 2011].

În lucrarea [Saulnier, 2009] a fost folosit un sistem BCI pentru a interacționa cu un robot Roomba. BCI a folosit semnalele bioelectrice ale participanților pentru a deduce nivelul lor de stres, care a fost apoi folosit pentru a ajusta comportamentul robotului. Rezultatele lor au indicat că nivelurile emoționale brute ar putea fi deduse din semnalele bioelectrice, care oferă controlul indirect al comportamentului robotului. Aceste rezultate sugerează ca BCI s-ar putea dovedi util pentru interacțiunea om-robot.

### **2.7. CONCLUZII**

În acest capitol au fost prezentate sintetic aspecte privind stadiul actual al interacțiunii om – robot și al mijloacelor de comunicare verbale sau non-verbale folosite, precum și aspecte despre roboții mobili folosiți în robotica de asistență. În toate aplicațiile de mai sus, deși există unele variații în ceea ce privește cerințele, un aspect cel puțin este comun: dorința ca roboții să fie capabili de o interacțiune naturală ușoară cu oamenii, suportând limbajul natural și comunicarea non-verbală, eventual ajutate și de alte mijloace.

În urma analizei bibliografice privind interacțiunea dintre oameni și roboți în medii domestice, s-a constatat faptul că s-au dezvoltat din ce în ce mai multe sisteme care permit o interacțiune ‘naturală’, ceea ce implică folosirea mai multor tipuri de ‘intrări’ din partea utilizatorilor, care sunt folosite în general în conversațiile inter-personale.

Comunicarea dintre utilizator și asistentul său robotic prin intermediul limbajului vorbit oferă multe posibilități utile. De asemenea, comunicarea non-verbală, prin intermediul gesturilor, poate secunda comunicarea verbală, îmbunătățind percepția mesajului și oferind posibilitatea de a transmite comenzi mai complexe.

Un sistem inteligent de interacțiune om-robot, bazat pe interacțiunea gestuală și vocală cu roboți de asistență în medii domestice reprezintă, astfel, un pas înainte spre dezvoltarea de noi sisteme avansate, la care pot fi incluse și alte modalități de interacțiune. Pe baza acestui sistem de interacțiune pot fi apoi realizate cercetări suplimentare.



În concluzie, în urma analizei stadiului actual al cercetărilor privitor la cooperarea și comunicarea om-robot, se evidențiază tendința de dezvoltare a acestora ca sisteme inteligente, cu capacități evolute de interacțiune.

## ***PREMISE HARDWARE ȘI SOFTWARE PENTRU STUDIILE TEORETICE ȘI EXPERIMENTALE***

În acest capitol se prezintă premisele necesare pentru dezvoltarea sistemului de interacțiune propus. Vor fi descrise componentele principale care au fost folosite pentru implementarea sistemului, atât cele software, cât și cele hardware.

Scopul inițial al cercetării a fost acela de a crea o interfață între un utilizator și un robot umanoid, prin care robotul să poată imita mișcările utilizatorului, fără a fi nevoie ca acesta să poarte un echipament suplimentar. Acest obiectiv a fost realizat cu ajutorul senzorului Kinect de la Microsoft. Apoi, funcțiile sistemului au fost extinse și s-a dorit realizarea unui sistem de interacțiune complex, care să integreze comenzi vocale și gestuale, folosind aceleași resurse hardware. Astfel, dispozitivul Kinect a fost folosit de data aceasta pentru recunoașterea gesturilor și, de asemenea, pentru recunoașterea comenzilor vocale ale operatorului uman.

### ***3.1. DISPOZITIVUL KINECT***

Întotdeauna a fost considerată o provocare dezvoltarea unei interfețe naturale de interacțiune. O interfață bazată pe gesturi umane, care nu necesită atașarea de alte dispozitive de corpul utilizatorului, va ajuta la imersarea acestuia în mod natural într-un mediu virtual sau în interacțiunea cu sisteme robotice.

Dispozitivul Microsoft Kinect (Fig. 3.1) oferă posibilitatea de a interacționa în mod natural cu interfețe grafice. Kinect a fost fabricat pentru a servi ca un controller pentru interfețe naturale. Senzorul a devenit foarte popular în domeniul cercetării. De la lansarea sa au existat numeroase

aplicații științifice care folosesc acest dispozitiv, dintre care multe sunt legate de domeniul roboticii. În cazul de față Kinect este foarte potrivit deoarece poate urmări gesturile unui utilizator și furnizează date, cum ar fi amplasarea în spațiul 3D a unei articulații a utilizatorului. Acesta lucru s-a dovedit a fi util întrucât în acest mod poate fi folosit pentru a calcula unghiurile membrelor sau localizarea brațului în spațiul 3D.



*Fig. 3.1. Dispozitivul Kinect*

Kinect este un dispozitiv de captură a mișcării dezvoltat inițial pentru consola de jocuri video Xbox și conține o camera video RGB-D pentru achiziția imaginii și o matrice de patru microfoane pentru captarea sunetului și localizarea sursei acestuia. Dispozitivul include și doi senzori de adâncime folosiți pentru a crea o hartă de profunzime a câmpului său vizual.

Două camere video în infraroșu (1 emițător și un senzor de profunzime) sunt folosite pentru captura datelor într-un mediu 3D. Datorită beneficiilor sale, Kinect a fost folosit în scopuri de cercetare, permițând interacțiunea fără atingere prin voce și gesturi. Utilizatorii se pot mișca liber, fără prea multe constrângeri.

Mai departe, va fi descrisă tehnologia Kinect și aplicațiile sale, precum și instrumentul de dezvoltare software folosit, care a fost ales datorită disponibilității documentației, a ușurinței de instalare și datorită faptului că oferă capacitatea de a urmări mișcările utilizatorului.

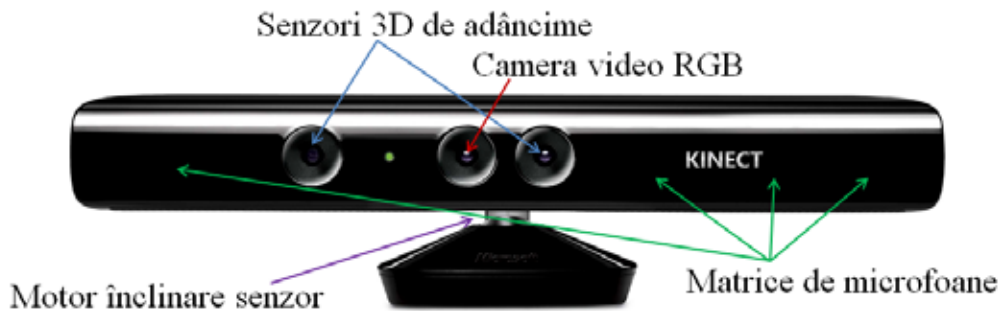
### **3.1.1. TEHNOLOGIA KINECT**

Kinect este o cameră produsă de Microsoft., care captează mișcările utilizatorului și le traduce în acțiuni de control pentru Xbox 360. Kinect a

fost lansat pentru prima dată în noiembrie 2010, numele său fiind un amestec dintre cuvintele „kinetic” și „connect”, care descrie aspecte cheie ale inițiativei.

Ceea ce este revoluționar în legătură cu dispozitivul Kinect este faptul că este primul proiect din lume care combină captarea 3D a mișcării, recunoaștere facială și vocală, cu software-ul dedicat, toate într-un singur dispozitiv. Combinația efectivă hardware și software conduce la un nou mod de a controla și de a interacționa. Nu este nevoie de niciun alt periferic (butoane, telecomenzi sau joystick-uri); utilizatorul trebuie doar să stea în fața aparatului Kinect și să folosească mișcările corpului, vocea și gesturile.

Platforma Kinect cuprinde ca tehnologii o cameră RGB, un senzor de adâncime 3D, un microfon multi-matrice și o înclinare motorizată, care sunt reprezentate în figura 3.2.



*Fig. 3.2. Componentele dispozitivul Kinect*

Camera video RGB oferă cele trei componente de culoare de bază, afișează informația video și ajută în procesul de recunoaștere facială. Ea furnizează informația video la o frecvență a cadrelor de 30 Hz și folosește o rezoluție maximă de  $640 \times 480$  pixeli, pe 32 de biți de culoare.

Senzorii 3D de adâncime constau dintr-un proiector laser în infraroșu care captează datele video în orice condiții de iluminare. Fascicolul laser este proiectat, iar senzorul este capabil să detecteze informațiile pe baza a ceea ce se reflectă înapoi la el. Proiectorul și senzorul creează împreună o hartă de adâncime. Astfel, camera de adâncime 3D oferă informații 3D detaliate despre mediul înconjurător. Altfel spus, determină cât de departe se află un obiect față de cameră. El are o limită practică variind între 1.2 și 3.5m distanță atunci când este utilizat cu software-ul Xbox. Camera în infraroșu (IR) este utilizată pentru urmărirea mișcării și adâncimii. Combinată cu un emițător IR, camera IR trimite în

scenă lumină invizibilă infraroșie. Astfel, ochiul nu vede lumina IR, iar condițiile de iluminare nu constituie o problemă pentru Kinect.

Matricea de microfoane permite recunoașterea de voci dintr-o cameră ale diferiților utilizatori, ignorând zgomotul ambiant. Cele patru microfoane sunt situate în partea de jos a dispozitivului Kinect, acestea fiind cele care au determinat dimensiunea și forma dispozitivului. Matricea de microfoane funcționează astfel încât fiecare canal procesează informația audio pe 16 biți la o rată de eșantionare de 16 kHz.

Înclinarea motorizată este folosită pentru ajustarea senzorului, permițând urmărirea utilizatorilor, chiar și atunci când aceștia se deplasează. Acesta este capabil de înclinare a senzorului până la 27° în sus sau în jos, în timp ce câmpul unghiular de vedere este de 57° pe orizontală și 43° pe verticală.

Kinect este capabil să urmărească simultan până la șase persoane și poate urmări 20 de articulații pentru fiecare utilizator în timp real.

Utilizarea inițială preconizată a Kinect a fost de a oferi divertisment gratuit

prin intermediul interfeței naturale cu utilizatorul. Dispozitivul Kinect a fost ușor de utilizat, nu numai pentru utilizatorii clinic sănătoși, dar, de asemenea, pentru copii cu diferite dizabilități sau care suferă de autism, paralizie cerebrală sau hidrocefalie. De exemplu, un utilizator stând într-un scaun cu roțile va fi detectat imediat ca o persoană mai scundă.

Ulterior, cercetătorii și comunități academice au început să fructifice noi utilizări interesante ale Kinect, nu numai pentru distracție și divertisment. Astfel, dezvoltatorii au creat drivere open-source care permit ca senzorul Kinect să lucreze pe alte dispozitive decât Xbox.

#### **3.1.2. APLICAȚII ALE DISPOZITIVULUI KINECT**

În continuare vor fi prezentate câteva noi utilizări ale dispozitivului Kinect, care sunt realizate de către dezvoltatori și cercetători universitari. Cele mai mult dintre aplicații și-au găsit aplicabilitatea în domeniul medicinei. Kinect permite chirurgilor să navigheze în mediul virtual, să manipuleze raze X și să scaneze în timpul operațiilor, doar cu un gest al mâinii, fără a atinge suprafața nesterilă a unui mouse sau tastatură [Ebert, 2013], [Wen, 2014]. Este un prim exemplu pentru domeniul înfloritor al interfețelor

naturale cu utilizatorul, care promite să schimbe relația noastră cu dispozitive omniprezente. Dar aplicațiile din domeniul medicinei se extind mult dincolo de sala de operație. Abilitățile de urmărire a corpului ale dispozitivului Kinect sunt exploatate și pentru evaluarea clinică a unei dizabilități motorii. Pacienții sunt urmăriți cu ajutorul dispozitivului în timp ce realizează o anumită activitate, sunt monitorizate mișcărilor lor, iar datele pot fi furnizate medicilor, care evaluează starea bolii [Su, 2014].

Alți cercetători exploatează modul de a folosi dispozitivul Kinect pentru a evalua daunele provocate de accidente vasculare cerebrale și de a crea și monitoriza exerciții de reabilitare bazate pe jocuri, dintre care multe pot fi efectuate de către pacienții cu accident vascular cerebral în propriile lor case. De asemenea, Kinect poate ajuta în diagnosticarea tulburărilor creierului și ale sistemului nervos, inclusiv Alzheimer și scleroza multiplă [Iarlari, 2014].

Aplicația dezvoltată prin proiectul KinectFusion [Izadi, 2011] permite unui utilizator să apuce și să deplaseze un dispozitiv Kinect standard pentru a crea rapid reconstrucții 3D detaliate ale unei scene. Sunt folosite numai datele de adâncime ale senzorului. Kinect este astfel folosit ca un scanner 3D pentru producerea de imagini 3D de înaltă calitate ale unor obiecte mici sau mari.

Dispozitivul Microsoft Kinect a depășit utilizarea prevăzută inițial. Idei uimitoare care înainte erau doar un vis, au devenit acum realitate. Au fost dezvoltate aplicații interesante, cum ar fi: sistem de supraveghere video [Benayed, 2014], recuperarea pacienților care au suferit accidente vasculare cerebrale [Webster, 2014], traducerea limbajului semnelor în limbă vorbită și scrisă, în timp real [Chai, 2013], controlul roboților prin intermediul mișcărilor corpului [El-laithy, 2012], evitarea coliziunilor sau detectarea altitudinii pentru elicoptere robotice [Stowers, 2011], interacțiune cu realitatea virtuală, utilizarea gesturilor pentru a manipula imagini [Jacob, 2012], teleconferințe 3D [Maimone, 2011], reconstrucție 3D, interacțiunea cu obiecte 3D prin intermediul tehnologiei de realitate augmentată și multe altele [Kitsunozaki, 2013], [Cruz, 2012].

### ***3.1.3. INSTRUMENTUL DE DEZVOLTARE SOFTWARE***

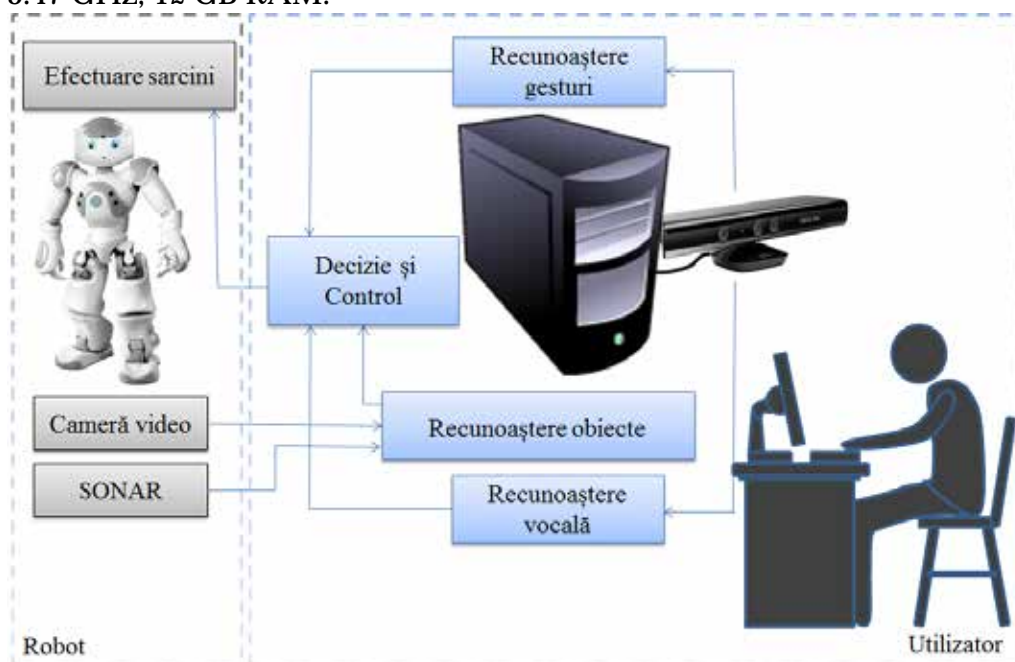
Există mai multe instrumente care definesc API-uri pentru scrierea de aplicații care folosesc interfețe naturale și care includ posibilitatea de a

activa, inițializa și comunica date cu dispozitivul Kinect. Cele mai importante dintre acestea sunt: OpenNI (Open Natural Interaction), Libfreenect (proiectul OpenKinect), middleware-ul NITE dezvoltat de PrimeSense, openni\_kinect (proiectul ROS OpenNI, care integrează Kinect cu ROS – Robot Operating System), middleware-ul FFAST (The Flexible Action and Articulated Skeleton Toolkit), Microsoft Kinect SDK.

În această lucrare a fost folosit Kinect for Windows SDK, care este un set de instrumente ce oferă o interfață pentru a interacționa cu dispozitivul. Acesta oferă biblioteci (API) pentru aplicații .NET și C/C ++ care rulează pe platforme Windows.

Kinect SDK urmărește coordonatele 3D ale 20 de articulații ale corpului, în timp real și pozițiile articulațiilor obținute sunt folosite pentru a recunoaște gestul sau postura care va reprezenta comanda dată robotului.

În figura 3.3 este prezentată arhitectura sistemului de interacțiune, având scopul de a evidenția rolul dispozitivului Kinect în cadrul interfeței propuse. Un calculator reprezintă unitatea principală de procesare și pe acest calculator rulează aplicația care permite recunoașterea gesturilor, recunoaștere vocală, un sistem de integrare voce/gest și comunicarea cu robotul. Computerul este echipat cu un procesor Intel core i7 X 990 CPU 3.47 GHz, 12 GB RAM.



**Fig. 3.3.** Arhitectura generală sistemului

În concluzie, Kinect este un periferic de urmărirea mișcării pentru consola Xbox care conține o combinație de camere și senzori. Acesta a fost proiectat în primul rând pentru distracție și divertisment social, dar oamenii de știință și cercetătorii au descoperit Kinect ca un dispozitiv uimitor care poate fi utilizat în multe alte domenii. Kinect a permis construirea de modele 3D a diferite medii, pentru a face ca roboții să răspundă la gesturile umane și pentru a evita coliziunile, pentru teleconferințe 3D și, de asemenea, pentru a ghida imagistica în timpul intervențiilor chirurgicale.

### **3.2. STRUCTURA FIZICĂ A ROBOTULUI UMANOID NAO**

Roboții au devenit o prezență obișnuită în medii industriale, cercetare, medicină și alte domenii, cum ar fi: divertisment, mediul domestic. Ne dorim ca ei să fie capabili să desfășoare activități complexe în medii dinamice. Datorită numărului mare de grade de libertate (DOF), roboții umanoizi sunt utilizați în aceste medii. Cu toate acestea, acest lucru face dificilă programarea roboților pentru o interacțiune om-robot intuitivă. Unii cercetători au propus conceptul PbD pentru a trata aceste neajunsuri. Ei au folosit conceptul PbD pentru transferul competențelor (abilităților) umane la roboți.

Robotul umanoid utilizat în sistemul dezvoltat în această lucrare este NAO [Aldebaran\_Robotics, 2012]. Acesta este un robot inteligent și a fost folosit în multe cercetări, cum ar fi: teleoperare [Stanton, 2012], îmbunătățirea stabilității mersului biped [Meriçli, 2012], reîncărcare autonomă pentru îmbunătățirea autonomiei energetice [Navarro-Guerrero, 2012], îmbunătățirea abilităților în jocul de fotbal [Cruz, 2011], interacțiunea cu copiii care suferă de autism, în scopul de a reduce problemele lor de comportament [Shamsuddina, 2012], cooperarea a doi roboți [Panfir, 2013], interacțiunea multimodală om-robot bazată pe folosirea vocii, privirii și gesturilor utilizatorului [Csapo, 2012b].

Caracteristica intrinsecă a unui robot mobil este capacitatea sa de a se mișca în mediul înconjurător. Există mai multe tipuri de roboți din punct de vedere al mișcării, pornind de la cei cu roți și terminând cu cei pășitori. Roboții umanoizi NAO, ca reprezentanți ai celor cu picioare, au dificultăți în a se deplasa cu precizie deoarece odometria mișcării este calculată din



mișcarea relativă a picioarelor, care adesea alunecă și astfel eroarea de odometrie este destul de mare și crește rapid.

În continuare vor fi prezentate caracteristicile principale și senzorii folosiți de către robotul NAO pentru interacțiunea cu mediul.

#### **3.2.1. COMPONENTELE ȘI SENZORII**

Robotul NAO este un robot umanoid conceput pentru scopuri educaționale de către compania franceză Aldebaran Robotics. NAO are 58 cm înălțime, o greutate de aproximativ 5 kg (Fig. 3.4a) și este echipat cu o varietate de senzori. Există senzori ultrasonici, microfoane, senzori cu infraroșu (IR), un senzor inerțial, senzori tactili și de presiune și două camere video. Capul robotului conține un procesor dual-core Atom 1.6GHz CPU, pe care rulează Linux și un software care controlează robotul. Există o mulțime de funcționalități expediate împreună cu software-ul integrat, cum ar fi detectarea feței, detectarea markerilor (NAOmark), mersul și alte comportamente simple.

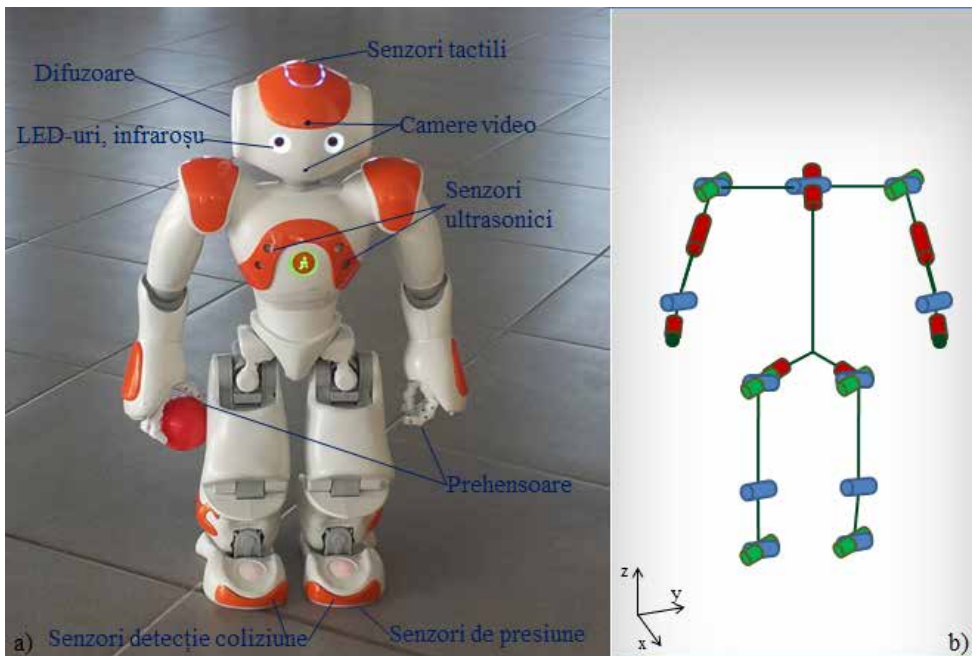
Robotul are un sintetizator de voce și un modul simplu de recunoaștere a vocii. Cele două camere ale robotului sunt plasate în zona frunții și, respectiv, a bărbiei. Deoarece câmpurile lor vizuale nu se suprapun, acestea nu pot fi utilizate ca o pereche stereo și, astfel, va fi utilizată numai una dintre ele. Camera superioară are 1.22 megapixeli și oferă o rezoluție VGA la rate de peste 15 fps (pe o conexiune Gigabit Ethernet). Rezoluția mare este disponibilă cu rată de cadre redusă.

În figura 3.4 sunt prezentate principalele caracteristici ale robotului umanoid. NAO are în total 25 de grade de libertate, 11 pentru partea inferioară care include picioarele și bazinul și 14 DOF pentru partea superioară a corpului, care include trunchiul, brațele și capul. Fiecare picior are 2 grade de libertate pentru gleznă, 1 la genunchi și 2 la șold. Un mecanism special compus din două articulații cuplate la fiecare șold echipează bazinul. Axele de rotație ale celor două articulații sunt înclinate la 45° față de corp. Această structură ajută la distribuirea mai bună a puterii între articulația șoldului și cea a bazinului și conferă un stil specific de mișcare a robotului [Gouaillier, 2009].

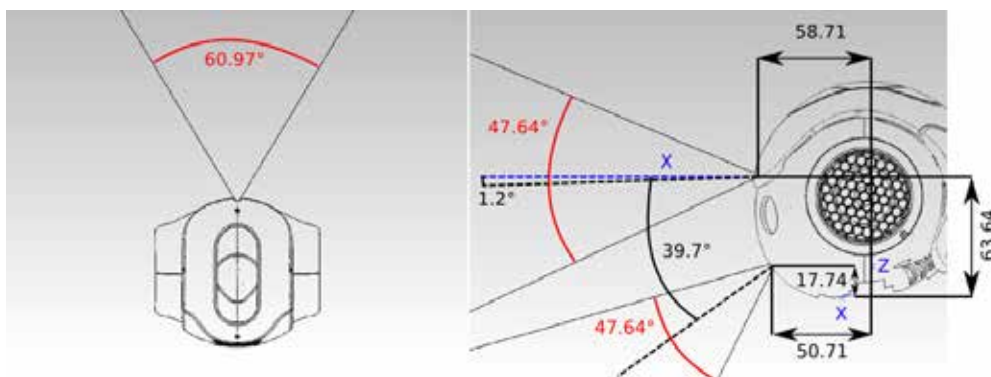
Fiecare braț are 2 grade de libertate la umăr, 2 la cot, unul la încheietura mâinii și unul suplimentar pentru prehensor. Capul se poate roti în jurul axelor y și z. Figura 3.4b oferă detaliile despre cuplele robotului.

Așa cum s-a menționat, NAO dispune de 2 senzori vizuali îmbunătățiți. Pentru activitatea de procesare a imaginii de pe calculatorul robotului, se pot utiliza până la 30 imagini/secundă la rezoluție HD. NAO poate deplasa capul cu  $239^\circ$  pe orizontală și de  $68^\circ$  pe vertical (Fig. 3.6.), iar câmpul vizual al camerei este de  $61^\circ$  pe orizontală și  $47^\circ$  pe vertical (Fig. 3.5.). Astfel, NAO are o mare capacitate de a simți mediul său și de a recunoaște o cantitate mare de obiecte, care pot fi salvate și recunoscute ulterior.

O altă caracteristică foarte utilă este recunoașterea și detecția feței. NAO poate detecta și învăța o față umană, în scopul de a o recunoaște data următoare. De asemenea, este capabil să vorbească până la 9 limbi. Recunoașterea vorbirii se află în centrul interacțiunii intuitive om-robot. NAO este capabil să recepționeze semnalul audio de la 2 metri (datorită microfoanelor amplasate în jurul capului), recunoaște o propoziție completă sau doar câteva cuvinte din propoziție, rezultând fluiditate și conversații naturale.



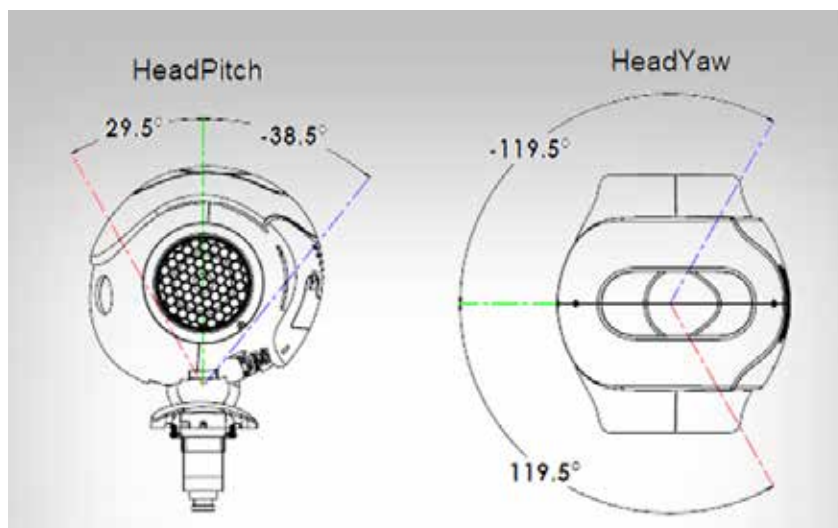
**Fig. 3.4.** Caracteristicile robotului NAO și articulațiile sale



**Fig. 3.5.** Câmpul vizual al camerelor video [Internet3.2]

Mișcarea lui NAO se bazează pe motoare de curent continuu (DC), iar robotul are o autonomie limitată la aproximativ 30 de minute.

NAO este echipat cu 2 dispozitive cu ultrasunete situate în zona pieptului, care oferă informații spațiale în limita a 1 metru distanță, dacă un obiect este situat la 30 grade de la pieptul robotului (60 de grade dacă se combină ambele dispozitive).



**Fig. 3.6.** Articulațiile capului

Senzorii de coliziune sunt situați în fața fiecărui picior și aceștia pot fi utilizați, de exemplu, pentru a ști dacă robotul intră în contact cu un obstacol.

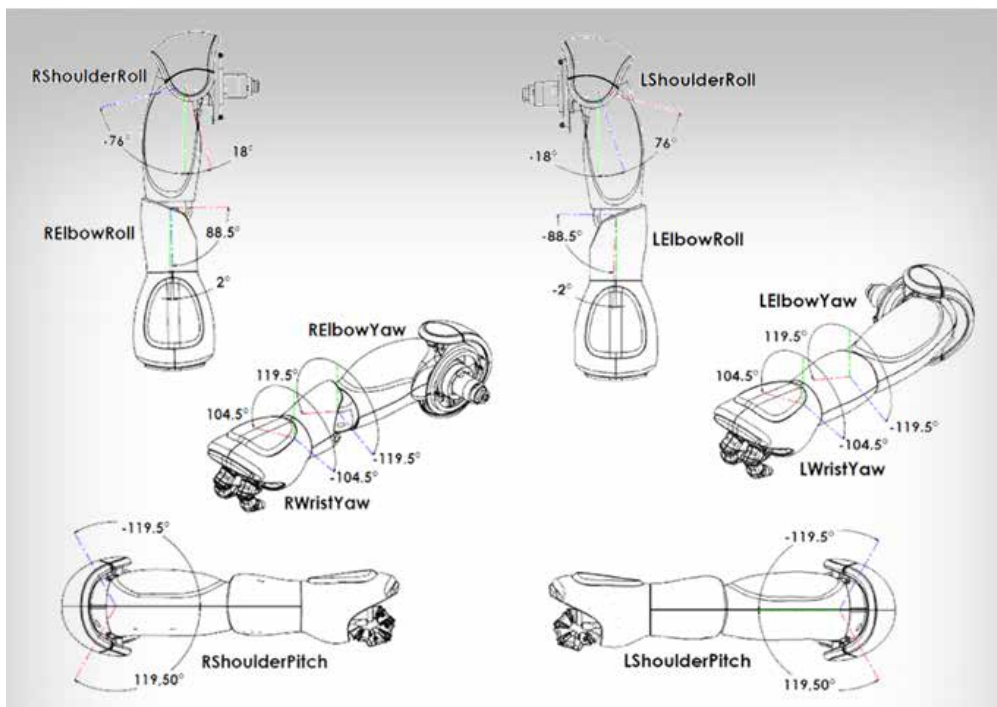
NAO are 8 rezistori de măsurare a forței (FSR) situați pe tălpile picioarelor. Există 4 rezistori pentru fiecare picior. Valoarea returnată de fiecare FSR este un timp necesar pentru un condensator de a se încărca, în

funcție de valoarea rezistorului FSR. Acesta nu este liniar ( $1/X$ ) și trebuie să fie calibrat. Când nu se aplică nici o forță, citirea senzorului este de 3000, iar când citirea senzorului este de 200 înseamnă că suportă aproximativ 3 kg. Acești senzori sunt utili atunci când se generează secvențe de mișcări pentru a ști când o poziție este un moment zero (ZMP) și senzorii pot fi complementați cu senzori de inerție. NAO este echipat cu 2 girometre și un accelerometru. Acești senzori ajută să știm dacă robotul este într-o poziție stabilă sau într-una instabilă atunci când se deplasează. De asemenea, acești doi senzori ajută la odometrie.

Alți senzori ai robotului sunt: 4 microfoane, 2 emițătoare și receptoare cu infraroșu, 9 senzori tactili, 36 de senzori cu efect Hall, LED-uri, difuzoare, conectivitatea include Ethernet și WiFi. Morfologia NAO pentru primele 4 grade de libertate ale membrelor superioare este echivalentă cu cea de la om. De aceea, transformările coordonatelor de la om la NAO sunt foarte simple, fiind nevoie de a lua în considerare doar limitele articulațiilor robotului pentru a evita coliziunile. Rutinele proiectate iau în considerare spațiul de lucru al lui NAO pentru a evita pozițiile inaccesibile și coliziunile.

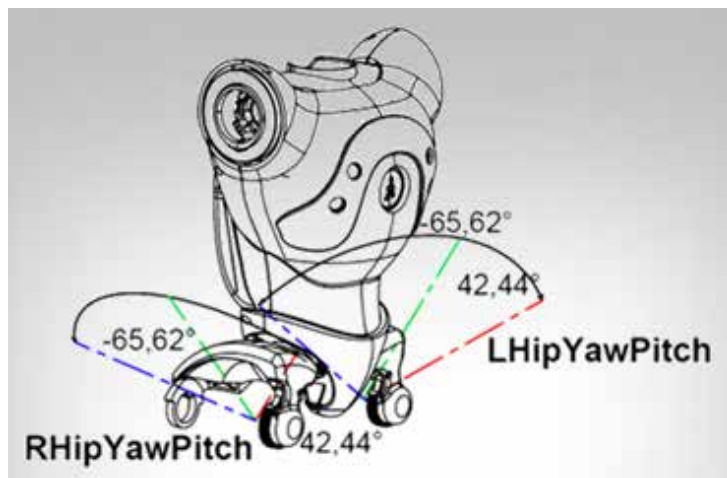
Caracteristicile prezentate se referă la versiunea NAO Next Gen H25, care a fost utilizată în această lucrare. Această versiune oferă și mai multe îmbunătățiri în partea software comparativ cu primele două generații ale robotului lansat în 2008, și 2010. Principalele îmbunătățiri: camerele video HD, cu o rezoluție crescută a imaginii maximă de 1280x960 pixeli, un procesor mai rapid (Intel Atom Z530), un sistem anti-coliziune complet reinventat, robustețe generală mai bună și traiectorii îmbunătățite pentru mișcări ale articulațiilor mai rapide.

În figurile următoare sunt prezentate intervalele de mișcare pentru articulațiile robotului. Imaginile sunt preluate din documentația robotului [Internet3.1].

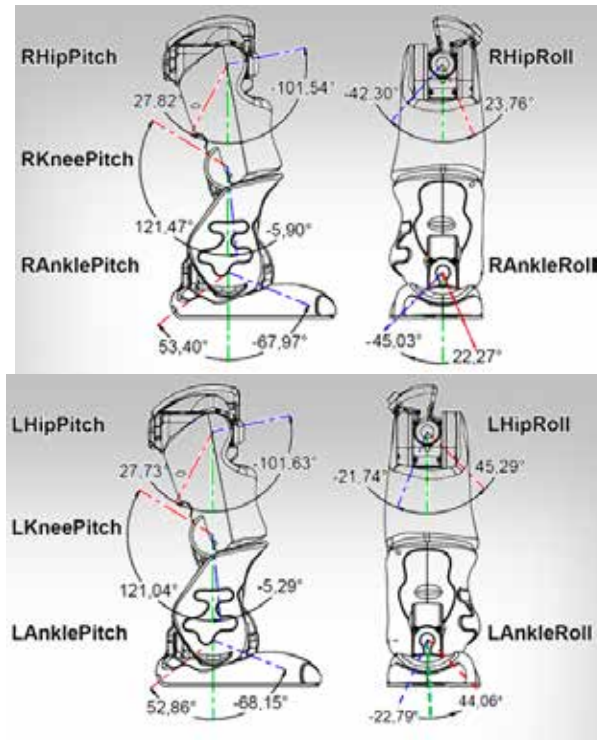


**Fig. 3.7.** Articulațiile: a) mâinii drepte; b) mâinii stângi

Odată ce au fost specificate partea hardware a robotului NAO și arhitectura mecanică, este necesară o introducere la arhitectura software pentru a cunoaște caracteristicile software ale sale.



**Fig. 3.8.** Articulațiile bazinului (șoldurilor)



**Fig. 3.9.** Articulațiile piciorului drept și ale celui stâng

### 3.2.2. COMPONENTELE SOFTWARE

Arhitectura software este numită NAOqi de către Aldebaran Robotics. NAOqi este software-ul principal care rulează pe robot și servește ca o platformă software ușor de utilizat, care permite modulelor utilizatorului să comunice unul cu celălalt și, de asemenea, cu diverse module intrinseci. Modulele pot rula direct pe robot sau ca aplicații de la distanță de la un calculator.

Această platformă permite comunicarea omogenă între diferite module (mișcare, audio, video), programare omogenă și schimbul de informații omogen. Poate rula pe Linux (pe procesorul robotului), dar și pe Windows sau Mac.

NAOqi oferă, ca și module de bază, interpretori pentru limbajele Python și Urbi, dar pot fi folosite și alte limbaje (C++, .NET), prin intermediul unor SDK-uri dedicate.

Software-ul poate fi dezvoltat în C++ și Python. În toate cazurile, metodele de programare sunt exact la fel, toate API-urile existente pot fi apelate din orice limbaj suportat: C++, Python.

Executabilul NAOqi care rulează pe robot este un agent (*broker*). Când este pornit, se încarcă un fișier de preferințe numit `autoload.ini` care definește ce biblioteci ar trebui să se încarce. Fiecare bibliotecă conține unul sau mai multe module pe care le folosește broker-ul pentru a apela metodele lor.

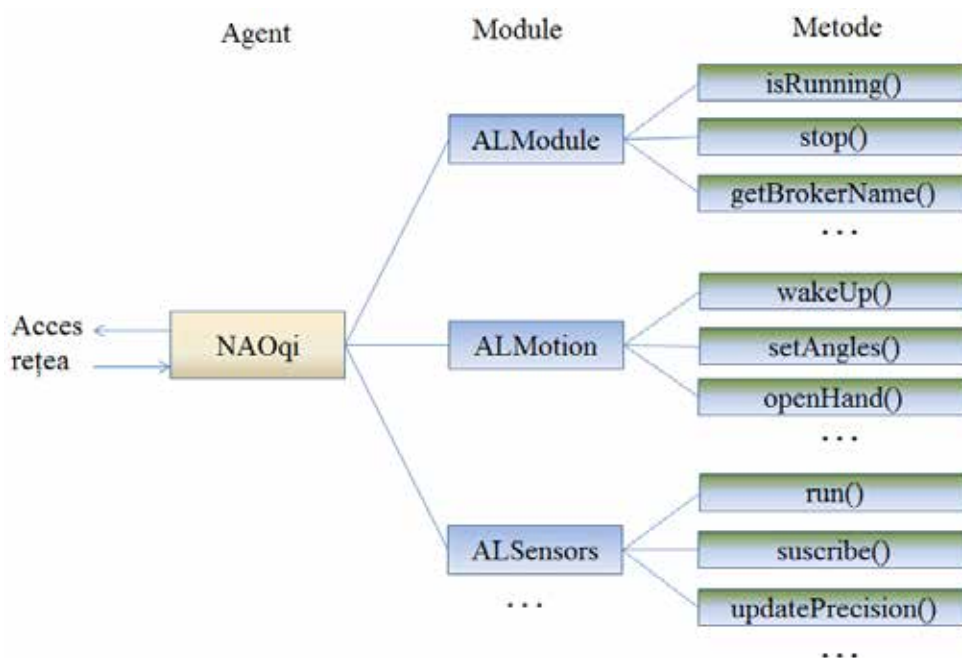
Agentul oferă servicii de căutare, astfel încât orice modul din arbore sau din întreaga rețea poate găsi orice metodă care a fost apelată.

Încărcarea modulelor formează un arbore de metode atașate la module, iar modulele atașate la un broker (Fig. 3.10).

Un *agent* este un obiect care oferă două roluri principale:

- oferă servicii de director: permite găsirea de module și metode;
- oferă acces la rețea: permite metodelor modulelor atașate să fie apelate din afara procesului.

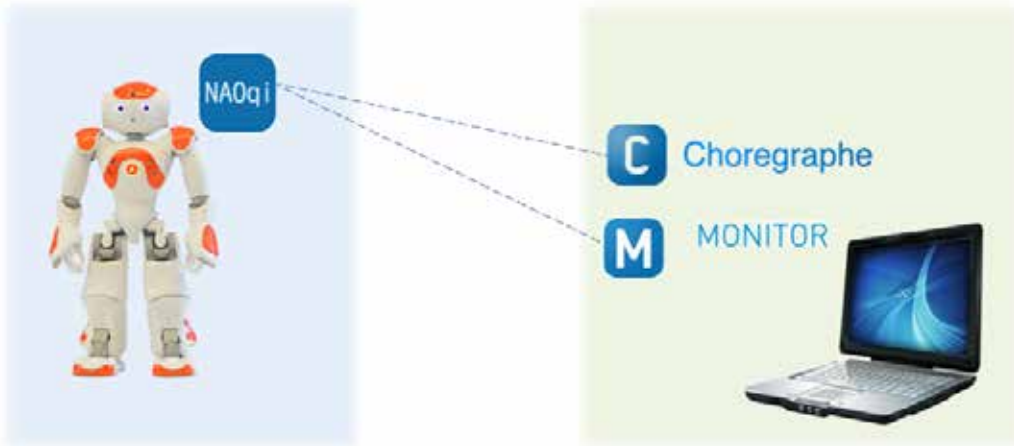
Acești agenți fac munca lor în mod transparent, permițând scrierea de cod, care va fi la fel pentru apeluri de "module locale" (în același proces) sau "module de la distanță" (într-un alt proces sau de pe altă mașină).



**Fig. 3.10.** Structura cadrului software NAOqi

Așadar, NAO are nevoie de două tipuri de programe pentru a putea fi controlat (Fig. 3.11):

- software încorporat, care rulează pe placa de bază situat în capul robotului și care permite comportamente autonome;
- software desktop, care rulează pe un computer și care permite crearea de noi comportamente, precum și controlul de la distanță al robotului.

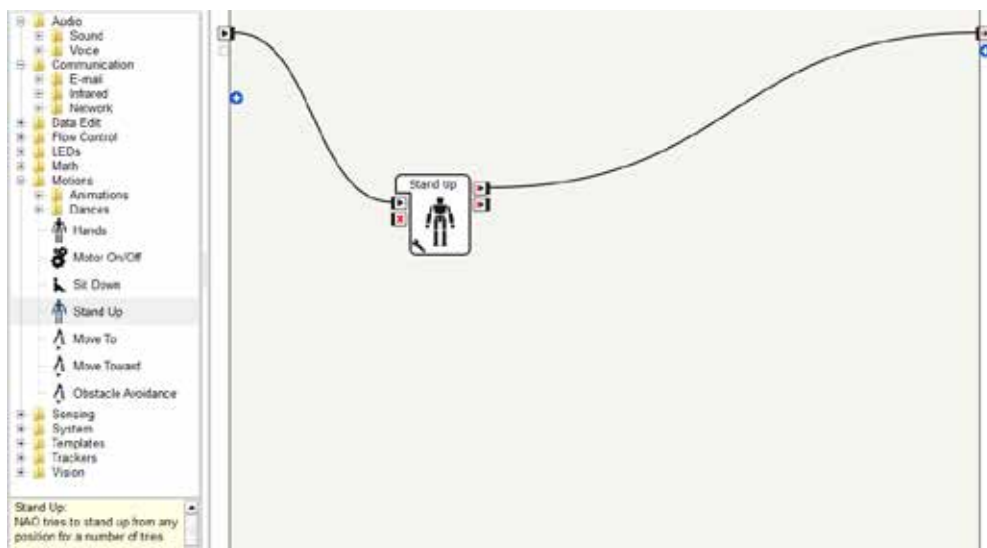


*Fig. 3.11. Modulele software ale robotului NAO*

**Choregraphe** – este un mediu grafic intuitiv care permite programarea simplă a robotului. Acest mediu rulează pe calculator și oferă acces la toate funcțiile oferite de NAOqi. Programarea constă în implementarea de comportamente, care la rândul lor sunt formate din mai multe secvențe. Există un set de comportamente predefinite, dintre care unele de nivel mai înalt (mers, ridicare, recunoașterea vocii), iar altele mai simple (aprinderea unui LED, deschiderea mâinii, etc.). Prin asamblarea acestor comportamente de bază, se pot crea alte comportamente originale. Oricine poate crea propriile „casete” (fig. 3.12) care pot fi adăugate la biblioteca existentă.

**Monitor** – este un software care permite vizualizarea datelor de senzori și de la motoarelor din articulațiile robotului în timp real. Acest mediu are două caracteristici: vizualizarea informațiilor de la senzorii robotului (girometru, accelerometru, acumulator) sau de la actuatori (poziție, temperatură, comandă) și vizualizarea fluxului video primit de la camera robotului. Astfel, această aplicație este utilă pentru monitorizarea robotului sau a sistemului în care este integrat.



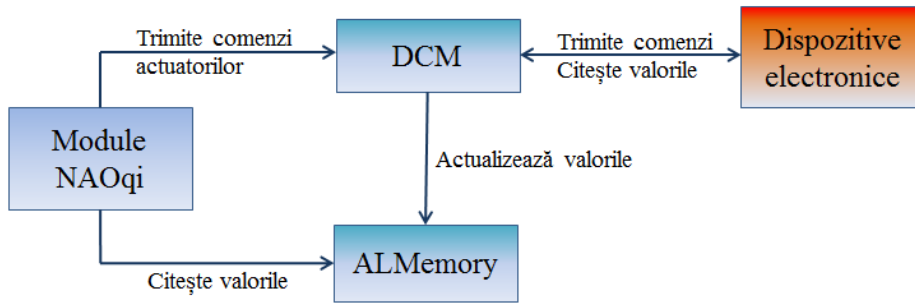


*Fig. 3.12. Interfața grafică a mediului Choregraphe*

Pentru comunicarea cu dispozitivele electronice ale robotului (senzori, actuatori, plăci electronice) este utilizat modul software DCM (Device Communication Manager), parte integrantă a NAOqi. Acest modul permite comunicarea cu toate dispozitivele electronice, în afară de camerele video și de componentele audio (difuzoare, microfoane). DCM gestionează principalele magistrale de comunicare între capul și corpul robotului, și, de asemenea, comunicarea în interiorul capului, făcând astfel legătura între software-ul „de nivel superior” (modulele NAOqi) și software-ul „de nivel inferior” (software-ul din plăcile electronice). Module cum ar fi ALMotion și ALLeds trimit direct comenzi către actuatori folosind acest modul, în timp alte module folosesc rezultatele primite de la senzori returnate de DCM în ALMemory (Fig. 3. 13). Întrucât DCM rulează unele elemente securitate software, utilizarea sa este obligatorie pentru a accesa partea hardware a robotului.

DCM are două scopuri principale:

- trimite comenzi către actuatori;
- actualizează valorile actuatorilor și ale senzorilor.



**Fig. 3.13.** Modulul DCM al NAOqi

Actuatorii și senzorii pot fi actuatori „fizici” (poziții ale articulațiilor, LED-uri, motoare) sau „virtuali” (de ex., coeficientul de rigiditate, care este un parametru de pe bucla de control).

API-ul DCM permite trimiterea de comenzi către actuatori. Se pot actualiza cu un apel unul sau mai mulți actuatori. Comenzile includ o informație de timp de când actuatorul trebuie să primească valoarea.

Pentru a obține valori de la actuatori și senzori, trebuie să fie citite în ALMemory. DCM le actualizează în măsura în care se modifică. ALMemory este o memorie centralizată utilizată pentru a stoca toate informațiile cheie legate de configurația hardware a robotului. Mai precis, ALMemory oferă informații despre starea actuală a dispozitivelor de acționare și a senzorilor. ALMemory poate fi, de asemenea, utilizată pentru a stoca și prelua valorile după nume și acționează ca un hub pentru distribuirea de notificări de evenimente.

### 3.2.3. INTERFAȚA DE COMANDĂ A ROBOTULUI

O interfață grafică a fost dezvoltată pentru a realiza o mapare între mișcările utilizatorului și cele ale robotului. Interfața ajută în faza de calibrare și de supraveghere a executării modulului de calcul a coordonatelor articulațiilor. Această interfață grafică (prezentată în Fig. 3.14) permite alegerea imaginii care să fie utilizată (RGB sau imaginea de adâncime) și o afișează. Interfața afișează articulațiile utilizatorului folosite pentru a comanda robotul real NAO, cât și a celui virtual din software-ul Coregraph, prin intermediul modulului de calcul al mișcării. Programul asociat cu interfața grafică stabilește conexiunea directă între simulatorul Choregraph și robotul NAO. În cazul în care platforma NAO nu este disponibilă, mișcările utilizatorului sunt reproduse în simulatorul

### CAPITOLUL 3

Choregraph prin intermediul NAOqi, care are o adresă locală a calculatorului atunci când se execută programul. Este, de asemenea, posibilă comanda robotului fără a rula simulatorul Choregraph.

Simularea în programul Choregraph este, așadar, folosită pentru a analiza aplicația fără a fi nevoie de robotul real NAO, evitând ciocniri cu mediul sau cu părți ale robotului, poziții singulare ale robotului sau poziții care nu se încadrează în spațiul de lucru al robotului.

Kinect permite comunicarea prin gesturi între om și robot, făcând posibilă comanda robotului prin intermediul mișcărilor umane. Este folosit Kinect for Windows SDK pentru a obține informații ale modelului-schelet al utilizatorului și pozițiile articulațiilor lui. Având aceste poziții, se pot calcula unghiurile articulațiilor umerilor. După aceasta, se pot transforma aceste unghiuri în unghiuri corespunzătoare robotului, așa cum sunt descrise în API-ul Aldebaran. După efectuarea acestei conversii, vor fi transmise valorile unghiurilor prin conexiunea wireless și atunci când robotul le primește, se utilizează funcția NAOqi (setAngleInterpolation) pentru a seta unghiurile umărului la valorile indicate, iar astfel brațul robotului se mișcă corespunzător.



**Fig. 3.14.** Interfața grafică utilizată pentru comanda robotului

Aplicația a fost implementată în limbajul C# și a fost gândită ca o arhitectură client-server. Un cod sursă rulează pe calculatorul gazdă la care este conectat dispozitivul Kinect și produce datele care vor fi utilizate de

către programul care rulează pe robot, așteptând să primească date. În funcție de datele primite, robotul va efectua mișcări conforme cu cele efectuate de către operatorul uman.

Interfața dintre cele două server și client este un socket TCP/IP. Software-ul care rulează pe calculator produce date folosind senzorul Kinect și apoi deschide un socket TCP, încercând să se conecteze la codul care rulează pe robot. Pe de altă parte, codul care rulează pe robot joacă rolul de server. El deschide un socket și așteaptă conexiuni. Odată ce o conexiune este stabilită, deschide un nou socket pentru a servi acestei conexiuni și continuă să asculte.

### **3.3. MODELAREA ROBOTULUI ȘI A MEDIULUI DE LUCRU**

Simularea reprezintă procesul de modelare 3D a robotului și a mediului, inclusiv rularea programelor care vor fi instalate pe modelul fizic al robotului. Acest proces permite utilizatorilor să testeze programele pe o reprezentare virtuală a scenei, economisind timp și având avantajul siguranței. Mediul de simulare poate fi folosit pentru a testa comportamente înainte de executarea lor pe un robot adevărat, cu unele modificări, deoarece robotul simulat și cel real sunt puțin diferiți, după cum se va vedea mai jos. Au fost considerate câteva prezumții legate de robotul NAO simulat.

#### **3.3.1. SUPOZIȚII**

Robotul umanoid NAO va funcționa într-un mediu de lucru în care se află obstacole statice. Sensorii și actuatorii săi trebuie să ofere aceleași funcții ca și pentru robotul real.

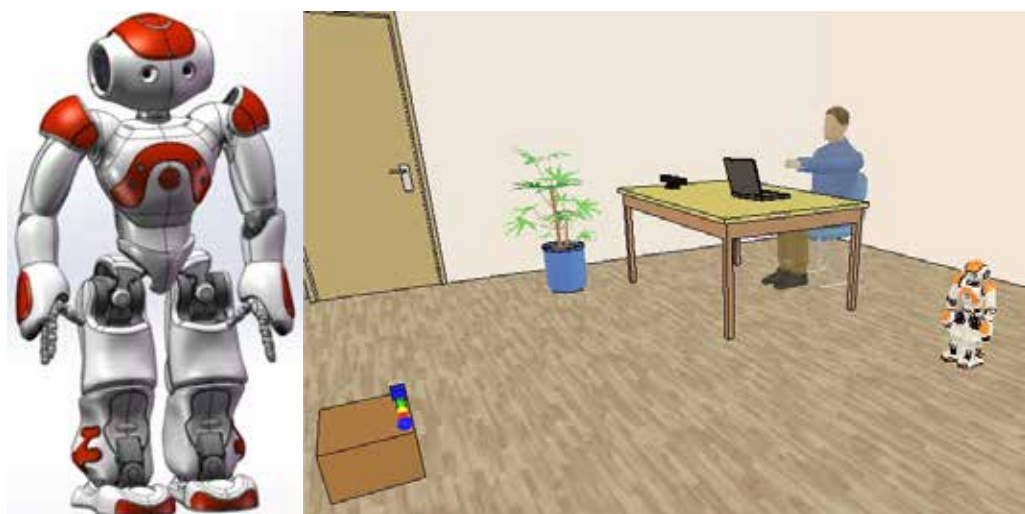
Așa cum s-a menționat anterior, robotul NAO real are 25 de grade de libertate. Robotul simulat are un grad de libertate (DOF) suplimentar deoarece șoldul stâng și cel drept pot fi controlate în mod independent, spre deosebire de robotul real, care are un singur motor care controlează cele două șolduri. Au fost făcute următoarele supoziții:

1. Valorile unghiurilor articulațiilor nu depășesc valorile minimă și maximă specificate în documentația NAO [Internet3.1];
2. Este prevăzut un comportament de evitare a coliziunilor;

3. Simulatorul nu detectează coliziunile dintre diferite părți ale robotului;
4. Robotul simulat nu are un controler pentru fiecare articulație, ca în cazul celui real, ceea ce înseamnă că trebuie implementat un controler.

### 3.3.2. MODELAREA ROBOTULUI ÎN MEDIUL VIRTUAL

Modelarea 3D poate fi folosită pentru a crea un model virtual al robotului și a mediilor virtuale care imită pe cele reale, permițând, de asemenea, programarea robotului în diferite limbaje de programare. Există mai multe instrumente destinate modelării și simulării, cum ar fi Microsoft Robotics Developer Studio (MRDS), Webots, Gazebo. În experimentele din această lucrare este folosit V-REP, un puternic simulator robotic 3D dezvoltat de Coppelia Robotics. Folosind acest instrument, un robot NAO simulat (fig. 3.15a) poate fi lansat într-o lume virtuală, oferind un loc sigur pentru a testa comportamente înainte de a le executa un robot real.



*Fig. 3.15. a) Versiunea simulată a robotului NAO; b) scenariul de test în V-REP*

Programele controlerului robotului pot fi scrise în C/C ++, Python, Java, Lua, Matlab, Octave sau Urbi. Software-ul oferă, de asemenea, posibilitatea de proiectare a robotului și a altor obiecte, folosind un software CAD, cum ar fi SolidWorks, CATIA, AutoCAD, 3D Studio Max și de a importa aceste

modele în V-REP (după ce au fost convertite într-un format cu plase triunghiulare). Proprietățile fiecărui obiect, cum ar fi forma, culoarea, textura, masa, frecarea, etc, pot fi alese de către utilizator. O gamă largă de senzori și actuatori simulați sunt disponibili pentru a dota fiecare robot și este prevăzută, de asemenea, și o interfață cu Kinect.

Cu ajutorul acestui simulator a fost creat un scenariu virtual (fig. 3.15b) asemănător cu cel real, în care a fost introdus robotul simulat. Cu toate acestea, acest scenariu virtual a fost mai puțin utilizat deoarece în condiții reale au fost întâmpinate o serie de dificultăți care nu au putut fi reproduse în mediul virtual.

### 3.4. DEPLASAREA ROBOTULUI

Deplasarea roboților mobili într-un mediu dinamic reprezintă o problemă importantă pentru cercetarea în domeniul roboticii. Pentru ca navigarea să se realizeze în condiții de siguranță în medii necunoscute, robotul trebuie să fie capabil să urmărească o anumită traiectorie, să evite obstacole, să fie conștient de poziția sa și să aibă capacitatea necesară de a se descurca în cazul în care apar situații noi.

Deplasarea robotului prin mediu până la un obiect, odată ce acesta a fost localizat, se face după o traiectorie liniară. Pentru deplasarea robotului au fost folosite de către cercetători diferite tehnici de localizare [Hornung, 2010], [Xu, 2012], [Osswald, 2010], [Fojtů, 2012], în această lucrare a fost utilizată o tehnică bazată pe markeri artificiali NAOmark (ca cel din Fig. 3.16), pe care robotul îi poate detecta și își calculează poziția în funcție de aceștia. În cazul în care un marker nu poate fi identificat, sunt cerute detalii utilizatorului și astfel localizarea autonomă nu este atât de critică, întrucât se realizează prin colaborare cu operatorul uman.

Sistemul de navigație ghidează robotul la locația unde se află obiectul care trebuie apucat. Acolo, prin intermediul sistemului vizual, el caută și apucă obiectul



**Fig. 3.16.** Marker de tip NAOmark

dacă este în raza de acțiune a brațului său. În caz contrar, robotul se apropie de obiect înainte de a-l apuca.

Doi markeri NAOmark sunt utilizați ca repere pentru calcularea poziției curente. Aceste repere sunt plasate pe doi pereți ai încăperii și sunt detectate cu ajutorul camerelor video ale robotului. Procedura pentru detectarea markerilor și localizare este următoarea:

- robotul se ridică în picioare și încearcă să identifice markerii, căutând în jurul său;
- dacă este detectat un singur marker, robotul se mișcă într-o direcție arbitrară și repetă căutarea;
- atunci când au fost detectați doi markeri, sunt utilizate distanțele până la aceștia și distanța cunoscută dintre cei doi markeri pentru a calcula orientarea robotului la un moment dat (față de o axă perpendiculară pe markerul plasat pe biroul utilizatorului);
- cunoscând orientarea, robotul se poate deplasa pe direcția indicată de utilizator.

Evitarea coliziunii cu obstacole fixe sau mobile (de exemplu, cu alți roboți mobili) aflate în spațiul de lucru al robotului se face prin folosirea senzorilor ultrasonici cu care este echipat robotul.

### **3.5. DETECȚIA OBIECTELOR ÎN MEDIU**

Oamenii folosesc întotdeauna obiecte în activitățile lor de zi cu zi. Când vorbesc despre un obiect în mediu, ei arată spre acesta folosind în același timp termeni specifici (de exemplu, „acela”) sau chiar descriind obiectul („mingea roșie”), în scopul de a atrage atenția cuiva față de el. Acesta este un comportament obișnuit folosit peste tot în lume.

După ce sistemul estimează direcția de deplasare, este rândul unității de procesare a imaginii să detecteze obiectele de interes. Robotul se deplasează înainte în direcția indicată pentru căutare de potențiale obiecte. În acest proces datele de la senzorii robotului (camere video, senzori ultrasonici) sunt prelucrate pentru a detecta obiecte sau obstacole. Biblioteca Emgu CV a fost utilizată pentru prelucrarea imaginilor, fiind o platformă care permite apelarea funcțiilor OpenCV în limbajul .NET. Biblioteca permite detectarea obiectelor după culoare și după formă. În această lucrare, a fost folosită o

implementare OpenCV a algoritmului Canny [Canny, 1986] pentru detectarea marginilor obiectelor din scenă. Apoi este utilizată tehnica Hough transform pentru a detecta linii și cercuri în imagine, iar pentru detectarea culorii obiectelor este folosit un algoritm care utilizează modelul de culoare HSV (*Hue, Saturation, Value*), care este considerat un mod mult mai natural decât modelul RGB (Red, Green, Blue) pentru descrierea senzațiile de culoare [Shi, 2013].



**Fig. 3.17.** Obiecte folosite pentru manipulare: sfere și cuburi

Algoritmul caută în imaginile preluate de camera robotului obiecte similare cu cele învățate anterior. Criteriile de comparație sunt forma, culoarea și dimensiunea. Un obiect este considerat a fi recunoscut atunci când imaginea capturată conține culori și forme asemănătoare cu cele ale obiectului care trebuie să fie recunoscut. Pentru atributul formă, este luată în considerare și distanța până la obiect, determinată cu ajutorul senzorilor SONAR.

Obiectele folosite pentru experimente sunt bile și conuri din polistiren și blocuri din lemn colorate în culori diferite (Fig. 3.17). Bilele au diametrul de 5 cm, iar cuburile au laturile de 4 cm lungime, pentru a permite robotului să le apuce. Există și obiecte cu dimensiuni mai mari, iar în acest caz, dacă se depășește o valoare de prag setată pentru fiecare tip de obiect, robotul va ști că acel obiect trebuie apucat cu amândouă mâinile.



### **3.6. CONCLUZII**

În acest capitol au fost prezentate elementele care formează baza pentru cercetarea de față, platforma robotică NAO, senzorul Kinect, precum și componentele software care au fost folosite.

A fost descrisă și interfața folosită pentru comandă și interacțiunea cu robotul mobil și s-au oferit detalii despre deplasarea acestuia în mediu și detecția obiectelor. Pentru navigare și localizare s-a optat pentru utilizarea de markeri NAOmark. Pentru detecția obiectelor s-a folosit un algoritm de detectare a marginilor și a culorii implementat într-o versiune proprie.

Interfața grafică este folosită pentru a realiza învățarea prin imitare. Deși metoda oferă o soluție bună, există încă multe probleme practice care trebuie rezolvate înainte ca această metodă să poate fi utilizată în mod eficient pentru învățare prin imitare sau pentru teleoperare bazată pe gesturi. Prima problemă este cea de limitare a articulațiilor. Din cauza limitării articulațiilor robotului, există multe ipostaze care nu pot fi imitate de către acesta. O altă problemă este cea de corespondență între robot și utilizatorul uman. Robotul are nevoie de timp pentru a mișca un membru al său dintr-o poziție în altă poziție. Acest lucru devine o problemă deoarece utilizatorul uman trebuie să aștepte până când termină robotul de executat gestul curent înainte de a trece la poziția următoare. Acest lucru nu este practic în contextul teleoperării bazată pe gesturi. O posibilă soluție pentru această problemă ar putea fi un filtru de predicție.

# 4

## *SISTEM DE INTERACȚIUNE NATURALĂ OM-ROBOT*

În acest capitol sunt descrise diferitele tehnici folosite pentru a interacționa cu robotul umanoid NAO, precum și abilitățile acestuia. Interacțiunea om-robot implică o mare varietate de provocări, iar una dintre cele mai importante este de a găsi o modalitate simplă, în care fiecare dintre cei doi parteneri de interacțiune să înțeleagă nevoile celuilalt. În ultimii ani a fost observat un interes tot mai mare pentru multimodalitate, ca mod de interacțiune cu roboți și cu alte dispozitive inteligente.

Interacțiunea necesită o comunicare între roboți și oameni și poate fi împărțită în interacțiune de la distanță și interacțiune în apropiere (în proximitate), în funcție de locația fizică în care se află omul și robotul [Goodrich, 2007]. Interacțiunea cu un asistent robotic intră în mod normal în a doua categorie, deoarece robotul și omul se află în aceeași cameră, dar poate fi, de asemenea, o interacțiune de la distanță, atunci când robotul este departe de operatorul uman, folosind tehnici de teleoperare.

Modalitățile de interacțiune om-robot sunt diverse, dar accentul este pus în această lucrare pe recunoaștere vizuală și vocală. Tehnici bazate pe vedere artificială au fost folosite pentru a prelua și interpreta gesturile umane și pentru a detecta obiecte, în timp ce comenzile vocale au fost folosite pentru schimbul de informații între om și robot. Utilizatorul poate da instrucțiuni robotului folosind atât gesturi, cât și vocea, la fel cum oamenii comunică unul cu altul în viața de zi cu zi.

## **4.1. INTERACȚIUNEA PRIN GESTURI**

Un gest este o mișcare corporală făcută în mod intenționat de către un om într-o conversație, cu scopul de a ajuta la o mai bună înțelegere a ceea ce spune. În lucrarea [McNeill, 1982], gesturile sunt clasificate în următoarele categorii: emblematice, iconice, metaforice, deictice și bruște. Deși o clasificare funcțională simplă a gesturilor este, de obicei, înșelătoare [Wagner, 2014], împărțirea în clase este utilă pentru a ușura recunoașterea gesturilor. În lucrarea de față, gesturile nu pot fi strict clasificate în una dintre aceste clase, ele având dimensiuni ale claselor emblematice, deictice, iconice, precum și din clasa celor metaforice. Având în vedere că metafora este omniprezentă în viața de zi cu zi a omului, nu doar în limbaj, dar și în gândire și în acțiune [Lakoff, 2003], este nevoie să apară și în interacțiunea om-robot, iar *dezvoltatorii de roboți trebuie să devină povestitori mai buni* [Bruemmer, 2007].

### **4.1.1. RECUNOAȘTEREA GESTURILOR UMANE**

În comunicarea umană gesturile mâinilor, ale capului și ale brațelor joacă un rol important. Gesturi cum sunt cele de indicare sunt adesea folosite de oameni atunci când doresc să îndrepte atenția cuiva spre un anumit lucru sau spre o anumită direcție. Gestul de indicare este un instrument cognitiv util încă din copilărie [Delgado, 2011]. Alte gesturi ilustrează unele aspecte care sunt dificil de exprimat verbal.

În cercetarea de față s-a plecat de la limbajul semnelor (sau limbajul mimico-gestual), care reprezintă un bun punct de pornire pentru recunoașterea gesturilor. Întrucât gesturile naturale sunt efectuate liber, fără constrângeri, pot apărea în orice moment și în orice ordine, este nevoie de o structurare a posibilităților de exprimare gestuală și de definirea unei gramatici de gesturi, astfel încât să poată fi recunoscute de către robot.

În continuare va fi prezentat algoritmul folosit pentru recunoașterea gesturilor, evidențiind principalele etape ce necesită a fi parcurse.

#### **4.1.1.1. Aspecte generale**

În recunoașterea gesturilor este esențială o comparație între două secvențe. Recunoașterea gesturilor dinamice conține de obicei două componente: segmentare și recunoaștere [Cheng, 2013]. Segmentarea este procesul de a localiza un gest dintr-o secvență video. În această lucrare a fost folosită tehnica de aliniere temporală dinamică (DTW) pentru a evalua similaritatea dintre două secvențe video obținute de la senzorul Kinect. Seriile de timp care vor fi menționate în secțiunea 4.1.2.4 sunt, în acest caz, vectorii de caracteristici extrași așa cum va fi descris mai jos. Pentru fiecare categorie de gesturi, există secvențe predefinite (de exemplu  $x_i$ ). Datele de intrare ( $y_j$ ) sunt comparate cu secvențele predefinite, iar cele două secvențe sunt aliniate pentru a determina calea de cost minim. Acest cost minim reprezintă alinierea optimă între cele două secvențe, ceea ce înseamnă că gestul corespunzător este considerat a fi recunoscut.

O problemă care apare este de a ști când să se înceapă procedura de recunoaștere a gestului, pentru că în cazul în care un gest diferă numai în poziția de început față de secvența predefinită, rezultatul alinierii cu tehnica DTW va fi foarte diferit. Pentru aceasta, s-a folosit o fază de inițializare, care constă într-un cuvânt simplu rostit de către utilizator, prin care robotul este avertizat că utilizatorul dorește să înceapă o interacțiune prin intermediul gesturilor. Când acest cuvânt predefinit este pronunțat, programul începe automat procesul de recunoaștere a gesturilor. Sfârșitul gestului este considerat atunci când brațele utilizatorului se opresc din mișcare.

Recunoașterea gesturilor este o sarcină importantă pentru că oferă o modalitate intuitivă de a comunica în mod natural cu roboții. Există o varietate de tehnici și tehnologii pentru detectarea gesturilor, dar cele mai multe abordări moderne sunt bazate pe sisteme de vedere artificială. Acestea implică una sau mai multe camere video care înregistrează mișcarea unor părți ale corpului uman.

În această lucrare a fost utilizat un dispozitiv Microsoft Kinect pentru a urmări mișcările brațelor utilizatorului. Bibliotecile puse la dispoziție de către Microsoft pentru Kinect oferă poziții ale articulațiilor corpului utilizatorului în timp real. Algoritmul de aliniere temporală dinamică a fost ales pentru recunoașterea gesturilor, deoarece este îndeajuns de robust și suficient de precis pentru nevoile aplicației pentru care a fost implementat.

Algoritmul a fost propus pentru a extinde vocabularul de gesturi, astfel încât să se asigure o interacțiune prin gesturi mai complexă.

Din cadrele video capturate de senzorul Kinect sunt obținuți vectorii de caracteristici care conțin pozițiile a diferite articulații ale corpului. O mișcare este exprimată ca un vector de caracteristici cu 27 de elemente, care conține pozițiile x, y și z ale articulațiile fiecărui braț (mijlocul umerilor, umăr, cot, încheietura mâinii și mână), exprimate în metri în sistemul de coordonate al camerei Kinect. Astfel, un gest este reprezentat de o succesiune de posturi ale brațului în timp, iar procesul de recunoaștere implică compararea secvenței curente cu unele secvențe înregistrate anterior, care formează o „gramatică” a gesturilor înregistrate.

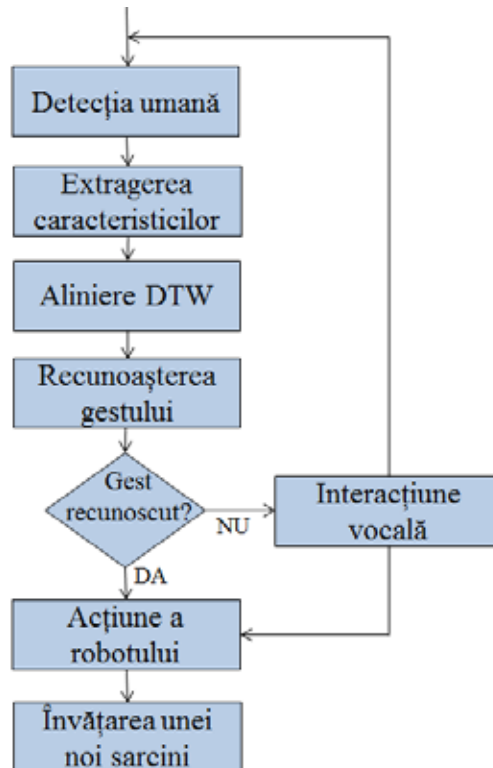
### **4.1.1.2. Structura algoritmului propus**

Diagrama algoritmului propus este prezentată în Fig. 4.1. După cum se poate observa, prima etapă este reprezentată de detecția utilizatorului în mediu. După aceea, sunt extrase caracteristicile, așa cum va fi descris în subcapitolul următor. Algoritmul DTW este aplicat vectorilor extrași și, dacă gestul este recunoscut, robotul va efectua acțiunea solicitată. În caz contrar, se va iniția o interacțiune prin sinteză vocală, prin care robotul îi cere utilizatorului mai multe detalii.

Algoritmul propus de recunoaștere a gesturilor implică așadar 4 etape: 1) detectarea utilizatorului uman, 2) extragerea caracteristicilor, 3) o etapă de aliniere, în care gesturile sunt comparate cu gesturi de referință, 4) recunoașterea gesturilor (Fig. 4.1).

În prima etapă, detecția umană este facilitată de către senzorul Kinect, care poate recunoaște o persoană care se află în fața sa. Senzorul trimite imaginea de profunzime capturată calculatorului, iar software-ul care rulează pe acesta conține logica necesară pentru a decoda informațiile și a recunoaște elementele din imagine cu forme care au caracteristici umane, fiind "pre-antrenat" cu o mare varietate de forme ale corpului. Se folosește alinierea diferitelor părți ale corpului, împreună cu modul în care se mișcă, pentru a le identifica și urmări, reconstruind modelul scheletic al utilizatorului. Astfel, Kinect nu necesită nicio configurare inițială, pentru că o mașină de învățare este instruită să recunoască scheletul. Pozițiile articulațiilor sunt obținute ca în Fig. 4.5. Pentru acest studiu, doar articulațiile brațului sunt relevante, în special articulațiile mâinii,

încheieturii mâinii și cotului (hand, wrist, elbow în Fig. 4.5). Coordonatele acestor articulații formează un vector de caracteristici.



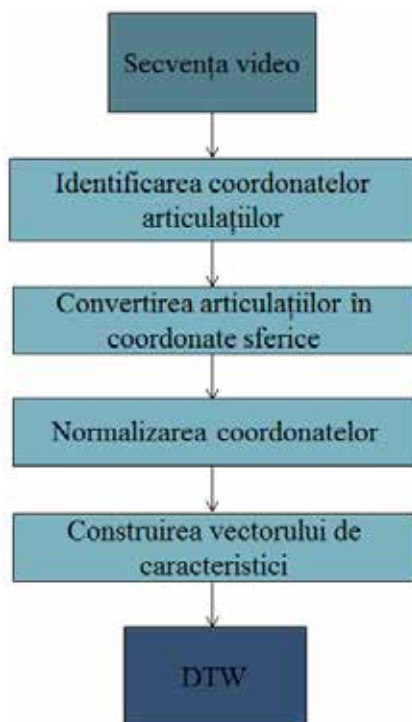
**Fig. 4.1.** Schema logică a algoritmului de recunoaștere a gesturilor

Pentru simplitate s-au făcut două supoziții: în primul rând, s-a presupus că o singură persoană este prezentă în fața sensorului și al doilea, că persoana are o postură șezândă. Inițiativa inițierii unei interacțiuni cu robotul aparține utilizatorului. După etapa de inițializare, sistemul este pregătit să recunoască gestul efectuat de către operatorul uman. Odată ce un gest este recunoscut, în funcție de semnificația sa, sistemul va decide ce sarcini ar trebui să fie efectuate de către robot.

#### 4.1.1.3. Extragerea caracteristicilor

Cea mai importantă caracteristică a unui gest este mișcarea membrilor. În acest caz membrele superioare sunt relevante deoarece sistemul a fost proiectat pentru oameni care stau pe scaun. Mișcarea unui braț poate fi descrisă de traiectoria sa în spațiu. Această traiectorie reprezintă o secvență în timp a pozițiilor brațului.

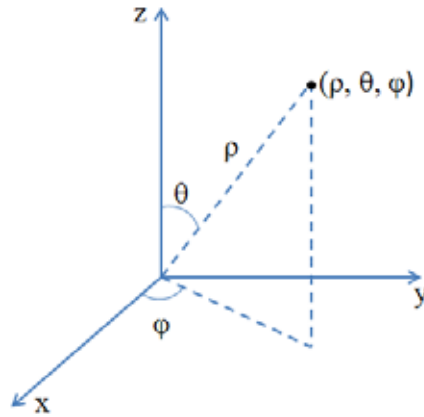
Vectorul de caracteristici capturat de către Kinect conține pozițiile  $x$ ,  $y$  și  $z$  ale fiecărei articulație a brațului (Fig. 4.4). Acest vector este apoi preprocesat în scopul de a-l pregăti pentru calculul DTW. Etapa de preprocesare include eliminarea datelor lipsă sau redundante și alte variații și stabilirea lungimii vectorului. Schema logică care descrie etapele de preprocesare a datelor este ilustrată în Fig. 4.2.



*Fig. 4.2. Schema logică de preprocesare a datelor*

Fiecare gest este reprezentat pe 33 de cadre. Durata gestului efectuat de către utilizator nu are un rol critic deoarece DTW poate compara secvențe de lungimi diferite. Algoritmul realizează alinierea în timp și normalizarea făcând o transformare temporală care permite să se realizeze potrivirea a două secvențe. Acest lucru înseamnă că algoritmul nu este sensibil la durata gestului.

Tot în faza de preprocesare mai sunt realizate două etape, care vor fi explicate în continuare. Înălțimea și dimensiunile fiecărui utilizator sunt diferite. Acest lucru are un impact destul de mare asupra performanței sistemului, motivul fiind coordonatele  $x$ ,  $y$  și  $z$  ale articulațiilor fiecărui utilizator, care pot diferi. Acest lucru se poate întâmpla și ca urmare a poziției diferite a utilizatorului față de Kinect. Așadar, este nevoie de o



**Fig. 4.3.** Coordonate sferice

normalizare a datelor pentru a crește acuratețea de recunoaștere a gesturilor. În acest scop, coordonatele articulațiilor sunt convertite din sistemul de coordonate carteziene (așa cum sunt atunci când sunt capturate de Kinect) în sistemul de coordonate sferice (Fig. 4.3), un sistem tridimensional definit de următoarele trei componente: **distanță radială ( $\rho$ )**, **unghiul azimutal ( $\theta$ )** și **unghiul zenit ( $\phi$ )**. Relațiile de transformare din sistemul de **coordoneate carteziene** în sistemul de **coordoneate sferice** sunt următoarele:

$$\rho = \sqrt{x^2 + y^2 + z^2} \quad (4.1)$$

$$\theta = \arctan \frac{y}{x} \quad (4.2)$$

$$\phi = \arccos \frac{z}{\sqrt{x^2 + y^2 + z^2}} \quad (4.3)$$

Motivul pentru care s-a făcut acest lucru este că în acest fel coordonatele sunt mai ușor de normalizat. Indiferent de dimensiunea utilizatorului, numai distanța de la origine va varia, în schimb toate coordonatele în sistemul cartezian și unghiurile vor rămâne constante. Acum este nevoie doar de normalizarea acestei distanțe.

În loc de a stoca direct coordonatele  $x$ ,  $y$  și  $z$ , toate coordonatele articulațiilor sunt normalizate față de poziția mijlocului umerilor. Această poziție rămâne mereu constantă de-a lungul cadrelor. Astfel, în loc de a

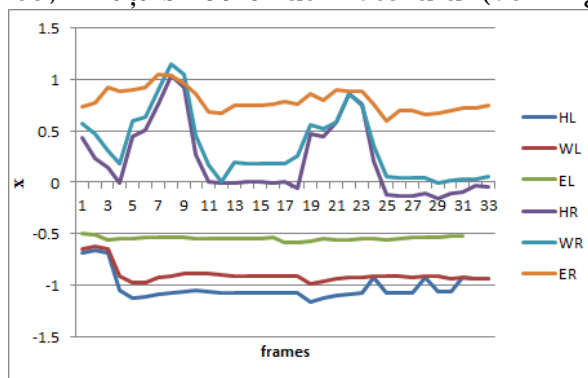


## CAPITOLUL 4

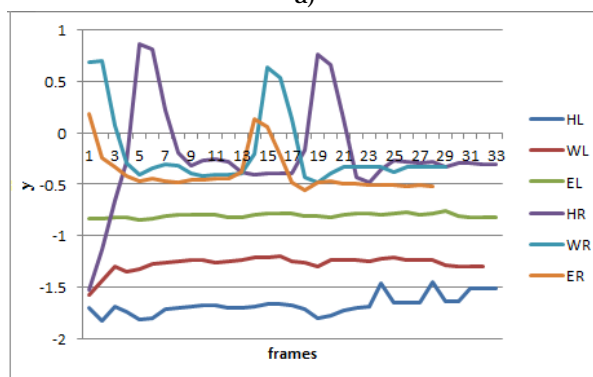
folosi coordonatele carteziene, sunt stocate coordonatele sferice considerând ca origine mijlocul umerilor utilizatorului.

Toate coordonatele sunt apoi normalizate cu distanța totală calculată ca suma distanțelor dintre articulațiile corpului pentru a reduce variațiile cauzate de dimensiunile diferite ale utilizatorilor.

Vectorii de caracteristici care sunt specifici pentru un gest de comandă sunt extrași și apoi stocați într-o bază de date. Distanța minimă de la dispozitivul Kinect pentru o detectare exactă este de 60 cm. Sensorul este plasat pe un trepied, în fața biroului utilizatorului (vezi Fig. 6.8).



a)



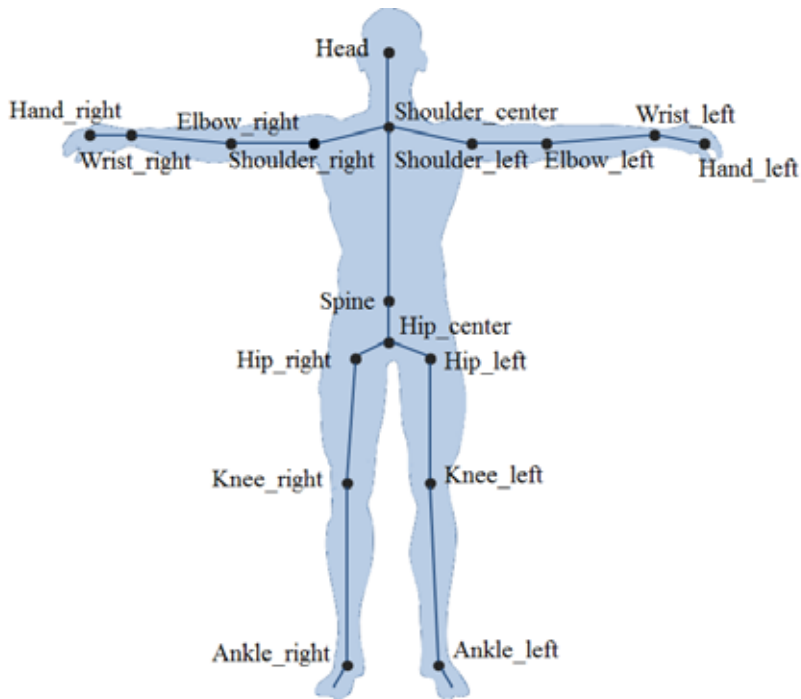
b)

**Fig. 4.4.** Valorile pozițiilor articulațiilor în timp pentru gestul „Triangle” din vectorul de caracteristici: a) coordonatele pe axa x; b) coordonatele pe axa y.

### 4.1.1.4. Metoda DTW (Dynamic Time Warping)

Pentru a detecta gesturile cu o cameră video, se poate utiliza *tehnica de potrivire de șablon* (pattern matching) sau alți algoritmi similari. Potrivirea

de șablon implică utilizarea de eșantioane înregistrate ale gesturilor care servesc ca modele pe baza cărora gesturile detectate pot fi comparate. Un exemplu de astfel de tehnică este alinierea temporală dinamică (DTW). Șablonul reprezintă în acest caz o secvență în timp de măsurători a mișcărilor corpului uman. DTW calculează distanța cumulată între fiecare pereche de valori ale celor două secvențe de timp, dând o măsură de similitudine între cele două secvențe de timp.



**Fig. 4.5.** Articulațiile corpului uman preluate de către dispozitivul Kinect

Astfel, alinierea temporală dinamică (DTW) este o tehnică robustă de căutare a similitudinilor pentru serii de timp [Zhou, 2011]. O privire de ansamblu a acestei metode este dată mai jos.

Având două secvențe reprezentate prin seriile de timp:  $x = x_1, x_2, \dots, x_i, \dots, x_n$  de lungime  $n$  și  $y = y_1, y_2, \dots, y_j, \dots, y_m$  de lungime  $m$ , se poate obține o matrice de  $n \times m$  elemente, unde fiecare element al matricei reprezintă distanța dintre două elemente ale seriilor de timp, denumit *cost al matricei*. Pentru a găsi cea mai bună potrivire între aceste două secvențe, trebuie găsită o cale prin matrice care minimizează distanța totală cumulată între elementele lor, așa cum este ilustrat în Fig. 4.6. Această cale este

## CAPITOLUL 4

denumită *cale de aliniere* (warping path)  $w = w_1, w_2, \dots, w_k, \dots, w_p$  și definește astfel maparea între elementele celor două serii de timp.

$$DTW(x, y) = \min \sum_{k=1}^p d(w_k) \quad (4.4)$$

, unde  $d(w_k)$  reprezintă distanța dintre elementele  $x_i$  și  $y_j$  ale seriilor de timp, adică

$$d(x_i, y_j) = |x_i - y_j| \quad (4.5)$$

Calea de aliniere DTW este constrânsă să urmeze unele restricții, cum ar fi monotonie, continuitate, fereastră de aliniere, limitare a pantei și condițiile de limită (boundary conditions) [Berndt, 1994].

Costul pentru alinierea optimă se obține recursiv prin:

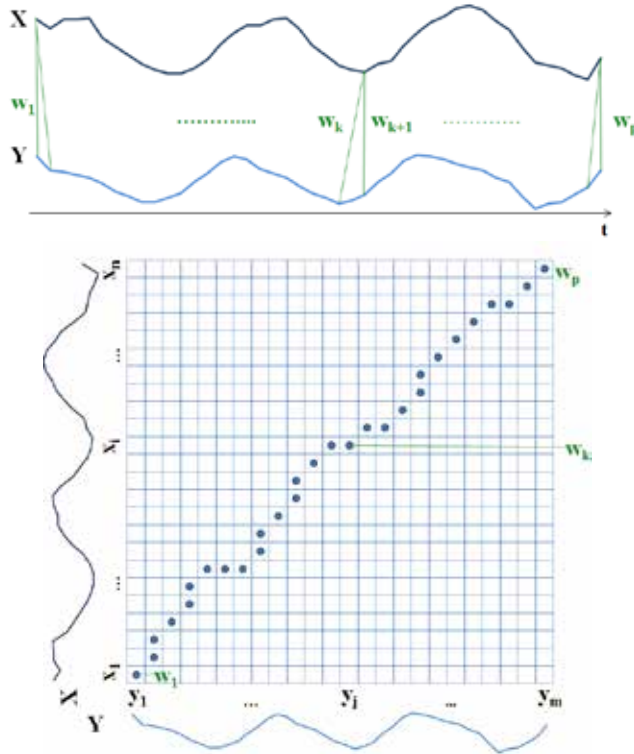
$$\gamma(i, j) = d(x_i, y_j) + \min[\gamma(i-1, j-1), \gamma(i-1, j), \gamma(i, j-1)] \quad (4.6)$$

, unde  $d(x_i, y_j)$  este distanța găsită în celula curentă, iar  $\gamma(i, j)$  este distanța cumulată dintre distanța  $d(x_i, y_j)$  și minimul distanțelor cumulative din trei celule adiacente

De-a lungul timpului au fost propuse diferite îmbunătățiri algoritmului DTW pentru a-l face mai eficient. Metodele folosite pentru a face tehnica DTW mai rapidă se încadrează în trei categorii [Salvador, 2007]:

- constângeri [Sakoe, 1971], [Itakura, 1975], [Yu, 2011], [Adwan, 2012]
- abstractizare a datelor [Salvador, 2007], [Chu, 2002]
- indexare [Yi, 1998], [Kim, 2004], [Keogh, 2005], [Sakurai, 2005], [Lemire, 2009], [Zhou, 2011].

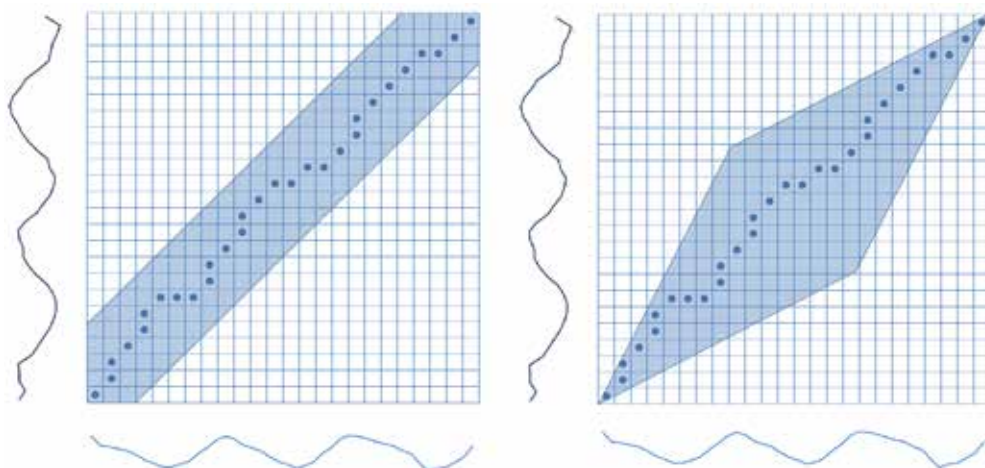
De asemenea există și alte abordări. În lucrarea [Chang, 2011] este prezentată o paralelizare a algoritmului inițial DTW, în scopul de a monitoriza mai multe fluxuri de date, folosind unități de procesare grafică (GPU). Mai recent, o abordare probabilistică a fost propusă în lucrarea [Bautista, 2013]. Unele dintre noile versiuni ale algoritmului DTW au fost dezvoltate pentru a îmbunătăți viteza, în timp ce altele au fost dezvoltate pentru îmbunătățirea preciziei.



**Fig. 4.6.** a) Alinierea a două serii  $(x, y)$  utilizând DTW; b) și maparea dintre ele – calea de aliniere (cu puncte albastre)

Prin tehnica de *constrângere*, calea de aliniere este limitată pentru a preveni abaterile prea mari de la diagonală [Keogh, 2005]. Astfel, calea de aliniere poate fi constrânsă să se încadreze într-un subgrup din matrice, care este numit *ferastră de aliniere* (warping window). Două dintre cele mai utilizate constrângeri sunt banda Sakoe-Chiba [Sakoe, 1971] și paralelogramul Itakura [Itakura, 1975] (Fig. 4.7).

Tehnica de *mărginire inferioară* (lower bound) pentru DTW a fost propusă pentru prima dată în lucrarea [Yi, 1998]. O funcție de mărginire inferioară pentru DTW este o funcție care returnează întotdeauna o valoare mai mică sau egală cu distanța curentă a căii de aliniere DTW. Cea mai cunoscută mărginire inferioară este LB\_Keogh, care folosește calea de aliniere pentru a calcula un *înveliș* (envelope) al acesteia. Versiuni îmbunătățite ale acestei tehnici au fost propuse în [Lemire, 2009], [Zhou, 2011]. În [Kim, 2014] este propusă o mărginire inferioară în două etape, folosind un calcul al matricei cu distanță redusă.



**Fig. 4.7.** a) Banda Sakoe\_Chiba; b) Paralelogramul Itakura

Învelișul de aliniere (warping envelope) a seriei de timp  $x$  este reprezentat de perechea  $U(x)$  și  $L(x)$ , unde:

$$U(x)_i = \max_k (x_k \mid |k-i| \leq \gamma), \quad (4.7)$$

$$L(x)_i = \min_k (x_k \mid |k-i| \leq \gamma), \quad i = 1, \dots, n, \text{ unde } \gamma \text{ este o constrângere locală} \quad (4.8)$$

Funcția de mărginire inferioară  $LB\_Keogh$  este definită astfel:

$$LB\_Keogh(x, y) = \sqrt{\sum_{i=1}^n \begin{cases} (y_i - U_i)^2, & y_i > U_i \\ (y_i - L_i)^2, & y_i < L_i \\ 0, & \text{otherwise} \end{cases}} \quad (4.9)$$

Abordarea  $FastDTW$  calculează aproximări ale căii de aliniere la diferite niveluri de granularitate [Salvador, 2007]. În [Chu, 2002] se propune o metodă de căutare bazată pe aproximarea distanței.

#### 4.1.1.5. Metoda DTW îmbunătățită

În scopul de a satisface cerințele unui sistem robust de recunoaștere a gesturilor pentru interacțiunea cu un robot mobil, s-a propus o versiune îmbunătățită a tehnicii  $DTW$ , care combină mai multe tehnici, așa cum va fi prezentat mai jos.

Complexitatea în timp a algoritmului DTW este  $O(n*m)$  pentru două secvențe de tipul celor prezentate mai sus, ceea ce face ca metoda să nu fie practică pentru serii de timp mai lungi. Deși secvențele folosite în sistemul de recunoaștere (reprezentate de vectorii de caracteristici) sunt mici, se va utiliza algoritmul prezentat în [Salvador, 2007], care este de complexitate  $O(n)$ , atât în timp, cât și în spațiu. Metoda prezentată, numită FastDTW, folosește o abordare pe mai multe niveluri, cu trei etape: înăsprire, proiecție și rafinare. În prima etapă, mărimea seriilor de timp este redusă, făcând o medie a perechilor adiacente de puncte și apoi o cale de aliniere este calculată pentru această rezoluție mai mică, care va fi folosită pentru a găsi calea de aliniere pentru rezoluții mai mari. În cele din urmă, calea de aliniere este rafinată, căutând calea optimă pe fiecare parte a căii proiectate, în funcție de un parametru numit *rază* (radius), care indică numărul de celule care vor fi evaluate.

Tehnica FastDTW a fost ușor modificată. După etapa de înăsprire, calea de aliniere cu distanța minimă a fost obținută folosind o altă tehnică, funcția de mărginire inferioară (lower bound) introdusă în [Lemire, 2009], care oferă o accelerare plauzibilă [Wang, 2013]. Având în vedere seriile de timp prezentate în secțiunea 4.1.2.4, LB\_Improved este definit ca:

$$LB\_Improved(x, y) = LB\_Keogh(x, y) + LB\_Keogh(y, H(x, y)) \quad (4.10)$$

unde  $H(x,y)$  este proiecția lui  $x$  pe  $y$ :

$$H(x, y)_i = \begin{cases} U(y)_i, & x_i \geq U(y)_i \\ L(y)_i, & x_i \leq L(y)_i \\ x_i, & \text{altfel} \end{cases}, i = 1, 2, \dots, n \quad (4.11)$$

Pe baza acestei limitări, accelerarea obținută este de aproximativ 1,4. Pseudo-codul pentru algoritmul modificat este prezentat mai jos.

---

### Algoritmul FastDTW modificat

---

- 1: Integer  $minTSsize = radius+2$
  - 2: IF ( $|X| \leq minTSsize$  OR  $|Y| \leq minTSsize$ )
  - 3: {
  - 4:     RETURN LB\_Improved(X, Y)
  - 5: }
  - 6: ELSE
-

---

```
7: {  
8:   TimeSeries shrunkX = X.reduceByHalf()  
9:   TimeSeries shrunkY = Y.reduceByHalf()  
10:  WarpPath lowResPath = FastDTW(shrunkX,shrunkY, radius)  
11:  SearchWindow window = ExpandedResWindow(lowResPath, X, Y,  
    radius)  
12:  RETURN LB_Improved(X, Y, window)  
11: }
```

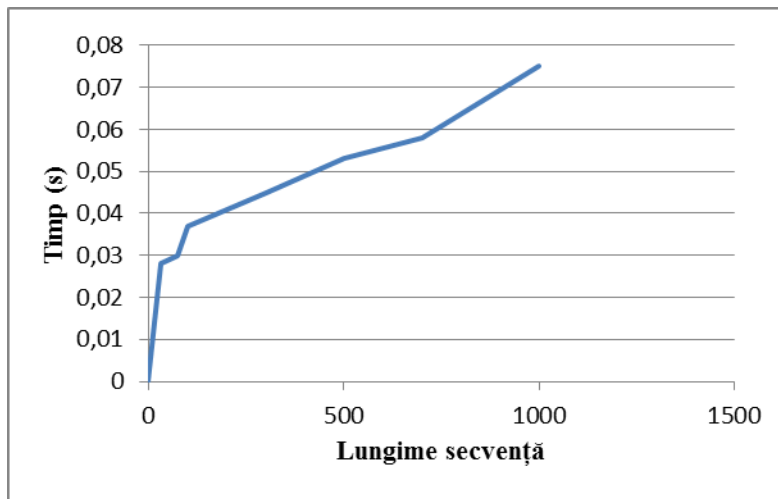
---

DTW compară secvența obținută pentru un gest necunoscut cu una sau mai multe modele sau șabloane de referință. Având mai multe astfel de șabloane, rata de recunoaștere va fi mai mare, dar timpul de calcul crește, de asemenea. Din acest motiv, este utilizată o abordare implementată în lucrarea [Zaharia, 2010] pentru recunoașterea vorbirii. Acest algoritm, numit DTW cuantificat, stochează un model de referință pentru fiecare gest și a fost adaptat pentru recunoaștere gesturilor.

DTW cuantizat, împreună cu FastDTW și LB\_Improved au fost combinate pentru a obține un algoritm de recunoaștere a gesturilor mai rapid și mai precis. Timpul de execuție pentru o secvență de timp de lungime mare, cu o rază de 10 este prezentată în Fig. 4.8.

#### 4.1.1.6. Concluzii

În literatura de specialitate există diferite implementări pentru recunoașterea gesturilor umane folosind tehnica de aliniere temporală dinamică. În prezentul capitol a fost prezentată o abordare care combină mai multe tehnici, contribuind la robustețea și eficiența algoritmului pentru o mai bună și mai sigură recunoaștere a gesturilor care vor fi folosite pentru comanda robotului umanoid.



**Fig. 4.8.** Timpul de execuție pentru DTW îmbunătățit

#### 4.1.2. BIBLIOTECA DE GESTURI

În domeniul HRI, întrucât roboții de asistență interacționează cu utilizatori neexperimentați, interfețele naturale sunt esențiale. La fel ca și în interacțiunea dintre oameni, un gest poate furniza informații sau poate exprima intenția utilizatorului pentru a fi înțeleasă de către robot. În acest sens a fost creat un set de gesturi, care reprezintă biblioteca de gesturi utilizată pentru interacțiunea cu robotul. În tabelul 4.1 sunt prezentate gesturile definite. O parte a gesturilor sunt inspirate din lucrarea [Kipp, 2004].

**Tabelul 4.1** Biblioteca de gesturi

Numele gestului	Abrev.	Descriere	Semnificație
Attention	A	O mână ridicată	„Hey!”
Big	B	Ambele brațe la nivelul capului cu o distanță mare între ele	„Un obiect mai mare!”
Break	Br	O mână plasată perpendicular pe cealaltă	„Pauză!”
Circle	O	Desenarea unui cerc în spațiu cu mâna	„Un obiect de forma aceasta”
Come	C	Mâna se mișcă în mod repetat din exterior spre corp	„Vino aici!”



## CAPITOLUL 4

Despair	D	Ambele mâini ridicate la nivelul capului	„Ce ai făcut?”
Doubt	Do	Brațele deschise într-un arc	„Nu știu”
Shrug		înspre exterior	
Head nod	HN	Mișcarea capului pe verticală de sus în jos o dată sau de mai multe ori	Acceptare
Head shake	HS	Capul întors spre stânga și dreapta în mod repetat	Negare
Left	L	Brațul stâng ridicat la nivelul umerilor, în partea stângă a corpului	„Ia-o la stânga!”
Object	Ob	Indicarea unui obiect cu brațul	„Acel obiect!”
Rectangle	R	Desenarea unui dreptunghi în spațiu cu mâna	„Un obiect de forma aceasta”
Refuse	Re	O mână deplasată spre exterior într-o mișcare de ștergere	Refuzare
Right	R	Brațul drept ridicat la nivelul umerilor în partea dreaptă a corpului	„Ia-o la dreapta!”
Small	S	Ambele mâini la nivelul capului cu o distanță mică între ele	„Un obiect mai mic”
Space	Sp	O mână indică un punct în spațiu	„Mergi acolo!”
Sway	Sw	Mâinile alternează într-o mișcare sus - jos	„Continuă să mergi!”
To-Fro	TF	Mâinile se mișcă dintr-o parte în alta a corpului	„Mută obiectul”
Triangle	T	Desenarea unui triunghi în spațiu cu mâna	„Un obiect de forma aceasta”
Turn left	TL	Ambele mâini imită rotația unui obiect în sens invers acelor de ceasornic	„Întoarce-te spre stânga cu 15°!”
Turn right	TR	Ambele mâini imită rotația unui obiect în sensul acelor de ceasornic	„Întoarce-te spre dreapta cu 15°!”
Wave	W	O mână întinsă, îndreptată în sus, efectuând mișcări laterale ușoare	Atrage atenția robotului
Wipe	Wi	Ambele mâini plecând din poziție apropiată una de alta și	Terminare

		îndepărtându-se într-o mișcare liniară	
X	X	Mâinile încrucișate	Închiderea aplicației

---

O atenție specială a fost acordată gestului de indicare pentru că este o cale mai ușoară de a atrage atenția robotului, indicându-i un obiect sau un loc în spațiu și este util pentru utilizatorii neexperimentați. Odată ce gestul a fost detectat, următorul pas este să se evalueze direcția de indicare. Pentru acest lucru, trebuie calculat unghiul dintre brațul utilizatorului și restul corpului. Trei articulații din scheletul trupului descriu acest gest: mijlocul umerilor, umărul și mâna (Figura 4.7). Unghiul estimat a fost calculat folosind următoarea formulă:

$$\alpha = \text{acos} \frac{v1.x*v2.x+v1.y*v2.y}{\sqrt{v1.x^2+v1.y^2}*\sqrt{v2.x^2+v2.y^2}} * \frac{180}{\pi} \quad (4.12)$$

,unde v1, v2 sunt următorii vectori de poziție:

$$v1 = \overline{s - h} \quad (4.13)$$

$$v2 = \overline{s_c - s} \quad (4.14)$$

, unde s = shoulder (left sau right) , h = hand, s<sub>c</sub> = shoulder\_center.

Gestul de indicare este folosit numai atunci când robotul este în poziția inițială, cunoscând orientarea și distanța până la Kinect. În caz contrar, el nu va ști în ce direcție să se deplaseze.

Toate gesturile sunt efectuate cu brațele, cu excepția a două: confirmare (head nod) și negare (head shake), care sunt efectuate prin mișcări ale capului. Aceste gesturi au fost folosite pentru că sunt cele mai frecvent utilizate în comunicarea interpersonală, atunci când oamenii acceptă sau resping un anumit lucru, o anumită idee. Pentru aceasta a fost utilizat un algoritm simplu de urmărire a capului pus la dispoziție de Microsoft Kinect, care ia în considerare rotația capului în plan sagital și transversal.

Unele gesturi au senzori diferite, în funcție de context. De exemplu, atunci când utilizatorul spune „Turn left” și robotul este în mișcare, se va schimba direcția de mers la stânga, cu 15°. Dacă robotul nu este în mișcare,

## CAPITOLUL 4

aceeași comandă se va referi la camera robotului, iar atunci el va roti capul cu 15° spre stânga.

### 4.2. INTERACȚIUNEA PRIN VOCE

Oamenii folosesc *vocea* pentru a interacționa între ei. Cercetătorii în domeniul roboticii au început să folosească de asemenea acest mod de interacțiune pentru comunicarea cu roboții. Recunoașterea vocii este procesul care transformă sunetele vocale în înțelesul acestor sunete, transformând limba vorbită în limba scrisă sau cunoștințe simbolice [Topoleanu, 2011].

În continuare, va fi prezentat procesul de identificare și recunoaștere a expresiilor vocale folosite pentru a putea comunica cu robotul mobil.

#### 4.2.1. METODA DE RECUNOAȘTERE A VOcii

Pentru captarea vocii umane a fost folosit același dispozitiv (Kinect). Kinect for Windows SDK permite utilizarea API-ului Microsoft.Speech recognition. Fluxul audio este capturat de către matricea de microfoane a dispozitivului Kinect, folosind un motor de recunoaștere a vocii. Microsoft.Speech a fost folosit pentru a crea gramatici, care pot recunoaște un singur cuvânt sau o frază scurtă. În funcție de contextul de interacțiune, există diferite gramatici implementate pentru fiecare situație în parte. Fiecare gramatică are mai multe reguli care definesc un model sau o secvență de cuvinte. Un exemplu pentru gramatica utilizată pentru comenzile inițiale Point-and-comand este următoarea:

<Nume Robot> <please (op)> <verb1> <deixis spațial> <**and**> <verb2>  
<articol (op)> <proprietate> <nume>

*Nume robot* sugerează numele robotului. Acesta poate fi NAO 1, NAO 2 sau alte nume dacă operatorul vrea să îi numească cu nume specifice și indică faptul că operatorul dorește să trimită o comandă robotului. În cazul în care expresia vocală nu conține acest cuvânt-cheie, aceasta va fi ignorată. *Verb1* conține un set de verbe ca „go” sau „look at”, iar *deixis-ul* spațial conține cuvinte ca: „there”, „that”. Conjuncția „and” este scrisă cu caractere

îngroșate doar pentru a sugera ideea Point-**and**-comand, care va fi prezentată în subcapitolul 4.4. Variabila *verb2* sugerează un set de acțiuni: „fetch”, „grasp”, „push”; articol este opțional (op) și poate fi un pronume demonstrativ ("that") sau un articol hotărât („an”, „a”); *proprietate* se referă la proprietatea unui obiect: culoare („red”, „blue”, „green”), formă („circle”, „square”), iar *nume* desemnează un obiect cunoscut sau un substantiv („sphere”, „cube”).

Sistemul de recunoaștere a vocii oferă, de asemenea, un *nivel de încredere* (confidence level) pentru fiecare cuvânt recunoscut. În cazul în care este detectat un cuvânt, dar are un nivel de încredere foarte scăzut, robotul va înștiința operatorul uman, cerându-i să repete comanda.

În blocul de management al dialogului ilustrat în Fig. 5.1 există mai multe sub-funcții. În afară de a gestiona dialogul dintre om și robot, acest bloc este responsabil, de asemenea, pentru combinarea cuvintelor rostite cu gestul de indicare, descifrând semnificația comezii complexe. Acest lucru este realizat printr-un motor de inferență. Mai mult decât atât, repartizarea sarcinilor se face, de asemenea, în acest bloc.

Interacțiunea dintre om și robot urmează un curs specific. Inițiativa de a începe o interacțiune aparține utilizatorului. El trebuie să pronunțe numele robotului sau să efectueze un gest (de exemplu, ‘wave’) cu scopul de a atrage atenția robotului. După aceea, conexiunea este stabilă, iar robotul este pregătit să primească următoarele comenzi.

#### 4.2.2. COMENZI VOCALE

La fel ca în cazul interacțiunii prin gesturi, a fost creat un set de comenzi vocale. În tabelul 4.2 sunt prezentate expresiile verbale de bază utilizate în interacțiune, dar nu este un tabel complet, deoarece unele dintre ele pot fi combinate pentru a forma expresii predefinite, așa cum s-a arătat în subcapitolul precedent.

Pentru a începe interacțiunea cu robotul, este necesar ca utilizatorul să rostească numele acestuia („NAO”), în scopul de a-l anunța că utilizatorul se adresează lui și de a lua în considerare doar comenzile vocale care îi sunt adresate, nu și alte expresii vocale sau sunete pe care le folosește utilizatorul în alte scopuri.

**Tabelul 4.2** Vocabularul pentru expresii vocale

Comandă	Abreviere	Semnificație
NAO!	N	Se dorește începerea unei interacțiuni
Stand up!	SU	Ridicarea robotului din poziția șezândă
Sit down!	SD	Robotul este așezat în poziția de repaus
Go there!	GT	Robotul se deplasează în direcția indicată
Go left!	GL	Întoarce-te cu 90° spre stânga și continuă deplasarea
Go right!	GR	Întoarce-te spre dreapta cu 90° și continua deplasarea
Turn left!	TL	Întoarce-te spre stânga cu 15°
Turn right!	TR	Întoarce-te spre dreapta cu 15°
Grab the object!	GO	Robotul apucă în mod autonom un obiect
Drop object!	DO	Robotul eliberează obiectul din mâna sa
Leave object!	LO	Robotul întinde brațul și deschide mâna
Let me control you!	C	Teleoperare prin gesturi
Learn this!	LT	Robotul învață o nouă mișcare
Bring it to me!	B	Robotul duce un obiect utilizatorului
Open hand!	OH	Robotul deschide mâna
Close hand!	CH	Robotul închide mâna
Thank you!	TK	Robotul așteaptă următoarea comandă
Yes!	Y	Acceptare
No!	N	Negare
Stop!	S	Robotul se oprește din acțiunea pe care o efectuează
Exit!	E	Oprește aplicația

Sistemul de comandă vocală a fost creat folosind Kinect for Windows SDK, combinat cu biblioteca Microsoft Speech Recognition (MSR) API.

Kinect SDK oferă diverse capacități audio, iar platforma Microsoft Speech oferă clase pentru a lucra cu recunoașterea semnalului vocal capturat de senzor Kinect, convertind cuvinte rostite în text scris. Dispozitivul poate detecta semnalul audio care se află între  $\pm 50$  grade în fața senzorului și, de asemenea, suportă până la 20 dB de anulare a zgomotului ambiental.

Microsoft Speech Recognition are gramatici și vocabulare avansate și nu necesită antrenare pentru modele. Utilizatorul trebuie să creeze gramatica sa folosind cuvintele cheie dorite. Un handler Kinect va inițializa fluxul audio și va începe captura audio. După ce motorul de recunoaștere a vorbirii este pornit, utilizatorul va încărca gramatica și din acel moment sistemul este pregătit să recepționeze date de la Kinect.

### **4.3. SISTEMUL DE INTEGRARE GESTURI/VOCE**

Modulele de recunoașterea a vocii și de recunoaștere a gesturilor rulează simultan. După comanda „Attention” (prin gest) sau „NAO!” (prin voce), sistemul așteaptă următoarea comandă care poate fi exprimată, de asemenea, prin gest sau prin voce. Există patru posibilități care rezultă din combinarea modalităților de interacțiune: doar gest (G), gest + voce (GV), voce + gest (VG), numai voce (V). Fiecare comandă este trimisă sistemului de decizie, care se bazează pe diferite reguli și, în conformitate cu aceste reguli, sarcina care trebuie să fie efectuată este identificată. În cazul în care comenzile sunt GV sau VG, sistemul decide dacă comanda prin gest este compatibilă sau nu cu cea vocală. Sarcinile sau acțiunile care au fost implementate sunt următoarele: navigare (N - navigation), aducere (F - fetch), apucare (G - grabbing), împingere (P - push) și teleoperare (T - teleoperation).

Motorul de inferență conține reguli care ajută la identificarea sarcinii pe care robotul trebuie să o efectueze. Intrările pentru motorul de inferență sunt reprezentate de comenzile vocale/gestuale, iar ieșirea este reprezentată de sarcina ce se dorește a fi executată de către robot. Așa cum s-a menționat la începutul capitolului, recunoașterea gesturilor și a vocii sunt procese separate, dar rulează simultan. După ce un gest sau o expresie vocală este recunoscută, numele său este stocat într-o variabilă, care va reprezenta intrarea pentru motorul de inferență. Astfel, în funcție de aceste două variabile de intrare, motorul de inferență returnează sarcina necesară, în conformitate cu setul de reguli.

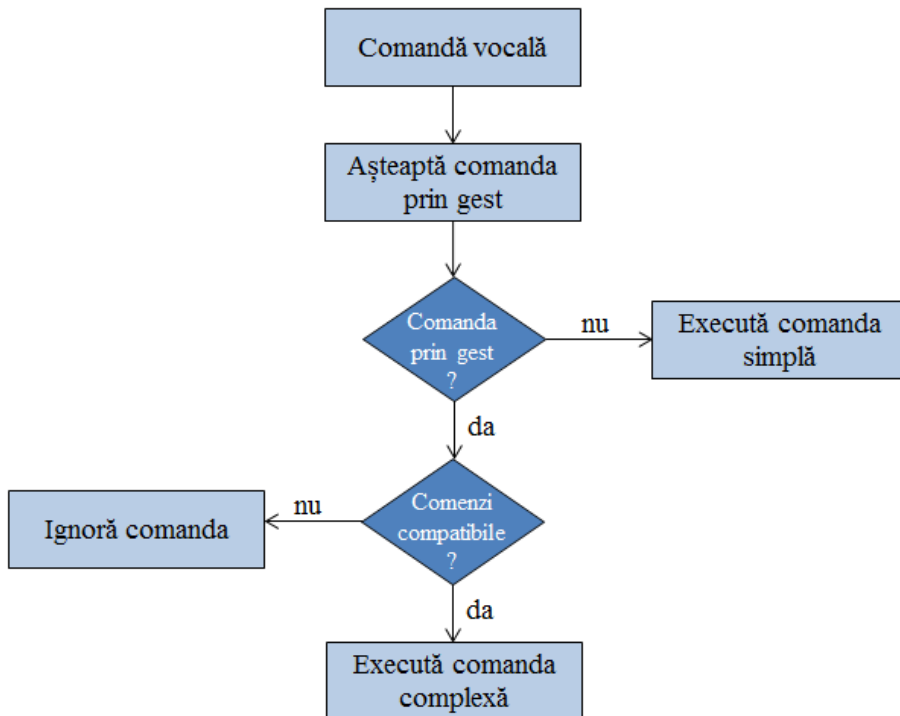
## CAPITOLUL 4

Există sarcini simple (de exemplu, navigație, apucare) sau sarcini mai complexe (cum ar fi fetching). Sarcinile simple pot fi solicitate prin intermediul unei singure modalități de interacțiune, dar sarcinile complexe au nevoie de fuziunea dintre cele două modalități, ca în cazul paradigmei Point-and-comand care va fi prezentată mai jos. După o comandă preliminară („NAO!“), sistemul de detecție „ascultă” continuu. Apoi, după o primă comandă, sistemul așteaptă un interval de timp stabilit la 4 secunde pentru o altă comandă, iar în cazul în care aceasta nu este dată, se ia în considerare doar prima comandă. În cazul în care comanda constă din combinații vorbire/gest, dar ele sunt incompatibile, comanda complexă este ignorată, iar robotul va răspunde prin comportamente predefinite (Fig. 4.15) sau prin sinteză vocală. În caz contrar, sistemul decide ce acțiune să efectueze robotul, în funcție de comanda dată prin una dintre cele două modalități sau prin ambele modalități simultan.

Modul de funcționare a sistemului descris mai sus este ilustrat în diagrama din Fig. 4.9. În figură se prezintă cazul în care prima comandă este vocală, dar același lucru se petrece și în cazul în care prima comandă este executată prin intermediul gesturilor.

Următoarele reguli constituie o parte a bazei de cunoștințe și exprimă modul în care sistemul trebuie să reacționeze:

```
IF <comandă gest> is C AND <comandă vocală> is GL THEN <sarcina> is N
IF <comandă gest> is Sp AND <comandă vocală> is B THEN <sarcina> is F
IF <comandă gest> is Ob și <comandă vocală> is GO THEN <sarcina> is G
IF <comandă gest> is A și <comandă vocală> is C THEN <sarcina> is T
```



**Fig. 4.9.** Diagrama de funcționare a sistemului de integrare gesturi/voce

Pentru sarcina de navigare, a fost implementat un sistem de inferență fuzzy, similar cu cel prezentat în lucrarea [Pires, 2002]. În cazul de față, variabilele de intrare sunt datele primite de la senzori și, dacă este necesar, și o variabilă de orientare. Există două variabile care păstrează date de la senzorii robotului, fiecare corespunzând unui senzor ultrasonic. Variabila de orientare este unghiul rezultat din gestul de indicare ( $\alpha$ ). Un set de ieșire fuzzy rezultă din procesul de inferență, care este direcția de deplasare, reprezentat sub forma unui unghi ( $\theta$ ).

Câteva reguli ale bazei de cunoștințe sunt următoarele:

IF  $d1$  is VN and  $d2$  is F and  $\alpha$  is R THEN  $\theta$  is R

IF  $d1$  is VN and  $d2$  is F and  $\alpha$  is L THEN  $\theta$  is F, unde VN = Very Near, F = Far, R = Right, iar  $d1$  și  $d2$  reprezintă distanțele până la obstacol obținute de la senzorii ultrasonici.

Datele necesare pentru identificarea de obiecte sunt obținute prin dialog, cum va fi arătat în capitolul 6.

Robotul poate învăța noi sarcini în felul următor: când utilizatorul rostește comanda „Learn this thing”, sistemul intră în modul de imitare. Utilizatorul îi demonstrează o mișcare și apoi va rosti un nume, care este



numele acelei mișcări. Sistemul va stoca secvența de mișcări și o va asocia cu numele dat de către utilizator. Apoi, utilizatorul îi poate indica sistemului comanda care va activa acea sarcină.

#### 4.4. PARADIGMA POINT-AND-COMMAND

Utilizarea gesturilor de indicare reprezintă unul dintre primele moduri de comunicare cu lumea pentru oameni [Colonnesi, 2010], fiind o piatră de temelie fundamentală a comunicării umane [Kita, 2003]. Indicarea prin gesturi a fost folosită în primele faze ale dezvoltării limbajului combinat cu vorbirea pentru a denumi obiecte, indicând o corespondență timpurie între cuvânt și gest [Roth, 2001]. Gesturile de indicare sunt gesturi deictice [McNeill, 1992] utilizate pentru a orienta atenția unui observator spre o locație sau asupra unui obiect sau pentru a indica o direcție sau un eveniment. Clasa de gesturi deictice include un set mai mare de gesturi, care sunt folosite pentru a atrage atenția asupra unui obiect [Sauppé, 2014a]. În această lucrare este utilizat gestul indicare, care este cel mai important mod de a comunica în timpul fazelor incipiente ale copilăriei, prin poziționarea mâinii spre o anumită locație.

Gesturile nu pot fi separate de vorbire. Acestea sunt *elemente sau instrumente ale limbajului*, cum spune renumitul filozof Wittgenstein [Wittgenstein, 1965]. Mai mult decât atât, gestul și expresiile vocale lucrează împreună pentru a transmite un singur mesaj integrat atât în timp, cât și în semnificație la adulți [Butcher, 2000] și gesturile de indicare sprijină învățarea limbii pentru copii [Goldin-Meadow, 2007]. Sunt identificate mai multe interpretări ale gesturilor de indicare de către psihologii de dezvoltare [Kita, 2003], dar două dintre cele mai evidente sunt gesturile de indicare imperative și cele declarative [Scassellati, 2003]. Ele sunt folosite în locul deicticelor vorbite sau pentru completarea lor. Este cunoscut faptul că indicarea unui obiect de referință este o metodă mult mai rapidă și convenabilă decât descrierea lui verbală.

În literatura de specialitate se folosește denumirea de *deixis* pentru a indica coordonatele spațiale și temporale care sunt implicate în exprimarea orală. Acest termen provine din limba greacă și este definit în felul următor [Lyons, 1977]:

Prin *deixis* se înțelege localizarea și identificarea de persoane, obiecte, evenimente, procese și activități despre care se vorbește, sau la care se face

referire, în raport cu contextul spațio-temporal creat și susținut prin actul de exprimare și participarea, de obicei, a unui singur vorbitor și cel puțin un destinatar.

În lucrarea [Kettebekov, 2001] deixis-urile sunt clasificate, în funcție de semantica lor, în două categorii: tranzitive și intransitive. Cele tranzitive cuprind actele gestuale care fac referire la obiecte concrete și nu este specificată nicio schimbare fizică a locației, iar cele intransitive cauzează deplasarea spațială. La rândul lor, deixis-urile tranzitive sunt clasificate în aceeași lucrare în funcție de informațiile verbale complementare, după cum urmează:

- nominale – referirea la un obiect se face printr-un substantiv sau pronume;
- spațiale – referirea la o zonă în care se află obiectul, completată de adverbe de loc;
- iconice – precizează și atribute ale obiectului (de exemplu, forma, culoarea).

Atunci când are de-a face cu roboți de asistență, utilizatorul ar trebui să fie în măsură să îi informeze despre locația unui obiect. Această interacțiune este crucială în mediile domestice, pentru că cele mai multe dintre sarcinile îndeplinite de robot necesită manipulare de obiecte [Boboc, 2014b]. În acest sens, *gestul de indicare* este recunoscut ca un mod esențial de interacțiune cu roboții. Comunicarea deictică ar trebui să fie bidirecțională. Operatorul folosește gestul de indicare combinat cu indicii verbale, iar robotul cere detalii suplimentare dacă ceva nu este clar.

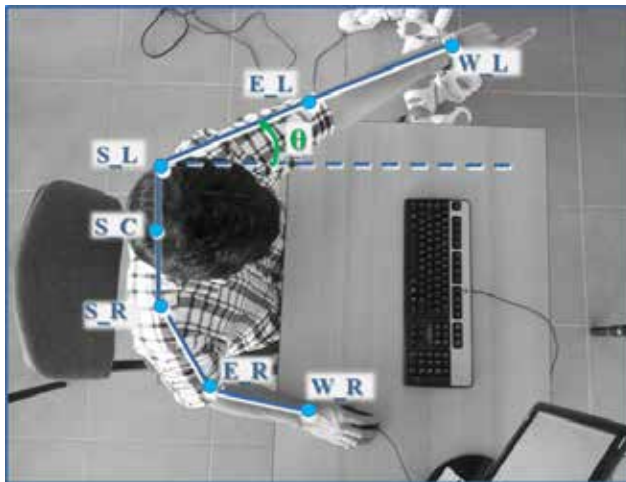
Au existat multe lucrări de cercetare legate de interacțiunea cu roboți folosind comunicarea deictică. În unele dintre ele un robot este folosit ca echipament de testare pentru a investiga aspecte specifice legate de dezvoltarea socială a oamenilor [Scassellati, 2003], [Nagai, 2005]. Alte lucrări au încercat să dea posibilitatea unui robot să genereze propriile sale gesturi de indicare [Sauppé, 2014a], [Chao, 2014], atât omul, cât și robotul folosind gesturi de indicare în interacțiune [Beuter, 2008], sau doi roboți folosind aceste gesturi [Hafner, 2005], dar cele mai multe lucrări se axează pe capacitatea roboților de a răspunde la gesturile de indicare efectuate de operatorii umani. Unele lucrări se concentrează pe metodologia de recunoaștere și estimare a direcției de indicare [Breuer, 2010], [Droeschel, 2011], pe partea de control a sistemului [Schiffer, 2012] sau pe procesul de

realizare a unei comunicări deictice naturale între roboți și oameni [Sugiyama, 2007]. Va fi propusă mai jos o soluție simplă de comunicare deictică, care îmbină recunoașterea vizuală și recunoașterea vocii folosind un singur dispozitiv, această paradigmă fiind denumită *Point-and-Comand* (PaC).

Există o tendință în creștere față de dezvoltarea de roboți care asistă oamenii în tot mai multe domenii de aplicații. Acest lucru se datorează în parte roboților personali care au invadat practic piața de cercetare și pe cea educațională și se presupune că vor invada viețile noastre în scurt timp. O mulțime de roboți lucrează deja în spațiul de locuit al omului, în spitale, în centre expoziționale, în universități sau muzee. Dar pentru a fi un asistent bun, robotul ar trebui să aibă competențe sociale și cognitive. De aceea, este important să fie construit astfel încât să fie capabil să interacționeze cu oamenii într-un mod specific acestora.

Oamenii folosesc gestul de indicare în comunicare ca un mecanism de indicare a unui obiect de referință sau pentru a informa ascultătorul cu privire la poziția acestuia. Cu alte cuvinte, gestul de indicare este folosit pentru a stabili identitatea sau locația spațială a unui obiect în cadrul domeniului de aplicare [Karam, 2005]. Gesturile de indicare pot fi utilizate, de asemenea, de către oameni pentru a furniza informația de direcție roboților [Bauer, 2009]. Ele sunt utile mai ales pentru utilizatorii care au nevoie să interacționeze cu roboți mobili, dar care nu au o cunoaștere a priori despre HRI.

În lucrarea de față a fost exploatată ideea de a combina gesturi de indicare umane cu expresii vocale pentru a comunica intențiile unui robot, într-o abordare centrată pe utilizator, numită *Point-and-comand* (PaC). În termeni simpli, ideea procesului de interacțiune este următoarea: o persoană indică un obiect (sau o locație) și rostește o expresie (de exemplu, „apucă acel obiect”), iar robotul se va deplasa în direcția indicată și va efectua sarcina (identifică obiectul și îl apucă). În [Sugiyama, 2007] sunt prezentate cinci procese care sunt implicate în comunicarea deictică naturală între roboți și oameni: focalizare pe context, sincronizarea atenției, recunoașterea obiectelor, stabilirea credibilității și indicarea obiectului.



**Fig. 4.10.** Articulațiile corpului operatorului obținute de la senzorul Kinect și unghiul estimat (cu verde)

În cazul de față, scenariul este conceput pentru profani, care au nevoie de o interfață ușor de utilizat. Un utilizator care este așezat la birou sau o persoană cu deficiențe de mobilitate poate folosi vocea și părțile superioare ale corpului pentru a cere unui robot mobil personal să îi aducă un obiect din cameră sau din spațiul de operare al acestuia. Prin urmare, ei trebuie să indice care obiect este avut în vedere sau care este punctul de interes (PoI) [Bauer, 2009]. Indicația nu este explicită; acesta este doar un indiciu gestual și vocal și, ca urmare a acestui fapt, acest mod de interacțiune este considerat a fi cu comenzi de nivel înalt, întrucât asumă un anumit nivel de autonomie din partea robotului [Jones, 2010].

Direcția de indicare este dată calculând unghiul dintre brațul operatorului și corpul său. Interfața software a dispozitivului Kinect oferă coordonatele a 20 de articulații (vezi Tabelul 4.3 și Fig. 4.5), dar în cazul de față doar 7 sunt relevante, care sunt ilustrate în Fig. 4.10. Coordonatele umărului și ale încheieturii mâinii au fost folosite pentru a crea o linie îndreptată spre țintă. Accentul nu este pe calculul foarte precis al unghiului de indicare deoarece această informație este folosită doar pentru a da o direcție inițială robotului, iar acesta va detecta ținta în mod automat sau semi-automat (cerând noi detalii). Acesta este motivul pentru care s-a ales să nu se folosească metode mai precise de urmărire (de exemplu, detectarea degetului arătător combinat cu detectarea feței, ca în [Breuer, 2010]).

**Tabelul 4.3.** Articulațiile corpului preluate de dispozitivul Kinect

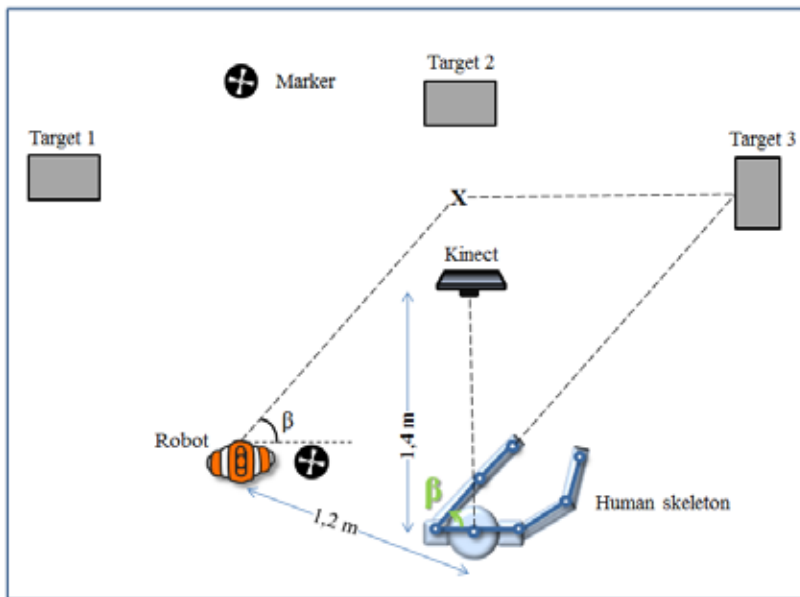
Denumire	Abreviere	Correspondent română engleză
Head	H	Cap
Shoulder_center	S_C	Mijlocul umerilor
Shoulder_left	S_L	Umăr stâng
Shoulder_right	S_R	Umăr drept
Elbow_left	E_L	Cot mâna stângă
Elbow_right	E_R	Cot mâna dreaptă
Wrist_left	W_L	Încheietura mâinii drepte
Wrist_right	W_R	Încheietura mâinii stângi
Hand_left	H_L	Mâna dreaptă
Hand_right	H_R	Mâna stângă
Spine	S	Coloana vertebrală
Hip_center	H_C	Șold
Hip_left	H_L	Șold stâng
Hip_right	H_R	Șold drept
Knee_left	K_L	Genunchi picior stâng
Knee_right	K_R	Genunchi picior drept
Ankle_left	A_L	Glezna picior stâng
Ankle_right	A_R	Glezna picior drept

Astfel, unghiul de indicare estimat se obține din coordonatele mijlocului umărului, ale umărului și ale încheieturii mâinii operatorului, care formează doi vectori și sunt proiectați în plan orizontal. Cu alte cuvinte, vectorul de la umăr la încheietura mâinii indică direcția în care arată operatorul. Articulația de la mijlocul umerilor este aproximativ perpendiculară pe senzorul Kinect și astfel unghiul format de linia umerilor cu linia brațului indică direcția pe care trebuie să se deplaseze robotul pentru a atinge ținta urmărită.

Există două cazuri pentru fiecare braț: atunci când unghiul este mai mare de  $90^\circ$ , obiectul este situat în aceeași parte cu brațul de indicare în raport cu corpul uman și unghiul este  $\theta$  (Fig. 4.10). Dacă unghiul este mai mic de  $90^\circ$ , obiectul este situat în partea opusă a corpului în raport cu

brațul îndreptat și unghiul este  $\beta$ , ca și în Fig. 4.11. Robotul are o poziție fixă, în locul în care se află și sursa de alimentare. El are inițial aceeași orientare ca și operatorul uman. Astfel, acesta va începe să se deplaseze pe direcția indicată de brațul uman, dar plecând din poziția sa. Știind orientarea sa inițială față de operator, acesta își poate rectifica orientarea sa prin rotirea capului și căutarea obiectelor după culoare.

În fig. 4.11 este prezentată o schemă experimentală care ilustrează ceea ce a fost descris mai sus. Robotul se deplasează pe direcția inițială (dată de unghiul  $\beta$ ) și apoi, când ajunge într-un punct intermediar (marcat cu X), își rotește capul încercând să identifice obiectul dorit. El cunoaște orientarea sa inițială față de utilizator și își rotește capul pe direcția corespunzătoare.



**Fig. 4.11.** Schema experimentală

Sistemul prevede, de asemenea, un mecanism de detectare a obstacolelor. În fig. 4.12 este prezentată schematic diagrama de flux a procesului Point-and-command, care ilustrează explicațiile de mai sus.

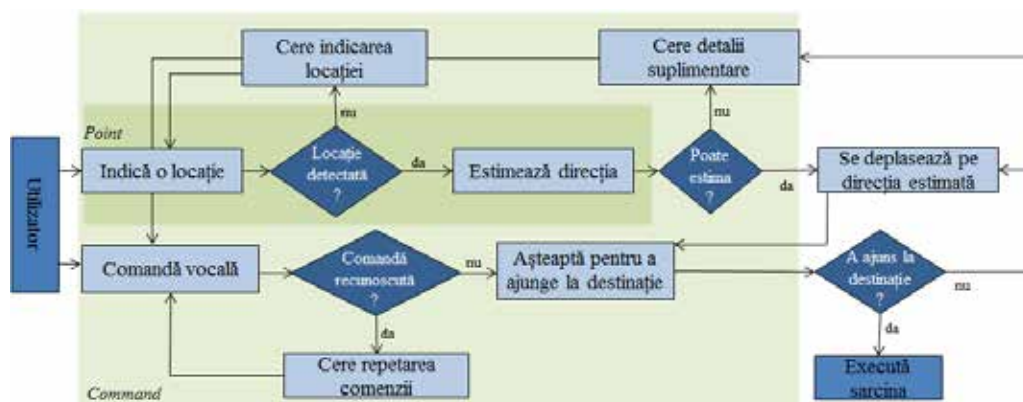


Fig. 4.12. Diagrama flux a paradigmei “Point-and-command”

#### 4.5. SARCINILE ROBOTULUI

Programarea robotului constă în planificarea traseului în funcție de o anumită sarcină. Astfel, o sarcină este definită în lucrarea de față ca deplasare într-un anumit loc, plus o manipulare simplă (două sub-programe). Fiecare activitate are astfel nevoie de 2 intrări esențiale: locație (unghiul de deplasare) și manipulare (numele sarcinii). Aceste două informații sunt obținute de către robot prin dialog: robotul întreabă prin sinteză vocală, iar operatorul uman răspunde prin una dintre metaforele menționate. Au fost utilizate sarcini simple, care sunt de obicei întâlnite în mediile domestice: push, fetch. De asemenea, s-au utilizat numai capacitățile de bază ale robotului, fără a fi sporite. Sistemul folosește un calculator extern pentru a efectua toate calculele referitoare la interacțiunea prin gest/voce, procesare video și așa mai departe.

Pentru sarcina de apucare a unui obiect a fost folosit un algoritm inspirat din [Jingwei, 2012], prin care este măsurată distanța până la obiect cu ajutorul camerelor video și a senzorilor ultrasonici.

**Tabelul 4.4.** Obiecte și atributele lor

Numele obiectului	Numele atributului	Valori
		Circle ( <i>cerc</i> )
Ball	Shape	Triangle ( <i>triunghi</i> )
	( <i>formă</i> )	Square ( <i>pătrat</i> )
Cube		Red ( <i>roșu</i> )
Cone	Color	Yellow ( <i>galben</i> )
	( <i>culoare</i> )	Blue ( <i>albastru</i> )
		White ( <i>alb</i> )
	Size	Small ( <i>mică</i> )
	( <i>dimensiune</i> )	Big ( <i>mare</i> )

Pentru simplitate, au fost folosite obiecte cu forme cunoscute: mingi, cuburi și conuri (v. Fig. 3.17). Fiecare obiect are anumite proprietăți particulare sau atribute care sunt prezentate în tabelul 4.4.

Atributul *formă* se referă la proprietatea volumetrică a obiectului (forma 2D). Software-ul asociază numele obiectului cu o reprezentare simplificată a sa, care corespunde proprietăților de formă, culoare și dimensiune.

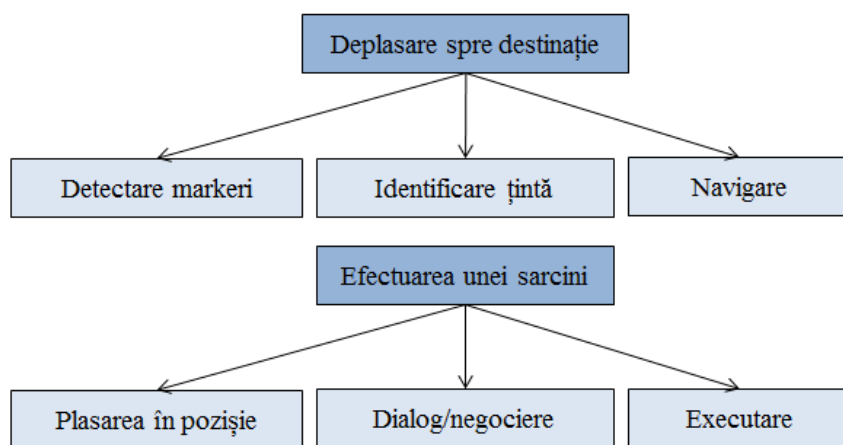
O imagine capturată de camera robotului este mai întâi segmentată cu ajutorul unui algoritm de detectare a culorii folosind OpenCV. În această operație, robotul încearcă să separe obiectul din scenă de fundal. Forma obiectelor este detectată folosind algoritmul de detectare a marginilor (edge detection) [Canny, 1986].

În această lucrare a fost folosit conceptul de model (pattern). Având în vedere ideea dezvoltată în [Kahn, 2008] și extinsă în [Sauppé, 2014b], au fost propuse câteva modele generale care sunt întâlnite frecvent în activitățile de asistență. Alte surse de inspirație sunt modelele de proiectare (design patterns), care valorifică aspectul de reutilizare, prezentate în [Taylor, 2004]. Au fost create câteva modele care au fost observate apărând în activitatea de interacțiune.

În fig. 4.13 sunt prezentate două dintre acestea, împreună cu componentele lor. În imagine se poate observa că șablonul *Move to Target* conține următoarele „sarcini secundare”: detectarea marker, identificarea țintă, navigare. Astfel, atunci când robotul primește comanda de a se



deplasa spre o țintă, trebuie să detecteze mai întâi un marker plasat în mediu pentru a-și determina poziția sa; apoi, trebuie să identifice ținta în funcție de direcția indicată de operator prin intermediul gestului de indicare; și, în cele din urmă, el trebuie să se deplaseze spre acel obiectiv. Identificarea țintei și navigarea se pot suprapune uneori. Când obiectivul este prea departe de robot, acesta trebuie să se deplaseze într-o direcție și apoi încearcă din nou procedura de identificare a țintei. În acest fel, sub-procedurile sunt apelate atunci când este dată o comandă. Unele sub-sarcini au propriile lor subrutine. De exemplu, o sub-sarcină de execuție cuprinde informațiile citite de la senzori și de control al servomotoarelor potrivit datelor obținute de la senzori.

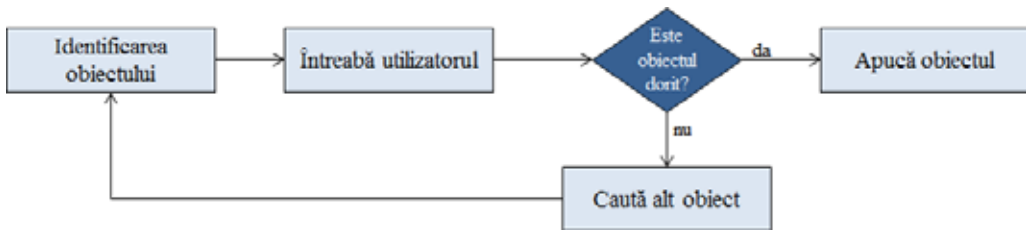


**Fig. 4.13.** Structura a două modele: *Move to Target* și *Perform a Task*

Sarcina ‘fetch-and-carry’ constă în următoarele sub-activități: detectarea obiectului de interes, prinderea obiectului și ridicarea lui.

Toate datele sunt prelucrate pe calculatorul gazdă. Fluxul video capturat de la camera robotului în timp real este procesat pe calculator și robotul primește doar comenzi de control pentru motoarele sale.

Robotul întrebă întotdeauna utilizatorul înainte de efectuarea unei acțiuni de manipulare. Schema bloc a acestui proces este prezentată în Fig. 4.14.



**Fig. 4.14.** Diagrama de flux pentru operațiunea de identificarea unui obiect țintă

Pentru a face interacțiunea realistă, au fost dezvoltate câteva comportamente ale robotului. Unele dintre ele, cum ar fi ridicarea din umeri (shrugging), confuzia (robotul se scarpină pe cap) sunt prezentate în Fig. 4.15.



**Fig. 4.15.** Comportamente ale robotului ca 'shrugging' și 'confused'

## 4.6. TELEOPERARE BAZATĂ PE GESTURI

Pe lângă interacțiunea prezentată mai sus, programul poate trece într-un mod de teleoperare bazată pe gesturi. În acest mod, robotul va imita gesturile umane și poate învăța unele mișcări noi [Boboc, 2013].

### 4.6.1. ASPECTE GENERALE

În ultimul deceniu au fost construiți o mulțime de roboți pentru a asista sau/și pentru a înlocui activitățile umane. Roboții autonomi sunt din ce în ce mai importanți în viața noastră de zi cu zi. Datorită evoluției noilor tehnologii, un număr tot mai mare de activități umane sunt efectuate de către roboți. La început, ei au fost folosiți doar în medii

industriale, însă de-a lungul timpului au fost integrați în medii dinamice, unde este de așteptat să interacționeze în mod natural cu oamenii. Datorită principalelor lor caracteristici, roboții pot fi integrați cu ușurință în diferite medii și permit oamenilor să interacționeze cu ei. Cu toate acestea, ei trebuie să aibă un grad ridicat de flexibilitate, pentru a se adapta la medii noi și necunoscute, învățând continuu sarcini personalizate [Asfour, 2006]. Din fericire, aceste tipuri de roboți au posibilitatea unică de a învăța demonstrații de la experți.

Astfel, Behnke a denumit *învățare prin imitare sau programare prin demonstrație* (PbD – programming by demonstration) [Behnke, 2008] această capacitate a roboților umanoizi. Punctul forte al conceptului PbD este reprezentat de rapiditatea de învățare într-un mod ușor și accesibil pentru utilizatori [Billard, 2008]. Conceptul PbD ar putea fi aplicat într-un sistem robotic, în conformitate cu următoarelor trei faze principale: (i) demonstrația operatorului uman trebuie să ofere toate exemplele unei acțiuni pentru o situație specifică; (ii) comportamentul uman este învățat de către robot; (iii) robotul umanoid generează un comportament similar cu cel al omului, în scopul de a finaliza secvențele de sarcini ale planului. Un studiu cuprinzător, bazat pe conceptul PbD este oferit în [Argall, 2009]. Autorii analizează conceptul PbD pentru mai multe sisteme robotice. Cu toate acestea, acest concept este dificil de aplicat în sisteme robotice bazate pe roboți umanoizi. Acest dezavantaj este reprezentat de părțile robotului, care implică multe grade de libertate pentru a imita mișcările umane.

Roboții umanoizi se găsesc adesea în diferite aplicații cu sisteme robotice care ajută omul. Pentru a atinge obiectivul utilizatorului, un robot trebuie să înțeleagă reacția umană și să anticipeze viitoarele acțiuni în contextul secvențelor de sarcini. Cu alte cuvinte, un sistem robotic necesită un cadru care să înțeleagă obiectivele și intențiile umane. Prin urmare, problema recunoașterii intenției este o problemă a cooperării om-robot într-un sistem robotic. Această problemă este încă important de a fi rezolvată pentru sistemele de interacțiune om-robot (HRI) bazate pe roboți umanoizi.

Mai mulți senzori inteligenți bazați pe tehnologii omniprezente sunt încorporați în robotul NAO. Astfel, acesta ar putea fi integrat cu ușurință într-un sistem robotic care poate fi capabil să asiste un om. Toate beneficiile robotului NAO au atras interesul cercetătorilor de a aplica diverse concepte și metode de recunoaștere a comportamentului uman. Modelele Markov ascunse (HMM) reprezintă o metodă populară care este aplicată pentru a

recunoaște gesturile și vorbirea umană. O altă metodă este GMM (Gaussian Mixture Model) și a fost utilizată pe scară largă în procesul de învățare a roboților prin imitare cu o reprezentare probabilistică a sarcinii.

Aceste metode sunt aplicate datelor care sunt înregistrate de către senzori în timp ce utilizatorul efectuează anumite sarcini. Chiar dacă aceste metode sunt bune pentru a fi aplicate în sistemele robotice, ele cer putere mare de calcul din partea robotului. Cu toate acestea, acest inconvenient nu reprezintă chiar o problemă pentru un robot umanoid care comunică cu stația de lucru prin conexiune wireless, cum este robotul NAO. Unii autori [Neto, 2009], [Stanton, 2012] au întrebuițat senzori suplimentari alături de roboții umanoizi. Sensorii sunt atașați corpului uman, în scopul de a recunoaște cu mare precizie gesturile. Datorită senzorilor atașați de corpul operatorului, mișcările lui pot fi obstrucționate. Aceste limitări pot fi îndepărtate cu ajutorul senzorului Kinect [Toma, 2012a], cum s-a încercat într-o abordare recentă pentru a urmări mișcările 3D umane ale unui conducător auto [Toma, 2012b], dar și pentru aplicații cu roboți umanoizi [Stanton, 2012]. Autorii au propus o metodă simplă bazată pe cinematica inversă pentru a transfera mișcarea umană la un robot umanoid.

În lucrarea de față s-a încercat realizarea unui sistem prin care un robot să îndeplinească o anumită sarcină luând în considerare problemele prezentate mai sus. Prin urmare, sistemul robotic propus trebuie să învețe continuu noi abilități și să se adapteze la abilitățile curente în noi contexte. Mai mult decât atât, s-a demonstrat prin experimente posibilitatea de a învăța noi sarcini folosind datele de mișcare capturate de la un senzor vizual.

#### **4.6.2. IMITAREA MIȘCĂRILOR UMANE**

O prezentare generală a arhitecturii sistemului robotic de teleoperare propus este ilustrată în Fig. 4.16. Principalele elemente sunt: senzorul de urmărire a mișcării Kinect de la Microsoft pentru urmărirea mișcărilor utilizatorilor, robotul umanoid produs de Aldebaran Robotics și un calculator pe care rulează aplicația personalizată. Funcțiile de bază ale aplicației sunt de a procesa datele de la senzorul Kinect și de a trimite comenzi robotului NAO.

Prima etapă de învățare prin imitare într-un sistem robotic este reprezentată de percepția comportamentului uman. Comportamentul uman este dat de secvența de mișcări urmărite cu ajutorul dispozitivului Kinect.

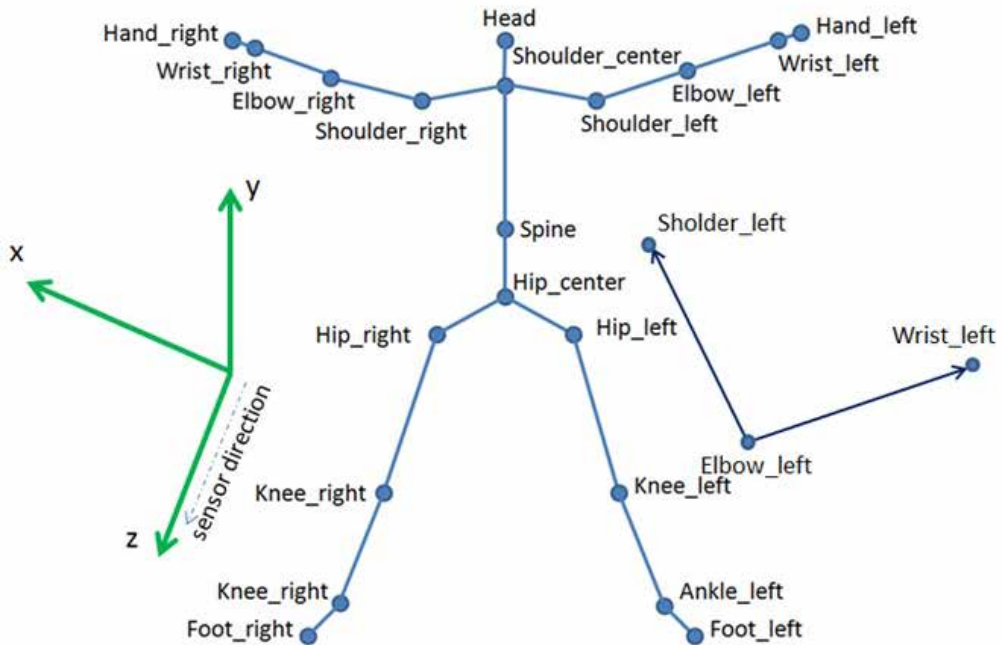


**Fig. 4.16.** Schema generală a sistemului de teleoperare bazată pe gesturi

S-a încercat așadar controlul mișcării de mers pe jos a robotului și manipularea brațelor sale pe baza mișcărilor pe care le efectuează operatorul. Prin urmare, teleoperarea prin gesturi propusă folosește toate gesturile corpului. Mișcările brațelor operatorului sunt folosite pentru a controla brațele robotului, iar poziția operatorului (a corpului său) definește acțiunea pe care acesta dorește să o efectueze, cum ar fi mersul înainte sau întoarcerea la stânga/dreapta.

Procesul de teleoperare are două etape: etapa de calibrare și faza de control. În primul rând, operatorul trebuie să realizeze calibrarea poziției sale în fața câmpului vizual al dispozitivului Kinect. Acesta este un pas necesar deoarece se verifică dacă dispozitivul Kinect a detectat persoana. Procesul de calibrare se realizează cu ajutorul Kinet SDK. Când procesul de calibrare este de realizat cu succes, sistemul este gata să primească comenzile prin gesturi.

Robotul trebuie să reproducă mișcarea cât mai fidel posibil. Prima dată toate pozițiile articulațiilor modelului-schelet sunt înregistrate prin intermediul senzorului Kinect. Acest dispozitiv produce o imagine detaliată a scenei, în care omul este detectat, iar poziția celor 20 de articulații este calculată. Pozițiile modelului-schelet sunt exprimate în coordonate  $x$ ,  $y$  și  $z$ . Aceste coordonate sunt exprimate în metri, cu axa  $z$  pe direcția senzorului, axa  $y$  pozitivă în sus, iar axa  $x$  pozitivă spre dreapta [Internet 4.1] (vezi Fig. 4.17, cu linii verzi).



**Fig. 4.17.** Pozițiile articulațiilor relativ la corpul uman

Comenzile pe care robotul le înțelege sunt: ridicare, ghemuire, oprire, deplasare înainte/ înapoi, deplasare stânga/dreapta și întoarcere la stânga/dreapta. Aceste gesturi trebuie să fie realizate dintr-o anumită poziție, deoarece sistemul este configurat pentru funcții specifice de utilizare și distanțe specifice. Procesul de teleoperare începe cu așezarea operatorului în poziția inițială, poziția ghemuit, pentru că aceasta este poziția de repaus a robotului NAO atunci când nu este utilizat. În cazul în care operatorul se ridică, robotul se ridică de asemenea și este gata să primească comenzile de mișcare.

Scopul acestui proces este de a oferi operatorului posibilitatea de a mișca brațele robotului de la distanță. Acest lucru poate fi foarte util dacă robotul este implicat într-o activitate în care trebuie să manipuleze obiecte.

Pentru a putea realiza imitarea, trebuie făcută o corespondență între mișcările umane și cele ale robotului. Brațele umane au 7 grade de libertate: trei la umăr și încheietura mâinii și unul la cot. În schimb, brațele robotului NAO au doar 5 grade de libertate, două la umăr (pitch și yaw) și cot (yaw și roll) și unul la încheietura mâinii (yaw) (așa cum s-a putut observa în Figura 3.7). Astfel, configurațiile de mișcare ale brațelor umane și ale

robotului diferă. Din acest motiv, mișcarea brațelor robotului se limitează la posibilitățile articulațiilor de care acesta dispune.

Imitarea începe cu percepția operatorului; acesta trebuie să efectueze poziția de calibrarea și apoi sistemul este gata pentru a efectua procesul de imitare. Pentru acest lucru, mai întâi sunt achiziționate mișcările operatorului, apoi datele sunt transformate în mod corespunzător în spațiul de coordonate al robotului, iar în cele din urmă, gestul este executat de NAO. Datele sunt achiziționate folosind trackerul pus la dispoziție de Microsoft și furnizează coordonatele carteziene ale articulațiilor corpului uman. Coordonatele carteziene sunt furnizate considerând ca referință dispozitivul Kinect, dar robotul are sistemul său propriu de referință, cu o origine diferită. Din acest motiv, informațiile articulațiilor Kinect trebuie să fie translatate.

Pentru exemplificare, se consideră brațul stâng al robotului. Acesta are 5 articulații, însă LElbowYaw și LwristYaw (vezi Fig. 3.7) nu au fost luate în considerare deoarece Kinect nu poate detecta mișcarea de rotație (yaw) a brațelor operatorului.

Pentru a calcula un unghi, vor fi creați mai întâi 2 vectori, definiți de legăturile dintre trei puncte (Fig. 4.17, cu linii de culoare albastru închis).

Astfel, pentru a calcula unghiul LShoulderRoll a fost folosit produsul scalar dintre vectorii  $\overline{S_c S_l}$  și  $\overline{S_l E_l}$ , unde  $\overline{S_c S_l}$  este vectorul dintre articulațiile Shoulder\_center și Shoulder\_left, iar  $\overline{S_l E_l}$  este vectorul dintre articulațiile Shoulder\_left și Elbow\_left.

Deoarece vectorii  $\overline{S_c S_l}$  și  $\overline{S_l E_l}$  sunt cunoscuți, unghiul dintre acești vectori se poate obține folosind produsul scalar. Înainte de a aplica ecuația, vectorii trebuie să fie normalizați. Unghiul ShoulderRoll poate fi calculată după cum urmează:

$$\text{LShoulderRoll} = \arccos(\overline{S_c S_l} \cdot \overline{S_l E_l}) \quad (4.15)$$

Unghiul LShoulderRoll se calculează în spațiul de coordonate al Kinect, de aceea, trebuie să fie transformat în spațiul de coordonate al NAO, prin rotirea acestuia cu  $-\pi/2$  radiani.

Unghiul LElbowRoll este calculat în același mod ca și unghiul LShoulderRoll. În acest caz, vectorii care trebuie să fie luați în considerare pentru a calcula unghiul sunt  $\overline{S_l E_l}$  și  $\overline{H_l E_l}$ , unde  $\overline{S_l E_l}$  este vectorul dintre

articulația umărului stâng și cot, iar  $\overline{H_1E_1}$  este vectorul dintre articulația mâinii stângi și cot. Din nou, unghiul ElbowRoll trebuie transformat în spațiul lui NAO, în acest caz, rotindu-l cu  $-\pi$  radiani.

Unghiul de mișcare LShoulderPitch, care se calculează luând în considerare numai vectorul  $\overline{S_1E_1}$  în raport cu axa z. Când brațul se extinde făcând un unghi de  $90^\circ$  cu trunchiul, se consideră pe axa  $z = 0$ . Prin urmare, dacă operatorul ridică brațul din această poziție unghiul va fi pozitiv, iar dacă îl coboară, va fi negativ.

După normalizarea vectorului LShoulderElbow, unghiul LShoulderPitch poate fi definit ca:

$$\text{LShoulderPitch} = \arcsin(\overline{S_1E_1} \cdot z) \quad (4.16)$$

,unde  $\overline{S_1E_1} \cdot z$  este coordonata z a vectorului  $\overline{S_1E_1}$ . Pentru a transforma unghiul LShoulderPitch la spațiul lui NAO trebuie să fie rotit  $2\pi - \alpha$  radiani.

#### 4.7. CONCLUZII

În cadrul acestui capitol au fost prezentate principalele componente ale sistemului de interacțiune propus: sistemul de interacțiune prin gesturi, sistemul de interacțiune prin comenzi vocale, un sistem care realizează sincronizarea comenzilor oferite prin cele două moduri.

Tehnica de aliniere temporală dinamică utilizată pentru recunoașterea gesturilor a fost tratată în detaliu în cadrul acestui capitol, prezentându-se diferitele sale implementări care au fost propuse de către diferiți autori de-a lungul timpului și propunând o nouă variantă care utilizează combinația a două dintre tehnicile prezentate. De asemenea, au fost definite comenzile utilizate pentru interacțiunea cu robotul mobil prin cele două biblioteci de gesturi și respectiv comenzi vocale. Un sistem de integrare permite trimiterea de comenzi combinate, prin cele două moduri de interacțiune.

Apoi a fost prezentat conceptul *Point-and-command* și sarcinile pe care le poate efectua robotul în cadrul sistemului. În finalul capitolului a fost prezentat sistemul robotic proiectat cu scopul de a înzestra robotul cu capacitatea de a imita mișcările corpului uman. A fost ilustrată metodologia de achiziție a mișcării pe care o efectuează utilizatorul și felul cum aceasta este reprodusă de către robot.



# 5

## *SISTEMUL DE LUARE A DECIZIILOR AL APLICAȚIEI DE INTERACȚIUNE NATURALĂ OM-ROBOT*

Un alt aspect care trebuie luat în considerare în HRI este procesul de luare a deciziilor. Atunci când un robot și un om împart același spațiu, aceleași obiecte sau au un obiectiv comun, robotul trebuie să aibă o percepție asupra mediului și să poată cere ajutorul uman atunci când este necesar. Pe de altă parte, și omul să aibă o idee despre stările interne ale robotului. În afară de aceste lucruri, omul și robotul trebuie să fie în măsură să își coordoneze acțiunile.

În această lucrare este de interes atât procesul de luare a deciziilor de către operatorul uman, cât și deciziile pe care trebuie să le ia robotul. Pentru luarea deciziilor de către robot a fost folosită de unii cercetători logica fuzzy (FL), atunci când informațiile sunt vagi, inconsistente sau incomplete (decizii în condiții de incertitudine), de exemplu, pentru detectarea și urmărirea persoanelor în medii aglomerate și nestructurate [Granata, 2013]. Alte modele stocastice, cum ar fi procesele de decizie Markov (MDP) [McGhan, 2012], procese de decizie Markov parțial observabile [Broz, 2013], rețele bayesiene (BN) [Patel, 2010] sau modele Markov ascunse (HMM) [Rossiter, 2011] sunt folosite în proiectarea sistemelor de asistare a utilizatorului pentru recunoașterea intenției sau alte aplicații în care este nevoie de luarea unei decizii. O abordare interesantă pentru luarea deciziilor de colaborare om-robot este prezentată în [Kaupp, 2010], făcând uz de reprezentări probabilistice.

## 5.1. SISTEME DE SUPORT AL DECIZIEI

Sistemele de suport pentru decizii (Decision Support Systems - DSS) sunt o clasă de sisteme informatice, incluzând sistemele bazate pe cunoștințe, care sprijină activitățile de luare a deciziilor. Scopul principal al unui DSS este de a ajuta factorii de decizie pentru a lua cea mai bună decizie atunci când au de a face cu situații și informații complexe [Cebi, 2010].

Termenul de *sistem de suport decizional* se referă și la interacțiunea dintre om și mașină în luarea deciziilor [Nof, 1999], acest sens fiind cel utilizat în această lucrare, iar un astfel de sistem nu este destinat să înlocuiască factorul de decizie uman, ci mai degrabă să îmbunătățească luarea deciziilor de către om.

Astăzi sistemele de suport decizional reprezintă o zonă largă de cercetare și practică. Potrivit DSSResources.com, un sistem de suport decizional este un sistem sau subsistem informatic interactiv bazat pe calculator destinat să ajute factorii de decizie să utilizeze tehnologii, date, documente, cunoștințe și/sau modele pentru a identifica și rezolva probleme, sarcini complete ale procesului de decizie și să ia decizii [Power, 2008]. Sistemul de suport al deciziei este un termen general pentru orice aplicație pe calculator care îmbunătățește abilitatea unei persoane sau a unui grup de a lua decizii.

Așa cum s-a precizat, sistemele de suport pentru decizii au fost asociate cu luarea de decizii strategice, manageriale, pe termen lung, însă disponibilitatea largă de resurse de calcul și nivelul crescut de sofisticare a DSS încurajează progrese suplimentare în cercetarea și dezvoltarea de noi instrumente care sporesc eficiența utilizatorilor în luarea deciziilor complexe, precum și de a avea acces la informații mai bune despre alternativele de decizie disponibile.

Sistemele de suport decizional sunt așadar o clasă de sisteme informatice care sprijină activitățile de luare a deciziilor. Conform clasificării lui Power, se pot identifica cinci tipuri de sisteme de suport decizional [Power, 2002]: (1) DSS orientat pe comunicații, (2) DSS orientat pe date, (3) DSS orientat pe documente, (4) DSS orientat pe cunoștințe (5) DSS orientat pe model.

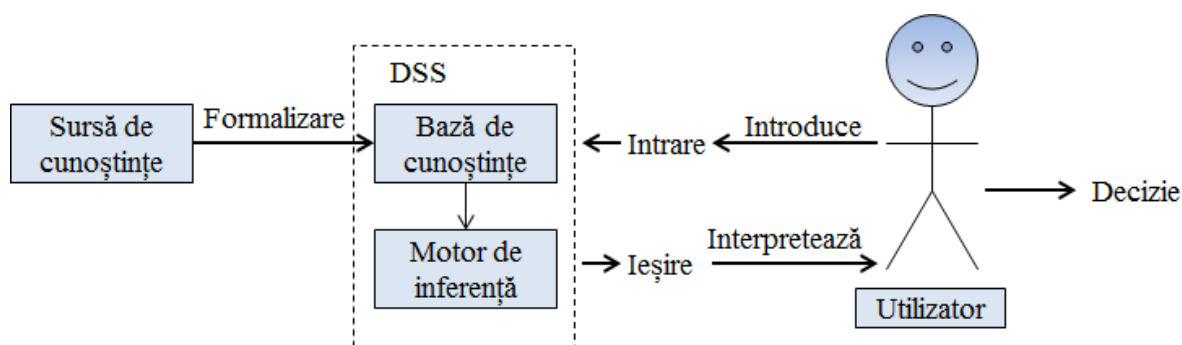
O altă clasificare a sistemelor de suport pentru decizii este oferită în lucrarea [Arnott, 2005], în care acest tip de sisteme sunt împărțite în:

- Sisteme de suport pentru decizii personale;
- Sisteme de suport pentru grupuri;

- Sisteme de suport pentru negociere;
- Sisteme inteligente de asistare a deciziilor;
- DSS bazate pe managementul cunoștințelor;
- Depozite de date;
- Sisteme de analiză și raportare pentru întreprinderi.

După cum se poate observa, DSSs sunt o clasă largă de sisteme informatice și pot lua diferite forme, putând fi utilizate în diferite moduri. Ele diferă în ceea ce privește utilizatorii vizați și modul în care sunt implementate. Unele se concentrează pe informații, unele pe modele, iar unele pe facilitarea comunicării și colaborare. Sistemele diferă, de asemenea, în funcție de domeniul de aplicare. Unele sunt destinate unui singur utilizator, iar altele sunt destinate mai multor utilizatori într-o organizație.

Arhitectura unui DSS, așa cum apare în lucrarea [Lamy, 2010], este ilustrată în Fig. 5.1. Acest tip de sisteme sunt formate de obicei dintr-o sursă de cunoștințe nestructurată (de exemplu, un ghid practic, un set de cazuri sau un grup de experți în domeniu), iar această sursă de cunoștințe este apoi structurată într-o bază de cunoștințe (de exemplu, un set de reguli sau o bază de date) și în cele din urmă, un motor de inferență se aplică bazei de cunoștințe, care determină ieșirea.



**Fig. 5.1.** Schema generală a unui DSS

Sistemele de suport decizional câștigă o popularitate crescută în diverse domenii, inclusiv afaceri, inginerie, armată și medicină. Acestea sunt deosebit de valoroase în situațiile în care cantitatea de informații disponibile este prea mică pentru intuiția unui factor decizional și în care precizia și eficiența sunt importante. Sistemele de suport decizional pot sprijini capacitățile cognitive umane prin integrarea de diverse surse de informare, oferind acces inteligent la cunoștințe relevante, ajutând și în

procesul de structurare a deciziilor. Ele pot folosi, de asemenea, metode de inteligență artificială pentru a aborda probleme care sunt greu de rezolvat prin tehnici tradiționale.

Există studii care fac uz de sisteme de suport decizional pentru aplicații robotizate. În lucrarea [Mouaddib, 2008], operarea unei flote de roboți este controlată cu ajutorul unor factori de decizie locali bazați pe procese de decizie Markov (MDP) încorporați în roboți, în combinație cu un DSS folosit de către operator pentru a decide dacă este nevoie să tele-opereze roboții atunci când aceștia se află într-o stare degradată. Un DSS a fost, de asemenea, implementat în lucrarea [Ding, 2009] pentru a controla un grup de vehicule aeriene (UAV) de către un singur pilot, sub forma unui mecanism de selecție a liderului. Unele lucrări se ocupă cu utilizarea de DSS pentru selecția roboților [Wang, 1991], [Tansel Îç, 2013] sau pentru selectarea de soluții pentru sisteme robotice autonome [Heikkilä, 2013]. În [Pernalet, 2007] a fost prezentat un sistem de suport decizional bazat pe logica fuzzy pentru îmbunătățirea coordonării ochi-mână la copii prin intermediul unei interfețe robotice haptice. Cu toate acestea, sistemele de suport al deciziei au fost foarte puțin sau deloc utilizate în aplicații de HRI.

În această lucrare a fost implementat un DSS pentru a selecta dintre doi roboți care să fie activ la un moment dat în procesul de asistare a unui operator uman. Un DSS inteligent utilizează tehnici de inteligență artificială pentru suport decizional [Arnott, 2008] și există mai multe modalități de punere în aplicare a unui astfel de sistem, cum ar fi utilizarea unui DSS împreună cu un sistem expert (ES) [Cheung, 2005]. DSS-ul propus în această lucrare se bazează pe logica fuzzy (FL) pentru a informa utilizatorul atunci când un robot devine inutilizabil.

## 5.2. LOGICA FUZZY

Cunoștințele pe care raționamentul uman se bazează sunt aproape întotdeauna afectate de o serie de incertitudini și inexactități. Cu toate acestea, în ciuda incertitudinii care caracterizează punctul nostru de vedere al lumii, calitatea de raționament a omului este remarcabilă. Fie în situații simple sau complexe, deciziile sunt în general foarte bune în comparație cu

caracterul vag și incertitudinea datelor problemei. Interpretarea sa impresionantă, având în vedere complexitatea și diversitatea situațiilor întâlnite, a interesat pe unii cercetători în teoria sistemelor, printre care și Lotfi Zadeh, fondatorul logicii fuzzy în jurul anilor 1960.

Logica fuzzy este o extensie a logicii booleene, pe baza teoriei matematice a mulțimilor fuzzy, care este o generalizare a teoriei clasice a mulțimilor [Zadeh, 1965]. Prin definiție, o mulțime fuzzy este o clasă de obiecte caracterizată printr-o *funcție de apartenență*, care atribuie fiecărui obiect un grad de apartenență cuprinsă între 0 și 1. Prin introducerea noțiunii de grad în verificarea unei condiții, permițând astfel unei condiții să fie într-o altă stare decât adevărat sau fals, logica fuzzy oferă un sprijin foarte valoros pentru raționament, ceea ce face posibil să se ia în considerare inexactități și incertitudini.

Principiile fuzzy au fost aplicate pentru o gamă foarte mare și diversă de probleme, iar cea mai mare parte a lucrărilor publicate a fost preocupată de controlul fuzzy, utilizând în general cele două metode mai cunoscute de implementare a inferenței fuzzy: metoda Mamdani și metoda Takagi-Sugeno, care au fost introduse pentru a rezolva aplicații de control [Ross, 2010].

Un avantaj al logicii fuzzy, în scopul de a formaliza raționamentul uman (vezi fig. 5.1), este că normele sunt stabilite în limbaj natural. Altfel spus, funcția de apartenență cuplează logica fuzzy de variabile lingvistice. Acest lucru face mai ușor de implementat un controler cu logică fuzzy, care permite luarea de decizii inteligente în timp real.

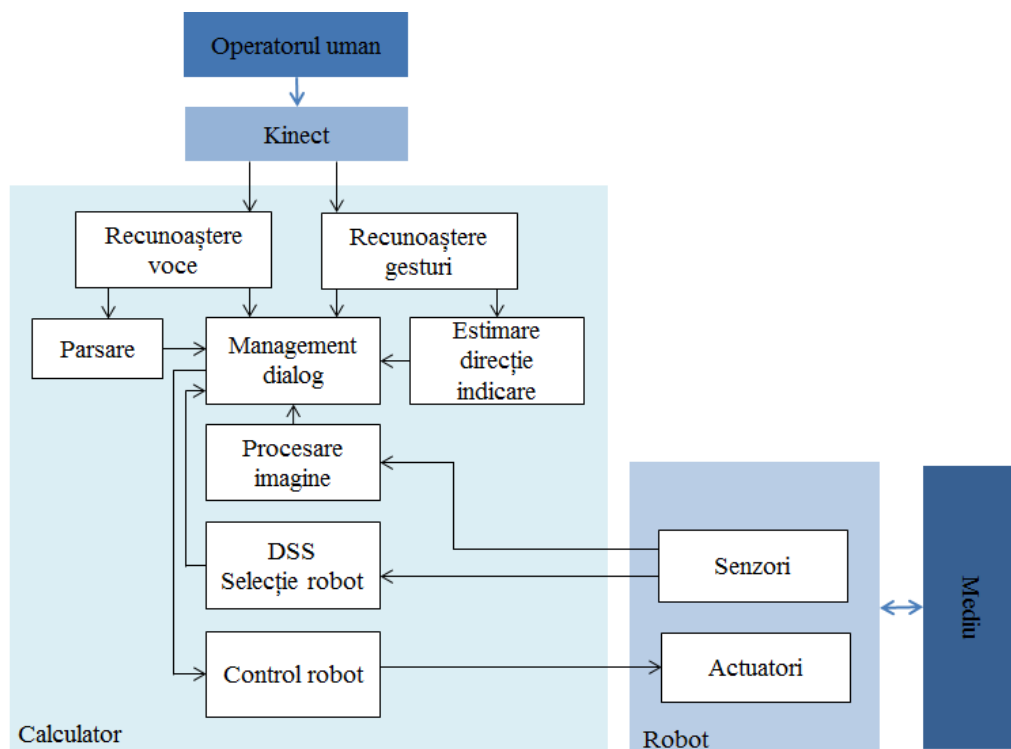
Prin urmare, logica fuzzy este potrivită pentru a fi utilizată la baza sistemelor expert, sisteme care au ca scop reproducerea mecanismelor cognitive ale unui expert într-un anumit domeniu. Logica fuzzy poate fi de asemenea utilizată pentru sisteme de suport decizional, în timpul fazei de analiză a datelor, de exemplu. Obiectivul principal al unor astfel de sisteme este, așa cum s-a specificat mai sus, să ajute la luarea deciziilor, iar logica fuzzy poate fi utilă în acest scop, fie pentru a descoperi reguli sau inferențe fuzzy pentru a înțelege mai bine datele, fie pentru a executa interogări fuzzy pe cunoștințe de specialitate.

### 5.3. ARHITECTURA SISTEMULUI DE INTERACȚIUNE OM-ROBOT

În această secțiune va fi prezentată arhitectura sistemului de interacțiune, în scopul de a stabili contextul pentru ilustrarea sistemului de suport pentru decizii propus.

Sistemul de interacțiune este proiectat pentru interacțiunea cu roboți mobili de asistență personali, care operează la domiciliu sau în birouri pentru efectuarea de sarcini plictisitoare, repetitive. În fig. 5.2 este reprezentată schema generală a sistemului, cu toate blocurile care sunt implicate în interacțiune. Operator uman și mediul sunt „punctele finale” ale procesului de interacțiune, în sensul că un operator interacționează cu mediul prin intermediul unui „asistent” al său, care este robotul mobil. El trimite comenzi gestuale și vocale robotului, care sunt capturate folosind dispozitivul Kinect, așa cum s-a prezentat în capitolele anterioare. Recunoașterea vocii și a gesturilor sunt două procese separate, dar ele sunt combinate în blocul de management al dialogului, cu scopul de a oferi comenzi mai complexe. Blocul DSS de selecție a roboților este responsabil pentru alegerea dintre doi roboți pe cel „activ”, care este pregătit pentru interacțiune. O unitate de control trimite comenzi robotului, care sunt generate de asemenea în blocul de management al dialogului.

În prima fază, sistemul recunoaște gesturile de indicare și indiciile verbale ale omului. Acestea sunt combinate și apoi sunt transformate în comenzi care sunt transmise robotului printr-o conexiune fără fir. Direcția de indicare este recunoscută folosind datele capturate cu dispozitivul Kinect și direcția estimată este apoi utilizată pentru a informa robotul cu privire la locația de interes. Așa cum s-a precizat mai sus, sistemul este proiectat pentru a ajuta omul în activitățile de zi cu zi și, prin urmare, spațiul de operare poate fi privit ca un mediu inteligent (cu senzori vizuali, markeri).



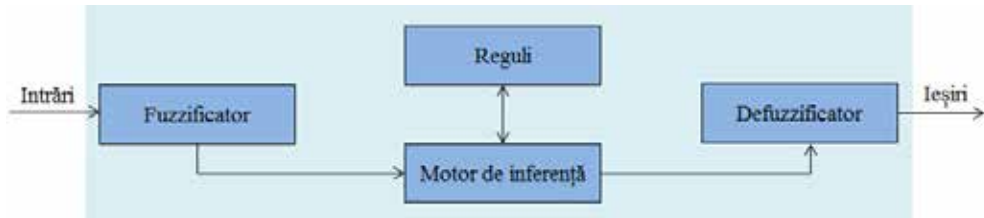
**Fig. 5.2.** Schema generală a sistemului [Boboc, 2015b]

### 5.3.1. BLOCUL DE DECIZIE

Una dintre principalele limitări pentru roboții mobili este reprezentată de autonomia lor energetică. În cazul platformei mobile utilizată în această cercetare, aceasta nu depășește 50-60 de minute de autonomie în modul normal de funcționare. Unii cercetători au propus o stație de ‘andocare’ pentru a evita acest inconvenient [Navarro-Guerrero, 2012], care nu necesită asistență umană. Soluția pare a fi bună, dar în cazul de față s-a propus o abordare dintr-o perspectivă diferită, fiind folosiți doi roboți de asistență identici. Atunci când bateria primului este pe sfârșite, operatorul uman poate decide să solicite un al doilea robot pentru a continua de îndeplinit sarcina începută de primul. În acest scop, a fost dezvoltat un sistem inteligent de suport al deciziei pentru a-l ajuta să decidă când să înceteze interacțiunea cu un robot și să înceapă interacțiunea cu celălalt.

DSS este bazat pe logica fuzzy și este proiectat astfel încât să sprijine operatorul în selectarea robotului cu care robot să interacționeze, luând în considerare mai multe variabile. Astfel, sistemul de suport decizional este

conceput ca un sistem cu logică fuzzy, cu principalele sale componente: fuzzificare, motor de inferență, bază de reguli și de defuzzificare (Fig. 5.3).



**Fig. 5.3.** Schema bloc a unui sistem cu logica fuzzy [Boumella, 2012]

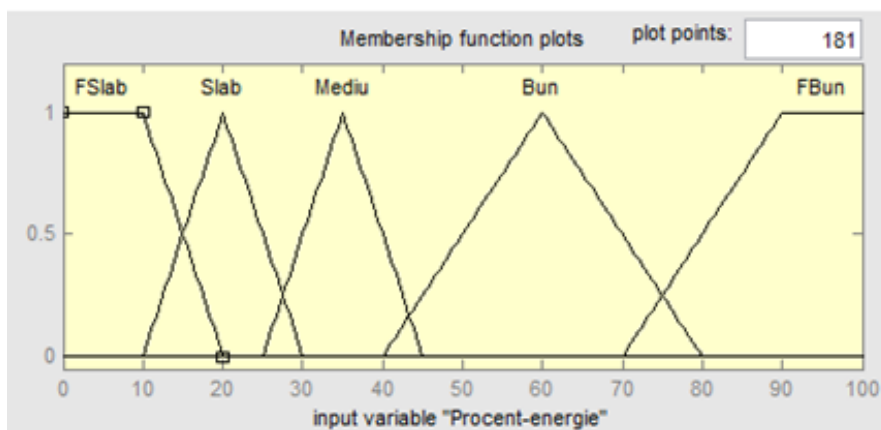
Vor fi descrise mai întâi variabilele de intrare cu funcțiile lor de apartenență, apoi variabila de ieșire, regulile sistemului și procesele de fuzzificare și defuzzificare. Simulările prezentate în acest capitol au fost realizate utilizând mediul de dezvoltare MATLAB, respectiv Fuzzy Logic Toolbox. Acest set de instrumente oferă funcții, aplicații și blocuri pentru analiza, proiectarea și simularea sistemelor bazate pe logica Fuzzy [Internet 5.1].

Există 4 variabile de intrare:

- procent de putere energie rămas;
- distanța dintre robot și sursa de alimentare;
- variabila care arată starea internă a robotului;
- timpul de funcționare.

Procentul de energie rămas este obținut prin intermediul proxy-ului ALMemory (vezi subcapitolul 3.2.2), care oferă informații despre nivelul de încărcare a bateriei. Acest procent este fuzzificat în următoarele mulțimi fuzzy: „foarte slab”, „slab”, „mediu”, „bun”, „foarte bun”. Funcțiile lor de apartenență sunt ilustrate în Fig. 5.4, iar intervalele mulțimilor fuzzy sunt date în Tabelul 5.1. Atunci când valoarea procentului de energie este sub 25%, nivelul bateriei este considerat a fi slab, iar un nivel bun al bateriei este considerat atunci când procentul de energie este peste 45%.



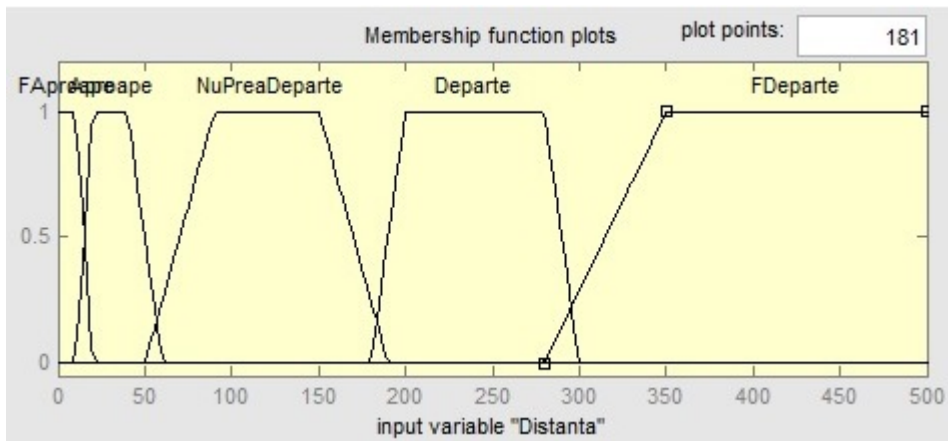


**Fig. 5.4.** Funcțiile de apartenență ale variabilei *Procent\_energie*

**Tabelul 5.1.** Mulțimile fuzzy ale variabilei *Procent\_energie*

Câmp de intrare	Interval [%]	Mulțime fuzzy
Procent_energie	$\leq 10$	foarte slab
	10 - 20	foarte slab sau slab
	20 - 30	slab sau mediu
	30 - 40	mediu
	40 - 45	mediu sau bun
	45 - 70	bun
	70 - 80	bun sau foarte bun
	$\geq 80$	foarte bun

Distanța dintre robot și locul în care se află sursa de alimentare este o variabilă determinată prin căutarea periodică a NAOmark-ului plasat în acea locație și extragerea distanței până la acesta. Seturi fuzzy ale acestei variabile sunt: „foarte aproape”, „aproape”, „nu prea departe”, „departe”, „foarte departe”. Aceste seturi s-au stabilit luând în considerare anumite distanțe tipice, corespunzătoare unei încăperi obișnuite în care se poate desfășura procesul de interacțiune cu robotul mobil. Distanțele sunt exprimate în centimetri. Funcțiile lor de apartenență sunt ilustrate în Fig. 5.5, iar intervalele mulțimilor fuzzy sunt date în Tabelul 5.2.

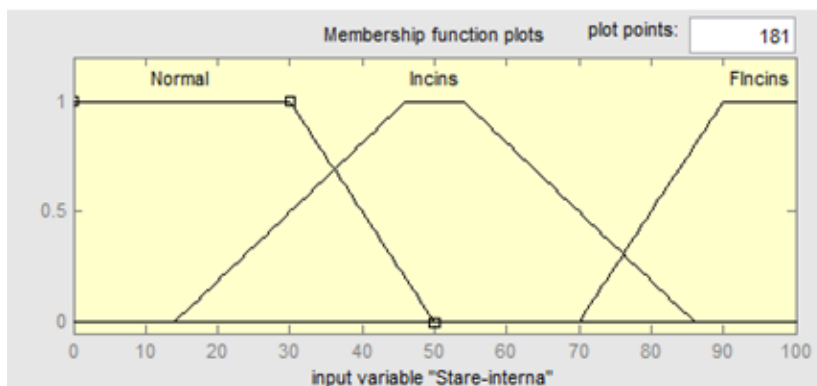


**Fig. 5.5.** Funcțiile de apartenență ale variabilei Distanță

**Tabelul 5.2.** Mulțimile fuzzy ale variabilei Distanță

Câmp de intrare	Interval [cm]	Mulțime fuzzy
Distanța	$\leq 10$	foarte aproape
	10 - 20	foarte aproape sau aproape
	20 - 50	aproape
	50 - 60	foarte aproape sau nu prea departe
	60 - 180	nu prea departe
	180 - 190	nu prea departe sau departe
	190 - 280	departe
	280 - 300	departe sau foarte departe
	$\geq 300$	foarte departe

Starea internă a roboților este dată de senzorii de temperatură și este fuzzificată ca: „normal”, „încins”, „foarte încins”. Funcțiile lor de apartenență sunt ilustrate în Fig. 5.6, iar intervalele mulțimilor fuzzy ale variabilei *Stare internă* sunt date în Tabelul 5.3.

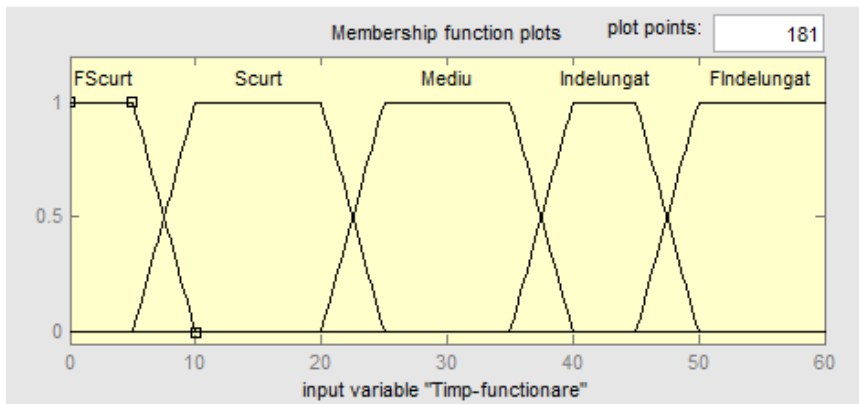


**Fig. 5.6.** Funcțiile de apartenență ale variabilei *Stare\_internă*

**Tabelul 5.3.** Mulțimile fuzzy ale variabilei *Stare\_internă*

Câmp de intrare	Interval [°C]	Mulțime fuzzy
Stare internă	$\leq 15$	normal
	15 - 50	normal sau incins
	50 - 70	încins
	70 - 85	încins sau foarte încins
	85 - 100	foarte încins

Variabila timp de funcționare reține timpul scurs de când NAO a început să funcționeze și include mulțimile: „foarte scurt”, „scurt”, „mediu”, „îndelungat”, „foarte îndelungat”. Valorile acestei variabile au fost stabilite luând în considerare timpul maxim de funcționare a robotului până la descărcarea completă a acumulatorului, care nu depășește 60 de minute, așa cum s-a constatat în timpul testelor efectuate. Funcțiile lor de apartenență sunt ilustrate în Fig. 5.7, iar intervalele mulțimilor fuzzy ale variabilei *Timp de funcționare* sunt date în Tabelul 5.4.

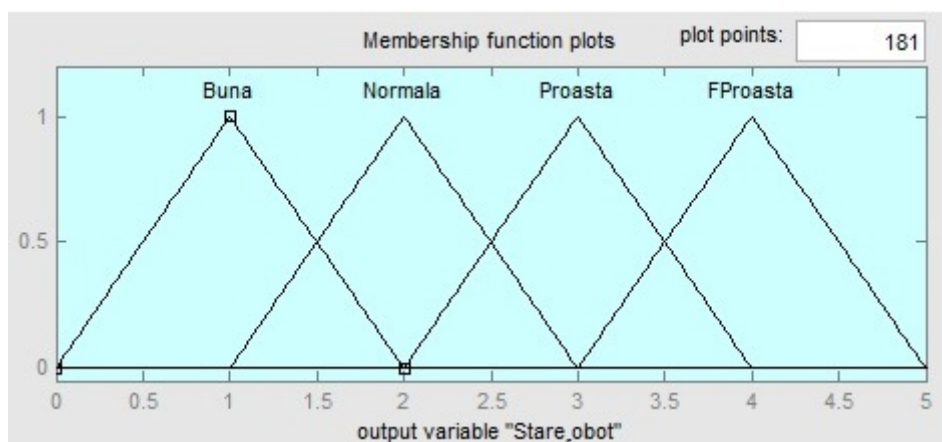


**Fig. 5.7.** Funcțiile de apartenență ale variabilei *Timp\_funcționare*

**Tabelul 5.4.** Mulțimile fuzzy ale variabilei *Timp\_funcționare*

Câmp de intrare	Interval [s]	Mulțime fuzzy
Timp_funcționare	$\leq 5$	foarte scurt
	5 - 10	foarte scurt sau scurt
	10 - 20	scurt
	20 - 25	scurt sau mediu
	25 - 35	mediu
	35 - 40	mediu sau îndelungat
	40 - 50	îndelungat sau foarte îndelungat
	50 - 60	foarte îndelungat

Scopul sistemului este de a identifica starea în care se află robotul la un moment dat. Ieșirea sistemului fuzzy este o valoare cuprinsă între 1 și 4, care reprezintă starea de funcționare a robotului: „bună”, „normală”, „proastă”, „foarte proastă”. În tabelul 5.5. sunt date aceste mulțimi fuzzy și intervalele lor, iar în Fig. 5.8. sunt ilustrate funcțiile lor de apartenență. Sunt considerate două cazuri: în cazul în care ieșirea sistemului este reprezentată de starea bună sau normală, atunci mesajul afișat operatorului este că robotul este operabil; în cazul în care starea sistemului este proastă sau foarte proastă, utilizatorul este atenționat că trebuie să înlocuiască robotul deoarece va deveni în scurt timp inoperabil.



**Fig. 5.8.** Funcțiile de apartenență ale variabilei de ieșire *Stare\_robot*

**Tabelul 5.5.** Mulțimile fuzzy ale variabilei de ieșire *Stare\_robot*

Câmp de intrare	Interval	Mulțime fuzzy
	0 - 2	bună
	1 - 3	normală
	2 - 4	proastă
<i>Stare_robot</i>	3 - 5	foarte proastă

Funcțiile de apartenență ale variabilelor fuzzy utilizate sunt de formă trapezoidală sau triunghiulară. Unele studii au arătat că aceste funcții dau o performanță mai bună față de alte forme, fiind preferate și datorită simplității lor matematice [Jin, 2002], [Monicka, 2011].

Baza de cunoștințe a sistemului conține unele reguli care determină starea folosită pentru a informa utilizatorul atunci când robotul este complet operabil și când nu. Baza de reguli constă din 375 de reguli care determină starea robotului (bună, normală, proastă sau foarte proastă) prin evaluarea variabilelor de intrare menționate mai sus. O parte din baza de reguli este redată mai jos:

1. *If (Procent\_energie is FSlab) and (Distanța is FAproape) and (Stare\_interna is Normal) and (Timp\_functionare is FScurt) then (Stare\_robot is Proasta)*
2. *If (Procent\_energie is FSlab) and (Distanța is FAproape) and (Stare\_interna is Normal) and (Timp\_functionare is Scurt) then (Stare\_robot is Proasta)*

3. If (Procent\_energie is FSlab) and (Distanta is FAproape) and (Stare\_interna is Normal) and (Timp\_functionare is Mediu) then (Stare\_robot is Proasta)
4. If (Procent\_energie is FSlab) and (Distanta is FAproape) and (Stare\_interna is Normal) and (Timp\_functionare is Indelungat) then (Stare\_robot is Proasta)
5. If (Procent\_energie is FSlab) and (Distanta is FAproape) and (Stare\_interna is Normal) and (Timp\_functionare is FIndelungat) then (Stare\_robot is Proasta)
151. If (Procent\_energie is Mediu) and (Distanta is FAproape) and (Stare\_interna is Normal) and (Timp\_functionare is FScurt) then (Stare\_robot is Buna)
154. If (Procent\_energie is Mediu) and (Distanta is FAproape) and (Stare\_interna is Normal) and (Timp\_functionare is Indelungat) then (Stare\_robot is Normala)
165. If (Procent\_energie is Mediu) and (Distanta is FAproape) and (Stare\_interna is FIncins) and (Timp\_functionare is FIndelungat) then (Stare\_robot is FProasta)
172. If (Procent\_energie is Mediu) and (Distanta is Aproape) and (Stare\_interna is Incins) and (Timp\_functionare is Scurt) then (Stare\_robot is Normala)
181. If (Procent\_energie is Mediu) and (Distanta is NuPreaDeparte) and (Stare\_interna is Normal) and (Timp\_functionare is FScurt) then (Stare\_robot is Buna)
183. If (Procent\_energie is Mediu) and (Distanta is NuPreaDeparte) and (Stare\_interna is Normal) and (Timp\_functionare is Mediu) then (Stare\_robot is Normala)
189. If (Procent\_energie is Mediu) and (Distanta is NuPreaDeparte) and (Stare\_interna is Incins) and (Timp\_functionare is Indelungat) then (Stare\_robot is Proasta)
212. If (Procent\_energie is Mediu) and (Distanta is FDeparte) and (Stare\_interna is Normal) and (Timp\_functionare is Scurt) then (Stare\_robot is Normala)
229. If (Procent\_energie is Bun) and (Distanta is FAproape) and (Stare\_interna is Normal) and (Timp\_functionare is Indelungat) then (Stare\_robot is Normala)
242. If (Procent\_energie is Bun) and (Distanta is Aproape) and (Stare\_interna is Normal) and (Timp\_functionare is Scurt) then (Stare\_robot is Buna)
251. If (Procent\_energie is Bun) and (Distanta is Aproape) and (Stare\_interna is FIncins) and (Timp\_functionare is FScurt) then (Stare\_robot is Proasta)
282. If (Procent\_energie is Bun) and (Distanta is Departe) and (Stare\_interna is FIncins) and (Timp\_functionare is Indelungat) then (Stare\_robot is FProasta)
294. If (Procent\_energie is Bun) and (Distanta is FDeparte) and (Stare\_interna is Incins) and (Timp\_functionare is Indelungat) then (Stare\_robot is Proasta)
309. If (Procent\_energie is FBun) and (Distanta is FAproape) and (Stare\_interna is Incins) and (Timp\_functionare is Indelungat) then (Stare\_robot is Normala)

## CAPITOLUL 5

324. *If (Procent\_energie is FBun) and (Distanta is Aproape) and (Stare\_interna is Incins) and (Timp\_functionare is Indelungat) then (Stare\_robot is Proasta)*

334. *If (Procent\_energie is FBun) and (Distanta is NuPreaDeparte) and (Stare\_interna is Normal) and (Timp\_functionare is Indelungat) then (Stare\_robot is Normala)*

346. *If (Procent\_energie is FBun) and (Distanta is Departe) and (Stare\_interna is Normal) and (Timp\_functionare is FScurt) then (Stare\_robot is Buna)*

359. *If (Procent\_energie is FBun) and (Distanta is FDeparte) and (Stare\_interna is Normal) and (Timp\_functionare is Indelungat) then (Stare\_robot is Normala)*

367. *If (Procent\_energie is FBun) and (Distanta is FDeparte) and (Stare\_interna is FIncins) and (Timp\_functionare is Scurt) then (Stare\_robot is Proasta)*

369. *If (Procent\_energie is FBun) and (Distanta is FDeparte) and (Stare\_interna is FIncins) and (Timp\_functionare is FIndelungat) then (Stare\_robot is FProasta)*

În cazul în care starea este „foarte proastă” sau chiar „proastă”, utilizatorul ar trebui să ceară robotului să revină în poziția inițială și apoi el va chema celălalt robot să continue sarcina începută de primul.

Simularea Matlab a rezultatelor experimentale a produs următoarele reguli (fig. 5.9) și suprafețele (fig. 5.10) pentru valorile testate. Suprafețele arată dependența ieșirii sistemului față de intrările determinate de funcționarea sa prin logică fuzzy.

Sistemul utilizat pentru simulare folosește modelul Mamdani pentru mecanismul de inferență, în care metoda AND este minimă.

Așa cum s-a spus mai sus, logica fuzzy utilizează variabile lingvistice în loc de variabile numerice. Procesul de conversie a unei variabile numerice (număr real) într-o variabilă lingvistică (număr fuzzy) este numit fuzzificare. Metoda de fuzzificare este max-min.

Procesul invers se numește defuzzificare. Utilizarea motorului de inferență produce ieșirea necesară într-o formă lingvistică. În conformitate cu cerințele din lumea reală, variabilele lingvistice trebuie să fie transformate în numere reale. Metoda centroid este cea mai bună metodă de defuzzificare bine-cunoscută pentru tipul de controler fuzzy Mamdani.

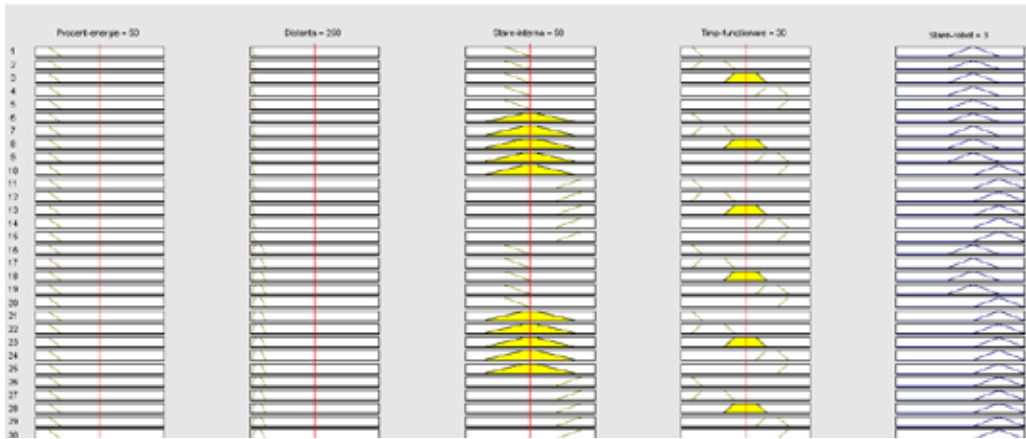


Fig. 5.9. Rezultate pentru anumite valorile testate

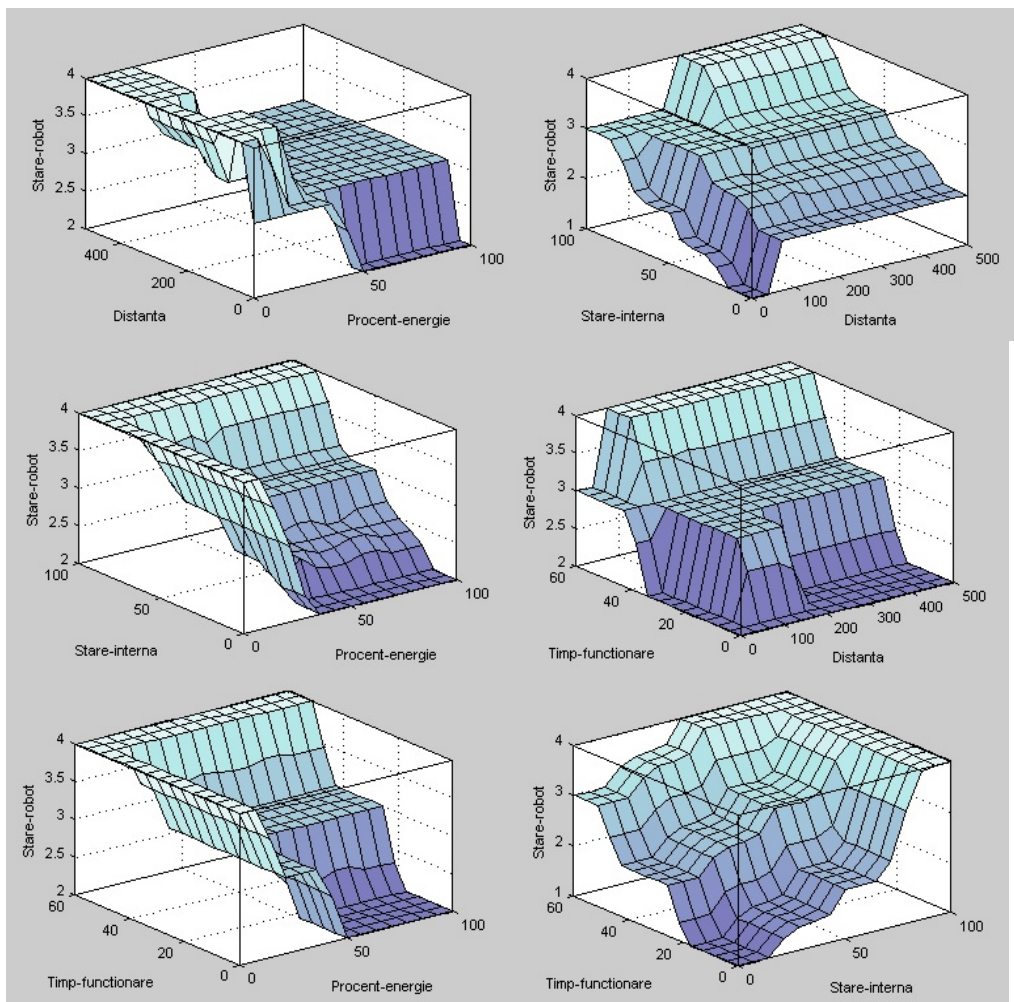


Fig. 5.10. Suprafețele rezultate pentru diferite variabile de intrare



### 5.3.2. Blocul de comunicare

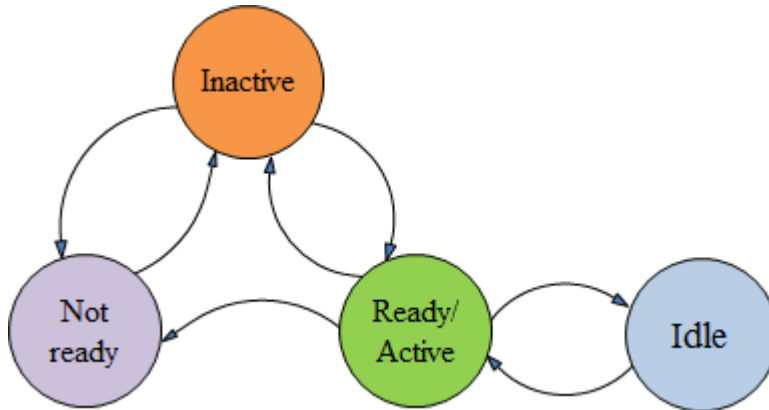
Blocul de comunicare este responsabil cu realizarea conexiunii între calculator și robotul mobil, precum și transmiterea comenzilor și sincronizarea datelor. S-a folosit o abordare client-server. Aplicația de pe calculator produce date folosind senzorul Kinect și apoi deschide un socket și încearcă să se conecteze la aplicația care rulează pe robot. Acesta deschide la rândul lui un socket și așteaptă conexiuni. Odată ce o conexiune este stabilită, deschide un nou socket pentru a servi acestei conexiuni și continuă să asculte. A fost definit un protocol de comunicare, care reprezintă un set de reguli pe baza cărora se desfășoară comunicarea.

Conectarea roboților la calculatorul gazdă se realizează prin protocolul TCP/IP, fiind adresați prin *IP* și *PORT*. O comandă sub formă de mesaj trimisă către server este direcționată către robotul conectat la sistem. Acesta primește mesajul, și execută comanda primită. Interpretarea unui mesaj se realizează pe calculatorul gazdă, care trimite robotului doar instrucțiunile pe care trebuie să le efectueze. De exemplu, datele preluate de senzorii video ai robotului sunt trimise către calculatorul gazdă, unde imaginile sunt prelucrate și apoi sunt date comenzile de acțiune.

### 5.3.3. Sistemului de control

Sistemul de control este numele generic dat procesului aflat în spatele aplicației, care hotărăște modul în care se află întreg sistemul la un moment dat. Sistemul de control este astfel conceput încât prezintă mai multe stări de funcționare. Unul dintre ele este atunci când este *inactiv* sau *închis*. În acest caz utilizatorul nu poate interacționa cu robotul, ceea ce înseamnă că unul dintre dispozitivele implicate în procesul de interacțiune (sau ambele – Kinect, platforma robotică) este inactiv. O altă stare a sistemului este *idle*, în care intră după o perioadă de timp în care nu a fost folosit. În această stare, este oprit fluxul audio și video de la dispozitivul Kinect, precum și comunicarea cu robotul mobil, până când este sesizată dorința de a utiliza din nou aplicația. Starea *activă* este cea în care aplicația este pornită, de asemenea și celelalte dispozitive (robotul și senzorul Kinect) sunt pregătite pentru interacțiune. Mai există cazul în care unul dintre aceste dispozitive nu este pornit sau nu se poate realiza conexiunea cu acesta, iar atunci utilizatorul va fi atenționat prin mesaje text. În Fig. 5.11 sunt prezentate

aceste stări în care se poate afla sistemul. Starea sistemului este stocată într-o variabilă, iar în funcție de aceasta sistemul alocă resursele necesare și informează utilizatorul.



**Fig. 5.11.** Stări ale sistemului

Aplicația păstrează într-o variabilă și starea în care se găsește robotul – pentru a ști ce operațiune execută acesta – detecția obiectelor, navigare, etc. Atunci când se dorește începerea unei interacțiuni sau trimiterea unei comenzi este necesar să se rostască numele robotului sau să se execute un gest predefinit.

#### 5.4. CONCLUZII

De-a lungul timpului, au fost dezvoltate diverse metode de luare a deciziilor, fiind implementate cu ajutorul unor tehnici ce provin din știința informației, psihologia cognitivă sau inteligența artificială, fie ca instrumente de sine stătătoare sau ca medii de calcul integrate. Acestor medii le-a fost dat numele comun de sisteme de suport pentru decizii. Acest concept este extrem de larg și definițiile sale variază în funcție de punctul de vedere al autorului. În lucrarea de față ele au fost definite ca sisteme interactive bazate pe calculator care ajută utilizatorii în activitățile de judecată și de alegere.

Așadar, un sistem de suport decizional este conceput pentru a promova procesul decizional. S-a arătat că logica fuzzy este un instrument util pentru a imita problema de luare a deciziilor întrucât, spre deosebire de logica clasică, este tolerantă la imprecizii, incertitudine și adevăr parțial.

## *CAPITOLUL 5*

În acest capitol a fost prezentată arhitectura generală a sistemului de interacțiune om-robot, evidențiind sistemul de suport decizional implementat, cu rolul și funcțiile lui.

# 6

## *TESTE ȘI EXPERIMENTE CU SISTEMUL DE INTERACȚIUNE NATURALĂ*

În cadrul acestui capitol vor fi prezentate experimentele efectuate utilizând sistemul de interacțiune cu roboți umanoizi. Experimentele au fost efectuate utilizând o echipă formată din doi roboți NAO.

Scopul experimentelor prezentate este reprezentat de efectuarea unei analize pentru evaluarea funcționalității sistemului. Accentul acestei evaluări va fi pus pe părțile de recunoaștere a gesturilor și interacțiune.

### *6.1. INTRODUCERE*

În vederea unei analize cât mai detaliate a modului de interacțiune pentru sistemul propus, au fost concepute câteva scenarii de lucru în care s-au realizat teste care au pus accent pe diferite aspecte ale interacțiunii prin gesturi și voce și a ratei de succes în îndeplinirea sarcinilor de către roboți. Au fost astfel analizate acuratețea sistemelor de recunoaștere a gesturilor și a vocii, precizia și timpul necesar pentru atingerea obiectivelor de către robot, precum și modul în care este realizată cooperarea cu utilizatorul în vederea înlocuirii robotului care nu mai poate fi utilizat, înainte de epuizarea energiei sale. Pentru fiecare experiment, au fost dezvoltate în prealabil replici virtuale în simulatorul V-REP.

Experimentele care s-au efectuat pentru testarea interacțiunii multimodale au respectat următoarele cerințe: utilizatorii au fost nevoiți să învețe expresiile pe care să le folosească pentru dialogul purtat între aceștia și robotul mobil. Cele două scenarii au avut ca scop testarea sistemului de recunoaștere a expresiilor vocale. De aceea, replicile sunt predefinite, dialogul purtat nu este unul spontan. Inteligența necesară pentru un astfel

## CAPITOLUL 6

de dialog este dincolo de ceea ce și-a propus această lucrare. Utilizatorul are impresia că interacțiunea este naturală, întrucât utilizează mijloacele pe care le folosește în viața de zi cu zi în interacțiunea cu semenii, dar expresiile sunt predefinite.

Pentru ca experimentele să se efectueze în bune condiții s-a ținut cont de următoarele lucruri: mediul de lucru a fost eliberat pentru a se evita ciocnirile dintre roboți și alte obiecte, suprafața de deplasare a fost curățată, s-au amplasat markeri pe pereții încăperii în mod corespunzător, dispozitivele au fost amplasate în mod optim, obiectele au fost amplasate pe suporturi pentru a putea fi apucate de către roboți, s-au supravegheat și monitorizat experimentele, s-a încercat optimizarea operațiilor pentru a evita timpii morți.

O problemă care a fost observată în timpul experimentelor și care este menționată și de alți cercetători este următoarea: robotul NAO este incapabil să se deplaseze pe o linie dreaptă, ci are tendința de a devia spre stânga, în condițiile în care este comandat să se deplaseze doar înainte. La început s-a considerat că acest lucru se datorează suprafeței podelei din mediul de laborator, care nu oferă suficientă aderență pentru picioarele din plastic ale robotului. Cu toate acestea, experimentele efectuate pe o altă suprafață au demonstrat că problema a rămas aceeași, chiar dacă abaterea spre stânga a fost mai mică. Această descoperire a incapacității de a merge în linie dreaptă chiar pe o distanță relativ scurtă a impus utilizarea de markeri pentru a ghida robotul.

### **6.2. DEFINIREA MEDIULUI DE LUCRU**

În vederea testării performanțelor rezultatelor sistemului de interacțiune descris în capitolele anterioare, au fost efectuate experimente în mediul real utilizând doi roboți umanoizi NAO. Pentru pregătirea roboților și a mediului de lucru în vederea efectuării experimentelor au fost realizate următoarele: delimitarea și pregătirea zonei de lucru, pregătirea dispozitivelor de interacțiune, punerea în funcțiune a sistemului wireless, încărcarea bateriilor roboților, amplasarea markerilor.

Spațiul de lucru pentru realizarea experimentelor are o lungime de 9 m și lățime de 5 m și se află în laboratorul Centrului de cercetare Informatică Industrială Virtuală și Robotică al Institutului de Cercetare al Universității Transilvania din Brașov. În cadrul acestui spațiu au fost amplasate cutii de

carton (servind ca suport) în poziții fixe, pe care se află diferite obiecte (Fig. 6.1).



*Fig. 6.1. Mediul de lucru pentru experimentele de interacțiune*

### **6.3. DEFINIREA SARCINILOR ȘI OPERAȚIILOR ROBOTULUI**

Pentru testarea sistemului de interacțiune s-au efectuat mai multe experimente, care vor fi prezentate în subcapitolele 6.6 și 6.7.

Sarcinile robotului au fost: navigația spre o anumită locație și realizarea unei manipulări simple. Așadar, operațiile au constat în deplasarea robotului în diferite puncte aflate la diferite distanțe față de poziția utilizatorului. Așa cum a fost prezentat în capitolul 4, punctele de interes sunt indicate de către operatorul uman.

Partea mai dificilă a constat în identificarea obiectelor de referință indicate, precum și așezarea în poziția convenabilă pentru a putea efectua sarcina de manipulare. Pentru acest lucru au fost efectuate foarte multe teste, astfel încât să se determine pozițiile prin care robotul poate apuca un obiect care se află în diferite poziții și la diferite înălțimi.

#### **6.4. EXPERIMENTUL DE IMITARE A MIȘCĂRILOR**

Pentru experimentul efectuat, sursa de inspirație a fost exemplul Avateering C# oferit de Microsoft [Internet 6.1], o aplicație care demonstrează modul în care un model 3D de avatar umanoid este animat folosind senzorul Kinect. În acest exemplu se folosește orientarea membrelor, permițând modelelor de orice dimensiune să fie animate. Pentru fiecare model-schelet dintr-un cadru sunt obținute pozițiile articulațiilor, iar orientările elementelor modelului-schelet sunt calculate și sunt folosite pentru a controla mișcările robotului. Instructorul face o demonstrație în fața senzorului Kinect, iar robotul va imita mișcarea, așa cum a fost descris în capitolul 4.

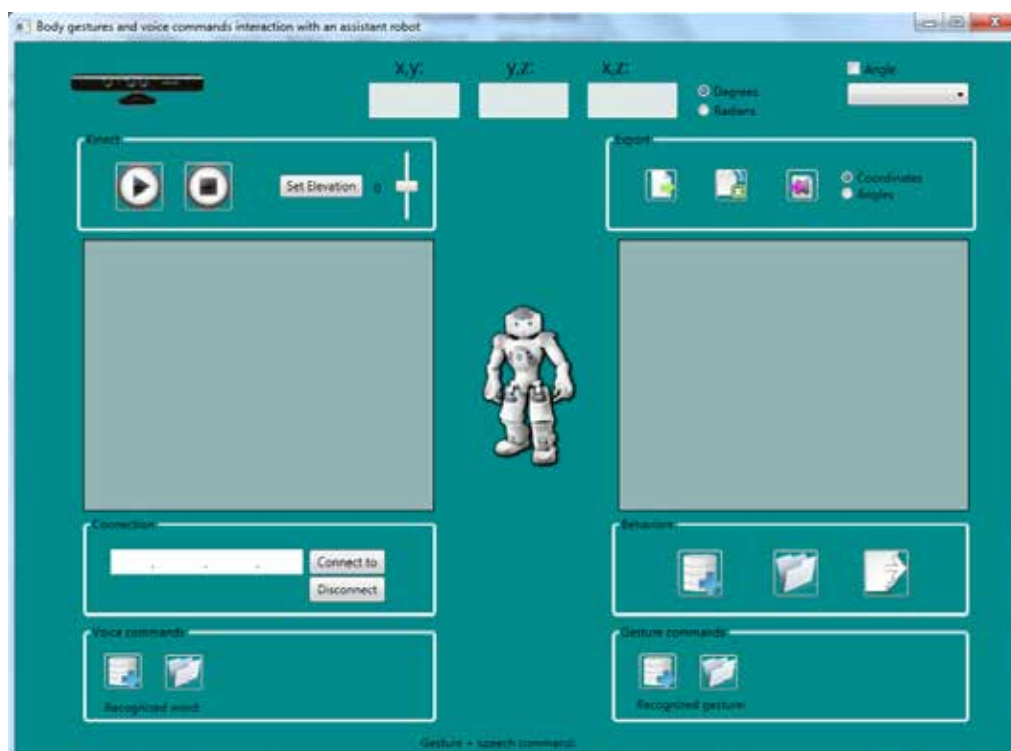
Valori diferite ale unghiurilor sunt obținute pentru diferiți utilizatori care efectuează același gest, dar acest lucru este evitat prin intermediul procesului de normalizare, așa cum este specificat în subcapitolul 4.1.1.3. După aceea, unghiurile sunt convertite (vezi subcap. 4.6.2) pentru a se potrivi cu valorile unghiurilor permise pentru fiecare articulație a robotului, care sunt descrise în documentația oferită de Aldebaran. În tabelul 6.1 sunt ilustrate unghiurile minim și maxim calculate pentru mai mulți utilizatori (fiecare unghi reprezentând o medie a unghiurilor obținute de la 3 utilizatori) și unghiurile din intervalul permis pentru robotul NAO. După conversie, valorile unghiulare sunt transmise robotului.

Pentru programare a fost folosit limbajul C#, fiind suportat atât de către dispozitivul Kinect, cât și de robotul NAO. Modelul-schelet este construit atunci când este urmărit corpul utilizatorului. Pozițiile articulațiilor sunt mai întâi filtrate și coordonatele articulațiilor detectate (vezi Fig. 4.17) sunt utilizate pentru a calcula unghiurile articulațiilor robotului. O corespondență între unghiurile articulațiilor utilizatorului și ale robotului trebuie făcută deoarece robotul simulat utilizat are mai puține grade de libertate decât persoana reală, cum a fost amintit și în capitolul 4, iar astfel unele posturi nu pot fi imitate deoarece gama de mișcare pentru articulații diferă de la om la robot. În Fig. 6.2 este prezentată interfața grafică implementată.

**Tabelul 6.1.** Intervalele valorilor unghiulare obținute și cele permise pentru articulațiile robotului

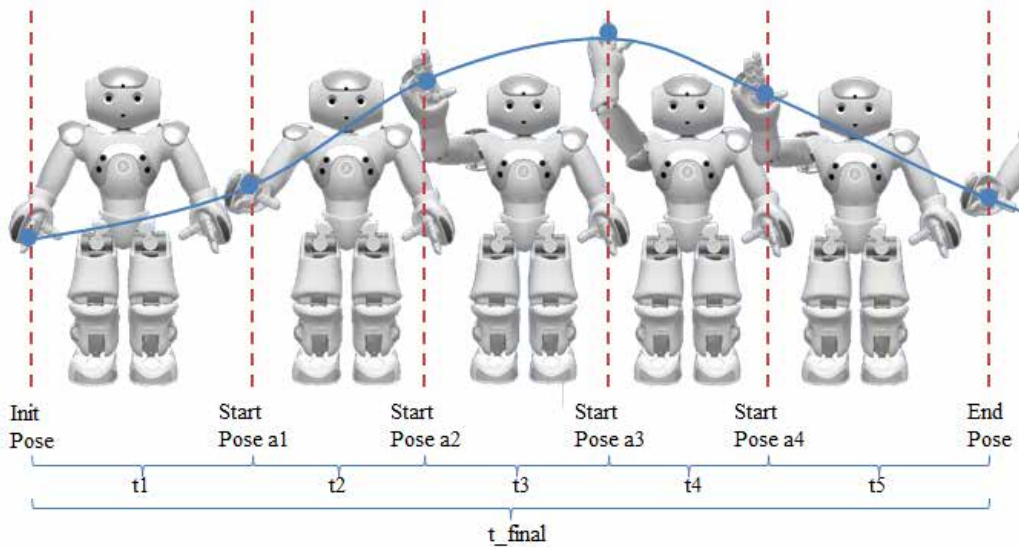
Unghi		Interval (grade)					
		x,y		x,z		y,z	
Head	Kinect	-20	20	-	-	-10	42
	NAO	-	-	-119,5	119,5	-38,5	29,5
Wrist_left	Kinect	-48	-20	-	-	-10	60
	NAO	-	-	-104,5	104,5	-	-
Wrist_right	Kinect	18	50	-	-	-18	71
	NAO	-	-	-104,5	104,5	-	-
Elbow_left	Kinect	-	-	-	-	-176	-35
	NAO	-	-	-119,5	119,5	-88,5	-2
Elbow_right	Kinect	-	-	-	-	35	176
	NAO	-	-	-119,5	119,5	2	88,5
Shoulder_left	Kinect	4	154	-	-	-90	150
	NAO	-18	76	-	-	-119,5	119,5
Shoulder_right	Kinect	-170	-10	-	-	-136	56
	NAO	-76	18	-	-	-119,5	119,5
Hip_left	Kinect	-75	20	-	-	-76	70
	NAO	-21,7	45,29	-	-	-101,6	27,73
Hip_right	Kinect	-60	15	-	-	-80	54
	NAO	-42,3	23,76	-	-	-101,5	27,82
Knee_left	Kinect	-	-	-	-	7	98
	NAO	-	-	-	-	-5,29	121,04
Knee_right	Kinect	-	-	-	-	5	104
	NAO	-	-	-	-	-5,9	121,47
Ankle_left	Kinect	-	-	-	-	-10	60
	NAO	-22,79	44,06	-	-	-68,15	52,86
Ankle_right	Kinect	-	-	-	-	8	56
	NAO	-45,03	22,27	-	-	-67,97	53,40





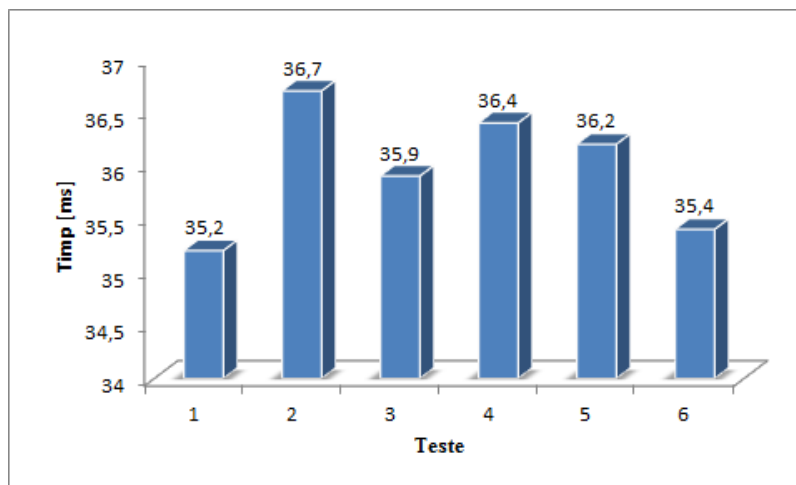
**Fig. 6.2.** Interfața grafică cu utilizatorul (GUI)

O altă problemă este reprezentată de sincronizarea dintre robot și utilizator în timp. Pentru a determina decalajul între mișcările umane și cele ale robotului, a fost creată o aplicație care calculează timpul scurs între diferite poziții care compun traiectoria de mișcare pentru o mișcare simplă (de exemplu, ridicarea mâinii drepte). În Fig. 6.3 este ilustrat modul în care sunt calculate intervalele de timp între anumite momente în timpul efectuării unei mișcări cu brațul. Aplicația creează un număr de stări sau poziții în care se află robotul în timp, pe parcursul mișcării mâinii (în acest caz 6 stări). Intervalele de timp scurse între stări sunt prezentate în tabelul 6.2, fiecare din ele fiind exprimat în ms, unde  $t-t$  reprezintă timpul total,  $mean$  reprezintă media valorilor de timp și  $sd$  este deviația standard. Fig. 6.4 prezintă decalajul de timp între mișcarea utilizatorului și mișcarea reprodusă de către robot, fiind exprimat, de asemenea, în milisecunde, pentru un număr de 6 încercări. Se poate observa că există o medie de aproximativ 36 ms întârziere, ceea ce este imperceptibil pentru ochiul uman.



**Fig. 6.3.** Diferite poziții ale brațului robotului în timpul unei mișcări

Învățarea unei abilități presupune învățarea de mișcări succesive. Dar unele dintre aceste mișcări sunt comune pentru mai multe abilități. Acest fapt a fost exploatat în felul următor: utilizatorul demonstrează o mișcare robotului, care îl imită în timp real și învață această mișcare, stocând-o într-o bază de date.



**Fig. 6.4.** Decalajul între mișcarea utilizatorului și cea a robotului

**Tabelul 6.2.** Duratele de timp între pozițiile brațului robotului pentru mișcarea ilustrată în Fig. 6.3

Teste	t1	t2	t3	t4	t5	t-t
1	418,3	390,0	371,6	365,0	408,3	1953,3
2	413,3	393,3	383,3	368,3	418,3	1976,6
3	420,0	381,6	378,3	358,3	420,0	1958,3
4	415,0	395,0	365,0	358,3	431,6	1965,0
5	410,0	391,6	388,3	346,6	413,3	1950,0
6	412,6	383,3	380,0	371,6	408,3	1965,0
mean	416,6	390,8	379,1	361,6	415,8	1961,6
sd	0,0040	0,0049	0,0076	0,0081	0,0080	0,0087

Utilizatorul îi dă un nume acestei mișcări și atunci robotul va fi capabil să combine diferite mișcări memorate și, astfel, să învețe un nou comportament. În Fig. 6.5 este prezentat modul în care robotul imită o mișcare simplă care descrie modul în care se ajunge în poziția pe care robotul ar trebui să o adopte atunci când încearcă să ridice un obiect. Acest comportament simplu, alcătuit dintr-o secvență de mai multe posturi ale brațelor, este demonstrat robotului și aceasta îl învață. Utilizatorul îi dă un nume și apoi această mișcare poate fi folosită pentru a efectua acțiuni mai complexe, care sunt necesare în procesul de învățare al unei abilități. În acest fel, o ontologie de mișcări de bază poate fi dezvoltată și apoi mai multe mișcări complexe ar putea fi generate într-o abordare bazată pe semantică.

**Fig. 6.5.** Demonstrarea unei poziții robotului

## 6.5. EXPERIMENTUL DE TESTARE A INTERACȚIUNII MULTIMODALE

Pentru a evalua performanța sistemului de interacțiune multimodală, au fost întreprinse două experimente, după scenarii predefinite: deplasarea robotului în mediu pentru a ajunge la ținta indicată de utilizator prin intermediul gesturilor și a manipula un anumit obiect descris prin mijloace verbale și non-verbale. Pentru fiecare tip de experiment au fost efectuate 3 teste, reținându-se eficiența de îndeplinire a sarcinii, precum și acuratețea de recunoaștere a gesturilor.

### 6.5.1. SCENARIUL 1

În acest experiment s-a propus un dialog între utilizator (U) și robotul umanoid NAO (N). Experimentul a fost realizat de către o singură persoană, fiind repetat de trei ori. Utilizatorul îi cere robotului să îi aducă o bilă roșie situată într-o anumită poziție în mediul înconjurător. Robotul va ‘culege’ mai întâi informațiile necesare pentru a putea identifica obiectul și apoi se deplasează în direcția indicată de utilizator. Software-ul aflat pe calculator procesează informația video primită de la robot și în momentul în care este detectat obiectul, se trece în modul de urmărire a obiectului, iar robotul se va deplasa spre acesta. Mai jos este prezentat întreg dialogul dintre utilizator și robot, iar în Fig. 6.6 și Fig. 6.7 sunt ilustrate câteva instantanee sugestive din timpul experimentului.

U: ‘NAO!’

N: ‘Yes, I hear you’

U: ‘Please, give me the red ball from there!’

N: ‘Can you show me how the red color looks like?’

(utilizatorul îi arată robotului o monstră colorată în roșu)

N: ‘What about the shape of the object?’

U: ‘The ball has this shape’ (utilizatorul îi arată un cerc desenat pe o hârtie sau prin gest – desenează un cerc în aer cu ajutorul brațului)

(robotul se deplasează în direcția indicată și atunci când identifică culoarea roșie, se apropie pentru a încerca să identifice forma)

N: ‘Is that the object?’

U: ‘No, I need a bigger one’

## CAPITOLUL 6

*(robotul va continua să caute până când va găsi o minge mai mare)*

U: 'Grab the object!'

*(robotul decide dacă poate apuca obiectul cu o mână sau cu amândouă mâinile)*

U: 'Bring it to me!'

*(NAO caută marker-ul din zona utilizatorului și se deplasează în acea direcție)*

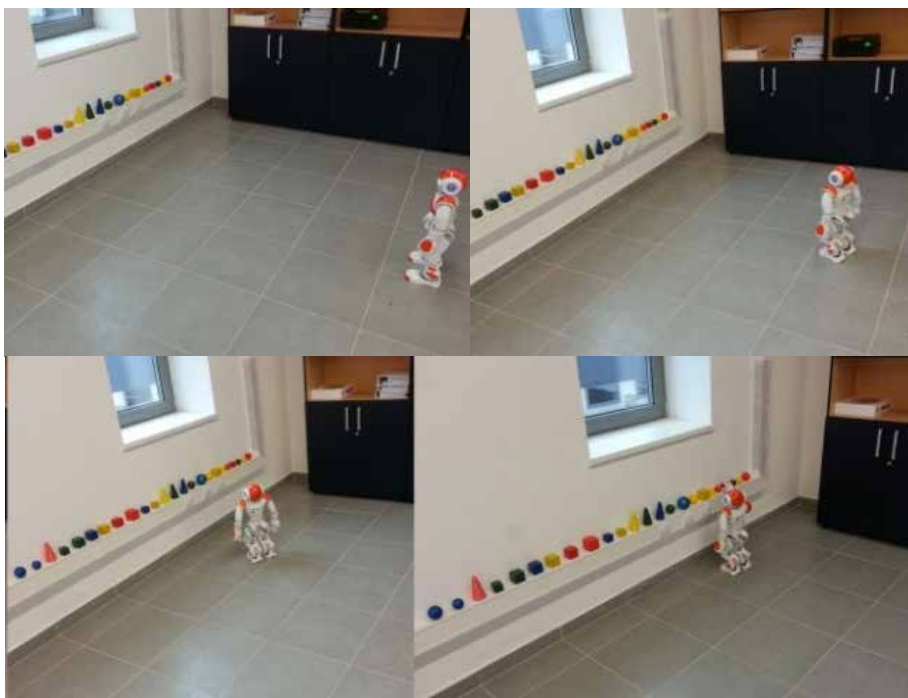
U: 'Leave it'

*(NAO lasă obiectul în mâna utilizatorului)*

U: 'Thank you!'

N: 'Do you want another thing?'

U: 'No' *(NAO se va deplasa în poziția inițială)*



**Fig. 6.6.** Robotul NAO căutând mingea roșie



**Fig. 6.7.** a), b) NAO apucând mingea cu o singură mână; c), d) Câmpul vizual al robotului

În acest scenariu, s-a considerat și etapa de ‘învățare’, prin care culoarea și forma obiectului sunt învățate de către robot. Această etapă are rolul de a familiariza utilizatorul cu modul de interacțiune și de a-i crea sentimentul de naturalețe. Întreg dialogul prezentat mai sus este alcătuit din pași predefiniți, dorindu-se în primul rând să se demonstreze performanța sistemelor de recunoaștere a gesturilor, a vocii și a obiectelor.

### 6.5.2. SCENARIUL 2

În acest scenariu utilizatorul trimite comenzi vocale robotului, cu scopul de a-l ghida să se deplaseze în direcția în care se află obiectele. El cere robotului să îi aducă o minge albastră. Comanda este trimisă atunci când robotul a ajuns la o distanță rezonabilă față de obiecte, astfel încât să le poată identifica cu ajutorul camerei video. În acel moment acesta va căuta obiectul solicitat și atunci când este identificat, se va plasa la o distanță potrivită pentru a fi capabil să apuce obiectul. Comanda „Let me control you!” va duce la posibilitatea de a controla brațele robotului și capul său. Algoritmii de urmărire a capului (vezi 4.1.2) permite utilizatorului să controleze capul lui NAO, cu scopul de a cunoaște câmpul vizual al acestuia (ca flux video afișat în interfața grafică) și apoi cu brațele sale conduce brațele robotului în poziția corectă, pentru a fi capabil să apuce obiectul. Folosind comenzi vocale, utilizatorul cere robotului să deschidă

## CAPITOLUL 6

sau închidă mâinile sale. În acest caz, robotul apucă mingea cu ambele mâini (Fig. 6.9).

O problemă este faptul că robotul nu sesizează dacă obiectul îi scapă din mână în timp ce se deplasează. Acest lucru se datorează faptului că prehensoarele sale nu sunt echipate cu senzori de contact sau alte tipuri de senzori cu ajutorul cărora să distingă dacă mâinile sale sunt sau nu ocupate.



**Fig. 6.8.** Mediul de testare. Utilizatorul indică locația prin gest de indicare și precizează verbal sarcina care trebuie efectuată



**Fig. 6.9.** Robotul NAO apucând o minge cu ambele mâini

Tot în cadrul acestui experiment au fost utilizate mai multe gesturi pentru a testa performanța algoritmului de recunoaștere, în special pentru sarcini de navigare. Matricea de confuzie pentru comenzile efectuate prin intermediul gesturilor pentru un număr de 4 utilizatori care au participat la experiment este prezentată în Tabelul 6.3. Din tabel, se poate determina că acuratețea algoritmului de recunoaștere obținută este de 94%. Cei 4 participanți au repetat fiecare gest de câte 10 ori.

**Tabelul 6.3.** Matricea de confuzie pentru sarcina de navigare

		Gestul recunoscut											
		B	Ca	C	L	R	Sp	Sw	TF	TL	TR	W	Necunoscut
Gestul efectuat	B	90%	0	0	0	0	0	0	0	0	0	0	10%
	Ca	0	100%	0	0	0	0	0	0	0	0	0	0
	C	0	0	90%	0	0	0	0	0	0	0	0	10%
	L	0	0	0	95%	0	0	0	0	0	0	0	5%
	R	0	0	0	0	90%	5%	0	0	0	0	0	0
	Sp	0	0	0	0	0	95%	0	0	0	0	0	0
	Sw	0	0	0	0	0	0	85%	0	0	0	0	15%
	TF	0	0	0	0	0	0	0	90%	0	0	0	10%
	TL	0	0	0	0	0	0	0	0	95%	0	0	5%
	TR	0	0	0	0	0	0	0	0	0	95%	0	5%
	W	0	0	0	0	0	0	0	0	0	0	90%	10%

În Fig. 6.10 este prezentată situația în care robotul se întoarce la utilizator, ținând în mână obiectul cerut de acesta. După ce ajunge la locația în care se află utilizatorul (orientându-se după markerul plasat în acel loc), îi întinde acestuia obiectul și așteaptă comanda care îi cere să elibereze obiectul. În acel moment el desface degetele și apoi așteaptă noi comenzi.



**Fig. 6.10.** NAO aduce obiectul utilizatorului (a) și îl înmânează acestuia (b)



## 6.6. EXPERIMENTUL DE TESTARE A EFICIENȚEI REALIZĂRII SARCINILOR DE CĂTRE ROBOT

În cadrul acestui experiment s-a dorit evaluarea sistemului de suport al deciziei și a performanței în ceea ce privește îndeplinirea cu succes de către robot a sarcinilor cerute de utilizator. În acest scop, a fost realizat un scenariu de test, după cum urmează: trei cutii pe care se află obiecte de diferite culori și forme au fost plasate în mediu în diferite poziții, așa cum este ilustrat în Fig. 6.14. Doi roboți NAO sunt poziționați în partea stângă a utilizatorului, la o anumită distanță de acesta, lângă sursa de alimentare (Fig. 6.11).



**Fig. 6.11.** Doi roboți NAO așteptând să primească comenzi

**Tabelul 6.4.** Pozițiile obiectivelor relativ la poziția robotului și intervalele de timp necesare pentru efectuarea unei sarcini

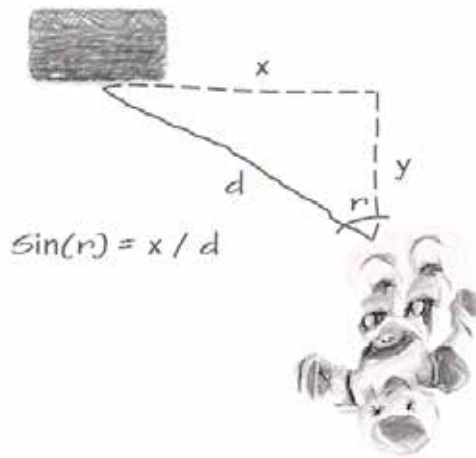
No.	Experiment 1			Experiment 2			Experiment 3		
	d [m]	r [°]	t [m]	d [m]	r [°]	t [min]	d [m]	r [°]	t [m]
1	3,02	26,6	3,5	3,15	5,1	4,04	2,88	51,4	3,23
2	3,81	45,7	4,87	3,51	50,2	4,6	4,15	12,6	5,9
3	3,71	14	5,41	5,49	55	7,12	4,78	41,2	6,41

Testele au fost efectuate de către 3 persoane, câte 3 seturi de experimente pentru fiecare participant. În prima rundă de teste, utilizatorul cere robotului NAO utilizând paradigma Point-and-command prezentată să îi aducă următoarele obiecte: o bilă roșie de pe cutia aflată în partea stângă, o

minge verde de pe cutia plasată în mijloc și un cub galben de pe cutia aflată în partea dreaptă relativ la poziția operatorului uman. Utilizatorii au fost inițial informați de funcționarea sistemului și l-au încercat de câteva ori pentru a se familiariza cu interacțiunea. Distanțele de la robot până la obiecte și media timpului scurs (pentru cei trei utilizatori) până când robotul efectuează cu succes sarcina sunt prezentate în tabelul 6.4, în care:  $d$  reprezintă distanța măsurată din locul în care se află inițial robotul până la suportul pe care se află obiectele;  $r$  reprezintă unghiul dintre direcția înaintea pe care se poate deplasa robotul și direcția în care se află obiectele față de poziția robotului (ilustrate ca două axe notate cu  $x$ , respectiv  $d$  în Fig. 6.12); iar  $t$  reprezintă timpul scurs până la efectuarea cu succes a unei sarcini. Câteva instantanee din timpul unuia dintre experimente sunt date în Fig. 6.15.

Testul a durat aproximativ 2 ore. În această perioadă, primul robot selectat pentru interacțiune a operat 35 de minute, după care sistemul de suport decizional a informat utilizatorul că este timpul să apeleze la cel de-al doilea robot. Al doilea robot a funcționat în continuare 7 minute, până la sfârșitul primei runde de teste. În runda a doua și a treia procentul de operare a fost similar, așa cum se arată în Fig. 6.13.

Acuratețea de recunoaștere a gestului de indicare a fost de 91%, iar pentru recunoașterea comenzilor vocale a fost atins un procent de 95%. Pentru estimarea unghiului de indicare, erorile medii sunt prezentate în tabelul 6.5. Aceste rezultate au fost prezentate în lucrarea [Boboc, 2015b].

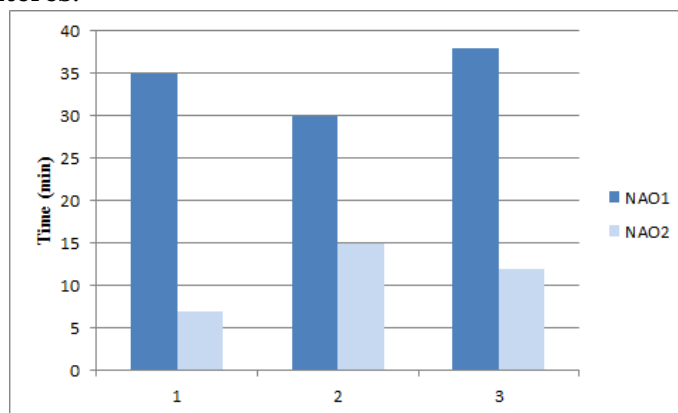


**Fig. 6.12.** Distanța și unghiul până la obiect

**Tabelul 6.5.** Eroarea medie a unghiului de indicare estimat

No.	Eroarea medie [ $^{\circ}$ ]		
	Experiment	Experiment	Experiment
	1	2	3
1	4.56	3.7	5.18
2	3.29	3.12	3.9
3	4.3	4.65	5.32

În ceea ce privește algoritmi de procesare vizuală, precizia de detectare și recunoaștere a obiectelor a fost, de asemenea, acceptabilă. Fiecare experiment presupune recunoașterea a 9 obiecte: 3 (pentru fiecare sarcină) x 3 (pentru fiecare utilizator). Astfel, în total a trebuit să fie recunoscute 27 de obiecte în toate cele trei experimente și dintre acestea numai de două ori robotul a identificat greșit un obiect, dar după ce utilizatorul i-a dat o indicație suplimentară (prin comenzi vocale), robotul a fost capabil să îl identifice pe cel dorit. Robotul a fost programat să întrebe întotdeauna înainte de a efectua o acțiune de apucare a unui obiect dacă acela este obiectul de interes.



**Fig. 6.13.** Timpul de operare pentru fiecare robot de-a lungul experimentelor

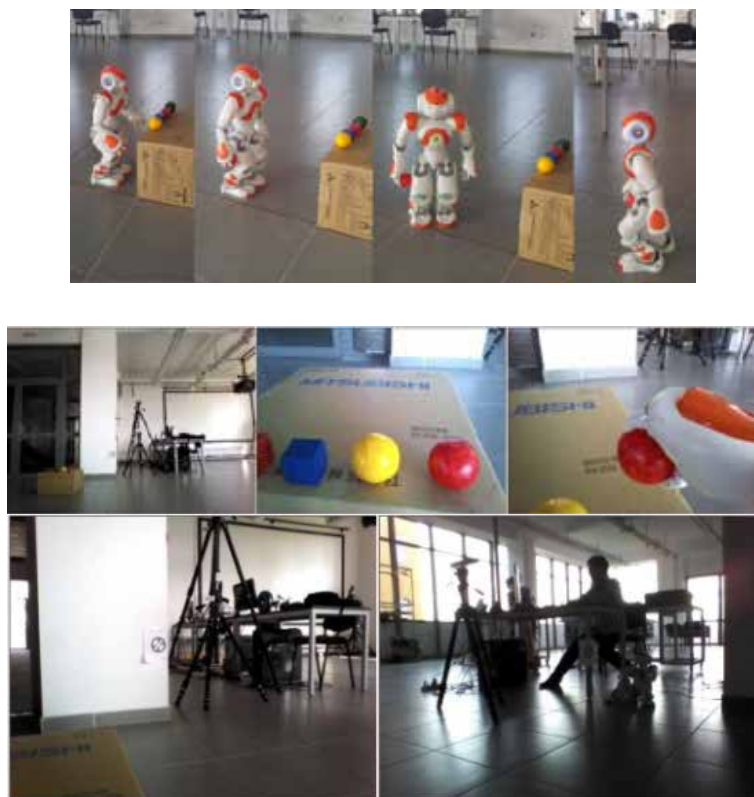
A fost considerat, de asemenea, cazul când două obiecte de aceeași culoare se află unul lângă altul. În cazul în care robotul se confruntă cu o astfel de situație, el trebuie să verifice care este cel dorit de operatorul uman, întrebându-l pe acesta. El ar trebui să răspundă folosind comenzile vocale (de ex: „stânga” sau „dreapta”). În cazul în care obiectele de aceeași culoare sunt foarte aproape unul de altul, robotul nu va fi capabil să le identifice separat.

O problemă tipică care a avut loc de mai multe ori în timpul experimentului a fost următoarea: atunci când robotul se deplasează cu un obiect în mână se poate întâmpla ca obiectul să îi scape ca urmare a contactului cu părți ale corpului său. În acest caz experimentele au continuat ca și cum nimic nu s-ar fi întâmplat, dar va trebui găsită o soluție pentru această problemă în viitor.



**Fig. 6.14.** Operatorul este așezat în fața senzorului Kinect. El indică prin gest (pointing) una dintre cele trei cutii și îi cere robotului să îi aducă un obiect colorat situat pe aceea





**Fig. 6.15.** a) NAO se deplasează pe direcția indicată, identifică mingea roșie, o apucă și se întoarce înapoi la utilizator; b) câmpul de vedere al robotului din poziții diferite

## 6.7. CONCLUZII

Odată cu dezvoltarea de roboți bipezi tot mai inteligenți, roboții personali încep să își facă apariția în case, în birouri, în instituțiile de învățământ sau de cultură. În acest context, cercetătorii ar trebui să ia în considerare dezvoltarea unor modalități noi și intuitive de interacțiune cu roboții, în scopul de a permite proprietarilor lor, care sunt, de obicei, profani, să își exprime intențiile lor cu ușurință și să se simtă confortabil în prezența lor. Comunicarea verbală și non-verbală joacă un rol important pentru acest tip de aplicații.

Scopul acestei lucrări a fost de a dezvolta o interfață pentru asistarea persoanelor în sarcinile lor de zi cu zi, folosind doi roboți umanoizi care navighează în spațiul de lucru al omului, fiind capabili să detecteze și să-i aducă anumite obiecte situate în diferite poziții din mediu. Utilizatorul

uman este informat și poate decide sfârșitul interacțiunii cu un robot și de a-l selecta pe celălalt pentru a continua sarcina începută de primul atunci când energia acestuia este pe sfârșite. Acest lucru este posibil datorită unui sistem de suport decizional inteligent, care colectează câteva informații despre starea internă a robotului și din mediu și informează utilizatorul din timp, în scopul de a permite robotului să se întoarcă în poziția inițială înainte de terminarea energiei.

Interacțiunea om-robot este bazată pe gesturi de indicare combinate cu comenzi vocale. Un concept de interacțiune „Point-and-command” a fost propus, cu scopul de a oferi utilizatorului o interfață ușor de utilizat și intuitivă și pentru a îmbunătăți eficiența colaborării om-robot. Sistemul nu posedă toate cunoștințele necesare pentru a îndeplini o sarcină, dar are capacitatea de a prelua informații de la utilizator.

Funcționalitatea sistemului a fost testată într-un mediu interior și rezultatele au dovedit că toți utilizatorii pot interacționa cu roboții umanoizi fără cunoștințe anterioare despre aceștia. Contribuția principală a lucrării este interacțiunea naturală realizată prin intermediul interfeței propuse.

Sistemul prezentat în această lucrare a avut mai multe limitări care vor fi avute în vedere în cercetările viitoare. Unele dintre ele au fost deja menționate mai sus, cum ar fi problema scăpării obiectelor din mâna robotului în timpul deplasării acestuia.

Unitatea de procesare principală folosită a fost un calculator și această unitate comunică cu robotul prin comenzi wireless. Unele din „cunoștințele” sistemului vor fi transferate în memoria robotului pentru a-l face mai inteligent și pentru a reduce cantitatea informațiilor transferate.

Teste similare cu mai mulți utilizatori vor fi, de asemenea, efectuate pentru a extrage mai multe rezultate și pentru a avea o bază prin care să se poată face comparația cu alte sisteme similare. Cercetări suplimentare includ îmbunătățiri în următoarele zone: comunicarea om-robot, estimarea unghiului gestului de indicare, localizarea și evitarea obstacolelor. De asemenea, capacitatea de învățare a robotului ar trebui să fie investigată. El trebuie să învețe și memoreze obiectele identificate.

## **CONCLUZII FINALE, CONTRIBUȚII ORIGINALE ȘI DIRECȚII VIITOARE**

Întrebarea principală la care prezenta lucrare răspunde este: *Cum se poate realiza o interacțiune naturală cu roboți mobili de asistență utilizați în mediile domestice?* Încă de la început, roboții de asistență au fost concepuți pentru realizarea de sarcini complexe în medii nestructurate. Folosirea roboților mobili în medii domestice, în procesele de navigație și manipulare de obiecte, este utilă întrucât oferă un sprijin persoanelor în vârstă, a celor care au diferite probleme locomotorii, dar și, în general, a celor care doresc să beneficieze de un asistent robotic care să îi ajute în activitățile zilnice. Roboții mobili de asistență sunt integrați în medii domestice pentru activități de manipulare și deplasare de diferite obiecte utilizate în mod curent de către oameni.

### **7.1. CONCLUZII FINALE**

Studiul literaturii de specialitate, prezentat în capitolele anterioare, a arătat faptul că deși s-au cercetat și dezvoltat un număr mare de interfețe de interacțiune naturală, idealul definit de către Valli [Valli, 2008] al conceptului de interfață naturală nu a fost încă atins.

În această lucrare a fost descris un cadru pentru comunicare și interacțiunea om - robot naturală și ușor de utilizat. În timp ce majoritatea sistemelor multimodale HRI propuse în literatura de specialitate se concentrează pe o singură modalitate, sistemul propus permite utilizatorilor să își exprime instrucțiunile lor ca și combinații de intrări prin gesturi și voce. Principalele puncte forte ale sistemului sunt: metoda îmbunătățită de detecție și recunoaștere a gesturilor, interacțiunea ușoară și

naturală prin gesturi și comenzi vocale și răspunsul prin gesturi și sinteză vocală oferit de robot.

Scopul interfeței este de a permite utilizatorilor experimentați sau neexperimentați să coopereze și să interacționeze cu un robot de asistență care operează într-un mediu domestic. O bibliotecă de gesturi și una de expresii vocale au fost implementate, iar comenzile pot fi trimise prin una sau prin ambele modalități de interacțiune. Astfel, un prim obiectiv al cercetării de față a fost de a înzestra robotul cu capacități de interacțiune socială, care sunt esențiale pentru aplicațiile cu roboți de asistență.

Testele experimentale efectuate cu roboți de asistență de tipul robotului umanoid NAO au dovedit o interacțiune om-robot fiabilă. Cea mai mare parte a lucrării s-a axat pe interacțiunea prin gesturi, în special recunoașterea gesturilor. O metodă îmbunătățită a DTW a fost implementată și testată, cu rezultate bune atât în precizie, cât și în eficiență. Așa cum s-a arătat în capitolul 6, acuratețea recunoașterii și eficiența metodei sunt satisfăcătoare și astfel metoda crește reactivitatea robotului la solicitările umane, sporind gradul de naturalitate al interacțiunii.

Al doilea obiectiv operațional a fost de a oferi robotului capacitatea de a învăța noi task-uri. A fost dezvoltată o tehnică simplă care permite robotului să imite gesturile umane. Aceasta, combinată cu abilitatea de exprimare a intențiilor cu ajutorul vocii/gesturilor a dus la o interfață versatilă, care facilitează paradigme de interacțiune robuste, cum este 'Point-and-command'.

Cu toate acestea, există unele probleme sau limitări care au fost întâlnite în timpul experimentelor și care trebuie să fie luate în calcul pe viitor. Precizia de recunoaștere este afectată în mod dramatic atunci când există condiții precare de iluminare sau zgomot în mediul de operare. Pe de altă parte, atunci când mai mulți oameni apar în raza vizuală a senzorului, sistemul are dificultăți în identificarea utilizatorului curent (care este implicat în procesul de interacțiune la un moment dat). În unele situații, robotul s-a aflat în imposibilitatea de a identifica markeri și și-a pierdut orientarea. De asemenea, unele obstacole mai mici nu au fost detectate și uneori robotul a căzut.

În cadrul sistemului a fost dezvoltată și o interfață de teleoperare, care permite robotului să imite mișcările utilizatorului uman. Senzorul Kinect a fost folosit pentru a urmări mișcarea umană, înregistrând poziția și orientarea fiecărei articulații a corpului, iar robotul reproduce traiectoriile



de mișcare ale operatorului și ar putea învăța noi abilități prin combinarea de mișcări succesive simple. Funcția aceasta de teleoperare poate fi extinsă astfel încât mișcările umane să poată fi recunoscute în mod automat de către robot.

Rezultatele preliminare arată că mișcările umane au fost bine imitate de către robot. Mai mult decât atât, a fost evaluată performanța sistemului în două moduri. În primul rând, a fost determinată performanța robotului atunci când acesta a efectuat o sarcină. În al doilea rând, timpul de întârziere între mișcarea umană și cea a robotului a fost calculat pe baza experimentului.

Intenția acestei lucrări nu este să ofere un răspuns imediat și exhaustiv pentru întrebarea care s-a pus la începutul acestui capitol, ci încearcă doar să abordeze acest domeniu al interacțiunii naturale și să ofere câteva contribuții. Totuși, se poate afirma pe baza cercetărilor efectuate în cadrul acestei lucrări că modalitatea de interacțiune contribuie foarte mult la realizarea unei interacțiuni naturale, iar vocea și gestul rămân două dintre cele mai potrivite modalități de interacțiune. De asemenea, se poate spune că nu s-a ajuns încă la un asemenea progres tehnologic încât omul să se poată raporta la un robot mobil ca și în cazul unui semen de-al său.

Pentru viitor, se pot aduce îmbunătățiri performanței sistemului propus prin extinderea aplicării pentru mai multe sarcini și domenii. De asemenea, poate fi introdusă o metodă prin care robotul să fie capabil să prezică intențiile utilizatorului.

### **7.2. CONTRIBUȚII ORIGINALE**

Prezenta lucrare conține un număr de contribuții originale ale autorului la domeniul de cercetare abordat, dintre care sunt enumerate următoarele:

1. Analiza studiilor teoretice și experimentale privind interacțiunea om-robot în aplicațiile de robotică asistivă, cu accent pe modalități de interacțiune naturală cu omul, precum și stadiul actual al sistemelor de interacțiune gestuală și vocală.
2. Algoritm de detectare a mișcărilor corpului uman folosind un dispozitiv Microsoft Kinect.

3. Realizarea unei aplicații în limbajul de programare C# care are ca scop preluarea, analiza și prelucrarea datele provenite de la senzorul Kinect.
4. Implementarea unei metode de învățare a robotului prin imitarea gesturilor utilizatorului.
5. Selectarea algoritmului utilizat pentru recunoașterea gesturilor, aplicarea și utilizarea acestuia în cadrul sistemului de interacțiune multimodală.
6. Implementarea unei variante îmbunătățite a algoritmului de aliniere temporală dinamică.
7. Implementarea unui algoritm de recunoaștere a limbajului vorbit, conform unor reguli gramaticale
8. Conceperea, dezvoltarea și implementarea unei paradigme de interacțiune multimodală cu roboți mobili de asistență, numită „Point-and-command”.
9. Conceperea, dezvoltarea, implementarea și testarea unui sistem de integrare a comenzilor vocale cu cele gestuale.
10. Conceperea și implementarea unei biblioteci de gesturi și a unui vocabular de expresii vocale folosite pentru interacțiunea cu robotul umanoid.
11. Conceperea, dezvoltarea și implementarea unei metode de detecție și identificare de obiecte în funcție de forma și culoarea acestora.
12. Implementarea și testarea unui algoritm de manipulare a obiectelor de către robot.
13. Conceperea, dezvoltarea și implementarea unui sistem de suport al deciziei pentru utilizator prin care acesta să fie informat despre starea roboților de asistență.
14. Conceperea, dezvoltarea și implementarea unei interfețe grafice cu uilizatorul care integrează toate funcțiile sistemului de interacțiune propus.
15. Programe și metodologii de experimentare pentru evaluarea și validarea sistemului de interacțiune și a rezultatelor teoretice obținute.

### 7.3. DIRECȚII VIITOARE DE CERCETARE

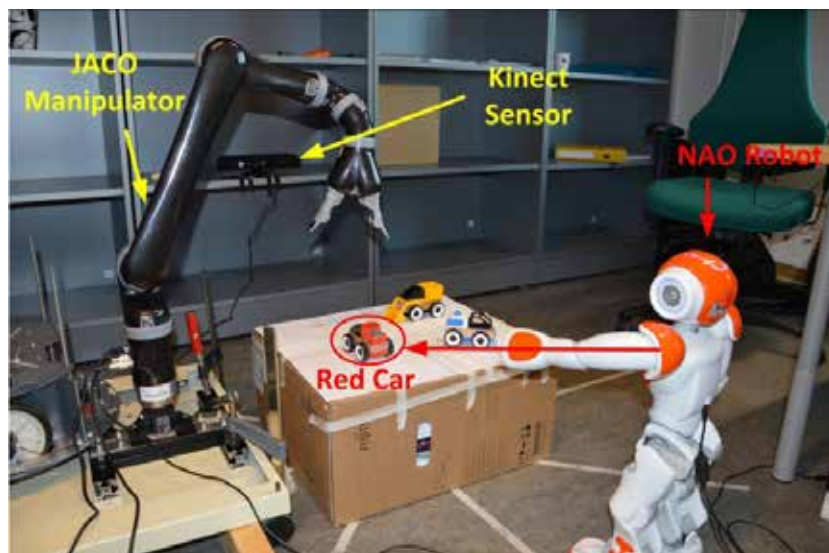
Rezultatele teoretice și practice ale cercetărilor dezvoltate în prezenta lucrare de doctorat deschid noi direcții de cercetare pentru domeniul interacțiunii naturale cu roboți mobili. Prin implementarea interfeței de interacțiune multimodală se realizează un pas important pentru înlocuirea sistemelor care presupun folosirea de diferite dispozitive pe care trebuie să le poarte utilizatorul uman sau care necesită pregătire și instruire preliminară îndelungată datorită complexității comenzilor. Sistemul propus trebuie validat prin aplicații de asistență concrete, cum este cazul asistării persoanelor în vârstă. Pentru acest lucru, trebuie avute în vedere o serie de dezvoltări care să fie abordate în cercetările viitoare, printre care se pot menționa:

- Extinderea bibliotecii de gesturi prin adăugarea de noi gesturi, care să fie mult mai intuitive și care să poată exprima mai bine dorințele utilizatorului;
- Personalizarea comenzilor vocale și pentru limba română;
- Implementarea unei metode mai avansate pentru a detecta obiectele cu forme și culori diferite și pentru manipularea obiectelor;
- Îmbunătățirea capacității robotului de a învăța noi abilități;
- Considerarea posibilității în care mai mulți utilizatori doresc să interacționeze cu robotul în același timp. În acest caz, sistemul trebuie să fie suficient de inteligent pentru a selecta utilizatorul care va interacționa cu robotul;
- Elaborarea de scenarii mai complexe cu o varietate de sarcini care trebuie să fie efectuate de către robot;
- Efectuarea mai multor teste și experimente, în scopul de a detecta unele situații posibile în care sistemul nu funcționează corect.

Sistemul implementat prezintă importanță pentru dezvoltarea de noi aplicații în care este necesară interacțiunea cu roboți sociali de asistență, putând fi adaptat pentru diverse scenarii. Două domenii de interes actual sunt: asistarea persoanelor în vârstă și a celor imobilizate și asistarea personalului în birouri. De asemenea, sistemul poate fi integrat în scenarii mai complexe care presupun și interacțiunea robot-robot, cum este cel din Fig. 7.1, în care un robot umanoid NAO îi indică unui braț robotic un anumit obiect, iar acela îl apucă și îl așează în altă poziție. Acesta este un

caz tipic de colaborare robot-robot, care extinde și completează interacțiunea om-robot.

După cum a mai fost specificat, lucrarea nu își propune să epuizeze problemele ridicate de interacțiunea naturală om-robot, ci numai să deschidă noi căi de reflecție asupra acestei teme.



**Fig. 7.1.** Scenariu de interacțiune robot-robot [Kondaxakis, 2014]

## BIBLIOGRAFIE

1. [Adwan, 2012] Adwan, S. and Arof, H., "On improving Dynamic Time Warping for pattern matching", *Measurement*, vol. 45, pp. 1609-1620, 2012.
2. [Aldebaran, 2012] Aldebaran\_Robotics, "NAO Datasheet", ed, 2012.
3. [Alibali, 2005] Alibali, M., "Gestures in spatial cognition: Expressing, communicating, and thinking about spatial information", *Spatial Cognition and Computation*, vol. 5, pp. 307-331, 2005.
4. [Alvarez-Santos, 2013] Alvarez-Santos, V., Iglesias, R., Pardo, X. M., *et al.*, "Gesture-based interaction with voice feedback for a tour-guide robot", *Journal of Visual Communication and Image Representation*, 2013.
5. [Antonya, 2007] Antonya, C. and Talabă, D., "Design evaluation and modification of mechanical systems in virtual environments," *Virtual Reality*, vol. 11, pp. 275-285, 2007.
6. [Antonya, 2008] Antonya, C., Butnaru, T., and Talabă, D., "Manipulation of Mechanisms in Virtual Environment," in *Product Engineering*, D. Talaba and A. Amditis, Eds., ed: Springer Netherlands, pp. 461-472, 2008.
7. [Antonya, 2011] Antonya, C., Barbuceanu, F. G., and Rusak, Z., "Path generation in virtual reality environment based on gaze analysis," in *AFRICON 2011*, pp. 1-4, 2011.
8. [Arkin, 2003] Arkin, R. C., Fujita, M., Takagi, T., *et al.*, "An ethological and emotional basis for human-robot interaction", *Robotics and Autonomous Systems*, vol. 42, pp. 191-201, 2003.
9. [Arnott, 2005] Arnott, D. and Pervan, G., "A critical analysis of Decision Support Systems research",

- Journal of Information Technology*, vol. 20, pp. 67-87, 2005.
10. [Arnott, 2008] Arnott, D. and Pervan, G., "Eight key issues for the decision support systems discipline", *Decision Support Systems*, vol. 44, pp. 657-672, 2008.
  11. [Asfour, 2006] Asfour, T., Gyarfas, F., Azad, P., *et al.*, "Imitation Learning of Dual-Arm Manipulation Tasks in Humanoid Robots", in *Humanoid Robots, 2006 6th IEEE-RAS International Conference on*, pp. 40-47, 2006.
  12. [Asoh, 1999] Asoh, H., Matsui, T., Fry, J., *et al.*, "A spoken dialog system for a mobile office robot," in *6th European Conference on Speech Communication and Technology*, Budapest, Hungary, p. 1139-1142, 1999.
  13. [Banu, 2012] Banu, S. M., Danciu, G. M., **Boboc, R. G.**, Moga, H. and Balany, C., A novel approach for face expressions recognition, in *Proceedings of the 10th IEEE Jubilee International Symposium on Intelligent Systems and Informatics (SISY 2012)*, Subotica, Serbia, pp. 537-541, 2012.
  14. [Bartneck, 2001] Bartneck, C. and Okada, M., "Robotic user interfaces", in *Human and Computer Conference*, pp. 130-140, 2001.
  15. [Bauer, 2009] Bauer, A., Wollherr, D., and Buss, M., "Information retrieval system for human-robot communication - Asking for directions", in *IEEE International Conference on Robotics and Automation (ICRA '09)*, pp. 4150-4155, 2009.
  16. [Bautista, 2013] Bautista, M. Á., Hernández-Vela, A., Victor Ponce, V., *et al.*, "Probability-Based Dynamic Time Warping for Gesture Recognition on RGB-D Data", *Advances in Depth Image Analysis and Applications*, vol. 7854, pp. 126-

## BIBLIOGRAFIE

- 135, 2013.
17. [Beer, 2011] Beer, J. M., Prakash, A., Mitzner, T. L., *et al.*, "Understanding robot acceptance", Georgia Institute of Technology, School of Psychology, Human Factors and Aging Laboratory, Atlanta, GA, 2011.
  18. [Behnke, 2008] Behnke, S., "Humanoid Robots - From Fiction to Reality?", *Künstliche Intelligenz*, vol. 22, pp. 5-9, 2008.
  19. [Bemelmans, 2012] Bemelmans, R., Gelderblom, G. J., Jonker, P., *et al.*, "Socially Assistive Robots in Elderly Care: A Systematic Review into Effects and Effectiveness", *Journal of the American Medical Directors Association*, vol. 13, pp. 114-120, 2012.
  20. [Benayed, 2014] Benayed, S., Eltaher, M., and Lee, J., "Developing Kinect-like Motion Detection System using Canny Edge Detector", *American Journal of Computing Research Repository*, vol. 2, pp. 28-32, 2014.
  21. [Beran, 2011] Beran, T. N., Ramirez-Serrano, A., Kuzyk, R., *et al.*, "Understanding how children understand robots: Perceived animism in child-robot interaction", *International Journal of Human-Computer Studies*, vol. 69, pp. 539-550, 2011.
  22. [Berndt, 1994] Berndt, D. J. and Clifford, J., "Using Dynamic Time Warping to Find Patterns in Time Series", in *KDD Workshop*, Seattle, Washington, USA, pp. 359-370, 1994.
  23. [Beuter, 2008] Beuter, N., Spexard, T., Lutkebohle, I., *et al.*, "Where is this? - gesture based multimodal interaction with an anthropomorphic robot", in *8th IEEE-RAS International Conference on Humanoid Robots (Humanoids 2008)*, pp. 585-591, 2008.
  24. [Billard, 2008] Billard, A., Calinon, S., Dillmann, R., *et al.*,

- "Robot Programming by Demonstration", in *Handbook of Robotics*, B. Siciliano and O. Khatib, Eds., ed: MIT Press, 2008.
25. [Billard, 2006] Billard, A., Robins, B., Nadel, J., *et al.*, "Building Robota, a Mini-Humanoid Robot for the Rehabilitation of Children with Autism", *RESNA Assistive Technology Journal*, vol. 19, pp. 37-49, 2006.
26. [Blake, 2011] Blake, J., *Natural user interfaces in .NET : WPF 4, Surface 2, and Kinect*, London: Manning Publications, 2011.
27. [Blanson Henkemans, 2013] Blanson Henkemans, O. A., Bierman, B. P. B., Janssen, J., *et al.*, "Using a robot to personalise health education for children with diabetes type 1: A pilot study", *Patient Education and Counseling*, vol. 92, pp. 174-181, 2013.
28. [Boboc, 2013] **Boboc, R. G.**, Toma, M.-I., Panfir, A. N., *et al.*, "Learning new skills by a humanoid robot through imitation", in *IEEE 14th International Symposium on Computational Intelligence and Informatics (CINTI 2013)*, pp. 515-519, 2013.
29. [Boboc, 2014a] **Boboc, R. G.**, Moga, H., and Talabă, D., "An Educational Humanoid Laboratory Tour Guide Robot", *Procedia - Social and Behavioral Sciences*, vol. 141, pp. 424-430, 2014.
30. [Boboc, 2014b] **Boboc, R. G.** and Talabă, D., Point-and-command Paradigm in Human - Robot Interaction, in *Proceedings of the the 6th Győr Symposium and 3rd Hungarian-Polish and 1st Hungarian-Romanian Joint Conference on Computational Intelligence*, Győr, Hungary, pp. 21-27, 2014
31. [Boboc, 2015a] **Boboc, R. G.**, Duguleană, M. and Talabă, D., Natural Interaction with an Assistive Humanoid Robot, *Applied Mechanics and Materials*, vol. 762, pp. 189-194, 2015.
32. [Boboc, 2015b] **Boboc, R. G.**, Antonya, C., Point-and-



- command Paradigm for Interaction with Assistive Robots, *International Journal of Advanced Robotic Systems*, ISSN: 1729-8814, acceptat spre publicare în 02.04.2015.
33. [Boden, 2006] Boden, M. A., *Mind as Machine: A History of Cognitive Science*: Oxford University Press, 2006.
34. [Böhme, 2003] Böhme, H.-J., Wilhelm, T., Key, J., *et al.*, "An approach to multi-modal human-machine interaction for intelligent service robots", *Robotics and Autonomous Systems*, vol. 44, pp. 83-96, 2003.
35. [Bolt, 1980] Bolt, R. A., "Put-that-there: Voice and gesture at the graphics interface", *SIGGRAPH Comput. Graph.*, vol. 14, pp. 262-270, 1980.
36. [Boulabiar, 2011] Boulabiar, M.-I., Burger, T., Poirier, F., *et al.*, "A Low-Cost Natural User Interaction Based on a Camera Hand-Gestures Recognizer", in *Human-Computer Interaction. Interaction Techniques and Environments*, vol. 6762, J. Jacko, Ed., ed: Springer Berlin Heidelberg, pp. 214-221, 2011.
37. [Boumella, 2012] Boumella, N., Figueroa, J. C., and Iqbal, S., *Enhancing Fuzzy Controllers Using Generalized Orthogonality Principle*, 2012.
38. [Breazeal, 2005] Breazeal, C., "Designing Socially Intelligent Robots", 2005.
39. [Breazeal, 2003] Breazeal, C., "Emotion and sociable humanoid robots", *International Journal of Human-Computer Studies*, vol. 59, pp. 119-155, 2003.
40. [Breazeal, 2004] Breazeal, C., Brooks, A. G., and J. G., *et al.*, "Tutelage and Collaboration for Humanoid Robots", *I. J. Humanoid Robotics*, vol. 1, pp. 315-348, 2004.
41. [Breuer, 2012] Breuer, T., Giorgana Macedo, G., Hartanto, R., *et al.*, "Johnny: An Autonomous Service

- Robot for Domestic Environments", *Journal of Intelligent & Robotic Systems*, vol. 66, pp. 245-272, 2012.
42. [Breuer, 2010] Breuer, T., Ploeger, P. G., and Kraetzschmar, G. K., "Precise pointing target recognition for human-robot interaction", presented at the Workshop on Domestic Service Robots in the Real World, SIMPAR, Darmstadt, Germany, 2010.
43. [Broekens, 2009] Broekens, J., Heerink, M., and Rosendal, H., "Assistive social robots in elderly care: a review", *Gerontechnology*, vol. 8, pp. 94-103, 2009.
44. [Broz, 2013] Broz, F., Nourbakhsh, I., and Simmons, R., "Planning for Human-Robot Interaction in Socially Situated Tasks: The Impact of Representing Time and Intention", *International Journal of Social Robotics*, vol. 5, pp. 193-214, 2013.
45. [Brueckmann, 2007] Brueckmann, R., Scheidig, A., and Gross, H., "Adaptive Noise Reduction and Voice Activity Detection for improved Verbal Human-Robot Interaction using Binaural Data", in *Robotics and Automation, 2007 IEEE International Conference on*, pp. 1782-1787, 2007.
46. [Bruemmer, 2007] Bruemmer, D. J., Gertman, D. I., and Nielsen, C. W., "Metaphors to Drive By: Exploring New Ways to Guide Human - Robot Interaction", *The Open Cybernetics and Systemics Journal*, vol. 1, pp. 5-12, 2007.
47. [Burridge, 2003] Burridge, R. R., Graham, J., Shillcutt, K., *et al.*, "Experiments with an EVA Assistant Robot", 2003.
48. [Butcher, 2000] Butcher, C. and Goldin-Meadow, S., "Gesture and the transition from one- to two-word speech: when hand and mouth come together", in *Language and Gesture*, D.

## BIBLIOGRAFIE

- McNeill, Ed., ed Cambridge, United Kingdom: Cambridge University Press, pp. 235-258, 2000.
49. [Caballero Morales, 2013] Caballero Morales, S. O., Enríquez, G. B., and Romero, F. T., "Speech-Based Human and Service Robot Interaction: An Application for Mexican Dysarthric People", *International Journal of Advanced Robotic Systems*, vol. 10, pp. 1-14, 2013.
50. [Canny, 1986] Canny, J., "A Computational Approach to Edge Detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, pp. 679-698, 1986.
51. [Carlisle, 2000] Carlisle, B., "Robot mechanisms," in *Proceedings of IEEE International Conference on Robotics and Automation (ICRA '00)*, vol.1, pp. 701-708, 2000.
52. [Carmona, 2012] Carmona, J. and Climent, J., "A Performance Evaluation of HMM and DTW for Gesture Recognition", in *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, vol. 7441, L. Alvarez, M. Mejail, L. Gomez, and J. Jacobo, Eds., ed: Springer Berlin Heidelberg, pp. 236-243, 2012.
53. [Cartmill, 2011] Cartmill, E. A., Demir, O. E., and Goldin-Meadow, S., "Studying gesture", in *The guide to research methods in child language*, E. Hoff, Ed., ed England: West Sussex, pp. 208-225, 2011.
54. [Cebi, 2010] Cebi, S. and Kahraman, C., "Developing a group decision support system based on fuzzy information axiom", *Knowledge-Based Systems*, vol. 23, pp. 3-16, 2010.
55. [Chai, 2013] Chai, X., Li, G., Lin, Y., *et al.*, "Sign Language Recognition and Translation with Kinect", in *IEEE International Conference on Automatic Face and Gesture Recognition*, Shanghai, China,

- 2013.
56. [Chang, 2011] Chang, J. and Yeh, M.-Y., "Monitoring Multiple Streams with Dynamic Time Warping using Graphic Processors", in *Parallel Data Mining Workshop (IPDM)*, Mesa, Arizona, USA, pp. 11-20, 2011.
  57. [Chao, 2014] Chao, F., Wang, Z., Shang, C., *et al.*, "A developmental approach to robotic pointing via human-robot interaction", *Information Sciences*, vol. 283, pp.288-303, 2014.
  58. [Chella, 2009] Chella, A. and Macaluso, I., "The perception loop in CiceRobot, a museum guide robot", *Neurocomputing*, vol. 72, pp. 760-766, 2009.
  59. [Cheng, 2013] Cheng, J., Bian, W., and Tao, D., "Locally regularized sliced inverse regression based 3D hand gesture recognition on a dance robot", *Information Sciences*, vol. 221, pp. 274-283, 2013.
  60. [Cheng, 2012] Cheng, J., Xie, C., Bian, W., *et al.*, "Feature fusion for 3D hand gesture recognition by learning a shared hidden space", *Pattern Recognition Letters*, vol. 33, pp. 476-484, 2012.
  61. [Cheung, 2005] Cheung, W., Leung, L. C., and Tam, P. C. F., "An intelligent decision support system for service network planning", *Decision Support Systems*, vol. 39, pp. 415-428, 2005.
  62. [Chida, 2004] Chida, K., Okuma, I., Isoda, S., *et al.*, "Development of a new anthropomorphic flutist robot WF-4", in *Proceedings of IEEE International Conference on Robotics and Automation (ICRA '04)*, vol. 1, pp. 152-157, 2004.
  63. [Christensen, 2008] Christensen, H. and Hager, G., "Sensing and Estimation", in *Springer Handbook of Robotics*, B. Siciliano and O. Khatib, Eds., ed: Springer Berlin Heidelberg, pp. 87-107, 2008.

## BIBLIOGRAFIE

64. [Chu, 2002] Chu, S., Keogh, E., Hart, D., *et al.*, "Iterative Deepening Dynamic Time Warping for Time Series", in *2nd SIAM International Conference on Data Mining*, Arlington, VA, USA, 2002.
65. [Colonnesi, 2010] Colonnesi, C., Stams, G. J. J. M., Koster, I., *et al.*, "The relation between pointing and language development: A meta-analysis", *Developmental Review*, vol. 30, pp. 352–366, 2010.
66. [Cruz, 2012] Cruz, L., Lucio, D., and Velho, L., "Kinect and RGBD Images: Challenges and Applications", in *25th SIBGRAPI Conference on Graphics, Patterns and Images Tutorials (SIBGRAPI-T 2012)*, pp. 36-49, 2012.
67. [Cruz, 2011] Cruz, L., Reis, L. P., and Rei, L., "Generic optimization of humanoid robots behaviours," presented at *the 15th Portuguese Conference on Artificial Intelligence*, Lisbon, Portugal, 2011.
68. [Csapo, 2012a] Csapo, A., Gilmartin, E., Grizou, J., *et al.*, "Multimodal conversational interaction with a humanoid robot", in *IEEE 3rd International Conference on Cognitive Infocommunications (CogInfoCom 2012)*, pp. 667-672, 2012.
69. [Csapo, 2012b] Csapo, A., Gilmartin, E., Grizou, J., *et al.*, "Speech, gaze and gesturing: multimodal conversational interaction with NAO robot", in *8th International Summer Workshop on Multimodal Interfaces ( eNTERFACE'12)*, Supélec, Metz, France, pp.. 7-12, 2012.
70. [Dautenhahn, 2007] Dautenhahn, K., "Socially intelligent robots: Dimensions of human-robot interaction", *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 362, pp. 679–704, 2007.
71. [Dautenhahn, 2005] Dautenhahn, K., Woods, S., Kaouri, C., *et al.*, "What is a robot companion - friend, assistant or butler?", in *IEEE/RSJ International*

- Conference on Intelligent Robots and Systems (IROS 2005)*, pp. 1192-1197, 2005.
72. [de Graaf, 2013] de Graaf, M. M. A. and Ben Allouch, S., "Exploring influencing variables for the acceptance of social robots", *Robotics and Autonomous Systems*, vol. 61., no. 12, pp. 1476–1486, 2013.
73. [Delaunay, 2010] Delaunay, F., de Greeff, J., and Belpaeme, T., "A study of a retro-projected robotic face and its effectiveness for gaze reading by humans", in *5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 39-44, 2010.
74. [Delgado, 2011] Delgado, B., Gómez, J. C., and Sarriá, E., "Pointing gestures as a cognitive tool in young children: Experimental evidence", *Journal of Experimental Child Psychology*, vol. 110, pp. 299-312, 2011.
75. [Ding, 2009] Ding, X. C., Powers, M., Egerstedt, M., *et al.*, "Executive decision support", *IEEE Robotics & Automation Magazine*, vol. 16, pp. 73-81, 2009.
76. [Droeschel, 2011] Droeschel, D., Stuckler, J., and Behnke, S., "Learning to interpret pointing gestures with a time-of-flight camera", in *6th ACM/IEEE International Conference on Human-Robot Interaction (HRI 2011)*, pp. 481-488, 2011.
77. [Duguleană, 2012] Duguleană, M., Barbuceanu, F. G., Teirelbar, A., *et al.*, "Obstacle avoidance of redundant manipulators using neural networks based reinforcement learning", *Robotics and Computer-Integrated Manufacturing*, vol. 28, pp. 132-146, 2012.
78. [Ebert, 2013] Ebert, L. C., Hatch, G., Thali, M. J., *et al.*, "Invisible touch—Control of a DICOM viewer with finger gestures using the Kinect depth camera", *Journal of Forensic Radiology and Imaging*, vol. 1, pp. 10-14, 2013.

## BIBLIOGRAFIE

79. [El-laithy, 2012] El-laithy, R. A., Jidong, H., and Yeh, M., "Study on the use of Microsoft Kinect for robotics applications", in *IEEE Position Location and Navigation Symposium (PLANS 2012)*, pp. 1280-1288, 2012.
80. [Engelberger, 2001] Engelberger, G., "NASA's Robonaut", *The Industrial Robot*, vol. 28, pp. 35-39, 2001.
81. [Ezer, 2008] Ezer, N., "Is a robot an appliance, teammate, or friend? Age-related differences in expectations of and attitudes towards personal home-based robots", Retrieved from ProQuest® Dissertations & Theses (3345975), 2008.
82. [Feil-Seifer, 2005] Feil-Seifer, D. and Mataric, M. J., "Defining socially assistive robotics", in *9th International Conference on Rehabilitation Robotics (ICORR 2005)*, pp. 465-468, 2005.
83. [Feil-Seifer, 2009] Feil-Seifer, D. J. and Mataric, M. J., "Human-Robot Interaction", in *Invited contribution to Encyclopedia of Complexity and Systems Science*, R. A. Meyers, Ed., ed New York: Springer, pp. 4643-4659, 2009.
84. [Fojtů, 2012] Fojtů, Š., Havlena, M., and Pajdla, T., "NAO robot localization and navigation using fusion of odometry and visual sensor data", presented at the *Proceedings of the 5th international conference on Intelligent Robotics and Applications*, Volume Part II, Montreal, QC, Canada, 2012.
85. [Fong, 2003a] Fong, T., Nourbakhsh, I., and Dautenhahn, K., "A survey of socially interactive robots", *Robotics and Autonomous Systems*, vol. 42, pp. 143-166, 2003.
86. [Fong, 2003b] Fong, T., Thorpe, C., and Baur, C., "Collaboration, Dialogue, Human-Robot Interaction", in *Robotics Research.*, vol. 6, R. Jarvis and A. Zelinsky, Eds., ed: Springer

- Berlin Heidelberg, pp. 255-266, 2003.
87. [Fridin, 2014] Fridin, M., "Storytelling by a kindergarten social assistive robot: A tool for constructive learning in preschool education", *Computers & Education*, vol. 70, pp. 53-64, 2014.
  88. [Fujita, 1998] Fujita, M. and Kitano, H., "Development of an Autonomous Quadruped Robot for Robot Entertainment", *Autonomous Robots*, vol. 5, pp. 7-18, 1998.
  89. [Fujita, 2004] Fujita, M., "On activating human communications with pet-type robot AIBO", *Proceedings of the IEEE*, vol. 92, pp. 1804-1813, 2004.
  90. [Fukui, 2005] Fukui, K., Nishikawa, K., Kuwae, T., *et al.*, "Development of a New Human-like Talking Robot for Human Vocal Mimicry", in *Proceedings of the 2005 IEEE International Conference on Robotics and Automation (ICRA 2005)*, pp. 1437-1442, 2005.
  91. [George, 2010] George, R. and Blake, J., "Objects, Containers, Gestures, and Manipulations: Universal Foundational Metaphors of Natural User Interfaces", presented at the CHI (*Conference on Human Factors in Computing Systems*), Atlanta, Georgia, USA, 2010.
  92. [Gockley, 2006] Gockley, R., Simmons, R., and Forlizzi, J., "Modeling Affect in Socially Interactive Robots", in *the 15th IEEE International Symposium on Robot and Human Interactive Communication, (ROMAN 2006)*, pp. 558-563, 2006.
  93. [Goldin-Meadow, 2007] Goldin-Meadow, S., "Pointing Sets the Stage for Learning Language - and Creating Language", *Child Development*, vol. 78, pp. 741-745, 2007.
  94. [Goodrich, 2007] Goodrich, M. A. and Schultz, A. C., "Human-Robot Interaction: A Survey," *Foundations and*



- Trends in Human-Computer Interaction*, vol. 1, pp. 203-275, 2007.
95. [Gouaillier, 2009] Gouaillier, D., Hugel, V., Blazevic, P., *et al.*, "Mechatronic design of NAO humanoid", in *IEEE International Conference on Robotics and Automation (ICRA '09)*, pp. 769-774, 2009.
96. [Graefe, 2009] Graefe, V. and Bischoff, R., "From ancient machines to intelligent robots - A technical evolution", in *the 9th International Conference on Electronic Measurement & Instruments (ICEMI '09)*, pp. 3-418-3-431, 2009.
97. [Granata, 2013] Granata, C., Bidaud, P., and, J. S., *et al.*, "Human activity analysis: a personal robot integrating a framework for robust person detection and tracking and physical based motion analysis", *PALADYN Journal of Behavioral Robotics*, vol. 4, pp. 131-146, 2013.
98. [Gunn, 2013] Gunn, T. and Anderson, J., "Dynamic Heterogeneous Team Formation for Robotic Urban Search and Rescue", *Procedia Computer Science*, vol. 19, pp. 22-31, 2013.
99. [Hafner, 2005] Hafner, V. and Kaplan, F., "Learning to Interpret Pointing Gestures: Experiments with Four-Legged Autonomous Robots", in *Biomimetic Neural Learning for Intelligent Robots*. vol. 3575, S. Wermter, G. Palm, and M. Elshaw, Eds., ed: Springer Berlin Heidelberg, pp. 225-234, 2005.
100. [Harris, 2010] Harris, J. and Sharlin, E., "Exploring emotive actuation and its role in human-robot interaction", presented at *the Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction*, Osaka, Japan, 2010.
101. [Heckel, 2008] Heckel, F. and Smart, W. D., "Mapping the Field of Human-Robot Interaction", presented at the *IEEE International Conference on Robotics and Automation*, Pasadena,

- California, 2008.
102. [Heckel, 2006] Heckel, F. and Smart, W. D., "Non-speech aural communication for robots", presented at the *AAAI Fall Symposium on Aurally Informed Performance: Integrating Machine Listening and Auditory Presentation in Robotics Systems*, Cambridge, MA, 2006.
103. [Heikkilä, 2013] Heikkilä, T., Dalgaard, L., and Koskinen, J., "Designing Autonomous Robot Systems - Evaluation of the R3-COP Decision Support System Approach", in *SAFECOMP 2013 - Workshop DECS (ERCIM/EWICS Workshop on Dependable Embedded and Cyber-physical Systems)*, Toulouse, France, 2013.
104. [Hewett, 1992] Hewett, T. T., Baecker, R., Card, S., *et al.*, "ACM Curricula for Human-Computer Interaction", 1992.
105. [Hinckley, 2012] Hinckley, K. and Wigdor, D., "Input Technologies and Techniques", in *The Human-Computer Interaction Handbook – Fundamentals, Evolving Technologies and Emerging Applications*, J. Jacko, Ed., ed: Taylor & Francis, 2012.
106. [Holland, 1997] Holland, O. E., "Grey Walter: The Pioneer of Real Artificial Life", in *Proceedings of the 5th International Workshop on Artificial Life*, Nara-shi, Japan, pp. 34-44, 1997.
107. [Hong, 2010] Hong, L. and Li, X., "A selection method of speech vocabulary for human-robot speech interaction", in *IEEE International Conference on Systems Man and Cybernetics (SMC 2010)*, pp. 2243-2248, 2010.
108. [Hornung, 2010] Hornung, A., Wurm, K. M., and Bennewitz, M., "Humanoid robot localization in complex indoor environments", in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2010)*, pp. 1690-1695, 2010.

## BIBLIOGRAFIE

109. [Internet 2.1] <http://whatis.techtarget.com/definition/natural-user-interface-NUI>
110. [Internet 2.2] [http://en.wikipedia.org/wiki/Natural\\_user\\_interface](http://en.wikipedia.org/wiki/Natural_user_interface)
111. [Internet 2.3] [http://en.wikipedia.org/wiki/William\\_Grey\\_Walter](http://en.wikipedia.org/wiki/William_Grey_Walter)
112. [Internet 2.4] <http://www2.derby.ac.uk/littlespaces/news/30-rousseau/179-cc-1b-turin>
113. [Internet 2.5] <http://www.ifr.org/service-robots/>
114. [Internet 2.6] <http://m.dexonline.ro/definitie/gest>
115. [Internet 3.1] [https://community.aldebaran-robotics.com/doc/1-14/family/robots/joints\\_robot.html#robot-joints](https://community.aldebaran-robotics.com/doc/1-14/family/robots/joints_robot.html#robot-joints)
116. [Internet 3.2] [http://doc.aldebaran.com/2-1/family/robots/video\\_robot.html](http://doc.aldebaran.com/2-1/family/robots/video_robot.html)
117. [Internet 4.1] <https://msdn.microsoft.com/en-us/library/hh973078.aspx>
118. [Internet 5.1] <http://www.mathworks.com/products/fuzzy-logic/>
119. [Internet 6.1] <https://msdn.microsoft.com/en-us/library/jj131041.aspx>
120. [Hussain, 2014] Hussain, A. T., Ahmed, S. F., and Hazry, D., "Tracking and replication of hand movements by teleguided intelligent manipulator robot", *Robotica*, vol. 33, no. 1, pp. 141-156, 2015.
121. [Iarlori, 2014] Iarlori, S., Ferracuti, F., Giantomassi, A., *et al.*, "RGB-D Video Monitoring System to Assess the Dementia Disease State Based on Recurrent Neural Networks with Parametric Bias Action Recognition and DAFS Index Evaluation", in *Computers Helping People with Special Needs*, vol. 8548, K. Miesenberger, *et al.*, Eds., ed: Springer International Publishing, 2014, pp. 156-163, 2014.
122. [Ierache, 2013] Ierache, J., Pereira, G., Sattolo, I., *et al.*, "Robot Control on the Basis of Bio-electrical Signals",

- in *Robot Intelligence Technology and Applications 2012*, vol. 208, J.-H. Kim, et al., Eds., ed: Springer Berlin Heidelberg, pp. 337-346, 2013.
123. [Ikeda, 2007] Ikeda, A., Mizoguchi, H., Sasaki, Y., *et al.*, "2D Sound Source Localization in Azimuth & Elevation from Microphone Array by Using a Directional Pattern of Element", in *IEEE Sensors*, pp. 1213-1216, 2007.
124. [Itakura, 1975] Itakura, F., "Minimum prediction residual principle applied to speech recognition", *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 23, pp. 67-72, 1975.
125. [Izadi, 2011] Izadi, S., Kim, D., Hilliges, O., *et al.*, "KinectFusion: real-time 3D reconstruction and interaction using a moving depth camera", presented at *the Proceedings of the 24th annual ACM symposium on User interface software and technology*, Santa Barbara, California, USA, 2011.
126. [Iztueta, 2011] Iztueta, E. J., "Advances towards behaviour-based indoor robotic exploration", Doctor of Philosophy in Computer Science, Department of Computer Science and Artificial Intelligence University of Basque Country, Donostia, Spain, 2011.
127. [Jacob, 2012] Jacob, M. G., Wachs, J. P., and Packer, R. A., "Hand-gesture-bases sterile interface for the operating room using contextual cues for the navigation of radiological images", *Journal of the American Medical Informatics Association*, vol. 20, pp. 183-186, 2012.
128. [Jaimes, 2007] Jaimes, A. and Sebe, N., "Multimodal human-computer interaction: A survey", *Computer Vision and Image Understanding*, vol. 108, pp. 116-134, 2007.
129. [Jin, 2002] Jin, Z. and Bose, B. K., "Evaluation of

- membership functions for fuzzy logic controlled induction motor drive", in *IEEE 28th Annual Conference of the Industrial Electronics Society (IECON '02*, vol.1, pp. 229-234, 2002.
130. [Jingwei, 2012] Jingwei, G. and Meng, M. Q. H., "Study on distance measurement for NAO humanoid robot", in *IEEE International Conference on Robotics and Biomimetics (ROBIO 2012)*, pp. 283-286, 2012.
131. [Jones, 2010] Jones, G., Berthouze, N., Bielski, R., *et al.*, "Towards a situated, multimodal interface for multiple UAV control", in *IEEE International Conference on Robotics and Automation (ICRA 2010)*, pp. 1739-1744, 2010.
132. [Jones, 2003] Jones, K. S., Middendorf, M., McMillan, G. R., *et al.*, "Comparing mouse and steady-state visual evoked response-based control", *Interacting with Computers*, vol. 15, pp. 603-621, 2003.
133. [Kahn, 2008] Kahn, P. H., Freier, N. G., Kanda, T., *et al.*, "Design patterns for sociality in human-robot interaction", presented at *the Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction*, Amsterdam, The Netherlands, 2008.
134. [Kajastila, 2013] Kajastila, R. and Lokki, T., "Eyes-free interaction with free-hand gestures and auditory menus", *International Journal of Human-Computer Studies*, vol. 71, pp. 627-640, 2013.
135. [Kanda, 2004] Kanda, T., Hirano, T., Eaton, D., *et al.*, "Interactive robots as social partners and peer tutors for children: a field trial", *Hum.-Comput. Interact.*, vol. 19, pp. 61-84, 2004.
136. [Kaneko, 2014] Kaneko, S.-i. and Capi, G., "Human-robot Communication for Surveillance of Elderly

- People in Remote Distance", *IERI Procedia*, vol. 10, pp. 92-97, 2014.
137. [Kang, 2013] Kang, J., Zhong, K., Qin, S., *et al.*, "Instant 3D design concept generation and visualization by real-time hand gesture recognition", *Computers in Industry*, vol. 64, pp. 785-797, 2013.
138. [Karam, 2005] Karam, M. and Schraefel, M. C., "A taxonomy of Gestures in Human Computer Interaction", University of Southampton, 2005.
139. [Karray, 2008] Karray, F., Alemzadeh, M., Saleh, J. A., *et al.*, "Human-Computer Interaction: Overview on State of the Art", *International Journal on Smart Sensing and Intelligent Systems*, vol. 1, pp. 137-159, 2008.
140. [Kaupp, 2010] Kaupp, T., Makarenko, A., and Durrant-Whyte, H., "Human-robot communication for collaborative decision making — A probabilistic approach", *Robotics and Autonomous Systems*, vol. 58, pp. 444-456, 2010.
141. [Kendon, 2004] Kendon, A., "Gesture: Visible action as utterance", Cambridge, UK: Cambridge University Press, 2004.
142. [Keogh, 2005] Keogh, E. and Ratanamahatana, C. A., "Exact indexing of dynamic time warping", *Knowl. Inf. Syst.*, vol. 7, pp. 358-386, 2005.
143. [Keren, 2014] Keren, G. and Fridin, M., "Kindergarten Social Assistive Robot (KindSAR) for children's geometric thinking and metacognitive development in preschool education: A pilot study", *Computers in Human Behavior*, vol. 35, pp. 400-412, 2014.
144. [Kerepesi, 2006] Kerepesi, A., Kubinyi, E., Jonsson, G. K., *et al.*, "Behavioural comparison of human-animal (dog) and human-robot (AIBO) interactions," *Behavioural Processes*, vol. 73, pp. 92-99, 2006.

## BIBLIOGRAFIE

145. [Kettebekov, 2001] Kettebekov, S. and Sharma, R., "Toward Natural Gesture/Speech Control of a Large Display", presented at *the Proceedings of the 8th IFIP International Conference on Engineering for Human-Computer Interaction (EHCO 2001)*, Toronto, Canada, 2001.
146. [Kim, 2014] Kim, D.-W., Lee, J., Lim, H., *et al.*, "Efficient dynamic time warping for 3D handwriting recognition using gyroscope equipped smartphones," *Expert Systems with Applications*, vol. 41, pp. 5180-5189, 2014.
147. [Kim, 2004] Kim, S.-W., Park, S., and Chu, W. W., "Efficient processing of similarity search under time warping in sequence databases: an index-based approach", *Inf. Syst.*, vol. 29, pp. 405-420, 2004.
148. [Kipp, 2004] Kipp, M., "Gesture generation by imitation: From human behavior to computer character animation", Ph. D, Saarland University, 2004.
149. [Kita, 2003] Kita, S., "Pointing: A foundational building block in human communication", in *Pointing: Where language, culture, and cognition meet*, S. Kita, Ed., ed Mahwah, New Jersey: NJ: Erlbaum, pp. 1-8, 2003.
150. [Kitsunezaki, 2013] Kitsunezaki, N., Adachi, E., Masuda, T., *et al.*, "KINECT applications for the physical rehabilitation", in *IEEE International Symposium on Medical Measurements and Applications Proceedings (MeMeA 2013)*, pp. 294-299, 2013.
151. [Kogan, 1998] Kogan, J. A. and Margoliash, D., "Automated recognition of bird song elements from continuous recordings using dynamic time warping and hidden Markov models: a comparative study", *Journal of the Acoustical Society of America*, vol. 103, pp. 2185-2196, 1998.

152. [Kollar, 2014] Kollar, T., Tellex, S., Roy, D., *et al.*, "Grounding Verbs of Motion in Natural Language Commands to Robots", in *Experimental Robotics*, vol. 79, O. Khatib, *et al.*, Eds., ed: Springer Berlin Heidelberg, pp. 31-47, 2014.
153. [Kollar, 2010] Kollar, T., Tellex, S., Roy, D., *et al.*, "Toward understanding natural language directions", presented at the *Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction*, Osaka, Japan, 2010.
154. [Kondaxakis, 2014] Kondaxakis, P., Pajarinen, J., and Kyrki, V., "Real-Time Recognition of Pointing Gestures for Robot to Robot Interaction", presented at the International Conference on Intelligent Robots and Systems (IROS14), Chicago, USA, 2014.
155. [Koppula, 2013] Koppula, H. S. and Saxena, A., "Anticipating human activities for reactive robotic response", in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2013)*, pp. 2071-2071, 2013.
156. [Kozima, 2007] Kozima, H., Nakagawa, C., and Yasuda, Y., "Children-robot interaction: a pilot study in autism therapy", in *Progress in Brain Research*, vol. 164, C. v. Hofsten and K. Rosander, Eds., ed: Elsevier, pp. 385-400, 2007.
157. [Kubota, 2013] Kubota, S., Nakata, Y., Eguchi, K., *et al.*, "Feasibility of Rehabilitation Training With a Newly Developed Wearable Robot for Patients With Limited Mobility", *Archives of Physical Medicine and Rehabilitation*, vol. 94, pp. 1080-1087, 2013.
158. [Lakoff, 2003] Lakoff, G. and Johnson, M., *Metaphors We Live By*: University of Chicago Press, 2003.
159. [Lamy, 2010] Lamy, J.-B., Ellini, A., Nobécourt, J., *et al.*, "Testing methods for decision support



- systems", in *Decision support systems*, ed Vukovar, Croatia: Intech, pp. 87–98, 2010.
160. [Lee, 2011] Lee, H.-T., Lin, W.-C., and Huang, C.-H., "Indoor Surveillance Security Robot with a Self-Propelled Patrolling Vehicle", *Journal of Robotics*, vol. 2011, 2011.
161. [Lemire, 2009] Lemire, D., "Faster retrieval with a two-pass dynamic-time-warping lower bound", *Pattern Recogn.*, vol. 42, pp. 2169-2180, 2009.
162. [Leroux, 2013] Leroux, C., Lebec, O., Ben Ghezala, M. W., *et al.*, "ARMEN: Assistive robotics to maintain elderly people in natural environment", *IRBM*, vol. 34, pp. 101-107, 2013.
163. [Li, 2013] Li, Y.-T. and Wachs, J. P., "Recognizing hand gestures using the weighted elastic graph matching (WEGM) method", *Image and Vision Computing*, vol. 31, pp. 649-657, 2013.
164. [Lim, 2006] Lim, H.-o., Hyon, S.-H., Setiawan, S. A., *et al.*, "Quasi-human biped walking", *Robotica*, vol. 24, pp. 257-268, 2006.
165. [Lin, 2011] Lin, P., Abney, K., and Bekey, G., "Robot ethics: Mapping the issues for a mechanized world", *Artificial Intelligence*, vol. 175, pp. 942-949, 2011.
166. [Linda, 2011] Linda, O. and Manic, M., "Self-Organizing Fuzzy Haptic Teleoperation of Mobile Robot Using Sparse Sonar Data", in *IEEE Trans. on Industrial Electronics*, pp. 3187-3195, 2011.
167. [Liu, 2010] Liu, R. and Wang, Y., "Azimuthal source localization using interaural coherence in a robotic dog: Modeling and application", *Robotica*, vol. 28, pp. 1013-1020, 2010.
168. [Loc, 2011] Loc, V.-G., Koo, I. M., Tran, D. T., *et al.*, "Improving traversability of quadruped walking robots using body movement in 3D rough terrains", *Robotics and Autonomous Systems*, vol. 59, pp. 1036-1048, 2011.

169. [Lyons, 1977] Lyons, J., *Semantics*: Cambridge University Press, 1977.
170. [Maimone, 2011] Maimone, A. and Fuchs, H., "Encumbrance-free telepresence system with real-time 3D capture and display using commodity depth cameras", in *the 10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR 2011)*, pp. 137-146, 2011.
171. [Martin, 2010] Martin, C., Steege, F.-F., and Gross, H.-M., "Estimation of pointing poses for visually instructing mobile robots under real world conditions", *Robotics and Autonomous Systems*, vol. 58, pp. 174-185, 2010.
172. [Martínez, 2013] Martínez, J. A., Úbeda, A., Iáñez, E., *et al.*, "Multimodal System Based on Electrooculography and Voice Recognition to Control a Robot Arm", *International Journal of Advanced Robotic Systems*, vol. 10, pp. 1-9, 2013.
173. [Mavridis, 2014] Mavridis, N., "A Review of Verbal and Non-Verbal Human-Robot Interactive Communication", *Robotics and Autonomous Systems*, vol 63, pp. 22-35, 2015.
174. [McGhan, 2012] McGhan, C. L. R., Nasir, A., and Atkins, E. M., "Human Intent Prediction Using Markov Decision Processes", presented at *the Infotech@Aerospace Garden Grove*, California, 2012.
175. [McNeill, 1982] McNeill, D. and Levy, E., "Conceptual representations in language activity and gesture", in *Speech, place, and action*, J. J. W. Klein, Ed., ed New York: John Wiley & Sons, pp. 271-295, 1982.
176. [McNeill, 1992] McNeill, D., *Hand and mind: What gestures reveal about thought*, Chicago, IL: University of Chicago Press, 1992.
177. [Meng, 2006] Meng, Q. and Lee, M. H., "Design issues for

- assistive robotics for the elderly", *Advanced Engineering Informatics*, vol. 20, pp. 171-186, 2006.
178. [Meriçli, 2012] Meriçli, Ç., Veloso, M., and Akin, H. L., "Improving biped walk stability with complementary corrective demonstration", *Auton. Robots*, vol. 32, pp. 419-432, 2012.
179. [Michaud, 2005] Michaud, F., Brosseau, Y., Cote, C., *et al.*, "Modularity and integration in the design of a socially interactive robot", in *IEEE International Workshop on Robot and Human Interactive Communication (ROMAN 2005)*, pp. 172-177, 2005.
180. [Millan, 2004] Millan, J. R., Renkens, F., Mourino, J., *et al.*, "Noninvasive brain-actuated control of a mobile robot by human EEG", *Biomedical Engineering, IEEE Transactions on*, vol. 51, pp. 1026-1033, 2004.
181. [Miller, 2008] Miller, D., Nourbakhsh, I., and Siegwart, R., "Robots for Education", in *Springer Handbook of Robotics*, B. Siciliano and O. Khatib, Eds., ed: Springer Berlin Heidelberg, pp. 1283-1301, 2008.
182. [Mitra, 2007] Mitra, S. and Acharya, T., "Gesture Recognition: A Survey", *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 37, pp. 311-324, 2007.
183. [Miwa, 2004] Miwa, H., Itoh, K., Matsumoto, M., *et al.*, "Effective emotional expressions with expression humanoid robot WE-4RII: integration of humanoid robot hand RCH-1", in *IEEE/RSJ International Conference on Intelligent Robots and Systems, (IROS 2004)*, pp. 2203-2208 vol.3, 2004.
184. [Mohamed, 2012] Mohamed, Z. and Capi, G., "Development of a New Mobile Humanoid Robot for Assisting

- Elderly People", *Procedia Engineering*, vol. 41, pp. 345-351, 2012.
185. [Monicka, 2011] Monicka, G. J., Sekhar, D. N. O. G., and Kumar, R. K., "Performance Evaluation of Membership Functions on Fuzzy Logic Controlled AC Voltage Controller for Speed Control of Induction Motor Drive", *International Journal of Computer Applications*, vol. 13, pp. 8-12, 2011.
186. [Mordoch, 2013] Mordoch, E., Osterreicher, A., Guse, L., *et al.*, "Use of social commitment robots in the care of elderly people with dementia: A literature review", *Maturitas*, vol. 74, pp. 14-20, 2013.
187. [Mouaddib, 2008] Mouaddib, A.-I., "Controlling and Sharing Authority in a Multi-Robot System", in *first conference on Humans Operating Unmanned Systems (HUMOUS'08)*, Brest, France, 2008.
188. [Mu, 2010] Mu, Y. H., "Task-Oriented Architecture for a Humanoid Robot", *Applied Mechanics and Materials*, vol. 40-41, pp. 228-234, 2010.
189. [Murray, 2005] Murray, J., Erwin, H., and Wermter, S., "A Hybrid Architecture Using Cross-Correlation and Recurrent Neural Networks for Acoustic Tracking in Robots", in *Biomimetic Neural Learning for Intelligent Robots*, vol. 3575, S. Wermter, *et al.*, Eds., ed: Springer Berlin Heidelberg, pp. 73-87, 2005.
190. [Mutlu, 2009] Mutlu, B., Shiwa, T., Kanda, T., *et al.*, "Footing in human-robot conversations: how robots might shape participant roles using gaze cues", presented at *the Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*, La Jolla, California, USA, 2009.
191. [Nagai, 2005] Nagai, Y., "Learning to Comprehend Deictic Gestures in Robots and Human Infants", in *the 2005 IEEE International Workshop on Robot*

- and *Human Interactive Communication*, Nashville, USA, pp. 217-222, 2005.
192. [Navarro-Guerrero, 2012] Navarro-Guerrero, N., Weber, C., Schroeter, P., *et al.*, "Real-world reinforcement learning for autonomous humanoid robot docking", *Robotics and Autonomous Systems*, vol. 60, pp. 1400-1407, 2012.
193. [Neto, 2009] Neto, P., Pires, J. N., and Moreira, A. P., "Accelerometer-Based Control of an Industrial Robotic Arm", in *18th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 1192-1197, 2009.
194. [Nieuwenhuisen, 2010] Nieuwenhuisen, M., Stuckler, J., and Behnke, S., "Intuitive multimodal interaction for service robots", in *the 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI 2010)*, pp. 177-178, 2010.
195. [Nof, 1999] Nof, S. Y., *Handbook of Industrial Robotics*: John Wiley, 2 edition, 1999.
196. [Norman, 2009] Norman, D., *The Design of Future Things*, New York, USA: Basic Books, 2009.
197. [Osswald, 2010] Osswald, S., Hornung, A., and Bennewitz, M., "Learning reliable and efficient navigation with a humanoid", in *IEEE International Conference on Robotics and Automation (ICRA 2010)*, pp. 2375-2380, 2010.
198. [Panfir, 2013] Panfir, A. N., **Boboc, R. G.**, and Mogan, G., "NAO Robots Collaboration for Object Manipulation", *Applied Mechanics and Materials*, vol. 332, pp. 218-223, 2013.
199. [Park, 2011] Park, C.-B. and Lee, S.-W., "Real-time 3D pointing gesture recognition for mobile robots with cascade HMM and particle filter", *Image and Vision Computing*, vol. 29, pp. 51-63, 2011.
200. [Patel, 2010] Patel, M., Valls Miro, J., and Dissanayake, G.,

- "Dynamic Bayesian Networks for Learning Interactions between Assistive Robotic Walker and Human Users", in *KI 2010: Advances in Artificial Intelligence*, vol. 6359, R. Dillmann, et al., Eds., ed: Springer Berlin Heidelberg, pp. 333-340, 2010.
201. [Pavlovic, 1997] Pavlovic, V. I., Sharma, R., and Huang, T. S., "Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, pp. 677-695, 1997.
202. [Pernalete, 2007] Pernalete, N., Gottipati, R., Grantner, J., et al., "Integration of an Intelligent Decision Support System and a Robotic Haptic Device for Eye-Hand Coordination Therapy", in *IEEE 10th International Conference on Rehabilitation Robotics (ICORR 2007)*, pp. 283-291, 2007.
203. [Pires, 2002] Pires, G. and Nunes, U., "A Wheelchair Steered through Voice Commands and Assisted by a Reactive Fuzzy-Logic Controller", *Journal of Intelligent and Robotic Systems*, vol. 34, pp. 301-314, 2002.
204. [Pollack, 2005] Pollack, M. E., "Intelligent Technology for an Aging Population: The Use of AI to Assist Elders with Cognitive Impairment", *AI Magazine*, vol. 26, no. 2, pp. 9-24, 2005.
205. [Postelnicu, 2011] Postelnicu, C.-C., Talaba, D., and Toma, M.-I., "Controlling a Robotic Arm by Brainwaves and Eye Movement", in *Technological Innovation for Sustainability*, vol. 349, L. Camarinha-Matos, Ed., ed: Springer Berlin Heidelberg, pp. 157-164, 2011.
206. [Postelnicu, 2012] Postelnicu, C.-C., Gîrbacia, F., and Talabă, D., "EOG-based visual navigation interface development", *Expert Systems with Applications*, vol. 39, pp. 10857-10866, 2012.

## BIBLIOGRAFIE

207. [Pourmehrer, 2013] Pourmehrer, S., Monajjemi, V. M., Vaughan, R. T., *et al.*, "'You two! Take off!': Creating, modifying and commanding groups of robots using face engagement and indirect speech in voice commands", presented at *the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Tokyo, Japan, 2013.
208. [Power, 2002] Power, D., *Decision Support Systems: Concepts and Resources for Managers*: Praeger, 2002.
209. [Power, 2008] Power, D. J., "Understanding Data-Driven Decision Support Systems", *Information Systems Management*, vol. 25, pp. 149-154, 2008.
210. [Quan, 2011] Quan, W., Niwa, H., Ishikawa, N., *et al.*, "Assisted-care robot based on sociological interaction analysis", *Computers in Human Behavior*, vol. 27, pp. 1527-1534, 2011.
211. [Rabiner, 1993] Rabiner, L. and Juang, B.-H., *Fundamentals of speech recognition*: Prentice-Hall, Inc., 1993.
212. [Ramos-Garijo, 2003] Ramos-Garijo, R., Prats, M., Sanz, P. J., *et al.*, "An autonomous assistant robot for book manipulation in a library", in *IEEE International Conference on Systems, Man and Cybernetics*, vol.4, pp. 3912-3917, 2003.
213. [Rascón, 2010] Rascón, C., Avilés, H., and Pineda, L., "Robotic Orientation towards Speaker for Human-Robot Interaction", in *Advances in Artificial Intelligence – IBERAMIA 2010*, vol. 6433, A. Kuri-Morales and G. Simari, Eds., ed: Springer Berlin Heidelberg, pp. 10-19, 2010.
214. [Rautaray, 2012] Rautaray, S. and Agrawal, A., "Vision based hand gesture recognition for human computer interaction: a survey", *Artificial Intelligence Review*, pp. 1-54, 2012.
215. [Richarz, 2007] Richarz, J., Scheidig, A., Martin, C., *et al.*, "A Monocular Pointing Pose Estimator for Gestural Instruction of a Mobile Robot",

- International Journal of Advanced Robotic Systems*, vol. 4, pp. 139-150, 2007.
216. [Roberts, 2002] Roberts, A., *Science Fiction*: Taylor and Francis, 2002.
217. [Rosheim, 2006] Rosheim, M. E., *Leonardo's Lost Robots*. Berlin: Springer, 2006.
218. [Ross, 2010] Ross, T. J., *Fuzzy Logic with Engineering Applications, 3rd Edition*: Wiley 2010.
219. [Rossiter, 2011] Rossiter, J., "Multimodal Intent Recognition for Natural Human-Robotic Interaction", Doctor of Philosophy, School of Electronic, Electrical and Computer Engineering, University of Birmingham, 2011.
220. [Roth, 2001] Roth, W.-M., "Gestures: Their Role in Teaching and Learning", *Review of Educational Research*, vol. 71, pp. 365-392, 2001.
221. [Ryu, 2006] Ryu, H. and Lee, W., "Where you point is where the robot is", presented at *the Proceedings of the 7th ACM SIGCHI New Zealand chapter's international conference on Computer-human interaction: design centered HCI*, Christchurch, New Zealand, 2006.
222. [Sakoe, 1971] Sakoe, H. and Chiba, S., "A dynamic programming approach to continuous speech recognition", in *7th International Congress on Acoustics*, Budapest, Hungary, 1971.
223. [Sakurai, 2005] Sakurai, Y., Yoshikawa, M., and Faloutsos, C., "FTW: fast similarity search under the time warping distance", presented at *the Proceedings of the twenty-fourth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, Baltimore, Maryland, 2005.
224. [Salvador, 2007] Salvador, S. and Chan, P., "Toward accurate dynamic time warping in linear time and space", *Intell. Data Anal.*, vol. 11, pp. 561-580, 2007.



## BIBLIOGRAFIE

225. [Sanna, 2013] Sanna, A., Lamberti, F., Paravati, G., *et al.*, "A Kinect-based natural interface for quadrotor control", *Entertainment Computing*, vol. 4, pp. 179-186, 2013.
226. [Saulnier, 2009] Saulnier, P., Sharlin, E., and Greenberg, S., "Using bio-electrical signals to influence the social behaviours of domesticated robots", presented at *the Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*, La Jolla, California, USA, 2009.
227. [Sauppé, 2014a] , Sauppé, A. and Mutlu, B., "Design patterns for exploring and prototyping human-robot interactions", presented at *the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Toronto, Ontario, Canada, 2014.
228. [Sauppé, 2014b] Sauppé, A. and Mutlu, B., "Robot deictics: how gesture and context shape referential communication", presented at *the Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, Bielefeld, Germany, 2014.
229. [Scassellati, 2003] Scassellati, B., "Investigating models of social development using a humanoid robot", in *Proceedings of the International Joint Conference on Neural Networks*, vol.4, pp. 2704-2709, 2003.
230. [Schiffer, 2012] Schiffer, S., Ferrein, A., and Lakemeyer, G., "Caesar: an intelligent domestic service robot", *Intelligent Service Robotics*, vol. 5, pp. 259-273, 2012.
231. [Scholtz, 2003] Scholtz, J., "Theory and evaluation of human robot interactions," in *Proceedings of the 36th Annual Hawaii International Conference on System Sciences (HICSS03)*, Hawaii, USA, 2003.

232. [Senin, 2008] Senin, P., "Dynamic time warping algorithm review", University of Hawaii at Manoa, Honolulu, USA, 2008.
233. [Shahid, 2014] Shahid, S., Krahmer, E., and Swerts, M., "Child-robot interaction across cultures: How does playing a game with a social robot compare to playing a game alone or with a friend?", *Computers in Human Behavior*, vol. 40, pp. 86-100, 2014.
234. [Shamsuddina, 2012] Shamsuddina, S., Yussof, H., Ismail, L. I., *et al.*, "Humanoid Robot NAO Interacting with Autistic Children of Moderately Impaired Intelligence to Augment Communication Skills", *International Symposium on Robotics and Intelligent Sensors 2012 (IRIS 2012)*, vol. 41, pp. 1533-1538, 2012.
235. [Shi, 2013] Shi, S., *Emgu CV Essentials*: Birmingham Packt Publishing, 2013.
236. [Shuyin, 2004] Shuyin, L., Kleinehagenbrock, M., Fritsch, J., *et al.*, "'BIRON, let me show you something': evaluating the interaction with a robot companion", in *IEEE International Conference on Systems, Man and Cybernetics*, vol.3, pp. 2827-2834, 2004.
237. [Sidner, 2005] Sidner, C. L., Lee, C., Kidd, C. D., *et al.*, "Explorations in engagement for humans and robots", *Artificial Intelligence*, vol. 166, pp. 140-164, 2005.
238. [Siegwart, 2003a] Siegwart, R., Arras, K. O., Bouabdallah, S., *et al.*, "Robox at Expo.02: A large-scale installation of personal robots", *Robotics and Autonomous Systems*, vol. 42, pp. 203-222, 2003.
239. [Siegwart, 2003b] Siegwart, R., Arras, K. O., Jensen, B., *et al.*, "Design, Implementation and Exploitation of a New Fully Autonomous Tour Guide Robot", in *ASER '03 1st International Workshop*

- on *Advances in Service Robotics*, Bardolino, Italy, 2003.
240. [Siegwart, 2011] Siegwart, R., Nourbakhsh, I., and Scaramuzza, D., *Introduction to Autonomous Mobile Robots*, Cambridge, Massachusetts: MIT Press, 2011.
241. [Stanton, 2012] Stanton, C., Bogdanovych, A., and Ratanasena, E., "Teleoperation of a humanoid robot using full-body motion capture, example movements, and machine learning", in *Australasian Conference on Robotics and Automation*, Wellington, New Zealand, 2012.
242. [Starețu, 2007] Starețu, I., "Actualități și perspective în robotica serviciilor", *Buletinul AGIR*, vol. 4, pp. 43-52, 2007.
243. [Staudte, 2009] Staudte, M. and Crocker, M. W., "Visual attention in spoken human-robot interaction", presented at *the Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*, La Jolla, California, USA, 2009.
244. [Stiefelhagen, 2004] Stiefelhagen, R., Fugen, C., Gieselmann, R., *et al.*, "Natural human-robot interaction using speech, head pose and gestures", in *Intelligent Robots and Systems (IROS 2004)*, vol.3, pp. 2422-2427, 2004.
245. [Sugiyama, 2007] Stowers, J., Hayes, M., and Bainbridge-Smith, A., "Altitude control of a quadrotor helicopter using depth map from Microsoft Kinect sensor", in *IEEE International Conference on Mechatronics (ICM 2011)*, pp. 358-362, 2011.
246. [Su, 2014] Su, C.-J., Chiang, C.-Y., and Huang, J.-Y., "Kinect-enabled home-based rehabilitation system using Dynamic Time Warping and fuzzy logic", *Applied Soft Computing*, vol. 22, pp. 652-666, 2014.
247. [Sugiyama, Sugiyama, O., Kanda, T., Imai, M., *et al.*,

- 2007] "Natural deictic communication with humanoid robots", in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2007)*, pp. 1441-1448, 2007.
248. [Suk, 2010] Suk, H.-I., Sin, B.-K., and Lee, S.-W., "Hand gesture recognition based on dynamic Bayesian network framework", *Pattern Recognition*, vol. 43, pp. 3059-3072, 2010.
249. [Surmann, 2008] Surmann, H., Holz, D., Blumental, S., *et al.*, "Teleoperated Visual Inspection and Surveillance with Unmanned Ground and Aerial Vehicles", *IJOE*, vol. 4, pp. 26-38, 2008.
250. [Susperregi, 2013] Susperregi, L., Arruti, A., Jauregi, E., *et al.*, "Fusing multiple image transformations and a thermal sensor with kinect to improve person detection ability", *Engineering Applications of Artificial Intelligence*, vol. 26, pp. 1980-1991, 2013.
251. [Takahashi, 2010] Takahashi, M., Suzuki, T., Shitamoto, H., *et al.*, "Developing a mobile robot for transport applications in the hospital domain", *Robotics and Autonomous Systems*, vol. 58, pp. 889-899, 2010.
252. [Takeda, 2006] Takeda, R., Yamamoto, S., Komatani, K., *et al.*, "Missing-Feature based Speech Recognition for Two Simultaneous Speech Signals Separated by ICA with a pair of Humanoid Ears", in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 878-885, 2006.
253. [Tanaka, 2006] Tanaka, F., Movellan, J. R., Fortenberry, B., *et al.*, "Daily HRI evaluation at a classroom environment: reports from dance interaction experiments", presented at *the Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, Salt Lake City, Utah, USA, 2006.

## BIBLIOGRAFIE

254. [Tansel 2013] Tansel İç, Y., Yurdakul, M., and Dengiz, B., "Development of a decision support system for robot selection", *Robotics and Computer-Integrated Manufacturing*, vol. 29, pp. 142-157, 2013.
255. [Tapus, 2007a] Tapus, A., Mataric, M. J., and Scassellati, B., "The grand challenges in socially assistive robotics", *IEEE Robotics And Automation Magazine*, vol. 14, pp. 35-42, 2007.
256. [Tavares, 2013] Tavares, T., Medeiros, A., de Castro, R., *et al.*, "The Use of Natural Interaction to Enrich the User Experience in Telemedicine Systems", in *HCI International 2013 - Posters' Extended Abstracts*, vol. 374, C. Stephanidis, Ed., ed: Springer Berlin Heidelberg, pp. 220-224, 2013.
257. [Taylor, 2004] Taylor, G. and Wray, R. E., "Behavior Design Patterns: Engineering Human Behavior Models", presented at *the Behavioral Representation in Modeling and Simulation Conference*, Arlington, VA, 2004.
258. [Thavasi, 2012] Thavasi, P. T. and Suriyakala, C. D., "Sensors and Tracking Methods Used in Wireless Sensor Network Based Unmanned Search and Rescue System -A Review", *Procedia Engineering*, vol. 38, pp. 1935-1945, 2012.
259. [Toma, 2012a] Toma, M.-I. and Datcu, D., "Determining Car Driver Interaction Intent through Analysis of Behavior Patterns", in *Technological Innovation for Value Creation*, pp. 113-120, 2012.
260. [Toma, 2012b] Toma, M.-I., Rothkrantz, L. J. M., and Antonya, C., "Car driver skills assessment based on driving postures recognition", presented at *the 3rd IEEE CogInfoCom*, Kosice, Slovakia, 2012.
261. [Topoleanu, 2011] Topoleanu, T.-S. and Mogan, G., "Automatic Speech Recognition: An Improved Paradigm", in *Technological Innovation for*

- Sustainability*, vol. 349, L. Camarinha-Matos, Ed., ed: Springer Berlin Heidelberg, pp. 269-276, 2011.
262. [Trafton, 2008] Trafton, J. G., Bugajska, M. D., Fransen, B. R., *et al.*, "Integrating vision and audition within a cognitive architecture to track conversations", presented at *the Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction*, Amsterdam, The Netherlands, 2008.
263. [Tur, 2011] Tur, G. and Mori, R. D., *Spoken Language Understanding: Systems for Extracting Semantic Information from Speech*: Wiley, 2011.
264. [Turk, 2014] Turk, M., "Multimodal interaction: A review", *Pattern Recognition Letters*, vol. 36, pp. 189-195, 2014.
265. [Valin, 2003] Valin, J. M., Michaud, F., Rouat, J., *et al.*, "Robust sound source localization using a microphone array on a mobile robot", in *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003)*, vol.2, pp. 1228-1233, 2003.
266. [Valli, 2005] Valli, A., Notes on Natural Interaction, 2005.
267. [Valli, 2008] Valli, A., "The design of natural interaction", *Multimedia Tools Appl.*, vol. 38, pp. 295-305, 2008.
268. [Van der Loos, 2008] Van der Loos, H. F. M. and Reinkensmeyer, D., "Rehabilitation and Health Care Robotics", in *Springer Handbook of Robotics*, B. Siciliano and O. Khatib, Eds., ed: Springer Berlin Heidelberg, pp. 1223-1251, 2008.
269. [van Osch, 2014] van Osch, M., Bera, D., van Hee, K., *et al.*, "Tele-operated service robots: ROSE", *Automation in Construction*, vol. 39, pp. 152-160, 2014.
270. [Varchavskaia, 2001] Varchavskaia, P., Fitzpatrick, P., and Breazeal, C., "Characterizing and processing

- robot-directed speech", in *IEEE/RAS international conference on humanoid robots*, ed. Tokyo, Japan, 2001.
271. [Wada, 2006] Wada, K., Shibata, T., Sakamoto, K., *et al.*, "Long-term Interaction between Seal Robots and Elderly People — Robot Assisted Activity at a Health Service Facility for the Aged", in *Proceedings of the 3rd International Symposium on Autonomous Minirobots for Research and Edutainment (AMiRE 2005)*, K. Murase, *et al.*, Eds., ed: Springer Berlin Heidelberg, pp. 325-330, 2006.
272. [Wagner, 2014] Wagner, P., Malisz, Z., and Kopp, S., "Gesture and speech in interaction: An overview", *Speech Communication*, vol. 57, pp. 209-232, 2014.
273. [Wang, 1991] Wang, M.-J. J., "A decision support system for robot selection", *Decision Support Systems*, vol. 7, pp. 273-283, 1991.
274. [Wang, 2013] Wang, X., Mueen, A., Ding, H., *et al.*, "Experimental comparison of representation methods and distance measures for time series data", *Data Mining and Knowledge Discovery*, vol. 26, pp. 275-309, 2013.
275. [Webster, 2014] Webster, D. and Celik, O., "Systematic review of Kinect applications in elderly care and stroke rehabilitation", *Journal of NeuroEngineering and Rehabilitation*, vol. 11, pp. 1-24, 2014.
276. [Wen, 2014] Wen, R., Tay, W.-L., Nguyen, B. P., *et al.*, "Hand gesture guided robot-assisted surgery based on a direct augmented reality interface", *Computer Methods and Programs in Biomedicine*, vol. 116, pp. 68-80, 2014.
277. [Wigdor, 2011] Wigdor, D. and Wixon, D., "Chapter 2 - The Natural User Interface", in *Brave NUI World*, D. Wigdor and D. Wixon, Eds., ed Boston:

- Morgan Kaufmann, pp. 9-14, 2011.
278. [Willems, 2007] Willems, R. M. and Hagoort, P., "Neural evidence for the interplay between language, gesture, and action: A review", *Brain and Language*, vol. 101, pp. 278-289, 2007.
279. [Wittgenstein, 1965] Wittgenstein, L., *The Blue and Brown Books (Preliminary Studies for the Philosophical Investigations)* New York: Harper Torchbooks, 1965.
280. [Wu, 2012] Wu, Y.-H., Fassert, C., and Rigaud, A.-S., "Designing robots for the elderly: Appearance issue and beyond", *Archives of Gerontology and Geriatrics*, vol. 54, pp. 121-126, 2012.
281. [Xu, 2012] Xu, J., Wei, C., Wang, C., *et al.*, "An Approach to Navigation for the Humanoid Robot NAO in Domestic Environments", in *Autonomous Robots and Multirobot Systems (ARMS) Workshop Affiliated with AAMAS*, Valencia, Spain, 2012.
282. [Xu, 2013] Xu, W. and Lee, E.-J., "A New NUI Method for Hand Tracking and Gesture recognition Based on User Experience", *International Journal of Security and Its Applications*, vol. 7, 2013.
283. [Yamamoto, 2005] Yamamoto, S., Valin, J. M., Nakadai, K., *et al.*, "Enhanced Robot Speech Recognition Based on Microphone Array Source Separation and Missing Feature Theory", in *Proceedings of the 2005 IEEE International Conference on Robotics and Automation( ICRA 2005)*, pp. 1477-1482, 2005.
284. [Yanco, 2002] Yanco, H. A. and Drury, J. L., "A Taxonomy for Human-Robot Interaction", in *AAAI Fall Symposium on Human-Robot Interaction*, Falmouth, Massachusetts, pp. 111-119, 2002.
285. [Yanco, 2004] Yanco, H. A. and Drury, J., "Classifying human-robot interaction: an updated



- taxonomy," in *IEEE International Conference on Systems, Man and Cybernetics*, vol.3, pp. 2841-2846, 2004.
286. [Yi, 1998] Yi, B. K., Jagadish, H. V., and Faloutsos, C., "Efficient retrieval of similar time sequences under time warping", in *Proceedings of the 14th International Conference on Data Engineering*, Orlando, Florida, USA, pp. 201-208, 1998.
287. [Yilu, 2007] Yilu, Z. and Juyang, W., "Task Transfer by a Developmental Robot", *IEEE Transactions on Evolutionary Computation*, vol. 11, pp. 226-248, 2007.
288. [Yu, 2011] Yu, D., Yu, X., Hu, Q., *et al.*, "Dynamic time warping constraint learning for large margin nearest neighbor classification", *Information Sciences*, vol. 181, pp. 2787-2796, 2011.
289. [Zadeh, 1965] Zadeh, L. A., "Fuzzy sets", *Information and Control*, vol. 8, pp. 338-353, 1965.
290. [Zaharia, 2010] Zaharia, T., Segarceanu, S., Cotescu, M., *et al.*, "Quantized Dynamic Time Warping (DTW) algorithm", in *the 8th International Conference on Communications (COMM 2010)*, pp. 91-94, 2010.
291. [Zhang, 2011] Zhang, D. and Gao, Z., "Hybrid head mechanism of the groundhog-like mine rescue robot", *Robotics and Computer-Integrated Manufacturing*, vol. 27, pp. 460-470, 2011.
292. [Zhou, 2011] Zhou, M. and Wong, M. H., "Boundary-based lower-bound functions for dynamic time warping and their indexing", *Inf. Sci.*, vol. 181, pp. 4175-4196, 2011.
293. [Zhu, 2009] Zhu, C. and Sheng, W., "Online Hand Gesture Recognition Using Neural Network Based Segmentation", in *the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, St. Louis, USA, 2009.

294. [Zhu, 2004] Zhu, Z. and Ji, Q., "Eye and gaze tracking for interactive graphic display", *Mach. Vision Appl.*, vol. 15, pp. 139-148, 2004.