

스크립트 기반의 로봇 제스처 자동생성 방법 및 집사로봇에의 적용

Automatic Generation of Script-Based Robot Gesture and its Application to Steward Robot

김현희, Heon-Hui Kim*, 이형욱, Hyong-Euk Lee**, 김용휘, Yong-Hwi Kim**,
박광현, Kwang-Hyun Park**, 변증남, Zeungnam Bien**

*인간친화 복지 로봇시스템 연구센터, **한국과학기술원 전자전산학과

요약 본 논문은 인간과 로봇간의 효과적인 상호작용을 위한 로봇제스처의 자동생성 기법을 다룬다. 이는 텍스트 정보만의 입력으로 의미 있는 단어에 대응되는 특정 제스처패턴이 자동적으로 생성되도록 하는 기법으로서 이를 위한 사전조사로 제스처가 출현하는 발화시점에서의 단어수집이 우선적으로 요구되었다. 본 논문은 이러한 분석을 위해 두 개 이상의 연속된 제스처 패턴을 효과적으로 표현할 수 있는 제스처 모델을 제안한다. 또한 제안된 모델이 적용되어 구축된 제스처 DB 와 스크립트 기법을 이용한 로봇제스처 자동생성 방법을 제안한다. 제스처 생성시스템은 규칙기반의 제스처 선택부와 스크립트 기반의 동작 계획부로 구성되고, 집사로봇의 안내기능에 대한 모의실험을 통해 그 효용성을 확인한다.

핵심어: Script-based, Robot Gesture Generation, TTG(Text-To-Gesture), HMI(Human-Machine Interaction)

1. 서론

최근 등장하고 있는 휴머노이드 로봇(Humanoid robots)이나 안드로이드 로봇(Android robots) 등은 인간과 유사한 팔이나 머리동작으로 음성출력과 동시에 자연스러운 제스처를 취함으로써 산업용로봇이나 청소로봇 등의 작업지향형 로봇과는 인간친화적인 관점에서 차별화된 존재로 여겨지고 있다. 로봇이 인간의 동작을 흉내 내도록 하기 위한 연구로는 연속된 동작에 대한 주요 자세(Key-postures)가 그려진 사진 등으로부터 각각의 관절 각들의 경유점들을 얻어낸 후 이를 추종하도록 하는 기법[1], 모션캡처(Motion Capture)등을 이용하여 연속된 인간의 동작에 대한 기하학적 정보를 순차적으로 기록한 뒤 이를 로봇에 복제하는 기법[2,3] 등을 들 수 있다. 이러한 제스처 생성기법은 인간의 자연스러운 동작구현 측면에서는 효과적일 수 있으나, 인간과 로봇간의 상호작용을 수행할 때 예비된 시나리오가 변경될 경우 이에 대한 동작을 재 생성하거나 재 계획해야 하는 번거로움이 따른다.

본 논문은 음성인식 및 출력기능을 가진 휴머노이드 형태의 로봇을 위한 제스처 자동생성 기법을 다룬다. 로봇제스처 자동생성이란 인간과 로봇이 대화 등의 상호작용을 수행할 때 발화-제스처(Speech-Gesture)간의 상관관계 등을 이용하여 음성출력과 함께 로봇의 팔이나 머리동작, 감정표현 등을 자동적으로 생성한다는 개념이다. 로봇의 음성은 문자정보를 발음으로 연결시켜주는 TTS(Text-To-Speech)기술의 발달로 텍스트 입력에 의해 비교적 쉽게 생성 가능하다[4].

유사한 개념을 적용하면, 음성에 동기화된 로봇제스처의 자동생성을 위해서는 로봇에의 적용에 앞서 텍스트를 제스처로 변환할 수 있는 TTG(Text-To-Gesture)기법에 관한 연구가 필요하다.

텍스트-제스처 변환에 관한 연구로는 컴퓨터 애니메이션을 통한 가상 캐릭터의 립싱크(Lip-sync) 및 얼굴표정의 생성에 관한 연구[5], 입력되는 텍스트 정보를 수화 애니메이션으로 변환하는 수화생성 시스템에 관한 연구[6] 등을 들 수 있다. 텍스트를 제스처로 변환하는 작업은 립싱크나 수화를 생성하는 것보다 까다롭다. 입술은 기호를 발음으로 연결시키는 주요 발음기관으로 텍스트와 밀접한 상관관계가 있으며 수화는 교본에 따라 정해진 규칙이 있는 반면 인간의 제스처는 개인의 성향, 감정상태, 청자의 지적 수준 등에 많은 영향을 받기 때문이다. 그렇지만 어떤 단어는 제스처와 밀접한 관련이 있는 경우가 있다. 예를 들어 “오른쪽” 등의 방향을 나타내거나 “아니오” 등의 응답은 각 단어와 제스처 간의 상관성이 높다고 볼 수 있다. 따라서 텍스트-제스처 변환시스템을 구현하기 위해서는 이러한 사례분석이 우선적으로 선행되어야 한다.

본 논문은 제스처 동작을 보다 효과적으로 분석하기 위한 제스처 모델을 제안한다. 또한 로봇에의 적용을 위한 스크립트 기반의 제스처 자동생성 시스템을 제안한다. 제스처 모델은 제스처 패턴의 분석을 위해 이용되고 분석결과로부터 단어-제스처 DB를 구성한다. 로봇 제스처는 입력되는 텍스트에 대하여 형태소 분석, 규칙기반 제스처선택, 스크립트기반

동작계획 과정을 통해 생성된다. 제안된 방법은 집사로봇에 의 적용을 위한 스크립트파일 생성결과를 통해 유효성을 확인한다.

본 논문의 구성은 다음과 같다. 먼저 2장에서는 제스처의 종류를 살펴보고 발화-제스처간의 상관성 분석에 대해 기술한다. 제 3장에서는 로봇제스처 자동생성 시스템에 대해 기술한다. 제 4장은 집사로봇에의 적용방안을 설명하고 마지막으로 제 5장에서 결론을 맺는다.

2. 제스처의 종류 및 분석

2.1 제스처의 역할 및 종류

제스처는 주로 몸의 상체 앞 공간에서 상징적인 표현을 위해 손과 팔을 움직이는 행위를 말한다[7]. 제스처는 대화도중이거나 혹은 말이 전혀 사용되지 않을 때 말을 대신 할 수 있으며 상호작용의 진행과 리듬을 조절하는 역할을 한다. 또한 발화시 주의력을 유지시킨다거나 말에 대한 강조 또는 명확성을 더해주기도 한다. 그뿐만 아니라 말의 내용에 특징을 주어 상대방이 이를 잊지 않도록 해주며 그 다음에 해야 할 말의 예보 역할을 하기도 한다[8]. 이러한 제스처는 각각의 역할이나 방법에 따라 분류가 가능하며 그림 1은 이러한 제스처의 종류 및 특징을 나타내고 있다.

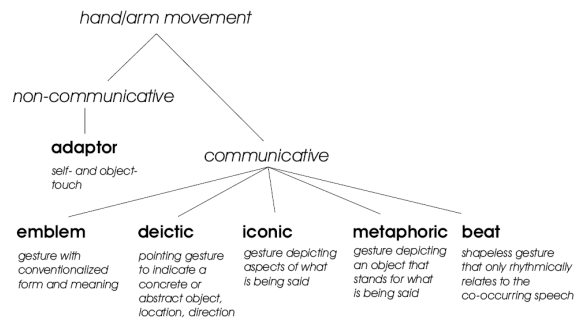


그림 1. 제스처의 종류 및 특징[8]

어댑터(Adaptor)라 불리는 제스처는 비 의사소통용 제스처로서 의사소통과는 무관하게 자신 혹은 특정 객체에 접촉하는 행위를 의미하는데, 펜으로 귀를 후비는 등의 행위가 이에 해당된다. 의사소통을 위한 제스처로는 엠블렘(Emblem), 지시적(Deictic), 사상적(Iconic), 은유적(Metaphoric), 비트(Beat)의 5가지로 분류할 수 있다[9]. 엠블렘은 발화와 독립적으로 사용될 수 있는 관습적인 형식이나 의미를 나타내는 제스처로서 곧바로 단어로 번역될 수 있는 특징이 있다. 지시적 제스처는 실재하는 사람 또는 객체의 위치, 방향을 지시하는 행위를 나타내며 여기에는 추상적이거나 보이지 않은 존재를 지시하는 행위도 포함된다. 사상적 제스처는 말하고자 하는 대상의 특징을 묘사하여 이를 동작으로 표현하는 제스처이다. 예를 들어 TV를 말할 때 두 손으로 허공에 사각형을 그려 보이는 행위를 사상적 제스처라 할 수 있다. 은유적 제스처는 사상적 제스처와 같이 그 특징을 직접 묘사할 수 없는 경우 제 3의 객체를 도입하여 은유적인 표현을 통해 말하고자 하는 바를 표현하는 행위를 말한다. 은유적 제스처의 예로 “잘 진행되고 있다”의 의미로 한 손을 들어 검지손가락을 편 채로 원을 그리며 손목을 돌리는 행위를 들 수 있다.

마지막으로 비트 제스처는 화자의 의도나 특정단어를 강조하기 위해 종종 사용되는 제스처로서 손을 반복적으로 두드리는 등의 행위를 나타낸다.

2.2 제스처 모델 및 분석전략

로봇제스처를 자동생성하기 위해서는 우선 발화와 제스처간의 관계를 기술할 수 있어야 한다. 문헌[8]에 의하면 제스처의 6가지 종류에 대한 각각의 패턴 별 의미, 관련단어 등이 잘 정리되어 있는데 이는 충분히 본 연구를 위한 응용가치가 높다. 그러나 제스처 표현은 인류 보편적인 표현이 있는 반면 대부분의 경우 민족이나 언어 문화권 별로 달리 표현되는 경우가 많다[10]. 따라서 한국인의 정서에 맞는 제스처의 패턴을 추출하고 각 패턴과 의미를 관계 짓는 것이 한국인에 맞는 제스처를 가장 잘 표현할 수 있을 것이다. 그러나 국내의 발화-제스처 관계가 잘 정리된 문헌은 찾아보기가 쉽지 않다. 따라서 본 논문에서는 일반적인 화제를 다루는 TV토크쇼를 통해 발화와 제스처 관계를 분석하고자 한다. McNeill은 제스처가 취해지는 형상을 상세히 기술할 수 있는 제스처 코딩 전략(Gesture coding scheme)을 제안하고 이를 이용하여 제스처의 유형을 분류하는 연구를 수행하였다[11]. 그러나 이 표현 방법은 단일 생성되는 제스처의 상세한 표현은 가능하나 두 개 내지 세 개로 표현되는 연속된 제스처 표현은 어려움이 있다. 따라서 본 논문은 기본제스처, 부가제스처, 최종제스처의 3단계로 표현 가능한 제스처 모델을 제안한다. 제안된 제스처 모델은 분석이 쉽고 TV토크쇼를 통해 수집된 총 86가지의 발화-제스처 표본 중 “홀러가다-물결모양제스처” 등을 제외한 84개 표본의 제스처 표현이 가능하였다. 또한 이 제스처 모델은 로봇에의 응용 시 동작에 따라 순차적으로 분류되어 있어 구현이 용이한 장점이 있다. 그림 2는 본 논문에서 제안한 제스처 모델을 나타낸다. 제스처 모델의 첫 번째 단계는 기본제스처(BG)를 기술하는 것으로 어떤 손이 사용되어(H) 어떤 형상의 팔 동작을 취하는지(Gs, Gt)를 나타낸다.

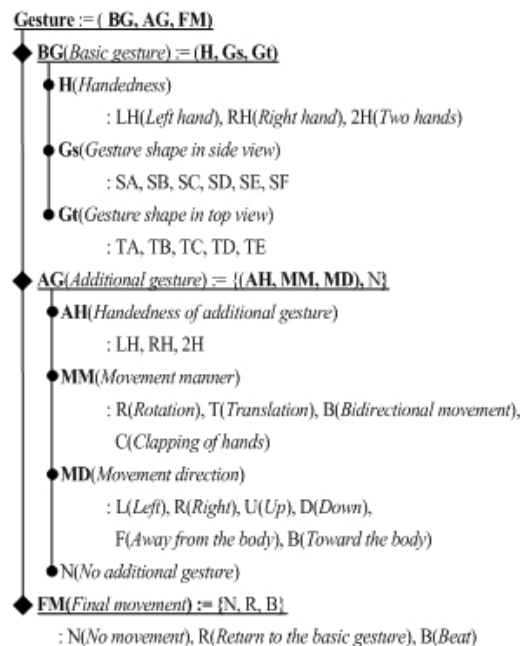


그림 2. 제안된 제스처 모델의 구조

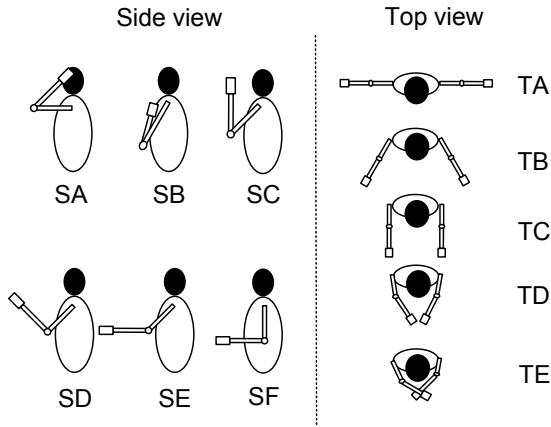


그림 3. 기본 제스처의 구분방법

여기서 Gs 와 Gt는 각각 측면과 윗면에서 바라본 형상들의 집합을 나타낸다. 각각의 제스처 패턴은 그림 3과 같다. 여기서 SA는 팔이 머리를 지지하고 SB는 가슴에 손을 대는 동작을 나타내며 TE의 경우에는 손을 교차시키는 행위를 나타낸다. 최종적인 형상은 두 가지 각도에서 바라본 동작의 조합결과로 얻어지게 된다. AG는 기본제스처 후 나타나는 부가 제스처로 여기에는 절대적인 위치보다는 상대적인 움직임을 나타낸다. 즉, 기본위치에서 어떤 손을, 어떤 방법으로, 어느 방향으로 동작하는지를 기술한다. 마지막 항목인 FM은 두 번째 제스처가 취해진 후 나타나는 최종동작을 의미한다. 일 예로 “보다 높은 하늘”을 말하면서 제스처를 취하면 두 손을 (2H) 이용하여 가슴위치에 손을 뻗는(SD, TC) 기본동작이 먼저 취해지며 이어서 두 손(2H)을 위쪽방향(U)으로 이(T)동을 하고 그 상태를 유지(N)한다. 이를 기호화 하면 ((2H, SD, TC), (2H, T, U), N))으로 표현할 수 있다.

2.3 발화-제스처 분석결과

제안된 제스처 모델을 이용하여 사회자 2명과 여자 게스트 1명이 출연하는 TV토크쇼(25분 분량)에서 게스트 1명에 대한 총 86장면의 발화-제스처 데이터를 수집하였고 제스처 모델로 표현이 가능한 84표본에 대한 통계적 분석을 시행하였다. 표 1은 수집된 데이터의 패턴 별 분류결과를 나타낸다. 이 결과에 의하면 피험자는 SD 와 TC 제스처(손을 가슴 앞으로 뻗는 제스처)를 가장 많이 취했으며 이 때 왼손보다는

표 1. TV토크쇼의 제스처 패턴 별 분류결과

항목	패턴 별 발생빈도
H	RH(44%), 2H(38%), LH(18%) ,
BG	Gs SD(49%), SF(18%), SE(15%), SC(10%), SB(7%), SA(1%)
	Gt TC(58%), TB(20%), TA(6%), TD(6%), TE(2%)
AG	AH RH(56%), 2H(34%), LH(9%)
	MM T(72%), R(13%), B(9%), C(3%)
FM	MD R(49%), L(19%), F(16%), D(9%), U(3%), B(1%)
	N(49%), B(38%), R(13%)

오른손이나 양손을 선호하여 제스처를 취하는 경향이 있음을 알 수 있다. 또한 부가제스처 없이 기본제스처만을 취하는 경우는 62%에 달했고 한 동작 이외에 동적으로 제스처가 취해지는 경우는 38%로 조사되었다. 표 1에서 AH, MM, MD, FM항목은 추가적인 제스처가 있을 경우의 조건부 확률에 대한 백분율을 나타낸다.

수집된 제스처-발화 표본을 발생빈도수 별로 분류한 결과 46개의 제스처 패턴을 얻을 수 있었다. 여기서 1회씩 나타나는 제스처 패턴을 제외하여 22개의 제스처 패턴을 추출하였다. 표 2는 각 패턴 별로 발생빈도수가 많은 순으로 분류한 후 추출한 상위 4개의 제스처 패턴과 이와 관련된 단어들을 나타내고 있다. 여기서 ((H, SD, TC), AG, N)는 이차적인 제스처가 발생하는 모든 경우에 대한 패턴을 나타낸다. 이는 전체 제스처 발생의 20%에 달하며 발화예제를 보면 “한번도”, “각자”, “빨리”, “아무리”, “절대로”, “왜” 등의 뒷단어를 강조하는 부사 단어가 눈에 띈다. 이 제스처는 비트 제스처, 손목을 돌리는 등의 이차적인 제스처가 발생하는 경향이 많았다. 또한 이 제스처 패턴에서 부가제스처 없는 ((H, SD, TC), N, N)패턴에서도 강조의 기능을 하는 단어가 출현하는 경우가 많았는데, 이는 SD와 TC가 조합된 제스처 패턴은 강조의 미를 가지는 단어와 밀접한 연관성이 있음을 의미한다. ((H, SF, TC), N, N)의 패턴은 주로 한 손을 이용하여 자신의 몸 앞으로 손목을 약간 굴리며 팔을 가볍게 뻗고 이차적인 제스처가 나타나지 않은 행위에 해당된다. 이 패턴에서는 문맥상 “~하는”의 행위를 암시하는 단어가 다수 나타났고 “그것을”, “다른 것” 등의 추상적인 대상을 지시하는 경향이 있었다. 또한 가슴에 손을 대는 SB패턴은 자신을 치장하거나 심정의 상태 등을 의미하는 단어들과 연관성이 높았다. 즉 이러한 결과는 제스처 공간 내에 존재하는 패턴들은 발화 중에 사용되는 특정 단어의 의미나 기능 등과 밀접한 관련이 있음을 보여주고 있다.

표 2. 제스처 패턴 별 발화예제

제스처 패턴	관련 단어 예
((H, SD, TC), AG, FM) (20%)	이것 좀 먹어라 하고, 한번도, 각자 나름대로, 애가 빨리 또 기운을, 근데 원망 안하덴데, 아무리 열심히 하셔도, 100 퍼센트 다 가야, 가정 외에 다른 일을, 절대로 할 수가 없어요, 왜 너 이 시간애...
((H, SF, TC), N, N) (17%)	애들이 편식하고 안먹어서, 항상 그렇게 좀 푸싱을 하는, 왜냐하면 그것을 받아주고, 그날 뒤도 안돌아보고, 일도하고 , 학생들 가르치고 , 어떤 다른거, 만나 는 인연들도...
((H, SD, TC), N, N) (9%)	그런데, 예를 들어서 남편에게, 100 퍼센트 가 있다면, 적극적, 찬성, 지지, 이런 거 해서 싫어할, 옛날에는 뭐 100 일 1주년 5주년
((H, SB, G), AG, N) (7%)	나, 얼마로서 느낌을, 저한테 서운해, 이해를 해줘야, 서로가 상대를 읽어가기가, 준비 를 해야

3. 로봇 제스처 자동생성 시스템

로봇제스처 자동생성 시스템은 텍스트정보를 입력 받아 의미 있는 단어를 찾아 이에 관련된 제스처를 선택하여 음성 출력과 함께 로봇의 팔과 머리 등의 동작제어를 수행토록 하는 시스템이다. 그림 3은 시스템의 전체 구성을 나타내는 것으로 시스템은 크게 제스처 DB를 활용한 규칙기반의 제스처 선택부와 스크립트 기반의 동작 계획부로 구분된다.

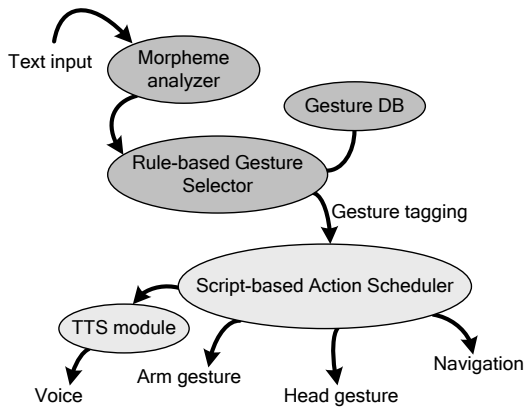


그림 4. 로봇 제스처 자동생성 시스템의 구성

제스처 생성부는 동작 제어부에 입력되는 스크립트 문서에 제스처코드를 삽입하는 모듈이다. 제스처 코드는 사전에 수집한 제스처 DB를 통해 단어의 앞에 삽입된다. 제스처 DB는 개념-단어-제스처의 구조로 구성되어 있으며 여기에 사용된 단어들은 2장에서 논의된 분석결과로부터 얻어졌다. 제스처 패턴에 대응되는 각각의 단어들을 일차적으로 분류하였고, 이어 이와 유사한 단어들을 추가하여 제스처-단어 DB를 구축하였다. 분석결과로부터 얻어진 의미 있는 제스처의 총 패턴 수는 22가지로, 여기에 사용되는 단어는 수집된 84개의 단어와 강조, 행위, 양, 지시 등의 의미분류를 통해 얻어진 10개의 개념 집합에 16개의 단어를 추가하여 총 100개의 단어가 DB에 수록되었다. 표 3은 제스처-단어 DB구성을 간략히 나타내고 있다.

표 3. 개념 별 단어-제스처 DB의 구성

개념	단어구성	제스처 패턴
강조	긍정 빨리, 각자	((RH, SD, TC), N, N)
	부정 한번도, 절대로	((RH,SD,TC), (RH,T,L), B)
행위	~하고, ~하는	((LH, SD, TC), N, N)
양	긍정 큰, 넓게	((2H, SD, TB), N, N)
	부정 작게, 좁은	((2H, SD, TD), N, N)
지시	오른쪽	((RH, SE, TB), N, N)

또한 한국어의 경우 어미 변화 등의 형태변화가 많으므로 이에 따른 데이터 손실을 줄이며 단어의 확장성을 위해 형태소 분석기를 이용하였다. 예를 들어 단어 제스처 DB에 사용되는 “넓게”와 유사한 단어가 입력될 경우 이들의 형태소 분석결과와 DB에 포함된 단어의 고유 의미의 차이가 같으면 이에 해당하는 제스처패턴을 선택하도록 한다. 다음은 “넓게”의 유사단어들과 형태소 분석 결과를 나타낸다. 여기서 “AJ”, “EM”, “SY”는 형용사, 어말어미, 부호를 각각 나타낸다[12].

넓고	넓/(AJ) + 고/(EM)
넓은	넓/(AJ) + 은/(EM)
넓다	넓/(AJ) + 다/(EM) + ./(SY)
넓어	넓/(AJ) + 어/(EM)

동작 계획부에서는 로봇이 보다 자연스러운 동작을 할 수 있도록 동작 시나리오에 따라 적절한 시점에서 팔, 머리, 이동 등의 동작명령을 생성한다. 동작제어 부에는 이러한 동작명령이 음성출력에 동기화 될 수 있도록 텍스트파일 형태의 입력방식이 이용되었다. 입력되는 텍스트 파일의 형식은 구분자 등의 정해진 규칙을 적용하여 음성출력을 위한 텍스트와 동작명령을 구분하였고, 각각 동작명령에 관련된 태그(Tag)들이 기록된다. 동작제어 부에 내장된 파서(Parser)는 입력된 스크립트 문서를 특정 이벤트 또는 시간에 따라 액션 테이블(Action table)을 작성하는 기능을 수행하며, 액션테이블에 기록된 제어명령들이 음성과 함께 순차적으로 출력된다.

4. 모의실험

4.1 대상 로봇시스템

집사로봇[13]은 노약자 및 장애인을 대상으로 개발된 서비스 로봇으로서, 사용자와 함께 거주하며 기본생활을 돕는 물리적 보조역할을 수행하거나 대화 등을 통한 정신적 보조역할을 수행한다. 또한 주거공간에는 사용자의 편의를 위한 보조로봇시스템과 다양한 가전기들이 위치해 있는데, 이들의 종류가 많을 뿐 아니라 그 조작방법이 쉽지 않은 경우가 대부분이다. 집사로봇은 이러한 장치와 사용자간의 매개체의 역할을 수행하는데, 사용자의 간단한 명령만을 입력 받은 로봇은 복잡하고 다양한 장치들을 순차적으로 작동시키는 역할을 수행한다. 로봇은 음성인식 및 TTS기능을 내장하고 있어 사용자의 명령을 지시 받을 수 있으며, 음성기능이 학습시스템, 감정모델 등과 연계되어 있어 경우에 따라 사용자와의 인터페이스 및 정보전달의 수단으로 이용된다. 집사로봇의 학습시스템은 주위의 온도, 습도, 조도 등의 환경정보와 선호하는 TV채널, 온도조절 상태 등의 사용자의 생활패턴을 지속적으로 관측하고 학습한다. 또한 이러한 학습결과를 통해 서비스를 수행할 때 사용자의 의도분석에 따른 능동적인 작업을 결정한다. 감정모델은 학습시스템을 통해 얻어진 결과를 토대로 눈, 머리, 팔 동작 등에 의해 감정의 표현을 결정한다.

집사로봇은 머리, 몸통, 팔로 구성된 서비스로봇 시스템이다. 머리 부는 팬/틸트의 2자유도의 구동부를 가지고 있어, 사용자를 응시하거나 “예”, “아니오” 등의 머리동작을 통해 사용자와의 친화적인 상호작용이 가능하도록 설계되었다. 머리 상단 부에는 초음파신호의 수신부가 장착되어 있다. 이는 로봇의 활동공간에 장착된 10개의 초음파 송신부로부터 주기적인 신호를 받고, 이를 거리로 환산한 후 삼각측량법에 의해 로봇의 위치추정 용도로 활용된다. 눈 부에는 LCD디스플레이가 장착되어 있으며 로봇의 감정모델을 통해 9가지의 감정 중 하나의 감정을 눈동자 애니메이션으로 보여준다. 몸통 부는 바퀴구동형 이동장치 및 초음파센서가 장착되어 있다. 로봇은 위치인식 기능과 활동공간에 대한 환경지도도를 내장하고 있으며, 장애물 감지 및 회피기능과 자체 경로계획기능을 통해 스스로 지정된 곳으로의 이동기능을 수행한다. 가슴 부에는 특정 객체를 인식하거나 추적기능을 가진 스테레오 비전 시스템이 장착되었다. 또한 로봇은 각각 6개의 구동 자유도와 그리퍼로 구성되어 있는 두 조의 팔을 가지고 있으며, 이를 통해 로봇은 특정 위치의 물건을 집거나 사용자와의 상호작용을 위한 제스처를 취하는 기능을 수행한다. 집사로봇의

하드웨어 시스템은 모듈화로 구성되어 중앙제어장치와 CAN (Control Area Network)을 통해 데이터를 송수신 할 수 있도록 설계되어 있다.

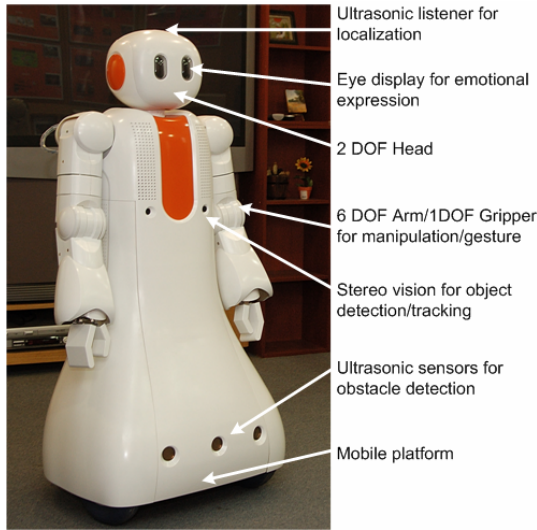


그림 5. 집사로봇의 하드웨어 구성

4.2 실험결과 및 검토

집사로봇의 기능 중에 주거공간에 위치한 다양한 로봇시스템을 소개하는 시나리오기반의 동작시퀀스가 예비되어 있다. 모의실험은 안내대본의 일부를 제스처 자동생성 시스템에 입력한 후 중간과정인 제스처 선택 부에서 제스처코드가 텍스트에 어떻게 삽입되는지를 알아보는 것으로 하였다. 제스처코드는 구현 상의 간략화를 목적으로 2장에서 설명한 제스처 표현법에서 괄호와 심표가 생략되어 표기된다. 또한 제스처가 없다는 의미인 'N'역시 생략되었는데, 예를 들어 ((LH, SC, TC), N, N)을 제스처 코드로 바꾸면 <LHSTC>로 표기된다. 로봇의 동작 계획부는 입력되는 스크립트파일 정보를 로봇의 각 하드웨어에 순차적인 제어명령을 분배하는 역할을 하며, 이는 기록된 태그정보와 일대 일 매핑관계가 성립하므로 이 과정은 생략하기로 한다. 다음은 테스트를 위한 주거공간 안내시나리오의 첫 부분을 나타낸다.

“안녕하세요. 지능형 주거공간에 오신 것을 환영합니다. 저는 이곳을 관리하는 집사로봇 조이입니다. 그럼 지금부터 안내를 시작하겠습니다.”



그림 6. 제스처 선택부의 출력파일

제스처 선택기에 안내스크립트를 입력한 결과 그림 6과 같은 출력을 얻었다. 출력파일을 확인한 결과 문장 중에 “~하다”.

“저는”, “이곳을”, “지금부터”의 단어에서 4가지 제스처 패턴이 각각 생성되었다. <RHSFTC>는 오른손을 가볍게 앞으로 뻗는 제스처 동작을 나타내며 <RHSB>는 오른손을 가슴 가까이 위치시키는 동작을 나타낸다. <RHSETCRHTR>은 오른손을 앞으로 뻗은 상태에서 다시 오른손 끝이 오른쪽으로 이동하는 동작을 의미하고 <RHSDTC>는 손을 가슴위치로 올리는 동작을 나타낸다. 모의실험 결과 제스처 패턴이 모두 오른손을 이용하는 데 이는 DB를 구축할 때 사용된 제스처 패턴이 오른손을 사용하는 사용자를 대상으로 얻어졌으며 그 결과 분류과정서 왼손을 사용한 “행위” 패턴은 22개의 제스처 패턴에서 제외되어 있었기 때문이다.

5. 결론

인간-로봇간의 상호작용 시 로봇의 제스처동작은 보다 명확한 의미전달의 역할을 하면서 인간에게 심리적으로 친근한 모습으로 다가설 수 있는 매개체가 될 수 있다. 본 논문에서는 이러한 개념을 로봇에 적용하기 위한 방법으로 텍스트 입력에 의한 자동 제스처생성방법을 다루고 있으며 구체적으로

첫째, 두 개 이상의 연속된 제스처를 효과적으로 표현할 수 있는 제스처 모델

둘째, 텍스트-제스처 변환 기능을 가진 스크립트 기반의 로봇제스처 자동생성 방법

을 제안하였다. 이에 제안된 제스처 모델이 84/86의 제스처 패턴을 표현할 수 있었고, 기존의 방법에 비해 연속된 제스처 동작의 분석이 가능하였다. 또 제안된 제스처 자동생성 방법은 집사로봇의 안내대본을 입력한 결과 의미 있는 단어에 대한 제스처 코드가 생성됨을 확인하였다. 그러나 제스처 DB가 한 명의 피험자로부터 얻어졌기 때문에 제스처 표현이 보편적이지 못하며 제스처 DB에 수록된 단어의 개수가 적어 표현상의 한계가 있었다. 따라서 추후의 연구로 이러한 한계를 보완하고 단어-제스처간의 상관관계의 학습법 등 보다 신뢰성 있는 제스처 생성에 관한 연구를 수행할 계획이다. 본 논문에서 제안한 기법은 구체적으로 인간과 로봇간의 의사소통 시 보다 친화적이며 효과적인 상호작용에 기여할 수 있을 것이라 기대된다.

감사의 글

본 연구는 과학기술부/한국과학재단 우수연구센터육성사업의 지원으로 수행되었음 (R11-1999-008)

참고문헌

[1] K. Noritake, S. Kato, T. Yamakita and H. Itoh, A motion generation system for humanoid robots - Tai Chi motion, Proceedings of 2003 International Symposium on Micromechatronics and Human Science, pp.265-269, 2003.

[2] D. Matsui, T. Minato, K. F. MacDorman and H. Ishiguro, Generating natural motion in an android by mapping human motion, 2005 IEEE/RSJ

- International Conference on Intelligent Robots and Systems, pp.3301-3308, 2005.
- [3] M. Ruchanurucks, S. Nakaoka, S. Kudoh and K. Ikeuchi, Humanoid robot motion generation with sequential physical constraints, Proceedings 2006 IEEE International Conference on Robotics and Automation, pp.2649-2654, 2006.
- [4] <http://www.voiceware.co.kr/>
- [5] H. H. S. Ip and C. S. Chan, Script-based facial gesture and speech animation using a NURBS based face model, Computers & graphics, v.20 no.6, pp.881-891, 1996.
- [6] 오영준, 박광현, 장효영, 김대진, 정진우, 변증남, 하이퍼 수화문장을 사용한 수화 생성 시스템, 제 25 회 한국정보처리학회 춘계학술발표대회 논문집, 서울, 제 13 권, 제 1 호, pp. 621-624, 2006.
- [7] D. McNeill, Speech and gesture integration: The nature and functions of gesture in children's communication, Jossey-Bass Publishers, San Francisco, pp. 11-27, 1998.
- [8] 김영순, 임지룡, 몸짓 의사소통적 한국어 교수법 모형, 이중언어학, pp. 1-24, 2002.
- [9] M. Kipp, Gesture Generation By Imitation: From Human Behavior To Computer Character Animation, Dissertation.com, Boca Raton, FL, USA, pp 31-48, 2004.
- [10] 성광수, 한국어에 있어서 몸짓과 발화의 관계, 이중언어학, 제 19 호, pp. 301-319, 2001.
- [11] D. McNeill, Hand and mind: What gestures reveal about thought. Chicago: University of Chicago Press, 1992.
- [12] <http://cs.sungshin.ac.kr/~shim/demo/tagset.htm>
- [13] H. E. Lee, Y. M. Kim, K. H. Park and Z. Z. Bien, Development of a Steward Robot for Intelligent Sweet Home, International Journal of Human-friendly Welfare Robotic Systems, vol. 6, no. 4, pp. 57-64, December 2005.