

지능형 로봇의 시각인식기술과 활용

정연구, 이재연, 소 정, 김계경, 조영조 (한국전자통신연구원 지능형로봇연구단)

1. 서론

로봇의 시각인식기술은 궁극적으로는 사람의 시각을 모델링하여 사람의 눈과 유사하게 정지 이미지 또는 비디오 이미지의 내용물을 인식하는 기술이다. 인식기술의 목적은 대상물의 형상을 인식하고, 사용자가 누구인가를 인식하고, 주위 상황을 인식하며, 목표 지점으로 안전하고 신속하게 이동하기 위함이다. 로봇 자신의 인식 결정을 필요로 하는 경우는 인간과의 상호작용을 하거나, 주어진 일을 달성하기 위하여 현장 상황을 이해하여 처리하는 일이 있을 수가 있다. 이러한 여러 종류의 일 중에서 현재 로봇이 할 수 있는 일은 극히 단순한 일만이 허용된다. 지능으로 말하면, 시각인식 기술의 경우에 유치원학생 정도의 수준이 될 것이다. 그러나, 제한적이며 적은 범위의 목적에 사용할 때에는 인간에게 상당히 유익하게 사용될 수가 있다.

로봇은 인간과 다르게, 감정에 지배를 받지 않고, 시간이 경과하여도, 피로하지 않고, 주어진 일을 일관성 있게 무한 시간 할 수가

있는 것이 장점이다. 그러나 단점으로는 주위의 상황이 변하거나, 대상물체의 형상에 변화가 있을 경우에 대처하는 능력이 현저하게 떨어진다. 주위 조명의 변화와 물체의 위치와 촬영각도의 변화에 대하여 적응력이 매우 낮은 것이 약점이다. 물체 중에서도 사람이 만든 조형 물이나 가공 형상 물은 형상이 일정하여 인식하기가 좋으나, 동물이나 사람은 계속해서 움직이기 때문에 형상이 약간씩 변화하여서 인식하기가 어렵다.

인간과 같이 높은 수준의 시각적 인식기능을 사용하는 로봇은 현재는 없고, 제한적인 상황 하에서 학습된 대상물을 인식하여 활용하고 있는 낮은 수준이라고 할 수가 있다. 산업용 로봇의 경우에는 눈이 없이도 고정된 위치에서 주어진 움직임 경로를 이용하여 작업을 하고 있는 경우가 많다. 그러나, 옥외의 일이나, 작업환경의 변화 등에 대처할 수 있는 일을 하기 위해서는 궁극적으로 사람과 같은 시각기능을 갖는 지능형 로봇이 필요하다.

지능형 로봇, 휴머노이드 로봇, IT 서비스 로봇 등은 사람의 기능을 모방하고, 사람과

유사한 모양을 갖게 하며, 기존의 공장에서 사용하는 딱딱한 기계와 같은 모습의 로봇이 아니고, 감정을 표현하며, 사람처럼 인식 능력을 갖게 하여 인간에게 편리하고, 친화적인 모습을 지향하고 있다. 이러한 형태의 로봇은 현재는 초보 수준이지만, 인간의 삶의 질을 높일 수가 있고, 일상생활에 활용이 될 수 있는 취지에서 기대가 된다. 이러한 로봇의 기술발전을 통하여 부가적으로 발생하는 기술의 효과는 많은 곳에서 자동화 또는 반자동화를 통하여 인간을 돕는 일에 사용되고, 인간의 삶의 질이 놀라울 정도로 향상되고 편리하여 질 것으로 보인다.

본 논문에서는 이와 같이 지능형 로봇의 시각인식 기술 중에서 당 연구실에서 개발하고 있는 사용자인식 기술과 문자인식에 대하여 살펴본다.

II. 사용자 인식기술

사용자 인식이란 말 그대로 사용자가 누구인지를 시스템이 알아보게 하는 기술을 총칭한다. 우리가 컴퓨터에 로그인할 때 ID와 패스워드를 입력함으로써 그 컴퓨터를 사용할 수 있는 권한을 획득하거나, 집에 들어갈 때 열쇠로 문을 열고 들어가거나, 아니면 최근에 점차 그 사용범위를 넓혀가고 있는 지문, 얼굴, 홍채 등의 특징을 이용한 본인확인(생체인식 기술에 의한 사용자 인식) 등이 모두 사용자 인식이다^{11,12}.

1. 로봇용 사용자 인식을 위한 요구사항 분석

로봇용 사용자 인식기술의 요구사항을 도

출하기 위하여, 기존의 사용자 인식기술과의 차이점을 다음의 세가지 측면에서 논의해 보고자 한다.

첫째로 기존의 사용자 인식이 일회적인데 반하여, 로봇에서의 사용자인식은 연속성이 있다. 로봇 환경에서는 로봇과 상호작용을 하는 사용자가 연속적으로 바뀔 수 있다. 방금 전 아버지가 시킨 일을 수행하던 로봇이 다음 순간 어머니에게 서비스를 제공해야 할 상황이 발생할 수 있고, 그 사이에는 로그아웃이나 로그인과 같은 단계가 존재하지 않는다. 심지어는 로봇이 인지해야 할 사용자가 동시에 두 사람 이상이 될 경우도 있을 수 있다.

두 번째 차이점은 기존의 사용자 인식은 엄격하게 규제된 환경과 사용자의 적극적인 협조를 가정할 수 있는 반면, 로봇에서는 자유롭게 움직이는 사용자를 대상으로 해야 한다는 점이다. 로봇에서는 사용자를 계속적으로 인지해야 하는 상황이 발생한다. 이러한 상황에서 로봇이 자기를 잘 알아 볼 수 있도록 협조한다는 가정은 비현실적이다. 따라서 로봇은 사용자에 관한 풍부한 정보를 얻을 수 있기는 하나, 로봇이 그 사용자를 인식하기에 적합한 형태의 데이터로 들어오는 것은 보장할 수는 없는 것이다.

셋째로, 기존의 사용자 인식은 일정한 위치에 설치되어 있는 컴퓨터나 출입문을 사용자가 찾아가는 형태로 이루어 지게 되지만, 로봇 환경에서는 로봇과 인식대상이 되는 사람이 모두 움직이고 있는 상황이라는 차이점을 가지고 있다. 이러한 환경에서는 로봇은 단지 그 사람이 누구인지를 인식하기만 하면 되는

것이 아니라, 그 사람은 어디에 있는가, 혹은 그 사람을 찾으려면 어느쪽으로 가야 하는가와 같은 질문에도 응답할 수 있어야 한다. 결국 로봇은 사용자의 신분과 위치 정보를 파악하는 기능을 갖출 수 있어야 하는 것이다.

이상에서 보인 바와 같이 로봇 환경에서의 사용자 인식은 보안의 관점에서 주로 연구되어 온 기존의 사용자 인식과는 요구사항 자체가 전혀 다르다는 점에서 새로이 설계되어야 할 것으로 생각된다.

기존의 사용자 인식 기술 중, 패스워드 혹은 카드와 같이 기억내용이나 소유물에 의존하는 사용자 인식 기술은 로봇 환경에서는 사용이 곤란하다. 일단 불편할 뿐 아니라 위에서 기술한 바와 같이 계속적으로 사용자가 바뀌어 가는 상황에서 그때마다 이러한 입력을 받는다는 것은 무리이기 때문이다.

2. 생체인식기술의 활용

생체인식기술이란 지문, 얼굴, 홍채, 목소리, 서명과 같은 인간의 신체적/행동적 특징에 기반 하여 개인이 누구인가를 인식하는 방법을 말한다⁶⁾. 로봇에서의 사용자 인식을 위한 기술로 가장 유망한 것은 사용자의 신체적인 특징을 이용하는 생체인식 방법이라고 볼 수 있다. 실제로 기존의 로봇 사용자 인식 시스템 중에는 얼굴인식기술을 이용하는 경우가 주를 이루고 있다. 여러 가지 생체 특징 중에서도 얼굴은 사용자와의 직접적인 접촉을 필요로 하지 않으며, 사용자에게 불편함을 주지 않고, 비교적 먼 거리에서도 인식이 가능하다는 장점을 가지고 있다.

그러나 이미 널리 알려진 바와 같이 얼굴인

식은 조명 및 자세의 변화에 취약하다는 약점을 가지고 있다. 즉 로봇이 얼굴인식기술을 이용하여 사용자 인식을 수행하기 위해서는 사용자의 정면얼굴이 로봇의 눈(카메라)에 정확하게 잡혀야 할 뿐 아니라, 그때의 조명상황도 등록 시와 유사하도록 조절될 수 있어야 한다는 뜻이다. 이는 로봇과 사용자가 각기 자유롭게 움직이고 있는 상황에서는 매우 까다로운 조건으로 사용자의 협조를 강제하지 않는 한은 만족시키기 어려운 조건이다.

이에 필자 등이 개발하고 있는 로봇용 사용자인식 시스템에서는 얼굴인식이 적용 가능한 상황에서는 얼굴인식 기술을 적용하되, 그것만으로 로봇환경에서 요구되는 사용자 인식기능을 구현할 수 없다는 인식 하에, 사용자와 관련된 다른 종류의 신체적인 특징들(이와 같은 특징을 Dynamic-Biometrics이라 부르기로 한다)을 적극적으로 활용한다. 또한 단발적인 인식에 그치지 않고 인식된 사용자를 추적/관리 하는 기능을 추가함으로써 앞서 언급한 요구조건을 충족시킬 수 있는 새로운 형태의 사용자 인식시스템 개발을 수행하고 있다. 여기서는 현재 진행 중인 연구내용의 소개를 통하여 향후 지능형 로봇이 갖추어야 할 사용자 인식 시스템에 대해 고찰 해 보기로 한다.

3. 동적생체인식 (Dynamic-Biometrics)

사용자의 신체특징을 안정적으로 획득하는 것은 단순하지 않다. 특히 로봇환경에서 로봇이 사용자를 인식할 필요가 있을 때 마다 지문을 찍으라고 하는 것은 비현실적이다. 심지어는 로봇에서 많이 이용되고 있는 얼굴

과 같은 경우도 이미 언급한 바와 같이 인식에 적절한 얼굴영상을 획득할 수 있는 기회란 그리 많지 않다. 옆 모습이나 심지어는 뒷 모습 밖에 보이지 않는 경우도 적지 않을 것이며, 너무 멀리 떨어져 있어 얼굴영상이 인식에 적당한 크기로 잡히지 않는 경우도 흔히 있을 것이다. 이와 같은 상황을 극복하기 위해서는 보다 다양한 상황에서 인식이 가능한 다양한 특징들을 사용자 인식 프레임워크에 포함시켜야 할 필요가 있다.

생체인식기술은 그 전제로 형상의 특징이 불변을 필요로 하지만, 현실적인 관점에서 보면, 전혀 변화되지 않는 특징은 없다. 본문에서 제안하는 동적생체인식이란 생체인식과 마찬가지로 사용자의 신분과 연계할 수 있는 특징을 말하지만, 다만 정상적인 생체인식 특징의 엄밀한 조건과는 달리 특징한 사용자 그룹 내에서와 일정한 시간 내에서만 불변한 특징을 말한다.

예를 들면, 사람의 키는 동적생체인식 특징이 된다. 어린이의 경우에 시간이 지남에 따라서 키가 자라므로 키는 불변한 특징량이라 볼 수도 없다. 성인도 아침에 측정한 키와 저녁때 측정한 키 사이에 차이가 있다는 것은 잘 알려진 사실이다. 따라서 키는 일반적인 의미에서의 생체인식 특징이 될 수 없는 것이다. 동적생체인식 특징의 다른 예로 옷 색깔을 들 수가 있다. 옷을 갈아입기 때문에 사용자가 입고있는 옷 색깔은 당연히 불변한 특징이 아니다. 이러한 동적생체인식은 그 특성상 대상이 되는 사용자 그룹에 의존적일 수 밖에 없다.

예를 들면, 어떤 가족에서는 아버지와 아들의 키가 거의 같은 경우도 있다. 이런 경우

키로는 아버지와 아들을 구별할 수 없고, 다른 특징을 찾아 구별하는 수 밖에 없다. 또 다른 예로 머리 색깔을 생각해 보자. 다른 식구들은 모두 검은색이지만, 중학생인 딸 만은 노란색으로 염색을 했다면, 머리 색깔이란 특징은 다른 식구들을 구별하는 데는 아무런 의미가 없지만, 염색한 그 딸인지 여부를 판정하는 데는 매우 유용한 특징이 될 수 있다. 이와 같이 동적생체인식은 사용자 그룹의 특성에 따라 적응적(Adaptive)으로 사용될 수 있어야 한다.

생체인식 특징에 비하여 제한조건이 현저히 완화된 만큼 동적생체인식의 특징은 매우 다양하다. 키, 옷 색깔, 머리 모양은 물론 몸무게, 체형 등이 모두 동적생체인식의 특징이 될 수 있다. 로봇 환경에서 이러한 특징들을 필요로 하는 이유는 이들을 측정할 수 있는 환경이 매우 다양하다는 점 때문이다. 키를 재기 위해서는 뒷모습이라도 상관이 없지만, 전신이 시야에 들어와야 하고, 옷 색깔이란 특징은 적어도 상체가 보일 때 사용이 가능하다. 머리 색깔은 머리부분이 카메라에 잡힌 상황에서 비로소 적용이 가능하다. 하지만, 이러한 여러 가지 조건을 결합하면, 로봇이 사용자와 접하는 다양한 상황에서 그 사용자를 인식할 수 있는 가능성을 향상시킬 수 있게 된다.

4. 사용자 추적 피드백 기술

사용자 추적 피드백 기술은 인간 친화적이고 편의성을 위하여 필요한 기술이다. 기존의 생체인식은 사람이 센서 앞으로 다가 가서, 요구 되는 자세를 취한다. 그러나, 로봇

의 경우에는 로봇과 사람이 서로 움직일 수가 있다. 즉, 카메라와 대상물체가 연속적으로 움직이는 상태에서 사용자를 인식해야 하는 것이 필수 사항이다. 이를 가능하게 하는 기술이 사용자 추적 피드백 기술이다.

기존의 사용자인식은 일반적으로 short-term memory를 가지지 않는다. 방금 전에 그 사람을 인식했다. 해도, 인식모듈은 그 결과를 기억하고 있지 않으며, 명령을 받으면 다시 인식을 위한 전체 프로세스를 반복하게 된다. 예를 들어, 얼굴인식 모듈이 한 사람을 성공적으로 인식한 후, 다음 순간에 사용자가 돌아서면, 얼굴인식 모듈은 이번에는 얼굴이 보이지 않으므로 당연히 인식에 실패하게 된다.

사람이라면 당연히 가지고 있는 이러한 능력을 로봇이 가지게 하기 위해 도입되는 것이 사용자 추적 피드백의 개념이다. 즉 단지 인식을 하는 단계에서 끝나는 것이 아니라, 그 인식결과를 기억하고, 이어지는 영상에서 그 사람을 연속적으로 추적함으로써 로봇은 사용자에게 대한 보다 풍부한 정보를 얻을 수 있으며, 동시에 로봇이 관리하는 사용자에게 관한 정보도 갱신할 수 있게 되는 것이다.

또한 사용자 추적 기술을 통하여 로봇은 사용자의 위치에 관한 정보를 연속적으로 갱신할 수 있다. 물론 로봇이 계속적으로 사용자를 추적할 수는 없지만, 적어도 추적을 통하여 얻어진 근사적인 위치 정보는 나중에 로봇이 그 사용자를 찾아야 할 상황이 발생했을 때 도움이 될 수 있다.

5. 로봇 사용자 인식 시스템

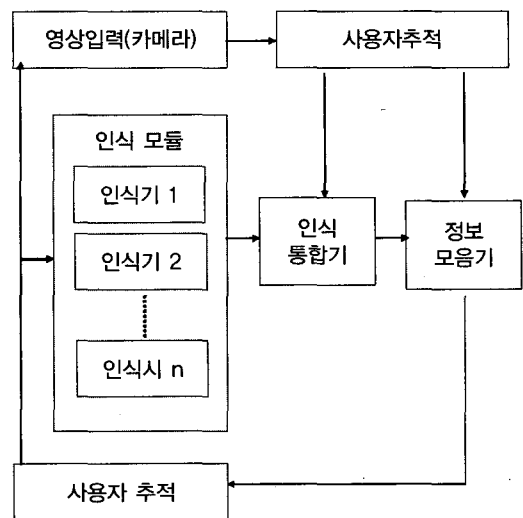
이 장에서는 현재 ETRI의 지능형 로봇 연

구단에서 개발하고 있는 로봇용 사용자 인식 시스템에 대한 기술을 통하여 향후 로봇이 갖추어야 할 사용자 인식 시스템의 구성 및 조건 등에 관하여 고찰해 보기로 한다.

(그림 1)은 로봇의 사용자 인식 시스템의 구성도이다. 그림에서 보는 바와 같이 이 구성도는 영상데이터의 입력이 되면, 그에 대해 결과를 출력하는 일반적인 전방향 처리형(feed forward type)이 아니라 폐쇄회로형(closed loop type)을 형성하고 있다.

이렇게 폐쇄회로형에 의하여 얻어진 정보가 피드백 됨으로써 사용자 인식시스템의 사용자들에 대한 지식을 지속적으로 갱신시켜 간다는 점을 보여주고 있다.

(그림 1)의 사용자 모델 관리자는 인식결과 및 획득된 정보를 저장하고 관리하는 역할을 수행한다. 사용자 모델 관리자가 관리하는 정보에는 장기간 지식(Long-term Knowledge)과 단기간 지식(Short-term knowledge)이 있는데, 장기간 지식은 명시적인 등록과정에 의



(그림 1) 지능형 로봇 사용자 인식 시스템의 구성도

하여 기록되게 되며, 단기간 지식은 자동적인 정보수집 처리에 의하여 갱신된다.

시스템의 동작 중에는 입력영상이 인식모듈과 사용자 추적 모듈로 전달된다. 사용자 인식 시스템은 앞서 언급된 바와 같이 복수의 인식기를 가지고 있어서 각 인식기는 주어진 입력영상에 대하여 판정을 내리는 데 유용한 정보인가를 판단한다. 예를 들어 두 눈이 정상적으로 검출되는 상황이라면 얼굴 인식 모듈이 작동하여 인식결과를 인식통합기로 전달한다. 두 눈의 검출이 불가능한 경우에는 얼굴 인식기는 입력 영상을 무시하고, 널(NULL) 값을 발생시킨다. 인식이 가능한 상황에서도 각각의 인식기는 그 사용자가 누구라는 최종적인 결론을 출력하는 것이 아니라, 그룹 내의 각 사용자에 대한 확률 리스트를 제공함으로써 다른 인식기의 결과와 통합하는 과정을 거치게 된다.

이와 같은 인식결과와 더불어, 통합 인식기는 사용자 추적기로부터 정보를 제공받게 된다. 시스템이 추적 모드에 있는 경우, 사용자 추적기는 입력된 영상에서도 추적이 여전히 유효한지 (즉, 앞서 추적하던 대상물을 놓치지 않았는지)를 검사한 후, 이전 루프에서 얻어진 사용자에 관한 확률 정보를 통합함으로써 사용자에 대한 확신도를 강화시키게 된다.

이와 같이 인식결과가 얻어지고 나면, 정보 모음기가 가동되어 주어진 입력영상으로부터 얻을 수 있는 새로운 정보를 수집하여 사용자 모델 관리자의 지식 정보를 갱신한다.

III. 다중 생체인식기술

사용자 인증을 위한 수단으로 위의 생체인

식과 함께 다중 생체인식 기술이 정확하고 다양한 인증 수단으로 사용된다. 이러한 기술들은 로봇의 다이나믹한 주위환경을 위하여 사용될 수가 있다. 멀티모달 생체인식 시스템은 여러 가지의 다른 형태로 구현될 수 있으나, 다음과 같이 다중센서와 다중 생체인식기술로 분류가 될 수가 있다^[4]:

- 다중 센서(multiple sensors): 하나의 생체 특징을 여러 개의 다른 방식의 센서로 획득한다. (예: 지문 획득을 위해 광학, 초음파, 반도체 방식 센서들을 사용)
- 다중 생체특징(multiple biometrics): 다수의 서로 다른 생체특징을 사용한다. (예: 얼굴과 지문을 함께 사용하여 생체인식을 수행)

1. 다중 센서

다중 센서 시스템은 여러 개의 센서가 필요하면서도 동일한 생체 특징을 사용하기 때문에 일반적으로 소요 비용에 대비하여 큰 성능 향상을 이루기 어렵다. 이 때문에 개념적으로는 가능하지만 실제 다수의 생체특징 센서를 적용한 잘 알려진 사례를 찾아볼 수 없었다. 가장 일반적인 적용 가능성은 지문 획득을 위해 광학식, 반도체식 등의 다중 센서를 사용하는 경우와 얼굴 획득을 위해 광학 영상(일반적인 영상)과 열상(thermogram)을 같이 사용하는 경우이다.

2. 다중 생체특징

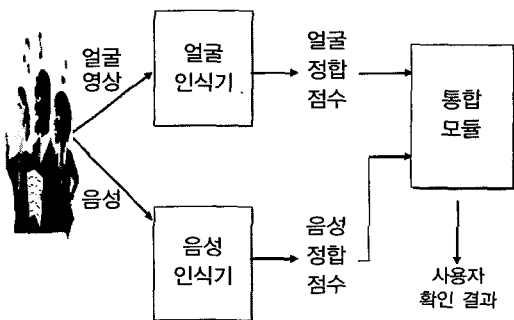
다중 생체특징 시스템은 현재 멀티모달 생체인식 연구에서 가장 자주 다루어지는 경우

이며, 일반적으로 멀티모달 생체인식이라고 하면 이 종류의 멀티모달 생체인식 시나리오를 의미하기도 한다. 이와 같은 시스템의 실제 연구개발 결과는 최근 몇 년 사이에 매우 많이 발표되었고, 그 중 대부분은 지문, 얼굴, 화자(speaker) 인식 등 오래 전부터 사용되어 온 생체특징들의 조합을 사용한다. 여기에서는 사용된 생체특징의 조합별로 구분하여 연구사례를 소개한다.

가. 얼굴 + 음성

얼굴 영상을 이용한 얼굴 인식과 음성 신호에 의한 화자 인식의 결합은 가장 일찍 시도되었고 현재까지도 가장 많이 사용되는 다중 생체특징 시스템의 구성 방식이다. 이 생체특징의 조합은 (그림 2)에 나타난 것처럼 비디오(얼굴)와 오디오(음성)의 결합, 또는 시간 측면에서 정적인 정보(얼굴)와 동적인 정보(음성)의 결합이라는 점에서 상호 보완성이 뛰어난 것으로 인식되고 있으며, 그에 상응하는 많은 개발 사례를 찾아볼 수 있다.

얼굴과 음성을 동시에 이용한 시스템 중 최초의 시도로 인정되는 것은 Brunelli와 Falavigna⁶⁾의 화자 인식과 얼굴 인식 모듈을



(그림 2) 얼굴과 음성을 이용한 다중 생체특징 시스템의 구조⁶⁾

통합하는 시스템이다. 이 시스템은 두 개의 화자 분류기(classifier)와 세 개의 얼굴 분류기를 두 가지 서로 다른 방법으로 통합하였다.

Kittler 등⁷⁾은 정면 얼굴, 얼굴 프로파일(profile), 음성 인식을 통합하는 방법을 발표하였다. 각 분류기의 사후 확률(a posteriori probability)을 통합하는 방법으로는 합, 곱, 최대값, 중앙값, 최소값을 제안하였다. 분류기 통합 실험 결과 동일에러율(equal error rate)은 정면 얼굴, 얼굴 프로파일, 음성 순으로 각각 12.2%, 8.5%, 1.4%였고, 통합 결과 동일 에러율은 합, 곱, 최대값, 중앙값, 최소값 순으로 각각 0.7%, 1.4%, 12.2%, 1.2%, 4.5%를 나타내었다.

Verlinde 등⁸⁾은 얼굴 프로파일, 정면 얼굴, 음성의 세 가지 생체특징을 사용하였다. 분류 결과의 통합은 여러 가지의 일반적으로 사용되는 방식을 실험하였다. 사용된 방식들은 maximum likelihood, quadratic classifier, linear classifier, AND-voting, OR-voting, k-nearest neighbor, binary decision tree 등으로서 각각의 통합 결과를 비교하였다.

나. 얼굴 + 음성 + 입술 움직임

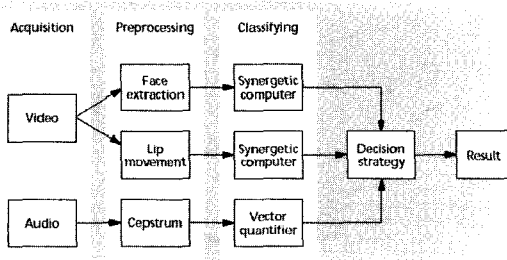
Dieckmann 등⁹⁾은 얼굴 정지 영상, 입술 움직임 비디오 시퀀스, 음성을 통합하여 동적인 정보와 정적인 정보를 함께 이용하는 SESAM(독일어: Synergetische Erkennung mittels Standbild, Aktusik und Motorik)이라는 시스템을 개발하였다. 분류 결과의 통합은 세 개 중 두 개가 일치하는 것을 택하는 2-from-3 방식과 모두 일치하는 것을 택하는 3-from-3 방식을 사용하였고, 모든 통합 방식의 결과가 단일 특징(얼굴, 입술 움직임, 음성)에 의한

결과보다 우수하다는 것을 보였다. 실험은 독일의 Fraunhofer Institute for Integrated Circuits의 출입문에서 수행하였다.

Frischholz와 Dieckmann^[10]은 (그림 3)과 같은 기능을 가지는 BioID라는 시스템에서 정면 얼굴, 입술 움직임, 음성을 함께 사용하였다. BioID는 시스템 관리자가 개별 인식에 의한 결정 결과를 통합하는 방식을 선택할 수 있도록 하였다. 가능한 통합 방식은 합, 2 out of 3, 3 out of 3가 있고, 합의 경우 각 인식 모듈의 가중치와 합의 임계치를 설정할 수 있게 하였다.

다. 음성 + 입술 움직임

Jourlin 등^[11]은 입술 움직임과 음성을 이용한 화자 인식 방법을 개발하였다. 입술 움직임은 active shape model에 의한 입술 추적으로 검출하고 음성 인식은 HMM을 이용하였다. 각 생체특징을 이용한 분류 결과의 통합은 weighted sum을 이용하였다.



(그림 3) BioID의 주요 기능^[10]

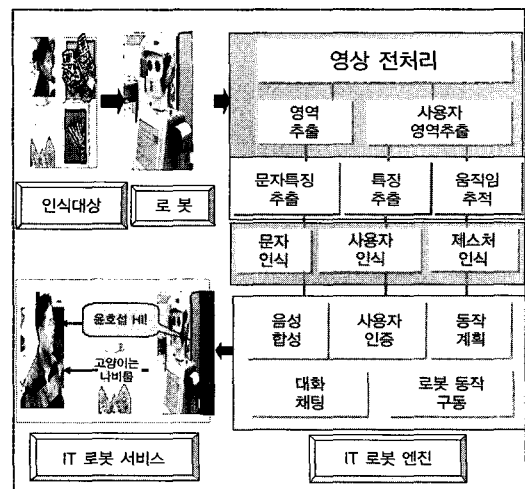
IV. 사용자 인식기술의 활용 사례

사용자 인식기술은 당 연구팀의 지능형 로봇에 활용되어 문자인식 기술, 제스처 인식 기술과 함께 사용되었다.

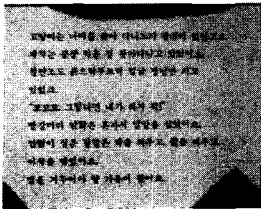
입력된 영상은 특징정보를 이용하여 사용자가 인식되고, 사용자에게 벽에 붙어 있는 책의 문장을 영상 전처리 과정과 영역해석을 통하여 문자영역을 추출하고, 추출된 문자영역으로부터 각 문자를 추출하며, 각 문자의 특징을 문자 인식기로 전달하고, 신경망기반의 문자 인식기는 문자를 인식한 후에 텍스트 정보로 변환한다.

변환된 문자는 음성합성기를 통하여 음성으로 전달하게 된다. 이러한 세부 과정은 (그림 4)와 같다. (그림 5)는 영상처리 과정중의 결과를 보여주고 있고, (그림 6)은 로봇이 벽에 있는 책을 읽고 있는 모습을 보여 주고 있다.

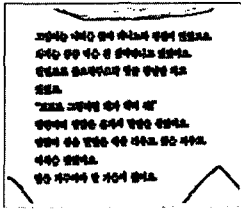
(그림 7)은 독일의 Karlsruhe 대학교에서 개발중인 멀티모달 휴머노이드 로봇을 보여주고 있다. 이 로봇은 학습기능, 시각인식 기능, 제스처 인식 및 트래킹 기능을 갖고 있으며, 이를 통하여 사람과 협업을 한다. 그림은 사람이 손짓을 하는 곳에 있는 물건을 사람의 지시대로 움직이는 일을 하는 모습을 보여 주고 있다.



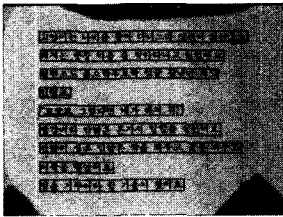
(그림 4) 사용자인식과 문자인식을 하는 책읽는 로봇의 엔진 구성도



(가) 입력 문서영상

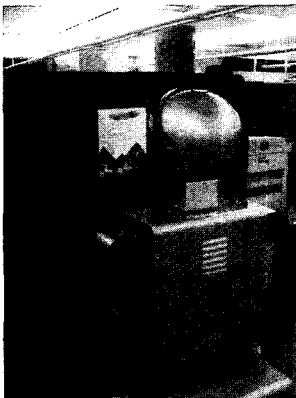


(나) 이진화 영상



(다) 문자영역 추출

(그림 5) 문자 추출 과정

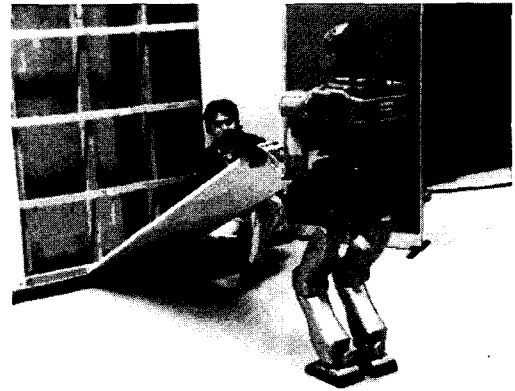


(그림 6) 책읽는 로봇



(그림 7) 다중 생체인식기능을 이용한 로봇

(그림 8)은 일본의 AIST와 동경대학교를 중심으로 개발중인 휴머노이드 로봇 HRP-2이며, 사람과 함께 합판을 드는 모습을 보여 주고 있다. 이외에도 우주에서 우주 비행사와 로봇이 장비를 설치하는 예가 있었다.



(그림 8) 사람과 물건을 함께 드는 HRP-2

이러한 과정에서 시각인식 기술은 필수적인 기술이 되고 있다. 지능형 로봇은 사물을 인식하고, 장비를 설치할 장소를 정확하게 인식하여 로봇이 위험한 곳이나, 무거운 물건을 운반하는 일 등에서 훌륭하게 인간을 도와 주게 된다. 위에서 보여준 예시에서 향후 개발될 로봇의 역할과 기대를 보여 주고 있다.

V. 결론

로봇 환경에서는 사용자와 로봇이 모두 자유롭게 움직이고 있어 사용자로 하여금 인식에 협조하도록 강요하기가 매우 곤란하다. 얼굴인식은 인간 친화적인 좋은 인식 기술이지만 움직이는 로봇에게는 사용자의 얼굴 자체가 보이지 않는 경우도 흔히 발생한다. 이

러한 상황에 대처하기 위해서는 사용자를 구별할 수 있는 보다 일반적인 특징을 도입하지 않으면 안 된다.

얼굴 자체가 보이지 않는 경우도 흔히 발생한다. 이러한 상황에 대처하기 위해서는 사용자를 구별할 수 있는 보다 일반적인 특징을 도입하지 않으면 안 된다.

이 글에서는 동적생체인식이라는 새로운 개념을 도입하고, 대표적인 동적생체인식 특징으로서 키와 옷 색깔을 사용자인식 시스템에 도입하였다.

더욱이 사용자 추적기능을 사용자 인식 프레임워크 안에 도입함으로써 이 글에서 소개된 사용자 인식 시스템은 계속적으로 이전의 인식결과 및 수집된 정보를 인지할 수 있도록 설계되었다.

지능형 로봇의 사용자 인식을 위한 다중 센서 및 다중 생체인식 방법에 대하여 여러 가지 기술을 소개하였다. 위에서도 얘기하였지만 로봇 환경에서 카메라가 연속적으로 움직이고, 사용자도 움직이는 열악한 환경에서, 사용자에게는 친화적인 자세를 견지해야 하므로 여러 가지 센서와 방법들을 퓨전하여 사용하는 기술을 소개하였다. 그리고, 마지막으로 이를 활용하는 응용 예로서, 당 연구실에서 구현한 바가 있는 책 읽는 로봇을 소개하였다. 책 읽는 로봇은 사용자를 인식할 뿐만 아니라, 사용자에게 벽에 걸려 있는 책의 내용을 음성으로 읽어주는 시나리오를 구현하였다. 또한, 시각인식기술의 응용 예로서, 독일과 일본에서 개발 중인 휴머노이드 로봇을 소개하였다. 아직은 일정한 시나리오에 의하여 제한적인 움직임과 기능을 보여주고 있으며, 실용화를 위해서는 인식기능의

발전이 절실히 요구되고 있다.

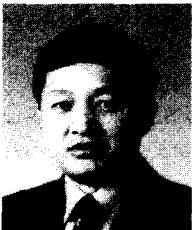
본 글에서는 시각인식기술 중에서 로봇의 사용자인식과 문자인식에 대해 논의하고, 현재 진행중인 연구에 근거한 기본적인 아이디어를 기술하였다. 여기에서는 간단한 내용이었지만, 앞으로도 계속적으로 열악하고 변화하는 환경에서 로봇이 정확하고 친절하게 서비스를 제공하기 위해서는 눈의 역할을 제대로 해야 할 것이며, 이를 위해서는 향후에도 많은 연구와 노력을 해야 할 것으로 생각된다.

참고문헌

- [1] J. Ashbourn, *Biometrics: Advanced Identity Verification*, Springer-Verlag, 2002, pp 15-44.
- [2] S. Nanavati, M. Thieme, R. Nanavati, *Biometrics: Identity Verification in a Networked World*, pp. 1-40.
- [3] K. Lee, H. Park, "A New Similarity Measure Base on Intra-class Statistics for Biometric Systems", *ETRI Journal*, vol. 25, no. 5, pp. 401-406, 2003.
- [4] N. Poh and J. Korczak, "Hybrid biometric person identification using face and voice features," *Proc. Third Int. Conf. on Audio- and Video-Based Biometric Person Authentication*, Halmstad, Sweden, June 2001, pp. 348-353.
- [5] S. Ben-Yacoub, Y. Abdeljaoued, and E. Mayoraz, "Fusion of face and speech data for person identity verification," *IEEE Trans. on Neural Networks*, vol. 10, no. 5, Sep. 1999, pp. 1065-1074.
- [6] R. Brunelli and D. Falavigna, "Person identification using multiple cues," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 17, no. 10, Oct. 1995, pp. 955-966.
- [7] J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas, "On combining classifiers," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, no. 3, March 1998, pp. 226-239.

- [8] P. Verlinde, G. Chollet, and M. Acheroy, "Multi-modal identity verification using expert fusion," *Information Fusion*, vol. 1, 2000, pp. 17-33
- [9] U. Dieckmann, P. Plankensteiner, and T. Wagner, "SESAM: A biometric person identification system using sensor fusion," *Pattern Recognition Letters*, vol. 18, 1997, pp. 827-833.
- [10] R. W. Frischholz and U. Dieckmann, "BioID: A multimodal biometric identification system," *Computer*, vol. 33, no. 2, Feb. 2000, pp. 64-68.
- [11] P. Jourlin, J. Luettin, D. Genoud, and H. Wassner, "Acoustic-labial speaker verification," *Pattern Recognition Letters*, vol. 18, 1997, pp. 853-858.

저자소개



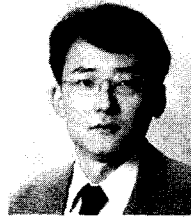
정 연구

1979년 2월 고려대학교 전자공학과 학사
 1986년 6월 미국, 클리블랜드 주립대학교 전자계산학과 석사
 1991년 5월 미국 웨인주립대학교 전산학 박사
 1978년 9월-79년 10월 대한전선 중앙연구소
 1991년 3월-현재 ETRI 지능형 로봇연구단 지능형작업제어연구팀 책임연구원
 주관심 분야 로봇비전, 영상인식, 영상처리 및 로봇 표준화



이 재 언

1984년 서울대학교 이학사(제어계측)
 1986년 한국과학기술원 공학석사(전기및전자)
 1992년 일본 東海대학 공학박사(광공학)
 1986년 3월-현재 한국전자통신연구원 재직, ETRI 지능형 로봇연구단 인간로봇 상호작용연구팀장
 주관심 분야 로봇비전, 영상인식, 멀티미디어



소 정

1986년 University of Wisconsin-Madison 전기컴퓨터공학과 학사
 1988년 SUNY at Buffalo 전산학과 석사
 1988년 SUNY at Buffalo 전산학과 박사
 1988년-1991년 Center of Excellence for Document Analysis and Recognition (CEDAR) 연구원
 1992년-2003년 한국전자통신연구원 컴퓨터소프트웨어연구소 책임연구원
 2004년-현재 한국전자통신연구원 지능형로봇연구단 지능형작업제어연구팀장
 주관심 분야 지능형로봇 응용 소프트웨어, 생체인식, 패턴 인식, 컴퓨터 비전.



김 계 령

1985년 3월-1989년 2월: 경북대학교 대학원 전자공학과 학사 과정(공학사)
 1990년 3월-1992년 2월: 경북대학교 대학원 전자공학과 대학원 석사 과정(공학석사)
 1993년 3월-1997년 2월: 경북대학교 대학원 전자공학과 대학원 박사 과정(공학박사)
 1998년 8월-2001년 2월: CENPARMI, Concordia Univ, Montreal, Canada, Post Doctoral Member & Visiting Scientist
 2001년 3월-현재 한국전자통신연구원 컴퓨터소프트웨어연구소 선임연구원
 주관심 분야 패턴인식, 문자인식, 영상처리, 신경회로망, 모바일 컴퓨팅



조 영 조

1989년 9월-1998년 2월: KIST 선임연구원 (분산제어시스템 설계)
 1993년 12월-1994년 12월: 일본통산성 기계기술연구소 로봇공학부 초빙연구원(가상현실기반 텔레로봇 제어)
 1997년 5월-1997년 8월: Univ. of Massachusetts at Amherst 초빙연구원 (로봇 제어구조 설계)
 1998년 3월-2001년 2월: KIST 책임연구원 (로봇지능제어)
 2001년 3월-2004년 1월: 아이콘트랩스 기술연구소장/상무(홈게이트웨이 및 빌딩용 DDC 개발)
 2004년 1월-현재: ETRI 지능형로봇연구팀장