보급형 u-로봇을 위한 스테레오 영상 칩을 탑재한 임베디드 시스템 및 이를 활용한 실시간 제스처 인식 알고리즘

신호철 | 김영근* | 조재일
한국전자통신연구원, 포스데이타*

요 약

제스처 인식은 로봇의 활용에 유용한 기능이지만 많은 계산 량이 요구되므로 보급형 u-로봇에는 탑재되기 어려운 단점이 있다. 따라서 u-로봇에서 일차적인 영상처리를 수행하고, 핵심 데이터만 u-로봇서버에 전송하여 다양한 응용을 수행하는 분산시스템을 구성할 필요가 있다.

본 연구에서는 스테레오 영상칩을 탑재한 ASIC과 임베디드 시스템을 활용하여 사용자의 상체 자세를 검출 할 수 있는 알고리즘을 개발하고 성능을 검증하였다. 개발된 시스템은 사용자의 다양한 상체자세를 최대 30fps까지 실시간으로 검출할 수 있으며, 80%전후의 검출률과 10cm전후의 손과 팔꿈치의 위치오차를 보였다.

I. 서 론

인간과 로봇의 상호작용은 서비스 로봇의 핵심기능이라고 할 수 있으며, 그 중 제스처 인식은 사용자에게 자연스러운 서비스 제공을 위해서는 필수적인 기능이라고 할 수 있다. 사람의 제스처 정보는 대부분 상체에서 발생하므로, 상체의 움직임을 실시간으로 검출하는 것이 필요하다.

이러한 제스처 인식을 위해서 마커 또는 센서를 몸에 부착하는 경우에는 [1,2,3,4] 사용자에게 번거로운 일이 되므로, 보급에 어려운 단점이 있으며, 열영상이나 단일 카메라를 이용하는 경우[5,6,7,8,9]에는 거리정보를 파악하기 어려우므로, 정확한 자세를 검출하기 힘들다는 단점이 있다. 한편 산업적 측면에서는 이러한 기술들이 저렴한 가격으로 제공되어야 할 필요가 있으나 기존의 연구들에서는 고가의 스테레오 카메라나, 열영상 카메라, 많은 계산량이 요구되는 알고리즘으로 인해 필연적으로 고가의 장비들이 필요하였다.

따라서 제스처 인식은 로봇의 활용에 유용한 기능이지만 보급형 u-로봇에는 탑재되기 어려운 단점이 있다. 이를 해결하기 위하여 영상 데이터를 전적으로 u-로봇 서버에 넘길 경우 네트워크 트래픽, 전송지연, u-로봇서버의 과부하 등 많은 문제점이 발생할 수 있다.

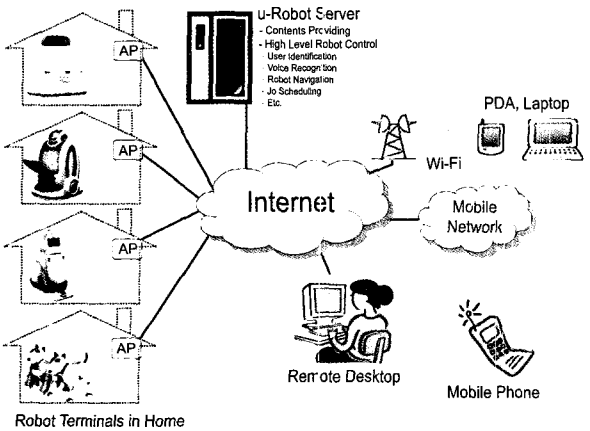
따라서 u-로봇에서 일차적인 영상처리를 수행하고, 핵심 데이터만 u-로봇서버에 전송하여 다양한 응용을 수행하는 분산시스템을 구성할 필요가 있다.

본 연구에서는 이러한 문제점을 해결하고자 사람에게 센서나 마커를 부착하지 않고서 저렴한 시스템 구성으로 사용자의 상체자세를 실시간으로 검출할 수 있는 알고리즘을 개발하고 검증하였다.

II. 본 론

1. u-로봇 시스템 구성

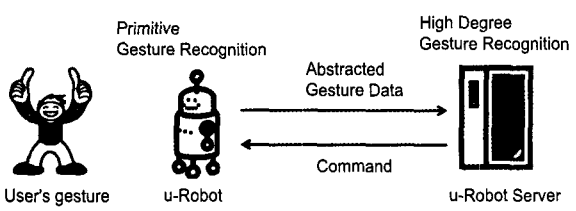
(그림 1)에서 u-로봇의 대략적인 구성을 볼 수 있다. u-로봇은 인터넷을 경유하여 u-로봇 서버, 원격지 사용자에게 연결된다. 보급형 u-로봇은 최소한의 기능을 탑재하고 있으며, 고도의 기능은 u-로봇 서버를 통하여 수행하도록 되어 있다.



(그림 1) u-Robot and u-Robot Server

앞서 설명한 바와 같이 제스처 인식은 로봇의 활용에 유용한 기능이지만 많은 연산 량이 필요하므로 보급형 u-로봇에는 탑재되기 어려운 단점이 있다. 이를 해결하기 위하여 영상 데이터를 전적으로 u-로봇 서버에 넘길 경우 네트워크 트래픽, 시간지연, u-로봇서버의 과부하 등 많은 문제점이 발생할 수 있다.

따라서 u-로봇에서 일차적인 영상처리를 수행하고, 핵심 데이터만 u-로봇서버에 전송하여 다양한 응용을 수행하는 분산시스템을 구성하는 것이 바람직하다 (그림 2).



(그림 2) Vision Data Abstraction

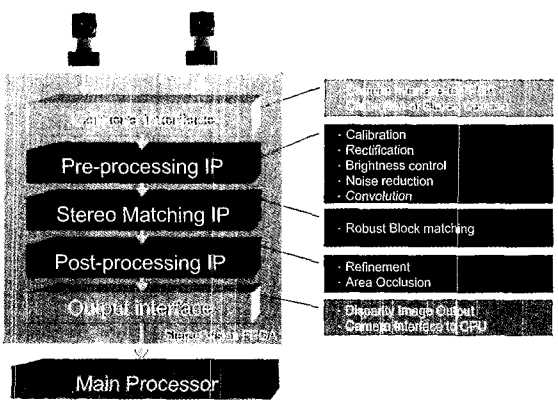
이를 위해서는 영상데이터를 ASIC(Application Specific Integrated Circuit)을 통하여 하드웨어적으로 처리하고, 임베디드 시스템에서 기초적인 영상처리를 수행하는 것이 필요하며 이에 적합한 제스처인식 알고리즘이 필요하다.

2. 하드웨어 구성

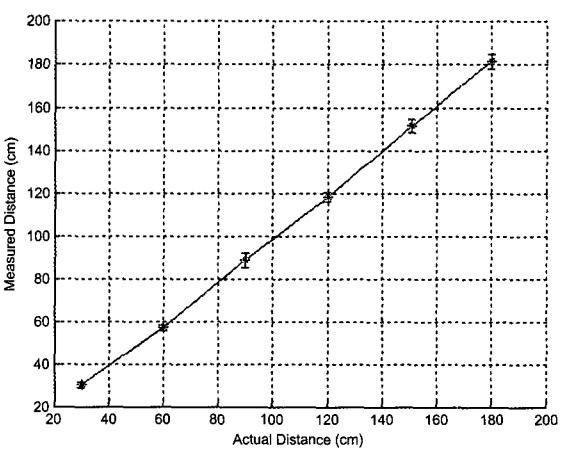
2.1 스테레오 영상 칩

스테레오 영상은 로봇의 주변공간을 3차원으로 파악할 수 있게 해주므로, 제스처 인식뿐만이 아니라 다양한 분야에서 활용될 수 있다. 단순 반복적인 계산이 요구되는 스테레오 영상은 ASIC으로 구성되는 것이 바람직 하다.

이를 위하여 parallel trellis DP algorithm[10]을 사용하여 QVGA(320x240)의 스테레오 영상을 30fps로 계산할 수 있는 FALCON으로 명명한 200만 게이트 가량의 ASIC을 개발하였다(그림 3). 본 AISC은 baseline 6/8/12cm의 CMOS카메라 쌍으로부터 입력을 받도록 설계되었다.



(그림 3) Stereo Vision ASIC Structure



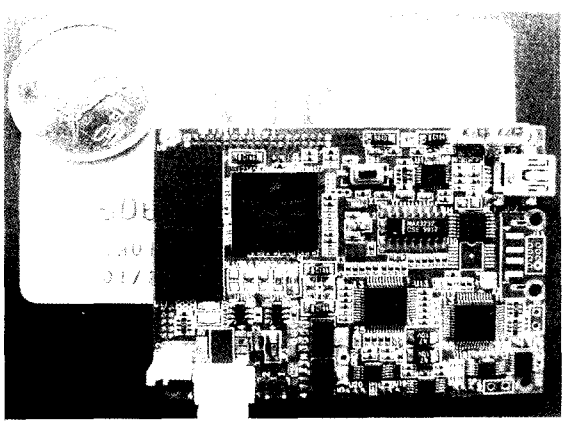
(그림 4) Stereo Vision Accuracy

본 ASIC은 정확도를 평가하기 위하여 실측거리와 depth map의 결과를 비교하였으며, 그 결과 180cm이내의 거리에서 평균 3cm이하의 오차를 보였다(그림 4).

2.2 임베디드 시스템

다양한 영상처리 및 기타의 응용프로그램을 수행하고 스테레오 영상 ASIC을 제어하기 위하여 350Mhz의 임베디드 프로세서와 128MB의 메모리를 장착한 임베디드 시스템을 개발하였다(그림 5).

스테레오 영상 ASIC은 임베디드 시스템에 장착 및 제어 되어 스테레오 영상을 제공하며, 임베디드 시스템은 영상처리를 수행할 수 있도록 구성되었다.



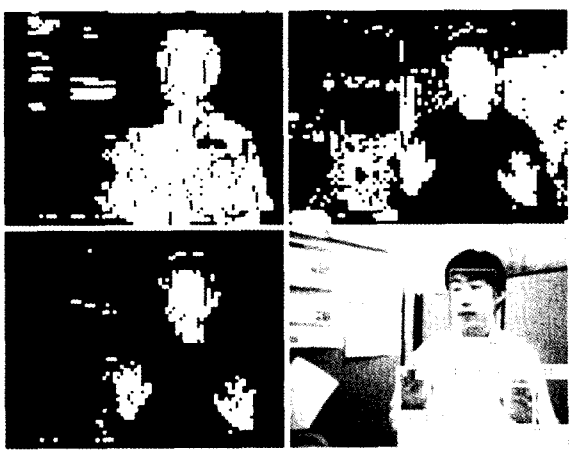
(그림 5) Embedded System for Image Processing and ASIC conrol

3. 실시간 사람의 상체 검출 알고리즘

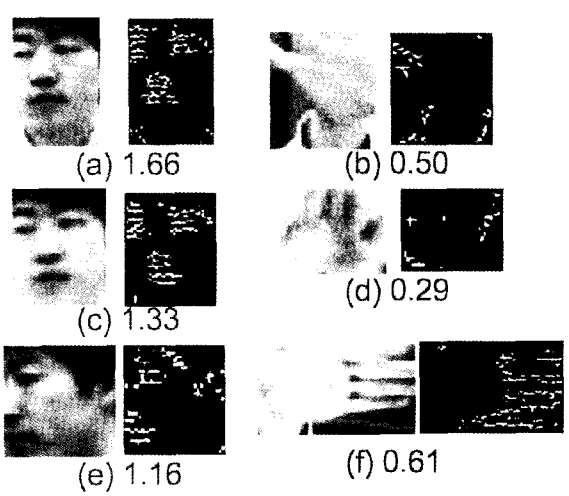
3.1 얼굴 및 손의 검출

피부색상은 계산량이 비교적 작으면서 얼굴 검출, 얼굴 인식, 제스쳐 인식 등 널리 활용되는 방법이나 광원과 배경등의 영향을 받는다는 단점이 있다. 스테레오 영상을 활용할 경우 배경을 제거함으로서 이러한 단점을 어느정도 해결할 수 있다. (그림 6)에서 배경 제거와 피부영역을 검출한 결과를 볼 수 있다.

일반적으로 얼굴은 손에 비해 높이가 높고, 가로방향의 패턴이 많으므로, 검출된 피부 영역에 대하여 높이와 가로방향 sobel mask 결과의 양을 곱한 값을 face score로 정의하고 (그림 7)에서 보는 바와 같이 검출된 영역 중 가장 큰 face score 값을 가지는 영역을 얼굴로 정의 하였다.



(그림 6) Skin Region Detection and Background Elimination

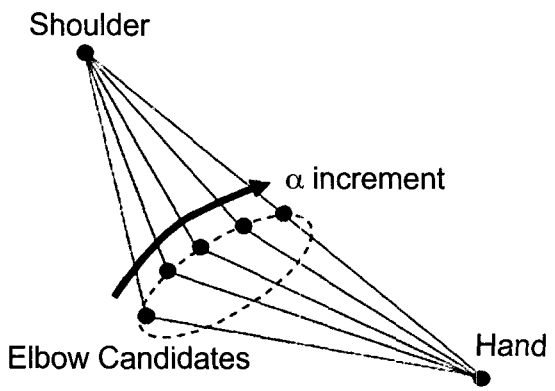


(그림 7) Face and Hand Classification

3.2 어깨 및 팔꿈치의 검출

임베디드 시스템이라는 제한된 환경에서 계산 량이 많은 알고리즘을 실행할 수 없다는 한계로 인해 사용자의 어깨와 팔꿈치 위치를 추정하기 위하여 다음과 같은 방법을 사용하였다. 사용자는 한국인의 20, 30대 표준체형[11]을 가지고 있다고 가정하고, 얼굴을 기준으로 어깨의 위치를 추정하고,

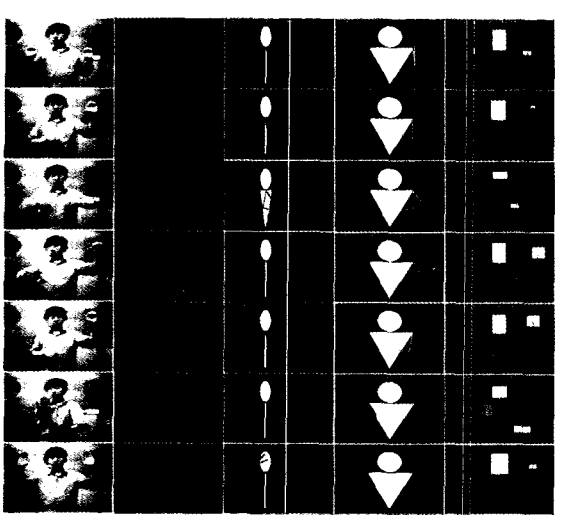
추정된 어깨와 검출된 손의 위치를 바탕으로 역기구학을 계산하여 스테레오 영상과 추정된 팔꿈치위치 오차를 최소화 할 수 있는 팔꿈치의 최적화된 위치를 찾았다(그림 8).



(그림 8) Elbow Position Optimization

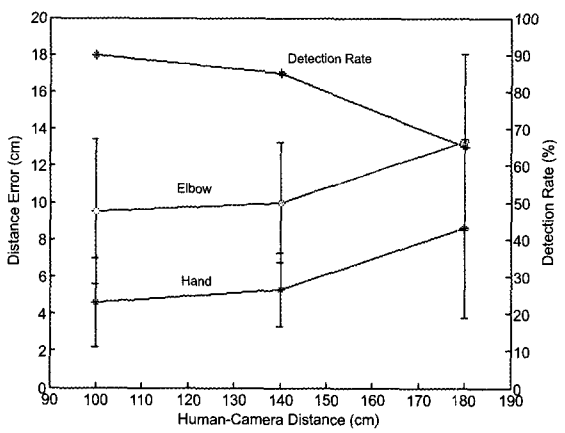
4. 성능 및 실험 결과

(그림 9)에서 제스처 검출의 결과를 볼 수 있다. 첫 번째 열부터 각각 원영상, 스테레오영상, 측면자세, 후면자세, 영상처리 결과이다. 자세의 검출은 최대 30fps의 속도로 이루어진다.



(그림 9) Gesture Detection Result

(그림 10)에서 제스처 검출의 성능을 볼 수 있다. 팔꿈치는 검출된 손의 위치를 바탕으로 계산되므로, 손에 비해 전반적으로 더 큰 오차를 보이고 있으며, 전반적으로 10cm전후의 오차를 보이고 있는 것을 알 수 있다. 로봇과 사용자의 거리가 멀어질수록 제스처 검출 성공률 또한 감소함을 알 수 있다.



(그림 10) Gesture Detection Performance

III. 결 론

제스처 인식은 로봇의 활용에 유용한 기능이지만 많은 계산 량이 요구되므로 보급형 u-로봇에는 탑재되기 어려운 단점이 있다. 이를 해결하기 위하여 영상 데이터를 전적으로 u-로봇 서버에 넘길 경우 네트워크 트래픽, 시간지연, u-로봇서버의 과부하 등 많은 문제점이 발생할 수 있다. 따라서 u-로봇에서 일차적인 영상처리를 수행하고, 핵심 데이터만 u-로봇서버에 전송하여 다양한 응용을 수행하는 분산시스템을 구성할 필요가 있다.

본 연구에서는 스테레오 영상 ASIC과 임베디드 시스템을 활용하여 사용자의 상체 자세를 검출 할 수 있는 알고리즘을 개발하고 성능을 검증하였다. 개발된 시스템은 사용자의 다양한 상체자세를 최대 30fps까지 실시간으로 검출할 수 있으며, 80%전후의 검출률과 10cm전후의 손과 팔꿈치의 위치오차를 보였다. 검출된 제스처는 서버에 전송되어 다양한

응용에 활용될 수 있다.

본 시스템은 저가의 임베디드 시스템과 스테레오 비젼 칩을 사용하므로, 로봇뿐만이 아니라 게임, 가전제품 등 다양한 분야에 적용될 수 있다

참 고 문 헌

[1] A. J. Ijspeert, J. Nakanishi, S. Schaal, "Movement imitation with nonlinear dynamical systems in humanoid robots", IEEE International Conference on Robotics and Automation, pp.1398-1403, 2002.

[2] C. A. A. Calderon, H. Hu, "Robot imitation from human body movements", AISB 2005.

[3] S. Nakaoka, A. Nakazawa, F. Kanehiro, K. Kaneko, M. Morisawa, K. Ikeuchi, "Task model of lower body motion for a biped humanoid robot to imitate human dances", IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 3157-3162, 2005.

[4] M. Riley, A. Ude, K. Wade, C. G. Atkeson, "Enabling real-time full-body imitation: a natural way of transferring human movement to humanoids", IEEE International Conference on Robotics and Automation, pp.2368-2374, 2003.

[5] L. Molina-Tanco, J.P. Bandera, R. Marfil, F. Sandoval, "Real-time human motion analysis for human-robot interaction", IEEE/RSJ International Conference on Intelligent Robots and Systems, pp.1402-1407, 2005.

[6] K. Ogaki, Y. Iwai, M. Yachida, "Posture estimation based on motion and structure models", Systems and computers in Japan, pp. 48-58, 2001.

[7] S. Iwasawa, K. Ebihara, J. Ohya, S. Morishima, "Real-time estimation of human body posture from monocular thermal images", Proc IEEE Computer Society Conference on CVPR, pp. 15-20, 1997.

[8] A. S. Micilotta, E. J. Ong, R. Bowden, "Real-time upper body detection and 3D pose estimation in monoscopic images", Lecture notes in computer science, v.3953, pp.139-150, 2006.

[9] P. Menezes, F. Lerasle, J. Dias, R. Chatila, "A single camera motion capture system dedicated to gestures imitation", 5th IEEE-RAS International Conference on Humanoid Robots, pp .430-435, 2005.

[10] H. Jeong, Y. Oh, "Parallel Trellis Based Stereo Matching Using Constraints", Lecture notes in computer science, pp.227-237, 2000

[11] http://sizekorea.kats.go.kr

약 력

신 호 철

1997년 KAIST 학사
1999년 KAIST 석사
2004 년 KAIST 박사
2004년 ~ 현재 한국전자통신연구원
2006년 ~ 현재 UST 겸임교수
관심분야 : 지능형로봇,군사/의료용 로봇

김 영 근

2005년 동국대 학사
2008년 UST 석사
2008년 ~ 현재 포스데이타
관심분야 : 영상처리, 로봇제어

조 재 일

1990년 성균관대 학사
1992년 성균관대 석사
1992년 ~ 현재 한국전자통신연구원
관심분야 : 지능형로봇,임베디드시스템, ASIC 및 SoC설계