

다해상도 변환을 이용한 실시간 능동 시각 시스템

이상웅, 최형철, 강성훈, 유명현, 이성환
고려대학교 컴퓨터학과/인공시각연구센터

Real-time active vision system using log-polar transform

Sangwoong Lee, Hyongchol Choi, Seonghoon Kang,
Myung-Hyun Yoo, Seong-Whan Lee
Dept. of Computer Science and Engineering
Center for Artificial Vision Research, Korea University

요약

KUeyes(Korea University's eyes)는 인간의 시각 정보처리 과정을 모델로 하여 고려대학교 인공시각연구센터에서 개발된, 스테레오 컬러 영상을 실시간으로 처리할 수 있는 능동 시각 시스템이다. 실시간 처리를 위하여 KUeyes는 병렬 처리가 가능하도록 개발된 영상 획득 및 영상 처리 모듈을 가지고 있으며, 다해상도 영상 변환 기법을 사용하여 입력 영상의 처리 속도를 증진시키고 있다. 이 시스템은 외형적으로는 처리된 영상에 따라 반응하여 움직이는 10-자유도의 헤드 아이 시스템으로 구성되어 있다. 본 연구에서는 영상의 칼라 정보와 움직임 정보 등을 실시간으로 분석하여 지능적이고 빠르게 개체를 탐지하고 추적하는 인간의 시각 반응 및 인식 모델을 KUeyes에 탑재하여 구현하고 실험하였다. 실험에서 얻어진 결과는 KUeyes가 인간의 시각, 인식시스템을 적절히 모델링하고 있음을 보여 주었다. 이는 KUeyes의 작동 방식과 거기에 탑재된 영상 처리 기법들이 인간의 시각 정보처리 과정을 이해하는데 매우 적합한 것임을 시사한다.

1. 서론

카메라 영상을 통해 정보를 획득하고 전달하는 방법은 센서나 음향 등의 다른 수단에 비해 정보가 다양하고 자세하기 때문에 매우 효과가 크다. 그러나 영상의 복잡성에 따라 정보 처리에 많은 노력이 소요되어 실시간 처리가 어렵기 때문에 실용적이지 못한 경향이 있다. 이에 비해 인간의 시각 및 인식 시스템은 비록 복잡하긴 하지만 매우 빠르며 정확하게 인지하고 반응하는 효율적인 시스템이다. 따라서 최근에는 인간의 시각 정보 처리 원리를 간략하게 모델링하여 영상 처리 및 로봇 제어에 이용하려는 연구들이 계속적으로 보고되고 있으며 주로 이 접근의 타당성은 헤드아이 시스템을 사용하여 검증되고 있다[4-8]. KUeyes 역시 이러한 접근법의 일환으로 고려대학교 인공시각연구센터에서 설

계하고 구현한 헤드아이 시스템이다. KUeyes는 영상 정보를 처리하는 과정에 인간의 시각 시스템이 영상을 처리하는 방식을 모사한 다해상도 변환기법과 병렬처리 기법을 이용하고 있다. 이는 정보의 선별적 추출을 통해 계산량을 감소시키고 효율적인 하드웨어를 사용함으로써 정보 처리 시간을 대폭적으로 줄여주는 장점이 있다.

본 논문에서는 영상 정보의 획득 과정부터 영상 처리 결과에 대한 반응 움직임까지를 실시간으로 모델링하여 로봇 시각으로 구현하는 것을 목표로 하였다. 각 단계별 영상 정보 처리 과정을 실험하고 검증하기 위해서 10-자유도를 가진 BiSight 헤드아이 시스템이 사용되었다. 이 시스템에는 또한 헤드아이 시스템을 운용하기 위한 하드웨어 제어 모듈과 인간의 목과 눈 운동을 모델링하여 실험하기 위한 모듈, 그리고 개체를 탐지, 인식하

고 추적하는 모듈 등이 포함되었다.

2. 관련 연구

인간의 시각 시스템을 모델링하여 실시간으로 작동하는 헤드아이 시스템의 설계에 있어 중요하게 고려할 점은 인간 시각 시스템과의 형태적 유사성과 시간적 유사성이다. 시간적 유사성이란 영상을 지각하여 판단하는 시간이 인간의 행동에서처럼 얼마나 자연스러운 지를 나타내며 이는 정보의 획득에서 반응까지 지속되는 시간이 주요 요소이다. 시간적 유사성을 극대화하기 위해 기존의 헤드아이 시스템들은 색상 정보나 차영상 등의 단순한 특징만을 주로 사용하거나 병렬 처리를 통해 수행 시간을 단축시키는 방법을 사용하고 있다[4-8]. KUeyes는 수행 시간의 단축을 위해 다해상도 변환을 통한 정보량의 감소와 병렬처리를 위한 구조적 모듈을 갖추는 방법을 사용하였다. 형태적 유사성을 판단하는 요소로서 주로 거론되는 것들은 줌, 순간 거리, 조리개, 버전스(Vergence) 등과 관련된 자유도이다. 순간 거리는 양 눈 사이의 거리를 말하며 인간의 경우 대개 6-7 cm[7] 정도 이나 실제 카메라의 경우 카메라의 크기 때문에 인간과 똑같이 설계하기는 매우 어려운 실정이다. 최근에 개발된 헤드아이 시스템 중에 비교적 인간과 유사한 국내의 KIST의 HECter[4]는 순간 거리 10 cm로 설계되었으며, 순간 거리를 조절할 수 있는 스웨덴의 CVAP의 KTH-Head, 중국의 Northern Jiaotong 대학의 NJU-Head 등도 설계가 되었다[6]. 그러나 순간 거리를 인간처럼 가깝게 설계한 경우 카메라의 크기가 작아져야 되므로 줌, 조리개, 초점조절 기능을 함께 탑재하기는 힘들다. 따라서 미국 펜실베니아 대학의 PennEyes, 스페인 Las Palmas 대학의 DESEO 등처럼 순간 거리는 비교적 멀지만 조리개, 초점조절 등 눈동자의 자유도를 가지고 있는 헤드아이 시스템과 인간과는 달리 광학적 특성을 이용한 줌을 이용하거나 인지력과 해상도 문제로 인해 카메라 3 개를 이용한 NIST의 TRICLOPS와 같은 헤드아이 시스템도 개발되었다[4-8]. KUeyes는 순간 거리

가 10"이며 목, 눈의 4-자유도와 눈동자 움직임의 6-자유도를 가지고 있는 BiSight 시스템이다.

3. 시스템 구조

3.1 하드웨어 구조

KUeyes는 다음과 같이 각 목적에 맞게 크게 3-부분으로 나뉘어져 있다. 첫 번째는, 영상을 획득하기 위한 카메라와 이를 구동하는 모터에 관련된 부분이고, 두 번째는 영상을 실시간으로 처리하기 위한 프로세서들이 있는 부분이다. 마지막으로 세 번째는, 영상을 처리하여 사람이나 개체의 얼굴을 인식하는 부분이다. 이들은 헤드아이 시스템과 두 대의 PC 안에 탑재되어 있으며 얼굴인식 부분은 계산량의 분산을 위해서 PC 한 대에 독립적으로 분리되어 설치되었다. 두 PC간 정보교환은 TCP/IP 프로토콜을 이용한다.

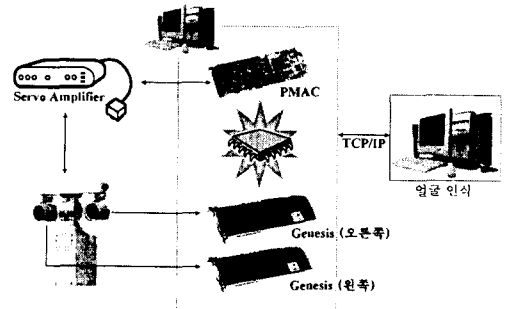


그림 1. KUeyes의 하드웨어 구성도

3.1.1 영상획득 및 처리

KUeyes가 실세계에서 실시간으로 영상을 획득, 처리하는 과정은 막대한 계산량을 요구한다. 주어진 계산량을 실시간으로 구현하기 위해서는 최적화된 프로그램의 개발과 함께 영상처리에 최적화된 하드웨어를 사용하는 것도 필수적이라 할 수 있다. 본 연구에서는 빠른 영상처리 속도를 위하여 Matrox사의 Genesis 보드 두 대와 Pentium III-600MHz PC 두 대를 이용하고 있다. Genesis 보드는

TI사의 TMS320C80 DSP를 이용한, 영상처리에 최적화된 명령어군을 포함하고 있어 영상 획득과 영상처리에 뛰어난 성능을 보여주고 있다. Genesis 보드와 PC의 성능을 최대한 발휘하기 위하여 각각의 하드웨어에 맞는 소프트웨어 MIL(Matrox Image Library)과 IPL(Intel Image Processing Library)을 사용하였다. 이렇게 구성된 영상처리 시스템은 좌, 우 카메라에서부터 각각의 영상을 동시에 획득하여 Genesis 보드에서 전처리 과정과 정보를 추출하는 과정을 수행한다. 이렇게 얻어진 정보들을 움직임 정보 추출용 PC와 얼굴 인식용 PC에서 각기 분석 처리한다.

낼 수 있다.

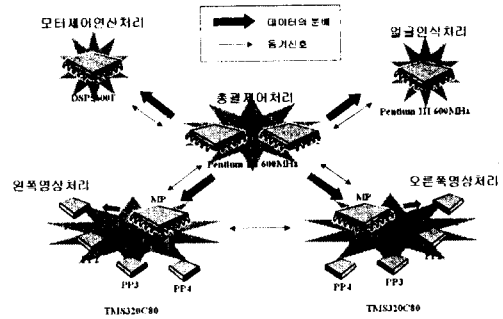


그림 2. 병렬처리의 흐름도

3.1.2 모터제어

Kueyes를 작동하는 8 개의 직류 모터들을 제어하기 위해서 Delta Tau사의 PMAC-PC 보드와 속도 향상을 위한 선택 사양인 DPRAM(Dual-Port RAM)을 이용하고 있다. 이 보드는 자체적으로 Motorola사의 DSP56001을 내장하고 있어서 각 축에 대한 벡터합 계산을 해주며, 또한 부분적으로 KUeyes의 각도를 제어하는 프로그램을 다운로드시켜서 주 프로세서와는 병렬적으로 수행하는 기능을 통해 시스템의 속도를 향상시킬 수 있다.

3.1.3 병렬처리

KUeyes에는 영상 정보를 실시간으로 처리하기 위하여 위에 소개된 여러 가지 프로세서들을 이용하여 병렬처리 시스템을 구성하고 있다. 주로 영상처리의 대부분을 담당하는 Genesis 보드가 헤드아이의 좌, 우 영상을 독립적으로 처리하고 있으며 PC 2대에서 전체적인 제어와 얼굴인식을 담당하고 있다. 또한 모터와 관계되어 있는 연산과 제어는 모터를 제어하는 PMAC 보드에 내장되어 있는 프로세서에서 담당하여 각 프로세서들이 서로 시간상의 동기를 맞추어 가며 영상 처리 계산량을 분산하여 수행하고 있다. 병렬처리상에서는 프로세서들간의 인터럽트와 공유 가능한 데이터 버스의 확보, 실행 순서와 동기 문제를 최적화하여야만 효과적인 결과를

3.1.4 헤드아이

양안 카메라를 장착할 수 있는 헤드아이 시스템으로 Helpmate사의 Unisight 위에 Bisight를 장착한 형태를 가지고 있다. 이 두대의 구성으로 4-자유도를 가지며 이 4-자유도는 목의 좌우-상하 움직임과 좌, 우측 카메라의 독립적인 좌, 우 회전으로 구성된다. KUeyes에 장착된 BiSight 시스템의 움직임은 인간과 대체적으로 유사하거나 뛰어나지만, 눈동자의 움직임이 상하로 되지 않는 점과 최대 가속력에서 인간에 비해서 떨어진다는 단점이 있다.

3.1.5 카메라

카메라는 획득된 영상의 질에 가장 중요한 영향을 미치는 요인이다. 본 시스템에서는 칼라영상을 처리하기 위해 PULNiX사의 TMC-7RGB CCD 카메라를 사용하였고 Fusinon사의 H10x11E-X41 렌즈를 장착하였다. 이 렌즈와 카메라는 각기 전동 모터를 내장하고 있어서 초점조절과 조리개, 줌을 사용자 임의대로 조절할 수 있으며 눈동자 움직임에서 3-자유도를 가진다. 한편, 카메라 제조시 생기기 마련인 양쪽 영상의 오차를 없애기 위하여 하드웨어적으로 제공되는 자동초점과 자동노출기능을 배제하였으며, 대신에 영상의 색상분포와 영상의 기울기 정보를 이용하여 두 카메라에서 동질의 영상이

나오도록 하는 방법을 사용하여 초점과 노출 정도를 조정하였다.

3.2 소프트웨어 구조

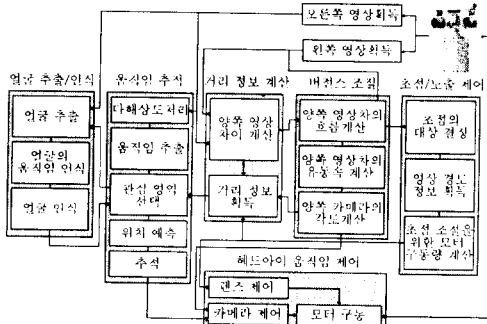


그림 3. KUeyes의 영상처리 흐름도

3.2.1 다해상도 처리

인간의 눈에는 영상을 처리하는 단위인 수용장이 균일하게 분포되어 있는 것이 아니라 관심을 가지고 있는 영상의 중앙에는 조밀한 분포를 보이며 주변부로 갈수록 밀도가 낮아지는 형태를 지니고 있다. 이러한 형태는 처리해야 하는 영상 정보를 감소시켜서 영상처리의 계산량을 줄일 수 있다는 장점이 있다. 이러한 인간의 시각 시스템에 근거하여 본 연구에서는 log-polar 영상 좌표계를 도입한 다해상도 변환 기법을 적용하였다. 이 방법은 중앙부로부터는 점의 크기를 작게 하여 밀도 있는 영상을 획득하는데 비해, 주변부에서는 점의 크기를 상대적으로 크게 설정하여 영상의 질을 저하시키는 대신 처리해야 하는 시간을 줄이도록 구성된다. 다해상도 변환에서는 영상좌표계를 변환하는 과정에서 시간이 걸리지만 변환 후 과정에서는 보다 빠른 영상처리를 행할 수 있어 결과적으로 속도 향상 효과를 얻을 수 있다[2].

3.2.2 움직임 추적

움직이는 물체의 추적을 위하여 광류(optical flow) 방법을 사용하였는데 이는 영상의 밝기의 기울기를 1차 차분하여 이용하는 Horn과 Schunk의 방법이다[1]. 카메라가 정지

한 상태에서는 광류가 측정되는 영역을 추출해 내어 정보로 이용하면 되지만 카메라가

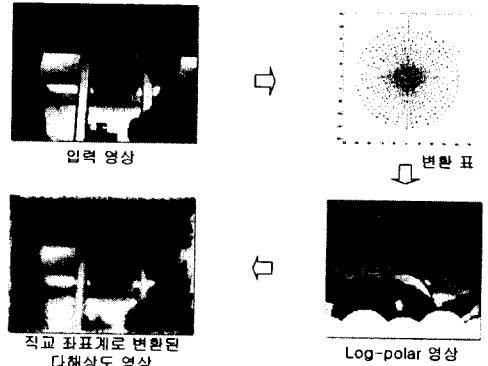


그림 4. 다해상도 영상변환 과정

이동하는 능동시각 시스템에서는 배경 영상이 함께 움직이게 되므로 전체 광류의 측정치를 실제 움직이는 개체의 움직임 정도에 보상하여 실질적인 움직임 개체의 추적정보를 추출하여 모터구동 부분으로 넘겨주게 된다.

3.2.3 눈동자 움직임 구현

스테레오 헤드아이에서의 눈동자 운동은 두 카메라의 응시점이 맞추어진 관심있는 물체까지의 거리를 측정할 수 있도록 해준다. 따라서 우리는 눈동자 운동의 빠르고 정확한 움직임을 위해 각 카메라로부터 획득된 영상 간의 disparity flow와 flux를 이용하는 제어 모델을 개발하였다. 이는 두 카메라 중 주요 추적 정보를 가지는 한 카메라가 움직이는 물체를 추적할 때 두 카메라의 각 영상의 disparity flow를 계산한 후 다해상도내의 관심영역을 지나는 disparity flow의 flux를 다시 계산해내어 flux의 값의 부호에 따라서 flux가 0에 가까워질 때까지 나머지 카메라를 flux의 반대 방향으로 이동하여 두 카메라의 응시점이 같아지도록 하는 방법이다.

3.2.4 KUeyes의 움직임 구현

움직임 추적을 통해 구해진 물체의 움직임 정도를 실제 KUeyes의 움직임 정도로 변환하기 위해서는 가상 공간상에서 화면의 점 단위를 미터법으로 바꿔주는 작업이 필요한데 이는 실험을 통하여 결정하였다. 실험을 통해 결정된 변환 상수 값과 카메라에서 개체까지의 거리 정보를 이용하여 양 쪽 카메라의 각도와 목의 움직임 각도의 관계를 삼각함수를 이용하여 구해낼 수 있다. 이것은 목이나 눈 어느 한 쪽이 움직일 때 응시점을 고정시키기 위하여 사용된다. 거리 정보는 disparity, 버전스(vergence) 각도, 초점조절 정보를 이용하여 구할 수 있는데 여기서는 속도와 효율을 고려하여 버전스 각도를 이용하였다. 다만 이 방법은 멀리 있는 물체에 대해서는 버전스 각의 움직임이 세밀함에 한계가 있어 정확한 거리를 구할 수 없다는 단점이 있다. 목과 눈의 움직임의 각도들이 결정된 후에 눈과 목의 움직임에 대한 일반적인 규칙들을 설정하고 여기에 맞추어 인간의 시각 시스템과 같이 눈과 목의 자연스런 움직임을 하도록 하였다.

3.2.5 얼굴 인식

얼굴 인식은 크게, 입력받은 영상에서 얼굴 영역을 추출하는 기술과 추출된 얼굴 영역을 인식하는 두 가지 기술이 사용된다. 얼굴 영역 추출 기법은 형판 정합 방법을 사용하였으며, 이때 탐색 공간을 축소하기 위해서 칼라 정보를 사용하였다. 즉, 칼라 정보를 이용하여 얼굴 영역이 나타날 가능성이 공간만을 탐색하여 계산량을 줄인다. 추출된 얼굴 영상은 인식기를 이용하여 입력 영상에 누가 존재하는지 판별할 수 있게 해준다. 이때 인식기에서 사용되는 기술은 주성분 분석 기반 방법과 SVM(Support Vector Machine) 기반 방법의 두 가지를 선별적으로 사용하였다 [3].

4. 구현 및 성능 분석

인간의 시각 정보처리 시스템의 모델링 구현을 위해서 실험은 움직이는 사람의 영상을

실시간으로 입력 받아 이루어 졌다. 본 실험은 카메라의 최대 해상도인 640x480 화소에서 실행되었으며, 각각의 칼라 영상을 좌, 우 카메라에서 획득하여 두 대의 Genesis 보드

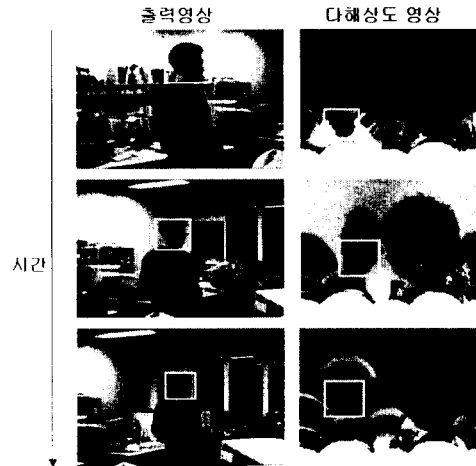


그림 5. 다해상도 영상내의 움직이는 사람 추적

에서 병렬적으로 영상을 디지털 신호로 변환하였다. 이 영상은 다시 다해상도로 변환을 통해 상대적으로 중요하지 않은 정보를 줄여 정보의 양을 약 1/50 정도로 감소시킨다. 얻어진 다해상도 영상으로부터 사람의 움직임 정보를 광류 방법으로 추출하여 추적한다. 이 추적 과정에서 추적 대상의 광류뿐만 아니라 조명과 배경의 변화에 의해 잡영으로 분류될 수 있는 광류들이 추출될 수 있는데, 이것들은 전 추적과정에서 생성된 추적 대상의 칼라 정보와 움직임 예측 등에 의해 제거된다. 광류의 움직임 정도에 따라 KUeyes는 능동적으로 사람을 추적하게 되며 목과 눈의 자연스러운 움직임을 통해 인간의 움직임을 흉내내도록 구현되었다. 이 과정은 영상의 입력부터 KUeyes의 움직임까지의 순환되는 구간의 시간이 빠를수록 더욱 자연스러워진다. 이 반복 시간을 줄이기 위해 이중 영상 저장(double buffering)과 같은 효과적인 병렬 처리 과정을 거쳤다. 위와 같은 방법들을 통해 640x480 해상도에서 초당 약 15 프레임 정도를 처리할 수 있는 수행 결과를 보였다. 사람이 뛰어나는 등의 순간적인 빠른 움직임

을 보여 영상의 중심부에서 벗어나지 않는 경우를 제외하면 추적결과는 만족할 만 하였다. 그림 5는 정지하고 있는 헤드아이의 입력영상을 다해상도로 변환한 후 움직임 정보를 이용하여 추적할 개체를 찾아내어, 일반 영상좌표계로 역변환하여 추적 개체에 사각형을 표시하여 주는 과정을 보여준다. 이 과정에서 얼굴 영역은 칼라 영역 추출 방법과 형판 정합 방법을 사용하여 계속적으로 검출이 되며, 만약 정면 얼굴이라고 판단이 되면 TCP/IP를 이용하여 얼굴 영역을 다른 PC로 전달하여, 가지고 있는 데이터베이스 안에서 얼굴 인식 과정을 수행한다. 얼굴 인식에서는 안경 착용 여부, 표정 변화 등과 같은 얼굴의 변화도 고려되었으며, 10명의 얼굴 데이터베이스에 대해 약 95%의 인식 성능을 보여 주었다.

5. 결론 및 향후 과제

본 논문에서는 인간의 시각 시스템처럼 영상을 획득하여 다해상도 기반으로 변환하고, 정보를 병렬처리하는 실시간 능동 시각 시스템인 KUeyes에 대하여 소개하였다. 실시간 수행을 위해 적용된 다해상도 변환 기법과 병렬처리 기법은 시간 단축에 많은 효과를 보여주었는 바 이는 인간 시스템의 정보처리 효율성을 시사하는 것이다. 본 실험은 비교적 간단한 특징들을 처리하도록 설계되었으나, 영상의 획득에서 처리 결과에 대한 반응까지 인간 시스템을 적절히 모델링하는데 성공하였다고 판단된다. 이는 KUeyes를 통해 인간의 시각 및 인지 시스템의 이해를 돕기 위한 실험 기반을 확보했음을 의미하는 것이며, 이를 이용하여 실세계의 복잡한 영상들을 인지하는 방법에 대하여 연구를 진행하는 것이 타당함을 보여준다. 현재의 자연스럽지 않은 눈동자 움직임과 속도를 개선하여, 앞으로는 영상에 대한 인간의 고차적인 판단과 반응까지를 모델링하는 과정으로 KUeyes의 이용을 확장할 수 있을 것이다.

감사의 말씀

본 연구는 과학기술부 창의적연구진흥사업의 연구비 지원을 받았음.

참고 문헌

- [1] 강성훈, 안창, 이성환, "망막-피질 매핑된 동영상에서의 움직임 추정 방법," *한국정보과학회 봄 학술발표논문집*, 목포, 제 26권 제 1호(B), 1999년 4월, pp. 481-483.
- [2] 강성훈, 이성환, "망막의 신경절 세포 특성에 기반한 다해상도 영상 표현 방법," *한국정보과학회 가을 학술발표논문집*, 제 25권 제 2호, 1998년 10월, pp. 351-353.
- [3] 김재진, 여창욱, 이성환, "Support Vector Machine을 이용한 다양한 조명 및 표정 하에서의 얼굴 인식," *한국정보과학회 봄 학술발표논문집*, 목포, 제 26권 1호(B), 1999년 4월, pp. 498~500.
- [4] M. Hwang-Bo, B.J. You, S.R. Oh, and C.W. Lee, "Design and Redundant Motion Control of Active Stereo Head-Eye System with Vergence," *Proceedings of 7th IEEE Workshop on Robot and Human Communication*, Takamatsu Japan, Sept 1998, pp. 135-141.
- [5] B.C. Madden and U.M. Cahn von Seelen, "PennEyes A Binocular Active Vision System," *Technical Report MS-CIS-95-37 / GRASP LAB 396*, 1995.
- [6] Z.Y. Zhang and B.Z. Yuan, "NJU Head: An Active Head-Eye System," *Proceedings of IEEE International Conference on Image Processing*, Vol. 2, 1997, pp. 1760-1764.
- [7] J.C. Fiala, R. Lumia, K.J. Roberts, and A.J. Wavering, "TRICLOPS: A Tool for Studying Active Vision," *International Journal of Computer Vision*, Vol. 12, No. 2-3, 1994, pp. 231-250
- [8] Y. Mae, Y. Shirai, J. Miura, and Y. Kuno, "Object Tracking in Cluttered Background Based on Optical Flow and Edges," *Proceedings of 13th International Conference on Pattern Recognition*, Vol.1, 1996, pp.196-200.