

## 영상 특징점 추출 기반의 임베디드 객체인식 시스템

이수현, 박찬일, 강철호, 이혁준, 이형근, 정용진  
광운대학교

### An Embedded Object Recognition System based on SIFT Algorithm

Su-Hyun Lee, Chan-Il Park, Cheol-Ho Gang, Hyukjoon Lee, Hyungkeun Lee, Yong-Jin Jeong  
Kwangwoon University

**Abstract** - 본 논문에서는 임베디드 환경을 위한 객체인식 시스템의 구조 및 실시간 처리를 위한 객체인식기의 하드웨어설계를 제안한다. 제안된 구조는 SIFT(Scale Invariant Feature Transform)를 이용하여 사물의 특징점을 추출하고, 비교하여 객체를 인식한다. SIFT는 영상의 크기 및 회전 등의 변화에 적용이 뛰어난 알고리즘이지만, 복잡한 연산이 반복되어 연산시간이 많은 특성상 임베디드 환경에서 실시간 처리가 어렵다. 따라서 해당 알고리즘을 하드웨어로 설계하여, 임베디드 사물인식 시스템에 적용한다. 사물인식의 빠른 처리와 인식영역의 구분을 위해 JSEG 영상분할 알고리즘을 활용하며, SIFT 특징점 추출 연산과 병렬 실행이 가능하도록 SIFT와 함께 하드웨어 구조로 설계한다.

#### 1. 서 론

영상기반의 객체인식기술은 카메라로부터 입력받은 영상을 처리하고, 분석하여 해당 영상 내에 존재하는 객체를 인식하는 기술로서 보안감시 시스템[1], 자율주행 시스템[2], 지능형 로봇의 시각 정보처리[2], 의학 영상 분석[3] 등 다양한 분야에서 활용되는 차세대 기술이다. 최근 영상인식 기술의 발달 및 활용분야의 다양화로 시스템의 고성능 및 소형화를 요구하고 있으며, 특히 지능형 로봇 분야에서는 로봇의 시각정보 처리를 위하여 임베디드 환경에서 실시간 처리를 할 수 있는 객체인식시스템을 필요로 한다[2]. 하지만 영상기반의 객체인식기술은 대용량의 영상 데이터를 처리하며, 복잡한 연산을 반복하는 특성상 임베디드 환경뿐만 아니라 일반 PC환경에서도 실시간 처리가 힘들다. 따라서 임베디드 환경의 객체인식시스템 설계를 위해서는 객체인식기의 하드웨어 설계가 필수적이다. 본 논문에서는 지능형 로봇 및 다양한 임베디드 환경을 위한 객체인식 시스템의 구조 및 객체인식기의 하드웨어설계를 제안한다.

#### 2. 본 론

##### 2.1 객체인식

본 논문에서는 객체인식을 구현을 위해 영상의 특징점을 추출하기 위한 SIFT(Scale Invariant Feature Transform) 알고리즘[4]과 영상의 영역 분할을 위한 JSEG 알고리즘[5]을 사용하며, 제안 하는 객체인식의 구조는 그림 1과 같다.

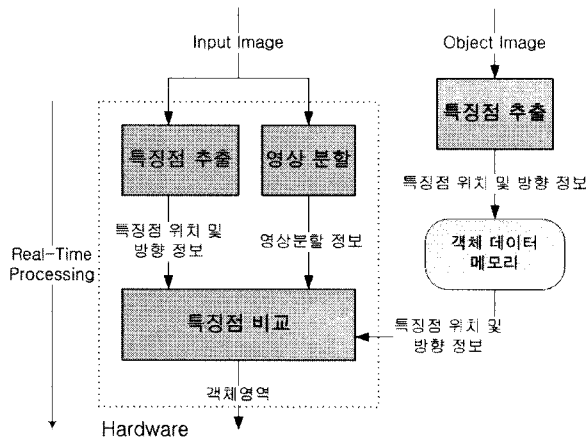


그림 1. 객체인식 구조

※ 본 연구보고서는 지식경제부 출연금으로 ETRI, SoC산업진흥센터에서 수행한 IT SoC 핵심인력 양성사업의 연구결과입니다.

제안하는 객체인식의 구조는 다음과 같다. 먼저 검출하기 위한 객체의 영상에서 특징점을 추출하여 미리 객체 데이터 메모리에 저장한다. 입력영상으로부터 특징점을 추출하고 영상을 분할한 후에 특징점을 비교한다. 검출하고자 하는 객체의 특징점과 일치하는 영역을 선별하여 객체영역으로 표시하여 출력한다.

##### 2.2 SIFT 알고리즘

SIFT(Scale Invariant Feature Transform) 알고리즘은 입력된 영상에서 벡터형식의 특징데이터를 추출하는 알고리즘으로, 영상의 크기 변환 또는 회전 등과 같은 영상 변형에 대해 강인한 특성을 가지고 있다. SIFT 알고리즘은 그림 1과 같이 크게 8가지 연산으로 이루어지며, 각각의 동작은 아래와 같다.

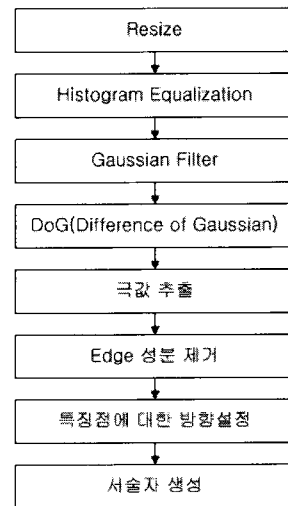


그림 2. SIFT 알고리즘 순서도

- (1) 특징점의 매칭단계에서 크기가 다른 객체의 인식성능을 높이기 위해 Resize를 통하여 여러 크기의 영상을 생성한다.
- (2) 입력 영상 픽셀의 최대값과 최소값을 이용하여, 전체 픽셀 값의 분포를 균등하게 만들기 위한 Histogram Equalization을 실행한다.
- (3) Gaussian Filter의 Sigma 값을 여러 단계로 조정하여 영상을 흐리게 만든다.
- (4) (3)에서 생성된 Gaussian 영상을 이용하여 단계별로 차영상을 만든다.
- (5) 연속된 3개의 DoG영상의 3x3x3 영역으로부터 극값(Maxima와 Minima)을 추출한다.
- (6) 극값이 위치한 좌표의 성분이 Edge에 해당하는지 체크한다.
- (7) Gaussian 영상과 극값의 좌표를 이용하여 특징점 주변 픽셀의 경사도를 추출한다.
- (8) 특징점 주변픽셀의 벡터성분을 추출한다.

##### 2.3 JSEG 알고리즘

영상에서 객체의 영역을 추출하기 위해서는 입력영상을 알맞게 분할할 필요가 있으며, 본 논문에서는 영상을 효율적으로 분할하기 위해 JSEG 알고리즘을 사용하였다. JSEG는 영상의 색상 분포에 기반을 둔 결감을 기준으로 영역을 분할하는 알고리즘으로 외곽선이나 칼라정보만을 이용한 분할보다 나은 성능을 보여준다. JSEG 알고리즘은 그림 3과 같이 크게 5단계의 연산으로 이루어지며, 각 단계별 내용은 아래와 같다.

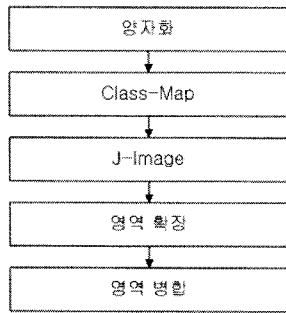


그림 3. JSEG 알고리즘 순서도

- (1) Peer Group Filtering을 통한 양자화를 실행한다.
- (2) 양자화된 각 색상마다 고유의 레이블을 부여하여, CCM(Color Class Map)을 만든다.
- (3) 영상의 질감을 상대적인 수치로 나타내는 J값을 이용하여 J-Image를 생성한다. J값은 식 (1)을 이용하여 구한다.

$$J = (S_T - S_W) / S_W$$

$$m = \frac{1}{N} \sum_{z \in Z} z$$

$$m_i = \frac{1}{N_i} \sum_{z \in Z_i} z_i$$

$$S_T = \sum_{z \in Z} \|z - m\|^2$$

$$S_W = \sum_{i \in Z} S_i = \sum_{i \in Z} \sum_{z \in Z_i} \|z - m_i\|^2$$

z : Color Class-map에서 각 레이블 들의 좌표  
 m : 각 z들의 위치 평균 m<sub>i</sub> : 각 클래스 마다의 위치 평균값  
 C : 클래스의 개수  
 S<sub>T</sub> : 위치 평균에서 각 레이블 간의 거리의 합  
 S<sub>W</sub> : 각 클래스별 S<sub>T</sub>값을 구해서 합한 값

- (4) 각 영역의 내부에서 가장 낮은 밝기값을 가지는 픽셀을 중심 픽셀로 결정하고, 중심 픽셀에서 이웃 픽셀로 한 픽셀씩 확장하며, 이웃 영역과 만나게 될 때 확장을 중지한다. 영역 확장을 통해 각 영역의 정확한 경계를 결정한다.
- (5) 색상 히스토그램을 이용하여 유사도가 비슷한 영역끼리 병합한다.

### 2.4 임베디드 객체인식 시스템

SIFT 알고리즘은 여러 장의 Gaussian과 DoG 영상의 반복 연산 및 특징점을 추출하기 위한 복잡한 연산으로 인하여 실시간 처리가 어려운 면이 있다. 따라서 임베디드 환경에서 실시간 처리를 위해 하드웨어 구현이 필수적이다. 표 1에 SIFT 알고리즘의 소프트웨어 구동시간과 현재까지 설계된 하드웨어의 속도와 비교 하였다.

〈표 1〉 SIFT 특징점 위치 결정 단계 까지 구동 시간 비교 [6]

연산 단계	소프트웨어 ARM9 400Mhz	하드웨어속도 FPGA 100Mhz
Resize 단계	941ms	220ms
Histogram Equalization 단계	201ms	210ms
Gaussian Filter 단계 (DoG 포함)	103,390ms	
특징점 위치결정 및 방향설정	68,763ms	
Total	154,956ms	-

표 1에서 테스트한 소프트웨어는 640x480픽셀 사이즈의 입력영상을 기준으로 고정수준 모델로 구현된 것으로 SIFT 알고리즘의 특징점 위치결정 단계까지 실행한 결과이다. 하드웨어 구조는 소프트웨어의 연산속도에 비하여 Resize 단계는 약 4.3배 빨라졌으며, Histogram Equalization 단계, Gaussian Filter 단계, 특징점 위치 결정 및 방향설정 단계는 파이프라인 구조로 하나의 모듈로 설계되어 약 820배 빨라졌다. SIFT 알고리즘에서 영상의 읽기와 쓰기 속도를 가장 많이 반복하는 Gaussian Filter 단계에서 병렬 처리 및 효율적인 메모리 관리가 가능한 하드웨어로 설계할 경우 표 1의 결과와 같이 그 연산 속도가 비약적으로 빨라짐을 볼 수 있다.

JSEG 알고리즘은 SIFT 특징점 추출 이후 특징점 매칭 단계에서 객체의 영역을 표시하기 위해 필요하므로, 빠른 연산을 위해서 SIFT 알고리즘과 병렬처리가 될 필요가 있다. SIFT 알고리즘을 하드웨어로 설계할 경우 CPU에서 JSEG를 처리하여도 객체인식 단일 프로세싱에서 문제가 없으나, 다양한 어플리케이션이 종합적으로 수행되어야 하는 임베디드 환경에서는 CPU의 연산 부담을 줄일 필요가 있다. 또한 두 알고리즘은 동일한 입력영상을 Input으로 필요로 하므로 효율적인 메모리 구조의 설계 및 관리를 위해서 병렬처리가 가능한 하드웨어 구조로 설계할 필요가 있다.

제안하는 시스템은 SIFT 알고리즘과 JSEG 알고리즘을 하나의 객체인식기로 통합하여 하드웨어로 설계한다. 그리고 AMBA Bus를 통하여 ARM9 CPU와 연결하고, 카메라로부터 들어오는 영상 저장하는 Frame Buffer와 검출 하고자하는 객체의 데이터를 저장하는 Object Data Memory를 객체인식기와 연결하여 임베디드 객체인식 시스템을 구성한다. SIFT 하드웨어의 연산시간을 비교한 표1에서 볼 수 있듯이 대용량의 영상 데이터를 다루는 영상 시스템은 연산이간을 주로 데이터를 읽고 쓰는 시간에 소모하게 된다. 따라서 시스템의 구조를 설계함에 있어서도 해당 사항을 고려하여, 많은 양의 데이터 전송이 필요한 Camera Image Buffer와 Object Data Memory를 객체인식기에서 직접 접근이 가능한 구조로 설계 한다. 아래의 그림 4에 시스템의 구조를 나타내었다.

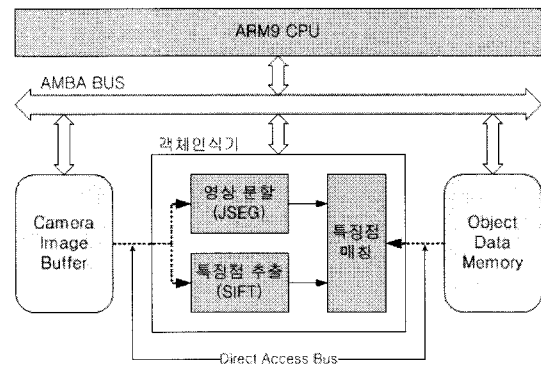


그림 4. 임베디드 객체인식 시스템 구조

### 3. 결 론

본 논문은 임베디드 환경을 위한 객체인식 시스템 구조를 설계하고, 실시간 처리를 위해 객체인식기를 하드웨어로 설계하는 것을 제안하였다. 객체인식기는 영상의 특징점을 추출하기위한 SIFT 알고리즘과 객체의 영역을 추출하기위해 입력 영상을 효율적으로 분할하는 JSEG 알고리즘으로 구성하였다. 그리고 많은 양의 영상데이터를 처리하는 시스템의 특성을 고려하여 객체인식기에서 직접적인 메모리 접근이 가능한 구조로 시스템을 설계 하였다.

현재 하드웨어 구조는 SIFT 알고리즘의 특징점 위치결정 및 방향설정 단계까지 설계 및 검증이 완료되었다. 차후 임베디드 검증 환경을 통해 해당 시스템의 구조 및 객체인식기 하드웨어를 전체적으로 검증할 계획이다. 또한 해당 시스템에 적용된 SIFT 알고리즘은 흑백영상을 기준으로 설계되어 색상이 다른 동일한 형태의 물체에 대하여 구별하지 못하는 단점이 있다. 따라서 테스트를 통하여 컬러정보를 비교 부분을 추가적으로 적용하고 보완할 계획이다.

### [참 고 문 헌]

- [1] 김현, 박찬호, 우종우, 두석배, "위치기반 감시 서비스를 위한 이동 객체 추적 및 인식", 대한전자공학회 2008년 하계종합학술대회, 2008. 6, pp. 1211~1212
- [2] 김영옥, 이종배, 백준기, "지능 로보틱스 영상처리 기술동향," 한국정보과학회 정보과학회지 제26권 제4호, 2008. 4, pp. 5~11
- [3] Warren Chung, Ghassan Hamarneh, "N-SIFT: N-DIMENSIONAL SCALE INVARIANT FEATURE TRANSFORM FOR MATCHING MEDICAL IMAGES", 2007 IEEE International Symposium on Biomedical Imaging. April 12-15, 2007.
- [4] D. G. Lowe. "Object recognition from local scale invariant features," In Proceedings of the Seventh International Conference on Computer Vision (ICCV'99), pages 1150-1157, Kerkyra, Greece, September 1999.
- [5] Y. Deng and B. S. Manjunath, "Unsupervised segmentation of color-texture regions in images and video," IEEE Trans. on PAMI, Vol. 23, No. 8, pp. 800~810, 2001.
- [6] 박찬일, 이수현 정용진, "SIFT알고리즘의 실시간처리를 위한 하드웨어 모듈 검증," 제21회 신호처리학술대회 논문집, Vol.21, No 1, 81p, 2008.