특집논문 (Special Paper)

방송공학회논문지 제18권 제4호, 2013년 7월 (JBE Vol. 18, No. 4, July 2013)

<http://dx.doi.org/10.5909/JBE.2013.18.4.535>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

지능형 자동차용 고성능 영상인식 엔진

여 준 기^{a)*}, 천 익 재^{a)}, 석 정 희^{a)}, 노 태 문^{a)}

High-Performance Vision Engine for Intelligent Vehicles

Chun-Gi Lyuh^{a)*}, Ik-Jae Chun^{a)}, Jung-Hee Suk^{a)}, and Tae Moon Roh^{a)}

요 약

본 논문에서는 고속 및 고인식률의 성능을 갖는 영상인식 엔진 구조를 제안한다. 본 엔진은 2단계의 특징점 추출 및 분류 알고리즘을 수행하여 자동차와 보행자를 인식할 수 있다. 엔진의 인식률을 높이기 위해 HOG 특징점 값과 LBP 특징점 값을 같이 사용하여 알고리즘을 구성하였으며, 처리 속도를 높이기 위해 병렬 구조를 개선하여 하드웨어를 설계하였다. 실험결과를 통해 설계한 엔진이 초당 90프레임의 인식 처리가 가능하며 FPPW 10^{-4} 하에서 97.7%의 보행자 인식률을 가짐을 보인다.

Abstract

In this paper, we proposed a advanced hardware engine architecture for high speed and high detection rate image recognitions. We adopted the HOG-LBP feature extraction algorithm and more parallelized architecture in order to achieve higher detection rate and high throughput. As a simulation result, the designed engine which can search about 90 frames per second detects 97.7% of pedestrians when false positive per window is 10^{-4} .

Keyword : intelligent vehicle, vision engine, vehicle recognition, pedestrian recognition

I. 서 론

최근 북미에서 후방 카메라 탑재 의무화가 발표되는 등,

자동차 분야에서의 안전시스템에 대한 관심과 요구가 커져가고 있으며, 지능형 자동차를 위한 자동차 및 보행자 인식과 관련하여 영상을 기반으로 한 다양한 연구들이 활발히 진행되고 있다. [1]은 부분적으로 가려진 보행자에 대한 알고리즘 연구를 수행하였으며, [2]에서도 보행자에 대한 인식을 개선을 위한 연구를 수행하였다. [3], [4], [5], [6] 등에서는 처리 속도를 높이기 위한 알고리즘 개선 연구를 수행하였으나, 이 연구들은 PC상에서의 알고리즘 수행에 대한 연구들이다. 영상인식 알고리즘의 임베디드 시스템 적용을 위한 하드웨어 설계와 관련하여서는 [7], [8]과 같은 연구들이 있으나 인식률과 전체 인식 하드웨어에 대한 통합에 대

a) 한국전자통신연구원 혼성신호처리연구실 (Electronics and Telecommunications Research Institute, Mixed Signal Processing Research Section)

* Corresponding Author : 여준기(Chun-Gi Lyuh)
E-mail: cglyuh@etri.re.kr
Tel: +82-42-860-1248

※ 본 논문은 지식경제부 및 한국산업기술평가관리원의 산업원천기술개발사업(정보통신)의 일환으로 수행하였음.[K1002162, 다중카메라 기반 고속 영상인식 SoC 플랫폼]

※ 이 논문의 연구결과 중 일부는 'IPIU 2013'에서 발표한 바 있음.

· Manuscript received May 13, 2013 Revised July 11, 2013 Accepted July 12, 2013

(a) EyeQ

(b) EyeQ2

그림 1. Mobileye사 EyeQ와 EyeQ2의 블록도

Fig. 1. Block diagrams of EyeQ and EyeQ2 of Mobileye

해서는 다루지 않고 있다.

기존에 연구되어 오던 자동차, 보행자 인식 알고리즘들이 시스템에 적용되어 개발된 것으로는 대표적으로 Mobileye사의 EyeQ 시리즈^[9]와 Renesas사의 IMAPCAR 시리즈^[10]가 있다. 그림 1은 EyeQ 시리즈의 블록도를 보여준다. EyeQ 시리즈는 인식전용엔진을 기반으로 영상인식

알고리즘을 수행한다. 먼저 그림 1(a)는 EyeQ 시리즈의 첫 제품인 EyeQ의 블록도를 보여주며, 영상인식을 위한 전용 엔진인 VCE(Vision Computing Engine) 코어 4개를 탑재하고 있음을 알 수 있다. 이 후, 그림 1(b)와 같이 영상인식 전용엔진의 개선과 함께 병렬 프로세서인 VMP(Vector Microcode Processor)를 추가하고, 클럭 속도를 향상시킨

그림 2. NEC사 IMAPCAR2의 블록도

Fig. 2. Block diagram of IMAPCAR2 of NEC

EyeQ2가 개발되었다.

그림 2는 IMAPCAR 시리즈의 두 번째 버전인 IMAPCAR2의 블록도를 보여준다. IMAPCAR 시리즈는 인식전용엔진을 기반으로 한 EyeQ 시리즈와는 병렬 프로세서를 기반으로 대용량 연산을 빠르게 수행하도록 함으로써 영상인식 알고리즘을 수행하도록 되어 있다. IMAPCAR는 SIMD (Single Instruction Stream Multiple Data Stream) 제어 방식의 병렬 프로세서로 구성되어 있으며, IMAPCAR2에서는 이 병렬 프로세서가 MIMD(Multiple Instruction Stream Multiple Data Stream) 제어 방식으로 전환이 가능하도록 개선하여 그 연산 효율을 높였다.

하지만, EyeQ 시리즈와 IMAPCAR 시리즈의 자료들에서는 처리속도와 인식률, 알고리즘의 처리 구조 등과 관련한 구체적인 사항들이 포함되어 있지 않아, 영상인식을 위한 임베디드 시스템 연구 및 구현에 활용되기는 어려운 점이 있다. 본 논문에서는 기존에 자동차에서의 응용을 고려하여 자동차와 보행자에 대해 높은 인식률과 함께 빠른 처리가 가능하도록 개발하였던 ETRI Vision Engine

(EVE)^[11]을 개선한 결과를 보임으로써 지능형 자동차용 영상인식 임베디드 시스템 연구 및 개발의 바탕을 제공하고 자 한다.

II. ETRI Vision Engine version 2(EVE2)

기존 개발된 EVE는 Haar 특징점 추출 및 Adaboost 특징점 분류 알고리즘^[12]과 HOG (Histograms of Oriented Gradients) 특징점 추출 및 SVM(Support Vector Machine) 특징점 분류 알고리즘^{[13][14]}을 기반으로 영상인식 처리를 수행한다. 칩으로 제작한 결과 200MHz로 동작하면서 초당 45프레임의 VGA 크기 영상을 처리할 수 있으며, FPPW (False Positive Per Window)가 10^{-4} 인 기준에서 자동차는 99.3%, 보행자는 91.7%의 인식률을 가지고 있다. 하지만 EVE는 자동차와 보행자에 대해 동시에 처리가 불가능하여, 자동차와 보행자를 모두 인식하기 위해서는 인식 알고리즘 수행을 위한 학습결과 데이터를 매번 재설정해야 하

그림 3. EVE와 EVE2의 영상인식 알고리즘 수행 구조 비교
Fig. 3. Comparison of Vision algorithm processes of EVE and EVE2

는 단점이 있었다. 또한 45프레임의 처리만으로는 두 개 이상의 카메라에 대한 실시간 처리가 불가능하다. EVE2에서는 이러한 EVE의 단점을 개선하여 그 처리속도와 인식률을 높이고, 2중 대상 동시 인식을 가능하게 하였다.

1. 고속 병렬처리 영상인식 엔진 구조

그림 3은 EVE를 개선하여 개발된 EVE2(ETRI Vision Engine version 2)의 영상인식 알고리즘 수행 구조를 EVE의 수행 구조와 비교하여 보여준다. EVE2는 EVE에서의 다소 낮은 보행자 인식률의 개선을 위하여 인식대상 판별 과정 중 2차 인식 알고리즘 수행에서는 [1]의 논문과 같이 HOG 특징점 추출과 함께 LBP(Local Binary Pattern) 특징점 추출 알고리즘을 동시에 사용하였다. 또한 자동차와 보행자를 동시에 인식하기 위해서 보행자 판별 하드웨어와 자동차 판별 하드웨어를 병렬로 구성하였다. 다만 영상 크기 변환, 인식 윈도우 추출은 자동차와 보행자에 동일하게 사용될 수 있으므로 하나의 하드웨어로 구성하였다.

그림 4는 EVE2의 하드웨어 구조를 보여 준다. 기존 EVE와의 가장 큰 차이는 LBP 특징점 추출 알고리즘에 대한 처리를 가능하게 한 것이다. 이를 위해 추가된 블록은 그림에서 굵은 실선으로 표시하였다. LBP 특징점 추출과 관련

된 하드웨어는 HOG 특징점 추출 관련 하드웨어와 병렬로 수행되도록 구성하였으며, 각 SVM 분류기의 결과는 제어기에서 한꺼번에 판단하여 HOG-LBP 알고리즘 결과를 도출할 수 있도록 하였다. 이 뿐만 아니라 각 SVM 분류기의 결과를 개별적으로 최종 판단에 사용함으로써 LBP 알고리즘을 Haar-Adaboost, HOG-SVM에 이은 3차 단계의 인식 알고리즘으로도 수행할 수 있도록 함으로써 알고리즘 구성의 유연성을 부여하였다. 자동차와 보행자의 동시인식 처리를 위해서는 처리속도 저하를 막기 위해 자동차와 보행자에 대한 각 특징점 추출 알고리즘 및 분류 알고리즘 처리 하드웨어를 구분하여 구성하였다. 다만 영상 축소기와 Haar 윈도우 버퍼는 자동차와 보행자에 대한 구분 없이 하나로 구성하였다.

EVE2에서는 자동차 인식을 위한 윈도우 크기는 32x32 픽셀, 보행자 인식을 위한 윈도우 크기는 32x64 픽셀을 사용하였다. 이렇게 구성할 경우 더 큰 윈도우 크기에 비해 인식률은 다소 저하되지만, 보다 먼 거리의 대상을 인식할 수 있다는 장점이 있으며, 하드웨어 절감에도 유리하다. 예를 들어 가로화각이 35도인 VGA 카메라를 사용하였을 경우 39m 거리에 있는 185cm 키의 보행자는 영상에서 48 픽셀의 높이로 나타난다.

영상 축소기는 외부 메모리에 저장된 원본영상을 32x480

그림 4. EVE2 하드웨어 구조

Fig. 4. Hardware architecture of EVE2

픽셀 단위로 읽어 64x480 픽셀 크기의 메모리에 저장하여 이를 20% 단계로 축소하는 과정을 거치며 각 축소된 이미지는 영상 축소기 내부에 각각 저장되도록 하였다. 그리고 이러한 원본크기의 이미지와 축소된 이미지는 윈도우 버퍼로 이동되면서 인식에 활용된다. 이러한 구조를 통해 EVE2는 입력영상의 가로폭에 대한 제한이 없이 인식이 가능하며, 다수개의 카메라 영상을 합성한 파노라마 영상에 대한 인식 등에도 활용이 가능하다.

Haar 특징점 추출을 위해서 Haar 윈도우 버퍼에는 누적영상이 계산되어 저장된다. 이 때 4x4 픽셀 블록에 대한 누적영상(Integral Image)을 사용하는 경우 보행자 인식을 위한 윈도우의 누적영상은 8x16 크기가 사용되지만, 다음 위치의 윈도우에 대해 연산을 수행하기 위해서는 이전 픽셀의 누적값을 저장하고 있어야 하며, 이를 위해 Haar 윈도우 버퍼는 8x17로 설계되었다. 또한 이 Haar 윈도우 버퍼는 자동차 인식과 보행자 인식에 동시에 사용된다. Haar 특징점 추출기는 Adaboost 메모리에서 각 Haar 필터의 크기, 위치, 종류에 대한 정보를 따라 연산을 수행하며, Adaboost 분류기는 Haar 필터 결과를 가지고 해당하는 약분류기 연산을 수행하며 이를 누적하여 최종 결과를 출력한다.

Adaboost 분류기에서 최종 강분류기의 결과에서 참으로

판단된 윈도우 정보는 태스크 큐에 저장되어, 해당 윈도우의 HOG-LBP 알고리즘 처리와 상관없이 Haar-Adaboost 처리가 진행될 수 있도록 하였다.

HOG 윈도우 버퍼는 4x4 픽셀 크기의 셀을 2x2 형태로 구성한 블록을 하나의 단위로 저장하기 위하여 12x12 픽셀 크기로 구성하였으며, 인접한 블록의 처리에 데이터를 재 활용할 수 있도록 설계되었다. HOG 특징점 추출기는 HOG 윈도우 버퍼에 저장된 블록에 대해 HOG 특징점 추출 연산을 수행하고, 이에 대해 HOG용 SVM 분류기는 SVM 메모리에 저장된 가중치 값을 읽어 가중합 연산과 함께 모든 블록에 대한 누적 연산을 수행하는 방법으로 최종 SVM 결과를 얻는다. LBP관련 블록은 HOG 블록과 유사하게 구성되어 있으며, 하나의 블록크기는 8x8로 구성하였다.

태스크 큐의 사용과 유사하게 HOG-LBP의 최종 결과가 참인 윈도우 정보는 인식결과 큐에 저장되어 외부의 입출력이 원활하게 수행될 수 있도록 설계하였다.

2. 특징점 추출 및 분류 하드웨어 최적화

LBP 알고리즘 및 자동차와 보행자의 동시인식을 가능하게 하기 위해서는 연산 및 제어 로직뿐만 아니라 Adaboost

그림 5. 정규화 연산 지연 하드웨어 최적화
Fig. 5. Hardware optimization by delayed normalization

메모리, SVM 메모리 등이 추가되어야 하기 때문에 하드웨어의 크기에 대한 부담이 커지게 된다. 이러한 부담을 줄이고자 특징점 추출 및 분류 알고리즘을 수행하는 하드웨어를 최적화하였다.

HOG 특징점 추출 알고리즘에는 각 블록에 대해 값을 정규화 하는 과정이 포함되어 있다. 본 엔진에 사용된 HOG 알고리즘에서는 블록당 36개의 특징값이 만들어 지며, 이 값을 정규화 한 후에 SVM 특징점 분류를 위한 가중합 연산을 수행하게 된다. 하지만 이 정규화 연산을 그림 5와 같이 SVM의 가중합 연산 이후로 옮기게 되면 수행되는 연산의 수를 줄일 수 있으며, 이를 통한 하드웨어 면적 절감이 가능하다. 또한 블록별로 만들어지는 36개의 특징값을 동시에 처리하던 기존과는 달리 사이클당 9개씩 생성하여 연산함으로써 지연시간의 추가 없이 하드웨어를 절감하였다. 이러한 방법은 LBP 특징점 추출기 및 분류기에도 적용하여 설계하였다.

III. 실험 결과 및 분석

EVE2는 Verilog HDL(Hardware Description Language)를 사용하여 구현되었으며, Global Foundry 65nm CMOS LPe 공정 표준 셀 라이브러리를 사용하여 Synopsys DesignCompiler에서 합성되었다. 이 때 사용한 클럭 속도는 333MHz이며, 실제 칩으로 제작될 경우 200MHz 수준의 클럭 속도로 동작이 가능하다. 다음 표 1은 기존 EVE와의 합성 결과를 비교하여 보여준다.

표 1. EVE2 인식엔진 합성 결과 비교

Table 1. Comparison of Synthesis Result

Name	EVE	EVE2
Synthesis tool	Synopsys DC Ultra	
Process	GF 65nm CMOS LPe	
Clock speed	333MHz	
Gate count with memory	1,492,877	2,075,856
Memory size	90.7KB	94.7KB

EVE2의 인식 처리속도 분석 과정에서는 엔진 자체의 개선 결과와 함께 시스템 상에서의 DMAC(Direct Memory Access Controller) 일부 기능의 개선을 반영하였으며, 이를 바탕으로 시뮬레이션 방법을 사용하였다. 이를 통해 EVE2는 초당 90 프레임의 VGA 영상을 처리할 수 있음을 확인하였다. 이는 카메라 3개의 입력에 대해 실시간으로 처리 가능한 수준이다.

보행자에 대한 인식을 분석을 위해서 INRIA 보행자 DB^[14]를 사용하였다. 그림 6은 기존 EVE와 EVE2의 보행자 인식을 성능을 비교하여 보여준다. 그림에서 가로축은 FPPW(False Positive Per Window), 세로축은 MR(Miss Rate)을 나타낸다. 그림에서와 같이 FPPW 10^{-4} 의 기준에서 보행자를 97.7%로 인식이 가능하였다.

그림 6. EVE2 인식엔진의 보행자 인식 성능

Fig. 6. Pedestrian recognition detection performance of EVE2

IV. 결론

본 논문에서는 지능형 자동차용 영상인식 임베디스 시스템 연구 개발의 바탕을 제공하고자 기존 개발된 영상인식 엔진을 개선하여 적은 하드웨어 부담으로 인식률과 처리속도를 높이고, 자동차와 보행자를 동시에 인식할 수 있게 한 하드웨어 엔진의 개발 결과를 소개하였다. 개발된 엔진은 LBP 특징점 추출 알고리즘을 HOG 특징점 추출 알고리즘과 동시에 처리하는 구조를 채택하였으며 이를 통하여

보행자를 97% 이상 인식할 수 있다. 또한 시스템적인 개선을 포함하여 초당 90프레임의 인식 처리가 가능하게 되었다.

향후에는 엔진 자체의 알고리즘 수행 외에 전처리와 후처리를 통한 엔진의 인식률 보완 및 처리 속도 개선, 보다 다양한 알고리즘의 빠른 적용을 가능하게 하는 병렬 프로세서 혼용구조에서의 알고리즘 구현에 대한 연구를 수행하고자 한다.

참 고 문 헌

- [1] X. Wang, T.X. han, and S. Yan, "An HOG-LBP human detector with partial occlusion handling," International Conference on Computer Vision, 32-39, 2009.
- [2] G. Gan and J. Cheng, "Pedestrian Detection Based on HOG-LBP Feature," International Conference on Computational Intelligence and Security, 1184-1187, 2011.
- [3] C. Zeng, H. Ma, and A. Ming, "Fast human detection using mi-SVM and a cascade of HOG-LBP features," IEEE International Conference on Image Processing, 3845-3848, 2009.
- [4] S.-H. Yoon, C.-G. Lyuh, I.-J. Chun, J.-H. Suk, and T. M. Roh, "A New Support Vector Compression Method Based on Singular Value Decomposition", ETRI Journal, vol. 33, no. 4, 652-655, Aug, 2011.
- [5] W.-J. Park, D.-H. Kim, S. Suryanto, C.-G. Lyuh, T. M. Roh, and S.-J. Ko, "Fast human detection using selective block-based HOG-LBP," IEEE International Conference on Image Processing, 601-604, 2012.
- [6] H. Cho, P.E. Rybski, A. Bar-Hillel, and W. Zhang, "Real-time pedestrian detection with deformable part models," IEEE Intelligent Vehicles Symposium, 1035-1042, 2012.
- [7] S. Lee, H. Son, J. C. Choi, and K. Min, "HOG feature extractor circuit for real-time human and vehicle detection," IEEE Region 10 Conference, 2012.
- [8] S. Kim, S. Park, S. Lee, S. park, and K. Cho, "Design of high-performance pedestrian and vehicle detection circuit using Haar-like features," IEEE Region 10 Conference, 2012.
- [9] <http://www.mobileye.com>
- [10] <http://www.nec.com>
- [11] C.-G. Lyuh, I.-J. Chun, J.-H. Suk, S.-H. Yoon, and T. M. Roh, "High-Speed Vision Engine for Intelligent Vehicles," Workshop on Image Prodesing and Image Understanding, 2012.
- [12] P. Viola and M. J. Jones, "Robust Real-Time Face Detection," International Journal of Computer Vision, vol. 57, no. 2, pp. 137-154, 2004.
- [13] N. Dalal, "Finding People in Images and Videos," PhD thesis, Institut National Polytechnique de Grenoble, 2006.
- [14] N. Dalal and B. Triggs, "Histogram of oriented gradient for human detection," IEEE Computer Vision and Pattern Recognition, 886-893, 2005.

저 자 소 개

여 준 기

- 1998년 2월 : 경북대학교 컴퓨터공학과 학사
- 2000년 2월 : 한국과학기술원 전자전산학과 전산학전공 석사
- 2004년 2월 : 한국과학기술원 전자전산학과 전산학전공 박사
- 2004년 5월 ~ 현재 : 한국전자통신연구원 선임연구원
- 주관심분야 : 컴퓨터비전, 재구성형 프로세서, 저전력 디지털 신호처리

천 익 재

- 1998년 2월 : 충남대학교 전자공학과 학사
- 2000년 2월 : 충남대학교 전자공학과 석사
- 2006년 2월 : 충남대학교 전자공학과 박사
- 2006년 9월 ~ 현재 : 한국전자통신연구원 선임연구원
- 주관심분야 : 디지털신호처리, 멀티미디어 SoC, 컴퓨터비전, 통신모델

저 자 소 개

석 정 희

- 2001년 2월 : 경북대학교 전자전기공학부 학사
- 2003년 2월 : 경북대학교 전자공학 석사
- 2007년 2월 : 경북대학교 전자공학 박사
- 2007년 2월 ~ 현재 : 한국전자통신연구원 선임연구원
- 주관심분야 : 컴퓨터비전, H.264 codec, 디지털 신호처리

노 태 문

- 1984년 2월 : 경북대학교 전자공학과 학사
- 1986년 8월 : 경북대학교 전자공학과 석사
- 1998년 2월 : 경북대학교 전자공학과 박사
- 1988년 2월 ~ 현재 : 한국전자통신연구원 실장(책임연구원)
- 주관심분야 : 영상인식, 디지털신호 처리, SoC 설계, 고신뢰 아날로그회로, ROIC, CMOS 소자/공정 및 신뢰성