

논문 2011-48SD-9-4

# 지능형 영상 보안 시스템의 얼굴 인식 성능 향상을 위한 얼굴 영역 초해상도 하드웨어 설계

( Hardware Design of Super Resolution on Human Faces for  
Improving Face Recognition Performance of Intelligent Video  
Surveillance Systems )

김 초 룡\*, 정 용 진\*\*

( Chorong Kim and Yong-Jin Jeong )

## 요 약

최근 카메라를 통해 입력된 영상정보로부터 실시간으로 상황을 인지하고 자율 대응할 수 있는 지능형 영상 보안 시스템의 수요가 증가함에 따라, 고성능의 얼굴 인식 시스템이 요구되고 있다. 기존의 얼굴 인식 시스템의 성능 향상을 위해서는 원거리에서 획득된 저해상도 얼굴 영상 처리를 위한 솔루션이 반드시 필요하다. 따라서 본 논문에서는 실시간 감시가 요구되는 지능형 영상 보안 시스템의 얼굴 인식 성능 향상을 위한 저해상도 얼굴 영상 복원 알고리즘을 하드웨어로 구현하였다. 저해상도 얼굴 영상 복원 방법으로는 학습 기반의 초해상도 알고리즘을 사용한다. 해당 알고리즘은 먼저 고해상도 영상으로 구성된 학습 집합에서 주성분 분석(PCA)을 활용하여 복원에 필요한 사전 정보들을 추출하고, 저해상도 영상과의 관계를 모델링하여 가장 적합한 고해상도 얼굴을 복원해내는 것이다. 저해상도 얼굴 영상 복원 알고리즘을 임베디드 프로세서(S3C2440A)를 사용하여 구현하였을 때, 약 25 초의 긴 연산 시간이 소요되었다. 이는 실시간으로 사람을 판별 및 인식하기 위한 지능형 영상 보안 시스템의 구축에는 어려움이 있다. 이를 해결하기 위하여 얼굴 영역 초해상도의 연산을 하드웨어로 구현하고 Xilinx Virtex-4를 이용하여 검증하였다. 약 9MB의 학습 데이터를 사용하였으며, 100 MHz에서 약 30 fps의 속도로 연산이 가능하다. 이러한 학습 기반의 얼굴 영역 초해상도 알고리즘을 단일 하드웨어 IP로 설계함으로써 임베디드 환경에서의 실시간 처리가 가능할 뿐만 아니라 기존의 다양한 얼굴 검출 시스템과의 통합이 용이하여 얼굴 인식 솔루션을 제공할 수 있을 것으로 판단된다.

## Abstract

Recently, the rising demand for intelligent video surveillance system leads to high-performance face recognition systems. The solution for low-resolution images acquired by a long-distance camera is required to overcome the distance limits of the existing face recognition systems. For that reason, this paper proposes a hardware design of an image resolution enhancement algorithm for real-time intelligent video surveillance systems. The algorithm is synthesizing a high-resolution face image from an input low-resolution image, with the help of a large collection of other high-resolution face images, called training set. When we checked the performance of the algorithm at 32bit RISC micro-processor, the entire operation took about 25 sec, which is inappropriate for real-time target applications. Based on the result, we implemented the hardware module and verified it using Xilinx Virtex-4 and ARM9-based embedded processor(S3C2440A). The designed hardware can complete the whole operation within 33 msec, so it can deal with 30 frames per second. We expect that the proposed hardware could be one of the solutions not only for real-time processing at the embedded environment, but also for an easy integration with existing face recognition system.

**Keywords :** Super Resolution, Face Hallucination, Face resolution enhancement, Embedded system

\* 학생회원, 광운대학교 임베디드소프트웨어공학과

(Department of Embedded Software Engineering, Kwangwoon University)

\*\* 정회원, 광운대학교 전자통신공학과

(Department of Electronics and Communications Engineering, Kwangwoon University)

※ 이 논문은 2011년 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임.(NRF-2010-0014557)  
접수일자: 2010년10월26일, 수정완료일: 2011년8월17일

## I. 서 론

초기의 영상 보안 기술은 카메라를 통하여 획득된 영상을 전달 받아 사람이 직접 감시하거나, 범죄가 이미 발생한 후의 저장된 영상을 분석하여 대응하는 수동적인 형태였다. 반면, 최근에는 IT 기술의 급격한 발전에 힘입어 입력된 영상 분석을 통해 실시간으로 범죄 상황을 인식하고 대응할 수 있는 지능형 영상 보안 시스템에 대한 요구가 꾸준히 증가하고 있다.<sup>[1]</sup>

특히, 기존의 보안 시스템에 접촉이나 강제성이 없는 얼굴 인식 기술을 접목하여, 원거리에서 용의자 또는 미아와 같이 특정 인물을 찾고 추적이 가능한 지능적인 보안 시스템에 대한 연구가 활발히 진행 되고 있다.

얼굴 인식의 경우 표정이나 자세, 조명의 변화에 따라 얼굴 영상의 차이가 크기 때문에 처리가 쉽지 않다. 뿐만 아니라, 먼 거리에서 촬영되어진 저해상도의 얼굴 영상일 경우 인식률이 크게 떨어진다. 따라서 저해상도 얼굴 영상의 처리가 얼굴 인식 시스템의 전체 성능을 좌우한다고 할 수 있다.<sup>[2~3]</sup>

저해상도 영상처리 문제의 해결을 위해, 다양한 초해상도(SR: Super Resolution)알고리즘의 연구가 진행되고 있다. SR 알고리즘은 저해상도 영상에 포함된 노이즈나 흐릿함의 제거뿐만 아니라 영상의 크기가 확대됨에 따라 기존에 존재하지 않던 영상 정보를 효과적으로 추정하여 보다 선명한 고해상도 영상을 복원하는 것을 목표로 한다.<sup>[4]</sup>

얼굴 영상을 대상으로 하는 초해상도 방법은 주로 미리 수집된 다수의 고해상도 얼굴 영상의 학습 집합(training set)을 이용한다. 학습 과정을 통해 얻어진 사전 정보는 저해상도 영상이 가지고 있는 고유 정보와 함께 고해상도 영상을 복원하기위해 사용된다.

본 논문에서는 한 장의 저해상도 얼굴 영상으로부터 고해상도 얼굴 영상을 복원하기 위한 알고리즘을 구현하여 분석하고, 그 결과를 바탕으로 임베디드 환경에 적합한 형태의 하드웨어 구조를 설계하고 검증하였다. 임베디드 환경을 위한 저해상도 얼굴 영상 복원 기술의 구현은 기존의 얼굴 인식 시스템의 성능 향상에 아주 중요한 역할을 할 수 있다.

본 논문의 구성은 다음과 같다. 먼저 II장에서 지능형 보안 시스템에 대한 설명과 학습 알고리즘 기반의 얼굴 영역 초해상도 방법을 소개한다. III장에서는 임베디드 환경에서의 실시간 동작을 위한 얼굴 영역 초해상

도 모듈의 하드웨어 구조 및 설계에 대해 설명한다. 그 다음, 설계한 모듈을 FPGA에서 검증한 결과 및 성능에 대한 분석을 제시한다. 마지막으로 IV장에서는 결론을 맺는다.

## II. 지능형 영상 보안 시스템을 위한 얼굴 영역 초해상도 복원 기술

### 1. 얼굴 인식 기반의 지능형 영상 보안 시스템

지능형 영상 보안 시스템의 가장 기본적인 목표는 카메라를 통해 획득된 영상 정보를 실시간으로 분석하여 자동으로 지정된 대상을 탐지, 추적, 식별하여 관찰된 객체의 행위나 상호 작용을 해석하는 것이다.

지능형 영상보안 시스템은 특정 응용영역에 관계없이 공통적으로 필요한 움직임 검출(change detection), 영역검출(region localization), 영역추적(region tracking) 모듈들과 응용에 따라 기능의 변화가 필요한 객체 분류(object classification), 객체 인식 및 추적(object identification and tracking)등 가변적 모듈들로 구분할 수 있다.<sup>[1,5]</sup>

이를 바탕으로 얼굴 인식 기능을 응용한 지능형 영상 보안 시스템을 그림 1과 같이 모델링 할 수 있다. 가장 기본적으로는 입력된 영상으로부터 얼굴을 관심 영역으로 추출하는 얼굴 검출모듈(그림1-(a))과 사전에 분류



(a) 얼굴 검출 모듈(ace detection module)  
 (b) 저해상도 영상 개선 모듈 (resolution enhancement module)  
 (c) 얼굴 인식 및 식별 모듈(face recognition module)

그림 1. 얼굴 인식 기능을 응용한 지능형 영상 보안 시스템 모델

Fig. 1. A model of intelligent video surveillance systems using face recognition.

및 정의된 얼굴 객체 모델들을 이용하여 대상을 분류하고 검출된 얼굴 영상이 누구인지 판별하는 인식모듈(그림1-(c))이 있다. 여기에, 영상 보안 시스템의 특성상 주로 원거리에서 촬영되는 저해상도 얼굴 영상에 대한 영상 처리 모듈(그림1-(b))이 인식 시스템의 성능 향상을 위해 반드시 필요하다.

다음 2절에서는 저해상도 얼굴 영상의 개선 방법 중 학습 기반의 초해상도 방법에 대해 소개한다.

## 2. 얼굴 영역 초해상도 복원 알고리즘

### 가. 학습 알고리즘의 개념

고해상도 얼굴 영상 복원을 위한 학습 방법으로 주성분 분석(PCA: Principle Component Analysis)을 이용한다. PCA를 이용하여 얼굴 집합으로부터 모든 얼굴 영상의 평균적인 특징을 가장 잘 나타내는 기저 벡터, 즉 고유 얼굴(eigenfaces)을 찾아낸다. 여기서, 얼굴 영상의 복원은 추출된 기저 벡터들이 구성하는 얼굴 공간(eigenspace)안의 모든 얼굴 영상은 이 고유 얼굴들의 선형 결합으로 표현될 수 있다는 점을 이용한다. 다음 식(1)은 PCA 주요 개념을 수식으로 나타낸 것이다.

$$\Delta\Gamma = \Gamma - \Psi \cong \omega_1 e_1 + \omega_2 e_2 + \dots + \omega_N e_N = \sum_{i=1}^N \omega_i e_i$$

$$\Gamma \cong \sum_{i=1}^N \omega_i e_i + \Psi \tag{1}$$

입력 얼굴 영상( $\Gamma$ )은 공분산 행렬(covariance matrix)에서 얻어진 고유 얼굴( $e_i$ )의 선형 결합으로 분해될 수 있다. 공분산 행렬은 각 얼굴 영상에서 평균 얼굴( $\Psi$ )을 뺀 차영상( $\Delta\Gamma$ )을 이용하여 구해진다. 이 때, 고유 얼굴은 학습 데이터의 수만큼 추출될 수 있지만 데이터의 수가 너무 많을 경우에는, 추출된 고유 얼굴 중 많은 얼굴 특징을 표현하는 N개를 선택하여 표현할 수 있다.<sup>[7]</sup>

그림 2는 공개된 얼굴 데이터베이스 영상<sup>[9~10]</sup>에서 데이터를 추출하여, 식(1)을 실제 얼굴 영상으로 표현한 것이다. 좌측의 입력 얼굴 영상은 우측의 학습을 통해 출력된 고유 얼굴의 선형 결합과 평균 얼굴 영상의 합으로 재구성될 수 있다.

### 나. 학습 기반 저해상도 얼굴 영상의 복원 방법

앞의 학습 알고리즘의 기본 개념을 바탕으로 한 장의 저해상도 얼굴 영상으로부터 고해상도 얼굴을 추정하는 것이 가능하다.

입력된 저해상도 얼굴 영상도 마찬가지로 저해상도 고유 얼굴의 선형 결합으로 재구성 할 수 있다. 이 과정에서 입력 영상의 특징 값인 각 고유 얼굴들의 계수( $\omega_i$ )를 얻는다. 이 계수를 다시 고해상도 고유 얼굴들에 적용하여 결합하면, 고해상도 영상을 얻을 수 있다.

C. Liu 등<sup>[6~7]</sup>은 최대 사후 확률(MAP: Maximum A Posterior)을 이용하여, 입력된 저해상도 영상( $\Gamma_L$ )을 고해상도로 복원한 것과 가장 유사한 값을 갖는 최적의 고해상도 영상( $\Gamma_H$ )을 찾는 문제로 정의하였다.

$$\Gamma_H^* = \arg \max_{\Gamma_H} p(\Gamma_L | \Gamma_H) p(\Gamma_H) \tag{2}$$

복원 대상인 고해상도 영상( $\Gamma_H$ )을 식(3)과 같이 정의하고, 저해상도 영상으로부터 최적의 특징 값을 추출하기 위한 모델을 식(4)를 통해서 구하게 된다.

$$\Gamma_H = E\Omega + \Psi$$

$$E = (e_1, e_2, \dots, e_N), \Omega = (\omega_1, \omega_2, \dots, \omega_N)^T \tag{3}$$

$$\Omega^* = (E_L^T E_L + \sigma^2 \Lambda^{-1})^{-1} E_L^T (\Gamma_L - \Psi_L) \tag{4}$$

고유 얼굴, 고유 값 그리고 평균 얼굴이 모두 고해상도 얼굴 영상을 학습시켜서 얻은 사전 정보에 해당한다. 또한,  $E_L^T E_L + \sigma^2 \Lambda^{-1}$ 의 역행렬은 특이값 재구성



그림 2. 주성분 분석을 통해 추출된 학습데이터를 이용하여 재구성된 입력 영상  
 Fig. 2. Reconstructed face image using training data extracted by PCA.

(SVD: Singular Vector Decomposition)으로 구한다. 최종적으로 식(5)를 통해 고해상도 얼굴을 복원한다.

$$\Gamma_H^* = E\Omega^* + \Psi \quad (5)$$

다. 초해상도 복원 알고리즘의 성능 검증

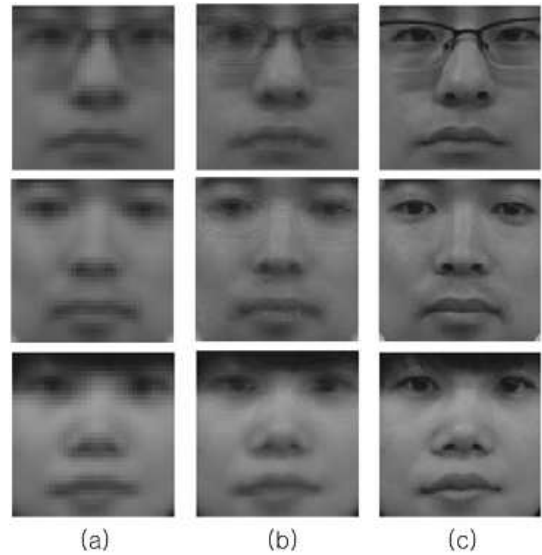
본 논문에서는 학습 기반 얼굴 영역 초해상도 복원 알고리즘 검증에 필요한 얼굴 데이터베이스를 직접 구축하였다. 공개된 얼굴 데이터베이스<sup>[9~10]</sup>들은 주로 외국인으로 구성된 것이 많아, 학습 집합의 내용에 큰 영향을 받는 알고리즘의 특성상 동양인 얼굴의 데이터베이스 구축이 필요하기 때문이다. 먼저 각각의 사람에 대해 그림 3과 같이 감정 변화에 따른 표정, 빛의 방향, 안경 등의 액세서리 착용에 따라 달라 질 수 있는 얼굴 영상을 실내에서 디지털 카메라로 촬영했다. 그 다음, 얼굴의 특징 영역인 눈, 코, 입을 기준 점으로 정하여 정렬하였다. 만약, 정렬과정이 생략될 경우 일그러지거나 사람의 얼굴이라고 판단할 수 없는 잘못된 결과를 얻게 된다.

입력 저해상도 영상은 얼굴 검출 알고리즘<sup>[8]</sup>을 통해 그림 4의 (a)와 같이 얻어졌다고 가정한다. 검증을 위한 저해상도 영상 (a)는 고해상도 영상을 다운샘플링(down-sampling)하고 블러링(blurring)하여 실제 카메라를 통해 얻은 영상과 유사하도록 생성한다. 또한, 입력 얼굴 영상의 사이즈는 실제 얼굴 검출에서 사용되는 필터 사이즈 중 성능을 보장할 수 있는 가장 작은 필터 사이즈를 기준으로 하여 35×35로 정하였다.

고해상도 얼굴 영상의 사이즈는 육안으로도 충분히 인식이 가능한 70×70으로 정하고, 총 520개의 영상을 학습 집합으로 구성하였다. 모든 영상 데이터는 gray scale의 8bits 명도 영상을 사용하였다.



그림 3. RTA 얼굴 데이터베이스  
Fig. 3. RTA face database.



(a) 입력 저해상도 영상(input low-resolution image),  
(b) 초해상도 복원 결과 영상  
(reconstructed high-resolution image),  
(c) 원본 고해상도 영상(original high-resolution image)

그림 4. 저해상도 얼굴 영상 복원 결과 영상  
Fig. 4. Experimental results on reconstructing high resolution face.

표 1. 저해상도 얼굴 영상 복원의 RMS와 PSNR 결과  
Table 1. RMS and PSNR for face hallucination algorithm.

	Man 1	Man 2	Woman 1
RMS	9.20	7.59	6.80
PSNR(dB)	28.85	30.53	31.48

그림 4는 해당 알고리즘을 검증하여 얻은 결과 영상이고, 표 1은 결과 영상의 정량적인 성능평가를 위해 (b)와 (c)의 RMSE(Root Mean Square Error)와 PSNR(Peak signal to noise ratio) 수치를 나타낸 것이다. 입력 저해상도 영상 (a)로부터 복원된 고해상도 영상 (b)는, 원본 고해상도 영상 (c)와 비교하였을 때 표 1의 높은 PSNR 수치를 나타낸다. 또한 사람의 육안으로 확인했을 때도 거의 차이가 나지 않는다는 것을 확인할 수 있다.

III. 얼굴 영역 초해상도 복원 모듈의 설계 및 검증

1. 임베디드 프로세서를 이용한 구현

그림 5는 학습 기반 저해상도 얼굴 영상 복원 알고리즘을 나타낸 블록도이다. 얼굴 영상 학습 과정은 얼굴 영상 집합으로부터 학습 데이터를 추출하는 과정으로, 최초 한 번 실행하고 그 결과를 저장하여 사용하게 된

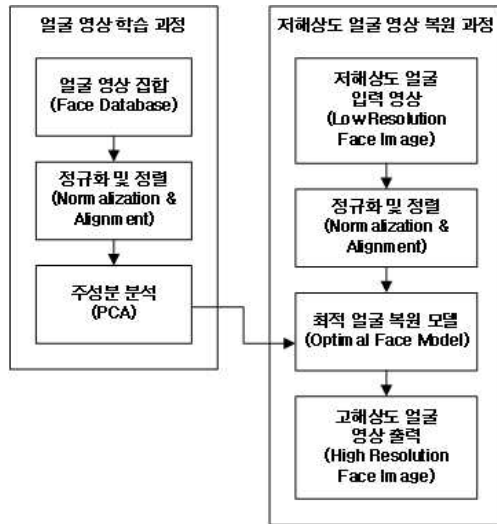


그림 5. 학습 기반 저해상도 얼굴 영상 복원 시스템 블록도

Fig. 5. Block diagram of the training-based super resolution on face image.

표 2. PC에서와 임베디드 환경에서의 연산 시간 비교

Tabel 2. Timing performance PC and ARM9-based embedded processor.

	PC	ARM 9	
	Intel® Core™ i5	ARM926EJ-S	S3C2440 (ARM920T)
연산 시간 (sec)	0.031	4.954	25.250

다. 저해상도 얼굴 영상 복원과정에서는 입력된 저해상도 얼굴 영상과 학습 데이터를 가지고 최적의 얼굴 모델을 연산하여, 최종적으로 고해상도 얼굴 영상을 복원한다. 이 과정은 저해상도 얼굴 영상이 입력될 때마다 연산이 일어나며, 실시간 동작이 가능하여야 한다.

구현을 고려한 입장에서 봤을 때, II장의 식 (4) 와 (5)의 행렬 연산을 풀어보면 지속적으로 업데이트 되는 저해상도 얼굴 입력 영상( $I_L$ )을 제외한 앞의 연산을 통합하여 학습데이터로 추출할 수 있다. 이 학습 데이터는 입력된 저해상도 영상의 해상도가  $35 \times 35$ 이고 고해상도 결과 영상의 해상도가  $70 \times 70$  일 때,  $(35^2 \times 70^2)$  사이즈의 32bit 부동 소수점 데이터를 갖는다.

얼굴 영역 초해상도 모듈의 전체 연산 중 가장 큰 부분을 차지하는 것은 32bit 부동소수점의 곱셈 및 덧셈 연산이  $(35^2 \times 70^2)$ 번 반복되는 부분이다.

표 2는 위의 소프트웨어로 구현된 저해상도 얼굴 영상 복원 알고리즘을 PC와 ARM9 기반의 임베디드 프로세서에서 실행하였을 때의 연산 시간을 측정된 것이

다. 이를 살펴보면, ARM920T기반의 32-bit RISC S3C2440A 프로세서를 사용하여 실행하였을 때 약 25초가 소모된다. 또 다른 종류의 ARM926EJ-S 프로세서에서도 약 5초의 연산 시간이 측정 되었다. 이는 임베디드 프로세서를 기반으로 하는 stand-alone 형태의 지능형 영상 감시 시스템에서 요구하는 실시간 감시 기능의 구현에 문제가 된다.

2. FPGA를 이용한 하드웨어 구현

임베디드 프로세서를 사용하여 구현된 얼굴 영역 초해상도 모듈의 느린 연산 속도는 얼굴 검출과 얼굴 인식을 포함한 전체 보안 시스템의 실시간 감시 기능 구현을 가로 막는다. 즉, 기존의 얼굴 검출 및 인식 기술과의 통합된 시스템 구성을 고려하여, 얼굴 영역 초해상도 알고리즘이 실시간으로 동작 할 수 있기 위해서는 임베디드 환경에서의 저해상도 얼굴 영상 복원 기술의 하드웨어 구현이 필요하다. 따라서 본 논문에서는 그림 6과 같은 하드웨어 설계 구조를 제안한다.

학습 기반의 저해상도 영상 복원 알고리즘은 미리 연산된 많은 양의 학습 데이터를 저장하고 있어야 하므로 외부 SDRAM을 사용한다. 때문에, 제안된 하드웨어 구조는 외부 SDRAM을 제어하기 위한 SDRAM 컨트롤러 부분을 포함하며, 저장된 학습 데이터를 이용하여 저해상도 얼굴 입력 영상으로부터 고해상도 얼굴 영상 복원 연산을 수행하는 FACE SR CORE 부분, 그리고 S3C2440A 프로세서와의 연결을 위한 INTERFACE 부분으로 구성된다.

FACE SR CORE의 구성을 살펴보면, 얼굴 영역 초해상도 알고리즘의 연산구조에 따라 크게 내부 메모리와 세 개의 모듈로 이루어져 있다.

내부 메모리의 경우, 평균 얼굴 데이터가 저장된

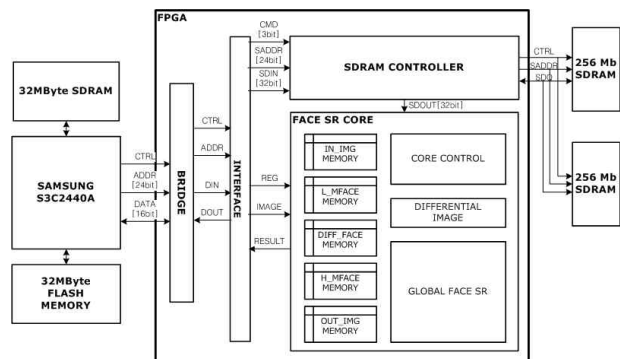


그림 6. FPGA를 이용한 하드웨어 구현

Fig. 6. Hardware Implementation on FPAG.

L\_MFACE와 H\_MFACE 메모리, 입력 영상이 저장되는 IN\_IMG 메모리, DIFFERENTIAL IMAGE 모듈의 중간 결과를 저장하는 DIFF\_FACE 메모리 그리고 최종 결과 영상이 저장되는 OUT\_IMG 메모리가 필요하다. 연산 모듈은 전체 동작을 컨트롤을 하는 CORE CONTROL 모듈과 입력된 저해상도 영상에서 평균 얼굴 영상을 제거하는 DIFFERENTIAL IMAGE 모듈 그리고 고해상도 영상을 복원하는 GLOBAL FACE SR 모듈로 구성되어 진다.

그림 7은 FACE SR CORE의 세부 구조도이다.

CORE의 실제 연산 동작에 앞서 S3C2440A 프로세서를 이용하여 최초의 한 번 미리 연산된 학습 데이터를 외부 SDRAM에 저장하고, 복원에 필요한 데이터인 저해상도 평균 얼굴 영상과 고해상도 평균 얼굴 영상의 데이터를 각각 L\_MFACE와 H\_MFACE 메모리에 입력한다. 저해상도 얼굴 영상이 IN\_IMG메모리에 저장이 된 후, CORE를 동작 시킨다.

CORE CONTROL 모듈은 외부 SDRAM 초기 셋업 및 DIFFERENTIAL IMAGE 모듈과 GLOBAL FACE SR 모듈의 동작을 제어한다.

DIFFERENTIAL IMAGE 모듈은 II장의 식 (4)에서

입력된 저해상도 영상( $T_L$ )과 저해상도 평균 얼굴 영상( $\Psi_L$ )의 차영상을 생성하는 부분에 해당하는 역할을 한다. 이 차영상은 DIFF\_FACE 메모리에 저장된다.

GLOBAL FACE SR 모듈은 식(4)와 (5)에 해당하는 연산을 수행한다. 외부 SDRAM에 저장된 학습 데이터

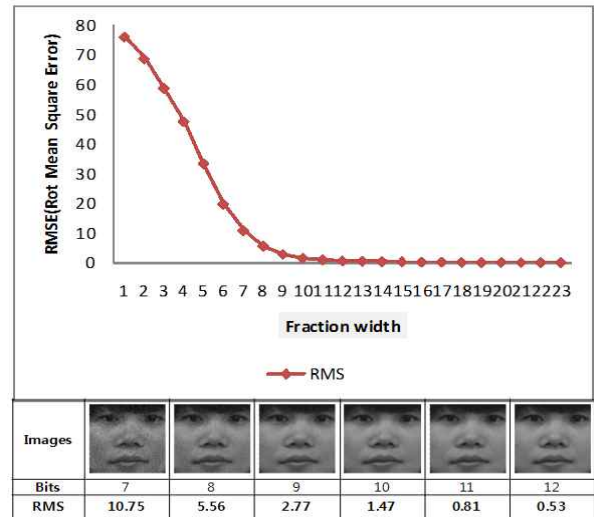


그림 8. 학습 데이터의 소수점 비트 수에 따른 오류율  
Fig. 8. Error in proportion to the fraction width training data.

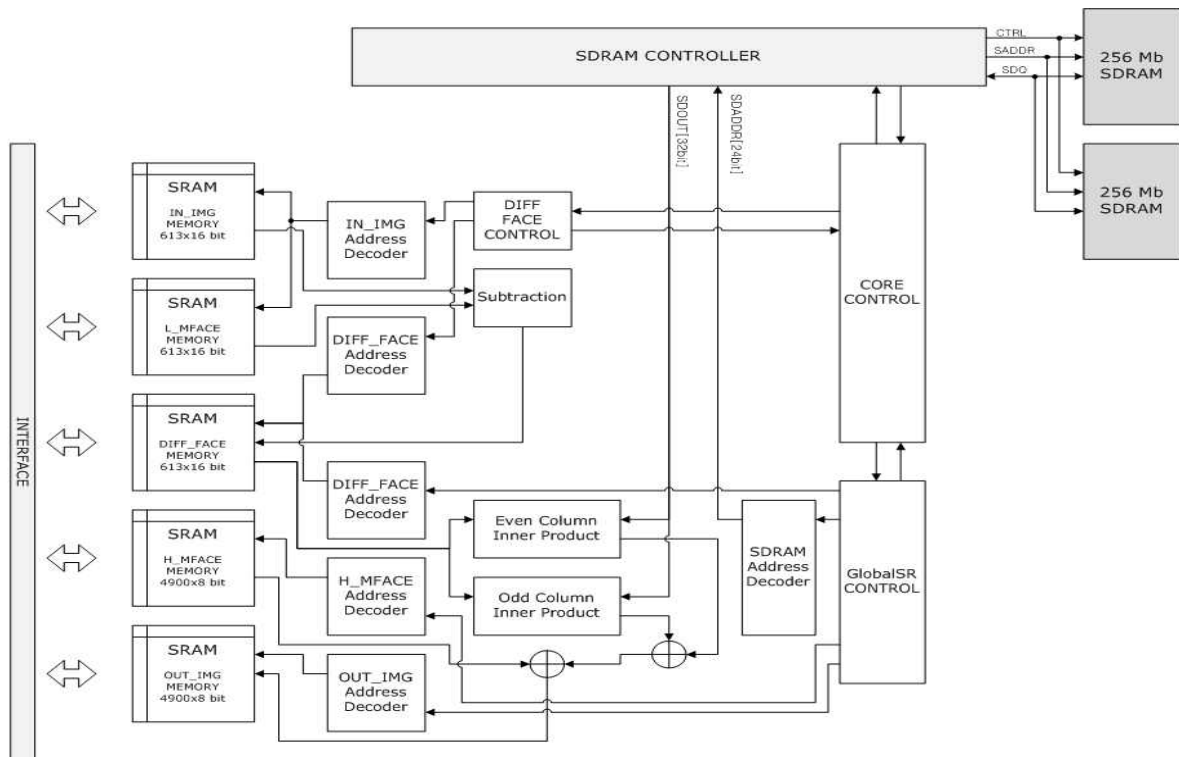


그림 7. 얼굴 영역 초해상도 모듈의 내부 구조  
Fig. 7. Detail Block Diagram of Face SR Core.

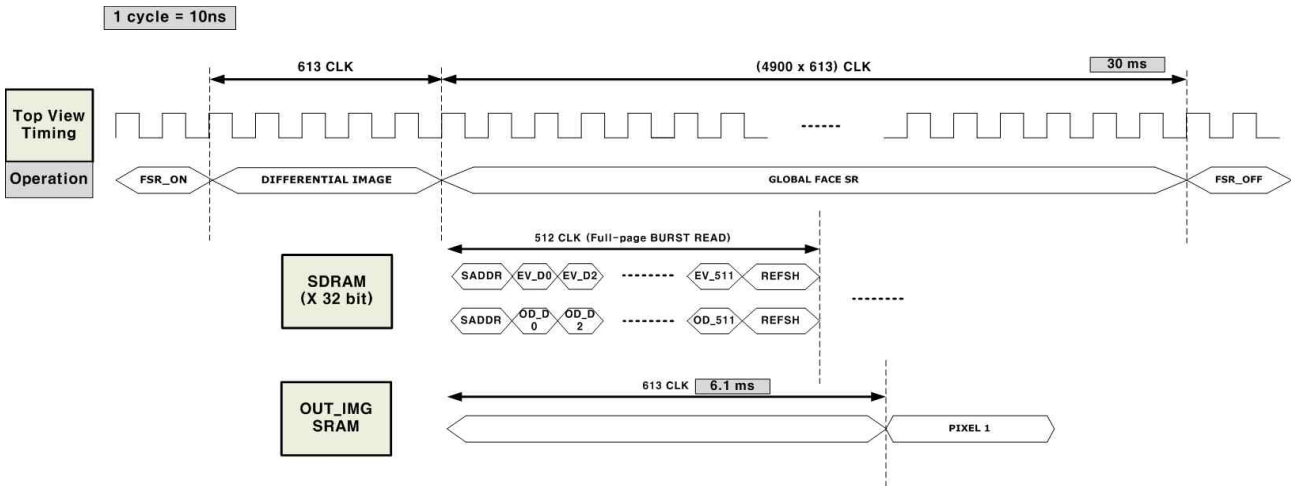


그림 9. 설계된 하드웨어 구조의 클럭 측정도  
 Fig. 9. Clock measurement of the proposed hardware design.

와 DIFF\_FACE 메모리에 저장된 차영상의 내적을 수행하고, 고해상도 평균 얼굴 영상을 합하여 최종 결과를 OUT\_IMG메모리에 저장하는 동작을 수행한다.

본 논문에서 구현하고자하는 학습 기반의 저해상도 영상 복원 알고리즘은 앞의 1절에서 언급한대로  $35^2 \times 70^2 \times 32$  bit의 부동 소수점 데이터인 학습 데이터가 전체의 연산에서 큰 부분을 차지한다. 그림 8은 소수점 이하의 비트 수에 따라 부동 소수점으로 얻은 결과와의 에러율(RMSE)을 나타낸 것으로, 이에 따라 정수 부분 1비트와 소수점 이하 12비트로 고정시켜 약 9MB의 학습 데이터를 SDRAM에 순차적으로 저장한다.

32bit SDRAM을 사용하여 한 클럭에 연속된 2개의 데이터를 동시에 처리 할 수 있도록 연산 로직을 구성하였다. 또한, SDRAM의 Full-page READ MODE를 사용하여 메모리의 한 페이지 단위로 데이터를 읽어 옴으로써 메모리 접근 시간을 최소화하였다.

그림 9는 전체 연산에 소요되는 클럭을 측정된 것이다.  $35 \times 35 \times 8$  bit 의 저해상도 입력 영상을 입력을 내부 메모리 L\_FACE에  $\times 16$ 으로 저장하는데 총 613 클럭이 소요된다. 다음  $70 \times 70 \times 35 \times 35 \times 16$  bit의 학습데이터를 SDRAM으로부터 순차적으로 32bit 씩 읽어 들여 해당 연산을 하는데 총 3,001,250 클럭이 소요된다. 여기에  $70 \times 70$ 의 영상을 출력하는데 소모되는 613 클럭을 더하면 전체 3,002,476 클럭을 소모하게 된다.

### 3. 결과 및 성능 분석

그림 10은 설계한 하드웨어와 소프트웨어를 검증하기 위한 FPGA 검증환경이다. 하드웨어로 구현된 얼굴



그림 10. FPGA 검증 환경  
 Fig. 10. FPGA verification environment.

표 2. FPGA 및 MPW 결과  
 Table 2. FPGA verification result and MPW result.

	FPGA (Xilinx Virtex-4)	MPW (Samsung 0.13 um)
Synthesis	Xilinx 사 ISE 10.1	Synopsys사 Design Compiler
	1,268 LUTs	141,389 gates
Memory	105 Kbits	160 Kbits
Operating Clock	100 Mhz	100 Mhz
Operation Speed	33 ms(약 30 fps)	30 fps (target)

영역 초해상도 모듈은 Xilinx Virtex-4에서 100 MHz의 동작 속도를 가지고, 1,268 LUTs와 약 105 Kbits의 블록 메모리를 사용하였다.

$35 \times 35$ 의 저해상도 입력 영상은 S3C2440A 프로세서를 이용하여 내부 메모리에 저장되고 하드웨어로 구현된 얼굴 영역 초해상도 모듈을 통해 복원된 최종 결과 영상은 다시 내부메모리에 저장된다. 이 결과 영상을

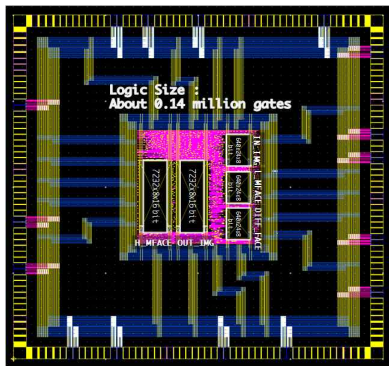


그림 11. 얼굴 영역 초해상도 하드웨어 칩 레이아웃  
Fig. 11. The chip layout of Face SR(Super-Resolution) hardware.

host로 전달 받아, 소프트웨어로 구현한 고정 소수점 알고리즘의 결과 영상과 하드웨어에서의 결과 영상이 동일함을 확인하였다.

연산 시간의 측면에서는, 저해상도 입력 영상으로부터 고해상도 복원 영상을 출력하는데 총 33 msec가 소요되며, 1초당 약 30프레임의 처리가 가능하다. 이는 실시간 감시를 위한 지능형 영상 감시 시스템이 요구하는 성능을 충분히 만족시킬 수 있다.

현재 본 논문에서 설계한 얼굴 영역 초해상도 모듈은 Samsung 0.13um CMOS 표준 라이브러리를 사용하여 MPW 공정으로 칩을 제작하여 테스트 중에 있다. Synopsys사의 Design Compiler를 이용하여 100 MHz의 동작 속도를 타겟으로 합성한 결과 141,389 gate count를 가지며, 160Kbits의 SRAM 메모리를 사용하였다. 표 2는 FPGA와 MPW의 제작 결과를 정리한 표이며, 그림 11은 제작된 칩의 Layout이다.

#### IV. 결 론

최근 공항, 항만 등과 같은 공공장소의 보안을 위한 원거리 감시 시스템은 실시간으로 용의자 또는 미아 등과 같은 특정 인물을 찾을 수 있는 능동적인 얼굴 인식 기술을 요구한다. 따라서 기존의 얼굴 인식 시스템들은 먼 거리에서 촬영된 저해상도 얼굴 영상 처리 문제에 따른 낮은 인식률에 대한 한계를 극복해야만 한다.

본 논문에서는 이를 해결하기 위해, 얼굴 영역에 대한 고해상도 이미지를 복원하는 방법을 분석하고, 소프트웨어를 통해 검증하였다. 그 다음, 기존의 얼굴 인식 기술과 통합된 보안 감시 시스템 구축을 고려하여 임베디드 환경에서의 성능을 측정하였다. 이를 바탕으로 실

시간 동작에 적합한 형태의 하드웨어 구조를 설계하고 ARM9기반 임베디드 프로세서 S3C2440 와 Xilinx Virtex-4에서 검증하였다. 그 결과, 제안된 하드웨어 구조는 100 MHz의 동작 속도에서 초당 30 프레임의 처리가 가능하여 실시간 동작에 무리가 없다.

향후 지능형 영상 보안 시스템에서의 실시간 처리를 위해서는 다음과 문제의 해결이 필요하다.

먼저, 실시간으로 획득된 저해상도 데이터의 경우, 동영상에서 입력되므로 정지된 형태로 촬영된 얼굴 데이터베이스와 차이가 생기게 된다. 이는 학습 기반 알고리즘의 특성상 복원 성능에 영향을 줄 수 있으므로, 실제 입력 영상과 유사한 정보를 많이 가지고 있는 얼굴 데이터베이스의 구축도 중요하다.

또한, 주성분 분석(PCA)을 사용하여 복원한 얼굴에서는 얼굴의 주를 이루는 성분들 이외의 부분은 제거되며, 입력된 얼굴 영상이 지니고 있던 국부적인 특징, 예를 들어, 얼굴 윤곽선, 머리카락, 눈썹 등의 정보가 손실된 상태이다. 그러므로 손실된 입력 얼굴 영상의 지역적인 특징들을 보완할 수 있는 알고리즘의 대한 연구가 필요하다.

더불어, stand-alone 형태의 임베디드 시스템을 위해 얼굴 검출단과 얼굴 인식을 통합하는 연구가 진행되어야 한다.

#### 참 고 문 헌

- [1] 유장희, 문기영, 조현숙 “지능형 영상보안 기술현황 및 동향,” ETRI 전자통신동향분석 제23권, 제4호, 476-486쪽, 2008년 8월
- [2] 박정선, “저해상도 얼굴 영상의 해상도 개선을 위한 영역 기반 복원 방법,” 정보과학회 논문지, 소프트웨어 및 응용 제34권, 제5호, 476-486쪽, 2007년 5월
- [3] 염석원, “포톤 카운팅 선형 판별법을 이용한 저해상도 얼굴 영상 인식,” 전자공학회 논문지, 제45권, SP편 제6호, 597-602쪽, 2008년 11월
- [4] S.C. Park, M.K Park, and M.G. Kang, “Super-Resolution Image Reconstruction: A Technical Overview”, IEEE Signal Processing Magazine, Vol. 20, No. 3, pp.425-434, 2005.
- [5] M.H. Sedky, M. Moniri, and C.C. Chibelushi, “Classification of Smart Video Surveillance Systems for Commercial Applications,” in Proc. of IEEE AVSS 2005, Sep. 2005, pp.638-643
- [6] S. Baker and T. Kanade, “Limits on



- super-resolution and how to break them,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 9, pp. 1167–1183, Sep. 2002.
- [7] C. Liu, H.Y. Shum, and WT Freeman, “Face Hallucination: Theory and Practice,” International Journal of Computer Vision, Volume 75, No. 1, Oct 2007.
- [8] 김윤구 외, “임베디드 시스템 적용을 위한 얼굴검출 하드웨어 설계”, 대한전자공학회, 제44권 SD편, 2007년 9월.
- [9] BioID face database  
<http://www.bioid.com/support/downloads/software/bioid-face-database.html>
- [10] The ORL Database of Faces  
<http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>

---

 저 자 소 개
 

---



김 초 롱(학생회원)  
 2005년 광운대학교 정보제어  
 공학과 학사 졸업.  
 2005년~2007년 하이닉스 반도체  
 연구원  
 2011년 광운대학교 임베디드  
 소프트웨어공학과 석사.

<주관심분야 : SoC 설계, 영상처리 및 인식, 임베디드 시스템>



정 용 진(정회원)  
 1983년 서울대학교 제어계측  
 공학과 학사  
 1983년~1989년 한국전자통신  
 연구원  
 1995년 미국 UMASS 전자전산  
 공학과 석사/박사

1995년 삼성전자 반도체 수석연구원  
 1999년~현재 광운대학교 전자통신공학과 교수  
 <주관심분야 : 무선통신, 정보보호, SoC 설계,  
 영상처리 및 인식, 임베디드 시스템>