DSP 기반 실시간 ATM 보안 시스템

DSP based real-time ATM security system

이태민, Tae-min Lee, 김용국, Yong-Guk Kim
세종대학교 컴퓨터 공학과

요약 은행에서 현금 지급기(ATM)를 부정한 목적으로 사용하는 사람들은 보통 마스크나 선글라스, 모자 같은 것으로 얼굴을 은폐하고 인출을 하는 경우가 많다. 그렇기 때문에 마스크나 선글라스, 모자로 얼굴을 가림으로써 특징을 검출하기 쉽지 않아 얼굴 인식을 통한 사람 판단이 어렵다. 본 논문에서는 차 영상과 Template Matching 을 통해 얼굴 영역을 추출하고 Adaptive Boost 를 통해 얼굴의 특징 점을 검출한 후 스킨 컬러 정보를 이용하여 현재 사람의 은폐 정보를 추정하는 방법을 제안한다. 제안된 방법은 영상신호처리에 강하고 비용이 적으며 적은 전력으로 동작하는 DSP 기반에 탑재 함으로써 ATM 기에 탑재하기 적합하고 또한 다른 형태의 검증 시스템에 적용할 수 있는 효율적인 구조를 제시한다.

핵심어: DSP(Digital Signal Processor) , Face Estimation with Template Matching, Adaptive Boost, Skin Color Algorithm

*이태민 : 세종대학교 컴퓨터 공학과 석사과정 e-mail: 002376@hanmail.net

**김용국 : 세종대학교 컴퓨터 공학과 교수 e-mail: ykim@sejong.ac.kr

1. Introduction

현재 사회에서 정보 보호를 위한 생체 인식 기술이 사회적 요구에 따라 급속도로 확대 되어 가고 있는 시점에서 소위 Ubiquitous 환경에 맞는 Embedded 형 보안 시스템의 연구가 활발히 진행되고 있다.

그 중에서 은행 ATM 기는 개인의 재산과 관련된 거래로 인하여 많은 보안적 요소가 요구된다[1]. 이러한 ATM 기에 설치된 카메라를 통해 범죄 유무를 식별할 때 마스크나 선글라스, 모자와 같은 얼굴 은폐물을 착용할 경우 개인을 구분 하기가 어렵고 계속 지켜보지 않는 이상 범죄 행위를 판단하기가 쉽지 않다. 본 논문에서는 사람의 은폐 정도에 따라 위험도를 결정하는 Embedded 형 보안 시스템을 제안한다.

이러한 시스템을 통한 개인의 위험도 추정을 통해 위험도가 높은 사람만 체크함으로써 ATM 기에서 일어날 수 있는 범죄 행위를 미연에 방지 하는데 효과적으로 사용될 수 있다. 뿐만 아니라 DSP 를 사용하여 소형화된 기기 구성이 가능하고 가격이 저렴하며 전력 소모도 크지 않기 때문에 보안이 많이 요구되는 장소에서도 이러한 시스템을 적용 시킬 수 있다.

2. Motion based method for estimating face region

차 영상은 두 영상 사이의 차이를 사용하여 움직임 정보를 획득 할 수 있는 방법이다 (1)

$$I_{out}(x,y) = I_1(x,y) + I_2(x,y) \qquad (1)$$

그러나 일반적인 차 영상은 계속해서 움직여야만 차 영상이 확보가 되는 단점이 있다. 따라서 움직임이 없을 경우 기존의 영상을 그대로 누적시켜 사용하는 방법을 사용하였다. 이러한 방법을 사용하여 획득한 영상을 1D Histogram 과 Labeling 을 통하여 얼굴 영역을 추정한다.

3. Feature based method for detecting face region.

3.1 Template Matching

Template Matching 은 영상 이미지와 Template 간의 유사도가 높은 부분을 추출해 내는 방식이다. Template Matching 의 가장 큰 장점은 간단한 구조를 가지고 있으며 환경 변화가 크지 않으면 정확도가 높다는 장점이 있다. 유사도의 측정은 상관계수를 측정 하고 상관계수가 가장 큰 부분을 Matching 영역으로 결정한다. (2)

$$T'(x',y') = T(x',y') - m_{T(x',y')}$$

$$I'(x+x',y+y') = I(x+x',y+y') - m_{I(x+x',y+y')}$$

$$Corr(x,y) = \frac{\sum_{(x',y')\in W} T'(x',y')I'(x+x',y+y')}{\sqrt{\sum_{(x',y')\in W} T'(x',y')^2 \sum_{(x',y')\in W} I'(x+x',y+y')^2}} \qquad (2)$$

(w: width, h: height, m :mean)

하지만 (2)와 같은 수식을 사용할 경우는 간단하지만 연산량이 많아 처리 속도가 매우 느리다. 따라서 Fast Fourier Transform 을 이용하여 보다 빠른 Template Matching 을 수행하는 방법[2]을 통해 연산 시간을 개선하였다. (3)

$$T'' = T'(-x',-y')$$

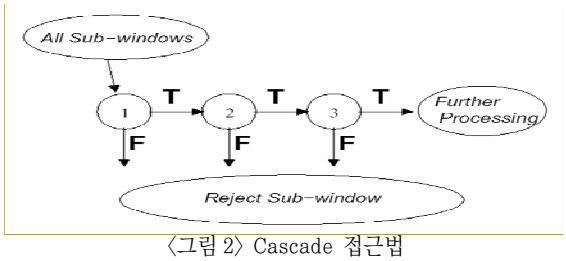
$$I * T'' = F^{-1}(F(I) \otimes F(T')) \qquad (3)$$

본 논문에서는 차 영상을 통해 나타내는 이진화 된 영상을 여러 개의 Face shape template mask〈그림 1〉를 사용하여 얼굴 영역을 검출하였다.

〈그림 1〉여러 개의 Face shape template mask

3.2 Boosted Cascade of Simple Features

Adaptive Boost 알고리즘은 n 개의 약한 분류기의 선형적인 결합을 통하여 최종적으로 높은 검출 성능을 가지는 강한 분류기를 구성하는 방법[3]이다. 환경 요소에 강인하면서 빠르게 연산이 가능한 Harr-Like Feature 를 사용하여 약한 분류기를 구성하고 〈그림 2〉과 같은 Cascade 구조를 통해 확률적으로 높은 데이터를 바탕으로 단계적으로 접근함으로써 원하는 오브젝트를 검출해 낼 수 있다.

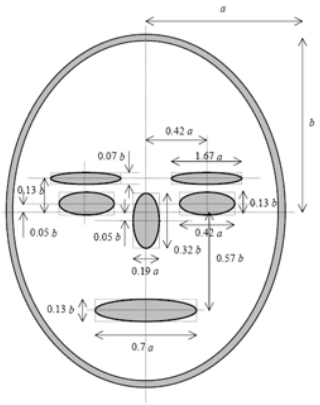


〈그림 2〉 Cascade 접근법

본 논문에서는 차 영상과 Template Maching 을 통해 추정된 얼굴영역 내에서 Adaptive Boost 알고리즘을 이용하여 눈, 코 객체를 검출함으로써 얼굴의 특징 점을 추출해 낸다.

4. Measure of hiding face region.

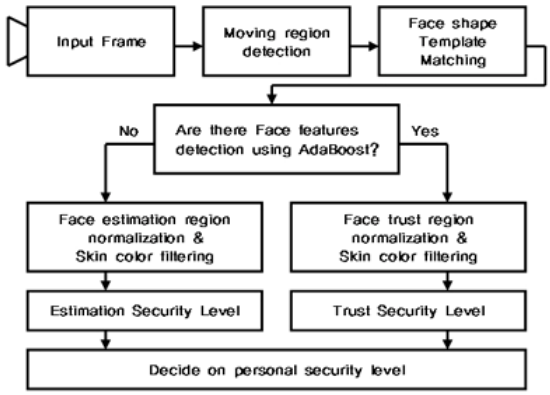
얼굴을 구성하는 특징점인 눈, 코 객체가 추출되면 위치 정보를 이용하여 정규화를 수행한다. 여기서 특징점으로 입 객체를 제외한 이유는 입 객체는 표정에 따른 변화가 심하여 연산 속도에 비해 효율적인 검출을 제공하지 못하기 때문이다. 정규화는 〈그림 3〉에 제시된 Mask Templete[4]의 좌표 비율을 통하여 구성한다.



〈그림 3〉 얼굴 영역의 정규화를 위한 Mask Template

얼굴 영역이 정규화 되면 YCbCr 공간에서 얼굴 영역 내의 Cb, Cr 정보를 이용하여 Skin-area 를 검출해 낸다. 여기서 Cb, Cr 정보를 사용하여 skin-area 를 검출해 내는 이유는 DSP 내에서 받는 영상 신호가 Y, Cb, Cr 로 수신되기 때문에 빠른 검출이 가능하고 다른 컬러 모델에 비해 보다 정확한 Skin-area 구성이 가능하기 때문이다. 이러한 Skin-area 의 분포 정보는 얼굴의 은폐 정도를 측정하는 기준이 된다.

5. Vision based ATM security system



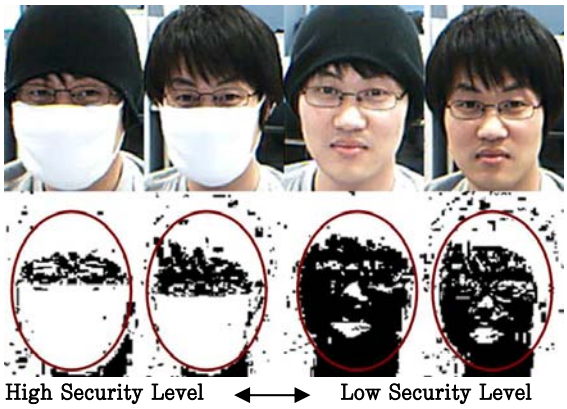
〈그림 4〉 Block Diagram of the proposed real-time ATM Security system

제안된 시스템은 Difference Image 를 사용하여 모션 정보를 획득하고 1D Histogram 과 Labeling 을 통하여 모션 영역을 잘라낸다. Difference Image 를 통한 모션 영역 정보는 얼굴 특징 점을 찾기 위한 연산 시간을 줄이고 개인이 ATM 을 사용하는 간격을 결정한다.

다음은 구해진 모션 영역 내에서 여러 개의 Face shape template mask 를 사용하여 Face Region 을 결정하고 Face Region 내에 있는 눈, 코 객체를 Adaptive Boost 를 통하여 추출한다. 하지만 선글라스나 마스크와 같은 은폐물을 착용하였을 경우 눈, 코 객체를 추출할 수 없는 경우가 발생하는데 그럴 경우 Face shape template mask 의 비율을 이용하여 얼굴 영역을 추정하게 된다.

본 논문에서는 다음과 같이 Adaptive Boost 를 통하여 정규화 된 얼굴 영역을 Trust Face Region 으로, 눈, 코 객체의 특징이 없어 Face shape template mask 의 비율로 정규화 된 얼굴 영역을 Estimation Face Region 으로 구분해서 위험도를 측정하는 방법을 제안하였다.

정규화된 얼굴 영역은 YCbCr 공간에 존재하는 Cb, Cr 의 Skin-area 를 통하여 스킨 정보를 획득하고 검출된 스킨 영역의 얼굴 영역 내에 차지하는 비율을 측정하여 보안 등급을 결정한다. 〈그림 4〉



〈그림 4〉 상황 별 스킨 영역을 통한 위험도 결정

Trust Face Region 에서 측정한 Trust Security Level 은 확률적으로 Estimation Face Region 에서 측정한 Estimation Security Level 보다 신뢰성이 높다. 따라서 Trust Security Level 의 가중치(W)를 Estimation Security Level 보다 값을 많이 주고, 각각에 대한 평균(m)과 분산(σ^2)을 고려하여 최종 Security Level 을 결정하게 된다. (4)

$$SecurityLv. = \frac{m_{trust} W_{trust} \sigma^2_{esti} + m_{esti} W_{esti} \sigma^2_{trust}}{W_{trust} \sigma^2_{esti} + W_{esti} \sigma^2_{trust}} \quad (4)$$

6. Environment

제안된 시스템은 Texas Instruments 에서 제공하는 DSP 64x[5]계열의 CPU 를 탑재한 Spectrum Digital 사의 EVM DM 642 Board 를 사용하여 실험하였다. 〈그림 5〉

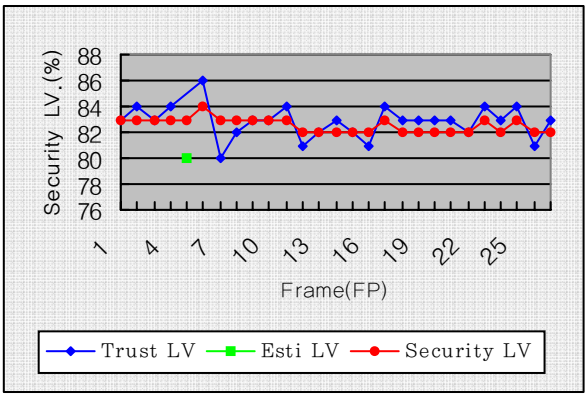
〈그림 5〉 Spectrum Digital 사의 DSP 64x 계열 Board

DSP 64x 계열은 Texas Instruments 에서 개발한 디지털 신호를 처리하기 위한 회로로 가격대 전력소모량 및 성능비가 우수 하며 고속의 반복적 연산 처리로 인해 현재 H264, MPEG4, MPEG2 와 같은 영상 압축용 인코더 및 디코더에 주로 사용한다.

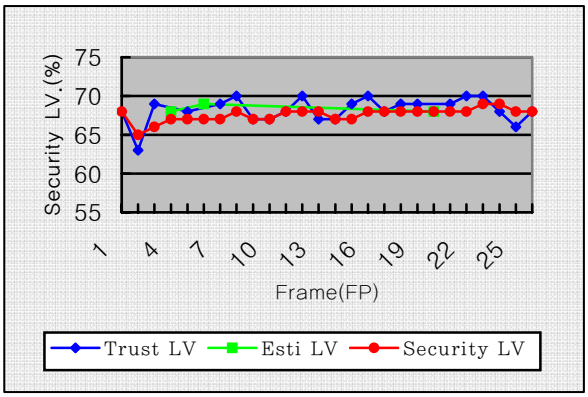
본 실험에 사용된 EVM DM 642 Board 의 사양은 TMS 6411 720 Mhz 의 CPU 와 64MB SDRAM 을 탑재하고 있으며 그 밖에 4MB Linear Flash, 2 개의 Video Decoder 와 1 개의 Video Encoder 를 제공하고 있다.

7. Experiments and Result

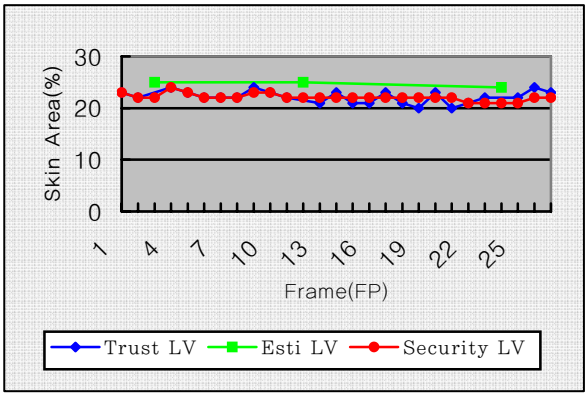
실험은 아무것도 착용하지 않은 상태, 모자를 착용한 상태, 마스크를 착용한 상태, 모자와 마스크를 착용한 상태를 가정하고 실험하였다. 모자 착용 시 에는 눈이 보이지 않도록 유의하여 실험에 참여하였다. Security Level 은 피부 영역을 백분율로 계산하여 측정하였다. 측정 시간은 1 분 동안 ATM 기를 사용한다는 가정 아래 실험을 수행하였다.



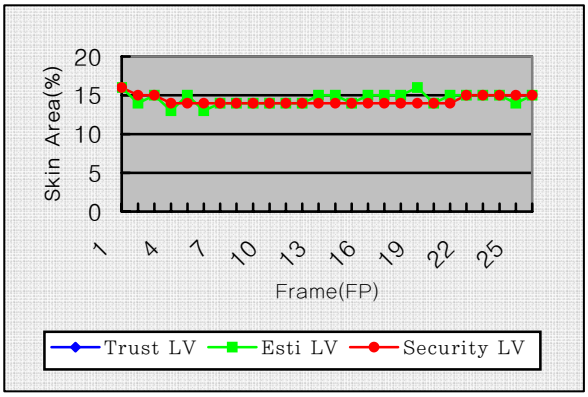
〈 표 1. 은폐물 미착용 시 Security Level〉



〈 표 2. 모자 착용 시 Security Level〉



〈 표 3. 마스크 착용 시 Security Level〉



〈 표 4. 마스크와 모자 착용 시 Security Level〉

실험 결과 모든 상황에 대해서 분당 25 Frame 에서 27Frame 의 연산 처리, 즉 2.2~2.4sec/frame 의 처리 성능을 보였다.

은폐물을 미착용 할 경우〈표 1〉에는 상대적으로 Trust Security Level 이 높은 것을 볼 수 있는데 얼굴이 특징을 은폐물을 통해 가리지 않았기 때문에 상대적으로 Trust Security Level 이 Estimation Security Level 보다 압도적으로 높을 수 있었다.

반면에 모자와 마스크를 모두 착용할 경우〈표 4〉에는 얼굴의 특징 점, 즉 눈, 코를 찾을 수 없기 때문에 Estimation Security Level 값으로만 최종 Security Level 을 결정한다.

모자 착용 시〈표 2〉와, 마스크 착용 시〈표 3〉에는 부분적으로 특징을 찾을 수 없지만 눈, 코의 특징은 각각 찾을 수 있기 때문에 Trust Security Level 이 높게 작용하는 것을 볼 수 있다.

10 번의 반복적인 실험에서 은폐물을 미 착용한 결과는 평균 80~83% 정도 Security Level 이 측정 되었고, 모자 착용 시에는 평균 64~68% 정도 Security Level 이 측정되었으며 마스크 착용시에는 평균 20~22%의 Security Level, 모자, 마스크를 전부 착용할 경우 평균 12~14%의 Security Level 이 각각 측정 되었다.

8. Conclusion and Future Work

본 논문에서는 DSP 기반의 모션 정보와 얼굴의 특징 점 및 스킨 컬러 영역 정보를 사용한 ATM 에서의 보안 시스템을 제안하였다.

제안된 시스템에 탑재된 알고리즘은 다른 인식 알고리즘에 비해 상대적으로 연산 량이 많지 않기 때문에 성능이 개인용 컴퓨터에 비해 떨어지는 Embedded System 에 탑재하기 적합하고 그에 따라 실시간 처리가 가능하다. 하지만 보다 정확한 처리를 위해서는 개인용 컴퓨터에 비해 현저히 느린 시스템의 성능 향상이 필수 요건이라 하겠다. 또한 환경 변화에 보다 강인한 시스템을 구축하는 것도 남겨진 문제라 할 수 있다.

앞으로 ATM 보안 시스템 알고리즘의 개선, DSP Module 최적화 구성을 통한 성능 개선을 통해 이러한 문제를 해결 하는데 중점을 둘 것이다.

참고문헌

[1] Bowyer, K.W, "Face recognition technology: security versus privacy," Technology and Society Magazine, IEEE,Volume,23, p9-19,2004

[2]Altmann, j. and Reitbock, H.h,P., A Fast Correlation Method for Scale-and Translation-Invariant Pattern Recognition, IEEE Trans. on PAMI, pp.46-57, 1984

[3]Paul Viola, Michael Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features" Conference on Computer Vision and Pattern Recognition 2001

[4] Dario Maio and Davide Maltoni "Real-Time Location on Gray-Scale Static Images"

[5]" http://focus.ti.com/lit/ml/ssdv004s/ssdv004s.pdf ",DSP Selection Guide, Texas Instruments, pp32-39,2007