

SVM 분류기에 의한 얼굴 특징 식별 시스템

박 강 령* · 김 재 희** · 이 수 연***

요 약

금융권의 주 5일제 근무에 따른 무인 현금 인출기의 사용 확대와 함께, 타인의 신용카드를 이용하여 무인 현금 인출기에서 돈을 인출하는 금융 범죄에 대한 원천적인 예방 대책이 필수적으로 요구되고 있다. 특히 무인 현금 인출기 부근에는 감시용 CCTV Camera가 설치되어 있으나, 지능적인 범죄자들은 이러한 사실을 인식하고 선글라스, 마스크 등을 착용하여 이러한 감시 시스템을 피해가고 있다. 본 논문에서는 이러한 문제점을 해결하기 위하여 무인 현금 인출기에 설치되어 있는 카메라를 통해 입력된 사용자의 얼굴 및 얼굴 특징점을 영상 신호처리 방법과 SVM(Support Vector Machine)으로 분석하여 향후 얼굴이 식별 가능한 경우에만 금융 거래를 할 수 있도록 하는 시스템을 개발하였다. 실험 결과, 학습 데이터에 대해서는 약 1%의 오 인식율과 2%의 오 거부율을 나타냈으며, Test 데이터에 대해서는 약 2.5%의 오 인식율과 1.43%의 오 거부율을 나타냈다.

Facial Feature Verification System based on SVM Classifier

Kang Ryoung Park* · Jaihie Kim** · Soo-youn Lee***

ABSTRACT

With the five-day workweek system in bank and the increased usage of ATM(Automatic Teller Machine), it is required that the financial crime using stolen credit card should be prevented. Though a CCTV camera is usually installed in near ATM, an intelligent criminal can cheat it disguising himself with sunglass or mask. In this paper, we propose facial feature verification system which can detect whether the user's face can be identified or not, using image processing algorithm and SVM(Support Vector Machine). Experimental results show that FAR(Error Rate for accepting a disguised man as a non-disguised one) is 1% and FRR(Error Rate for rejecting a normal/non-disguised man as a disguised one) is 2% for training data. In addition, it shows the FAR of 2.5% and the FRR of 1.43% for test data.

키워드 : 얼굴 특징 검증(Facial Feature Verification), SVM

1. 서 론

최근 금융권의 주 5일제 근무에 따른 무인 현금 인출기의 사용 확대와 함께, 타인의 신용카드를 이용하여 현금 인출기에서 돈을 인출하는 금융 범죄에 대한 원천적인 대책이 필수적으로 요구되고 있다. 현금 인출기가 설치되어 있는 장소는 은행 지점 내에 설치되어 있는 유인 코너와 365일 자동화 코너와 같은 무인 코너로 나뉜다. 이중, 유인 코너는 은행원이나 청원 경찰이 업무 시간 내에 항상 대기하고 있으므로 비교적 안전한 상태에 있지만, 무인 코너는 은행원이나 청원 경찰이 아무도 없는 ATM 기기만 설치되어 있는 곳이므로 범죄의 주 대상이 되고 있다. 현금 인출기 범죄에는 침입자가 들어와 기기를 물리적으로 파손한 뒤 돈을 찾

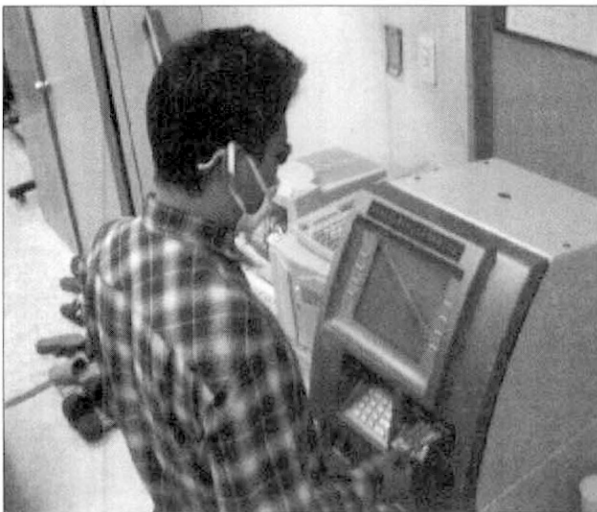
아가는 유형과 타인의 카드를 훔치거나 위조하여 돈을 찾아가는 유형 등이 있다. 전자의 경우는 지점 내에 각종 물리적인 파손을 감지하는 센서들이 부착되어 있기 때문에 보안회사에서 바로 출동할 수 있도록 되어 있어 비교적 안정성이 보장되어 있다. 그러나 타인의 카드를 훔치거나 위조하여 돈을 찾아가는 유형의 경우는 기기에서 동작중인 센서들에서 비정상적인 거래로 판단할 수 없기 때문에 예방하기 어렵다. 이러한 이유로 최근 후자에 대한 범죄율이 급속히 증가되고 있다. 물론 감시용 카메라가 설치되어 있으므로 사후에 조사 자료로는 활용될 수 있으나, 지능적인 범죄자들은 이러한 사실을 인식하고 선글라스, 마스크 등을 착용하여 이러한 감시 시스템을 피해가고 있다. 이와는 별도로 최근 생체 인식 시스템(지문, 홍채 등)을 이용하여 이러한 범죄를 예방하고자 하는 움직임이 일고 있으나, 기본적으로 이러한 생체 인식 시스템을 사용하기 위해서는 자신의 생체 정보를 등록해야하는 별도의 과정이 필요하며, 또한 기존에 설치되어 있는 현금인출기의 교체 작업에 따른 막대한 비용 증가

* 이 논문은 2003년도 광운대학교 교내연구비를 지원받았음.
 † 정 회 원 : 상명대학교 소프트웨어대학 미디어학부 전임강사
 ** 정 회 원 : 연세대학교 전기·전자공학과 교수
 *** 정 회 원 : 광운대학교 컴퓨터공학부 교수
 논문접수 : 2003년 3월 10일, 심사완료 : 2004년 5월 10일

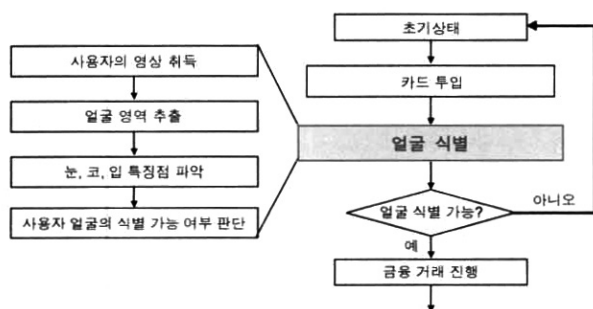
의 문제점이 있다. 본 논문에서는 이러한 문제점을 해결하기 위하여 무인 현금 인출기의 교체 작업 없이 내부에 손쉽게 설치할 수 있는 소형 핀홀 카메라(pinhole camera)를 통해 입력된 사용자의 얼굴 및 얼굴 특징을 영상 신호 처리 방법과 SVM 분류기(Support Vector Machine Classifier)로 분석하여, 향후 얼굴을 식별 가능한 경우에만 금융 거래를 할 수 있도록 하는 시스템을 개발하였다.

2. 현금 인출기에서의 얼굴 식별 시스템

본 논문에서 제안하는 현금 인출기에서 얼굴 특징 식별 시스템(그림 1)이 작동되는 단계별 과정은 (그림 2)와 같다.



(그림 1) 현금 인출기에서의 얼굴 특징 식별 시스템



(그림 2) 시스템의 개요

(그림 2)를 보다 자세히 설명하면 다음과 같다. 사용자가 카드를 투입하는 순간, RS-232C 인터페이스를 통해 현금 인출기내의 프로세서로 거래 시작 신호를 보낸다. 이와 같은 거래 시작 신호를 받게 되면 현금 인출기내의 프로세서에서 동작중인 얼굴 특징 식별 프로그램이 현금 인출기내에 부착된 소형 핀홀 카메라를 통해 사용자의 얼굴 영상을 취득한다. 그리고 취득된 영상 내에서 사용자의 얼굴 영역을 추출하고, 얼굴 영역 내의 눈·코·입 특징점을 영상 처리

방법 및 SVM(Support Vector Machine)으로 판단하여 모두 제대로 보이는 경우에는 사용자가 정상적으로 현금 인출기를 사용하는 상태로 판단한다. 반면, 선글라스, 마스크, 모자 등을 착용하여 얼굴 특징점들이 제대로 보이지 않는 경우는 비정상적인 경우로 간주하여 금융 거래를 차단시킨다.

이처럼 얼굴이 정상적인 상태인지를 판단하는 것을 얼굴 특징 식별이라 한다. 얼굴 특징 식별에서 무엇보다도 중요한 기술은 입력영상에서 얼굴부분을 정확하게 추출하고 이로부터 눈·코·입 특징점의 존재 가능 부위를 정확하게 결정하는 것이다. 이로부터 결정된 영역 내에 눈, 코, 입 특징점이 제대로 인지 가능하지 여부를 판단하는 것이다. 현금 인출기에서 사용하는 카메라는 다양한 범위(현금 인출기 앞 10cm~100cm)에 존재하는 다양한 키(신장 150cm~190cm)의 사용자의 얼굴을 모두 취득하기 위해서 -50~+50도 이상의 광각 렌즈(어안 렌즈)를 사용해야 한다. 그러므로 사용자는 신장 및 사용 위치 등에 따라 전체 영상 영역 중 어느 곳에서든지 얼굴이 존재 가능하다. 또한, 대부분 줌렌즈 없이 단초점 렌즈를 사용하므로, 사용자 얼굴 크기 역시 사용 위치에 따라 다양한 크기로 나타나게 된다. 그리고 사용자의 불편감을 유발하지 않기 위해서는 사용자가 미처 감지하지 못할 정도로 작은 크기의 렌즈(pinhole lens)를 사용해야 한다. 일반적으로 렌즈 직경은 Iris 직경과 비례관계에 있으며, Iris 직경이 작아지면 카메라 sensor로 입사되는 조명의 양이 작아져서 충분한 밝기를 내기 위해서는 카메라의 AGC(Automatic Gain Control)를 상승시키게 된다. 이처럼 AGC의 상승은 카메라에 잡음 성분을 증가시켜서 입력 영상 내의 영상 처리를 더욱 어렵게 만든다. 그러므로 충분한 광량을 내기 위해서는 추가 조명을 설치해야 하나 880nm 이하의 많은 양의 가시광선 조명을 사용자 정면에 설치하게 되면 사용자에게 눈부심으로 인한 많은 불편감을 유발하여 사용하기에 적절하지 않다. 반면, 880nm 이상의 적외선 조명은 사람이 눈에 의해 감지가 되지 않으므로 사용자에게 불편감을 유발하지는 않으나, 일반적인 CCD Sensor의 경우 적외선대의 감도가 가시광선 대에 비해 상대적으로 많이 떨어진다. 그러므로 적외선으로 충분한 양의 광량을 내기 위해서는 상당히 많은 양의 적외선 조명을 설치해야 하므로 비용 상승의 원인이 된다. 또한 가시광선과 적외선의 감도가 비슷하다고 알려진 CMOS Sensor의 경우, 역시 전체적인 감도가 CCD Sensor의 50~60% 정도 밖에 나오지 못하며, 또한 외부광 차단을 위해 필수적으로 요구되는 전기 셔터(electric shutter)기능이 포함되어 있지 않아서 기계적 셔터(mechanic shutter)에만 의존해야 한다. 최근 일부 회사[1]에서 CCD Sensor와 감도가 비슷하고 전기 셔터 기능을 포함하고 있는 CMOS Sensor를 출시하고는 있으나 아직까지 그 성능이 검증되지 않은 상태이다. 그러므로 본 논문에서는 이러한 문제점들을 해결하기 위하여, 사용자에게 불편

감을 주지 않을 정도의 최소한의 은은한 가시광선을 추가 조명으로 사용하였으며, 이를 바탕으로 영상처리 방법과 SVM 분류기(Support Vector Machine classifier)에 의한 얼굴 특징점 식별 시스템을 개발하였다.

3. 얼굴 영역 추출

얼굴 영역을 추출하는 데에 무엇보다도 가장 큰 어려움은 얼굴 모습의 다양한 변화라 할 수 있다. 카메라로부터의 거리, 얼굴의 회전(rotation), 이동(translation), 크기(scale)의 변화에 따라 얼굴의 모습이 변할 뿐만 아니라 안경, 턱수염, 콧수염, 머리 스타일에 따라 얼굴의 모습 자체가 다양하다는 것이 일반적인 얼굴 영역의 추출에 어려움을 준다. 또한 복잡한 배경, 다양한 조도 환경 하에서 얼굴 영역의 추출이 이루어져야 한다는 제약이 있다. 현금 인출기 시스템은 위에서 언급한 문제 이외에 사용자가 서비스를 받는 짧은 시간 내에 얼굴영역을 추출하고, 식별할 수 있어야 한다는 시간적인 제약이 있다. 일반적인 현금 인출기의 경우 내부에 Pentium-II 500Mhz 이상 CPU 및 WindowNT기반의 Embedded PC를 내장하고 있다. 현금 인출기 개발 업체의 조사에 따르면(그림 2)의 “카드 투입”에서 “금융 거래 진행” 단계(현금 인출기에 금융 거래 메뉴창이 뜨는 단계)까지 최대 1초 이내에 이루어져야 사용자가 지루함을 느끼지 않는 것으로 알려져 있다. 이와 같은 사용자 요구 분석(User Requirement Spec.)에 따라 본 논문에서는 1초 이내에 얼굴 특징 식별이 끝나는 알고리즘을 개발하였다.

입력 영상에서 얼굴 영역을 추출하기 위한 기존의 많은 연구들이 수행되었다. 입력 영상에서 얼굴 영역을 추출하는 방법으로는 크게 정지 영상에서 추출하는 방법과 연속 영상에서 추출하는 방법으로 나눌 수 있다. 정지 영상에서 얼굴을 추출하는 방법은 타원 허프 변환(elliptic Hough transform)[3-5], 수평 및 수직 투영 해석(horizontal and vertical projection analysis)[6, 7], 스네이크(snake)[8-10], 가변 템플릿(deformable template)[11, 12] 및 신경망[2]을 이용하는 방법 등이 있으며, 각 방법들은 다음과 같은 장·단점을 지니고 있다[2]. 타원 허프 변환 방식은 처리 시간이 비교적 오래 걸리고 배경의 에지 성분이 많은 경우 제대로 얼굴 영역을 찾지 못한다는 단점이 있다. 투영 해석법에 의한 방법은 영상의 밝기 변화에 대하여 비교적 영향을 덜 받으며 계산량이 적지만, 배경이 복잡한 영상에서는 그대로 적용하기 어려운 단점이 있다. 스네이크를 이용하는 방법은 앞에서 설명한 얼굴 영역 추출 방법과는 달리 얼굴 영역의 분할(segmentation) 및 추적(tracking)이 동시에 가능하다. 스네이크는 물리적인 힘을 바탕으로 자율적으로 움직이는 가변 곡선으로서, 이때 곡선을 움직이는 외부 에너지(External Energy)와 내부 에너지(Internal Energy)를 정의하여 에너지의 합이 최소가 되는 방향

으로 움직이면서 얼굴 영역을 검출한다. 이러한 스네이크에 의한 얼굴 영역 추출은 비정형적인 움직임에 대한 적응력과 실시간 처리의 강점을 지니고 있으나, 초기 위치에 대한 의존이 크므로 전역 탐색이 불가능하며, 잡음에 약한 단점이 있다. 이외에도 가변 템플릿에 의한 방법은 기존의 템플릿 정합 방법과 같이 고정된 템플릿에 대해 가장 유사한 입력 패턴을 찾는 것이 아니라 주어진 패턴에 일치하도록 템플릿을 변형시키는 방법이다. 가변 템플릿은 비교적 정형화된 형태를 가진 눈, 입 등의 얼굴 특징점을 찾는 데는 우수한 성능을 나타내지만, 사람에 따라 형태가 다양한 얼굴 영역을 찾는 데는 정확도가 많이 떨어진다. 또한 전역 탐색이 이루어질 수 없고, 계산량이 많아 실시간 처리를 하는 응용 분야에는 적합하지 않다. 신경망을 이용하여 얼굴 영역을 찾는 방법은 다양한 얼굴 표정 및 회전 각도를 지닌 얼굴들을 역전파 신경망(back-propagation neural network)에 의해 학습시키고, 입력 영상으로부터 사용자의 얼굴을 둘러싸는 사각형을 찾는 방법이다. 앞의 방법들보다 비교적 다양한 배경과 조도를 지닌 복잡한 영상에서 얼굴을 정확하게 찾는다는 장점은 있으나, 탐색 영역이 넓은 경우 처리시간이 오래 걸릴 수 있다는 단점이 있다. 이외에도 얼굴 색상의 RGB를 그대로 이용하거나 혹은 RGB 화소 값을 다른 칼라 좌표계로 변환하여 얼굴 영역을 검출하는 연구들이 있다. 두 장 이상의 연속 영상을 이용하여 얼굴 영역을 검출하는 방법에는 미리 저장된 배경영상과 현재의 입력 영상과의 차이를 이용하여 얼굴 영역을 검출하는 방법이 있다[13]. 이 방법은 미리 저장된 배경영상과 현재의 입력 영상과의 조도 변화가 생기거나 혹은 배경에 다른 고정 물체가 생기는 경우 제대로 얼굴 영역이 검출되지 않는 단점이 있다. 이외에도 2장의 연속 영상의 차영상을 이용하는 방법이 있다. 이는 사용자의 얼굴이 두 장의 연속 영상에서 약간이라도 움직임이 있다는 가정 하에 손쉽게 사용할 수 있는 방법이다. 정지영상에서만 얼굴을 찾는 것은 일반적으로 두 장 이상의 영상 정보를 사용하는 것보다는 처리 시간이 오래 걸리고 얼굴 영역을 찾기가 어렵다.

이 논문에서는 위와 같은 얼굴 추출 연구들 중에서 빠른 시간 내에 정확하게 얼굴을 찾을 수 있도록 얼굴의 움직임 정보를 이용한 방법, 얼굴의 색깔 정보를 이용한 방법들을 혼합하여 사용하였다.

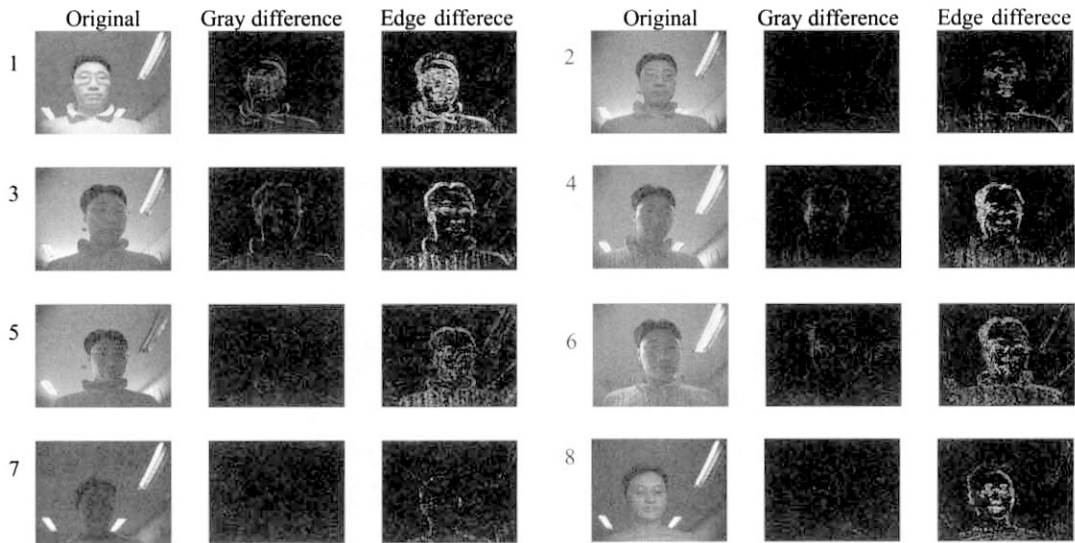
3.1 얼굴의 움직임 정보를 이용한 방법

얼굴의 움직임 정보를 이용하는 방법은 사용자가 현금 인출기에 카드를 투입할 때 취득된 연속된 두 영상의 차이로부터 움직임이 있는 부분을 얻고 이를 이용하여 얼굴영역을 추출하는 방법이다. 입력된 두 영상의 차이를 이용할 때, 일반적인 그레이 영상(Gray Image)의 차이를 이용할 수도 있고 먼저 영상의 에지 부분(Edge Region)추출하여 두 경계영상의 차이를 이용할 수도 있다. 조도가 밝을 때에는 그레이 영

상의 차이를 이용한 방법과 에지 영상의 차이를 이용한 방법 모두 좋은 성능을 발휘하며 움직이는 물체를 추출할 수 있다. 그러나 조도가 어두울 때는 영상의 그레이 값이 전체적으로 낮아져서 연속된 두 프레임의 그레이 값의 차를 구할 경우 그 값이 매우 작게 나타나며, 이러한 경우 영상의 잡음에 의한 것과 움직임에 의해서 생기는 부분과의 차이가

거의 나지 않는다. 이와 같이 조도가 어두운 경우에는 그레이 값의 차이를 이용한 방법보다 에지 영상의 차이를 이용한 방법이 훨씬 더 좋은 성능을 나타내는 것을 알 수 있었다.

다음 (그림 3)은 외부 조명 및 현 논문에서 사용하는 추가 조명의 밝기를 변화시켜 가면서 그레이 영상 차이 및 에지 영상 차이를 나타낸 것이다.



(그림 3) 조명 변화에 따른 영상 차이

(그림 3)의 1과 8은 사용자 정면에 아주 밝은 조명을 직사했을 경우이고, (그림 3)의 2는 사용자에게 불쾌감을 주지 않을 정도의 최소한의 조명을 직사했을 경우이다. 이외의 경우는 추가 조명을 사용하지 않았을 경우이다. 그림에서 알 수 있듯이 연속된 그레이 영상의 차이를 이용하는 것보다는 연속된 에지 영상을 이용하는 것이 모든 경우에서 얼굴 영역 추출에 더 좋은 성능을 보인다. 특히 어두운 경우에는 연속된 그레이 영상을 이용해서 얼굴 영역을 추출하기 어렵다는 결과를 보인다.

그런데, 주변광으로 많이 사용하는 형광등은 시간이 변함에 따라 약간의 밝기 변화가 생기는데 이는 에지 영상의 차에 더욱 두드러지게 노이즈 성분으로 나타나는 문제점이 생긴다. 그러므로 본 논문에서는 이와 같이 연속된 에지 영상의 차영상에서 노이즈 문제점을 해결하기 위해서 노이즈 제거 필터(3×3 median filter)를 사용하였다. 이외에도 너무 짧은 시간 간격동안 취득된 영상에서는 사용자의 얼굴 움직임이 발생하지 않을 수 있으므로, 취득 영상의 시간 간격을 적응적으로 변화시켜가면서 정해진 임계치이상으로 움직임이 발생하는 두 영상을 선택하여 얼굴 영역 추출에 사용하였다. 또한 카메라 설치 위치를 사용자 얼굴보다 낮은 곳에 설치하여 얼굴을 올려다보듯이 영상을 취득함으로써, 배경에 사용자 이외에 움직이는 물체의 영향을 최소한으로 줄일 수 있었다.

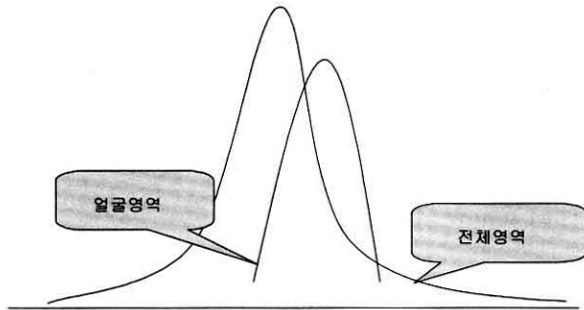
(그림 3)과 같이 얻어진 에지 영상의 차영상에 노이즈 제

거 필터를 적용한 후, 차영상의 수직/수평 히스토그램 투영법을 이용하여 얼굴 박스의 위치를 구한다. 진술한 바와 같이 본 시스템에서는 카메라를 사용자 얼굴 보다 낮은 곳에 설치하여 얼굴을 올려다보듯이 영상을 취득함으로써, 배경 특히 얼굴 윗부분에는 차영상에서의 움직임을 나타내는 배경 물체가 거의 없게 된다. 이러한 특성을 이용하여 차영상으로부터 먼저 Y축(수직) 방향의 차영상 히스토그램을 구하고, 영상의 제일 상단으로부터 히스토그램 값을 검사하다가 일정값이상 커지는 부분을 얼굴(머리)의 맨 윗 위치를 찾는다. 이후 이를 기준으로 얼굴의 왼쪽 및 오른쪽 경계 위치를 X축(수평) 방향의 차영상 히스토그램을 이용하여 같은 방식으로 찾는다. (그림 3)에서 알 수 있듯이 차영상 정보에서는 턱 선의 위치가 잘 나타나지 않으므로, 본 논문에서는 찾아진 얼굴의 윗부분/왼쪽/오른쪽 위치 등을 기본으로, 일반적인 얼굴의 가로대 세로 비 정보를 바탕으로 대략적인 얼굴 박스 위치를 설정하게 된다.

3.2 얼굴의 색깔 정보를 이용한 방법

얼굴의 색깔 정보를 이용하는 방법은 얼굴영상에서 사람의 피부색에 해당하는 칼라범위를 많은 사람들의 얼굴영상 샘플로부터 구한 후, 입력영상에서 각 픽셀들이 얼굴에 해당하는지 여부를 판단하여 얼굴영역을 추출하는 방법을 말한다. 일반적으로 색상을 표현하는 방식중의 하나인 YIQ

방식에서 I성분은 살색에 민감한 특성을 지니고 있다. 따라서 본 논문에서는 RGB칼라를 YIQ방식으로 변환해서 살색 검출에 유리한 I성분을 이용해 얼굴 영역을 검출하였다 [14].



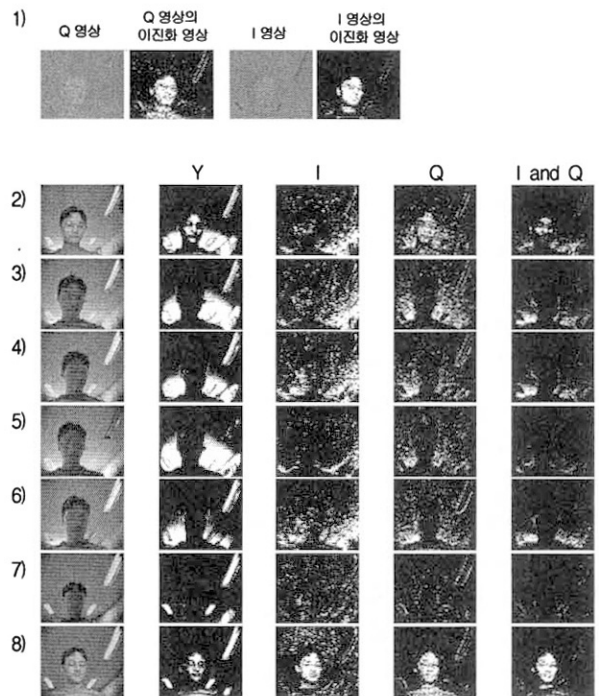
(그림 4) 얼굴의 I 성분 분포

다양한 조도에서 얼굴 영역의 I 성분의 분포를 살펴보면 약 130~150 정도의 분포(그림 4)를 가지고 있다. RGB 칼라를 YIQ 칼라로 변경하게 되면 영상의 밝기 정보는 Y성분에, 그리고 칼라 정보는 I와 Q성분에 포함된다. 이중, I는 오렌지/붉은색 계열의 정보를 주로 나타내며, Q성분은 청록색 계열의 정보를 나타내게 된다[14]. 본 시스템에서 대상으로 하는 동양인의 얼굴색은 일반적으로 붉은색 계열을 많이 나타내므로, (그림 4)와 같이 I 성분 분포에서 다른 전체 영역에 비해 상대적으로 큰 값을 나타내게 된다.

(그림 5)는 다양한 실험 환경에서 I 영상을 이용한 얼굴 영역 추출 실험 결과를 나타낸다. (그림 5)의 1)은 RGB 칼라 형태의 입력 영상을 YIQ로 변경한 후, Q와 I 영상 및 각각의 영상을 이진화한 영상의 예를 나타낸 것이다. 그리고 (그림 5)의 2)~8) 역시 입력 영상을 YIQ로 변경한 후, Q와 I 영상 및 각각의 영상을 이진화한 영상을 나타낸 것이다. 이때 I and Q 영상은 각각 이진화 된 I 영상과 Q영상을 AND 시켜서 얻은 영상을 나타낸 것이다.

(그림 5)의 2)는 사용자에게 불쾌감을 주지 않을 정도의 최소한의 조명을 직사했을 경우이며, (그림 5)의 1)과 8)은 사용자 정면에 아주 밝은 조명을 직사했을 경우이다. 이외의 경우는 추가 조명을 사용하지 않았을 경우이다. 실험 결과 (그림 5)의 2)와 같이 추가조명이 조금 어두운 경우 I영상만으로는 얼굴 영역이 거의 나타나지 않는 문제가 있었다. 이를 보완하기 위해서 YIQ성분에서 지금까지 사용한 I성분에 Q성분을 합해서(AND) 실험을 하였다. I성분에 Q성분을 합한 실험 결과는 거의 모든 경우에서 I성분만을 이용한 것보다 얼굴 영역 검출이 더 잘 되는 결과를 나타냈다. 이러한 원인에 대해 분석해 보면 다음과 같다. 전술한 바와 같이 I 성분은 오렌지/붉은색 계열의 정보를 포함하고 있으나, (그림 5)의 2), 3), 4), 6)과 같이 사용자 얼굴 뒤에 주변광(태양광)이 들어올 때에는 이 역시 오렌지/붉은 계열의 정보를 포

함하게 되므로, 사용자의 얼굴과 구별이 어렵게 되는 현상을 나타낸다. 그러나 이때 Q성분을 같이 고려하여 (비록 Q 성분 자체가 청록 계열의 색을 많이 나타내지만) 사용자 얼굴에 대한 Q성분 범위(하향 및 상향 임계치)를 넓게 잡으면, 이 범위 내에는 얼굴 정보는 포함되지만 주변광(태양광) 정보는 많이 제거될 수 있게 된다((그림 5)의 2), 3), 4), 6) 참조). 이는 동양인의 얼굴색이 주변광(태양광)보다는 더 많은 Q성분을 포함하고 있기 때문이다. 그러므로 I성분과 Q성분을 합쳐서 처리한 결과, 보다 정확한 얼굴 영역을 추출할 수 있는 결과를 나타냈다.



(그림 5) I 및 Q를 이용한 얼굴 검출

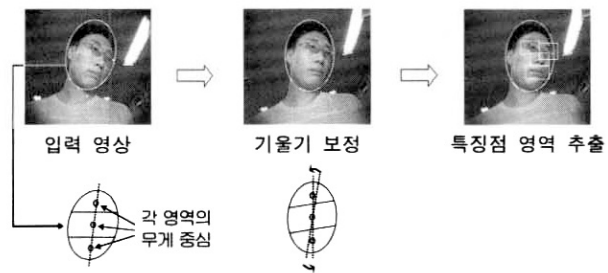
이때 얼굴과 같은 색의 옷을 입고 있는 경우에는 I나 Q영상에서 얼굴 영역이랑 유사하게 검출되는 현상을 나타낸다. 이러한 문제점을 해결하기 위하여, 본 시스템에서는 (그림 5)의 1), 2) 및 8)과 같이 보조 조명을 이용하는 방법을 사용한다. 일반적으로 사람 얼굴 표면의 반사율(reflectance ratio)은 옷에서의 반사율보다 높기 때문에, 조명을 비추었을 때 (그림 5)의 1), 2) 및 8)와 같이 얼굴 표면에서의 I나 Q값이 옷에서의 값보다 상대적으로 높게 나타난다. 그러나 전술한 바와 같이 사용자의 얼굴 정면에 조명을 비추게 되면 눈부심 현상과 함께 불쾌감을 야기하기 때문에, 본 시스템에서는 입력 영상의 I 및 Q성분의 AND영역을 처리하여 얼굴 살색과 유사한 영역이 일정 임계치 이상 크게 나타나는 경우(살색형태의 옷 등을 입은 경우)에만 추가 조명을 On시킴으로써 정확한 얼굴 영역을 추출하는 방법을 사용하였다.

3.3 얼굴의 움직임 정보 및 색깔 정보를 혼합한 얼굴 영역 추출

3.2에서 설명한 얼굴의 움직임 정보에 의해 얼굴로 검출된 부분과 IQ 칼라 좌표계에서 얼굴로 검출된 부분에 대한 공통부분(AND)을 택함으로써 입력되는 얼굴 영상으로부터 빠르고 정확하게 얼굴영역을 검출할 수 있다. 움직임 정보 영상과 차영상은 각각 영상 확장(dilation)과 라벨링(labeling)을 거쳐 고립된 노이즈 성분(isolated noise)을 제거한 영상으로 변환 후, 수직, 수평 에지 투영법에 의해 얼굴을 포함한 사각형 영역을 추출하였다[14].

4. 기울기 보정 및 얼굴 특징점 추출

일반적으로 현금 인출기 시스템에서 사용 시, 사용자가 전면부의 메뉴를 쳐다보는 경우에는 얼굴이 기울어지지 않지만, 카드 투입구를 쳐다보는 경우에는 얼굴이 기울어지게 된다. 본 논문에서는 얼굴 및 얼굴 특징의 기하학적인 특성을 바탕으로 눈, 코, 입 등의 탐색 영역을 설정하므로 기울어진 얼굴의 기울기 보정은 필수적으로 요구된다. 얼굴의 기울기 보정을 위해서는 다음과 같은 처리 과정을 거친다. 앞에서 설명한대로 차영상과 칼라 정보를 통해 구해진 얼굴 박스 영역 내에서, 차영상정보를 이용하여 보다 정확한 얼굴 타원 영역을 찾는다. 이후, (그림 6)과 같이 추출된 얼굴 타원 영역을 수직 방향으로 3등분하여 각각의 무게중심 지점을 연결하는 선의 기울기를 구한 후, 이 선분을 수직으로 만드는 형태로 얼굴 타원 영역을 회전시킴으로써 얼굴에 대한 기울기 보정을 하게 된다. 얼굴 영역의 기울어짐을 보정한 후 얼굴 특징점의 존재 가능 영역을 선택한다. 다음 (그림 6)은 얼굴의 기울기 보정 및 얼굴 특징점의 존재 가능 영역을 나타낸 것이다.



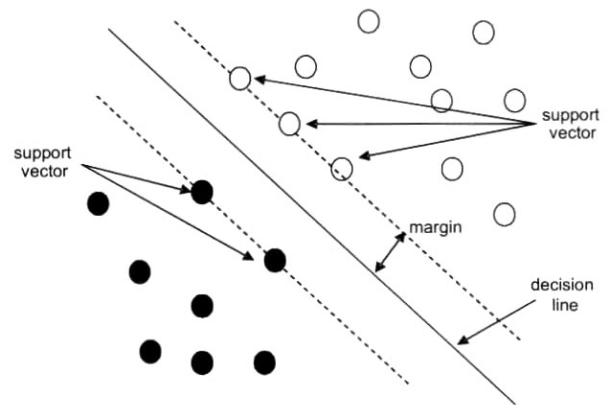
(그림 6) 얼굴의 기울기 보정 및 얼굴 특징점의 존재 가능 영역 설정

5. SVM을 이용한 얼굴 특징점 식별 가능 여부 판단

설정된 얼굴 특징점의 존재 가능 영역으로부터 추출된 I와 Q영상 및 SVM(Support Vector Machine)을 이용하여 식별 가능 여부를 판단하게 된다. 본 논문의 경우 전술한 바

와 같이 전면부에 약간의 조명을 사용하였으므로, 얼굴 특징점 부근에 그림자가 생기는 경우가 거의 발생하지 않아서 조명 정규화가 별도로 필요하지는 않았다. 게다가 입력 칼라 영상에서 조도에 민감한 Y성분을 제거한 I와 Q영상을 이용하였으므로, 외부 조명의 영향이 더욱 낮아짐을 알 수 있었다.

일반적으로 SVM은 패턴 인식 문제를 해결하기 위해 비교적 최근에 제안된 방법으로써[15-17], (그림 7)과 같이 Support Vector라고 불리는 training set point들에 의해 결정되는 결정 경계(decision boundary)를 결정함으로써 2개의 classification 문제에 주로 많이 사용되고 있다.



(그림 7) 2차원 공간에서의 SVM의 결정경계영역

이때, 결정 경계는 두 클래스간의 최대 거리를 유지하도록 결정된다. SVM의 기본 아이디어는 구조적 리스크 최소화화를 통해 벡터공간에서의 최적의 결정경계영역을 찾아내는 것으로 이진분류문제를 푸는 방법으로 이용되고 있다. 일반적으로 SVM의 분류 성능(classification performance)은 기존에 많이 사용하는 다층 신경망(Multi-layered Perceptron)보다 우수한 것으로 알려져 있다. 특히, positive와 negative sample의 수가 충분치 않고 입력 데이터에 잡음 성분을 많이 포함하고 있는 경우에 더욱 우수한 분류 성능을 나타내고 있다. 또한 다층 신경망의 경우 사용자가 경험적으로 선정해야 할 많은 초기 설정 파라미터들을 포함하고 있으며, 이러한 파라미터들의 설정에 따라 분류 성능이 달라지는 결과를 나타낸다.

본 논문에서는 얼굴 특징 탐색 영역을 위해 I와 Q영상으로부터 Gray영상을 취득한 후, 크기 정규화를 통해 각각 30×30 픽셀 크기(눈 영역), 10×10 픽셀 크기(코 영역) 그리고 60×20 픽셀 크기(입 영역)의 영상을 사용하였다. 일반적으로 본 논문에서와 같이 SVM에서 입력 데이터의 차원이 큰 경우 비선형 분리 문제(non-linearly separable problem)가 되어 고차원 커널을 사용해야 하는 것으로 알려져 있기 때문에, 본 논문에서는 5차원 다항식 커널(polynomial kernels of degree 5)을 사용하였다.

기타, 다른 커널(RBF, MLP, Splines, B-Splines)들의 사

용이 support vector의 분류 성능에는 영향을 주지 않는 것으로 알려져 있으며[17], 본 논문에서도 다른 종류의 커널들을 사용했을 때와 성능 비교를 해본 결과 거의 차이를 나타내지 않는 것으로 조사되었다. SVM의 출력 노드는 정상적인 얼굴 특징과 비정상적인 얼굴 특징(선글라스 등으로 가려진 눈, 마스크 등으로 가려진 코 및 입)의 2 클래스 문제로 정하였다. 반면, C값은 SVM의 성능에 영향을 많이 미치므로, 본 논문에서는 실험 결과 가장 우수한 성능을 나타내는 값인 10000을 C값으로 사용하였다.

6. 실험 결과

학습 데이터를 위해서는 다양한 사람들(20명분)로부터 정상적인 경우와 비정상적인 경우(선글라스, 마스크 등을 착용한 경우 포함)를 모두 포함한 약 200개의 영상을 취득하여 사용하였으며, test 데이터는 학습에 참가하지 않은 사람들(10명분)로부터 150개의 영상을 취득하여 사용하였다. 입력 영상은 외부광의 영향을 알기 위하여 다양한 시간대에 취득하였다.

실험 결과, 12개의 positive support vectors와 51개의 negative vector들이 선택되었다. 일반적으로 support vectors는 학습 데이터들 중에서 분류되기 어려운 데이터를 의미하므로, 이 실험 결과 입력 데이터가 많은 잡음 성분을 포함하고 있으며 본 논문에서의 문제가 상당히 분류하기 어려운 문제라는 것을 의미한다. 실험 결과는 크게 두 가지로 측정하였다. 오인식율(FAR: False Acceptance Rate)과 오거부율(False Rejection Rate)이다. 오인식율은 비정상적인 눈인데 정상적인 눈으로 잘못 받아들이는 경우를 나타내며, 오거부율은 정상적인 눈을 비정상적인 눈으로 거부하는 에러를 나타낸다. 다음 <표 1>은 눈, 코, 입 각각의 영역에 대한 실험 결과이다.

<표 1> SVM을 이용한 실험 결과

	학습 데이터(200개)		Test 데이터(150개)	
	오인식율	오거부율	오인식율	오거부율
눈 영역 판단	0/100 = 0%	1/100 = 1%	1/80 = 1.25%	1/70 = 1.43%
코 영역 판단	0/100 = 0%	0/100 = 0%	0/80 = 0%	0/70 = 0%
입 영역 판단	1/100 = 1%	1/100 = 1%	1/80 = 1.25%	0/70 = 0%
전체 판단	1/100 = 1%	2/100 = 2%	2/80 = 2.5%	1/70 = 1.43%

다음 <표 2>에서는 기존의 다층 신경망(MLP)을 이용한 학습 및 test 결과이다.

<표 1>과 <표 2>에서 알 수 있듯이 모든 경우에서 MLP 보다는 우수한 성능을 나타냄을 알 수 있다. 또한 SVM을

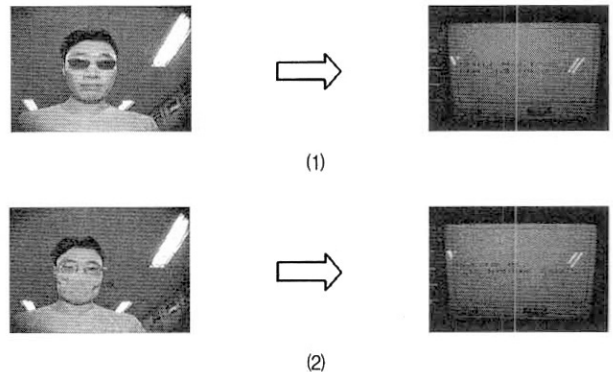
이용한 분류 시간 역시 Pentium-II 550Mhz에서 약 13ms정도 소요됨을 알 수 있었다. 영상 입력 장치로는 SONY 1/3" CCD가 포함된 Pinhole 카메라(SCM-2AXI)와 Video Blaster Image Capture Board를 사용했다.

<표 2> 다층 신경망을 이용한 실험 결과

	학습 데이터(200개)		Test 데이터(150개)	
	오인식율	오거부율	오인식율	오거부율
눈 영역 판단	2/100 = 2%	1/100 = 1%	1/80 = 1.25%	2/70 = 2.86%
코 영역 판단	1/100 = 1%	2/100 = 2%	1/80 = 1.25%	1/70 = 1.43%
입 영역 판단	2/100 = 2%	1/100 = 1%	2/80 = 2.5%	1/70 = 1.43%
전체 판단	5/100 = 5%	4/100 = 4%	4/80 = 5%	4/70 = 5.71%



현금 인출기 메인 화면



(그림 8) 현금 인출기에서의 얼굴 식별 구동 예

각각의 얼굴의 특징점의 식별 여부를 판단한 후 현금 인출기 시스템에 금융 거래의 가능 여부를 전달한다. (그림 8)은 실제 현금 인출기에 본 논문에서 제안한 얼굴 식별 시스템을 탑재하여 구동시킨 예를 나타낸 것이다. (그림 8)의 제일 상단에 나타나 있는 화면은 초기 현금 인출기의 메인 화면을 나타낸 것이고, (1)과 (2)는 각각 선글라스나 마스크 등으로 얼굴을 가린 경우에 처리 화면의 예를 나타낸 것이다.

즉, 눈 식별이 불가능하면 현금 인출기 시스템은 (그림 8)(1)의 오른쪽 화면과 같은 메시지("모자나 선글라스를 벗고 다시 거래하여 주십시오")를 내보내며, 코·입의 식별이 불가능하면 (그림 8)(2)의 오른쪽 화면과 같은 메시지("마스크를 벗고 다시 거래하여 주십시오")를 보낸다. 본 시스템에서는 이러한 얼굴 특징점 중 하나라도 제대로 보이지 않으면 얼굴이 식별 불가능한 것으로 간주한다. 물론 이러한 사

향은 향후 field test 및 은행의 시스템 관리자와의 협의하여 조정 가능하다.

7. 결 론

최근 무인 현금 인출기의 사용 확대와 함께, 타인의 신용 카드를 이용하여 무인 현금 인출기에서 돈을 인출하는 금융 범죄에 대한 예방 대책이 필수적으로 요구되고 있다. 특히 무인 현금 인출기 부근에는 감시용 CCTV Camera가 설치되어 있으나, 지능적인 범죄자들은 이러한 사실을 인식하고 선글라스, 마스크 등을 착용하여 이러한 감시 시스템을 피해가고 있다. 본 논문에서는 이러한 문제점을 해결하기 위하여 무인 현금 인출기에 설치되어 있는 카메라를 통해 입력된 사용자의 얼굴 및 얼굴 특징점을 영상 신호처리 방법과 SVM (Support Vector Machine)으로 분석하여 향후 얼굴을 식별 가능한 경우에만 금융 거래를 할 수 있도록 하는 시스템을 개발하였다. 실험 결과, 학습 데이터에 대해서는 약 1%의 오인식율과 2%의 오거부율을 나타냈으며, Test 데이터에 대해서는 약 2.5%의 오인식율과 1.43%의 오거부율을 나타냈다.

참 고 문 헌

[1] <http://www.toshiba.com>.
 [2] 김정환, "다층 퍼셉트론을 이용한 얼굴 영역 분할 및 특징부 추출에 관한 연구", 연세대학교 대학원 전자공학과 석사학위 논문, 1996.
 [3] Saad Ahmed Sirohey, "Human face segmentation and identification," Technical Report(CAR-TR-695) of Center for Automation Research, University of Maryland, Nov., 1993.
 [4] G. Chow, X. Li, "Towards a system for automatic facial feature detection," Pattern Recognition, Vol.26, No.12, pp.1739-1755, 1993.
 [5] 이미숙, 김남진, 이 호, 임희석, 이성환, "얼굴 영상의 이해를 위한 얼굴 특징의 자동 추출", HCI '96 학술대회발표논문집, pp.205-219, 1996.
 [6] T. Kanade, "Picture processing by computer complex and recognition of human faces," Technical Report of Dept. Inform. Science, Kyoto Univ., 1973.
 [7] Roberto Brunelli, Tomaso Poggio, "Face Recognition : Features versus Templates," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.15, No.10, pp.1042-1052, Oct., 1993.
 [8] D. Terzopoulos, Keith Waters, "Analysis and Synthesis of Facial Image Sequences Using Physical and Anatomical Models," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.15, No.6, pp.569-579, June, 1993.
 [9] Akikazu Takeuchi, Steven Frank, "A Rapid Face Construction Lab.," Technical Report(SCSL-TR-92-010) of

Sony Computer Science Laboratory Inc., May, 1992.
 [10] R. Magnolfi, P. Nesi, "Analysis and Synthesis of Facial Motions," Proc. of International Workshop on Automatic Face and Gesture Recognition, Zurich, pp.308-313, 1995.
 [11] A. Yulle, P. Hallinan, "Deformable Templates," Active Vision, The MIT Press, London, England, 1993.
 [12] Marcus E. Hennecke, K. Venkatesh Prasad, David G. Stork, "Using Deformable Templates to Infer Visual Speech Dynamics," Technical Report of the Dept. of Electrical Engineering, Stanford University.
 [13] Jun Rekimoto, "A Vision-Based Head Tracker for Fish Tank Virtual Reality -VR without Head Gear-," Technical Report of Sony Computer Science Lab., 1995.
 [14] William L. Pratt, "Digital Image Processing," 1st Edition, Addison Wesley, 1993.
 [15] Vapnik, "The Nature of Statistical Learning Theory," New York : Springer Verlag, 1995.
 [16] Vapnik, "Statistical Learning Theory," Wiley-Interscience publication, 1998.
 [17] Saunders, "Support Vector Machine User Manual," RHUL, Technical Report, 1998.



박 강 령

e-mail : parkgr@smu.ac.kr

1994년 연세대학교 전자공학과(학사)

1996년 연세대학교 전자공학과(공학석사)

2000년 연세대학교 전기·컴퓨터공학과

(공학 박사)

2000년~2003년 LG전자기술원 홍채인식팀

(선임연구원)

2003년~현재 상명대학교 소프트웨어대학 미디어학부 전임강사
 관심분야 : 영상처리, 컴퓨터 비전, 컴퓨터 그래픽스

김 재 희

1982년 Case Western Reserve Univ. Electrical Eng. 석사

1984년 Case Western Reserve Univ. Electrical Eng. 박사

1984년~현재 연세대학교 전기·전자공학과 교수

2002년~현재 생체 인식 연구 센터(BERC) 소장

관심분야 : Biometric 영상처리, 패턴인식, 컴퓨터 vision

이 수 연

e-mail : leesy@daisy.kw.ac.kr

1969년 광운대학교 정보통신공학과 학사

1977년 연세대학교 전자공학석사

1983년 일본교토대학교 정보공학박사

1973년~현재 광운대학교 컴퓨터공학부 교수

관심분야 : SGML/XML, XML, 전자상거래