

Multi-Camera를 이용한 인터넷 기반의 지능적 감시 시스템

정도준⁰, 이창우, 김향준
경북대학교 컴퓨터공학과

Intelligent Surveillance System with Multi-Camera on the Internet

Do Joon Jung,⁰ Chang Woo Lee, Hang Joon Kim
Dept. of Computer Engineering, Kyungpook National University
{djjung⁰, cwlee, kimhj}@ailab.knu.ac.kr

요약

본 논문에서는 multi-camera를 이용한 인터넷 기반의 지능적 감시 시스템을 제안한다. 제안된 시스템은 두 종류의 카메라, static camera와 pan-tilt camera, 를 이용하여 출입구를 감시하고, 비인가자를 추적한다. static camera는 출·입을 검출하고 출입자를 인가자와 비인가자로 분류하는데 이용되고, pan-tilt camera는 비인가자로 분류된 출입자를 추적하는데 이용된다. 제안된 시스템은 세가지 단계: 출입구 감시, 출입자 검출 및 분류(인가자/비인가자), 비인가자 추적으로 구성된다. 출입구 감시는 출입문의 밝기값 변화를 이용한다. 출입자 검출 및 분류는 skin color 모델과 얼굴 크기, 위치와 관련된 휴리스틱을 이용하여 얼굴을 검출하고, PCA(Principal Component Analysis)를 이용한 eigenspace상에서의 유클리디언 디스턴스로 템플릿 얼굴과 입력 얼굴의 유사도를 계산하여 인가자인지 비인가자인지 분류한다. 비인가자 추적은 pan-tilt 카메라를 이용하여, static camera에서 분류된, 비인가자의 움직임을 검출하고 카메라를 제어함으로써 추적한다. 제안된 시스템은 무인 감시 상황에서 비인가자의 출입시 감시자에게 경고 신호를 제공하고, 감시지역에서 사건 발생시, 사건의 개요를 파악하는 중요한 정보를 빠른 시간에 제공할 수 있다는 장점을 가진다.

I. 서론

보안 감시 시스템은 범죄 활동의 가능성이 큰 교도소, 은행, 현금자동출납기, 주유소, 지하주차장, 엘리베이터 등에서 보안의 목적으로 많이 사용되고 있으며, 의심스러운 행동을 관찰할 수 있도록 영상을 실시간으로 보여주고,

녹화한다. 보안 감시자는 감시 시스템의 모니터를 통해 종종 몇몇 장소를 동시에 감시 해야 하므로, 이런 일은 노동 집중적이고 비효율적이며, 단순 루틴의 반복으로 감시가 소홀해 질 수 있다.

기업 및 국가 부설 연구소, 국방부 및 군 주요 시설, 은행 대여 금고 등과 같이 출입이 허가된 사람과 그렇지 않은 사람이 명확히 나누어 지는 곳에서는 허가되지 않은 사람의 출·입에 대해서만 경고 신호를 줄 수 있다면, 보안 감시자의 임무를 크게 줄여주고, 의심스러운 활동이 검출 되었을 때, 빠르고 효과적인 조치 활동을 가능하게 한다. 또한, 지능적인 원거리 모니터링 시스템은 사용자가 멀리 떨어진 위치에서 감시가 가능하도록 해주며, 몇몇 장소에서 보안 감시가 동시에 필요할 때도 특히 유용할 수 있다.

S.Stillman은 pan-tilt zoom camera를 이용하여 움직이는 여러 사람을 추적하는 시스템을 개발하였고 [1], J.Yang은 하나의 camera를 이용하여 움직이는 사람을 추적하는 시스템을 개발 하였다 [2]. A.C.M.Fong은 감시 시스템을 웹에 올리려는 노력을 하였다 [3]. S.Stillman이나 J.Yang 의 시스템은 원거리 감시가 곤란하고, A.C.M.Fong의 시스템은 감시자가 항상 화면을 모니터링 하여야 한다는 단점이 있다.

이런 단점을 해결 하고자, 본 논문에서는 multi-camera를 이용한 인터넷 기반의 지능적인 감시 시스템을 제안한다. 제안된 시스템은 카메라가 설치된 원격지에서 일어나는 출·입 상황과 비인가자에 대한 추적상황을 인터넷 상에서 볼 수 있으며, 출·입 감지된 시점의 영상과 시간을 저장하여 제공함으로써, 무인 감시 상황에서 사건 발생시 효과적인 조치가 가능할 수 있다.

II. 제안된 시스템

제안된 시스템은 제한된 공간의 출·입 상황과 비인가자에 대한 추적상황을 인터넷상에서 실시간으로 볼 수 있고, 녹화영상도 검색하여 볼 수 있는 시스템이다. static camera와 pan-tilt camera로부터 입력된 영상과 출입 감지된 시점의 영상 및 시간은 그림 1 에서와 같이 서버로 전송되어 인터넷상에서 접속하는 client 이용자에게 제공된다. 제안된 시스템은 출입구 감시, 출입자 분류, 비인가자 추적의 3단계로 구성된다.

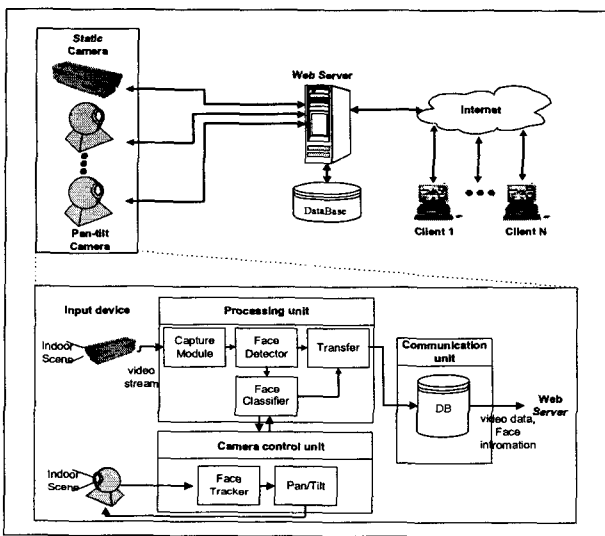


그림 1. 제안된 시스템의 개요도

1. 출입구 감시

출입구 감시 단계에서는 고해상도의 static camera로부터 입력된 영상에서 사람의 출·입을 검출하고, 출입구로 들어오고 나가는 모든 출입자 영상과 그때의 시각을 기록한다. 출입구에서 사람의 출·입을 구분하기 위해서 그림 2와 같이 출입구에 2개의 가상 센서를 두고(위와 아래), 가상 센서의 밝기값 변화를 감지하여 출입자의 출·입을 구분 하며 각각의 영상과 시간을 저장한다.

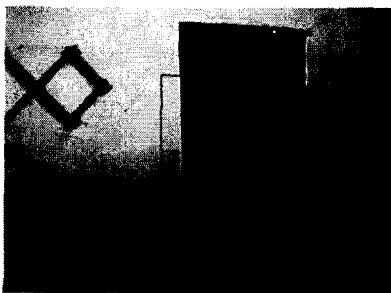


그림 2. 출입문의 가상센서2개(위, 아래 사각형)와 출입자 출현 후보 위치(가운데 사각형)

2. 출입자 검출 및 분류

출입자 검출 및 분류 단계에서는 출입구로 들어오는 사람의 얼굴을 검출하고 인가자인지 비인가자인지 분류한다. 출입구 감시 단계에서 사람이 들어온다고 감지된 경우, 출입자 출현 후보 위치 영역 내에서 skin color 모델을 이용하여 얼굴을 검출한다. 사람 얼굴의 칼라 분포는 색채 공간의 작은 영역에서 군집을 이루고, 2D-Gaussian 분포로 근사화 될 수 있다 [2]. 따라서 skin color 모델은 색채 칼라 공간에서 얼굴 피부 색깔의 r, g 성분이 2D-Gaussian 분포를 따른다고 가정하고 근사화 시킨 모델이며, 아래 수식 (1) 과 같이 표현된다.

$$r = \frac{R}{R+G+B}, \quad g = \frac{G}{R+G+B}$$

$$N(m, \Sigma^2), \quad \text{where } m = (\bar{r}, \bar{g}) \quad (1)$$

$$\bar{r} = \frac{1}{S} \sum_{i=1}^S r_i, \quad \bar{g} = \frac{1}{S} \sum_{i=1}^S g_i, \quad \text{and } \Sigma = \begin{bmatrix} \sigma_r^2 & \rho_{X,Y} \sigma_r \sigma_g \\ \rho_{X,Y} \sigma_r \sigma_g & \sigma_g^2 \end{bmatrix}$$

따라서, 얼굴은 Skin color 모델을 이용하여 입력 영상을 skin color pixel과 그렇지 않은 픽셀로 분류하고, skin color pixel의 connected component를 찾아서 검출한다. 검출된 얼굴을 인가자와 비인가자로 분류하기 위하여 eigenspace상에서 각 영역의 가중치 벡터를 사용한다.

$N \times N$ 크기의 이미지는 N^2 차원의 이미지 공간에서 한 점으로 표현할 수 있고, 이때 각각의 차원은 한 픽셀에 대응된다. 이러한 이미지를 심각한 정보의 손실 없이 낮은 차원으로 표현하기 위하여 PCA 방법을 사용한다. M 개의 학습 이미지들로부터 M' 개의 고유얼굴($E_1, E_2, E_3, \dots, E_{M'}$)를 구하고, 특정 입력 이미지는 직교공간에서 M' 개의 고유얼굴들의 선형 조합(linear combination)으로 표현될 수 있다 [4, 5].

고유얼굴을 구성하는 방법은 다음과 같다. M 개의 학습 이미지는 고유공간을 구성하기 위해서 열 벡터(Column vector) 인 $I_1, I_2, I_3, \dots, I_M$ 로 나타내고, 이들 입력 이미지의 평균 이미지인 $A = \frac{1}{M} \sum_{n=1}^M I_n$ 를 구한다. 각 학습이미지와 평균 이미지와의 차이 벡터인 $\Phi_i = I_i - A$ 들로 이루어진 새로운 벡터집합 $Y = [\Phi_1, \Phi_2, \dots, \Phi_M]$ 를 구한다. 얼굴 이미지의 분포를 최적으로 표현하는 M 개의 직교(Orthogonal) 벡터 u_k 를 구하기 위하여, 공분산 행렬(Covariance matrix)을 다음과 같이 구한다.

$$C = \frac{1}{M} \sum_{i=1}^M \Phi_i \Phi_i^T = YY^T \quad (2)$$

C 행렬의 크기는 $N^2 \times N^2$ 으로 계산상의 효율성을 위하여

수식 (3)과 같이 C' 을 사용한다.

$$C' = \frac{1}{M} \sum_{i=1}^M \Phi_i^T \Phi_i = Y^T Y \quad (3)$$

C' 행렬로부터 구한 M 개의 고유벡터(Eigenvector) v_k 와 고유값(Eigenvalue) λ_k 를 이용하여 u_k 를 다음과 같이 구한다.

$$u_k = \frac{Y \times v_k}{\sqrt{\lambda_k}}, k = 1 \dots M \quad (4)$$

λ_k 의 값에 따라 u_k 중에서 중요한 M' 개를 선택하여 m 개의 학습 이미지에 대해서 고유벡터의 선형 조합, $W_m^{training} = [w_1, w_2, \dots, w_{M'}]$ 을 구한다

$$w_k = u_k^T \Phi_i, k=1 \dots M \quad (5)$$

얼굴 영역의 분류를 위하여 k 개의 검출된 영역을 직교 공간으로 투영하여 얻은 $W_k^{candidate}$ 값을 학습이미지들의 인가자 얼굴 클러스터의 평균 $W^{authorized}$ (수식 (6))과 인가자가 아닌 클러스터의 평균 $W^{unauthorized}$ (수식(7))을 수식 (8)과 같이 비교한다. 비교한 값이 인가자 학습 얼굴 이미지에 가까우면 검증된 얼굴 영역을 인가자로 분류하고 그렇지 않으면 비인가자로 분류한다.

$$W^{authorized} = \frac{1}{N_1} \sum_{i=1}^{N_1} W_i^{authorized} \quad (6)$$

$$W^{unauthorized} = \frac{1}{N_2} \sum_{i=1}^{N_2} W_i^{unauthorized} \quad (7)$$

$$\min (| W_k^{candidate} - W^{authorized} | , | W_k^{candidate} - W^{unauthorized} |) \quad (8)$$

3. 비인가자 추적

비인가자 추적 단계에서는 static camera에 의해 비인가자로 분류된 출입자를 추적한다. static camera가 출입자를 비인가자로 분류한 경우, pan-tilt camera는 입력영상에서 motion을 검출하여 가장 큰 motion 영역이 화면의 중앙에 오도록 camera를 제어하여 비인가자의 영상을 항상 화면에 보일수 있게 한다. motion은 아래 수식(9)으로 계산하고 threshold값 θ 는 15를 사용하였다.

$$M_t(x, y) = \begin{cases} 1, & |I_t(x, y) - I_{t-1}(x, y)| > \theta \\ 0, & otherwise \end{cases} \quad (9)$$

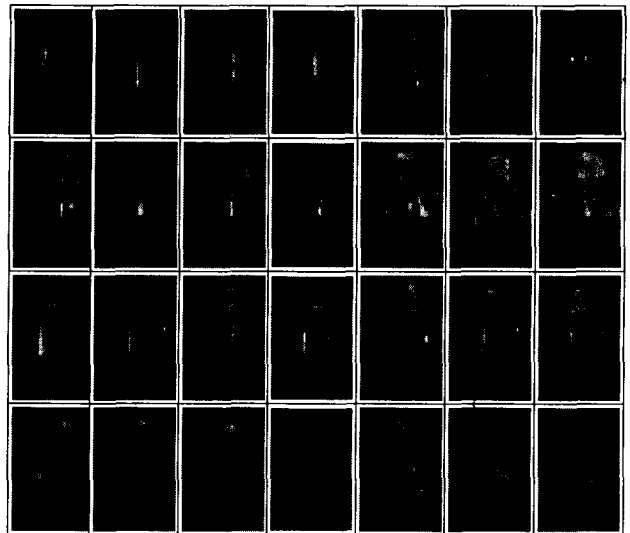
III. 실험결과

제안된 시스템을 연구실에 설치하여 실험하였다. 실험 결과 출·입 감지율은 100%, 감시 대상자가 들어오는 상황은 96 %, 감시 대상자 나가는 상황은 97%의 정확성을 보였다. 들어온 감시 대상자의 인가자/비인가자 분류의 정확성은 87%였다. 제안된 시스템은 감시자의 편의성을 위해 그림 3과 같이 인터넷 브라우저를 통해 인터페이스 할 수 있도록 구현하였다.



그림 3. 제안된 시스템의 사용자 인터페이스

출입자 분류에 사용된 학습 이미지는 10명을 5명씩 인가자와 비인가자로 나누어 각각 다른 각도에서 캡처한 40×60 그레이 얼굴 이미지 10장씩을 사용하였다. 그림 4는 사용된 얼굴 학습 이미지 중에서 일부를 보여준다.



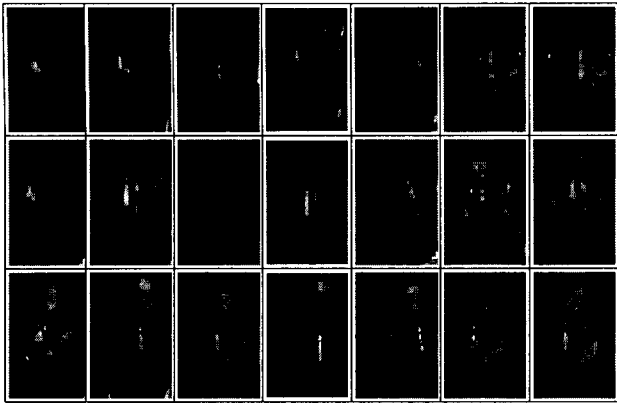


그림 4. 얼굴 학습 이미지

본 시스템에서 이용한 pan-tilt camera는 Smilecam의 Su-320이고, 사용된 PC는 Pentium 1.8GHz이며, Widnow 2000 professional에서 MS사의 Visual C++ 6.0, java 2.0 을 사용하여 구현하였다.

표 1. 시스템 사양

Hardware	Camera	Smilecam Su-320
	CPU	Pentium 1.8GHz
	Main Memory	256 Mbytes
	Network Adapter	100Base-T Ethernet Card
Software	OS	MS Windows 2000 Professional
	Web Server	Ms Internet Information Service
	Camera Access	Video for Windows Library
	Network Protocol	TCP/IP Protocol
	Developed Language	MS Visual C++ 6.0, Java 2.0

IV. 결론

본 논문에서는 제한된 영역을 multi-camera를 이용하여 감시하고, 인터넷 상에서 감시 상황과 결과를 볼 수 있는 시스템을 제안하였다. 제안된 시스템은 원격지 실시간 모니터링, 감시영상 디지털 녹화 및 검색, 비인가자에 대한 경고 신호 발생, 비인가자 추적의 기능을 가지며, 로그인 기능을 주어 사전에 인증된 사용자에게만 정보를 전송하고, 출·입 상황을 각각 분리하여 보여준다. 제안된 시스템은 출입자 감시의 효율성을 증가 시킬 뿐만 아니라, 영상처리 기법을 필요로 하는 응용분야에 기본 기술을 제공하며, 고인건비의 관리 인원을 대체할 수 있다는 장점을 가진다.

참고문헌

- [1] S.Stillman, R.Tanawongsuwan and I.Essa, "Tracking Multiple People with Multiple Cameras", PUI 1998 Workshop, 1998.
- [2] J.Yang and A.Waibel, "A Real Time Face Tracker," IEEE Workshop Appl. Comput. Vision, pp.142-147, 1996.
- [3] A.C.M.Fong and S.C.Hui, "Web-based intelligent surveillance system for detection of criminal activities", IEE Comp & Cont Engg J, Vol. 12/6, pp.263-270, 2001.
- [4] M.Turk, A.Pentland, "Face Recognition Using Eigenfaces", IEEE Proceedings CVPR '91, pp.586-591, 1991.
- [5] A.X.Guan, H.H.Szu, "A local face statistics recognition methodology beyond ICA and/or PCA", International Joint Conference on Neural Network, pp.1016-1021, 1999.