

폐쇄회로 카메라에서 운동에너지를 이용한 모션인식과 의상색상 및 얼굴인식을 통한 특정인 추적 알고리즘

이 인 정[†]

요 약

특정인을 추적하는 기술은 인간처럼 행동하는 로봇기술에서 가장 많이 등장하는 기술이다. 이 기술은 세 가지 영역에서 접근하고 있는데 첫째가 특정인의 의상 색상이고 두 번째가 특정인의 얼굴과 그 표정이며 세 번째가 특정인의 제스처나 머리의 움직임이다. 그러나 로봇은 센서를 통해 색상이나 제스처를 감지할 수 있기 때문에 폐쇄회로 카메라를 통해 획득한 영상만으로 특정인을 추적하는 것과는 다르다. 폐쇄회로 카메라에서 가장 큰 문제점은 시스템 속도인데 입력된 영상에서 다시 계산에 의해 특정인을 추적하기 위해서는 계산수를 줄여야한다. 시스템 속도를 높이기 위해 색상 추적은 통계치를 사용하는 것이 좋고 얼굴인식은 고유 얼굴을 사용하는 것이 바람직하다. 색상과 얼굴인식만으로는 추적에 어려움이 있기 때문에 모션 분석이 필요하다. 기존의 모션 분석이 주어진 영상의 전체 영역에서 형상을 바탕으로 이루어지기 때문에 속도가 느리고 인식률도 떨어진다. 본 논문에서는 얼굴 인식 시 찾아진 얼굴영역에 대한 모션분석을 계산속도가 빠른 운동에너지를 써서 인식률과 인식 속도를 높였다. 본 논문이 제안한 알고리즘과 Girondel, V. 등이 제시한 방법을 같은 동영상에서 실험한 결과 동일한 인식률을 얻었으며 인식 속도는 제안한 알고리즘이 더 빨랐으며 LDA를 사용할 경우 속도는 비슷하나 인식률은 더 나은 결과를 얻었으며 특정인을 찾는 것은 제안한 알고리즘이 더 효과적이었다.

키워드 : 특정인추적, 움직임에너지, 얼굴인식

A Tracking Algorithm to Certain People Using Recognition of Face and Cloth Color and Motion Analysis with Moving Energy in CCTV

In-Jung Lee[†]

ABSTRACT

It is well known that the tracking a certain person is a vary needed technic in the humanoid robot. In robot technic, we should consider three aspects that is cloth color matching, face recognition and motion analysis. Because a robot technic use some sensors, it is many different with the robot technic to track a certain person through the CCTV images. A system speed should be fast in CCTV images, hence we must have small calculation numbers. We need the statistical variable for color matching and we adapt the eigen-face for face recognition to speed up the system. In this situation, motion analysis have to added for the propose of the efficient detecting system. But, in many motion analysis systems, the speed and the recognition rate is low because the system operates on the all image area. In this paper, we use the moving energy only on the face area which is searched when the face recognition is processed, since the moving energy has low calculation numbers. When the proposed algorithm has been compared with Girondel, V. et al's method for experiment, we obtained same recognition rate as Girondel, V., the speed of the proposed algorithm was the more faster. When the LDA has been used, the speed was same and the recognition rate was better than Girondel, V.'s method, consequently the proposed algorithm is more efficient for tracking a certain person.

Key Words : Tracking a Certain Person, Moving Energy, Face Recognition

1. 서 론

특정인을 추적하는 기술은 인간처럼 행동하는 로봇기술

[1]에서 가장 많이 등장하는 기술이다. 이 기술은 세 가지 영역에서 접근하고 있는데 첫째가 특정인의 의상 색상이고 두 번째가 특정인의 얼굴과 그 표정이며 세 번째가 특정인의 제스처나 머리의 움직임이다. 그러나 로봇은 센서를 통해 색상이나 제스처를 감지할 수 있기 때문에 폐쇄회로 카메라를 통해 획득한 영상만으로 특정인을 추적하는 것과는

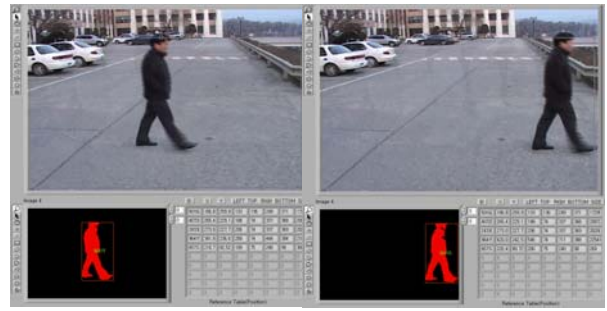
[†] 정 회 원 : 호서대학교 컴퓨터공학부 교수
논문접수 : 2008년 1월 16일
수 정 일 : 1차 2008년 2월 22일, 2차 2008년 3월 4일, 3차 2008년 3월 10일
심사완료 : 2008년 3월 25일

매우 다른 상황이다. 가장 큰 문제점은 시스템 속도인데 입력된 영상에서 다시 계산에 의해 필요한 목적을 이루기 위해서는 계산수를 줄여야한다. 본 논문에서는 시스템 속도를 높이기 위해 기존의 방법이라 할지라도 속도가 빠른 것을 선택하거나 계산식을 변경하여 추적에 활용하고자 한다.

먼저 특정인의 의상을 기준으로 색상을 추적하는 방법을 살펴보자. 색감을 인식하는 감성인식 방법은 이미 오래전부터 연구되어 잘 알려진 사실이지만[7, 16, 17, 19, 20], 특정 의상을 입은 사람을 다른 사람과 빠른 시간 내에 구별하기 위해서는 빠른 비교 방법이 필요하다. 이를 위해 계산이 빠른 통계치를 구하고 임계값을 찾아 비교하는 방법을 사용하기로 한다.

두 번째는 특정인의 얼굴과 그 표정을 인식하여 추적하는 것인데 표정이 바뀌거나 조명이 바뀌어도 잘 변하지 않으면서 속도가 빠른 알고리즘이 요구된다. 본 논문에서는 고유 얼굴을 찾아 특정인을 추적하는 방법을 사용한다. 고유 얼굴은 계산수가 적고 속도가 빠르며 조명의 변화에 강하고 얼굴이 정면상이 아니어도 추적범위가 넓기 때문이다. 얼굴 인식(face recognition)에 대한 기술은 두 가지로 나누어 생각해야 하는데 먼저 얼굴이라고 생각되는 얼굴 영역인식[13]이고, 그 다음은 얼굴의 패턴을 인식하는 것이다. 얼굴의 패턴을 인식하는 방법으로는 부분인식[21]과 고유 얼굴 인식[18, 22]이 있다. 또한 얼굴의 색상을 구별하여 추적하는 시스템이 소개된바있으며[5], 3차원 형상을 획득하여 이용하는 방법도 있다[4]. 특히 3차원 형상은 특정인이 얼굴을 돌려서 다른 데를 본다든지 또는 뒷모습만 보일 때도 추적하기위한 방법이다. 이 중에서 부분인식은 각 영역을 인식하기 위해 계산수가 늘어나며 얼굴색상인식은 환경변수에 너무 크게 영향을 받고 3차원 형상을 획득하여 이용하는 경우는 많은 시간을 요한다.

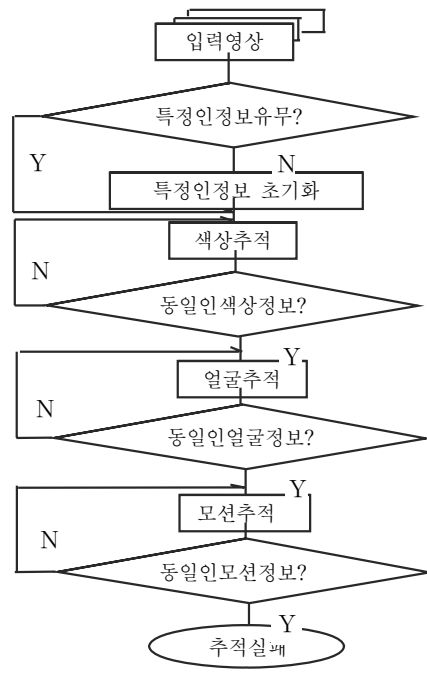
세 번째는 특정인의 제스처나 머리의 움직임을 인식하여 추적하는 방법이다. 센서를 사용하지 않고 폐쇄회로 카메라를 통해 전송된 영상에서 움직임을 감지하는 시스템으로는 차량추적[11,13,15]시스템을 예로 들 수 있으나 이 시스템은 고정식으로 검지영역을 설정하여 그 영역에서만 추적하므로 전 영역으로 확장 시 속도가 매우 느려진다. 또한 선택된 개별 물체의 이동을 추적하는 알고리즘[2]은 움직임 물체수가 많아지면 구별이 어려운 단점을 보인다. Javier-Ruiz-del-Solar 외 2인이 발표한 논문[14]에서는 모션분석을 위해 State-of-the-art methodologies를 사용했으나 얼굴인식에 집중되어 있어서 얼굴영역이 확실치 않으면 추적이 어렵다. 최근에 발표된 논문[3]중 형상 정보와 모션 정보 융합을 통한 물체인식이 있다. 보통의 경우 배경영상이 고정적일 때 움직임을 통해서 형상을 찾아내는 것은 어려운 일이 아니다 (그림 1). 그러나 움직이는 개체가 혼재해 있을 때는 영상 겹침이 많아 이 논문에서 제시하는 방법으로는 특정인을 추적하기에는 어려움이 많다. 더구나 이 논문에서 제시한 방법은 학습을 시켜야 할뿐 아니라 입력된 영상의 전체 영역에서 움직이는 물체를 모두 찾아야 하므로 속도가 느리다.



(그림 1) 움직임을 통해 형상을 찾음

뿐만 아니라 찾은 형상 속에서 특정인을 다시 구별해 내야 하는 문제가 생기며 많은 사람이 통행하는 광장에서 폐쇄회로 카메라로 전송된 영상에서는 적용이 어렵다.

본 논문에서는 얼굴인식에서 이미 찾아진 얼굴 영역을 이용하여 필요한 모션만 분석하며 계산수가 적고 인식에 변별력이 있는 운동에너지를 구하여 특정인을 추적하고자 한다. 대개의 경우 폐쇄회로 카메라는 통행인의 위쪽에 설치되므로 움직이는 개체가 겹쳐도 얼굴영역이 겹치는 경우는 더 적어진다. 또한 많은 사람들이 고개를 움직이는 모션이 특징적이어서 특정인을 구별해 내기가 쉬워진다. 모션정보와 얼굴인식을 통해 다수의 사람을 추적하는 방법을 제안한 논문으로 Girondel et al[12]이 제안한 방법을 들 수 있다. 색상인식을 제외하고는 본 논문이 제안한 방법과 비슷한 점이 많으나 특정인을 주목하여 추적한다는 점에서는 매우 다르다. 더구나 이 논문에서는 영상 겹침을 해결하기위해 Kalman 필터링을 사용하는 것을 주제로 다루고 있다. 본 논문에서는 얼굴부분에서 일어나는 영상 겹침이 대부분 오



(그림 2) 특정인을 추적하는 순서

래 지속하지 않는 점을 고려하여 LDA를 통해 특징인을 구분하여 다시 추적해 나가는 방법을 제안한다.

본 논문이 제안하는 알고리즘의 순서는 다음과 같다. 먼저 의상을 기준으로 색상을 인식하며 다수의 동일색상이 나타나면 얼굴인식을 시도하고 동일 얼굴이 두 개 이상 나타나면 모션인식을 하여 특징인을 추적한다(그림 2). 이때 특징인에 대한 정보는 미리 알고 추적하는 것을 원칙으로 한다.

2절에서는 색상인식을 위해 계산이 적은 통계량과 임계값을 계산하는 식을 새우고 3절에서는 얼굴인식과 LDA를 소개한다. 4절에서는 모션분석을 위해 얼굴영역의 운동에너지를 소개한다. 차영상을 구하는 과정은 논문[3]의 방법으로 접근하나 필터 값은 확률변수를 사용하여 조정한다. 다음 본 논문이 제안하는 방법이 특징인을 인식하는데 변별력이 있음을 보인다. 5절에서는 본 논문이 중요시 여기는 속도에 대한 실험결과를 제시하며 특히 비교분석을 위해 같은 동영상에서 Girondel, V.가 제시한 방법과 비교한다.

2. 의상 색상 인식 및 추적

특정인이 의상색상을 빠르게 추적하기위해 통계적 수치와 거리개념을 사용한다. 그러나 먼저 생각해야 할 것은 추적하기 위한 특정인의 의상색상에 대한 자료가 없을 때이다. 이때는 시스템에서 인위적으로 지시해주어야 할 필요가 있다. 폐쇄회로를 통해 전송된 영상에서 원하는 사람의 중심 부위에서 마우스 포인터로 원하는 색상을 얻기로 하자. 이때 포인터로부터 반경을 키워가며 평균을 계산하고 동일색상을 벗어날 때 평균값의 변화가 크게 나타나는 것을 활용한다.

반경을 r_0, r_1, \dots, r_n 이라 하고 이 각각의 반경이 만드는 영역을 R_0, R_1, \dots, R_n 이라 하자.

$$E_i^Y = \frac{1}{|R_i - R_{i-1}|} \int_{R_{i-1}}^{R_i} I^Y(x) dx \quad (1)$$

$$E_i^U = \frac{1}{|R_i - R_{i-1}|} \int_{R_{i-1}}^{R_i} I^U(x) dx \quad (2)$$

$$E_i^V = \frac{1}{|R_i - R_{i-1}|} \int_{R_{i-1}}^{R_i} I^V(x) dx \quad (3)$$

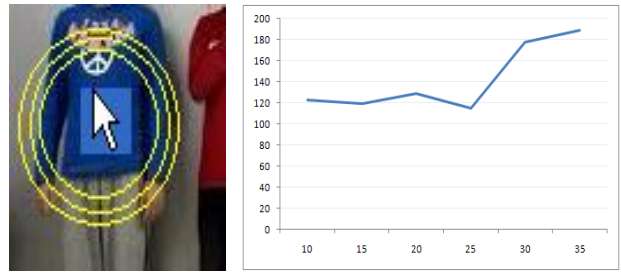
여기서 $I^Y(x)$, $I^U(x)$, $I^V(x)$ 는 영역내의 각각의 YUV 색상 정보이다. $0 \leq k \leq n$ 에 대해서 각각의 평균을

$$E(Y) = \frac{1}{k} \sum_{i=1}^k E_i^Y \quad (4)$$

$$E(U) = \frac{1}{k} \sum_{i=1}^k E_i^U \quad (5)$$

$$E(V) = \frac{1}{k} \sum_{i=1}^k E_i^V \quad (6)$$

라 하고 분산을 각각 $\sigma^2(Y)$, $\sigma^2(U)$, $\sigma^2(V)$ 라 하자. 만일



(그림 3) 마우스 포인터로부터 반경을 키워가며 구한 색상의 평균값과 그 변화

$E_{k+1}^Y, E_{k+1}^U, E_{k+1}^V$ 중 하나가 다음 부등식인 식(7)의 범위를 벗어나면 $E(Y), E(U), E(V)$ 값을 원하는 색상으로 선택한다. $H = \{Y, U, V\}$ 라 하고 $\mathcal{J} \in H$ 에 대해서

$$E(\mathcal{J}) - t_{0.025} \sigma(\mathcal{J}) \sqrt{\frac{2}{k}} \leq E_{k+1}^{\mathcal{J}} \leq E(\mathcal{J}) + t_{0.975} \sigma(\mathcal{J}) \sqrt{\frac{2}{k}} \quad (7)$$

여기서 $t_{0.025}$ 와 $t_{0.975}$ 는 95% 유의수준으로 자유도가 k 인 t 분포 확률변수이다.

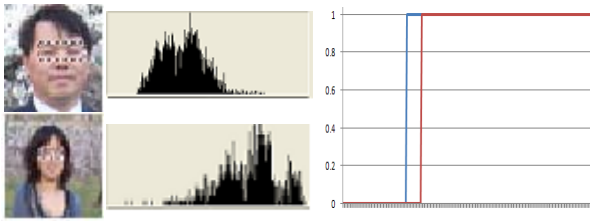
다음 (그림 3)은 최초 10픽셀 반경에서 시작하여 5픽셀씩 키워가며 YUV중 V값을 계산한 결과이다. 마우스 포인터로부터 의상을 벗어나기 시작하는 부분이 30픽셀정도인데 30픽셀 보다 커지면 평균값에 큰 차이를 보인다.

추적을 위해 10*10 마스크를 만들고 그것을 M^A 라 하자. 회선방법에서처럼 한 픽셀씩 이동하며 계산한다. 카메라로부터 입력된 영상에서 마스크에 대응되는 영상의 색상평균을 계산하여 각각 $E(M^Y), E(M^U), E(M^V)$ 하고 분산을 각각 $\sigma^2(M^Y), \sigma^2(M^U), \sigma^2(M^V)$ 라 한다. 특정인의 색상이 70%이상 마스크 내에 들어오면 추적에 성공한 것으로 판단한다. 즉, 다음 식(8)을 만족하는 것으로 선택한다.

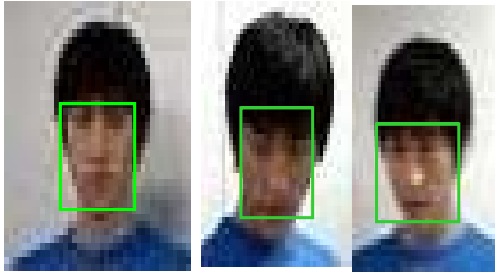
$$\sum_{\mathcal{J} \in H} (E(\mathcal{J}) - E(M^{\mathcal{J}}))^2 \leq \sum_{\mathcal{J} \in H} \sigma(M^{\mathcal{J}}) \quad (8)$$

3. 얼굴 인식

얼굴인식은 특징인을 추적하기위해 계산수가 적고 속도가 빠르며 조명의 변화에 강하고 얼굴이 정면상이 아니어도 추적범위가 넓은 고유 얼굴을 적용한다. 얼굴영역 검출이 먼저 이루어져야 하는데 얼굴의 지역적 특성(Local Feature)에 근거한 방법[8]과 얼굴의 전체 형상(Template based)에 기반을 둔 방법[10]중 여기서는 속도가 빠른 지역적 특성을 사용한다. 지역적 특성 중 눈의 동공은 밝기가 매우 낮기 때문에 이진화 과정을 통해 비교적 쉽게 눈 영역을 찾을 수 있다. 다음 (그림 4)에서 보듯이 이진화하지 않고 두 개의 영상에서 눈의 영역의 밝기 분포를 보면 동질성을 찾기 어렵다. 그러나 이진화한 영상에서 0과 1의 비율은 비슷한 향을 보인다.



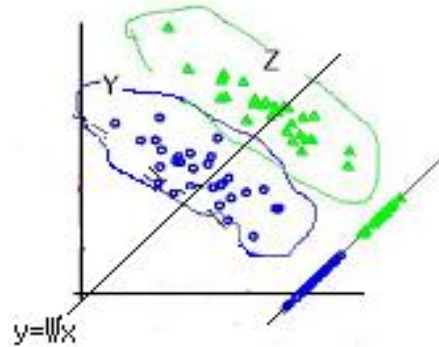
(그림 4) 눈 부위의 밝기 분포와 이진화한 후 0과 1의 수 비교



(그림 5) 고유 얼굴(Eigen-face) 매치

이렇게 찾아진 얼굴영역에서 특징인을 추적하기 위해 얼굴의 특징을 보관한다. 얼굴의 특징을 보관하고 검색하고 비교하는데 계산수가 적은 방법이 고유 얼굴 방법이다. 개개인들의 얼굴 분포는 각 얼굴마다 확률분포함수를 갖는다. 이러한 확률분포로부터 고유치를 찾기 위해 PCA(Principal Component Analysis)를 사용한다. 이 고유치들의 공간이 얼굴만을 위한 고유 공간(Eigen-space)이고 그 공간에 대응하는 벡터들을 고유벡터라 하는데 이 고유벡터를 고유 얼굴(Eigen-face)이라 부른다. 속도를 높이기 위해 변수를 점도와 왜도로 줄이고 고유 얼굴을 찾은 후 20개의 샘플 중에서 비슷한 고유 값을 찾은 결과를 다음 (그림 5)에서 보였다.

이미 학습되어진 자료에서 추출한 특징인에 대한 고유 얼굴 정보라면 문제가 없지만 추적중인 동영상에서 특징인 정보를 만들어 추적해야 한다면 2절에서처럼 마우스를 사용하여 영상속의 특징인에 대한 고유 얼굴을 찾는다. 그런 다음 다른 사람의 고유 얼굴과 비교하여 일치 여부를 결정한다. 이때 고유벡터들 사이의 거리가 작으면 두 얼굴이 구별되지 않는 경우가 생긴다. 이 경우 계속되는 이미지 시퀀스에서 고유 벡터를 찾아서 저장한 후 저장된 자료들을 잘 분리해 내면 특징인과 다른 사람을 구분하게 되는데 잘 알려진 LDA(Linear Discriminant Analysis) 방법을 사용한다. LDA 방법은 Identity의 변화와 그 밖의 다른 요인에 의한 변화를 구분할 수 있도록 하는 방법이다. 그래서 특징인이 등록 시와 다른 조명이나 다른 각도에서 인식을 시도하였을 경우에도 Identity는 동일하다고 판단을 내릴 수 있는 것이다. 여기서는 변환 후 서로 다른 Identity를 가지는 집단에 속하는 영상간의 분산은 최대로 하고 Identity가 같은 집단 내부의 영상간의 분산은 최소가 되도록 하는 변환을 구하여 해결하기로 한다. 즉, 특징 벡터들로부터 분류를 위한 가장 효과적인 주축을 찾는 것으로 d 차원의 n 개의 데이터 샘플 x_1, x_2, \dots, x_n 이 두 개의 클래스로 분리된다고 했을 경우, x_i



(그림 6) LDA를 사용하여 분류한 두 자료

의 컴포넌트에 대한 선형결합으로서 다음과 같은 식9를 얻는다. W 를 $n \times n$ 행렬이라고 할 때

$$\vec{y} = W\vec{x} \tag{9}$$

이것에 대응하는 n 개의 샘플 y_1, y_2, \dots, y_n 이 두개의 집합으로 분류된다. 기하학적으로는 \vec{y} 의 방향을 가진 축에 x_i 데이터가 투사된 것이 y_i 샘플이다. 여기서 y_i 데이터를 두 개의 클래스로 잘 분리할 수 있는 W 를 찾는 것이 중요하다. W 를 찾기 위해 x_i 를 두 집단 D_1, D_2 로 나누고 D_1, D_2 의 농도를 각각 n_1, n_2 라 하자.

$$S_{same} = \sum_{k=1}^2 S_k \quad S_k = \sum_{x \in D_k} (x - m_k)(x - m_k)^t \tag{10}$$

$$\text{여기서 } m_k = \frac{1}{n_k} \sum_{x \in D_k} x \quad k=1,2.$$

$$S_{diff} = \sum_{k=1}^2 n_k (m_1 - m_2)(m_1 - m_2)^t \tag{11}$$

$$\mathcal{J}(W) = \frac{W^{-1} S_{diff} W}{W^{-1} S_{same} W} \tag{12}$$

라 할 때 \mathcal{J} 를 최대로 하는 W 를 찾으면 된다. 2차원 상에 분포하는 두 종류의 자료로부터 LDA를 사용하여 분류한 것을 다음 (그림 6)에 나타내었다.

4. 모션 인식

일반적으로 폐쇄회로 카메라는 움직이는 물체보다 위쪽에 설치하는 경우가 많다. 속도를 높이기 위해 이미 찾은 얼굴영역에서 머리의 움직임을 추적하는 것이 바람직하다. 머리의 움직임으로부터 움직임 벡터를 찾고 속도와 그 운동에너지를 계산하여 비교함으로써 특징인을 추적한다. 움직임 벡터를 얻어 운동에너지를 계산하기 위해서 먼저 차영상을 찾는다. 차영상은 배경영상을 설정하고 이 배경영상과 현재 영상의 차를 이용한다. 그런데 배경영상은 항상 일정하지 않고 상황에 따라 변하게 된다. 따라서 이러한 단점을 보완

하면서 보다 정확한 움직임 영역을 검출하기위해 Gaussian Mixture Learning[9] 방법이나 적응형 배경 차 방법[3]을 사용하는데 여기서는 적응형 배경 차 방법에 확률적 필터를 사용하기로 한다.

B_n 을 시간 $t=n$ 에서 배경영상, B_{n+1} 은 갱신된 배경영상, I_n 은 $t=n$ 에서 입력영상을 나타낸다고 하자. 그리고 배경영상과 입력영상에 적절한 가중치를 부여하여 움직이는 물체를 배경에서 제외하는 역할을 하는 필터를 α 라 한다.

먼저 아무런 움직이는 물체가 없다고 보고 필터 값 $\alpha=0$ 으로 하여 식13을 만든다.

$$B_{n+1} = \alpha I_n + (1-\alpha)B_n \quad (13)$$

두 번째 단계는 갱신된 배경에 움직이는 물체가 검출되었을 때, 배경영상과 입력영상과의 차 영상을 구한 후 움직이는 물체만을 구하기 위하여 이진화 시키는 과정으로 아래의 식 14, 15를 구한다.

$$Bm(D_n) = \begin{cases} 1 & \text{if } D_n > \tau \\ 0 & \text{otherwise} \end{cases} \quad (14)$$

$$\tau = m(D_n) + \sigma(D_n) \quad (15)$$

위의 식에서 D_n 은 입력영상과 배경영상과의 차를 나타내고 τ 는 이진 영상을 검출하기위한 임계값을 나타낸다. 또한 $Bm(D_n)$ 은 차영상에 의한 이진영상, $m(D_n)$ 은 차영상 D_n 에 $N \times N$ 사이즈의 윈도우를 사용하여 얻은 평균을 나타내고, $\sigma(D_n)$ 은 차영상의 표준편차를 나타낸다. 위의 식14, 15에 의하여 얻어진 결과를 식13에 적용하여 움직임 픽셀이면 다음 배경영상의 픽셀 값을 갱신하고 움직임이 없는 픽셀은 다음 배경영상의 픽셀 값을 이전 배경영상의 픽셀 값으로 그대로 유지한다. 이때 사용하는 필터 α 는 움직이는 물체가 나타날 확률 값에 따라 적용하는데 $\alpha = prob(Bm(D_n) = 1 | B_n \neq I_n)$ 이 적절하며 이는 확률이 클수록 α 값이 커져서 입력영상이 배경영상에 크게 영향을 미친다. 이를 정리하면 다음 식16과 같다.

$$B_{n+1} = \begin{cases} \alpha I_n + (1-\alpha)B_n & \text{if } Bm(D_n) = 1 \\ B_n & \text{otherwise} \end{cases} \quad (16)$$

이렇게 구한 배경영상을 적용하여 아래 (그림 7)에서처럼 차영상을 구한다.

정해진 영역 내에서 구해진 차영상의 이진 영상을 구하고 잡음을 제거한다. 이렇게 구해진 영상은 움직인 흔적을 나



(그림 7) 차영상 과정 : 좌측 두 영상 차와 확대한 영상

타내게 되는데 이 움직임들의 크기를 속도로 보고 에너지를 계산한다. 속도는 밝기 값의 변화율로부터 구할 수 있는데 픽셀 $p(x,y)$ 에서의 밝기를 $I(x,y)$ 라 하면 속도

$$\begin{aligned} \vec{v} &= \frac{dI}{dt} \\ &= \frac{\partial I(x,y)}{\partial x} \frac{dx}{dt} + \frac{\partial I(x,y)}{\partial y} \frac{dy}{dt} \\ &= v_x + v_y \end{aligned} \quad (17)$$

이다. 시간을 일정한 크기로 한정하면

$$\begin{aligned} \vec{v} &= \frac{dI}{dt} \\ &\approx \frac{\Delta I_x + \Delta I_y}{\Delta t} \\ &\approx \Delta x + \Delta y \end{aligned} \quad (18)$$

이고 여기서 Δ 는 변위의 차를 나타낸다. 즉 카메라로부터 보내온 이미지의 프레임에서 정해진 개수의 프레임사이의 변위를 계산한다. 다시 말해서 차영상에서 얻어진 변위의 차를 속도 벡터로 사용하여 에너지를 계산한다(그림 8). 이때 벡터의 방향은

$$\theta = \arctan \frac{|\frac{\partial I(x,y)}{\partial x}|}{|\frac{\partial I(x,y)}{\partial y}|} \quad (19)$$

이다. 에너지를 계산하는 방법은 다음과 같다.

식17로부터 속도 벡터

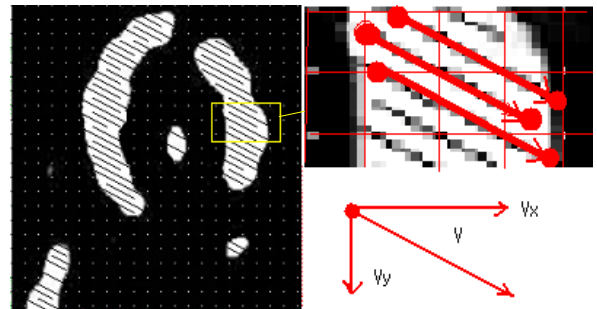
$$\vec{v} = \vec{v}_x + \vec{v}_y \quad (20)$$

이고 한 포인트에서의 에너지 값은

$$E = \frac{1}{2} m(\vec{v})^2 = \frac{1}{2} m(\vec{v}_x + \vec{v}_y)^2 \quad (21)$$

이며 임의의 주어진 영역 Ω 에서의 에너지 총합은

$$E_{sum} = \iint_{\Omega} \frac{1}{2} m(\vec{v}_x + \vec{v}_y)^2 dx dy \quad (22)$$



(그림 8) 움직임벡터를 구하는 과정 : 움직임을 확대하여 벡터로 표시함

이다.

그런데 식 (18)로부터 에지에서 변위의 크기가 속도 벡터를 대신 적용하여도 문제가 없음을 보인다. 그러므로 에너지 총합은

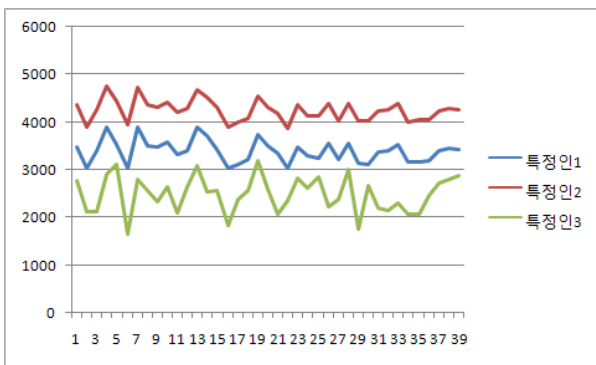
$$Esum = \frac{1}{2} \sum_{s \in S} (\Delta x^2 + \Delta y^2)^{\frac{1}{2}} \quad (23)$$

이 되며 여기서 S 는 에지이다.

다음 (그림 9)는 임의의 동영상에서 6프레임 차영상으로 에너지를 계산하고 2초간의 에너지 평균을 계산하여 특정인 세 사람이 카메라에서 사라질 때까지 추적한 결과이다. 여기서 세 사람의 에너지 평균이 서로 차이를 보이고 있음을 알 수 있다. 같은 실험을 두 사람에게 적용하여 40초 동안 추적한 후 20개의 자료를 가지고 t-검정을 한 결과 유의수준 95%로 유의한 결과를 얻었다<표 1>.

카메라 방향으로부터 일 직선상에 두 사람 이상이 움직이면 영상 겹침 현상이 생기는데 이 상황이 오래 동안 지속된다면 추적은 어려움을 겪게 된다. 여기서 1초나 2초 후에 해소되는 것으로 가정한다. 비슷한 에너지 값이나 빈번한 영상 겹침이 발생하면 LDA방법으로 분류한 다음 특징인을 찾는다.

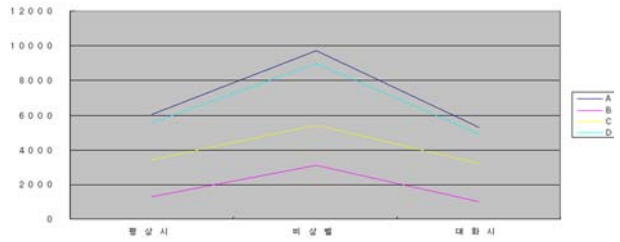
다음 (그림 10)은 작은 광장에서 카메라에 4명의 사람이 보이기 시작하여 카메라에서 사라질 때까지 16초 동안의 영상에서 4명의 운동에너지 값을 나타낸 그래프이다. 상황별 변화를 보기위해 평상시와 같이 걸어서 통과할 때와 비상벨이 울렸을 때와 서로대화하고 있을 때로 나누어 에너지 값을 조사하여 평균값을 비교한 것이다. 각 사람의 에너지 값은 상황이 바뀌어도 차이를 보이고 있다.



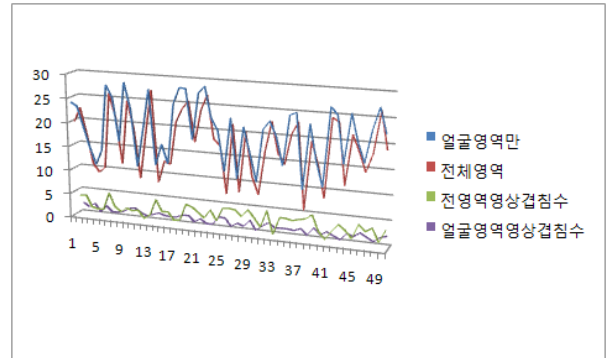
(그림 9) 특정인 세 사람의 에너지 추적 값

<표 1> 두 사람의 에너지 값에 대한 t-test

| 특정인 A와B | Paired Differences | | | t | df | Sig. (2-tailed) |
|---------|--------------------|----------------|-----------------|-------|----|-----------------|
| | Mean | Std. Deviation | Std. Error Mean | | | |
| | 1038.90 | 457.40 | 102.28 | 10.16 | 19 | .000 |



(그림 10) 네 사람의 에너지를 상황별로 비교한 그래프



(그림 11) 전체영역에서와 얼굴영역에서의 속도비교 및 영상 겹침 수 비교

위에서 보았듯이, 운동에너지가 각 사람마다 서로 다르므로 특징인을 인식하는데 변별력이 있다. 본 논문이 제안하는 얼굴영역에서만 운동에너지를 찾는 것은 속도 면에서도 빠름을 보인다. 다음 (그림 11)은 논문[3]에서처럼 전체영역에서 움직임 개체를 찾은 것과 본 논문이 제안한 얼굴영역에서 찾은 것의 속도를 비교하였다.

5. 특징인 추적 과정과 실험결과

먼저 추적 과정에서 발생하는 동일한 판정에 대해 생각해 보자. 의상색상의 경우는 YUV값이 각각 0에서 255까지의 값을 가지며 얼굴인식에서 고유 얼굴에 대한 고유벡터는 왜도가 -255에서 +255값을 갖고 첨도는 0에서 255*255 사이 값을 갖는다. 운동에너지는 그 크기가 최대 네 자리 정수이다. 통계적으로 모분산이 동일하다는 가정 하에 95% 유의수준으로 표본이 n 개일 경우 동일한 판정의 범위 R 은 다음 식24와 같다.

$$\mu_1 - \mu_2 + t_{0.025} \sigma \sqrt{\frac{2}{n}} \leq R \leq \mu_1 - \mu_2 + t_{0.975} \sigma \sqrt{\frac{2}{n}} \quad (24)$$

다음은 특징인을 추적하는 과정이다. 먼저 동일한 판정 기준으로 추적하고 자료를 저장하면서 LDA 방법을 사용한다.

$E(Y), E(U), E(V)$: 특징인 색상정보

$\vec{\lambda}(s, \kappa)$: 특징인 고유 얼굴

$Esum$: 특징인 모션에너지

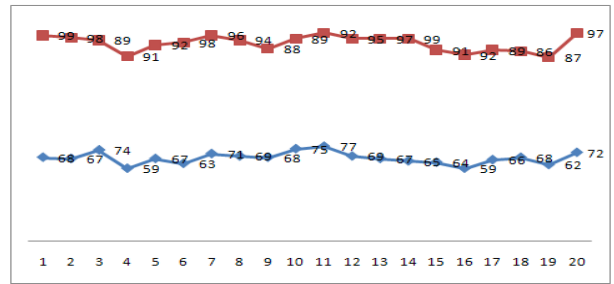
```

For( ; E(Y), E(U), E(V),  $\vec{\lambda}(s, \kappa)$ , Esum are not null; )
    scan E(Y), E(U), E(V),  $\vec{\lambda}(s, \kappa)$ , Esum;
While( Moving sequence )
{
For( ; Same Color Cloth  $\geq 2$ ; )
{
    tracking Same Color Cloth;
    save Same Color data;
}
For( ; Same Face  $\geq 2$ ; )
{
    tracking Same Face on Same Color data;
    save Same Face;
    LDA();
}
For( ; Same Motion Energy  $\geq 2$ ; )
{
    tracking Same Motion Energy on Same Face data;
    save Same Motion Energy;
    LDA();
}
Exit;
}
Return(Same Color or Same Face or Same Motion Energy);
    
```

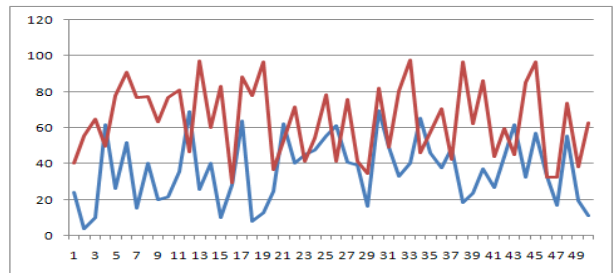
이 알고리즘에서 색상추적의 속도를 비교하기 위해 색상센서와 추적속도를 비교하여 (그림 12)에 보였다. 대여섯 명의 사람이 움직이는 현장에서 센서와 폐쇄회로 카메라가 동시에 추적한 결과 센서의 70%정도 성능을 보이는 것으로 나타났다.

얼굴인식에서 부분인식은 정확도는 높으나 인식 속도는 고유 얼굴이 더 나은 것으로 보인다. 다음 (그림 13)은 눈과 코 그리고 입을 인식하는 부분인식과 고유 얼굴의 인식속도를 비교한 것이다. 실험영상은 정면영상만으로 한정하였으며 영상의 크기도 동일한 것으로 하였다.

운동에너지의 계산 수는 다른 어떤 방법보다 작은 것이 확실하다. 인식률에 있어서도 제안한 알고리즘이 다른 방법



(그림 12) 색상센서와 추적속도비교 : 위가 색상센서 속도이고 아래는 폐쇄회로에서의 추적속도이며 시스템 실행 시간을 백을 기준으로 조정함

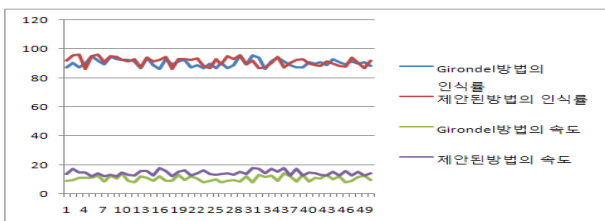


(그림 13) 부분인식과 고유 얼굴인식의 인식속도비교: 위가 고유 얼굴인식속도이며 아래가 부분인식 속도로 백을 기준으로 조정함

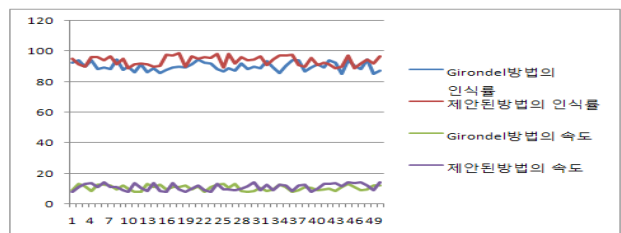
과 비교하여 뒤지지 않으며 LDA를 사용할 경우 인식률은 더 좋아지는 것으로 나타났다. 다음 (그림 14)에서 (a)는 제안된 방법과 Girondel, V.의 방법을 동일한 상황에서 실험하여 인식률과 속도를 비교하였으며 (b)는 LDA를 사용할 경우 제안된 방법이 인식률에서 앞서며 속도는 거의 같음을 보였다. 여러 사람을 추적해도 특정인을 찾기 위해서는 다시 인식을 시도해야 하는 점을 감안할 때 제안된 방법이 특정인을 추적하는데 효과적인 방법임을 알 수 있다.

6. 결론

폐쇄회로 카메라로부터 전송된 영상에서 특정인을 추적하는 데 있어 가장 중요한 것은 인식속도이다. 로봇기술에서는 센서를 사용할 수도 있지만 이미 전송된 영상에서만 인식을 수행해야 하기 때문에 계산수를 줄여야한다. 본 논문



(a) Girondel.V.의 방법과 제안된 방법의 인식률과 속도비교



(b) Girondel. V.의 방법과 제안된 방법에서 LDA를 고려할 때의 인식률과 속도비교

(그림 14) 제안된 방법과 Girondel. V.의 방법의 인식률과 속도비교

에서는 추적시스템에 필요한 세 가지 요소인 의상색상인식과 얼굴인식과 모션인식을 모두 수행하면서 수행속도가 빠른 알고리즘을 제안하였다. 빠른 처리를 위해 색상과 얼굴인식에서는 통계적 수치를 활용하였다. 특별히 본 논문에서는 계산수가 적은 운동에너지를 모션인식에 사용하였다. 또한 영상의 전체영역에서 인식을 시도하지 않고 얼굴 인식 시 찾아진 얼굴영역에 대한 모션분석을 수행하여 인식률을 높였다. 본 논문이 제안한 알고리즘과 Girondel, V. 등이 제시한 방법을 같은 동영상에서 실험한 결과 동일한 인식률을 얻었으며 인식속도는 제안한 알고리즘이 더 빨랐으며 LDA를 사용할 경우 속도는 비슷하나 인식률은 더 나은 결과를 얻었다. Girondel, V. 등이 제시한 방법에서처럼 여러 사람을 추적해도 특징인을 찾기 위해서는 다시 인식을 시도해야 하는 점을 감안할 때 특징인을 찾는 면에서는 제안한 알고리즘이 더 효과적이다.

참 고 문 헌

[1] 광근창, 지수영 편. "인간과 로봇이 자연스럽게 교감하는 상호 작용기술", ETRI CEO Information 제37호, March 2006.

[2] 김성일, "선택적 개별물체의 이동추적 알고리즘 개발에 관한 연구," 정보처리학회논문지, Vol.8-B, No.1, pp.50-58, 2001.2.

[3] 김정호, 고한석. "형상정보와 모션 정보 융합을 통한 움직이는 물체인식", 전자공학회논문지, Vol.43, No.5, pp.416-425, 2006.

[4] 박도영, 심연숙, 변혜란. "얼굴의 움직임 추적에 따른 3차원 얼굴합성 및 애니메이션." 정보과학회 논문지: 소프트웨어 및 응용, 제27권 제6호, 2000.

[5] 장기정, 권 인소. "적응 칼라 모델을 이용한 실시간 얼굴추적." 제12회 영상처리 및 이해에 관한 워크샵 발표논문집. 2000.

[6] 조용진, 양현승, "Snakes모형을 이용한 얼굴 윤곽선 추적 시스템." 정보과학회논문지(B), Vol.24, No.1, pp.62-72, 1997.

[7] A. Hanjalic, L. Xu. "User-oriented Affective Video Content Analysis." Proc. IEEE Workshop on Content Based Access of Image and Video Library., Kauai, HI, pp.50-57, Dec., 2001

[8] David G. Lowe. "Distinctive Image Features from Scale-Invariant Key points." International journal of computer vision, January 5, 2004.

[9] Dar-Shyang Lee, "Effective Gaussian Mixture Learning for Video Background Subtraction." IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.27, No.5, pp.827-832, May, 2005.

[10] Dufour, R.M. Miller, E.L. Galatsanos, N.P. "Template matching based object recognition with unknown geometric parameters." Image Processing, IEEE Transactions on Image Processing Vol.11 No.12 pp.1385-1396 Dec., 2002.

[11] E. D. Dickmanns, B. Mysliwetz, T. Christians, "An Integrated Spatial-Temporal Approach to Automatic Visual

Guidance of Autonomous Vehicles," IEEE Trans. on Syst. Man, Cybern., Vol.20, No.6, pp.1274-1284, 1990.

[12] Girondel, V. Caplier, A. Bonnaud, L. "Real time tracking of multiple persons by Kalman filtering and face pursuit for multimedia applications." Image Analysis and Interpretation, 6th IEEE Southwest Symposium on Image Processing, 2004.

[13] G. Luca, V. Murino, C. Regazzoni, "Vehicle Recognition and Tracking from Road Image Sequences," IEEE Trans. on Vehicular Technology, Vol.48, No.1 pp.301-318, Jan., 1999.

[14] Javier-Ruiz-del-Solar, Alon Shats, Rodrigo Verschae, "Real Time Tracking of Multiple Persons." ICIAP'03, pp109, 2003.

[15] J. M. Menendez, L. Salgado, E. Rendon, N. Garcia., Motion Parametric Modeling for Very Low Bit-Rate Video Coding. Proceedings of SPIE, Vol.3309, 1998.

[16] J. Itten. The Art of Color, Wiley, 1973.

[17] J. Itten. The Element of Color, Wiley, 1970.

[18] J. Zhang, Y Yan, M Lades. "Face Recognition: Eigenface, Elastic Matching, and Neural Nets." Proceedings of the IEEE, 1997.

[19] M. Cooper. Color Smart: How to use color to enhance your business and personal life. Pocket Book, 2000.

[20] S.N.Choi, Y.S.Lee, B.H.Lee, "본소지제조와 그것을 이용한 붉은 색소지 개발", Journal of Korean Ceramic Society. Vol.43, No.10, pp.601-606, 2006.

[21] Tanaka, J. W., Farah, M. J. "Parts and wholes in face recognition". Quarterly Journal of Experimental Psychology, Vol.46A, pp.225-246, 1993.

[22] Turk, M., Pentland, A. "Eigenfaces for recognition." Journal of Cognitive Neuroscience, Vol.3. pp.71-86. 1991.



이 인 정

e-mail : leej@office.hoseo.ac.kr

1981년 전남대학교 수학과 이학사

1984년 중앙대학교 수학과 이학석사

1990년 중앙대학교 수학과 이학박사

2002년 아주대학교 전자공학과 공학박사

1992년~현재 호서대학교 컴퓨터공학부 교수

관심분야 : 영상신호처리, 수치해석, 인공지능