

연속 영상에서의 얼굴표정 및 제스처 인식

한영환·홍승홍*

상지대학교 전자계산공학과, *인하대학교 전자공학과
(1999년 3월 18일 접수, 1999년 7월 8일 채택)

Recognizing Human Facial Expressions and Gesture from Image Sequence

Y.H. Han, S.H. Hong*

Dept. of Computer Science & Engineering, Sangji University
Dept. of Electronic Engineering, Inha University
(Received March 18, 1999, Accepted July 8, 1999)

요약: 본 논문에서는 흑백 동영상을 사용하여 얼굴 표정 및 제스처를 실시간으로 인식하는 시스템을 개발하였다. 얼굴 인식분야에서는 형판 정합법과 얼굴의 기하학적 고찰에 의한 사전지식을 바탕으로한 방법을 혼합하여 사용하였다. 혼합 방법에 의해 입력 영상에서 얼굴 부위만을 제한하였으며, 이 영역에 오픈리플 플로우를 적용하여 얼굴 표정을 인식하였다. 제스처 인식에서는 엔트로피를 분석하여 복잡한 배경영상으로부터 손 영역을 분리하는 방법을 제안하였으며, 이 방법을 개선하여 손동작에 대한 제스처를 인식하였다. 실험 결과, 입력 영상의 배경에 크게 영향을 받지 않고서도 동일 영상에서 움직임이 큰 부위를 검출하여 얼굴의 표정 및 손 제스처를 실시간 적으로 인식할 수 있었다.

Abstract: In this paper, we present an algorithm of real time facial expression and gesture recognition for image sequence on the gray level. A mixture algorithm of a template matching and knowledge based geometrical consideration of a face were adapted to locate the face area in input image. And optical flow method applied on the area to recognize facial expressions. Also, we suggest hand area detection algorithm from a background image by analyzing entropy in an image. With modified hand area detection algorithm, it was possible to recognize hand gestures from it. As a results, the experiments showed that the suggested algorithm was good at recognizing one's facial expression and hand gesture by detecting a dominant motion area on images without getting any limits from the background image.

Key words: Facial expression, Gesture, Recognition, Optical flow, PIM

서 론

문자와 말의 중요성과 더불어 사람의 손짓, 몸짓, 얼굴 표정 등도 의사전달에 있어서 중요한 역할을 담당하고 있다. 따라서 이들을 통하여 상대방의 의사를 인식하려는 연구가 활발히 진행되고 있으며, 구체적으로 얼굴 인식, 표정 인식, 시선 인식, 인술의 움직임, 몸짓 및 손짓 인식 등이 있다[1-2].

인간의 얼굴영상은 복잡하고 다차원적이며 보는 각도 및 조명

등에 따라 수시로 변하기 때문에 처리가 매우 까다롭다. 얼굴 인식을 위해 사용되고 있는 방법으로는 신경망(neural network), 신축성이 있는 형판정합(elastic template matching), KL expansion, 대수적 모멘트(algebraic moment), Isodensity line 등이 있다[3-6]. 그러나 이상의 방법들은 많은 처리시간을 필요로 하고 별도의 하드웨어적인 장치가 필요한 단점이 있으며 얼굴영상의 인식자체에 목적을 둔 연구들이다.

얼굴 영상의 표정 인식에서 연구되고 있는 부분은 정면 얼굴의 농담영상이나 칼라 영상으로부터 눈, 코, 입 등의 형상이나 기하학적 위치관계에 근거한 인식수법이 가장 많이 연구되고 있다. 눈, 코, 입 등의 형상을 찾기 위해서는 형판정합법과 특징에 기반한 정합(feature based matching)법이 사용되고 있다. 특징에 기반한 정합은 인식의 속도가 높고 메모리 효율이 좋으나

이 논문은 1997년 한국학술진흥재단의 공모과제 연구비에 의하여 연구되었음

통신저자: 한영환, (220-702) 강원도 원주시 우산동 660
상지대학교 이공과대학 전자계산공학과
Tel. (0371)730-0485 Fax. (0371)730-0480

문턱치와 같은 설정값에 많은 영향을 받는다. 형판정합법은 추가의 메모리 사용이 있으나 단순하며 좋은 결과를 보여주고 있다. 또 다른 연구 방법으로는 신경망을 통한 표정인식 수법과 얼굴면의 근육의 움직임을 옵티컬 플로우(optical flow)를 이용하여 해석하고, 인식하려는 연구도 진행되고 있다[7-8].

한편, 제스처는 단순히 의사 소통의 보조수단으로 활용되었으나 이를 인식하여 상용 제품에서 이용하는 등 여러 관련 기술에서 활용하려는 연구가 주를 이루고 있다. 그러나 대부분의 연구들이 입력 영상에 대하여 여러 가지 제약을 가하거나 비교적 적은 데이터 집합에 대해서만 실험을 하고 있다. 제스처 인식에 대한 지금까지의 연구는 대개 데이터 클러브를 사용하였다. 데이터 클러브를 사용한 방법은 장갑에 센서를 부착하여 손동작에 대한 3차원 공간상의 정보를 실시간으로 입력받아 사용하므로 비교적 용이하게 손짓의 분석과 인식에 이용되고 있다. 그러나 이 기구는 인식 대상자에게 부자유스러움과 심리적 부담을 주게 된다.

이에 반해 영상처리에 의한 손동작의 이해는 인식 대상자의 동작을 카메라 영상을 통하여 처리하기 때문에 부자유스러운 제약을 요구하지 않는다. 지금까지 영상처리에 의한 몸짓 및 손짓 인식의 연구에서는 표식(marker)에 의한 추적으로 동작을 인식하거나 RGB 칼라 카메라를 이용한 피부색 분석이나 특정 색깔의 장갑을 착용하고 칼라 분석에 의한 이동 궤적이나 손동작을 인식하려는 연구가 진행되고 있다[9]. 그러나 이들 방법들은 표식의 모양이나 색깔에 따른 제한점이 생길 수 있으며, 칼라 영상에서 배경내에 특정 색깔이 중복되거나 동일한 피부색인 얼굴과 손이 겹쳐질 경우 손동작의 추적이 어려워지는 문제점이 있다.

따라서 본 논문에서는 부자유스러운 센서 등을 피하고 사용자 친화성과 범용성 및 실시간 구현을 도모하기 위하여 흑백 동영상을 사용하여 얼굴 표정 및 제스처를 인식하였다.

본 논문에서는 얼굴 인식분야의 두 방법을 혼합하여 사용하였다. 먼저, 연속적으로 변화하는 얼굴영상에서 배경에 크게 영향 받지 않는 형판정합법에 의하여 눈 부위의 후보점을 찾았다. 이 후보점을 기준으로 얼굴의 특징과 기하학적 고찰을 통한 사전지식을 바탕으로 입력 영상에서 얼굴 부위에 대한 대략적인 영역을 제한하였다. 제한된 영역 내에서 옵티컬 플로우를 적용하여 얼굴 표정을 인식하였다.

또한 제스처 인식에서는 표정인식에 사용한 옵티컬 플로우를 사용하지 않았다. 옵티컬 플로우를 연속적으로 변화하는 영상에 적용하여, 제스처를 실시간으로 인식하기 위해서는 많은 처리시간이 필요하므로 별도의 하드웨어적인 장비가 필요하다. 따라서 엔트로피를 분석하여 복잡한 배경영상으로부터 손 영역을 분리하는 방법을 제안하였으며, 이 방법을 개선하여 손동작에 대한 제스처를 인식하였다.

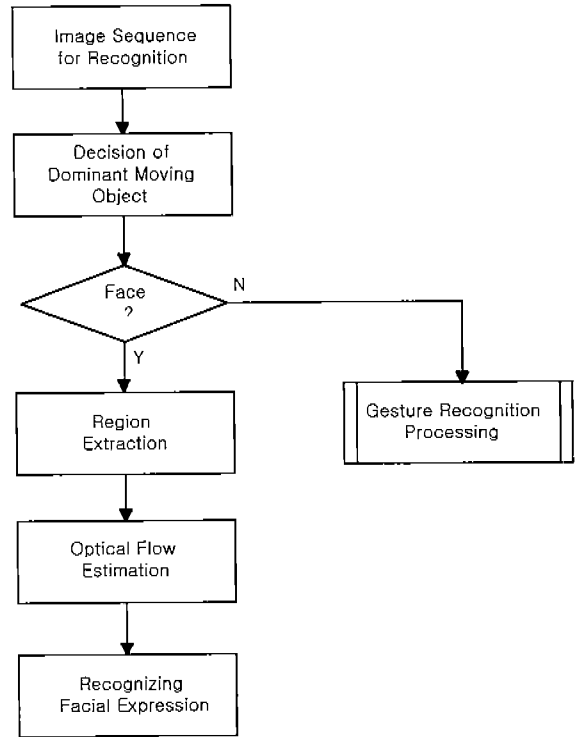


그림 1. 얼굴 표정 인식과정의 흐름도
Fig. 1. Flowchart of recognition process of facial expression

옵티컬 플로우(optical flow)에 의한 얼굴 표정 인식

옵티컬 플로우에 의하여 얼굴 표정을 인식하기 위한 처리과정은 그림 1과 같다. 먼저 연속적으로 입력되는 영상에서 움직임이 많은 부분이 어디인가를 결정한다. 움직임이 많은 부분이 얼굴부분이면 옵티컬 플로우를 적용하여 표정을 인식하고, 얼굴 부분이 아니면 제스처를 인식하는 과정으로 진행된다.

1. 얼굴 영상의 기하학적 특징

얼굴인식을 위해서는 눈, 코, 입으로 이루어진 얼굴 성분들의 거리, 코와 입으로부터 얼굴 측면까지의 거리와 같은 특징들이 사용된다. 그러나 얼굴의 각 성분들에 대한 특징값들은 그림자, 잡음, 머리카락들 때문에 정확하게 분리하기가 어렵다.

얼굴 영상의 기하학적 고찰을 위해 얼굴 모양을 타원으로 근사화하면 개인에 따라 차이가 있으나 그림 2와 같이 장축과 단축의 비가 약 7 : 5 인 타원의 형태로 근사화 할 수 있다.

측정에 의해, 입력 화면에서 얼굴을 인식하기 위한 최대 얼굴 크기와 최소 얼굴크기의 범위를 그림 2와 같이 나타낼 수 있다.

2. 형판 정합에 의한 눈의 위치 설정

형판 정합은 단순성의 이점이 있으므로 입력 영상에서 얼굴의 위치를 찾는 데 사용한다. 사용한 기준형판은 눈에 대한 기준 영

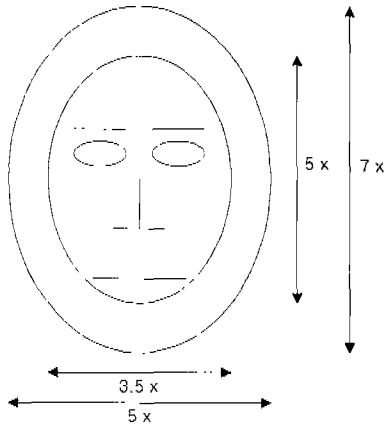


그림 2. 얼굴 인식에 사용되는 범위
Fig. 2. Range for face recognition

상만을 사용하였다. 형판 정합에 의한 눈의 위치 설정은 영상을 구성하는 모든 화소에 대해 많은 연산을 행하게 되므로, 동영상에 적용할 경우에는 많은 시간이 요구되어 실시간적 인식에는 적합하지 않다. 그러므로 초기 입력 영상에서만 형판정합을 사용하여 눈의 위치를 설정하고 이 위치에서 표정을 인식하기 위한 ROI(region of interest) 영역을 설정한다. 형판 정합법으로는 입력영상과의 절대차의 합(sum of absolute difference)을 구해 가장 유사한 값을 갖는 구간을 정합 구간으로 설정하였다.

형판 정합에 의해서 눈의 위치가 짐출되면, 이 위치를 기준으로 ROI 영역내에서 표정을 인식하기 위한 옵티컬 플로우를 계산한다.

3. 옵티컬 플로우 추정

옵티컬 플로우는 영상의 변화를 감지하여 이동 물체를 짐출하거나, 이동 변화량을 알아내는 방법으로 널리 사용되고 있다. 옵티컬 플로우 알고리즘은 영상의 세기변화에 대한 2차원적인 흐름을 추정하는 것으로 2차원 속도의 장(field)으로 정의된다. 2차원 속도는 3차원 공간 속도의 투영이다. 3차원 공간상에서 움직이는 물체를 옵티컬 플로우로 표현하면 그림 3과 같이 표시된다. 옵티컬 플로우는 영상 표면에 2차원 벡터로서 표시되므로 z축에 대한 이동 성분은 무시된다.

옵티컬 플로우 알고리즘은 일반적으로 4개의 부류로 나누어지는데, 미분 방법, 영역에 기반한 성합 방법, 에너지에 기반한 방법, 위상에 기반한 방법 등이다[6]. 각각의 방법들 사이의 경계는 명확하지 않다. 미분 방법은 영상 세기의 시공간 도함수로부터 속도를 계산하는 것으로 1차와 2차 미분 계산에 대한 방법들이 고안되었다. 영역에 기반한 방법은 영역들 사이의 유사도나 거리추정에 의하여 가장 적합함을 가져오는 변위로서 정의된다. 에너지에 기반한 방법은 주파수 영역에서 속도 동조(velocity-tuned) 필터 에너지로부터의 출력을 사용하여 옵티컬 플로우를 계산한다. 위상에 기반한 방법은 대역 필터 출력의 위

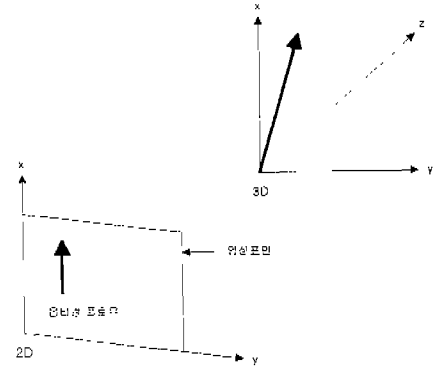


그림 3. 물체의 움직임과 옵티컬 플로우
Fig. 3. Object moving and optical flow

상에 대한 향으로 속도를 정의한다.

어떤 형태의 개체 또는 움직임에 대해, 각각의 점에 대한 정확한 움직임 정보의 추출은 필요하지 않다. 상대적으로 커다란 영역이나 전체 영상에서의 정보 사용으로 발생된 특징이 영역에 기반한 방법으로 참조된다. 영역에 근거한 정합 방법이 미분 방법보다 잡음에 더 견실하며 입력이 비월주사(interlace)인 경우에도 잘 적용된다. 또한 프레임들 사이의 조명 변화에 덜 민감하다.

본 논문에서는 영역에 기반한 방법을 사용하여 옵티컬 플로우를 추정하였다. 영역에 기반한 방법에서 사용하는 특징으로는 여러 가지가 있으나 본 논문에서는 절대 차의 합을 사용하였다. 절대 차의 합 특징은 계산이 가장 간단하기 때문에 다른 특징을 이용하는 것보다 속도 면에서 유리하기 때문이다.

입력 영상의 크기를 $I \times I$ 정합에 사용할 패턴 일부분의 반경을 P , 입력 영상의 탐색 반경을 R 이라고 하자. 화소 좌표 (x, y) 와 부등(disparity) (dx, dy) 에 대한 입력영상 I_1 과 I_2 사이의 유사도를 다음 식 (1)과 같이 정의한다.

$$D(x, y, dx, dy) = \sum_{i=-P}^P \sum_{j=-P}^P | I_1(x+i, y+j) - I_2(x+i+dx, y+j+dy) | \tag{1}$$

좌표 (x, y) 에서 I_1 과 I_2 사이의 흐름을 결정하기 위하여 $(dx, dy) \in [-R, R] \times [-R, R]$ 에 대한 $D(x, y, dx, dy)$ 를 최소화한다.

PIM에 의한 제스처 인식

일반적으로 연속되는 동영상에 대한 움직임을 추출하기 위해서는 프레임간의 차(difference) 영상을 주로 이용한다. 본 논문에서는 주로 DB 검색에 있어서 응용되던 PIM(picture

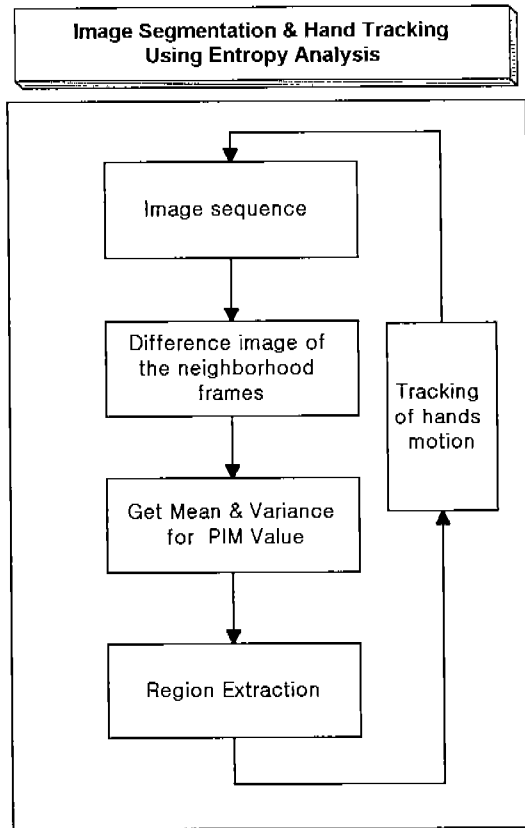


그림 4. 제스처 인식을 위한 손 영역 분할 및 추적 흐름도
 Fig. 4. Flowchart of segmentation and tracking of hand region for gesture recognition

information measure) 방법 중 Chang's PIM 방법을 개선하여 차 영상에 적용함으로써 움직임이 큰 부위를 추적하여 제스처를 인식하였다[10].

움직임이 큰 부위는 PIM의 평균값과 분산을 이용하여 검출하였으며, 손 영역의 제스처에서 영역을 검출하고 추적하기 위한 알고리즘은 그림 4와 같다.

1. PIM에 기초한 움직임 검출

흔히 화상 정보 시스템을 설계할 때에는 주어진 영상의 정보 내용을 측정할 필요가 생긴다. 영상에 포함된 정보량이 적으면 보통 이 정보는 저장되지 않거나 압축되어 저장되어야 한다. 화상 통신의 경우를 생각해 보면, 정보가 적은 영상은 원래 형태로 전송해서는 안되며 압축을 해서 전송을 해야 한다. 따라서 영상의 정보량을 결정해주는 측정 기구를 갖는 것이 중요하다.

본 논문에서는 영상 데이터의 특징을 얻기 위해서 화소간의 엔트로피를 이용하고, 영상에서 엔트로피를 정량화하기 위해 Chang이 제안한 PIM을 사용하였다. 본 논문에서 사용한 PIM을 수식으로 표현하면 다음 식 (2)와 같다.

$$PIM = \sum_{i=0}^{L-1} h(i) - \text{Max},h(i) \tag{2}$$

여기서 h(i)는 각 영상 또는 해당 블록의 히스토그램(histogram)을 의미한다. PIM값은 해당 블록의 전체 화소수와 히스토그램 중 가장 큰 도수의 히스토그램 값과의 차이로 계산된다.

블록 내에 있는 화소의 값들이 모두 같은 경우, 즉 블록의 엔트로피가 '0'인 경우에는 식 (3)과 같이 되어 PIM = 0의 최소값을 갖는다.

$$\sum_{i=0}^{L-1} h(i) = \text{Max},h(i) \tag{3}$$

블록 내에서 화소의 각 레벨 값들이 균일하게 분포하는 경우 즉, 엔트로피가 큰 경우에는 Max,h(i)이 작은 값을 갖기 때문에 큰 PIM값을 갖게 된다.

PIM의 최대값을 구하기 위하여 Chang이 제안한 것에 다음의 내용을 추가하여 알고리즘을 개선하였다.

For M×N size, L-gray level image
 if M×N ≤ L,
 Max(PIM) = M×N-1
 else
 Max(PIM) = M×N- [M×N/L]
 ([n] : n을 넘지 않는 최대의 정수)

PIM은 해당 블록이 많은 정보를 갖는 경우에는 큰 값을 얻을 수 있고, 적은 정보량을 갖는 경우에는 작은 값을 얻을 수 있다. 또한 정규화 PIM(Normalized PIM)은 식(4)와 같이 정의되며 n(f)는 화소의 총 수이다.

$$NPIM = 1 - \text{Max}, \left[\frac{h(i)}{h(f)} \right] \tag{4}$$

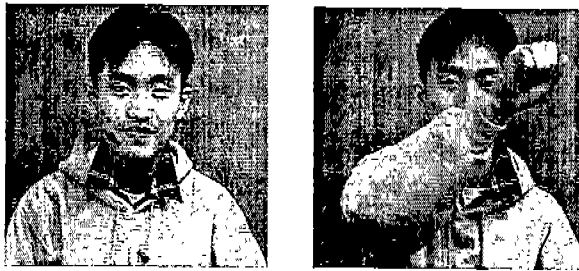
PIM_k는 임의의 영상을 k 그레이 레벨을 갖는 것으로 변환하기 위한 그레이 레벨 변화의 최소수로서 정의된다. 일반적으로 PIM_k, NPIM_k는 각각 식(5), 식(6)과 같이 정의된다.

$$PIM_k = \sum_{i=0}^{L-1} h(i) - \sum_{i=0}^k b(i) \tag{5}$$

$$NPIM_k = 1 - \sum_{i=0}^k p(i) \tag{6}$$

실험결과 및 고찰

본 논문에서 목적으로 하는 얼굴표정 및 제스처의 인식을 위한 시스템 구성은 다음과 같다. 연속 영상에서 실시간 인식 시스템을 구현하기 위하여 펜티엄II 266MHz의 컴퓨터를 사용하였으며 Visual C++ 6.0 으로 프로그램 하였다. 본 논문에서는 다음과 같은 가정을 하였다.



#1_1

#2_1



(a)

(b)

그림 6. 얼굴영역의 ROI
Fig. 6. ROI of face



#1_2

#2_2



(a)

(b)

그림 7. 옵티컬 플로우의 적용 결과
Fig. 7. Applied result of optical flow



#1_10

#2_10

(a)

(b)

그림 5. 입력 영상의 예
Fig. 5. Example of input image

- 배경은 상대적으로 정적이다.
- 조명의 변화가 작다.
- 영상내에서 움직임이 있는 부분은 얼굴 또는 동작에 의한 제스처 부분이다.

제안된 얼굴표정 및 제스처 인식 시스템의 타당성을 조사하기 위하여, 각각 10개의 연속적인 프레임으로 구성되는 5명의 피검자에 대한 영상을 획득하여 실험하였다. 이중 한 사람의 입력 영상데이터의 예를 그림 5에 나타내었다. 그림에서 (a)는 움직임이 큰 부분이 얼굴영역인 경우에 사용하는 10개의 입력 영상

이며, (b)는 움직임이 큰 부분이 손동작에 대한 제스처인 경우에 사용되는 10개의 입력 영상이다.

본 논문에서 제안한 방법에서는 얼굴 표정을 인식하기 위하여 첫 번째 입력영상에 형판정합을 적용하여 얼굴부위의 ROI 영역을 찾았다. 형판정합에 의하여 안경을 쓴 경우나 얼굴내의 수염 등과 같은 것이 존재하는 경우에도 얼굴의 ROI 영역을 잘 찾을 수 있었다. 안경을 착용하지 않은 경우와 착용한 경우에 형판정합에 의해 찾아진 시작점 및 ROI 영역을 그림 6(a), (b)에 나타내었다. 그림 6의 (a), (b)에서 찾아진 ROI 영역이 그 이후에 입력되는 연속영상에서 표정을 인식하기 위한 영역으로 사용된다. 입력 영상 전체를 대상으로 하지 않고, 제한된 영역내에서만 옵티컬 플로우를 계산하므로써 상당한 계산량을 줄일 수 있으며 빠르게 얼굴 표정을 인식할 수 있다.

표정을 인식하기 위해 찾아진 ROI 영역을 그림 7(a)에 나타내었으며 옵티컬 플로우를 적용한 결과를 그림 7 (b)에 나타내었다. 그림 7에서 볼 수 있듯이 입 주변에서 옵티컬 플로우의 크기가 변화되었으며, 볼 주변에서도 플로우의 흐름변화가 있는 것을 알 수 있다. 따라서 그림 7 (b)의 플로우 변화는 웃음에 의하여 변화된 것으로 짐작할 수 있다.

그림 8은 배경이 존재하는 영상에서 움직임이 큰 부분이 얼굴이 아니고 제스처인 경우에 실험결과를 구하기 위한 중간과정을 나타낸 것이다. 두 입력 영상의 차영상을 구하고, 이 차영상에 PIM을 적용하여 손 영역만을 분리(segmentation)한 결과를 보여준다. 그림 8의 (a)와 (b)는 두 개의 입력 영상이고, 이 두

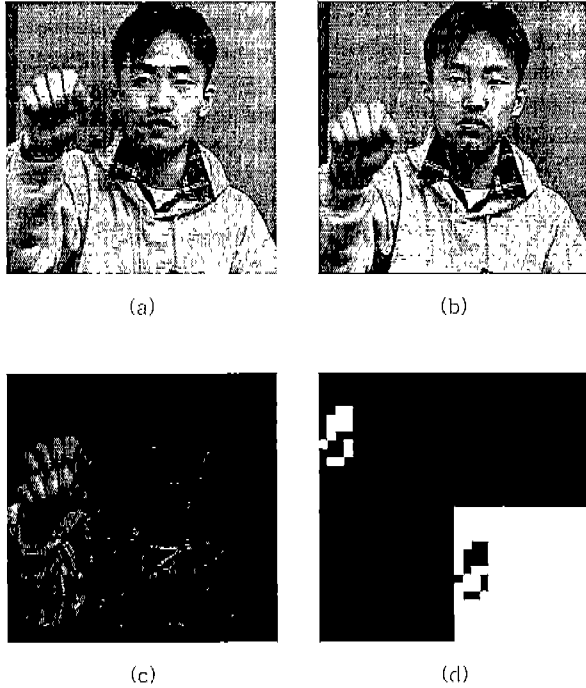


그림 8. PIM에 의한 손 영역의 추출
Fig. 8. Extraction of hand region using PIM

영상의 차 영상 결과가 (c)이며 PIM을 적용한 결과가 (d)이다. (c)영상을 블록으로 나누어서 처리하였기 때문에 (d)영상에서 움직임이 가장 큰 손 영역만이 PIM에 의해 검출되는 것을 알 수 있다. (d)의 우측 하단은 입력 영상을 가로, 세로 각각 1/2로 줄여서 처리한 결과를 나타내었다.

그림 9에는 PIM을 적용하여 제스처를 인식한 최종 결과 화면을 나타내었다. 다양하고 복잡한 제스처의 효과적인 인식을 위하여 제스처의 시작에 대한 정의를 하였다. 움직임은 주로 손 동작에 의한 개별적인 제스처를 사용하여 인식을 행하였다.

화면에서 볼 수 있듯이 한글 자음 8개(ㄱ, ㄴ, ㄷ, ㄹ, ㅁ, ㅂ, ㅅ, ㅇ)를 정의하여 손동작에 의한 인식을 행하였다. 화면에서 총 프레임 수는 하나의 제스처를 형성하는 영상의 수로 10개의 입력 영상이 사용되었으며, 시작점 및 끝점의 좌표는 왼쪽 하단의 좌표를 (0, 0)으로 하였을 때의 시작점 및 끝점에 대한 좌표이다. 각은 2개의 입력영상에서 움직임이 있는 블록간의 각도를 표시한 것으로 10개의 입력 영상에서 9개의 간격에 대한 각도를 표시한 것이다. 한 예로 'ㄱ'에 대한 동작을 나타낼 때, 7번과 8번째 영상 사이에서 각도 차이가 많은 것을 알 수 있다. 최종 인식 결과는 한글 자음 8개중에서 'ㄱ'에 대한 동작이라는 것을 우측의 버튼에 의해 표시하였다.

표 1에는 5사람에 대한 한글 자음 8가지에 대한 손 제스처의 인식결과를 정리하여 나타내었다. 인식 결과 'ㄹ', 'ㅂ'에서는 좋지 않은 결과를 나타내었는데, 이는 영상의 획득과 손 제스처에 대한 속도가 맞지 않은 원인에 의한 것이다. 예를 들어 'ㄱ'에

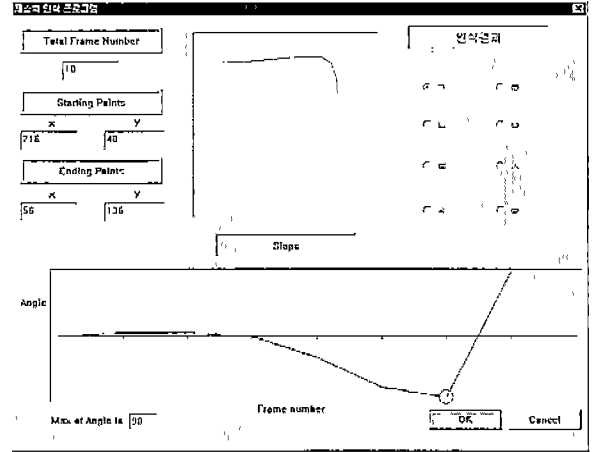


그림 9. 제스처 인식의 결과
Fig. 9. Result of gesture recognition

표 1. 제스처의 인식결과
Table 1. Recognition results of gesture

		(단위 : 회수)							
인식결과	제스처	ㄱ	ㄴ	ㄷ	ㄹ	ㅁ	ㅂ	ㅅ	ㅇ
	ㄱ		5						
ㄴ			5						
ㄷ				5					
ㄹ					3				
ㅁ						5	2		
ㅂ							3		
ㅅ								5	
ㅇ									5
오류					2				

대한 동작을 하는데 움직임 속도가 너무 빨라 1, 2 프레임 내에서 동작을 마쳤을 때 나머지 3~10 프레임에서는 움직임에 대한 내용이 없게 된다. 따라서 1, 2 프레임 내에 있는 움직임 정보만을 사용해서는 'ㄱ'에 대한 정확한 인식을 할 수 없다. 또한 움직임 속도가 너무 느린 경우는 1~10 프레임 내에 'ㄱ'에 대한 완전한 동작이 입력되지 않고 'ㅡ'으로 입력되므로 정확하게 인식을 할 수 없다.

다른 원인으로서는 움직임 동작에서 동작의 크기가 너무 작은 경우, 마지막 동작이 화면 밖에서 종료된 경우에도 정확하게 인식하지 못하는 결과를 보였다.

결론

본 논문에서는 연속영상에서의 얼굴표정 및 제스처를 실시간으로 인식하기 위한 시스템을 제안하였으며, 실현에 의해 얻어

진 결과는 다음과 같다.

제안된 시스템은 기존의 방법과 다르게 문턱치와 같은 설정값에 영향을 받지 않으며, 초기 영상에만 형판정합을 사용하여 인식범위를 제한하는 등 계산이 간단한 방법을 사용하여 실시간적인 인식이 가능 하였다.

얼굴 표정을 인식하기 위한 기존의 방법들에서는 입력 영상의 획득시, 영상의 대부분이 얼굴이거나 배경을 제한하는 등의 제약조건이 있었으나 본 논문에서의 방법은 이러한 제한이 없었으며 안경을 쓴 경우나 수염이 있는 경우에도 사용할 수 있다.

5명의 피검자를 대상으로 한글자음 8개에 대한 제스처 인식 실험을 한 결과 ‘ㄱ’, ‘ㅂ’ 에서만 올바르게 인식하지 못했다. 고찰에서 언급한 내용이 보완된다면, 언어로 의사소통이 불가능한 환자와의 의사소통 등에 적용될 수 있을 것으로 기대된다.

앞으로 얼굴표정의 인식에서 생면 얼굴 영상이 아니고 얼굴이 회전된 경우에도 인식될 수 있는 알고리즘과 연속적인 제스처를 더 자연스럽게 인식하고 표현하기 위한 연구가 필요할 것이다.

참 고 문 헌

1. R.Chellappa, C.H.Wilson, S. Sirohey, "Human and Machine Recognition of Faces:A Survey", Proc. of the IEEE, Vol.83, No.5, pp.705-740, 1995
2. 한영환 외, "환자 감시장치를 위한 눈의 개폐상태 인식에 관한 연구", 대한전자공학회 논문지, 제32권 11호, 1995
3. R.Brunelli, T.Poggio, "Face Recognition:Features versus Templates", IEEE Trans., PAMI, Vol.15, No.10, pp.1042-1052, 1993
4. Kirby M.and Sirobich L, "Application of the Karhunen-Loeve procedure for the Characterization of Human Faces", IEEE Trans., PAMI, Vol.12, No.1, pp.103-108, 1990
5. Hong Z.Q, "Algebraic Feature Extraction of Image for Recognition", Pattern Recognition, Vol.24, No.3, pp.211-219, 1991
6. Nakamura O,Mathur S and Minami T, "Identification of Human Faces Based on Isodensity Maps", Pattern Recognition, Vol.26, No.3, pp.263-272, 1991
7. Y.Yacoob, L.S.Davis, "Recognizing Human Facial Expressions from Long Image Sequences Using Optical Flow", IEEE Trans., PAMI, Vol.18, No.6, pp.636-642, 1996
8. S.S.Beauchemin, J.L.Barron, "The Computation of Optical Flow", ACM Computing Surveys, 27(3), 433-467, 1995
9. T.J.Darrell, I.A.Essa, "Task-Specific Gesture Analysis in Real-Time Using Inter-polated Views", IEEE Trans., PAMI, Vol.18, No.12, pp.103-108, 1996
10. S.K.Chang, Principles of Pictorial Information Systems Design, Prentice-Hall, 1989