

손의 구조적 정보를 이용한 지문자 영상의 인식

최성현^o 양윤모
고려대학교 전자정보공학과
(shchoi, ymyang)@hard.korea.ac.kr

A Hand-Gesture Recognition Using Structural Information of Hand

Sung-Hyun Choi Yun-Mo Yang
Dept. of Electronics & Information Engineering, Korea University

요 약

본 논문은 동작자가 표현하는 수화 동작 중, 지문자 영상을 2차원 영상처리에 의하여 인식하는 방법을 제안한다. 손은 개인에 따라 변량이 존재하고 같은 동작을 표현하더라도 항상 일정하지 않기 때문에, 지문자 인식을 위하여 손의 구조적인 정보를 이용한다.

추출된 손 영역을 입력으로 하여 손의 외곽 정보를 이용한 MAT(Medial Axis Transform)를 수행한다. 여기에서 나온 골격의 변화에 따라 각각 손가락에 해당하는 2번 속(Shock)과 손 바닥에 해당하는 4번 속의 관계에 의하여, 지문자를 인식한다. 이러한 구조적인 정보를 이용함으로써 개인에 의해 나타나는 표현의 차이를 제거할 수 있고, MAT를 이용할 때 나타나는 외곽영역의 잡음에 의한 구조 변화에도 안정적으로 대처할 수 있다.

제안한 알고리즘을 이용하여 31개의 단순 지문자에 대한 270개의 입력영상을 실험한 결과 81.1%, 모양이 흡사한 지문자를 통합하여 26개의 지문자로 식별할 경우에는 91.1%의 인식율을 나타내었다.

1. 서론

동작자가 표현하는 수화 인식을 위해 다각도로 연구가 진행되고 있다. 그러나 대부분의 방법은 인식상의 어려움으로 인하여 영상 이외의 외부 입력을 이용한다.

외부 입력을 받기 위하여 사용되는 전기적 장치 또는 도구들은 사용자의 편의성과 사용상의 범용성을 고려할 때 많은 한계를 나타내고 있다.

본 논문에서 제안하는 방법은 손의 구조적 정보와 2차원 영상처리 기법만을 이용하여 그림 1과 같이 지문자 영상 인식을 수행한다.

2. 전처리

본 논문은 240×180 화소, 256 계조 영상으로 표현되는 31개의 단순 지문자(자음 14개, 모음 17개)에서 MRF(Markov Random Field)와 LPM(Line Process Model) 의해 추출한 손 영역을 입력 영상으로 하였다 [1]. 실제 실험에 사용되는 손의 크기는 100×100 이내로 고르지 못한 외곽 영역을 갖는 저 품질 영상을 사용하였다[그림 2-가].

지문자 인식에서 사용할 특징값 추출을 위한 전처리 단계로서, 원 영상을 이진화 하여 골격(skeleton)을 검출한다.

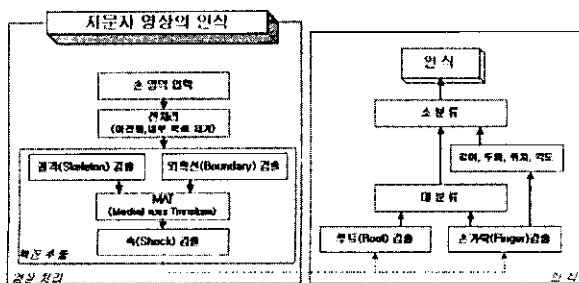
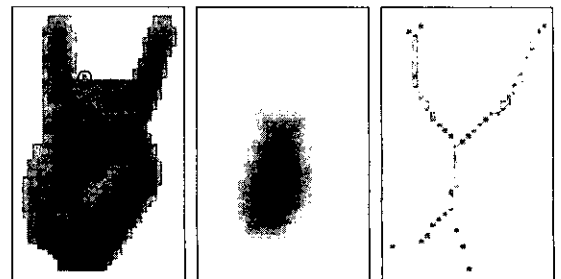


그림 1. 지문자 인식 흐름도



(가) 원영상 (나) Distance Map (다) ridge검출
그림 2. 지문자 "ㅈ"

골격(skeleton)검출을 위해 본 논문에서는 입력 영상의 크기, 골격화 알고리즘의 계산량과 속도 등을 고려하여 거리변환(Distance Transform)을 이용한다.

손의 외곽선에 존재하는 작은 변화가 골격 검출에 큰 영향을 끼칠 수 있다[2]. 따라서, 손의 원 구조와 관계없는 픽셀의 작은 변화를 제거할 필요가 있다. 내부에 존재하는 배경 픽셀은 골격 구조에 지대한 영향을 끼칠 수 있고, [그림 2-가]의 작은 원 안에 있는 한 픽셀 두개의 픽셀은 거리 변환에 의한 골격 검출 시 의도하지 않는 구조를 생성하기에 제거할 필요가 있다.

$$D_c = |x_2 - x_1| + |y_2 - y_1| \quad (1)$$

위 수식(1)은 영상의 좌표 (x_1, y_1) 과 (x_2, y_2) 사이의 City-Block 거리를 나타낸다.

[그림 2-나]는 이진 영상에 Distance Transform을 적용한 영상의 Distance Map이다.

골격(skeleton)검출을 위하여 Distance Map에서 Ridge에 해당하는 부분을 검출한다. 중심 픽셀이 상-하 또는 좌-우 또는 사선 방향의 픽셀 값들보다 크면 Ridge로 검출한다[그림 2-다].

이산 영상의 경우 [그림 2-다]와 같이 Ridge검출만으로는 골격 선이 끊어진 경우가 발생한다. 이것은 중심 픽셀을 중심으로 최대 픽셀 값이 하나 이상 존재할 때 나타나게 된다. 분리된 영상을 연결하기 위해서 많이 사용하는 방법으로는 최대 픽셀 값 중 임의의 픽셀을 선택하여 골격을 연결하는 방법이 있다. 그러나 위와 같은 방법은 골격 선이 떨어진 정도가 클 경우 의도하지 않은 방향으로 나아갈 수가 있고, 인식에서 이용할 특징 값으로 골격의 방향성에 많이 의존하게 되는데 인식에 좋지 않은 영향을 끼치게 된다.

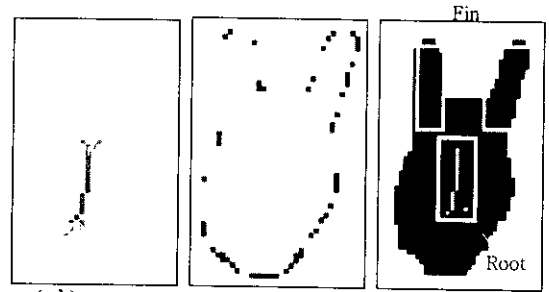
본 논문에서는 이러한 단점을 보완하기 위하여 이진 골격과 방향 관계를 이용하여 끊어진 골격을 연결한다.

먼저, 골격의 연결을 시도하기 위하여 서로 연결된 픽셀을 같은 픽셀 값으로 치환하는 Numbering과정을 수행하고, 같은 픽셀값을 갖는 골격의 시작과 끝 부분에서 연결을 시도한다. 연결 시도는 (1)거리 변환(Distance Transform)을 수행한 영상(Distance Map)에서 연결을 시도하는 픽셀에 대응하는 픽셀의 8방향 이웃 중 가장 큰 값을 갖는 픽셀들과 (2)이전 골격픽셀과의 방향성을 고려하여 다음 골격으로 예상되는 예상 골격픽셀의 위치를 선택한다. 예상 골격픽셀에서는 (1)픽셀의 8방향 주변의 0이 아닌 픽셀의 개수, (2)8방향 주변 픽셀의 패턴 변화를 이용하여 골격 픽셀의 진위 여부를 가른다.

3. Medial Axis Transform(MAT)과 속(shock)을 이용한 특징값 추출

전처리 과정에서 검출한 골격과 원 영상의 외곽 정보를 이용하여 MAT를 계산할 수 있다. MAT는 골격에서 원의 반지름을 점진적으로 증가 시켜 나가면서 2개 이상의 외곽선 영역에서 원이 접하게 되는 반지름 값을 갖는 골격들로 구성되어지고, 픽셀 값은 원의 반지름으로 표현된다 [그림 3-가][2]. 보다 정확한 MAT계산을 위하여 작은 크기의 이산 입력 영상일 경우, 골격에서 나온 원과 외곽선의 두 접점이 일정한 거리 이상이 되어야 한다.

또한 MAT를 이용하여 원 영상의 복원이 가능하고 복원을 통하여 추출한 MAT값이 원 영상을 정확하게 표현하고 있는지 알 수 있다.[그림 3-나]에서 밝게 표현된 부분이 원 영상이고 검은 부분이 MAT에 의해 복원된 영상과의 차이가 된다. 이러한 차이는 골격에서 그런 원이 2개 이상의 외곽선 영역과 접촉할 때, 원의 반지름을 MAT값으로 하면서 발생하지만, 특징 값 추출과 인식에 있어서는 영향을 끼치지 않는다. 외곽선과 접촉한 원의 두 접점의 길이를 조절하면 차이를 줄일 수 있다.



(가) MAT (나) 복원 (다) 특징 추출
그림 3. 특징 값 추출

평면내의 폐곡선에 적용되는 curve evolution 과정에서 본 논문의 특징 값으로 사용하고자 하는 속을 추출할 수 있다[2].

$$C_t = (1 + \alpha k)N \quad (2)$$

$$C(s, 0) = C_0(s)$$

$C(s, t)$ 는 곡선의 벡터 좌표를, $N(s, t)$ 는 내부를 향하는 법선 벡터(normal vector), s 는 path 매개변수, t 는 전개 시간을 나타낸다. 위 식(2)에서 $\alpha = 0$ 일 경우, 전개 방정식은 쌍곡선(hyperbolic)형태를 지니게 되고 여기에서 속이 검출된다.

즉, 골격과 외곽선과의 거리(=MAT)의 변화량을 조사하여 닫힌 폐곡선 영역을 위의 4가지 속으로 표현 가능하다[4].

동작자가 표현하는 지문자는 손의 형태를 보고 대상이 인식하게 되고, 손의 형태를 표현하는 데는 손가락과 손바닥이 중요한 역할을 한다. 따라서, 본 논문에서는 입력 영상으로부터 손가락과 손바닥에 해당하는 영역을 이용하여 지문자 영상을 인식한다.

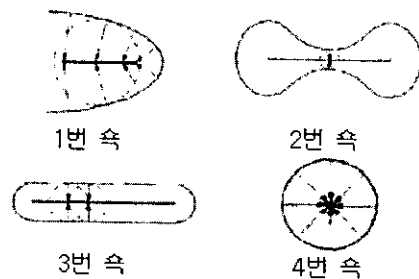


그림 4. 4가지 형태의 속(Shock)

4번 속은 최대 MAT값을 갖는 픽셀을 나타내고, 손바닥 영역 내의 한 픽셀로 나타난다. 4번 속과 연결된 일정한 범위 이내의 MAT를 손바닥 영역으로 추출하기 위하여 문턱 값을 두었다. 본 논문에서는 최대 픽셀 값의 15%를 문턱 값으로 사용하였다.

지문자 영상의 인식을 위하여 다음의 *Fin* 과 *Root* 의 두 가지 특징값을 이용한다. 각 특징값은 해당 영역의 MAT(위치+반지름)값이다. [그림 3-다]에 추출된 특징 값들이 원 영역보다 밝게 표시되어 있다.

- Fin* : 3번 속. 입력 영상에서 손가락 영역에 해당.
- Root* : 4번 속 - 문턱 값. 입력 영상에서 손바닥 영역에 해당.

4. 지문자 인식

31개의 단순 지문자에는 각 지문자마다 미리 정의된 표현 방식이 존재한다. 이 각각의 지문자는 아래의 특징량을 이용하여 검출 할 수 있다.

표1. 특징량 분류

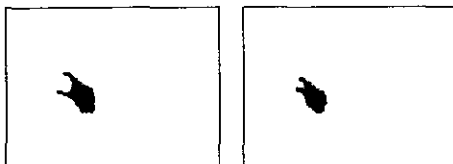
1차 특징량	Fin의 개수, Root의 방향, Fin의 방향, Fin의 접촉
2차 특징량	Fin의 위치, Fin의 길이, Fin의 두께, 1번 속 검출

표2. 1,2차 특징량에 의해 분류 가능한 지문자

1차 특징량에 의한 인식	ㄱ, 一, ㄴ, ㄷ, ㅍ, ㅑ, ㅓ, ㅕ, ㅗ, ㅛ, ㅜ, ㅠ, ㅡ, ㅝ, ㅞ, ㅟ, ㅠ, ㅡ, ㅢ
2차 특징량에 의한 인식	(ㄱ, ㅛ, ㅜ), (ㅓ, ㅕ, ㅗ, ㅛ, ㅜ), (ㅓ, ㅕ), (ㅓ, ㅕ, ㅗ, ㅛ)

31개의 단순 지문자 가운데 1차 특징량만을 이용하여 인식 가능한 지문자는 16개이다[표 2]. 아래 [그림 5]는 1차 특징량 -Fin의 개수(2개), Root의 방향(NorthWest), Fin의 방향(NorthWest), Fin의 접촉(No)-의 같이 동일하여 인식을 위하여 추가 정보가 필요한 경우이다. 31개의 지문자 가운데 1차 특징량을 이용하여 인식가능한 16개의 지문자를 제외한 나머지 15개의 지문자의 경우, 2차 특징량을 이용하여 인식한다.

위의 알고리즘을 이용하여 270개의 입력 영상에 대하여 인식을 수행하여 81.1%의 인식율을 얻을 수 있었다. 1,2차 특징량에 의한 인식결과, 오인식된 지문자중 많은 지문자가 다른 하나의 지문자로 인식되는 결과를 나타내었다. 이러한 오인식의 경향은 2차원 형태정보만을 이용할 경우 인식과정을 수행함에 있어 정보가 부족한 경우와 손영역을 분리함에 있어 일괄되지 않기 때문이다.



(가) 지문자 "귀" (나) 지문자 "극"

그림 5. 1차 특징량으로 인식 불가능한 지문자

표3. 1,2차 특징량에 의한 인식결과(정인식/실험데이터)

ㄱ	7/10	ㅓ	5/6	ㅗ	10/10	ㅛ	10/11
ㄴ	10/11	ㅕ	7/8	ㅜ	8/8	ㅠ	10/11
ㄷ	10/10	ㅗ	3/6	ㅛ	5/6	ㅜ	7/8
ㄹ	8/10	ㅝ	8/9	ㅞ	4/8	ㅟ	8/8
ㅁ	0/10	ㅞ	10/10	ㅟ	8/8	ㅠ	1/2
ㅂ	6/11	ㅟ	3/10	ㅡ	6/6	ㅢ	10/11
ㅅ	5/8	ㅡ		ㅢ	6/8	ㅣ	11/11
ㅇ	5/7	ㅣ	11/11	ㅤ	9/9	ㅥ	8/8

지문자 'ㅁ', 'ㅂ'의 경우 오인식된 지문자의 대부분이 'ㅞ'으로 인식되었다. 형태상 유사하고 정보의 부족으로 인하여 발생하는 오인식을 해결하기 위하여 인식보다는 유사한 지문자를 따로 분류할 필요가 있다.

표4. 유사 지문자 세트

'ㅞ' 군	"ㅛ", "ㅜ", "ㅟ"
'ㅓ' 군	"ㅎ", "ㅓ"
'ㅗ' 군	"ㅗ", "ㅛ"
'ㅕ' 군	"ㅕ", "ㅋ"

표4의 유사 지문자 세트- 입력 지문자가 'ㅛ', 'ㅞ', 'ㅟ' 일 경우 'ㅞ' 군으로 인식- 까지 인식하였을 경우 인식율은 91.1%이다.

5. 결론 및 발전방향

본 논문에서는 MRF를 이용하여 검출된 손 영역 영상에 대하여 인식을 수행하였다.

영상처리만을 이용하여 입력 지문자에 대한 인식을 수행하였고, 잡음에 의한 특징 값의 변화를 3,4번 속만을 이용함으로써 제거하였다. 31개 지문자에 대한 270개의 입력영상에 대하여 81.1%의 인식율을 나타내었고, 모양이 흡사한 지문자를 통합하여 26개의 지문자로 식별할 경우 91.1%의 인식율을 나타내었다.

보다 나은 인식 성능을 위하여 정확한 손영역의 분리, 특징량 추출 후 해석적 인식 알고리즘이 필요하다.

6. 참고 문헌

[1] 송효섭, 양운모, " MRF를 이용한 수화 동영상에서의 효율적인 손 형상 추출", 정보 과학회 추계 학술대회 논문집, 2000, pp.395-397

[2] Jonas August, Kaleem Siddiqi and Steven W.Zucker. " Ligature Instabilities in the Perceptual Organization of Shape". Proc. CVPR(1999), vol.2, pp.42-48

[3] Kaleem Siddiqi and Benjamin B. Kimia " A Shock Grammar for Recognition", CVPR, 1996

[4] Kaleem Siddiqi, Ali Shokoufandeh, Sven J. Dickinson and Steven W. Zucker, " Shock Graphs and Shape Matching", Computer Vision, 1998. Sixth International Conference on , 1998 pp.222-229