

# 얼굴과 얼굴 특징점 자동 검출을 위한 탄력적 특징 정합

박호식<sup>\*</sup> · 손형경<sup>\*</sup> · 정연길<sup>\*</sup> · 배철수<sup>\*</sup>

<sup>\*</sup>관동대학교

## A Flexible Feature Matching for Automatic Face and Facial Feature Points Detection

Ho-Sik Park<sup>\*</sup> · Hyung-Kyung Son<sup>\*</sup> · Youn-Gil Jung<sup>\*</sup> · Cheol-Soo Bae<sup>\*</sup>

<sup>\*</sup>Kwandong University

E-mail : mediana@unitel.co.kr

### 요 약

본 논문에서는 자동적으로 얼굴과 얼굴 특징점을 검출하는 시스템을 제안하였다. 얼굴은 Gabor 특징에 의하여 지정된 특징점의 교점 그래프와 공간적 연결을 나타내는 에지 그래프로 표현하였으며, 제안된 탄력적 특징 정합은 모델과 입력 영상에 상응하는 특징을 취하였다. 정합 모델은 국부적으로 경쟁적이고 전체적으로 협력적인 구조를 이룸으로서 영상공간에서 불규칙 확산 처리와 같은 역할을 하도록 하였다. 복잡한 배경이나 자세의 변화, 그리고 왜곡된 얼굴 영상에서도 원활하게 동작하는 얼굴 식별 시스템을 구성함으로써 제안된 방법의 효율성을 증명하였다.

### ABSTRACT

An automatic face and facial feature points(FFPs) detection system is proposed. A face is represented as a graph where the nodes are placed at facial feature points(FFPs) labeled by their Gabor features and the edges are describes their spatial relations. An innovative flexible feature matching is proposed to perform features correspondence between models and the input image. This matching model works likes random diffusion process in the image space by employing the locally competitive and globally corporative mechanism. The system works nicely on the face images under complicated background, pose variations and distorted by facial accessories. We demonstrate the benefits of our approach by its implementation on the face identification system.

### 키워드

Gabor Feature, Face Representation, Flexible Feature Mapping , Facial Feature Point(FFP)

## 1. 서 론

얼굴은 개략적으로 동일한 기하학적 구성으로서 눈썹, 눈, 코, 입, 턱과 같은 얼굴 특징(Facial Features: FFs)으로 구성된다. 이러한 사실은, 얼굴과 얼굴 특징의 위치, 지역적 특징 정보와 공간적 연결 같은 얼굴 구조를 이해하는데 있어 중요한 요소이다. 그러므로, 영상에서 자동적으로 얼굴부분을 결정하는데 있어서는 얼굴 특징과 얼굴 특징점(얼굴 특징을 나타낸 지점)은 가장 기본적인 요소이다.

많은 이전의 연구에서 템플릿 매칭과 에지 투영 그리고 변형 모델 같은 방법으로 얼굴 특징점

을 찾고자 하였다. 템플릿 매칭은 얼굴 표정에 의한 기하학적 변화, 왜곡, 변형에 민감하고, 좌우 대칭 정보를 얼굴과 얼굴 특징에 위치시킨 에지 투영은 Brunelli[7]에 의하여 제안되었으나 조명에 민감하게 반응함으로써 에지 특징 정보만을 이용하고 있다. 얼굴 특징점을 찾기 위한 능동적 윤곽선 모델(스네이크)은 Kass[5]에 의하여 단일 스네이크방법이, Wu[1]에 의해 복합 스네이크와 칼라 정보를 사용하는 방법이 제안되었다. 스네이크는 활동적 작용을 최소화하기 위하여 평탄한 표면을 필요로 하였으며, 이러한 접근 방식은 왜곡에 민감하고 고화질의 영상을 필요로 하였다.

Lades [6]과 Wiskott[4]는 Gabor 특징으로 분류된 그래프를 이용하여 얼굴을 표현하였고, 분류 그래프 정합으로 표현한 모델은 입력 영상과 조화를 이루었다.

본 논문에서는 복잡한 배경, 자세의 변화, 그리고 안경이나 모자 등에 의해 왜곡된 영상에서도 얼굴과 얼굴 특징점을 검출할 수 있는 새로운 시스템을 제안하고자 한다. 얼굴은 얼굴 특징점이 위치한 교점에서 그래프로서 표현되어지고, 얼굴 특징점의 위치 정보는 Gabor 웨이블릿 변환(GWT: Gabor Wavelet Transformation)에 의하여 대응되는 지점에서 추출한다. 얼굴의 모델 데이터 베이스는 상이한 사람들의 얼굴을 표현한 그래프의 모음인 얼굴 모델 그래프(FMGs: Face Model Graphs)를 통하여 변형된다. 얼굴 모델 그래프(FMGs)는 얼굴과 얼굴특징점 검출, 얼굴 식별을 위한 시스템에서 지식으로서 사용되어진다. 본 논문에서는 얼굴과 얼굴 특징점의 효율적인 자동 검출을 위하여 탄력적 특징 정합을 제안하였다. 얼굴 식별 시스템을 구성함으로써 제안된 방법의 효율성을 증명하였다.

### II. 시스템 구성

본 논문에서는 얼굴과 얼굴특징점 자동 검출 시스템을 제안하였다. 시스템은 얼굴모델그래프(FMGs) 발생기, 얼굴 검출, 얼굴특징점 검출 절차로 구성되어있으므로, 결과적으로 효율적인 얼굴 식별방법을 제안하였다. 시스템 구성도를 그림 1에 나타내었다.

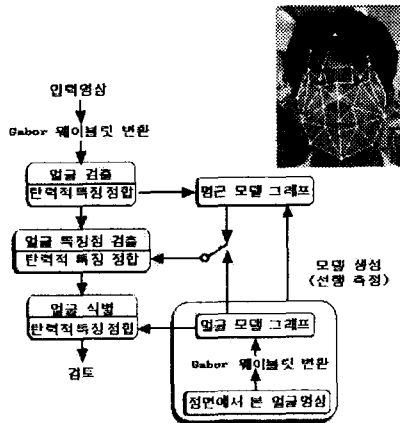


그림 1. 시스템 구성도

### III. 얼굴 표현

Gabor 웨이블릿을 이용한 전처리는 국한된 형

상[2]과 실험적 습득[3]으로 선택되어진다. 영상의 Gabor 웨이블릿 변환(GWT)은 영상과 Gabor 커널 사이의 컨벌루션 과정이다.

$$\phi_{k,\theta} = \frac{K^2}{\sigma^2} \exp\left(-\frac{k^2(x^2+y^2)}{2\sigma^2}\right) \dots \left[\exp(i k(x\cos\theta + y\sin\theta)) - \exp\left(-\frac{\sigma^2}{2}\right)\right] \quad (1)$$

$$\theta = \frac{k\pi}{N_\theta} \text{ with } k=0, \dots, N_\theta-1 \quad (2)$$

식(1)에서 첫 번째 요소는 정규화 된 가우스 창이다. 가우스 창의 너비는 변수  $\sigma=2\pi$  에 의하여 조절되어지고, 두 번째 요소는 웨이블릿의 사인곡선 성분이다.  $(x, y)$  항은  $\theta$  방향으로의 복소지수를 결정하는데 사용된다. 커널은 식(1)의 괄호 안의 두 번째 항에 의하여 DC 값에 적합하게 조정된다. 방위에 대한 민감도는 방위방향  $\theta$ 에 달려 있고,  $N_\theta$  는 지수  $k$  에 대한 방위의 수이다.

Gabor 커널에 대한 영상의 컨벌루션은 매 중심 지점과, 방위, 비율에 대하여 복소 특징을 산출한다. 지점에 대해서는 점진적으로 변하므로 진폭에 대한 정보만을 연산에 사용한다.

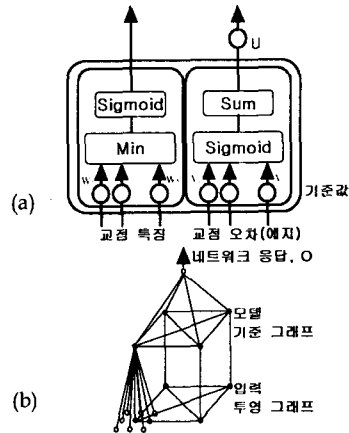


그림 2. 탄력적 정합 모델 (a) 단일 셀 (b) 정합 메카니즘

얼굴은 얼굴 특징점에 의해서 지정된 교점의 그래프로서 표현된다. 모든 교점은 그래프 안에서 각각 연결된다. 그래프 안에서 각 교점은 2개의 신경망 셀(그림 2a) 같이 구성된다. 첫 번째 셀은 얼굴 영상으로부터 Gabor 웨이블릿 변환에 확장된 지역정보를 유지하고, 두 번째 셀은 그래프 안에서 다른 교점의 거리를 유지한다.

정면에서 그래프를 작성하고 얼굴 모델 그래프를 수집하였다. 얼굴 식별대신에, 얼굴 모델 그래프를 얼굴과 얼굴 특징점 검출을 위하여 기준 그래프로 사용할 수 있다. 평균 모델 그래프는 평

균 정보와 모델 그래프에서 교점과 에지에 상응하는 교점 거리를 취하여 그래프를 생성한다. 모든 얼굴 그래프 모델의 원점은 코의 중심으로 표현된 교점으로 정규화 된다.

#### IV. 탄력적 특징 정합

탄력적 특징 정합 모델은 기준값을 이용하여 사용한다. 신호가 기준 값을 통과한다면 출력은 입력과 기준값 사이에 유클리드 거리법으로서 정의된다. 모델 그래프의 교점은 그림 2a 에서와 같은 구조를 가진다. 모델 그래프는 Gabor 특징 벡터와 에지 벡터(다른 교점과의 거리)로 표현된 기준치 벡터  $W$ 와  $V$ 로 구성된다.

정합 메카니즘은 그림 2b에서 나타내었다. 기준 그래프는 얼굴 모델 그래프에서 선택하였다. 투영 그래프는 입력 영상에서 Gabor 웨이블릿 변환 후 기준 그래프를 사상함으로써 생성할 수 있다. 탄력적 정합이 기준 그래프의 교점이 투영 영상의 대응되는 교점의 중심으로부터 9개의 인접한 연결이 이루어진다. 연결 사이에 하나만이 동작을 한다. 기준그래프에서 교점으로 근접한 Gabor 특징을 가진 연결이 동작하도록 지역 경쟁을 실행한다. 이 방법은 새로운 투영그래프의 후보를 생성하고, 변형된 투영 그래프를 만든다. 단일 네트워크의 전체적인 결함을 최소로 하게되면 투영 그래프의 후보는 새로운 투영 그래프가 된다. 전체적인 결함은 네트워크 응답 ( $O$ ) 에 의해 측정될 수 있다

$$O = \sum_{i=1}^{N_{in}} [S(\min_i(\|W_i - I_i\|)) + \sum_{j=1}^{N_{in}} S(\|v_{ij} - e_{ij}\|)] \quad (3)$$

여기서  $S(\cdot)$ 는 시그모이드 함수로 완만한 경사의 변화를 갖는다.  $W_i$ 는 기준 모델의 기준 값이고, 그값은 노드  $i$ 로부터의 Gabor 특징 벡터이다.  $I_i$ 는 노드  $i$ 로부터의 Gabor 특징 벡터 또는 가장 근접한 Gabor 특징벡터  $W_i$ 를 갖는 투영 그래프 인접한것이다.  $v_{i,j}$ 는 기준 그래프상의 노드  $i$ 와  $j$ 사이의 거리 정보이고.  $e_{i,j}$ 는 투영 그래프상의 거리정보이다.  $u_i$ 는 에지 셀에서 배수로 증가하는 가중치이다.  $u_i$ 단순 감소 함수이다.

#### V. 얼굴 검출

탄력적 특징 정합은 얼굴 검출시에 임의의 얼굴 검색에 사용되어진다. 임의의 검색에 대한 연산 효과를 고려하여 우선적으로 탄력적 특징 정합을 사용하였지만 투영그래프의 변형은 이루어지지 않았다. 그래프 상의 공간적 관계의 강성률

은 지역적으로 유사한 복잡한 배경의 입력 영상에서 얼굴 같은 패턴을 찾아내는데 있어 중요한 요소이다.

평균 모델 그래프를 산출함으로써 얼굴 검출 작업이 시작된다. 약 3~4개의 픽셀 간격을 갖는 입력 영상의 gabor 응답은 래스터 스캔을 통하여 적용되어지고, 각각의 지점에서의 네트워크 응답  $O$ 를 연산함으로써 기준그래프에서 가장 작은  $O$ 의 지점을 찾아 낼 수 있다. 가장 작은  $O$ 가 있는 지점을 얼굴 위치로 추정한다. 얼굴 검출 과정은 그림 3a, b와 c에서 나타내었다.

#### VI. 얼굴 특징점 검출

얼굴 검출 과정에서, 얼굴의 추정 위치를 찾을 수 있다. 이 단계에서 얼굴특징점의 위치가 정확한 위치에서 이동되어진다. 이는 탄력적으로 변형된 위치를 찾는 것을 필요로 한다. 변형이 되어진 그래프와 함께 탄력적 특징 정합을 적용한다.

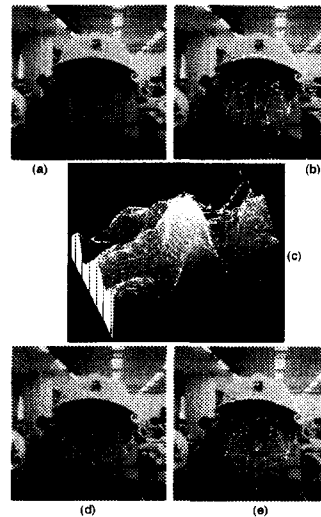


그림 3. 얼굴과 얼굴 특징점 검출

- (a) 입력영상
- (b) 얼굴 영역에서의 그래프 배치로 보여준 얼굴 검출 결과
- (c) 역 네트워크 응답 지도, 낮은  $O$ 는 지도에서 높고 밝게 나타냄
- (d) 탄력적 특징 정합을 처리하는 동안의 교점의 이동
- (e) 얼굴특징점 검출 결과

알고리즘은 다음으로 요약될 수 있다:

1. 입력 공간에서 초기 투영 그래프를 생성하는

얼굴 검출 과정의 결과를 사용한다. 모델 그래프(평균 모델을 포함하는)를 검색하여 네트워크 응답  $O$ 로 투영 그래프와 가장 근접한 모델을 찾는다. 그리고 이 모델 그래프를 기준 모델로 사용한다.

2. 투영 그래프와 기준그래프의 교점에서 활성화된 인접 연결에 의하여 새로운 투영 그래프의 후보를 생성한다. 후보 투영 그래프의 새로운 거리  $e_{i,j}$ 를 연산한다
3. 수식 (3)과 같이 네트워크 응답  $O$ 를 연산한다. 만일 네트워크 응답이 감소되면, 후보 그래프가 새로운 투영 그래프가 된다.
4.  $u_i$ 의 증가하는 가중값을 감소시킨다

2에서 4의 과정은  $O$ 가 최소화 될 때까지 반복된다. 인접된 연결의 간격이 3~4픽셀 정도이고 정합단계가 증가함에 따라 감소된다.

얼굴특징점 검출 과정의 예를 그림 3(b), (d) 그리고 (e)에 나타내었다. 이러한 방식의 흥미 있는 점은 시스템의 로버스트에 있다. 시스템은 모델 얼굴에서 동일한 얼굴이 없을지라도 얼굴특징점을 검출 할 수 있다. 얼굴특징점은 단지 평균 얼굴 모델을 사용함으로써 검출해 낼 수 있다. 시스템은 코 밑 수염과 안경을 착용한 입력 영상에서도 무난하게 검출 할 수 있다.



그림 4. 얼굴 특징점 검출의 예

### Ⅶ. 얼굴 식별

얼굴특징점은 얼굴 영상을 이해하는데 있어 기본적인 요소이다. 얼굴특징점의 위치는 여러가지로 유용하게 사용될 수 있으며, 본 논문에서는 얼굴특징점 검출 결과를 응용하여 얼굴인식을 수행하고자 하였다.

투영그래프에서 입력 영상의 표본을 감안함으로써 얼굴 특징점은 검출되고 투영그래프의 모든 정보는 기본 그래프로 입력된다. 모든 모델은 기본 그래프로써 하나씩 제공받는다. 네트워크 응답  $O$ 는 모든 모델을 위해 계산되어진다. 만일 모델이 최소 네트워크 응답값이고, 그 값이 임계치 낮다면 입력 얼굴을 수용하고, 그 외의 값들은 모두 무시하게 된다. 이러한 방식은 모든 모델에 대하여 탄력적 정합을 할 필요가 없기 때문에 데이터베이스 검색시간을 단축시킬수 있다.

실험을 위하여 얼굴 데이터베이스 300개를 구성하였다. 구성된 데이터베이스는 총 30명으로부

터 5개의 자세에 대해 2개씩의 영상을 얻었다. 실험 결과를 표 1에서 나타내었다.

표 1. 얼굴데이터베이스에 대한 성능

실험 영상	성공	비율
60 개의 정면 영상	58	96.67
120개의 오른쪽에서 왼쪽으로 회전된 영상	110	91.67
120개의 위쪽에서 아래쪽으로 회전된 영상	100	83.33

### Ⅷ. 결 론

본 논문에서는 GWT와 탄력적 특징 정합을 기반으로 한 얼굴식별 시스템을 표현하였다. 제안된 시스템은 복잡한 배경이나 왜곡된 얼굴 영상에서도 자동적으로 얼굴 특징점을 검출 할 수 있었다. 탄력적 특징 정합을 위하여 안경, 모자 등을 착용한 얼굴에서도 정합을 수행 할 수 있는 모델을 제안하였다. 얼굴 특징점 정보의 유용성을 이용함으로써, 효과적인 얼굴 식별을 위한 연산과정을 구할 수가 있었다.

정면영상과 움직임이 있는 영상두가지 형태로 이루어진 실험에서 정면 영상일 경우 대상 60개의 영상중 58개의 영상을 식별하여 96.67%의 성공률을 보여 주었고 움직임이 있는 영상에서는 좌우 움직임의 경우 91.67% 상하움직임의 경우에는 83.33%의 인식율을 보여 주었다.

제안한 시스템은 간단하고, 학습 과정을 필요로 하지 않는다. 새로운 얼굴과 연관된 관련 정보는 다른 학습 과정 없이 모델에 직접 추가할 수 있었다. 얼굴 특징점 검출 시스템의 성능은 모델의 수를 증가시킴으로 보다 안정적으로 동작하였다.

### 참고문헌

- [1] H. Wu, T. Yokoyama, D.Pramadiahanto, and M.Yachida. Face and facial feature extraction from color image. *Proc. of the Int. Worksh. on Autom. Face-and Gesture Recogn.*, 1996.
- [2] J.Daugman. Complete discrete 2-d gabor transform by neural networks for image analysis and compression. *IEEE Trans. on Acoust., Speech, Signal Process.*, 36(7): 1169-1179, 1988.
- [3] J.P.Jones and L.A.Palmer. An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex. *Jour. of Neurophys.*, 58(6):1233-1258,1987.
- [4] L.Wiskott, J.M.Fellous, N.Kruger, and C. der Malsburg. Face recognition and gender

- determination. *Proc. of the Int. Work. on Autom. Face and Gesture Recogn.*, pages 92-97, 1995.
- [5] M.Kass, A.P.Witkin, and D.Terzopoulos. Snakes: Active contour models. *Int. Jour. of Computer Vision*, pages 321-331, 1988.
- [6] M.Lades, J.C.Vorbruggen, J.C. Buhmann, R. C. von der Malsburg, and W.Konen. Distortion invariant object recognition in the dynamic link architecture. *IEEE Trans. on Computers*, 42(6):300-311, 1993.
- [7] R.Brunelli and T.Poggio. Face recognition: Features versus templates. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 15(10):1042-1052, 1993.