

# 복잡배경의 영상에서 NTGST를 이용한 효과적인 얼굴 검출

이재근, 김종화, 서경석, 박은진, 최흥문

경북대학교 대학원 전자공학과

## An Effective Face Detection for the Images with the Complex Backgrounds Using NTGST

Jae-Geun Lee, Jong-Hwa Kim, Kyung-Suk Seo, Eun-Jin Pak Heung-Moon Choi

Department of Electronics, Graduate School, Kyungpook National University

E-mail: [poet@pdp.knu.ac.kr](mailto:poet@pdp.knu.ac.kr)

### 요 약

본 논문에서는 NTGST(noise-tolerant generalized symmetry transform)[1]를 이용하여 복잡배경 영상으로부터 효과적으로 여러 얼굴을 검출할 수 있는 알고리즘을 제안하였다. 먼저 NTGST를 이용하여 얼굴이 존재할 가능성이 있는 관심영역(region of interest: ROI)을 찾고, 각각의 관심영역 내에서 얼굴의 주된 특징인 눈, 코, 입을 부각시킨 Fovea 영상으로부터 대칭변환의 국부 최대치(local maximum)를 구한다음, 이들간의 관계를 기하학적 상관관계로 분석 확인함으로써 사람 얼굴만을 검출 하도록 하였다. 여러 얼굴을 포함하는 복잡한 배경 영상에 대해 제안한 알고리즘을 적용한 결과 89.7%의 검출율을 얻을 수 있었다.

### 1. 서론

최근, 출입통제시스템 등의 분야에서는 지문, 홍채, 망막, 또는 얼굴 등과 같은 개인의 생체특징 정보를 이용하여 개인 식별을 하고 있다. 이중 얼굴인식을 이용하는 방법은 지문이나 홍채인식 등에 비해 사용자의 직접적인 접촉을 요구하지 않으므로 거부감이 적다는 장

점 등으로 인해 현재 많은 연구가 진행 되고 있다. 얼굴 영역의 검출은 얼굴 인식을 위한 전처리 과정으로서 얼굴인식에 선행되어야 할 중요한 부분이다.

기존의 얼굴 검출 방법 중에서 먼저 형태 기반 시스템(shape-based system)은 얼굴의 모양, 표정의 변화가 크기 때문에 복잡한 배경을 가진 영상에서 얼굴 영역의 탐색에 어려움이 있다[2]. 특징 기반 시스템(feature-based system)은 얼굴의 주된 특징인 눈, 코, 입 등의 상관관계를 이용하므로 방향의 변화에 효과적이거나 조명의 변화에서 매우 민감하고, 처리시간이 길다는 단점이 있다[3,4]. 그리고, 패턴기반 시스템(pattern-based system)에서는 신경망(neural network)을 이용하여 비교적 좋은 성능을 나타내지만 영상의 상태에 따라 알고리즘의 일반화가 어렵다[5]. 또한 대칭성을 기반으로 하는 시스템(symmetry-based system)에서는 얼굴의 신축, 표정변화, 회전에 효과적이거나 조명의 변화에 민감하다[6].

제안한 알고리즘은 복잡배경의 영상에 대해 잡음에 강건한 주목 연산자인 NTGST를 이용하여 얼굴의 후보영역을 찾고, 각 후보영역에 대해 얼굴의 주된 특징점인 눈, 코, 입 등을 찾아 이들의 기하학적 위치관계를 비교함으로써 얼굴만을 검출하였다. 그림 1에 제안한 알고리즘의 전체적인 흐름도를 나타내었다.

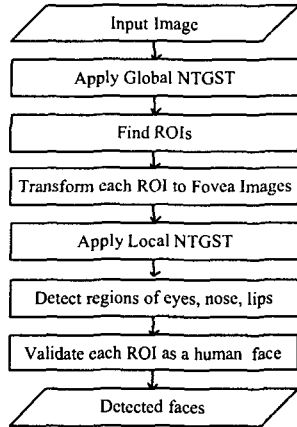


그림 1. 제안한 알고리즘의 흐름도  
Fig. 1. Flowchart of the proposed algorithm

## 2. 얼굴 후보 영역의 검출

인간의 얼굴은 일반 물체등과 같이 나름대로의 대칭성을 가진다. Reisfeld 등[6]은 임의 형태의 물체와 얼굴에 대해 그 대칭성을 부각시킬 수 있는 일반화 대칭변환(generalized symmetry transform: GST)을 주목 연산자로 제안하였다. GST의 경우 대칭 화소쌍의 명도변화의 크기와 방향으로부터 대칭크기를 반영하는 대칭기여도(symmetry contribution)를 정의하고 이를 계속 누적하여 영상 내에서 검출대상에 대한 사전정보 없이도 사람의 얼굴이나 얼굴 내에서의 눈, 코, 입 등의 관심영역을 검출하였다. 그러나, 이 방법에서는 대칭 화소쌍의 대칭 크기만을 고려하므로 물체에 의한 대칭기여도 뿐만 아니라 불규칙한 잡음이나 복잡한 배경 영상에 의한 기여도도 누적되기 때문에 물체만을 검출하기가 어렵다. Park 등[1]은 대칭 화소쌍의 명도변화의 크기와 방향으로부터 대칭크기 뿐만 아니라 극성도 정의하여 대칭기여도에 반영함으로써 잡음에 강건한 NTGST를 제안하였다.

NTGST에서는 두 화소의 명도 변화 방향의 대칭성을 부각시키며 수렴과 발산을 구분 할 수 있도록 위상 가중함수(phase weight function)를

$$P(i, j) = \sin\left(\frac{\theta_j + \theta_i}{2} - \alpha_y\right) \times \sin\left(\frac{\theta_j - \theta_i}{2}\right) \quad (1)$$

와 같이 정의하여 두 화소의 명도 변화 방향이 수렴하면  $P(i, j)$ 는 음극성(-)을 가지고, 발산하면 양극성(+)을

갖도록 하여 두 화소의 명도 변화 방향의 수렴과 발산을 구분하면서도 대칭성을 잘 부각할 수 있도록 하였다. 여기서  $\theta_i$ 와  $\theta_j$ 는 각 화소의 명도 변화 방향이고,  $\alpha_y$ 는 두 화소를 연결하는 직선이 수평선과 이루는 각도이다. 또한, 두 화소간의 거리가 가까울수록 큰 값을 갖도록 가우시안 거리 가중함수(distance weight function)를

$$D_\sigma(i, j) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{|P_i - P_j|}{2\sigma}\right) \quad (2)$$

와 같이 정의하였다. 여기서,  $\sigma$ 는 대칭 변환할 영역의 크기를 나타낸다. 화소의 명도 변화 크기  $r_k$ 와 위상가중함수 및 거리가중함수를 이용하여 두 화소의 대칭 정도를 나타내는 대칭기여도를

$$C(i, j) = D_\sigma(i, j)P(i, j)r_i r_j \quad (3)$$

와 같이 정의하였다.

이와 같이 구성된 NTGST를 복잡한 배경에 얼굴이 존재하는 영상에 적용할 경우 복잡한 배경영역에서는 대칭기여도의 극성을 다르게 하여 상쇄시키고, 얼굴 및 물체에 의한 대칭기여도만을 누적하여 복잡한 배경의 영상에서 얼굴 후보 영역을 효과적으로 검출 할 수 있다.

따라서, 이러한 NTGST로부터 대칭도 맵을 구하여 국부 최대점들을 이용하여 얼굴 후보 영역을 설정하였다. 이 때의 임계되는 그림 2(a)에 따라 마스크 크기별로 적응적으로 가변하였다.

## 3. 얼굴 검출 및 판정

본 논문에서는 얼굴 내에서의 눈, 코, 입 등을 제외한 부분에 대해 누적되는 대칭기여도를 줄이기 위해 Fovea 영상을 이용한다. Fovea 영상은 사람이 한 곳을 집중하게 되면 집중되는 부분은 뚜렷하게 보이지만 그 외의 영역은 흐려지는 현상을 이용한 것이다. 즉, 관심영역에서 얼굴 내에서는 원래의 이미지를 그대로 유지하고, 이를 벗어나는 영역의 화소값을 감쇄시켜 지워버렸다. 관심영역으로부터 구한 Fovea 영상에 국부적 NTGST를 적용하여 얼굴의 특징인 눈을 먼저 찾았다. 여기서 임계값은 그림 2(b)에서 보듯이 마스크 크기별로 적응적으로 가변시켰다.

눈은 얼굴의 특징 중에서 가장 명도의 변화가 크고 대칭성이 강하여 누적 대칭도가 가장 높게 나타나기 때

문에 쉽게 검출할 수 있다. 검출한 두 눈의 중심을 지나는 대칭축에서 코, 입의 영역을 추정하고 영역 안에 나타나는 국부적 최고점을 찾는다. 얼굴 영역에서는, 대칭도 맵의 국부 최대점에서 두 눈의 대칭도가 다른 국부 최고점의 2배 이상의 대칭도를 가짐을 확인하였다. 두 눈과 코, 입의 기하학적 위치 관계는 그림 3 과 같으며 여기서  $P$ 는 특징점들의 좌표를 나타내고,  $D$ 는 두 눈사이의 거리를 나타내고,  $\theta$ 는 두 눈을 연결하는 직선라인과 수평축과의 각이다.

$$P_{nose}(x, y) = \left( \frac{P_{eye\_right}(x) + P_{eye\_left}(x)}{2}, \frac{P_{eye\_right}(y) + P_{eye\_left}(y)}{2} - \beta_1 \right) \quad (4)$$

where,  $\alpha_1 = \frac{7}{12} D_{eye} \cos \theta, \quad \beta_1 = \frac{7}{12} D_{eye} \sin \theta$

$$P_{lip}(x, y) = \left( \frac{P_{eye\_right}(x) + P_{eye\_left}(x)}{2}, \frac{P_{eye\_right}(y) + P_{eye\_left}(y)}{2} - \beta_2 \right) \quad (5)$$

where,  $\alpha_2 = \frac{13}{12} D_{eye} \cos \theta, \quad \beta_2 = \frac{13}{12} D_{eye} \sin \theta$

이와 같이 정의된 위치에 국부 최대치가 존재함을 확인함으로써 얼굴을 판정하였다.

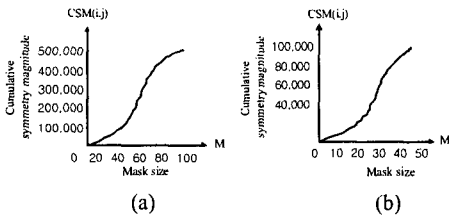


그림 2. (a) 얼굴 전체 영역에서 얼굴에 대한 마스크 크기별 평균 누적대칭도 (b) 얼굴 특징점 영역에서 특징점에 대한 마스크 크기별 평균 누적 대칭도

Fig. 2. Average cumulative symmetry map according to the mask size in (a) a face region and (b) a face feature region.

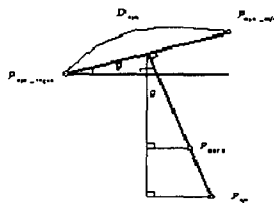


그림 3. 두 눈과 코, 입의 기하학적 위치관계  
Fig. 3. Geometrical relation among eyes, a nose, and a lip.

#### 4. 실험결과

본 논문에서는 CCD 카메라로 획득한 복잡배경의 정면 얼굴 영상에 신축, 회전 및 조명의 변화가 있는 영상 70장을 실험 데이터베이스 A, CMU 데이터베이스, 181장의 영상을 실험 데이터베이스 B로 두고 실험을 수행하였다. 그림 4에 단계적 얼굴검출 알고리즘의 결과를 나타내었고, 그림 5는 다양한 영상에 대한 얼굴 실험 결과이다.

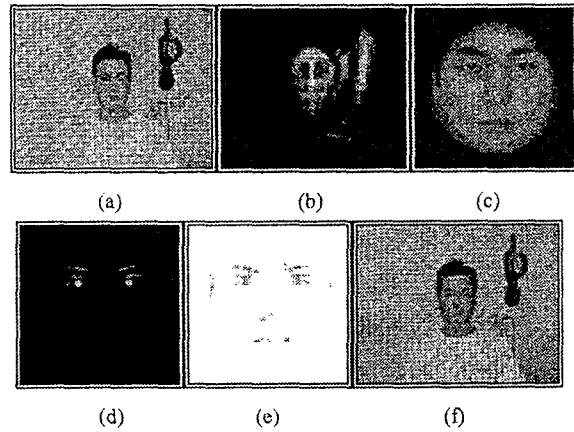


그림 4. (a) NTGST를 이용한 얼굴 검출 입력 영상, (b) NTGST를 적용한 누적 대칭도 맵, (c) Fovea 영상, (d) (c)의 영상에 NTGST를 적용한 경우의 누적 대칭도 맵, (e) (d)의 이진화 영상, (f) 제안한 방법으로 얼굴을 검출한 결과

Fig. 4. Face detection using NTGST (a) Input image, (b) Cumulative symmetry map of the NTGST, (c) Fovea image, (d) Cumulative symmetry map of the NTGST in the image (c), (e) The binarized image of (d), (f) The result of face detection using the proposed method.

표 1 실험 set A, B에 대한 얼굴 검출 결과

Table 1. The results of the face detection in the experiment set A and B.

	No. of images	No. of faces	No. of Detected faces	No. of False detection	Face detection rates
set A	60	131	119	12	90.8%
set B	120	527	467	60	88.6%

표 1에 실험 set A, B에 대한 실험결과를 나타내었다.

실험 set A의 경우 60개 영상의 131개의 얼굴로부터 90.8%, 실험 set B의 경우 120개 영상의 527개 얼굴로부터 88.6%의 검출율을 보여 전체적으로 89.7%의 검출율을 나타내었다. 그림 5에서 (a)~(f),(h)는 모두 검출한 결과를 나타내고 있고, (g)는 한 얼굴에서 특징점이 일부 가려져 검출하지 못하였다.

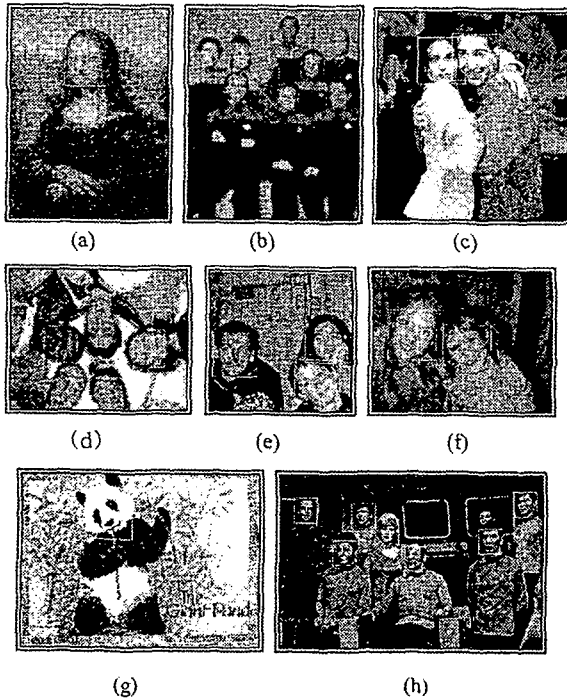


그림 5. 다양한 영상에서 얼굴을 검출한 결과  
Fig. 5. The results of the face detection for various images.

## 5. 결론 및 추후 연구

본 논문에서는 NTGST를 이용하여 얼굴의 후보영역을 찾고, 다시 NTGST를 적용하여 얼굴의 특징점들 중 눈, 코, 입 등의 기하학적 위치관계를 비교함으로써 얼굴을 검출 확인하였다. 복잡배경의 영상에서 회전 및 다양한 크기의 650개 얼굴이 포함된 180개 영상에 대한 실험을 통해 제안한 방법을 이용하여 효과적으로 얼굴을 검출할 수 있음을 확인하였다. 제안한 방법은 얼굴의 특징점들을 모두 포함하는 영상에 제한을 두었다. 따라서, 추후로 얼굴의 특징점이 일부 가려진 영상으로부터의 능률적인 얼굴 검출 방법에 대하여 연구를 계속 진행하고자 한다.

## 참고 문헌

- [1] 박창준, "물체 위치검출을 위한 잡음에 강건한 주목 연산자," *경북대학교 박사학위 논문*, 2000.
- [2] Lixin Fan and Kah Kay Sung "Face detection and pose alignment using colour, shape and texture information," *Proceedings of Third IEEE International Workshop on Visual Surveillance*, pp. 19-25, 2000.
- [3] R. Brunelli and T. Poggio, "Face recognition: features versus templates," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, Issue 10, pp. 1042-1052, Oct. 1993.
- [4] V. Vogelhuber and C. Schmid, "Face detection based on generic local descriptors and spatial constraints," *Proceedings of 15th International Conference on Pattern Recognition*, vol. 1, pp. 1084-1087, 2000.
- [5] S. Lawrence, C. L. Giles, Ah Chung Tsoi, and A. D. Back, "Face recognition: a convolutional neural-network approach," *IEEE Transactions on Neural Networks*, vol. 8, Issue 1, pp. 98-113, Jan. 1997.
- [6] D. Reisfeld, H. Wolfson, and Y. Yeshurun, "Context-free attentional operators: The generalized symmetry transform," *IJCV*, vol. 14, pp. 119-130, 1995.