

논문 2004-41SP-6-14

복잡한 배경에서 계층적 주목 연산자를 이용한 다중 얼굴 검출

(Multi-face Detection from Complex Background Using Hierarchical Attention Operators)

이 재근*, 김복만**, 서경석**, 최흥문**

(Jae-Geun Lee, Bok-Man Kim, Kyung-Seok Seo, and Heung-Moon Choi)

요 약

본 논문에서는 단일 문맥자유 주목 연산자 (context-free attention operator)만을 계층적으로 적용하여 복잡배경의 인물 영상으로부터 여러 얼굴을 효과적으로 동시 검출하는 알고리즘을 제안하였다. 입력 영상을 피라미드 구조로 변환하고, 잡음에 강건한 주목 연산자를 전역적으로 적용하여 먼저 고속으로 얼굴이 존재할 후보영역들을 찾고, 다시 이 영역들에 국한하여 지역적인 주목 연산자를 적용하여 얼굴이 갖는 특징을 확인함으로써 얼굴임을 검증하였다. 복잡 배경 속에 여러 얼굴을 포함하는 여러 가지 인물 영상 데이터베이스에 대해 제안한 알고리즘을 적용한 결과 93.5%의 검출율을 얻을 수 있었다.

Abstract

An efficient multi face detection technique is proposed based on hierarchical context-free attention operators in which multiple faces are efficiently detected from a noisy and complex background. A noise-tolerant generalized symmetry transform (NTSGT) is applied hierarchically, as a context free attention operator, to the input pyramidal image for the high speed global location of the regions of face candidates (ROFCs) with a single mask. For the face verification, local NTGST is applied within each ROFC to confirm the existence of the detailed facial features. First, by globally applying NTGST which introduces the average pyramid method and focusing to the input image with complex background, ROFCs with recognizable resolution are detected robustly. Morphological operations are applied only to the each detected ROFCs to emphasize the facial features like eyes and lips. Then, eyes are detected by locally applying NTGST to the ROFCs and only faces are detected by verifying the existence of the geometrical features of the faces relatively to the location of eyes. The experimental results show that the proposed method can efficiently detect multiple faces from a noisy or complex background with 93.5% detection rate.

Keywords : 얼굴검출, 주목 연산자, 복잡한 배경

I. 서 론

여러 가지 생체인식 기법 중에서 얼굴의 검출과 인식은 사용시의 편의성, 친밀성 및 정확성을 보장하는 인식 기법으로서 컴퓨터와 사용자 간의 보다 현실감 있는 인터페이스, 각종 신분증에서의 사진 인식, 통제구역 출입제어, 정보 보호를 위한 접근제어 등 매우 광범위한

분야에 걸쳐 응용되고 있다. 정확한 얼굴 검출은 얼굴 인식 시스템의 전처리 과정으로 필수적이며 그 효율적인 운용을 위해 중요하다. 그러나, 20여 년간 이 분야의 많은 연구가 진행되어 왔음에도 불구하고 얼굴은 그 표정이나 방향 및 자세의 변화 뿐 아니라 조명이나 카메라 각도에 따른 이미지의 상태 및 주위의 복잡 배경 때문에 얼굴만의 검출과 인식은 그리 쉽지가 않다^[1].

기존의 얼굴 검출 방법에는 지식 기반 (knowledge-based)법, 템플릿 정합 (template matching)법, 모습 기반 (appearance-based)법 및 특징 불변적 접근 (feature invariant approach)법 등이 있다. Yang 등^[2]은 인간의 전형적인 얼굴을 구성하는 사전 지식에 기반을

* 정회원, 주식회사 디보스
(Diboss Co., LTD.)

** 정회원, 경북대학교 전자전기컴퓨터학부
(School of Electrical Engineering and Computer Science, Kyungpook National University)

접수일자: 2004년8월26일, 수정완료일: 2004년10월29일

두어 얼굴이 주위 배경보다 평균적으로 밝다는 특징을 가지고 ROFC (region of face candidate)를 찾고, 수직 수평으로 명도값을 투영하여 눈과 입, 코 등을 판별하여 얼굴을 검출하였으나, 이 방법에서는 얼굴에 대한 정보를 바탕으로 정확한 판별을 위한 잘 정의된 규칙을 만들기 어렵다. Sakai 등^[3]은 얼굴의 특징들을 나타낼 수 있는 몇 개의 특정 패턴만을 저장하여 놓고 입력 영상과 비교하는 템플릿 정합법을 제안하였으나, 얼굴의 크기, 자세 및 외모의 변화에 대해 효과적으로 대응할 수 없다는 단점이 있다. 한편 모습 기반법으로 Mener 등^[4]은 PCA (principal component analysis) 학습에 의해 얼굴 영상들만의 통계적 특성을 분석하여 기저 벡터를 만들어 비교함으로써 얼굴을 검출하였으나, 이는 얼굴인식단계에서 사용되는 PCA기법을 그대로 얼굴 검출에 사용하므로 기저 벡터들이 다양한 얼굴들의 공통적인 특징을 잘 반영하지 못하는 단점이 있다. 특징 불변적 접근법으로 Reisfeld 등^[7]은 얼굴의 특징인 눈, 코 및 입 등이 대칭성이 강하다는 점을 이용하여 화소들의 대칭기여도 (symmetry contribution)를 계산하여 얼굴의 눈, 코 및 입에서 큰 대칭도를 누적하는 일반화대칭 변환 (generalized symmetry transform: GST)을 이용하여 얼굴을 검출하였으나, 이 방법에서는 객체뿐만 아니라 객체 주위의 불규칙한 잡음이나 복잡배경에 의한 대칭 기여도도 같이 누적되기 때문에 복잡 배경의 실영상에는 적용하기가 어렵다. 이러한 방법들을 종합해 볼 때, 실영상에서와 같이 복잡배경으로부터도 얼굴의 특징을 잘 반영하고 얼굴 크기에 관계없이 정확하게 여러 얼굴들을 판별해 내기 위한 잘 정의된 규칙의 알고리즘에 관한 연구가 필요하다.

본 논문에서는 잡음에 강건한 일반화 대칭변환 NTGST (noise tolerant generalized symmetry transform)^[8]를 문맥 자유 주목 연산자 (context-free attention operator)로 사용하여 단일 알고리즘을 계층적으로 적용함으로써 얼굴이 갖는 전역적 및 지역적 특징을 찾아 복잡배경의 영상으로부터 효과적으로 여러 얼굴을 검출하는 알고리즘을 제안하였다. 먼저 입력된 복잡 배경 영상에 평균 피라미드 기법과 포커스(focus)를 도입하여 NTGST를 전역적으로 적용함으로써 얼굴후보 영역 ROFC를 더욱 강건하게 검출하였다. 검출된 ROFC에 국한하여 지역적으로 NTGST를 적용하여 두 눈을 찾고 이를 기준으로 얼굴의 기하학적 특징들을 검증함으로써 얼굴만을 검출하였다. CMU 얼굴 데이터베이스^[1]를 비롯하여 총 160개의 얼굴들이 포함된 복잡 배경

인물 영상 데이터베이스에 대해 실험한 결과, 93.5%의 검출율을 얻어 기존의 GST에 기반한 방법에 비해 효과적이고 강건한 방법임을 확인하였다.

본 논문의 II장에서는 제안한 방법의 전반적인 알고리즘에 대해서 서술하고 III장에서는 실험과 그 고찰을 나타내고 IV장에서 결론을 맺는다.

II. 본 론

본 논문에서는 그림 1과 같이 입력 영상에 대한 피라미드 영상에 잡음에 강건한 주목연산자 NTGST를 계층적으로 적용하여 먼저 전역적으로 ROFC 만을 검출한 다음, 각 ROFC 내에 지역적으로 NTGST를 적용하여 두 눈을 지나는 수평선을 찾고, 이를 기준으로 눈과 입으로 구성되는 일반적인 얼굴 특징을 확인하여 얼굴들을 검출하였다.

1. 얼굴 후보영역의 검출

영상으로부터 크기가 다양한 미지의 여러 얼굴들을 효과적으로 추출 해내기 위해서는 먼저 단일 마스크를 이용하여 각 얼굴에 대한 구체적인 사전 정보 없이도 얼굴의 추출이 가능해야 한다.

본 논문에서는 먼저 입력 영상을 피라미드 구조로 변환하고 여기에 NTGST를 계층적으로 적용함으로써 얼굴 크기에 관한 사전 정보 없이 단일 마스크로도 여러 얼굴들의 위치를 효과적으로 검출할 수 있도록 하였다.

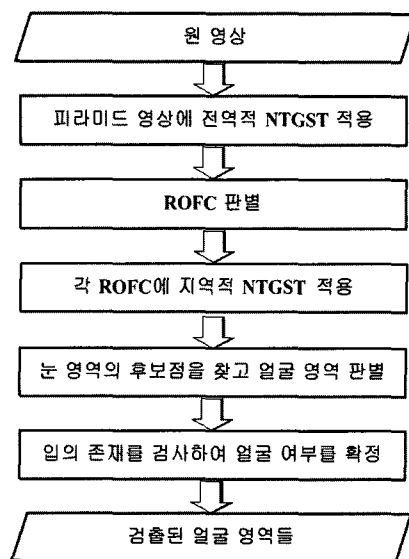


그림 1. 제안한 얼굴 검출 알고리즘의 흐름도
Fig. 1. Flow chart of the proposed face detection algorithm.

평균 피라미드 (mean-pyramid) 기법을 사용하여 하위 층의 4개의 화소들로부터 상위층의 명도 $p_{i,j}^k$ 를

$$p_{i,j}^k = \frac{1}{4} \sum_{x=0}^1 \sum_{y=0}^1 p_{(2i+x), (2j+y)}^{k+1} \quad k = 0, \dots, L-1 \quad (1)$$

로 평균하여 상위 층의 1개의 화소로 사상시켰다. 식 (1)에서 L 은 피라미드 층의 개수를 의미한다. 따라서, NTGST를 이용하여 피라미드 하위 층에서는 각 물체의 국부적 대칭성을 반영하고 상위층으로 갈수록 전역적인 대칭성을 반영한 대칭 크기 맵을 고속으로 얻을 수 있다.

NTGST에서는 대칭도 맵을 얻기 위해서 먼저 각 화소 p 에서의 명도 변화 크기 r 와 이의 방향 θ 를 구한 후 위상 가중 함수 (phase weight function)를

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

와 같이 정의하였고, 여기서 α_{ij} 는 두 화소와의 연결하는 직선이 수평선과 이루는 각도이다. 두 화소의 거리가 가까울수록 큰 값을 갖도록 거리 가중 함수 (distance weight function)를

$$D_\sigma(i, j) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{p_i - p_j}{2\sigma}\right) \quad (3)$$

와 같이 정의하였으며 여기서 σ 는 대칭 변환할 영역의 크기를 나타낸다. 두 화소의 대칭 정도를 나타내는 대칭 기여도를

$$C(i, j) = D_\sigma(i, j)P(i, j)r_i r_j \quad (4)$$

같이 정의하여 두 화소의 거리가 가까울수록, 명도 변화 방향이 대칭일수록, 명도 변화 크기가 클수록 큰 대칭 기여도를 나타내도록 하여, 이 대칭 기여도를 대칭 화소쌍의 중심 위치에 누적하여 대칭도 맵을 얻었다.

각 피라미드 영상으로부터 NTGST를 계층적으로 적용하여 누적된 대칭도 맵에서 얼굴 등 객체가 존재할 가능성을 가진 국부 영역을 찾아 ROFC로 선정한다. 그림 2(a)의 원영상에서 각 축소된 피라미드 영상을 만들들로부터 ROFC들을 찾은 후, 음극성을 갖는 ROFC들과 각각에 NTGST를 적용하여 누적 대칭도 맵을 구한 다음, 누적된 대칭도 맵에서 국부 최고점 (local peak) 만을 얼굴 후보로 선정하였다. 일반적으로 얼굴은 머리

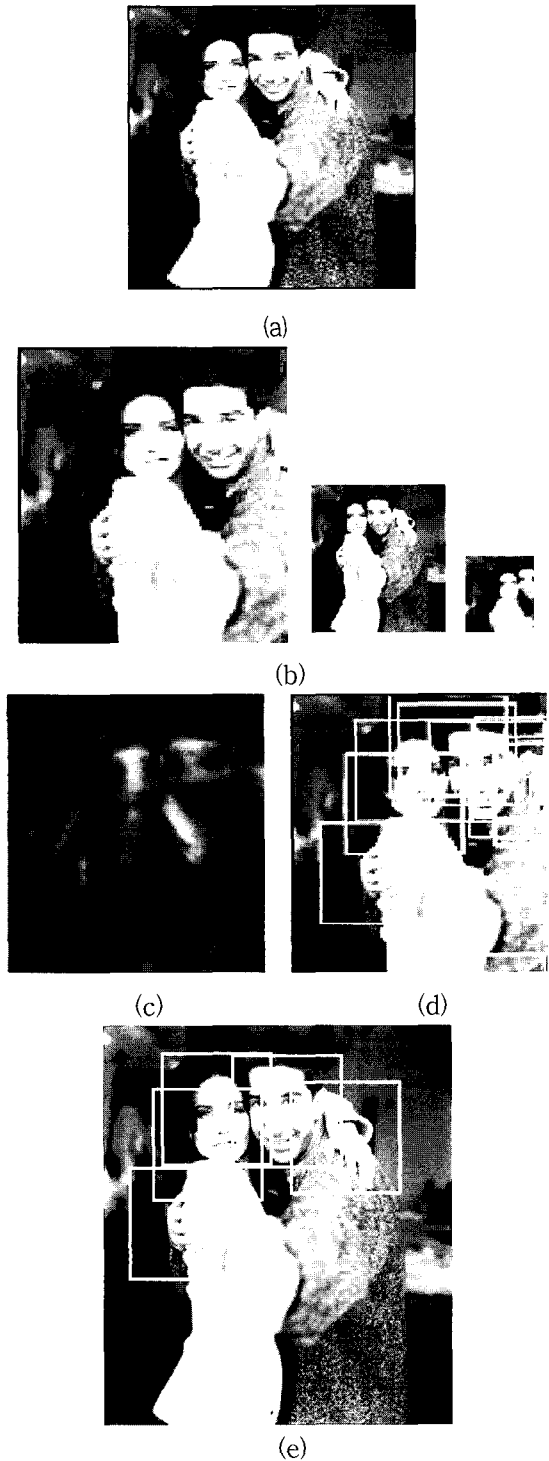


그림 2. ROFC들의 결정: (a) 원영상, (b) 피라미드 영상, (c) 상위층에서의 NTGST 대칭도 맵(16배 확대), (d) 검출된 ROFC 및 (e) 대칭 극성을 이용한 ROFC 수의 감축

Fig. 2. Detection of ROFC's: (a) Original image, (b) pyramidal images, and (c) symmetry map of the NTGST at the upper layer(16 times enlarged), (d) the detected ROFCs, and (e) reduction of the number of ROFCs using the symmetry polarity.

와 조명에 의한 그늘 및 그림자로 인해 그 주위에 비해 명도가 높아 얼굴을 중심으로 한 두 화소의 명도변화가 수렴하는 방향이어서 얼굴 영역에서 NTGST의 누적 대칭도는 음극성을 가지므로 양극성의 누적 대칭도를 갖는 국부최고점들을 제외함으로써 처리대상 ROFC의 개수를 축소하였다.

2. 얼굴 검증

전술한 바와 같이 피라미드 영상에 전역적으로 NTGST를 적용하여 대칭도가 높은 영역들을 얼굴 후보 영역들로 검출해 내고, 지역적 NTGST를 적용하여 국부 최대점을 중심으로 두 눈과 입 등 얼굴이 갖는 기하학적 특징들을 확인하여 얼굴들을 검증하였다.

먼저 그림 3(a)와 같은 선택된 ROFC에 대해 형태수리학적 부식(erosion) 연산자를 적용하여 얼굴 영역에서 명도가 낮은 눈, 코, 입 등의 전역적인 얼굴 특징들을 먼저 부각시켰다. 그런 다음 대비를 강조한 후 NTGST 대칭도 맵으로부터 얼굴에서 가장 명확한 특징을 갖는 두 눈의 위치를 추출하였다. 눈은 얼굴의 특징들 중에서 가장 명도의 변화가 크고 대칭성이 강하기 때문에 ROFC내에 지역적 NTGST를 적용하면 안정적으로 그 위치를 검출할 수 있다. 이때 두 눈은 주위의 피부색에 비해 그 명도가 낮아 NTGST 적용시 누적대칭기여도는 항상 양(+)으로 강하게 누적되므로 양의 극성을 갖는 국부최대치만을 조사함으로써 효과적으로 얼굴 요소를 검출할 수 있을 뿐만 아니라, 근접한 두 물체 사이에 나타나는 거짓물체(ghost object)¹⁸⁾의 영향도 줄일 수 있다. 결정된 두 눈의 중심을 지나는 대칭축에서 그림 3(f)와 같이 입의 중심과 입술영역을 찾아 확인함으로써 사람 얼굴을 최종 확인하였다.

III. 실험결과 및 고찰

제한한 알고리즘을 잡음 및 복잡배경 내에 여러 얼굴이 있는 영상데이터 베이스에 대해서 실험하였다. 본 병렬분산처리 연구실(PDP Lab)에서 CCD 카메라로 직접 획득한 복잡 배경에 여러 정면 얼굴이 존재하는 영상 30장을 PDP 데이터베이스로 하고, 미국 카네기 멜론 대학의 CMU 데이터베이스¹¹⁾에서 여러 정면 얼굴영상을 포함하는 영상 50장을 포함하여 총 80장의 영상에 포함된 248개의 얼굴을 대상으로 제안한 알고리즘을 실험하였다.

얼굴 검출 결과들을 일부 예시한 그림 4에서 보는바와 같이 대부분의 얼굴들을 효과적으로 검출함을 알 수 있

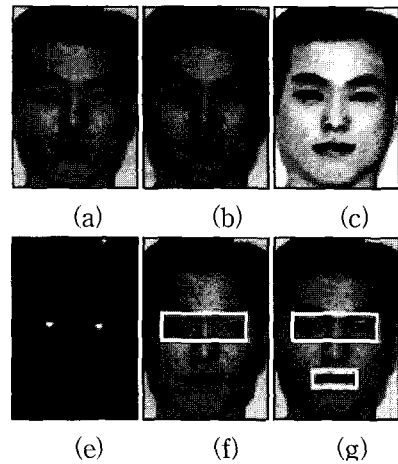


그림 3. 얼굴 검증: (a) ROFC, (b) 형태수리학적 부식 결과, (c) 대비 향상, (d) 누적 대칭도맵, (e) 두 눈의 검출 및 (f) 입의 검출

Fig. 3. Face verification: (a) ROFC, (b) morphological erosion, (c) contrast enhancement, (d) the symmetry map of the NTGST, (e) the detected eyes, and (f) a detected lip.



그림 4. 얼굴 검출 예
Fig. 4. Face detection examples.

다. 다만, 끝에 있는 영상에서 흑인 얼굴과 같이 주위 배경보다 얼굴이 더 어두우면 NTGST에서 대칭기여도가 양극성을 나타내므로, ROFC의 개수를 줄이기 위해 음극

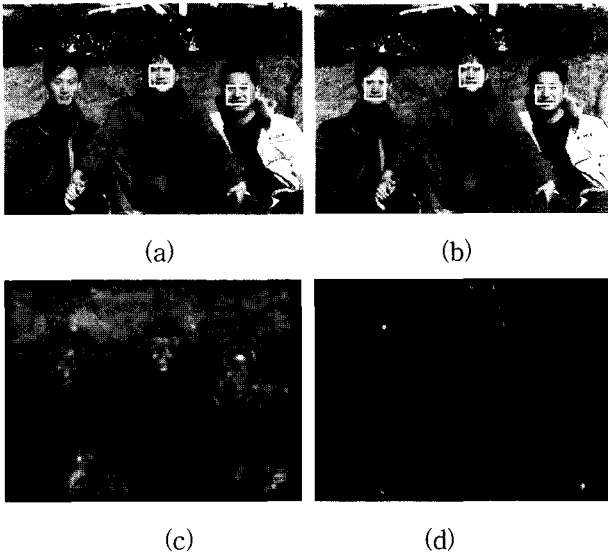


그림 5. 얼굴 검출 비교: (a) GST 적용시, (b) NTGST 적용시, (c) (a)의 대칭도 맵 및 (d) (b)의 누적 대칭도 맵
 Fig. 5. Comparison on face detection: (a) GST-based method, (b) NTGST-based method, (c) symmetry map of (a), and (d) symmetry map of (b).

표 1. PDP 및 CMU 얼굴 데이터베이스에 대한 얼굴 검출 결과
 Table 1. The results of face detection for PDP and CMU face databases.

	영상 수	얼굴 수	검출된 얼굴 수	검출율
PDP	30	95	91	95.7%
CMU	50	153	141	92.2%
Total	80	248	232	93.5%

표 2. 복잡 배경에서의 얼굴 검출에 있어서 GST와 제안한 NTGST와의 성능 비교
 Table 2. Comparison of the GST-based and the proposed NTGST-based face detections from complex backgrounded images.

	영상 수	얼굴 수	검출된 얼굴 수	검출율
GST-based	80	248	226	91.1%
Proposed	80	248	232	93.5%

성을 선택하는 현재의 알고리즘으로는 검출에 실패할 수 있으며, 이를 보완하기 위한 알고리즘은 현재 강구 중이다. 한편, 그림 5에서 보는 바와 같이 GST 기반의 얼굴 검출법은 잔디 및 수풀이 존재하는 배경의 영상에서는 얼굴부분보다 배경의 대칭 기여도가 크게 누적되어 그림 5(a)에서처럼 배경과 대비가 강하지 못한 얼굴을 검출하지 못하는 경우도 있지만 이러한 복잡배경으로부터도 NTGST 기반의 얼굴 검출은 좀 더 강건하여 얼굴만을 검출함을 알 수 있다.

표 1은 제안한 알고리즘의 얼굴 검출 능력을 실험한

결과를 정리한 것이다. 30장의 PDP 데이터베이스와 50장의 CMU 데이터베이스를 합하여 80개 영상이 포함되어 있는 총 248개 얼굴에 대해 제안한 알고리즘을 적용한 결과, 복잡배경 내의 여러 정면얼굴들에 대해 93.5%의 검출율을 확인하였다. 또한, 얼굴이 아닌데 얼굴로 검출한 오검출율은 0.1% 이내이다. 표 2는 동일한 인물 영상 데이터베이스에 대해 제안한 NTGST기반 알고리즘을 GST기반 알고리즘과 비교 실험한 결과이다. 배경이 복잡한 실영상에서의 얼굴 검출에는 본 연구에서 제안한 NTGST기반 얼굴 검출이 더 효과적임을 알 수 있다.

IV. 결 론

본 논문에서는 단일 주목연산자로서 NTGST만을 계층적으로 적용하여 복잡배경의 인물 영상으로부터 여러 얼굴을 효과적으로 검출하는 알고리즘을 제안하였다.

피라미드 구조의 입력영상에 전역적으로 NTGST를 적용하여 먼저 잡음에 강건하면서도 고속으로 ROFC를 찾고, 다시 이 영역들에 국한하여 지역적인 NTGST를 적용함으로써 얼굴이 갖는 일반적인 특징을 확인하여 사람 얼굴들만을 검출하였다. 여러 얼굴을 포함하는 복잡한 배경 영상에 대해 제안한 알고리즘을 적용한 결과 93.5%의 검출율을 얻을 수 있었다. NTGST 한가지만을 사용하여 얼굴을 검출하도록 함으로써 앞으로 NTGST의 성능 개선에 따라 얼굴 검출 성능의 향상을 기대할 수 있으며, 현재 고속화 및 컬러 영상으로의 확장 등을 통해 배경보다 어두운 얼굴의 검출 등 검출율 개선에 관해 폭넓게 연구 중이다.

참 고 문 헌

- [1] M. H. Yang, D. J. Kriegman, and N. Ahuja, "Detection Faces in Images: A Survey," IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 24, no. 1, Jan, 2002.
- [2] G. Yang and T. S. Huang, "Human Face Detection in Complex Background," Pattern Recognition, vol. 27, no. 1, pp. 53-63, 1996.
- [3] T. Sakai, M. Nagao, and S. Fujibayashi, "Line Extraction and Pattern Detection in a Photograph," Pattern Recognition, vol. 1, pp. 233-248, 1969.
- [4] B. Mener and F. Muller, "Face Detection in Color Images Using Principal Components Analysis," Image Processing and its Applications, IEE 1999.

- [5] S. A. Sirohey, "Human Face Segmentation and Identification," Technical Report CS-TR-3176, Univ. of Maryland, 1993.
- [6] T. K. Leung, M. C. Burl, and P. Perona, "Finding Faces in Cluttered Scenes Using Random Labeled Graph Matching," Proc. Fifth IEEE Int'l Conf. Computer Vision, pp. 637-644, Cambridge, MA, Jun 1995.
- [7] D. Reissfeld, H. Wolfson, and Y. Yeshurun, "Context-free attentional operators: The generalized symmetry transform," IJCV, vol. 14, pp. 119-130, Jan. 1995.
- [8] C. J. Park, S. H. Cho, and H. M. Choi, "An Implementation of Noise-Tolerant Context-free Attention Operator and its Application to Efficient Multi-Object Detection," IEEE Transaction on Signal Processing, vol. 38SP, no. 1, pp. 89-96, Jan. 2001.

 저 자 소 개



이 재 근(정회원)
 2000년 경일대학교 전자공학과
 학사 졸업.
 2002년 경북대학교 전자공학과
 석사 졸업.
 <주관심분야: 영상처리, 컴퓨터
 비전, 신호처리>

서 경 석(정회원)
 제38권 SP편 3호 참조



김 복 만(정회원)
 1999년 경일대학교 전자공학과
 학사 졸업.
 2002년 경북대학교 전자공학과
 석사 졸업.
 <주관심분야: 영상처리, 병렬분산
 처리, 임베디드 프로세서>

최 흥 문(중신회원)
 제38권 SP편 3호 참조