

예측맵을 이용한 얼굴탐색의 가속화기법

조경식*, 구자영**

An Acceleration Method of Face Detection using Forecast Map

Kyoung-sik Cho*, Ja-young Koo**

요 약

본 논문은 주성분 분석에 의한 특징 탐색 기법의 가속화 방법을 제안하고 있다. 특징 탐색이란 주어진 영상이 찾고자 하는 특징을 포함하고 있는지의 여부를 판단하고, 만일 그 특징이 포함되어 있다면 그 위치와 영역을 결정하는 방법이다. 탐색 대상으로 하는 얼굴 또는 특징의 위치와 스케일을 미리 알 수 없으므로 모든 위치에서 다양한 스케일의 특징에 대한 존재 가능성을 계산해야하는데 이것은 방대한 공간에서의 탐색문제이다. 본 논문에서는 다단계 예측맵과 윤곽선 예측맵을 이용함으로써 탐색공간을 축소하고 빠른 얼굴 및 특징 탐색을 가능케하는 방법을 제안하고 있다. 실험결과, 제안된 방법이 기존의 전역탐색방법에 비하여 계산량을 10%이하로 줄일 수 있었다.

Abstract

This paper proposes an acceleration method of PCA(Principal Component Analysis) based feature detection. The feature detection method makes decision whether the target feature is included in a given image, and if included, calculates the position and extent of the target feature. The position and scale of the target feature or face is not known previously, all the possible locations should be tested for various scales to detect the target. This is a search problem in huge search space. This paper proposes a fast face and feature detection method by reducing the search space using the multi-stage prediction map and contour prediction map. A proposed method compared to the existing whole search way, and it was able to reduce a computational complexity below 10% by experiment.

*, ** 단국대학교 정보·컴퓨터학부
본 연구는 2001년 단국대학교 교내 연구비에 의해서 연구되었음

I. 서론

본 논문에서 다루고 있는 얼굴탐색 문제는 주어진 영상이 사람의 얼굴을 포함하는지의 여부를 판단하고, 만일 얼굴이 포함되어 있다면 그 위치와 영역을 결정하는 방법을 다룬다. 얼굴탐색은 자동적인 얼굴인식 시스템을 위한 필수적인 전반기 처리과정이며, 또한 인간-컴퓨터 상호작용의 다양한 응용이나 무인감시 시스템을 위한 필수 요소로 사용될 수 있다.

얼굴탐색을 어렵게 하는 것은 대부분의 얼굴들에서 눈, 코 입 등의 특징들이 유사한 구조를 가지고 있지만 사람마다 질감과 크기 및 모양이 같지 않다는 점이다. 이러한 각 사례에 연관된 변형들을 흡수하여 인식률을 높이기 위한 정합 방법들로 correlation template를 이용하는 방법 [1][2], deformable template를 이용하는 방법[3], image invariant를 이용하는 방법[4] 등이 주로 사용되어 왔다.

본 논문에서 사용하는 주성분 분석 방법(PCA : Principal Component Analysis)은 correlation template와 유사한 방법으로서, 다차원 자료 집단의 주성분을 분석하여 낮은 차원의 벡터로 표현하여 분류하는 방법이다. 특히 영상과 같이 다차원의 정보로부터 특정 대상을 탐색하고자 할 때 그 성능이 뛰어나 얼굴인식, 얼굴탐색과 같은 특징 탐색분야에서 널리 사용되고 있다[2][5].

탐색 대상으로 하는 얼굴 또는 특징의 위치와 스케일을 미리 알 수 없으므로, 주성분 분석에 기반한 탐색방식에서는 주어진 이미지의 모든 위치에서 다양한 스케일의 부분 영상을 취하여 각 부분 영상이 나타내는 패턴 벡터가 주성분 부분 공간상에서 탐색 대상 클래스에 속할 가능성을 계산하여 가장 큰 값을 산출하는 위치와 스케일을 얼굴 또는 특징의 영역으로 결정한다[6]. 그러나 기존의 방법에서는 주어진 영상에서 여러 스케일에 대해서 반복적으로 부분영상을 추출하고 유사도를 계산해야 하기 때문에 과도한 계산을 유발하고 이것은 빠른 응답을 요구하는 시스템에서 실제적인 활용을 어렵게 한다.

본 연구에서는 얼굴을 탐색하기 위해서 다양한 스케일

의 눈을 탐색하고 찾아진 눈의 위치에 의해서 얼굴 전체의 스케일을 결정하는 방법을 사용한다. 이 때 앞에 언급한 방대한 계산을 줄이기 위해서 눈이 있다고 보기 어려운 영역을 윤곽선 맵으로부터 판단하여 탐색공간을 줄임으로써 효율적으로 얼굴탐색을 하는 새로운 방법을 제안하고 있다.

II. Maximum Likelihood를 이용한 특징 탐색

벡터로 표현된 탐색의 대상이 되는 특징의 샘플들의 집합을 $\{ \mathbf{x} \}$, 그 평균과 공분산을 각각 \mathbf{m} 과 \mathbf{C}_x 라고 할 때 주어진 입력 벡터 \mathbf{x} 가 주어진 클래스 Ω 에 속할 유사도는 식(1)과 같다[7].

$$p(\mathbf{x}|\Omega) = \frac{\exp\left(-\frac{1}{2}(\mathbf{x}-\mathbf{m})^T\mathbf{C}_x^{-1}(\mathbf{x}-\mathbf{m})\right)}{(2\pi)^{N/2}|\mathbf{C}_x|^{1/2}} \dots (1)$$

식(1)의 유사도의 특성을 결정하는데 중요한 통계치는 위쪽 괄호에 포함되어 있는 Mahalanobis distance $d(\mathbf{x})$ 이다[8].

$$d(\mathbf{x}) = \mathbf{x}'^T\mathbf{C}_x^{-1}\mathbf{x}' \dots \dots \dots (2)$$

식(2)에서 $\mathbf{x}' = (\mathbf{x}-\mathbf{m})$ 이다. \mathbf{C}_x 의 고유벡터 \mathbf{e} 와 고유값 λ 를 이용하여 $d(\mathbf{x})$ 를 식(3)과 같이 구할 수 있다.

$$d(\mathbf{x}) = \mathbf{x}'^T\mathbf{C}_x^{-1}\mathbf{x}' = \mathbf{x}'^T[\mathbf{e}\lambda^{-1}\mathbf{e}^T]\mathbf{x}' = \mathbf{y}^T\lambda^{-1}\mathbf{y} \dots \dots \dots (3)$$

식(3)에서 구해진 $d(\mathbf{x})$ 를 합의 항으로 표현하면 식(4)와 같이 된다.

$$d(\mathbf{x}) = \sum_{i=1}^N \frac{y_i^2}{\lambda_i} \dots \dots \dots (4)$$

식(4)에서 M개의 주성분을 사용하기 때문에 정확한 $d(\mathbf{x})$ 의 거리는 알 수 없고 그 값을 식(5)와 같이 추정할 수 있다.

$$d'(\mathbf{x}) = \sum_{i=1}^M \frac{y_i^2}{\lambda_i} + \frac{1}{\rho} \left(\sum_{i=M+1}^N y_i^2 \right) \dots\dots\dots (5)$$

$$= \sum_{i=1}^M \frac{y_i^2}{\lambda_i} + \frac{e^2(\mathbf{x})}{\rho}$$

변수 ρ 는 cost function $J(\rho)$ 를 이용하여 최적화할 수 있다. $J(\rho)$ 는 $P_F(\mathbf{x}|\Omega)$ 와 $P_{F'}(\mathbf{x}|\Omega)$ 사이의 차이를 줄이는 Kullback-Leibler방법에 근거를 두고 있다.

$$J(\rho) = E \left(\log \frac{P(\mathbf{x}|\Omega)}{P'_{F'}(\mathbf{x}|\Omega)} \right) \dots\dots\dots (6)$$

Mahalanobis 거리 $d(\mathbf{x})$ 와 추정치 $d'(\mathbf{x})$ 그리고 $E[y_i^2] = \lambda_i$ 라는 것을 이용하여 위 식을 다시 표현하면 식(7)과 같다.

$$J(\rho) = \frac{1}{2} \sum_{i=M+1}^N \left(\frac{\lambda_i}{\rho} \right) - 1 + \log \frac{\rho}{\lambda_i} \dots\dots\dots (7)$$

ρ 의 최적치 ρ^* 는 이 cost function의 값을 최소화하는 식(8)과 같이 ρ 를 찾으므로 결정할 수 있다. 이는 J 를 ρ 로 편미분하여 $\frac{\partial J}{\partial \rho} = 0$ 을 얻을 수 있다.

$$\rho^* = \frac{1}{N-M} \sum_{i=M+1}^N \lambda_i \dots\dots\dots (8)$$

위에서 설명한 것을 식(1)의 Mahalanobis 유사도 형태로 바꾸면 식(9)와 같이 된다.

$$P'(\mathbf{x}|\Omega) = \left(\frac{\exp \left(-\frac{1}{2} \sum_{i=1}^M \frac{y_i^2}{\lambda_i} \right)}{(2\pi)^{M/2} \prod \lambda_i^{1/2}} \right) \cdot \left(\frac{\exp \left(-\frac{e^2(\mathbf{x})}{2\rho} \right)}{(2\pi\rho)^{(N-M)/2}} \right) \dots (9)$$

$$= P_F(\mathbf{x}|\Omega) P_{F'}(\mathbf{x}|\Omega)$$

$P_F(\mathbf{x}|\Omega)$ 는 F 공간에서의 실제 경계 밀도이고, $P_{F'}(\mathbf{x}|\Omega)$ 는 F' 공간에서의 추정 경계 밀도이다.

$P'(\mathbf{x}|\Omega)$ 는 원하는 특징을 탐색하기 위하여 유사도 맵을 구성하기 위해서 사용된다. 유사도 L 은 입력 영상에 있는 각 공간 좌표의 유사도 추정치를 식(10)과 같이 계산함으로써 얻어진다.

$$L(i, j, \Omega) = P'(\mathbf{x}^j | \Omega) \dots\dots\dots (10)$$

식(10)에서 \mathbf{x}^j 는 입력영상에서 (i, j)위치가 중심인 부분영상으로부터 얻어진 검사용 영상의 벡터화된 영역이다. 얻어진 유사도 맵으로부터 선택되어지는 목적 템플릿은 식(11)로 구할 수 있다.

$$(i^*, j^*) = \max_{i,j} L(i, j, \Omega) \dots\dots\dots (11)$$

여기서, (i^*, j^*) 는 찾은 특징의 중심 위치이다. 비슷한 방법으로 다양한 크기의 특징 탐색에 이러한 방법을 확장한다면 식(12)와 같다.

$$\text{Scale } \sigma = \{\sigma^1, \sigma^2, \dots, \sigma^k, \dots, \sigma^n\}$$

$$(i^*, j^*, k^*)^{ML} = \max_{i,j,k} L(i, j, k, \Omega) \dots (12)$$

스케일 σ 는 입력 영상의 크기 변환을 지정한다. 이 경우, 입력 영상을 스케일 σ^k 로 크기변환 한 후에 (i, j)위치에서 템플릿과 같은 크기의 영역을 얻어 유사도를 구한다. 이렇게 얻어진 유사도 맵의 유사도 값들 중 가장 큰 값을 갖는 (i, j)위치의 k번째 크기가 특징의 영역으로 선택되게 된다(9).

III. 예측맵을 이용한 탐색공간 축소

영상 내부에 얼굴이 존재하는지 여부가 사전에 알려지지 않았고, 존재함을 안다고 해도 그 위치와 스케일을 알 수 없으므로 2장에서 기술한 탐색 방법을 모든 위치에서 모든 스케일에 대해서 유사도를 계산해야한다. 그러나 이러한 방법은 방대한 계산을 필요로 하므로 빠른 반응을 필요로 하는 시스템에 대해서는 적용이 불가능하다. 본 연구에서 제안하는 것은 특징의 위치와 스케일을 결정하는 탐색공간을 현격히 줄일 수 있는 방법에 관한 것이다. 이것은 다음과 같은 두 가지 전제로부터 출발한다.

- (1) 어떤 위치에서 하나의 스케일에 대한 유사도가 현격히 낮았다면 인접 스케일의 특징이 그 위치에서 존재할 가능성도 낮다.
- (2) 어떤 위치에서 윤곽선 분포가 탐색 대상의 윤곽선 분포와 현격히 다르다면 특징의 존재 가능성이 낮다.

〈그림 1〉은 첫 번째 전제를 적용한 것으로서, 탐색대상 특징을 얼굴 전체로 한 경우의 예를 보인다. 예측맵1은 하나의 스케일에서 각 위치에서의 계산된 얼굴의 존재 가능성을 음영으로 나타낸 것이다. 유사도 값이 현격히 낮은 점들을 대상에서 제외하면 다음 스케일에서는 유사도의 계산량이 줄어들 것이다. 예측맵 2와 3은 각기 전단계의 예측맵에서 최대값과 최소값의 중간값을 임계치로하여 얻어진 예측맵이다.

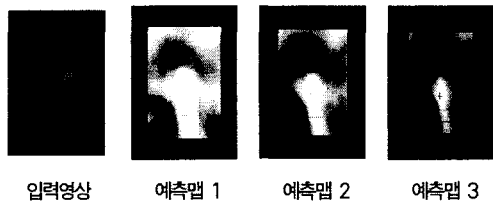


그림 1. 다단계 예측맵을 이용한 얼굴탐색
Fig. 1 Face detection using multilevel forecast map

본 연구에서는 얼굴 전체를 탐색대상으로 하지 않고 다양한 스케일의 눈을 대상으로 하여 탐색하며 탐색된 양쪽 눈들로부터 얼굴 전체의 위치와 스케일을 결정하는 방법을 택한다. 따라서 주성분 분석으로 눈의 영역을 찾기 위해서는 여러 종류의 눈들과 그것들로부터 구해진 평균 눈이 필요하다. 〈그림 2〉는 좌측과 우측 눈들의 평균 눈을 보인다.



그림 2. 평균눈 영상
Fig. 2 Average eye images

앞에 언급한 첫 번째 전제를 이용한 방법으로 눈 영역을 찾기 위해서는 〈그림 3〉과 같이 좌측 눈에 의한 좌측 예측맵과 우측 눈에 의한 우측 예측맵을 이용할 수 있다.

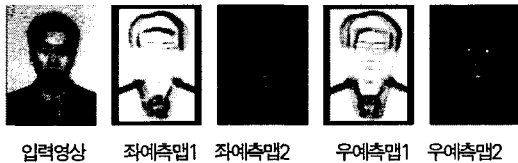


그림 3. 다단계 예측맵을 이용한 눈탐색
Fig. 3 Eye detection using multilevel forecast map

〈그림 3〉에서 보이는 것과 같이 첫 번째 템플릿을 이용한 예측맵1들을 이용하여 두 번째 예측맵2들은 상당히 빠른 탐색을 하는 것으로 보인다.

〈그림 4〉는 두 번째 전제를 사용한 방법을 예시한 것으로, 입력영상의 윤곽선을 추출한 후에 추출된 윤곽선을 이용하여 예측맵을 생성한 예를 보인다.



그림 4. 윤곽선 예측맵의 생성
Fig. 4 Generation of Contour forecast map

입력된 검사 영상의 특정 위치에서의 유사도는 식(13)과 같이 구할 수 있다.

$$P(x|\Omega) = \left(\frac{\exp\left(-\frac{1}{2} \sum_{i=1}^M \frac{y_i^2}{\lambda_i}\right)}{(2\pi)^{m/2} \prod_{i=1}^M \lambda_i^{1/2}} \right) \cdot \left(\frac{\exp\left(-\frac{e^2(x)}{2\rho}\right)}{(2\pi\rho)^{(N-M)/2}} \right) = P_F(x|\Omega) P'_F(x|\Omega) \dots (13)$$

그러나 식(13)은 계산 과정이 복잡하고, 수행시간이 많이 걸리며 또한 매번 계산을 행할 필요가 없는 부분도 있다. 특징을 탐색하는 데는 크기만을 비교하게 됨으로 실제적으로 특별한 의미를 갖지 않는다. 그러므로 식(13)은 식(14)와 같이 변형할 수 있다.

$$P(x|\Omega) = \text{Exp}\left(-\frac{1}{2} \left(\sum_{i=1}^M \frac{y_i^2}{\lambda_i} + \frac{e^2(x)}{\rho} + N \log 2\pi + (N-M) \log \rho + \sum_{i=1}^M \log \lambda_i \right)\right) \dots (14)$$

식(14)에 나오는 알 수 없는 값들은 ρ , $e^2(x)$ 이다. ρ 는 구하지 못한 λ 값들의 산술평균이며 이는 식(15)와 같이 구할 수 있다.

$$\rho = \frac{1}{N-M} \sum_{i=M+1}^N \lambda_i = \frac{1}{N-M} \left(C_s - \sum_{i=1}^M \lambda_i \right) \dots (15)$$

식(15)에서 C_s 는 훈련 벡터들에 대한 공분산 행렬의 대각축의 값들의 합으로 이 값은 PCA를 수행하면서 구할 수 있다. $e^2(x)$ 는 식(16)과 같이 계산할 수 있다.

$$e^2(x) = (x-x')^T(x-x') - \sum_{i=1}^M y_i^2 \dots\dots\dots (16)$$

식(16)은 다음과 같이 주성분 분석 결과가 주어졌을 때 계산할 수 있는 식(17)의 A와 매 위치마다 계산해야 하는 식(18)의 B로 구분할 수 있다.

$$A = N \log 2\pi + (N-M) \log \rho + \sum_{i=1}^M \lambda_i \dots\dots\dots (17)$$

$$B = \sum_{i=1}^M \frac{y_i^2}{\lambda_i} + \frac{e^2(x)}{\rho} \dots\dots\dots (18)$$

IV. 실험 및 결과

예측맵을 사용한 주성분 분석 얼굴 탐색의 실험은 170개의 템플릿을 이용하여 구성된 주성분 분석자료를 이용하여 실행되었다. 템플릿 영상들은 위치와 크기 빛의 효과를 정규화 시켰으며 이를 이용하여 탐색 공간의 단축을 실험하였다. 광학적 정규화에서는 각 영상마다 조명된 빛의 양이 다르므로 조명효과를 제거하는 방법인 Homomorphic 필터링(10)을 사용하였다. 사용된 필터링은 입력된 영상을 주파수 공간으로 변환하여 조명의 효과로 발생하는 낮은 주파수를 제거하는 방법으로 매개변수 γ_L 과 γ_H 를 사용하며 여기에서는 $\gamma_L=0.6$, $\gamma_H=1.2$ 를 사용하였다.

〈그림 5〉는 예측맵을 이용하여 좌측 눈과 우측 눈을 찾은 다음에 두 눈의 기울기 보정에 의한 위치정규화, 크기 정규화를 한 후에 미리 만들어 놓은 마스크를 이용하여 얼굴 영역을 추출하는 과정을 보인다.



그림 5. 예측맵을 이용한 얼굴영역 추출 과정
Fig. 5 Face detection process with forecast map

위의 예제의 경우 233×303 크기의 실험영상에서 탐색공간의 크기가 70599픽셀에서 6357픽셀로 축소됨으로써 제안된 방법이 기존의 전역탐색방법에 비하여 필요한 계산량을 10% 이하로 줄일 수 있었다.

V. 결론

본 논문에서는 Maximum likelihood방법을 사용하여 얼굴을 탐색하는 경우에 대상의 위치와 스케일에 대한 사전정보가 없어 전체 영상을 탐색함으로 인한 방대한 계산량을 줄이는 방법을 제안하였다. 어떤 위치에서 하나의 스케일에 대한 유사도가 현격히 낮았다면 인접 스케일의 특징이 그 위치에서 존재할 가능성도 낮다는 첫 번째 전제를 사용하여 PCA기반의 다단계 예측맵에 의한 템플릿의 탐색공간을 축소하였다. 이와 더불어, 어떤 위치에서 윤곽선 분포가 탐색 대상의 윤곽선 분포와 현격히 다르다면 특징의 존재 가능성이 낮다는 두 번째 전제를 사용하여 윤곽선 예측맵을 이용함으로써 추가적인 탐색공간의 축소를 달성하였다.

참고문헌

- [1] M. Bichsel, Strategies of Robust Objects Recognition for Automatic Identification of Human Laboratory, Massachusetts Institute of Technology, 1993.
- [2] R. Brunelli and T. Poggio, Face Recognition : Features bersus Templates. IEEE Transactions on Pattern Analysis and Machine Intelligence, 15(10) : 1042 - 1052, 1993.
- [3] A. Yuille, P. Hallinan, and D. Cohen. Feature Extraction from Faces using Deformable

- Templates. International Journal of Computer Vision, 8(2) : 99 - 111, 1992.
- [4] P. Sinha. Object Recognition via Image Invariants : A Case Study. In Investigative Ophthalmology and Visual Science, volume 35, pages 1735 - 1740, Sarasota, Florida, May 1994.
- [5] M. C. Burl, U. Fayyad, P. Perona, P. Smyth, and M. P. Burl. A Trainable Tool for Finding Small Volcanoes in SAR Imagery of Venus. Technical Report CNS TR 34, California Institute of Technology, October 1993.
- [6] Shiguang Shan, Wen Gao, Xilin Chen, Jiyong Ma, "Novel Face Recognition Based on Individual Eigen - subspaces," 000 5th International Conference on Signal Processing Proceedings - Volume 3, 1522 - 1525, Aug. 2000.
- [7] Ara Y. Nefian and Monson H. Hayes III, "Maximum Likelihood Training Of The Embedded Hmm For Face Detection And Recognition," Proceedings of the 2000 International Conference on Image Processing - Volume 1 , 33 - 36 , Aug. 2000.
- [8] Kamei, T., "Face retrieval by an adaptive Mahalanobis distance using a confidence factor," Proceedings of the IEEE International Conference on Image Processing - Volume 1, 153 - 156, Aug. 2002.
- [9] Baback Moghaddam and Alex Pentland, "A Subspace Method for Maximum Likelihood Target Detection," M.I.T Media Lab., PCST Report No. 335
- [10] Refeal G. gonzalez, Richard E. Woods, 'Digital Image Processing,' Addison - Wesley Publishing Company, 1993.

저 자 소 개



조 경 식

1986 단국대학교 계산통계학과
학사
1993 단국대학교 전산통계학과
석사
2001 단국대학교 전산통계학과
박사 수료
1999년 ~ 현재 강원관광대학
컴퓨터정보계열 전임강사



구 자 영

1977 서울대학교 전자공학과 학사
1980 한국과학기술원 전기 및 전
자공학부 석사
1986 한국과학기술원 전기 및 전
자공학부 박사
1986년 ~ 현재 단국대학교
정보·컴퓨터학부 교수