

얼굴 인식을 위한 기술어 추출의 고속화 알고리즘

*이재용, **김지은, ***정광수, ****오승준
 광운대학교 방통융합 플랫폼 센터
 *ljy6225@media.kw.ac.kr

Fast descriptors extraction algorithm for face recognition

*Lee, Jea-Yong **Kim, Ji Eun ***Chung, Kwang-Sue ****Oh, Seung-Jun
 BnC Convergence Platform Center, Kwangwoon University

요 약

기존 얼굴 인식 알고리즘은 단일 특징 기반의 전역적 방식이었다. 정확도를 향상시키기 위해 복수의 특징점을 이용하는 방법들이 제안되었으나 이는 알고리즘의 복잡도가 증가하고, 계산 속도가 느린 단점이 있다. 본 논문에서는 대표적인 특징점 추출 알고리즘인 SURF (Speeded Up Robust Features)를 이용한 얼굴 인식 방법을 제안한다. SURF 를 통해 기술어를 추출하고, Gabor 특징과 LBP 특징을 이용해 해당 특징점에서 기술어를 추출함으로써 기존 알고리즘보다 경량화할 수 있고, 수행시간을 줄일 수 있다. 잘 알려진 ORL 데이터베이스에서의 실험에서 제안한 방법이 정합시간을 포함한 수행 시간에서 약 16%의 감소를 보였고, 정확도 또한 약 34% 향상되었다.

1. 서론

얼굴인식은 다양한 응용이 가능한 주제이지만 해결하기 어려운 문제이기 때문에 지난 10년간 많은 관심을 받았다[1]. 데이터 집합의 통계적 특성을 분석하여 적절한 차원의 특징을 추출하는 부분공간분석 방법인 Eigenface 방법과 Fisherface 방법, Gabor 웨이블릿 변환(wavelet transform)을 이용한 방법은 잘 알려진 얼굴 인식을 위한 접근 방법들이다[2-3]. 하지만 이러한 전역적 특징 추출 방법은 사용하는 벡터의 차원 수가 높기 때문에 연산 시간의 증가 뿐 아니라 인식률을 감소시키는 차원의 저주 (Curse of dimensionality)와 같은 문제를 유발하기도 한다. 이를 해결하기 위해, 벡터의 차원을 축소시키는 PCA(Principal Component Analysis), LDA (Linear Discriminant Analysis) 및 ICA (Independent Component Analysis) 등의 방법이 제안되었다[4-5]. 그리고 하나의 특징을 이용한 단일 특징 기반의 방법들의 성능을 향상시키기 위해 다양한 특징을 이용한 복수 특징 기반의 방법들 또한 제안되었다. 그림 1은 단일 특징 기반과 복수 특징 기반 방법을 보여준다.[6] 하나 이상의 특징을 이용하여 기술어를 추출하면 좋은 인식 성능과 안정성을 확보할 수 있는 반면, 알고리즘의 복잡도가 증가하고, 계산 속도가 느려지게 된다.

반면 영상의 객체 인식 분야에서 주로 사용되는 지역적 특징 추출 방법은 일정 지역에 해당하는 화소를 이용하기 때문에 벡터의 차원 수는 전역적 특징 추출 방법보다 낮아진다. 따라서 연산시간이나 차원의 저주와 같은 문제에 상대적으로 영향을 덜 받는다. 하지만 해당 화소를 선택하고 그에 따른 기술어를 추출함에 있어서 기술어는 영상의 주요

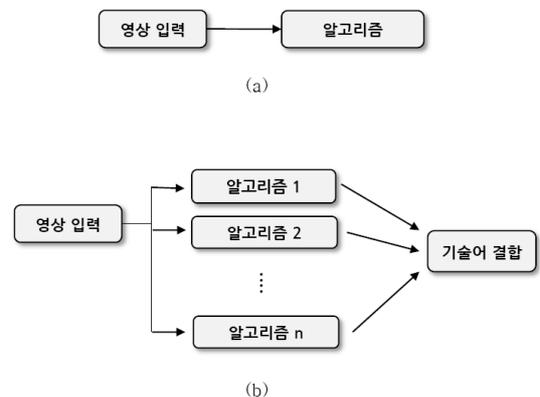


그림 1. 기술어 추출 알고리즘:

(a) 단일 특징 기반 방법, (b) 복수 특징 기반 방법

부분을 충분히 잘 표현할 수 있도록 하는 구별성 (distinctness)이 보장되어야 하며, 동시에 환경 변화에 강인해야 한다[7].

대표적인 지역적 접근 방법으로는 SIFT (Scale Invariant Feature Transform)와 SURF (Speeded Up Robust Feature) 등이 있으며 이들은 물체인식, 위치인식 등에 사용되고 있다. 최근에는 얼굴 인식과 관련된 연구에도 응용되는 추세이다 [8-9].

일반적으로 SURF 는 물체를 인식하는데 응용되어 왔다. 서로 다른 물체의 경우에는 형태 및 구조에 많은 차이가 있어 특징점 간에 분별력이 있다. 반면에, 서로 다른 사람의

경우에는 얼굴의 형태 및 구조가 유사해 물체의 경우보다 분별력이 떨어진다.

본 논문에서는 SURF 를 이용해 추출한 특징점에서 Gabor 특징과 LBP 특징을 사용해 기술어를 추출하는 방법을 제안한다. 얼굴 전체가 아닌 특징점으로 추출된 화소에서만 기술어를 추출함으로써 수행시간을 줄이고, 알고리즘을 경량화할 수 있다. 또한 다양한 크기와 방향성을 반영한 Gabor 특징과 특징점 화소값과 주변 화소값들의 상대적인 차이를 반영하여 조명 변화에 강한 LBP 특징을 결합함으로써 더욱 강한 기술어를 추출할 수 있다.

본 논문의 구성은 다음과 같다. 2 장에서는 영상에서 기술어를 추출하기 위해 기존에 사용된 기술어 추출 방법을 설명하고, 3 장에서는 제안하는 방법에 대하여 설명한다. 이어서 4 장에서는 실험 환경을 제시하고 실험을 통해 제안하는 방법을 입증한다. 마지막으로, 5 장에서 본 논문에 대한 결론을 맺는다.

2. 기존의 기술어 추출 방법

전역 방식 중 한 방법인 PCA(Principal Component Analysis)는 패턴인식, 통계학, 신호처리 분야 등에서 많이 쓰이는 패턴 분류 방법으로, 고차원의 입력 데이터에서 원하는 수만큼의 주성분 요소를 추출하여 이를 이용해 데이터를 표현하면 데이터의 차원을 축소한 분석이 가능하다[9]. PCA 는 다차원 데이터 집합의 차원 축소뿐만 아니라, 주성분 요소들이 데이터 집합에서 차례로 변동이 큰 방향의 벡터들이기 때문에 가장 큰 분산을 갖는 부분공간(subspace)을 보존하는 최적의 선형변환이라는 특징을 갖는다. 일반적으로 고차원인 얼굴 영상 데이터를 저차원의 데이터로 줄여서 수행 시간을 단축시키고 데이터 저장 공간을 절약할 수 있기 때문에 얼굴 인식 분야에서 많이 사용되고 있는 방법 중 하나이다. 하지만 얼굴 영상 전체 지역의 화소를 이용하여 고유의 얼굴 성분 벡터를 구하기 때문에 환경변화에 민감하다.

하나의 특징을 이용한 방법은 성능 개선에 한계가 있기 때문에 얼굴 인식의 성능을 높이기 위해 다양한 특징을 이용한 복수 특징 기반의 기술어 추출 방법들이 제안되었다. Tan 과 Triggs 는 그림 2 와 같이 대표적인 얼굴 인식 특징인 Gabor 특징과 Texture 인식에서 많이 사용되던 LBP 특징을 동시에 사용하였다. 각각의 특징 벡터는 PCA 를 통과하면서 차원을 낮춘 뒤 정합하게 된다[10].

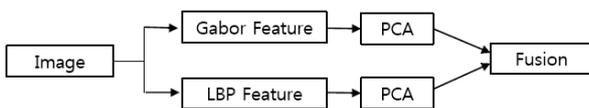


그림 2. Tan and Triggs의 얼굴 인식 블록도

이러한 방법들은 좋은 인식 성능과 함께 안정성을 확보할 수 있지만, 알고리즘의 복잡도가 증가하거나 계산 속도가 느려지게 된다.

특징점 추출 방법인 SIFT 는 최근 얼굴인식에도 적용되고 있다. SIFT 는 공간영역에서 선택된 특징점 후보들 중에 다양한 스케일에서 스케일 변화에 영향을 받지 않을 특징점을 선택하기 때문에 방향 및 크기의 변화에 강인하다는 장점을 갖는다. 먼저 특징점을 추출하기 위해 scale-space DoG(Difference of Gaussian)를 구성한다. DoG 함수는 가우시안 평활화(Gaussian smoothing)와 부분 샘플링을

통하여 구한 각기 다른 스케일을 가지는 이미지 피라미드(image pyramid)간의 차로 구해지며, 다음 식 1 로 정의될 수 있다[8].

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \quad (1)$$

$$= L(x, y, k\sigma) - L(x, y, \sigma)$$

위 식에서 k 는 상수이고 $I(x, y)$ 는 영상의 (x, y) 화소 값, 그리고 $G(x, y, \sigma)$ 는 분산 σ 를 가지는 2 변량 가우시안 함수이다. DoG 에서의 극대, 극소점이 SIFT 의 특징 후보점이 되며, 이후 선정된 특징점 후보 픽셀들 중에 밝기의 대비가 낮거나 에지에 걸쳐있는 후보들을 제거하면 최종 특징점이 정해진다. 특징점이 확정되면 특징점 부근의 4x4 영역 내부 픽셀의 기울기 값을 8 방향의 히스토그램으로 나타내고 이를 기술어로 사용한다. 각 4x4 영역에서의 기울기의 크기 와 방향 는 각각 다음과 같이 정의한다. 특징점이 확정되면 특징점 부근의 4x4 영역 내부 화소의 기울기 값을 8 방향의 히스토그램으로 나타내고 이를 기술어로 사용한다. 각 4x4 영역에서의 기울기의 크기 $m(x, y)$ 와 방향 $\theta(x, y)$ 는 식 (2), 식 (3) 과 같이 정의한다[11].

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (2)$$

$$\theta(x, y) = \tan^{-1} \left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \right) \quad (3)$$

위의 식으로 계산한 기울기는 주 방향(principle direction)에 맞춰 정렬하며, 이러한 정렬을 통해 회전변화에 강한 128 차원의 기술어를 얻는다. SIFT 는 크기, 방향, 회전변화에 강한 기술어를 얻을 수 있지만 기술어를 추출하는데 많은 양의 특징점 정보를 계산하고 저장하기 때문에 시간이 많이 소요된다. 이를 개선한 방법으로 SURF 가 있다. SURF 는 SIFT 와 비슷한 알고리즘으로 유사한 성능을 보이면서 연산 속도 측면에서 향상된 방법이다.

3. 제안하는 방법

기존의 전역 방식의 얼굴 인식 방법들은 정확도와 안정성은 높지만 시간이 얼굴 영상 전체에서 기술어를 추출하기 때문에 시간이 매우 오래 걸린다. 본 논문에서는 대표적 특징점 추출 방법인 SURF 를 이용해 특징점을 추출하고, 해당 특징점에서만 기술어를 추출하는 방법을 제안한다. 기존의 SURF 를 이용한 얼굴 인식 방법은 다른 방법들에 비해 속도는 빠르지만 정확도가 떨어진다. SURF 기반의 기술어 추출 방법에서는 특징점을 검출하면서 정해진 스케일에 따라 하나의 주 방향과 영역의 크기가 결정되게 된다. 결정된 해당 특징점의 주변영역에서 방향의 기울기와 방향의 기울기 성분의 합을 이용해 기술어를 구성하는데, 정해진 하나의 방향과 크기가 기술어를 구성함에 있어 큰 영향을 미친다.

반면, Gabor 특징을 이용한 기술어 추출 방법에서는 특징점에서 여러 방향 성분과 여러 스케일의 조합으로 이루어진 커널을 컨벌루션하여 기술어를 추출한다. 다양한 방향과 스케일을 적용하기 때문에 Gabor 특징 기반의 기술어는 SURF 기반의 기술어보다 더 강한 특성을 갖는다. 그리고 중앙값과 주변 값들의 상대적인 값 차이로 표현하는 LBP

특징은 조명 변화에 강인한 성질을 보여준다. 서로 다른 특징을 가지는 두 특징을 이용해 각각 하나의 특징을 이용한 기술어 추출 방식보다 좋은 성능 개선을 이끌어 낼 수 있다.

본 논문에서는 그림 3 과 같이 SURF 방법을 이용해 특징점을 검출한 후, 해당 특징점에서 Gabor 특징과 LBP 특징을 이용해 기술어를 추출하고 정합하는 방법을 제안한다.



그림 3. 제안하는 얼굴 인식 블록도

4. 실험 결과

이 절에서는 제안된 방법의 유용성을 검증하기 위하여 다양한 변화를 지니는 ORL 데이터베이스를 이용한 비교실험을 수행한다. 사용한 데이터는 Cambridge 의 Olivetti Research Laboratory 에서 제작된 ORL 데이터베이스이다. 영상의 크기는 92x112 이고 40 명에 대하여 빛의 위치 변화, 안경 착용 유무, 표정 변화(눈의 변화, 웃음 유무) 등에 따라 각 10 장씩으로 구성되어 있다. 전체 데이터베이스에 일정하게 나타나는 규칙은 없고, 모든 얼굴 영상은 검은 배경을 바탕으로 한다. 안경, 수염 등이 있는 얼굴들이 많이 분포 되어 있고 좌우 회전을 통해 변형된 얼굴 영상도 다소 존재하기 때문에 난이도가 있는 얼굴 데이터베이스라고 할 수 있다[12].



그림 4. ORL 데이터베이스의 예

ORL 데이터베이스에 대하여 기존의 기술어 추출 방법인 Gabor, SIFT, SURF128, SURF64 와 제안하는 방법의 성능을 비교하였다. 실험의 정량적인 평가를 위해서 MPEG-7 검색 성능 평가 지수인 ANMRR(Average Normalized Modified Retrieval Rank)를 이용한다[13].

다음 표 1 은 기존 방법들과 제안한 방법의 수행시간을 나타낸 것이다. T_e 는 각 방법에 의해 기술어가 추출되는 시간이고, T_m 은 정합하는데 걸리는 시간을 의미한다. 최종적으로 기술어 추출시간과 정합시간을 더한 시간이 각 방법들의 얼굴 인식 수행시간이 된다. 얼굴 인식은 많은 데이터베이스와 정합해야 하기 때문에 오랜 수행시간은 불리할 수밖에 없다. 표 1 을 통해 제안된 방법의 수행시간이 PCA 를 제외한 방법들 중 가장 적게 걸리고, 기존 SURF64 방법과 비교해 약 15.7%의 효율을 얻을 수 있는 것을 확인할 수 있다.

표 1. 기존 방법들과 제안된 방법의 수행시간 비교

Method	PCA	Gabor	SIFT	SURF128	SURF64	제안하는 방법
T_e (ms)	5.76	498.36	21.94	4.11	4.07	189.98
T_m (ms)	58.50	384479.68	3920.77	2275.12	1255.61	872.49
$T_e + T_m$ (ms)	64.26	384978.04	3942.72	2279.23	1259.68	1062.47

SIFT 와 SURF 를 이용한 방법과 비교했을 때 기술어를 추출하는 시간은 증가하였지만 정합시간에서 많은 감소를 보였다. 이는 영상 전체가 아닌 특징점에서만 기술어를 추출하여 정합하기 때문이다. 뿐만 아니라 한 특징점에서의 기술어의 수가 기존 방법들은 128 개거나 그 이상, 또는 64 개인데, 제안하는 방법은 그보다 적은 42 개이므로 정합하는 시간에서 많은 시간적 효율을 얻을 수 있다.

얼굴 인식은 시간적 효율뿐만 아니라 성능도 중요하다. 아래 그림 5 는 기존의 다른 방법들과 본 논문에서 제안하는 방법의 성능을 비교한 그래프이다.

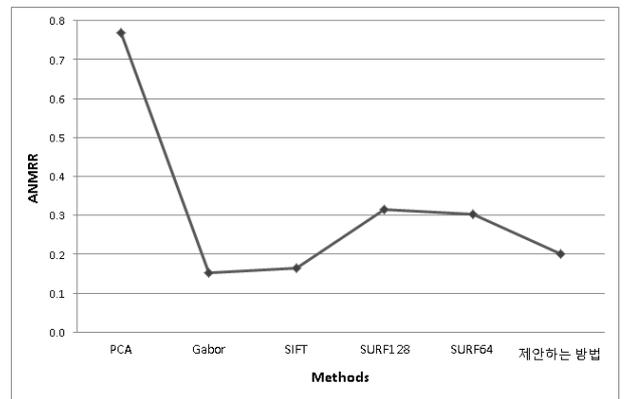


그림 5. 기존 방법들과 제안한 방법의 성능 비교

표 1 에서 PCA 는 수행시간이 상당히 적게 걸리지만 아래 그림 5 에서 보면 다른 인식 방법들에 비해 정확도가 떨어지는 것을 확인할 수 있다. 반면, Gabor 와 SIFT 방법은 성능은 좋지만 수행시간이 상당히 오래 걸린다. 제안하는 방법의 성능은 Gabor 와 SIFT 방법에는 못 미치지만 SURF 방법보다 약 34% 향상된 것을 확인할 수 있다. 따라서 SURF 보다 성능이 좋고, 다른 방법과 비교해 수행시간도 가장 빠르기 때문에 제안하는 방법이 인식률과 수행시간을 동시에 고려할 때 효율적인 방법이라고 볼 수 있다.

5. 결론

본 논문은 얼굴 인식에 관한 연구로 SURF 알고리즘의 특징점 추출과 Gabor 특징, LBP 특징을 이용한 기술어를 결합하여 새로운 얼굴 인식 시스템을 제안하였다. 제안하는 방법의 유용성을 판단하기 위해 ORL 데이터베이스를 이용하여 다양한 변화를 가진 얼굴을 인식하는 실험을 하였다. 실험 결과, 기존의 SURF 를 이용한 얼굴 인식보다 빠른 수행속도를 보여주며, 동시에 높은 성능을 얻을 수 있었다.

감사의 글

본 연구는 지식경제부와 한국산업기술진흥원의 전략기술 인력양성사업으로 수행된 결과임

참 고 문 헌

- [1] W. Zhao, R. Chellappa, P.J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey", *ACM Comput. Surv.*, vol. 35, No. 4, pp. 399-458, 2003.
- [2] M. Turk and A. Pentland, "Eigenfaces for recognition", *Journal of cognitive neuroscience*, vol. 3, No. 1, pp. 71-86. 1991.
- [3] D. Gabor, "Theory of Communication", *J. Inst. Elect. Eng.*, vol. 93, no 3, pp. 429-457, 1946.
- [4] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711- 720, Jul.1997.
- [5] M. S. Bartlett, J. R. Movellan, and T. J. Sejnowski, "Face recognition by independent component analysis", *IEEE Trans. Neural Netw.*, vol.13, no. 6, pp. 1450- 1464, Jun. 2002.
- [6] 황원준, 김준모, "Face Recognition Grand Challenge (FRGC) 및 조명 변화에 강인한 얼굴 인식 기술 개발 동향", *전자공학회지*, 제 39 권 2 호, pp.108-116, Feb. 2012.
- [7] R. Verschae, J. Ruiz-del-Solar, and M. Correa, "Face recognition in unconstrained environments: a comparative study", *Proc. of ECCV Workshop on Faces in Real-Life Images*, 2008.
- [8] D.G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints", *Int. J. Comput. Vision*, vol. 60, No.2, pp. 91-110, 2004.
- [9] H. Bay, Beat Fasel, and Luc Van Gool, "Interactive museum guide: Fast and robust recognition of museum objects", *In First international workshop on mobile vision*, 2006.
- [10] X. Tan and B. Triggs, "Fusing Gabor and LBP feature set for kernel-based face recognition", *IEEE International Workshop on Analysis and Modeling of Face and Gestures*, pp.235-249, 2007.
- [11] D. Lowe, "Object recognition from local scale-invariant features", *In ICCV*, 1999.
- [12] Ferdinando Samaria, Andy Harter. "Parameterisation of a Stochastic Model for Human Face Identification", *Proceedings of 2nd IEEE Workshop on Applications of Computer Vision*, Sarasota FL, December 1994
- [13] V. V. Vinod and B. S. Manjunath, "Report on AHG of color and texture", *ISO/IEC/JTC1/SC29/WG11*, Doc. M5560, Maui, December 1999. 332