

파티클 필터에 기반한 강인한 얼굴추적을 위한 텍스처 특징 추출에 관한 연구

김동규, 이승호, 김형일, 노용만

카이스트 전기 및 전자공학부

e-mail : {d.kim, leesh09, hyungil.kim, ymro}@kaist.ac.kr

Texture Feature for Robust Particle Filter Based Face Tracking

Dongkyu Kim, Seung Ho Lee, Hyung-Il Kim, Yong Man Ro

School of Electrical Engineering, KAIST

요 약

파티클 필터 기반 얼굴추적은 비교적 빠른 속도와 구현의 용이성으로 널리 사용되고 있으나 조명이나 포즈변화가 있는 영상에서 드리프트(drift) 현상에 의해 얼굴추적의 정확도가 급격히 저하된다. 본 논문에서는 앞에 언급한 얼굴의 다양성에 강인한 얼굴 텍스처 특징을 제안한다. 제안방법은 인접한 픽셀들 간의 관계를 고려한 텍스처 패턴을 정의할 때 인접한 픽셀들의 평균(average)을 적용하여 조명변화에 강인하다. 또한 얼굴의 구조적 정보를 반영한 블록 기반의 텍스처 패턴 풀링(pooling)에 의해 포즈변화에 강인하다. 실제 감시환경을 가정해 CCTV 카메라로 자체 제작한 비디오 영상에서 Local Binary Pattern(LBP)와 같은 대표적인 특징들과 비교 실험을 수행하였다. 실험결과, 드리프트(drift) 폭이 적어 더 높은 얼굴추적 정확도를 보였으며 초당 28 프레임의 매우 빠른 처리속도를 보였다.

1. 서론

얼굴 추적 기술은 시간에 따른 관심 얼굴 영역의 위치 변화를 추적해 나가는 기술로서 기본적인 화상 회의, e 러닝, 감시시스템, 출입구 관리 시스템, 프라이버시 보호 등 다양한 응용분야 있다. 특히, 대용량·대규모 감시시스템에서는 군중 등 복수의 얼굴의 추적을 위해 보다 효율적인 얼굴추적 기술이 요구된다.

통계적 방법의 하나인 파티클 필터(particle filter)는 비교적 좋은 성능과 구현의 용이함, 빠른 계산속도로 인해, 많은 얼굴추적 방법 에 사용되고 있다[1]. 파티클 필터 기반의 얼굴추적은 각 프레임에서의 얼굴의 상태(state)를 결정하기 위해 관측에 대한 확률 모델을 이용하는데, 초기 상태에서 추출한 특징과 해당 프레임의 파티클에서 추출한 특징을 비교를 수행한다. 여기서, 보다 강인한 얼굴 추적을 하기 위해 다양한 얼굴 특징들 가운데 분별력이 좋은 특징을 선택하여 관측 모델에 적용하는 것이 중요한 요소 중에 하나로 알려져 있다.

적은 연산량으로 빠른 얼굴 추적에 용이한 특징인 칼라 히스토그램과 국부이진패턴(Local Binary Pattern, LBP)[2]은 각각 급격한 조명변화 및 포즈변화에 취약해 쉽게 드리프트(drift)가 일어나게 된다.

본 논문에서는 파티클 필터 기반 얼굴추적에 적합한 얼굴 특징을 제안한다. 제안방법은 얼굴추적에서 포즈나 조명 변화에 강인함과 고속처리를 목적으로 한다. 제안방법의 특징점은 다음과 같이 두 가지로 요약된다.

- 1) 제안하는 얼굴 텍스처 특징은 인접한 픽셀들 간의 관계를 고려한 텍스처 패턴을 정의할 때 인접한 픽셀들의 평균(average)을 적용하여 조명변화에 강인하다. 또한 얼굴의 구조적 정보를 반영한 블록 기반의 텍스처 패턴 풀링(pooling)에 의해 포즈변화에 강인하다.
- 2) 제안하는 얼굴 텍스처 특징은 낮은 특징 차원수로 적은 메모리로 고속 처리가 가능하다. 따라서 대규모 군중들의 얼굴 데이터를 실시간으로 추적하고 보관해야 하는 지능형 감시시스템에 적합하다.

실제 감시환경을 가정해 자체 제작한 비디오 영상에 대한 실험에서 포즈 및 조명변화에도 다른 특징들에 비해 작은 중심 위치 오차와 함께, 초당 28 프레임의 빠른 처리를 확인하였다. 본 논문의 구성은 다음과 같다. 2 장에서는 제안하는 얼굴 텍스처 특징 추출 방법을 설명하고 3 장에서는 이 특징과 파티클 필터를 이용한 얼굴 추적 방법에 대해 설명한다. 그리고 4 장에서는 실험 결과를 보이고, 5 장에서는 결론을 맺는다.

2. 제안하는 얼굴 텍스처 특징 추출 방법

제안하는 얼굴 텍스처 특징 추출은 얼굴 분석(face analysis)에서 대표적인 텍스처 특징 중의 하나인 LBP에 기반한다. LBP 텍스처 패턴 연산은 사전에 정의된 파라미터 P 와 R 에 대해, 다음 수식과 같이 정의된다 [2].

n_0	$n_1(g_0)$	n_2
$n_3(g_3)$	$n_4(g_c)$	$n_5(g_1)$
n_6	$n_7(g_2)$	n_8

그림 1 제안하는 방법에서 LBP 텍스처 패턴 추출 시 중심 픽셀과 이웃 픽셀 표기 정의 ($P=4, R=1$ 인 경우)

$$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p, \quad (1)$$

수식 (1)과 같이 중심 픽셀(center pixel) 값과 이웃 픽셀(neighborhood pixel) 값의 크기를 비교하여 0 또는 1의 2진수로 부호화 된 후에 십진수로 변환함으로써 LBP 값을 얻게 된다. 이 때, P 는 비교가 되는 이웃 픽셀의 개수, R 은 중심 픽셀과 이웃 픽셀 간의 거리, g_c 는 중심 픽셀의 값을, g_p 는 p 번째 이웃 픽셀의 값을 뜻한다. 그리고 함수 $s(x)$ 는 x 가 0 보다 클 경우에만 1 을 반환하는 함수이다. LBP 는 단조로운(monotonic) 조명변화가 있는 경우에도 정확한 추적이 가능하게 하는 장점이 있지만, 단조롭지 않은 조명변화 또는 심한 포즈변화가 있는 경우 문제를 겪게 된다. 제안하는 방법은 위에 언급한 LBP 의 두 가지 변화에 대한 취약성을 보완하고자 한다.

먼저 조명변화에 대한 강인성을 향상시키기 위해 수식 (2)와 같이 새로운 텍스처 패턴 $LBP_{P,R}^{revised}$ 을 정의한다.

$$LBP_{P,R}^{revised}(x_c, y_c) = \sum_{p=0}^{P-1} s(g'_p - g'_c) 2^p, \quad (2)$$

$$g'_p = \frac{1}{2}(g_p + g_c), p = 0, 1, \dots, P, \quad g'_c = \frac{1}{K} \sum_{k=0}^{K-1} n_k, \quad (3)$$

여기서 K 는 R 픽셀 이내에 있는 전체 픽셀의 개수를 가리킨다. 본 연구에서는 P, R 을 각각 4, 1 로 두고 사용하였다. 그림 1 인 이 때 픽셀 표기의 정의를 보여준다. 수식 (3)에서와 같이 중심 픽셀과 이웃 픽셀의 값을 정의할 때 인접한 픽셀들과의 평균을 취함으로써 단조롭지 않은 조명변화에 대한 강인성을 향상시킬 수 있다.

두 번째로 포즈변화에 대한 강인성을 향상시키기 위해 수식 (2)와 (3)을 이용하여 추출된 바이너리 형태의 텍스처 패턴을 그림 2 와 같이 16×16 픽셀로 리사이징 후 양쪽 눈과 입에 해당하는 세 개의 블록들에서 각각 LBP 히스토그램[2]을 계산하여 풀딩한다. 얼굴의 구조적 정보를 얼굴추적에 이용하기 위해 세 개의 블록들에서 계산된 LBP 히스토그램을 연결(concatenate)하여 융합한다. 이 때, 각 블록으로부터 나온 특징의 차원 수(히스토그램 빈 수)는 15 이며, 최종적으로 45 차원의 특징이 된다.

3 장에서는 융합된 LBP 히스토그램을 이용한 파티클 필터 기반 얼굴추적 방법을 설명한다.

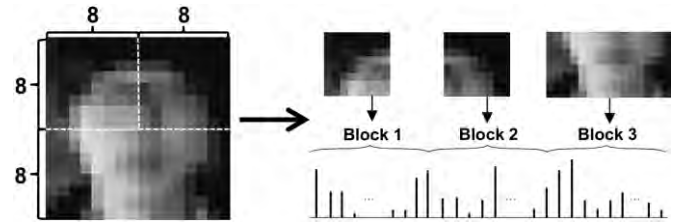


그림 2 얼굴 구조 정보를 활용한 블록 구성 방법[3]

3. 제안하는 얼굴 텍스처 특징과 파티클 필터를 이용한 얼굴 추적

파티클 필터는 일종의 베이지안 순차 중요 표본추출(sequential importance sampling) 방법으로 유한한 가중 파티클 집합을 가지고 표본추출을 반복함으로써 사후 확률 분포를 추정하는 기법이다[1]. 파티클 필터 기반의 얼굴 추적 기법은 이전 프레임에서의 관측에 의해 현재 프레임에서의 얼굴의 상태를 예측하여 추정하는 것이다. 얼굴 추적에서 상태는 위치 좌표(x, y)와 크기(너비 및 높이)로 표현된다. 관측은 생성된 파티클들의 가중 합으로 정의되며, 가중치는 특징 비교에 따른 우도(likelihood)로 결정된다.

본 연구에서는 초기에 검출된 얼굴로부터 2 장에서 제안한 텍스처 특징(초기 상태 특징)을 추출하고 우도(또는 가중치) 계산 시 비교 대상으로 정한다. 즉, 이전 프레임에서 얻은 추적 결과로부터 예측된 위치로 산란된 서로 다른 상태를 갖는 파티클 패치에 대해 제안하는 특징을 추출하고, 이를 초기 상태 특징과 비교해 가장 높은 우도를 갖는 파티클의 상태를 현재 프레임에서의 얼굴의 상태로 결정한다.

4. 실험 결과

실험은 실내 감시환경을 가정해 실제 CCTV 카메라로 자체 제작한 비디오 시퀀스에 대해 실시하였다. 비디오 시퀀스의 프레임 사이즈는 640×480 픽셀이고 총 204 프레임으로 구성된다. 그림 3 에서 볼 수 있듯이 비디오 시퀀스에서 시간이 지남에 따라 얼굴 모습(appearance)의 변화가 발생한다. 구체적으로, 출입문에 들어오기 이전에 어두운 조명의 얼굴이 출입문에 들어온 이후 급격히 밝아진다. 또한 등장인물이 CCTV 카메라에 근접함에 따라 카메라가 얼굴을 바라보는 각도가 달라지고 얼굴 틀어짐(head tilting)이 발생한다.

파티클 필터 기반의 얼굴 추적 방법에 가장 많이 사용되고 있는 칼라 히스토그램[4]과 LBP 특징을 3 장에서 제안한 얼굴 텍스처 특징과 비교하였다. 칼라 히스토그램은 적색, 녹색, 청색에 대해 각각 5 개씩, 총 15 개의 빈으로 구성하였고, LBP 는 제안하는 방법과 동일하게 16×16 픽셀로 리사이징 후 P 와 R 을 각각 4 와 1 로 두고, 8×8 픽셀 크기의 블록으로 나누어 추출하였다. 얼굴 추적 성능 비교를 위해 그림 3 과 같이 얼굴 추적 결과를 도시하고, 얼굴의 실제 위치(ground truth) 정보에 대한 중심 위치 오차(center location error)를 측정하였다. 측정 결과는 그림 4 에

감사의 글

이 논문은 2014 년도 정부(교육부)의 재원으로, 한국연구재단의 지원을 받아 수행된 기초연구사업임(No. 2011-0011383).

참고문헌

- [1] M. S. Arulampalam, et al. "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Processing*, Vol. 50, No. 2, pp. 174-188, 2002.
- [2] T. Ojala, et al. "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 7, pp. 971-987, 2002.
- [3] H.-I. Kim, et al. "Face detection for low power event detection in intelligent surveillance system," *IEEE International Conference on Digital Signal Processing (DSP)*, pp. 562-567, 2014.
- [4] D. Kim, et al. "A Comparative Study of Facial Features Robust to Varying Environment in Particle Filter-based Face Tracking," *Proceedings of Korea Multimedia Society* Vol. 17, No. 2, pp. 93-96, 2014.



그림 1 얼굴 추적 결과 비교 영상. # 뒤에 숫자는 비디오 시퀀스에서 프레임 번호에 해당함

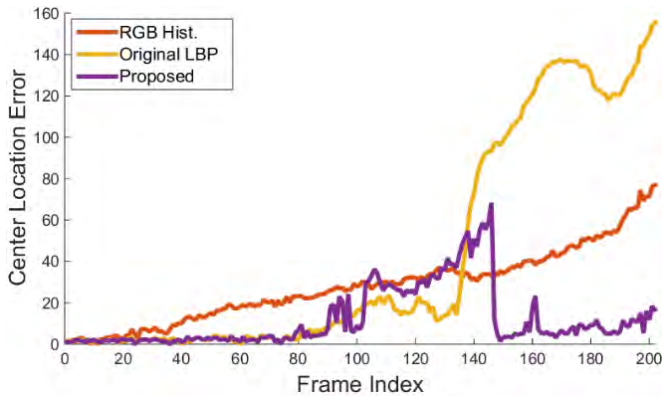


그림 2 중심 위치 오차 그래프 비교

나타냈다. 얼굴의 초기 상태가 어두울 때 특징을 추출함에 따라, 칼라 히스토그램 특징을 사용한 경우 머리카락 쪽으로 드리프트가 된 채 추적되어 중심 위치 오차가 증가하였다. LBP 특징은 약간의 얼굴 틀어짐에 의해 배경 쪽으로 드리프트 되어 매우 급격하게 중심 위치 오차가 증가하였다. 한편, 제안하는 방법에서는 비교하는 두 가지의 특징들에 비해 드리프트의 폭이 적어서 작은 중심 위치 오차를 유지하면서 효과적으로 얼굴 추적을 수행하였다. 한편, 제안하는 특징을 비롯한 모든 특징들은 초당 평균 27 프레임에서 30 프레임 사이의 빠른 처리 속도를 보였다.

5. 결론

본 논문에서는 기존의 파티클 필터 기반 얼굴추적의 강인함을 증대시키기 위해 새로운 얼굴 텍스처 특징을 제안하였다. 제안 방법은 인접한 픽셀의 평균(average)을 이용한 조명에 강인한 텍스처 패턴 추출과 얼굴의 구조적 정보를 반영하여 포즈에 강인한 텍스처 패턴 풀링(pooling)방법을 사용하였다. 실험결과 제안하는 얼굴 텍스처 특징이 비교하는 얼굴 특징에 비해 추적 정확도가 높았고 적은 차원 수에 의해 얼굴 추적 속도가 매우 빨랐다.