

다중 크기 블록 지역 이진 패턴을 이용한 랜덤 포레스트 기반의 머리 방향 분류 기법

강민주, 이하연, *강제원
이화여자대학교

joo2950@ewhain.net, lhy94@naver.com, *jewonk@ewha.ac.kr

Head Pose Classification using Multi-scale Block LBP and Random Forest

Minjoo Kang, Hayeon Lee, *Je-Won Kang
Ewha Womans University

요 약

본 논문에서는 다중 크기 지역 이진 패턴(Multi-scale Block LBP, MB-LBP) 특징과 랜덤 포레스트에 기반한 새로운 기법의 머리 방향 분류 기법을 제안한다. 제안 기법에서는 occlusion 과 조명의 변화에 강인한 분류 정확도를 얻기 위해서 랜덤화된 트리를 학습하는 것을 목표로 한다. 우선, 얼굴 이미지로부터 많은 MB-LBP 특징을 추출하고, 얼굴 영상들을 랜덤하게 입력하고 MB-LBP 크기 파라미터와 같은 랜덤 특징과 블록 좌표들을 사용하여 트리를 생성한다. 게다가 각 노드에서 정보 이득을 최대화 하는 트리의 내부 노드를 생성하기 위해서 uniform LBP 의 특성을 고려한 분할 함수를 개발한다. 랜덤화된 트리는 랜덤 포레스트에 포함되어 있으며 마지막 결정단계에서 Maximum-A-Posteriori criterion 으로 최종 결정을 한다. 실험 결과는 제안 기법이 다양한 조명, 자세, 표현, occlusion 상황에서 기존의 방법보다 개선된 성능으로 머리 방향을 분류 할 수 있음을 보여준다.

1. 서론

머리 방향 분류는 이미지나 비디오 시퀀스에서 사람의 머리의 방향을 나타내는 기법으로 얼굴 분석 어플리케이션에서 전처리 알고리즘으로 중요하다. 방향 예측을 위하여 다양한 이미지 특징 서술자들이 얼굴 데이터와 관련 있는 특징들을 표현하도록 고려되었다. Fanelli et al 은 머리 방향 추정을 위해서 이미지의 깊이 지도를 사용하는 알고리즘을 제안했다[1]. 이러한 접근법은 우수한 성능을 보이는 반면 2 차원 이미지를 사용하는 실제 시나리오에서는 이용하기 어렵다는 단점이 있다. LBP 기반의 특징들은 다양한 조명 변화[2]에서도 안정적인 성능을 보여 머리 방향 추정에 많이 사용되었다. Ma et al 은 Local Gabor Binary Pattern[3]을, Kim et al 은 Binary pattern run length matrix[4]를 특징으로 사용했다. Moore et al 은 LBP 특징들을 사용해 다 시점에서의 얼굴 표현들을 나타냈다[5]. 그러나 LBP 기반의 특징은 얼굴에 액세서리나 안경과 같은 부분적인 occlusion 이 있을 때 성능을 상당히 저하시키는 단점이 있었다.

랜덤 포레스트는 결정 트리의 집합으로 기계학습 기법 중에서도 분류에 효과적이라고 알려져 있으며[6] 컴퓨터 비전 어플리케이션에서 방향 추정 문제[7-9]에도 적용되었다. 본 논문에서는, 다중 클래스 문제와 run-time 효율성 측면에서의 정확도와 얼굴 occlusion 에 강인한 개선된 LBP 특징 문제를 해결하기 위하여 랜덤 포레스트를 기반의 방향 추정 알고리즘을 제안한다. 이전의 연구와 달리 랜덤 포레스트에서

트리의 각 내부 노드에 사용되는 효율적인 분할 함수를 개발하고, uniform LBP 특징을 랜덤 하게 선택함으로써 중요한 얼굴 패턴을 구분한다. MB-LBP 특징들은 지역 잡음이나 occlusion 에 너무 민감해서 부정확한 추정을 할 수 있으므로 각 노드에서 결정 트리를 생성할 때 특징 추출에서 단계에서 사용되는 다 블록 사이즈를 랜덤 하게 선택된 인자로 고려한다.

본 논문의 구성은 다음과 같다. 2 절에서는 LBP 특징을 살펴보고, 3 절에서는 MB-LBP 를 사용한 제안 알고리즘을 설명한다. 4 절에서는 본 논문에서 제안하는 기법을 설명하고, 5 절에서는 본 논문의 결론을 맺는다.

2. 지역 이진 패턴

2.1 다중 크기 블록 지역 이진 패턴

기존 LBP 연산자는 3 x 3 크기의 블록을 2 진 string 으로 부호화한다. 먼저 중심 픽셀과 주변 픽셀을 비교하여 주변 픽셀이 크면 1, 작으면 0 으로 변환하여 8bit 이진 스트링을 생성한다. LBP 가 개별의 픽셀을 비교하는 MB-LBP 연산자[10]는 sub-blocks 의 평균 값을 비교하여 블록의 multi-resolution 을 해석한다.

2.2 Uniform LBP 다중 크기 블록 지역 이진 패턴

Uniform LBP 는 최대 2 번의 bit 단위의 0->1 변화 혹은 1->0 의 변화가 발생하는 이진 패턴들을 말한다. Uniform

이 논문은 2016 년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업 (No.2014R1A1A22056587)

LBP 는 제안 기법과 관련하여 다음과 같은 유용한 특성들이 있다. 첫째로, 패턴들은 공간적으로 작은 구조에 대한 템플릿처럼 행동한다. 예를 들면 밝은 부분은 0, 경계나 코너는 1~7, 평평한 영역은 8 과 같은 9 의 템플릿이 있을 수 있다. 둘째로 uniform LBP 패턴이 다수의 텍스처에서 나타난다[11]. 따라서 uniform LBP 를 사용하면 전체 LBP 를 사용하는 것보다 특징 차원을 감소 할 수 있다. 셋째로 uniform LBP 는 회전에 강인하기 때문에 유사한 uniform LBP 는 upper bound 와 lower bound 의 2 개의 제약 파라미터로 압축적으로 표현될 수 있다.

3. 제안기법

머리 방향 분류는 이미지 내에서의 얼굴 검출과 방향 추정 의 두 단계로 나눌 수 있지만, 본 논문에서는 후자의 문제만 다루기로 한다. 즉 머리는 이미 검출되어 영상의 중심으로 정렬되었다고 가정한다. 학습할 얼굴 샘플들은 이미 정의된 yaw 방향이 머리 방향 각도(ex, -90, -75, ..., 0, ..., 75, 90)가 라벨링 되어 있다.

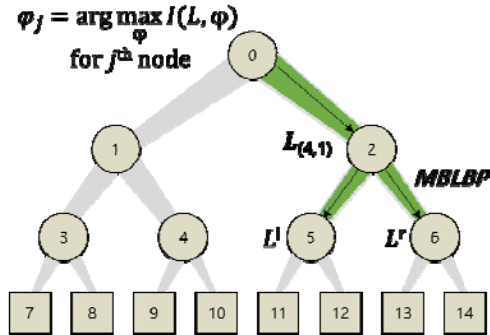


그림 1. 각 노드에서 샘플을 학습하며 구성되는 랜덤포레스트의 예시

그림 1 에서 볼 수 있듯이, 제안 기법에서는 각 노드에서 최적화된 분할 함수 파라미터로 정보이득을 최대화 하도록 랜덤 트리를 학습한다. 결정 단계에서 각 트리들은 랜덤 포레스트에서 함께 사용된다.

트리 **T** 는 먼저 전체 샘플 중 랜덤 하게 샘플들을 선택하고, 선택한 이미지 ($W \times H$)에서 겹치지 않는 sub-block 에 대해 계산된 scaled LBP 특징의 집합을 사용하여 학습된다. i 번째 MB-LBP 특징 샘플을 $L_{s,k}(i)$ 라 하고 s 는 MB-LBP 특징 공간의 크기 파라미터로서 sub-block 의 사이즈이고, k 는 sub-block 의 인덱스이다. s 는 4, 12, 36 이 될 수 있으며, 그에따라 제안 기법에 적용된 3 개의 MB-LBP 특징 공간이 생성된다.

현재 랜덤 특징들을 왼쪽 자식 노드 혹은 오른쪽 자식 노드로 나누기 위해서 각 노드에서의 분할 함수 $h_\phi(L_\phi)$ 을 정의한다. ϕ 는 (s, k, τ_u, τ_l) 로 정의된 모든 분할 파라미터의 공간이다. τ_u 와 τ_l 은 2 개의 제약 문턱 값이다. 노드에서 입력 L 을 2 진수 출력 (1,0)으로 사상시키고 왼쪽 자식

노드로 보낼지 오른쪽 자식 노드로 보낼 지 결정한다. 이 때 $h_\phi(L_\phi)$ 는 다음과 같다.

$$h_\phi(L_\phi) := \begin{cases} 1, & \tau_l < L_\phi < \tau_u \\ 0, & \text{otherwise} \end{cases} \dots(1)$$

식 1 에서 문턱값은 MB-LBP 의 10 진수 표현으로 upper bound 와 lower bound 이며 최대 2 bit 변환이 있기 때문에 유사한 uniform LBP 를 압축적으로 군집시키기에 충분하다. 트리를 최적화 하는 기법은 일반적인 랜덤 포레스트 최적화 기법을 따랐다[6].

테스트 과정에서는 이전에 보지 않은 샘플들이 트리의 뿌리 노드를 지나 학습된 내부 노드를 통과해 최종적으로 잎사귀 노드에 도착한다. 각 노드에서 학습된 분할 함수는 테스트 샘플들을 분할 함수의 만족 유무에 따라 왼쪽 자식 노드 혹은 오른쪽 자식 노드로 보낸다. 샘플이 잎사귀 노드에 도착하면 샘플의 머리 방향이 추정된다. 최종적으로 모든 트리의 잎사귀 노드에서 가장 많이 투표를 받은 클래스로 결정된다.

4. 실험 결과

4.1 실험 구성과 데이터 셋

우리는 다양한 얼굴 방향, 조명 조건, 얼굴 표현을 포함하고 있는 CMU Multi-PIE head pose image database [11]에 대해서 다양한 조명과 얼굴방향을 포함하는 이미지 셋을 이용하여 제안 기법의 성능을 평가하였다.

랜덤 트리를 학습하기 위해서 트리의 최대 깊이는 9, 트리의 학습을 멈추기 위한 각 노드의 최소 샘플의 수는 5 로 설정하였고, 최대 15 개의 트리를 학습하여 랜덤 포레스트를 생성하였다. 학습에서 5-fold 교차 검증들 사용하여 과 적합을 방지하였다. 또한 전체 데이터 셋에서 랜덤하게 샘플들을 선택하여 한 트리를 학습하는 bagging 을 사용하였다. 모든 실험은 Intel i7 @3.60GHz CPU and 8GB memory 실험 환경에서 진행되었다.

4.2 성능 평가

표 1 에서서 서로 다른 머리 방향들의 클래스의 수에 따른 제안 기법의 예측 정확도를 보여준다. 제안 기법과 비교하기 위해서 “Original LBP” 와 “MB-LBP” 에 대해 실험하였다. “Original LBP” 와 “MB-LBP” 는 오직 original LBP 혹은 MB-LBP 만을 사용하였으며, 제안기법과 달리 분할 함수에서의 문턱 값이 1 개만 존재한다.

제안 기법은 서로 다른 수의 머리 방향에 대해 실험했을 때 기존 알고리즘 기법 대비 개선된 예측 정확도를 보여주고 있다. 예를 들면, 제안 기법은 95%까지 분류 정확도를 제공한다. 분류하는 머리방향에 대한 클래스의 수가 증가하면 성능이 단조적으로 감소함에도 불구하고 제안 알고리즘의 성능이 95~74%로 가장 안정적이고 MB-LBP 는 90.2%~52%, original LBP 는 75.8%~28% 의 상대적으로 낮은 성능을 보여준다. 클래스의 개수가 3 개 혹은 5 개 일 때는 MB-LBP 성능이 제안 기법에 비교할 만 하지만 클래스의 개수가 7 개 혹은 9 개로 늘어날수록 약 7~22%로 성능 차이가 더 커짐을 확인할 수 있다.

단위 (%)		3 poses	5 poses	7 poses	9 poses
Original LBP	(1)	75.8	63.7	43	28
	(2)	77.3	63.9	41.6	29.6
Mb-LBP	(1)	90.2	82.2	75	52
	(2)	87.7	81.3	63.9	48.5
Proposed Algorithm	(1)	95	87.2	82	74
	(2)	93.5	88.9	78.3	67

표 1. 클래스의 수에 따른 제안 기법과 기존 기법의 머리 방향 분류 정확도

앞에서 언급한대로, 몇몇의 얼굴 샘플은 안경을 쓰고 있거나 머리카락으로 얼굴이 가려져 있는 등의 occlusion 이 존재한다. 제안 기법이 occlusions 에도 안정적이라는 것을 증명하기 위해서 우리는 데이터베이스를 occlusion 이 있는 데이터와 없는 데이터로 나누어 실험을 진행해 보았다. (1)이 표기된 실험은 occlusion 이 없는 데이터로 실험한 결과이며 (2)가 표기된 실험은 occlusion 이 있는 데이터로 실험한 결과이다. 제안 알고리즘의 정확도는 비교 알고리즘과 비교하여 occlusion 에 더 강인하다는 것을 결과를 통해 확인 할 수 있다. (2) 데이터에 대해서 클래스의 개수가 3 일 때와 9 일 때의 성능 차이는 26.5%이다. 하지만 original LBP 와 MB-LBP 로 실험했을 때는 47.7%, 39.2%로 점점 더 큰 성능 차이가 발생한다.

5. 결론

본 논문은 랜덤 포레스트와 다중 스케일 블록 지역 이전 패턴 특징을 기반의 효과적인 머리 방향 예측 기법을 제안하였다. 제안 기법에서 랜덤 트리는 다양한 조명 조건 과 얼굴 occlusions 에서 정확한 분류 성능을 제공하기 위해 학습되었다. 본 논문에서는 정보 이득을 최대화 하기 위해 트리의 각 내부 노드를 위해 사용된 분할 함수를 개발하였다. 실험 결과에서는 제안 기법이 다양한 조건의 조명과 occlusion 에서도 상당히 개선된 95~74%의 분류 정확도를 제공함을 확인하였다.

6. 참고문헌

[1] G. Fanelli, J. Gall, and L. Van Gool, "Real time head pose estimation with random regression forests," , in Computer Vision and Pattern Recognition (CVPR). IEEE, 2011, pp. 617- 624.

[2] D. Huang, C. Shan, M. Ardabilian, Y. Wang, and L. Chen, "Local binary patterns and its application to facial image analysis: a survey," , Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, vol. 41, no. 6, pp. 765- 781, 2011.

[3] B. Ma, W. Zhang, S. Shan, X. Chen, and W. Gao, "Robust head pose estimation using lgbps," , in International Conference on Pattern Recognition. IEEE,

2006.

[4] M. Sohn D. Kim H. Kim, S. Lee, "Illumination invariant head pose estimation using random forests classifier and binary pattern run length matrix," , Human-centric Computing and Information Sciences, 2014.

[5] S. Moore and R. Bowden, "Local binary patterns for multi-view facial expression recognition," , Computer Vision and Image Understanding, vol. 115, no. 4, pp. 541- 558, 2011.

[6] Leo Breiman, "Random forests," , Machine learning, vol. 45, no. 1, pp. 5- 32, 2001.

[7] Y. Li, S. Wang, and X. Ding, "Person-independent head pose estimation based on random forest regression," , in Image Processing (ICIP), 2010 17th IEEE International Conference on. IEEE, 2010, pp. 1521- 1524.

[8] S. Min, P. Kohli, and J. Shotton, "Conditional regression forests for human pose estimation," , in Computer Vision and Pattern Recognition (CVPR). IEEE, 2012, pp. 3394- 3401.

[9] J. Gall, A. Yao, N. Razavi, L. Van Gool, and V. Lempitsky, "Hough forests for object detection, tracking, and action recognition," , IEEE Trans. on Pattern Analysis and Machine Intelligence, 2011.

[10] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," , IEEE Trans. on Pattern Analysis and Machine Intelligence, 2002.

[11] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker, "Multi-pie," , Image and Vision Computing, vol. 28, no. 5, pp. 807- 813, 2010.