

거리 정보를 이용한 빠른 얼굴검출방법

이초일, 김병수, 김희율
한양대학교

cilee@vision.hanyang.ac.kr, bskim@vision.hanyang.ac.kr, wykim@hanyang.ac.kr

Rapid face detection using depth information

Cho-Il Lee Byeoung-su Kim Whoi-Yul Kim
Hanyang University

요 약

얼굴검출기술의 발전으로 인하여, 다양한 분야에 얼굴 검출을 활용한 기술이 이용되고 있다. 최근 Viola 와 Jones 의 얼굴검출 방법이 신뢰도 있는 검출률과 빠른 연산속도로 인하여 주로 이용되고 있다. 하지만 고해상도 이미지와 제한된 하드웨어를 사용하는 시스템의 경우, 실시간 처리가 어려워지는 문제가 있다. 본 논문에서는 이와 같은 문제를 해결하고자 거리 정보를 이용한 빠른 얼굴검출방법을 제안한다. 속도 개선을 위해 먼저 거리 정보를 이용하여 영상의 불필요한 부분을 제거하고, 피부색상정보를 이용하여 관심영역을 설정한다. 또 크기에 대응하기 위해 피라미드 이미지를 이용하는 방법 대신, 거리 정보를 이용하여 얼굴의 크기를 추정한다. 마지막으로 검색창 내의 거리 분산을 계산하여, 평평하거나 굴곡이 심한 영역을 제거함으로써 얼굴 검출 속도를 개선하였다. 실험결과 기존 방법에 비해 더 빠른 검출속도와 유사한 검출성능을 확인할 수 있었다.

1. 서론

얼굴검출기법은 수십 년 동안 컴퓨터비전 분야에 중요한 연구분야로 활발히 연구되어왔으며, 다양한 분야에 널리 활용되어왔다. 주된 사용범위로는 얼굴 인식을 통한 개인 인증 보안, 디지털 카메라의 얼굴 포커싱, 무안경 3D 시스템, 운전자 졸음감지 등이 있다. 특히 최근에는 신뢰도를 요하는 금융 및 의료분야에 까지 적용되며, 그 응용범위가 점차 넓어지고 있다. 이에 따라, 빠르고 신뢰도 있는 얼굴검출에 대한 연구가 요구되고 있다.

얼굴 검출을 위해서 Viola 와 Jones 의 방법 [1]이 일반적으로 이용된다. 이 방법은 빠른 연산속도를 위해 단순 명암 값의 차이를 특징점으로 이용하고, 명암값을 누적한 적분영상으로 계산 복잡도를 감소시킨다. 또 얼굴이 아닌 영역을 초반에 빠르게 제거하여 연산속도를 향상시킨다. 이러한 특징 때문에 빠르고 신뢰도 있는 얼굴검출이 가능하게 된다. 하지만 고해상도 이미지에서는 제한된 성능의 하드웨어를 사용하는 시스템의 경우, 실시간 처리가 어려워진다.

실시간 얼굴 검출을 위한 저 복잡도 얼굴검출 방법으로는 피부색상영역에서만 얼굴을 검출하여 검출속도를 향상시키는 방법 [2], 대상영역의 에지성분을 검출한 후, 얼굴 모델과의 유사도를 비교하여 얼굴을 검출하는 방법 [3] 그리고 거리 정보에 따라 검색창의 크기를 변화하여 이미지 피라미드를 만들지 않고 검색하는 방법 [4] 등이 있다. 이러한 방법들의 경우, 일반적으로 좋은 얼굴 검출속도를 보여준다. 하지만 영

내 피부와 유사한 색상이 많거나 배경이 복잡할 경우, 검출속도가 개선되지 않거나 검출률이 감소하는 단점이 있다.

본 연구에서는 거리 정보를 이용하여 빠르게 얼굴을 검출하는 방법을 제안한다. 검출성능은 유사하면서 빠른 얼굴검출을 위해, 얼굴검출이 가능한 거리를 계산하여 불필요한 배경을 제거한다. 배경이 제거된 영상과 피부색상영역의 모폴로지 연산을 통해 관심영역을 설정한 후, 검색창 내의 거리 정보에 분산값을 검사한다. 이때 평평하거나 굴곡이 심한 영역을 빠르게 제거하여, 얼굴검출속도를 개선하였다.

본 논문의 2 절에서는 거리 정보를 바탕으로 빠르게 얼굴을 검출하기 위한 방법을 소개하였고, 3 절에서는 실험을 통하여 제안하는 시스템의 성능을 평가하였으며, 마지막으로 4 절에서 결론을 맺었다.

2. 제안하는 방법

본 논문에서는 빠르게 얼굴을 검출하기 위해 거리 정보를 활용하였다. 거리 정보를 바탕으로 불필요한 영역을 제거하고, 피부색상 정보를 이용하여 관심영역을 설정한다. 그리고 검색창안에서 거리 정보의 분산을 검사하여 얼굴이 아닌 영역을 제거한다. 마지막으로 다양한 얼굴크기에 대응하기 위해 검색창의 크기를 변화시키며 얼굴검출을 위한 분류기에 통과시켜 최종 결과를 얻는다. 그림 1 은 제안하는 방법의 전체 흐름도를 보여준다.

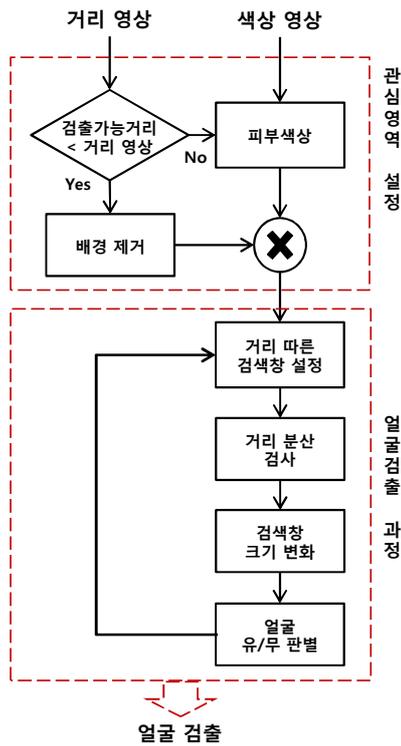


그림 1. 제안하는 방법의 흐름도

2.1. 관심영역 설정



그림 2. 색상 영상과 거리 영상

영상에서 얼굴검출에 불필요한 영역을 빠르게 배제할 경우, 탐색영역이 줄어들므로 검출속도가 향상 되며, 본 논문은 거리 정보를 활용하여 불필요한 영역을 제거한다.

영상에서 얼굴검출이 가능한 최대거리인 $ValidDist$ 는 영상 내의 얼굴크기가 훈련을 위해 사용한 얼굴영상의 크기와 동일해질 때의 거리로 구할 수 있으며, 다음 식으로 나타낼 수 있다.

$$ValidDist = f \times W_{size} / T_{size} \quad (1)$$

이때, f 는 카메라의 초점거리이며, W_{size} 는 평균얼굴에 가로길이, T_{size} 는 훈련에 사용한 얼굴영상의 가로픽셀이 된다. 검출 가능 최대거리를 벗어난 경우, 얼굴의 크기가 특징값을 추출하기 위한 최소크기보다 작아짐으로 얼굴검출을 할 수 없다. 그러므로, 얼굴검출이 가능한 최대거리를 벗어나는 영역을 제거하여도 얼굴 검출에 영향을 주지 않는다.

그림 3 의 (b)와 같이 불필요한 부분이 제거된 영역과

피부색상 영역이 겹치는 부분을 최종 관심영역으로 설정한다. 구해진 관심영역에서만 얼굴 유무를 판별함으로써 빠른 얼굴검출이 가능하게 된다.



(a) 피부색상 영역 (b) 최종 관심영역

그림 3. 관심영역 설정

2.2. 얼굴 검출

거리변화에 상관없이 얼굴을 검출하기 위해서는 영상의 크기를 조금씩 바꾼 여러 장의 이미지를 만들고 모든 이미지에 대해 검사를 수행하여야 한다. 하지만 거리 정보가 주어질 경우, 대략적인 얼굴의 크기를 알 수 있기 때문에 원본 크기의 이미지에서만 얼굴을 검출하면 된다. 대략적인 얼굴의 크기는 얼굴을 찾기 위한 검색창의 크기와 같으며, 다음 식을 통해 구할 수 있다.

$$I_{size} = f \frac{W_{size}}{Z} \quad (2)$$

이때, I_{size} 는 검색창의 가로 길이가 되며, Z 는 거리 값, f 는 카메라의 초점거리, W_{size} 는 일반적인 얼굴의 가로길이가 된다.

얼굴 검출을 위해 관심영역의 모든 위치를 이동하며, 거리 정보를 바탕으로 검색창의 크기를 계산한다. 검색창 내의 원본영상에서 특징 값을 추출한 후, 분류기를 통과하여 얼굴 유무를 판별한다. 분류기의 경우, 얼굴과 비 얼굴을 가장 잘 나누는 다수의 특징 값들로 이루어진다. 일반적으로 신뢰도 있는 분류기일수록 많은 수의 특징 값들로 구성되기 때문에 얼굴의 유무를 판별하기 위해 많은 연산이 요구된다. 이러한 연산을 줄이기 위해, 분류기를 통과하기에 앞서 검색창 범위 안에 거리의 분산을 먼저 검사한다. 계산의 효율성을 위해 먼저 적분영상을 만들고, 다음 식을 통해 분산을 계산한다.

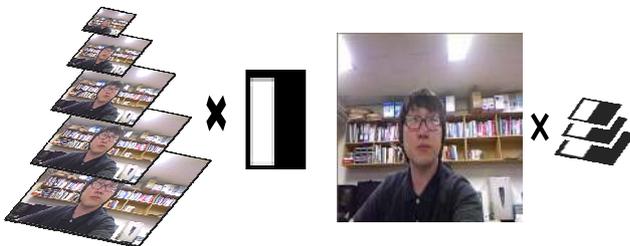
$$E[W] = \left(\frac{I[x, y] + I[x + H, y + V]}{-I[x + H, y] - I[x, y + V]} \right) / H \times V \quad (3)$$

$$Var[W] = E[W^2] - (E[W])^2 \quad (4)$$

이때, I 는 거리의 적분영상이며, H 와 V 는 검색창의 가로, 세로 길이가 된다. $E[W]$ 는 검색창 W 의 거리 평균값을

나타내며, $Var[M]$ 는 검색창 내 거리의 분산값을 나타낸다. 거리 분산값이 너무 작은 경우, 굴곡이 없는 물체임으로 얼굴이 아니라 판별하며 너무 큰 경우 역시 얼굴 유무 검사를 수행하지 않는다.

거리 정보로 계산된 검색창은 일반적인 얼굴의 크기를 기준으로 구해졌기 때문에, 얼굴이 작은 사람이나 큰 사람일 경우 정확하게 얼굴에 맞지 않는다. 때문에 계산된 검색창의 크기에서 일정범위의 크기 변화를 주면서 얼굴을 판별하여야 얼굴 크기변화에 무관하게 얼굴을 검출 할 수 있게 된다. 본 논문에서는 그림 4 의 (b)처럼, 계산된 검색창의 크기에서 $\pm 25\%$ 의 범위를 9 개의 크기로 나누어서 각 검색창마다 분류기를 통과한다. 이때 분류기는 Viola&Jones 의 방법과 동일 한 방법을 사용하였다. 최종적으로 분류기를 통과하여 검출된 결과의 무게중심이 유사한 것끼리 하나로 묶어서 얼굴 유무를 판별하게 된다.



(a) Viola&Jones 의 방법 (b) 본 논문의 방법
그림 4. 얼굴크기변화에 대응하기 위한 방법

3. 실험 결과

최근 구조광 방식의 kinect 카메라 [6]에 상용화로 손쉽게 고해상도의 거리 정보를 얻을 수 있게 되었다. 본 논문에서도 거리 정보와 색상 정보를 동시에 얻기 위하여, Kinect 카메라를 이용하였다. 실험을 위해 사용된 프로그램은 Visual studio C++을 이용하여 작성되었고, 실험 환경은 Intel dual core CPU 2.5Ghz, 4GB RAM, Windows XP 에서 실험 되었다. 실험영상은 카메라를 바라보며 얼굴의 위치를 변화시키는 640x480 크기에 동영상상을 이용하였다.

알고리즘의 속도를 비교하기 위해, 실험영상을 이용해 얼굴을 검출하는데 까지 Viola&Jones 의 방법과 제안하는 방법 걸리는 시간을 측정하였다. 그림 5는 속도비교 결과이다.

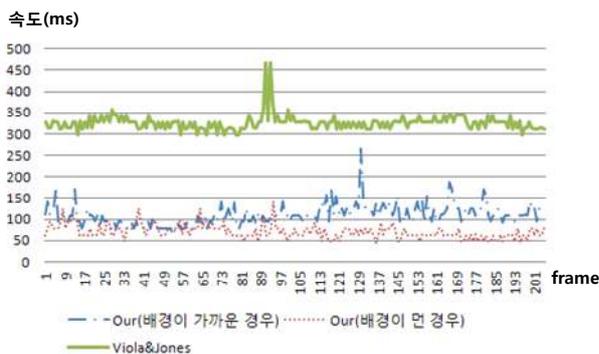


그림 5. 속도 비교 결과 그래프

<표 1> 속도 측정 결과

방법		속도(ms)
Viola&Jones		329.53
제안하는 방법	배경이 가까운 경우	109.28
	배경이 먼 경우	66.94

표 1 에서 볼 수 있듯이, 제안하는 방법은 배경이 얼굴검출 가능 거리보다 멀리 위치할 때 평균 66.94ms 의 속도가 소요되었고, 배경이 얼굴검출 가능거리에 있을 경우 109.28ms 가 소요됐다. 반면 Viola 와 Jones 의 방법은 329.53ms 가 소요되었으며 제안하는 방법이 약 3~5 배 빠른 속도를 보여주었다. Viola 와 Jones 의 경우, 640x480 크기의 이미지에서 모든 얼굴 크기에 대응하기 위해서는 25%씩 크기를 줄이며 만든 19 개의 이미지를 모두 검색하여야 한다. 반면 제안하는 방법의 경우, 거리 정보를 바탕으로 얼굴의 크기를 추정 할 수 있기 때문에 추정된 얼굴 크기를 기준으로 $\pm 25\%$ 범위만 검사하여 얼굴을 검출한다. 또 거리 정보를 이용하여 관심영역을 설정하고 거리분산검사를 수행하여 빠른 연산이 가능하였다.

얼굴검출 성능을 비교하기 위해 검출률과 false alarm 을 측정하였고, 정확도(precision)과 재현율(recall)을 측정한 후, f-measure 를 측정하였다. F-measure 는 정확도와 재현율의 조화 평균값을 나타내며, 검출성능을 측정할 때 정량적 수치로 이용된다[7]. 측정 결과는 그림 6 과 표 2 와 같다.

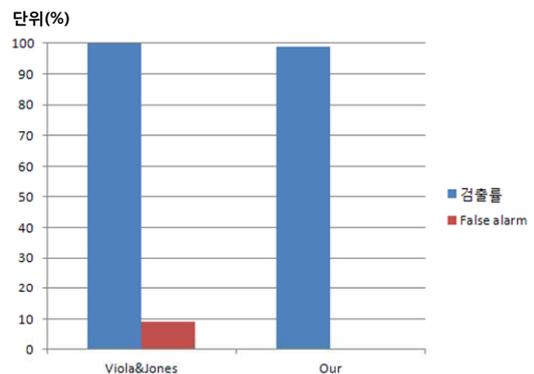


그림 6. 검출률과 false alarm 비교 결과

<표 2> 검출성능 측정 결과

	Viola&Jones 의 방법	제안하는 방법
Precision (%)	91.6	99.8
Recall (%)	99.9	98.7
F-measure (%)	95.6	99.2

검출률의 경우, 99.9%와 98.7%로 Viola&Jones의 방법이 제안하는 방법에 비하여 조금 더 높은 결과를 보였으나, false alarm의 경우, 9.08%와 0.33%로 제안하는 방법이 더 좋은 성능을 보였다. 제안하는 방법의 경우, 거리에 따라 가능한 얼굴의 크기와 거리분산 범위를 만족해야 얼굴을 검출하기 때문에 false alarm이 줄어든 것을 확인할 수 있다. 마지막으로 f-measure를 이용한 측정결과, 제안하는 방법이 Viola&Jones의 방법보다 더 좋은 성능을 보임을 확인할 수 있었다.

Wiley-Interscience Publication, pp. 164-174, 2001.

4. 결론 및 향후 과제

본 논문에서는 거리 정보를 이용하여 빠른 얼굴검출방법을 제안하였다. 빠른 검출속도를 위해, 거리 정보와 색상정보를 활용하여 관심영역을 설정하였다. 또 거리 정보를 이용하여 얼굴에 맞는 탐색창의 크기를 정하였고, 거리 분산을 검사하였다. 실험 결과 Viola&Jones의 방법과 유사한 얼굴 검출 성능을 보이면서, 약 3~5 배 빠른 동작속도를 보였다. 또 유사한 검출률과 더 낮은 false alarm 결과를 확인할 수 있었다.

거리 정보를 활용할 경우, 검출을 위한 다양한 정보를 얻을 수 있으므로 성능을 향상 시킬 수 있다는 것을 확인하였다. 뿐만 아니라 거리 정보는 조명 및 색상에 강인하기 때문에 인종과 조명에 독립적인 얼굴 검출방법을 구현 할 수 있다.

향후 피부색상정보를 사용하지 않고 관심영역을 설정할 수 있는 방법에 대한 연구가 필요할 것이며, 거리 정보를 이용한 인종과 조명에 강인한 얼굴 검출 알고리즘에 대한 연구를 진행할 예정이다.

참고문헌

- [1] P. Viola, M. Jones, "Robust real-time face detection," *Computer Vision*, vol. 2, pp. 747-747, 2001.
- [2] R.-L. Hsu, M. Abdel-Mottaleb, and A.K. Jain, "Face detection in color images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 696-707, 2002.
- [3] K.J. Kirchberg, O. Jesorsky, R.W. Frischholz, "Genetic model optimization for hausdorff distance-based face localization," *Lecture notes in computer science*, pp. 103-111, 2002.
- [4] H. Wu, K. Suzuki, T. Wada and Q. Chen, "Accelerating face detection by using depth information," *PSIVT2009*, pp. 657-667, 2009.
- [5] E. Parzen, "On estimation of a probability density function and mode," *Annals of Mathematical Statistics*, vol. 33, pp. 1065-1076, 1986.
- [6] Jeffrey L. Wilson, "Microsoft kinect for Xbox 360," *PCMag.com*, Retrieved November 10, 2010.
- [7] Richard O. Duda, Peter E. Hart and Davie G. Stork, "Pattern classification second edition," A