거리순위를 이용한 얼굴검출

박재희*, 김성대**
한국과학기술원 전자전산학과 전기및전자공학전공

Face Detection using Distance Ranking

Jae Hee Park* and Seong Dae Kim**
Div. of EE, Dep. of EECS, Korea Advanced Institute of Science and Technology
E-mail : *yscclo@sdvision.kaist.ac.kr, **sdkim@ee.kaist.ac.kr

Abstract

In this paper, for detecting human faces under variations of lighting condition and facial expression, distance ranking feature and detection algorithm based on the feature are proposed. Distance ranking is the intensity ranking of a distance transformed image. Based on statistically consistent edge information, distance ranking is robust to lighting condition change. The proposed detection algorithm is a matching algorithm based on FFT and a solution of discretization problem in the sliding window methods. In experiments, face detection results in the situation of varying lighting condition, complex background, facial expression change and partial occlusion of face are shown

I. 서론

얼굴검출이란 주어진 영상 내에 사람의 얼굴이 존재하는지를 확인하고, 존재할 경우 그 크기와 위치를 찾는 일련의 과정을 의미하며, 화상감시, Human Computer Interface 등의 분야에서 널리 활용되고 있다[1].

많은 얼굴검출기법은 이동 창(sliding window) 기법을 이용한다. 이동 창 기법이란 찾고자 하는 얼굴크기의 창(window)을 정하고 그 창을 영상 내에서 이동시키면서 창에 보여지는 영상영역이 얼굴인지 아닌지를 확인하는 방법이다. 살색 정보 등 부가적인 정보를 사용하지 않을 경우, 크기가 일정하지 않은 얼굴을 검출하기 위해서는 다른 크기의 창을 이용한 별도의 독립적인 검출과정을 수행하게 된다. 이때, 이동 창의 이동간격, 크기변화 간격이 작을 경우에는 검출소요시간이 매우 길어지며, 간격이 클 경우에는 정확한 얼굴검출이 불가능하다. 이러한 문제를 이산화문제 (discretization problem)라고 한다.

얼굴검출을 위한 특징(feature)은 크게 밝기기반 특징과 형태기반 특징으로 나누어 생각할 수 있다. 이때 형태기반 특징은 비교적 조명 등 밝기변화에 강인한 것으로 알려져 있으며, 대표적으로 영상의 에지(edge) 정보 등을 생각할 수 있다.

본 논문에서는 조명변화 및 얼굴의 표정변화에 강인한 얼굴검출기법을 제안한다. 제안기법은 이산화문제의 근본적인 문제인 많은 계산량을 효율적으로 줄인다.

II. 제안기법

2.1 거리변환(Distance Transform)

거리변환이란 이진영상을 거리영상으로 변환하는 과정을 말한다[2]. 여기서 이진영상은 그림 1 (a)와 같이 에지 화소와 에지가 아닌 화소로 이루어져 있으며, 거리영상은 각 화소의 값이 가장 가까운 에지 화소까지의 거리값을 나타내는 영상이다. 그림에서와 같이 거리영상에서 에지 또는 에지 화소로부터 가까운 화소는 작은 밝기값을 가지며, 먼 화소는 큰 밝기값을 가진다.

(a) (b)

그림 1 에지 이진 영상(a)과 대응 거리영상(b)

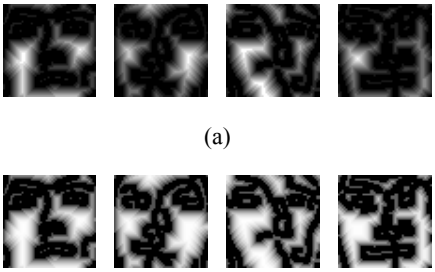
두 장의 에지 영상이 주어져 있을 때 두 영상 간의 유사도를 측정하기 위해서는 복잡한 계산과정이 필요하였다. 그러나 에지 영상을 거리영상으로 변환할 경우 기존의 영상기반 유사도 측정방법을 활용할 수 있다.

2.2 거리순위(Distance Ranking)

거리순위란 거리영상에 밝기순위 (intensity ranking) 특징추출 기법[3]을 적용한 것이다. 이때, 각 화소에 대한 거리순위(dr_{xy})값은 거리 영상 내 화소들 중 그 거리값(d_{lm})이 자신의 거리값(d_{xy})보다 같거나 작은 화소들의 비율로 정의되며 식(1)과 같이 계산된다.

$$dr_{xy} = \frac{1}{XY}\sum_{l=0}^{X}\sum_{m=0}^{Y} u(d_{xy} - d_{lm}), \quad u(t) = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0 \end{cases} \qquad (1)$$

where, X, Y: image size



(a)

(b)

그림 2 여러 얼굴에서 획득한
거리영상(a)과 대응 거리순위영상(b)

순위특징은 히스토그램 균등화(Histogram Equalization)의 개념으로 입력의 정규화를 위해 널리 사용되고 있다.

그림 2는 여러 얼굴에서 획득한 거리영상과 이에 대응되는 거리 순위영상을 나타내고 있다. 그림에서와 같이 거리영상에서는 조명환경 및 개인 차에 따라 발생하는 에지 검출 유무 및 에지의 위치에 따른 밝기값의 변화를 확인할 수 있지만, 순위영상의 경우 그러한 변화가 많이 줄어듦을 확인할 수 있다.

2.3 수정된 거리순위(Modified Distance Ranking)

거리순위는 비교적 얼굴의 형태 정보를 동질성 있게 잘 표현하는 특징이지만 일부 화소의 거리순위 변화가 다른 부분의 거리순위에도 크게 영향을 주는 단점을 가지고 있다. 특히, 조명의 방향이 극단적으로 변할 경우 사람의 코 부분으로부터 추출되는 에지의 존재유무 및 위치는 크게 변하게 되며, 그 결과 코 주위뿐만 아니라 다른 부분의 거리순위 역시 변하게 된다.

본 논문에서는 이러한 문제점을 해결하기 위하여 학습영상들로부터 얼굴 내 각 화소의 위치 별로 거리순위의 항상성을 계산하고 이를 기반하여 거리순위에 위치 별로 신뢰도를 부여한 수정된 거리순위 특징을 제안한다.

우선 학습영상들로부터 거리순위를 계산하고 학습영상들 내 (x,y) 에 위치한 화소들의 거리순위의 분산 $\sigma^2_{dr_{xy}}(x,y)$ 을 구한다. 이를 이용한 각 (x,y) 에서의 거리순위의 신뢰도, dw_{xy} 는 식(2)와 같다.

$$dw_{xy} = \alpha\left(\exp(-\sigma^2_{dr_{xy}}(x,y) \Big/ \sigma_{th}{}^2)\right) \qquad (2)$$

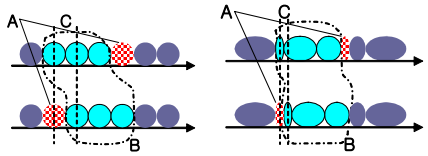
for const. $\alpha, \sigma_{th}{}^2$

이때 각 신뢰도 값은 $\sum_{l=0}^{X}\sum_{m=0}^{Y} dw_{lm} = 1$ 를 만족하도록 상수를 결정한다.

이때, 제안하는 (x,y) 에서의 수정된 거리순위 $\widetilde{dr}_{xy}$ 는 자신보다 거리값이 같거나 작은 화소들의 신뢰도의 합으로 정의되며 식(3)과 같다.

$$\widetilde{dr}_{xy} = \sum_{d=0}^{d_{xy}} \sum_{\{(l,m)|d_{lm}=d\}} dw_{lm} \qquad (3)$$

그림 3(a)는 거리순위 특징에서 특정화소(A)의 순위변동에 따른 주변화소(B)들의 순위변동(C)을 나타낸 것으로 그림 3(b)와 같이 수정된 거리순위에서는 그 순위변동 폭을 줄일 수 있다.



(a) (b)

그림 3 거리순위(a) 및 수정된 거리순위(b)

2.4 거리순위 특징을 이용한 얼굴식별

사람의 얼굴은 눈, 코, 입 등 비슷한 곳에서 에지가 검출된다. 이를 이용하여 각 학습영상으로부터 수정된 거리순위를 예측하고, 실제 입력이 예측된 순위특징에 잘 부합하는가를 확인하여 후보영상들 중에서 사람의 얼굴을 식별해내고자 한다.

우선 크기 정규화되고, 수동으로 지정된 눈의 위치를 이용하여 정렬된 다수의 얼굴 학습영상들로부터 각 화소에서의 거리순위의 신뢰도를 계산하고, 각 학습영상의 수정된 거리순위를 획득한다. 그리고 획득한 수정된 거리순위의 화소별 순위값을 각각 N 개 등급으로 균일하게 양자화(quantization)하여 각 등급(level)이 포함하는 화소의 위치를 구한다. 이때, n 번째로 큰 순위를 가지는 등급의 화소들을 모은 n 번째 거리등급 벡터를 각각 $\mathbf{dLv}_n$ 이라 하고 다음과 같이 정의한다.

$$\mathbf{dLv_n} = (dLv_{n00}, dLv_{n10}, ..., dLv_{nXY})^T$$
$$dLv_{nxy} = \begin{cases} dw_{xy} & n-1/N \le \widetilde{dr}_{xy} \le n/N \\ 0 & otherwise \end{cases} \quad (4)$$
where, $1 \le n \le N$

그림 4 거리등급 영상의 예 ($N=5$ 인 경우)

입력영상의 거리변환 영상 $\mathbf{D} = (d_{00}, d_{10}, ..., d_{XY})^T$ 과 거리등급 벡터와의 내적으로 각 등급에 해당되는 후보영상 내 영역에서의 평균거리값(dm_n)을 구할 수 있다.

$$dm_n = \mathbf{dLv_n} \cdot \mathbf{D} \quad (5)$$

만약 입력영상이 얼굴이라면 두 자연수 $a, b\,(a>b)$ 에 대하여 $dm_a > dm_b$ 이어야 한다. 만약 높은 등급영역의 평균 거리값이 낮은 등급영역의 평균값보다 더 작다면 얼굴이 아닐 가능성이 크다. 따라서 본 논문에서는 평균 거리값의 크기가 순위역전 된 양을 후보영상과 모델과의 거리에러로 정의하고 식(6)과 같이 계산하였다.

$$der = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} (dm_j - dm_i) u(dm_j - dm_i) \quad (10)$$

2.5 FFT 를 이용한 빠른 얼굴검출

고정된 크기의 얼굴을 검출하기 위하여 그림 5에서와 같이 왼쪽 위 꼭지점의 좌표가 (x,y) 이고 크기가 X_C, Y_C 인 후보영역을 앞 절에서 제안한 에러함수를 기반으로 얼굴 또는 비얼굴로 식별하고자 한다. 이때, 후보거리영상을 $\mathbf{C}^{xy} = \left(c^{xy}(0,0), ..., c^{xy}(X_C-1, Y_C-1)\right)^T$ 라 하면 $c^{xy}(m,n) = d(m+x, n+y)$ 인 관계가 성립한다. 여기서 $d(x,y)$ 는 입력영상의 (x,y) 에서의 거리값이다.

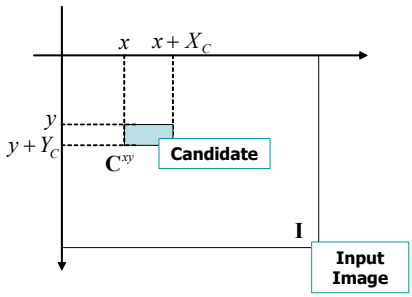


그림 5 (x,y)에 위치한 후보영역

이때, 후보영역 $\mathbf{C}^{xy}$ 에 대한 k 번째 등급의 평균 거리값 dm_k^{xy} 는 식 (11)과 같이 정의된 함수 $dlv_k^R(m,n)$ 을 이용하여 식 (12)와 같이 계산할 수 있다.

$$dlv_k^R(m,n) = \begin{cases} dlv_k(X_C-1-m, Y_C-1-n) & 0 \le m \le X_C-1 \,\&\, 0 \le n \le Y_C-1 \\ 0 & elsewhere \end{cases} \quad (11)$$

$$\begin{aligned} dm_k^{xy} &= \mathbf{dLv}_k \cdot \mathbf{C}^{xy} \\ &= \sum_{n=0}^{Y_C-1} \sum_{m=0}^{X_C-1} dlv_k(m,n) \times c^{xy}(m,n) \\ &= \sum_{a=-\infty}^{\infty} \sum_{b=-\infty}^{\infty} dlv_k^R(X_C-1+x-a, Y_C-1+y-b) \times d(a,b) \\ &= dlv_k^R(m,n) * d(m,n)\big|_{m=x+X_C-1, n=y+Y_C-1} \\ &= cv_k(x+X_C-1, y+Y_C-1) \end{aligned} \quad (12)$$
(* : convolution)

즉, dm_k^{xy} 는 거리변환된 입력영상과 $dlv_k^R(m,n)$ 와의 컨볼루션 결과 $cv_k(x,y) = dlv_k^R(x,y) * d(x,y)$ 에서 하나의 화소값으로 표현된다. $dlv_k^R(x,y)$ 와 $d(x,y)$ 에 충분한 0 삽입(zero-padding)을 가정할 때, 이산퓨리에변환(DFT) $\Im(\cdot)$ 및 역변환 $\Im^{-1}(\cdot)$ 을 이용하여 식 (13)과 같이 계산할 수 있으며 고속퓨리에변환(FFT) 알고리즘을 이용하여 빠른 시간 내에 함수 $cv_k(x,y)$ 를 획득할 수 있다. $\mathbf{C}^{xy}$ 의 얼굴에 대한 에러는 이와 같이 획득한 $cv_k(x,y)$ 를 이용하여 계산된다.

$$cv_k(x,y) = \Im^{-1}\left\{\Im\left(lv_k^R(x,y)\right) \times \Im\left(i(x,y)\right)\right\} \qquad (13)$$

이러한 과정을 통하여, 모든 후보영역에 대하여 별도의 특징추출 및 식별과정을 수행하지 않고, 입력영상을 거리변환한 후 $cv_k(x,y)\,(1 \le k \le N)$ 를 획득하여 검출 과정을 수행한다. 이는 인접한 후보영역 간의 계산의 중복성을 줄여 전체검출 과정을 빠르게 수행할 수 있게 한다.

III. 실험결과

제안하는 거리순위 특징 및 식별기법을 이용하여, 실제 얼굴 검출기를 구성하여 실험하였다.

검출 대상은 조명의 방향 및 밝기가 다양한 상황에서 촬영된 50×50(pixel^2) 이상의 크기를 가지는 사람의 정면 또는 준정면 얼굴을 포함한 320×240(pixel^2) 크기의 흑백영상이다. 이때 각 얼굴은 표정이 극단적으로 바뀌며, 손 등으로 얼굴의 일부가 가려진 경우를 다수 포함하였다. 학습영상은 50×50(pixel^2)의 총 4 명, 16 장의 정면얼굴 영상으로 구성하였으며, 가벼운 조명변화 및 표정변화를 포함시켰다.

실험은 50×50(pixel^2) 크기의 얼굴을 가정한 상태에서, 원 입력 영상에 대하여 수행한 후, 10%씩 입력 영상의 크기를 줄여가며 반복적으로 수행하였다.

그림 6은 다양한 조명 환경에서 촬영된 영상에 대한 얼굴 검출결과이다. 얼굴의 표정변화가 다양하거나 얼굴의 부분적인 가려짐이 있더라도 정확한 얼굴검출결과를 보인다. 또한 정확히 정면 얼굴이 아닌 상하좌우 방향의 자세변화에 대한 강인함을 관찰할 수 있었다. 크기가 고정된 입력영상에 대한 검출과정에는 약 0.2~0.9 초의 시간이 소요되었으며 입력영상의 크기변화를 고려한 전체검출과정에는 약 2.1 초의 시간이 소요되었다. 이는 기존의 전역특징 기반 얼굴검출기법의 처리속도에 비하여 3 배 이상 빠른 속도이다.

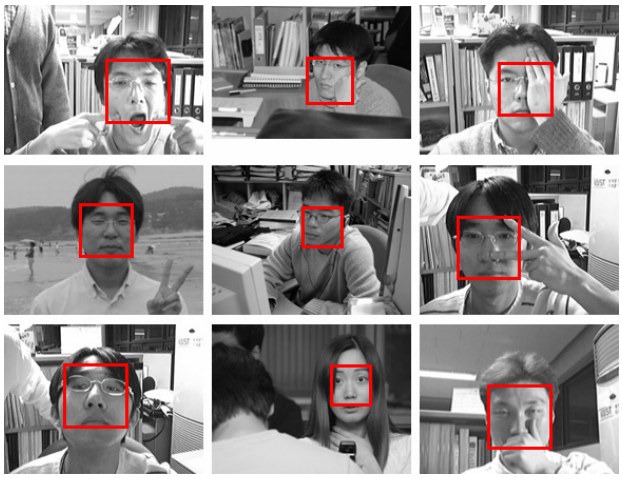


그림 6 얼굴검출 결과

IV. 결론

본 논문에서는 빠르고 정확한 사람의 얼굴 검출을 위하여 거리변환 영상에 대하여 밝기순위 개념을 도입한 거리순위 특징을 제안하였다. 그리고 제안한 순위특징에 적합한 양자화된 거리영역 기반 검출기법과 고속계산기법을 제안하였다. 실험에서는 제안기법이 빠르고, 얼굴의 가려짐에 현상에 강인하며, 기존의 얼굴검출기법들에 비하여 검출결과의 정확도가 우수함을 확인하였다. 제안기법은 비교적 높은 해상도의 영상을 활용할 수 있으며 빠른 처리속도와 높은 정확도가 요구되는 얼굴 인증(face authorization)등의 분야에서 효율적으로 사용될 수 있다.

참고문헌

[1] Jae Hee Park, Hae Chul Choi, and Seong Dae Kim, "Bayesian face detection in an image sequence using face probability gradient ascent," *Proceedings of ICIP 2005*, Genoa, Italy, pp. II-346-349, Sep. 2005

[2] Gunilla Borgefors, “Distance transformations in digital images,” *Computer Vision, Graphics, and Image Processing*, vol. 34, issue 3, pp344-371, 1986

[3] 박재희, 최학훈, 김성대, “밝기순위 특징을 이용한 적외선 정지영상 내 물체검출기법”, 대한전자공학회 논문지 제 42 권 SP 편 제 2 호, pp.37~48, 2005. 3