

영상압축이 얼굴인식에 미치는 영향

김창한, 김지훈, 이철희
연세대학교 전기전자공학부

Effects of image compression on face recognition

Changhan Kim, Jihoon Kim, Chulhee Lee
Dept. Electrical and Electronic Engineering, Yonsei University
E-Mail : gomajor@yonsei.ac.kr

Abstract - 얼굴인식은 카메라로부터 영상을 취득하여, 이미 등록되어 있는 영상 중에 가장 높은 유사성을 보이는 영상을 취득 영상의 대상자로 판단하는 것이다. 정확한 판단을 위해서 일반적으로 얼굴인식에 사용되는 영상은 취득 당시 영상의 정보를 모두 가지고 있는 영상, 즉 압축하지 않은 영상을 사용한다. 하지만 대용량 테이터 얼굴인식 시스템에서는 저장 공간의 이유로 영상에 압축을 해야 하는 상황이 발생할 수 있다.

본 논문에서는 영상압축이 얼굴인식에 미치는 영향에 대해서 고찰한다. 영상압축을 위해 사용되는 압축 형식은 JPEG, JPEG2K, SPIHT이다. 동일은 형식의 얼굴인식의 알고리즘이라도 취득된 얼굴영상의 형식에 따라 인식률에 차이가 발생한다. 얼굴의 조명(illumination), 표정(expression), 자세(pose)는 인식률에 영향을 미치는 대표적인 요인이다. 따라서 얼굴인식이 압축에 미치는 영향을 설명하기 위해 영상별로 조명조건이 차이가 나는 데이터를 사용하여 실험하였다.

1. 서 론

얼굴인식에서 얼굴 영상의 최종 판단 형식은 크게 2가지로 나뉜다. 첫 번째는 얼굴 검증(face verification) 방식이다. 이 방식은 2개의 얼굴 영상이 임계치보다 높은 유사성을 가지고 있으면 동일인물로 판단을 하고, 임계치보다 낮은 유사성을 보이면 동일인물이 아닌 것으로 판단하는 방식이다. 얼굴 검증 방식에서는 임계치를 어느 정도로 설정하느냐에 따라 인식률에 큰 차이를 가져오게 된다. 두 번째 얼굴인식 방법은 얼굴 식별(face identification) 방식이다. 얼굴 식별은 한 개의 얼굴 영상을 입력 받아서 이미 입력되어 있는 영상 중에 가장 유사성이 높은 영상으로 영상을 식별하는 것이다.

영상압축이 얼굴 검증에 미치는 영향에 대해서는 여러 연구자들에 의해 연구 되었다 [1]. 본 논문에서는 영상압축이 얼굴 식별에 미치는 영향에 초점을 맞추기로 한다. 따라서 추후에 논의되는 얼굴인식은 얼굴 식별 방법을 사용한 것이다. 본론에서는 압축 형식과 압축률에 따른 화질을 알아보고, 실제 압축된 영상의 압축률에 따른 얼굴인식률의 변화에 대해 논의한다.

2. 본 론

2.1 Face database와 영상압축의 화질

본 논문에서 사용된 영상 압축 형식은 JPEG, JPEG2K, SPIHT이 사용되었다 [2][3]. JPEG2K와 SPIHT의 경우 bpp(bits per pixel)값을 조절하여 영상압축의 파일크기를 정할 수 있다. 하지만 JPEG의 경우 품질(quality)값을 입력 받아서 압축률을 결정하기 때문에 같은 품질이어도 영상에 따라 압축률이 달라지게 된다. 따라서 JPEG의 경우 database에 12가지 품질값으로 압축을 취하게 되는데, 각 품질별로 파일의 평균 크기는 JPEG의 압축률로 계산하였다. 아울러 JPEG2K, SPIHT의 압축률은 JPEG와 비교를 위해 JPEG 품질값에 의해 산출된 압축률과 동일한 압축률을 사용하였다.

2.1.1 Yale Face database B

Yale Face database B는 2001년 다양한 조건으로 얼굴영상을 취득한 데이터베이스다.[4] 이 데이터베이스는 10명의 사람에 대한 5760개의 조명, 자세 변화에 따른 영상을 포함하고 있다. 본 논문에서는 조명 조건에 따른 영상을 알아보기 위해서 눈의 위치를 기준으로 100×100 사이즈의 크기의 영상으로 사용하였다.

2.1.2 Orl Face database

Orl Face database는 40명을 대상으로 10가지 다른 표정, 자세의 영상을 담아내고 있다 (<http://www.cam-orl.co.uk>). 각 사람은 1992~1994년 사이의 다른 시간에 촬영되었으며, 원본 데이터베이스가 사이즈는 92×112이다.

<표 1>은 압축률에 따른 데이터베이스들의 PSNR을 보여준다. M×N 영상의 원본 픽셀값을 $f(i,j)$, 압축 픽셀값을 $f'(i,j)$ 라고 정의하면, PSNR의 계산 방법은 다음과 같다.

$$MSE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (f(i,j) - f'(i,j))^2 \quad (1)$$

$$PSNR = 10\log_{10}\left(\frac{(255)^2}{MSE}\right) dB \quad (2)$$

<표 1>에서 파일의 크기는 원본 영상의 파일 크기를 1로 계산하였으며, 파일 크기 0.2는 5:1 압축률을 의미한다.

Yale Face database B			Orl Face database				
파일크기	JPEG	JPEG2K	SPIHT	파일크기	JPEG	JPEG2K	SPIHT
0.208	36.38	40.26	44.52	0.227	36.35	39.29	42.80
0.167	34.82	38.14	42.01	0.184	34.84	37.48	40.39
0.142	33.82	36.67	40.29	0.156	33.82	36.13	38.81
0.126	33.12	35.70	39.09	0.138	33.08	35.22	37.71
0.111	32.44	34.69	37.86	0.121	32.34	34.23	36.60
0.095	31.63	33.53	36.55	0.103	31.44	33.05	35.33
0.076	30.47	31.82	34.83	0.081	30.13	31.30	33.51
0.065	29.63	30.64	33.68	0.069	29.19	30.07	32.28
0.053	28.35	28.94	32.25	0.054	27.79	28.30	30.66
0.037	25.80	25.76	29.87	0.036	25.09	25.04	27.85
0.030	23.56	23.16	28.46	0.028	23.00	22.19	26.36
0.028	22.94	21.96	27.94	0.026	22.42	20.93	25.81

<표 1> 데이터베이스의 압축에 따른 화질(dB)

2.2 실험 결과

2.2.1 Yale Face database B 실험 결과

Yale Face database B는 한 사람당 1장의 영상을 학습영상(train image)로 사용하였으며, 나머지 63장의 영상을 실험영상(test image)로 사용하였다. 즉, 64장의 영상 모두 학습영상으로 한 번씩 사용되었다. 일반적으로 조명 조건이 다른 영상에 대해서는 영상에 대해서 히스토그램 이퀄라이제이션(Histogram Equalization) 과정을 거쳐서 어두운 부분을 보정한 뒤에 사용하게 된다. 본 논문에서는 보다 나은 성능을 위해서 영상의 에지(edge)를 강조하는 방법을 사용하는 영상 전처리 기술을 사용하여 보았다. 이는 조명 변화에 의해 나타난 영상들이 기본적으로 에지의 위치가 유사한 점에 착안한 것이다. <표 2>에 나온 바와 같이 4단계에 의해 이미지 전처리를 해준다.

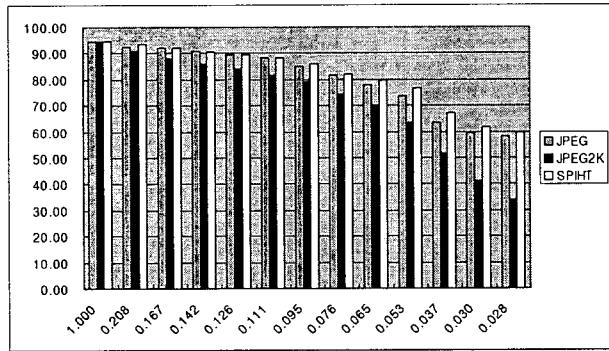
순서	전처리 기술형식	방법		
		어둡고, 밝은값 보정	Laplacian filtering (최적 A=0.0)	
1	Histogram Equalization		$\begin{array}{ccc} -1 & -1 & -1 \\ -1 & A+8 & -1 \\ -1 & -1 & -1 \end{array}$	
2	에지 강조			
3	Histogram Equalization	어둡고, 밝은값 보정		
4	픽셀값 정규화(Normalize)	$\text{Img_pixel} = \frac{(X - X_{mean})}{\text{standard-deviation}}$		

<표 2> Yale Face database B의 영상 전처리

전처리 이후, 얼굴인식을 위해 알고리즘으로 PCA를 사용하였다 [5]. 그리고 얼굴인식 최종 판단 방식으로는 유clidean 거리(Euclidean Distance)가 가장 짧은 것을 결과로 정하였다. <표 2>의 전처리 과정을 거친 후, 영상 압축 형식과 압축률에 따른 Yale Face database B의 인식률 변화는 <표 3>, <그림 1>과 같다.

파일크기	JPEG	JPEG2K	SPIHT
1.0(원본)	94.36	94.36	94.36
0.208	92.56	90.68	93.29
0.167	92.04	87.91	92.04
0.142	90.76	85.80	90.62
0.126	89.58	83.95	89.37
0.111	88.05	81.64	88.17
0.095	84.99	78.87	86.00
0.076	81.72	74.33	81.96
0.065	78.06	69.92	79.55
0.053	73.80	63.47	76.68
0.037	63.46	51.52	67.22
0.030	59.46	40.99	61.94
0.028	58.10	33.93	59.40

<표 3> Yale Face database B의 다양한 압축에 따른 인식률



<그림 1> Yale Face database B의 다양한 압축에 따른 인식률 그래프

위에서 보는 바와 같이 3가지 압축 형식 모두 압축률을 5:1 정도의 경우에서 90%를 넘을 정도의 높은 인식률을 보였다. 특히 SPIHT의 경우 원본과 큰 차이가 나지 않는 인식률을 보여주었다. 주목할 만한 점은 JPEG2K보다 상대적으로 PSNR이 떨어지는 JPEG이 더 높은 인식률을 기록했다는 점이다. 이는 JPEG이 JPEG2K보다 애지가 선명했기 때문에 발생되는 결과로 분석된다.

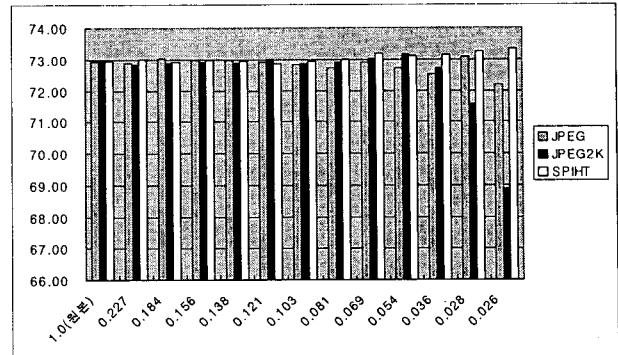
2.2.2 Orl Face database 실험 결과

Orl Face database에서도 한 사람당 1장의 영상을 학습영상(train image)로 사용하였으며, 나머지 9장의 영상을 실험영상(test image)로 사용하였다.

Orl Face database는 Yale Face database B와 달리 얼굴영상의 표정과 자세가 다른 영상을 사용한 것이다. Face database B와 달리 영상마다 위치의 위치가 어느 정도 차이를 보인다. 실제로 예지를 강조하는 영상 전처리 기술을 사용하면 영상의 서로 다른 부분이 강조가 되는 상황이 발생하여 오히려 전처리를 하지 않은 영상보다 인식률이 떨어지는 것을 관찰할 수 있었다. 그러므로 Orl Face database는 Face database B과 다른 방식의 알고리즘을 적용해 보았다. 일반적으로 얼굴의 표정과 자세 변화에서도 높은 인식률이 보인다고 알려진 알고리즘을 적용하였다. 그 알고리즘은 $(PC)^2A$ 과 2DPCA이다. $(PC)^2A$ 은 영상의 가로축과 세로축 성분을 강조하여 가로와 세로로 움직이는 변화에 민감하도록 하는 영상 전처리 기술이다 [6]. 2DPCA는 기존의 PCA가 영상을 1개의 벡터로 다루어서 에너지가 같은 방향의 직교축을 만드는 것과 달리 2차원의 영상 자체를 벡터로 이용하여 직교축을 만들어준다 [7]. 실험 결과 원본영상을 그대로 PCA를 취한 것보다 $(PC)^2A$ 는 1.1%, 2DPCA는 6.1%의 성능 향상 효과를 보여준다. 본 논문에서는 Orl Face database의 영상압축에 따른 얼굴인식률 변화를 살펴보기 위해 2DPCA 알고리즘을 사용하였다. 그 결과는 <표 4>, <그림 2>와 같다.

파일크기	JPEG	JPEG2K	SPIHT
1.0(원본)	72.94	72.94	72.94
0.227	72.89	72.83	73.00
0.184	73.03	72.89	72.92
0.156	72.97	72.92	72.97
0.138	72.97	72.89	72.94
0.121	72.92	73.00	72.86
0.103	72.83	72.86	72.94
0.081	72.72	72.92	73.00
0.069	72.92	73.03	73.19
0.054	72.72	73.17	73.11
0.036	72.53	72.72	73.14
0.028	73.08	71.56	73.25
0.026	72.19	68.89	73.33

<표 4> Orl Face database의 다양한 압축에 따른 인식률



<그림 2> Orl Face database의 다양한 압축에 따른 인식률 그래프

Orl Face database의 인식률 변화에서 가장 눈에 띄는 특징은 높은 압축률에서도 얼굴인식 성능이 떨어지지 않는다는 것이다. 오히려 원본의 얼굴인식률보다 높은 인식률을 보이는 압축 영상도 관찰되었다. 이러한 현상은 기존의 PCA 방식, $(PC)^2A$ 방식 등과 같은 다른 얼굴인식 알고리즘을 Orl Face database에 적용했을 때에도 동일하게 발견되는 현상이었다.

3. 결 론

얼굴인식의 경우 조명 변화보다 표정, 자세 변화를 다루기가 어렵다. 실제로 대부분의 표정, 자세 변화가 있는 영상들이 조명 변화가 있는 영상에 비해 얼굴인식률이 낮게 나온다. 이를 극복하기 위해 표정, 자세 변화가 있는 영상은 학습영상을 많이 쓰는 방법을 통해서 인식률을 올릴 수 있지만, 기본적으로 같은 숫자의 학습 영상을 사용할 경우 조명 변화 영상보다 인식률이 저하할 수 밖에 없다.

이는 조명 변화가 있는 얼굴영상의 경우 학습영상과 실험영상의 에지성분이 비슷한 위치에 있어서 에지 성분을 이용할 경우 높은 인식률을 얻을 수 있다. 반면 표정, 자세 변화의 경우 에지 성분보다는 영상의 영역별의 유사도를 이용해서 학습영상과 실험영상을 매칭 시키기 때문에 오히려 블러(blur) 효과가 있는 영상에서 높은 인식률을 나타내는 현상도 관찰되었다. 결론적으로 에지 정보를 많이 사용하게 되는 알고리즘의 경우 표정, 자세 변화가 있는 데이터베이스를 압축할 때 압축이 높아질수록 성능이 떨어질 것을 전망된다.

참 고 문 헌

- [1] Wolfgang Funk, Michael Arnold, Christoph Busch, Axel Mund e, "evaluation of image compression algorithms for Fingerprint and Face Recognition systems", Proceedings of the 2005 IEEE Workshop on Information Assurance and Security United States Military Academy, pp. 72-78, 2005
- [2] David Taubman, kakadu software version 6.0, <http://www.kakadusoftware.com/>
- [3] P.-E. Gougelet, "NConvert image conversion tool," 2003. www.xnview.com
- [4] A. Georgiades, P. Belhumeur and D. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," IEEE Trans. Pattern Anal. Machine Intell., vol.23,no.6,pp.643-660,2001.
- [5] M. Turk and A. Pentland, "Eigenfaces for recognition," J. Cognit. Neurosci., vol. 3, no. 1, pp. 71-86, 1991.
- [6] I. Wu, and Z.H. Zhou, "Face recognition with one training image per person", Pattern Recognition Letters, vol. 23, pp. 1711-1719,2002.
- [7] J. Yang, D. Zhang, A.F. Frangi, and J.Yang, "Two-Dimensional PCA: A New Approach to Appearancebased Face Representation and Recognition", in PAMI 2004