

얼굴 포즈 추정을 이용한 다중 RGB-D 카메라 기반의 2D - 3D 얼굴 인증을 위한 시스템*

김 정 민,[†] 이 성 철, 김 학 일[‡]
인하대학교

2D - 3D Human Face Verification System based on Multiple RGB-D Camera using Head Pose Estimation*

Jung-min Kim,[†] Shengzhe Li, Hak-il Kim[‡]
Inha University

요 약

현재 영상감시 시스템에서 얼굴 인식을 통한 사람의 신원 확인은 정면 얼굴이 아닌 관제로 매우 어려운 기술에 속한다. 일반적인 사람들의 얼굴 영상과 입력된 얼굴 영상을 비교하여 유사도를 파악하고 신원을 확인 하는 기술은 각도의 차이에 따라 정확도의 오차가 심해진다. 이런 문제를 해결하기 위해 본 논문에서는 POSIT을 사용하여 얼굴 포즈 추정을 하고, 추정된 각도를 이용하여 3D 얼굴 영상을 제작 후 매칭 하여 일반적인 정면 영상끼리의 매칭이 아닌 rotated face를 이용한 매칭을 해보기로 한다. 얼굴을 매칭 하는 데는 상용화된 얼굴인식 알고리즘을 사용하였다. 얼굴 포즈 추정은 10°이내의 오차를 보였고, 얼굴인증 성능은 약 95% 정도임을 확인하였다.

ABSTRACT

Face recognition is a big challenge in surveillance system since different rotation angles of the face make the difficulty to recognize the face of the same person. This paper proposes a novel method to recognize face with different head poses by using 3D information of the face. Firstly, head pose estimation (estimation of different head pose angles) is accomplished by the POSIT algorithm. Then, 3D face image data is constructed by using head pose estimation. After that, 2D image and the constructed 3D face matching is performed. Face verification is accomplished by using commercial face recognition SDK. Performance evaluation of the proposed method indicates that the error range of head pose estimation is below 10 degree and the matching rate is about 95%.

Keywords: Multiple-RGBD Camera, 3D Face Data, Head pose Estimation, Face Verification

1. 서 론

최근 우리나라는 날이 갈수록 강력 범죄 발생이 급증하는 등 사회적으로 불안감이 고조되고 있다. 특히 성폭력, 학교폭력, 가정폭력, 유괴 등이 심각한 사회 문제로 대두하고 있고, 생활의 불안원인으로 지적되고 있다. 이를 해결하기 위하여 ICT (Information & Communication Technology)를 이용한 범죄예방

접수일(2013년 10월 15일), 수정일(1차: 2014년 3월 10일, 2차: 2014년 8월 6일), 게재확정일(2014년 8월 7일)

* 본 연구는 미래부가 지원한 2014년 정보통신·방송(ICT) 연구개발사업의 연구결과로 수행되었음.

[†] 주저자, jmkim@vision.inha.ac.kr

[‡] 교신저자, hikim@inha.ac.kr(Corresponding author)

이 해결책으로 부상하고 있다[17]. 그중에서도, CCTV를 이용한 각종 영상감시기술이 주목을 받고 있는데, CCTV는 범죄 예방의 효과뿐만 아니라, 첨단 IT 기술을 적용한 감시시스템과 현장 예방 순찰을 결합하여 시너지 효과를 거둘 수 있기 때문이다. 그리고 이를 이용한 영상감시기술 중의 하나는 영상내 사람의 얼굴을 인식하고 그 것을 이용하여 신원을 판별하는 기술이다.

현재 상용화된 얼굴인식 기술은 얼굴의 영상에서 특징점을 추출하고, 서로 비교하여 결과를 도출하는 것이 대부분이다. 인식을 위한 얼굴 데이터의 경우 정면을 바라보는 것이 가장 좋지만, CCTV는 일반적으로 사람의 손이 닿지 않는 높은 곳에 설치되어 있기 때문에 정면 얼굴의 데이터를 취득하기 쉽지 않다. CCTV 영상에서는 카메라의 시야각으로 인해 얼굴에서 필요한 특징들을 찾지 못하게 되기 때문에, 얼굴 인식을 하는 게 힘들어지게 된다. 이를 해결하기 위하여 얼굴의 포즈를 파악하고 그를 통한 인식방법을 개발해야 할 필요가 있다.

3차원 얼굴 모델을 생성하는 연구로는 Won[1]이 제안한 템플릿 스네이크를 이용한 방법이 있다. 이 방법은 정면과 측면이 직교하는 2장의 2차원 얼굴 영상 입력받아 주요 특징을 템플릿 스네이크를 이용하여 추출하고, 그 특징점을 이용하여 일반적인 3차원 얼굴 모델을 생성하는 방법이다. 또한, Jang[2]은 템플릿 스네이크를 이용하여 특징 추출 후에 사용자 얼굴 형태에 적합한 얼굴 모델을 생성하는 방법을 제안하였다. 다른 방법으로 Bronstein[5] 이 제시한 Bending invariant canonical representation 을 기반으로 하는 방법은 얼굴의 표정과 자세로부터 모델 변형 결과에 관한 표준 영상을 생성하였다. Chang[6]은 밝기와 범위 영상을 합친 결과를 PCA(Principal Component Analysis)로 나누어 수행하여 2D-3D 영상을 매칭하는 방법을 제시하였고, Tsalakanidou[7]은 깊이 영상과 컬러 영상에 고유얼굴(eigenface)을 이용하여 3차원 얼굴을 생성하는 방법을 제안하였다. 또한, Aiptayev[15]나 Tong[16]는 2차원이 아닌 3차원의 정보를 받아 1대의 RGB-D 센서를 이용하여 3차원 얼굴 모델을 생성하였다.

얼굴 각도와 위치를 추정하는 방법으로 Kim[3]은 Modified Census Transform-based Adaboost 를 이용하였는데, 특징을 추출하고 이용하여 하나의 2차원 영상으로만 3차원 각도의 위치를 파악하였다.

Joo[4]는 AAM(Active Appearance Model)을 이용하여 표정변화를 학습하고 정면과 측면의 얼굴 포즈를 갖는 3장의 영상으로부터 얼굴의 깊이 정보를 추정하여 3차원 정보를 피팅하는 방법을 사용하였다.

본 연구는 정면이 아닌 얼굴들을 인증하는 방법으로 POSIT 알고리즘을 이용하여 얼굴의 각도를 추정한 영상과 다중 키넥트를 이용하여 생성한 3차원 얼굴영상을 이용하여, 다중 템플릿(multiple template)을 생성하고 매칭하는 방법을 제안한다. 2장에서는 얼굴의 검출 및 얼굴의 pose 추정을 설명하고, 3장에서는 3차원 얼굴 영상을 생성하는 방법을 소개하기로 한다. 4장에서는 2, 3 장의 결과를 이용한 실험 방법을 설명하고, 5장에서는 실험 결과를 나타낸다.

II. 얼굴 검출 과 얼굴 포즈 추정

2.1 Haar 특징을 이용한 얼굴 검출

영상에서 사람의 얼굴을 찾기 위해 널리 알려진 Viola-Jones[8]의 haar 특징을 이용한 얼굴 검출을 한다. Viola-Jones의 알고리즘은 학습에는 오랜 시간이 걸리지만, 얼굴 검출속도가 매우 빠르기 때문에 널리 유용하게 사용되는 알고리즘이다. 영상이 입력되면 입력 영상을 그레이 영상으로 변환한 후 haar 특징 기반의 얼굴 검출을 수행한다. 이를 적용하기 위해 학습된 cascade 분류기는 수천 장의 얼굴 이미지(Positive)와 그 두 배의 수에 달하는 얼굴이 아닌 이미지(Negative)로 생성되었다. 네모 모양으로 생성되는 haar 특징을 생성하여 cascade 연산을 통해

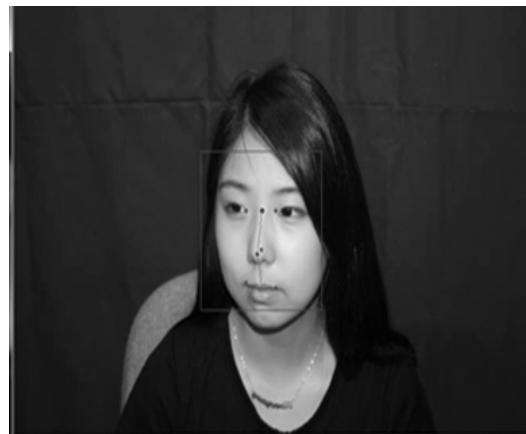


Fig. 1. Head pose estimation result using POSIT algorithm

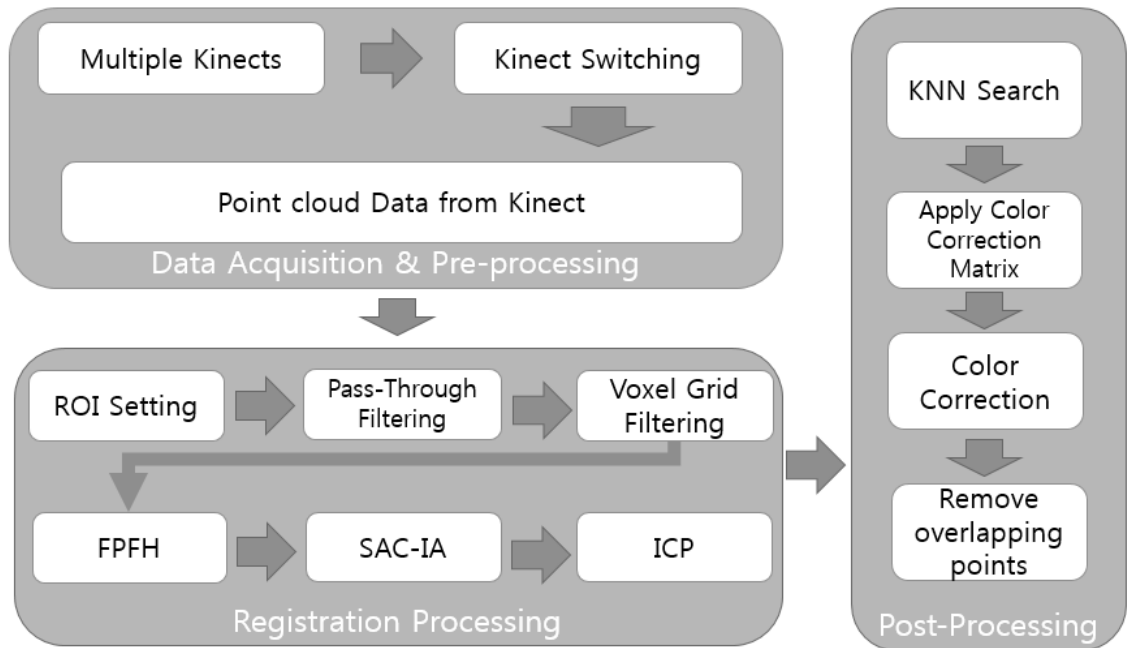


Fig. 2. 3D Face Scanning process using Multiple-Kinect.

얼굴 영상이 아닌 부분을 제거하고, cascade 연산을 반복하여 얼굴로 판별되는 부분을 검출하는 과정을 거치게 된다.

2.2 POSIT 알고리즘

영상에서의 얼굴을 찾은 후, POSIT(Pose from Orthography and Scaling with Iterations) 알고리즘을 통해 얼굴의 각도를 추정하게 된다[9]. 이 알고리즘은 6DOF(Degree of Freedom)를 이용하여 얼굴 포즈를 빠르고 정확하게 찾는 알고리즘이다. POSIT의 실행을 위해 2차원 이미지들과 3차원 객체의 일치하는 점(correspondence point)가 제공되어야 하며, 점들이 동일 평면상에 존재하지 않을 때 동작하게 된다. 동일 평면상에 위치하는 경우, 점들의 각도를 통한 연관성을 찾을 수 없게 되므로 3차원 정보를 얻을 수가 없게 된다. Fig. 1.은 POSIT을 이용하여 얼굴의 각 특징점을 찾아서 포즈를 추출한 결과를 나타낸다.

III. 다중 RGB-D 카메라를 이용한 3차원 얼굴 영상 스캔

3.1 데이터 취득 환경 설정

RGB-D 카메라를 사용하여 3차원 얼굴 영상을 생성하는 데 사용할 수 있는 방법은 단일 RGB-D 카메라를 사용하여 생성하는 방법과 두 대 이상의 카메라를 세워놓고 동시에 촬영하여 영상들을 정합하는 방법이 있다. 본 논문에서는 전 세계적으로 상용화되어 널리 쓰이는 RGB-D 카메라의 하나인 키넥트[10]를 이용하여 3차원 얼굴 영상을 생성하였다.

한 대의 키넥트나 RGB-D 카메라를 이용하여 3차원 객체를 생성하는 몇몇 연구들은 단일 키넥트를 움직이면서 대상을 스캔하고 재구성하는 방법을 사용하였다[11]. 그러나 이 경우 사람의 움직임과 촬영자의 움직임에 따라 얼굴 영상이 바르지 못하게 생성될 수 있다. 그러므로, 얼굴영상 촬영 시스템을 구성할 때 키넥트를 고정할 수 있는 프레임을 제작하고, 일정한 거리와 각도를 두어 정면만이 아닌 측면까지 촬영할 수 있도록 하였다. Fig. 2.는 3차원 얼굴 스캔을 위한 데이터 처리의 전반적인 프로세스들을 나타낸다.



Fig. 3. Test environment for Face capture using multiple-Kinect

3.2 데이터 취득을 위한 전처리 과정 및 데이터 취득

다중-키넥트를 사용하는 데 있어 가장 먼저 해결해야 하는 부분은 적외선 센서끼리의 간섭현상(interference)이다. 키넥트를 이용하여 얻는 거리(depth) 정보는, 적외선 프로젝터에서 일정한 패턴의 적외선을 투사하고, 그 투사한 적외선이 돌아오는 반사 패턴을 이용하여 얻어낸다. 그런데 다중-키넥트를 사용할 경우 각 키넥트 별 적외선 센서끼리 패턴의 간섭현상에 의해 정확한 거리 정보 획득이 불가능하게 된다. 거리 정보의 획득이 불가능해지면, 점들의 RGB 값과 거리를 매칭할 수 없어서 정확한 점들의 위치 파악이 불가능하다. 간섭현상을 해결하기 위하여 우리는 키넥트 센서를 키넥트 스위칭 방식을 사용하기로 한다. 키넥트 스위칭은 포인트 클라우드 라이브러리에 있는 OpenNIGrabber class의 start()/stop() 메소드를 이용하여, 차례대로 키넥트를 작동/정지 하여 데이터를 취득하는 방식으로, 간섭현상을 받지 않기 위해 개발한 방식이다. 연구에서 키넥트를 여러대 사용하여 데이터를 취득하지만, 실질적으로 1대를 순차적으로 사용하므로, 다중 키넥트 사용 시 생기는 간섭현상을 피할 수 있게 된다.

키넥트 스위칭을 통하여 키넥트 간의 간섭현상을 해결하고 Fig. 3.의 키넥트로부터 간섭현상이 일어나지 않은 데이터를 취득하게 된다. 본 연구에서는 키넥트로부터 데이터를 취득하기 위해 OpenNI 라이브러리와 포인트 클라우드 라이브러리를 사용하기로 한다. 데이터 취득 시, 각 키넥트로부터 데이터를 취득하는데 평균적으로 약 10초가 걸린다.

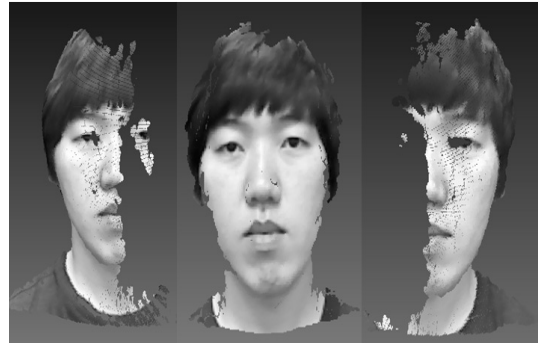


Fig. 4. Apply the pass-through filter to data from each Kinect.

3.3 포인트 클라우드 정합

정합 과정은 전체적인 과정 중 가장 중요한 부분으로, 정합의 정확성은 최종적으로 만들어지는 3차원 얼굴영상에 매우 큰 영향을 미친다. 얼굴을 3차원 영상으로 만들 때 나타나는 특징은 다음과 같다.

- 1) 데이터의 크기가 다른 연구에서 키넥트로 취득하는 다른 영상에 비해 상대적으로 낮음 [15][16].
- 2) 얼굴의 모양이 복잡하여 데이터의 정확도가 중요
- 3) 각 키넥트로 얻은 데이터는 겹치는 부분이 적음.

이러한 특징들을 고려하여, 얻은 데이터들을 다음과 같은 5가지 방법을 사용하여 처리하고 정합 과정을 진행하도록 하였다.

3.3.1 통과 필터링

키넥트를 이용하여 처음 얻은 데이터는 영역의 제한 없이 배경과 물체의 정보를 같이 얻게 된다. 이를 제한하여 필요한 정보만을 얻기 위해 x, y, z 좌표를 제한해서 데이터를 취득하여야 한다. 본 연구에서는 얼굴에 해당하는, 약 $40 \times 40 \times 40 \text{cm}^3$ 의 해당하는 영역만을 ROI(Region of interest) 영역으로 잡아서 통과 필터링(Pass-through filtering)을 하였다. Fig. 4.는 필터를 적용하여 배경을 제거한 후의 영상이다.

3.3.2 복셀 그리드 필터링

통과 필터링 이후에도 거처도 영상 내부에 존재하

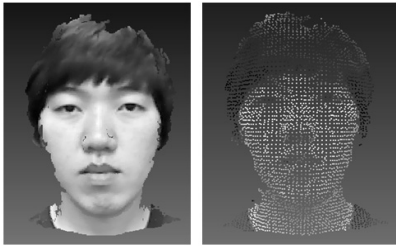


Fig. 5. Before/After image of voxel grid filtering process

는 점의 수는 수만개에 달한다. 이 모든 점을 처리하기에는 너무 많으므로, 점의 수를 줄이는 복셀 그리드 필터링(voxel grid filtering)을 거쳐야한다. 필터링 이후에는 데이터가 격자무늬(grid)처럼 일정하게 위치하게 된다. 필터링 이후 점들의 개수는 약 1/20정도로 줄어들게 된다. Fig. 5.는 복셀 그리드 필터링을 적용하기 전/후의 영상이다.

3.3.3 FPFH 알고리즘 적용

복셀 그리드 필터링을 한 이후에, 각 점으로부터 특징점을 추출해 내어 매칭에 필요한 특징들을 찾아내야 한다. 그 특징을 찾아내기 위하여 FPFH(Fast Point Feature Histogram)을 사용한다[12]. FPFH는 점들의 특징의 히스토그램을 계산하는 PFH(Point Feature Histogram) 프로세스를 이용하는데, PFH 프로세스에서 k 가 한 점에서 근처의 점의 개수라고 할 때, $O(nk^2)$ 에 해당하는 속도를 가지고 있어서 느린 편이다. FPFH는 PFH 에서 느린 속도를 빠르게 하고자 고안한 방법이다. FPFH를 적용하면 PFH와 같은 특징점을 찾아내면서 $O(nk)$ 의 속도로 줄어들게 된다.

3.3.4 Sample Consensus Initial Alignment

데이터들을 정합시키기 위해서는 하나의 기준이 필요하다. 데이터들의 정합을 위해 SAC-IA(Sample Consensus Initial Alignment) 을 사용하였다. 포인트 클라우드 객체 중 기준을 잡은 후, 기준이 되는 객체로 이동과 회전을 통해 맞추고 합쳐주는 프로세스를 뜻한다. 이 연구에서는 가운데에 있는 키넥트에서 추출된 포인트 클라우드 데이터를 기준으로 삼고, 양 옆에 있는 키넥트의 포인트 클라우드 데이터를 가운데에 있는 포인트 클라우드 데이터로 가지고 와서



Fig. 6. The concept of SAC-IA Process. Side image data move to center image data.

맞추는 작업을 하고 결합하였다. 이를 통해 각 부분의 데이터를 통하여 정합이 되면 얼굴의 형태를 띄게 된다. Fig. 6.은 SAC-IA 프로세스의 개념을 보여준다.

3.3.5 Iterative Closest Points

SAC-IA 방법을 통하여 정합하였지만, 점들이 겹쳐있는 것이므로 이제 비슷한 위치에 있는 점들을 하나의 점으로 만드는 과정이 필요하다. 정합된 점들을 더욱 정밀하게 만들기 위하여 ICP(Iterative Closest Points)를 통하여 점들 사이의 거리를 구하여, 서로 가까이 있는 점들을 하나로 합치는 과정을 통해 불필요한 데이터를 줄이는 과정을 거친다. 이 과정까지 마치면, 얼굴의 형태가 나오게 된다.

3.4 Post-Processing

정합이 정확하게 이루어졌지만, 색의 차이 때문에 얼굴의 피부색이 부자연스러운 것을 발견할 수 있다. 그 이유로는 각 키넥트의 방향이 다르므로, 영상은 서로 다른 밝기 값을 가지기 때문이다. 이를 해결 위해, 두 대의 키넥트간의 같은 위치의 대응점을 구하는 K-NN(K-Nearest Neighbor) 검색을 통하여 문제가 되는 부분의 위치를 찾아낸다. 또한, 서로 다른 데이터 간의 색 차이를 줄이기 위해 색 보정 행렬을 만들어 색의 차이를 보정해준다. 다음은 키넥트 간 색을 보정하는 식이다.

$$M_{cc} * [r_p b_p g_p]^T = [r_q b_q g_q]^T \quad (1)$$

M_{cc} 는 Color Correction 행렬, $[r_p g_p b_p]$ 와 $[r_q g_q b_q]$ 는 점 p와 q 의 RGB값이다. 이를 통해 데이터 간의 색의 차이의 변화가 줄어들게 된다. 하지만 색 변환 과정을 수행했음에도 자연스럽게 않은 부분이



Fig. 7. Experimental result of 3D Face construction and Pose Estimation.
Left side : Our result / Right side : Test Sample

남아있게 되는데, 이 부분은 포인트들의 품질 가중치 연산(quality weighting)을 통하여 가중치가 낮은 부분의 제거를 해준다. 품질 가중치 연산에 필요한 Q_p 는 점 p 에서 반지름 r 이내에 있는 점의 개수를 나타낸다. Q_p 가 적은 부분을 제거하여, 가장자리 부분을 제거한다. 이 과정을 통해 가장자리의 울퉁불퉁한 부분이 제거된다. Fig. 7.은 모든 과정을 거친 후 나온 영상들의 예시이다. 좌측 열은 실제 모든 과정 이후 나온 결과이고, 우측 열은 일정치 않은 각도를 가진 사진들을 나타낸다.

IV. 실험 및 결과

4.1 실험 환경

3대의 키넥트를 이용한 실험환경을 구축하고, 각 키넥트 간의 각도는 45° , 영상은 Fig. 3.과 같이 사무실과 강의실에 배경천막을 설치하여 가능한 조명 환경의 영향을 덜 받고 얼굴 영상에 잡음이 끼지 않는 환경을 설정하였다. 3차원 영상은 키넥트를 통하여 취득하여 수집하였고, 2차원 영상은 DSLR을 이용하여 랜덤하게 yaw와 pitch의 변화를 주어가며 취득하였다. 실험자의 수는 총 30명이였다. 얼굴 인식을 위한 알고리즘은 상용 얼굴 인식 알고리즘인 VeriLook[13]을 사용하였고, 포즈 추정에는 POSIT 알고리즘[14]을 사용하였다.

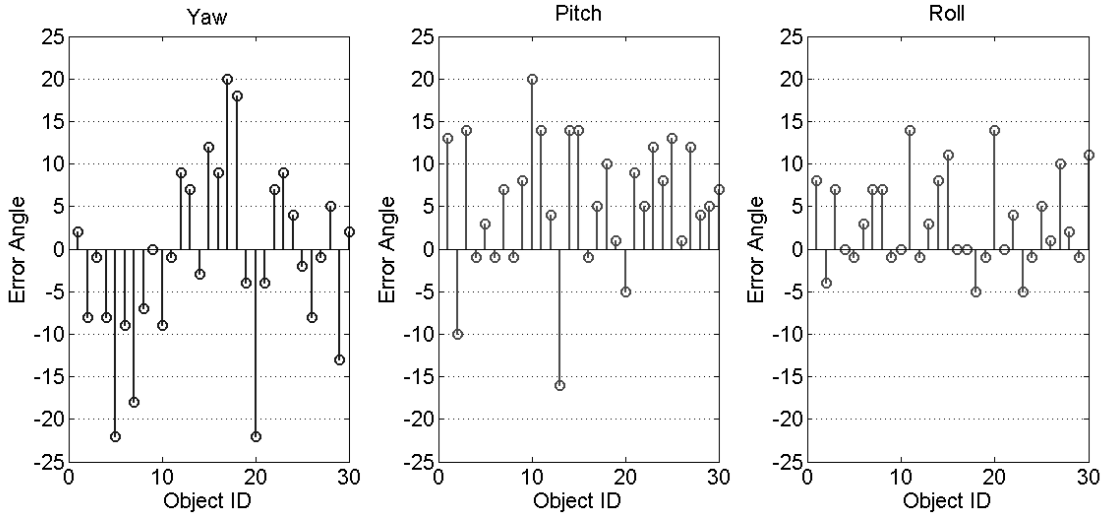


Fig. 8. Mean error of each test

Table 1. Mean error of angle value

	Yaw(y)	Pitch(x)	Roll(z)
Mean of Error(°)	-1	6	3
Std	10.39	7.71	5.47

Table 2. Range of estimated angle

Direction	Estimated angle in 2D	Estimated angle in 3D
Yaw	-56° ~ 51°	-47° ~ 50°
Pitch	-54° ~ 29°	-56° ~ 31°
Roll	-31° ~ 15°	-34° ~ 32°

4.2 실험 방법

실험은 일차적으로 POSIT을 이용하여 30명의 얼굴 영상에서의 각도를 추정하는 프로세스를 진행한다. 그 결과로 나오는 추정된 각도들의 정확성을 구하기 위해, 추정된 각도로 만들어진 3차원 얼굴 영상의 각도를 다시 한 번 추정하여, 실제로 차이가 얼마나 나는지를 확인해보도록 한다. 그 이후, 추정된 각도만큼 돌아간 3차원 얼굴 영상과 디지털카메라로 촬영한 영상과 VeriLook을 이용한 다중 얼굴 템플릿 생성 후 얼굴 인증을 하는 방법을 사용한다. 성능향상을 확인하기 위하여 회전된 2차원 영상과 3차원 얼굴의 정면 영상을 매치해보고, 다시 회전된 2차원 영상과 회전된 3차원 얼굴 영상을 매치하였다.

4.3 실험 결과

4.3.1 포즈 추정

원본 영상으로 포즈 추정을 한 결과와 3차원 생성 영상을 포즈 추정된 결과, 차이의 평균값은 다음과 같다. Table 1.에 나온 것과 같이 Yaw 방향으로 -평균 1°, pitch 방향으로는 평균 6°, roll 방향으로는 평균 3° 정도의 차이가 있는 것을 확인할 수 있었다. Fig. 8.은 각 방향의 오차 값을 나타내는 그래프이다. Roll 방향의 오차가 가장 적게 보이는데, 실험 시나리오 상 roll의 방향으로 움직이는 것은 가정 하지 않았기 때문이다. 표준편차는 각도별로 10.39, 7.71, 5.44로 나타났다. 각도가 크게 변화였던 Yaw 부분의 표준편차가 크게 나타났는데, 실제로 확인해보면 Pitch보다 Yaw의 각도 변화가 Fig. 8.에서도 큰 것을 확인할 수 있었다.

4.3.2 2차원 영상과 3차원 영상 매칭

앞서 설명한바와 같이, 회전된 2차원 영상과 3차원 얼굴의 정면 영상을 매칭한 후, 회전된 2차원 영상과 회전된 3차원 영상을 매칭 하였다. 실험에 사용된 각 방향의 범위는 Table 2.와 같다. 실제 얼굴에서 추정된 결과 값을 이용하여 매칭은 한 사람의 전체 2차원 영상과 전체 3차원 영상을 1 : 1 매칭하는 인증방식으로, 개인별로 본인 인증을 진행하였다. 매칭에 사용

Table 3. Parameters of VeriLook

VeriLook Parameter	value
Min eye of distance	16 pixel
Max eye of distance	4000 pixel
Max roll angle deviation	90°
Max yaw angle deviation	45°
Face confidence threshold	40
Face quality confidence	40

하기 위한 VeriLook 얼굴 인증 알고리즘의 매개변수는 다음 Table 3.과 같다. 인증 실험에 사용된 FAR (False Acceptance Rate)의 값은 0.1%이었다.

이를 이용한 실험 결과를 Table 4.에 나타내었다. 처음 실험한 회전된 2차원 영상과 3차원 얼굴의 정면 영상은 89.7%의 인증률을 보였다. 매칭이 되지 않는 경우, 3차원 정면 영상의 특징점과 회전된 2차원 영상의 특징이 매칭되지 않고, 포즈에 따른 차이로 인하여 매칭을 못하는 경우가 많았다.

회전된 2차원 영상과 회전된 3차원 영상을 매칭은 96.6%의 본인 인증률을 보여 약 6.9%의 상승률을 보였다. 추정된 각도를 이용하여 3차원 영상을 생성하면 높은 확률로 매칭되는 것을 볼 수 있었다.

하지만 그럼에도 불구하고 인증이 안 된 경우는, 특징점을 추출해야 하는 눈, 코, 입이 없을 정도로 얼굴의 각도가 기준범위를 벗어난 경우거나, 얼굴의 품질이 낮아서 템플릿을 생성하지 못하여 인증에 실패하는, FTE(Failure To Enroll)였다.

V. 결 론

본 논문에서는 POSIT 알고리즘을 이용하여 얼굴의 얼굴 포즈 추정을 수행하고, 그 수행된 결과를 3차원 영상에 적용, 회전된 3차원 영상을 생성하여 매칭하는 과정을 보였다. 일정하지 않은 포즈에서 얼굴이 향하는 각도를 찾는 것을 해결하기 위해 POSIT 알고리즘을 적용하여 3차원 좌표로 나오는 값들을 이용하여 각도를 추정하였고, 3차원 얼굴 데이터를 실제로 추정된 각도로 만들었다. 더 나아가, 그것들을 인증과정을 거쳐 실제로 얼마나 일치하고 인증률이 증가하는지를 보았다.

본 연구에서 사용한 3차원 얼굴 데이터를 만드는 방법은 기존 3D 스캐너와 비교하여 키넥트 센서 3대를 만들면 되기 때문에 매우 빠르다. 게다가, 낮은 해상도에도 불구하고, 얼굴 인증에 사용할 수 있을 만큼

Table 4. Experimental result

Experiment method	Verification rate(%)
3D frontal face : 2D rotated face	89.65
3D rotated face : 2D rotated face	96.55

의 3차원 얼굴 데이터 생성에 성공한 것을 확인할 수 있었다. 하지만 키넥트 센서의 특성상 조명의 영향을 매우 많이 받게 되고, 또한 POSIT 알고리즘을 이용한 추정이 제대로 안 된 경우엔 인증 시도 자체가 무산될 수도 있다.

추후 좀 더 정확하면서도 촬영 환경에 강인한 얼굴 데이터 스캔 방법을 연구해야 할 것이다. 또한, 얼굴에서 더 나아가 사람의 신체 데이터를 스캔하여, soft-biometrics를 사용해 신체의 특징을 활용하게 된다면, 좀 더 정교한 사람인식이 가능해질 것이다. 또한, 이를 이용하여 영상의 인물 감지나 특정 사람 찾기와 같은 많은 인력이 필요한 힘든 일을 쉽게 할 수 있게 될 것이다.

References

- [1] Sun-hee Won and Gye-young Kim, "Texture Mapping and 3D Face Modeling using Two Views of 2D Face Images," Journal of KIISE : Computer Systems and Theory, 15(9), Sept. 2009.
- [2] Seok-woo Jang, Gye-young Kim, "Synthesizing Faces of Animation Characters Using a 3D Model," Journal of The Korea Society of Computer and Information, 17(8), pp. 31-40, Aug. 2012.
- [3] Woo-Won Kim, Jin-kyu Hwang and Sang-youn Lee, "Estimation of Head Angle and Position from 2D images," 2011 Korean Society of Broadcast Engineers Summer Conference, pp. 1-3, July 2011.
- [4] Myung-ho Joo and Hang-bong Kang, "3D Face Model Fitting Method on Active Appearance Model with 3D depth Estimation," Journal of KIISE : Software and Applications, 39(2), pp. 109-117, Feb. 2012.

- [5] A.M. Bronstein, M.M. Bronstein, and R.Kimmel, "Expression-invariant 3D face recognition," *Audio-and Video -Based Biometric Person Authentication*, 2688, pp. 62-70, Jan. 2003.
- [6] K. Chang, K. Bowyer, and P. Flynn, "Face recognition using 2D and 3D facial data," *ACM Workshop on Multimodal User Authentication*, pp 25-32, Dec. 2003.
- [7] F. Tsalakanidou, D. Tzovaras, and M. Strintzis, "Use of depth and colour Eigenfaces for face recognition," *Pattern Recognition Letters*, vol. 24, no. 9, pp. 1427 - 1435, Jun. 2003.
- [8] P. Viola and M.J. Jones, "Robust Real-Time Face Detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137-154, May 2004.
- [9] D.F. Dementhon and L.S. Davis, "Model-Based Object Pose in 25 Lines of Code," *International Journal of Computer Vision*, vol. 15, no. 1-2, pp. 123-124, June 1995
- [10] Microsoft, Kinect for Windows, <http://www.microsoft.com/en-us/kinectforwindows>
- [11] Y.W. Lim, H.Z. Lee, N.E. Yang, and R.H. Park, "3-D Reconstruction Using Kinect Sensor and Its Application to a Visualization System," *Systems, Man, and Cybernetics*, 2012 IEEE International Conference on, pp. 3361-3366, Oct. 2012.
- [12] R.B. Rusu, N. Blodow, and M. Beetz, "Fast Point Feature Histograms (FPFH) for 3D Registration," *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3212-3217, May 2009.
- [13] Neurotechnology, VeriLook, http://neurotechnology.com/vl_sdk.html
- [14] Seong-jun Lee, Seong-joo Lee and Hak-il Kim, "Head angle estimation of ststic feature using POSIT algorithm," 2013 28th ICROS Annual Conference, pp. 528-529, May 2013.
- [15] K. Aitpayev and J. Gaber. "Creation of 3D Human Avatar using Kinect," *Asian Transactions on Fundamentals of Electronics, Communication & Multimedia*, vol. 1, no. 5, pp. 3-5. Jan. 2012.
- [16] J. Tong, J. Zhou, L. Liu, Z. Pan and H. Yan. "Scanning 3D Full Human Bodies Using Kinect," *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 4, pp. 643-650. April 2012
- [17] Young-sun Lee, "ICT application plan for crime prevention response system," National Information Society Agency, Dec. 2013.

〈 저자 소개 〉



김 정 민 (Jung-min Kim) 학생회원
 2012년: 인하대학교 정보통신공학부
 2012년 ~ 현재: 인하대학교 정보통신공학과 석사과정
 <관심분야> 영상처리, 생체인식



이 성 철 (Shengzhe Li) 학생회원
 2008년: Beihang University, China, 소프트웨어 공학과
 2009년 ~ 현재: 인하대학교 대학원 정보통신공학과 통합과정.
 2012년 ~ : 비전인 CTO.
 <관심분야> 영상처리, 바이오인식, 지능형 비디오감시.



김 학 일 (Hak-il Kim) 종신회원
 1983년 2월: 서울대학교 제어계측공학과 졸업
 1985년 2월: Purdue Univ. 전기/컴퓨터공학과 석사
 1990년 2월: Purdue Univ. 전기/컴퓨터공학과 박사
 現 인하대학교 정보통신공학부 교수, 한국생체인식포럼 부의장, 한국정보보호학회 바이오인
 증연구회 회장, 한국 정보보호진흥원 바이오인식기술 자문의원, IOS/IEC JTC1-Sc37
 (Biometrics) 국내 및 국제 전문의원, ITU-T/SG17/WP2/Q.9(Telebio-metrics)
 Rapporteur 등
 <관심분야> 패턴인식, 컴퓨터비전, 바이오인식, 초분광영상처리, 로봇비전, 의료영상처리