

포즈 인식에서 효율적 특징 추출을 위한 3차원 데이터의 차원 축소

경 동 욱[†] · 이 윤 리^{**} · 정 기 철^{***}

요 약

사용자 포즈의 3차원 데이터 생성을 통한 3차원 포즈 인식은 2차원 포즈 인식의 문제점을 해결하기 위해서 많이 연구되고 있지만, 3차원 표면 데이터의 방대한 양으로 포즈 인식에서 중요한 특징 추출(feature extraction)이 어렵고 수행 시간이 많이 걸리는 문제점을 가지고 있다. 본 논문에서는 3차원 포즈 인식의 두 가지 문제점인 특징 추출의 어려움과 느린 처리속도를 개선하기 위해서 3차원 형상복원 기술로 모델의 3차원 표면 점들로 구성된 데이터를 2차원 데이터로 변환하는 차원 축소(dimension reduction) 방법을 제안한다. 실린더형 외곽점을 이용한 메쉬없는 매개변수화(meshless parameterization) 방법은 방대한 데이터인 3차원 포즈 데이터를 2차원 데이터로 변환하여 특징 추출과 매칭과정의 연산 속도를 향상 시키며, 특징 추출의 효율성 검증을 위해 간단한 환경에서 실험이 가능한 손 포즈 인식 및 인간 포즈 인식에 적용하였다.

키워드 : 차원 축소, 특징 추출, 포즈 인식, 메쉬없는 매개변수화

3D Data Dimension Reduction for Efficient Feature Extraction in Posture Recognition

Dongwuk Kyoung[†] · Yunli Lee^{**} · Keechul Jung^{***}

ABSTRACT

3D posture recognition is a solution to overcome the limitation of 2D posture recognition. There are many researches carried out for 3D posture recognition using 3D data. The 3D data consist of massive surface points which are rich of information. However, it is difficult to extract the important features for posture recognition purpose. Meanwhile, it also consumes lots of processing time. In this paper, we introduced a dimension reduction method that transform 3D surface points of an object to 2D data representation in order to overcome the issues of feature extraction and time complexity of 3D posture recognition. For a better feature extraction and matching process, a cylindrical boundary is introduced in meshless parameterization, its offer a fast processing speed of dimension reduction process and the output result is applicable for recognition purpose. The proposed approach is applied to hand and human posture recognition in order to verify the efficiency of the feature extraction.

Keywords : Dimension Reduction, Feature Extraction, Posture Recognition, Meshless Parameterization

1. 서 론

컴퓨터 기술의 발전으로 자연스럽게 편리한 인터페이스에 대한 요구가 날로 증가하고 있다. 기존의 사용자 인터페이스 장치인 키보드 및 마우스 등은 사람들이 일상적으로 경험하기 어려운 환경을 직접 체험하지 않고서도 그 환경에 들어와 있는 것처럼 보여주고 조작할 수 있는 가상환경의

입력 장치로 사용하기에는 많은 불편함을 가진다. 이러한 불편함을 해결하고자 사용자의 직관적인 제스처 및 포즈를 인터페이스의 수단으로 사용하고자 하는 연구가 활발하게 진행되고 있다. 제스처 및 포즈는 일상 생활에서 인간의 모션을 이용한 의사소통의 한 형식을 말한다. 즉, 우리는 택시를 멈추기 위해서 손을 흔들며, 동의할 때에는 머리를 끄덕이며, 또한 관심을 끌기 위해서 손을 들어 올린다. 이러한 모든 것을 제스처 또는 포즈라 부르며 인간의 신체는 의사소통을 위한 다양한 제스처 및 포즈로 표현할 수 있다.

제스처 및 포즈 인식은 처리하는 정보의 입력 방식에 따라 센서 기반의 방법과 컴퓨터 비전을 기반으로 하는 방법으로 나눌 수 있다. 전자는 사용자의 관절 부위에 센서를 부착하여 제스처 인식을 수행한다[1-3]. 예를 들어, 인체 중

* 이 논문은 2007년 정부(교육인적자원부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구임(KRF-2007-357-D000230)

† 정 회 원 : 스위스 ETH Zurich 계산과학과

** 준 회 원 : 송실대학교 IT대학 미디어학과 박사과정

*** 종신회원 : 송실대학교 IT대학 미디어학과 교수

논문접수: 2007년 8월 9일

수정일: 1차 2008년 1월 16일, 2차 2008년 4월 25일

심사완료: 2008년 5월 30일

요 부분의 위치를 추적하는 모션 트래킹(motion tracking)[1], 손의 관절 정보를 제공하는 데이터 글러브(data gloves)[2], 기존의 2차원 데이터를 제공하는 마우스에서 3차원 환경에 사용할 수 있는 3차원 마우스(3D mice)[3] 등이다. 이러한 장치들은 정확한 움직임과 위치를 제공하지만 센서의 사용으로 높은 비용과 항상 번거로운 장치를 착용해야 하므로 불편하고 자연스럽지 못하다. 후자는 한 대 이상의 카메라 영상으로부터 컴퓨터 비전 기법을 사용하여 제스처 인식을 수행한다[4-6]. 2차원 데이터인 카메라 영상을 이용한 포즈 인식은 처리시간이 적게 소요되지만 인간의 포즈를 정확하게 예측하기는 어렵다. 그 이유는 카메라 시점에 따라 자기 자신에게 가려지는 문제와 시점에 따라 포즈가 다르게 캡처되기 때문이다[7]. 그리고 최근 가상환경의 발달로 2차원 포즈 인식만으로는 정확하고 자연스러운 인식을 수행할 수 없기 때문에 보다 정확한 모션 정보를 가지는 3차원 포즈 데이터를 제스처 인식에 사용하고자 연구되고 있다. 3차원 데이터는 기존의 2차원 데이터보다 많은 양의 데이터를 가지기 때문에 3차원 제스처 인식에서는 효율적인 특징 추출(feature extraction)과 빠른 연산 속도를 요구한다[8,9].

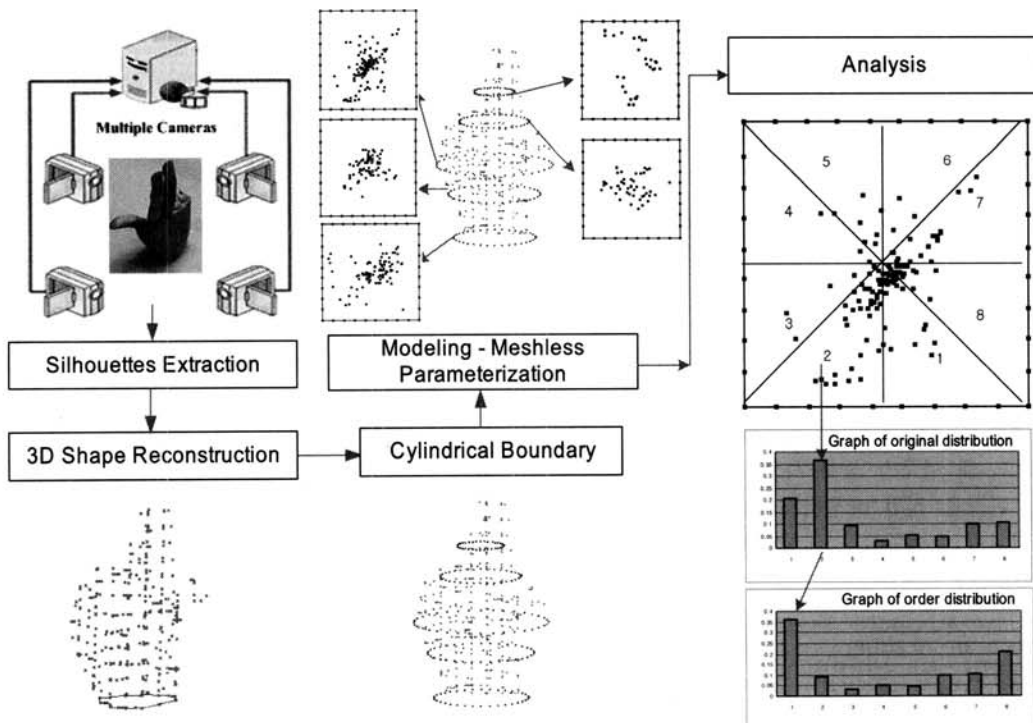
본 논문에서는 3차원 포즈 인식에서 중요한 특징 추출과 빠른 연산 속도를 위해서 3차원 데이터를 2차원 데이터로 변환하는 실린더형 외곽점을 이용한 매쉬없는 매개변수화 방법을 제안한다. 매쉬없는 매개변수화(meshless parameterization)는 컴퓨터 그래픽(computer graphics)에서 3차원 모델을 2차원 모델로 변환하여 텍스처 매핑(texture mapping)을 수행하는 방법으로써, 포즈 모델링 과정의 특징 추출의 빠른 연산을 위해서 적용되었다[10].

(그림 1)은 손 포즈 인식을 위한 특징 추출 과정으로써, 4대의 카메라를 이용하여 인공 손 모델의 표면점인 3차원 데이터를 사용한다. 본 논문에서는 인공 손 모델과 인간 모델을 이용하여 실험하였다.

2. 관련 연구

3차원 제스처 및 포즈 인식은 많은 데이터 양과 고차원 데이터 표현으로 특징 벡터의 선택 및 특징 추출이 어렵다. 특징 벡터는 제스처 모델의 예측 및 인식에 많은 영향을 주는 요소로써, 대부분의 연구에서는 위치, 방향, 속도, 가속도 등을 사용한다. 특징 벡터의 선택이 3차원 제스처 인식에서 중요함을 보여주는 연구[11]로써, 3차원 제스처 인식을 위해 10개의 다른 특징 벡터를 3차원 손 및 머리의 추적 데이터(tracking data)를 근간으로 18개의 T'ai Chi 제스처를 테스트하였다.

특징 추출은 입력 패턴으로부터 대상을 표현하는 공통적인 성질을 추출하는 것으로써, 포즈 인식의 정확도 및 처리 속도에 많은 영향을 미치는 중요한 요소이다. 하지만 3차원 포즈 인식에서는 3차원 데이터의 방대한 양으로 특징 추출에 많은 어려움을 가진다[20,21]. Aguiar[20] 등은 사용자의 포즈를 3차원 데이터로 생성 및 인간의 중요 관절 부위를 특징 벡터로 인식을 수행하지만, 특징 추출시 많은 연산량으로 실시간 응용에 적합하지 않다. Theobalt[21] 등은 오프라인으로 3차원 포즈 데이터를 생성 및 포즈를 인식하는 연구로써, 연속된 3차원 포즈데이터에서 이전 데이터와의 차이점 분석을 위해 3차원 광류(optical flow)를 사용한다. 3차원 광류는 사용자의 움직임 정보를 추출하기 위해 사용되지



(그림 1) 3차원 형상 복원으로부터 특징 추출까지의 처리과정

만 느린 처리속도로 실시간 처리에 적합하지 않다. 이와 같이 3차원 포즈 인식에서의 특징 추출은 많은 연산 량으로 느린 처리속도를 가진다.

3차원 제스처 및 포즈 인식에서의 데이터 표현(data representations)은 2가지 형식으로 나눈다. 하나는 객체기반(object-based) 또는 모델기반(model-based)으로 점(point), 박스(box), 실루엣(silhouettes)과 같은 형태로 표현하는 방식이다. 다른 하나는 이미지기반(image-based)으로 공간(spatial), 시간적공간(spatio-temporal), 모서리(edge) 또는 특징(feature)으로 표현하는 방식이다[12].

객체기반의 연구[11,13,14]는 인간의 포즈를 3차원으로 캡처하여 생성된 데이터를 사용한다. Malassiotis[13] 등은 포즈 인식을 위해서 3차원 센서를 이용하여 생성된 데이터를 통해서 강건하게 또는 효과적으로 성능을 향상시키지만, 3차원 센서 또는 특별한 카메라의 사용으로 다른 제스처 인식 시스템 보다 많은 비용이 요구된다. Shin[14] 등은 제스처 인식을 위해서 3차원 모션 히스토리 모델(3D motion history model)을 사용하였다. 이 방법은 3차원 상에서 연속적인 스테레오 입력 데이터를 사용하여 2차원 모션 데이터가 가지는 시점과 확장성의 한계를 극복하고자 많이 사용되지만[8,15], 2차원 모션 히스토리 영상(Motion History Image)보다 처리속도가 느린 단점을 가진다. Fablet and Black[15]는 인간의 모션을 자동 검출 및 추적을 위해서 주 성분 분석법(principal component analysis)을 이용한 학습으로 새로운 저차원의 시공간적 모델 표현법을 개발하였다. 하지만 주 성분 분석법은 방대한 데이터를 저차원 데이터로 변환하여 특징 추출이 빠르지만, 많은 정보 손실을 유발하기 때문에 정확한 제스처 인식에 효율적이지 못하다.

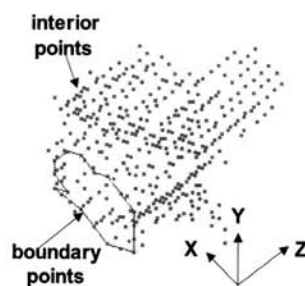
기존의 연구들은 3차원 데이터를 사용하여 제스처 및 포즈 인식을 수행하지만 처리속도와 정보손실 그리고 비용의 한계를 가지고 있다. 본 논문에서 제안한 방법은 3차원 포즈 데이터의 정보 손실을 최소화하여 2차원 데이터로 차원 축소하는 방법으로, 기존 연구의 문제점인 특징 추출 및 연산 속도를 향상시킨다.

3. 포즈 모델링

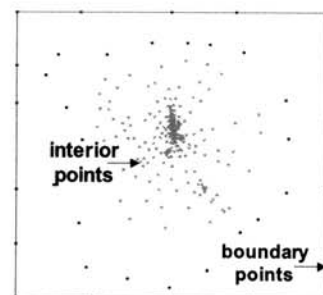
포즈 모델링은 포즈 인식에 필요한 특징 추출을 수행하는



(a) 모델 영상



(b) 모델의 3차원 데이터



(c) 2차원 데이터

(그림 2) 메쉬없는 매개변수화의 데이터 변환과정

과정으로써, 3.1장은 기존의 메쉬있는 매개변수화 방법을 소개하며, 3.2장은 3차원 포즈 인식을 위해 개선된 메쉬없는 매개변수화 방법을 기술한다.

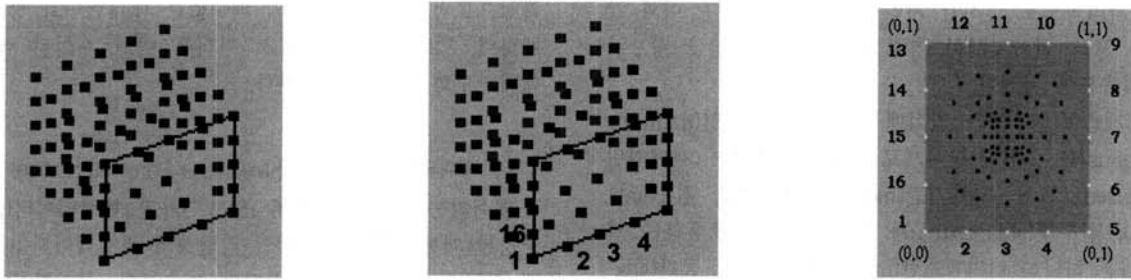
3.1 메쉬없는 매개변수화(Meshless Parameterization)

메쉬없는 매개변수화[16-18]는 3차원 복셀데이터(x,y,z)로 구성된 객체를 2차원 도메인 평면상의 픽셀데이터(u,v)로 점대 점(point to point) 변환하는 기술으로써, 복잡한 3차원 데이터 구조를 단순한 2차원 구조로 변환한다(그림 2). 3차원 데이터는 모델(그림 2(a))의 표면점들로 구성된 데이터로써, Z축이 가장 낮은 표면점들의 집합을 외곽점(boundary points)이라 하며 외곽점을 제외한 표면점들의 집합을 내부점(interior points)이라 한다(그림 2)(b). 외곽점과 내부점으로 명명한 이유는 3차원 데이터를 2차원 데이터로 차원 축소시, 외곽점들은 2차원 데이터의 외곽에 위치하며, 내부점들은 2차원 데이터의 내부에 존재하기 때문이다(그림 2)(c).

메쉬없는 매개변수화 방법은 N개의 3차원 데이터로 이루어진 객체를 집합 $P = \{p_1, p_2, \dots, p_n, p_{n+1}, p_{n+2}, \dots, p_N\}$ 로 표현할 때, 집합 P는 Z축을 기준으로 가장 낮은 표면점들 N-n개의 외곽점 $P_B = \{p_{n+1}, p_{n+2}, \dots, p_N\}$ 와 외곽점을 제외한 n개의 내부점 $P_I = \{p_1, p_2, \dots, p_n\}$ 로 분류한다. 분류된 외곽점과 내부점을 이용하여 2가지 단계로 3차원 데이터를 2차원 평면으로 변환한다. 정렬되지 않은 연속적인 벡터들로 이루어진 3차원 데이터 집합 P는 외곽점 집합 P_B 를 분류하기 위해서 오직 하나의 외곽으로 이루어져있는 개방 표면(open surface) 모델로 가정한다.

첫 번째는 분류된 외곽점 P_B 를 정렬 및 시계 반대 방향으로 도메인 평면의 외곽점 $U = \{u_{n+1}, u_{n+2}, \dots, u_N\}$ 로 매핑하는 단계이다. 두 번째는 도메인 평면에 3차원 내부점들을 $U = \{u_1, u_2, \dots, u_n\}$ 로 매핑하는 단계이다.

외곽점을 2차원 도메인 평면에 매핑하기 위해서는 분류된 3차원 데이터의 외곽점을 정렬 및 둘레의 길이를 계산해야 한다. 외곽점의 정렬은 boundary-following[19] 알고리즘을 통해서 외곽점을 순서에 맞게 정렬한다(그림 3)(b). 2차원 도메인 평면은 (0,0)을 기준으로 가로,세로 길이가 1인 정사각형 영역을 가지며, 3차원 데이터의 정렬된 외곽점들은 각 길이의 비율에 맞게 2차원 도메인 평면 둘레에 시계 반대 방향으로 매핑된다[10]. 즉, (그림 3)(b)의 정렬된 외곽점 16



(a) 3차원 데이터 및 외곽점 (b) 정렬된 외곽점 (3) 2차원 도메인 평면

(그림 3) 3차원 모델의 차원 축소 결과

개는 정사각형의 2차원 도메인 평면 둘레에 시계 반대 방향으로 1번부터 16번까지 비율에 맞게 매핑된다(그림 3)(c).

두 번째는 내부점 집합 P_i 의 각 점 p_i 마다 이웃하는 점들의 집합 N_i (점 p_i 와의 거리가 반지름 r 을 만족하는 점들의 집합)를 구성하며, 이웃점들의 상호간의 거리 가중치(reciprocal distance weight)을 이용하여 2차원 도메인 평면에 매핑된다. 이웃점들의 가중치(weight)는 2차원 도메인 평면의 내부점 $U = \{u_1, u_2, \dots, u_n\}$ 로 n 개 수식의 선형 시스템으로 계산되며[16-18], (그림 3)(c)와 같이 정사각형 영역 내부에 매핑된다.

각 내부점 p_i 의 이웃점 결정과 이웃점들 사이의 가중치 λ_{ij} 을 계산하기 위해서 상호간의 거리 가중치(reciprocal distance weight) 방법과 이웃점을 결정하기 위해서 반지름 r 을 이용한 볼 이웃(ball neighborhood) 방법을 사용한다[17]. 수식 (1)은 이웃점 선택과 가중치 계산을 위한 수식이다.

$$N_i = \{j : 0 < \|p_j - p_i\| < r\} (r > 0),$$

$$0 < \|p_j - p_i\| < r\} (r > 0), \lambda_{ij} = \frac{1}{\|p_j - p_i\|} \Big/ \sum_{k \in N_i} \frac{1}{\|p_k - p_i\|} \quad (1)$$

가중치 λ_{ij} 은 $j = N_i$ 일때 $\sum_{j \in N_i} \lambda_{ij} = 1$ 을 만족하며, 내부점 p_i 와 대응대는 n 개의 점 $U = \{u_1, u_2, \dots, u_n\}$ 를 계산하기 위해서 n 개 수식의 선형 시스템을 계산한다. 이때 p_i 는 u_i 에 대응되며, p_i 의 이웃점 집합 u_j 로 볼록결합을 위해서 모두 동일한 값을 가진다.

$$u_i = \sum_{j \in N_i} \lambda_{ij} u_j, \quad i = 1, \dots, n. \quad (2)$$

수식 (2)로부터 선형 시스템 $Au=b$ 으로 재정의 할 수 있다. 즉, 가중치 A 는 $n \times n$ 행렬이며, u 는 2차원 도메인 상의 점이며, b 는 u 의 이웃점들의 합을 나타낸다. 다음은 선형 시스템을 행렬로 표현한 것이다.

$$\text{where } A = \begin{bmatrix} 1 & -\lambda_{12} & \dots & -\lambda_{1n} \\ -\lambda_{21} & 1 & \dots & -\lambda_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -\lambda_{n1} & -\lambda_{n2} & \dots & 1 \end{bmatrix}, a_{ij} = \begin{cases} 1 & , j=i \\ -\lambda_{ij} & , j \in N_i \\ 0 & , \text{otherwise} \end{cases}$$

$$\text{and } u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}, b = \begin{bmatrix} \sum_{j \in N_1} \lambda_{1j} u_j \\ \sum_{j \in N_2} \lambda_{2j} u_j \\ \vdots \\ \sum_{j \in N_n} \lambda_{nj} u_j \end{bmatrix}. \quad (3)$$

수식 (3)에서 가우스 소거법(Gauss Elimination)으로 행렬 A 의 역행렬을 계산 및 이용하여 u 를 계산한다. 이와 같은 방법으로 3차원 모델 데이터를 2차원 데이터로 차원 축소를 수행한다. (그림 4)는 3차원 손 모델을 2가지 단계로 구성된 메쉬없는 매개변수화로 수행하였을 때 결과이다.

3.2 실린더형 외곽점을 이용한 메쉬없는 매개변수화

메쉬없는 매개변수화 방법은 3차원 제스처 인식의 차원 축소 방법으로 적합 하지 못하다. 그 이유는 3차원 포즈 데이터의 생성 방법에 따라 크고 작은 에러가 발생하며, 메쉬 없는 매개변수화 방법은 작은 에러에도 결과에 많은 영향을 미치기 때문이다. (그림 5)는 차원 축소시 외곽점이 어떠한 영향을 미치는지 살펴보기 위한 실험 데이터로써, 동일한 3차원 포즈 데이터에서 외곽점만 조금 다른 데이터이다. (그



(그림 4) 기존의 매개 변수화를 이용한 3차원 손 포즈의 2차원 표현 결과



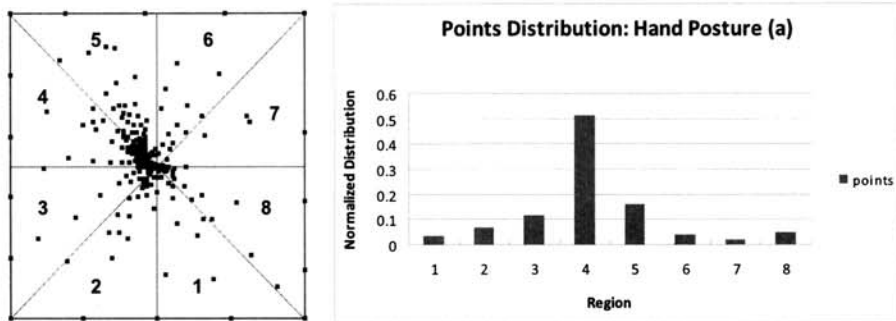
실험 데이터 (a)

실험 데이터 (b)

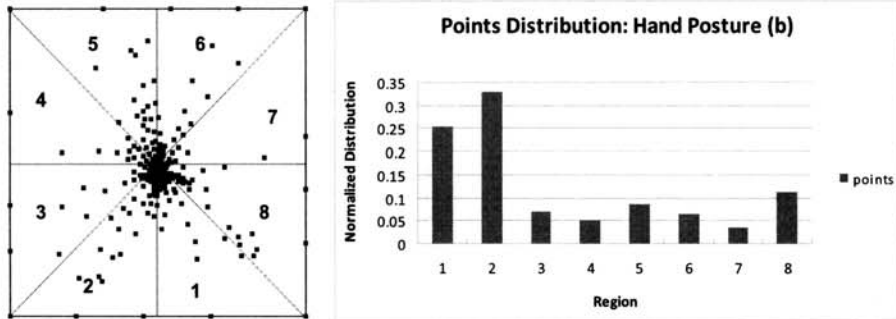
(그림 5) 동일한 3차원 데이터에서 외관점만 다른 실험 데이터

림 6(a)의 왼쪽과 (그림 6(b)의 왼쪽은 (그림 5(a)와 (그림 5(b)의 차원 축소 결과이며, (그림 6(a)의 오른쪽과 (그림 6(b)의 오른쪽은 (그림 6(a)와 (그림 6(b)의 차원 축소 결과를 8등분하여 데이터 분포를 표현한 그래프이다. (그림 6(a)와 (그림 6(b)의 데이터 분포를 비교하였을 때, 외관점의 작은 차이가 차원 축소 데이터 분포에 많은 영향을 끼치므로, 데이터 분포의 정확성을 요구하는 제스처 인식에 적용할 수 없는 문제를 가진다.

이러한 문제를 해결하기 위해서 실린더형 가상의 외관점을

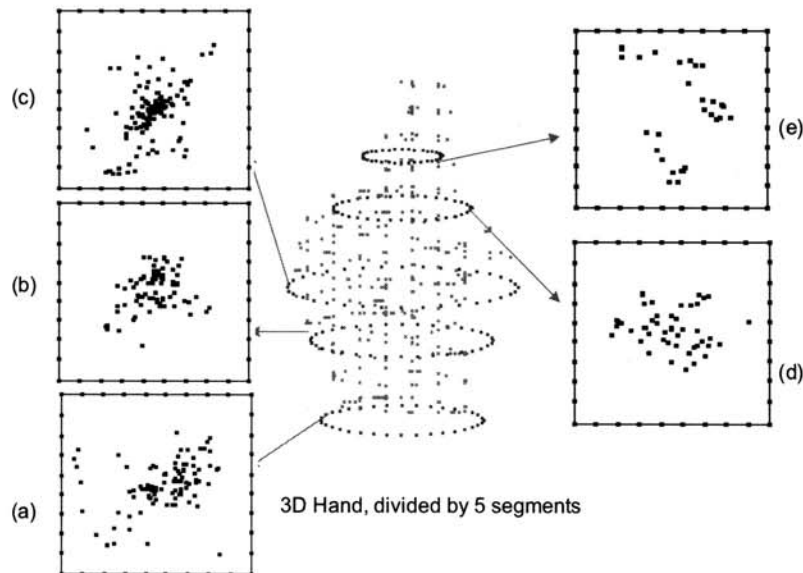


실험 데이터 a의 차원 축소 결과 및 데이터 분포도(a)



실험 데이터 b의 차원 축소 결과 및 데이터 분포도(a)

(그림 6) 실험 데이터의 결과



(그림 7) 실린더형 외관점을 이용한 5개의 3차원 복셀 데이터 영역으로 분할:(a-e) 각 실린더형 외관 영역의 3차원 데이터를 2차원 데이터로 변환 결과

매쉬없는 매개변수화 방법에 적용하였다. 즉, 가상 외곽점은 어떠한 3차원 포즈 데이터라도 동일한 외곽점으로 2차원 도메인 데이터를 계산하기 때문에 에러를 포함한 데이터라도 보다 정확한 데이터 분포로써 제스처 인식에 사용될 수 있다.

제안된 방법의 주요 아이디어는 3차원 포즈 데이터를 하나의 실린더로 표현하며, 실린더를 z축을 기준으로 5개의 영역을 분할 및 가상의 외곽점을 이용하여 매쉬없는 매개변수화 방법을 수행한다. 가상의 외곽점은 실린더 중심에서 각 영역마다 가장 먼 거리에 있는 점의 거리를 반지름으로 36개의 외곽점을 10도의 간격으로 생성 및 매쉬없는 매개변수화 방법을 사용하여 5개의 3차원 데이터를 각 2차원 도메인 평면으로 변환한다. (그림 7)는 3차원 손 모델을 실린더형 외곽으로 적용한 결과로써, 5개의 2차원 도메인을 생성한다.

4. 포즈 매칭(Posture Matching)

본 장에서는 생성된 5개의 2차원 도메인에서 특징 추출과 포즈 인식을 위한 매칭 방법으로, 4.1장은 5개의 도메인에서 특징 추출을 위한 방법을 소개하며, 4.2장은 매칭시 각 데이터 방향을 일치 시키기 위한 시작영역 동기화 및 매칭 방법을 기술한다.

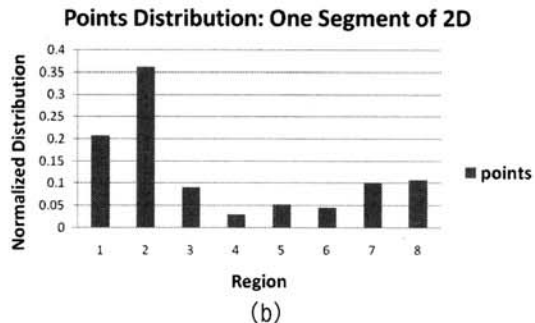
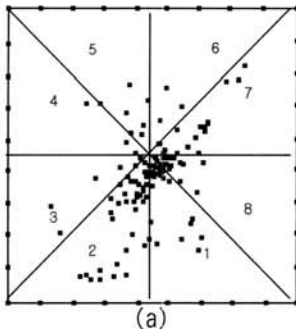
4.1 멀티 도메인을 이용한 특징 추출

특징 추출은 각 2차원 도메인을 8개의 영역으로 분할하여 영역별 데이터 분포 수를 계산하는 것으로써, 각 포즈를 쉽고 간단하게 구분할 수 있다. 그 이유는 3차원 데이터는 포즈마다 분포되는 데이터가 다르며 차원 축소한 2차원 데이터 분포 또한 다르기 때문이다.

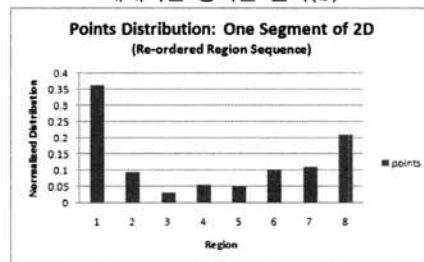
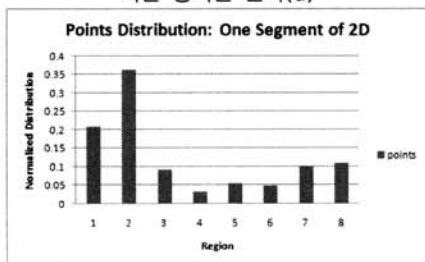
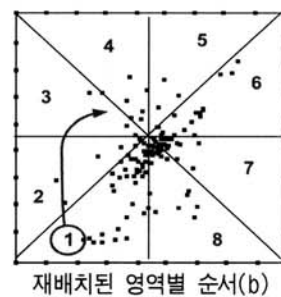
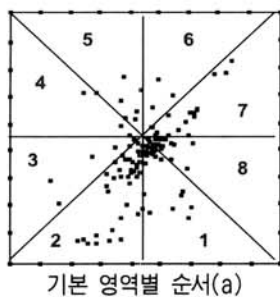
각 2차원 데이터는 8개의 영역으로 분할하여 영역별 분포량을 특징 벡터로 사용한다. (그림 8)(a)는 실제 하나의 2차원 도메인을 분할한 그림으로써 분포량을 그래프로 표현하면 (그림 8)(b)와 같다. 5개의 2차원 도메인은 각각 8개의 영역으로 분할 및 분포를 분석하여 비교할 포즈마다 데이터를 매칭함으로써 쉽고, 빠르게 포즈를 구분할 수 있다.

4.2 2차원 도메인의 시작영역 동기화 및 매칭

제안된 방법은 3차원 포즈 데이터를 어떤 방향 순서로 계산하느냐에 따라 같은 모션이라도 2차원 도메인 집합 U 가 다르게 분포된다. 그러므로 같은 포즈의 2차원 도메인 데이터라도 방향에 따라 다른 포즈라고 인식하는 문제가 발생하기 때문에 매칭되는 데이터의 방향을 일치시켜야 한다. 즉, 2차원 도메인은 각각 비교할 대상의 가장 많은 분포를 가지는 영역을 기준으로 데이터 분포를 재 정의 하는 과정을 시작영역 동기화라 한다. (그림 9)는



(그림 8) 하나의 2차원 도메인의 8개 영역으로 분할 및 데이터 분포



기본 영역별 데이터 분포(c)

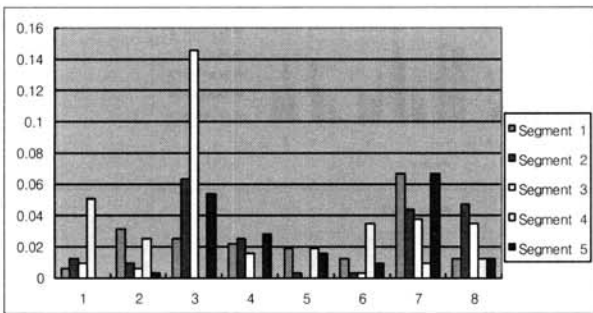
재배치된 영역별 데이터 분포(d)

(그림 9) 시작영역 동기화 과정

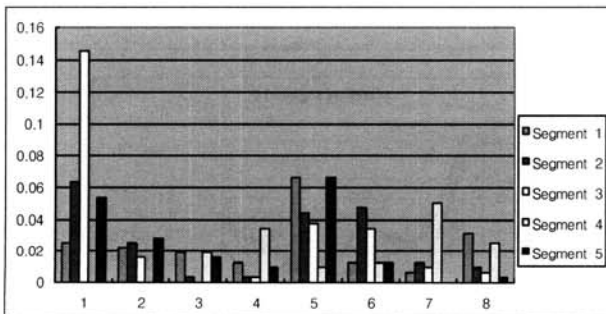
하나의 2차원 도메인을 이용하여 시작영역 동기화 과정을 보여주는 예로써, (그림 9)(a)에서 가장 많은 데이터를 가지는 영역을 선택하여 (그림 9)(b)와 같은 순으로 재배치 하는 것을 말한다. 보다 정확한 결과는 (그림 9)(c)에서 (그림 9)(d)의 데이터 분포 변화를 살펴보면 알 수 있다.

(그림 10)은 시작영역 동기화 수행 전의 데이터 분포를 그래프로 나타낸 그림으로써, x축은 5개 도메인의 8개 영역 데이터 분포를, y축은 각 영역 데이터를 총 데이터 수로 나눈 값을 나타낸다. 이때 가장 많은 데이터를 가지는 영역은 3번째 영역으로써, 이 영역을 첫 번째 영역으로 재배치된 결과를 보여준다(그림 11).

매칭 과정은 시작영역 동기화 과정이 완료된 2개의 제스처 데이터를 각 영역에 맞게 순서대로 차이를 계산하는 것이다. 2차원 도메인 데이터는 포즈에 따라 분포되는 데이터가 다르기 때문에 차이값이 크면 클수록 서로 다른 포즈이며, 작으면 작을수록 비슷한 포즈이다. 보다 자세한 매칭 과정은 제 5장의 실험 결과에서 예를 통해서 기술한다.



(그림 10) 시작영역 동기화 수행 전 데이터 분포



(그림 11) 시작영역 동기화 수행 후 데이터 분포

5 실험 결과

방대한 데이터를 가지는 3차원 데이터를 2차원 데이터로 축소하여 포즈 인식에 중요한 매칭과정의 계산량을 줄이고 처리속도를 높이고자 제안된 방법으로써, 효율성 검증을 위해서 간단한 손 포즈 인식 및 인간의 포즈 인식에 적용하였다.

5.1 인공 손 모델을 이용한 포즈 인식

손 포즈 인식에 사용하기 위한 데이터는 3차원 형상 복원

기술로써 4개의 포즈를 3차원 데이터로 생성하였다. 생성된 데이터는 다음 사이트에서 다운받을 수 있다[http://hci.ssu.ac.kr/dwkyoung/dwkyoung_research.html]. 포즈 인식을 위해서 분류하고자 하는 손 모델을 제안된 방법으로 2차원 데이터를 생성 및 데이터베이스에 저장한다. 실제 포즈 인식을 위해서 생성된 3차원 데이터는 제안된 방법으로 2차원 데이터로 변환하여 데이터베이스에 저장된 데이터와 비교하여 가장 적은 차이를 가지는 데이터를 이용하여 포즈를 인식한다. 실제 검증을 위해 손 포즈 인식에 사용하였다.

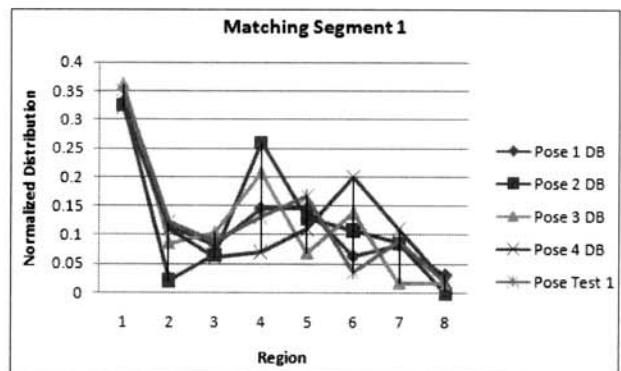
(표 1)은 데이터베이스에 저장될 손 포즈 데이터이며, (표 2)는 실제 매칭을 위한 테스트 데이터이다. (표 1과 표 2)는 손 포즈 인식에 사용된 포즈영상과 3차원 데이터를 제안된 방법을 사용하여 2차원 도메인 데이터로 구성되며, 그래프 x축의 값은 2차원 도메인의 8개의 영역을 나타내며, y축은 데이터 분포를 나타내며, segment1~segment5는 5개의 2차원 데이터를 의미한다.

포즈를 분류하기 위한 과정은 (그림 12,13)과 같다. (그림 12)은 데이터베이스에 저장된 전체 포즈 데이터와 테스트 포즈 1번 데이터를 비교하는 과정으로써, 하나의 2차원 도메인에서 영역별 분포의 차이를 보여준다. (그림 13)은 (그림 12)과 같이 나머지 2차원 도메인의 분포도를 보여준다.

(그림 12,13)과 같은 과정으로 테스트 포즈와 데이터베이스의 포즈와의 차이값을 구할 수 있다. (그림 14)는 테스트 포즈 1번과 데이터베이스에 저장된 모든 포즈와의 차이값을 그래프로 표현하였으며, 이와 같은 결과를 이용하여 테스트 포즈 1은 5개의 2차원 데이터 차이값이 가장 작은 포즈 1로 분류된다.

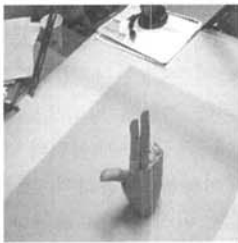
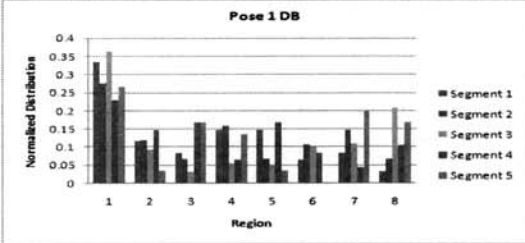

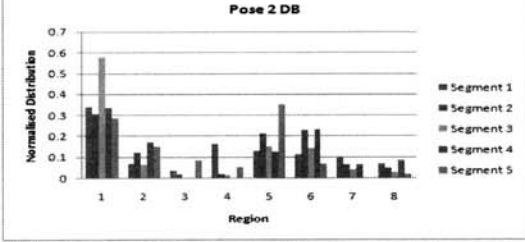

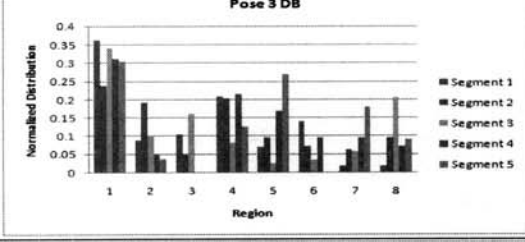

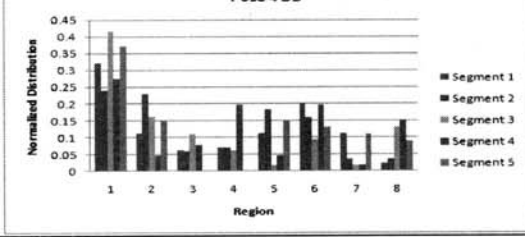
(그림 15)는 테스트 포즈 2번과 데이터베이스에 저장된 모든 포즈와의 차이값을 그래프로 표현하였으며, 이와 같은 결과를 이용하여 테스트 포즈 2는 5개의 2차원 데이터 차이값이 가장 작은 포즈 4로 분류된다.

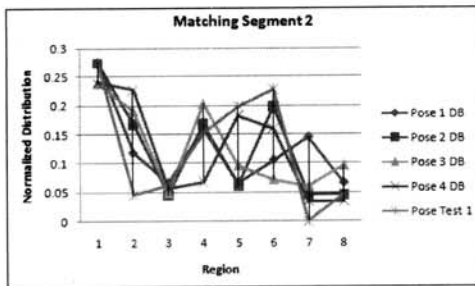
제안된 방법은 미리 정의된 포즈와 생성된 사용자 포즈를 매칭하여 포즈를 구분한다. 즉, 동일할수록 데이터의 차이는 작아지며, 동일하지 않을수록 데이터 차이는 점점 크게 된다. 이와 같이 방대한 양의 데이터를 제안된 차원 축소 방법을 통해서 2차원 데이터로 변환하여 쉽게 특징 추출을 수



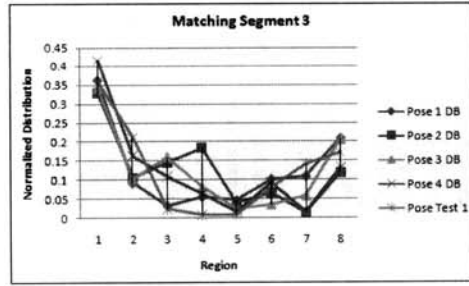
(그림 12) 테스트 포즈 1번과 데이터베이스의 포즈 데이터의 첫 번째 2차원 도메인 분포도

〈표 1〉 손 모델의 데이터베이스 및 시각영역 동기화 과정을 수행한 분포도

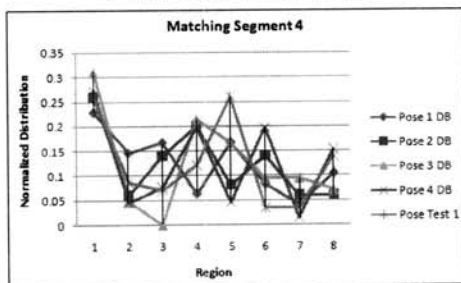
No.	Pose DB	Re-ordered Distribution of 2D Graph
1		
2		
3		
4		



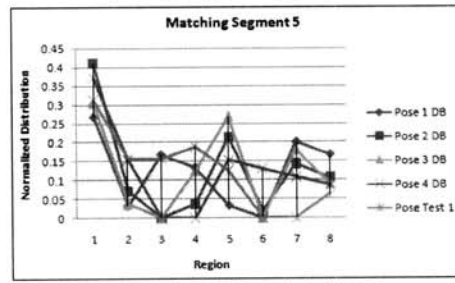
두 번째 2차원 도메인 분포도



세 번째 2차원 도메인 분포도




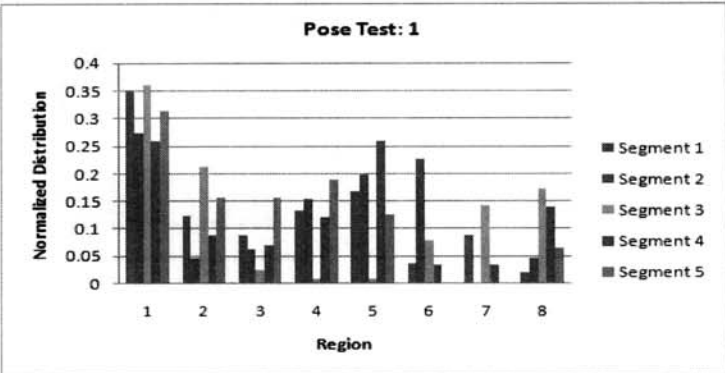
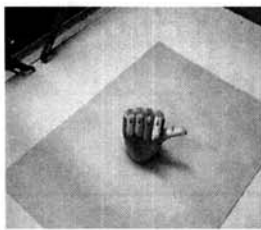
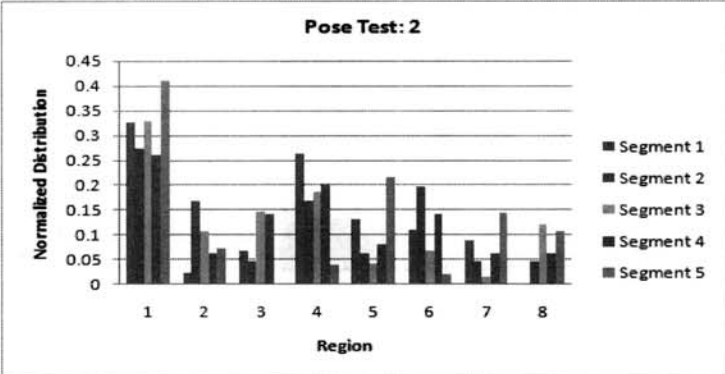
네 번째 2차원 도메인 분포도

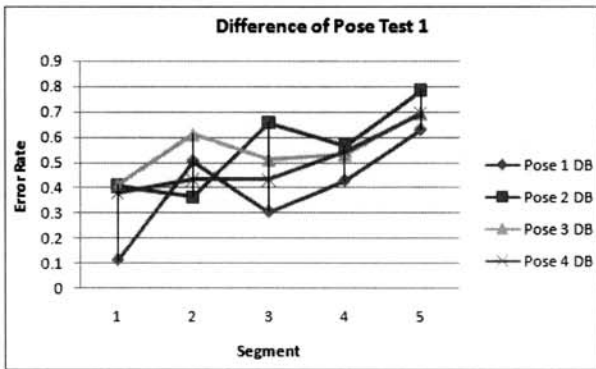


다섯 번째 2차원 도메인 분포도

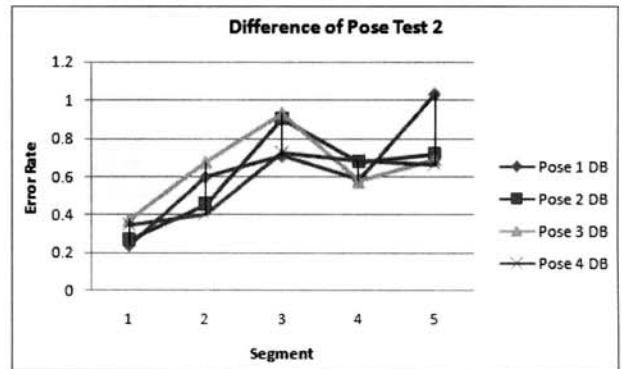
(그림 13) 테스트 포즈 1번과 데이터베이스의 포즈 데이터의 2차원 도메인 분포도

〈표 2〉 테스트를 위한 손 모델과 시작영역 동기화 과정을 수행한 분포도

No.	Pose Test	Re-ordered Distribution of 2D Graph
1		 <p>Pose Test: 1</p> <p>Normalized Distribution</p> <p>Region</p> <p>Legend: Segment 1, Segment 2, Segment 3, Segment 4, Segment 5</p>
2		 <p>Pose Test: 2</p> <p>Normalized Distribution</p> <p>Region</p> <p>Legend: Segment 1, Segment 2, Segment 3, Segment 4, Segment 5</p>



(그림 14) 테스트 포즈 1번 데이터와 데이터베이스의 데이터 차이값 분포도



(그림 15) 테스트 포즈 2번 데이터와 데이터베이스의 데이터 차이값 분포도

행하며, 제스처 인식에서 단순한 데이터 양을 비교함으로써 빠르게 포즈를 구분할 수 있다.

5.2 인간 모델을 이용한 포즈 인식

인간 모델을 이용한 포즈 인식은 인간 모델의 3차원 형상 복원, 3차원 데이터의 차원 축소 그리고 데이터 베이스에 저장된 데이터와의 매칭과정으로 구성된다. 본 실험에서는 실시간으로 생성되는 인간 모델의 3차원 데이터를 2차원으로 차원 축소하여 데이터베이스에 저장된 4가지 포즈의 2차원 데이터를 비교하여 포즈 인식을 수행한다.

인간 모델을 이용하기 위한 3차원 형상 복원은 멀티 카메라에서 획득되는 영상의 동기화 과정, 획득한 영상에서 오

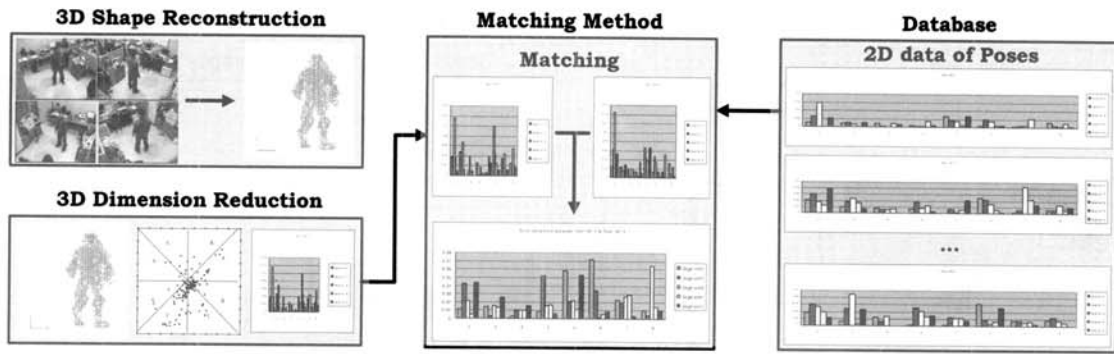
브젝트를 추출하는 실루엣 추출(silhouette extraction) 과정, 그리고 평면 기반의 볼륨 인터섹션(plane-based volume intersection) 방법을 이용한 3차원 포즈 데이터 생성과정으로 구성된다.

(그림 17)과 같이 인간 모델로 3차원 형상 복원을 이용하여 생성된 사용자의 3차원 데이터를 3.2장의 방법을 이용하여 5개의 2차원 데이터를 계산한다(그림 18).

(그림 19)은 2차원으로 차원 축소된 5개의 데이터를 8개의 영역으로 분할하여 데이터 량을 그래프로 나타내었다.

이와 같은 방법으로 아래의 정의된 4가지 포즈를 2차원 데이터 생성 및 데이터베이스에 저장한다(그림 20)

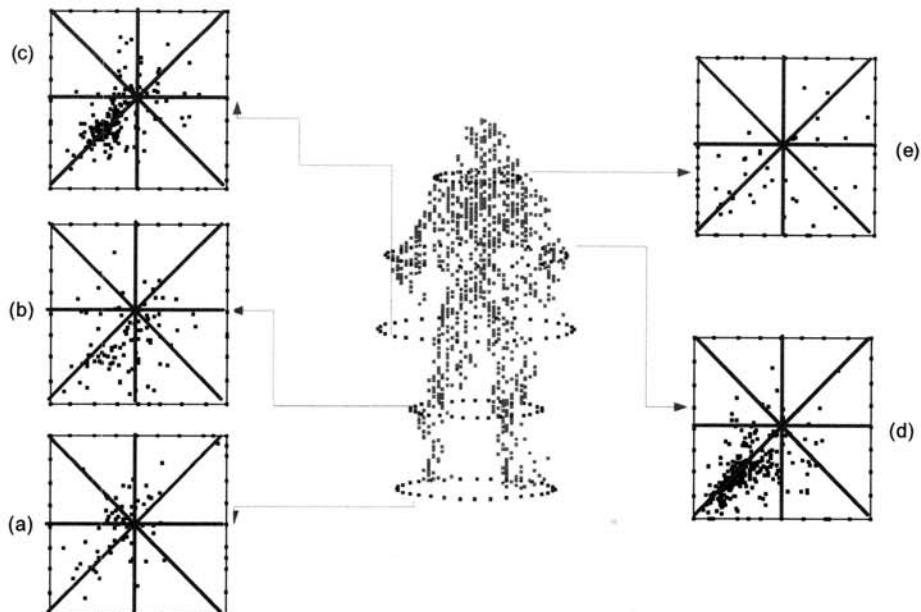
(그림 21)은 4가지 포즈에 대한 데이터베이스에 저장되는



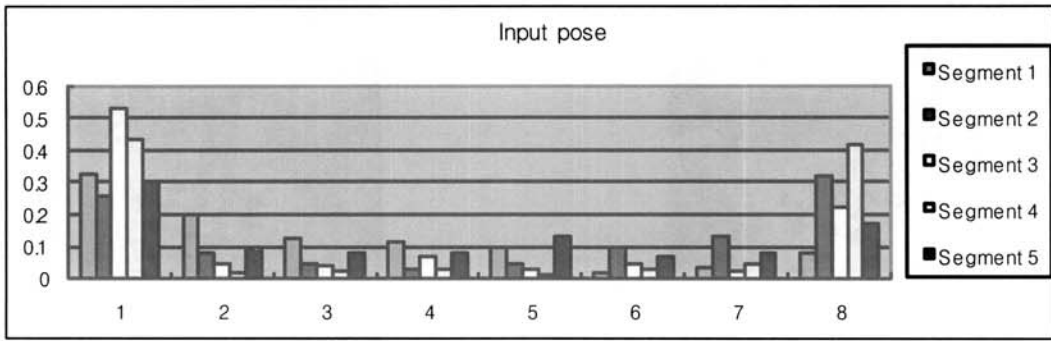
(그림 16) 인간 모델을 이용한 포즈 인식의 처리과정



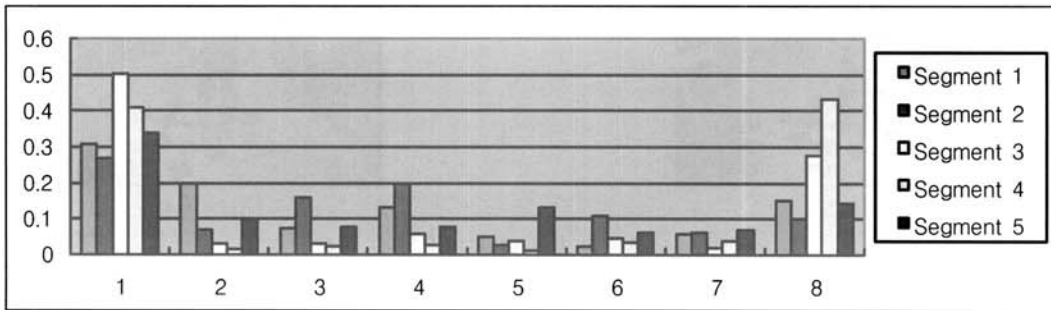
(그림 17) 4대의 카메라를 이용한 캡처 이미지



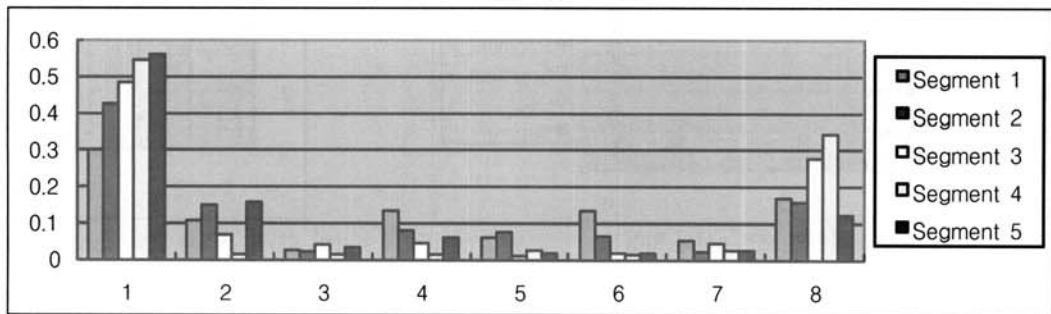
(그림 18) 3차원 형상 복원 및 차원 축소의 결과영상: (a-e) 각각의 실린더형 외곽점을 기반으로 3차원 복셀 데이터를 2차원 데이터로 변환



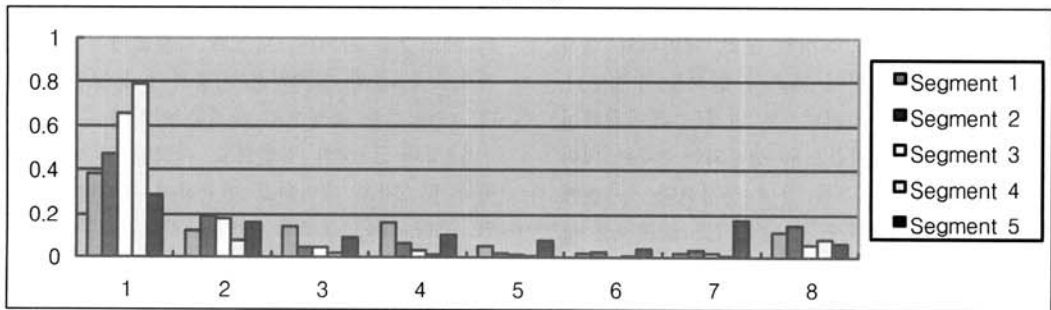
(그림 19) 각 영역의 2차원 데이터를 그래프로 통합한 결과



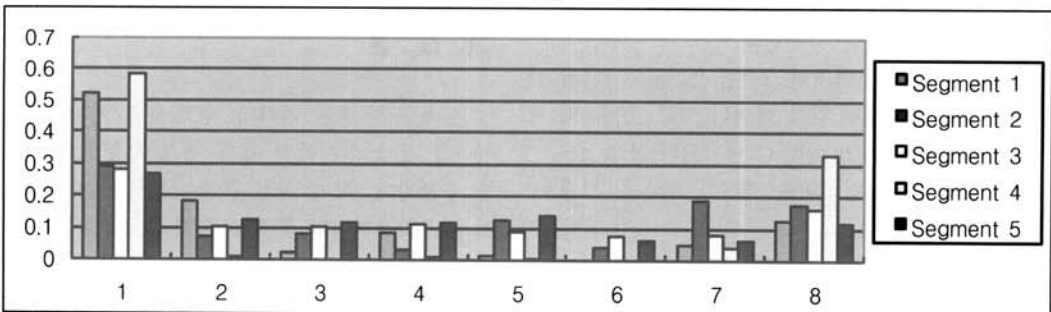
포즈 A



포즈 B

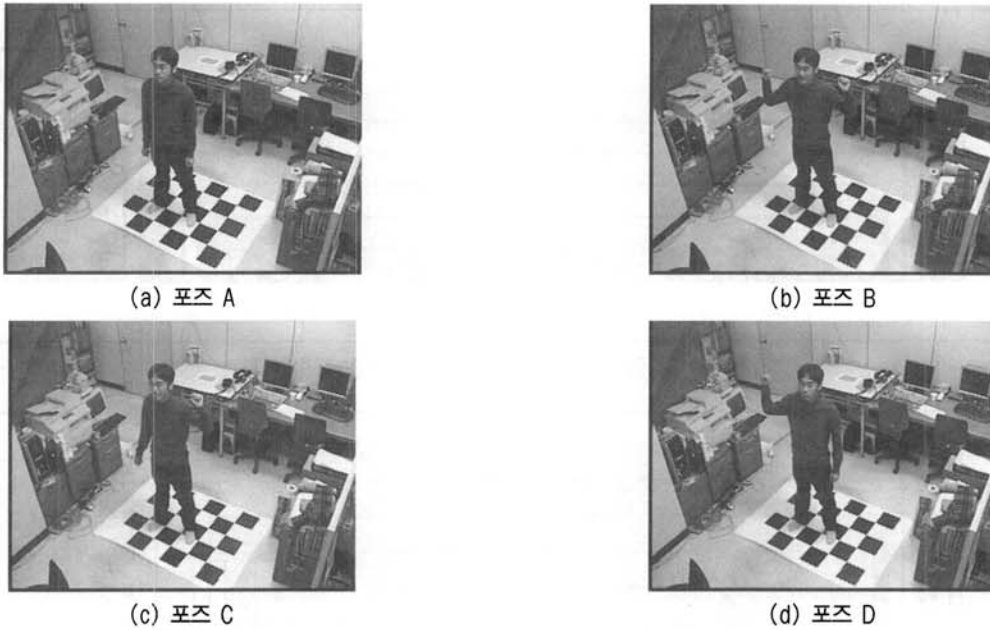


포즈 C

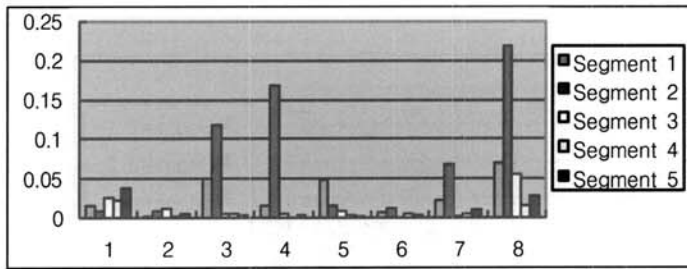


포즈 D

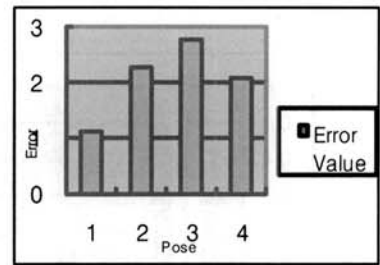
(그림 20) 정의된 4가지 포즈



(그림 21) 정의된 포즈 데이터의 그래프 표현



(a) 입력 데이터와 포즈 a와의 차이 값



(b) 4가지 포즈와 입력 포즈와의 매칭 결과

(그림 22) 입력 데이터와 데이터베이스의 매칭 결과

데이터를 그래프로 표현한 그림이다.

본 시스템은 데이터베이스에 저장된 포즈 데이터와 실시간으로 사용자의 포즈를 매칭하여 제스처 인식을 수행한다. 매칭 방법은 2가지 단계로 구성되며, 첫 번째는 데이터베이스의 모든 2차원 데이터와 실시간으로 생성되는 2차원 데이터를 비교하여 에러가 적은 포즈를 하나 선정한다. 두 번째는 선정된 포즈가 정확한지 정확하지 않은지를 임계치를 통해서 결정한다. 실제 시스템은 실시간으로 생성되는 포즈와 일치되지 않아도 가장 작은 에러값을 가지는 포즈가 선정되므로 오작동의 위험이 너무 크다. 그러한 오작동의 에러를 최소화 하기 위해서 임계치를 사용한다.

(그림 22)는 매칭 결과를 보여주며, (그림 22)(a)는 (그림 21)(a)의 데이터베이스 데이터와 입력 데이터(그림 20)를 매칭한 결과를 보여준다. (그림 22)(b)는 (그림 21)의 포즈 a~d의 데이터와 입력 데이터(그림 19)를 매칭한 결과를 그래프로 표현한 것이다.

(그림 22)(b)처럼 비슷한 포즈 일수록 매칭시 차이값이 적고, 포즈가 다를수록 차이값이크다. 선정된 포즈는 입력된 포즈와 다를 수 있기 때문에 임계치를 이용하여 동일한 포즈인지를 확인한다. 본 시스템에서 임계치는 1.5로 사용하였

다. 그 이유는 (그림 22)(b)에서 동일한 포즈일때에는 약 1의 에러값을 가지며, 포즈가 다를때에는 2보다 큰 값을 가짐으로 1.5값의 임계치를 사용하여 동일한 포즈인지 결정한다. 이와 같은 과정으로 포즈를 인식할 수 있다.

5.1장과 5.2장의 실험으로 다양한 포즈를 사용자로부터 획득 및 2차원 데이터로 변환하여 데이터베이스에 저장한다면, 보다 많은 포즈를 인식할 수 있을뿐만 아니라, 제안된 포즈 인식 방법을 시간상으로 분류하여 제스처 인식에 사용될 수 있다.

6. 결 론

본 논문은 많은 데이터 및 정보를 가지는 3차원 포즈 데이터를 포즈 인식에서 보다 쉽게 사용하기 위해 차원 축소를 수행한다. 이때 차원 축소 시에는 3차원의 정보의 손실이 최소화하여야만 정확한 특징 추출 및 인식에 사용할 수 있기 때문에 정확하면서도 빠른 처리가 가능한 실린더형 외곽점을 이용한 매쉬없는 매개변수화 방법을 제안하였다. 제안된 방법은 3차원 데이터를 2차원 데이터로 변환하여 쉽게 특징 추출을 수행한다. 즉, 2차원 도메인 평면의 데이터 분

포즈 카운트 하는 것으로 특징 추출이 이루어 지기 때문에 빠른 처리 및 쉬운 매칭이 가능하다. 하지만 실린더형 외곽의 매쉬없는 매개변수화는 2차원 데이터 분포의 랑을 특징 벡터로 사용하기 때문에 세밀한 포즈를 구분하기는 힘들다. 앞으로는 보다 미세한 포즈를 구분하기 위한 연구가 필요하다.

참 고 문 헌

- [1] E. Foxlin, "Motion Tracking Requirements and Technologies," Handbook of Virtual Environments: Design, Implementation, and Applications. K. Stanney (Ed.), Lawrence Erlbaum Associates, pp.163-210, 2002.
- [2] Thomas S. Huang and Vladimir I. Pavlovic, "Hand Gesture Modeling, Analysis, and Synthesis," Proceedings of International Workshop on Automatic Face-and Gesture-Recognition, Zurich, pp.26-28, 1995.
- [3] R. Zeleznik, J. LaViola, D. Acevedo and D. Keefe, "Pop-Through Bottons for Virtual Environment Navigation and Interaction," Proceedings of IEEE Virtual Reality, pp.169-134, 2002.
- [4] Qiulei Dong, Yihong Wu and Zhanyi Hu, "Gesture Recognition Using Quadratic Curves," Proceedings of the 15th International Conference on Pattern Recognition, pp.817-825, 2006.
- [5] Mori G., Xiaofeng Ren, Alexei A. Efros and Jitendra Malik, "Recovering Human Body Configurations: Combining Segmentation and Recognition," Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol.2, pp.326-333, 2004.
- [6] Pengyu Hong, Matthew Turk and Thomas S. Huang, "Gesture Modeling and Recognition Using Finite State Machines," Proceedings of IEEE Conference on Face and Gesture Recognition, pp.410-415, 2000.
- [7] Chi-Wei Chu and Isaac COHEN, "Posture and Gesture Recognition using 3D Body Shapes Decomposition," Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp.69-69, 2005.
- [8] Guangqi Ye, Jason J. Corso and Gregory D. Hager, "Gesture Recognition Using 3D Appearance and Motion Features," Proceeding of Computer Vision and Pattern Recognition Workshops, pp.160-160, 2004.
- [9] Yoichi Sato, Makiki Saito and Hideki Koike, "Real-time Input of 3D Pose and Gestures of a User's Hand and Its Applications for HCI," Proceeding of IEEE Virtual Reality Conference, pp.79-86, 2001.
- [10] Y. Lee, D. Kyoung, E. Han and K. Jung, "Dimension Reduction in 3D Gesture Recognition Using Meshless Parameterization," Proceedings of Pacific-Rim Symposium on Image and Video Technology, pp.64-73, 2006.
- [11] Lee W. Campbell, David A. Becker, Ali Azarbayejani, Aaron F. Bobick and Alex Pentland, "Invariant Features for 3-D Gesture Recognition," Proceedings of the Second International Conference on Automatic Face and Gesture Recognition, pp.157-162, 1996.
- [12] Thomas B. Moeslund and Erik Granum, "A Survey of Computer Vision-Based Human Motion Capture," Computer Vision and Image Understanding, pp.231-268, 2001.
- [13] Malassiotis S., Aifanti N. and Strintizis M.G., "A Gesture Recognition System Using 3D Data," Proceedings of the First International Symposium on 3D Data Processing Visualization and Transmission, pp.190-193, 2002.
- [14] Ho-Kuen Shin, Sang-Woong Lee and Seong-Whan Lee, "Real-Time Gesture Recognition Using 3D Motion History Model," Proceedings of ICIC 2005, pp.888- 898, 2005.
- [15] R.Fablet and M.J. Black, "Automatic Detection and Tracking of Human Motion with a View-Based Representation," Proceedings of 7th European Conference on Computer Vision, pp.476-491, 2002.
- [16] Floater M. S., "Parameterization and B-spline Surface Approximation," The Mathematics of Surfaces IX, R. Cipolla and R. Martin (eds.). Springer-Verlag, pp.1-18, 2000.
- [17] Floater M. S. and Reimers M., "Meshless Parameterization and Surface Reconstruction," Computer Aided Geometric Design, pp.77-92, 2001.
- [18] Floater M. S. and Hormann K., "Surface Parameterization: a Tutorial and Survey," Advances in Multiresolution for Geometric Modelling, pp.157-186, 2004.
- [19] R. Jain, R. Kasturi and B. G. Schunck, "Machine Vision," McGraw-Hill, 1995.
- [20] E. D. Aguiar, C. Theobalt, M. Magnor, H. Theisel and H. Seidel, "Marker-free Model Reconstruction and Motion Tracking from 3D Voxel Data," Proceedings of The 12th Pacific Conference on Computer Graphics and Applications, pp.101-110, 2004.
- [21] C. Theobalt, J. Carranza, M. A. Magnor and H. Seidel, "Combining 3D flow fields with silhouette-based human motion capture for immersive video," Graphical Models, vol.66, pp.333-351, 2004.



경 동 욱

e-mail : kyoungd@inf.ethz.ch

2002년 동의대학교 산업공학과(공학사)

2005년 부산대학교 대학원 컴퓨터공학과
(공학석사)

2007년 숭실대학교 대학원 IT대학 미디어
학부 미디어공학(공학박사)

2007년 숭실대학교 IT대학 미디어학부 박사후연구원

2007년 11월~현재 CVG Lab, ETH Zurich, Switzerland, 박사
후연구원

관심분야 : 컴퓨터 비전, HCI, 영상처리, 인공지능, 하드웨어 설
계, 정보보호



정 기 철

e-mail : kcjung@ssu.ac.kr

1996년 경북대학교 컴퓨터학과 (공학석사)

1999년 방문 연구원 Intelligent User Inter-
face Group, DFKI (The German
Research Center For Artificial In-
telligence GmbH), Germany

1999년 방문 연구원 Machine Understanding Division, Electro
Technical Laboratory, Japan

2000년 경북대학교 컴퓨터학과 (공학박사)

2000년~2002년 PRIP Lab. Michigan State University, U.S. 박
사후연구원

2003년~현재 숭실대학교 IT대학 미디어학부 교수

관심분야 : HCI, Interactive Contents, 영상처리/컴퓨터비전, 패
턴인식, 증강현실, 인공지능



이 윤 리

e-mail : yunli@ssu.ac.kr

2002년 Multimedia 대학교 IT학과(학사)

2004년 동서대학교 대학원 소프트웨어학
과(공학석사)

2004년~2006년 Multimedia 대학교 IT학
부 강사

2006년~현재 숭실대학교 대학원 미디어학과 박사과정

관심분야 : 컴퓨터비전, 패턴인식