

모션 데이터의 계층적 가시화에 의한 3차원 아바타의 표정 제어

김 성 호* · 정 문 렬**

요 약

본 논문은 사용자 하여금 계층적 가시화 기법에 의한 표정들의 공간으로부터 일련의 표정을 선택하게 함으로써 3차원 아바타의 표정 제어 기법을 기술한다. 본 시스템에서는 2,400여개의 표정 프레임에 의하여 2차원 표정공간을 구성하였으며, 3차원 아바타의 표정 제어는 사용자가 표정공간을 항해하면서 수행한다. 그러나 표정의 수가 너무 많아 사용자가 항해를 하는데 어려움이 많기 때문에 계층적 가시화 기법이 필요하다. 표정공간을 계층적으로 분할하기 위해, 퍼지 클러스터링을 이용한다. 초기 단계에서는 2,400여개의 표정들을 이용하여 약 11개의 클러스터 센터를 가지도록 클러스터링한다. 클러스터 센터들은 2차원 평면에 표시되며 후보 키 프레임으로 사용된다. 사용자는 후보 키 프레임들 중에서 특정 키 프레임들을 선택하여 초기 항해경로를 생성한다. 사용자가 줌 인(이산적인 단계를 가지고 있음)을 하면 상세한 단계를 보기를 원한다는 의미므로 줌 인된 단계에 적합한 표정 클러스터들을 생성한다. 단계가 증가될 때 마다 클러스터의 수를 두 배로 하고, 클러스터의 수만큼 표정들을 클러스터링한다. 사용자는 현재 단계에서 그진 단계의 항해경로를 따라 새로운 키 프레임(클러스터 센터)들을 선택하여 항해경로를 갱신한다. 줌 인을 최대로 한 마지막 단계에서 항해경로를 갱신하면 표정 제어 설정이 끝난다. 사용자는 언제든지 줌 아웃을 통해 그 전단계로 돌아가서 항해경로를 수정할 수 있다. 본 논문은 본 시스템이 어떤 효과가 있는지를 알기 위해 사용자들로 하여금 본 시스템을 사용하여 3차원 아바타의 표정 제어를 수행하게 하였으며, 그 결과를 평가한다.

Facial Expression Control of 3D Avatar by Hierarchical Visualization of Motion Data

Sung-Ho Kim* · Moon-Ryul Jung**

ABSTRACT

This paper presents a facial expression control method of 3D avatar that enables the user to select a sequence of facial frames from the facial expression space, whose level of details the user can select hierarchically. Our system creates the facial expression space from about 2,400 captured facial frames. But because there are too many facial expressions to select from, the user faces difficulty in navigating the space. So, we visualize the space hierarchically. To partition the space into a hierarchy of subspaces, we use fuzzy clustering. In the beginning, the system creates about 11 clusters from the space of 2,400 facial expressions. The cluster centers are displayed on 2D screen and are used as candidate key frames for key frame animation. When the user zooms in (zoom is discrete), it means that the user wants to see more details. So, the system creates more clusters for the new level of zoom-in. Every time the level of zoom-in increases, the system doubles the number of clusters. The user selects new key frames along the navigation path of the previous level. At the maximum zoom-in, the user completes facial expression control specification. At any moment, the user can go back to previous level by zooming out, and update the navigation path. We let users use the system to control facial expression of 3D avatar, and evaluate the system based on the results.

키워드 : 얼굴 모션 캡처(Facial Motion Capture), 얼굴 표정 제어(Facial Expression Control), 항해(Navigation), 퍼지 클러스터(Fuzzy Clustering), 계층적 가시화(Hierarchical Visualization), 표정 제어(Facial Expression Control)

1. 서 론

인간은 자신의 감정을 말보다는 얼굴 표정을 통해서 더 잘 표현하며, 상대방의 얼굴 표정을 보고 상대방의 현재 감정 상태를 파악한다. 그런 연유로 지금까지 3차원 컴퓨터 그래픽스 기법을 통해서 인간의 얼굴 표정을 표현할 수 있는

많은 방법이 연구[1-7]되었다. 특히 인터넷 시스템과 멀티미디어의 기능이 고급화됨으로써 가상공간 속에 사용자 자신을 대신하는 3차원 아바타의 필요성이 대두되었고, 지금은 많은 멀티미디어 분야에서 널리 사용되고 있다. 이런 가상공간에서 사용자들이 3차원 아바타를 표현할 때 가장 관심을 가지는 것은 역시 3차원 아바타의 얼굴 표정이다. 왜냐하면 인간은 표정의 변화를 보고 가장 빠르게 상대방의 감정을 파악하고, 아울러 자신의 감정을 상대방에게 표정을 통하여 실시간으로 보여주기 때문이다.

이에 가상공간에서 3차원 아바타의 표정을 사용자가 제어

* 본 논문은 과학재단 특장기초 연구과제(R01-2002-000-00311-0)의 수행결과 일부임.

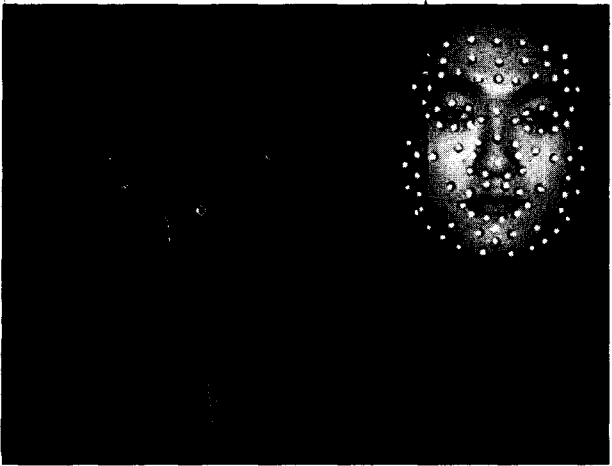
† 준 회원 : 숭의여자대학교 정보통신계열 멀티미디어 콘텐츠전공 교수

** 정 회원 : 서강대학교 미디어공학과 교수

논문접수 : 2004년 3월 17일, 심사완료 : 2004년 6월 28일

할 수 있는 기법이 필요하다. 표정 제어 방법으로는 많은 표정 데이터를 이용하여 생성한 표정공간을 사용자가 항해하는 방법이 유력한데, 이 방법은 [13]에서 사용하였다. 그러나 다량의 표정 데이터들로 구성된 표정공간을 사용자가 항해하면서 적당한 표정을 선택하기가 어렵기 때문에 이 표정공간을 계층적으로 가시화하는 것이 필요하다. 본 논문은 이와 같이 표정공간을 계층적으로 가시화하고 표정을 제어하기 위한 방법을 기술한다. 본 논문의 주제인 계층적 가시화에 의한 표정 제어와 관련하여 모션 캡처 데이터를 사용한 연구는 아직 제시된 바가 없다. 그러므로 본 논문에서는 다량의 얼굴 모션 캡처 데이터를 수 개의 단계를 두어 2차원 평면에 수 개에서 수천 개까지 계층적으로 분포시키고, 사용자가 적당한 계층의 공간을 항해하면서 원하는 얼굴 표정들을 선택하여, 3차원 아바타의 얼굴 표정을 제어하는 방법을 기술한다.

표정 데이터를 사용하여 계층적 가시화 기법에 의한 표정 제어를 수행하기 위해서는 다음과 같은 사전 작업이 필요하다. 먼저 얼굴 표정을 전문적으로 연출하는 배우의 도움을 받아 (그림 1)과 같이 광학식 모션 캡처 시스템을 사용하여 얼굴 표정을 캡처한다. 표정을 캡처할 때, 배우는 얼굴 주 근육 부분에 작은 반사 마커 100개를 부착한다. 그런 다음 배우로 하여금 서로 다른 10개의 얼굴 모션을 연출하게 하고 초당 60 프레임으로 캡처한다. 모든 얼굴 모션은 처음 무표정 상태에서 출발하여 특정 표정까지 진행한 다음, 다시 무표정 상태로 되돌아오는 방법으로 진행한다. 한 개의 마커는 3개의 좌표 값으로 표현되므로 100개의 마커 위치로 표현되는 한, 표정은 300차원의 데이터이다.



(그림 1) 마커 100개 부착위치 및 광학식 모션 캡처 시스템을 사용한 얼굴 모션 캡처 장면

그리고 얼굴 모션 캡처 작업을 수행한 배우의 정면 사진을 이용하여 3차원 얼굴 모델링 작업을 한다. 또한 100개 마커를 3차원 아바타의 얼굴 각 부위(모션 캡처시 마커 부착 위치)에 연결하여 표정 재생이 가능하게 한다. 그리고 2,400

여개의 얼굴 표정 데이터에서 임의의 두 표정간의 최단경로를 결정하여 표정공간을 생성[11]하고, 이 공간을 다차원 스케일링(Multidimensional Scaling, 이하 'MDS')[12] 기법으로 2차원 평면에 투영시킨다. 본 논문에서는 이 과정들을 미리 구축하였으며, 2차원 평면에 투영된 2,400여개의 MDS 결과 데이터를 실험에 그대로 사용한다.

계층적 가시화는 먼저 2,400여개의 표정들을 사용하여 수 개의 클러스터 센터들을 가지도록 클러스터링하여 대표적인 얼굴 표정만을 2차원 평면에 표시한다. 그리고 매 단계마다 클러스터의 수를 두 배로 증가시키고 사용자로 하여금 줌 인 줌 아웃으로 한 단계를 선택하도록 한다. 계층적 가시화의 최대 단계는 전체 표정의 수에 따라 자동적으로 결정된다. 본 논문에서는 계층적 가시화에 사용될 클러스터 센터들을 구하기 위해 퍼지 클러스터(Fuzzy C-Mean 클러스터링, 이하 'FCM') 기법[8-10]을 사용한다. 2차원 평면에 분포된 클러스터 센터들은 얼굴 표정 제어를 위한 후보 키 프레임들로 설정하고, 사용자로 하여금 특정 키 프레임을 선택하게 하여 선택한 순서대로 항해경로를 설정한다.

2. 얼굴 표정공간의 생성 및 MDS

표정공간을 생성하기 위해서는 각각의 얼굴 표정상태를 수치적으로 표현해야 하고, 표정상태는 얼굴에 부착된 마커들 사이의 거리에 의해 결정된다. 표정상태의 표현은 표정들 간의 상대적인 거리 관계를 잘 표현하는 것으로서, 임의의 두 마커 간의 상호거리를 표현하는 "거리행렬"을 이용하여 표정상태를 표현한다. 표정상태를 거리행렬로 표현하면 표정간의 직선거리는 두 거리행렬간의 직선거리로 표현된다. 거리행렬로 표현된 표정상태들의 공간은 임의의 두 거리행렬간의 거리를 결정함으로써 결정된다. 본 논문에서는 거리행렬을 하나의 벡터로 보고, 이 벡터간의 직선거리를 거리행렬간의 직선거리로 사용한다. 즉 임의의 두 마커 간의 거리가 서로 비슷한 두 표정은 서로 인접한 표정으로 간주한다. 표정공간은 임의의 두 표정간의 거리를 두 표정거리행렬간의 직선거리로 정의할 수 있는 벡터공간이 아니다. 한 표정에서 다른 표정으로 옮겨가는 과정은 얼굴의 여러 가지 제약조건으로 말미암아, 복잡한 경로를 거치게 되기 때문이다. 표정공간은 구면과 같은 다양체(Manifold) 공간이고, 이 공간상에서의 거리는 두 점간의 거리를 한 점에서 이 공간을 벗어나지 않으면서 다른 점까지 도달하는 최단경로의 길이로 정의한다. 본 논문에서는 이 다양체 공간을 근사적으로 표현한다. 이를 위해 먼저, 두 거리행렬간의 직선거리가 일정 값 이하인 경우 이 직선거리가 두 표정간의 최단 거리에 대한 근사치라고 간주한다. 이 조건을 만족하는 두 표정을 "인접표정"이라고 한다. 인접표정이 주어지면 한 표정에서 다른 표정까지 바로 이동할 수 있다고 본다. 두 표정

이 인접해 있지 않은 경우에 한 표정에서 다른 표정으로 바로 이동할 수 없고, 그 사이에 있는 인접한 표정들을 통해서만 이동할 수 있다고 가정한다. 그러나 인접표정을 결정하는 한계거리를 미리 알기는 쉽지 않다. 따라서 이는 실험을 통해서 좋은 결과를 내는 한계 값을 정한다. 인접표정을 결정하는 인접거리 임계값은 실험에 의해 최적의 값을 결정해야 한다. 다만, 인접거리 임계값은 임의의 표정에서 다른 임의의 표정으로 이동하는 데 필요한 충분한 수의 인접표정들이 나오도록 설정되어야 한다. 본 논문에서는 실험을 통하여 최적의 인접거리 임계값으로 230mm를 찾았다. 인접표정들이 결정되면, 인접하지 않은 두 표정상태간의 거리는 그 사이에 있는 인접 표정들 간의 거리들을 합하여 구한다. 이를 위해 최단거리를 구하는 플로이드 알고리즘[11]을 이용한다. 이렇게 임의의 두 표정간의 최단거리가 구해지면 해당 다양체 공간이 결정된다. 본 논문에서는 모두 2,400여개의 얼굴표정을 사용하여 다양체 공간을 형성하였다. 그러나 얼굴표정의 다양체 공간은 300차원 공간이다. 따라서 이 공간을 사용자가 행해하면서 원하는 표정을 선택하여 제어할 수는 없다. 그러므로 원래 표정공간의 구조를 근사적으로 표현하는 2차원 또는 3차원으로 공간을 구하여 이 공간을 향해하는 방법을 사용한다. 이를 위해서는 고차원 공간을 저차원 공간으로 차원 축소하여 투영하는 방법이 필요하며, 여기에는 PCA(Principal Component Analysis), MDS, LLE(Locally Linear Embedding) 등이 있다. PCA는 투영된 분산을 최대화하는 인접 구조 보존 방식이고, LLE는 모든 데이터 점으로부터 가장 가까운 인접 데이터의 수를 임의로 정하고 이를 보존하는 인접 데이터 보존 방식이다. 그리고 MDS는 데이터 점들 사이의 Pairwise distance 보존을 사용하는 거리 보존 방식이다. 그러므로 본 논문에서는 표정들 사이의 최단거리를 보존하면서 저차원 공간에 투영하기 위해 MDS 기법[12]을 사용한다. MDS는 고차원 데이터들 사이의 거리가 주어지면, 이 거리들의 분포를 대변하는 좌표들의 집합을 구하는 방법이다. 이때 이 좌표들의 차원은 필요에 따라 미리 정하는데, 본 연구에서는 구한 좌표들을 시각적으로 표현해야 되기 때문에, 2차원 좌표를 사용했다. 그리고 거리 보존 방식인 MDS는 직선거리(Euclidean 거리)를 기본적으로 사용하는데, 본 논문에서는 표정들 사이의 거리가 인접 표정을 거쳐서 계산되기 때문에 다양체 거리를 사용한다.

표정상태를 나타내는 n 개의 다차원 다양체 공간상의 점들을 x_1, \dots, x_n , 임의의 표정상태 x_i 와 x_j 사이의 다양체 거리를 d_{ij} , $i, j=1, \dots, n$ 이라고 하자. d_{ij} 로 구성된 거리행렬을 $\{d_{ij}\}$ 로 표시한다. 행렬 $D = \{d_{ij}\}$ 는 대각 원소들이 0인 대칭행렬이다. 이를 MDS 이론에서 보통 비유사성 대칭행렬이라고 한다. 다차원 표정공간을 근사적으로 나타내는 p 차원 공간상의 점들을 y_1, \dots, y_n 라 하자. 각 y_i 를 상태

(Configuration) 벡터라고 한다. 상태 벡터들은 시각화하여 보여주는 것을 목적으로 하는 경우가 많으므로 p 는 보통 2 또는 3이다. p 차원 점 y_i 와 y_j 간의 거리를 δ_{ij} 이라고 하자. 일반적으로 MDS를 적용하면, p 차원 거리들의 집합 $\{\delta_{ij}\}$ 가 다차원 다양체 거리들의 집합 D 와 가장 근사한 분포를 가지도록 하는 p 차원 평면상의 점들의 집합 $Y = \{y_i\}$ 을 구할 수 있다.

본 논문에서는 2차원 평면상의 점들의 집합 $\{y_i\}$ 을 구하기 위해 Matlab V6.5를 사용하였다. Matlab의 cmdscale 함수에 비유사성 대칭행렬 D 와 차원의 수 2를 입력으로 사용하여 MDS를 구하고, 이를 집합 $\{y_i\}$ 로 사용하였다. 사용자는 이 공간을 향해하여 다양한 표정 제어를 수행한다. 그러나 2차원 표정 공간에 분포된 표정들의 수가 너무 많아 사용자가 향해하기에는 너무 복잡하고 난해하다. 그러므로 본 논문에서는 2,400여개의 얼굴 표정들을 FCM을 사용하여 수 개의 단계로 구분시킨 계층적 가시화 기법을 적용한다. 계층적 가시화 초기 단계에서는 수 개의 표정, 중간 단계에서는 수백 내기 수천 개의 표정, 마지막 단계에서는 2,400여개의 모든 표정들의 분포가 나타나도록 계층적인 가시화를 구축한다. 또한 계층적 가시화 기법에 의하여 사용자는 얼굴 표정 제어를 수행한다.

3. FCM 클러스터링

클러스터링은 n 개의 데이터가 주어졌을 때, 데이터들의 유사성을 기반으로 하여 그들을 그룹(Cluster)으로 지정하는 것이다. 클러스터링 알고리즘에는 계층적(Hierarchical) 클러스터링[15, 16]과 K-Means[14], FCM 등이 있으며, 본 논문에서는 논문의 목적과 가장 잘 부합되는 FCM을 사용한다. 계층적 클러스터링은 상부 트리로 계층이 상승될수록 연결(Link)에 의해 가지(Branch)의 수는 점차 줄어들게 되며, 최초 잎(leaves)으로 표현된 점들의 정보를 상실하게 되어 본 논문의 목적을 위해서는 적합하지 않는다. 그리고 K-Means 알고리즘은 계산량이 너무 많아서 큰 데이터 집합을 위해서는 유용하지 않고, 전역 최소화를 찾기도 부적합하다. 그러므로 K-Means는 본 논문에서 사용하고자 하는 모션 데이터를 Clustering 하기에는 적합하지 않다.

본 논문에서 사용한 클러스터링 알고리즘 FCM[8-10]은 클러스터 지정에 있어서 불확실성을 가지고 있으며, 각 데이터에 주어진 각 클러스터에 속할 확률을 부여하여 데이터를 클러스터링하는 방법이다. 그리고 하나의 데이터가 두 개의 클러스터에 속하게 하는 것을 허용하는 클러스터링 방법 중의 하나이다. FCM은 Dunn[10]에 의해서 개발되었으며, Bezdek[8]에 의해서 개선되어져 패턴 인식 분야에 자주 사용되어지고 있다.

클러스터하기 위한 얼굴 표정들의 집합을 $X = \{x_i\}$, $i = 1$

..., n 라고 하자. X 의 퍼지 부분집합인 클러스터는 'Membership 함수'라고 부르는 매핑 함수 $U: X \rightarrow [0, 1]$ 로 표현된다. 여기서 $U_j(x_i)$ 는 얼굴 표정 x_i 의 클러스터 j 에 대한 Membership 계수를 나타낸다. 즉, $U_j(x_i) = 1$ 은 표정 x_i 가 클러스터 j 의 센터를 기준으로 완전한 Membership을 가지고 있다는 것이다. $U_j(x_i) = 0$ 는 표정 x_i 가 클러스터 j 의 센터를 기준으로 Membership이 없다는 것이다. 그리고 $U_j(x_i) = 0.8$ 은 표정 x_i 가 클러스터 j 의 센터를 기준으로 불완전하거나 부분적이지만 높은 Membership이 있다는 것을 가리킨다. 비록 하나의 얼굴 표정이라도 하나의 클러스터에 유일하게 속하지 않고, 다른 클러스터와의 임의의 위치 혹은 중앙 부근에 놓여질 수도 있다. 즉, 모든 표정들은 지정된 수만큼의 각 클러스터 센터로부터 측정된 서로 다른 Membership 계수를 가진다.

FCM은 최적화 문제로 귀결된다. FCM의 목적은 목적 함수를 최소화하는 Membership 함수의 집합 U 를 결정하는 것이다. U 는 다음 식 (1)을 만족하는 $U = \{U_1, \dots, U_c\}$ 으로 표현된다.

$$\sum_{j=1}^c U_j(x_i) = 1 \quad (1)$$

Membership 함수 U_j 는 $U_j(x_i) = u_{ij}$ 로 표시하고, 여기서 u_{ij} 는 클러스터 j 에서의 표정 x_i 의 Membership 계수를 나타낸다. 그리고 c 는 $2 \leq c \leq n$ 의 범위를 가지는 클러스터 센터의 수를 의미한다. 식 (1)은 하나의 표정 x_i 에 대한 모든 클러스터의 Membership 계수의 합이 1이라는 것을 말한다. 이것은 퍼지 클러스터 안에 있는 각 표정의 Membership이 모든 클러스터에 적절히 분리되어 할당되어진다는 것을 말한다. 예를 들어, $U_1(x_i) = 0.8$ 은 클러스터 1에 표정 x_i 의 Membership 계수 0.8이 할당되어져 있고, 계수 0.2는 다른 클러스터에 할당되어져 있다는 것이다.

퍼지 분할을 위한 FCM의 최적화 함수는 표정 x_i 로부터 클러스터 센터 c_j 까지 거리를 구하고, 이를 퍼지 Membership 함수와의 곱으로 정의하며, 다음 식 (2)와 같이 표현된다.

$$J_m(U) = \sum_{i=1}^n \sum_{j=1}^c [U_j(x_i)]^m \|x_i - c_j\|^2 \quad (2)$$

여기서 m 은 $1 < m < \infty$ 의 범위를 가지는 임의의 실수로써 일반적으로 1.25나 2의 값을 가지는데, 본 논문에서는 $m=2$ 를 사용하였다. 그리고 c_j 는 클러스터 센터로서 다음 식 (3)과 같이 정의된다.

$$c_j = \frac{\sum_{i=1}^n (u_{ij})^m x_i}{\sum_{i=1}^n (u_{ij})^m}, \quad j = 1, \dots, c \quad (3)$$

이와 같은 최적화 문제는 퍼지 Membership 행렬 U 의

초기값을 필요로 한다. 행렬 U 의 초기값은 $u_{ij} \in [0, 1]$ 및 $\sum_{j=1}^c u_{ij} = 1$ 와 같은 조건을 만족하는 $(n \times c)$ 크기이며, 임의로 정한다. 행렬 U 의 초기값을 입력으로 사용한 최적화는 최적화 함수가 최소가 될 때까지 클러스터 센터와 표정들 사이의 거리 $d_{ij} = \|x_i - c_j\|^2$ 에 따라서 퍼지 Membership 행렬 U 를 반복적으로 갱신한다. 만약 $d_{ij} > 0$ 이면, u_{ij} 는 다음 식 (4)과 같고,

$$u_{ij} = \left(\sum_{k=1}^c \left(\frac{d_{ij}}{d_{ik}} \right)^{\frac{2}{m-1}} \right)^{-1} \quad (4)$$

만약 $d_{ij} = 0$ 이면, 얼굴 표정 x_i 는 클러스터 센터 c_j 와 일치하고, 완전한 Membership이 $u_{ij} = 1$ 로 설정되어질 수 있다. 행렬 U 의 변화가 주어진 허용오차 ϵ 보다 작을 때까지 반복한다. 여기서 ϵ 은 0과 1 사이의 값을 가지는 알고리즘의 반복 종료 판정기준으로서, 일반적으로 0.01에서 0.001 사이의 값을 가지는데, 본 논문에서는 정밀도를 높이기 위하여 0.00001을 사용한다.

본 논문에서는 Membership 함수의 집합 U 와 클러스터 센터 c_j 의 집합으로 구성된 행렬 $C = \{c_j\}$ 를 구하기 위해 Matlab V6.5를 사용하였다. Matlab의 fcm 함수에 2차원 얼굴 표정 공간과 각 계층별 임의로 설정한 클러스터 센터의 수 c 를 입력으로 사용하고, 이로부터 행렬 U 와 행렬 C 를 구하였다.

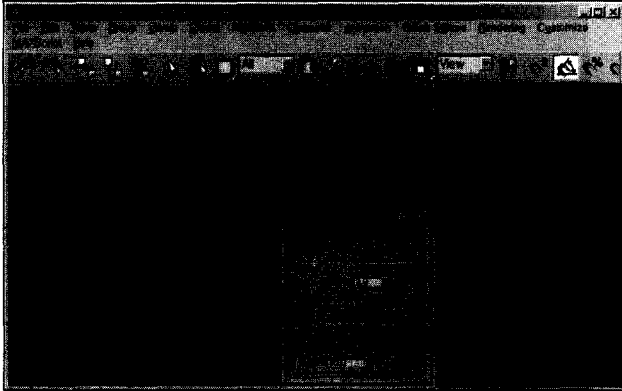
4. 계층적 가시화에 의한 표정 제어 실험

본 논문에서 사용하고자 하는 얼굴표정 공간은 2차원 공간이며, 사용자는 이 공간을 향해하여 다양한 표정 제어를 수행한다. 그러나 2차원 표정 공간에 분포된 얼굴 표정들의 수가 너무 많아 사용자가 향해하기에는 너무 복잡하고 난해하다. 그러므로 본 논문에서는 2,400여개의 얼굴 표정들을 계층적으로 가시화한다. 그리고 줌 인 줌 아웃을 통해 한 단계를 선택하여 향해하고 향해경로를 갱신하도록 한다.

4.1 사용자 인터페이스

계층적 가시화에 의한 표정 제어를 위한 사용자 인터페이스는 (그림 2)와 같으며, 3D Studio MAX R5.1의 MAXScript로 구현하였다. 계층적 가시화 초기단계에서는 수 개의 대표적인 얼굴 표정만을 사용자 인터페이스의 향해공간에 표시하고, 사용자는 이들 중 특정 표정들을 선택하여 초기 향해경로를 생성한다. 계층적 가시화를 위해서는 매 단계마다 클러스터의 수를 두 배로 한다. 그러면 계층적 가시화를 위한 최대 단계는 총 표정의 수에 따라 자동적으로 결정된다. 또한 매 단계마다 클러스터의 수만큼 클러스터링 한다. 그리고

사용자는 항해경로의 근처에 있는 키 프레임(클러스터 센터)들을 선택하여 항해경로를 갱신한다.



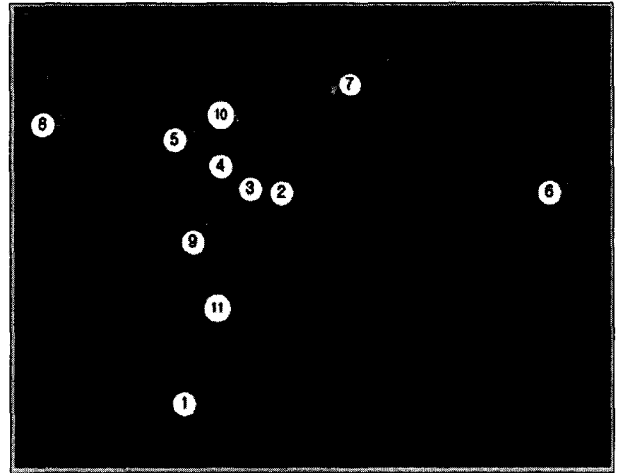
(그림 2) 사용자 인터페이스(좌: 항해공간과 스크립트 메뉴, 우: 3차원 얼굴 모델), 사용자가 2차원 항해공간에 분포된 각 얼굴 표정상태를 대표하는 작은 점을 마우스로 선택하면서 항해를 하면, 선택된 점에 해당되는 표정이 3차원 얼굴 모델에 보여짐.

4.2 표정들 사이의 연결 관계

본 논문에서는 실험을 위해 약 500개의 모션 데이터를 사용하려고 하였으나, 실험결과 가장 최신의 성능을 가진 PC 및 슈퍼컴퓨터를 사용하더라도 지금의 기술로는 불가능하다는 것을 확인하였다. 그러므로 500개의 모션을 모두 실험에 사용하기 위해서는 슈퍼컴퓨터를 기반으로 병렬처리 기법을 향후 연구하여야 할 것이다.

본 논문에서는 PC에서 계산이 가능한 적당한 모션의 수 10개를 선정하였다. 계층적 가시화의 초기 단계는 10개의 모션으로 구성된 2,400여개의 표정을 FCM을 사용하여 무표정 상태를 포함한 11개의 클러스터 센터를 가지도록 클러스터링한다. 그리고 클러스터 센터들을 후보 키 프레임들로 설정하여 그림 2와 같이 항해공간에 디스플레이 한다. 클러스터 센터의 표정은 항해공간에서 가장 가까이 있는 표정과 같다고 간주하고 그것을 사용자에게 보여준다. 사용자는 후보 키 프레임들 중에서 얼굴 표정 제어에 사용할 키들을 선택하여 초기 항해경로를 생성한다. 이때 사용자는 표정의 계층적 분포도에서 표정들 사이의 이웃하는 연결 관계를 미리 알아야 한다. 왜냐하면 임의의 표정에서 특정 표정으로의 직접적인 항해경로가 존재하지 않는 경우가 있기 때문이다. 또한 이는 실험에 사용한 모션 데이터의 수와 종류에 따라 달라진다. 그러므로 표정의 계층적 분포도는 단순히 표정만을 분포시키는 것이 아니라 (그림 3)과 같이 표정들 사이의 연결 관계까지 점선으로 표시를 하여 사용자가 한 눈에 파악할 수 있도록 알려준다. 표정들 사이의 연결 관계는 2,400여개의 모든 표정들을 분포시킨 상세 분포도에 포함되어있는 정보인데, (그림 3)과 같은 초기 단계인 요약 분포도에도 표시를

해 주어야 한다. 표정의 상세 연결 관계로부터 클러스터 센터들의 연결 관계를 찾아내는 방법은 다음과 같다. 먼저 임의의 두 클러스터 센터간의 최단경로를 찾는다. 여기서 최단경로는 인접한 표정들을 거쳐서 가거나 거치지 않고 바로 가는 최단거리를 의미한다. 그러나 최단경로가 다른 클러스터 센터들이나 센터들 주변을 지나가면, 이것은 두 클러스터 센터간의 직접경로가 아니라고 본다. 즉, 표정들 사이의 연결 관계는 임의의 두 클러스터 센터 간에 직접 경로만이 존재하여야 한다. 임의의 두 클러스터 센터간의 최단경로가 다른 클러스터 센터들이나 센터들 주변을 지나가는지는 다른 클러스터 센터를 중심으로 가상의 원을 그리고 그 원의 영역을 최단경로가 통과하는지를 확인하면 된다. 그리고 클러스터 센터를 중심으로 한 가상의 원의 반지름은 적당한 임계값을 설정하여 실험으로 찾는다.

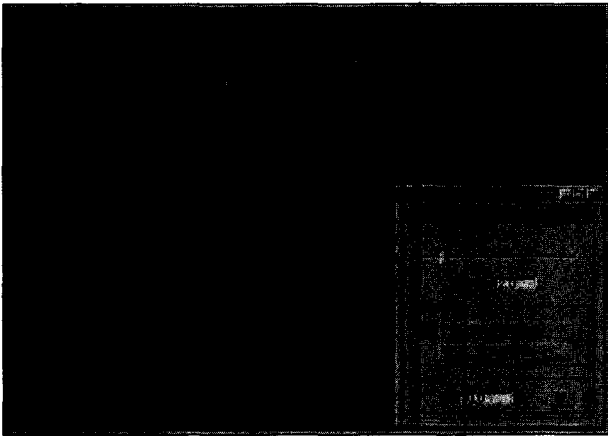


(그림 3) 계층적 가시화 초기 단계 - 약 11개 후보 키 프레임들의 분포와 표정들 간의 연결 관계. 표정은 설명의 편리를 위해 항해공간에 표시함.

본 논문에서는 이와 같은 방법으로 표정들 간의 연결 관계를 구하고 (그림 3)과 같이 점선으로 표시하였다. (그림 3)에서 연결 관계가 존재하지 않는 후보 키 프레임들 사이에서는 항해를 할 수가 없고, 항해를 하기 위해서는 점선으로 연결된 인접한 후보 키 프레임들을 거쳐서 항해해야만 한다. 예를 들어 키 프레임 1에서는 키 프레임 9, 11로 항해를 할 수 있지만, 키 프레임 1에서 키 프레임 8이나 5로 항해를 하고자 할 경우에는 인접한 키 프레임 11 혹은 키 프레임 9를 거쳐야 가능하다. 이는 표정이 급격하게 변하는 경우가 없고 인접한 표정을 거쳐서 항해를 해야 한다는 가정 하에서 구해진 표정들 간의 연결 관계이다. 또한 (그림 3)은 사용자가 2차원 표정공간에 있는 약 11개의 모든 후보 키 프레임들을 항해하면서 나타나는 얼굴 표정을 표시한 것으로서, 원래 얼굴모델은 그림 2에서와 같이 사용자 인터페이스의 오른쪽 창에 표시되지만, 설명의 편리를 위해 항해공간에 표시하였다.

4.3 계층적 가시화에 의한 표정 제어 실험

(그림 3)을 이용하여 얼굴 표정 제어를 수행하기에는 부적합한데, 이는 사용자가 선택한 키 프레임과 인접한 키 프레임을 보간 시켜주는 항해경로가 전혀 없기 때문이다. 그러므로 이들 사이에 적당한 개수의 새로운 후보 키 프레임들을 분포시켜서 세부적인 항해경로를 생성할 수 있도록 한다. 이를 위해서는 계층적 가시화가 필요하며, 매 단계마다 클러스터의 수를 두 배로 하면서 클러스터링한다. 그러면 모든 표정이 가시화되기까지의 최대 단계는 자동적으로 결정된다. 그리고 사용자로 하여금 줌 인 줌 아웃으로 적당한 단계를 선택하게 하고, 항해경로 근처에 있는 키 프레임을 선택하여 항해경로를 갱신한다.

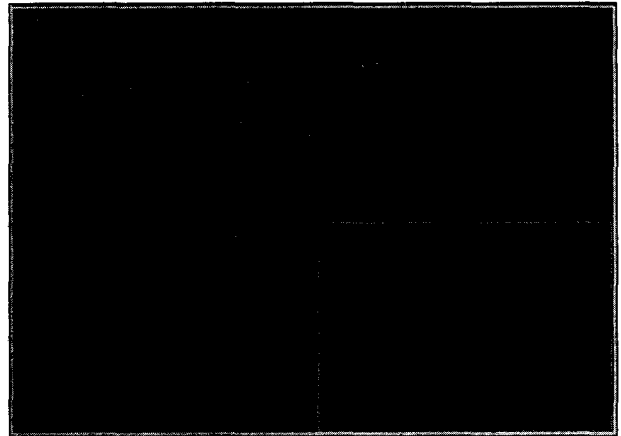


(그림 4) 계층적 가시화 초기 단계 - 약 11개 후보 키 프레임들 중에서 4개(6-2-9-1) 키 프레임을 선택한 초기 항해경로

예를 들어, 사용자가 계층적 가시화 초기 단계인 약 11개의 후보 키 프레임들로 분포된 항해공간에서 키 프레임으로 6, 2, 9, 1을 (그림 4)와 같이 스크립트 메뉴의 'Start Stroking'을 선택하여 초기 항해경로를 생성하였다고 하자. 계층적 가시화를 위해서 매 단계마다 클러스터의 수를 두 배로 하고 클러스터링하는데, 2,400여개의 표정을 위한 줌 인의 최대 단계는 최소한 8단계($11 \times 2^8 = 2,816$)가 필요하다. 본 논문에서는 이와 같이 수 개의 단계를 두어 사용자로 하여금 줌 인 줌 아웃으로 적당한 표정공간을 선택하여 항해하고 항해경로를 갱신할 수 있도록 하였다. (그림 5)의 작은 창은 (그림 4)와 같이 초기 단계에서 생성한 항해경로를 그대로 보여주고, 사용자가 줌 인 줌 아웃을 사용하여 선택한 단계 3의 후보 키 프레임들을 보여주고 있다.

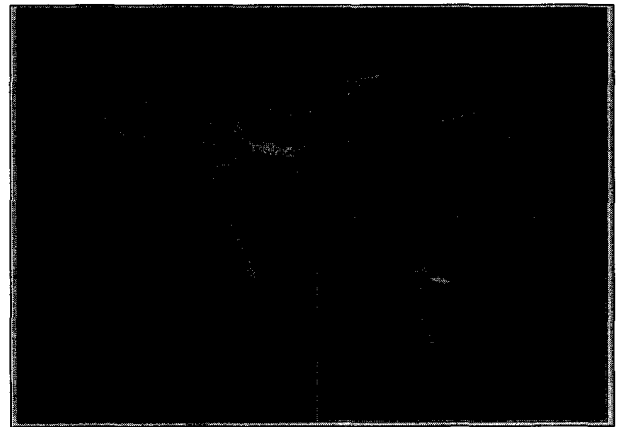
또한 (그림 5)는 계층적 가시화 초기 단계에서 생성한 항해경로와 인접한 특정 키 프레임들을 선택하여 갱신한 항해경로를 보여준다. 사용자는 계층적 항해공간을 줌 인 줌 아웃을 사용하여 항해하는데, 각 단계마다 항해경로를 갱신하는 것이 아니라 각 단계별로 표정들의 분포만 나타날 뿐, 사용자가 마음에 드는 단계를 선택하여 항해경로를 갱신한다. 그리고 현재의 항

해경로가 마음에 들지 않거나 그 근처에 있는 점들이 세부적인 항해경로의 점으로서 적합하지 않으면, 다시 줌 아웃 하여 새로운 경로를 정한 다음 다시 줌 인 하여 갱신하도록 한다.



(그림 5) 계층적 가시화 단계 3 - 초기 단계에서 선택한 키 프레임을 기반으로 줌 인으로 선택한 단계에서 새로운 후보 키 프레임들을 표시(작은 창)하고, 항해경로와 가장 인접한 특정 키 프레임을 선택하여 갱신한 항해경로

계층적 가시화의 마지막 단계는 사용자가 줌 인을 최대 하였을 때이고, 2,400여개의 표정들이 모두 분포되어지는 단계이다. 사용자는 마지막 단계에서 기존의 항해경로 근처에 표시된 후보 키 프레임들 중에서 기존의 항해경로와 가장 인접한 특정 키 프레임을 선택하고, (그림 6)과 같이 수정 및 갱신한다.



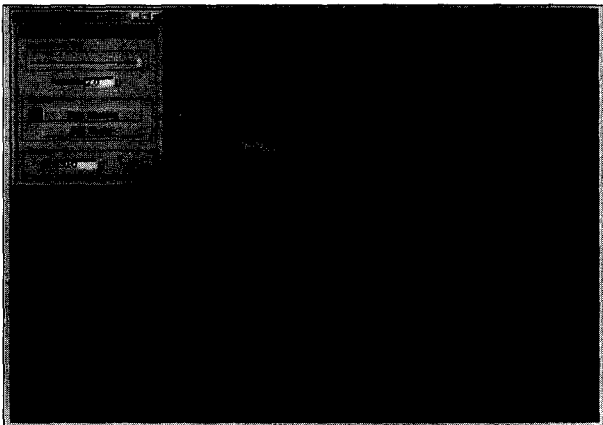
(그림 6) 계층적 가시화 마지막 단계 - 줌 인으로 마지막 단계의 새로운 후보 키 프레임들을 표시(작은 창)하고, 기존의 항해경로와 가장 인접한 특정 키 프레임을 선택하여 갱신된 항해경로

사용자의 표정 제어를 위한 항해 과정이 끝나면 사용자의 항해경로에 해당되는 얼굴 표정들을 연속적으로 보고 얼굴 표정의 변화를 확인할 필요가 있다. 이때에는 사용자의 항해 경

로를 처음부터 끝까지 자동으로 반복 항해하여 연속된 얼굴 표정의 변화를 3차원 얼굴 모델이 바로바로 보여주게 된다.

(그림 7)은 2,400여개의 얼굴 표정들이 분포된 2차원 표정 공간에서 3차원 아바타의 얼굴 표정 제어를 위해서 마우스를 사용하여 입의의 방향으로 항해경로를 생성할 때, 3차원 아바타의 얼굴 표정 변화를 즉시 보여주고 있다. 3차원 아바타의 얼굴 표정 제어는 (그림 7)의 스크립트 메뉴 'Start Stroking'을 사용하여 항해공간을 항해하는 사용자에게 의해서 수행되어진다. 즉, (그림 7)과 같이 하나의 표정에서 다른 표정으로 표정의 변화를 줄 때, 인접한 표정들을 항해하면 유사한 표정들이 나타나고 거리가 멀어질수록 특정 표정으로 변화가 진행된다.

일반적인 애니메이션이나 동영상은 렌더링 과정을 거친 각 프레임들을 이용하여 약 30fps의 재생 속도를 가진다. 본 논문에서는 그림 7의 스크립트 메뉴 'Replay Menu'을 사용하여 사용자가 사용 목적에 따라 재생 속도를 마음대로 설정하여 표정 제어를 반복적으로 확인할 수 있도록 하였다. (그림 7)은 재생 시 (그림 6)의 항해경로상의 적십자를 따라가면서 나타나는 렌더링 과정을 거친 대표적인 얼굴 표정들을 보여준다. 사용자는 이와 같이 계층적 가시화가 가능한 사용자 인터페이스를 이용하여 3차원 아바타의 표정을 제어할 수 있다.



(그림 7) 계층적 가시화 마지막 단계에서 최종 갱신된 항해경로상의 대표적인 얼굴 표정들을 설명의 편리를 위해 항해공간에 표시함(항해경로에 존재하는 적십자는 재생 시 항해 경로를 따라가는 현재의 경로 위치를 가리켜 주기 위한 지시자).

본 논문은 본 시스템이 어떤 효과가 있는지를 알기 위해 애니메이션 업계에서 재직 중인 약 30명의 사용자들로 하여금 약 30분 동안 20회 이상씩 본 시스템을 사용하여 3차원 아바타의 표정 제어를 수행하게 하였다. 그 결과 실험에 참여한 사용자들의 70%정도는 표정의 변화가 빠르고 급격한 표정들의 제어뿐만 아니라 표정의 변화가 정밀하면서 유사한 표정들을 거처서 다른 표정으로의 변화를 제어할 수 있는 효과가 있다고 판단하였다. 그러나 약 30%정도의 사용자들은

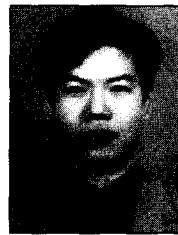
모든 표정들이 점으로만 표현되어져 있어서 계층을 변경할 때마다 대표적인 표정들의 위치를 기억하고 있어야 한다는 문제점이 있다고 하였으나, 수번의 사용 경험을 통하여 표정들의 분포를 알 수 있었다고 하였다. 이 문제는 향후 사용자 인터페이스를 보완하여 본 시스템을 처음 사용하는 사용자라도 표정들의 분포를 쉽게 파악할 수 있도록 해결해야 한다.

5. 결 론

본 논문에서는 적당한 공간에 분포시킨 다량의 얼굴 모션 데이터를 계층적 가시화 기법에 의하여 사용자가 이 공간을 항해하면서 원하는 얼굴 표정들을 선택하여 표정을 제어하는 방법을 기술하였다. 계층적 가시화를 위해 초기 단계인 약 11개 후보 키 프레임을 시작으로 마지막 단계인 2400여개의 후보 키 프레임으로 구분된 항해 공간을 구성하였으며, 이를 위해서 FCM을 사용하였다. 각 단계별 FCM에 의해서 생성한 클러스터 센터들은 얼굴 표정 제어를 위한 후보 키 프레임으로 설정하고, 이들 후보 키 프레임들 중 사용자로 하여금 선택하게 하여 키 프레임으로 설정하였다. 얼굴 표정 제어를 수행하기 위해서 개발한 사용자 인터페이스는 단계별로 설정된 계층적 가시화 및 사용자에게 의해 생성된 항해 결과를 확인하는데 유용하게 사용되었다. 또한 계층적 가시화에 의한 사용자 인터페이스의 활용에서는 계층적 가시화 초기 단계에서 약 11개 정도의 후보 키 프레임만을 표시하고 표정들 사이의 연결 관계를 참조하여 항해하게 함으로써 사용자가 표정의 분포를 보다 쉽게 파악할 수 있었다. 그리고 사용자가 좀 인 좀 아웃으로 선택한 단계에서는 초기 단계에서 생성한 항해경로를 가능한 유지하면서 기존의 항해 경로와 가장 인접한 점들을 지나도록 항해 및 갱신하였기 때문에 보다 자연스러운 3차원 아바타의 표정 제어가 가능하였다. 본 논문의 연구 결과는 기존의 관련 연구와 비교할 때, 모션 데이터의 계층적 가시화 기법을 사용한 3차원 아바타 얼굴 표정의 제어가 가능하다는 것이 가장 큰 특징이다. 특히 사용자가 선택한 단계가 낮을수록 2차원 평면에 분포된 연속된 표정의 수가 작아 급격한 표정의 변화를 보여줄 수 있고, 높을수록 자연스럽게 연속된 표정의 변화를 보여줄 수 있다는 것이 장점이다. 예를 들어 문자 위주의 채팅, 메신저 및 아바타를 사용한 인터넷 응용 프로그램에서 상대방과 의사소통을 하는 도중에 자신의 감정을 3차원 아바타를 사용하여 실감나게 표현해줄 필요가 있는데, 이런 경우 본 논문의 연구 결과는 매우 적합하다. 최근 메신저를 사용하는 네티즌들은 이모티콘을 사용하여 자신의 감정을 표현하는데, 네티즌들로부터 매우 큰 호응을 얻고 있으며, 3차원 아바타를 사용하여 본 논문의 결과를 적용할 경우 더욱 더 효과적인 것으로 본다. 그러므로 본 논문의 연구결과는 3차원 아바타의 얼굴 표정 제어가 필요한 각종 응용 프로그램 분야에서 매우 유용하게 사용될 것이다.

참 고 문 헌

- [1] Demetri Terzopoulos, Barbara Mones-Hattal, Beth Hofer, Frederic Parke, Doug Sweetland, Keith Waters, "Facial animation : Past, present and future," Panel, SIGGRAPH97.
- [2] Frederic I. Parke, Keith Waters, "Computer facial animation," A K Peters, 1996.
- [3] Brian Guenter, Cindy Grimm, Daniel Wood, Henrique Malvar, and Frederic Pighin. "Making Faces," In SIGGRAPH 98 Conference Proceedings. ACM SIGGRAPH, July, 1998.
- [4] Won-Sook Lee, Prem Kalra, Nadia Magnenat Thalmann, "Model based face reconstruction for animation," Proc. MMM'97, Singapore, pp.323-338, 1997.
- [5] Won-Sook Lee, Nadia Magnenat-Thalmann, "From real faces to virtual faces : Problems and solutions," Proc. 3IA'98, Limoges, pp.5-19, 1998.
- [6] Cyriaque Kouadio, Pierre Poulin, and Pierre Lachapelle, "Real-time facial animation based upon a bank of 3D facial expressions," Proc. Computer Animation 98, June, 1998.
- [7] Wonseok Chae, Yejin Kim, Sung Yong Shin, "An Example-based Approach to Text-driven Speech Animation with Emotional Expressions," EUROGRAPHICS 2003, Vol. 22, No.3.
- [8] J. C Bezdek, "Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York, 1981.
- [9] R. O. Duda and P. E. Hart, "Pattern Classification and Scene Analysis," Wiley, New York, 1973.
- [10] J. C. Dunn, "A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters," Journal of Cybernetics, 3, pp.32-57, 1973.
- [11] R. W. Floyd, "Algorithm 97 : Shortest Path," CACM Vol. 5, p.345, 1962.
- [12] T. Cox and M. Cox, "Multidimensional Scaling," Chapman & Hall, London, 1994.
- [13] 김성호, 정문렬, "모션 데이터를 이용한 3차원 아바타 얼굴 표정 제어", 정보처리학회논문지A, 제11-A권 제6호, 2004.
- [14] Seber, G. A. F., "Multivariate Observations," Wiley, New York, 1984.
- [15] Ying Zhao and George Karypis, "Evaluation of hierarchical clustering algorithms for document datasets," pp.515-524. ACM Press, 2002.
- [16] S. C. Johnson, "Hierarchical Clustering Schemes," Psychometrika, 2, pp.241-254, 1967.



김 성 호

e-mail : kimsh1204@hotmail.com

1996년 상지대학교 이공과대학 전산학과 (학사)

1998년 숭실대학교 일반대학원 컴퓨터학과 (공학석사)

2001년 숭실대학교 일반대학원 컴퓨터학과(박사수료)

1997년~1999년 숭실대학교 전자계산원 시간강사

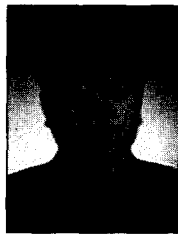
1999년~2000년 숭실대학교 컴퓨터학부 시간강사

1999년~2002년 숭의여자대학 인터넷정보과 시간강사

숭의여자대학 컴퓨터게임과 시간강사

2002년~현재 숭의여자대학 정보통신계열 멀티미디어콘텐츠 전공 겸임교수

관심분야 : 컴퓨터 그래픽스, 모션 캡처 애니메이션, 가상현실, Web3D, 멀티미디어 등



정 문 렬

e-mail : moon@mail.sogang.ac.kr

1980년 서울대학교 계산통계학과(학사)

1982년 KAIST 전산학과(공학석사)

1992년 펜실베니아 대학교 전산학과

(공학박사)

1992년~1994년 일본 구주공과대학교

정보공학부 조교수

1994년~1999년 숭실대학교 컴퓨터학부 부교수

2000년~현재 서강대학교 영상대학원 미디어공학과 부교수

관심분야 : 컴퓨터 그래픽스, 디지털 방송 등