

모션캡처 애니메이션을 위한 거리 측정방법

이 희 만[†] · 서 정 만^{††} · 정 순 기^{†††}

요 약

본 논문에서는 모션캡처를 위해 컬러 스테레오 카메라를 이용한 거리 측정 알고리즘을 제안하였다. 현실공간의 연기자 각 관절 부위에 컬러 마커를 부착시키고 이를 컬러 스테레오 카메라를 이용해 촬영한 후, 본 논문에서 제안한 컬러 매칭방법과 컬러 영역중심의 분석방법에 의해 카메라로부터 마커까지의 거리를 계산한다. 스테레오 영상에서 마커의 컬러영역을 추출하기 위해 국부지역내의 각 픽셀의 RGB(red, green, blue) 컬러 정보를 CIE(Commission Internationale de l'Eclairage) 컬러 공간으로 변환시켜 컬러의 파장을 계산하고, 국부지역의 우월 파장이 마커의 컬러 파장과 일치하는지를 검색한다. 가상공간에서의 캐릭터의 움직임은 시간에 따른 마커의 위치 변화정보를 처리하는 프로그램에 의하여 제어된다.

Distance Measuring Method for Motion Capture Animation

Heeman Lee[†] · Jeong-Man Seo^{††} · Soon-Key Jung^{†††}

ABSTRACT

In this paper, a distance measuring algorithm for motion capture using color stereo camera is proposed. The color markers attached on articulations of an actor are captured by stereo color video cameras, and color region which has the same color of the marker's color in the captured images is separated from the other colors by finding dominant wavelength of colors. Color data in RGB (red, green, blue) color space is converted into CIE (Commission Internationale de l'Eclairage) color space for the purpose of calculating wavelength. The dominant wavelength is selected from histogram of the neighbor wavelengths. The motion of the character in the cyber space is controlled by a program using the distance information of the moving markers.

키워드 : 모션캡처(Motion Capture), 거리측정(Distance Measuring)

1. 서 론

가상현실(virtual reality)이란 실제로 존재하지 않는 인위적인 세계를 인간의 오감에 전달하여 가상의 환경에서 상호작용을 가능케 함으로써 마치 실제로 존재하는 환경에서와 같은 현실감을 갖도록 하는 기술이라 정의할 수 있다. 가상현실을 위한 3차원 캐릭터의 애니메이션 제작에 소요되는 핵심기술로는 모션캡처 기술을 들 수 있다. 모션캡처 기술을 이용하여 가상공간 캐릭터의 각 관절부를 움직이는데 필요한 정보를 동시에 수집할 수 있다. 현실공간에 있는 연기자의 신체 여러 부분에 부착된 센서를 통하여 연기자의 동작을 파악하고, 이를 이용해 가상공간의 캐릭터를 동일하게 움직이게 하는 실시간 캐릭터 애니메이션의 제작분

야 뿐만 아니라 재활의학, 산업디자인, 스포츠 등 다양한 분야에서 모션캡처 기술이 활용되고 있다[1].

현재 모션캡처 기술의 문제점으로 지적되고 있는 것 중의 하나는, 움직이는 연기자의 신체에 부착된 트랙커의 움직임을 정밀하게 측정할 수 없기 때문에 캐릭터의 자연스러운 동작에 필요한 데이터의 수집에 한계가 있다는 점이다. 그러므로 연기자 신체의 각 부분이 움직이는 위치와 캐릭터가 움직이는 위치가 조금씩 다른 경우 캐릭터의 동작이 부자연스럽게 연출되는 경우가 발생한다. 현재 널리 이용되고 있는 모션캡처 시스템의 종류는 <표 1>과 같다.

본 연구에서는 모션캡처의 핵심 기술로 평가되고 있는 거리 측정방법에 관한 것이다. 모션캡처에 이용되고 있는 기술은 기계적, 음향학적, 관성학적, 광학적 기술 및 마그네틱 기술로 분류할 수 있으며, 일반적으로 널리 사용되고 있는 광학적 기술과 마그네틱 기술을 비교하면 <표 2>와 같다.

† 정 회 원 : 서원대학교 컴퓨터정보통신공학부 교수
 †† 정 회 원 : 한국재활복지대학 컴퓨터게임개발과 교수
 ††† 정 회 원 : 충북대학교 전기전자컴퓨터공학부 교수
 논문접수 : 2001년 3월 5일, 심사완료 : 2002년 1월 3일

〈표 1〉 모션캡처 제품 비교

제품명	특징	장점	단점
Gypsy	데이터 추출을 위해 신체의 주요 관절에 포텐셜 미터를 부착, 사용. 포텐셜미터와 컴퓨터의 시리얼포트에 연결된 A/D 컨버터로 구성.	어느 환경에서도 동작이 가능, 특별한 셋업이 필요 없고, 연결에 시리얼 케이블을 사용.	부피가 크므로 자유스러운 움직임에 다소 불편이 따름.
Cyber Glove/Suit	무선 전신 모션캡처용 복장으로 신체의 움직임을 포착하기 위해 위치를 측정.	무선 전신 모션캡처용 복장으로 전신을 연결에 관계없이 자유롭게 움직일 수 있다.	무선 복장의 값이 고가
Expert Vision Hires	연기자의 정교한 몸 동작 모션 데이터를 처리	마커가 가려질 경우 소프트웨어에 의한 자동 보상 가능	데이터 처리에 5-8분이 소요됨.

〈표 2〉 모션캡처 기술 비교

기술 비교	내용
마그네틱 모션캡처	자기장을 측정할 수 있는 센서를 연기자의 각 관절 부위에 부착시킨다. 연기자가 자기장 발생 공간내에서 움직일 때 각 센서에서 측정되는 자기장의 변화를 공간적인 변화량으로 변환시켜 연기자의 움직임을 측정한다.
광학적 모션캡처	광선, 카메라 및 반사점을 이용하여 3차원 공간에서 관절의 움직임을 분석하며, 카메라가 읽어들이는 데이터를 프로그램에 의해 처리, 3차원 데이터를 추출한다.

마그네틱 모션캡처 방법에서는 자기장을 감지할 수 있는 수신기를 사용하며, 연기자의 미세한 움직임의 변화에 따라 변화되는 자기장의 강도(세기)를 유무선을 통해 포착한다. 마그네틱 기술은 데이터를 수집하는 동안 잡음이 발생되지 않는 특수한 작업공간을 필요로 한다. 작업공간 주변의 각종 구조물들(실내 벽, 철근 구조물 등)은 자기장을 발생시키거나, 왜곡시킬 수 있는 잡음 발생의 원인을 제공한다. 마그네틱 기술을 사용하는 몇몇 유선 모션캡처 시스템들은 센서와 컴퓨터를 연결하는 케이블의 길이가 고정되어 있기 때문에, 연기자의 움직임이 제한된 공간을 벗어나지 않도록 하기 위해 작업공간 상부에 보조장치의 설치를 필요로 한다. 마그네틱 기술은 광학 기술이 도입되기 전에 가장 널리 이용되던 기술로 회전 동작과 같은 복잡한 동작을 포착할 수 없다는 단점이 있으나, 광학 기술에 비해 가격이 저렴하다는 장점이 있다. 광학 기술에서는 연기자의 관절 부위에 발광 재료로 제작된 마커를 부착시키고 마커에 특수 광선(예, 적외선 등)을 발사한다. 그리고 마커로부터 반사되는 광선을 여러 대의 카메라로 촬영하여 카메라의 모델과 기하학적 배치에 따라 발광되는 물체까지의 거리를 계산, 3차원 공간상의 물체의 위치를 파악한다. 신체의 각 부위에 부착된 센서에 케이블을 연결하는 대신에 발광체로 된 마커를 부착함으로써 연기자는 보다 자연스러운 동작을 수행할

수 있다. 또한 절대 좌표 값을 구할 수 있으므로 동작 데이터의 수집 시 에러가 적게되고, 에러의 보정이 용이하므로 얼굴 표정과 같은 세밀한 움직임의 측정이 가능하다. 그러나 광학 기술을 이용한 기존 모션캡처 시스템은 유지보수 비용이 많이 들며, 마커가 카메라의 사각지대에 위치하게 되는 경우에는 위치 계산이 어렵게 된다는 단점이 있다.

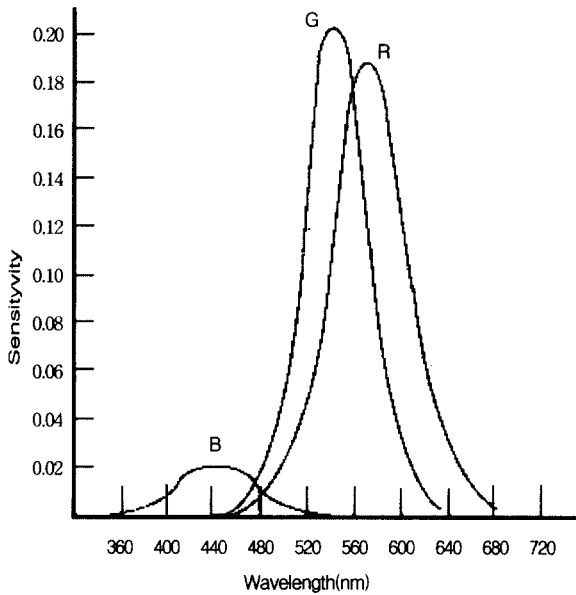
본 연구에서는 2대의 일반 컬러 카메라를 이용하여 연기자의 움직임을 이미지 그래픽으로 포착하고, 연기자의 각 관절 부위에 부착된 컬러 마커의 이미지 상의 위치를 검색한 후 카메라의 기하학적인 배치에 따라 거리를 계산한다. 그리고 각 마커의 위치변화를 계속 추적하여 연기자의 움직임을 추적한다. 기존 광학 시스템에서 사용하는 적외선 카메라 대신에 저가의 일반 컬러 카메라를 사용하여 컬러 마커의 위치를 추적, 연기자의 움직임을 감지할 수 있는 방법을 제안하였으며, 카메라의 측정단위 매김(calibration)을 단순화할 수 있는 기법을 개발하여 시스템의 유지보수 비용의 감소는 물론 사용자의 편리성을 향상시키고자 하였다.

제 2장에서는 본 연구의 기초가 되는 컬러 이론에 대해 논의하였고, 제 3장에서는 컬러의 파장 계산방법을 도입하였다. 제 4장에서는 본 연구에서 사용한 스테레오 컬러 카메라의 구조에 대해 설명하였다. 제 5장에서는 모션캡처 시스템의 구조에 대해 설명하였으며, 제 6장과 제 7장에서는 각각 실험 결과와 결론을 기술하였다.

2. 컬러 이론

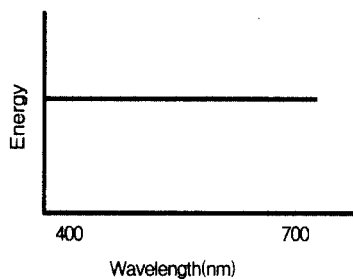
컬러는 여러 광선의 파장에 대한 인간 눈의 반응을 통해 인식된다. 즉, 컬러는 인간의 신경 시스템과 빛의 상호작용에 의해 인식되는 현상이다. 동일한 광선이라도 조건에 따라 사람마다 각자 다른 컬러로 인식할 수 있다. 우리는 눈의 망막에 있는 감광기관(photoreceptor)을 통해 빛을 감지한다. 감광기관은 간상체(rod)와 원추체(cone)로 구성된다[2]. 간상체는 야시경(night-vision)과 같이 빛의 세기에만 반응하며, 컬러에 대해서는 반응하지 않는다. 원추체는 간상체에 비하여 빛에 대한 감도(sensitivity)는 떨어지나, 빛의 파장에 따라 각각 다른 반응 일으켜 컬러를 인식하게 한다. (그림 1)은 3종류 광색소의 감도를 나타낸다[3]. 눈에 입사되는 단일 파장의 빛은 3종류의 원추체에서 각각 다른 반응을 일으킨다. 450nm 파장의 경우 청색 광색소와 녹색 광색소가 각각 반응을 일으키며, 540nm 파장의 경우 녹색 광색소와 적색 광색소가 반응을 일으킨다. 그러나 3종류의 광색소가 같은 비율로 존재하는 것이 아니므로 같은 세기의 광원이라도 각각 다른 밝기로 인지된다. 망막의 중심에는 대부분 녹색 광색소가 존재하며, 그 주변에는 적색 광색소가 존재하고, 그리고 가장자리에는 청색 광색소가 존재한

다. 눈의 망막에 있는 광색소에서 감지된 빛은 전기 신호로 변환되지만, 신호는 즉시 뇌신경으로 전달되지 않고 중간에 컬러 신호처리기(color-processing unit)인 뉴럴 네트워크를 통해 뇌신경으로 전달된다[4].



(그림 1) Cones의 파장에 따른 감도변화

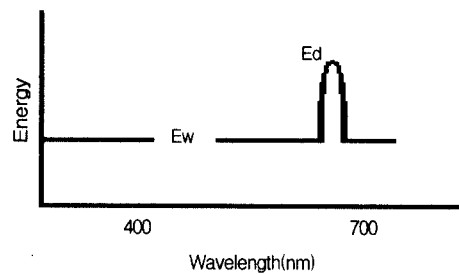
모든 파장의 광선이 균일하게 혼재되어 있는 경우 사람은 백색으로 인지하게 된다. 그러므로 태양 광선은 백색 광원으로 인지된다. 이상적인 백색 광원은 (그림 2)와 같이 인간이 감지할 수 있는 영역의 모든 파장에 해당하는 스펙트럼을 균일하게 가지고 있다.



(그림 2) 이상적인 백색광

우리의 눈은 혼재된 여러 광원 중에서 세기가 가장 우월한 빛의 파장을 그 빛의 실제 컬러로 인지한다. 백색 광원이 물체에 투사되면 물체에 따라서 특정한 파장의 빛을 흡수하거나 반사하게 되며, 이를 통해 물체의 컬러를 인식하게 된다. 만약 백색광원에 대하여 모든 파장의 빛을 반사한다면 백색으로 보이게 되며, 모두 흡수한다면 검은색으로 보이게 된다. 예를 들어 430nm 파장이 다른 파장에 비하여 상대적으로 반사율이 높다면 청색으로 보이게 된다. 이때

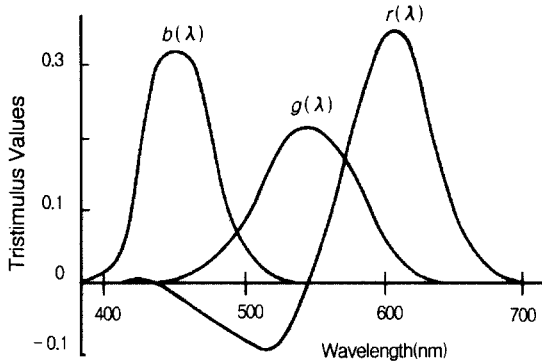
430nm 파장을 우월 파장(dominant wavelength)이라고 하며, 컬러 또는 휴(hue)라고 부른다. (그림 3)은 가시광선 영역의 빛의 파장에 대한 에너지 분포에 대한 예이다. Ed는 우월 파장의 에너지이며 Ew는 우월파장을 제외한 일반광원의 평균 에너지이다. 인간이 컬러 정보를 인지하는 데 필요한 3가지 종류의 지각변수로는 색상(hue), 채도(saturation) 및 명도(brightness)를 들 수 있다. (그림 3)에서 명도는 그래프 곡선의 하위 면적으로 나타내며, 컬러의 채도(또는 purity)는 $(E_d - E_w)$ 식으로 계산된다. 이는 백색으로 희석되지 않는 컬러의 정도를 의미한다.



(그림 3) 색상(Hue)의 스펙트럼

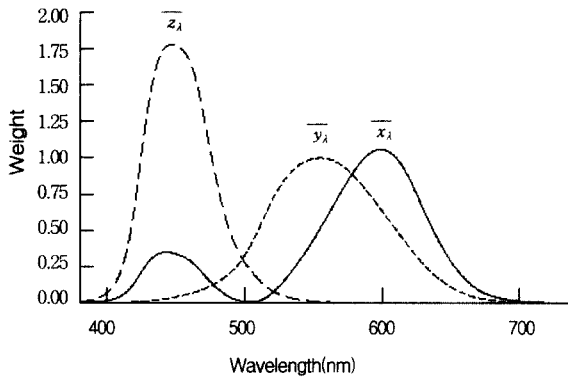
즉, 100% 순수한 색상은 $E_w = 0$ 인 경우이며, $E_d - E_w = 0$ 인 경우는 0% 채도를 의미한다. 그러므로 컬러의 표현 정보는 채도와 색상(hue, dominant frequency)에 의해 추출된다. 빛의 채도와 색상은 각각 다른 2개 이상의 광원에 의하여 다양한 색광을 표현할 수 있다. 2개의 광원을 혼합하여 백색 광으로 인식되는 경우 이를 보색 광이라고 한다. 적색과 청록색(cyan), 녹색과 자홍색(magenta), 청색과 노란색은 보색 관계가 된다. 동일한 컬러를 표현하는 방법은 매우 다양하여 많은 조합이 가능하다. 이를 메티머스(metemers)라고 부른다. 그러나 우리는 공학적으로 표준화된 3개의 광원 RGB(red, green, blue)을 사용하여 컬러의 영역(color gamut)을 표현하고 있으며, 이 3개의 광원을 기본 광원(primary colors)라고 부른다. CIE(Commission Internationale de l'Éclairage)에서는 인간이 인지하는 모든 파장에 대한 색상을 표시할 수 있는 컬러 좌표 국제 표준안을 실험적인 방법에 의하여 제정하였다. 즉, 단일 파장 컬러의 색상을 보여주고, 이 컬러와 동일한 컬러가 되도록 RGB 광의 세기를 조절하는 단자를 사용, RGB의 매칭 함수를 구하였다. (그림 4)는 RGB를 사용한 매칭 함수의 그래프이다[3].

(그림 4)에 의하면 RGB에 의해서 100% 채도된(saturated) 광원을 표시하기 위해서는 음(negative)의 값을 사용하여야 한다. 그러므로 음의 값을 사용하지 않는 한 기본 광원(RGB)만으로는 인간이 인지할 수 있는 가시 영역내의 모든 파장에 대한 색상을 표시할 수 없다. 그러므로 TV 화면이나 컬러 모니터에서 모든 색상을 표현할 수 없다.



(그림 4) RGB 컬러 매칭함수

음의 값을 피하기 위해서 CIE에서는 비실현적인 방법이긴 하지만, 실험을 통해 경험적인 3가지 기본 광원을 표준으로 정하였다. (그림 5)은 CIE에서 표준으로 정한 기본 광원에 대한 매칭 함수이다[3]. CIE에서는 가시 영역에서 380nm-780nm 파장의 빛에 대한 표본 집단 사람들의 색상 인지도를 측정하여 평균값을 구해 컬러 매칭 함수를 도출하였다.



(그림 5) CIE의 컬러 매칭함수

빛의 스펙트럼을 $P(\lambda)$ 라 정의하고, λ 를 가시 영역내의 빛의 파장(380nm-780nm)이라고 할 때, $P(\lambda)$ 에서 CIE의 X, Y, Z값을 계산하기 위해서는 다음의 수식 (1)을 이용한다. 수식에서 자체 발광하는 광원의 k 값은 680이다. CIE의 Y 값은 광원의 밝기를 측정할 값이 되며 X, Z는 스펙트럼의 컬러 정보를 나타낸다.

$$\begin{aligned}
 X &= k \sum_{\lambda=380}^{780} \bar{x}_\lambda P(\lambda) \\
 Y &= k \sum_{\lambda=380}^{780} \bar{y}_\lambda P(\lambda) \\
 Z &= k \sum_{\lambda=380}^{780} \bar{z}_\lambda P(\lambda)
 \end{aligned} \tag{1}$$

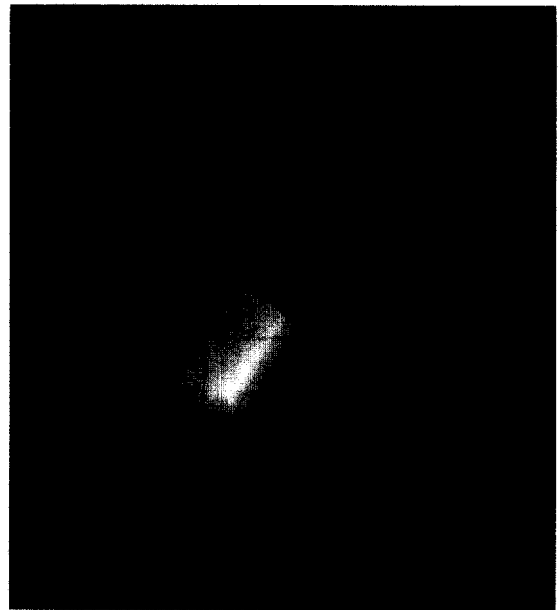
컬러 정보는 절대 값보다는 상대적인 값을 사용한다. 정규화된 컬러 정보는 다음의 식 (2)과 같이 계산된다.

$$\begin{aligned}
 x &= \frac{X}{(X+Y+Z)} \\
 y &= \frac{Y}{(X+Y+Z)} \\
 z &= \frac{Z}{(X+Y+Z)}
 \end{aligned} \tag{2}$$

식 $x+y+z=1$ 로부터 $z=1-(x+y)$ 를 구할 수 있으므로, 컬러는 일반적으로 x, y 값으로 만 표현한다. (그림 6)은 정규화된 x, y 좌표 공간상에 SMPTE의 RGB 기본광원을 표시한 것이다[5]. 혀모양의 가장자리 부분은 채도가 100%인 순수한 컬러를 의미한다. CIE 좌표내의 삼각형 영역은 RGB 기본 광원에 의해 표현 가능한 색상의 영역을 나타낸 것으로, 모든 색상을 표현할 수 없으며, 특히 채도가 100%인 컬러의 표현은 불가능하다. <표 3>은 NTSC, EBU, SMPTE에서 표준규격으로 사용하는 기본 광원(red, green, blue)과 백색 광에 대한 CIE 좌표(x, y)를 정리한 것이다[5].

<표 3> Primary Color 색상의 CIE좌표

system	Red		Green		Blue		White Point	
	x_r	y_r	x_g	y_g	x_b	y_b	x_{wh}	y_{wh}
NTSC	0.67	0.33	0.21	0.71	0.14	0.08	0.3101	0.3162
EBU(PAL/SECAM)	0.64	0.33	0.29	0.60	0.15	0.06	0.3127	0.3291
SMPTE	0.630	0.340	0.31	0.595	0.155	0.07	0.3127	0.3291



(그림 6) CIE 컬러공간

3. 컬러의 파장 계산

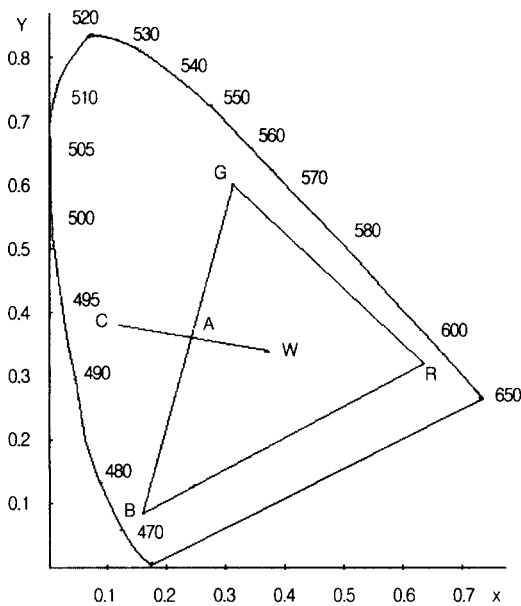
컬러는 여러 광선의 파장에 대한 망막의 반응을 통해 뇌

신경 시스템과 빛의 상호작용에 의해 인지된다. 이러한 현상을 공학적인 기법을 이용해 분석한다면 컬러는 표준화된 3개의 기본 광원(RGB)의 조합에 의해서 표현된다. 즉, 컴퓨터 모니터와 TV 브라운판에서는 기본 광원을 조합하여 컬러를 재현한다. 그러나 기본 광원을 사용하는 장치는 가시광선 영역내의 모든 컬러를 표현할 수 없다. 3개의 RGB광원을 사용하여 표현할 수 있는 컬러영역은 (그림 6)의 삼각형 내부이다.

RGB 좌표와 CIE좌표 값의 관계는 식 (3)과 같으며, R, G, B는 각각 0~1 사이의 값을 갖는다. 현재 사용하고 있는 장비의 규격에 따라 <표 3>을 이용하여 R, G, B 값으로부터 CIE의 x, y, z를 산출할 수 있다.

$$\begin{aligned} x_r R + x_g G + x_b B &= x \\ y_r R + y_g G + y_b B &= y \\ z_r R + z_g G + z_b B &= z \end{aligned} \quad (3)$$

RGB 좌표를 사용하는 컬러 CCD 카메라는 CIE의 모든 영역의 컬러를 표현하지 못하므로 가장 유사한 RGB 컬러 값으로 변환하여 표현한다. CIE의 x, y, z 값을 R, G, B 값으로 변환시킬 때 R, G, B 값이 음수가 되는 경우가 이에 해당한다. 이 경우는 (그림 7)의 삼각형 영역 밖의 좌표 값을 의미한다.



(그림 7) 색상의 보정

이 경우 CCD 카메라는 영역 밖의 좌표 C와 백색광의 좌표 W를 직선으로 연결할 때 교차되는 좌표 A의 색상으로 표현한다. 좌표 A의 색상은 채도는 감소되지만 C 좌표의 색상과 가장 근접한 색상으로 우월 파장은 동일하다. 하나의 컬러를 표현하기 위해서 여러 파장을 갖는 광원의 다양한 조합이 가능하다. 그러므로 단일 색상으로 보이는 컬러

광선을 분광기로 분석한다면 다양한 스펙트럼이 존재한다는 것을 파악할 수 있다. 그러나 컬러 비디오 카메라로 촬영한 이미지의 컬러는 이미 RGB 값으로 변환되었기 때문에 전체 스펙트럼을 파악할 수 없다. 다만 특정 컬러에 대한 우월 주파수를 CIE 좌표로부터 구할 수 있다. 그러나 인간이 인지할 수 있는 모든 영역의 컬러에 대해 우월 주파수가 존재하는 것이 아니므로, 모든 컬러에 대해서 우월 주파수를 구할 수는 없다.

컬러의 우월 주파수를 계산하는 과정은 다음과 같다.

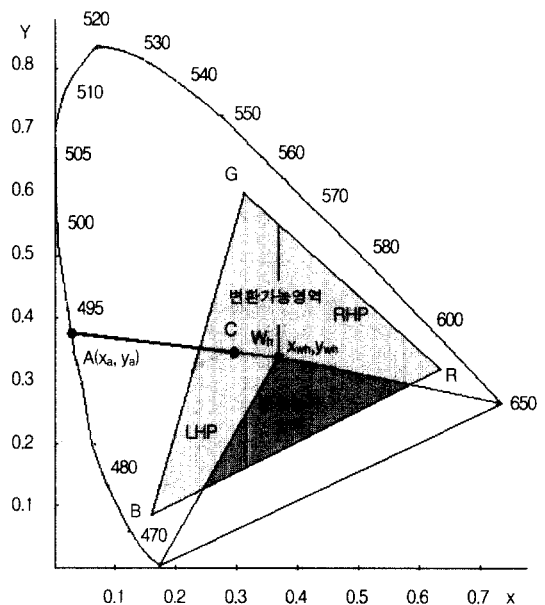
제 1단계: RGB의 색상 좌표를 수식 (3)을 이용하여 CIE의 x, y, z 값으로 변환시킨다. 24비트를 사용하는 완전한 컬러(full color)의 경우 각각의 R, G, B 값은 0~255이므로, 이를 정규화 하기 위하여 0~1의 값으로 변환한다.

제 2단계: 변환된 CIE의 좌표 x, y, z 값으로부터 우월 주파수가 있는 변환 가능 영역인가를 확인한다. (그림 8)는 변환 가능 영역과 불가능 영역을 나타내고 있으며, 변환 불가능 영역에는 우월 주파수가 존재하지 않는다.

제 3단계: 변환 가능 영역이 존재한다면, 채도가 100%인 가장 길이(wavelength)를 구한다. 이를 위해 (x_{wh}, y_{wh}) 좌표와 (x_c, y_c) 좌표를 연결하는 1차 함수를 구한다. 1차 함수는 다음 식 (4)과 같이 구할 수 있다.

$$y = ax + b \quad (4)$$

where $a = \frac{(y_{wh} - y_c)}{(x_{wh} - x_c)}$, $b = y_{wh} - ax_{wh}$



(그림 8) 컬러 파장

위에서 계산된 1차 함수 식에 채도가 100%인 컬러의 좌표 값을 대입한다. 채도가 100%인 컬러의 좌표 값을 매칭 함수로부터 구할 수 있다. 이때 y 값이 음수에서 양수로, 또는 양수에서 음수로 바뀌는 좌표를 검색하면 그 지점이 우월 주파수가 된다.

4. 스테레오 컬러 카메라

여러 대의 카메라로부터 촬영된 영상에서 서로 대응되는 이미지의 좌표를 확인할 수 있다면, 카메라의 기하학적 위치에 대응되는 영상과 카메라 위치간의 거리를 계산할 수 있다.

카메라 위치를 어떻게 배치하더라도 두 대 이상의 카메라로부터 상호 대응되는 좌표를 찾을 수 있다면 카메라의 기하학적 위치에 따라 3차원 데이터를 계산할 수 있다. 그러나 계산의 편리성을 위하여 스테레오 모델은 수렴(convergence) 카메라 모델과 평행 카메라 모델이 많이 사용된다. 수렴 카메라 모델은 인간의 눈과 같이 카메라의 두 광축이 하나의 고정 점에서 만난다. 수렴 카메라 모델에 대한 설명은 참고문헌 [6]을 참조한다. 대응되는 좌표의 위치검색이 용이하고, 또한 대응점으로부터 거리 계산이 용이하기 때문에 대부분의 스테레오 비전은 평행 카메라 모델을 사용한다(그림 9). 실제로 카메라는 핀홀(pinhole)이 아니므로, 보다 정확한 거리 계산을 위해서는 보정방법을 사용해야 한다. (그림 9)에서와 같이 2대의 카메라는 기본 선(base line) 위에 설치되며, 카메라를 기본 선과 수직으로 전방을 향하도록 한다. 왼쪽과 오른쪽 카메라를 연결하는 기본 선의 1/2 지점을 월드(world) 좌표의 원점(O_w)으로 하고, 두 카메라 광축사이의 거리를 b mm라고 하자. 카메라의 초점 길이는 f mm이며, 카메라의 영상평면은 기본 선으로부터 전방 f mm에 위치한다고 가정하자. 카메라의 영상 평면은

실제로 기본 선으로부터 f mm 후방에 위치하지만, 개념상의 이해를 돕기 위해서 전방향 위치로 모델링한다. 왼쪽과 오른쪽 카메라에 투영된 각각의 영상 좌표의 중심은 (O_l, O_r)을 기준으로 한다.

임의의 월드 좌표 $P(x_w, y_w, z_w)$ 에 있는 광선은 왼쪽 카메라 $P_l(u_l, v_l)$ 와 오른쪽 카메라 $P_r(u_r, v_r)$ 영상평면 좌표에 각각 투영된다. 투영되는 영상 평면의 좌표는 투영 변환(perspective transform) 식 (5)에 의하여 계산된다. 식 (5)에서와 같이 평행 카메라 모델의 수직축 변위(disparity) $D_v = v_l - v_r = 0$ 이 되므로, 임의의 월드 좌표는 투영 평면상의 동일한 수직 축에 존재한다. 즉, 대응되는 좌표가 동일한 수직 축에 존재하므로, 대응되는 위치의 검색이 다른 카메라 모델에 비하여 용이하다는 장점이 있다.

$$\begin{bmatrix} u_l \\ v_l \end{bmatrix} = \begin{bmatrix} \frac{f(x_w + b/2)}{z_w} \\ \frac{fy_w}{z_w} \end{bmatrix}, \quad \begin{bmatrix} u_r \\ v_r \end{bmatrix} = \begin{bmatrix} \frac{f(x_w - b/2)}{z_w} \\ \frac{fy_w}{z_w} \end{bmatrix} \quad (5)$$

수평축의 변위 (D_h : disparity)는 식 (6)과 같이 거리 z_w 에 반비례한다. 즉, 영상의 좌표로부터 D_h 를 구하면 측정대상의 거리 z_w 를 계산할 수 있다. 측정 대상이 무한 거리에 존재하는 경우 변위 $D_h = 0$ 이 된다.

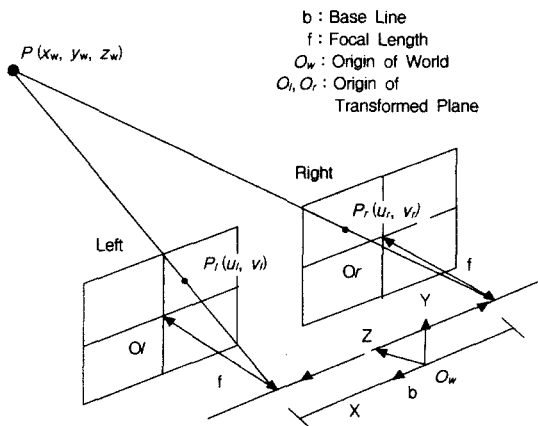
$$D_h = u_l - u_r = \frac{fb}{z_w} \quad (6)$$

그러나 영상이 도립 형태이고, 기본 선의 후방에 위치하는 경우 변위 $D_h = u_r - u_l$ 를 사용하여야 한다. 그러므로 두 카메라 영상의 대응 좌표로부터 월드 좌표 $P(x_w, y_w, z_w)$ 를 구하면, 최종적으로 식 (7)과 같이 된다. 본 연구에서는 컬러 비디오 카메라로부터 컬러 마커의 위치를 구하고, 상호 대응되는 좌표를 구하여 거리를 계산하는 방법을 사용한다.

$$\begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} = \begin{bmatrix} \frac{b}{2} \frac{u_l + u_r}{D_h} \\ \frac{b}{2} \frac{v_l + v_r}{D_h} \\ \frac{bf}{D_h} \end{bmatrix} \quad (7)$$

5. 모션캡처 시스템

본 연구에서 제안한 모션캡처 방법의 원리는 다음과 같다. 신체의 각 관절 부위에 컬러 마커를 부착하고 연기자를 움직이도록 한다. 연기자의 움직임을 스테레오 컬러 비디오 카메라로 촬영하고, 비디오 신호를 디지털 신호로 변환시킨다. 디지털화된 영상에서 컬러 마커를 검색한다. 영

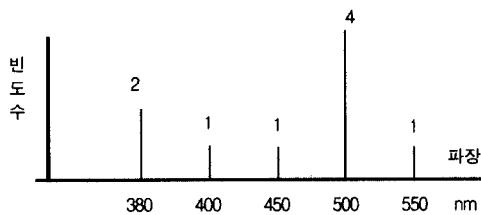


(그림 9) 평행 카메라 모델

상 정보는 RGB 신호로 표현되므로 특정 컬러를 검색한다는 것은 쉽지 않다. 그러므로 RGB 신호를 CIE 좌표로 변환시킨 후 컬러의 파장을 계산하여 원하는 컬러를 검출한다. 스테레오 카메라를 사용하므로, 좌측 화면에서의 컬러 마커 위치와 우측 화면에서의 동일한 컬러 마커 위치를 검색한 후 식 (7)을 이용하여 카메라로부터 컬러 마커의 현재 위치를 계산할 수 있다. 컬러 마커의 위치를 시간에 따라 계속 추적한다면 연기자의 움직임을 파악할 수 있다. 컬러 마커의 위치 추적 데이터를 이용하여 3차원 애니메이션을 구현한다.

5.1 컬러 마커의 위치검색

스테레오 컬러 카메라로부터 포착한 좌/우 영상으로부터 컬러 마커의 중심 좌표를 이용하여 거리를 계산한다. 그러나 RGB 컬러 공간을 사용하는 영상 정보에서 컬러 마커를 추출하는 것은 매우 어렵다. 그 이유로는, 비디오 카메라로 획득한 영상에서 컬러 마커 부분의 컬러는 우월 파장에 의하여 결정되지만, 주변 픽셀에는 우월 파장이 아닌 다양한 파장의 컬러가 공존하고 있으므로 RGB 컬러 공간상의 좌표만으로는 원하는 컬러의 영역 분리가 어렵기 때문이다. 본 연구에서는 영상에서 국부영역의 우월 파장을 계산하여 컬러 영역을 탐색한다. 컬러의 우월 파장을 계산하기 위해 주변 픽셀 각각의 컬러에 대한 파장을 구하고, 그 중에서 가장 빈도 수가 높은 파장을 우월 파장으로 결정한다. (그림 10)은 3×3 픽셀 윈도우에서의 파장에 대한 빈도 수의 예를 나타내며, 우월 파장은 500nm가 된다.



(그림 10) 우월 파장

연기자의 동작을 촬영한 하나의 영상에는 여러 개의 동일한 컬러를 갖는 마커가 존재하므로, 각 마커에 대한 영역의 중심을 계산하여야 한다. 마커의 중심을 찾는 방법은 K-Means 알고리즘과 유사하며, 알고리즘의 처리과정은 다음과 같다. 영상의 모든 픽셀에 대하여 우월 파장을 구하고, 우월 파장이 마커의 컬러와 동일하면 마커테이블(그림 11)에 저장한다. 기존 마커 테이블에 저장되어 있는 위치 정보의 데이터 값과 비교하여 일정거리 내에 있다면 테이블에 있는 좌표와 현재 좌표의 평균 좌표 값을 계산하여 테이블을 갱신하며, 거리의 차이가 일정거리 이상이 될 경우에는 마커 테이블에 현재 위치를 추가로 삽입한다.

Index	Horizontal	Vertical
0	L_{h0}	L_{v0}
1	L_{h1}	L_{v1}
⋮	⋮	⋮
i	L_{hi}	L_{vi}

(그림 11) 좌측 마커 테이블

이러한 방법을 이용하여 영상 내에 있는 모든 픽셀에 대한 거리를 계산한다면 영상에 있는 마커의 모든 중심을 구할 수 있다. (그림 11)는 좌측 영상에 대한 마커 테이블의 한 예이며, 마커테이블은 좌/우 각각 1개씩 2개를 작성한다.

컬러 마커의 중심을 검색하는 알고리즘의 처리 단계는 다음과 같다.

- step 1. 영상에서 하나의 픽셀(I_h, I_v)을 중심으로 이웃 픽셀 n개에 대한 파장을 계산한다.
- step 2. 우월 파장을 계산하고, 우월 파장이 마커의 컬러로 판명되면 step 3을 수행하고, 아니면 step 7로 간다.
- step 3. 마커 테이블에 있는 좌표와 현재 위치 좌표와의 거리를 계산하고, 그 중에서 가장 짧은 거리 d_{min} 을 계산한다. 거리 d_i 는 다음 식 (8)에 의해 구한다. (I_h, I_v)는 영상의 좌표이며, (L_{hi}, L_{vi})는 마커 테이블에 저장된 좌표이다.

$$d_i = \sqrt{(I_h - L_{hi})^2 + (I_v - L_{vi})^2} \tag{8}$$

- step 4. 이미지상의 현재 위치(I_h, I_v)와 마커 테이블에 기록된 위치 정보를 비교하여 가장 짧은거리 d_{min} 가 일정거리 d_{thr} 보다 작으면 step 5로 가고, 아니면 step 6으로 간다. d_{thr} 값은 마커와 카메라 사이의 거리와, 카메라의 초점거리 및 마커들 사이의 공간적 거리에 따라 경험적으로 산출한다.
- step 5. 다음 식 (9)를 계산하여 마커 테이블을 갱신한 후 step 7로 간다.

$$L_{h\ new} = \frac{(L_{h\ old} + I_h)}{2}$$

$$L_{v\ new} = \frac{(L_{v\ old} + I_v)}{2} \tag{9}$$

- step 6. 마커 테이블에 이미지상의 현재 위치 (I_h, I_v)를 추가로 저장한다.
- step 7. 모든 픽셀에 대한 처리가 완료 되었으면 종료하고, 그렇지 않은 경우는 step 1로 간다.

5.2 컬러 마커의 거리계산

좌/우측 영상으로부터 각각의 마커 테이블을 작성한 후 2개의 테이블을 이용하여 마커의 거리를 계산한다. 마커 테이블에는 마커의 중심 좌표 정보가 기록되어 있지만, 좌/우측에 대응되는 마커의 위치를 파악해야 만 카메라로부터 마커까지의 거리 ($P(x_w, y_w, z_w)$)를 계산할 수 있다. 이를 위해 좌/우측 영상의 2개의 마커 테이블로부터 <표 4>와 같은 링크(link) 테이블을 유도한다. (그림 11)과 같은 좌/우측의 마커 테이블에 저장된 위치 정보를 이용하여 거리가 가장 가깝고, 카메라의 기하학적 위치에 따른 제약조건 ($R_k > L_k$)을 만족시키는 상대 위치를 검색하여 링크 테이블을 만든다. 링크 테이블상의 marker_id는 임의의 일련 번호를 의미하며, left index와 right index는 마커 테이블의 인덱스 값을 의미한다. 링크 테이블로부터 컬러 마커의 좌우 좌표를 파악할 수 있으므로 식 (6)과 식 (7)을 이용하여 카메라로부터 컬러 마커까지의 거리를 계산할 수 있다. 컬러 마커의 위치를 시간에 따라 계속 추적함으로써 현실공간 연기자 움직임 포착할 수 있다.

<표 4> 링크테이블

marker_id	Left	Right
id number	left index	right index
⋮	⋮	⋮

6. 실험 및 토론

본 논문에서 제안한 모션캡처 알고리즘은 Visual C++ 6.0으로 구현하였으며, 인텔 펜티엄 III 500MHz, Windows Me 환경에서 실험하였다. Pico-Pro(image grabber)를 사용하여 640×480 ($N = 640, M = 480$)의 영상을 획득하였다. 2대의 카메라는 NTSC방식의 초점 길이 $f = 8mm$, CCD 소자의 가로길이 $a = 4.8mm$, 세로길이 $\beta = 3.6mm$ 이며, 카메라는 컬러 PULNiX를 사용하였으며, 기본 선의 길이는 $b = 240mm$ 로 설정하였다.

디지털 영상평면에서의 픽셀 변위 (D_{hd} disparity)와 영상 평면의 변위 D_h 와의 관계는 (10)과 같다.

$$D_{hd} = \lfloor \frac{N}{\alpha} D_h \rfloor = \lfloor \frac{N}{\alpha} \frac{bf}{z_w} \rfloor \quad (10)$$

그러므로 $D_{hd} \geq 1$ 조건을 만족하여야 거리측정이 가능하며 D_{hd} 는 정수 값만이 존재한다. 그러므로 1픽셀에 의한 거리오차는 마커의 거리에 따라 달라지게 된다. <표 5>는 본 실험에서 사용한 시스템의 거리 측정 오차를 분석한 결과를 나타낸다.

<표 5> 실험장치의 픽셀변위와 거리와의 관계

D_{hd} (pixel)	30	31	32	33	34	...	57	58	59	60
z (mm)	8533	8258	8000	7758	7529	...	4491	4414	4339	4267

D_{hd} 값이 30에서 31로 1픽셀 차이가 날 때 거리는 275 mm오차를 갖지만 D_{hd} 값이 59에서 60으로 1픽셀 차이가 날 때는 72 mm거리 계산 오차를 갖는다. 본 논문에서는 카메라와 마커까지의 거리를 다양하게 변경하면서 실험하였다. 수식 (11)은 D_{hd} 값이 1픽셀 변할 때 거리에 따른 분해능을 계산한 것이다. 수식 (11)에서 거리 분해능을 늘려 오차를 좀 더 줄이기 위해서는 b, N 값을 높여야 함을 알 수 있다.

$$\Delta z_w = \frac{z_w^2}{\frac{N}{\alpha} bf - z_w} \quad (11)$$

이미지 획득으로부터 거리를 계산하고, 컬러 마커의 위치 계산까지 소요되는 시간은 약 5초로 실시간 모션캡처 이용에 적합하지 않은 것으로 판명되었지만 향후 전용 DSP 칩을 이용한다면 속도개선이 되어 사용이 가능할 것이다.



(그림 12) 캐릭터 애니메이션

현실공간 연기자 의 각 관절 부위에 부착된 컬러 마커의 움직임 추적을 통하여 가상공간의 캐릭터 애니메이션을 수행한다. 3차원 캐릭터 애니메이션의 수행을 위해 Viscap사의 VRT® 소프트웨어를 사용하였다. 각 관절의 위치 정보로부터 관절의 퍼짐과 굽힘 각도를 계산하여 SCL(script control language) 프로그램의 변수로 전달하며, 프로그램

수행에 의해 캐릭터의 움직임이 제어된다. (그림 12)은 캐릭터 애니메이션 수행의 한 예를 보여주는 화면이며, (그림 13)은 SCL로 작성된 프로그램의 예이다.

```

long      maxclimb ;
objnum   Hit ;
short    walk = 1, crunch, turn, sturn,
         amount, witch = -1 ;
long      xval, yval, zval, flag = 0, i, sum ;
resume   (1, 1) ;
if (flag == 0)
{
    property (me, "m_ixpos") = xpos (me) ;
    property (me, "m_iypos") = ypos (me) ;
    property (me, "m_izpos") = zpos (me) ;
    flag = 1 ;
    xpos (me) = property (me, "m_ixpos")
              + property (me, "m_xpos") ;
    ypos (me) = property (me, "m_iypos")
              + property (me, "m_ypos") ;
}
zpos (me) = property (me, "m_izpos")
          + property (me, "m_zpos") ;
xbend (1, me) = property (me, "m_thigh_x") ;
xbend (2, me) = property (me, "m_knee") ;
xbend (3, me) = property (me, "m_ankle") ;
waitf ;
    
```

(그림 13) SCL 샘플프로그램

7. 결 론

본 연구에서는 모션캡처를 위해 컬러 스테레오 카메라를 사용하는 거리측정 알고리즘을 제안하였다. 연기자의 각 관절 부위에 컬러 마커를 부착시키고, 이를 컬러 스테레오 카메라로 촬영한 후 본 논문에서 제안한 컬러 매칭방법과 중심영역 계산방법에 의해서 카메라로부터 마커까지의 거리를 계산한다.

스테레오 영상에서 마커의 컬러영역을 추출하는 방법은 RGB 컬러 정보를 CIE 컬러 공간으로 변환시킨 후 컬러의 파장을 계산한다. 국부지역의 우월 파장을 구하여 우월 파장이 마커의 컬러와 일치하는 영역을 추출한 후 마커 영역의 중심을 구하고, 카메라로부터 마커까지의 거리를 계산한다. 시간에 따른 마커의 위치 변화를 추적하여 캐릭터 애니메이션을 수행한다. 캐릭터 애니메이션은 Superscape사의 VRT®를 사용하였으며, SCL프로그램 이용하여 캐릭터의 움직임을 제어하도록 하였다. 시스템의 구현 결과로서, 스테레오 영상의 획득으로부터 마커의 거리 계산까지는 많은 시간이 소요되기 때문에 실시간 모션캡처 및 애니메이션에 이용이 불가능하였으나, 영상의 획득, 거리의 계산 및 마커

위치의 추적 과정 등을 전용 DSP 칩을 사용하여 구현한다면 실시간이용이 가능할 것이다. 본 연구에서는 적외선 카메라를 이용하지 않고도 비교적 저렴한 가격의 일반 컬러 카메라를 이용해서 모션캡처를 수행할 수 있다는 가능성을 제시하였다

참 고 문 헌

- [1] 윤중선, "운동제어를 위한 운동 포착 및 재현 시스템", 한국정밀공학회지, Vol.14, No.7, 1997.
- [2] Wyszecki, G., & Stiles, W. S., "Color Science(2nd ed.)," New York : Wiley, 1982.
- [3] Andrew S. Glassner, "Principles of Digital Image Synthesis," Morgan Kaufmann Pub., 1995.
- [4] G. Murch, IEEE CG&A, "Physiological Principles for the Effective Use of Color," pp.49-54, Nov. 1984.
- [5] <http://ceiba.cc.ntu.edu.tw/thu-02/Color Rendering of Spectra.htm>.
- [6] 이적식, "교차시각 스테레오카메라 모델해석", 전자공학회논문지, Vol.33 b, No.10, 1996.
- [7] Andre Redert, "Correspondence Estimation in Image Pairs," IEEE Signal Processing Magazine, Vol.16, No.3, pp.29-46, 1999.
- [8] J. Liu, "Stereo and Motion Correspondence in a Sequence of Stereo Images," Signal Processing : Image Communication, Vol.5, pp.305-318, 1993.
- [9] J. K. Aggarwal, "Human Motion Analysis : A review," Proc. IEEE Nonrigid and Articulated Motion Workshop, Piscataway, NJ pp.90-102, 1997.
- [10] Yoshifumi Oyama, "A Stereo Matching Algorithm Using Cross-Iteration Fusion Process for Crossing and Occlusion," 電子情報通信學會論文誌, Vol.J80-D-II, No.6, pp. 1941-1449, 1997.
- [11] Shigeru Kimura, "A Convolver-Based Read Time Stereo Machine (SAZAN)," IEEE, pp.457-463, 1999.



이 희 만

e-mail : hlee@seowon.ac.kr

1984년 고려대학교 전자공학과 졸 BS

1986년 한국과학기술원 전기 및 전자공학과 졸 MS

1994년 Texas A&M Electrical Eng. Ph.D

1996년~현재 서원대학교 컴퓨터정보통신 학부 조교수

관심분야 : 가상현실, 멀티미디어, 컴퓨터그래픽스



서 정 만

e-mail : seojm@unitel.co.kr

1985년 충북대학교 컴퓨터공학과 졸업
(공학사)

1988년~1993년 엘지전자 컴퓨터연구소
주임연구원

1993년~1999년 삼성중공업 중앙연구소
선임연구원

1998년 충북대학교 컴퓨터공학과 졸업(공학석사)

2001년 충북대학교 박사수료

2000년~2002년 극동정보대학 컴퓨터게임과 교수

2002년~현재 한국재활복지대학 컴퓨터게임개발과 교수

관심분야 : 데이터베이스, 실시간처리, 컴퓨터게임, 문화 콘텐츠



정 순 기

e-mail : soonkey@cbucc.chungbuk.ac.kr

1983년~1985년 한국 데이터통신, 책임
연구원

1990년~1994년 Rijksuniversiteit Gronin-
gen, computing science

1986년~현재 충북대학교 전기전자컴퓨터
공학부, 교수

관심분야 : 데이터베이스, 실시간처리, 소프트웨어공학