

논문-03-08-2-08

## 얼굴 모션 캡처 애니메이션을 위한 추출 및 추적 알고리즘

이문희\*, 김경석\*\*

### Extracting & Tracking Algorithm for Facial Motion Capture Animation

Mun-Hee Lee\* and Kyoung-Sok Kim\*\*

#### 요 약

본 논문은 얼굴 모션 캡처 애니메이션을 위한 빠르고 정확한 추출 및 추적 알고리즘을 제안한다. 제안한 알고리즘은 추출 및 추적의 두 단계로 구성된다. 먼저 신경회로망 기반의 영역 병합 기법을 이용하여 입력 영상으로부터 다중 마크를 분리한다. 그 다음, 신경회로망 기반의 추적 알고리즘을 사용하여 각각의 프레임에서 추출된 다중 마크들을 추적한다. 실험의 결과는 추출단계에서는 노이즈를 제거하고 처리 시간을 줄일 수 있었다. 또한 낮은 프레임율에서도 성능이 좋은 추적 결과를 얻을 수 있었다.

#### Abstract

In this paper, we propose fast and precise extracting & tracking algorithm based on general camera and frame grabber for facial motion capture animation. Proposed algorithm consists of two steps, extracting and tracking. The former is to separate multiple markers from input image using region merging based on neural networks. The latter is to track extracted multiple markers at each frame using tracking algorithm based on neural networks. In the experiment, we could remove noise and reduce processing time in the step of extraction. In addition, we could have good tracking results in the low frame rates.

#### I. 서 론

일반적으로 3D 애니메이션 도구를 이용하여 3D 애니메이션 영화, SF 영화 그리고 컴퓨터 게임 등을 만들 때 키 프레임 애니메이션 또는 동작 제어 애니메이션 기술을 적절히 혼합하여 수작업으로 캐릭터를 설정하고 제어함으로써 3D

캐릭터의 움직임을 하나하나 직접 만들어 낸다. 따라서 각각의 동작을 수작업으로 수행하여야 되기 때문에 작업량이 매우 많아 많은 인력과 작업시간을 필요로 하였으며 또한 캐릭터의 자유롭고 유연하며 섬세한 동작표현은 매우 어려웠다.

모션 캡처 시스템은 바로 이러한 문제를 해결하기 위해 사용되는 장비로써 최근 3D 캐릭터의 제작시 거의 필수적인 요소로 등장하고 있다. 이 장비는 하이테크 멀티미디어 콘텐츠 개발에 있어서 가장 진보된 기술이며 가장 고가의 장비에 해당하고 실제 액터의 얼굴 또는 각 관절에서 움직임을 포착하여 그 특징점을 추출한 후 생성된 정보를 3D 그래픽 제너레이터와 연결함으로써 실제 배우와 같은 섬세하고 유연한 동작표현의 고품위 3D 애니메이션을 제작할

\* 대구산업정보대학 정보통신계열  
Group-Dept of Informaion and Comm., Daegu Polytechnic College

\*\* 부산대학교 정보컴퓨터공학부  
Dept. of Computer Science and Engineering, Pusan National Univ

※ 본 논문은 2002년도 대구산업정보대학 연구구조성비에 의하여 수행되었습니다.

수 있게 해준다<sup>[8][10]</sup>.

최근에는 애니메이션 영화, 컴퓨터 게임, 기타 캐릭터 산업에서 액터의 얼굴표정과 입술모양의 표현에 있어서 많은 발전이 있었으며 이는 첨단 영상 처리기술의 활용으로 가능할 수 있었다. 현재, 우리 나라 3개 방송사에서도 모션 캡처 시스템을 이용한 사이버 캐릭터를 제작하여 방송을 하고 있으며 그 외 게임사 및 영화사에서도 특수효과제작 및 캐릭터 개발에 활발하게 적용하고 있다. 모션 캡처 시스템은 크게 세 가지의 방식으로 나눌 수 있다. 첫째로는 액터의 관절에 여러 개의 자석센서를 부착하고 유선 또는 무선으로 컴퓨터에 연결하여 모션을 캡처하는 자기(Magnetic)방식이 있고 둘째로는 각 관절에 슬라이더 엔코더(slide encoder)를 부착하여 컴퓨터로 각도를 전달하는 기계(Mechanic)방식이 있으며 마지막으로 골프공 모양의 적외선 반사물질이 발린 반사체를 액터의 각 관절에 부착하여 두고 적외선을 쬐이면 반사된 적외선 마커를 6 ~ 32대 가량의 적외선 비디오 카메라로 촬영하여 여러장의 2차원 이미지를 얻은 후 서로 비교하여 3차원 공간상의 좌표를 파악하는 광학(Optical)방식이 있다. 자기방식은 가격이 저렴한 반면 마그네틱 센서의 부정확한 특성과 많은 감지선 및 송신기를 몸에 부착함에 따른 불편함 그리고 간단한 동작을 제외한 복잡한 동작에는 사용하기가 힘이 든다는 문제가 있다. 마찬가지로 기계방식은 각 관절에 기계장치를 달고 있어서 매우 불편하다<sup>[1][3]</sup>. 반면 광학방식은 전술한 방식들보다는 섬세하고 유연하며 또한 빠른 동작에서도 움직임을 더 잘 감지할 수 있다. 현재 많은 상업화된 모션 캡처 시스템이 개발되었고, 컴퓨터 캐릭터 애니메이션에 좋은 성능을 보이고 있지만, 기존의 시스템에서는 고속의 카메라, 프레임 그레버(frame grabber) 및 컴퓨터를 이용하여 240 ~ 1000 frame/sec 정도의 프레임을 이용한 추적기법을 적용하여 다중 마커(marker)를 추적하는 고속 하드웨어 기반 모션 캡처 시스템을 구현하고 있다. 또한 대부분의 시스템에서 많은 시간을 소모하는 후처리 단계인 모션데이터의 수정작업을 거쳐야만 한다. 이는 정확하고 안정적인 2/3차원좌표의 획득과 획득된 2/3차원좌표의 신뢰성있는 궤적을 추적하기가 어렵다는 점에 기인한다.

본 논문에서는 고가의 고속 카메라를 이용한 기존의 광학방식 모션 캡처 시스템에 비해 일반 저가의 카메라를 이용하여 얼굴 움직임을 빠르고 효과적으로 추적할 수 있는 얼굴 애니메이션용 광학방식 모션 캡처 시스템 기술을 제시하고자 한다. 관련연구, 얼굴 움직임의 추출 및 추적 알고리즘, 실험 결과 및 분석 그리고 결론으로 끝을 맺고자 한다.

## II. 관련 연구

모션캡처 기술은 비교적 최근에 시작된 기술분야로 아직은 해결해야할 많은 문제점이 존재하지만 특히 마커의 추출 및 추적과 차폐(occlusion)현상의 해결이다. 최근 추적기술의 발달로 정확도가 많이 향상되고 있는 추세이기는 하나, 아직 만족할만한 기술 수준은 아니다. 특히, 차폐 문제는 광학식 시스템의 가장 중요한 단점중의 하나이므로 반드시 해결되어야 할 문제점이다. 이의 해결을 위한 연구로는 칼만 필터(Kalman filter)를 이용한 마커추적기술, 스테레오 컬러 카메라를 이용한 거리 측정 방법을 이용한 마커 추적 기술에 관한 연구 등이 있다<sup>[13]</sup>. 여기서는 전신용(full body) 및 얼굴용(facial) 모션 캡처 애니메이션을 위한 기존의 마커 추적 기술을 보이고자 한다.

### 1. 전신용 모션 캡처 애니메이션을 위한 3차원 좌표 마커 추적 기술

#### 1.1 칼만 필터를 이용한 마커 추적

시간의 흐름에 따른 3차원 좌표의 추적을 위해서 가장 간단하게는 현재의 프레임에서 이전 프레임의 특정 좌표의 위치와 가장 가까운 곳의 특징점을 추적하는 방법이 있다 (nearest neighborhood method)<sup>[14]</sup>. 그러나, 이 방법은 마커의 움직임이 많을때는 적당하지 않으며, 캡처된 움직임에 노이즈가 포함되므로, 잡음이 존재하는 시스템에서 안정적으로 추적할 수 있는 방법이 필요하다. 따라서 노이즈가 존재하는 물리적 시스템의 상태를 예측, 측정, 추적, 제어하는데 널리 사용되는 칼만 필터를 이용한다<sup>[6][15]</sup>. 칼만 필터는 가우시안 노이즈를 가진 선형 동적 시스템에 대한 최적의 예측 방법을 제공하므로 동작 예측 분야에서 가장 널리 알려진 기법중의 하나이다. 칼만 필터는 예측 단계와 갱신 단계의 두 단계로 구성된다. 예측 단계는 현재의 상태를 적용하고 다음 프레임의 상태 모델과 측정 모델에 대한 사전 예측을 수행한다. 갱신 단계는 피드백을 수행한다. 즉, 향상된 사후 측정 벡터를 획득하기 위해 실제 측정 벡터와 사전 측정 벡터를 통합한다. 향상된 측정 벡터는 교대로 갱신 단계에 입력되고, 예측과 갱신 단계는 반복한다. 즉, 칼만 필터는 실시간에서의 동작의 예측치와 측정치의 비교를 통해 추적을 제어한다. 그러나 실제 측정치가 없는 경우, 예측값에 대응되는 특징점이 없는 경우에는 현재 상태의 올바른 갱신이 이루어지지 않는다<sup>[6][7]</sup>.

## 1.2 스테레오 컬러 카메라를 이용한 거리 측정에 의한 마커 추적

여러대의 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 카메라 모델과 평행 카메라 모델이 많이 사용된다. 수렴 카메라 모델은 인간의 눈과 같이 카메라의 두 광축이 하나의 고정 점에서 만난다. 대응되는 좌표의 위치검색이 용이하고, 또한 대응점으로부터 거리 계산이 용이하기 때문에 대부분의 스테레오 비전은 평행 카메라 모델을 사용한다. 실제로 카메라는 핀홀(pinhole)이 아니므로, 보다 정확한 거리 계산을 위해서는 보정방법을 사용해야한다<sup>[2]</sup>. 마커의 움직임을 스테레오 컬러 카메라로 촬영하고, 디지털화된 영상에서 컬러 마커를 추출한다. 스테레오 카메라를 사용하므로, 좌측 화면에서의 컬러 마커 위치와 우측 화면에서의 동일한 컬러 마커 위치를 검색한 후 카메라로부터 컬러 마커의 현재 위치를 계산할 수 있다. 컬러 마커의 위치를 시간에 따라 계속 추적한다면 액터의 움직임을 알 수 있다. 그러나 이 방법은 스테레오 영상의 획득으로부터 마커의 거리 계산까지는 많은 시간이 소요되기 때문에 실시간 모션 캡처 및 애니메이션에 이용이 불가능하다<sup>[2]</sup>.

## 2 얼굴용 모션 캡처 애니메이션을 위한 2차원 좌표 마커 추적 기술

### 2.1 마커 영역화 기법

마커의 영역화 기법은 정지영상에서 배경과 마커 또는 다중 마커들 사이에서 원하는 마커를 가능한 정확하게 추출하여 다음단계인 추적시스템(정합신경회로망)에 제공하기 위한 것이다.

영역화 기법으로는 경계선 기반(edge based) 영역화 기법과 영역확장(region growing) 영역화 기법으로 나뉘어 질 수 있다<sup>[4][12]</sup>.

#### 2.1.1 경계선 기반의 영역화 기법

경계선 기반의 영역화 기법은 정지영상 내에서 경계선을

추출한 후 이 경계선 정보를 이용하여 마커영역을 추출해내는 기법이다. 이 기법은 영역의 다양한 정보를 제공해 주는 반면 노이즈에 민감하고 정확한 경계선의 추출이 특성상 어려우며 많은 계산량을 필요로 한다.

#### 2.1.2 영역확장 영역화 기법

영역확장 영역화 기법은 입력영상의 각 화소들로부터 특징을 추출하는것이 아니라 일정한 화소집단, 즉 화소집단이 갖는 유사성을 고려하여 영역을 추출하는 기법이다. 또한 영역확장 영역화 기법은 잡영(노이즈)이 많고 물체의 세밀한 정보를 요구하지 않는 경우에 매우 좋은 성능을 발휘한다.

### 2.2 마커 추적 기법

얼굴 움직임 추적을 위한 기존의 추적기법으로는 크게 중심점 추적기법과 상관 추적기법이 있다<sup>[4][5]</sup>.

#### 2.2.1 중심점 추적 기법

중심점 추적 기법은 현재 프레임의 정지영상에서 마커를 배경으로부터 분리한 후 추출된 각 마커들의 중심점을 이전 프레임과 비교하여 추적하는 기법이다. 이 기법은 간단한 계산을 수행하므로 처리시간이 빠르며 구현이 쉽지만 물체의 중심점만을 비교하기 때문에 프레임간의 마커의 위치 차가 클 경우, 마커를 놓칠 가능성이 많다. 그리고 중심점 추적기법을 모션캡처에 사용할 경우 고속의 동작을 캡처해 내기 위해서는 프레임 비율이 높은 카메라와 프레임그레버가 필요하다.

#### 2.2.2 상관 추적 기법

상관추적기법은 동영상의 시간적 공간적 상관관계를 이용하여 움직임 벡터를 추정하고 추정된 움직임 벡터를 표적의 추적에 활용하는 방식이다. 현재 프레임의 마커의 위치에서 적당한 크기의 영역을 정의하고, 정의된 영역과 다음 프레임의 검색 영역과의 상관도를 계산하여 상관도가 가장 높은 영역으로 마커가 이동했다고 추정한다. 실제로 상관추적 기법을 이용하여 이동 마커의 움직임을 추정하기 위해서는 상관영역의 크기를 정하는 방법과 사용할 상관 함수를 결정해야한다. 이 기법은 탁월한 추적 성능을 보여주고 있지만 중심점 추적기법보다 계산량이 많다는 단점이 있다<sup>[4][9][11]</sup>.

### Ⅲ. 얼굴 움직임의 추출 및 추적 알고리즘

얼굴용 광학방식 모션 캡처 시스템의 카메라 입력 영상은 대부분 마커와 배경이 뚜렷한 대조를 이루고 있으며, 조명이나 얼굴, 눈, 치아 등 신체일부에 의한 반사로 입력 영상에 노이즈(noise)가 발생하기도 하며 효율적인 움직임 추적을 위해 빠른 처리시간과 정확성을 필요로 하고 있다.

본 연구에서는 이러한 점을 최대한 고려하기 위하여 2.2.2에서 설명한 일반적인 상관 추적 기법을 변형하여 우선 영상을 신경회로망을 이용한 분할 및 병합 기법(Split & Merge)으로 영역화한 후 마커를 추출하여 전후 프레임에서 추출된 마커들에 대하여 상관도를 계산하는 방식의 다중 마커 추적 기법을 제시하였다. 따라서 제안한 방법은 미리 마커를 추출하여 추출된 마커들에 대하여 상관도를 계산하기 때문에 상관 추적 기법의 추적 성공률을 유지하면서 계산량은 획기적으로 줄일 수 있다.

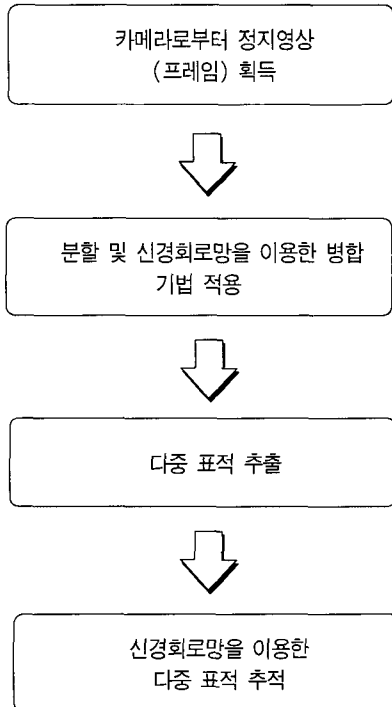


그림 1. 얼굴용 광학 모션 캡처 시스템 흐름도  
Fig. 1. Facial Optical Motion Capture System flow diagram

실제 사람의 얼굴에 다수 마커들을 부착하고, 그림 2의 얼굴 움직임 추출 전용 헤드셋(headset)을 이용하여 영상을 입력하고, 신경회로망을 이용한 분할 및 병합 기법을 적용하여 다중 마커들을 추출한 후, 신경회로망을 이용하여 다수 마커들을 추적한다.

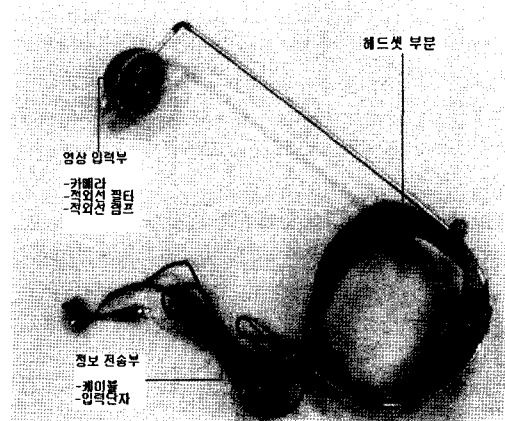


그림 2. 얼굴용 광학 모션 캡처 시스템의 헤드셋  
Fig. 2. Headset of Facial Optical Motion Capture System

#### 1. 신경회로망 기반의 분할 및 병합(Split & Merge) 기법을 이용한 다중 마커 추출

본 연구에서는 영상에서 마커를 추출하는 기법으로 신경회로망을 이용한 분할 및 병합기법을 사용하였다.

새로운 프레임이 입력되면 먼저 영역을 균등분할(예: 가로 32, 세로 24)하여 이진화와 동시에 분할된 공간 내에서의 마커 픽셀수를 구한 후 마커 픽셀수가 기준값 이상일 경우 해당마커를 활성화 시키고 기준값 이하의 경우 노이즈로 처리한다.

본 연구에서 제안한 신경회로망을 이용한 마커 병합에 사용할 입력패턴의 생성방법은 다음과 같다. 적외선을 사용하여 반사되어 오는 영상을 분석해보면 일반영역에서의 픽셀들은 대개 0에서 200이하의 값들을 가지고 있는데 반해 마커 픽셀들은 240에서 255의 값들을 가진다. 따라서 본 연구에서 이진화에 적용한 마커 threshold(MarkThreshold)값을 230으로 정하여 실험하였다.

해당마커를 활성화 또는 비활성을 정함에 있어서 기준값(ActThreshold)은 임의의 값으로 정할 수 있으며 너무 클 경우 마커를 잃어버릴 수 있으며 반대로 너무 적으면 노이

스로 인한 오류 및 계산량의 증가를 초래할 수 있다. 따라서 많은 실험을 거친 후 경험적으로 적당한 기준값을 정하여야 한다.

한 프레임내의 마커 패턴들은 각각 이웃한 다른 마커들과 매우 가까운 거리에 모여서 하나의 마커를 형성하고 있다. 따라서 이러한 현상들을 반영하여 입력되는 마커 패턴이 이웃한 마커들과 하나의 마커집단으로 병합되기 위하여는 최대한 떨어져 있을 수 있는 거리 경계값  $\rho$ 를 정하여 두어야 한다. 이 경계값 또한 프레임내의 마커패턴을 분석하여 경험적으로 정하여 지며  $\rho$ 값이 너무

클 경우 전혀 다른 마커패턴이 서로 잘못 병합되어 실제보다 적은 수의 마커를 만들어 낼 수 있으며 반대로 너무 적은  $\rho$ 값은 노이즈에 약하고 실제 마커 수보다 더 많은 마커를 만들어 낼 수 있으며 불필요한 계산을 초래할 수 있다.

## 2. 신경회로망을 이용한 다중 마커 추적

마커의 추적은 연속적인 프레임들 속에서 수행되는 단계이다. 단일 프레임에서 추출된 마커들을 각각 저장하고 계

### [알고리즘 1] 신경회로망을 이용한 마커병합에 사용할 입력패턴의 생성 알고리즘

```

for(x=0;x<XMax;x++){ //영역내의 마커픽셀수를 구한다
    for(y=0;y<YMax;y++){
        if (Pixel(x,y) >= MarkThreshold) Block(i,j)++; //MarkThreshold = 230(범위:0~255)
    }
}
for(i=0;i<BxMax;i++){ //해당영역을 마커패턴으로 승인할 것인지를 정한다
    for(j=0;j<ByMax;j++){
        if (Block(i,j) >= ActThreshold) BlockAct(i,j)=1; //ActThreshold = 20(범위:0~99)
        else BlockAct(i,j)=0;
    }
}

```

- \* 주 :
1. Pixel(x,y)는 프레임 내의 화소값으로 0~255의 그레이레벨값을 가진다.
  2. Block(i,j)는 해당블록에서의 마커 픽셀수이다.
  3. BlockAct(i,j) 해당블록의 활성화 여부를 체크하여 입력패턴으로 사용한다.

### [알고리즘 2] 신경회로망을 이용한 다중 마커 추출 알고리즘

각 프레임에 대하여 STEP 1 ~ STEP 5을 반복 수행

STEP 1: 실시간 입력 동영상에 대하여 입력패턴 생성

STEP 2: 신경회로망을 이용한 병합을 수행한다.

(1)새로운 블록 패턴Xi가 주어지면 최소 거리의 승자 뉴런SAMj( 이전프레임의 해당마커)를 선택한다.  
 $Distance(Minimum) = ||Xi - SAMj||$

(2)병합 경계값( $\rho$ )을 비교한다.

if(Distance(Minimum) <  $\rho$ ) Goto STEP 4 ;허용거리( $\rho$ )이내의 경우 정합에 성공  
 else Goto STEP 3 ;가깝지만 동일 패턴으로 보기 힘들다

STEP 3: 새로운 뉴런을 생성한다.

$SAMj = Xi$

STEP 4: 승자 뉴런 블록 패턴 Xi 와 SAMj 를 병합한다.

**[알고리즘 3] 신경회로망을 이용한 다중 마커 추적 알고리즘**

각 프레임에 대해서 STEP 1 ~ STEP 6을 반복 수행

STEP 1: 입력 동영상에서 추출된 마커들을 입력패턴  $X_i$  로 한다.

STEP 2: 각 입력 패턴에 대해서 최소거리의 승자뉴런  $TRA_j$ (이전프레임에서 추적된 마커들의 집합)를 선택한다.  
 $Distance(Minimum) = ||X_i - TRA_j||$

STEP 3: 정합 경계 값( $\rho$ )을 검사한다. 승자 뉴런  $TRA_j$ 와 입력패턴  $X_i$ 의 최소거리가 경계값  $\rho$  보다 작으면 동일 마커로 결정한다.  
 if(Distance(Minimum) <  $\rho$ ) Goto STEP 6  
 else Goto STEP 4

STEP 4: 정합되지 못한 입력 패턴  $X_i$  에 대해 다시 정합을 수행한다.  
 if(!  $X_i.Match$  <  $\rho$ ) Goto STEP 5  
 else Goto STEP 1

\*  $X_i.Match$ : 1차 정합에서 입력패턴의 정합여부를 의미한다. 만약  $X_i.Match$ 가 TRUE 라면 정합되었음을 의미한다.

STEP 5: 정합 경계값( $\rho$ )을 일정 크기만큼 증가시킨다.

$$\rho = \rho + \alpha$$

Then Goto STEP 3

STEP 6: 입력 패턴  $X_i$ 와  $TRA_j$ 을 정합한다.

속적으로 연결되는 프레임에서 추출된 마커들과 저장된 마커들을 여러 가지 추출된 정보를 이용하여 관련시키고 예상하여 추적하는 단계이다.

본 논문에서는 마커의 추적을 위해서 기존의 상관 추적 기법을 변형하여 인공신경회로망을 이용한 다중 마커 추적 알고리즘을 사용하였다. 추출된 마커들로부터 이전 프레임의 마커들과의 중심점의 위치, 각 마커들의 운동벡터, 마커들간의 상대적 위치 관계등을 고려하여 상관도를 계산한 후 정합 하는 방법을 통해 추적을 수행하였다. 마커 추적 단계를 수행한후 마커의 히스토리에 현재 마커의 수정된 벡터, 크기, 그리고 위치들을 갱신한다.

여기서 정합을 위한 경계값 $\rho$ 는 상기한 프레임간에 상관도를 합산하여 수치적으로 나타낸 값으로 많은 실험을 거친 후 경험적으로 정한다.

병합에서와 마찬가지로  $\rho$ 값이 너무작으면 실제보다 많은 뉴런이 만들어져 2차 또는 2차의 정합을 해야하는 오류가 발생할 수 있으며 반대로 너무 크면 잘못된 병합으로 인한 치명적인 마커 손실을 초래할 수 있다.

**IV. 실험결과 및 분석**

**1. 실험환경**

시스템 사양은 Pentium IV 1.7GHz, RAM 512MB, Windows 2000 Server에서 Visual C++ 6.0을 이용하여 알고리즘을 구현하였다. 실험에는 적외선필터를 장착한 전용 헤드셋(그림 3)과 프레임 그레버를 사용하여 320×240 정지 영상(프레임)을 초당 15 프레임/ 30 프레임으로 처리하였으며 실제 얼굴에는 20개의 마커를 부착하여 사용하였다.



그림 3. 얼굴용 광학 모션 캡처 시스템  
 Fig. 3. Facial Optical Motion Capture System

2. 결과 및 분석

2.1 신경회로망 기반의 분할 및 병합기법을 이용한 다중 표적 추출 결과

일반적으로 카메라를 1대를 이용하면은 고개를 들거나 내리거나 하는 행동을 취했을 경우 카메라의 광각에서 벗어나게 되어 마커를 추출할 수 없게 된다. 따라서 이러한 단점을 보완하고자 그림 3처럼 카메라를 얼굴 전면의 움직임과 함께 움직일 수 있도록 자체 카메라 헤드셋을 제작하여 본 시스템에 적용하였다. 본 논문에서 제시한 신경회로

표 1. 제안된 추출기법의 적용 결과  
Table 1. The application results of the proposed extracting method

제안된 추출 기법의 적용결과	얼굴부분						비고
	이마	눈	볼	턱	입술		
다중 표적 추출 결과(단위:%)	96	98	97	96	98		대부분
	추출 성공률/전체 회수						손실 없음.

망 기반의 분할 및 병합기법을 적용한 결과, 그림 4에서와 같이 카메라 광각 내에 포착된 모든 마커에 대해서 빠른 시간 내에 다중 마커들을 추출함으로써 실시간 처리에 적합한 좋은 결과를 나타내었다.

2.2 신경회로망을 이용한 다중 표적 추적기법 실험 결과

본 논문에서 제시한 추적기법의 실험결과는 그림 5에서 볼 수 있듯이 이마, 눈, 볼 그리고 턱 부분의 경우는 마커의 변화량이 일정하여 프레임간의 추적을 잘 수행하였다. 그러나 입술부위의 경우는 다른 부위에 비해 추적의 결과가 좋지 못한데, 그 이유는 입술 부위 자체가 가지는 근육의 물리적인 특징 때문에 움직임의 다양한 변화가 급격하게 발생하기 때문이다. 이러한 급격한 큰 변화량이 발생한 경우, 신경회로망에서 상대적으로 변화량이 적은 부분에 적용하였던 단일 경계값( $\rho$ )만을 적용함으로써 입술부위에 대해 몇몇 마커의 추적을 실패하는 경우가 발생하였다. 따

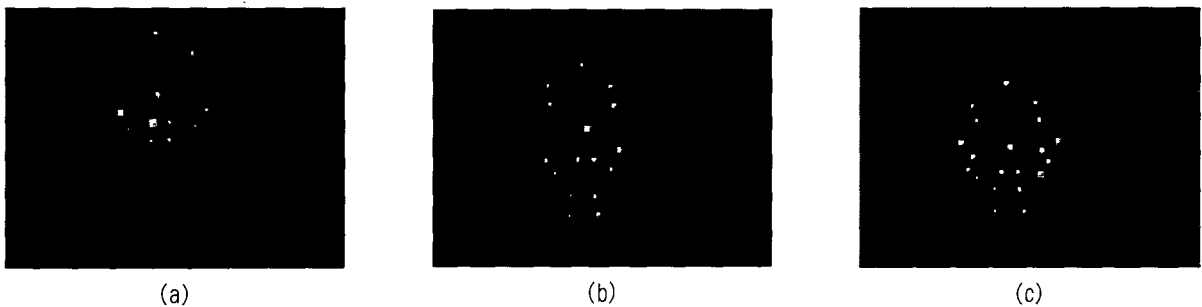


그림 4. 신경회로망을 이용한 분할 및 병합기법 결과 (a) 추출 결과 I (12개의 마크) (b) 추출 결과 II (16개의 마크) (c) 추출 결과 III (22개의 마크)  
Fig. 4. Results of Split & Merge method using Neural Networks (a). Extracting result I (12 Mark) (b). Extracting result II (16 Mark) (c). Extracting result III (22 Mark)

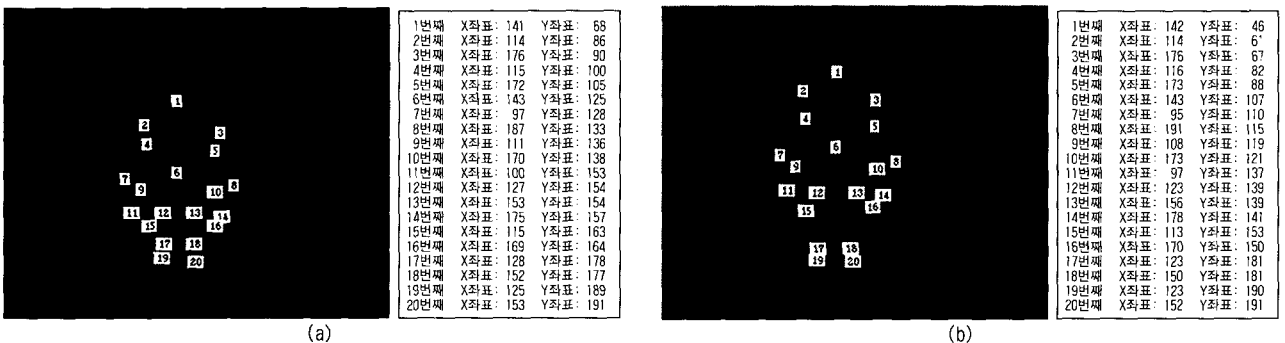


그림 5. 신경회로망을 이용한 추적기법의 결과 (a) 시간 t일 때 마커들의 위치 (b) 시간 t+1일 때 마커들의 위치  
Fig. 5. Results of tracking method using Neural Networks (a) Position of a markers at time t (b) Position of a markers at time t+1

표 2. 제안된 추적기법의 적용 결과 (a) 15 프레임 (b) 30 프레임  
 Table 2. The application results of the proposed tracking method (a) 15 frames (b) 30 frames

제안된 추적 기법의 적용 결과		얼굴부분				
		이마	눈	볼	턱	입술
다중 표적 추적 결과(단위:%)		97	98	96	94	84
		추적 성공률 / 전체실험회수				

제안된 추적 기법의 적용 결과		얼굴부분				
		이마	눈	볼	턱	입술
다중 표적 추적 결과(단위:%)		100	100	99	98	97
		추적 성공률 / 전체실험회수				

라서 본 논문에서는 추적에 실패한 부분에 대해서 경계값을 변화하여 즉, 다중 경계값을 적용하여 추적되지 못한 마커들에 대해서 부가적인 추적을 함으로써 그에 대한 정확도를 높일 수 있었다. 먼저 정확하고 빠른 추적은 일정한 프레임률(frame rate)의 획득과 빠르고 정확한 다중 물체 추출이 가능하도록 하는 것이 선행되어야 된다. 또한 얼굴의 움직임 추적은 신체의 다른 부위보다는 공간적 시간적 변화량이 적고 부분별로 다른 특징을 가지고 있기 때문에 좀더 섬세하고 완벽한 추적기법을 요구한다.

실험 결과는 다음과 같으며 전체실험회수는 20회이며, 제시한 추적 성공률은 평균수치이다. 그림 5에서 각 마커들은 서로 다른 번호를 가지고 있으며 프레임간에 동일한 번호로 표현됨으로써 추적의 결과를 나타내고 있다. 또한 각 마커들에 대한 X, Y 좌표상의 위치를 보여준다. 표 2는 제안된 추적 기법에서 프레임율이 높을수록 입술부근의 마커 추적 성공률이 우수함 보여주며, 전체적으로도 추적 성공률이 향상되었음을 보여준다. 표 3은 기존의 중심점 추적 기법과 제안한 상관 추적 기법의 실험 결과를 비교하였다. 표 3에서 볼 수 있듯이 제안한 상관 추적 기법은 추적 결과 발생하는 미 추적 마커들에 대해 재 추적과정을 수행함으로써 향상된 추적결과를 얻을 수 있었다. 또한 한 프레임을 처리하여 마커를 추적하는데 소요된 시간은 150 $\mu$ s로 기존의 중심점 추적 기법의 소요 시간 1200 $\mu$ s보다 상당히 우수함을 알 수 있었다. 이는 추적단계에서 중심점 추적 방식에 비해 상관 추적 기법이 다소 많은 계산량을 필요로 하지만,

추출단계에서 일반 영영확장 기법에 비해 제안한 신경회로망을 이용한 영역 분할 및 병합 기법이 약 70배에서 100배 정도 빠르게 나타나기 때문에 전체적으로 소요되는 시간은 기존의 추출 및 추적 방식에 비해 제안된 추출 및 추적 방식이 약 8배 정도 빠르다. 본 논문에서 제시한 신경회로망 기반의 추적 기법은 간단하고 효과적인 실시간 처리에 용이한 추적 기법으로 사려된다.

### V. 결 론

본 논문에서는 얼굴 모션 캡처 애니메이션을 위한 효과적인 추출 및 추적알고리즘을 제시하였다. 비교적 마커와 배경과의 대조가 잘 나타나고 노이즈가 발생하는 입력영상에 대해 영역병합기법과 신경회로망을 이용한 상관 추적기법을 적용한 결과, 다중 마커를 정확하고 빠르게 추적하여 얼굴용 광학방식 모션 캡처 시스템에 적용 가능하다고 본다. 얼굴의 경우, 마커(표적)의 움직임이 한정된 공간상에서 이동하지만 섬세한 움직임의 추적이 필수적이고 얼굴 부위에 따라 움직임 정도가 다양하기 때문에 한정된 공간(얼굴) 내에서 부분에 따라 적응적으로 추적할 수 있는 기법에 관한 연구가 요구되어진다. 낮은 성능의 프레임 그레버와 저가의 일반 카메라를 이용하여 제시한 알고리즘으로 실험한 결과 대부분 추적에 성공하였다.

향후 연구 과제로 좀더 섬세하고 완벽한 얼굴 움직임을 추적하기 위한 얼굴의 각 부분에 대한 특징을 고려한 추적 기법이 제시되어야 될 것이다. 또한 2대 이상의 카메라를 이용하여 얼굴 전반에 대한 움직임을 추적하고 동시에 얼굴 움직임의 3차원 정보의 생성에 대해 연구가 이루어져야 되겠다. 또한 이러한 기술을 바탕으로 얼굴뿐만 아니라 전신(Full-body)용 모션 캡처 시스템의 연구개발도 병행되어야 될 것이다.

표 3. 기존의 알고리즘과의 비교  
 Table 3. Comparing with the previous study

	기존의 중심점 추적기법	제안한 상관 추적 기법
15프레임	82%	94%
30 프레임	86%	99%
프레임 당 추적 소요 시간	1200 $\mu$ s	150 $\mu$ s



## 참 고 문 헌

- [1] 김운순, 김영수, "3차원 캐릭터 애니메이션 기술 동향", 정보과학회지, 제17권, 제2호, pp.48~59, 02, 1999.
- [2] 이희만, 서정만, 정순기, "모션캡처 애니메이션을 위한 거리 측정 방법", 한국정보처리학회논문지, 제9-B권, 제1호, pp.129~138, 2, 2002.
- [3] 이문희, 김철기, 김경석, "Facial Animation을 위한 다중 마커의 추적", 한국멀티미디어학회 춘계학술발표논문집, 제4권, 1호 pp.553~557, 06, 2001.
- [4] 서울대학교, "실시간 표적인식 및 추적기법 연구", 자동제어특화연구센터, pp.8~18, 02, 1996.
- [5] 김형재, 이문희, 남양일, 이도엽, 차의영, "신경회로망을 이용한 동적 표적의 추적에 관한 연구", 한국정보처리학회 99 춘계학술발표논문집, pp.1355~1358, 04, 1999.
- [6] 이동훈, 추창우, 김성진, 정순기, "광학식 동작 포착에서 동작 모델을 이용한 신뢰성있는 3-D 좌표 추정 및 추적", 한국정보과학회논문지: 시스템 및 이론, 제27권, 제10호, pp.825~834, 10, 2000.
- [7] S. Ilaykin, "Adaptive Filter Theory," Prentice-Hall, Englewood Cliffs, New Jersey, 1986.
- [8] George Maestri, "Digital Character Animation," Brain&Computer, 1997.
- [9] D. A. Montera, S. K. Rogers, D. W. Ruck, and M. E. Oxley, "Object tracking through adaptive correlation," Optical Engineering, Vol. 33, No 1, pp.294~302, 01, 1994.
- [10] Michael J. Potel, "The Mystery of Motion Capture," IEEE Computer Graphic and Application, pp.14~19, Sept./Oct. 1998.
- [11] S. U Lee, S. Y. Chung, and R. H. Park, "A comparative performance study of several thresholding techniques for segmentation," Computer Vision, Graphic, and Image Processing, Vol. 52, pp.171~190, 1990.
- [12] M. Sonka, V. Hlavac and Roger Boyle, "Image Processing Analysis and Machine Vision," Chapman&Hall, pp.507~547, 1993.
- [13] S. K. Jung and K. Y. Wohn, "Tracking and Motion Estimation of the Articulated Object: A Hierarchical Kalman Filter Approach," Journal of Real-Time Imaging, 3(6), pp.415~432, 1997.
- [14] I. J. Cox, "A Review of Statistical Data Association Techniques for Motion Correspondence," International Journal of Computer Vision, Vol. 10, pp.53~65, 1993.
- [15] Greg Welch and Gray Bishop, "An Introduction to the Kalman Filter," TR 95-041, University of North Carolina at Chapel Hill Department of Computer Science.

## 저 자 소 개



## 이 문 희

- 2001년 2월 : 부산대학교 대학원 멀티미디어학과 공학박사 수료
- 1998년 6월~2002년 2월 : (주)인트빔 부설 멀티미디어 시스템 연구소 선임연구원
- 2002년 3월~현재 : 대구산업정보대학 정보통신계열 교수
- 주관심분야 : 모션 캡처 시스템, 인터넷 방송, 멀티미디어 응용, 디지털 영상 처리



## 김 경 석

- 1977년 2월 : 서울대학교 학부 졸업
- 1979년 2월 : 서울대학교 대학원 졸업
- 1988년 10월 : 미국 일리노이 주립대학교 (어바나-샴페인) 전산학박사
- 1988년 9월~1992년 2월 : 미국 노스다코타 주립대학교 컴퓨터학과 교수
- 1992년~현재 : 국제 표준화 기구 (ISO) 및 대응 국내위원회 활동
- 1992년 3월~현재 : 부산대학교 정보컴퓨터공학부 교수
- 주관심분야 : 데이터베이스, 인터넷 응용, 한글 정보 처리