

MPEG-4 FAP 기반 세분화된 얼굴 근육 모델 구현

이인서¹, 박운기², 전병우³
성균관대학교 전기전자 및 컴퓨터 공학부

Subdivided Facial Muscle Modeling based on MPEG-4 FAP

Insuh Lee¹, Woongi Park², Byeungwoo Jeon³
School of Electrical & Computer Engineering, Sungkyunkwan University

E-mail: {¹shui, ²rook, ³bjeon}@media.skku.ac.kr

Abstract

In this paper, we propose a method for implementing a system for decoding the parameter data based on Facial Animation Parameter (FAP) developed by MPEG-4 Synthetic/Natural Hybrid Coding (SNHC) subcommittee. The data is displayed according to FAP with human muscle model animation engine. Proposed model has the basic properties of the human skin specified by be energy funtional for realistic facial animation.

Keywords: MPEG-4, facial animation, multimedia communication, virtual reality

1. 서 론

인간의 얼굴은 가장 친밀감을 주며, 표정을 통해 많은 정보를 전달할 수 있는 의사전달 수단이다. 현재 활발한 기술개발이 이루어지고 있는 가상현실 분야에서는 사용자를 표현한 대리인이 행위의 주체가 되기 때문에, 대리인의 모델링 기술은 중요한 위치를 차지하고 있으며, 특히 얼굴은 가장 많은 정보를 전달할 수 있는 객체이기 때문에, 보다 사실적인 모델링, 애니메이션 기술¹이 적용되어야 한다.

MPEG-4(Moving Picture Experts Group, Phase 4) SNHC(Synthetic/Natural Hybrid Coding) 분과에서는 상호작용을 바탕으로한 자연영상과 합성영상의 통합 부호화 방법을 표준화하고 있으며, 첫번째 표준안에서 가상현실에서 사용할 수 있는 얼굴 객체의 부호화 방식으로 얼굴을 정의하는 FDP(Face Description Parameter)와 움직임을 정의하는 FAP(Facial Animation Parameter)를 제안하고 있다[1].

하지만 MPEG-4의 FAP는 얼굴 각 요소들의 움직임에 대한 정의만을 포함하고 있기 때문에, 사실적인 얼굴의 표정변화 및 Lip Synchronization등을 애니메이션 하기 위해서는 모델링된 얼굴 객체 안에 실제 사람 얼굴의 근육역할을 할 수 있는 동작 시스템이 구현되어야 한다

[2]. 본 논문에서는 FAP를 입력변수로 받아들여 사실감 있는 얼굴의 동작을 표현하기 위한 얼굴 근육 모델을 구현하는 방법을 제안한다. 다양한 얼굴 객체 모델에 적용되어 FAP를 통해 실시간으로 사실적인 표정 변화를 보여줄 수 있는 시스템의 구현이 본 연구의 목적이다.

2. 기존의 얼굴 모델 구현 방법

2.1 얼굴 모델

기존의 얼굴 애니메이션에 이용된 얼굴 모델은 대부분 3차원 객체를 표현하기 위한 vertex로 구성된 mesh 구조를 통해 얼굴 형태를 구성하고, vertex의 위치가 달라짐에 따라 변화하는 mesh의 형태를 통해서 얼굴 객체를 애니메이션 시키고 있다. 애니메이션 시키는 동작 방식은 크게 surface movement의 제어를 통한 방법과 사랍의 골격 및 근육구조를 모델링한 동작 시스템을 가지고 있는 방식으로 나눌 수 있다. 현재 많이 연구되는 분야가 얼굴 골격과 근육을 이용한 애니메이션 시스템이지만, 다양한 얼굴모델에 적용되기 힘든 문제를 가지고 있다. 본 논문에서는 기존의 얼굴 근육모델을 기반으로하여 다양한 얼굴 모델에 적용될 수 있고, MPEG에서 제안하고 있는 얼굴 애니메이션 파라미터인 FAP를 적용할 수 있는 얼굴 애니메이션 시스템을 제안한다. 제안된 얼굴 근육 모델은 크게 다음 두 부분으로 나눌 수 있다. 사람의 근육을 모델링한 근육 벡터 부분과 피부의 움직임을 제어하는 피부 움직임 함수 부분으로 전체 시스템이 이루어진다, 얼굴 각 부분의 움직임은 각각 근육모델의 조합을 통하여 다양한 얼굴 표정을 애니메이션할 수 있다.

2.2 얼굴 동작 변수(FAP)

MPEG-4에서는 얼굴 객체에 개성 있는 형태와 움직임을 부여하는 변수로 FDP와 FAP를 제안하고 있다. 얼굴 모델에 FDP가 입력으로 주어지면 일반 모델은 FDP가 갖고 있는 특정한 모습의 얼굴 모델로 변형되고, FDP에 이어 입력으로 주어지는 FAP를 토대로 얼굴

본 연구는 한국과학재단 1999 특성기초연구(1999-2-515-001-5) 지원에 의하여 수행되었음.

모델은 사용자의 의지에 따라 다양한 표정을 연출할 수 있게 된다[7].

얼굴 동작 변수(FAP)는 얼굴 표정 연출과 얼굴 내의 근육 모델에 관한 지식을 바탕으로 구성되고, 얼굴의 동작을 제어하는 변수로 사용되어 얼굴 모델이 자연스러운 표정을 지을 수 있도록 설정되었다[1]. 하지만 MPEG-4에서는 얼굴 객체 모델에 관한 표준을 정하지 않고, 단지 얼굴의 기본적인 움직임만을 표현한 FAP를 동작 변수의 표준으로 정하고 있기 때문에 FAP를 입력으로 한 얼굴 객체 모델은 다양한 방식으로 구현될 수 있다.

FAP는 TTS(text to speech) 등과의 연동에 필요한 음운에 따른 움직임과 표정에 따른 움직임을 표현하는 2개의 하이 레벨 변수와 눈, 입 등의 각 구성요소의 세부적인 움직임을 표현하는 66개의 로우 레벨 변수로 구성되어 있다. 하이 레벨 변수는 viseme와 expression의 2가지로 나뉘어 진다. Viseme는 음소(phoneme)에 따른 입의 움직임을 표현하며, viseme rendering을 통해 말하고 있는 사람의 입 모양을 표현할 수 있다. expression은 얼굴 모델이 가질 수 있는 즐거움, 슬픔, 화남, 두려움, 싫음, 놀람의 6개의 기본 표정을 표현하는 변수로, 구현된 얼굴 모델의 감정 표현을 전달하는데 효율적으로 활용될 수 있다.

로우 레벨 변수는 viseme와 expression을 이용하여 표현할 수 없는 얼굴 모델의 다양한 표정을 연출하기 위해 사용된다. 눈, 입, 턱 등 얼굴의 구성요소의 움직임을 표현하는 변수이므로 얼굴의 회전이나, 눈동자의 초점, 턱관절의 움직임 등 얼굴모델을 사실적이고 자연스럽게 표현하기 위한 여러 움직임을 표현하기 위해 사용된다.

기존에 연구 결과로는 Parameterized Approach를 적용한 IST face[5]등이 있다. IST face의 경우는 파라미터에 기반하여 근육 모델의 적용이 없어 사실감 있는 얼굴 표정의 구현이 어려웠다. 이 후에 발표된 추상 근육 동작(AMA)을 사용한 MIRALab의 Miraplayer[6]의 경우는 객체 모델의 움직임이 실제사람의 근육을 근사화 시킨 함수에 따라 표현되었기 때문에 이전의 연구결과에 비해 향상된 얼굴표정의 표현이 가능했다. 하지만 Miraplayer의 경우에 적용된 근육모델은 단순히 얼굴 모델을 구성하는 vertex의 이동을 근육 함수로 사용했기 때문에 불과 입술의 움직임 같은 세부적인 움직임에서는 사실적이지 못한 결과를 나타내었다. 또한 기존의 Computer Graphics에서 연구결과들로 인체 해부학을 바탕으로한 근육 모델들의 경우[2, 3, 4], 다양한 모델에 적용되기 어렵고, FAP를 이용한 애니메이션에 적합하지 못한 부분이 있다.

3. 개선된 얼굴 근육 모델 구현

FAP에 의해 움직이게 되는 얼굴 모델은 MPEG-4의 제안에 포함되어 있지 않으므로 표준화의 호환성 문제 없이 다양한 모델의 구현이 가능하다.

본 논문에서는 FAP에 기반한 사실적인 얼굴의 움직임을 구현하기 위해, 사람의 실제 근육을 모델링한 가상 근육 함수와 피부의 움직임을 만들기 위한 탄성 에너지 함수를 제안한다. 제안된 얼굴 근육 모델은 다양한 얼굴 모델에 적용될 수 있도록 실제 얼굴 골격 및 근육의 위치와는 관계없이, FAP에 의해 움직이게 되는 부분에 배치된 세분화된 근육 벡터와 피부의 움직임을 사실적으로 애니메이션시키기 위해, 제어함수로 사용되는 피부 움직임 함수로 구성된다. 또한 고차원의 얼굴

표정, 예를 들면 입술이 움직일 때의 볼의 움직임등을 구현하기 위해 배치된 근육 모델을 유기적으로 조합하여 전체 애니메이션 시스템을 구성하였다.

3.1 세분화된 근육 모델

FAP에 따른 얼굴 모델의 움직임을 위해서 세분화된 근육 모델을 제안한다. 본 연구에 적용된 근육 모델은 해부학적인 얼굴 근육의 역할이 아닌, FAP에 의해 제어되는 표정 연출을 위해 사용된다. 따라서 각 표정 연출 및 입술의 세밀한 움직임을 표현하기 위한 기본적인 근육 벡터를 만들고 근육 벡터간의 유기적인 상호작용을 통해 동작하는 세분화된 근육 모델로 얼굴의 움직임을 표현하였다. 얼굴 한 부분의 움직임은 얼굴 전체에 영향을 미치므로, 세분화된 근육 벡터의 유기적인 조합은 중요한 위치를 차지한다. 특히 립싱크시의 입술의 움직임은 대단히 정교하면서, 얼굴 전체의 움직임을 좌우하기 때문에 입술 주변의 FAP 제어점을 중심으로 세분화된 근육 모델을 배치함으로써 입의 움직임을 사실적으로 제어할 수 있다. 또한 하이 레벨 변수에 의한 표정변화를 제어하기 위한 근육 모델, 예를 들면, 웃음 근육은 입술과 눈, 코 등의 모습에 영향을 미치게 되므로 웃음에 의해 영향을 받는 구성 요소들의 근육 벡터를 조합하여 웃음 근육 모델을 구성하면 각 벡터들의 상호동작에 의해 웃음이라는 표정을 연출할 수 있다. 세분화된 근육 모델은 움직임이 작은 여러 근육 벡터의 조합이기 때문에, 특정한 모델이 아닌 다양한 얼굴 모델에도 쉽게 적용되어 얼굴의 사실적인 움직임을 만들 수 있다.

또한 MPEG-4 SNHC에서 제안하고 있는 BAP(Body Animation Parameter)에 의한 고차원적인 인체 움직임에도 적용될 수 있다. BAP에서는 인체 각 관절의 움직임을 제어하지만 근육모델의 유기적인 조합, 예를 들면 F 제안된 얼굴 근육 모델의 웃음근육과의 조합을 통해 웃고있을 때의 걸음걸이를 애니메이션 시킬 수 있다.

3.2 얼굴 근육 벡터

얼굴 애니메이션에서 FAP에 의해 움직임이 결정되 얼굴 근육 벡터는 근육을 이루는 근섬유의 움직임으로 표시될 수 있다. 기존의 해부학적인 근육 모델에 기반한 방법에서는 근육 벡터를 실제 사람의 얼굴 근육에 맞추어 여러가지로 다르게 모델링하여 적용하고 있다[2, 3]. 하지만 본 논문에서는 FAP에 기반한 얼굴의 움직임을 구현하기 위하여, FAP에 따른 움직임을 표현하기 위해 배치된 각 제어 vertex의 움직임을 선형적인 근육의 움직임으로 표현하는 근육 벡터를 적용하고, 제어점의 움직임을 통해 움직이는 주변 피부를 사실적으로 애니메이션 시키기 위해 피부 움직임 함수를 사용하였다. 적용된 근육 벡터는 [2]에서 제안된 Linear muscle model을 FAP에 따른 움직임을 표현할 수 있도록 수정하여 적용하였다. 근육 모델은 수축과 팽창 정도와 방향을 받아 들어 정도를 조절함으로써 다양한 표현이 가능하도록 설계되었다.

적용된 근육벡터는 그림 (1)과 같은 콘(cone) 형태로 모델링 된다. 뼈에 붙어서 근섬유의 움직임을 지지하는 점 v_1 이 벡터의 시작점이 되고, 피부 조직 속에 있는 점 v_2 는 끝점이 된다. 근육의 영향력이 미치는 부분은 ϕ 로 정한다.

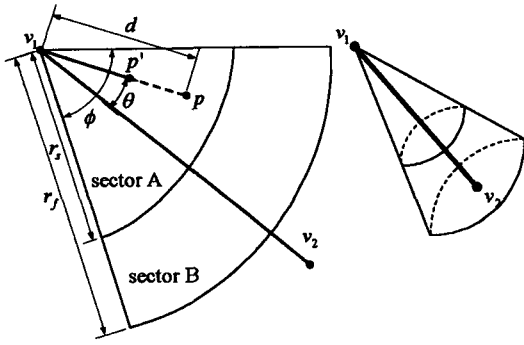


그림 1. 근육벡터의 정의

근육벡터의 시작점은 이동하지 않으며, 근육의 힘은 A와 B영역에 걸쳐서 영향을 준다. 근육의 힘이 미치는 영역 안에서 피부 조직의 한 vertex, p 의 움직임은 다음의 식과 같이 표현된다.

$$\Delta p = k \cdot r(d) \cdot a(\theta) \left(\frac{p \cdot v_1}{|p \cdot v_1|} \right) \quad (1)$$

위 식에서 k 는 근육의 탄성 상수를 나타내고, r 는 radial displacement parameter로써 식(2)에서 정의되며, a 는 angular displacement parameter로써 식(3)과 같이 정의된다. 따라서 p 가 Δp 만큼 이동한 p' 는 angular displacement parameter의 함수로 표현된다.

$$h(d) = \begin{cases} 1 & \text{for } d > r_j \\ 1 - \frac{d}{r_i} & \text{for } 0 \leq d \leq r_i \\ \frac{d - r_i}{r_j - r_i} & \text{for } r_i \leq d \leq r_j \end{cases} \quad (2)$$

$$r(d) = \cos\left(\frac{\pi}{2} \cdot h(d)\right) \quad (3)$$

$$a(\theta) = \cos\left(\frac{\pi}{2} \cdot \frac{2\theta}{\phi}\right)$$

3.3 피부 움직임 함수

세밀한 피부의 움직임을 표현하기 위해서는 얼굴 모델의 표면을 구성하고 있는 mesh의 자유로운 변형이 가능해야 한다. 본 논문에서는 세밀한 피부의 변화를 표현하기 위하여 energy functional의 적용을 통한 mesh의 변형방법을 적용하였다[4]. Energy functional은 근육 벡터에 의해 움직이게 되는 피부 영역의 수축과 팽창 정도를 제어한다. FAP의 변화에 따라 움직임이 정해지는 근육 모델의 제어점을 기준으로, 표정을 객체 표면의 영역을 피부 영역으로 생각하여 피부영역을 설정하고 피부영역의 움직임을 제압된 함수에 의해 animation 시킴으로써 사실적인 움직임을 만들어 낼 수 있다.

피부의 움직임은 질량이 제어점에 집중되어 있는 mass-spring의 변화에 따른 mesh들의 변형으로 모델링할 수 있다. Mass-spring의 움직임을 제어하는 energy functional은 근육이 움직이는 방향의 mesh edge에 적용

하였다. 적용된 energy function은 fixed rest-length spring의 움직임으로 모델링되었다. Fixed length의 조절을 통해 세분화된 피부 모델에 있어 각 위치에 따른 다양한 변화를 주기 위해서 Fixed rest-length spring을 적용하였다. 예를 들면 움직임의 영역이 넓은 입술 주변과 볼 주변은 작게, 영역이 좁은 이마부분은 크게 설정함으로써 기존의 spring model에서는 볼 수 없는 사실적인 움직임을 만들 수 있다.

실제 동작은 p_1 에서 p_2 까지의 mesh edge에 적용되는 energy term, $k_s E_s(p_1, p_2)$ 에 의해 결정된다. k_s 는 spring constant, E_s 는 식 (4)와 같다.

$$E_s(p_1, p_2) = \frac{1}{2} \left(\frac{|p_1 - p_2|}{|p_1^* - p_2^*|} - 1 \right)^2 \quad (4)$$

식 (4)의 p_1^*, p_2^* 는 두 vertex의 rest position을 나타낸다.

Energy function을 통한 mesh 변형을 통해 예상치 못한 피부의 뒤틀림이 발생할 수 있다. 이러한 뒤틀림을 방지하기 위해, mesh들의 집합으로 이루어진 큰 사각형에서 p_1 과 p_2 가, p_3 와 p_4 가 서로 사각형의 대각선을 따라 마주보고 있다고 하면 vertex에 대해 뒤틀림 방지 energy term, $k_d E_d(p_1, p_2, p_3, p_4)$ 을 적용한다. E_d 는 식 (5)와 같다.

$$E_d(p_1, p_2, p_3, p_4) = E_d(p_1, p_2) E_s(p_3, p_4) \quad (5)$$

또한 실제 사람의 피부와는 다르게 발생할 수 있는 bending 현상을 피하기 위해 연속된 vertex들의 열로 정의되는 virtual thread를 따라 anti-bending energy term, $k_b E_b(p_1, p_2, p_3)$ 을 더하였다. E_b 는 식 (6)과 같이 정의된다.

$$E_b(p_1, p_2, p_3) = \frac{1}{2} \left(\frac{|p_3 - p_2|}{|p_3^* - p_2^*|} - \frac{|p_2 - p_1|}{|p_2^* - p_1^*|} \right)^2 \quad (6)$$

각 energy term에 적용된 spring constant는 앞서 설명한 세분화된 근육 벡터에 따라 다르게 적용되어서 사실적인 얼굴 동작을 표현할 수 있도록 가능하게 하였다.

3.4 Direct Parameterized Approach (DPA)

위에서 언급한 근육 벡터는 모든 표정 변화에 적용될 수 없다. 예를 들어 눈꺼풀의 움직임을 이용하여 눈을 깜빡일 때나, 눈동자의 움직임, 동공의 팽창과 턱의 움직임 등은 근육 벡터로는 표현이 어렵다. 따라서 본 논문에서는 변수의 변화에 따라 직접 그 움직임을 전달하는 Direct Parameterized Approach(DPA)를 적용한다. 이는 해부학적인 움직임이 아닌 단순히 원하는 얼굴 모델의 특정한 변화를 제공한다.

(1) 눈의 움직임 표현

눈의 깜빡임은 구형이고 비선형적인 동작이기 때문에 수학적인 모델링이 쉽지 않다. 따라서 이 구형 동작을 표현하기 위해, 구형 매핑 방법[2]을 적용하였다.

눈동자의 움직임은 눈의 깜빡임과 마찬가지로 근육

벡터를 사용하여 표현할 수 없다. 눈동자의 움직임을 표현하기 위해 DPA를 이용하여 변수의 변화에 따라 동공의 팽창과 눈동자의 움직임을 직접 렌더링 한다. 동공의 움직임을 표현하기 위해서는 반지름의 변화를 전달해야 할 것이다. 식 (7)과 그림 (3)는 눈동자를 표현하기 위한 관계식을 보여준다.

$$\begin{aligned} \phi_i &= \sin^{-1}(R_i / R_e), \quad \phi_p = \sin^{-1}(R_p / R_e) \\ n &= \frac{\pi}{2} - \phi_i \\ R_w &= R_e \sin\left(\phi_i + \frac{n}{2}\right), \quad X_w = R_e \cos\left(\phi_i + \frac{n}{2}\right) \\ X_i &= \sqrt{R_e^2 + R_i^2}, \quad X_p = \sqrt{R_e^2 + R_p^2} \end{aligned} \quad (7)$$

(2) 턱의 움직임 표현

턱의 움직임과 같은 변화는 얼굴 근육 벡터로는 그 움직임을 표현하기에는 무리가 따른다. 왜냐하면 턱의 움직임에 따라 얼굴 하부구조가 전체적으로 큰 변화를 갖기 때문이다. 턱의 움직임을 표현하기 위해서는 모든 vertex들이 턱 뒤쪽을 기준으로 같은 방향으로 회전을 하도록 해야한다. 이를 구현하기 위해 턱의 상하 및 좌우 이동, 턱 끝을 내리는 등의 표현을 DPA를 이용하여 처리하였다.

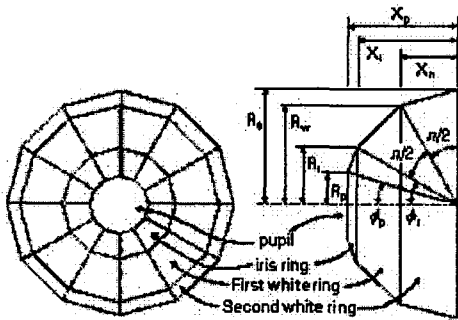


그림. 2. 눈을 움직이는 근육 모델의 구성

4. 실험결과

실험에 사용된 얼굴 객체 모델은 1990년 Compaq Cambridge Research Laboratory에서 Keith Waters에 의해 개발된 모델[2]을 바탕으로 하여 MPEG-4에서 실험모델로 사용된 Oscar를 제안된 방법으로 얼굴의 움직임을 표현하기 위해 수정한 모델을 사용하였다. 모델의 총 vertex의 수는 512개이며 렌더링 엔진으로는 OpenGL을 사용하였다. Texture mapping을 적용하지 않은 상태에서 smooth shading만을 적용하였다.

그림 (3)은 하이레벨의 FAP 변수를 입력으로 받아 렌더링 시킨 결과로써 기존의 FAP Player 보다 사실적인 결과를 얻을 수 있었다. 고차원적인 얼굴 표정은 근육 모델간의 조합에 의해 표현되기 때문에 이를 응용하여 보다 복잡한 표정을 쉽게 만들어 낼 수 있다.

제안된 얼굴 애니메이션을 통해 특히 세밀한 얼굴의 움직임을 효과적으로 나타낼 수 있었다. 그림 (5)에서 볼 수 있듯이 말하고 있는 얼굴 객체를 표현할 때 제안

된 방법을 통해 입술의 세밀한 움직임을 재현할 수 있고, 근육 모델의 조합을 통해 입술에 따라 움직이는 볼 근육의 움직임도 효과적으로 재현 할 수 있었다.

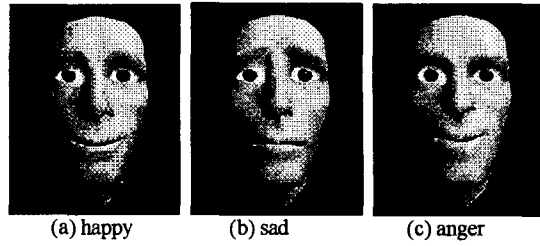


그림. 3. 하이레벨 FAP 표현

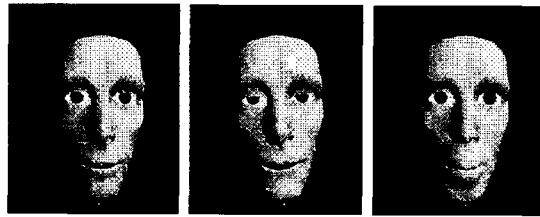


그림. 4. Lip Synchronization의 구현

5. 결 론

본 논문에서는 MPEG-4에서 제안한 FAP 변수에 의한 사실적인 얼굴 애니메이션 시스템을 제안하였다. 제안된 방법을 통해 가상환경에서 많은 정보를 전달하는 얼굴을 사실적으로 제어함으로써 사용자에게 친밀감을 줄 수 있는 시스템을 구축할 수 있었다. 또한 특정 모델에 제한되지 않는 모델을 제안함으로써 다양한 환경에 적용될 수 있다.

앞으로 BAP와의 조합을 통한 전체 인체 애니메이션 시스템을 구축한다면 가상환경에서 보다 다양한 용도로 사용될 수 있을 것이다.

참고문헌

- [1] MPEG-4 SNHC, 'Information Technology-Generic Coding of Audio-Visual Object Part 2: Visual', ISO/IEC JTC1/SC29/WG11 N2688, Seoul, Oct 1999.
- [2] Frederic I. Parke and Keith Waters. *Computer Facial Animation*. A K Peters, Wellesley, Massachusetts, 1996.
- [3] Torbjorn Soderman. *Face the World*. PhD thesis, University of Umea, Nov., 1998.
- [4] Pixar Animation Studio, 'Subdivision Surfaces in Character Animation', 1998 Conf. Proc. of ACM SIGGRAPH, pp85-94, 1998.
- [5] Gabriel Antunes Abrantes. *ISTface: MPEG-4 player*. Instituto Superior Tecnico, Portugal, 1998.
- [6] MIRALab. *Miraplayer: MPEG-4 player*. MIRALab, Switzerland, 1997.
- [7] 이주상, 유지상, MPEG-4 SNHC 기반 얼굴 객체의 구현, Telecommunication Review 제 8 권 3 호 1998. 5-6 호