

한글 입술 움직임과 얼굴 표정 동기화를 위한 얼굴 애니메이션 편집기

송 미영, 조 형제
동국대학교 컴퓨터공학과

Face Animation Editor for the Korean Lip_Sync and Face Expression

Miyoung Song, Hyungje Cho
Dept. of Computer Engineering Dongguk University

요 약

본 논문은 한글 단어에 따른 한글 발음에 적합한 입술의 움직임을 자동 생성하며 또한 단어에 적절한 얼굴 표정을 생성할 수 있는 입술 움직임과 얼굴 표정을 동기화하는 3차원 얼굴 애니메이션 편집기를 구축하였다.

얼굴 애니메이션 편집기에서 얼굴 표정은 근육 기반 모델 방법으로 정의된 각 얼굴 부위별 근육에 따라 가중치를 조절하여 생성하며 입술 움직임은 텍스트 구동 방법으로 음소에 따른 정의된 입모양 연속적으로 표현하여 동작한다. 또한 이렇게 생성된 얼굴 표정을 저장관리한다.

따라서 3차원 얼굴 애니메이션 편집기는 6가지의 기본 얼굴 표정을 자동적으로 생성할 수 있으며 또한 입력 단어에 적합하도록 각 얼굴 부위별의 근육 움직임을 편집할 수 있다.

이렇게 생성된 얼굴 표정들은 데이터베이스에 저장관리할 수 있으며 컴퓨터 대화시 자동적으로 입력 단어에 적합한 입술의 움직임과 얼굴 표정을 동기화하여 자연스러운 3차원 얼굴 애니메이션을 표현할 수 있다.

1. 서 론

사람의 얼굴은 상호간의 자연스러운 대화를 위한 가장 중요한 보조 수단 중 하나로서 상대방의 얼굴 표정으로부터 감정을 전달받기도 하고, 대화시에는 음성뿐 아니라 상대의 입술 모양의 변화도 대화 내용을 이해하는데 중요한 정보가 된다.

따라서 사람들은 서로의 얼굴을 마주보며 생각을 교환하기 때문에, 컴퓨터를 사용할 때는 기계라는 거부감을 느끼게 됨으로 이러한 측면에서 사람과 컴퓨터와의 인터페이스는 사람을 닮은 즉, 음성, 표정, 입술 움직임, 몸짓을 자연스럽게 할 수 있어야 한다.

얼굴 애니메이션은 Parke의 초창기 연구를 시작으로 지속되고 있다. 기본적으로 얼굴 애니메이션에 관한 연구는 3D 메쉬 디자인, 3D메쉬 디지털라이징, 3D 메쉬 애니메이션의 3영역으로 구분된다.

이중 3D메쉬 애니메이션은 메쉬노드들을 움직여서 얼굴 표정과 입술의 움직임등을 표현하는 방법이다.

이에 관한 연구들은 파라미터화 모델(parameterized model), 제어점 모델(control-point model), 운동학적 근육 모델(kinematic muscle mode), 동적 근육 모델(dynamic muscle mode), 스플라인 모델등이 대표적이라 할 수 있다[7].

또한 입술 움직임 동기화 연구는 음성의 출처에 따라 텍스트구동(text-driven) 방식과 음성구동(speech-driven) 방식의 두가지로 분류된다.

텍스트 구동 방식은 TTS(Text-to-Speech)와 비슷한 방법으로 텍스트로부터 음소(phoneme)정보를 추출하고 음소를 기반으로 하는 입모양 DB로부터 연속된 화면을 생성하는 방식이다.

음성 구동 방식은 녹음된 음성 신호로부터 동기화된(synchronized) 애니메이션을 생성하는데 필요한 일

굴제어(control)정보를 생성한다.

이밖에도 이들 두방식을 혼합하여 적용한 하이브리드 방식이 있다.

이 논문의 3차원 얼굴 애니메이션은 파라미터 모델을 기반으로 얼굴 표정을 생성하며 또한 입술 움직임의 표현은 텍스트 구동방법을 사용하여 자연스러운 입술 움직임과 얼굴 표정을 동시화 할 수 있는 3차원 얼굴 애니메이션 편집기를 구현하였다.

이 논문의 구성은 2장에서는 얼굴 애니메이션 편집기의 시스템 구성도에 대해 설명하고 3장은 얼굴 표정 편집기 모듈, 4장은 한글 단어 처리 모듈, 그리고 5장은 입술 모양 움직임 모듈들을 검토하고 6장에서는 실험 결과 및 결론을 맺는다.

2. 3차원 얼굴 애니메이션 편집기 시스템

이 논문은 한글 단어의 음소에 따른 입술 움직임 및 얼굴 표정을 편집하여 서로 동기화하기 위한 3차원 얼굴 애니메이션 편집기이다.

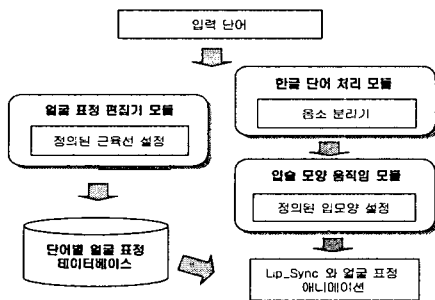


그림 1 시스템 구조

얼굴 애니메이션 편집기는 크게 얼굴 표정 편집기 모듈과 한글 단어 처리 모듈, 입술 모양 움직임 모듈과 입력 단어에 대한 적합하게 생성한 얼굴 표정을 저장하는 데이터베이스등으로 구성되어 있다.

얼굴 표정 편집기 모듈은 입력되는 한글 단어에 따라 적합한 각 얼굴 부위별을 편집하여 얼굴 표정을 생성하는 부분이다.

한글 단어 처리 모듈은 입술 움직임 모양을 설정하기 위해 입력되는 한글 단어의 초성, 중성, 종성등의 음소들로 분석하는 부분이다.

입술 모양 움직임 모듈은 분석된 한글단어의 음소들에 따라 입술모양을 연속으로 표현할 수 있도록 처리하는 부분이다.

단어별 얼굴 표정 데이터베이스 모듈은 입력 단어에 적합하게 생성한 얼굴 표정에 대한 얼굴 각 부위별 가중치를 저장 관리하는 부분이다.

따라서 얼굴 애니메이션 편집기는 입력 단어에 대한 음소를 분석하여 입술 움직임을 표현하고 이에 따른 얼굴 표정을 생성하여 서로 동기화함으로써 자연스러운 얼굴 애니메이션을 표현할 수 있다.

3. 얼굴 표정 편집기 모듈

얼굴 애니메이션 편집기 시스템은 일정한 감정을 나타내는 얼굴 표정은 사람마다 얼굴의 각 부위의 변화하는 정도가 거의 비슷하다는 사실에 기반하여 가장 일반적인 얼굴 표정 6가지(기쁨, 슬픔, 분노, 공포, 놀람, 혐오)를 미리 정의하고, 각 얼굴 부위별을 편집하여 자연스러운 얼굴 표정을 생성할 수 있다.

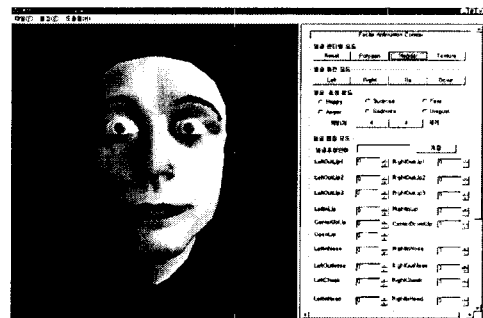


그림 2 편집기 화면 인터페이스

얼굴 애니메이션 편집기의 화면은 그림2와 같이 얼굴 렌더링 모드, 얼굴 회전 모드, 얼굴 기본 표정 모드, 얼굴 표정 편집 모드 그리고 단어 입력 모드등으로 구성되어있다.

얼굴 렌더링 모드는 얼굴을 초기화하거나 다각형 모델, 또는 렌더링 모델로 선택할 수 있도록 구성되어 있으며 얼굴 회전 모드는 얼굴 모델을 상·하·좌·우로 움직여 원하는 시점에서 모델을 볼 수 있다.

얼굴 기본 표정 모드는 일반적인 6가지 표정인 기쁨, 슬픔, 분노, 공포, 놀람, 혐오등의 얼굴 표정을 선택하고 이 해당 표정의 가중치를 변화시켜 얼굴 표정을 표현할 수 있다.

얼굴 표정 편집 모드는 입력되는 한글 단어에 적합한 얼굴 표정을 생성할 수 있도록 각 얼굴 부위별 변화 가중치를 편집할 수 있다.

얼굴의 각 부위별의 기본 근육선은 그림3과 같이 설정하고 근육선의 변화정도를 편집하여 얼굴 표정등

을 생성한다.

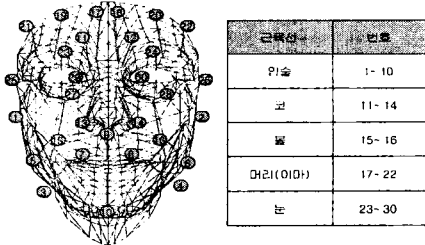


그림 3 얼굴 근육선

따라서 자연스러운 얼굴 표정을 생성하기 위해 다음의 식(1)에 따라 근육들을 움직인다.

근육의 움직임을 위해 필요한 정보들은 미리 정의된 얼굴 근육선의 시작좌표점(Mu), 끝좌표점(Mu0), 근육움직임의 거리범위(Md1 ~ Md0), 근육움직임의 각도(Mθ) 그리고 모든 얼굴의 좌표점(Pv)등이다.

$$if (Md0 < (Pv - Mu0) < Md1) ,$$

$$newPv' = Pv \times CosV \times Cos\left(\frac{(Pv - Mu0) - Md0}{Md1 - Md0}\right) \times \frac{\pi}{2}$$

$$else newPv' = Pv \times CosV$$

-----식(1)

여기서, α는 미리정의된 근육선의 끝점을 기준으로 시작점과 임의의 얼굴 좌표점이 이루는 각도를 의미하며 $CosV = W \times (1.0 - \frac{Cos \alpha}{Cos(M\theta)})$ 이고, W는 각 근육의 가중치를 의미한다.

위 식(1)은 Cosα가 Cos(Mθ)보다 크고 |Pv - Mu0|보다 |Mu1 - Mu0|이 큰 경우에만 적용하여 자연스러운 얼굴 표정을 생성한다.

4. 한글 단어 처리 모듈

입술 모양의 설정하기 위해 우선 입력되는 한글 단어에 대해 초성, 중성, 종성등의 음소로 분류하여 이에 해당되는 입술 모양을 연속적으로 표현하게 한다.

한글 단어 처리 모듈에서는 크게 받침이 없는 경우와 받침이 있는 경우로 나눌수 있다. 받침이 없는 경우는 초성과 중성으로 자음과 모음으로 이루어져 있으며 받침이 있는 경우는 초성, 중성에 자음이 덧붙여진 종성으로 이루어져 있다.

(1) 초성(자음)

한글의 자음은 조음자리에 따라 표1과 같이 분류한다[5].

음절에 따른 입모양패턴은 대부분 모음에 의존하지만 입술소리(/ㅂ, ㅃ, ㅍ, ㅍ/)는 입술 위치에서 소리가 나기 때문에 시각적으로 입술이 다물어지게 한다.

그러나 다른 조음자리의 자음소리는 입술과는 상관없는 자리에서 소리가 나기 때문에 시각적으로는 차이가 없다.

조음자리		자음
입술소리		ㅂ, ㅃ, ㅍ, ㅍ
기타소리	치(조)음	ㄷ, ㅌ, ㄸ, ㅊ, ㅍ, ㄱ, ㅋ
	경구개음	ㅈ, ㅊ, ㅉ
	연구개음	ㄱ, ㅋ, ㆁ, ㅇ
	성문음	ㅎ

표 1 한글 자음 분류

따라서 입술소리(/ㅂ, ㅃ, ㅍ, ㅍ/)와 조합되는 경우에는 발음초기의 입모양이 다르므로 음절에서 자음의 영향을 받아 표현되는 입모양은 입술소리와 그의 기타 소리에 의한 입모양으로 분류한다[6].

(2) 중성(모음)

한글 모음은 기본모음과 조합형모음으로 분류한다. 기본모음인 /ㅏ, ㅑ, ㅓ, ㅕ, ㅗ, ㅛ, ㅜ, ㅠ/의 입 모양만을 정의하고 /ㅜ, ㅠ, ㅡ, ㅣ/의 4개 모음은 각각 /ㅏ, ㅑ, ㅓ, ㅕ/의 입모양 패턴과 같게하고 발음시간을 지연한다.

그리고 조합모음인 /ㅘ, ㅙ, ㅚ, ㅜ, ㅠ, ㅡ, ㅣ/는 기본형 모음의 조합으로 하여 처리한다.

분류		모음
기본형 모음	기본	ㅏ, ㅑ, ㅓ, ㅕ, ㅗ, ㅛ, ㅜ, ㅠ, ㅡ, ㅣ
	기타 모음 (ㅜ, ㅠ, ㅡ, ㅣ)	ㅜ, ㅠ, ㅡ, ㅣ
조합형 모음	ㅘ, ㅙ, ㅚ, ㅜ, ㅠ, ㅡ, ㅣ	기본형 모음의 조합

표 2 한글 모음 분류

(3) 받침(중성)

받침으로 오는 자음들 또한 모음에 의해 입모양은 거의 영향을 받으나 최종 생성되는 입모양은 표1과 같이 이미 분류한 입술소리와 그의 기타 소리의 입모양으로 다르다.

따라서 입술 소리를 제외한 자음들은 모음에 영향을 거의 미치지 못하며 모음에 의해서 크게 지배받기

때문에 모두 비슷한 모습을 보인다.

6. 입술 모양 움직임 처리 모듈

음소에 따른 입술 움직임을 위한 근육선은 그림3에서의 입 부위의 입술 근육선(1~10)에 가중치를 설정하여 동작한다. 입술 모양은 표3에서 정의된 모음과 자음에 대해서 그림 4와 같이 정의한다.

음소분류		입모양 근육
자음	입술소리 (ㄴ, ㅍ, ㅁ, ㅂ)	기본 입술 애니메이션
모음	ㅏ, ㅑ	9,10, Open_jaw(Large)
	ㅣ	5, 6, Open_jaw(Small)
	ㅓ	7,8,9,10, Open_jaw(Small)*2
	ㅕ, ㅖ	10, Open_jaw(Middle)
	ㅗ	7,8, Open_jaw(Small)*2
	ㅡ	5, 6, 9, Open_jaw(Small)
	ㅜ	9, 10, Open_jaw(Middle)

표 3 입모양 분류를 위한 음소

따라서 입력되는 단어의 음소를 분석하여 각 음소에 해당되는 입술 모양을 연속적으로 표현하여 입술 움직임을 동작한다.

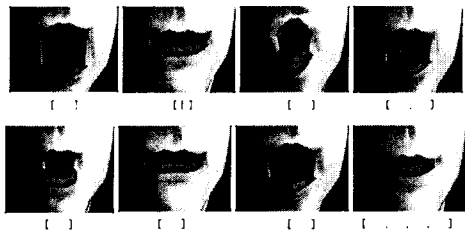


그림 4 음소발음별 입모양

7. 실험결과 및 결론

이 논문은 사용자간의 컴퓨터상에서 대화를 위한 친근감있는 인터페이스를 제공하기 위한 연구로 입술 움직임 및 얼굴의 표정생성을 위한 3차원 얼굴 애니메이션 편집기에 대해서 기술하였다.

시스템 환경은 IBM 호환 PentiumII-333 시스템을 기본으로 하여 OpenGL 라이브러리와 Microsoft Visual C++ 6.0으로 구현하였다.

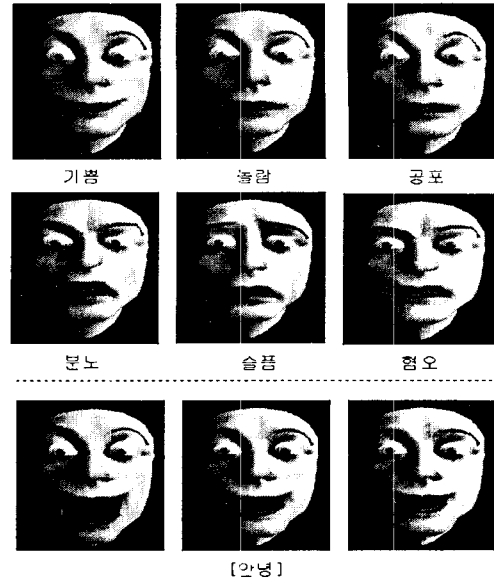


그림 5 기본 얼굴 표정과 “안녕”의 발음 움직임

그림5는 일반적인 6가지 얼굴 표정인 기쁨, 슬픔, 분노, 공포, 놀람, 혐오에 대한 가중치를 다르게 하여 표현한 결과이고 “안녕”이라는 단어에 대한 얼굴의 표정과 입술 움직임을 동기화하여 표현한 결과이다.

앞으로의 연구에서는 좀더 섬세하고 자연스러운 입술 모양의 합성이 연구되어야 하며 이러한 3차원 얼굴 애니메이션 편집기를 통해 다양한 표정과 입술 움직임을 동기화는 영화 및 오락, 교육, 가상현실 등 3차원 얼굴 애니메이션을 표현하는 여러 분야에 응용할 수 있을 것이다.

[참고문헌]

- [1] Parke, F. I., "Parametrized models for facial animation.", IEEE Computer Graphics, 2(9), pp.61-68, 1982
- [2] Frederic I.Parke, Keith Waters, " Computer Facial Animation" A K.Peters, 1996
- [3] Kong K.S., Kim C.H., "Modeing Coarticulation in Korean Visual Speech", Prceedings if the first CGIM conference,PP.207-210, 1998
- [4] J.A.Provine, L.T.Bruton, "Lip Synchronization in 3-D Model Based Coding dor Video-conferencing" ISCAS'95, pp.453-456
- [5] 공관식, 김창현, "입술 애니메이션을 위한 한글 발음의 동시조음 모델", 정보과학회논문지(A)'99 제 26 권 제 9호,
- [6] 양현승, 한태우, 이주호, " 실감있는 얼굴 표정 생성을 위한 3차원 얼굴 모델링과 애니메이션",Telecommunications Review 제 8 권 3호 1998.5~6
- [7] 정보과학회지 "CG 애니메이션" 제 17 권 제 2호 통권 제 117호 1999.2