

휴먼인터페이스를 위한 한글음절의 입모양합성

正會員 李 龍 東* 正會員 崔 昌 石** 正會員 崔 甲 石**

Lip Shape Synthesis of the Korean Syllable for Human Interface

Yong Dong Lee*, Chang Seok Choi**, Kap Seok Choi** *Regular Members*

요 약

얼굴을 마주보며 인간끼리 대화하는 것처럼 인간과 자연스럽게 대화할 수 있는 휴먼인터페이스를 실현하기 위해서는 음성의 합성과 얼굴영상의 합성이 필요하다. 본 논문은 얼굴영상의 합성을 대상으로 한다. 얼굴영상의 합성에서는 표정변화와 입모양의 변화를 3차원적으로 실현하기 위하여 얼굴의 3차원 형상모델을 이용한다. 얼굴의 3차원 모델을 얼굴 근육의 움직임에 따라 변형하므로써 다양한 얼굴표정과 음절에 어울리는 입모양을 합성한다. 우리말에서 자모의 결합으로 조합가능한 음절은 14,364자에 이른다. 이 음절에 대한 입모양의 대부분은 모음에 따라 형성되고, 일부가 자음에 따라 달라진다. 그러므로, 음절에 어울리는 입모양의 변형규칙을 정하기 위해, 이들을 모두 조사하여 모든 음절을 대표할 수 있는 입모양패턴을 모음과 자음에 따라 분류한다. 그 결과, 자음에 영향을 받는 2개의 패턴과 모음에 의한 8개의 패턴, 총 10개의 패턴으로 입모양을 분류할 수 있었다. 나아가서, 분류된 입모양패턴의 합성규칙을 얼굴근육의 움직임을 고려하여 정한다. 이와같이 분류된 10개의 입모양패턴으로 모든 음절에 대한 입모양을 합성할 수 있고, 얼굴근육의 움직임을 이용하므로써 다양한 표정을 지으면서 말하는 자연스런 얼굴영상을 합성할 수 있었다.

ABSTRACT

Synthesizing speech and facial images is necessary for human interface that man and machine converse naturally as human do. The target of this paper is synthesizing the facial images. In synthesis of the facial images a three-dimensional(3-D) shape model of the face is used for realizing the facial expression variations and the lip shape variations. The various facial expressions and lip shapes harmonized with the syllables are synthesized by deforming the three-dimensional model on the basis of the facial muscular actions. Combinations with the consonants and the vowels make 14,364 syllables. The vowels dominate most lip shapes but the consonants do a part of them. For

*明知大學校 電子工學科
Dept. of Electronics, Myong Ji Univ.

**明知大學校 情報通信工學科
Dept. of Information, Myong Ji Univ.
論文番號 : 93172
接受日字 : 1993年 9月 3日

determining the lip shapes, this paper investigates all the syllables and classifies the lip shapes pattern according to the vowels and the consonants. As the result, the lip shapes are classified into 8 patterns for the vowels and 2 patterns for the consonants. In advance, the paper determines the synthesis rules for the classified lip shape patterns. This method permits us to obtain the natural facial images with the various facial expressions and lip shape patterns.

1. 서론

현재, 눈부신 과학기술의 발달에 따라 기계는 성능이 다양화, 고기능화되어 가고있다. 그러나, 이 고도의 기계들은 너무나 복잡하여 전문지식이 없는 대부분의 일반이용자는 제품이나 기계의 고도의 능력을 심분 활용할 수 없는 경우가 많아, 방대한 인원과 장비를 투입하여 애써 개발한 기술이 제역활을 발휘하지 못하는 결과를 초래하게 된다. 이것은 이제까지의 기술이 기계의 성능향상 중심으로는 발전되어 왔지만, 이용자가 사용하기 쉽게 기계의 높은 성능을 심분 발휘할 수 있도록 하기위해 인간과 기계사이의 대화를 원활히 하는 맨머신인터페이스(man-machine interface) 기술이 상대적으로 낙후 되었음을 의미한다고 생각한다.

이러한 관점으로부터, 인간끼리의 대화와 같은 감각으로 기계가 대화하기 위한 맨머신 인터페이스를 검토하고 있다. 현재, 인간과 기계가 대화하는 맨머신 인터페이스의 형태는 기계가 처리하기 쉬운 문자와 숫자를 이용한 기계 중심의 인터페이스이다. 그리고 최근에 선을 보이고 있는 도형이나 그래픽을 이용한 GUI(Graphical User Interface)는 인간의 감각에 좀더 부합된 형태라고 생각되고 있다. 그러나, 맨머신 인터페이스의 궁극적인 형태는 숫자, 문자, 도형뿐만 아니라 인간이 구사하기 쉬운 언어, 얼굴표정, 제스처 등을 종합적으로 사용하는 인간중심의 인터페이스, 즉 휴먼인터페이스(human interface)가 되어야 할 것으로 생각된다. 이러한 형태의 휴먼인터페이스 기술을 이용하면, 사용자의 물음에 대한 기계측의 응답을 보다 인간의 감각에 맞게 표현할 수 있을 것이다. 즉, 컴퓨터 모니터의 한 모서리에 사용자가 좋아하는 얼굴이 나타나 우리에게 상냥한 표정을 지으며 말을 걸어와 곤란에 처한 이용자에게 도움을 줄 수 있을 것이다.

이와같은 휴먼인터페이스를 실현하기 위한 핵심기술은, 기계의 응답을 나타내기 위한 얼굴영상합성기술과 음성합성기술, 그리고 인간의 감정 및 명령을

이해하기 위한 얼굴표정인식기술과 음성인식기술이 필수 기술이 된다. 최근에 얼굴영상의 합성기술, 음성합성기술은 상당한 발전을 보이고 있고, 음성인식과 얼굴영상인식기술도 중요성이 점점 인식되고 있다.

음성과 얼굴표정을 이용한 휴먼인터페이스의 구상은 Parke나 Lippman등에 의해 산발적으로 제안되었다[1-3]. 그러나 얼굴의 합성법이 단순한 그래픽기술에 의존하였기 때문에, 현실감이 떨어진 만화와 같은 얼굴영상이 얻어져, 현실감이 요구되는 휴먼인터페이스에 대한 체계적인 연구로는 연결되지 못했다. 그 이후, 일본 동경대학의 原島 博교수를 중심으로한 연구그룹에 의해 2차원 얼굴사진을 이용하여 3차원 얼굴영상을 분석하고 합성하는 방법이 제안되어 이 분야의 연구가 활발해 졌다[4-9]. 이 방법은 얼굴사진을 기본으로 하여 현실감이 뛰어나 얼굴영상을 합성할 수 있기 때문에, 다방면에 응용이 가능할 것으로 보여 크게 주목을 받고 있다. 同 그룹의 森島는 이러한 연구의 연장선상에서 일본어에 대한 입모양합성법을 검토하였다[6]. 이 연구에서는 일본어의 모음과 자음에 대해서 자소단위로 입모양 변형규칙을 정하고 있다. 그러나 얼굴근육의 움직임을 고려하지 않고 입모양만에 주목하여 입모양변형규칙에 정하고 있기 때문에 합성된 얼굴영상은 부자연스러워 진다. 또한, 입모양변형뿐만 아니라 감정을 나타내기 위해 표정을 동시에 변화시키는 경우, 양자간에 변형규칙의 친화성이 없기때문에 자연스런 얼굴영상을 얻기가 곤란하다. 한편, 同그룹이었던 필자의 한사람은 얼굴근육의 움직임을 고려한 얼굴표정의 분석과 합성법을 제안하여, 자연스럽게 다양한 얼굴표정을 합성할 수 있음과 얼굴 동영상을 분석하여 두부의 3차원운동 패턴, 얼굴표정변화패턴을 분석하는 방법을 제시했다[7-9].

이와같은 연구가 이루어졌음에도 불구하고 우리말을 이용한 휴먼인터페이스를 실현하기 위한 충분한 연구가 이루어 졌다고는 보기가 곤란하다. 왜냐하면 일본어의 50음에 비해서 한글을 조합하는 경우 14,364자에 이르며, 현재 실용되는 한글한자 표준코드(KSC 5601)만도 2350음을 규정하고 있기 때문에, 한글의

입모양에 대한 새로운 검토가 필요하다. 또한, 입모양도 얼굴근육의 움직임에 따라 변화하는 것이므로 자연스런 얼굴 영상을 합성하기 위해서는 얼굴근육을 고려한 입모양의 합성법이 개발되어야 한다고 생각된다.

본 논문에서는 우리말을 이용한 휴먼인터페이스를 실현하기 위해서 얼굴근육의 움직임을 고려한 입모양합성법을 제안한다[12-13]. 먼저, 우리말의 모든 음절에 대한 입모양을 합성하는데 있어서 기본이 되는 입모양을 분류한다. 분류된 입모양에 대한 합성규칙을 정하고, 합성규칙에 따라 얼굴영상을 합성하고자 한다. 또한, 이 방법은 표정변화의 합성규칙과도 친화성이 좋기때문에 표정을 지으면서 입모양이 변하는 얼굴영상의 합성이 용이할 것으로 생각된다. 이하, 2장에서는 얼굴영상을 3차원적으로 합성하기 위한 얼굴의 3차원 모델의 구성법을 소개하고, 3장에서는 얼굴근육의 움직임에 고려하여 얼굴의 3차원 형상 모델을 변형하여 얼굴영상을 합성하는 방법을 요약한다. 4장에서는 한글의 모든 음절에 따른 대표적인 입모양패턴을 분류하고, 입모양패턴에 대한 합성규칙을 정한다. 그리고 다양한 표정을 지닌 입모양합성도 시도한다. 마지막으로 결론을 정리한다.

II. 얼굴의 3차원 모델[5]

얼굴영상은 3차원 얼굴의 2차원 화면상에 투영이라는 관점에서 볼때, 얼굴영상의 합성은 3차원적으로 이루어져야 한다. 따라서, 여기에서는 얼굴형상을 점과 선으로 구성된 삼각형으로 근사한 와이어프레임 모델(wire frame model)을 준비하고, 이것을 일반적인 얼굴의 3차원 형상모델로 하였다. 이것을 그림 1에 나타낸다. 이 3차원 형상모델을 2차원 대상인물(그림 2)의 정면상에 정합한다. 이때, 눈썹, 눈, 코, 입등과 같은 특징적인 부위는 3차원 아핀변환(affine transform)과 윤곽조정을 통하여 보다 정밀하게 정합한다. 이렇게 하여 얻어진 와이어프레임모델을 개인얼굴의 3차원 형상모델로 한다. 그리고 정합된 형상모델의 표면에 얼굴영상의 휘도정보를 투영하므로써 얼굴의 3차원모델이 생성된다.

얼굴의 3차원모델의 구성은 2차원 얼굴영상에서 3차원 구조로 복원하는 작업이므로, 모델이 일단 구성되면 표정변화를 용이하게 실현할 수 있을 뿐 아니라, 음절에 동기되는 입부분의 합성도 가능하다.

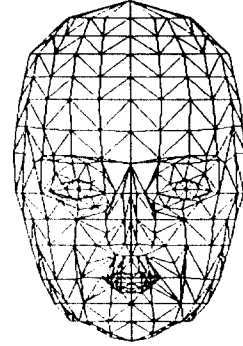


그림 1. 얼굴의 일반적인 3차원 형상모델.
Fig 1. A general three dimensional shape model of the face.



그림 2. 대상인물의 정면상.
Fig 2. A full face image.

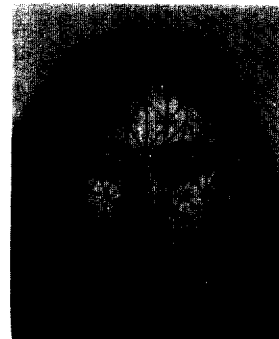


그림 3. 정합된 얼굴의 3차원 모델.
Fig 3. Ajusted the three dimensional mdoel of the face.

III. 얼굴표정의 합성

얼굴표정의 합성은 얼굴형상모델의 변형과 변형된 형상모델에 얼굴영상의 휘도치를 부여하므로써 이루어진다. 필자등은 얼굴형상모델의 변형에 있어서는 얼굴근육의 움직임은 고려하고, 휘도치부여에 있어서는 텍스처매핑[10]을 이용하여 다양한 얼굴표정을 고품질로 자연스럽게 합성하는 방법을 제시했다[7-8]. 얼굴근육의 움직임을 고려하기 위해서 FACS(facial action coding system)[11]을 참고로 했다. FACS에 의하면 얼굴근육의 움직임은 "Action Unit(AU)"라는 44개의 기본동작으로 분해되고, 여러가지 표정은 AU조합으로 기술될 수 있다고 한다. 필자등은 AU에 대한 얼굴형상모델의 변형규칙을 컴퓨터내에서 정하였다. 변형규칙이 정해진 AU의 일람은 표 1과 같다.

표 1. AU목록.
Table 1. AU List.

AU no.	AU명	AU no.	AU명
1	눈썹 내측을 올린다.	20	입술양단을 옆으로 끈다.
2	눈썹 외측을 올린다.	23	입술을 강하게 다문다.
4	눈썹을 내린다.	24	입술을 상하로 누른다.
5	윗눈꺼풀을 올린다.	25	턱을 내리지 않고 아래입술을 내린다.
6	뺨을 올린다.	26	턱을 내리면서 아래입술을 내린다.
7	눈꺼풀을 긴장시킨다.	27	입을 크게 벌린다.
8	입술을 서로 접근시킨다.	28	입술을 뺨라들인다.
9	코에 주름을 잡는다.	29	아래턱을 내린다.
10	윗입술을 올린다.	30	턱을 좌우로 이동시킨다.
11	鼻脣溝를 깊게 한다.	32	입술을 깨문다.
12	입술양단을 끌어 올린다.	35	불을 빨아드린다.
13	입술양단을 예리하게 끌어올린다.	41	윗눈꺼풀을 내린다.
14	보조개를 만든다.	42	눈을 가늘게 뜬다.
15	입술양단을 내린다.	43	눈을 감는다.
16	아랫입술을 내린다.	44	눈을 작게 뜬다.
17	턱을 올린다.	45	눈을 깜빡인다.
18	입술을 좁힌다.	46	윙크
			안구의 회전

음절에 따른 입모양의 변화도 표정변화의 연장으로 보면, AU의 변형규칙이 입모양의 변형규칙에도 그대로 적용가능하다. 다음 장에서는 AU를 이용하여 음절에 따른 입모양 변형규칙을 정하고자 한다.

또한, 입모양변환규칙에 있어서 AU를 이용하면, 얼굴에서 감정을 나타내는 표정변화와 함께 입모양을 변형하는 것도 용이하게 실현가능하리라 생각된다. 실제 대화에 있어서 입모양만이 변화하는 것이 아니라 감정표현은 대부분 표정을 이용하고 있기 때문에 표정변화와 입모양변형을 동시에 자연스럽게 합성할 수 있는 방법은 대단히 중요한 의미를 갖는다고 생각된다.

IV. 음절에 따른 입모양분류

한글은 조합가능한 음절이 총 14,364개에 이르고 있으나, 일반적으로 모든 음절에 대한 입모양을 정할 필요가 없다고 생각한다. 그것은 음절에는 같은 입모양을 나타내는 同型異音, 즉 음절의 음이 달라도 같은 형태의 입모양패턴이 많이 존재하기 때문이다. 또한, 음절 발생시 입모양은 복수의 입모양의 연결이다. 따라서, 모든 음절에 대해서 입모양을 정할 것이 아니라 대표적인 음절에 대해서만 입모양을 정하여, 이 입모양의 결합으로 모든 음절의 입모양을 나타내는 것을 생각한다. 이렇게 하므로써, 비교적 적은 수의 입모양의 변형규칙을 정하여도 얼굴동영상의 합성이 가능하다고 생각되기 때문이다. 본 장에서는 한글의 구조와 모든 음절에서 표현될 수 있는 입모양을 분류하고자 한다.

4.1 한글의 구성형태

한글은 자모결합에 의해서 이루어진다. 한글 표준어의 자음은 중자음의 포함한 21개이고, 모음은 중모음을 포함한 19개로 구성된다. 각 자모는 음절을 이루는 조합형태에 따라 초성, 중성, 종성으로 나눌 수 있다. 이러한 한글의 구성형태는 표 2와 같다.

표 2. 한글의 구성형태.
Table 2. Composition of the Hangeul.

초성	단자음	ㄱ, ㄴ, ㄷ, ㄹ, ㅁ, ㅂ, ㅅ, ㅇ, ㅈ, ㅊ, ㅋ, ㆁ, ㆁ
	중자음	ㄲ, ㄴ, ㄷ, ㄹ, ㅁ, ㅂ, ㅅ, ㅇ, ㅈ, ㅊ, ㅋ, ㆁ, ㆁ
중성	단모음	ㅏ, ㅑ, ㅓ, ㅕ, ㅗ, ㅛ, ㅜ, ㅠ, ㅡ, ㅣ
	중모음	ㅘ, ㅙ, ㅚ, ㅜ, ㅝ, ㅞ, ㅟ, ㅠ, ㅡ, ㅢ, ㅣ, ㅤ, ㅥ, ㅦ, ㅧ, ㅨ, ㅩ, ㅪ, ㅫ, ㅬ, ㅭ, ㅮ, ㅯ, ㅰ, ㅱ, ㅲ, ㅳ, ㅴ, ㅵ, ㅶ, ㅷ, ㅸ, ㅹ, ㅺ, ㅻ, ㅼ, ㅽ, ㅾ, ㅿ, ㅿ
종성	단자음	ㄱ, ㄴ, ㄷ, ㄹ, ㅁ, ㅂ, ㅅ, ㅇ, ㅈ, ㅊ, ㅋ, ㆁ, ㆁ
	중자음	ㄲ, ㄴ, ㄷ, ㄹ, ㅁ, ㅂ, ㅅ, ㅇ, ㅈ, ㅊ, ㅋ, ㆁ, ㆁ

4.2 입모양 분류

한글은 자소(초, 중, 종성)들의 조합으로 생성되는 문자는 총 14,364자에 이르며, 현재 실용되는 한글 한자 표준코드(KSC5601)에서의 한글은 2350字이다 [14]. 그러므로 2350자를 포함한 모든 음절에 대해서 한글의 표준 발음법[15]에 준하여 입모양을 분류하므로써 모든 음절에서 표현될 수 있는 입모양을 만들 수 있다. 한글의 구조형태에서 받침이 있는 음절과 없는 음절로 나눌 수 있는데 이 두가지로 나누어서 각각 입모양을 분류한다. 먼저, 받침이 없는 음절에 따른 입모양을 분류하고, 이 분류된 입모양에 받침이 결합하는 경우를 고려하면 모든 음절을 분류하게 된다.

[1] 받침이 없는 경우(초성+중성)

받침이 없는 음절에 따른 입모양을 분류하기 위해서 자음과 모음 각각이 입모양에 미치는 영향을 살펴보고, 이 영향에 따른 입모양을 분류하고자 한다. 이것은 입모양이 발음초기에는 자음인 초성에 의해 결정되지만, 점차로 중성인 모음에 의해 지배된다고 생각되기 때문이다.

(1) 자음(초성)에 의한 영향

음절에 따른 입모양패턴은 대부분 모음에 의존하지만 입술소리(ㄱ, ㅋ, ㆁ)와 조합되는 경우에는 발음초기의 입모양이 다르다. 예를들면 자음이 모음 「ㅏ」와 결합할 경우에 「나」, 「까」등과 같이 입술소리를 제외한 다른소리와 결합할 때는 입모양의 시작위치가 거의 동일하지만 「파」등과 같이 입술소리와 결합할 때는 닫혀진 모양에서 시작된다. 따라서 음절에서 자음의 영향을 받아 표현되는 입모양은 입술소리와 그외소리에 의한 두개의 입모양으로 분류할 수 있다. 표 3은 조음부위별 자음의 발음을 분류한 것이다.

표 3. 조음부위별 자음의 발음.
Table 3. Pronunciation of the consonants according to modulation region.

	입술소리	혀끝소리	구개음	연구개음	목청소리
예사소리	ㅂ	ㄷ, ㅌ	ㅈ	ㅊ	ㅎ
기센소리	ㅍ	ㅌ	ㅊ	ㅊ	
된소리	ㅃ	ㄸ	ㅆ	ㅈ	ㅊ
비음	ㅁ	ㄴ		ㅇ	
유음		ㄹ			

(2) 모음(중성)에 의한 분류

자음은 발음초기에는 영향을 미치지만 입모양패턴을 지배하는 것은 모음이므로 본 절에서는 모음에 대한 분류를 하고자 한다. 여기에서 모음은 중모음과 단모음으로 구분할 수 있지만 발음시에는 대부분의 중모음은 단모음의 연결이기 때문에, 그 입모양도 단모음에 대한 입모양이 순차적으로 변한다고 생각되어 동영상합성시 고려하기로 한다. 그러나, 「ㅘ」, 「ㅙ」와 같은 중모음은 하나의 입모양으로 발음되므로 입모양분류의 편의를 위해 다음의 발음 정의한다.

- 기본형모음: 발음에 있어서 기본이 되는 입모양 패턴을 형성하는 모음
- 조합형모음: 두개의 기본형모음의 결합으로 이루어진 모음

이 정의에 따라 모음을 분류한 것을 표 4에 나타낸다. 단, 여기에서 포함되지 않는 4개모음 「ㅚ」, 「ㅙ」, 「ㅘ」, 「ㅙ」는 각각 「ㅘ」, 「ㅙ」, 「ㅘ」, 「ㅙ」의 입모양 패턴을 간으나, 발음지속시간이 다르기 때문에 기본적인 입모양패턴으로는 분류하지 않는다. 발음지속 시간차는 동영상합성시 고려하기로 한다. 그리고 각 모음에 대한 입모양패턴을 살펴보면, 기본형모음은 그림 4의 입술의 가로와 세로로 벌어진 정도에 따라 8개로 분류할 수 있다. 이에 반해 조합형모음의 입모양패턴은 기본형모음의 입모양패턴이 순차 조합된 것으로 나타난다. 예를들면, 「과」의 입모양은 「고」와 「ㅏ」의 입모양패턴의 조합이다.

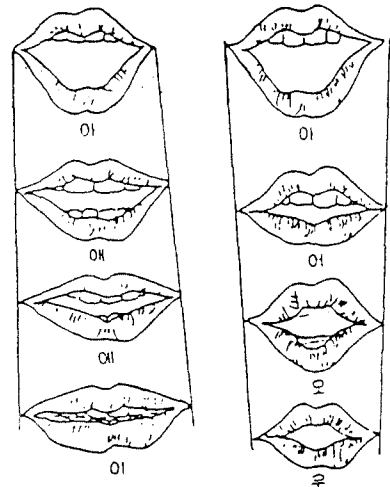


그림 4. 입모양의 비교.
Fig 4. Comparison of the lip shapes.

표 4. 입모양에 따른 모음분류.

Table 4. Vowel classification according to the lip shapes.

모음	기본형모음	ㅏ(ㅑ), ㅓ(ㅕ), ㅗ(ㅛ), ㅜ(ㅠ), ㅡ, ㅣ, ㅐ, ㅓ
	조합형모음	ㅘ, ㅙ, ㅚ, ㅜ, ㅠ, ㅡ, ㅣ

이것과 같이 자음과 모음에 대해서 입모양의 생성에 미치는 영향을 살펴보면, 총 10개의 입모양패턴으로 받침이 없는 모든 음절을 표현할 수 있음을 알 수 있다.

[2] 받침이 없는 음절

한글의 표준발음법에서 받침소리로는 ‘ㄱ, ㄴ, ㄷ, ㄹ, ㅁ, ㅂ, ㅇ’의 7개 자음으로만 발음한다. 즉, 한글의 구성형태에서 어느 자음이 받침으로 올지라도 발음은 7가지로만 발음한다. 따라서 받침이 있는 경우의 입모양을 분류하기 위한 음절은 7개의 자음이 받침으로 오는 경우만 고려하면 된다. 이 경우 또한 모음에 의해 입모양은 거의 지배되거나 최종생성되는 입모양은 이미 분류한 입술소리와 그외소리의 입모양으로 된다.

4.3 입모양의 변형규칙

음절에 따른 입모양 합성시스템을 실현하기 위해 음절에 따른 입모양을 분류하였다. 그 결과, 모든 음절의 입모양은 8개의 기본형 입모양과 자음에 의한 2개의 입모양으로 표현할 수 있었다. 이처럼 분류된 10개의 입모양에 대해 그림 4의 표준패턴과 같은 모양이 되도록, 문헌[7]을 참고로 AU를 선택하고 AU의 강도를 정하였다. 이 결과를 표 5에 나타낸다. 이와같이 얼굴근육을 고려하여 입모양변형규칙을 정하면 얼굴영상의 크기 및 개인차가 있어도 같은 규칙을 적용할 수 있다. 표 5의 변형규칙에 따라 얻어진 입모양의 합성영상을 그림 5에 나타낸다. 또한, 동영상 고려하기 위해 입술소리와 기본형 모음 「ㅏ」가 결합하여 입모양의 생성되는 과정을 그림 6에 나타낸다. 나아가서, 표정합성과 친화성을 살펴보기 위해 다양한 표정을 지으면서 말하는 얼굴영상도 합성하였다. 이것을 그림 7에 나타낸다.

그림 5~그림 7은 그림 2에 나타난 1장의 사진으로 합성한 영상으로, 여러가지 음절에 대한 입모양과, 동영상, 다양한 표정을 고품질로 자연스럽게 합성할 수 있음을 나타내고 있다. 이와같이 얼굴영상을 자음

표 5. 분류한 입모양의 AU조합과 강도.

Table 5. AU combinations and intensities of classified the lip shapes.

입모양	AU	AU의 강도
「ㅏ」	AU10	0.3
	AU20	0.2
	AU26	0.6
「ㅓ」	AU10	0.3
	AU26	0.3
「ㅗ」	AU10	0.3
	AU18	0.5
	AU26	0.2
「ㅜ」	AU10	0.1
	AU18	0.6
	AU26	0.1
「ㅛ」	AU10	0.2
	AU20	0.3
	AU26	0.3
「ㅠ」	AU20	0.3
	AU26	0.3
「ㅡ」	AU20	0.4
	AU26	0.2
「ㅣ」	AU10	0.2
	AU20	0.5
	AU25	0.2
	AU26	0.1
입술소리	AU18	0.3
그외소리	AU26	0.2

자재로 합성한다면, 다방면에서 휴먼인터페이스의 실현이 가능하리라 생각된다.

V. 결 론

인간과 기계가 자연스럽게 대화하는 맨머신 인터페이스의 궁극적인 형태로서 휴먼인터페이스를 실현하기 위해, 얼굴표정합성과 음절에 따른 입모양의 합성방법에 대해서 검토하였다. 음절의 발음에 부합된 입모양패턴을 한글의 모든 음절을 대상으로 입모양

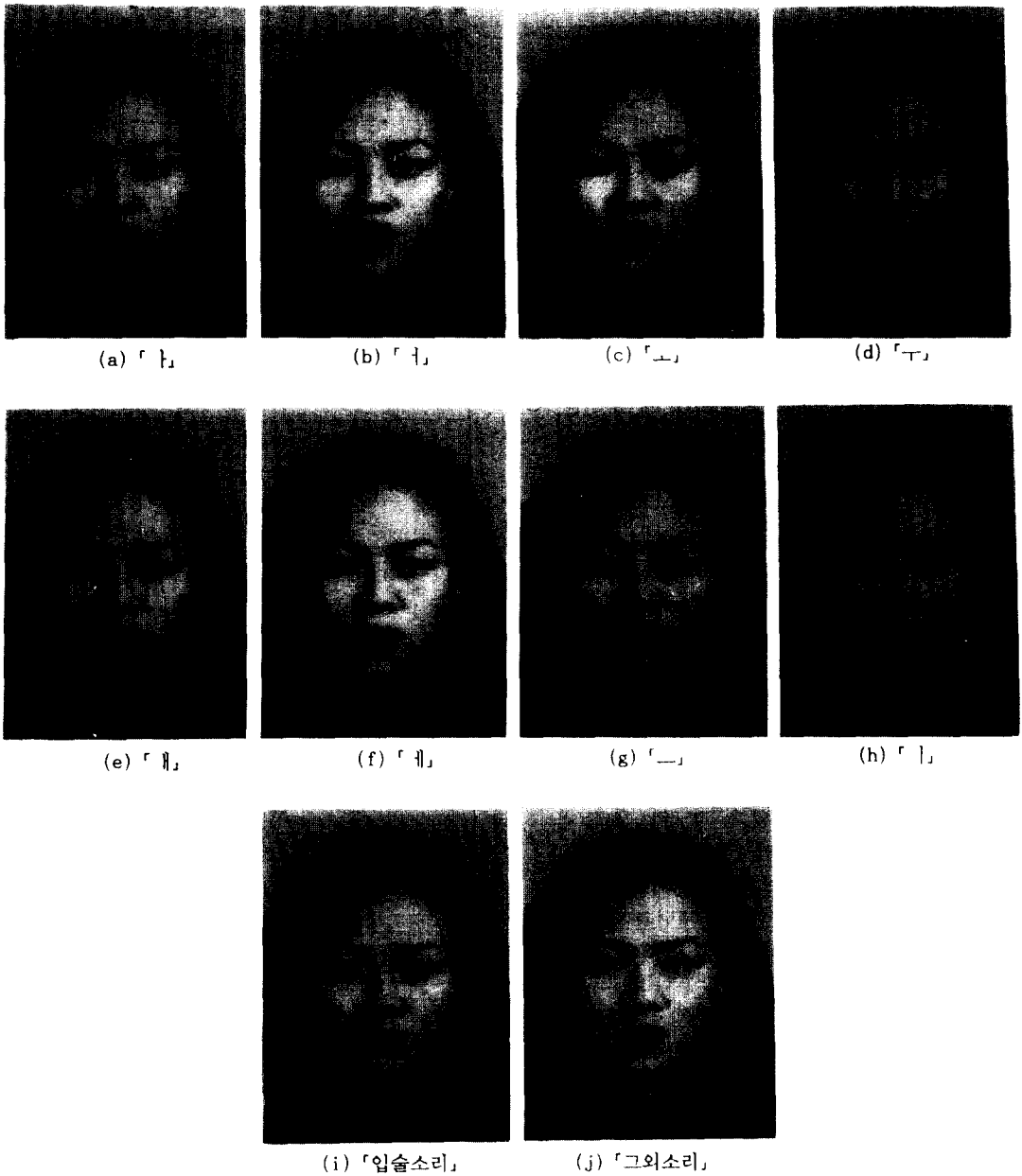


그림 5. 입모양패턴.
Fig 5. The lip shape patterns.

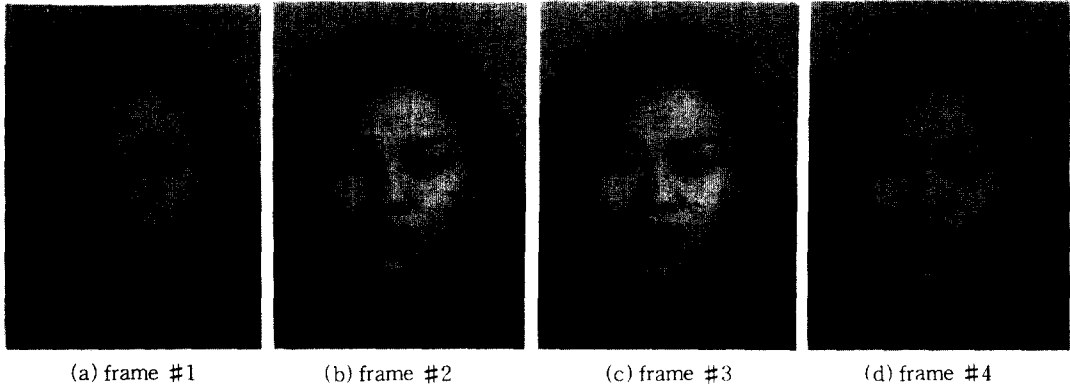


그림 6. 「바」에 대한 동영상의 예.
Fig 6. Image sequence for 「Ba」.

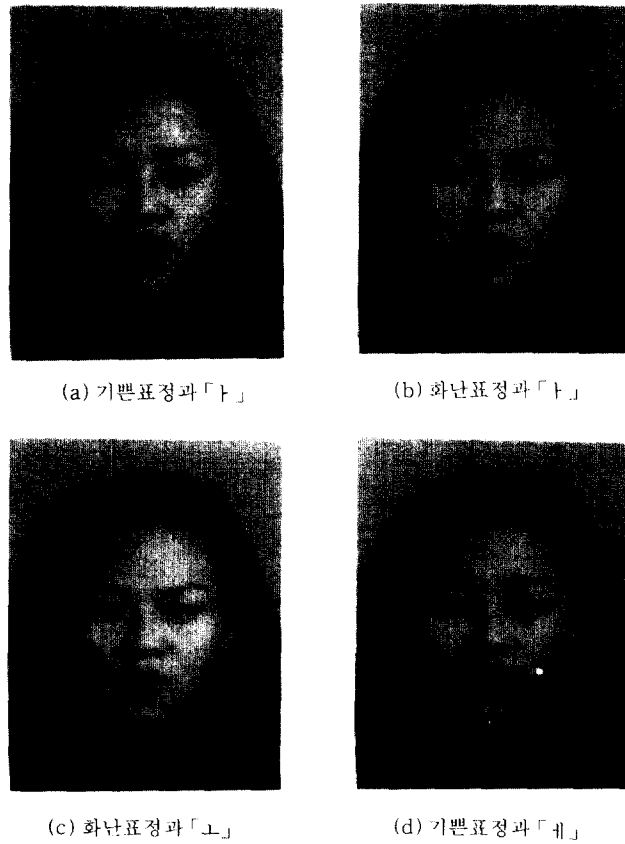


그림 7. 얼굴표정을 가진 입모양패턴.
Fig 7. The lip shape patterns with the facial expression.

의 가로와 세로로 벌어진 형상을 토대로 분류하였다. 입모양패턴의 분류에서는 음절에서 자음이 입모양에 미치는 영향과 모음이 미치는 영향을 고려하였다. 그 결과, 자음에 의한 입모양패턴은, 입술소리와 그외소리(혀끝소리, 구개음, 연구개음, 목청소리)로 2개, 모음에 의한 입모양패턴은 기본형모음 8개로 분류되었다. 이와같이 음절에서 자모에 의한 영향을 고려한 결과, 10개의 입모양패턴으로 모든 음절을 표현할 수 있음을 알았다. 또한, 얼굴근육의 움직임도 고려하여 입모양패턴을 합성한 결과, 얻어진 입모양은 실제의 사진과 같이 현실감있게 표현되었다. 또한, 다양한 표정을 지으면서 말하는 얼굴영상도 비교적 자연스러운 것을 합성할 수 있다.

참 고 문 헌

1. F. Parke, "A Parameterized Model for Facial Animation," IEEE Computer Graphics & Applications, 2, 9, pp.61-68, Nov., 1982
2. J. P. Lewis and F. I. Parke, "Automated Lip-Synch and Speech Synthesis for Charactor Animation," CHI+GI 1987 Conf. Proc., pp. 143-147, Toronto, April, 1987
3. A. Lippman, "Semantic Bandwidth Compression :Speech-maker," Picture Coding Symposium (PCS81), pp. 29-30, 1981
4. H. Harashima, K. Aizawa and T. Satio, "Model-Based Analysis Synthesis Coding Videotelephone Images," Trans. IEICE Japan, E72, 5, pp. 452-459, 1989
5. K. Aizawa, H. Harashima and T. Satio, "Model-Based Analysis Synthesis Image Coding System for a Person's Face," Image Comm, 1, 2, pp. 139-152, 1989
6. 森島 繁生, 岡田 信一, 原島 博, "知的 イソタフェ -スのための顔の表情合成法の一検討," 日本電子情報通信學會論文誌, J73-D-II, 3, pp.351-359, 1990.
7. 崔昌石, 原島 博, 武部 幹, "顔の3次元モデルに基づく表情の記述と合成," 日本電子情報通信學會論文誌, J73-A, 7, pp.1270-1280, 1990.
8. C. S. Choi, H. Harashima and T. Takebe, "Analysis and Synthesis of Facial Expressions in Knowledge-Based Coding of Facial Image Sequences," IEEE International Conf. Acoust. Speech & Signal Process. (ICASSP), M9.7, pp.2737-2740, 1991
9. 崔昌石, 原島 博, 武部 幹, "顔の3次元モデルを用いた顔面表情の分析," 日本電子情報通信學會論文誌, D II, vol. J74 D II, 6, pp766-777, 1991
10. P.S. Heckbert, "Survey of Texture Mapping," IEEE CG&A, pp.56-67, 1986
11. P. Ekman and W.V. Friesen, "Facial Action Coding System," Consulting Psychologist Press, 1977
12. 이용동, 최창석, 최갑석, "휴먼인터페이스를 위한 얼굴표정과 입부분합성에 관한 기초검토," 통신학회 하계학술대회논문집, pp.430-434, 1992
13. 이용동, 최창석, 최갑석, "휴먼인터페이스를 위한 음절분류에 따른 얼굴영상의 합성," 제5회 신호처리합동학술대회 논문집, pp.433-437, 1992
14. "情報時代," 月刊マイクロソフト웨어, 3月号, pp.72-73, 1989
15. 문교부 고시·문화부 공고, "국어 어문 규정집," 대한교과서주식회사, 1991



李 龍 東(Yong Dong Lee) 正會員
1954年 6月 16日生
1991년 2월 : 명지대학교 전자공학과(공학사)
1993년 2월 : 명지대학교 대학원 전자공학과(공학석사)
※주관심분야 : 영상부호화, 영상처리, 컴퓨터비전, 음성합성



崔 昌 石(Chang Seok Choi) 正會員
1954年 7月 15日生
1978년 2월 : 홍익대학교 전자공학과 졸업(공학사)
1988년 3월 : 일본 가나자와(金澤)대학원 전기정보공학(공학석사)
1991년 3월 : 일본 가나자와(金澤)대학원 전기정보공학과(공학박사)

1988년 4월 ~ 1991년 3월 : 동경대학 연구원
1984년 1월 ~ 1992년 2월 : 산업기술정보원(국, KIET) 전자전기실장

1992년 3월 ~ 현재 : 명지대학교 정보통신공학과 조교수
1993년 : 일본 전자정보통신학회 논문상수상

※주관심분야 : 영상부호화, 영상처리, 컴퓨터비전, 컴퓨터 그래픽스

崔 甲 石(Kap Seok Choi)

正會員

1963年 5月 6日生

현재 : 명지대학교 정보통신공학과 교수(제16권 1호 참조)

※주관심분야 : 디지털신호처리, 음성인식, 음성합성, 패턴인식