
3차원 모델을 이용한 입모양 인식 알고리즘에 관한 연구

배 철 수*

A study on the lip shape recognition algorithm using 3-D Model

Cheol-Soo Bae*

요 약

최근 통신 시스템의 연구와 발전 방향은 목소리의 음성 정보와 말하는 얼굴 영상의 화상 정보를 함께 적용하되 음성 정보만을 제공하는 경우보다 높은 인식율을 제공한다. 따라서 본 연구는 청각장애자들의 언어 대체수단 중 하나인 구화(speechreading)에서 가장 시각적 변별력이 높은 입모양 인식을 일반 퍼스널 컴퓨터상에서 구현하고자 한다. 본 논문은 기존의 방법과 달리 말하는 영상 시퀀스에서 입모양 인식을 행하기 위해 3차원 모델을 사용하여 입의 벌어진 정도, 턱의 움직임, 입술의 돌출과 같은 3차원 특징 정보를 제공하였다. 이와 같은 특징 정보를 얻기 위해 3차원 형상 모델을 입력 동영상에 정합시키고 정합된 3차원 형상 모델에서 각 특징점의 변화량을 인식파라미터로 사용하였다. 그리고, 인식단위로 동영상을 분리하는 방법은 3차원 특징점 변화량에서 얻어지는 강도의 기울기에 의하여 이루어지고, 인식은 각각의 3차원 특징벡터를 이산 HMM 인식기의 인식 파라미터로 사용하였다. 본 논문에서는 한국어 10개 모음에 대하여 인식실험하여 비교적 높은 인식율을 얻을 수 있는 것으로 보아 본 연구에서 사용한 특징 벡터를 시간적 변별 요소로서 사용할 수 있음을 제시하였다.

Abstract

Recently, research and developmental direction of communication system is concurrent adopting voice data and face image in speaking to provide more higher recognition rate than in the case of only voice data. Therefore, we present a method of lipreading in speech image sequence by using the 3-D facial shape model. The method use a feature information of the face image such as the opening-level of lip, the movement of jaw, and the projection height of lip. At first, we adjust the 3-D face model to speaking face image sequence. Then, to get

* 관동대학교 전자통신공학과

접수일자 : 1998년 9월 225일

a feature information we compute variance quantity from adjusted 3-D shape model of image sequence and use the variance quality of the adjusted 3-D model as recognition parameters. We use the intensity inclination values which obtaining from the variance in 3-D feature points as the separation of recognition units from the sequential image. After then, we use discrete HMM algorithm at recognition process, depending on multiple observation sequence which considers the variance of 3-D feature point fully. As a result of recognition experiment with the 8 Korean vowels and 2 Korean consonants, we have about 80% of recognition rate for the plosives and vowels. We propose that usability with visual distinguishing factor that using feature vector because as a result of recognition experiment for recognition parameter with the 10 korean vowels, obtaining high recognition rate.

1. 서론

언어의 자동 인식에 관한 연구는 1950년대부터 활발히 진행되어져 왔다. 초기에는 음성 정보만을 이용하여 처리되었지만 90년대에 이르러 처리할 수 있는 하드웨어의 급속한 발전에 힘입어 마이크로폰을 이용한 음성 정보와 CCD카메라 등과 하드웨어의 급속한 발전에 힘입어 마이크로폰을 이용한 음성정보와 CCD 카메라 등과 같은 영상입력장비의 영상정보를 이용하여 특수한 실험환경이 아닌 일반적인 환경에서 음성 정보와 영상 정보를 동시에 처리하여 화자의 언어 인식률을 향상시켰다. 따라서, 독순(lipreading)은 구화에서 변별이 가장 높은 시각정보로서 컴퓨터로 구현하기 위해서는 세가지 문제점을 해결해야 한다. 첫째, 입력영상에서 정확한 특징점을 추출하는 문제이고, 둘째, 입술의 움직임을 충실히 표현할 수 있는 인식파라미터를 사용하는 것, 마지막으로 입력동영상에서 인식단위로 영상의 프레임을 분리하는 문제이다. 이러한 문제점들에 대한 기존의 연구들[1~4]은 주로 2차원적인 형태와 특정점의 움직임 변위에 대해서만 고려하였기 때문에 입력영상에서 약간의 두부의 움직임만 있어도 인식파라미터가 변하게 된다. 또한, 동영상에서 인식단위로 분리하는 과정은 거치지 않아 자동화하기가 어렵다.[5~6]

본 논문에서는 위의 문제점들을 해결하기 위해 말하는 동영상에서 입술의 움직임에 부합하는 글자를 인식하는 방법과 음절단위로 입모양을 인식하기 위해서 동영상에 음절을 효과적으로 분리할 수 있는 간단한 방법제안과 한글에 대해 고찰을 한다. 그

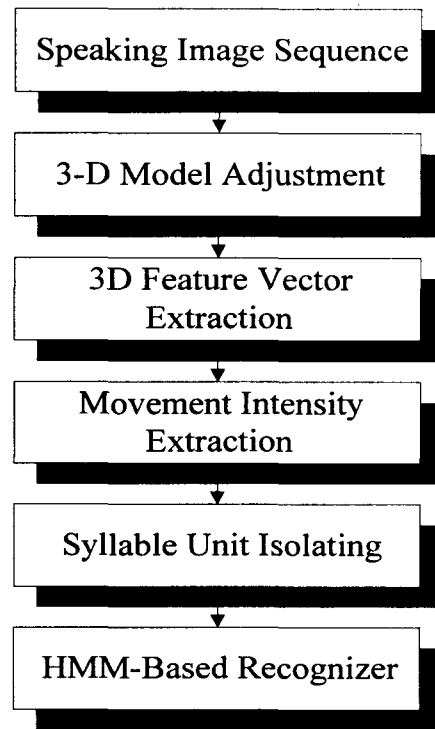


그림 1. 전체 영상처리 과정
Fig 1. Image Processing

림 1은 동영상 입력에서부터 특징벡터의 추출까지의 과정이다. 인식에서는 각 입력특징벡터에 대한 출력확률을 평균벡터로 하는 정규분포로 가정하고, 다차원(multi-dimensional), 다단계 라벨링방법을 사용하여 3차원 특징벡터를 입력벡터로 한 이산 HMM을 사용하였다.[7~8]

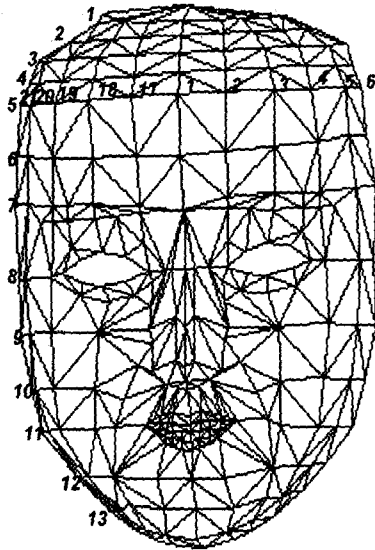


그림 2. Wire Frame Model
Fig 2. Wire Frame Model

II. 입력 영상의 3차원 모델 정합

2차원영상은 3차원영상의 투영이라는 관점에서 입술의 움직임은 3차원적으로 추정하여야 한다. 입력동영상으로부터 입술움직임에 대한 3차원 특징 벡터를 얻기 위해서는 얼굴형상을 충실히 표현할 수 있는 3차원모델이 필요하다. 따라서 본 논문에서는 얼굴형상 및 얼굴표면을 표현하기 위해 3차원 위치좌표 점(x, y, z)과 선으로 근사한 삼각형 구조의 Wire Frame Model (그림. 2)을 사용한다.

WFM을 영상의 각 프레임에 정합하여 입력영상을 표현하므로 각 프레임당 입술움직임에 대한 WFM 정점의 이동변위를 구할 수 있다.

본 연구에서는 특징점 검출을 위해 영상의 휘도치 분포를 다단계로 임계화하는 방법을 사용한다. 입의의 대상 영상들로부터 획득된 확률·통계적 분석에 의해 다단계 임계화된 구간을 수직·수평으로 투영하여 얼굴부위에 해당하는 휘도치분포의 경계값을 결정한다. 이 결정된 임계값에 따라 얼굴 이외의 성분으로부터 얼굴을 분할하고, 분할된 얼굴부위를 바탕으로 안면 요소 특징점들(눈, 코, 입 등)을 추출한다.

III. 3차원 정보의 특징 파라미터 검출 알고리즘

독수는 화자의 입술의 움직임을 관찰하여 음성언어를 이해하는 보조수단으로서 음성의 지각단위인 음소와 같이 시각적 지각 요소를 가질 수 있다. 이러한 시각적 지각요소를 시각소(visual phoneme)라고 하고 이것들은 음소에 대한 입술의 독특한 움직임을 가지고 있다. 대표적인 시각소는 입술의 움직임, 입의 벌어진 정도, 턱의 움직임, 입술의 돌출, 뺨의 움직임 등과 같은 복합적인 움직임으로 이루어진다. 그러므로 이들의 복합적인 움직임 고려하여야 하지만은 입술의 돌출과 뺨의 움직임 등은 2차원적인 기존의 접근방법으로는 곤란하다. 이때, 얼굴 영상의 특징점의 깊이 좌표 z_i 는 3차원 모델의 깊이 좌표 z_i 를 기본으로 추정할 수 있다. 그 추정식은 식(1)과 같다.

$$z'_i = \frac{w_{face}}{W_{WFM}} z_i \dots\dots\dots \text{식(1)}$$

윗식에서 w_{face} 는 대상에 대한 얼굴의 폭이고 W_{WFM} 은 얼굴의 일반적인 3차원 형상 모델에 대한 얼굴의 폭이다. 따라서 본 연구에서는 입력된 동영상에 정합된 3차원모델로부터 시각소를 표현할 수 있는 3차원 특징 벡터를 그림 3과 같이 12개의 특징점을 사용하여 4개의 특징벡터를 추출한다. 4개의 특징벡터는 다음과 같다.

- 입술점 특징 벡터
- 윗 입술 움직임 벡터(UP) :
VL = P₂ - P₀
- 아랫 입술 움직임 벡터(Down) :
HL = P₃ - P₀
- 입의 벌어진 정도 :
- TL = P₅ - P₆
- 턱의 상하움직임
- JL = P₇ - P₁₀

이와같은 움직임벡터를 정량적으로 취급하기 위해서는 개인차를 흡수할 수 있는 기본량이 필요하다. 여기에서는 입의 중형폭을 이용하여 기준량으로 정한다. 움직임벡터의 프레임당 이동변위는 특징점의 이동량을 기준량으로 정규화한 것이다.

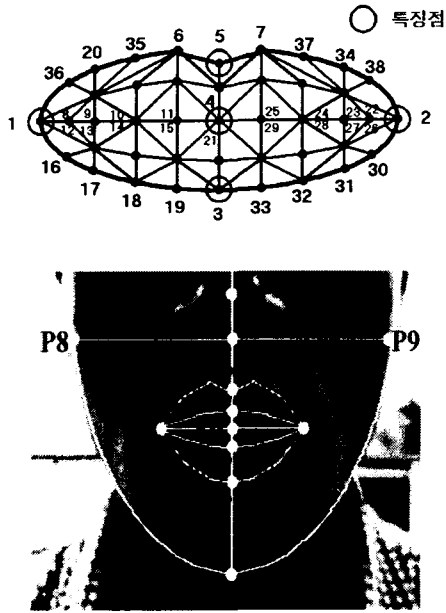
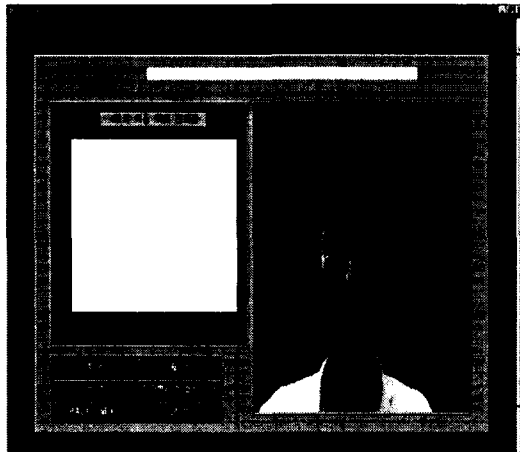
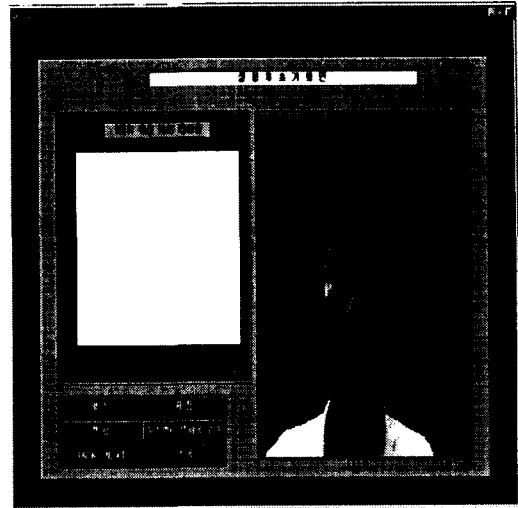


그림 3. 특징점
Fig 3. Feature Points

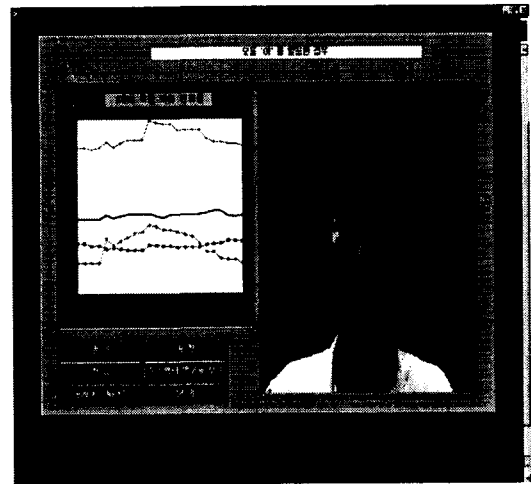
그림 4는 본 연구의 입모양 인식 알고리즘을 실행한 예이다. 그림 4(a)는 입력 영상의 초기화면이고 그림 4(b)는 입력 영상에 3차원 형상 모델을 정합한 후를 나타낸다. 또한, 그림 4(c)는 모음 '아'를 발음하였을 때 각 4개의 특징 벡터 변화율과 영상 정보를 함께 나타내고 있다.



(a) 정합 전
(a) Before Adjustment



(b) 정합 후
(b) After Adjustment



(c) 모음 '아'
(c) Vowel /a/

그림 4. 3차원 모델 정합 예 'a/'
Fig 4. A examples of 3-D model's adjustment

IV. 입력 동영상의 음절분리

입모양의 인식은 음절단위로 행해진다. 따라서 말하는 영상 시퀀스를 음절단위로 구분할 필요가 있다. 본 연구에서 입력 동영상의 음절분리는 특징

벡터의 기울기의 굴곡점을 검출하여 분리할 수 있다. 3차원모델로 부터 얻어지는 특징벡터의 강도는 [코리아]를 발성했을 경우 그림 5에 나타난 것처럼 선형적으로 증가하다가 다음 음절의 발성으로 갈 때 입모양패턴의 끝점으로 연결된다. 이때 특징벡터의 시간에 대한 기울기는 변하게 된다. 이 기울기가 변하는 점을 음절의 끝부분으로보고 분리한다. 이 분리는 시간 t의 모든 구간에 대해 좌미분계수와 우미분계수를 구하여 이룰수 있다.

시간 t₁에서 미분 계수는

$$\lim_{h \rightarrow 0, h < 0} \frac{f(t_1+h) - f(t_1)}{h} \neq \lim_{h \rightarrow 0, h > 0} \frac{f(t_1+h) - f(t_1)}{h} \dots \dots \dots \text{식(2)}$$

과 같이 되어 좌미분계수와 우미분계수가 다르게 된다. 이때, 시간 t₁은 음절 구분점이다. 실제 이 방법은 음절분리에 좋은 결과를 주었다. 그러나 이 방법은 동일한 모음이 계속될때 좋은 결과를 얻을 수 없다. 그것은 같은 패턴이 계속될 때 입모양의 특징벡터의 기울기는 변하지 않고 같은 입모양형상이 계속되기 때문이다. 예를 들면 [아가가]라고

발성하였을 때 자음에 대한 입모양의 영향은 없었고 [가]의 입모양패턴이 세음절의 발생시간동안 계속되기 때문이다. 따라서 동일한 입모양패턴이 계속되는 경우 음절발성시간의 평균시간에 따라 음절분리를 행한다.

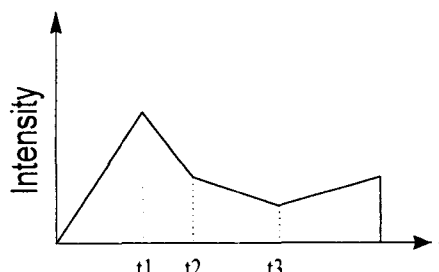


그림 5. 시간에 따른 특징벡터의 강도변화
Fig 5. The change of the feature vector intensity for the time

V. 이산분포 HMM의 인식 알고리즘

Speaking image sequence로부터 추출된 각 특징점의 변화량은 multiple observation sequence로

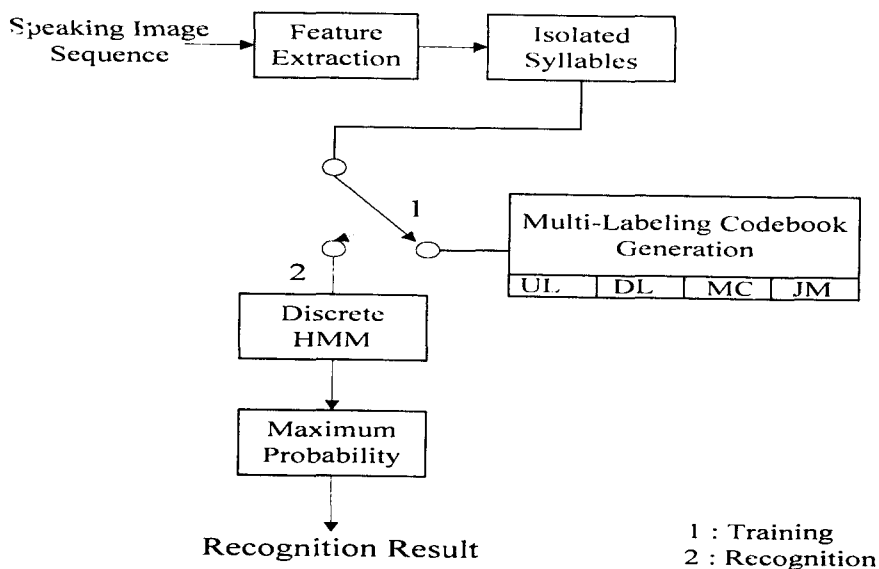


그림 6. 이산HMM 인식기
Fig 6. Discrete HMM Recognizer

HMM의 입력 파라미터로 사용한다. 그림 3에서 나타낸 4개의 특징벡터는 모든 모델 파라미터의 신뢰성있는 평가를 충분히 수행한다.

먼저, 다음과 같이 4개의 관측 sequence의 집합을 정의한다.

$$O=[O^{VL}, O^{HL}, O^{TL}, O^{LL}] \dots\dots\dots\text{식(3)}$$

우리는 각각의 관측 sequence가 매 다른 관측과 독립으로 가정하고, 우리의 목표는 모델 λ 의 파라미터를 식(4)과 같이 최대로 하기위해 조정하는 것이다.

$$P(O|\lambda) = \prod_{K=1}^K P(O^{(K)}|\lambda) \dots\dots\dots\text{식(4)}$$

$$= \prod_{K=1}^K P_K$$

채산정 수식은 다양한 사건의 빈번한 발생에 기초하기 때문에, 다중 관측 sequence의 채산정 수식은 각각의 sequence에 대한 개개의 빈번한 발생과 함께 더해져서 수정된다.

이러한 multiple observation sequence를 특징 벡터로 이용하여 이산 HMM의 입력벡터로 사용하였다. 그림 6에 HMM인식기의 구성도를 나타내었다.

VI. 실험 및 고찰

1) 실험 시스템 구성

본 논문에서는 말하는 동영상 데이터는 형광등 및 창문을 통한 자연빛을 조명으로 하고 카메라와 대상간의 거리는 1.5m으로 일정한 간격에서 벽면을 배경으로 하여 비디오 캠코더로 녹화하였다. 녹화된 영상을 조크셔틀을 이용하여 매 프레임을 512×480×24bit로 입력을 받았다. 동영상 데이터베이스는 5인×10개×5회=250개로 구성된다. 이 중 3인×10개×5회=150개를 학습데이터로 사용하였고, 나머지 2인이 발생한 영상 데이터는 HMM인식 실험 데이터로 사용하였다. 이와 같은 실험 시스템을 그

림 7에 나타내었다. 이 시스템은 윈도우즈 95환경에서 영상처리 통합환경을 구축하였다.

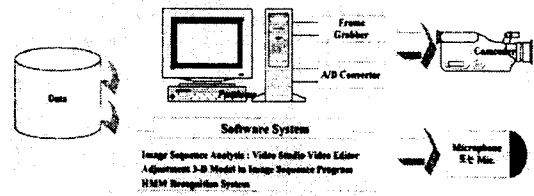


그림 7. 실험 시스템
Fig. 7. Experiment System

2) 실험 및 결과

본 논문에서 입모양 인식에 사용한 영상데이터는 학습에 사용될 학습데이터와 인식실험에 사용될 시험데이터로 구분한다. 실험에 사용된 한국어 10개 모음 「가」, 「나」, 「다」, 「라」, 「마」, 「바」, 「사」, 「자」, 「차」, 「카」 음절을 사용하여 인식실험 하였다. 인식 실험을 위하여 입력영상을 12개의 특징점을 바탕으로 4개의 특징벡터를 추출하여 그 변화량을 표 1과 같이 나타내었다.

또한, 그 특징 벡터의 변화량에 의한 각 모음에 대한 4개의 특징 벡터 변화를 그림 8~17에 각각 나타내었다.

그림에서 볼 수 있듯이 각 음절에 따른 특징 벡터의 차이를 볼 수 있다. 인식은 입력 특징 벡터를 벡터 양자화하여 이산 HMM의 입력벡터로 하였다. 인식 실험한 결과 모음 10개의 평균 인식률이 55%를 각각의 인식률을 그림 18에 나타내었다.

여기서 「나」의 인식률이 좋지 않은 것은 「오」와 혼동이 발생한다고 사료된다. 또한 국어에서는 입술의 실물크기 사진을 가지고 인식 실험한 결과 모음 10개의 평균 인식률이 47.7%하 하였다. 그 각각의 모음에 대한 인식 결과를 그림 19에 나타내어 본 논문과 비교, 분석한 결과 본 논문에서 향상된 인식률을 보였다.

표. 1 4개의 벡터 변화량
Table 1. Change value for four

모음 \ 벡터값	VL	HL	JL	PL
아				
1 Frame	68.11754546	28.06799427	133.9295848	46.17216847
2 Frame	68.11754546	28.06799427	133.9295848	46.17216847
3 Frame	68.11564472	27.75465481	132.3252167	43.70546272
4 Frame	68.11754546	28.08932955	133.9295848	44.23177411
5 Frame	69.12797537	44.23878562	139.6506639	41.23672457
6 Frame	71.02816343	48.28534345	134.7316409	42.05521961
7 Frame	73.01931515	51.58692637	137.9402736	41.17598666
...
24 Frame	68.11754546	27.6523179	137.4370453	46.17216847
25 Frame	68.11754546	27.6523179	135.797562	46.17216847
어				
1 Frame	71.94976594	24.31251321	138.4560635	47.53730007
2 Frame	71.06335202	29.15656599	138.4560635	48.04579802
3 Frame	70.11419257	36.12624171	138.4560635	47.54365055
4 Frame	65.06919394	42.16757926	138.4560592	46.65188858
5 Frame	65.06919394	45.33807799	144.2437607	44.54197551
6 Frame	65.06919394	46.15778717	144.2437607	44.54197551
7 Frame	65.07273746	51.14615553	144.2437628	43.82678385
...
27 Frame	72.24740159	27.77466776	27.77466776	46.21853025
28 Frame	72.24740159	27.09285304	27.09285304	46.21791714
오				
1 Frame	71.94976594	24.3461434	138.4560613	47.53729378
2 Frame	71.06335202	29.15656599	138.4560613	47.53730007
3 Frame	71.06335202	34.13363942	138.4560635	47.42781576
4 Frame	71.06335202	34.13393942	138.4560613	48.21200792
5 Frame	71.06335202	36.12624171	138.4560635	48.21201829
6 Frame	71.06335202	37.08241282	138.4560635	47.3328502
7 Frame	71.06335202	37.08241282	138.4560613	47.33284175
...
20 Frame	72.0624729	27.16809416	138.4560613	48.21200792
21 Frame	72.0624729	27.16809426	138.4560635	48.21201829
우				
1 Frame	71.94976594	25.12181005	138.4560613	47.53729378
2 Frame	71.94976594	25.18144834	138.4560635	47.20591814
3 Frame	71.94976594	27.16809416	138.4560613	47.25884158
4 Frame	71.94976594	27.16809416	138.4560635	47.25884793
5 Frame	71.94976594	30.08496868	138.4560657	47.25885428
6 Frame	72.11102551	35.0728576	138.4560678	47.25886063
7 Frame	68.0661443	35.0728576	138.45607	47.25886698
...
28 Frame	66.2721661	26.81397896	137.7263749	48.02655992
29 Frame	66.2721661	23.5633623	137.7263786	48.02655992
유				
1 Frame	49.01020302	23.11071916	141.8826818	50.06667993
2 Frame	66.12110102	27.09782833	138.4560613	47.53729378
3 Frame	70.11419257	29.15656599	138.4560613	48.40192813
4 Frame	70.17834424	32.14195607	138.4560635	50.61025299
5 Frame	58.03447251	36.08827865	141.8826818	49.80360135
6 Frame	53.33854141	32.14597729	141.882686	50.06667993
7 Frame	53.23532662	30.23844996	141.882686	50.06668774
...
27 Frame	70.17834424	25.42368691	141.8826777	50.06668774
28 Frame	70.17834424	27.09914862	141.8826818	50.06667994

모음 \ 벡터값	VL	HL	JL	PL
이				
1 Frame	72.1110255	27.0966856	140.130874	48.1299632
2 Frame	75.1065909	27.0966856	140.130871	48.1199565
3 Frame	76.1051904	32.0972019	140.130874	48.1199632
4 Frame	76.1051904	35.131615	140.130871	48.1199565
5 Frame	76.0591875	37.0840986	140.130874	48.1199632
6 Frame	76.0591875	37.0706133	140.130874	48.1199632
7 Frame	78	37.0706113	140.130871	48.1199565
...
25 Frame	72.0277724	30.1534471	140.130874	49.3555085
26 Frame	72.0277724	30.1036604	140.130871	49.355505
애				
1 Frame	71.1758386	28.16212599	138.4560613	48.21200792
2 Frame	71.11258679	28.09101885	138.4560635	48.21201829
3 Frame	71.11258679	27.25995855	138.4560613	48.21200792
4 Frame	71.11258679	26.0979949	138.4560635	48.21201829
5 Frame	71.11258679	34.08966618	138.4560613	49.24832289
6 Frame	71.1758386	40.11365528	138.4560635	49.24833711
7 Frame	71.06335202	40.17592986	138.4560635	49.80360135
8 Frame	70.00714249	41.36550907	138.4560635	49.80360135
...
30 Frame	74.56052776	29.10831738	140.9778917	47.49530544
31 Frame	74.56052776	28.11217069	140.9778917	49.41240981
에				
1 Frame	71.06335202	25.12181005	138.4560635	47.53730007
2 Frame	71.06335202	27.11282612	138.4560613	47.53729378
3 Frame	71.07722149	29.60161633	140.0864615	48.08458558
4 Frame	69.6464695	33.13949688	140.0864615	48.08458558
5 Frame	69.02897942	37.109061	140.0864615	48.08458558
6 Frame	69.01578059	38.08352132	141.780853	48.68397234
7 Frame	69.01578059	40.11676203	141.7808557	48.68397894
8 Frame	72.24973011	41.07742197	141.780853	48.68397234
...
28 Frame	75.05997602	30.56906124	144.2845763	49.63156523
29 Frame	72.11102551	30.56906124	141.8102623	48.78267274
외				
1 Frame	70.11419257	24.46379822	140.9559102	48.40192813
2 Frame	69.11584478	24.46379822	140.9559076	48.4019215
3 Frame	69.11584478	28.25406697	140.9559076	48.4019215
4 Frame	69.11584478	28.16544515	140.9559076	48.4019215
5 Frame	67.11929678	33.21584411	140.9559076	48.4019215
6 Frame	67.11929678	35.30286533	140.9559076	48.4019215
7 Frame	67.11929678	36.07342928	140.9559102	48.40192813
8 Frame	63.19810124	37.28662361	140.9559076	48.4019215
...
29 Frame	72.08557853	21.807690761	141.8265637	45.69236939
30 Frame	72.08557853	27.86345289	141.8265637	45.69236939
위				
1 Frame	70.11419257	24.46379822	140.9559076	48.4019215
2 Frame	70.11419257	24.46379822	140.9559076	48.4019215
3 Frame	70.11419257	26.10157659	140.9559102	48.40192813
4 Frame	70.11419257	26.10157659	140.9559102	48.40192813
5 Frame	64.19501538	26.12072549	140.9559076	48.4019215
6 Frame	64.19501538	26.10157659	140.9559076	50.23543343
7 Frame	58.03447251	27.09782833	140.9559076	53.34923778
8 Frame	58.03447251	29.09110346	140.9559076	53.34923778
...
29 Frame	67.0298441	25.18516032	140.9559076	48.88539221
30 Frame	67.0298441	25.18516032	140.9559076	48.88539221

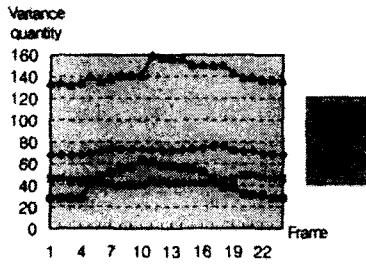


그림 8. /a/의 벡터 변화량
Fig. 8. Vector variance quantity of /a/

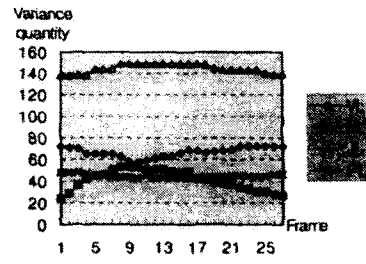


그림 9. /a/의 벡터 변화량
Fig. 9. Vector variance quantity of /a/

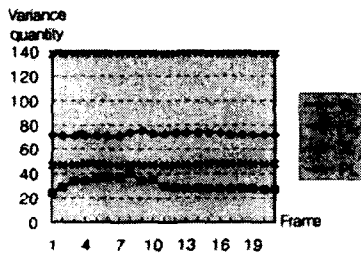


그림 10. /i/의 벡터 변화량
Fig. 10. Vector variance quantity of /i/

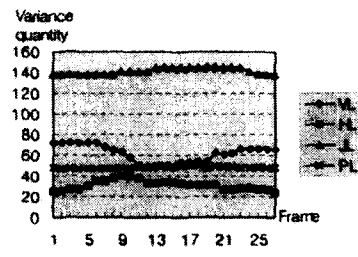


그림 11. /o/의 벡터 변화량
Fig. 11. Vector variance quantity of /o/

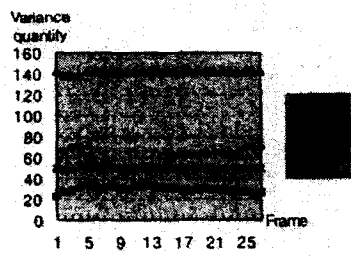


그림 12. /u/의 벡터 변화량
Fig. 12. Vector variance quantity of /u/

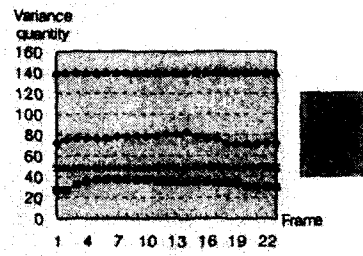


그림 13. /i/의 벡터 변화량
Fig. 13. Vector variance quantity of /i/

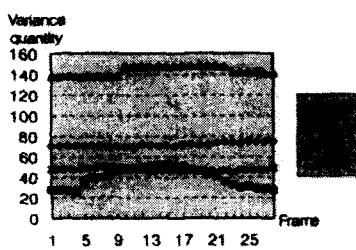


그림 14. /ε/의 벡터 변화량
Fig. 14. Vector variance quantity of /ε/

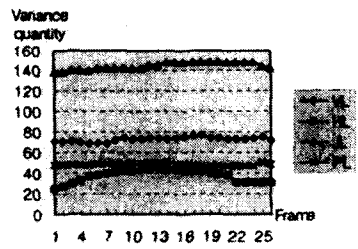


그림 15. /e/의 벡터 변화량
Fig. 15. Vector variance quantity of /e/

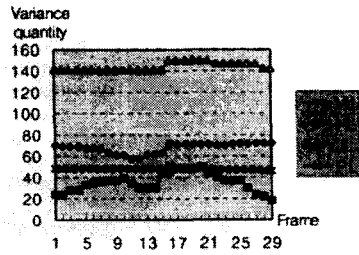


그림 16. /ø/의 벡터 변화량

Fig. 16. Vector variance quantity of /ø/

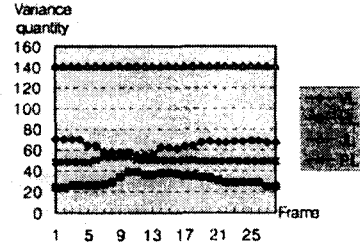


그림 17. /y/의 벡터 변화량

Fig. 17. Vector variance quantity of /y/

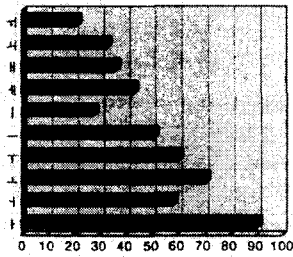


그림 18. 인식 실험 결과

Fig. 18. Experiment System

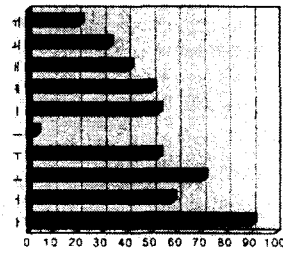


그림 19. 국어학에서 실물사진으로 인식

Fig. 19. Experiment result of the korean language study by a life-size photo

Ⅶ. 결 론

본 논문은 3차원 모델을 사용하여 입모양을 인식하는 독순 방법을 제안 하였고, 입력 동영상을 인식단위로 분리하는 간단한 방법을 제시하였다. 모음 10개에 대하여 인식 실험한 결과 향상된 인식률을 얻었다.

입력 동영상에서 음절단위를 효과적으로 분리하였으며, 3차원 모델의 움직임변위에 따른 특징벡터를 얻을 수 있었다. 인식은 입력 특징벡터를 벡터 양자화하여 이산 HMM을 사용하였다. 인식실험은 한국어 10개 모음을 음절단위로 행하여 모음[10개]의 경우 정상적인 조건하에서 55%의 인식율을 보였다. 국어학에서는 모음 평균 인식률인 47.7%보다 향상된 인식률을 보였다. 그러나, 3차원 모델을 정확히 정합하기 위한 특징점 검출 알고리즘의 개선과 두부움직임에 따른 움직임 추정에 따른 정확

한 움직임 변위를 추출하는 문제가 앞으로의 과제로 남아있다.

따라서 이러한 문제점을 해결하게 되면 크게 인식률을 향상시킬 수 있을 것으로 기대된다.

참고문헌

- [1] E.Petajan, B.Bischoff, D.Bodoff, and N. M. Brooke, " An Improved Automatic Lipreading System to enhance Speech Recognition." In ACM SIGCHI, 1988.
- [2] Mase and A.Pentland."LIP Reading. Automatic Visual Recognition of Spoken Word". Proc. Image Understanding and Machin Vision, Optical of America, June. 1989
- [3] K. E. Finn and A. A. Montgomery. Automatic Optically-Based Recognition of Speech. *Pattern*

Recognition Letters, 8:159 - 164, 1988.

- [4] K. Mase and A. Pentland. Lip Reading: Automatic Visual Recognition of spoken Words. Technical Report 117, M.I.T. Media Lab Vision Science, 1989.
- [5] Danial Reisfeld and Yehezkel Yeshurun, "Robust Detection of Facial Features by Generalized Symmetry" , Proc. ICPR, pp.117-120 , 1992 .
- [6] Young Dong Lee, Chong Seak Choi, Kap Seak Choi, "Lip Shape Synthesis of Korean Syllable for Human Interface". Korea Institut Communication, vol 19, pp.614-623.
- [7] L.R.Raider, "Mathematical Foundations of Hidden Markov Models", Recent Advances in speech understanding and Digital systems,
- [8] L.R.Raider and B.H.Juang, "An Introduction to Hidden Markov Models", IEEE ASSP Magazine Vol. 3, No.1, 99.4 - 16, Jan 1986.



배 철 수(CHEOL-SOO BAE)

1950년 3월 14일생

1979년 2월 명지대학교 전자공학과 졸업(공학사)

1981년 2월 명지대학교 대학원 전자공학과 졸업(공학석사)

1988년 8월 명지대학교 대학원 전자공학과(공학박사)

1989년 11월~현재 한국통신학회 강원지부 지부장

1991년 3월~현재 산업자원부 공 기반 평가위원 및 심의위원

1982년 3월~현재 관동대학교 전자통신공학과 정교수

1993년 3월~1998년 2월 전자,정보통신공학부 학부장

1995년 3월~1999년 2월 관동대학교 대학원 교학부장

1999년 3월~현재 관동대학교 공과대학 학장