

시간특성을 고려한 음성신호의 발성율 검출에 관한 연구

김익성, 서지호, 배명진
송실대학교 정보통신공학과
lawyer24@hanmail.net

요약

발성율은 일정한 시간동안 발생되는 음성신호 내에 몇 개의 음절이 포함되어 있는 지를 나타낸다. 발성율은 화자마다 다르고 각 음소들의 특징에 따라 변화할 수 있다. 발성율의 사전 측정이 이루어 진다면 음성부호화 측면에서도 중요한 정보로 사용될 수 있다. 기존의 음성부호화기는 발성율에 관계없이 고정적인 분석구간을 정하여 전송율을 결정하고 있다. 따라서, 발성율을 미리 측정한다면, 발성율이 느린 부분과 빠른 부분에 각기 다른 부호화 방법을 적용하여 음질을 향상할 수도 있고 전송율을 가변적으로 적용할 수도 있게 된다.

정확한 발성율을 측정하기 위해서는 음절의 변화를 추정하여야 한다. 음절의 변화를 추정하기 위한 방법으로 음성신호의 에너지 포락선 추정법과 LSP를 이용한 추정법이 각각 제안된 바 있으나, 본 논문에서는 위 두 가지 방법을 혼합한 방법을 사용하였다. 에너지 변동은 음성신호의 시간영역 처리방법으로 LSP 파라미터는 음성신호의 선형예측 분석에 의해 구해질 수 있다.

1. 서론

음성신호의 발성율은 화자식별이나 음성인식에서 중요한 파라미터로 사용될 수 있다. 기존의 발성율의 측정은 파형에너지의 변동이나 음운특성의 LSP 파라미터를 별도로 사용하고 있다.

본 논문에서는 변환영역의 LSP 파라미터를 사용하여 발성율을 측정하고자 할 때 시간특성의 에너지 파라미터를 이용하여 발성율 변동의 범위를 제한함으로써 그 측정의 정확성을 개선해 보자고 한다. 제2장에서는 발

성율에 대하여 간략하게 설명하고, 제3장에서도 기존에 제안된 발성율 측정방법을 설명한다. 제4장에서는 제안한 측정법을 설명하고 제5장에서는 실험 및 결과를 기술하고, 제6장에서는 결론을 맺는다.

2. 발성율

발성율이란 화자가 발화할 때 말의 속도와 비교되는 개념으로 초당 발화되는 음절수로 표현된다. 음성통신 분야에 있어서 발성율이라는 파라미터는 음성인식이나 음성 합성에 있어서 중요하게 쓰이는 요소이다. 또한, 발성율은 음성부호화 측면에서도 중요한 정보로 사용될 수 있다.

기존의 음성부호화기는 발성율에 관계없이 고정적인 분석구간을 정하여 전송율을 결정하고 있다. 따라서, 발성율을 미리 측정한다면, 발성율이 느린 부분에 각기 다른 부호화 방법을 적용하여 음질을 향상시킬 수도 있고, 전송율을 가변적으로 적용할 수도 있게 될 것이다.

3. 기존의 발성율 측정법

발성율을 측정하기 위해서는 일정한 시간동안의 음성신호 내에 몇 개의 음절이 포함되어 있는지를 추정하여야 한다. 기존에 제안된 LSP 파라미터법은 음소가 변화하게 되면 각 음소의 포먼트 주파수가 변화하게 되고, 그에 따라 스펙트럼 쌍의 위치가 변화하는 정도를 수치화 하여 발성율을 측정한다.

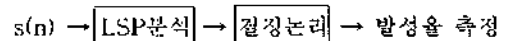


그림3-1 LSP분석법

에너지 측정법은 음성신호의 에너지 포락선을 사용하여 시간당 음성 전이구간 변화수를 발성율을 측정한다.

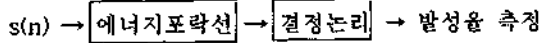


그림3-2 에너지포락선법

4. 제안한 측정법

본 논문에서 제안하는 발성율 측정법은 기존에 제안된 방법인 에너지 변화와 LSP 파라미터만을 동시에 사용한 하이브리드 방법이다.

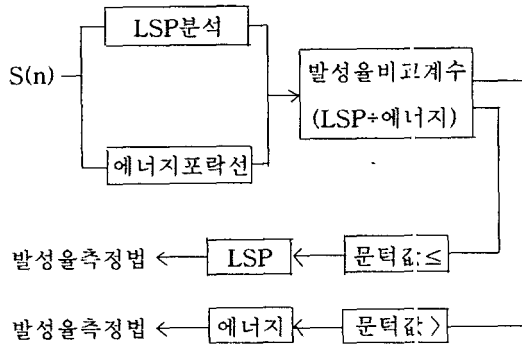


그림4-1 혼성측정법

그림 4-1에서와 같이 발성율 비교계수가 문턱값보다 크거나 같을 경우에는 LSP 분석법을, 작을 경우에는 에너지 포락선법을 이용한다.

본 논문에서 제안하는 문턱값은 1.5로 정하였다. 이는 음성신호가 모음과 자음의 반복구조 되어 있는데다가, 사람이 발성할 경우 허파를 이용해서 공기를 내뿜으면서 발성하기 때문에 지속적인 날성이 불가능하고 통계상 에너지포락선에 의해 측정된 발성율이 실제 발성율보다 1.5배이상 나올 수 없는 것으로 알려져 있기 때문이다.

5. 실험 및 결과

제안한 방법을 실험하기 위해서 IBM PC(400MHZ)에 마이크로 입력이 가능한 A/D변환기로 인터페이스 하였다. 음성자료는 5명이 연구실 환경에서 발성한 음성을 8KHZ로 표본화하고 16bit로 양자화 하여 사용하였다.

그림 5-1, 5-2에서(a)는 음성을 시간축에 대해서 본 파형이



그림 5-1. 작르게 발성한 시도에 대한 발성속도 / 아아아아오오우우이이 /
(a) 음성파형 (b) 에너지 변화도
(c) LSP 거리의 변화도 (d) 발성속도

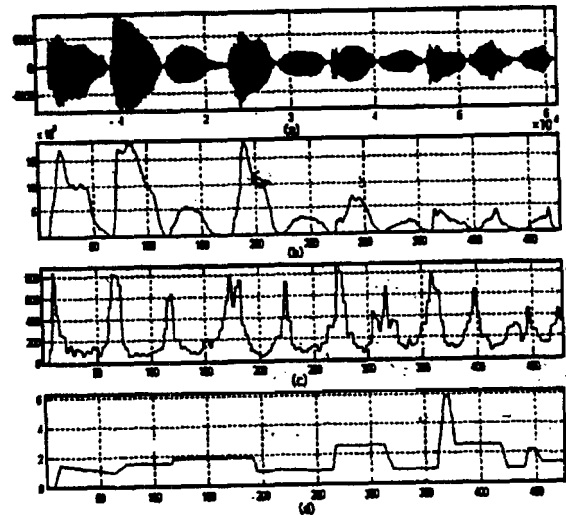


그림 5-2. 높리게 발성한 시도에 대한 발성속도 / 아아아아오오우우이이 /
(a) 음성파형 (b) 에너지 변화도
(c) LSP 거리의 변화도 (d) 발성속도

며, (b)는 파형에 대한 에너지를 나타낸다. (c)는 LSP거리의 변화도를 나타내며, (d)는 발성속도를 나타낸다. 빠른 발성의 경우에는 에너지 포락선에 의해 음절을 추정하는 것 보다는 LSP의 거리 변화도에 따라 음절을 추정하는 것이 보다 정확

하고, 느린 발성의 경우에는 LSP 거리변화도 보다는 에너지 포락선이 "0"에 근접할 때마다 음절이 존재한다고 추정하여 음절을 계산하는 것이 더 정확하다는 것을 확인할 수 있다.

6. 결론

표 1. 발성을 측정표

구분	발성률(음절/sec)			
	제안안측정법	LSP	에너지	수동측정
빠른발성	3.25	3.25	2.97	3.36
느린발성	1.50	1.05	1.49	1.35

표 1.에서 보는 바와 같이 빠른 발성의 경우는 LSP를 느린 발성의 경우에는 에너지를 이용해서 측정한 발성율이 실제 수동으로 측정한 발성율에 가까움을 알 수 있다. 따라서, 제안한 방법에서처럼 일정한 문턱값에 따라 발성을 측정방법을 달리 사용함으로써 보다 정확한 발성율의 측정이 가능하다는 것을 확인할 수 있었다.

참고문헌

1. 장경아, 배명진, "LSP 파라미터를 이용한 발성속도 측정에 관한 연구"에 대한 전자공학회, 추계종합학술대회, 논문집, Vol.23, No.2, 2000. 11. 25.
2. 장경아, 배명진, "CELP 보코더 전송을 감소를 위한 발성속도 측정법에 관한연구", 제3회 신호처리합동 학술대회 Vol.13, 2003. 9. 30.
3. 배명진, "디지털 음성부호화" pp.39-44, 동영출판사 1996. 11.
4. 한진수, 음성신호처리, 오성미디어, 2000