

## 화자검증에서 통로변동에 안정한 한가지 특징량추출방법

림현수, 주일혁

경애하는 최고령도자 김정은동지께서는 다음과 같이 말씀하시였다.

《현시대는 과학기술의 시대이며 과학기술의 발전수준은 나라의 종합적국력과 지위를 규정하는 징표로 됩니다.》

본문에서는 생체인증 특히 화자검증에서 통로차에 안정하면서도 높은 성능을 보장하는 한가지 특징량추출방법을 연구하였다.

화자검증체계가 실용적인것으로 되자면 통로차, 시기차, 잡음에 안정하면서 높은 성능을 보장하여야 한다.

멜주파수케프스트람결수(MFCC)는 화자검증에서 보편적으로 리용되는 특징량으로서 다음과 같이 정의된다.[2]

음성신호를  $x(t)$ , 그것의 푸리에변환을  $X(\omega)$ 라고 할 때 케프스트람결수는

$$Cep(\tau) = F^{-1}(\log |(X(\omega))|) \quad (1)$$

이다. 주파수영역에서 멜척도변환

$$Mel(f) = 2595 \log_{10} 10(1 + f/700.0)$$

를 진행하여 얻어지는 케프스트람결수를 멜주파수케프스트람결수라고 한다.

MFCC결수는 성도의 주파수특성을 나타내는 특징량으로서 여기에는 음성정보뿐만 아니라 통로특성, 잡음 등도 들어있으므로 이 특징량은 통로변동에 안정하지 못한 결함이 있다.

선행연구[1]에서는 케프스트람평균정규화(CMN)수법을 리용하여 통로변동을 제거한 특징량추출방법을 제기하였다.

선행연구[1]에서 제기한 MFCC+CMN방법은 단시간의 음성에 대해서는 케프스트람평균을 정확히 얻을수 없으므로 케프스트람평균정규화과정에 통로특성뿐만 아니라 개인의 음성정보도 함께 잃게 될수 있다는 결함이 있다. 또한 선행연구 [1]에서는 통로변동이 스펙트럼변동보다 상대적으로 작다는 전제를 가지고있으므로 통로변동이 큰 경우들에는 효과적이지 못하다.

본문에서는 선행연구[2]의 모형을 그대로 리용하면서 선행연구[2]에서 리용한 MFCC 특징량을 주파수대역에서 통로변동에 보다 안정한 국부정규화된 케프스트람결수(LNCC)로 바꾸어 화자검증체계를 구성하였다.

LNCC특징량추출과정은 다음과 같다.

① 음성신호를 프레임길이가 25ms이고 중첩도가 50%가 되도록 프레임들로 분할한다. 매 프레임에 대하여 리산푸리에변환을 진행하여 전력스펙트럼을 얻는다.

② 아래의 주파수응답을 가진 려파기군을 통과시킨다.

$$F(f) = \frac{Num(f)}{Den(f)} \quad (2)$$

$$Num(f) = \begin{cases} -\frac{2}{B}|f - f_i^c| + 1, & |f - f_i^c| \leq \frac{B}{2} \\ 0, & \text{기타} \end{cases} \quad (3)$$

$$Den(f) = \begin{cases} \frac{2}{B}(1 - d_{\min})|f - f_i^c| + d_{\min}, & |f - f_i^c| \leq \frac{B}{2} \\ 0, & \text{기타} \end{cases} \quad (4)$$

여기서  $f_i^c$ 는 주파수대역의 중심주파수,  $B$ 는 주파수대역의 너비이다. 논문에서는  $d_{\min} = 0.0001$ 로 주었다.

식 (3)은 특정한 주파수를 중심으로 하는 대역통과려파기와 같은 려파기로서 선행연구[2]에서와 같이 MFCC결수를 얻는 삼각려파기와 본질적으로 같다. 식 (4)는 주파수대역의 양끝점에서의 에너지를 나타내는 항으로서 일종의 정규화역할을 한다. 즉 식 (2)로 주어지는 려파기를 통과하여 통로에 의한 스펙트르기울어짐을 제거할수 있다.

식 (2)에 로그를 취하면 로그변환의 성질로부터 식 (2)로 표시되는 려파기의 주파수 응답의 로그식은 식 (3), (4)의 로그변환의 차로 표시되므로 통로에 의한 스펙트르기울어짐이 제거된다.

LNCC려파기군을 통과하여 얻어진 스펙트르기울어짐곡선을 MFCC려파기군을 통과하여 얻어진 스펙트르기울어짐곡선과 비교하여 보여주었다.(그림)

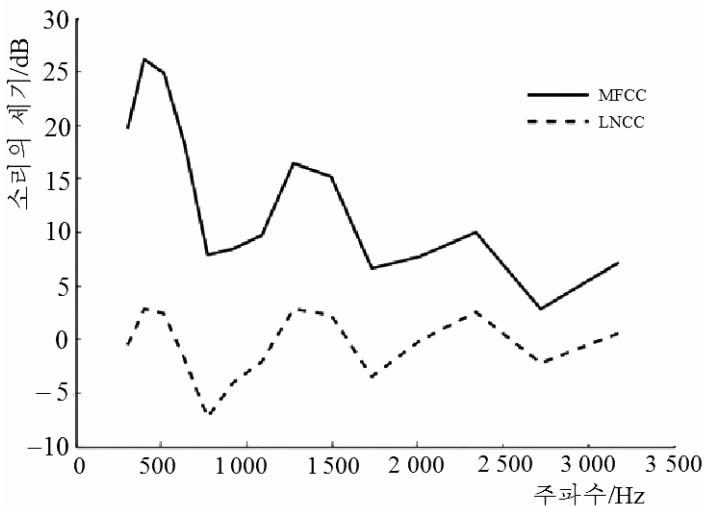


그림. LNCC려파기군을 통과하여 얻어진 스펙트르포락

을 평가하기 위하여 138명(남성 106명, 여성 32명)의 음성자료를 리용하여 자료기지를 구성하였다.

이 자료기지에 대하여 MFCC, MFCC+CMN, MFCC+RASTA특징량을 리용한 화자검증체제와 논문에서 연구한 LNCC, LNCC+CMN특징량을 리용한 화자검증체제의 성능을 비교하였다.

실험결과는 표와 같다.

그림에서 알수 있는바와 같이 LNCC려파기군을 통과하여 얻어진 스펙트르기울어짐곡선은 스펙트르피크점과 같이 스펙트르기울어짐의 모양은 보존하면서 통로변동에 의한 스펙트르기울어짐을 제거하였다는것을 알수 있다.

③ 얻어진 결수에 로그변환을 진행하고 DCT변환을 실시하여 LNCC결수를 얻는다.

LNCC특징량을 리용한 화자검증체제의 성능평가실험은 다음과 같다.

논문에서는 LNCC에 의한 본문의존형화자검증체제의 성능

표. 실험결과

음성자료	방 법				
	MFCC	MFCC+CMN	MFCC+RASTA	LNCC	LNCC+CMN
스펙트르기울어짐이 없는 음성	0.561	0.579	0.673	0.714	0.819
-3dB/octave	1.05	0.66	0.765	1.05	0.65
-6dB/octave	4.03	1.02	1.12	2.02	0.99
-9dB/octave	13.4	4.62	4.62	6.56	4.12

실험결과로부터 통로조건이 좋을 때에는 LNCC특징량을 리용하는것이 MFCC특징량을 리용할 때보다는 등오유률이 높지만 -6dB/octave, -9dB/octave의 스펙트르기울임조건에 서는 보다 낮은 등오유률이 나왔다. 이로부터 LNCC특징량을 리용하는것이 통로변동에 보다 안정한 특징량이라는것을 알수 있다.

## 참 고 문 헌

- [1] A. Y. Nakano et al.; IEICE Trans. Inform. Syst., 2451, 2010.
- [2] S. Tranter et al.; IEEE Trans. Audio Speech Language Process, 14, 1557, 2006.

주체108(2019)년 12월 15일 원고접수

## A Future Extraction Robust to Changes in Channels in Speaker Verification

*Rim Hyon Su, Ju Il Hyok*

In this paper, we study a future extraction that is robust to changes in channels in speaker verification. We refer these locally-normalized cepstral coefficients to LNCC.

And, we propose an algorithm to calculate the LNCCs.

Keyword: LNCC