

본문독립형발성자인증에서 인자분석법의 한가지 방법

한철진, 김덕호

위대한 령도자 김정일동지께서는 다음과 같이 교시하시였다.

《정보산업을 빨리 발전시키고 인민경제의 모든 부문을 정보화하여야 합니다.》

(《김정일선집》 증보판 제20권 380페이지)

지금까지 고전적인 MAP적응이 본문독립발성자인식에서 발성자모형화의 가장 보편적인 형태로 리용되었으나 등록자료가 충분하지 않은 경우 고유목소리에 의한 MAP적응이 훨씬 더 효율적이며 이로부터 공동인자분석법을 리용한 발성자모형화방법이 제안되었다.[1, 2] 공동인자분석은 2개의 공간 즉 발성자공간과 통로공간에서 정의되는 발성자인자와 통로인자에 기초하고있다. 그런데 통로인자들이 일반적으로 통로효과만이 아니라 발성자정보도 포함한다는것이 실험을 통하여 밝혀졌다. 이로부터 우리는 발성자변동공간과 통로변동공간을 구별하지 않는 새로운 형태의 인자분석방법을 제안한다.

1. 종합변동공간에 의한 인자분석

종합변동공간은 종합변동공분산행렬에서 가장 큰 고유값들을 가지는 고유벡토르들을 포함하는 변동행렬에 의해 정의된다. 종합변동공간에 의한 새로운 인자분석에서는 GMM초벡토르공간에서 발성자 및 통로효과들사이에 구별을 두지 않는다.

종합인자분석에서는 주어진 발성에 대하여 발성자 및 통로의존GMM초벡토르를 다음과 같이 정의한다.

$$M = m + Tw \quad (1)$$

여기서 m 은 발성자 및 통로독립초벡토르(UBM초벡토르), T 는 변동행렬(저위수행렬), w 는 $N(0, I)$ 에 따르는 우연벡토르이다. 즉 M 이 평균 m 과 공분산행렬 TT^T 를 가지는 정규분포에 따른다고 가정하며 w 의 성분들이 종합인자로 된다.

주어진 발성에 대한 바움-웰프통계량으로부터 종합인자 w 는 다음과 같이 계산된다.

$$w = (I + T^T \Sigma^{-1} N(u) T)^{-1} T^T \Sigma^{-1} \tilde{F}(u) \quad (2)$$

여기서 $N(u)$, $\tilde{F}(u)$ 은 주어진 발성에 대하여 계산한 바움-웰프통계량, Σ 는 종합변동행렬 T 에 포착되지 않은 오차변동을 모형화하는 공분산행렬이다.

2. 점수계산 및 정규화방법

1) 점수계산

앞에서 본것처럼 종합인자분석은 CF차원 GMM초벡토르를 저차원공간으로 넘기는 넘기기처럼 보인다. 그러므로 발성자모형 그자체는 저차원의 벡토르로 표현되며 따라서 단순한 거리척도를 리용하여 점수를 계산할수 있다. 즉 등록발성자료로부터 추출한 종합인

자와 인증발성자료로부터 추출한 종합인자사이의 거리를 리용하여 점수를 계산할수 있다.

이렇게 되면 점수계산이 다른 방법들에 비해 속도가 빨라지고 복잡성이 덜어진다. 논문에서는 이러한 거리척도로 코시누스거리를 리용한다. 즉 점수는 다음과 같이 계산된다.

$$\text{score}(w_1, w_2) = \frac{\langle w_1, w_2 \rangle}{\|w_1\| \|w_2\|} \quad (3)$$

여기서 w_1, w_2 는 각각 등록발성과 인증발성으로부터 추출한 종합인자이다.

코시누스거리를 리용하면 두 벡토르사이의 진폭을 고려하지 않고 각만을 리용하므로 시기차나 통로영향과 같은 비발성자정보의 영향을 줄일수 있다.

2) 시기차변동상쇄

공동인자분석에서는 GMM초벡토르공간에서 통로효과를 상쇄한다. 그러나 종합인자분석에서는 통로효과의 상쇄를 공동인자분석과는 달리 종합변동공간 즉 저차원공간에서 진행함으로써 계산비용을 줄일수 있게 한다.

우리는 통로효과의 상쇄를 위해 공동인자분석의 발성자인자공간에 적용되어 효과성이 나타난 클라스내공분산정규화(WCCN: Within Class Covariance Normalization)방법을 리용하였다.

코시누스거리에 WCCN알고리즘을 적용하면 다음과 같은 식이 얻어진다.

$$k(w_1, w_2) = \frac{(B^T w_1)^T (B^T w_2)}{\sqrt{(B^T w_1)^T (B^T w_1)} \sqrt{(B^T w_2)^T (B^T w_2)}} \quad (4)$$

여기서 W 는 학습환경에서 모든 침입자들에 대하여 계산된 클라스내공분산행렬이며 행렬 B 는 행렬 $W^{-1} = BB^T$ 의 콜스키(Cholesky)분해로 얻어진다.

우리는 이 값을 턱값과 비교하여 발성자인증을 진행하였다.

3. 평 가 실 험

실험에서는 특징벡토르를 추출하기 위하여 표본화주파수 22.05kHz, 양자화비트 16bit, 창문너비 20ms, 프레임간격 10ms를 리용하였으며 그로부터 에네르기를 포함한 20차원 MFCC 파라미터를 추출하고 그것의 Δ, Δ^2 도 포함하여 60차원 특징벡토르를 추출하였다.

실험에서 리용된 코퍼스는 남자 526명의 발성자료 8 049개로 이루어졌다. 그리고 UBM 모형의 혼합수는 512개로 설정하였다.

실험에서 비교체제로 리용된 공동인자분석체제의 인자수는 성능이 제일 높은 발성자인자 300개, 통로인자 50개로 설정하였으며 논문에서 제안한 종합인자분석의 성능평가를 위하여 종합인자수를 200, 300, 400개로 변경시키면서 실험을 진행하였다.(표)

표에서 보는것처럼 종합인자분석을 리용한 경우 GMM-UBM이나 공동인자분석을 리용한 경우에 비하여 등오유률이 훨씬 개선되었다는것을 알수 있다.

표. 종합인자분석체제의 성능평가

체제	등오유률/%
GMM-UBM	2.251
공동인자분석	1.826
종합인자분석(200)	1.48
종합인자분석(300)	1.201
종합인자분석(400)	1.27

맺는말

본문독립형발성자인증의 성능개선을 위한 새로운 인자분석법을 제안하고 실험을 통하여 그 효과성을 검증하였다.

실험결과 기존의 체계에 비하여 등오유률이 상대적으로 각각 46.6, 34.2% 개선되었다.

참고문헌

[1] P. Kenny et al.; ICASSP, 5, 37, 2004.

[2] X. Zhao et al.; IEEE Trans. Audio. Speech. Lang. Process., 20, 3, 1032, 2012.

주체105(2016)년 2월 5일 원고접수

A Factor Analysis Method in Text-Independent Type Speaker Verification

Han Chol Jin, Kim Tok Ho

We propose a new type of factor analysis method that doesn't separate speaker variability space and channel variability space.

We achieved a relative improvement of 34.2% on the EER as compared to classical joint factor analysis by introducing a factor analysis method using total variability space.

Key words: speaker verification, factor analysis, total factor analysis