

프레임과 가우스분포의 동시선택에 의한 본문 독립발성자인증속도 개선에 대한 연구

김덕호, 한철진

GMM-UBM체계는 본문독립형발성자인식에서 가장 널리 쓰이는 방법으로서 점수계산량이 많아지는 결함이 있는데 이것을 극복하기 위한 여러가지 방법들이 제안되었다.[1-4] 최근에는 정렬된 가우스혼합모형(SGMM)이라고 부르는 새로운 가우스분포선택방법이 제안되었는데[5] 이 방법에서는 UBM(통합배경모형)평균벡터의 스칼라점수화를 리용하여 모든 가우스분포들에 대한 탐색을 하지 않고 UBM으로부터 우세한 혼합성분들을 찾아 점수계산에 리용한다.

논문에서는 이 방법을 2차원점수화의 리용 즉 프레임과 가우스분포의 동시선택으로 확장하여 그 성능을 보다 개선할수 있게 하는 새로운 방법을 논의한다.

이를 위해 다음과 같은 2개의 정렬함수를 리용하여 정렬된 가우스혼합모형의 개념을 2차원으로 확장한다.

$$\begin{aligned} s_1^t &= f_1(\mathbf{x}_t) = \langle \mathbf{a}, \mathbf{x}_t \rangle \\ s_2^t &= f_2(\mathbf{x}_t) = \langle \mathbf{b}, \mathbf{x}_t \rangle \end{aligned} \quad (1)$$

여기서 \mathbf{x}_t 는 t 번째 특징벡터이며 \mathbf{a} 와 \mathbf{b} 는 각각 무게벡터로서 직교에 가깝도록 설계되고 립자무리최량화(PSO: Particle Swarm Optimization)알고리즘을 리용하여 최량화된다.[6]

정렬된 가우스혼합모형에서는 대응하는 정렬값들이 중심혼합의 정렬값들에 관하여 취한 립점값들의 모임에 놓이는 혼합들에 대하여서만 평가에서 논의되도록 한다. 따라서 2차원에서 대응하는 정렬값들이 $|S_1^{UBM} - s_1^i| < \alpha R_1$ 과 $|S_2^{UBM} - s_2^i| < \alpha R_2$ 에 의하여 지정된 4각형 근방에 존재하는 혼합들을 고려하여 평가되는 혼합들을 결정할수 있다. 여기서 α 는 근방을 지적하기 위한 조종과라메터이다. 그리고 S_1^{UBM} 과 S_2^{UBM} 은 $f_1(\cdot)$ 과 $f_2(\cdot)$ 에 따르는 UBM평균들의 정렬값이고 R_1, R_2 는 대응한 S_1^{UBM} 과 S_2^{UBM} 의 범위들이며 s_1^i, s_2^i 는 미지 입력특징벡터정렬값 s_1^i 와 s_2^i 가 S_1^{UBM} 과 S_2^{UBM} 에 의하여 스칼라량자화된 값들이다.

$f_1(\cdot)$ 과 $f_2(\cdot)$ 에 의하여 지정되는 탐색구역의 교차구역에 위치한 혼합성분들은 점수평가에 참가하며 만일 이 영역에 어떤 혼합성분도 없다면(이런 경우는 α 가 작은 값일 때 나타나게 된다.) 그 특징벡터는 가우스분포평가에서 제외된다.

새로운 정렬된 가우스혼합모형에 요구되는 기억용량은 GMM-UBM의 경우에 비해 $(2D+3)/(2D+1)$ 배이다. 여기서 D 는 파라메터차원수이다. 일부 프레임들이 점수평가에 참가하지 않는 경우가 있는것으로 하여 속도갱신률은 검사토막들에서 서로 다르다. 그러므로 전체 검사토막에 대한 평균속도갱신률을 계산하여 종전의 정렬된 가우스혼합모형방법과 비교분석하여야 한다.

앞에서 언급한것처럼 새로운 2차원정렬된 가우스혼합모형에 대하여 PSO는 \mathbf{a} 와 \mathbf{b} 로 표시되는 두 벡토르들을 최량화한다.

선행연구[5]에서 정의된 적합도함수와 2차원정렬된 가우스혼합모형의 정의를 고려하여 새로운 적합도함수를 다음과 같이 정의한다.

$$F'(\mathbf{a}, \mathbf{b}) = F(\mathbf{a}) + F(\mathbf{b}) - |(\mathbf{a} \cdot \mathbf{b}' / |\mathbf{a}| |\mathbf{b}'|)| \quad (2)$$

여기서

$$F(\cdot) = \sum_{t=1}^D \frac{E\{(\mathbf{x}'_t - E(\mathbf{x}'_t))(s' - E(s'))\}}{\sqrt{E\{(\mathbf{x}'_t - E(\mathbf{x}'_t))^2\}E\{(s' - E(s'))^2\}}} \quad (3)$$

는 선행연구[5]에서 정의된 적합도함수이며 마지막항은 무게벡토르사이의 각의 코시누스 값의 절대값으로서 방향이 반대되는($\pi/2$ 와 $3\pi/2$ 사이의 각) 벡토르들을 고려한다는것을 의미한다.

우리는 제안한 방법의 효과성검증을 위한 실험을 다음과 같은 조건에서 진행하였다.

우선 1인당 1개의 등록음성과 3개의 검사음성으로 이루어진 500명분의 음성자료를 리용하였다. 매 등록음성은 30~60s이고 검사음성은 3~5s이며 모든 음성파형은 22.05kHz, 16bit, Mono형식으로 준비하였다.

다음 특징파라미터를 추출하기 위한 창문너비는 25.6ms, 프레임간격은 10ms로 하고 26차원MFCC를 특징파라미터로 리용하였다. 배경모형의 가우스혼합수는 1 024로 하였다.

여기에 기초하여 진행한 실험결과는 다음과 같다.

등오유률이 2.1이고 2차원정렬된 가우스혼합모형의 조종파라미터 α 가 각각 2, 3, 5, 10, 15, 20%일 때 평균적인 속도갱신률은 각각 157:1, 85:1, 37:1, 11:1, 5:1, 3:1이었다.

실험결과는 새로운 정렬된 가우스혼합모형을 리용한 경우 발성자 인식속도가 보다 개선되었다는것을 보여준다.

맺는 말

정렬된 가우스혼합모형의 개념을 2차원으로 확장하여 발성자 인식속도를 보다 개선할 수 있는 방법을 제안하고 실험을 통하여 그 효과성을 검증하였다.

참고 문헌

- [1] H. R. S. Mohammadi et al.; In Proc. ICASSP'07, 4, 309, 2007.
- [2] M. Roch; Speech Commun., 48, 85, 2006.
- [3] B. Xiang et al.; IEEE Trans. Audio, Speech and Lang. Process., 11, 5, 447, 2003.
- [4] Z. Xiong et al.; Speech Commun., 48, 1273, 2006.
- [5] T. Kinnunen et al.; IEEE Trans. Audio, Speech and Lang. Process., 14, 1, 277, 2006.
- [6] H. Aronowitz et al.; IEEE Trans. Audio, Speech and Lang. Process., Special Issue on Speaker and Language Recognition, 15, 7, 2033, 2007.

Improvement of Text-Independent Speaker Verification Speed by Jointly Selecting of Frame and Gaussain

Kim Tok Ho, Han Chol Jin

We proposed the improvement method of text-independent speaker verification speed using 2nd order Sorted Gaussian Mixture Model(SGMM-2).

We introduced the concept of 2nd order Sorted Gaussian Mixture Model, defined fitness function for weight optimization, and verified the effectiveness of the proposed method empirically.

Key words: sorted Gaussian mixture model(SGMM), fitness function, speaker verification