

사용자 적응을 통한 한국 수화 인식 시스템의 개선

Improvement of Korean Sign Language Recognition System by User Adaptation

정성훈*, 박광현**, 변증남***

Seong-Hoon Jung*, Kwang-Hyun Park**, Zeungnam Bien***

Abstract - This paper presents user adaptation methods to overcome limitations of a user-independent model and a user-dependent model in a Korean sign language recognition system. To adapt model parameters for unobserved states in hidden Markov models, we introduce new methods based on motion similarity and prediction from adaptation history so that we can achieve faster adaption and higher recognition rates comparing with previous methods.

Key Words : 한국 수화 인식, 사용자 적응, 동작 유사도, 적응 추이

1. 서론

최근 들어 손 제스처를 인간-기계 인터페이스에 활용하는 연구가 많아지고 있다. 그 중에서도 수화 인식은 청각장애인과 일반인 사이의 원활한 의사소통을 돕는 인터페이스로서 중요성이 날로 더해가고 있다. 하지만 기존의 수화 인식 연구는 한 명의 사용자만을 대상으로 하거나 다수 사용자를 위한 모델을 사용하기 때문에 사용자 개개인의 수화 동작의 차이를 고려하지 않았다[1, 2]. 전자의 경우에는 한 명의 사용자의 데이터에 기반한 것으로 이로부터 얻어진 모델을 수화자 종속 모델이라고 하고 후자의 경우에는 다수의 사용자의 데이터에 기반한 것으로 이로부터 얻어진 모델을 수화자 독립 모델이라고 한다. 현실적으로 한 명의 사용자로부터 충분한 양의 학습 데이터를 얻는 것은 힘들기 때문에 실제 상황에서 수화자 종속 모델을 사용하는 것은 어렵다. 다수의 사용자로부터 학습 데이터를 얻고 이로부터 확률에 기반한 모델을 생성하는 수화자 독립 모델은 구축이 용이한 반면 수화 동작의 개인차로 인해 인식률이 상대적으로 낮다. 이러한 한계를 극복하기 위해서는 수화자 독립 모델을 기반으로 특정한 사용자의 동작 특성을 반영할 수 있도록 모델을 변형시키는 사용자 적응이 필요하다.

수화 인식에서 은닉 마르코프 모델을 사용한 방법이 인식률이 높은 것으로 알려져 있는데, 사용자 적응을 위해 MLLR (Maximum Likelihood Linear Regression) 혹은 MAP (Maximum A Posteriori) 방법을 사용할 수 있다[3, 4]. 하지만, 사용자 적응을 위해 입력되는 데이터가 문장 단위로 이루어지는데 하나의 문장에 포함되는 단어의 수는 한정되기 때

문에 은닉 마르코프 모델에 관측되지 않은 상태가 존재하게 된다. 이 경우 비관측된 상태의 모델 파라미터를 수정할 수 없어서 전체 인식률이 떨어뜨리게 되므로 비관측된 상태의 모델 파라미터도 같이 수정할 수 있는 방법이 필요하다.

본 논문에서는 수화자 독립 모델이 주어질 때 시간의 경과에 따라 제공되는 사용자의 수화 문장을 사용하여 점진적으로 적응하는 것을 다루며, 비관측된 상태의 모델 파라미터를 같이 적응시키는 방법을 제안한다.

전체 논문의 구성은 다음과 같다. 먼저 2절에서는 기존의 사용자 적응 방법의 한계를 지적하고 3절에서 이를 해결하기 위한 새로운 적응 방법을 제안한다. 4절에서는 실제 청각장애인들로부터 획득한 수화 데이터베이스를 사용하여 실험한 결과를 보이고, 5절에서 결론을 맺는다.

2. 기존의 사용자 적응 방법과 한계

기존의 사용자 적응 방법은 주로 음성인식 분야에서 활발히 연구되었는데, 특히 MLLR과 MAP 방법이 널리 사용되어 왔다[3, 4]. 비관측된 상태를 적응시키기 위해 고안된 방법들은 모델 파라미터 공간에서 거리가 가까운 상태들이 유사한 적응 특성을 보인다는 가정을 바탕으로 한다. 하지만, 수화 인식에는 이러한 가정이 맞지 않다는 것을 관찰하였다.

관측된 상태들만 적응시키는 경우에는 사용자의 특성이 부분적으로만 반영이 되고, 모델들 간의 구분도가 나빠져 인식률이 떨어질 수 있다. 그림 1은 MAP을 이용하여 관측된 상태들에 대해서만 적응을 수행할 때 인식률의 변화를 보인다. (GDM: 수화자 종속 모델, GIM: 수화자 독립 모델)

본 논문에서는 이와 같이 사용자 적응 과정에서 발생하는 비관측된 상태의 문제를 다루고, 이를 위해 다음 절에서 동작 유사도와 변화 추이를 이용한 추정 방법에 대해 서술한다.

저자 소개

* 정성훈: 한국과학기술연구원 인지로봇연구단 연구원

** 박광현: 한국과학기술원 전자전산학과 조빙교수

*** 변증남: 한국과학기술원 전자전산학과 교수

본 연구는 과학기술부/한국과학재단 우수연구센터육성사업의 지원으로 수행되었음 (R11-1999-008)

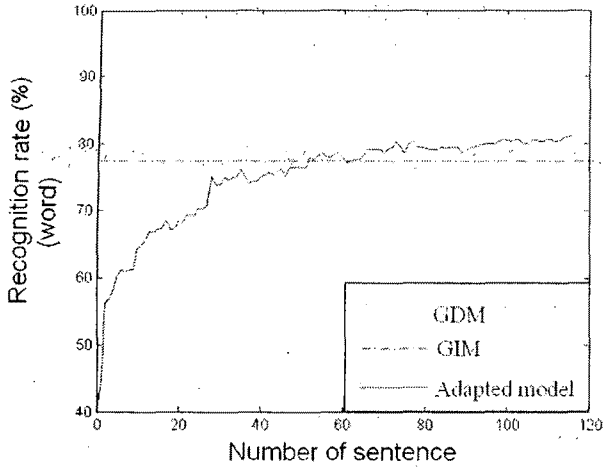


그림 1. 관측된 상태에 대해서만 적응을 수행한 경우

3. 제안하는 사용자 적응 방법

3.1 동작 유사도에 의한 추정

수화에서 나타나는 손 운동은 직선과 원 운동과 같은 기본 운동의 조합으로 표현된다. 실제 청각 장애인들의 수화 동작을 관찰한 결과 동작이 유사한 수화의 경우 적응되는 특성도 유사하다는 것을 발견하였다. 다시 말해 동작이 유사한 경우 해당 동작의 길이와 속도 변화의 경향이 유사하다는 사실을 관찰하였다. 여기서 동작은 오른손의 움직임 궤적으로 표현하였다. 왼손은 오른손의 보조 수단으로 사용되고 대상 단어에서 왼손만 단독으로 수화를 표현하는 경우는 없었기 때문에 제외하였다. 이러한 관찰을 바탕으로 관측되지 않은 상태의 파라미터는 관측된 상태 중에서 동작 유사도가 높은 상태들에 대한 데이터를 사용하여 적응하였다. 수화 동작들 사이의 유사도는 Dynamic Time Warping (DTW) 방법을 이용하여 식 (1)과 같이 측정하였다[5].

$$Sim(S, S^*) = \frac{1}{1 + \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N DTW(A_i, A_j)} \quad (1)$$

모든 상태들 간의 동작 유사도를 계산한 후 특정 상태 S가 현재 데이터에서 관측되지 않는 경우, S와의 동작 유사도가 특정 문턱 값을 넘는 관측된 상태 중에서 가장 유사도가 높은 S*를 찾고 S*에 대한 적응 데이터를 사용하여 S의 파라미터를 적응시킨다. 실험을 통해 동작 유사도가 높은 상태들은 유사한 적응 방향을 가진다는 것을 확인하였다.

3.2 변화 추이를 이용한 추정

어떤 특성이 변화하는 추이를 관찰하면 앞으로의 변화에 대해 추정을 할 수 있다. 사용자 적응 과정에서 각각의 상태들의 적응 추이를 관찰할 수 있으며 비관측된 상태에 대해서는 추정을 이용하여 모델 파라미터를 적응시킬 수 있다.

관찰을 통해 모델 파라미터의 변화 추이는 대체적으로 갑작스러운 변화가 없이 완만하게 변화하는 특성을 보였다. 이처럼 일관되게 변화하는 경우에는 2차 랜덤 마르코프 과정으로

로 변화 특성을 모델링할 수 있다. 즉, 상태 벡터 x 를 추정하기 위한 2차 동적 방정식은 식 (2)와 같이 표현되며, A_0 와 A_1 은 식 (3)을 통해 구할 수 있다.

$$\begin{bmatrix} x_n - \bar{x} \\ x_{n+1} - \bar{x} \end{bmatrix} = \begin{bmatrix} 0 & I \\ A_0 & A_1 \end{bmatrix} \begin{bmatrix} x_{n-1} - \bar{x} \\ x_n - \bar{x} \end{bmatrix} \quad (2)$$

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$$

$$S_{20} - A_0 S_{00} - A_1 S_{10} = 0 \quad (3)$$

$$S_{21} - A_0 S_{01} - A_1 S_{11} = 0$$

$$S_{ij} = \sum_{n=1}^{K-2} x_{n-1+i} x_{n-1+j}^T, \quad i, j = 0, 1, 2$$

파라미터의 변화 추이가 일관된 경우에는 그림 2와 같이 추정의 정확도가 높은 반면, 변화 추이가 일관되지 않은 경우에는 그림 3과 같이 오차를 수반하여 추정의 정확도가 낮아진다. 이와 같이 오차가 크게 발생하는 경우에는 모델 파라미터의 적응이 부정확하게 이루어질 수 있다. 이를 위해 추정에 대한 신뢰도를 측정하였다.

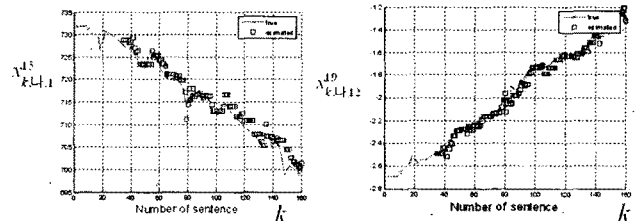


그림 2. 변화 추이가 일관된 경우

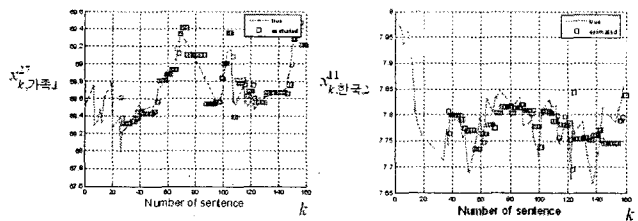


그림 3. 변화 추이가 일관되지 않은 경우

현재의 추정에 대한 신뢰도는 가장 최근에 수행한 M번의 추정의 정확도에 근거하여 결정하였다. 식 (4)와 같이 가장 최근의 추정에 대한 평균 오차를 구한 후, 시간 k에서의 추정 신뢰도는 식 (5)과 같이 지수 함수로 표현하였다. (w: 단어 인덱스, i: 모델의 상태 인덱스, j: 벡터의 요소 인덱스)

$$e_{k,w,i}^j = \frac{\sum_{m=1}^M (x_{k-m,w,i}^{j,true} - x_{k-m,w,i}^{j,predicted})}{M} \quad (4)$$

$$\beta_{k,w,i}^j = e^{-\rho |e_{k,w,i}^j|} \quad (5)$$

감쇄 파라미터 ρ_j 를 결정하기 위해 모든 데이터에 대한 추정을 수행한 후 발생한 오차를 계산하였다. 오차의 분포가 가우시안이라는 가정 하에 10%의 특이값을 제외한 최대 오차 e_j^* 를 구하고, 추정 오차가 e_j^* 가 되었을 때의 신뢰도가 0.1이 되도록 ρ_j 를 결정하였다.

3.3 제안한 방법의 결합

비관측된 상태의 모델 파라미터를 수정하기 위해 제안한 두 가지 방법은 각각 장단점을 가지고 있다. 동작 유사도에 기반한 방법은 데이터가 적은 적응 초기에도 이용할 수 있다는 장점이 있지만, 동작 유사도가 높은 상태가 모두 관측되지 않아 이를 이용할 수 없는 경우가 있다. 적응 추이에 기반한 방법은 일정한 시간이 지나면 항상 사용할 수 있지만 적응 추이가 일관되지 않은 경우에는 신뢰도가 낮아진다는 단점이 있다. 따라서 각각의 방법을 단독으로 사용하기 보다는 상호 보완적으로 두 가지 방법을 결합한 형태로 사용하는 것이 필요하다. 이를 고려한 방법의 절차는 다음과 같다.

단계 1. 모든 관측된 상태의 모델 파라미터를 MAP을 이용하여 적응시킨다.

단계 2. 모든 비관측된 상태에 대해

단계 2-1. 동작 유사도에 의한 추정치 $x_{k,w,i}^{j,motion}$ 를 구한다.

단계 2-2. 적응 추이에 의한 추정치 $x_{k,w,i}^{j,history}$ 를 구한다.

단계 2-3. 적응 추이에 의한 추정치의 신뢰도를 계산한 후 이를 이용하여 가중합을 구한다.

$$x_{k,w,i}^{j,estimation} = (1 - \beta_{k,w,i}^j)x_{k,w,i}^{j,motion} + \beta_{k,w,i}^j x_{k,w,i}^{j,history}$$

단계 3. 계산한 가중합을 이용하여 비관측된 상태의 모델 파라미터를 수정한다.

$$x_{k+1,w,i}^j = \alpha x_{k,w,i}^j + (1 - \alpha)x_{k,w,i}^{j,estimation}$$

이 때, α 는 실험을 통해 인식률의 증가가 최대가 되는 0.65로 설정하였다.

4. 실험 결과

실제 청각장애인을 대상으로 구축한 수화 데이터베이스에 대해 실험을 수행하였다. 70%를 적응 데이터로 사용하고 나머지 30%를 테스트 데이터로 사용하는 실험을 30번 반복하여 평균을 구하였다. 관측된 상태의 파라미터만 적응시키는 MAP 방법과 본 논문에서 제안한 방법을 비교하였다.

그림 4는 사용자 1에 대한 적응의 결과를 보인다. 적응 방법의 우수성은 독립 모델의 인식률로부터 얼마나 빠르게 종속 모델의 인식률까지 도달할 수 있는지로 이해할 수 있다. 비관측된 상태를 다루지 않는 MAP의 경우에는 초반에 인식률이 독립 모델보다 오히려 떨어지고 인식률의 증가도 느리다. 제안한 방법은 비관측된 상태의 모델 파라미터를 적응시키기 때문에 인식률의 증가가 MAP에 비해 빠르다는 사실을 알 수 있다.

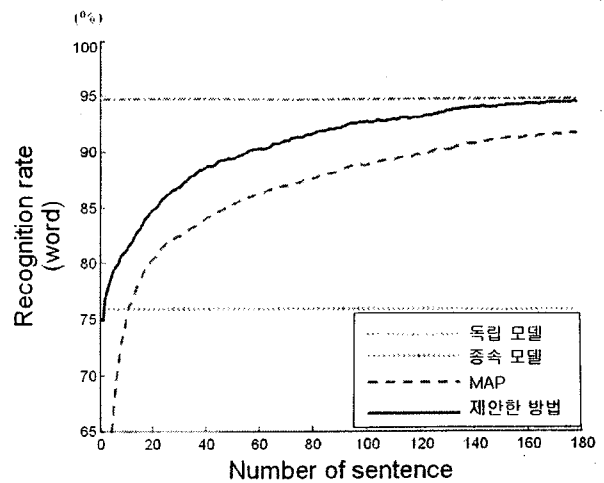


그림 4. 사용자 1에 대한 적응 결과

5. 결론 및 추후 과제

본 논문에서는 기존의 수화 인식에서 다루지 않았던 사용자 적응 문제를 다루었다. 점진적으로 이루어지는 사용자 적응에서 발생하는 비관측된 상태를 효과적으로 적응시키기 위해 동작 유사도와 적응 추이에 기반한 방법을 제안하였고 이들을 결합한 형태 또한 제안하였다. 실제 청각장애인들로부터 획득한 수화 데이터베이스를 사용하여 비관측된 상태를 다루지 않는 기존의 방법보다 제안한 방법이 인식률을 빠르게 향상시키는 것을 보였다. 제안한 방법을 대용량의 수화 데이터베이스에 적용하고 결과를 분석하는 것이 앞으로 필요하다.

참 고 문 헌

- [1] Starner, T., Petland, A., "Real-Time American Sign Language Recognition Using Desk and Wearable Computer Based Video," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, no. 12, pp. 1371-1375, December 1998.
- [2] Kim, J.-B., A Study on the Continuous Hand Gesture Recognition System for the Korean Sign Language, Ph.D. Thesis, KAIST, 2004.
- [3] Gauvain, J., Lee, C.H., Maximum a Posteriori Estimation for Multivariate Gaussian Mixture Observation of Markov Chains," IEEE Transactions on Speech and Audio, vol. 2, pp. 291-298, April 2004.
- [4] Leggetter, C., Woodland, P., "Maximum Likelihood Linear Regression for Speaker Adaptation of Continuous Density HMMs," Proceedings of Communication and Technology, vol. 2. pp. 171-185, 1995.
- [5] Keogh, E.J., Pazzani, M.J., "Derivative Dynamic Time Warping," Proceedings of the First International Conference on Data Mining, 2001.