

개선된 2차원 필기 인식 모델을 이용한 3차원 온라인 필기인식

김대환⁰, 이택헌, 김진형
한국과학기술원 전자전산학과
(kimdh⁰, three, jkim)⁰@ai.kaist.ac.kr

3D Online Handwriting Character Recognition with Modified 2D Handwriting Recognition Model

Dae Hwan Kim⁰, Taik Heon Rhee, Jin-Hyung Kim
Dept. of Electrical Engineering & Computer Science, KAIST

요 약

본 연구에서는 3차원 온라인 필기의 효과적인 인식 방법을 제안한다. 3차원 필기 시 pen-up/pen-down 정보의 구분이 없이 입력하도록 하여 사용자가 편리하게 필기하도록 하고 구분의 부정확함으로 인해 발생하는 오류를 줄인다. 또한, 기존의 2차원 필기 인식 모델을 개선하여 3차원 필기 데이터의 특성을 반영하게 함으로써 경제적이며 안정적인 인식이 가능하다. 실험 결과 제안된 인식 방법을 통해 pen-up/pen-down 정보의 구분이 없는 3차원 필기 숫자에 대해 91.6%의 인식 성능을 얻었으며, 특히 인식 모델의 개선을 통해 여러 획으로 이루어진 글자의 경우 높은 인식 성능의 향상을 보임을 확인하였다.

1. 서 론

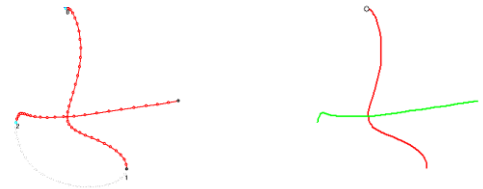
3차원 온라인 필기 인식이란, 3차원 입력장치를 입력 수단으로 하여 공간에 필기한 문자를 인식하는 것을 말한다. 펜 등을 이용한 2차원 필기 입력 방식과는 달리 3차원 필기 입력 방식은 입력 시 공간상의 제약이 없어 최근 활발히 연구가 진행되고 있다.

이 입력 시스템에 사용되는 3차원 입력장치란 센서를 내장하여 센서 신호로부터 3차원 공간에서 움직인 궤적을 복원시켜주는 장치를 일컫는다. 초기에는 주로 3축 가속도 센서를 이용하여 개발[1][2]되어 왔으나, 최근에는 센서의 오차를 효과적으로 관리하기 위해 3축 가속도 센서와 3축 각속도 센서를 함께 활용[3][4][5]하는 방향으로 연구가 진행되고 있다.

3차원 입력장치로부터 들어온 필기 문자를 인식하는데에는 다음과 같은 어려움이 따른다.

첫째, 일반 사용자들이 3차원 필기 시 pen-up/pen-down 상태를 정확하게 구분해주기 어렵다. Pen-up/pen-down이란 펜을 이용한 필기 인식에서 각각 펜을 들고 움직이는 부분, 펜을 바닥에 대고 글자를 쓰는 부분을 의미한다. 이는 펜 디바이스 등의 2차원 필기 시에는 자연스럽게, 3차원 입력장치를 이용한 경우에는 사용자가 버튼과 같은 별도의 장치를 통해 구분해주어야 한다. 따라서 익숙하지 못한 사용자는 필기 입력이 부자연스럽고, 의도하는 것과 상이한 궤적을 그리게 된다. 그림 1은 익숙하지 못한 사용자가 숫자 '5'의 필기 시 버튼을 사용함으로써 인해 발생하는 문제점을 보여준다. 사용자는 그림 1(a)와 같이 '5'의 궤적을 그려내려 하였으나, '5'의 아래쪽 곡선을 그리면서 버튼을 일찍 떼어 곡선의

일부가 pen-up 상태로 입력이 되었다. 그로 인해 그림 1(b)와 같이 pen-down 궤적만 보면 '5'의 형태와 많이 상이함을 알 수 있다.



(a)pen-up/pen-down 모든 궤적 (b) pen-down 궤적만
그림 1. 부정확한 pen-up/pen-down의 예(숫자 '5')

둘째, 많은 양의 3차원 필기 데이터를 수집하기 어렵다. 아직은 많은 사용자들에게 3차원 필기가 익숙하지 않고, 필기 시 공간 상에서 팔을 움직이므로 힘이 많이 들기 때문에 인식 모델의 훈련을 위한 충분한 양의 데이터를 모으는데 2차원 필기에 비해 시간과 비용이 많이 든다. 따라서, 직접 3차원 필기 데이터를 수집하는 것에 비해, 다량의 2차원 필기 데이터로 훈련된 안정된 모델에 3차원 필기의 특징을 반영하는 것이 보다 경제적이다.

본 논문에서는 효과적인 3차원 필기 문자 인식을 위해 두 가지 방법을 제안하였다.

첫째, 한 문자가 입력되는 동안 중간의 pen-up/pen-down 구분을 두지 않음으로써 사용자가 편리하게 필기 입력을 하도록 하였다. 2차원 필기는 pen-down 상태만으로 인식하는 것이 가능하지만 3차원 필기에서는 그림 1의 예와 같이 pen-up/pen-down의 부정확성으로 인해 두 궤적을 동시에 보아야만 인식할 수 있는 경우가 많다.

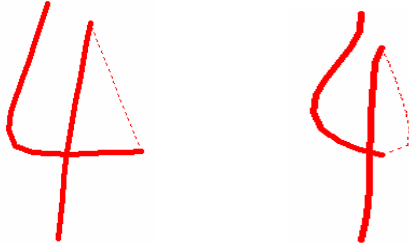
따라서 3차원 필기인식에서는 pen-up/pen-down을 구분하지 않고 이용하는 것이 인식에 도움이 된다고 볼 수 있다.

둘째, 기존에 훈련된 2차원 필기 인식 모델에 3차원 필기의 특징을 반영하도록 수정하여 효과적인 3차원 필기 인식이 가능토록 하였다. Pen-up/pen-down이 부정확하거나 또는 그 구분이 없는 입력의 경우 2차원 필기 인식 모델 내에서 이를 보완하는 과정이 필요하다.

본 논문의 구성은 다음과 같다. 먼저 2장에서는 3차원 필기의 특징을 살펴본다. 3장에서는 기존 인식 모델에 대한 설명과 함께 개선된 인식 시스템에서의 필기 인식 과정을 설명한다. 4장에서는 제안 방법을 이용한 실험 결과를 분석하며, 마지막으로 5장에서 결론 및 향후 과제를 논의하고자 한다.

2. 3차원 필기의 특징

3차원 필기는 2차원 필기에 비해 pen-up/pen-down 상태를 정확하게 구분하기 어렵지만, 3차원 입력장치를 통해 공간 상에서 연속적인 입력을 받을 수 있으므로, 2차원 필기와 달리 pen-up 상태에서도 궤적의 움직임을 알 수 있다. 그림 2는 2차원 필기와 3차원 필기 시 각각의 궤적을 보여준다. 2차원 필기에서는 pen-up 상태의 움직임을 입력 받을 수 없기 때문에 일반적으로 그림 2(a)와 같이 이전 획의 끝점과 그 다음 획의 시작점을 잇는 직선으로 나타낸다. 반면, 3차원 필기에서는 그림 2(b)와 같이 pen-up 상태에서도 실제 궤적의 움직임을 얻을 수 있다. 그러나, 기존의 2차원 필기 인식 모델은 2차원 필기를 바탕으로 하므로 pen-up 상태의 경우 모두 직선 형태로 학습되게 된다. 따라서, 3차원 필기의 pen-up 부분을 그대로 입력하면 직선 형태로 학습된 모델에서는 형태가 매우 상이하여 낮은 유사도를 주게 되고, 이는 인식률의 저하를 가져오므로, 이에 대한 기존 모델의 개선이 필요하다.



(a) 2차원 필기 (b) 3차원 필기
그림 2. 2차원 필기와 3차원 필기의 궤적 모양

어 진다고 가정하며, 문자 내의 획과 획의 관계, 획 안의 점과 점의 관계를 베이지안 네트워크를 이용하여 계층적으로 표현한다. 문자가 입력되면 모델은 획의 개수에 맞추어 하나의 문자를 여러 개의 획으로 분할하고, 다시 각각의 획을 여러 개의 점으로 분할 한다. 분할된 획과 점 각각의 발생 확률을 계산하여 각 모델의 점수를 산출하고 가장 점수가 높은 모델의 레이블을 인식 결과로 반환한다. 베이지안 네트워크 기반 인식 모델은 필기문자의 변이를 흡수하면서도 획 간의 위치 정보를 잘 모델링하고 있어, 안정적인 인식 성능을 보여준다.

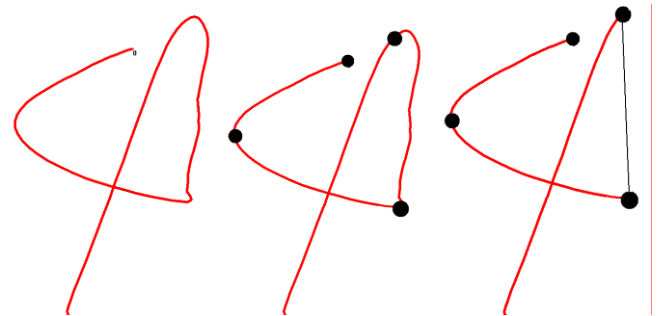
3.2 개선 시스템의 인식과정

3.2.1 획 분할

입력으로 필기 궤적이 들어오면 해당하는 모델이 가지고 있는 획의 수만큼 궤적을 분할한다. 그림 3(a)와 같이 숫자 4의 필기 궤적이 입력되는 경우, 숫자 모델 4는 4개의 획을 가지므로 그림 3(b)와 같이 분할된다.

3.2.2 연결획을 직선획으로 변경

모델상의 pen-up부분은 pen-down 획의 끝점과 그 다음 pen-down 획의 시작점을 연결하여 주는데, 이를 연결획이라고 한다. 기존의 모델은 연결획에 대해 대개 직선 획으로 학습이 되어있으나 3차원 필기는 연결획에 해당하는 부분 또한 사용자의 움직임에 따라 다양한 형태의 궤적이 발생한다. 이를 그대로 인식에 사용할 경우 연결획의 확률을 산출하는 부분에서 낮은 점수를 받게 되므로 본 연구에서는 분할된 획 중에서 모델 상의 연결획에 해당하는 획을 찾아 이 획을 직선획으로 변경하여 모델의 점수를 계산한다. 즉, 연결획의 세부적인 모양을 보지 않고 전체적인 방향과 길이만을 봄으로써, 연결획의 변이를 흡수하는 효과를 가져온다. 그림 3(a)의 입력의 경우, 그림 3(c)와 같이 연결획에 해당하는 부분을 직선 획으로 변경하여 연결획의 변이를 흡수한다.



(a) 입력 궤적 (b) 획의 분할 (c) 연결획의 직선화
그림 3. 획의 분할과 연결획의 직선화

3. 3차원 필기의 특징을 활용한 기존 모델의 개선

본 연구에서는 2차원 필기인식에서 안정적인 성능을 검증 받은 베이지안 네트워크 기반 인식 모델[6]을 개선하여 3차원 필기 인식에 이용하였다.

3.1 기존 모델 소개

베이지안 네트워크 기반 인식 모델은 필기 문자가 여러 개의 획으로 이루어지고 획은 여러 개의 점으로 이루어

4. 실험 및 분석

본 연구에서는 3차원 필기를 위해 ㈜마이크로인피니티 [7]에서 개발한 3차원 입력장치를 사용하였다. 이 장치는 3축 가속도 센서와 3축 각속도 센서를 이용하여 3차원 상의 움직임을 2차원상에 투영시킨 것과 같이 궤적을 복원[5]하여 준다.

성능 테스트에는 14명의 사용자로부터 수집한 2100개의 숫자 데이터를 이용하였다. 개선된 시스템의 성능을 분석하기 위해 기존의 모델에 수정하지 않고 사용한 경우(Original)와 개선한 경우(Modified)에 대하여 실험하였다.

표 1은 각 경우의 클래스별 인식률과 전체 인식률을 보여준다.

	0	1	2	3	4	5
Original	86.2	99.5	100.0	97.6	80.5	88.1
Modified	85.7	99.5	99.5	90.5	85.2	98.1

	6	7	8	9	Total
Original	91.4	82.4	90.5	94.3	91.0
Modified	91.4	85.7	88.1	91.9	91.6

표 1. 연결획을 직선화 하지 않았을 때와 했을 때의 비교



(a) 3->5 (b) 6->0 (c) 7->1 (d) 4->0
그림 4. 잘못 인식된 경우

실험 결과, 개선된 인식 모델은 개선 전 모델보다 인식률이 향상됨을 볼 수 있으며, 4, 5, 7과 같이 획의 개수가 여러 개인 문자에 대해서는 높은 인식률의 향상을 보였다. 여러 획으로 이루어진 문자에 대하여 인식률이 상승한 것은 연결획의 세부 모양을 고려하지 않는 인식 방법이 효과가 있음을 보여준다.

그러나, 하나의 획으로 이루어진 문자에 대해서 각각 조금씩 낮은 인식률을 보임을 확인하였다. 이는, 여러 획을 가지는 문자 모델의 확률을 계산할 때 연결획 부분의 세부 모양에 상관없이 확률이 높게 계산됨에 기인한다. 그림 4는 오인식의 대표적인 예이다. 그림 4(a)의 경우, 개선된 모델 중 5에 해당하는 모델의 연결획 부분에 대한 변이가 상대적으로 줄어들어 오인식이 발생한 예이다. 그림 4(b) 및 그림 4(c)는 0과 6, 1과 7의 형태상 모호함으로 인한 오인식의 예이다. 그림 4(d)는 펜 상태를 고려하지 않음으로 인해 4와 0의 구분이 어려워 오인식이 발생한 예이다.

5. 결론

본 연구에서는 3차원 스마트 입력장치로부터의 효과적인 필기 인식 방법을 제안하였다. 3차원 필기 입력 시 발생하는 부정확한 pen-up/pen-down으로 인한 오류를 줄이기 위해 그 구분을 두지 않음으로써 사용자가 편하게 입력을 할 수 있도록 하였다. 또한, 기존의 베이지안 네트워크 기반의 2차원 필기 문자 인식 모델을 개선하여 사용하였다. 연결획의 큰 변이를 흡수하기 위하여 3차원

필기의 연결획 부분을 직선화하여 변이를 줄임으로써 인식 성능이 향상되었다.

개선된 인식 모델은 비록 하나의 획을 가지는 문자에 대해서는 조금씩 낮은 확률을 보였으나, 여러 획을 가진 문자에 대하여 높은 성능을 보였다. 이는 앞으로 한글이나 한자와 같이 다수의 획을 가지는 문자를 다루게 될 때에도 본 연구의 개선방법이 유효함을 말해 준다.

본 연구에서는 3차원 궤적의 특징을 반영하여 인식 모델을 개선하였으나, 기존 모델은 2차원 필기 데이터를 기반으로 학습되었으므로 3차원 필기를 인식하는데 여전히 미흡한 점이 있다. 따라서, 다량의 2차원 필기 데이터와 새로 수집된 3차원 필기 데이터를 같이 사용하여 모델을 학습할 경우 좀 더 나은 인식성능의 향상을 기대할 수 있다.

감사의 글

본 연구는 정보통신부에서 총괄하는 선도기반기술개발사업 중 (주)마이크로인피니티가 주관하는 3차원 스마트 입력장치의 필기 문자 인식기법 개발(2차년도) 과제에 지원을 받아 수행되었습니다.

참고 문헌

[1] D. Reynaerts and H. Van Brussel, "Design of an Advanced Computer Writing Tool," Proc. of the 6th Int. Symposium on Micro Machine and Human Science, pp. 229-234, 1995

[2] 조성정, 이택현, 최현일, 김진형, "3차원 가속도 신호의 문자 인식 방법에 관한 연구," 14회 영상처리 및 이해에 관한 워크샵 발표 논문집, pp. 289-292, 2002

[3] Sung-Jung Cho, Jong Koo Oh, Won-Chul Bang, Wook Chang, Eunseok Choi, Yang Jing, Joonkee Cho and Dong Yoon Kim, "Magic Wand: A Hand-Drawn Gesture Input Device in 3-D Space with Inertial Sensors," Proc. of the 9th Int. Workshop on Frontiers in Handwriting Recognition, pp. 106-111, 2004

[4] Jong K. Oh, Sung-Jung Cho, Won-Chul Bang, Wook Chang, Eunseok Choi, Yang Jing, Joonkee Cho and Dong Yoon Kim, "Inertial Sensor Based Recognition of 3-D Character Gestures with an Ensemble of Classifiers," Proc. of the 9th Int. Workshop on Frontiers in Handwriting Recognition, pp. 112-117, 2004

[5] 송진우, 장현기, 이정환, 안시홍, 정학영, "MEMS 관성센서를 이용한 3차원 스마트 입력장치 설계," 한국정보기술학회 학회지 2004

[6] Sung-Jung Cho and Jin H. Kim, "Bayesian Network Modeling of Strokes and Their Relationships for On-line Handwriting Recognition," Pattern Recognition, vol. 37, no.2, pp. 253-264, 2004

[7] (주)마이크로인피니티, <http://www.minfinity.com>