

멀티모달 사용자 인터페이스를 위한 펜 제스처인식기의 구현

오 준 택*, 이 우 범**, 김 옥 현*

*영남대학교 컴퓨터공학과 **대구과학대학 컴퓨터공학과

전화 : 053-810-1503 / 핸드폰 : 016-819-7604

Implementation of Pen-Gesture Recognition System for Multimodal User Interface

Juntaek Oh*, Woobeom Lee**, Wookhyun Kim*

*Dept. of Computer Engineering, Yeungnam University

**Dept. of Computer Engineering, Taegu Science College

E-Mail : ojt@cse.yu.ac.kr, beomlee@mail.taegu-c.ac.kr, whkim@yu.ac.kr

Abstract

In this paper, we propose a pen gesture recognition system for user interface in multimedia terminal which requires fast processing time and high recognition rate. It is realtime and interaction system between graphic and text module. Text editing in recognition system is performed by pen gesture in graphic module or direct editing in text module, and has all 14 editing functions. The pen gesture recognition is performed by searching classification features that extracted from input strokes at pen gesture model. The pen gesture model has been constructed by classification features, ie, cross number, direction change, direction code number, position relation, distance ratio information about defined 15 types.

The proposed recognition system has obtained 98% correct recognition rate and 30msec average processing time in a recognition experiment.

I. 서론

펜 제스처인식은 멀티 모달 인터페이스의 핵심요소로서 이동 환경에서 사용자의 다양한 요구사항을 처리하는 지능형 단말기의 구현을 위해 필수적으로 개발되어야 할 과제이다. 그러나 대부분의 기존 연구는 인식률의 향상만을 위해 복잡한 획 해석과 처리과정에 의해 적은 개수의 펜 제스처만을 인식하므로 다양한 기능과 실시간 처리를 요구하는 멀티미디어 단말기에 적

합하지 못하며 미흡한 실정이다. 따라서 본 논문은 멀티미디어 단말기에서의 적용을 목적으로 정의된 펜 제스처의 분류특징을 기반으로 한 높은 인식률과 빠른 처리과정을 가지는 펜 제스처인식시스템의 구현을 목적으로 한다.

제안하는 인식시스템은 일반적으로 사용하는 15가지의 펜 제스처를 정의한 후 분류특징을 추출하여 인식을 위한 정보로 이용한다. 펜 제스처인식은 그래픽 모듈에서 수행되고 펜 제스처에 따른 14가지 기능은 인터페이스에 의해 그래픽 모듈과 텍스트 모듈에서 실시간으로 수행된다. 펜 제스처인식은 정의된 펜 제스처들에 대한 분류특징을 기반으로 구축된 펜 제스처모델과의 비교·검색에 의해 수행된다. 펜 제스처인식을 위한 기존의 복잡한 획 해석이나 인식과정과는 달리 어느 정도 변형된 정보라도 수용할 수 있도록 민감한 획 정보가 아닌 획의 교차수, 방향변화, 방향코드개수, 위치관계정보, 거리 비율정보를 이용함으로써 빠른 처리과정에 의해 수행된다.

II. 펜 제스처인식시스템

인식시스템은 정의된 펜 제스처기본요소의 인식부분인 그래픽 모듈, 펜 제스처에 의한 기능부여나 텍스트 모드에서의 기능을 수행하기 위한 텍스트 모듈, 모듈간의 데이터 공유를 위한 인터페이스로 구성된다. 그림 1은 제안하는 인식시스템의 처리과정을 보여준다.

2.1 그래픽 모듈

그래픽 모드에서 출력된 텍스트를 대상으로 펜 제스

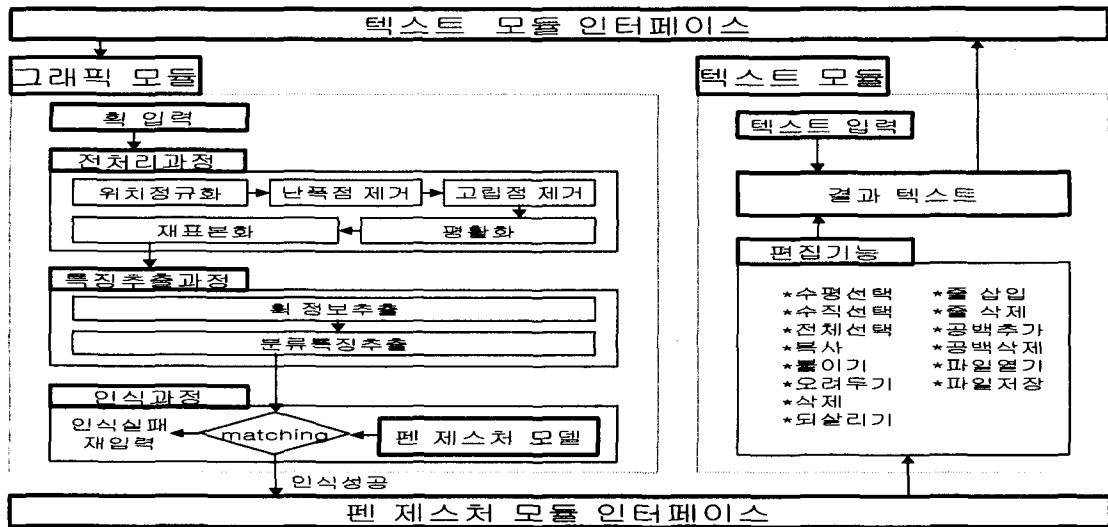


그림 2. 펜 제스처인식시스템의 전체처리과정

처인식을 하는 부분으로 전처리과정, 특징추출과정, 인식과정으로 구성된다.

2.1.1 전처리과정

정확한 분류 특징을 추출하기 위해 입력 획의 좌표열로부터 잡음 제거 및 개선 효과를 가지기 위한 과정으로써 위치정규화, 난폭점 제거, 고립점 제거, 평활화, 재표본화과정으로 구성된다.

가. 크기정규화

입력 펜 제스처의 크기를 일정한 크기로 규정하는 과정이다. N개의 획 S_i 로 구성되어 있으며 M_i 개의 점 $P_{i,j}$ 로 구성된 획이라면 아래의 식에 의해 정규화 되어진다.

$$W_{i,j} = \text{Max}X_{i,j} - \text{Min}X_{i,j} \quad (1)$$

$$H_{i,j} = \text{Max}Y_{i,j} - \text{Min}Y_{i,j}$$

$$X_{i,j} = \text{Norm}W \times \frac{X_{i,j} - \text{Min}X_{i,j}}{W_{i,j}}$$

$$Y_{i,j} = \text{Norm}H \times \frac{Y_{i,j} - \text{Min}Y_{i,j}}{H_{i,j}}$$

식 (1)에서의 $\text{Max}X_{i,j}$, $\text{Min}X_{i,j}$, $\text{Max}Y_{i,j}$, $\text{Min}Y_{i,j}$ 는 입력 획의 좌표열에서 최대, 최소값을 나타내며 입력 획의 크기를 300×300 로 규정하기 위해 NormW와 NormH를 300으로 설정하여 입력 획의 좌표열 데이터 $X_{i,j}$, $Y_{i,j}$ 를 변환한다.

나. 난폭점 제거

기계적인 결함에 손의 움직임에 제약을 받아 가속도가 큰 점을 교정 또는 제거하는 과정으로써 아래의 식

(2)을 이용하여 세 점을 대상으로 점들간의 거리 차에 따라 대상점을 난폭점으로 규정하고 이를 이웃하는 점들간의 평균값으로 대치함으로써 난폭점을 제거한다.

$$d_1 = \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2} \quad (2)$$

$$d_2 = \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}$$

If $d_1 \geq a \times d_2$, then (x_i, y_i) is 난폭점

where $a = 0.8$

다. 고립점 제거

미숙한 필기자들에 의해 삽입되는 불필요한 정보를 제거하기 위한 과정으로써 식(3)을 이용하여 제거한다.

$$d_1 = \sqrt{(sx - ex)^2 + (sy - ey)^2} \quad (3)$$

$$W = \text{Max}X - \text{Min}X, \quad H = \text{Max}Y - \text{Min}Y$$

If $(d_1, W, H) \leq a$, then S_i is 고립점

where $a = 15$

식 (4)에서 d_1 는 획의 시작점(sx, sy)와 끝점(ex, ey)간의 거리를 나타내고 W 와 H 는 획을 포함하는 최소외접사각형의 너비와 높이를 나타낸다.

라. 평활화

평활화는 원래 패턴의 특성을 잘 살려서 인식에 용이하도록 하기 위해 이웃하는 점들과의 평균을 구하여 점들의 위치를 고르게 하는 과정으로 아래의 식 (4)를 이용하여 대상점(x_i)으로부터 인접한 두 개의 점(x_{i-1}, x_{i+1})과의 평균에 의해 대상점을 변환한다.

$$x_i = \frac{1}{4} x_{i-1} + \frac{1}{2} x_i + \frac{1}{4} x_{i+1} \quad (4)$$

마. 재표본화

획을 쓰는 속도 차에 의한 불규칙한 점들간의 간격을 일정한 간격으로 정규화시키는 과정으로 DDA (Dot Density Algorithm)를 사용하였다. 그러나 하드웨어의 발전과 많은 처리시간의 소모에 의해 선택적인 전처리과정으로 사용한다.

2.1.2 특징추출과정과 인식과정

펜 제스처인식을 위한 과정으로써 특징추출과정에서는 획 정보추출과 분류특징추출과정으로 구성되며 인식과정에서는 추출된 분류특징을 이용하여 구축된 펜 제스처모델과의 비교·검색에 의해 인식한다.

2.1.2.1 특징추출과정

획 정보추출은 분류특징을 추출하기 이전에 필요로 하는 특징들을 추출하는 과정으로 특징점, 특징벡터, 위치관계벡터를 추출한다. 특징점은 입력 획의 시작점, 끝점, 굴곡점으로 정의하였으며 굴곡점은 인접한 점의 각도의 변화와 전 참조점(특징점)과의 각의 변화가 임계치(40°)보다 큰 점으로 정의하였다. 또한 재표본화 처리를 하지 않은 경우 특징점은 각도의 변화가 여러 번 있을 수 있는 획의 시작과 끝 부분에서 집중적으로 여러 개가 생성될 수 있으므로 이는 불필요한 정보가 된다. 이에 고리, 장식선과 인접한 특징점 제거과정에 의해 반드시 필요로 되는 획의 특징점을 추출한다. 특징벡터는 추출된 특징점간의 진행방향에 대한 정보로써 8방향 체인 코드를 사용하여 추출한다. 또한 획에 대한 구조적인 위치관계를 나타내기 위해 각 획에 대한 구조적인 위치관계정보인 이전 획의 마지막 점과 현재 획의 시작점에 대한 방향 정보(Tail to Head), 이전 획의 시작점과 현재 획의 시작점에 대한 방향 정보(Head to Head), 전체 획의 마지막 획의 끝점과 전체 획의 시작 획의 시작점에 대한 방향 정보(Last Tail to First Head)를 추출하여 여러 획으로 구성된 펜 제스처일 경우 이용한다.

인식을 위해 사용되는 분류특징은 정의된 15가지의 펜 제스처를 기반으로 어느 정도 변형된 입력 획도 인식 가능하기 위해 민감한 획 정보가 아닌 아래의 6가지 특징을 추출한다.

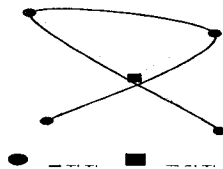


그림 2 교차점 추출 예

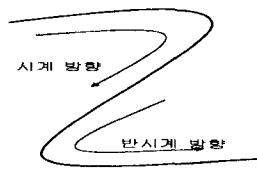


그림 3 방향변화 추출의 예

가. 획 내의 교차수

획 내에서 존재하는 특징점간에 이루는 직선의 교차수에 대한 정보이다. 그림 2는 한 획에서의 교차점 추출에 대한 예로써 교차수 1을 가진다.

나. 방향 벡터

특징벡터로써 단일 획이나 여러 개의 단일 획으로 구성된 펜 제스처를 인식하기 위한 진행방향정보이다.

다. 방향 코드의 수

방향 코드의 수가 많이 발생함은 입력 획에 대해 굴곡이 많음을 나타내므로 직선성분의 획과 곡선성분의 획을 구별하기 위한 정보로써 이용된다.

라. 획 내의 방향변화

특징벡터를 이용하여 입력 획에 대한 진행방향변화에 대한 정보로써 시계 방향, 반시계 방향, 무 방향으로 나누어진다. 시계방향은 진행방향이 반시계 방향보다 시계방향으로 써 가는 횟수가 많은 경우이고 반시계 방향은 시계 방향보다 반시계 방향으로 써 가는 횟수가 많은 경우이며 무 방향은 시계 방향 변화와 반시계 방향변화가 똑같은 경우이거나 전혀 방향변화가 없는 경우를 나타낸다. 그림 3은 'z'에 대한 방향변화정보를 추출하는 예로써 무 방향정보를 가진다.

마. 위치관계정보

한 획이 이루는 전체 범위에서의 방향 변화와 획 간에 이루어지는 위치관계에 대한 방향정보이다.

바. 시작점과 끝점간의 거리에 대한 비율정보

복사, 줄 삭제 제스처인 'c'와 파일열기 'o'를 구별하기 위한 추가적인 정보로써 이용된다.

2.1.2.2 인식과정

입력 획으로부터 분류특징을 추출한 후 구축한 펜 제스처모델과의 비교·검색에 의해 수행된다. 정의된 15개의 펜 제스처모델은 정확한 패턴매칭을 위한 민감한 획 정보가 아닌 특징을 분류특징으로 이용하였으며 변형이 가미된 펜 제스처에 대한 다양한 분류특징으로 구축되어 어느 정도 변형이 가미된 펜 제스처들도 인식이 가능하다. 정의된 15개의 펜 제스처모델은 그림 4와 같다. 펜 제스처의 인식여부에 따라 수행되어지는 인터페이스는 펜 제스처에 따른 기능을 수행한다.

2.2 텍스트 모듈

텍스트 모듈에서는 텍스트 입력, 기능수행, 결과 텍

수평 선택	수직 선택	전체 선택	복사	붙이기	오려 두기	삭제	되살 리기
-		+	C	V	ℓ	X	Z
문 입력	문 삭제	공백 삽입	공백 삭제	파일 열기	파일 저장		
□	C	>	□	□	O	S	

그림 4. 펜 제스처 유형

스트로 구성된다. 텍스트 입력부분은 펜 제스처인식과 기능수행을 위해 텍스트를 입력하는 부분이며 기능수행부분은 펜 제스처인식이나 텍스트 모드에서의 사용자 인터페이스에 의해 수행되며 편집기능 및 추가적인 기능 등 모두 14가지의 기능을 수행한다. 그림 4에서는 인식시스템이 제공하는 펜 제스처에 따른 기능의 종류를 보여준다. 결과 텍스트부분은 기능수행에 의해 편집되거나 수정된 결과를 보여주는 부분이다.

2.3 편집 기능에 의한 인터페이스

모들간의 상호작용을 위한 인터페이스부분으로써 그래픽 모듈과 텍스트 모듈간의 데이터 공유를 위해 텍스트 모듈 인터페이스와 그래픽 모듈 인터페이스로 구성된다. 텍스트 모듈 인터페이스는 텍스트 모듈에서의 결과 텍스트부분이 기능수행에 의해 변경되었을 경우 그래픽 모듈로 새로운 결과 텍스트부분의 값을 부여하는 기능을 수행하며 그래픽 모듈 인터페이스는 그래픽 모드에서 펜 제스처의 인식에 따른 텍스트를 선택하거나 기능 수행을 위해 인식된 펜 제스처에 따른 기능수행을 부여한다. 텍스트를 대상으로 기능을 수행하기 위해 그래픽 모드에서 나타난 텍스트에 대해 펜 제스처인식을 수행하며 인식과 함께 그래픽 모듈 인터페이스에 의해 텍스트 모듈에서의 결과 텍스트부분의 텍스트를 대상으로 편집기능을 수행한다.

III. 실험 및 결과

본 논문에서 제안하는 펜 제스처인식시스템은 Accecat III 5" x 5" tablet(2540 LPI)를 입력장치와 펜터엄 PC의 윈도우 환경에서 자바 언어를 사용하였다. 입력 텍스트로는 일정한 크기의 텍스트나 다른 높이와 너비를 가진 텍스트를 이용하였으며 타블렛에 어느 정도 훈련을 시킨 사용자들을 대상으로 실험한 결과 98%의 펜 제스처인식률과 30msec의 전체 평균 처리속도를 보였다. 그림 5는 인식시스템에서 수평, 수직 선택과 삭제명령어를 가진 펜 제스처를 사용한 실행화면을 보여준다. 기존의 펜 제스처인식시스템과는 달리 전처리과정에서 많은 작업을 수행함과 펜 제스처유형에 대한 다양한 분류특징으로 펜 제스처를 구축함으로써 높은 인식률을 얻었다. 또한 15가지 펜 제스처의

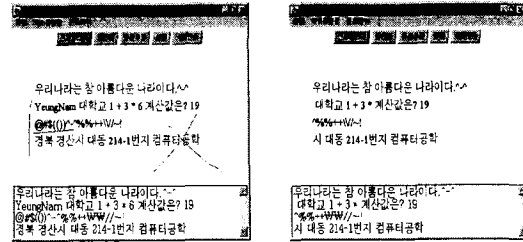


그림 5. 펜 제스처인식시스템의 실행화면

빠른 분류와 인식을 위해 제안한 6가지의 분류특징만을 이용함으로써 빠른 처리속도를 보였다.

오인식되거나 미인식된 펜 제스처는 회전이 가미된 경우이거나 변형이 많이 가미된 펜 제스처들로 펜 제스처 모델을 이용한 인식이 불가능한 경우였다.

IV. 결론

본 논문은 15가지 펜 제스처인식에 의한 14가지 텍스트 편집 기능을 부여한 펜 제스처인식시스템을 구현하였다. 제안한 인식시스템은 실시간으로 텍스트와 그래픽 모듈에서 상호작용하며 펜 제스처인식을 위해 교차수, 방향 변화, 방향 벡터, 방향 코드의 개수, 위치관계, 전체 획에 대한 시작점과 끝점간의 거리에 대한 비율에 대한 정보를 분류특징으로 이용하여 평균 30msec의 빠른 처리에 의한 98%의 높은 펜 제스처인식률을 획득하였다. 또한 펜 제스처모델에서의 분류특징은 다양한 펜 제스처에 대한 정보로 구축되어 어느 정도 변형된 펜 제스처도 인식 가능하였다.

제안한 펜 제스처인식시스템은 다양한 사용자 인터페이스와 빠른 처리과정에 의한 높은 인식률을 요구하는 멀티미디어 단말기로의 적용이 가능함을 보였으며 향후 변형이 가미된 불규칙한 문자나 도형의 처리나 추가적인 펜 제스처인식을 위한 분류특징에 대한 연구가 기대된다.

[참고문헌]

- [1] A. Chris Long, James A. Landay, Lawrence A. Rowe and Joseph Michiels, "Visual similarity of pen gestures" CHI 2000 conference, Pages 360 - 367, April. 2000
- [2] David Goldberg and Aaron Goodisman, "Stylus user interfaces for manipulating text", Proceedings of the fourth annual ACM symposium on User interface software and technology, Pages 127-135, 1991