

## (19) 대한민국특허청(KR) (12) 등록특허공보(B1)

(51) Int. Cl.<sup>6</sup>  
G06K 9/62

(45) 공고일자 1999년06월 15일

(11) 등록번호 10-0204618

(24) 등록일자 1999년03월29일

(21) 출원번호	10-1996-0027872	(65) 공개번호	특1997-0012219
(22) 출원일자	1996년07월11일	(43) 공개일자	1997년03월29일
(30) 우선권주장	95-216249	1995년08월24일	일본(JP)

(73) 특허권자	인터내셔널 비지네스 머신즈 코포레이션	포맨 제프리 엘
	미국 10504 뉴욕주 아몬크	
(72) 발명자	야마사끼 가즈따가	
	일본 도쿄도 마찌다시 쓰루마 157-3	
(74) 대리인	장수길	

심사관 : 오홍수

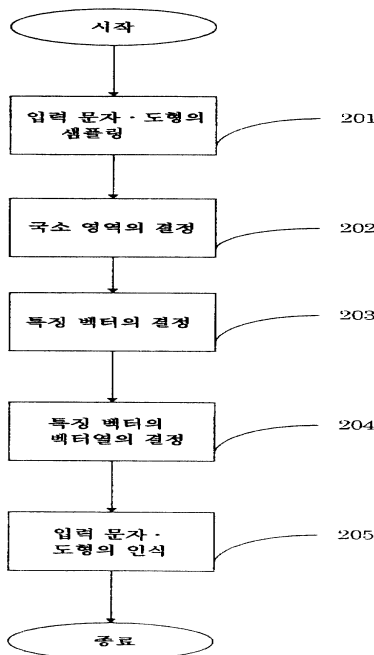
### (54) 문자 또는 도형의 인식 방법 및 시스템

#### 요약

온라인으로 입력된 문자의 필기 순서이외의 정보의 정보를 이용하여 문자인식을 행하는 방법을 제공하는 데 있다.

온라인으로 기입영역상에 입력된 문자 또는 도형을 인식하는 방법에 있어서, 온라인으로 입력된 문자 또는 도형을 샘플링하여 샘플링 정보를 추출하는 단계와, 샘플링 정보에 기초하여 입력된 문자 또는 도형으로부터 복수의 국소 영역을 결정하는 단계와, 각각의 국소 영역마다 특징 벡터를 구하는 단계와, 각각의 특징 벡터의 기입 영역상의 배치에 기초하여 벡터열을 구하는 단계와, 벡터열에 기초하여 입력된 문자 또는 도형을 인식하는 단계를 갖는 방법.

#### 대표도



#### 명세서

##### 도면의 간단한 설명

제1도는 본 실시예에 있어서의 수서 인식 시스템(the handwriting-recognition system)의 블록도.

제2도는 본 실시예에 있어서의 수서 인식 방법의 순서도.

제3도는 특징 벡터의 벡터 열을 구하는 점을 상술한 순서도.

제4a도 및 b도는 태블릿(tablet)의 기입영역에 「と」가 입력된 경우의 점의 집합 및 국소 영역을 도시하는 도면.

제5도는 각각의 국소 영역에 있어서의 구성하는 점을 도시하는 표.

제6도는 각각의 국소 영역의 y좌표를 도시한 표.

\* 도면의 주요부분에 대한 부호의 설명

102 : 수서 인식 프로그램	104 : 프런트 엔드
108 : 모델링 요소	110 : 컴퓨터 플랫폼
112 : 오퍼레이팅 시스템	114 : 마이크로 명령코드
116 : 하드웨어 요소	118 : 랜덤 액세스 메모리(RAM)
120 : 중앙처리장치(CPU)	122 : 입출력 인터페이스
124 : 단말기	126 : 전자 입력 태블릿
128 : 데이터 기억 디바이스	130 : 프린터

## 발명의 상세한 설명

### 발명의 목적

#### 발명이 속하는 기술분야 및 그 분야의 종래기술

본 발명은, 문자 또는 도형의 인식방법 및 시스템에 관한 것으로, 특히 온라인으로 입력된 수서 한자 등의 인식방법 및 시스템에 관한 것이다.

온라인으로 입력된 수서 문자 또는 도형을 인식하는 자동 시스템으로서, 많은 수법이 종래 제안되고 있다. 예를 들어, 일본국 특원평4-220410호(미합중국 특허 제5,343,537호)는 한 개 또는 몇 개의 특징 벡터 공간, 각 공간에 있어서의 가우스·모델링 및 모든 공간에 있어서의 연관된 모든 프로토타입의 기여를 감안한 혼합 복합에 있어서 수서의 적당한 표시에 기초하여 수서된 텍스트의 자동 인식을 행하는 방법 및 장치를 개시하고 있다.

구체적으로는, 전자 태블릿 상의 스타일러스에 의한 필기자의 필기에 응답하여, 필기자로부터의 수서 입력을 포함한 공지된 문자를 샘플링한다. 이 샘플링된 공지된 문자의 수서 공간에 있어서의 파라미터·벡터 표시를 행하는, 샘플링된 공지된 문자의 수서 공간에 있어서의 파라미터·벡터 표시의 제공에 응답하여, 수서 프로토타입을 제공한다. 그리고, 태블릿 상의 스타일러스에 의한 필기자의 필기에 응답하여, 필기자로부터 인식될 수서 입력을 포함한 미지의 문자를 샘플링하여 샘플링된 미지의 문자의 수서 공간에 있어서의 파라미터·벡터 표시를 행한다. 후보 문자의 리스트를 작성하는 것을 포함하는, 수서 프로토타입과 상기 샘플링된 미지의 문자의 수서 공간에 있어서의 파라미터·벡터 표시와의 탄도 비교(ballistic comparison)에 기초하여 수서 프로토타입의 적어제1도개가 미지의 문자로서 인식되는 후보 문자인 유사성을 평가된다. 그리고, 후보 문자 리스트의 탄도 분석을 행하여, 인식될 수서 입력을 포함한 샘플링된 미지의 문자를 인식하고 있다.

또한, 일본국 특원평4-328128호는, 감춰진 마코브 모델(a hidden Markov model)을 이용하여 수서 인식을 행하는 방법 및 시스템을 개시하고 있다.

일반적으로, 온라인 수서 문자 인식은, 문자가 정확한 필기 순서로 쓰여져 있다고 가정할 수 있는 경우에 대해서는 높은 확률로 문자를 인식할 수 있다. 상술한 종래 기술도 이와 같은 필기 순서, 즉 스트로크 순서(stroke order)라고 하는 시간적인 정보에 기초하여 문자 인식을 행하고 있다. 그러나, 한자 등의 화수(strokes)가 많은 문자에서는 쓰는 사람에 따라 필기 순서가 달라지기 때문에, 동일한 문자이면서 시간적인 정보도 서로 다르다. 따라서, 같은 문자도 그 필기 순서에 의해서는 잘못으로 인식되는 경우가 있거나, 또는 문자 인식은 입력된 문자로부터 추출·작성된 모델을 문자 인식 시스템속의 사전에 기억된 다수의 모델과 비교하여 각각의 스코어를 구함으로써 달성된다. 따라서, 상술한 스트로크의 순서가 다른 것을 고려하여 임의의 1개 문자에 대해 예상되는 몇 개 패턴의 모델을 사전에 기억하여 두면, 사전 내용이 필연적으로 팽창하게 된다. 이것은, 사전 용량의 증대뿐 아니라, 인식 속도의 저하에도 관련되기 때문에 바람직하지 않다.

#### 발명이 이루고자 하는 기술적 과제

본 발명은 상술한 문제점을 감안하여 이루어진 것으로, 그 목적은 입력 필기 순서와 화수에 제한되지 않은 인식 수법을 제안하는 것이다. 이와 같은 수법은 문자의 화수가 영어 등과 비교하여 많은 일본어, 특히 한자의 인식에 대해 효과적이다.

또한, 본 발명의 다른 목적은, 문자의 필기 순서 이외의 정보, 즉, 시간적인 정보이외의 정보를 이용하여 문자 인식을 행하는 것이다.

### 발명의 구성 및 작용

본 발명은, 기입 영역상에 입력된 문자 또는 도형을 인식하는 방법에 있어서, 입력된 문자 또는 도형을 샘플링하여 샘플링 정보를 추출하는 단계; 샘플링 정보에 기초하여 입력된 문자 또는 도형으로부터 복수의 국소 영역을 결정하는 단계; 각각의 국소 영역마다 특징 벡터를 구하는 단계; 각각의 특징 벡터의 기

입 영역상의 배치에 기초한 벡터 열을 구하는 단계; 및 벡터 열에 기초하여 입력된 문자 또는 도형을 인식하는 단계를 갖는 방법을 제공한다. 여기서, 기입 영역이라는 것은, 인식하고자 하는 문자 또는 도형이 존재하는 영역을 말한다. 그리고, 이 영역상의 국소 영역의 위치에 기초하여 특징 벡터를 순서에 따라 나열함으로써 특징 벡터의 벡터 열을 구할 수 있다. 이 기입 영역은, 온라인 문자 인식에 대해서는 일반적으로 문자가 입력되는 프레임 영역이나 하선 영역(a frame region or underline region)등이다.

여기서, 상기 벡터 열을 구하는 단계는, 기입 영역상의 한쪽 방향으로부터 대향하는 방향에 있어서의 국소 영역이 배치된 순서에 기초하여 결정되도록 하는 것이 바람직하다. 보다 구체적으로는,  $x$ ,  $y$ 좌표로 표현되는 기입 영역에 대해, 벡터 열에 대응하는 국소 영역을 참조하여 그  $y$ 좌표가 큰 것으로부터 순차적으로 나열함으로써 구하는 것이 바람직하다.

또한, 다른 발명은 온라인으로 기입 영역상에 입력된 문자 또는 도형을 인식하는 방법에 있어서, 온라인으로 입력된 문자 또는 도형을 샘플링하여 샘플링 정보를 추출하는 단계; 샘플링 정보에 기초하여 입력된 문자 또는 도형으로부터 복수의 국소 영역을 결정하는 단계; 각각의 국소 영역마다 특징 벡터를 구하는 단계; 온라인으로 기입된 문자 또는 도형의 스트로크의 순서에 기초하여 복수의 특징 벡터로 이루어지는 제1 벡터 열을 구하는 단계; 각각의 특징 벡터의 기입 영역상의 배치에 기초하여 복수의 특징 벡터로 이루어지는 제2 벡터 열을 구하는 단계; 및 제1 벡터 열 및 제2 벡터 열에 기초하여 입력된 문자 또는 도형을 인식하는 단계를 갖는 방법을 제공한다.

여기서, 상기 제2 벡터 열을 구하는 단계는, 기입 영역상의 한쪽 방향으로부터 대향하는 방향에 있어서의 국소 영역이 배치된 순서에 기초하여 결정되도록 하는 것이 바람직하다.

또한, 상기 인식하는 단계는, 감춰진 마코브 모델법 또는 DP 매칭법을 이용하여 상기 입력된 문자 또는 도형을 인식하는 것이 바람직하다.

또한, 다른 발명은 온라인으로 기입 영역상에 입력된 문자 또는 도형을 인식하는 단계에 있어서, 온라인으로 입력된 문자 또는 도형을 샘플링하여 샘플링 정보를 추출하는 수단; 샘플링 정보에 기초하여 입력된 문자 또는 도형으로부터 복수의 국소 영역을 결정하고 각각의 국소 영역마다 특징 벡터를 구하는 수단; 각각의 특징 벡터의 기입 영역상의 배치에 기초한 벡터 열을 구하는 수단; 벡터열에 기초하여 입력된 문자 또는 도형을 인식하는 수단을 갖는 시스템을 제공한다.

여기서, 상기 벡터 열을 구하는 수단은, 기입 영역상의 한쪽 방향으로부터 대향하는 방향에 있어서의 상기 국소 영역이 배치된 순서에 기초하여 벡터 열을 구하는 것이 바람직하다.

이하, 본 발명의 바람직한 실시예로서, 온라인 문자 인식을 예로 설명한다. 제1도는 본 발명에 있어서의 상기 인식 시스템의 블록도이다. 이 시스템은 컴퓨터 플랫폼(110)을 포함한다. 컴퓨터 플랫폼(110)은 랜덤 액세스 메모리(RAM)(118), 중앙처리장치(CPU)(120) 및 입출력 인터페이스(122) 등으로 이루어지는 하드웨어 요소(116)를 갖는다. 컴퓨터 플랫폼(110)은, 오퍼레이팅 시스템(112)을 가지며, 마이크로명령 코드(114)를 가질 수도 있다.

컴퓨터 플랫폼(110)에는, 온라인으로 기입 영역상에 입력된 문자 또는 도형을 샘플링하여 샘플링 정보를 추출하는 수단, 예를 들어 전기 기입 테블릿(126)이 접속되어 있다. 이 테블릿(126)은, 사용자가 입력 펜을 사용하여 기입 영역에 원하는 문자나 도형을 기입하기 위한 것이다. 단말기(124), 데이터 기억 디바이스(128) 및 프린터(130) 등의 많은 주변장치도 접속되어 있다.

플랫폼 디바이스(110)에서는, 수서 인식 프로그램(102)이 작동한다. 수서 인식 프로그램(102)에 의해 프론트 에지(104), 재배치 처리 기구(106) 및 모델링 요소(108)가 작동한다. 여기서, 프론트 에지(104)는 샘플링 정보에 기초하여 입력된 문자 또는 도형으로부터 복수의 국소 영역을 결정함과 동시에, 각각의 국소 영역마다 특징 벡터를 구하기 위한 것이다. 재배치 처리 기구(106)는 각각의 특징 벡터의 기입 영역상의 배치에 기초한 벡터 열을 구하기 위한 것이다. 이것은, 예를 들어 기입 영역상의 한쪽 방향으로부터 대향하는 방향에 있어서의 국소 영역이 배치된 순서에 기초하여 구할 수 있다. 또한, 모델링 요소(108)에는 벡터 열에 기초하여 사용자에게 의해 입력된 문자 또는 도형을 인식하기 위한 것이다.

제2도는 본 실시예에 있어서 수서 인식 방법의 순서도이다. 우선, 온라인으로 입력된 문자 또는 도형을 샘플링하여 샘플링 정보를 추출한다(단계 201). 이것은, 사용자가 전자 입력 테블릿(126)의 기입 영역상에 기입한 문자 또는 도형이 샘플링의 대상으로 된다.

이 샘플링 정보에 기초하여, 입력된 문자 또는 도형으로부터 복수의 국소 영역이 결정된다(단계 202). 즉, 단계(201)에 의해 추출된 입력의 각 샘플점은, 기입 영역에 있어서의 좌표( $X_n, Y_n$ )로 정의되는 점이다. 이들 점의 간격은 필기자의 필기 속도가 반드시 일정하지 않기 때문에 통상은 같지 않게 사용자가 기입한 속도의 함수로서 표시된다. 여기서, 각 샘플점은 정규화하여 등간격인 점  $p(X_m, Y_m)$ 로 한다. 이와 같이, 테블릿(126)에 의해 포획된, 시간에 의존하는 각 샘플점은 모든 점이 등간격인 시간에 의존하지 않는 표현으로 변환된다.

다음에, 정규화된 등간격인 점  $p$ 에 기초하여 국소 영역이 결정된다. 여기서 국소 영역이라는 것은, 임의의 문자, 도형을 인식하는 경우에 필요로 되는 특징적인 일부분을 말한다. 일반적으로, 국소 영역은 스트로크의 시점, 종점 또는  $x$ ,  $y$ 좌표의 극대값·극소값을 포함하는 영역인 경우가 많다. 국소 영역은 같은 수의 점(예를 들어  $2k+1$ 개)으로 구성되도록 결정되어 있고, 문자·도형은 복수의 국소 영역을 갖는다.

단계(202)에 의해 구해진 복수의 국소 영역의 각각에 대해, 대응하는 특징 벡터를 구한다(단계 203). 또한, 이 단계는 단계(202)와 함께 제1도에 도시한 시스템의 프론트 엔드(104)에 대해 행해진다.

여기서, 특징 벡터라는 것은 임의의 국소 영역의 특징, 예를 들어 그 영역내의 각 점의 좌표나 스트로크의 곡선과 같은 점간의 관계 등의 파라미터를 갖는 벡터를 말한다. 국소 영역내의 임의의 점이 같은 영역내의 다른 점으로부터 어느 정도 변위하고 있는지를 도시하는 파라미터 등이 포함되어 있다.

또한, 임의의 문자에 있어서의 특징 벡터의 수가  $N$ 개인 경우는,  $N$ 개의 특징 벡터의 순서를 도시하는 벡터

열도 이 단계에서 결정되게 된다. 그러나, 이 벡터 열의 순서는 이하에 서술되는 바와 같이 기입 영역상의 배치에 기초하여 순서의 교체가 행해진다. 따라서, 이 단계에서 구해진 것은 초기 벡터 열(the primary vector series)이기 때문에, 그 벡터의 순서는 어느 순서라도 관계없다. 여기서 구해지는 초기 벡터 열은 일반적으로는 온라인으로 입력된 문자나 도형의 스트로크 순서, 즉 시간적인 순서에 기초하여 나열되어 있다.

구해진 복수의 특징 벡터를 기입 영역상의 배치에 기초하여 순서에 따라 벡터 열을 구한다(단계 204). 단계(203)에서, 스트로크의 순서에 따라 이미 벡터 열이 구해져 있는 경우에는 그 순서를 기입 영역상의 배치에 기초하여 교체하는 경우도 포함된다. 이 단계에 대해서는 후에 상술한다. 또한, 이 단계는 재배치 처리 기구(106)에 대해 행해진다.

마지막으로, 벡터 열에 기초하여, 입력된 문자 또는 도형을 인식한다(단계 205). 단계(204)에 의해, 기입 영역상의 배치에 기초한 벡터 열을 감춰진 마코브 모델(HMM)등으로 모델화함으로써 입력된 문자·도형을 시스템이 인식한다. 즉, 특징 벡터의 벡터 열을 시스템의 사전에 기억되어 있는 임의 문자의 감춰진 마코브 모델등과 순차적으로 비교하여 스코어를 계산하고, 스코어의 가장 양호한 것을 특정함으로써 문자 등의 인식이 달성된다. 즉, DP 매칭법을 이용하여 인식하는 것도 물론 가능하다. 이와 같이, 본래적으로 필기 순서에 기초하여 문자 인식을 행하는 HMM 등에 본 발명을 적용함으로써 올바른 필기 순서로 쓰여져 있지 않거나 또한 연속하여 쓰여져 있는 문자에 대해서도 높은 확률로 인식할 수 있게 된다.

상기 단계(204)에 대해 제3도를 더욱 상세히 설명한다. 우선, 단계(202)에서 구해진 임의의 문자·도형중 한 개의 국소 영역의 y좌표를 구한다(단계 301). 그러나, 임의의 국소 영역을 구성하는 점은 복수개로, 그 y좌표도 소정의 범위를 갖고 있기 때문에, 국소 영역의 y좌표를 한 개로 특정하기 위해서는, 몇 개의 알고리즘이 필요하다. 그래서, 임의의 특징 벡터  $f_n$ 에 대응한 국소 영역의 y좌표  $y_n$ 를 특정하기 때문에, 국소 영역이  $2K+1$ 개의 점으로 구성되어 있는 경우,  $K+1$ 번째의 점을 국소 영역의 y좌표로 정의한다. 예를 들어, 국소 영역이 5개의 점으로 구성되는 경우에 있어서, 3번째의 점  $p$ 의 y좌표를 국소 영역의 y좌표로 한다. 이와 같이 정의함으로써 용이하게 임의의 국소 영역의 y좌표를 특정할 수 있다.

또한, 국소 영역의 y좌표를 이와 같이 정의하는 것은, 특징 벡터의 벡터 열을 기입 영역의 배치에 기초하여 교체하기 위한 평가 기준을 부여하기 때문이다. 따라서, 본 발명은 이 정의 이외에도 여러 가지 방법이 고려될 수 있다. 예를 들어  $2K+1$ 개의 점의 각각의 y좌표를 구하고, 그 평가값을 국소 영역의 y좌표로 정의하여도 좋다.

또한, 벡터 열의 교체는, 본 실시예에서는 y좌표를 기준으로 하지만, x좌표에 기초하여도 물론 좋다. 본 발명은 온라인으로 기입된 문자의 스트로크 순서에 기초한 특징 벡터의 벡터 열 대신에, 기입 영역상의 배치에 기초한 벡터 열을 가지고 문자 인식을 행하는 것이 중요한 특징이다. 따라서, 기입 영역상의 한쪽 방향에서 대향하는 방향까지 국소 영역의 좌표의 대소 관계가 파악되는 한, 이 방법에 한정되어 있지 않다.

모든 국소 영역의 y좌표를 구할 것인지를 판단한다(단계 302). 「아니오」의 경우에는, 단계(301)로 복귀하여 모든 국소 영역의 y좌표가 결정될 때까지 이것을 반복한다. 이 일련의 단계에 의해 문자 또는 도형이  $N$ 개의 국소 영역으로 구성되어 있는 경우에, 이들 국소 영역의 모든 y좌표가 결정된다.

단계(302)에서 「예」로 된 경우, 즉 모든 국소 영역의 y좌표가 구해지면 배열 A가 작성된다(단계 303). 배열 A는 특징 벡터와, 이것에 대응하는 국소 영역의 y좌표의 조( $f_n, y_n$ )를 파라미터로 하고 있다.

국소 영역의 y좌표의 큰 순서로 배열 A를 교체한다(단계 304). 이 소트는 기존의 정렬 알고리즘으로 행할 수 있다.

벡터 열을 결정한다(단계305). 소트된 배열 A로부터 그 순서로 특징 벡터를 추출하여 열거한 것을 벡터 열이라고 한다.

이상과 같은 순서(단계 301에서 305)에 의해 구해진 벡터 열을 상기 단계(205)의 시스템의 사전 기억 내용과 비교함으로써 문자를 인식할 수 있다.

다음에, 구체적으로 문자 인식의 예로서, 제4도에 도시하는 바와 같이, 테블릿의 기입 영역에 「と」가 입력된 경우에 대해 설명한다. 제4a도와 같이, 기입 영역에 입력된 문자는, 도시하지 않은 횡방향인 x좌표 및 종방향인 y좌표에 따라 특정된 34개의 점의 집합으로서 정규화된다. 시스템은 정규화된 34개의 점의 집합을 근거로 제4b도에 도시하는 바와 같이 6개의 국소 영역( $r_1$ 에서  $r_6$ )을 결정한다. 국소 영역은 5개의 점을 포함하도록 구성되어 있고, 입력된 문자의 각 스트로크의 시점·종점이나 x좌표값 또는 y좌표값이 극대·극소로 되는 영역을 포함하도록 결정된다.

제5도는 이와 같이 특정된 6개의 국소 영역( $r_1$ 에서  $r_6$ )과 국소 영역을 구성하는 점  $p_n$ 과의 관계를 도시하고 있다. 또한, 여기서,  $r_1, r_2, r_3,$  및  $r_6$ 에는, 제4a도에 도시되어 있지 않은 점  $p_0$ 가 포함되어 있다. 이와 같은 가공의 점을 설정한 것은, 스트로크의 시점·종점이 국소 영역의 중심이 되도록 함과 동시에, 시점·종점을 포함한 국소 영역을 특정하는 알고리즘과 이것 이외의 영역을 알고리즘을 구별하지 않고, 통일적으로 취급하도록 하기 때문이다. 따라서, 제5도로부터도 이해할 수 있는 바와 같이, 국소 영역의 y좌표는  $n+1$ 번째의 점, 즉  $n=2$ 이기 때문에 실질인 3번째의 점이므로, 3번째의 점이 특정되어 있는 한, 가공의 점  $p_0$ 의 구체적인 값은 문제가 되지 않는다.

국소 영역  $r_1$ 에서  $r_6$ 에 각각 대응한 특징 벡터를  $f_1$ 에서  $f_6$ 를 구하여 입력 스트로크의 순서에 기초한 벡터 열을( $f_1, f_2, f_3, f_4, f_5, f_6$ )이라 한다. 이 경우에 각각의 국소 영역의 y좌표를 제6도에 도시한다. 따라서, 특징 벡터와 국소 영역의 y좌표의 요소로 이루어지는 배열 A는 이하와 같이 된다.

배열 A : ( $f_1, 100$ ), ( $f_2, 59$ ), ( $f_3, 73$ ), ( $f_4, 27$ ), ( $f_5, 0$ ), ( $f_6, 19$ )

이것을 배열 알고리즘을 이용하여 y좌표의 큰 것 순서로 소트한다. 이하와 같이 된다.

배열 A :  $(f_1, 100)$ ,  $(f_3, 73)$ ,  $(f_2, 59)$ ,  $(f_4, 27)$ ,  $(f_6, 19)$ ,  $(f_5, 0)$

이와 같이 하여, 특징 벡터만을 추출하여 y좌표의 크기에 기초하여 벡터 열  $(f_1, f_3, f_2, f_4, f_6, f_5)$ 가 얻어진다.

이 얻어진 벡터 열은 제4a도의 「 $\tau$ 」의 스트로크의 순서가 역인 경우에 있어서도, 모두 동일하다. 종래에는, 동일한 문자라도 스트로크의 순서에 따라 별도의 벡터 열이 생성되기 때문에, 1개의 문자에 대해서도 순서에 따른 복수의 모델을 사전에 준비하지 않으면 안되었다. 그러나, 이 방법에서는 스트로크의 순서에 관계없이 동일한 특징 벡터의 벡터 열이 얻어지기 때문에, 사전에 갖는 모델은 1개이어도 좋다. 따라서, 사전의 기억 용량을 작게 할 수 있어 인식 시간을 단축할 수 있다.

또한, 실제의 문자 인식에 있어서는 필기 순서가 작은 글자 종류와 많은 종류가 혼재하는 경우가 많다. 이와 같은 장면에서는 기입 영역의 배치에 기초한 본 발명의 인식 방법과, 스트로크의 순서에 기초한 인식 방법이라는 두가지 방법을 이용하여 양쪽의 스코어를 참작함으로써 문자 인식율을 향상시킬 수도 있다. 또한, 필기 순서가 사람에 따라 다를 가능성이 높은 문자에 대해서는 본 발명의 방법에 따라 얻어지는 스코어를 중시하고, 가능성이 낮은 문자에 대해서는 스트로크 순서에 기초한 스코어를 중시한다고 하는 방법도 효과적이다. 이 경우에는, 온라인으로 기입된 문자 또는 도형의 스트로크 순서에 기초하여 특징 벡터가 배열된 벡터 열과 기입 영역상의 국소 영역(특징 벡터에 대응)의 배치에 기초하여 특징 벡터가 배치된 벡터 열의 양쪽을 구하고, 각각에 대해 모델과 비교하여 스코어를 더할 필요가 있다.

### 발명의 효과

본 발명에서는, 사용자가 어떠한 필기 순서로 문자를 입력하여도 동일한 특징 벡터가 얻어지기 때문에, 1개의 문자 모델을 삭감할 수 있다. 따라서, 모델이 기억된 사전의 용량을 감소시킬 수 있어 고속으로 문자 인식을 행할 수 있게 된다.

### (57) 청구의 범위

#### 청구항 1

문자 또는 도형을 인식하는 방법에 있어서, 임의의 영역에 존재하는 문자 또는 도형을 샘플링하여 샘플링 정보를 추출하는 단계; 상기 샘플링 정보에 기초하여 상기 문자 또는 도형에 있어서의 복수의 국소 영역을 결정하는 단계; 각각의 상기 국소 영역마다 특징 벡터를 구하는 단계; 각각의 상기 특징 벡터를, 상기 문자 또는 도형이 존재하는 영역상의 상기 국소 영역의 위치에 기초하여 순서에 따라 나열함으로써 상기 특징 벡터의 벡터 열을 구하는 단계; 및 상기 벡터 열에 기초하여 상기 문자 또는 도형을 인식하는 단계를 갖는 것을 특징으로 하는 문자 또는 도형의 인식 방법.

#### 청구항 2

제1항에 있어서, 상기 벡터 열을 구하는 단계는, 상기 문자가 존재하는 평면 영역상의 한쪽 방향에서 대향하는 방향까지 각각의 상기 국소 영역의 위치 순서에 근거하고 있는 것을 특징으로 하는 문자 또는 도형의 인식 방법.

#### 청구항 3

온라인으로 입력된 문자 또는 도형을 인식하는 방법에 있어서, 기입 영역에 입력된 문자 또는 도형을 샘플링하여 샘플링 정보를 추출하는 단계; 상기 샘플링 정보에 기초하여 입력된 상기 문자 또는 도형으로부터 복수의 국소 영역을 결정하는 단계; 각각의 상기 국소 영역마다 특징 벡터를 구하는 단계; 각각의 상기 특징 벡터의 상기 기입 영역상의 배치에 기초한 상기 벡터 열을 구하는 단계; 및 상기 벡터 열에 기초하여 상기 문자 또는 도형을 인식하는 단계를 갖는 것을 특징으로 하는 문자 또는 도형의 인식 방법.

#### 청구항 4

제3항에 있어서, 상기 벡터 열을 구하는 단계는, 상기 기입 영역상의 한쪽 방향에서 대향하는 방향까지 상기 국소 영역이 배치된 순서에 근거하고 있는 것을 특징으로 하는 문자 또는 도형의 인식 방법.

#### 청구항 5

제1항 또는 제2항에 있어서, 상기 문자는 한자인 것을 특징으로 하는 문자 또는 도형의 인식 방법.

#### 청구항 6

온라인으로 기입 영역상에 입력된 문자 또는 도형을 인식하는 방법에 있어서, 입력된 문자 또는 도형을 샘플링하여 샘플링 정보를 추출하는 단계; 상기 샘플링 정보에 기초하여 상기 입력된 문자 또는 도형으로부터 복수의 국소 영역을 결정하는 단계; 각각의 상기 국소 영역마다 특징 벡터를 구하는 단계; 온라인으로 기입된 문자 또는 도형의 스트로크 순서에 기초하여 복수의 상기 특징 벡터로 이루어지는 제1 벡터 열을 구하는 단계; 각각의 상기 특징 벡터의 상기 기입 영역상의 배치에 기초하여 복수의 상기 특징 벡터로 이루어지는 제2 벡터 열을 구하는 단계; 및 상기 제1 벡터 열 및 상기 제2 벡터 열에 기초하여 상기 입력된 문자 또는 도형을 인식하는 단계를 갖는 것을 특징으로 하는 문자 또는 도형의 인식 방법.

#### 청구항 7

제6항에 있어서, 상기 제2 벡터 열을 구하는 단계는 상기 기입 영역상의 한쪽 방향에서 대향하는 방향까지 상기 국소 영역이 배치된 순서에 근거하고 있는 것을 특징으로 하는 문자 또는 도형의 인식 방법.

**청구항 8**

제1항 또는 제2항에 있어서, 상기 인식하는 단계는, 감춰진 마코브 모델법(a hidden Markov model method) 또는 DP 매칭법을 사용하여 상기 입력된 문자 또는 도형을 인식하는 것을 특징으로 하는 문자 또는 도형의 인식 방법.

**청구항 9**

기입 영역상에 입력된 문자 또는 도형을 인식하는 시스템에 있어서, 입력된 문자 또는 도형을 샘플링하여 샘플링 정보를 추출하는 수단; 상기 샘플링 정보에 기초하여 상기 입력된 문자 또는 도형으로부터 복수의 국소 영역을 결정하고, 각각의 상기 국소 영역마다 특징 벡터를 구하는 수단; 각각의 상기 특징 벡터의 상기 기입 영역상의 배치에 기초한 상기 벡터 열을 구하는 수단; 및 상기 벡터 열에 기초하여 상기 입력된 문자 또는 도형을 인식하는 수단을 갖는 것을 특징으로 하는 문자 또는 도형의 인식 시스템.

**청구항 10**

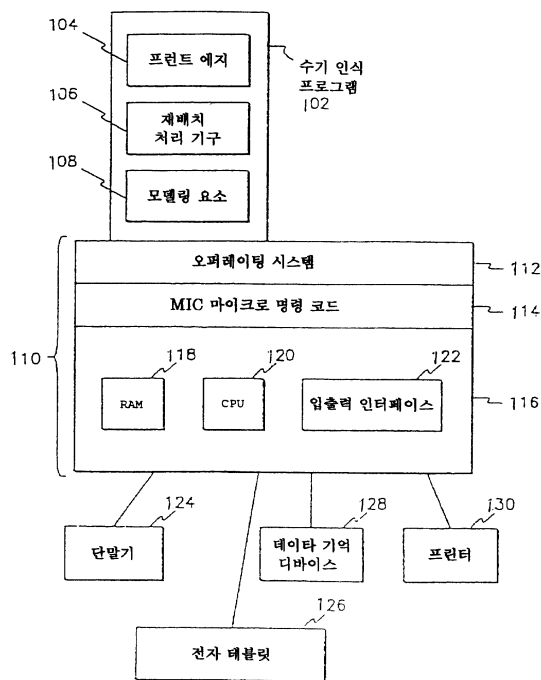
제9항에 있어서, 상기 벡터 열을 구하는 수단은, 상기 기입 영역상의 한쪽 방향에서 대향하는 방향까지 상기 국소 영역이 배치된 순서에 기초하여 상기 벡터 열을 구하고 있는 것을 특징으로 하는 문자 또는 도형의 인식 시스템.

**청구항 11**

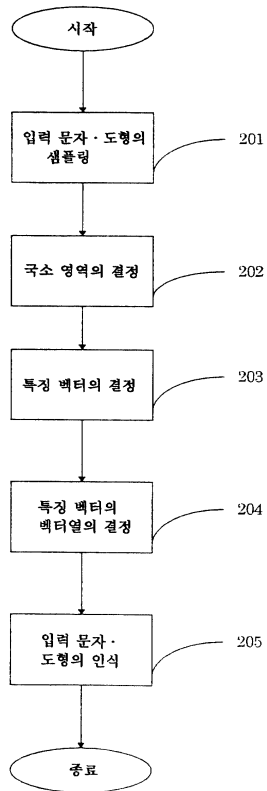
제9항에 있어서, 상기 인식하는 수단은, 감춰진 마코브 모델 또는 DP 매칭을 이용한 인식 수단인 것을 특징으로 하는 시스템.

**도면**

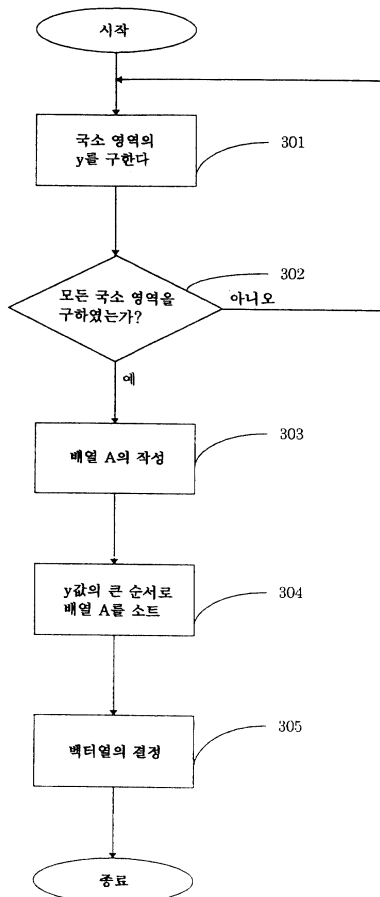
**도면1**



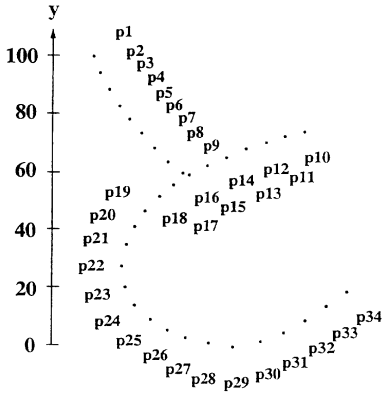
## 도면2



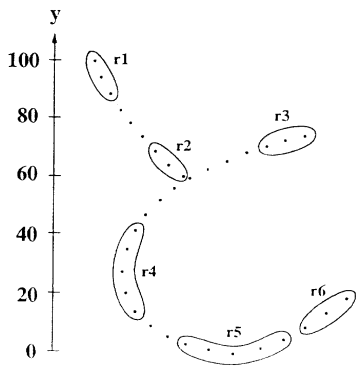
## 도면3



도면4a



도면4b



도면5

국소 영역	국소 영역을 구성하는 점
r1	p0 p0 p1 p2 p3
r2	p7 p8 p9 p0 p0
r3	p0 p0 p10 p11 p12
r4	p20 p21 p22 p23 p24
r5	p27 p28 p29 p30 p31
r6	p32 p33 p34 p0 p0

도면6

국소 영역의 y좌표	좌표값
y1	100 (p1의 좌표값)
y2	59 (p9의 좌표값)
y3	73 (p10의 좌표값)
y4	27 (p22의 좌표값)
y5	0 (p29의 좌표값)
y6	19 (p34의 좌표값)