

Dynamic Time Warping을 이용한 자기파형 인식

배윤지^o, 김황수

경북대학교 컴퓨터 과학과

tkatka@hanmail.net, hsk@knu.ac.kr

Magnetic Ink Character Recognition using Dynamic Time Warping

YoonJi Bae^o, HwangSu Kim

Dept. of Computer Science, Kyungpook National University

요 약

본 논문은 수표 하단에 인쇄된 자기 데이터를 읽어 들여 수표를 인식할 때 오독률을 최소화하는 효율적인 방법을 소개하고자 한다. 수정된 Dynamic Time Warping 기법으로 왜곡되거나 손상된 파형의 인식률을 높이고, class-selective rejection 방법을 통해 오독률을 최소화할 수 있는 효과적인 방법을 제안한다.

1. 서 론

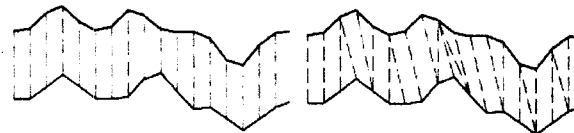
MICR(Magnetic Ink Character Recognition)은 수표의 정보를 감지하기 위해 수표 하단에 인쇄된 자기성분 데이터를 인식하는 것이다. ATM기에 입금된 수표로부터 MICR head가 읽어 들인 연속된 파형을 샘플링 하고, 글자의 파형으로 분리하여 각각의 문자를 인식 알고리즘에 의해 판독하여 해당수표의 정보를 알아낸다. 이러한 MICR 데이터는 사용자의 부주의로 쉽게 훼손되어 왜곡되거나 변형될 수 있는데, 이러한 요인으로 발생할 수 있는 오독을 최소화하는 효과적인 인식 알고리즘을 개발하는 것이 목적이다.

본 논문에서는 DTW (Dynamic Time Warping)기법 [1][2][3] 으로 14가지 템플릿과의 매칭을 통해 각 문자를 인식 하는데, 적절한 매칭을 위하여 수정된 DTW와, class selection rejection rule을 통하여 효율적인 방법으로 부독을 판정하여, 위험부담을 줄이면서 인식률을 높이는 방법을 제안한다. 2장에서는 인식에 적용된 DTW 매칭 기법에 대해서 소개하고, 3장에서는 부독 판정을 위한 class-selective rejection rule을 설명하며, 4장에서는 실제 ATM기

에서 오독 또는 부독 판정된 수표 데이터에 적용한 인식 결과를 보여주고, 5장에서는 결론 및 향후 과제에 대해 논의한다.

2. DTW 매칭 기법의 적용

Dynamic Time Warping (DTW) 기법은 음성인식, 생물정보공학, 온라인 필기체 문자 인식 등에서도 널리 사용되고 있는 알고리즘으로, 두 time series간의 distance를 최소화 하는 방향으로 dynamic하게 매칭 시켜 각 템플릿과의 누적 distance를 계산하여 최소가 되는 클래스로 인식한다. 아래 그림을 비교해서 살펴보면 DTW로 매칭했을 경우 단순 Euclidean distance로 매칭했을때와 달리 부분적으로 왜곡되거나 shift된 파형에 대해서도 적절한 매칭이 가능하다는 것을 알 수 있다.



(a)

(b)

그림 1 단순 매칭(a)과 DTW매칭(b)

본 연구는 ㈜효성과 BK사업단의 지원으로 수행되었음.

아래와 같이 각각 길이가 m, n 인 두 개의 time series X, Y $X = x_1, x_2, \dots, x_m, Y = y_1, y_2, \dots, y_n$ 가 주어졌을 때, 두 점 x_i 와 y_j 간의 거리 $d(x_i, y_j)$ 는 Euclidean distance를 사용한다. Warping path W 는 X 와 Y 사이의 매핑을 정의하는데, boundary condition, continuity, monotonicity의 세가지 조건[2]을 만족해야 하며, 이 조건을 만족하는 여러 가지 warping path 가운데 warping cost를 최소로 하는 path를 발견하는 것이 목적이다.

$$DTW(X, Y) = \min \left\{ \frac{1}{K} \sqrt{\sum_{k=1}^K w_k} \right\}$$

$$W = w_1, w_2, \dots, w_k \quad \max(m, n) \leq K < m + n - 1$$

k 번째 warping path w_k 를 (i, j) 의 매핑이라고 하면 $d(i, j)$ 는 Euclidean distance를 사용하며, (i, j) 번째까지의 누적 distance $D(i, j)$ 는 다음과 같이 정의할 수 있고, 다이내믹 프로그래밍 방법을 사용하면 효과적으로 구해낼 수 있다.

$$D(i, j) = d(x_i, y_j) + \min \begin{cases} D(i-1, j-1) \\ D(i-1, j) \\ D(i, j-1) \end{cases}$$

$i=j=0$ 에서부터 시작하여 $d(x_{i+1}, y_j), d(x_i, y_{j+1}), d(x_{i+1}, y_{j+1})$ 중 최소의 distance를 가지는 방향으로 진행하면서 누적 distance를 구해나간다. 일반적인 DTW에서 매칭은 Euclidean distance를 사용하며, 1차원 데이터의 경우 $d(x_i, y_j) = \sqrt{(x_i - y_j)^2}$ 로 정의할 수 있으나, 좀 더 자연스럽게 적절한 매칭을 위하여 아래와 같이 각 경우에 대하여 두 가지 변형을 가하였다.

i) 현재지점 i, j 에서 $i+1, j$ 로 진행할 경우

$$d(x_{i+1}, y_j) = \sqrt{\{(i+1) - j\}^2 + \left\{ \frac{x_i + x_{i+1}}{2} - \frac{y_j + y_j}{2} \right\}^2}$$

ii) 현재지점 i, j 에서 $i, j+1$ 로 진행할 경우

$$d(x_i, y_{j+1}) = \sqrt{\{i - (j+1)\}^2 + \left\{ \frac{x_i + x_i}{2} - \frac{y_j + y_{j+1}}{2} \right\}^2}$$

iii) 현재지점 i, j 에서 $i+1, j+1$ 로 진행할 경우

$$d(x_{i+1}, y_{j+1}) = \sqrt{\{(i+1) - (j+1)\}^2 + \left\{ \frac{x_i + x_{i+1}}{2} - \frac{y_{j+1} + y_{j+1}}{2} \right\}^2}$$

i)에서는 수평방향으로 지나치게 벗어나서 매칭되지 않도록 인덱스 차에 대한 패널티 항을 추가하였고, ii)에서는 진행방향의 값만을 사용하지 않고 현재 지점과 진행방향을 함께 고려하여 평균값을 사용하도록 수정하였다.

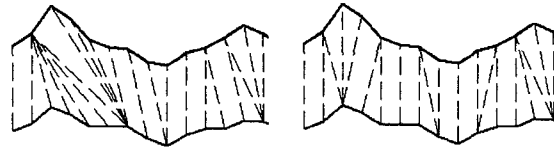


그림2 기본 식으로 매칭한 것(a)과 변형된 식으로 매칭한 것(b)

최종적으로 테스트 파형에 대해서 각 템플릿 파형과 매칭하여 DTW의 누적 distance를 최소로 하는 템플릿의 클래스로 인식하게 되는데, 그림2에서 기존의 식으로 매칭한 것과 비교해보면 변형된 식으로 매칭한 것이 수평방향으로 지나치게 벗어나지 않으면서 자연스럽게 매칭되었음을 확인할 수 있고, 이 결과는 곧 인식률에 영향을 미치게 된다.

3. class-selective rejection rule

수표인식 MICR에서는 잘못된 판정에 대한 위험부담이 크기 때문에, 실제로는 판정하기 곤란하다고 파악되면 잘못된 판정을 하기 보다는 부득으로 판정하여 reject 하는 것이 바람직하다. 부득판정을 위해 제안된 방법으로 class selective rejection[4]을 적용하였다.

이 기법은 인식과 동시에 부득을 판정할 수 있는 기법인데, 기존의 DTW 매칭기법을 적용하면서 class를 판정할 때 클래스의 subset으로 분류하는 기법으로 각 클래스에 대한 신뢰도가 높다면 정확히 하나의 클래스로 분류하겠지만, 그 외의 subset으로 판정된 것은 모두 부득으로 판정하는 rejection rule이다.

예를 들어, {1}, {2}, {3} 세개의 클래스로 구분하는 문제라면, {1}, {2}, {3}, {1, 2}, {1, 3}, {2, 3}, {1, 2, 3}의 7개의 클래스로 구분하는 문제로 바꾸어 생각할 수 있다., {1}, {2}, {3}으로 판정된다면 판정결과대로 판정하되 나머지 subset으로 판정된 경우는 모두 reject한다.

최적의 subset으로 판정하기 위하여 판정 위험률을 나타내는 loss function을 정의하고 각 클래스에 속할 확률과 그 클래

스를 선택했을 때 loss function의 값의 합이 최소가 되는 subset으로 판정하게 된다. 테스트 데이터 X에 대한 최적의 subset판정을 $\delta^*(X)$ 라 할 때, $\delta^*(X)$ 는 다음과 같이 정의한다.

$$\delta^*(X) = \arg \min_{a_j \in A} \left\{ \sum_{i=1}^N L(i, j) \cdot P(i/X) \right\}$$

$P(i/X)$ 는 X가 i번째 클래스에 속할 확률이고, a_j 는 가능한 subset이며, A는 가능한 모든 subset의 전체 집합이다. L(i, j)는 i번째 클래스에 대해 j번째 subset으로 판정한 경우의 Loss로 정의하고 i번째 클래스를 포함하지 않은데 대한 Loss L_m 과 subset의 size가 지나치게 큰데 대한 Loss L_n 의 합으로 다음과 같이 정의한다.

$$L(i, j) = L_m(i, j) + L_n(j)$$

$$L_m = \begin{cases} 0 & \text{if } i \in a_j \\ C_m & \text{otherwise ; } C_m > 0 \end{cases}$$

$$L_n = C_n \cdot |a_j| \quad ; C_n > 0$$

| |는 subset a_j 의 size(set cardinality)로 정의한다.

상수 C_m , C_n 은 잘못된 판단과 subset의 size중에서 어디에 가중치를 두어 판정할 지를 결정하는 상수이다.

4. 실험결과

손상되지 않은 MICR 수표 데이터와, 사용자의 부주의에 의해 훼손되어 실제 ATM기에서 부득 판정된 데이터에 대해서 각각 DTW와 개선된 DTW 그리고 역전파(Back Propagation) 알고리즘으로 인식결과를 테스트 해보았다.

표 1 개선된 DTW인식률 비교

	비 훼손 데이터	훼손 데이터
(a)기존 DTW	100% (4200/4200)	94.21% (358/380)
(b)개선된 DTW	100% (4200/4200)	97.89% (372/380)
(c)역전파 알고리즘	98.22% (4125/4200)	96.05% (365/380)

개선된 DTW(b)가 가장 높은 인식률을 나타냈으나, 여전히 오독이 존재하므로, 다음에는 class-selective rejection rule을 적용하여 오독가능성이 있는 데이터는 부득 처리 하여 보았다.

다음 표2에 class-selective rejection rule을 적용하였을 경우(b)와 각 클래스에 대한 confidence value를 thresholding하여 부득판정을 한 경우(a)를 비교하였다.

표 2 부득 처리결과 비교

	인식률	부득
(a)Confidence value	100%(298/298)	82
(b)class-selective rejection rule	100%(357/357)	23

오독 없이 즉, 인식률 100%를 유지할 때 reject되는 개수를 비교해보면 class-selective rejection rule을 적용한 경우가 훨씬 효율적인 결과를 나타낸다는 것을 알 수 있다.

5. 결론 및 향후 과제

본 논문에서는 사용자에게 의해 훼손된 수표 데이터를 개선된 DTW 알고리즘으로 인식률을 향상시키고, class-selective rejection rule에 의해 부득처리를 함으로써 낮은 부득률을 유지하면서 오독을 없애는 방법을 제안하였다. 수표인식에서는 오독이 치명적인 결과를 초래할 수 있기 때문에 부득처리가 반드시 필요하나, 부득률을 낮추는 것도 해결해야 할 과제로 남아있다. class-selective rejection rule에 의해 reject된 데이터는 subset으로 분류가 되어있기 때문에 그 정보를 가지고 subset내에서 전문적으로 분류할 수 있는 인식모듈을 마련하여 재판정한다면, 부득률을 낮추는 더 좋은 결과를 가져올 것으로 예상되며, 앞으로의 연구과제로 남아있다.

참 고 문 헌

[1] D. Sankoff and J. B. Kruskal: Time Warps, String Edits, and Macromolecule: The Theory and Practice of Sequence Comparison. Addison-Wesley, Reading, MA, 1983

[2] Eamonn J. Keogh and Michael J. Pazzani. Scaling up Dynamic Time Warping to Massive Datasets. In PKDD99, number 1704 of LNAI, Prague, Czech Republic, pages 1-11, 1999.

[3] T. M. Rath and R. Manmatha: Word Image Matching Using Dynamic Time Warping. In: Proc. of the Conference on Computer Vision and Pattern Recognition (CVPR), Madison, WI, vol. 2, pp. 521-527, June 18-20, 2003

[4] Ha, T.M.:The optimum class-selective rejection rule Pattern Analysis and Machine Intelligence, IEEE Transactions on , Volume: 19 Issue: 6, Pages: 608 -615 June 1997.

[5] Claudio De Stefano, Carlo Sansone, and Mario Vento: To Reject or Not to Reject: That is the Question, IEEE Transactions on systems, man, and cybernetics, VOL. 30, Feb 2000.