The Modified DTW Method for On-Line Automatic Signature Verification

Dong Uk Cho, Young Lae J. Bae, and Il Seok Ko

Dept. of Information & Communications Engineering,
Chungbuk Provincial University of Science & Technology, Chungbuk, Korea
`{ducho, yljb, isko}@ctech.ac.kr`

Abstract. Dynamic Programming Matching (DPM) is a mathematical optimization technique for sequentially structured problems, which has, over the years, played a major role in providing primary algorithms in pattern recognition fields. Most practical applications of this method in signature verification have been based on the practical implementational version proposed by Sakoe and Chiba [1], and is usually applied as a case of *slope constraint p = 0*. We found, in this case, a modified version of DPM by applying a forward seeking implementation is more efficient, offering significantly reduced processing complexity as well as slightly improved verification performance.

1 Introduction

Dynamic Programming Matching (DPM) is a mathematical optimization technique for sequentially structured problems, which has, over the years, played a major role in providing primary algorithms for automatic signature verification [1], [2], [3], [4]. In the pattern recognition field, it has been particularly used to eliminate the timing differences between two differently originating pattern signals. Hence it is called as the *Dynamic Time Warping* (DTW) method owing to its non-linear time-normalization function. Most practical applications of this method in signature verification [2], [5], [6] have been based on the practical implementational version proposed by Sakoe and Chiba [1], which is an analytical optimization method unlike others' rather heuristic approaches[1]. For practical use in signature verification, it is usually applied as a case of *slope constraint p = 0* as, apart from the fact that this provides the simplest and the fastest implementation owing to the least constraint (see Fig. 1), the slope constraint on the warping function has been noted to be merely time-consuming. The problem in the DPM application to signature verification was that many writers have an unstable pattern of signature writing, which confuses the DTW mechanism. A different approach from the opposite perspective to investigate the DTW function is performed by applying a forward seeking implementation of DTW under the assumption that the applied patterns satisfy the preconditions for the DTW function, i.e., the patterns have only a monotonic and continuous shift on the time axis. Thus a modified version of

---

[1] Other elastic matching methods include *the peak matching technique*, a finite state machine approach and *regional correlation*.

DPM in this context is developed. To verify the proposed method, experiments are applied under the same conditions and using the same data base to standardize and simplify the test for both conventional and proposed DTW methods. The results have proved the proposed method to be efficient, offering significantly reduced processing complexity as well as slightly improved verification.

2 DPM for Signature Verification

2.1 DPM Basics [1, 2]

Consider two different signals as sequences of feature vectors:

$$A = a_1, a_2, ..., a_i, ..., a_I \qquad (1)$$
$$B = b_1, b_2, ..., b_j, ..., b_J \ .$$

These two patterns, A and B, can be depicted in an i-j plane as shown in Fig. 2, where two patterns are represented along the i-axis and j-axis, respectively, and their matching stages are by a sequence of points $S(k)$, where $S(k) = (i(k), j(k))$.

To normalize these two signals with a N-stage decision process, a sequence of decision functions can be expressed as:

$$D_{(k, x_k)} = \sum_{k=1}^{N} C_k (q_k, x_k) . \qquad (2)$$

where Ck is a contribution function at kth stage for the decision vector qk and the state vector $xk(ai, bj)$.

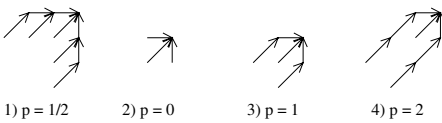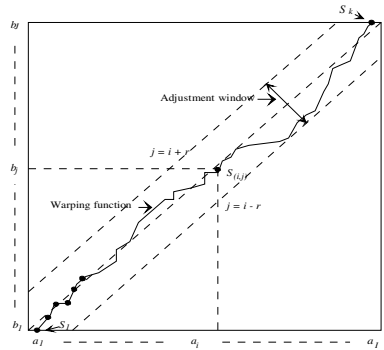| 1) p = 1/2 | 2) p = 0 | 3) p = 1 | 4) p = 2 |

Fig. 1. DTW slope constraint

Fig. 2. DTW mechanism for time alignment

DP matching seeks to find the optimum function $D(k, xk)$ at the kth stage:

$$D_{(k, x_k)} = \underset{q_k}{Optimum} \ [D_{(k-1, x_{k-1})} + C_k (q_k, x_k)] . \qquad (3)$$

In the context of the DTW algorithm, this problem of determining the optimal sequence corresponds to finding a minimum sequence of warping function $F(i(k), j(k))$, which is normally composed of two components:

$$F_k = d_{(i_k, j_k)} * w_{(i_k, j_k)} .$$
(4)

where $d(i_k, j_k)$ is the k_{th} occupancy cost and $w(i_k, j_k)$ is the corresponding weight.

Then the optimal objective function at the k_{th} stage, Df_k, is given as:

$$D_{f_k} = \underset{F_k}{Min} \left[D_{f_{k-1}} + F_k \right] .$$
(5)

The optimal value of this function will be the result of the sequence of recursive functions:

$$D_{f_k} = \underset{F}{Min} \sum_{l=1}^{k} \left[\frac{\sum_{l=1}^{k} d_{i_l j_l} * w_{i_l j_l}}{\sum_{l=1}^{k} w_l} \right] .$$
(6)

This is expanded as follows.

1. Initial condition:

$$D_{f_1} = d(f_1) * w(f_1).$$
(7)

2. DP-equation:

$$D_{f_k} = \underset{F_k}{Min} \left[D_{f_{k-1}} + d(f_k) * w(f_k) \right].$$
(8)

3. Time-normalized distance:

$$D_{i_k j_k} = \tfrac{1}{N} D_{f_k} .$$
(9)

where $N = \sum_{l=1}^{k} w(f_l)$

2.2 DPM Implementation

Sakoe and Chiba [1] provided a practical solution for Equation (6), which originally was proposed for speech recognition. Since then, this method has been extended for use in signature verification and has been widely accepted for practical applications.

Restrictions on the warping function
To provide a safeguard against unusual deviations during the warping process and to keep a desirable warping gradient, two conditions are imposed on the warping function:

1. Adjustment window (see Fig. 2)

$$| i(k) - j(k) | \le r.$$
(10)

where r is an adequate value for the window size.

This is to prevent unusual deviations from the warping function, which is based on the assumption that the normal time-axis fluctuation does not cause an excessive timing difference.

2. Slope constraint

An appropriate slope constraint is imposed to keep the warping gradient from an undesirable time warping (see Fig. 1).

Let the pattern at kth stage, (i_k, j_k), be a simplified term, (i, j), then Equation (9) becomes:

1) $p = 1/2$

$$D_{ij} = Min \begin{bmatrix} D_{(i-1,j-3)} & + & 2d_{(i,j-2)} & + & d_{(i,j-1)} & + & d_{(i,j)} \\ D_{(i-1,j-2)} & + & 2d_{(i,j-1)} & + & d_{(i,j)} \\ D_{(i-1,j-1)} & + & 2d_{(i,j)} \\ D_{(i-2,j-1)} & + & 2d_{(i-1,j)} & + & d_{(i,j)} \\ D_{(i-3,j-1)} & + & 2d_{(i-2,j)} & + & d_{(i-1,j)} & + & d_{(i,j)} \end{bmatrix}. \tag{11}$$

2) $p = 0$

$$D_{ij} = Min \begin{bmatrix} D_{(i,j-1)} & + & d_{(i,j)} \\ D_{(i-1,j-1)} & + & 2d_{(i,j)} \\ D_{(i-1,j)} & + & d_{(i,j)} \end{bmatrix}. \tag{12}$$

3) $p = 1$

$$D_{ij} = Min \begin{bmatrix} D_{(i-1,j-2)} & + & 2d_{(i,j-1)} & + & d_{(i,j)} \\ D_{(i-1,j-1)} & + & 2d_{(i,j)} \\ D_{(i-2,j-1)} & + & 2d_{(i-1,j)} & + & d_{(i,j)} \end{bmatrix}. \tag{13}$$

4) $p = 2$

$$D_{ij} = Min \begin{bmatrix} D_{(i-2,j-3)} & + & 2d_{(i-1,j-2)} & + & 2d_{(i,j-1)} & + & d_{(i,j)} \\ D_{(i-1,j-1)} & + & 2d_{(i,j)} \\ D_{(i-3,j-2)} & + & 2d_{(i-2,j-1)} & + & 2d_{(i-1,j)} & + & d_{(i,j)} \end{bmatrix}. \tag{14}$$

For practical use in signature verification, it is usually applied as a case of *slope constraint* $p = 0$ as in Equation (12) as, apart from the fact that this provides the simplest and the fastest implementation owing to the least constraint (see Fig. 1), the slope constraint on the warping function has been noted to be merely time-consuming.

Sakoe and Chiba [1] gave an example of practical implementation of DTW. The flow of the DTW solution for Equation (6) is diagrammed from the initialization according to Equation (7) to the time-normalization as in Equation (9). Unlike Equation (6), which uses variable "k", for indexing from the first stage, 1, to the final stage, "K", this implementation uses two indices, "i, j", to iterate "J" times the DP-equation (8) (see Fig. 2) for the sequential solution. The adjustment window size is applied as variable "r".

2.3 Experimentation

An experiment was performed to investigate how the nature of signatures affects the performance of DTW. It was relevant to the issue about the vulnerability of the DTW mechanism to relatively variable signature patterns. For this experiment, the data base consists of two contrasting types of signature sample groups:

1. Group I has the members who have relatively "stable" signature patterns.
2. Group II members have relatively "unstable" patterns in signature writing.

Group I has a membership of 15 writers and Group II 24. A total of 50 signatures was collected from each member in five sessions. Each individual donated ten signatures in each session. Random forgeries, i.e., signatures generated by others, were used for the forgery samples, on the same grounds. To eliminate effects arising from the variation of magnitude and orientation, a precise normalization process in the spatial domain was performed. The performance in terms of the equal error rate was measured as a function of the adjustment window size applying the $f(x,y)$ function. Fig. 3 is the DPM performance result from Group I and Fig. 4 is from Group II.
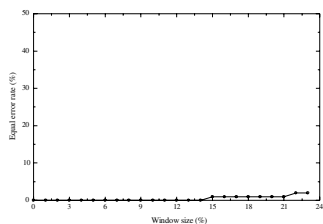
Fig. 3. Group I DPM result
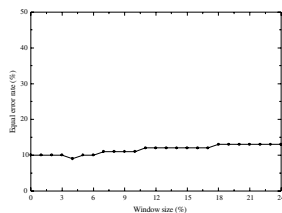
Fig. 4. Group II DPM result

From these results, it has been observed that the nature of signature samples has a considerable effect on the DTW performance:

1. For Group I, in which each member has a stable signature pattern, DTW has ideally functioned at zero error rates with smaller window sizes. Increasing the window size over 14% has caused the degradation of the error rate performance.
2. For Group II, in which most members have variable signature patterns, the DPM performance has been considerably degraded. The window size of 4% has recorded the best result at the equal error rate of 9%, which is slightly better than the results of 10% with neighbouring window sizes.

3   Development of Modified DPM

The problem in the DPM application to signature verification in the preceding sections, which applied the implementational version proposed by Sakoe and Sato [1], was that many writers have an unstable pattern of signature writing, which confuses the DTW mechanism. In this section, a different approach from the opposite perspective to investigate the DTW function is performed by applying a heuristic (forward seeking) implementation of DTW under the assumption that the applied patterns satisfy the preconditions for the DTW function, i.e., the patterns have only a monotonic and continuous shift on the time axis. Under such ideal conditions, there is little necessity of DTW functioning for all cases at the preceding stage (see Equation (8)) as the function is continuously increasing.

Algorithm

If the optimal objective function at the $k\text{-}1_{th}$ stage, $Df_{k\text{-}1}$, has been correctly selected, and the function satisfies the necessary conditions of continuity and monotonicity for DTW [1] and it does not have an abnormal (excessive) fluctuation[2] on the time axis, then Equation (8) can be alternatively expanded as:

$$D_{f_k} = D_{f_{k-1}} + \mathop{Min}_{F_k}\left[d\left(f_k\right)*w\left(f_k\right)\right] \tag{15}$$

A *slope constraint* then can be imposed as in Fig. 5 to maintain a normal time warping gradient, which corresponds to the slope constraint for Sakoe's version as in Fig. 1.

1) p = 1/2   2) p = 0   3) p = 1   4) p = 2
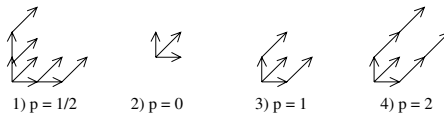
Fig. 5. MDTW *slope constraint*

For the practical application, it is implemented as follows:

1) $p = 1/2$

$$D_{i_k j_k} = d_{i_{k-1} j_{k-1}} + Min \begin{bmatrix} d(i_{k-1}+3, j_{k-1}+1) + d(i_{k-1}+2, j_{k-1}+1) + d(i_{k-1}+1, j_{k-1}+1) \\ d(i_{k-1}+2, j_{k-1}+1) + 2d(i_{k-1}+1, j_{k-1}+1) \\ 3d(i_{k-1}+1, j_{k-1}+1) \\ d(i_{k-1}+1, j_{k-1}+2) + 2d(i_{k-1}+1, j_{k-1}+1) \\ d(i_{k-1}+1, j_{k-1}+3) + d(i_{k-1}+1, j_{k-1}+2) + d(i_{k-1}+1, j_{k-1}+1) \end{bmatrix} \tag{16}$$

2) $p = 0$

$$D_{i_k j_k} = D_{i_{k-1} j_{k-1}} + Min \begin{bmatrix} d(i_{k-1} & , & j_{k-1}+1) \\ d(i_{k-1}+1 & , & j_{k-1}+1) \\ d(i_{k-1}+1 & , & j_{k-1}) \end{bmatrix} \tag{17}$$

3) $p = 1$

$$D_{i_k j_k} = D_{i_{k-1} j_{k-1}} + Min \begin{bmatrix} d(i_{k-1}+1 & , & j_{k-1}+2) \\ d(i_{k-1}+1 & , & j_{k-1}+1) \\ d(i_{k-1}+2 & , & j_{k-1}+1) \end{bmatrix} \tag{18}$$

4) $p = 2$

$$D_{i_k j_k} = D_{i_{k-1} j_{k-1}} + Min \begin{bmatrix} d(i_{k-1}+2 & , & j_{k-1}+3) \\ d(i_{k-1}+1 & , & j_{k-1}+1) \\ d(i_{k-1}+3 & , & j_{k-1}+2) \end{bmatrix} \tag{19}$$

---

[2] This was assumed for DTW mechanism in Sakoe and Chiba [1] and became the ground for implementing the adjustment window condition.

Equation (15) in the modified DPM (MDPM) version, firstly, has a strong point compared to Equation (8) in the conventional DPM (CDPM) as it requires only one DTW process at each decision stage while the conventional one requires this process as many times as the window size. Hence, this alternative method can reduce the computational complexity.

4   Experimentation

To compare the performances of both DPM methods, the same error rate performance tests were applied to the modified DPM (MDPM) for the two groups. Fig. 6 is the result for Group I and Fig. 7 is for Group II.
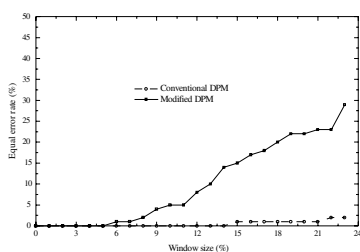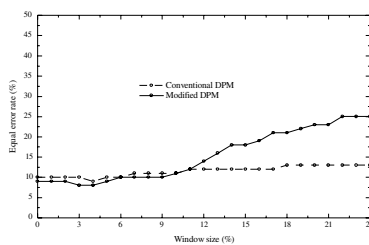
Fig. 6. Group I MDPM result

Fig. 7. Group II MDPM result

For both of the two groups, the modified DPM (MDPM) method has shown an equal or better performance compared to the conventional DPM (CDPM) method with smaller window sizes while it has a considerably degraded performance with larger window sizes. For Group I, MDPM as well as CDPM has recorded a zero error rate: for stable signature patterns, MDPM performs well as CDPM does. But its performance becomes degraded as the window size increases. For Group II, the best performance has been recorded by MDPM with the window size of 4 percent: for unstable patterns, MDPM has a slightly better performance than CDPM with smaller window sizes. Through all experiments, MDPM has shown equal or better performance than CDPM.

5   Conclusion

During the experiments for CDPM, it was observed that applying precise normalization such as preprocessing results in both an improvement in error rate performance and a smaller optimal window size. Accordingly, it was thought that the time domain fluctuation can also originate from the attitude variation during signature collection as the normalization process mainly reduces this geometrical variation. The results from the MDPM, which has been proposed for stable patterns satisfying the preconditions for DPM, applied under the same conditions as for CDPM, have also confirmed these implications as all the results have corresponded to the previous results for CDPM.

Some results have even emphasized the assumed trends, e.g., if normalization is more precisely carried out, the optimal window size is reduced.

The results from the experiments have shown that:

1. The temporal variations are ideally applied to DTW. Patterns which are affected only by these variations produce a good DTW result. (See the experimental results for Group I.)

2. The geometrical variations due to attitude change can be removed by using precise normalization, which correspondingly improves the error rate performance.

3. The random variations cannot be corrected. Patterns which are severely affected by these variations produce the worst DTW results. (See the experimental results for Group II.) Their influence can be minimized by reducing the adjustment window size.

References

1. H. Sakoe and S. Chiba: Dynamic Programming Algorithm Optimization for Spoken Word recognition, IEEE Trans. on ASCP, Vol.26, No.1 (1978)
2. M.C. Fairhust, S. Ng: Management of access through biometric control: A case study based on automatic signature verification, Universal Access in the Information Society, vol. 1, no. 1 (2001) 31-39
3. C. Lee:, Applications of Dynamic Programming to Speech and Language Processing, AT & T Technical Journal, May/June (1989) 115-130
4. H. F. Silverman and D. P. Morgan: The Application of Dynamic Programming to Connected Speech Recognition, IEEE ASSP Magazine, Jul (1990) 6-24.43
5. Parizeau and R. Plamondon: A Comparative Analysis of Regional Correlation, Dynamic Time Warping and Skeletal Tree Matching for Signature Verification, IEEE Trans. on PAMI, Vol.12, No.7 (1990)
6. Y. Sato and K. Kogure: On-Line Signature Verification based on Shape, Motion and Handwriting Pressure, Proc. 6th Int. Conf. on Pattern Recognition, Vol.2, München (1982) 823-826