Handwritten Character Recognition Using Piecewise Linear Two-Dimensional Warping

Mohammad Asad RONEE
Graduate School of Information Science and Electrical Engineering
Kyushu University
6-10-1 Hakozaki, Higashi-ku, Fukuoka-shi, 812-8581 Japan
ronee@human.is.kyushu-u.ac.jp

Seiichi UCHIDA
Faculty of Information Science and Electrical Engineering
uchida@is.kyushu-u.ac.jp

Hiroaki SAKOE
Faculty of Information Science and Electrical Engineering
sakoe@is.kyushu-u.ac.jp

Abstract

In this paper, the effectiveness of piecewise linear two-dimensional warping, a dynamic programming-based elastic image matching technique, in handwritten character recognition is investigated. The present technique is capable of providing compensation for most variations in character patterns while its computation remains tractable. The superiority of the present technique over several conventional two-dimensional warping techniques in providing deformation compensation is justified by experimental results with English alphabet. Another comparison with monotonic and continuous two-dimensional warping, a more flexible matching technique, reveals that the present method takes far less computation than the latter, yet provides almost the same recognition accuracy for most categories.

1 Introduction

Image pattern matching is one of the most widely used techniques in character recognition. For handwritten character images, inconsistent shape variations are often found in different samples of the character images of the same category. Here arises the necessity for elastic matching, or two-dimensional warping (2DW), which is defined as a pixel-to-pixel mapping with the minimum feature distance between two given images. This minimum feature distance is expected to remain stable against the shape variations present in the handwritten character images.

In this paper, we investigate the effectiveness of dynamic programming (DP)-based piecewise linear 2DW (PL2DW) method [1] on handwritten character recognition for its following features: 1) flexibility to provide compensation for most shape variations due to translation, rotation, scaling, skewness, and uneven local deformations, 2) isomorphic mapping to prevent excessive warping (a phenomenon when two character images from two different categories fit each other), 3) certainty in providing the optimal solution by virtue of using DP, 4) feasible computational complexity, 5) versatility with mapping constraints, and 6) data-driven bottom-up approach to the solution. In PL2DW, the mapping of each column of one image into another image is given by the linear interpolation of the mapping of several prespecified points (*pivots*) on that column. Thus the mapping is controlled by a small number of points. Though the linearization imposes a limit on the flexibility to some extent, PL2DW still provides sufficient fitting ability for a wide range of character images. In our investigation, we focus mainly on the flexibility and complexity issues, where the use of additional constraints is also considered.

The organization of this paper is as follows. In Section 2, a brief review of conventional DP-based 2DW methods for handwritten character recognition is made. The basic principle of PL2DW is discussed in Section 3. A character recognition experiment using PL2DW is discussed and a performance comparison of PL2DW with several 2DW methods is presented in Section 4. Finally, a summary is provided in Section 5.

2 Related work

The optimization strategy to determine 2DW plays an important role in the performance of a 2DW method. Among the optimization strategies, DP is used for 2DW problems for its several advantageous features – global optimality, versatility with criterion functions and constraints, and computational stability.

DP-based 2DW methods for handwritten character recognition can be classified into two groups. For conventional methods of the first group, pixels of one column(row) of an image are restricted to map only on the same column(row) of the target image. Nakano et al. [2] employed DP in optimizing the orthogonal mapping of rows and columns of the peripheral images of Chinese characters. Tanaka et al. [3] proposed a DP-based dynamic directional matching method to recognize handwritten Chinese characters where the optimal 2DW was provided as a collection of independently optimized one-dimensional warping of each column of the directional images of the target character. Agazzi et al. proposed an HMM-based method [4], combining above methods to find an optimized mapping sequence of columns(rows) where mapping of each column(row) was one-dimensionally optimized. Methods $\mathbf{C}_1$, $\mathbf{C}_2$, and $\mathbf{C}_3$, of Fig. 1 show possible types of warping from [2], [3], and [4], respectively. All these methods are practical in the sense that their complexities are polynomial order of the image size. However, inability to cope with some common variations of character image patterns, such as rotation, remains as their limitation.

For the second group, where monotonic and continuous 2DW (MC2DW) [5] (method $\mathbf{C}_4$ of Fig. 1) is the only candidate, no such restriction described above is applied on the mapping of pixels of the same row. Though MC2DW is flexible enough to overcome the limitation of the previous group, the computation for MC2DW is in the exponential order of the image size making the mehtod infeasible for application in real handwritten character recognition system.

3 Piecewise linear two-dimensional warping

3.1 Formulation of PL2DW

For $N \times N$ reference and input images $\boldsymbol{A} = \{a(i,j)$
$|i,j = 1,2,...,N\}$ and $\boldsymbol{B} = \{b(x,y)|x,y = 1,2,...,N\}$ respectively, consider a problem of optimizing 2D-2D mapping $\{x = u(i,j), y = v(i,j)\}$ from $\boldsymbol{A}$ to $\boldsymbol{B}$. If the mapping is optimized so that a pixel on $\boldsymbol{A}$ is mapped to its corresponding pixel

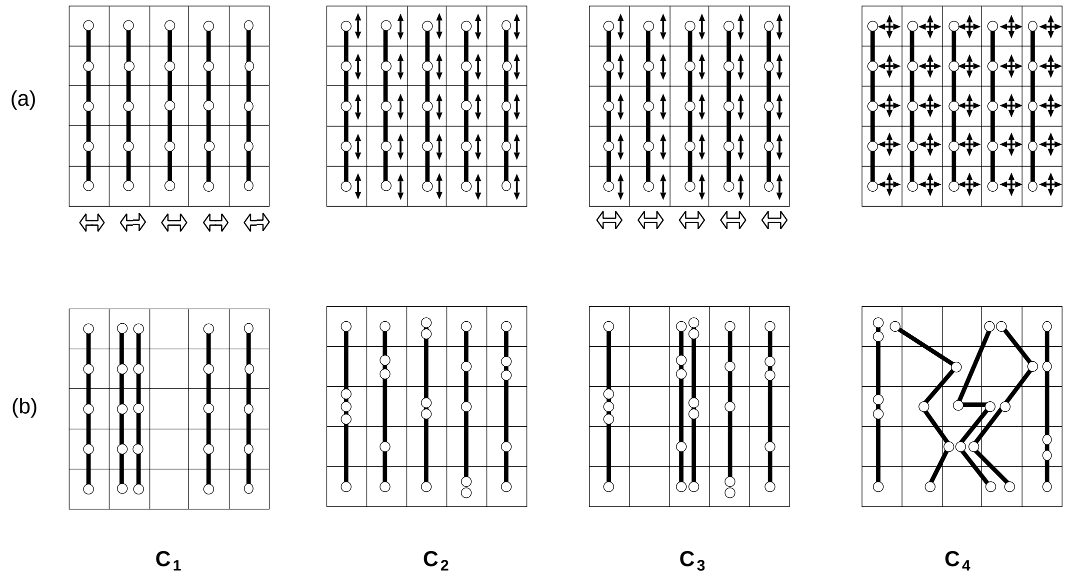

Fig. 1. Types of DP-based 2DW:(a) allowable direction of mapping, (b) warping examples.

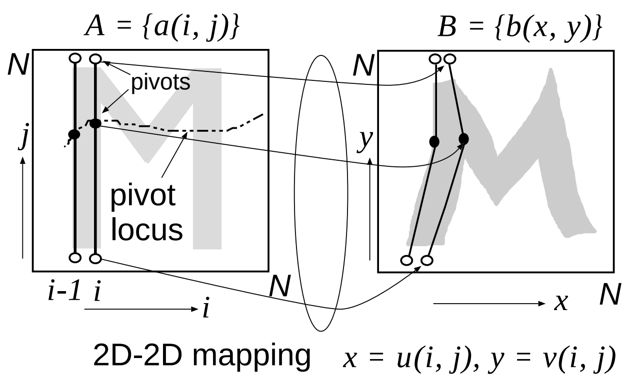

Fig. 2. PL2DW for K = 3.

on $\boldsymbol{B}$, the distance

$$D(\boldsymbol{A},\boldsymbol{B}) = \min_{u(i,j),v(i,j)} \sum_i \sum_j |a(i,j) - b(u(i,j),v(i,j))| \quad (1)$$

between the two images is expected to be invariant of any $\boldsymbol{B}$ if $\boldsymbol{B}$ belongs to the same category as $\boldsymbol{A}$ (that is, if different samples of the same category are used as input images).

In PL2DW, only K ($\leq N$) points, called pivots, of each column of $\boldsymbol{A}$ control the mapping of $\boldsymbol{A}$ on $\boldsymbol{B}$ (Fig. 2). The mapping of the remaining (non-pivotal) points of $\boldsymbol{A}$ is determined by linear interpolation, that is, each column of $\boldsymbol{A}$ is mapped as connected line segments on $\boldsymbol{B}$. Examples of image pattern matching using PL2DW are given in Fig. 3.

Pivots are given as K continuous non-intersecting lines (pivot loci) on $\boldsymbol{A}$. Since each column of $\boldsymbol{A}$ can only bend at pivots while being mapped on $\boldsymbol{B}$, it is desirable that the pivots should be arranged in such a way that their mapping can correspond to the bending and stretching points on $\boldsymbol{B}$, giving a fair matching between image patterns of the same category. On the other hand, poor matching between two character images of different classes is also desirable as well. Pivots are arranged artificially considering these two factors.

Monotonicity and continuity constraints are imposed on the pivot mapping in order to realize isomorphic mapping. These two constraints ensure that the mapping of each of the K continuous

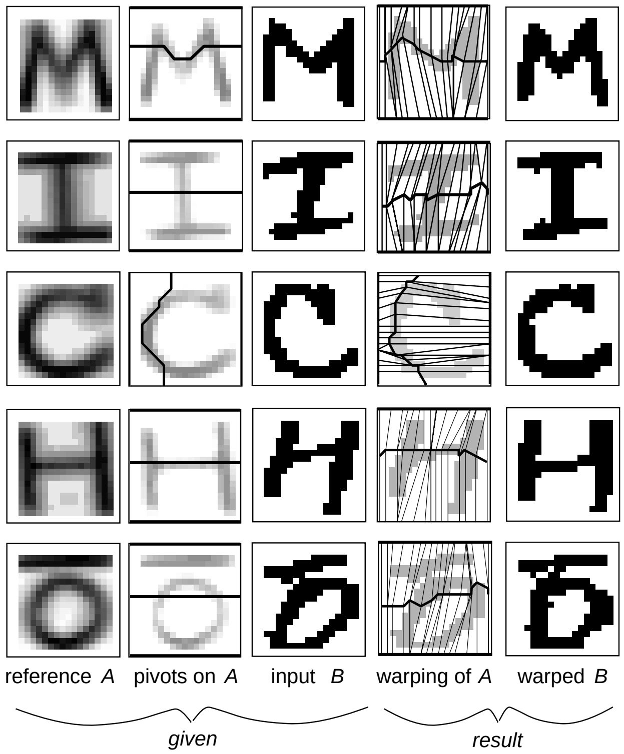

Fig. 3. Matching examples using PL2DW.

pivot loci on $\boldsymbol{A}$ will be (approximately) continuous, and will not intersect with each other. Boundary conditions are also imposed on the mapping to ensure that the boundary of $\boldsymbol{A}$ will be mapped on the boundary of $\boldsymbol{B}$. Furthermore, pivot mapping is limited to a small range $w(> 0)$ called warp range. Finally, a weighted penalty which evaluates the degree of local deformation is imposed on the pivot mapping in order to further suppress excessive warping.

PL2DW is considered as strong against variations in handwritten character images due to rotation, skewness, and uneven local deformations (Fig. 4). Compensation for variations due to global translation and scaling using PL2DW needs a relatively large value of K. Nevertheless, compensation for these variations can be provided easily during the preprocessing phase of input character images. Therefore, PL2DW can be applied to handwritten character recognition with small K.

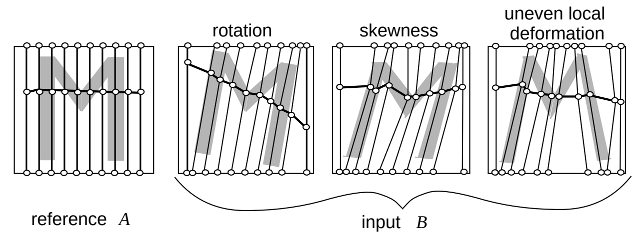

Fig. 4. Compensation for variations in handwritten character images by PL2DW.

3.2 DP algorithm

The *globally* optimal PL2DW, which minimizes the criterion function of Eq.(1) satisfying above constraints, is obtained by a DP-based algorithm, where the PL2DW between $\boldsymbol{A}$ and $\boldsymbol{B}$ is specified as the optimal state transition sequence in a multistage decision process. See Appendix for details on this algorithm. In order to reduce the computation in the case of higher w, less promising search paths are pruned off with the help of beam search technique incorporated into the algorithm.

For PL2DW, computational complexity is exponential order of the number of pivots K and polynomial order of the image size N. Therefore, the computation of PL2DW is tractable with small K, which is sufficient to provide compensation for variations present in most character images. That is why PL2DW is considered as feasible to implement in handwritten character recognition systems.

3.3 Additional constraints for individual categories

Since DP is used as the optimization strategy in PL2DW, incorporation of additional constraints into the above algorithm in order to reduce excessive warping is not difficult. Character images of several specific categories, due to their resem-

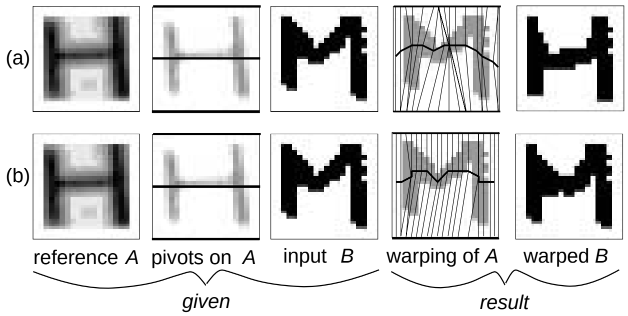

Fig. 5. The effect of additional constraint: (a) excessive warping with conventional PL2DW, (b) suppression of excessive warping with the use of an additional local constraint.

blance to the ones of some other categories, are vulnerable to excessive warping. For example, it is found experimentally that the reference pattern for character 'H' shows the tendency to mismatch with input pattern of character 'M' (Fig. 5(a)) due to excessive warping. In order to prevent such warping, additional constraints can be imposed both globally/locally when those character images are concerned. As a remedy to the problem in above example, the pivot mapping of character image 'H' is constrained such that the images of any two consecutive columns near the vertical axis of character image 'H' maintain a piecewise parallel property (Fig. 5(b)).

4 Experimental result and analysis

4.1 Experimental data

English capital letter patterns from the ETL6 database, a database of Japanese and English letters supplied by Electrotechnical Laboratory, were used where each 64×63 pattern was binarized, filtered out from noises, and normalized to fit in a 64×63 frame. Directional features [6], [7] in the directions of [—, \, |, /] were extracted from the contour of each normalized pattern. Then both the intensity pattern and the directional patterns

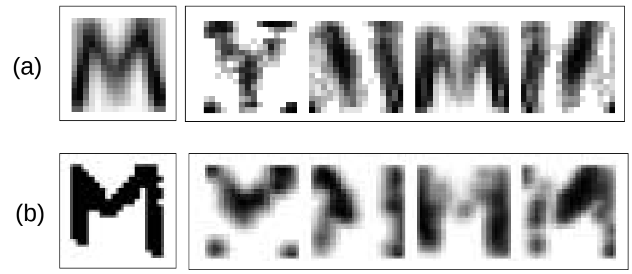

Fig. 6. An example of intensity and directional feature patterns for each of the directions [—, \ ,|, /] :(a) reference pattern, (b) input pattern.

Table 1. Recognition accuracy(%) of PL2DW

warp range w				
0	1	2	3	4
97.4	98.3	**98.6**	98.3	98.0

were down-sized to 16×16 rectangular frames centered on a 20×20 pattern plane. Finally, blurring and histogram equalization operations were then carried out only for directional patterns.

Each reference pattern $\boldsymbol{A}$ was prepared by histogram equalization after taking the average of the first 100 samples for each category. Each of the next 1000 samples for each category was used as input $\boldsymbol{B}$. The total number of input patterns stood at 25895 after samples with poor preprocessing were omitted from the experiment. Fig. 6 gives an example of a reference and an input pattern.

The positions and the direction of pivots on each reference pattern (Fig. 7) were decided considering the results of preliminary experiments.

4.2 Recognition result of PL2DW

A recognition experiment was conducted in order to measure the basic performance of PL2DW for different warp range w (that is, no additional constraint was used). The minimum distance

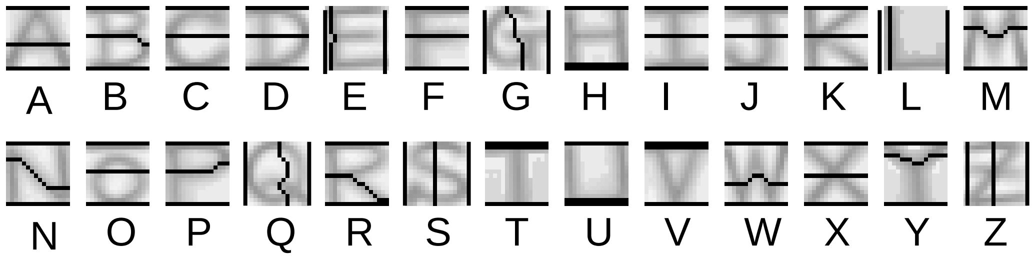

Fig. 7. Arrangement of pivot loci on reference patterns.

Table 2. Recognition accuracy (%) of three different 2DW methods and PL2DW

$\mathbf{C}_1$	$\mathbf{C}_2$	$\mathbf{C}_3$	PL2DW ($w = 2$)
97.6	97.5	97.8	**98.6**

$D(\boldsymbol{A}, \boldsymbol{B})$ between input $\boldsymbol{A}$ and reference $\boldsymbol{B}$ was used as the discrimination function for each input pattern. The result of the experiment is summarized in Table 1. In the table, the recognition accuracy corresponds to rigid matching when $w = 0$. It is clear from the table that best recognition accuracy was achieved when $w = 2$.

Patterns misrecognized with the present method were investigated for the possible cause of misrecognition. Among the 386 misrecognized patterns, about 59% have been identified as the result of excessive warping. The remaining 41% have been identified as the result of insufficient matching.

4.3 Comparison with other DP-based 2DW methods

In order to verify the effectiveness of the present method in deformation compensation, recognition experiments have been conducted using three different 2DW methods whose flexibility is illustrated as $\mathbf{C}_1$, $\mathbf{C}_2$, and $\mathbf{C}_3$ of Fig. 1. Penalty weight, warp range, and the orientation (row-wise or column-wise) are optimized for each of these methods. It should be noted that in order to evaluate the effect of warping flexibility the same criterion function in Eq.(1) and the same image features (intensity and directional) were used for each of these methods. Results are summarized in Table 2. From the result it is clear that PL2DW establishes its superiority over any of the compared methods.

The performance of PL2DW was also compared with that of MC2DW, which is nearly equivalent to the PL2DW with $K = N$. MC2DW, with optimized warp range and penalty, gives slightly better recognition accuracy (98.9%) than PL2DW (98.6%). However, the time required to recognize single character for MC2DW (110.8 sec) is far longer than that for PL2DW (3.4 sec) on a Sun Ultra2 (SPECint_95:12.3, SPECfp_95:20.2) computer. Furthermore, a comparison in recognition accuracy for each category (Fig. 8) unveils that the performance of PL2DW is close to that of MC2DW for recognition of most of the character categories except for categories 'E', and 'S'. Thus, it can be said that PL2DW, despite its somewhat reduced flexibility comparing with MC2DW, has the ability to provide compensation for variations present in most character images.

4.4 Effect of additional local constraints

From the error analysis of Section 4.2, excessive warping is identified as the principal cause of misrecognition. Additional local constraint technique, discussed in Section 3.2, is considered to

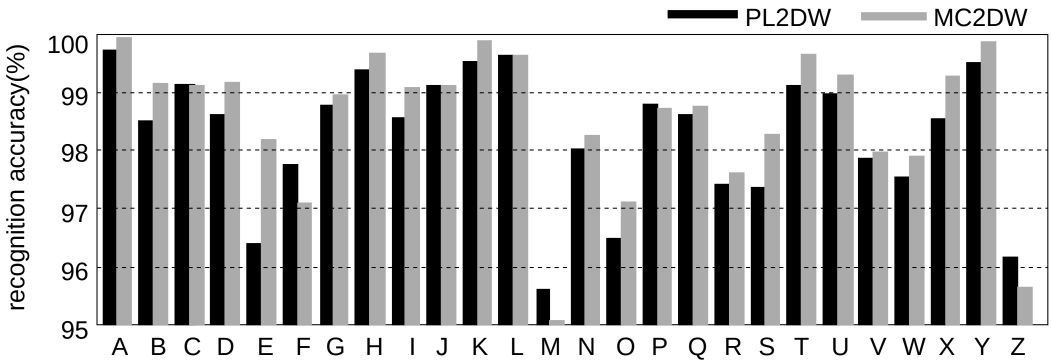

Fig. 8. Category-wise recognition accuracy of PL2DW and MC2DW

solve the problem. The effectiveness of this technique was experimentally investigated. Local constraints were imposed on the reference patterns of categories 'H' and 'Y' since these two categories show obvious tendencies to mismatch with input patterns of 'M' and 'T', respectively. The effectiveness of this technique is verified by the fact that about 20% and 30% of previously misrecognized 'M' and 'T' have been recognized. The overall recognition accuracy improved to 98.8%.

5 Conclusion

We have investigated the effectiveness of piecewise linear two-dimensional warping (PL2DW) in handwritten character recognition. Experimental results clearly indicate that the present method has more flexibility than conventional DP-based 2DW methods to provide compensation for deformations in character images. Another comparison with a more flexible DP-based 2DW method (monotonic and continuous 2DW) shows that PL2DW, despite its reduced flexibility, has the ability to compensate deformations almost as same as that of the former; moreover the latter takes considerably less computational time.

Additional local constraints have been imposed to control excessive warping successfully. Introduction of this technique also shows the advantage of using DP as the optimization method which allows to incorporate such constraints.

PL2DW, being inherently a structural analysis process, provides information on pixel-to-pixel correspondence between input and reference images. Some post-processing using this information is expected as a promising extension to the present framework.

References

[1] S. Uchida and H. Sakoe. Piecewise linear two-dimensional warping. *Proc. 15th IAPR Int. Conf. on Pattern Recognition*, Vol. 3 of 4, pp. 538–541, 2000.

[2] Y. Nakano, K. Nakata, Y. Uchikura, and A. Nakajima. Improvement of Chinese character recognition using projection profiles. *Proc. of 1st Int. Joint Conf. on Pattern Recognition*, pp 172–178, 1973.

[3] N. Tanaka, M. Shiono, H. Sanada, and Y. Tezuka. Recognition of handprinted Kanji characters by dynamic directional matching method. *Trans. Inst. Electron. & Commun. Eng.*, Vol. J68-D, No. 1, pp. 56–63, 1985. (in Japanese)

[4] O. Agazzi, S. Kuo, E. Levin, and R. Pieraccini. Connected and degraded text recog-

nition using planar hidden Markov models. *Proc. IEEE Int. Conf. on Acoustics, Speech & Signal Processing*, pp. V113–116, 1993.

[5] S. Uchida and H. Sakoe. Handwritten character recognition using monotonic and continuous two-dimensional warping. *Proc. 5th Int. Conf. on Document Analysis and Recognition*, pp. 499–502, 1999.

[6] T. Saito, H. Yamada, and K. Yamamoto. An analysis of handprinted chinese characters by directional pattern matching method. *Trans. Inst. Electron. & Commun. Eng.*, Vol. J65-D, No. 5, pp. 550–557, 1982. (in Japanese)

[7] S. Mori, H. Nishida, and H. Yamada. *Optical character recognition.* John Wiley & Sons, Inc., Sec. 11.4.1, pp. 469–472, 1999.

Appendix: Detailed DP algorithm

PL2DW between images $\boldsymbol{A}$ and $\boldsymbol{B}$ can be viewed as an optimal state transition problem in a multistage decision process. A DP algorithm that searches for the optimal state transition sequence on the multistage decision process is given in Fig. A·1. Here, each column i of image $\boldsymbol{A}$ constitutes a stage in the decision process, and the mapping of the pivots of column i of $\boldsymbol{A}$ denoted as $\boldsymbol{xy}_K(i)$ or $\boldsymbol{xy}$ for short, satisfying boundary conditions constitutes a state in that stage. The set of all such $\boldsymbol{xy}$ is denoted as $\boldsymbol{XY}_K(i)$, or simply $\boldsymbol{XY}$. In addition, the set of $\boldsymbol{xy}_K(i-1) \in \boldsymbol{XY}_K(i-1)$ satisfying monotonicity and continuity conditions is denoted as $\overline{\boldsymbol{XY}}_K(\boldsymbol{xy})$(Fig. A·2). Transition to $\boldsymbol{xy}$ is valid only from any state $\overline{\boldsymbol{xy}}$ of $\overline{\boldsymbol{XY}}_K(\boldsymbol{xy})$.

The transition cost to state $\boldsymbol{xy}$, denoted as $d(\boldsymbol{xy}|i)$, is defined as

$$d(\boldsymbol{xy}|i) = \sum_{j=j_{i,1}}^{j_{i,K}} |a(i,j) - b(u(i,j), v(i,j))|$$

/* Initialization */
for each $\boldsymbol{xy} \in \boldsymbol{XY}_K(1)$ **do**
  $g(\boldsymbol{xy}|1) := d(\boldsymbol{xy}|1)$

/* Recursion */
for $i = 2$ **to** N **do**
  for each $\boldsymbol{xy} \in \boldsymbol{XY}_K(i)$ **do**

$$g(\boldsymbol{xy}|i) := d(\boldsymbol{xy}|i) + \min_{\overline{\boldsymbol{xy}} \in \overline{\boldsymbol{XY}}_K(\boldsymbol{xy})} g(\overline{\boldsymbol{xy}}|i-1)$$

/* Termination */

$$D(\boldsymbol{A}, \boldsymbol{B}) := \min_{\boldsymbol{xy} \in \boldsymbol{XY}_K(N)} g(\boldsymbol{xy}|N)$$

Fig. A·1. A DP algorithm for PL2DW

where $u(i,j), v(i,j)$ is given by $\boldsymbol{xy}$. The quantity $g(\boldsymbol{xy}|i)$ refers to the cumulative state transition cost up to stage i with the final state being $\boldsymbol{xy}_K(i)$.

The algorithm terminates at stage N, where the lowest $g(\boldsymbol{xy}|N)$ gives the distance $D(\boldsymbol{A}, \boldsymbol{B})$. If necessary, the optimal warping, that is optimal pivot mapping sequence can be obtained from backtracking operation (not present in the above algorithm).

It should be noted that in the experiments, another slightly complex DP algorithm with even reduced computational complexity was used where each pivot is considered as a stage. Nevertheless, these two algorithms differ in the definitions of stage and states, and eventually lead to the same solution for two given images.

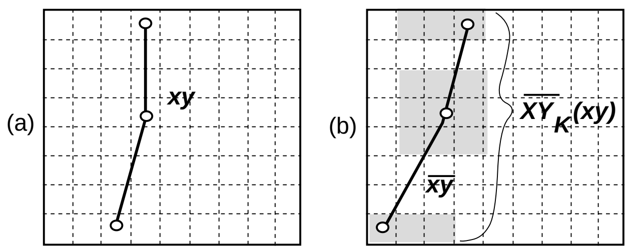

Fig. A·2. An example of (a) state, (b) corresponding set of previous states from where transition is possible.