A Monotonic and Continuous Two-Dimensional Warping Based on Dynamic Programming

Seiichi Uchida and Hiroaki Sakoe
Graduate School of Information Science and Electrical Engineering, Kyushu Univ.
6-10-1 Hakozaki, Higashi-ku, Fukuoka-shi, 812-8581, Japan
{uchida, sakoe}@human.is.kyushu-u.ac.jp

Abstract

A novel two-dimensional warping algorithm is presented which searches for the optimal pixel mapping subject to continuity and monotonicity constraints. These constraints enable us to preserve topological structure in images. The search algorithm is based on dynamic programming (DP). As implementation techniques, acceleration by beam search and excessive warp suppression by penalty and/or range limitation are investigated. Experimental results show that this method provides successful warpings between images.

1. Introduction

Two-dimensional warping, a pixel-to-pixel mapping technique between images is one of the most challenging problems in pattern recognition and image processing areas. Many researchers in these fields have been attracted to the two-dimensional warping from their pure theoretical interests and/or expectations for its wide prospective applicability to real problems. Several two-dimensional warping methods based on dynamic programming (DP) have been proposed [1, 2, 3, 4] motivated by successful applications of DP to the time warping problem in speech recognition. Among them, the method proposed by Levin and Pieraccini[4] is the most promising one, we considered. It searches all possible *monotonic* pixel-to-pixel mappings for the optimal warp with minimum residual error. For practical use, however, there remain two serious problems with their method. First, it often yields considerably unrealistic warp. This means that monotonicity condition alone is insufficient for preserving two-dimensional topological structure in images. Second, since a large number of such unrealistic warps are included in its search space, the amount of computational resources required is prohibitively large.

This paper presents a two-dimensional warping method where a *continuity* condition is newly employed along with the monotonicity condition. The warp obtained by the proposed method approximately preserves the topological structure in images. The optimal warp is searched for by a DP-based algorithm. Due to the continuity condition, its computational complexity is remarkably lower than that of Levin and Pieraccini. As implementation techniques, acceleration by beam search and a penalty-based excessive warp suppression technique are also investigated.

2. Monotonic and continuous two-dimensional warping

2.1. Problem formulation

Consider two images $\boldsymbol{A} = \{a(i,j)\}$ $(i, j = 1, \ldots, N)$ and $\boldsymbol{B} = \{b(x,y)\}$ $(x, y = 1, \ldots, M)$. The optimal monotonic and continuous two-dimensional warp between $\boldsymbol{A}$ and $\boldsymbol{B}$ is defined by the warping function $(x(i,j), y(i,j))$ which minimizes the following criterion function

$$\sum_{i=1}^{N}\sum_{j=1}^{N} |a(i,j) - b(x(i,j), y(i,j))|, \tag{1}$$

subject to the following monotonicity and continuity conditions

$$0 \leq y(i,j) - y(i,j-1) \leq 2, \tag{2}$$
$$|x(i,j) - x(i,j-1)| \leq 1, \tag{3}$$
$$0 \leq x(i,j) - x(i-1,j) \leq 2, \tag{4}$$
$$|y(i,j) - y(i-1,j)| \leq 1. \tag{5}$$

Vertical and horizontal monotonicity and continuity relations between a pixel and its 4-adjacent pixels are

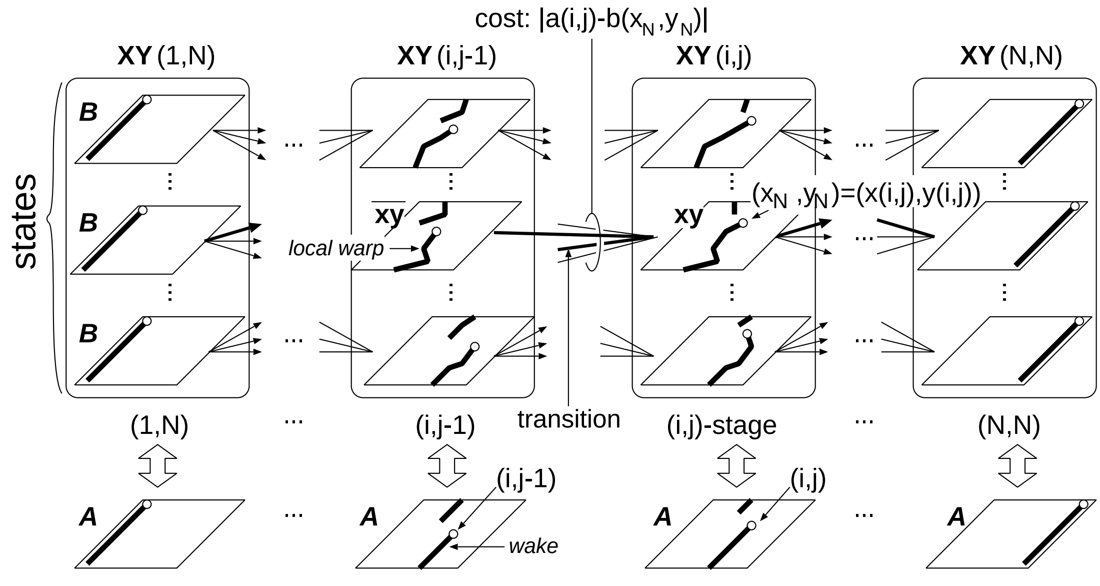

Figure 1: Multi-stage decision process for the two-dimensional warping problem.

preserved by these conditions. Thus, the topological structure in images is preserved after warping. In addition to these conditions, boundary conditions, i.e.,

$$x(1,j) = y(i,1) = 1, \quad x(N,j) = y(i,N) = M \qquad (6)$$

are also used. Let $D(\boldsymbol{A},\boldsymbol{B})$ denote the minimum of the criterion function (1), i.e.,

$$D(\boldsymbol{A},\boldsymbol{B}) = \min_{\substack{x(i,j)\\ y(i,j)}} \sum_{i=1}^{N}\sum_{j=1}^{N} |a(i,j) - b(x(i,j),y(i,j))| \, . \qquad (7)$$

The quantity $D(\boldsymbol{A},\boldsymbol{B})$ gives a superimposing distance between $\boldsymbol{A}$ and optimally deformed $\boldsymbol{B}$, and can be interpreted as the residual error between $\boldsymbol{A}$ and $\boldsymbol{B}$ yielded by the warping. From the view point of structural analysis, the optimal warping function $(x(i,j), y(i,j))$ gives an interpretation of the image $\boldsymbol{A}$ according to the generation model $\boldsymbol{B}$.

2.2. Multi-stage decision process formulation of warping

Figure 1 shows a multi-stage decision process [5] for the monotonic and continuous two-dimensional warping problem. In the following, its detailed specifications, i.e., stage, state, state transition and transition cost, are discussed.

Stage: At stage (i,j), the mapping of the pixel (i,j), i.e., $(x(i,j), y(i,j))$ is decided. Stages are arranged in the order of vertical raster-scanning. In order to satisfy the monotonicity and continuity conditions, the mapping $(x(i,j), y(i,j))$ should be decided depending on not only $(x(i,j-1), y(i,j-1))$ but also $(x(i-1,j), y(i-1,j))$ decided N stages before. For dealing with this Nth-order Markovian nature, it is convenient to consider that stage (i,j) is accompanied by a *wake*, which is a sequence of N consecutive pixels in $\boldsymbol{A}$, i.e.,

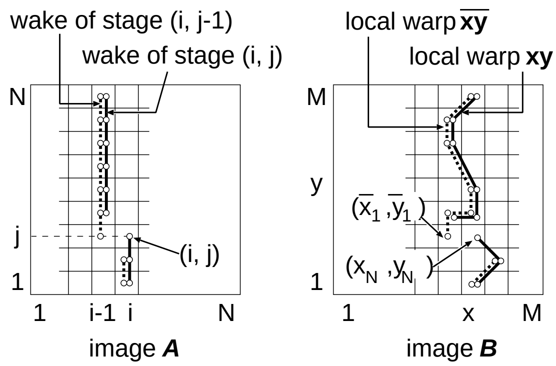

Figure 2: Wakes and local warps of consecutive states.

$$[(i-1,j+1),\ldots,(i-1,N),(i,1),\ldots,(i,j)]\, . \qquad (8)$$

Figure 2 shows the wakes of stages $(i,j-1)$ and (i,j).

State: Each state in stage (i,j) is specified by a unique mapping of the wake of stage (i,j). This mapping, i.e.,

$$\boldsymbol{xy}(i,j) = [(x(i-1,j+1), y(i-1,j+1)), \ldots, (x(i-1,N), y(i-1,N)), (x(i,1), y(i,1)), \ldots, (x(i,j), y(i,j))] \quad (9)$$

(often abbreviated as $\boldsymbol{xy}$) is called *local warp.* (See Fig.2.) Let $\boldsymbol{XY}(i,j)$ denote the set of local warps $\boldsymbol{xy}(i,j)$. Monotonicity and continuity conditions (2) and (3) hold in the individual local warps, while conditions (4) and (5) hold between consecutive local warps $\boldsymbol{xy}(i,j-1)$ and $\boldsymbol{xy}(i,j)$ as discussed in the next part. The local warp $\boldsymbol{xy}$ should also satisfy boundary conditions (6).

State transition: For notational simplicity, we denote local warps $\boldsymbol{xy} \in \boldsymbol{XY}(i,j)$ and $\overline{\boldsymbol{xy}} \in \boldsymbol{XY}(i,j-1)$ in the following manner.

$$\boldsymbol{xy} = [(x_1,y_1), \ldots, (x_k,y_k), \ldots, (x_N,y_N)] \quad (10)$$

$$\overline{\boldsymbol{xy}} = [(\overline{x}_1,\overline{y}_1), \ldots, (\overline{x}_k,\overline{y}_k), \ldots, (\overline{x}_N,\overline{y}_N)]. \quad (11)$$

If the following conditions are satisfied, the transition from state $\overline{\boldsymbol{xy}}$ to $\boldsymbol{xy}$ is allowed.

monotonicity and continuity (from (4), (5)):

$$0 \leq x_N - \overline{x}_1 \leq 2, \quad |y_N - \overline{y}_1| \leq 1, \quad (12)$$

and transition consistency:

$$\overline{x}_k = x_{k-1}, \quad \overline{y}_k = y_{k-1} \qquad 2 \leq k \leq N. \quad (13)$$

Figure 2 shows an example of local warps $\overline{\boldsymbol{xy}}$ and $\boldsymbol{xy}$ of consecutive states. Let $\overline{\boldsymbol{XY}}(\boldsymbol{xy})$ denote the set of states preceding to $\boldsymbol{xy}$. From (12) and (13), its size $|\overline{\boldsymbol{XY}}(\boldsymbol{xy})|$ becomes at most 9.

Transition cost: Again, the mapping of the pixel (i,j), i.e, $(x(i,j), y(i,j))$ is decided at stage (i,j). Therefore, by considering the criterion function (1), the cost

$$d(i,j,\boldsymbol{xy}) = |a(i,j) - b(x_N, y_N)| \quad (14)$$

is imposed to the transition to $\boldsymbol{xy}$.

2.3. DP algorithm

The optimal decision sequence of the foregoing multi-stage decision process can be obtained by a DP-based algorithm. Its DP-equation is described as follows

Initialization:

$$g(1,N,\boldsymbol{xy}) = \sum_{k=1}^{N} |a(1,k) - b(x_k,y_k)| \quad (15)$$

Recursion:

$$g(i,j,\boldsymbol{xy}) = d(i,j,\boldsymbol{xy}) + \min_{\overline{\boldsymbol{xy}} \in \overline{\boldsymbol{XY}}(\boldsymbol{xy})} \begin{cases} g(i,j-1,\overline{\boldsymbol{xy}}) & j \neq 1 \\ g(i-1,N,\overline{\boldsymbol{xy}}) & j = 1 \end{cases} \quad (16)$$

Termination:

$$D(\boldsymbol{A},\boldsymbol{B}) = \min_{\boldsymbol{xy} \in \boldsymbol{XY}(N,N)} g(N,N,\boldsymbol{xy}) \quad (17)$$

where $g(i,j,\boldsymbol{xy})$ denotes the accumulated cost along the optimal path to $\boldsymbol{xy} \in \boldsymbol{XY}(i,j)$. The optimal decision sequence, i.e., the optimal warp is obtained by backtracking.

The time complexity of the algorithm is proportional to the total number of possible transitions in the decision process. This total number of transitions is given by the product of the numbers of stages, states belonging to a stage, and possible transitions from a state. Hence, in the case of $N = M$, the time complexity of proposed algorithm is $O(N^3 9^N)$ which is the product of $O(N^2)$, $O(N 9^N)$ and $O(1)$. This order is remarkably lower than $O(N^{4N})$ of Levin and Pieraccini, coming from the product of $O(N)$,$O(N^{2N})$ and $O(N^{2N})$.

3. Implementation techniques

3.1. Acceleration by beam search

As shown in Section 2.3, the time complexity of the proposed algorithm is lower than that of Levin and Pieraccini, but it is still exponential in N. Here, a polynomial-time approximation algorithm with beam search [6, 7] is introduced. Beam search is a technique based on eliminating searching paths with less possibilities for the optimal path (called "pruning"). The solution becomes sub-optimal because the pruning is based on a local criterion. However, it will be shown from experimental results in Section 4 that sub-optimal solutions are sufficient in most cases.

It is simple to apply beam search to the algorithm described in Section 2.3. The process at stage (i,j) consists of the following two steps. First, R states with smaller accumulated cost $g(i,j-1,\overline{\boldsymbol{xy}})$ are taken into account in the further search. The number R is called beam size. Second, accumulated costs $g(i,j,\boldsymbol{xy})$ are calculated from their $g(i,j-1,\overline{\boldsymbol{xy}})$. Clearly, the number of states whose costs are calculated is at most $9R$ in every stage and therefore the time complexity of this algorithm is $O(N^3 R)$, i.e., polynomial order.

3.2. Excessive warp suppression by penalty and/or range limitation

In some cases warps given by the proposed algorithms yield excessive deformation, it is feared. This will be due to 1) taking no account of the degree of global deformation, and/or 2) side effects of the pruning. Use of a penalty to control the transition $\overline{\boldsymbol{xy}} \rightarrow \boldsymbol{xy}$, is a simple practical solution to the problem. Here, the following penalty is imposed to the transition cost $d(i, j, \boldsymbol{xy})$

$$P(i, j, \overline{\boldsymbol{xy}}, \boldsymbol{xy}) = |x_N - x_{N-1}| + |x_N - \overline{x}_1 - 1| + |y_N - y_{N-1} - 1| + |y_N - \overline{y}_1|. \quad (18)$$

Now the DP-equation becomes

$$g(i, j, \boldsymbol{xy}) = d(i, j, \boldsymbol{xy}) + \min_{\overline{\boldsymbol{xy}} \in \overline{\boldsymbol{XY}}(\boldsymbol{xy})} [g(i, j-1, \overline{\boldsymbol{xy}}) + \alpha \cdot P(i, j, \overline{\boldsymbol{xy}}, \boldsymbol{xy})] \quad (19)$$

where α is a positive constant for controlling penalty gain.

Another excess warp suppression technique is warp range limitation. This is defined as

$$|x(i, j) - i| \leq w, \quad |y(i, j) - j| \leq w, \quad (20)$$

where w is a positive integer for limiting warp range.

4. Experimental results

Figure 3 shows experimental results of five sub-optimal warps between 32×32 8bit images. Beam size R, penalty gain α and warp range w were fixed at 1000, 10 and 5, respectively. It is shown from deformed mesh patterns representing the warping functions $(x(i, j), y(i, j))$ that monotonicity and continuity are preserved after warping. Deformed images $b(x(i, j), y(i, j))$ show that all $\boldsymbol{B}$'s are successfully deformed to match $\boldsymbol{A}$'s in spite of the approximation algorithm used.

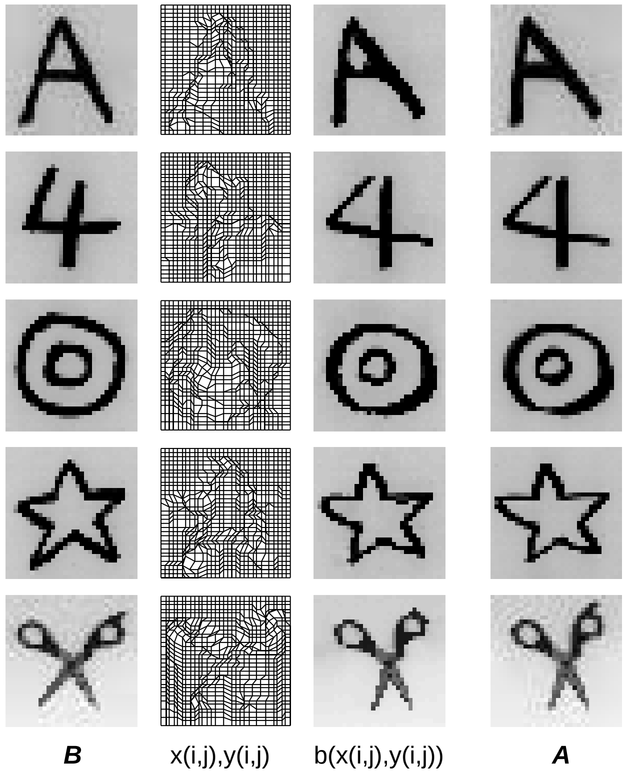

Figure 3: Experimental results of warping.

5. Conclusion

A monotonic and continuous two-dimensional warping method was presented. The problem was formulated as a pixel-to-pixel correspondence optimization subject to two-dimensional monotonicity and continuity conditions. The optimal warp is obtained by DP as the optimal decision sequence of an Nth-order Markovian multi-stage decision process. The time complexity of the proposed algorithm was remarkably reduced from that of Levin and Pieraccini by newly incorporated continuity conditions. The efficiency of the proposed algorithm was improved by applying the beam search technique and the penalty function. Experimental results indicated the validity of the proposed method.

References

[1] R. K.Moore. A dynamic programming algorithm for the distance between two finite areas. *IEEE Trans. Pattern Anal. Machine Intell.*, PAMI-1(1):86–88, 1979.

[2] E. Tanaka and Y. Kikuchi. A metric between pictures. *Trans. IEICE Jpn*, J63-D(12):1018–1025, 1980.

[3] C.M.Wu, R.M.Owens, and M.J.Irwin. Distortion processing in image matching problems. *Proc ICASSP*, pages 2181–2184, 1990.

[4] E. Levin and R. Pieraccini. Dynamic planar warping for optical character recognition. *Proc ICASSP*, pages III 149–152, 1992.

[5] R. Bellman. *Adaptive Control Processes.* Princeton University Press, New Jersey, 1961.

[6] H. Sakoe, H. Fujii, K. Yoshida, and M. Watari. A high-speed dp-matching algorithm based on frame synchronization, beam search and vector quantization. *Systems and Computers in Japan*, 20(11):33–45, 1989.

[7] H. Ney, D. Mergel, A. Noll, and A. Paeseler. Data driven search organization for continuous speech recognition. *IEEE Trans. Signal Proc.*, 40(2):272–281, 1992.