Analytical Dynamic Programming Matching

Seiichi Uchida, Satoshi Hokahori, and Yaokai Feng

Kyushu University, Fukuoka, Japan

Abstract. In this paper, we show that the truly two-dimensional elastic image matching problem can be solved analytically using dynamic programming (DP) in polynomial time if the problem is formulated as a maximum a posteriori problem using Gaussian distributions for the likelihood and prior. After giving the derivation of the analytical DP matching algorithm, we evaluate its performance on handwritten character images containing various nonlinear deformations, and compare other elastic image matching methods.

1 Introduction

Elastic matching is one of the most fundamental tools for pattern recognition and computer vision. For one-dimensional (i.e., sequential) patterns, elastic matching detects and compensates, for example, any temporal fluctuations. For two-dimensional patterns (i.e., images), elastic matching also detects and compensates various geometric deformations. Owing to these abilities, elastic matching has been applied to recognition tasks, deformation analysis, pattern alignment, image compression, stereo, and so on.

Elastic matching is formulated as an optimization problem of a warping function between two patterns. As such, the property of elastic matching is determined by the problem formulation and the optimization method. These two factors are not independent — the optimization method is selected according to the formulation. For example, if the problem is formulated as a combinatorial optimization problem, a combinatorial optimization method will be used.

Throughout this paper, we focus on dynamic programming (DP) as the optimization method for elastic image matching. As detailed later, DP has promising properties for elastic matching. In fact, DP has undoubtedly been the most established optimization method for sequential patterns since the late 1960s.

For elastic *image matching*, however, DP has not been fully utilized. Consider a combinatorial optimization problem of elastic image matching with a truly two-dimensional warping ability. Unfortunately, as this is an NP-hard problem, DP (as well as other optimizers) cannot solve the problem in polynomial time. Consequently, the warping ability is severely restricted in reducing the computation. This limitation results in so-called pseudo 2D elastic matching that cannot compensate vertical and horizontal deformations simultaneously. In other words, it cannot even compensate rotation.

In this paper, a truly two-dimensional polynomial-time DP matching method, called *analytical DP matching*, is presented, which is very different from the

conventional combinatorial DP matching. The performance thereof is evaluated qualitatively and quantitatively on handwritten character images. The key idea is to formulate the elastic image matching problem as a maximum a posteriori (MAP) problem with a Gaussian likelihood and a Gaussian prior. The quadratic nature of the Gaussian distributions helps us to derive an analytical solution using DP. Consequently, analytical DP matching can obtain the globally optimal solution of the MAP problem with a polynomial-time computation while retaining the truly two-dimensional warping ability.

1.1   Related Work

DP [1–3] has been applied to various tasks in pattern recognition and computer vision. Nowadays, DP is considered a classic optimization method and there are several fascinating alternatives, such as graph cut, belief propagation (or message passing), and so on [4]. Nevertheless, many researchers still choose DP for their current tasks because of its conciseness, versatility, and ability to obtain the globally optimal solution. In fact, DP has recently been employed in studies on tracking [5], stereo [6–8], and elastic image matching [9].

Elastic matching is a typical application of DP. As previously noted, the DP-based matching algorithm, called DP matching or dynamic time warping, has been widely and successfully applied to sequential pattern recognition tasks since the late 1960s. In fact, DP matching (and its stochastic extension, i.e., Hidden Markov Models) is a standard in speech recognition [10, 11] and on-line character recognition [12].

It is quite natural to try to extend the sequential DP matching algorithm to a two-dimensional one. Several researchers [13, 14] have developed DP algorithms for truly two-dimensional elastic image matching, but have encountered the inherent NP-hardness of the problem [15].

Because of this computational intractability, conventional DP-based elastic image matching algorithms employ various approximation strategies, the most popular of which is the limitation of matching flexibility. In fact, we can find many pseudo 2D elastic matching algorithms, such as [9]. Another strategy is partial omission of the mutual dependency between 4-adjacent pixels (e.g., the tree representation in [6, 8, 16]). It is also popular for introducing local search techniques, such as pruning (or beam-search) and coarse-to-fine strategies [17], at the cost of global optimality. Notwithstanding these strategies, there is currently no practical DP algorithm that can provide both globally optimal and truly two-dimensional elastic matching.

All the conventional DP-based elastic matching algorithms (apart from four exceptions noted later) have always used DP as a combinatorial breadth-first search method, i.e., a combinatorial optimization method. This is confirmed by the fact that a very recent survey [3] reported only combinatorial (i.e., discrete) DP algorithms. Even if an optimization problem is originally formulated as a continuous variational problem, it is discretized and then solved by DP as a combinatorial optimization problem [2, 10].

This fact is somewhat peculiar since DP was originally developed as a continuous optimization method to obtain the solution efficiently using an *analytical* strategy [1]. Such analytical solutions have rarely been utilized even in other computer vision and pattern recognition problems. To the best of our knowledge, there are only four studies in which DP has been utilized as an analytical solver. Angel [18] used analytical DP for smooth interpolation. Serra and Berthod [20] and Munich and Perona [21] used it for nonlinear alignment of one-dimensional patterns. Finally, Uchida et al. [22] used it for object tracking.

1.2   Our Contribution

Our main contribution is regenerating the classic DP-based elastic matching algorithm, the combinatorial breadth-first search method that has not been considered since its introduction in the late 1960s. The proposed algorithm, called analytical DP matching, does not include the combinatorial search that becomes intractable for elastic image matching problems. Instead, it utilizes DP to provide an analytical solution, thereby successfully reducing the computational complexity from an exponential order to a polynomial one.

The analytical DP matching algorithm is derived by formulating the elastic matching problem as a MAP problem with Gaussian distributions. The quadratic nature of the Gaussian distributions enables the globally optimal solution of the MAP problem to be obtained with $O(I^4)$ computations for $I \times I$ images. Since this is an analytical solution, it is not necessary to consider either convergence or the initial value issue. It is noteworthy that this algorithm has the potential to be combined with other optimization methods based on a sequential decision process, such as tree-reweighted message passing [19].

Since the problem is formulated as a quadratic optimization problem, it can be solved analytically by a more popular closed-form solution, precisely like least-mean-square (LMS) problems. In this case, the solution requires $O(I^6)$ computations to deal with a large $(O(I^2) \times O(I^2))$ matrix. In contrast, our analytical DP matching algorithm utilizes a column-wise recursive formulation that provides a more efficient solution with $O(I^4)$ computations and a far smaller $(O(I) \times O(I))$ matrix.

2   Formulation of Elastic Image Matching Problem

2.1   Elastic Image Matching as a MAP Problem

Elastic image matching between a pair of $I \times I$ images [1], $\boldsymbol{X} = \{x_{i,j}|i,j = 1,\ldots,I\}$ and $\boldsymbol{Y} = \{y_{u,v}|u,v = 1,\ldots,I\}$, is an optimization problem of the warping function $\boldsymbol{W} = \{\boldsymbol{w}_{i,j}\}$, where $\boldsymbol{w}_{i,j} = (u_{i,j}, v_{i,j})^T$ denotes that pixel $(i,j)^T$ on $\boldsymbol{X}$ corresponds to $(u_{i,j}, v_{i,j})^T$ on $\boldsymbol{Y}$. Hereafter, we assume the boundary conditions, $u_{1,j} = 1, u_{I,j} = I, v_{i,1} = 1, v_{i,I} = I$.

---

[1] We can easily extend the following discussion to arbitrary size images.

We formulate the optimization problem of $\boldsymbol{W}$ as the following MAP problem:

$$\overline{\boldsymbol{W}} = \underset{\boldsymbol{W}}{\operatorname{argmax}} \, P(\boldsymbol{W}|\boldsymbol{X}, \boldsymbol{Y}). \tag{1}$$

According to Bayes' rule,

$$\underset{\boldsymbol{W}}{\operatorname{argmax}} \, P(\boldsymbol{W}|\boldsymbol{X}, \boldsymbol{Y}) = \underset{\boldsymbol{W}}{\operatorname{argmax}} \, P(\boldsymbol{X}, \boldsymbol{Y}|\boldsymbol{W})P(\boldsymbol{W}), \tag{2}$$

where $P(\boldsymbol{X}, \boldsymbol{Y}|\boldsymbol{W})$ is a likelihood that evaluates the similarity between $\boldsymbol{X}$ and $\boldsymbol{Y}$ under the warping function $\boldsymbol{W}$, and $P(\boldsymbol{W})$ is a prior of $\boldsymbol{W}$.

We assume the following Gaussian likelihood $P(\boldsymbol{X}, \boldsymbol{Y}|\boldsymbol{W})$:

$$P(\boldsymbol{X}, \boldsymbol{Y}|\boldsymbol{W}) = \prod_{i,j} \mathcal{N}(\boldsymbol{w}_{i,j}; \, \boldsymbol{\mu}_{i,j}, \boldsymbol{\Sigma}_{i,j}), \tag{3}$$

where $\mathcal{N}(\cdot)$ is a two-dimensional Gaussian distribution and $\boldsymbol{\mu}_{i,j}$ and $\boldsymbol{\Sigma}_{i,j}$ are its mean vector and covariance matrix, respectively. Each pixel $(i, j)^T$ on $\boldsymbol{X}$ has its own Gaussian distribution on $\boldsymbol{Y}$, which evaluates the pixel-wise similarity between $(i, j)^T$ and $\boldsymbol{w}_{i,j} = (u_{i,j}, v_{i,j})^T$. The parameters $\boldsymbol{\mu}_{i,j}$ and $\boldsymbol{\Sigma}_{i,j}$ are estimated before optimizing $\boldsymbol{W}$. As the prior $P(\boldsymbol{W})$, we use a smoothness function [2],

$$P(\boldsymbol{W}) = \prod_{i,j} \mathcal{N}(\boldsymbol{w}_{i,j} - \boldsymbol{w}_{i,j-1}; \, \boldsymbol{0}, \lambda\boldsymbol{I}) \cdot \mathcal{N}(\boldsymbol{w}_{i,j} - \boldsymbol{w}_{i-1,j}; \, \boldsymbol{0}, \lambda\boldsymbol{I}). \tag{4}$$

The first and second Gaussian distributions in (4) evaluate intra- and inter-column smoothness, respectively. If the constant λ is set to a larger value, the smoothing effect by the prior becomes weaker. By taking the logarithm of (2), our objective function $F(\boldsymbol{W})$ is derived as follows:

$$F(\boldsymbol{W}) \equiv \lambda \left[\log P(\boldsymbol{X}, \boldsymbol{Y}|\boldsymbol{W}) + \log P(\boldsymbol{W}) \right] \tag{5}$$

We now introduce the *matrix-vector formulation* [24], or column-wise formulation, where $\boldsymbol{W}$ is treated as a sequence of $2I$-dimensional vectors, $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_i, \ldots, \boldsymbol{w}_I$, where $\boldsymbol{w}_i = (\boldsymbol{w}_{i,1}, \boldsymbol{w}_{i,2}, \ldots, \boldsymbol{w}_{i,j}, \ldots, \boldsymbol{w}_{i,I-1}, \boldsymbol{w}_{i,I})^T$. Then the objective function $F(\boldsymbol{W})$ becomes

$$F(\boldsymbol{W}) = F(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_I) = \lambda \sum_{i=1}^{I} d_i(\boldsymbol{w}_i) + \sum_{i=1}^{I} \eta(\boldsymbol{w}_i) + \sum_{i=2}^{I} \rho(\boldsymbol{w}_i, \boldsymbol{w}_{i-1}). \tag{6}$$

The first term of (6) is the log likelihood and is defined as

$$d_i(\boldsymbol{w}_i) = \sum_{j=1}^{I} \boldsymbol{w}_{i,j}^T \boldsymbol{P}_{i,j} \boldsymbol{w}_{i,j} + \boldsymbol{q}_{i,j}^T \boldsymbol{w}_{i,j} + r_{i,j} = \boldsymbol{w}_i^T \boldsymbol{P}_i \boldsymbol{w}_i + \boldsymbol{q}_i^T \boldsymbol{w}_i + r_i, \tag{7}$$

---

[2] It is possible to use an arbitrary covariance matrix in these priors instead of $\lambda\boldsymbol{I}$. It is also possible to use different covariance matrices at different $(i, j)^T$. Although such priors can represent various deformation tendencies, we use the most general prior of (4) throughout this paper to simplify our algorithm derivation.

where $\boldsymbol{P}_i = \text{diag}[\boldsymbol{P}_{i,1}, \ldots, \boldsymbol{P}_{i,j}, \ldots, \boldsymbol{P}_{i,I}]$, $\boldsymbol{q}_i = (\boldsymbol{q}_{i,1}, \ldots, \boldsymbol{q}_{i,j}, \ldots, \boldsymbol{q}_{i,I})^T$, and $r_i = \sum_{i=1}^{I} r_{i,j}$, with these coefficients derived from $\log P(\boldsymbol{X}, \boldsymbol{Y}|\boldsymbol{W})$, i.e., $\boldsymbol{\Sigma}_{i,j}$ and $\boldsymbol{\mu}_{i,j}$. Since $\boldsymbol{\Sigma}_{i,j}$ and $\boldsymbol{\mu}_{i,j}$ are pre-determined as noted above, the coefficients $\boldsymbol{P}_i$, $\boldsymbol{q}_i$, and r_i are also pre-determined. The second and third terms of (6) are the log priors for intra- and inter-column smoothness, respectively:

$$\eta(\boldsymbol{w}_i) = \boldsymbol{w}_i^T \boldsymbol{H} \boldsymbol{w}_i, \tag{8}$$

$$\rho(\boldsymbol{w}_i, \boldsymbol{w}_{i-1}) = (\boldsymbol{w}_i - \boldsymbol{w}_{i-1})^2, \tag{9}$$

where $\boldsymbol{H}$ is a constant matrix.

The minimization problem of (6) is a quadratic (i.e., convex) problem and can thus be analytically solved like LMS problems using a system of $O(I^2)$ linear equations derived by partial differentiation of (6) by each of $2I^2$ variables $\{(u_{i,j}, v_{i,j})\}$. The solution, however, requires $O(I^6)$ computations, because we must deal with a large non-diagonal $O(I^2) \times O(I^2)$ coefficient matrix for the system of linear equations. In the following section, we derive a more efficient $O(I^4)$ algorithm based on DP.

3   Analytical DP Matching

3.1   Derivation of DP Recursion

Similar to Angel [18], we introduce function $f_i(\boldsymbol{w}_{i-1})$, which is defined as

$$f_i(\boldsymbol{w}_{i-1}) = \min_{\boldsymbol{w}_i, \ldots, \boldsymbol{w}_I} \sum_{k=i}^{I} [\lambda d_k(\boldsymbol{w}_k) + \eta(\boldsymbol{w}_k) + \rho(\boldsymbol{w}_k, \boldsymbol{w}_{k-1})]. \tag{10}$$

Note that the minimum value of F is represented using f_2 as follows:

$$\min F(\boldsymbol{W}) = \min_{\boldsymbol{w}_1}[\lambda d_1(\boldsymbol{w}_1) + \eta(\boldsymbol{w}_1) + f_2(\boldsymbol{w}_1)] = \lambda d_1(\overline{\boldsymbol{w}}_1) + \eta(\overline{\boldsymbol{w}}_1) + f_2(\overline{\boldsymbol{w}}_1), \tag{11}$$

where $\overline{\boldsymbol{w}}_1$ denotes $\boldsymbol{w}_1$ giving the minimum of the first equation.

According to the principle of optimality [1], (10) can be rewritten as the following recursive equation, known as *DP recursion.*

$$f_i(\boldsymbol{w}_{i-1}) = \min_{\boldsymbol{w}_i} [\lambda d_i(\boldsymbol{w}_i) + \eta(\boldsymbol{w}_i) + \rho(\boldsymbol{w}_i, \boldsymbol{w}_{i-1}) + f_{i+1}(\boldsymbol{w}_i)]. \tag{12}$$

This recursion indicates that the two-dimensional optimization problem of elastic image matching can be solved as the sequential optimization problem $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_i, \ldots, \boldsymbol{w}_I$ by virtue of the column-wise formulation.

If we use the conventional combinatorial DP algorithm, we first need to discretize $\boldsymbol{w}_i$ as a $2I$-dimensional integer vector, and then calculate the recursion (12) for *all possible* $\boldsymbol{w}_i$ from $i = 1$ to I. Clearly, this is computationally intractable because the number of possible $\boldsymbol{w}_i$ is an exponential order of I. Instead, we use DP as an analytical solver while fully utilizing the fact that the terms of (12) are quadratic in nature and thus differentiable with respect to $\boldsymbol{w}_i$.

Input: Coefficients: $\{\boldsymbol{P}_i, \boldsymbol{q}_i, r_i \mid i = 1, \ldots, I\}$ and $\lambda \in \Re^+$.
Output: $\overline{\boldsymbol{W}} = \overline{\boldsymbol{w}}_1, \ldots, \overline{\boldsymbol{w}}_i, \ldots, \overline{\boldsymbol{w}}_I$ and $\min F$.
Step 1: Initial condition   Obtain $(\boldsymbol{A}_I, \boldsymbol{b}_I, c_I)$.
Step 2: DP recursion
     For $i = I - 1$ downto 2: Obtain $(\boldsymbol{A}_i, \boldsymbol{b}_i, c_i)$ from $(\boldsymbol{A}_{i+1}, \boldsymbol{b}_{i+1}, c_{i+1})$ by (15).
Step 3: Termination
     Obtain $\overline{\boldsymbol{w}}_1$ from $\partial[\lambda d_1(\boldsymbol{w}_1) + \eta(\boldsymbol{w}_1) + f_2(\boldsymbol{w}_1)]/\partial \boldsymbol{w}_1 = 0$.
     $\min F = \lambda d_1(\overline{\boldsymbol{w}}_1) + \eta(\overline{\boldsymbol{w}}_1) + f_2(\overline{\boldsymbol{w}}_1)$.
Step 4: Backtrack   For $i = 2$ to I: Obtain $\overline{\boldsymbol{w}}_i$ by (14) with $\boldsymbol{w}_{i-1} = \overline{\boldsymbol{w}}_{i-1}$.

Fig. 1. Pseudo-code for analytical DP matching.

3.2   Solution using Analytical DP

The most important fact in deriving the proposed method is that all the d_i, as well as η and ρ are quadratic functions of $\boldsymbol{w}_i$ as indicated by (7), (8), and (9), respectively, and thus differentiable with respect to $\boldsymbol{w}_i$. This means that $f_i(\boldsymbol{w}_{i-1})$ is also a quadratic function [3] and can therefore be represented as

$$f_i(\boldsymbol{w}_{i-1}) = \boldsymbol{w}_{i-1}^T \boldsymbol{A}_i \boldsymbol{w}_{i-1} + \boldsymbol{b}_i^T \boldsymbol{w}_{i-1} + c_i, \tag{13}$$

where $\boldsymbol{A}_i$ is a $2I \times 2I$ matrix, $\boldsymbol{b}_i$ is a $2I$-dimensional vector, and c_i is a scalar, all of which are determined by optimizing $\{\boldsymbol{w}_i\}$. By substituting (13) into (12), and then differentiating with respect to $\boldsymbol{w}_i$, the optimal $\boldsymbol{w}_i = \overline{\boldsymbol{w}}_i$, which gives the minimum of (12), is derived as:

$$\overline{\boldsymbol{w}}_i = [\lambda \boldsymbol{P}_i + \boldsymbol{A}_{i+1} + \boldsymbol{H} + \boldsymbol{I}]^{-1} (\boldsymbol{w}_{i-1} - (\lambda \boldsymbol{q}_i + \boldsymbol{b}_{i+1})/2), \tag{14}$$

where $\boldsymbol{I}$ is the identity matrix.

By substituting $\overline{\boldsymbol{w}}_i$ into (12) and then comparing with (13), we have the recursive procedure for obtaining $(\boldsymbol{A}_i, \boldsymbol{b}_i, c_i)$ from $(\boldsymbol{A}_{i+1}, \boldsymbol{b}_{i+1}, c_{i+1})$ as follows:

$$\left.\begin{array}{l} \boldsymbol{A}_i = \boldsymbol{I} - [\lambda \boldsymbol{P}_i + \boldsymbol{A}_{i+1} + \boldsymbol{H} + \boldsymbol{I}]^{-1} \\ \boldsymbol{b}_i = [\boldsymbol{I} - \boldsymbol{A}_i](\lambda \boldsymbol{q}_i + \boldsymbol{b}_{i+1}) \\ c_i = -(\lambda \boldsymbol{q}_i + \boldsymbol{b}_{i+1})^T \boldsymbol{b}_i/4 + c_{i+1} + \lambda r_i \end{array}\right\}. \tag{15}$$

The complete algorithm for analytical DP matching is summarized in Fig. 1. After finding the initial value $(\boldsymbol{A}_I, \boldsymbol{b}_I, c_I)$, the value $(\boldsymbol{A}_i, \boldsymbol{b}_i, c_i)$ is calculated from $(\boldsymbol{A}_{i+1}, \boldsymbol{b}_{i+1}, c_{i+1})$ for $i = I - 1$ downto 2 according to the recursive procedure (15). While the derivation of $(\boldsymbol{A}_I, \boldsymbol{b}_I, c_I)$ is not detailed here, it can be derived by comparing (10) with (13) at $i = I$. The optimal correspondence $\overline{\boldsymbol{w}}_i$ is calculated repeatedly by (14) from $\overline{\boldsymbol{w}}_1$. It should be noted that the optimal correspondence $\{\overline{\boldsymbol{w}}_i\}$ is the globally optimal solution of (6). The computational complexity of analytical DP matching is $O(I^4)$, i.e., a polynomial order of I, and is dominated by the $O(I^3)$ computations to obtain the $2I \times 2I$ inverse matrix in (15). Since
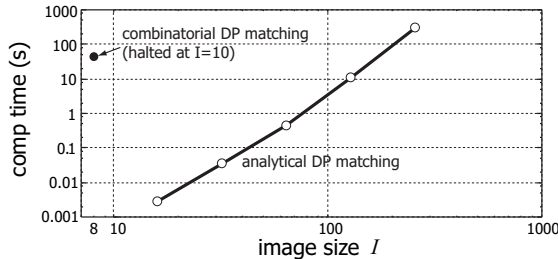
---

[3] This can be proved inductively.

Fig. 2. Comparison of computation times for analytical DP matching and conventional combinatorial DP matching.

there are generally fewer than $O(I^4)$ computations for $\boldsymbol{P}_i$, $\boldsymbol{q}_i$, and r_i, the total number of computations is still $O(I^4)$.

Figure 2 shows the average computation times for analytical DP matching at $I = 16, 32, \ldots, 256$ on a personal computer. This graph coincides with the theoretical computational complexity, i.e., $O(I^4)$. The conventional truly two-dimensional DP matching algorithm [14] required 41 s at $I = 8$ and halted at $I = 10$ owing to lack of memory. These comparative results show that analytical DP matching is far more efficient than the conventional DP matching algorithm.

4   Performance Evaluation

For a qualitative and quantitative performance evaluation, several experiments were carried out using handwritten digit images from MNIST [23]. MNIST comprises 60,000 training samples and 10,000 test samples. Each image is gray-scale with size 28×28.

There are several benefits of using handwritten character images for performance evaluation. (i) Elastic matching of character images is often more ambiguous and difficult than general object images because character images are binary patterns showing only curves (i.e., strokes). (ii) Through a character recognition experiment using the elastic matching distance (min F), the "over-fitting" phenomenon can be strictly observed. This is because character images from *different* classes often become similar to each other through over-fitting, and this can be detected as a misrecognition result. (iii) Since handwritten character images have typically been the target of elastic image matching, there are many past results of recognition experiments using the same database, especially MNIST.

Etohfs method [25] was used to determine $\boldsymbol{P}_{i,j}$, $\boldsymbol{q}_{i,j}$, and $r_{i,j}$. Figure 3(a) shows the pixel-wise log-likelihood functions (i.e., $\boldsymbol{w}_{i,j}^T \boldsymbol{P}_{i,j} \boldsymbol{w}_{i,j} + \boldsymbol{q}_{i,j}^T \boldsymbol{w}_{i,j} + r_{i,j}$) for a pair of "2". It is noteworthy that the major axis direction is often similar to the direction of the character stroke. For example, consider the correspondence illustrated by the thick orange arrow in Fig. 3(a). This indicates that point $(i,j)^T$, which lies on the "/"-shaped stroke on $\boldsymbol{X}$, has a high probability of being matched to point $(u,v)^T$ along the "/"-shaped stroke of $\boldsymbol{Y}$.

Figure 3(b) shows the results of analytical DP matching on several handwritten character image pairs. The images at either end are $\boldsymbol{Y}$ and $\boldsymbol{X}$, while
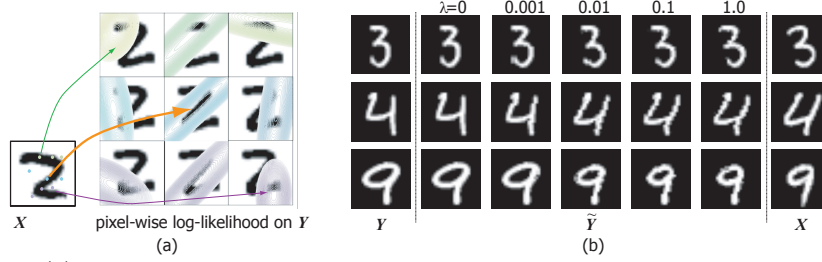
Fig. 3. (a) Pixel-wise log-likelihood for a pair of $\boldsymbol{X}$ and $\boldsymbol{Y}$. For better visibility, only nine functions are plotted separately. (b) Matching results for different λ values.
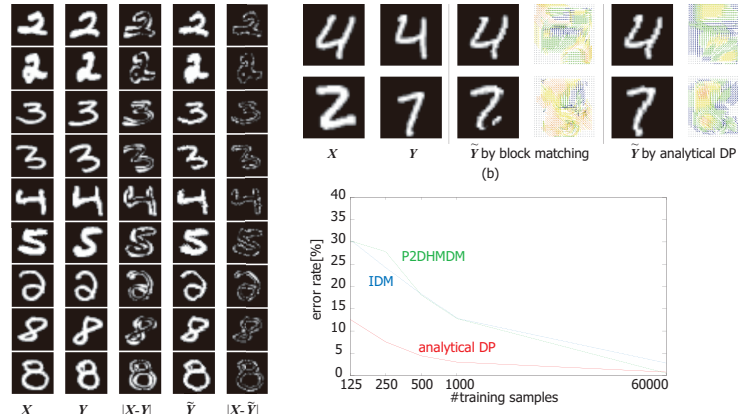

Fig. 4. (a) Visualization of matching accuracy using difference image. (b) Effect of global optimization. (c) Recognition rate of 10,000 handwritten digit images.

the other five images are the matching results $\widetilde{\boldsymbol{Y}} = \{y_{\overline{\boldsymbol{w}}_{i,j}}|i,j = 1,\ldots,I\}$ for five different values of λ. When $\lambda = 0$, the objective function is governed by the smoothness prior η, ρ, and the boundary conditions. Thus, $\widetilde{\boldsymbol{Y}} = \boldsymbol{Y}$. As λ increases, $\widetilde{\boldsymbol{Y}}$ becomes more similar to $\boldsymbol{X}$.

Figure 4(a) visualizes the accuracy of the proposed algorithm using difference images between $\boldsymbol{X}$ and $\widetilde{\boldsymbol{Y}}$. The fact that $\widetilde{\boldsymbol{Y}} \sim \boldsymbol{X}$ indicates that $\boldsymbol{Y}$ was appropriately fitted to $\boldsymbol{X}$. It also indicates that the matching flexibility is truly two-dimensional.

Figure 4(b) compares the global optimization result using the proposed algorithm with a local optimization result obtained by block matching. The latter result is equivalent to $(\hat{u}_{i,j}, \hat{v}_{i,j})^T$, which was determined individually at each block. (Thus, it is equivalent to the result obtained by IDM [9].) In Fig. 4(b), the result of matching two images from the same class ("4") is shown. Although the warped images are similar to each other, the pixel correspondences are different. The correspondence by block matching is somewhat scattered and excessive, whereas that by the proposed algorithm is smooth. This result is also confirmed by the other result, where block matching caused over-fitting between different classes ("2" and "7"). For the quantitative performance evaluation, a recognition experiment was carried out. Each of the 10,000 test samples ($\boldsymbol{X}$) was matched

to all the training samples ($\boldsymbol{Y}$) and their matching distances $\min F$ were used for discrimination. The recognition result of $\boldsymbol{X}$ was finally determined using the 3-nearest neighbor method. The number of training samples was changed from 125 to 60,000, that is, all the training samples in MNIST.

For a comparative evaluation, IDM and P2DHMDM [9] were also used as other promising elastic image matching methods. It was reported in [23] that these methods achieved the best recognition performance on MNIST of all the elastic image matching methods. Their discrimination was done under the same condition [4] as the proposed algorithm. Figure 4(c) shows the recognition rates using IDM, P2DHMDM, and the proposed algorithm. The proposed algorithm outperforms the others especially with fewer training samples. This result indicates that the proposed algorithm can match images with large differences, while at the same time avoiding over-fitting.

5   Conclusion

An analytical DP matching algorithm was proposed for elastic image matching. The proposed algorithm was derived by formulating the matching problem as a MAP problem with a Gaussian likelihood and Gaussian priors. By virtue of the quadratic nature of the Gaussian distributions, DP can be used as an analytical solver that obtains the globally optimal solution of the MAP problem with $O(I^4)$ computations for $I \times I$ images. On the other hand, if DP is used as a conventional combinatorial solver, it requires an exponential number of computations. The computational efficiency of the proposed algorithm was shown through experimental results.

As the discussion in this paper is somewhat general, some specialization would be necessary for each specific image matching problem. In particular, we can apply the proposed algorithm to various images other than handwritten character images. Use of a more sophisticated and less ambiguous pixel value, such as a SIFT image [26], instead of the simple gray-scale value, would be useful to obtain a more reliable pixel-wise likelihood. As noted previously, it is possible to elaborate the prior to incorporate a pattern specific deformation tendency in its covariance matrix. It is also possible to incorporate hard constraints to fix $\boldsymbol{w}_{i,j}$ at arbitrary pixels, like the boundary conditions, by virtue of a property of DP. In other words, given sparse pixel-to-pixel correspondences (by, for example, SIFT matching), the proposed method can provide the remaining correspondences optimally.

References

1. R. Bellman and S. Dreyfus, *Applied Dynamic Programming*, Princeton University Press, 1962.

---

[4] Keysers et al. [9] indicated using Sobel filtered images for better performance. Although this would also be beneficial for the proposed algorithm, gray values were used here for the sake of simpler analysis.

2. A. A. Amini, T. E. Weymouth, and R. C. Jain, "Using dynamic programming for solving variational problems in vision," *PAMI*, 12(9), 1990.
3. P. F. Felzenszwalb, and R. Zabih, "Dynamic programming and graph algorithms in computer vision," *PAMI*, 33(4), 2011.
4. R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother, "A comparative study of energy minimization methods for Markov random fields," *ECCV*, 2006.
5. A. Buchanan and A. Fitzgibbon, "Interactive feature tracking using K-D trees and dynamic programming," *CVPR*, 2006.
6. O. Veksler, "Stereo correspondence by dynamic programming on a tree," *CVPR*, 2005.
7. J. C. Kim, K. M. Lee, B. T. Choi, and S. U. Lee, "A dense stereo matching using two-pass dynamic programming with generalized ground control points," *CVPR*, 2005.
8. C. Lei, J. Selzer, Y. -H. Yang, "Region-tree based stereo using dynamic programming optimization," *CVPR*, vol. 2, pp. 2378-2385, 2006.
9. D. Keysers, T. Deselaers, C. Gollan, and H. Ney, "Deformation models for image recognition," *PAMI*, 29(8), 2007.
10. H. Sakoe and S. Chiba, "A dynamic programming algorithm optimization for spoken word recognition," *IEEE Trans. ASSP*, 26(1), 1978.
11. H. Ney and S. Ortmanns, "Progress in dynamic programming search for LVCSR," *Proc. IEEE*, 88(8), 2000.
12. C. -L. Liu, S. Jaeger, and M. Nakagawa, "Online recognition of Chinese characters: the state-of-the-art," *PAMI*, 26(2), 2004.
13. E. Levin and R. Pieraccini, "Dynamic planar warping for optical character recognition," *ICASSP*, 1992.
14. S. Uchida and H. Sakoe, "A monotonic and continuous two-dimensional warping based on dynamic programming," *ICPR*, 1998.
15. D. Keysers and W. Unger, "Elastic image matching is NP-complete," *Pattern Recog. Lett.*, 24(1–3), 2003.
16. V. Mottl, S. Dvoenko, A. Kopylov, "Pattern recognition in interrelated data: the problem, fundamental assumptions, recognition algorithms," *ICPR*, 2004.
17. H. Lester and S. R. Arridge, "A survey of hierarchical non-linear medical image registration," *Pattern Recog.*, 32(1), 1999.
18. E. Angel, "Dynamic programming for noncausal problems," *IEEE Trans. AC*, 26(5), 1981.
19. A. Shekhovtsov, I. Kovtun, and V. Hlavac, "Efficient MRF deformation model for non-rigid image matching," *CVPR*, 2007.
20. B. Serra and M. Berthod, "Subpixel contour matching using continuous dynamic programming," *CVPR*, 1994.
21. M. E. Munich and P. Perona, "Continuous dynamic time warping for translation invariant curve alignment with applications to signature verification," *ICCV*, 1999.
22. S. Uchida, I. Fujimura, H. Kawano, and Y. Feng, "Analytical dynamic programming tracker," *ACCV*, 2010.
23. http://yann.lecun.com/exdb/mnist/
24. E. Angel and R. Bellman, *Dynamic Programming and Partial Differential Equations*, Academic Press, 1972.
25. M. Etoh, "Promotion of block matching: parametric representation for motion estimation," *ICPR*, 1998.
26. C. Liu, J. Yuen, and A. Torralba, "SIFT Flow: dense correspondence across scenes and its applications" *PAMI*, 33(5), 2011.