Chapter II
Elastic Matching Techniques for Handwritten Character Recognition

Seiichi Uchida
Kyushu University, Japan

ABSTRACT

This chapter reviews various elastic matching techniques for handwritten character recognition. Elastic matching is formulated as an optimization problem of planar matching, or pixel-to-pixel correspondence, between two character images under a certain matching model, such as affine and nonlinear. Use of elastic matching instead of rigid matching improves the robustness of recognition systems against geometric deformations in handwritten character images. In addition, the optimized matching represents the deformation of handwritten characters and thus is useful for statistical analysis of the deformation. This chapter argues the general property of elastic matching techniques and their classification by matching models and optimization strategies. It also argues various topics and future work related to elastic matching for emphasizing theoretical and practical importance of elastic matching.

INTRODUCTION

In handwritten character recognition, it is important to tackle geometric deformations of characters. The geometric deformations are classified into the following four types: fluctuation of stroke thickness, linear deformations (e.g., translation, scaling, shear, and rotation), nonlinear and topology-preserving deformations, and deformations changing topology. Those deformations will be caused by many factors; for example, writing material, writer’s habit, writing speed, writing style (especially cursive style), character size, inherent character shape, and noise and geometric transformation at character image acquisition.

The purpose of this chapter is to overview various elastic matching (EM) techniques for handwritten character recognition. EM is also called deformable template (Trier, Jain & Taxt, 1996), flexible matching (Mori, Yamamoto &
Elastic Matching Techniques for Handwritten Character Recognition

Yasuda, 1984), or nonlinear template matching (Mori, Suen & Yamamoto, 1992), and has been developed by many researchers to tackle the geometric deformations. EM has been employed in not only handwritten character recognition but also many other image pattern matching problems, such as face recognition, fingerprint recognition, gesture recognition, medical image analysis, automatic image morphing, computer vision (e.g., stereo), and motion analysis. For more general surveys, see Glasbey and Mardia (1998), Jain, Zhong, and Dubuisson-Jolly (1998), Lester and Arridge (1999), Redert, Hendriks, and Biemond (1999), and Zitová and Flusser (2003).

EM is formulated as an optimization problem of planar matching, or 2D-2D mapping between two character images, A and B. From another viewpoint, EM treats a character image A like a “rubber sheet” and fits it to another character image B as closely as possible. Hereafter, this 2D-2D mapping from A to B is called 2D warping (2DW). Note that we can consider EM based on 1D-2D mapping, where a 1D-stroke model is fitted to input image, although this chapter mainly concerns EM techniques based on 2DW. In a later section, we will briefly review these EM techniques based on 1D-2D mapping.

For handwritten character recognition, EM possesses two merits. The first merit is that the distance evaluated under the optimized 2DW is deformation-invariant. Thus, by using the EM distance as a discriminant function, we can realize character recognition systems robust to the geometric deformations. The range of the invariance depends on the definition of 2DW; that is, the more flexible 2DW becomes, the more invariant the EM distance becomes. Thus, EM has a potential to provide more intuitive and robust recognition frameworks than other deformation-invariant techniques, such as invariant feature (e.g., the horizontal projection profile by Nakata, Nakano, and Uchikura (1972)) and shape normalization (Lee & Park, 1994; Liu, Nakashima, Sako, & Fujisawa, 2004).

The second merit is that the optimized 2DW itself describes the deformation of subjected characters. This fact shows that EM possesses a useful property of structural analysis techniques. Furthermore, EM can be linked to statistical and stochastic frameworks by the merit. Active shape models (Cootes, Taylor, Cooper & Graham, 1995; Shi, Gunn, & Damper, 2003; Uchida & Sakoe, 2003a) and (pseudo-) 2D HMMs (Agazzi, Kuo, Levin, & Pieraccini, 1993; Kuo & Agazzi, 1994; Levin & Pieraccini, 1992; Park & Lee, 1998) are two good examples.

The remaining part of this chapter is organized as follows. First, EM is formulated as an optimization problem of 2D-2D mapping (i.e., 2DW). General properties of EM and the EM distance are also described. Second, EM techniques are classified according to their specific formulations of 2DW and optimization strategies. It will be emphasized that there is a strong relation between the formulation and the optimization strategy. Third, several related topics are discussed, such as incorporation of category-dependent deformation tendency. Fourth, EM techniques based on 1D-2D mapping are briefly reviewed, which is another type of EM used in handwritten character recognition. Finally, conclusions are presented after listing various future tasks for EM.

OUTLINE OF ELASTIC MATCHING

Formulation of EM

As described before, EM is formulated as an optimization problem of 2D-2D mapping (i.e., 2DW) between two character images, A and B. Let \( a_{i,j} \) and \( b_{x,y} \) denote pixel values (e.g., intensity values) or pixel feature vectors (e.g., RGB vectors) at pixel \((i,j)\) on \(A\) and \((x,y)\) on \(B\), respectively. While we can deal with the matching between two images of arbitrary sizes, we hereafter assume \(N \times N\) images for simplicity.