Calligraphic Character Boundary Coding with Rational B-spline Based on Energy Minimization Using Genetic Algorithm

Paul Ge Bao and Stephen Wang Cheung Lam
Department of Computing, Polytechnic University of Hong Kong,
Hung Hom, Kowloon, Hong Kong.
Email: csbao@comp.polyu.edu.hk Tel: (852) 2766 7266

Traditional salient point based approaches fail in coding calligraphic characters as noisy boundaries make the extraction of the salient points a difficult task. We propose an alternative solution based on genetic algorithm which searches through the space of possible parameter values until a global optimal solution is found.

The objective function we employed is a modified version of the total energy function found in the active contour literature. The equations defining the objective functions are shown as follows:

\[
E_{\text{total}}(\mathbf{v}(s)) = \int_0^1 [E_{\text{int}}(\mathbf{v}(s)) + \gamma E_{\text{ext}}(s)] ds
\]

\[
E_{\text{int}}(\mathbf{v}(s)) = [\alpha \| v_s(s) \|^2 + \beta \| v_{ss}(s) \|^2 ] / 2
\]

\[
E_{\text{ext}}(\mathbf{v}(s)) = \delta(v(s) - B(s))
\]

where \( E_{\text{total}}, E_{\text{int}} \) and \( E_{\text{ext}} \) are the total, the internal and the external energy of the rational B-spline curve respectively. The internal energy specifies the smoothness requirement while the external energy is a potential energy which attracts the fitting surface toward the data points. Function \( v(s) \) is the vector of the spline curve on the point \( (x(s),y(s)) \) and \( s \) is the normalized parameter of the curve by its length \( l \). \( v_s(s) \) and \( v_{ss}(s) \) are the 1st and the 2nd differentiations with respect to \( s \). The constants \( \alpha, \beta \) and \( \gamma \) are the weights for each term and \( \delta(s) \) is the Euclidean distance function between the spline point and its closest object boundary point. Unlike the conventional external energy which usually involves gray level gradient, we compute the similarity between the fitted curve and the object boundary. The parameters are selected to make the above objective functions minimal.

Our approach works as follows. Each curve fitted is represented as an individual chromosome and represented uniquely within the system by a string. A new generation of chromosomes are generated by simulating the dynamics of population development by employing some genetic operators. Finally, the fittest survivor is obtained and it represents the optimal curve used to model the boundary of a calligraphic character. Figure 1 summaries the mechanism of our approach.

![Diagram](image)

**Figure 1.** Overall structure the reproductive plan.