1-28-2010

Segmentation of thermographic images of hands using a genetic algorithm

Payel Ghosh
Portland State University

Judith Gold

Melanie Mitchell
Portland State University

Let us know how access to this document benefits you.

Follow this and additional works at: https://pdxscholar.library.pdx.edu/compsci_fac

🔗 Part of the Other Analytical, Diagnostic and Therapeutic Techniques and Equipment Commons, and the Theory and Algorithms Commons

Citation Details

This Article is brought to you for free and open access. It has been accepted for inclusion in Computer Science Faculty Publications and Presentations by an authorized administrator of PDXScholar. For more information, please contact pdxscholar@pdx.edu.
Segmentation of thermographic images of hands using a genetic algorithm

Payel Ghosh\textsuperscript{a}, Melanie Mitchell\textsuperscript{b}, Judith Gold\textsuperscript{c}

\textsuperscript{a}Dept. of Electrical Engineering, Portland State University, Portland, OR, USA; \textsuperscript{b}Dept. of Computer Science, Portland State University, Portland, OR, USA and Santa Fe Institute, Santa Fe, NM, USA \textsuperscript{c}Department of Public Health, Temple University, Philadelphia, PA, USA

ABSTRACT

This paper presents a new technique for segmenting thermographic images using a genetic algorithm (GA). The individuals of the GA also known as chromosomes consist of a sequence of parameters of a level set function. Each chromosome represents a unique segmenting contour. An initial population of segmenting contours is generated based on the learned variation of the level set parameters from training images. Each segmenting contour (an individual) is evaluated for its fitness based on the texture of the region it encloses. The fittest individuals are allowed to propagate to future generations of the GA run using selection, crossover and mutation.

The dataset consists of thermographic images of hands of patients suffering from upper extremity musculo-skeletal disorders (UEMSD). Thermographic images are acquired to study the skin temperature as a surrogate for the amount of blood flow in the hands of these patients. Since entire hands are not visible on these images, segmentation of the outline of the hands on these images is typically performed by a human. In this paper several different methods have been tried for segmenting thermographic images: Gabor-wavelet-based texture segmentation method, the level set method of segmentation and our GA which we termed LSGA because it combines level sets with genetic algorithms. The results show a comparative evaluation of the segmentation performed by all the methods. We conclude that LSGA successfully segments entire hands on images in which hands are only partially visible.

Keywords: Segmentation, level sets, shape, texture, genetic algorithms, thermographic image processing.

1. INTRODUCTION

Infrared thermography of hands reveals skin temperature which is largely determined by the amount of blood flowing in the hands. This characteristic of thermographic imaging is useful for studying distal upper extremity musculoskeletal disorders (UEMSD) (such as tendinitis or carpal tunnel syndrome) which may be caused due to reduced blood flow\textsuperscript{[3][11][19]}. The data for this analysis was obtained from Temple University’s Ergonomics and Work Physiology Lab where far-infrared images of hands were acquired using ThermaCAM AM40 thermographic camera (FLIR Systems, Wilsonville, OR) with a sampling rate of 7 HZ. Subjects were given a nine minute typing challenge and images were taken at the following intervals: before typing, 0-2 minute (post-typing), 3-5 minute (post-typing), and 8-10 minute (post-typing). Figure 1 (upper panel) shows the hands of four subjects with UEMSD before starting the typing challenge. Figure 1 lower panel shows the hands of the same subjects after 9 minutes of typing. Since, the fingertips of these patients became cold, the entire hand is not visible in these images. In some of these images a near infrared spectroscopy (NIRS) probe was adhered to the skin above the first dorsal interosseous muscle of their right hand (between the thumb and index finger). Therefore, the boundary of the hand touching the probe is not visible. Typically these images are manually segmented by a human who has the prior knowledge of the shape of the hand. Once an image is segmented, a score of the mean temperature of the hand is generated and compared with the temperature of the hand at successive time steps to determine if and where the blood flow is reducing in the hand with time.

The goal of this work was to develop a method for automating the segmentation of hands in these images. The segmentation task is challenging because the subjects tend to move their hands during the imaging process. Thus the automatic segmentation method needs to incorporate all of the following to perform successful segmentation: known shape, shape variation, movement and texture of the hands.
1.1. Automatic segmentation methods

Segmentation is performed by demarcating an object on an image with a boundary using pixel-level or object-level properties of the object. These properties can be edges, texture, pixel intensity variation inside the object, shape, size, orientation, and so on. Texture is usually defined as a region consisting of mutually related elements. Various approaches for textural feature extraction exist including co-occurrence matrices [8], filtering methods such as Gabor filters [7], Fourier transform methods, texture element finding methods such as textons [17][28], to name a few. One major drawback of all pixel-based segmentation algorithms is that regions outside the object can also be classified as part of the object and there is no notion of shape in these methods. Segmentation using object-level features involves quantifying object characteristics such as shape, pose, and relative position with respect to objects as well as region-based properties of the object. Shapes are generally represented using contours, transforms, or regions. In this work, shape has been represented using an active contour in which a deformable contour is used to segment an object.

1.1.1. Active contour method and the level set method of segmentation

Active contour method of segmentation is a shape-based procedure in which an energy function is minimized to perform segmentation. The energy function typically consists of region-based terms such as edges, mean pixel intensity, and/or object-level terms such as curvature of the object and size. In these methods the initial contour is usually placed randomly in the region of interest. During the curve evolution process, minimization of the energy function drives the curve towards the boundary of the object.

Figure 1 Thermographic images of hands of four subjects taken before typing (upper panel); after 9 minutes of typing (lower panel).

One approach for curve evolution is the marker point method [26] in which the segmenting curve $C$ is parameterized by points on the curve along a certain orientation. The front can be interpolated from these marker points as either line segments in two-dimensions (2D) or triangles in three-dimensions (3D). Another approach to active shape modeling is the level set method introduced by Sethian [23][27]. In this approach the evolving boundary (interface) is implicitly embedded as the zero isocontour of some function. In the level set method, the equation of motion of the curve is defined using a simple convection equation (the level set equation) such as:

$$\dot{\phi} + \vec{V} \cdot \nabla \phi = 0,$$

where $\dot{\phi}$ is the temporal partial derivative of the implicit function $\phi$, $\vec{V} = \langle u, v, w \rangle$ is the velocity field ($u$, $v$, $w$ are components of the velocity field in the $x$, $y$ and $z$ directions respectively), and $\nabla$ is the spatial gradient operator. The level set function $\phi$ may be defined in terms of the signed distance function. The signed distance function takes any point in the plane and returns the Euclidean distance between the pixel and the closest point on the interface. Pixels outside the interface have positive distance while the pixels inside have negative distance values assigned to them. The zero level set is the interface itself, i.e., the set of all points whose distance to the interface is zero. Curve evolution in the level set method is stable and small errors in approximation are not amplified with time.
Level set methods have been used extensively for image segmentation\cite{6,14}. Some of the popular methods are by Leventon et al.\cite{19}, Tsai et al.\cite{29} and Chan and Vese\cite{5}. Leventon et al. introduced the concept of shape representation by principal component analysis (PCA) on signed distance functions. They also incorporated statistical shape priors into their geodesic active-contour model to generate maximum a posteriori estimates of pose and shape. Chan and Vese introduced a region-based energy function based on Mumford-Shah segmentation techniques\cite{25} in order to detect features with diffuse boundaries. Tsai et al. derived a shape-based level set function. Tsai et al.’s goal was to find the parameters of this function that produce a good model of the object shape based on priors from the training data. Tsai et al. derived these parameters via an optimization procedure that used statistics defined over local regions in a set of training images. We have adopted this level-set function for the LSGA because it is a scheme for modeling known shapes as is required by the current problem. Here, a GA has been used to combine the shape-based level set function with texture feature for performing segmentation.

1.2. Genetic algorithms

Genetic algorithms (GA)\cite{11,15} simulate the learning process of biological evolution using selection, crossover and mutation. In contrast with traditional optimization methods, a GA uses a randomized parallel search to reach the optimum solution and so is less likely to be stuck in a local maximum. Individuals of the GA are candidate solutions and are typically encoded as bit strings or vectors defined based on the application. The GA searches the space of candidate solutions to identify the best solution for the problem at hand. A fitness function is used to evaluate individuals and compare them based on a score. This score is used in the selection process to determine which individuals get to produce offsprings and propagate their “genes” (bits/vector elements) to future generations. Selection can be performed in a number of different ways. The method implemented here is the rank selection method. In rank selection candidate solutions are sorted according to their fitness score and the probability of choosing higher ranked individuals is more than lower ranked individuals. After selecting two individuals from the current population, the crossover operator is applied to produce two new offsprings that are members of the next generation. Crossover is performed using a crossover mask which copies the bit position $i$ from one of the two parents into the bit position $i$ of the offspring (figure 2 shows a single point crossover)\cite{24}. The mutation operator chooses a single gene at random and changes its value. The new population thus generated is evaluated using the fitness function and the process of selection crossover and mutation is repeated until an offspring with an acceptable value is produced.

![Parent strings Crossover Mask Offspring]

\textbf{Figure 2. Single-point crossover}

1.2.1. Significance of this work

Genetic algorithms have been used for segmentation by\cite{4,2,22,13}. A model-based image analysis technique using a GA is described in\cite{21}. The method uses a Hough transform scheme with Parisian evolution to model the visible surfaces of an object using a collection of 3D points. Cagnoni et al.\cite{4} used a GA for segmenting images by evolving parameters of an active contour model called “snakes”\cite{18}. They optimized an energy function based on low-level features such as smoothness of the curve, curvature and image gradient.

Unlike the explicit energy function used in these methods, the LSGA optimizes an implicit level set function. An individual in the LSGA population is a vector of parameters of a level set function and is referred to as a chromosome of the GA. The GA adapts the parameters of the level set function to produce fit individuals or good segmentations of the given image using the information encoded in its fitness function. The algorithm terminates by finding a reasonable segmentation within the bounds of known shape and texture of the hands. Thus, the main contribution of this work is the use of GAs to optimize a level set function thereby combining multiple features (such as shape and texture) for segmentation. In future, features such as relative position between objects would be included to perform segmentation on images containing multiple objects.
2. METHOD
The GA consists of two stages: the training stage and the segmentation stage. In the training stage, shape, shape variability and texture information of the region of interest are derived from manually segmented images. In the segmentation phase the genetic algorithm evaluates candidate contours for segmenting the desired object in the test image using a fitness measure, and iterates over successive generations until the fitness exceeds a threshold.

2.1. Training: Deriving shape information
The shape representation is derived from the mean and variance of all manually drawn contours in a training set. The manually drawn contours from the training data are first converted into signed distance functions, \( \psi_i \) (1 to \( n \), is the number of training contours). The level set function is a linear combination of the mean shape and weighted shape variances in the signed distance domain. The mean shape is defined for \( n \) contours as:

\[
\overline{\Phi} = \frac{1}{n} \sum_{i=1}^{n} \psi_i.
\] (2)

Mean offset functions are derived by subtracting the mean from the signed distance representations of the training contours (\( \bar{\psi}_i = \psi_i - \overline{\Phi} \)). The columns of the mean offset functions (size \( N = N_1 \times N_2 \) the same as the training images) are then successively stacked on top of one another to form one large column vector (\( \beta \)) of size \( 1 \times N \). A new matrix \( S \) (size \( N \times n \)), called the shape variability matrix, is formed from \( n \) such column vectors

\[
S = [\beta_1, \beta_2, ..., \beta_n].
\] (3)

The variance in shape is then computed by an eigenvalue decomposition on this shape variability matrix as,

\[
\frac{1}{n} SS^T = U\Sigma U^T.
\] (4)

Here \( U \) is an \( N \times n \) matrix whose columns represent \( n \) orthogonal modes of shape variation and \( \Sigma \) is a diagonal matrix of eigenvalues. By rearranging the columns of \( U \) to form an \( N_1 \times N_2 \) structure, the \( n \) different eigleshapes can be obtained \( \{\Phi_1, \Phi_2, ... , \Phi_n\} \). The mean shape and shape variability are used to define a level set function,

\[
\Phi[w] = \overline{\Phi} + \sum_{j=1}^{k} w_j \Phi_j.
\] (5)

Here \( w \) are the weights for linearly combining the \( k \) principal eigleshapes. By incorporating pose parameters into this level set framework a new level set function is obtained that can handle object shapes with different sizes and orientation. Pose is defined using an affine transform which is the product of three matrices, the translation matrix, the scaling matrix and the rotation matrix respectively,

\[
\begin{bmatrix}
\tilde{x} \\
\tilde{y}
\end{bmatrix} = \begin{bmatrix}
1 & 0 & a \\
0 & 1 & b \\
0 & 0 & 1
\end{bmatrix} \begin{bmatrix}
h & 0 & 0 \\
0 & h & 0 \\
0 & 0 & 1
\end{bmatrix} \begin{bmatrix}
\cos \theta & -\sin \theta & 0 \\
\sin \theta & \cos \theta & 0 \\
0 & 0 & 1
\end{bmatrix} \begin{bmatrix}
x \\
y
1
\end{bmatrix}.
\] (6)

Here, \( x \) and \( y \) are the pixel coordinates of the input image and \( \tilde{x} , \tilde{y} \) are the pixel coordinates of the affine transformed image. Note that a homogeneous coordinate system is used here. Using this homogeneous coordinate system allows the translation operation in equation (6) to be represented with a matrix multiplication.

This new level set function is defined as [29]:

\[
\Phi[w,p](x,y) = \overline{\Phi}(\tilde{x}, \tilde{y}) + \sum_{j=1}^{k} w_j \Phi_j(\tilde{x}, \tilde{y}).
\] (7)

Here \( p = [a, b, h, \theta] \), \( a, b \) are \( x, y \) translation parameters, \( h \) is the scale factor and \( \theta \) is the angle of rotation. The zero level of this level set function gives the segmenting contour and its parameters are evolved by the GA.

Before deriving the mean shapes and shape variance from the training data, the images need to be aligned for pose. Gradient descent is used to minimize the difference between pairs of binary images with respect to their pose parameters. The transformed image based on pose is given by:

\[
\tilde{I} = T[p] * I,
\] (8)
where, \( T[p] \) is the 2D transformation matrix of equation (6). The energy functional used to minimize the difference between two images is given by:

\[
E_{\text{align}} = \frac{\int_{\Omega} \left( I^i - I^j \right)^2 \, d\Omega}{\int_{\Omega} \left( I^i + I^j \right)^2 \, d\Omega}, \quad i \neq j.
\]  

(9)

Here, \( \Omega \) is the image domain. The area normalizing term in the denominator is employed to prevent the images from shrinking to improve the cost function. The initial pose parameters of one of the shapes are kept fixed and the pose parameters of the second image are calculated to minimize the pose differences.

### 2.2. Training: Deriving Texture Information

The textural priors are derived from training images using Gabor wavelet-based texture segmentation method [1][7]. Gabor wavelets are based on the Gabor elementary function given by the modulation of the Gaussian with a complex exponential function (equations 10, 11).

\[
h(x, y) = g(x, y) \exp \left[ \frac{j2\pi(Ux + Vy)}{\sigma_x \sigma_y} \right]
\]

(10)

\[
g(x, y) = \frac{1}{2\pi\sigma_x \sigma_y} \exp \left\{ -\frac{1}{2} \left( \frac{x}{\sigma_x} \right)^2 + \left( \frac{y}{\sigma_y} \right)^2 \right\}.
\]

(11)

Gabor wavelets are derived from several translations and dilations of the mother wavelet \( h(x, y) \). The method of Gabor wavelets assumes that local texture regions are spatially homogeneous and the mean and standard deviation of the transform coefficients are used to represent regions for classification. The Gabor wavelet is given by equation (12)

\[
h_{mn}(x, y) = \alpha H(x', y'), \quad a > 1,
\]

\[
m, n = \text{integers},
\]

\[
a' = a^{-m} (x\cos \frac{\pi n}{k} + y\sin \frac{\pi n}{k}),
\]

\[
y' = a^{-m} (-x\sin \frac{n\pi}{k} + y\cos \frac{n\pi}{k}).
\]

(12)

Here, \( k \) is the number of orientations. The Gabor wavelet transform is given by

\[
W_{mn}(x, y) = \int I(x_1, y_1) h^*_{mn}(x - x_1, y - y_1) \, dx_1 \, dy_1.
\]

(13)

Here, \( h^* \) is the complex conjugate of \( h \). The mean and standard deviation of transform coefficients for each known region are the derived texture priors.

### 2.3. Segmentation using LSGA

Segmentation is performed using LSGA by optimizing a population of segmenting contours for shape, texture of the enclosed region, location and pose. Each individual in the GA population is defined as a fixed-length string of real-valued genes.

\[
I = [w_1, w_2, w_3, w_4, \alpha, b, h, \theta].
\]

(14)

Four weight parameters are used for deriving the weighted \( \pm \sigma_1 \) and \( \pm \sigma_2 \) variation (where \( \sigma^2 \) is the eigenvalue corresponding to the principal eigenshapes) of the mean shape and \( \alpha, b, h, \) and \( \theta \) are pose parameters as defined in equation (6). Each individual \( I \) in the population represents a unique segmenting contour. This segmenting contour produces a binary image \( (B_i) \) at its output. The fitness of an individual \( I \) is calculated based on the degree to which the contour encloses the region of interest (ROI) in a test image (a new image not in the training set). Binary image \( B_2 \) is generated by the textural classification of pixels on a new test image using the mean and variance of the saved wavelet coefficients. The fitness is calculated by comparing the two binary images \( B_1 \) and \( B_2 \). The fitness \( F \) is a function of the detection rate \( (DR) \) and the false alarm rate \( (FA) \) as

\[
F = 500(DR + (1- FA)).
\]

(15)
Note that $B_2$ is the so-called “truth plane” used by the fitness function and is not the ground truth (derived from manual segmentation). The ground truth images are only used for evaluating the final segmentation results. Rank selection and fixed-length crossover have been implemented here. Mutation rate is defined as the probability of a single gene to be mutated. Similarly, the crossover rate defines the probability of a crossover to occur between two individuals.

The dice similarity coefficient \[^{30}\] and the partial Hausdorff distance \[^{16}\] were used for evaluating the performance of the algorithm. The dice similarity coefficient provides a measure of the degree of overlap between two segmentations as:

$$DSC(A, B) = \frac{2|A \cap B|}{|A| + |B|},$$ (16)

A $DSC$ of 1 indicates a perfect match and 0 indicates no match. The partial Hausdorff distance is derived between the boundary points of two contours. If $A = \{a_1, \ldots, a_p\}$ and $B = \{b_1, \ldots, b_q\}$ be finite sets of points on two contours, then the partial Hausdorff distance between them is defined as:

$$H(A, B) = \max(h(A, B), h(B, A)),$$

where $h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\|$. (17)

The function $h(A, B)$ takes each point in $A$ and finds the closest point in $B$ from that point. It then ranks the points in $A$ based on the distance values and finds the point with the greatest “mismatch”. Thus, the partial Hausdorff distance is a measure of the distance by which two contours i.e., the final segmentation outcome and the ground truth differ.

3. RESULTS

Segmentation was performed on test images separate from the training images of each subject using the following methods: LSGA, Gabor wavelet based segmentation algorithm, and the Chan & Vese algorithm. Figure 3 shows a manually segmented test image. The outcome of the Gabor wavelet-based segmentation algorithm on the sample test image is shown in figure 3 second panel. Figure 3 third panel shows the result of applying the Chan & Vese algorithm to the same hand image. Since both these methods do not have the notion of a known shape they could not segment the entire hand on the test image. Finally, the LSGA was used to segment the same hand image (figure 3 fourth panel). The parameters used by the GA to evolve the segmenting contour are shown in table 1.

![Figure 3. Segmentation outcomes on a test image using: (From left) manual segmentation, Gabor-wavelet method, Chan & Vese method and LSGA.](image)

**Table 1. LSGA parameters**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population size</td>
<td>50</td>
</tr>
<tr>
<td>Mutation rate</td>
<td>10 % per gene</td>
</tr>
<tr>
<td>Crossover rate</td>
<td>50% single-point</td>
</tr>
<tr>
<td>Selection criteria</td>
<td>Rank selection</td>
</tr>
<tr>
<td>Weights for eigen shapes, $w$</td>
<td>(1-5) integers</td>
</tr>
<tr>
<td>Translation parameters $a, b$</td>
<td>Integer (0-30)</td>
</tr>
<tr>
<td>Rotation parameter $\theta$</td>
<td>-90° to +90°</td>
</tr>
<tr>
<td>Scale parameter $h$</td>
<td>1</td>
</tr>
</tbody>
</table>
Figure 4 shows some candidate segmenting contours in the GA population. The LSGA found the optimal location and pose of the hand in the image from a population of 50 segmenting contours. The segmenting contour is shown on top of the test image to show the segmentation outcome here. The fitness of the final segmenting contour for this image was 828. The figure clearly shows that the LSGA outperforms the other methods. This is also confirmed by visual analysis of the results.

Figure 4. Three candidate segmenting contours in the GA population. Here, F denotes the fitness of each individual defined in equation 15.

The average $DSC$ and $H$ were computed from all the segmentation outcomes for each subject. Table 2 shows the average values of $DSC$ and $H$ obtained for all the images in the validation set for each of the four subjects. Note that in the case of subject 4 the $DSC$ is low and for subject 1 the $H$ is relatively high even though a visual inspection of the results shows satisfactory segmentation. This is due to the limitation in acquiring accurate ground truth as the fingertips are completely invisible in the test images.

Table 2. Validation criteria: comparison with ground truth

<table>
<thead>
<tr>
<th></th>
<th>$DSC$</th>
<th>$H$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject 1</td>
<td>0.8</td>
<td>6.7</td>
</tr>
<tr>
<td>Subject 2</td>
<td>0.9</td>
<td>3.4</td>
</tr>
<tr>
<td>Subject 3</td>
<td>0.85</td>
<td>3.5</td>
</tr>
<tr>
<td>Subject 4</td>
<td>0.6</td>
<td>3.5</td>
</tr>
</tbody>
</table>

4. CONCLUSION

The LSGA performs derivative-free optimization of a level set function for image segmentation. Replacing the energy function with a fitness function makes the GA implicitly explore the feature space using payoff values. This brings the flexibility to incorporate any feature for level set curve evolution. In this paper two features, texture and shape, have been explored for segmenting thermographic images of hands. Although these images had visible textural areas separating the hand region from the background, the knowledge of known shape was needed to segment them. The LSGA combined shape with texture to achieve the desired segmentation. Other features such as relative position, context etc. would be incorporated in future for more complex segmentation tasks [9][10].

REFERENCES


