

- [4] R. Lienhart, "Indexing and retrieval of digital video sequences based on automatic text recognition," in *Proc. ACM Int. Conf.*, Boston, MA, Nov. 1996.
- [5] A. K. Jain and B. Yu, "Document representation and its application to page decomposition," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 20, pp. 294–308, Mar. 1998.
- [6] J. Ohya, A. Shio, and S. Akamatsu, "Recognizing characters in scene images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 16, pp. 215–220, Feb. 1994.
- [7] H. M. Suen and J. F. Wang, "Text string extraction from images of color-printed documents," *Proc. Inst. Elect. Eng. Vis., Image, Signal Process.*, vol. 143, no. 4, pp. 210–216, 1996.
- [8] L. Wang and T. Pavlidis, "Direct gray-scale extraction of features for character recognition," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 15, pp. 1053–1067, Oct. 1993.
- [9] Y. Zhong, K. Karu, and A. K. Jain, "Locating text in complex color images," *Pattern Recognit.*, vol. 28, no. 10, pp. 1523–1535, 1995.
- [10] F. M. Wahl, K. Y. Wong, and R. G. Casey, "Block segmentation and text extraction in mixed text/image documents," *Comput. Graph. Image Process.*, vol. 20, pp. 375–390, 1982.
- [11] J. Serra, *Image Analysis and Mathematical Morphology*. New York: Academic, 1982.
- [12] D. S. Bloomberg, "Multiresolution morphological analysis of document images," *Proc. SPIE Visual Communication Image Processing*, vol. 1818, pp. 648–662, 1992.
- [13] R. Kasturi and M. M. Trivedi, *Images Analysis Applications*. New York: Marcel Dekker, 1990.
- [14] K. K. Chin and J. Sanjie, "Morphological processing for feature extraction," *Proc. SPIE*, vol. 2030, pp. 288–302, 1993.
- [15] S. U. Lee, S. Y. Chung, and R. H. Park, "A comparative performance study of several global thresholding techniques for segmentation," *Comput. Vis. Graph. Image Processing*, vol. 52, pp. 171–190, 1990.
- [16] Y. M. Y. Hasan and L. J. Karam, "Morphological reversible contour representation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 22, pp. 227–240, Mar. 2000.
- [17] S. E. Umbaugh, Y. Wei, and M. Zuke, "Feature extraction in image analysis," *IEEE Eng. Med. Biol. Mag.*, vol. 16, no. 4, pp. 62–73, 1997.
- [18] E. Gose, R. Johnsonbaugh, and S. Jost, *Pattern Recognition and Image Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1996.
- [19] P. Yang and P. Maragos, "Morphological systems for character recognition," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, 1993, pp. 97–100.
- [20] B. Yu, A. Jain, and M. Mohiuddin, "Address block location on complex mail pieces," in *Proc. 4th Int. Conf. Document Analysis Recognition*, Ulm, Germany, 1997.

A Wavelet-Frame Based Image Force Model for Active Contouring Algorithms

Hsien-Hsun Wu, Jyh-Charn Liu, and Charles Chui

Abstract—This paper proposes a directional image force (DIF) for active contouring. DIF is the inner product of the zero crossing strength (ZCS) of wavelet frame coefficients, and the normal of a snake. By representing strength and orientation of edges at multiple resolution levels, DIF markedly improves the immunity of snakes to noise and convexity.

I. ACTIVE CONTOURING ALGORITHMS

The objective of active contouring, i.e., *snake* algorithms, is to find the closest contour around the natural boundary of an object. Ter-

Manuscript received January 25, 1999; revised June 5, 2000. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Kannan Ramchandran.

H.-H. Wu is with the Cytokinetics Inc., San Francisco, CA 94080 USA.
J.-C. Liu is with the Department of Computer Science, Texas A&M University, College Station, TX 77843 USA (e-mail: liu@cs.tamu.edu).
C. Chui is with the TeraLogic, Inc., San Jose, CA 95101 USA.
Publisher Item Identifier S 1057-7149(00)09395-7.

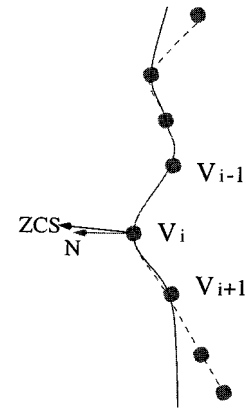


Fig. 1. Relationship between ZCS and the normal of the snake.

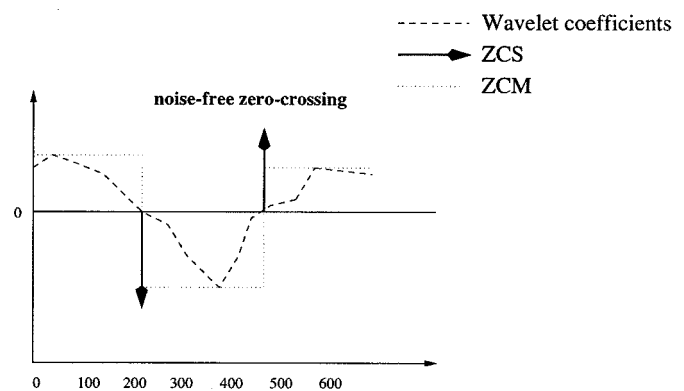


Fig. 2. Zero crossing representation.

zopoulos and Kass [1] proposed the first snake algorithm, with several algorithms proposed later trying to solve various problems such as noise trap [5], snake initialization [2], energy function optimization [4], [5], image object extraction [6], concavity issues [3], [7], etc.

Let the line segments that connect adjacent pairs of the n snaxels on a snake be denoted as $v(0), v(1), v(2), \dots, v(n)$, where $|v(s)|$ is the arc length. Snake shape deformation is controlled by an *energy function* that consists of the *internal* and *external* energy,

$$E_{snake}(v(s)) = \int_0^1 [E_{int}(v(s)) + E_{ext}(v(s))] ds. \quad (1)$$

The internal energy is commonly defined as $E_{int}(v(s)) = 1/2(w_1(s)|v'(s)|^2 + w_2(s)|v''(s)|^2)$, where $v'(s) \equiv (\partial v(s)/\partial s)$, and $v''(s) \equiv (\partial^2 v(s)/\partial s^2)$. $E_{ext}(v(s))$ is usually derived from lines or corners surrounding the snake. Ideally, a snake should stop the deformation process when its energy function is minimized, or $(\partial E_{snake}/\partial v) = 0$, where v is the discretized snake, provided that the spot noise does not trap a snaxel.

In this study, we adopt the expansion (ballooning) approach, where the external energy is the weighted sum of the image force and the balloon force [5]: $\int_0^1 E_{ext}(v(s)) = \int_0^1 E_{image}(v(s)) + \int_0^1 E_{balloon}(v(s))$. By using Green's theory, we have $\int_0^1 E_{balloon}(v(s)) = w_4(s) \iint_R dA(s)$, where R is the snake region and $w_4(s)$ is a weight factor. This newly proposed image force, called the *directional image force* (DIF), is the inner product of ZCS and the contour normal (see Fig. 1), $E_{image-dir}(v(s)) = w_3(s)N(v(s)) \cdot ZCS(v(s))$, where $w_3(s)$ is a weight constant, and ZCS is the *polarized strength* of a *zero*

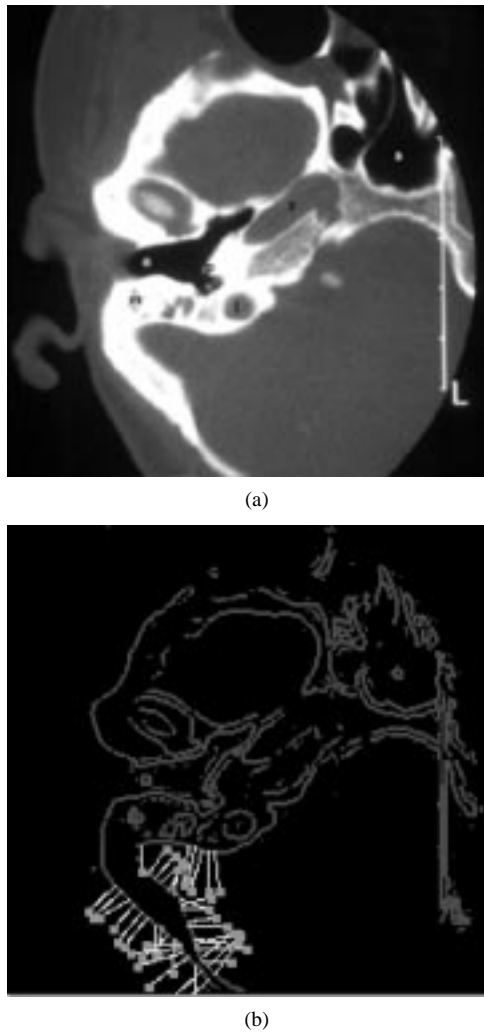


Fig. 3. Orientation of edge pixels derived from ZCS. (a) Medical image and (b) orientation.

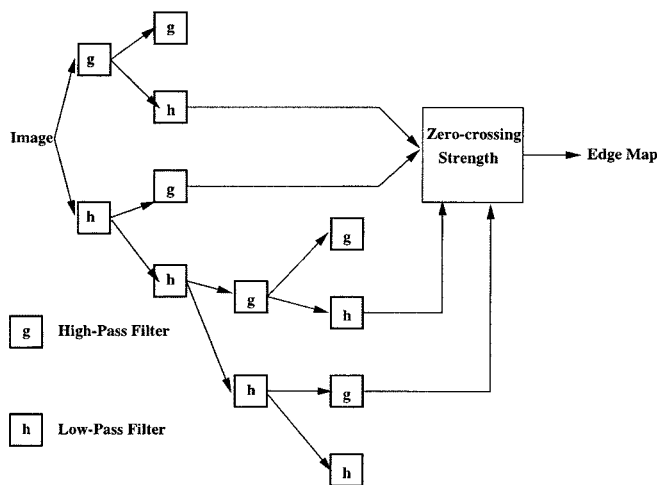


Fig. 4. Multiresolution property of the ZCS.

crosspoint (ZCP), meaning that ZCS represents the orientation and strength of the ZCP. ZCPs are detected from the discrete wavelet transform (DWFT) coefficients along the x and y axes. A DWFT coefficient is considered a ZCP if its adjacent neighbors have different signs, and ZCS of a ZCP is the difference between the sums

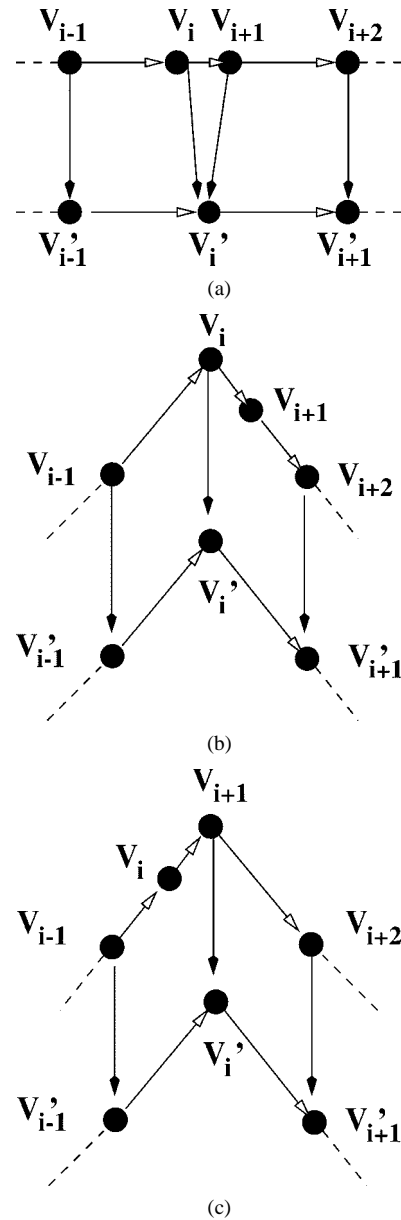
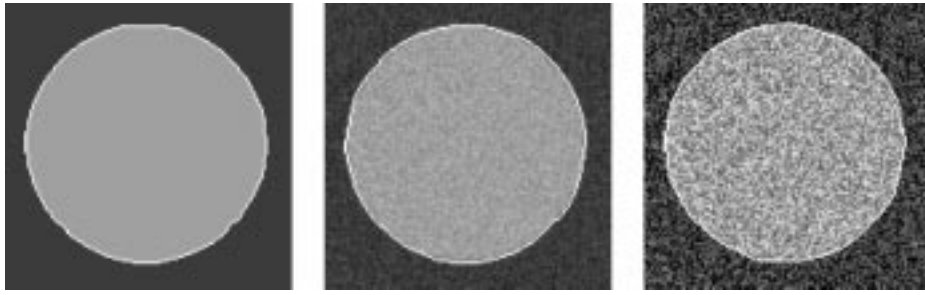
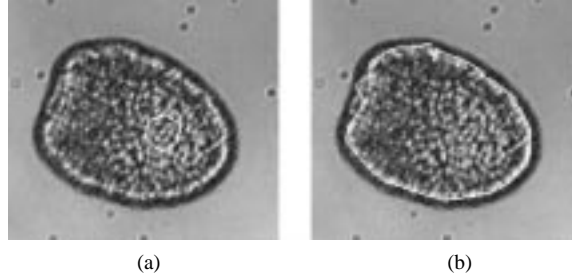


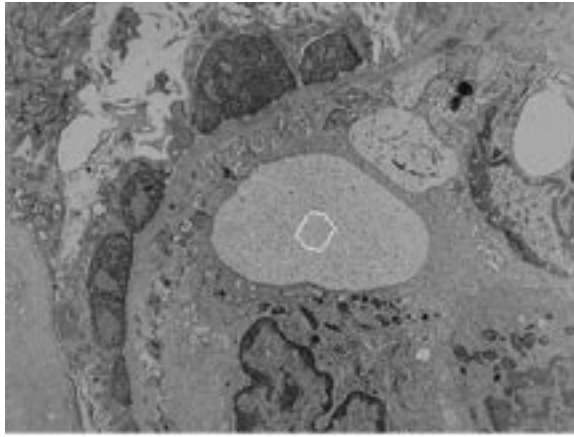
Fig. 5. Deleting a snaxel during resample. (a) v_i and v_{i+1} are not corners. (b) v_i is corner. (c) v_{i+1} is corner.

of DWFT coefficients between its two sides (see Fig. 2). For example, let P_i denote a ZCP for a sequence of DWFT coefficients extracted along a row/column of a two dimensional DWFT matrix, ZCS of P_i , $ZCS(i) = \sum_{r=1}^n f(i+r) - \sum_{\ell=0}^e f(i-\ell)$, if P_i is a ZCP. Otherwise, $ZCS(i) = 0$. Here, $f(i)$ is the DWFT value of P_i , $n(e)$ is the number of consecutive nonzero, same-signed DWFT coefficients that lay at the right (left) hand side of P_i . For example, when P_0, P_1, \dots, P_6 , have values of 1.3, 0.3, 2.1, 0.1, -5.3, -7.3, -2.8, $ZCS(i) = -19.2 = -2.8 - 7.3 - 5.3 - (0.1 + 2.1 + 0.3 + 1.3)$. If ZCS is positive (negative), then the intensity changes from dark(bright) to bright(dark), i.e., the *polarity* of the intensity change, along the scan direction of the ZCS calculation. The two-dimensional ZCS measure of a pixel is a vector, (ZCS_x, ZCS_y) , whose values are taken from ZCS_x and ZCS_y measures of the pixel along x and y axes. The orientation of a ZCS observation is orthogonal to the edge, and its magnitude is proportional to edge strength [8]. The bright lines with dotted tips in Fig. 3(b) mark the strength and orientation of some selected pixels in Fig. 3(a), by using the ZCS measure.

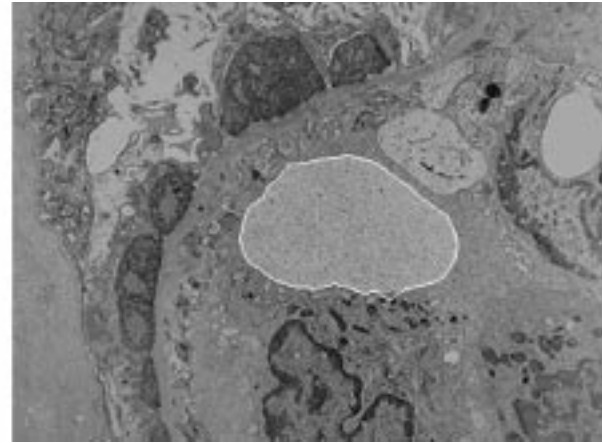
Fig. 6. Experimental results for (a) SNR = ∞ , (b) SNR = 20, and (c) SNR = 10.

(a)

(b)



(c)



(d)

Fig. 7. (a) Cell_1 image and its initial contour, (b) final contour, (c) cell_2 image and its initial contour, and (d) final contour.

For multiresolution image force calculations, one can take weighted vectors of ZCS components from different wavelet frames (see Fig. 4). Empirical results suggested that the symmetric, three-tap Daubechies biorthogonal wavelet has the best performance, because its compactness makes it sensitive to weak edges. A DWFT frame is essentially a subband produced from a particular combination of the lowpass and band-pass filters used for filtering along the x and y -axes. Typically, the “high-low” (HL) and “low-high” (LH) frames are adequate for our applications. For every ZCP on HL and LH frames, we calculate and threshold its ZCS value, based on a localized mean-variance technique to eliminate weak and false edges, and mark the ZCS values that exceed the thresholds in edge maps M_v and M_h , respectively. The final edge map M_f is generated by $|M_f(x, y)| = \sqrt{|M_h(x, y)|^2 + |M_v(x, y)|^2}$, and $\tan \theta \xrightarrow{M_f(x, y)} = M_v(x, y)/M_h(x, y)$.

The multiresolution property of the ZCS in DIF can help reduce curvature trap, which can occur at two conjunctive edges, or a partial (convex or concave) contour formed by a group of edges. When a snake marches to a convex area, the balloon force first pushes it into the door in the first few iterations. Then, the coarser version of the ZCS measure, which covers a large area than its finer counterpart, would

trap snaxels expanded along the convexity area. Empirical results suggest that a two-level decomposition implementation usually suffices for most images.

II. IMPLEMENTAION AND EXPERIMENTS

For computer implementation, equation (1) can be rewritten as

$$E_{snake}(v_1, v_2, \dots, v_n) = \sum_{i=1}^n [E_{int}(v_i) + E_{image_dir}(v_i) + E_{balloon}(v_i)] \quad (2)$$

where v_i is a snaxel. For the internal energy, the elastic force at v_i can be approximated by using $|(\partial v_i / \partial s)|^2 \approx (|v_i - v_{i-1}|)^2$, and the bending force by $|(\partial^2 v(s) / \partial s^2)|^2 \approx |v_{i-1} - 2v_i + v_{i+1}|^2$. The balloon force at snaxel $v_i = (x_i, y_i)$ can be derived using Green’s theorem, $\Delta v_{i-1} v_i = x_{i-1} * y_i - x_i * y_{i-1}$. Let $\overline{v_{i-1} v_i}$ and \overline{N}_i , respectively denote the line segment from v_{i-1} to v_i , and its normal. DIF at v_i is defined as the sum of the DIF measure for points on $\overline{v_{i-1} v_i}$. Let ZCS of each element $j \in \overline{v_{i-1} v_i}$ be denoted as \overline{ZCS}_j , the sum of the directional image force for segment $\overline{v_{i-1} v_i}$ becomes

$$DIF_i = \sum_{j \in \overline{v_{i-1} v_i}} \overline{ZCS}_j \cdot \overline{N}_i = \sum_{j \in \overline{v_{i-1} v_i}} |\overline{ZCS}_j| |\overline{N}_i| \cos \alpha_j$$

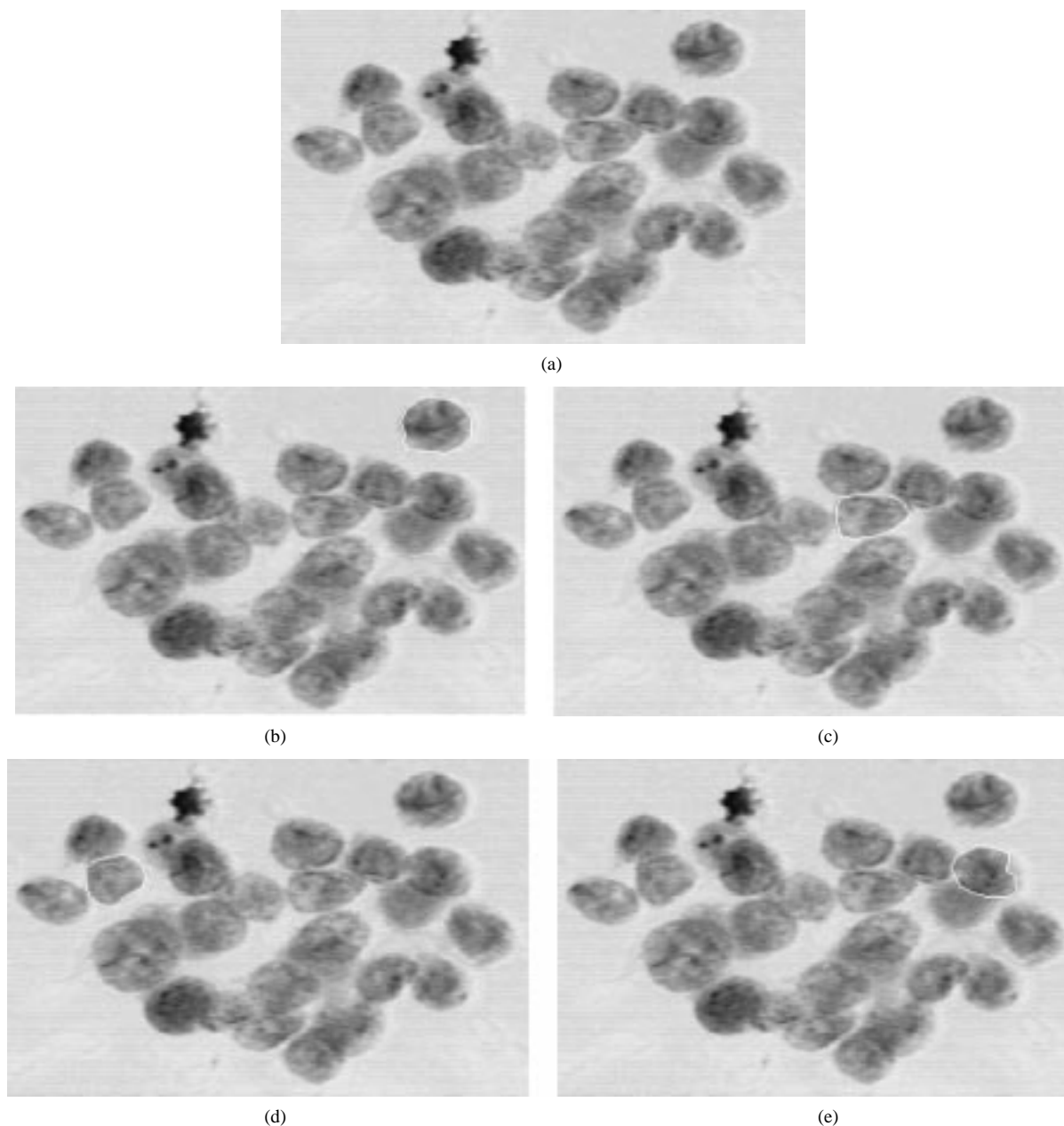


Fig. 8. Cell_3 experimental results: (a) Cell_3 image, (b) Cell_3 result1, (c) Cell_3 result2, (d) Cell_3 result3, and (e) Cell_3 result4.

where α_j is the angle between $\overline{ZCS_j}$ and \overline{N} , and $\overline{v_{i-1}v_i}$ is considered to be a real edge segment when its value is greater than a threshold value. By taking into account all the points on $\overline{v_{i-1}v_i}$, DIF_i represents the energy measure of the line, $\overline{v_{i-1}v_i}$, not that of a point.

To keep the average internal energy less sensitive to the snake size [9], snaxels are resampled when their distance exceeds the range of $[\delta_\ell, \delta_u]$, where $\delta_\ell = D/4$, $\delta_u = 1.5D$, and D is a constant. When the distance of $\overline{v_i v_{i+1}}$ is larger than $1.5D$, it is simple to add one additional snaxel between v_i and v_{i+1} . When deleting excess snaxels, however, one needs to keep snaxels close to corners or convexity of the local deformation [see Fig. 5(a)–(c)].

Minimizing (2) through exhaustive search is costly, i.e., $O(m^n)$, where n is the number of snaxels, and m the size of the search space for each snaxel. To reduce the computing costs, we adopted the dynamic programming approach proposed by Amini *et al.* [4], $E_{snake}(v_1, v_2, \dots, v_n) = \min_{v_{n-1}} [E_{snake}(v_1, v_2, \dots, v_{n-1}) + E(v_{n-1}, v_n)] E_{snake}(v_1, v_2, \dots, v_{n-1})$. Knowing that the energy

function at each snaxel is based only on local image features, the energy of a snake with n snaxels is

$$E_{snake}(v_1, v_2, \dots, v_n) = E_0(v_0, v_1, v_2) + E_1(v_1, v_2, v_3) + \dots + E_{n-1}(v_{n-1}, v_0, v_1) \quad (3)$$

where $E_{i-1}(v_{i-1}, v_i, v_{i+1})$ is the energy at each snaxel v_i , and $E_{i-1}(v_{i-1}, v_i, v_{i+1}) = E_{int}(v_{i-1}, v_i, v_{i+1}) + E_{ext}(v_i)$. S_i is updated at each iteration, based on the information on its two adjacent snaxels

$$S_i(v_i, v_{i+1}) = \min_{v_{i-1}} [S_{i-1}(v_{i-1}, v_i) + E_{i-1}(v_{i-1}, v_i, v_{i+1})] \quad (4)$$

where the initial condition is set as $S_0(v_0, v_1) = 0$, and the search process starts from S_1 . An *energy matrix* is used to store the minimization value of optimal functions in the neighborhood of v_i . After S_{n-1}

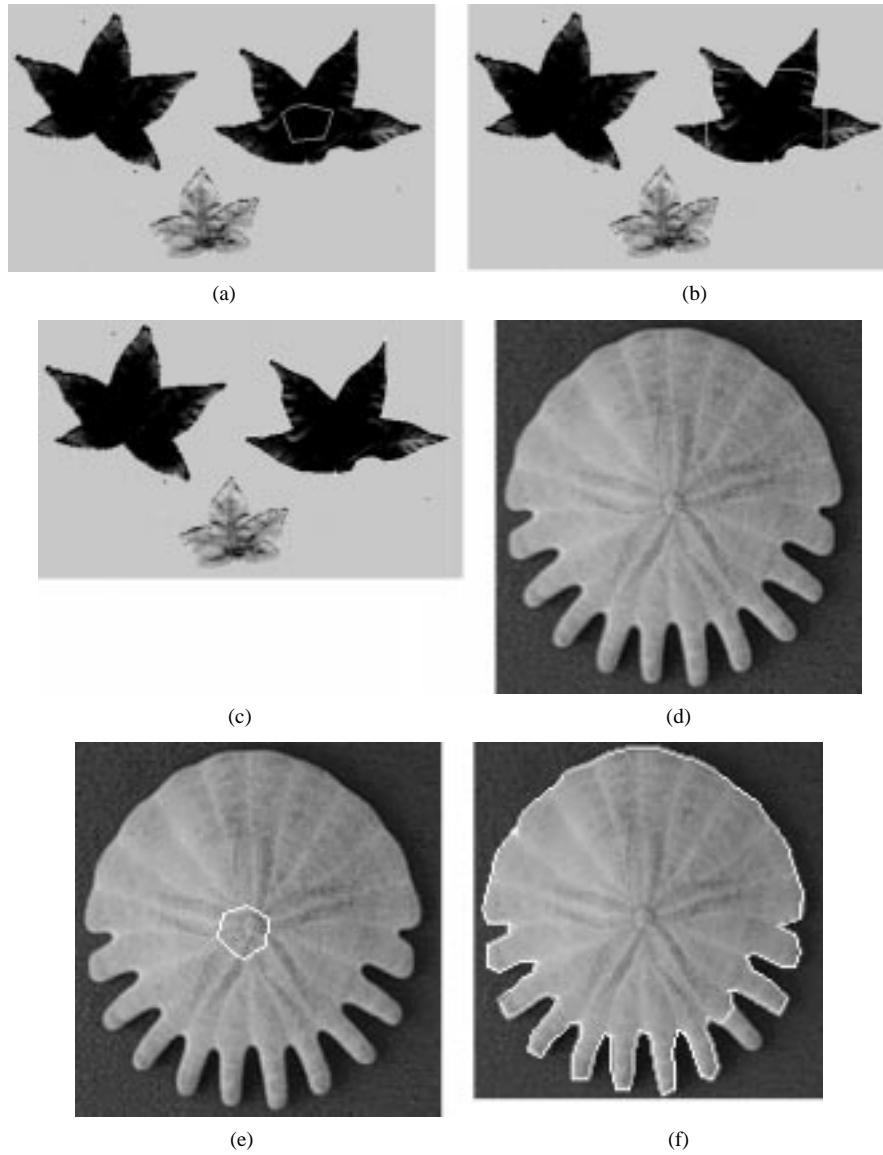


Fig. 9. (a) Leaf image and the snake after ten iterations, (b) 30 iterations, (c) final contour; (d) conch image, (e) initial snake, and (f) final contour.

is calculated, the snake with the minimum energy in an iteration is obtained by back-tracking of the energy matrix of snaxels. The snake algorithm terminates at iteration t , when the energy E_{snake} normalized over the snake length is minimized

$$E_{avg}^t(v) = \frac{E_{snake}(v_1, v_2, v_3 \cdots v_n)}{\sum_{i=1}^n |(v_i - v_{i-1})|}$$

Values of $E_{avg}^t(v)$ often oscillate in a few iterations before some snaxels can move to a new region. To overcome these problems, we keep track of variations of $E_{avg}^t(v)$ every T iterations. If the combined variations of $E_{avg}^t(v)$ is lower than a threshold in T iterations, we consider that the snake has reached its steady state. $|\sum_{i=0}^{T-1} (E_{avg}^{t-i}(v) - E_{avg}^{t-i-1}(v))| < \epsilon$, $t \pmod{T} = 0$, where ϵ is a predetermined threshold, and T is usually equal to five.

Let n and m , respectively, denote the number of snaxels, and the search space of each snaxel. For the $O(m^2)$ combinations between v_i and v_{i-1} , the complexity of calculating the elastic energy in one iteration is $O(nm^2)$. The time complexity of calculating the bending

energy is $O(nm^3)$ for each snaxel in an iteration. Calculating the balloon force for v_i needs to refer to its neighbor v_{i-1} , and the time complexity is $O(nm^2)$. The time complexity of calculating the image force is $O(nDm^2)$, where D is the number of pixels on the link between two snaxels. Since m is usually greater than D , or $O(nm^3) > O(nDm^2)$, the complexity of our scheme becomes $O(nm^3)$.

III. EXPERIMENTAL RESULTS AND SUMMARY

In our experiments, the distance between two consecutive snaxels is set to eight pixels for all cases. To test robustness of our scheme, the first experiment was performed on a synthetic circle, whose pixel value is 160 within the circle, and 60 otherwise, with added Gaussian noise. Our scheme works perfectly when the signal-to-noise ratio (SNR) is greater than 10 dB (see Fig. 6), but its performance deteriorates when the SNR falls lower than 5 dB. For real images, the algorithm worked well for Cell_1 and Cell_2 (see Fig. 7). The Cell_3 image (Fig. 8) is challenging, because of many blurred and overlapped cell boundaries. While over 90% of the snaxels marched to the real contour, in some cases, e.g., Fig. 8(e), the snake did not converge.

The next experiment is to test our algorithm for convexity adaptation. Referring to Fig. 9, this a very challenging case, because of hidden sharp edges, which produce strong image forces, in parallel with the leaf profile. Without any special calibration, our algorithm did find the correct convex shape, without being trapped by the (invisible) lines. In the Conch image, the snake successfully marched into all but one convex area, whose door width is very close to the minimal snaxel distance in the experiments. Reducing the distance between snaxels could push the snake into this most narrow door, but the computing times will increase. Snake deformation terminates in 20 to 40 iterations for all experiments. These experimental results clearly showed that our scheme can effectively detect corners and convexity on boundaries. The multiresolution solution approach has good boundary convergence and is robust to a much higher level of noise than conventional schemes.

REFERENCES

- [1] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," in *Proc. 1st Int. Conf. Computer Vision*, London, U.K., 1987, pp. 259–268.
- [2] M. Atkins and B. Mackiewicz, "Fully automatic segmentation of the brain in MRI," *IEEE Trans. Med. Imag.*, vol. 17, pp. 98–107, Feb. 1998.
- [3] C. Xu and J. Prince, "Snakes, shapes, and gradient vector flow," *IEEE Trans. Image Processing*, vol. 7, pp. 359–369, Mar. 1998.
- [4] A. Amini, T. Weymouth, and R. Jain, "Using programming for solving variational problems in vision," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, pp. 855–867, Sept. 1990.
- [5] L. Cohen and I. Cohen, "Finite-element methods for active contour models and balloons for 2-D and 3-D images," *IEEE Pattern Anal. Machine Intell.*, vol. 15, pp. 1131–1147, Nov. 1993.
- [6] A. Abrantes and J. Marques, "A class of constrained clustering algorithms for object boundary extraction," *IEEE Trans. Image Processing*, vol. 5, pp. 1507–1521, Nov. 1996.
- [7] P. Yuen, Y. Wong, and C. Tong, "Enhanced snakes algorithm for contour detection," in *Proc. Southwest Symp. Image Analysis Interpretation*, San Antonio, TX, Sept. 1996, pp. 54–59.
- [8] Z. Cvetkovic and M. Vetterli, "Discrete-time wavelet extrema representation: Design and consistent reconstruction," *IEEE Trans. Signal Processing*, vol. 43, pp. 681–693, Mar. 1995.
- [9] R. Samadani, D. Mihovilovic, C. Clauler, G. Wiederhold, J. Craven, and L. Frank, "Evaluation of an elastic curve technique for finding the auroral oval from satellite images automatically," *IEEE Trans. Geosci. Remote Sensing*, vol. 28, pp. 590–596, July 1990.

Image Data Compression Using Cubic Convolution Spline Interpolation

T. K. Truong, L. J. Wang, I. S. Reed, and W. S. Hsieh

Abstract—A new cubic convolution spline interpolation (CCSI) for both one-dimensional (1-D) and two-dimensional (2-D) signals is developed in order to subsample signal and image compression data. The CCSI yields a very accurate algorithm for smoothing. It is also shown that this new and fast smoothing filter for CCSI can be used with the JPEG standard to design an improved JPEG encoder-decoder for a high compression ratio.

Index Terms—Cubic convolution spline interpolation (CCSI), image compression, JPEG.

I. INTRODUCTION

Interpolation estimates the intermediate values of a set of discrete samples. It is used extensively in image data compression to magnify or reduce images and correct spatial distortion [1], [2]. In general, the process of decreasing the data rate is called decimation and the process of increasing data samples is called interpolation [3]. Interpolation functions [4] such as linear interpolation, cubic convolution interpolation (CCI) [1], [4], cubic B-spline interpolation [3], [5]–[7], etc., can be used in the image data compression process. The disadvantage of these interpolation schemes is that in general they are not designed to minimize the error between the original image and its reconstructed image. In 1981, Reed [8], and in 1998, Reed and Yu [2]¹ developed a linear spline interpolation scheme for resampling discrete image data. This linear spline interpolation is based on the least-squares method with a special linear interpolation function. Using an extension of the ideas of Reed in [2], a considerably modified version of the linear spline interpolation algorithm, called the cubic convolution spline interpolation algorithm, is developed here for the subsampling of discrete image data. The new CCSI scheme combines the least-squares method with a cubic convolution interpolation formula [1] for the decimation process. In addition, it is shown by computer simulation that the CCSI scheme, that is quite different from both cubic B-spline interpolation and cubic convolution interpolation, obtains a better quality of reconstructed image than all other interpolation methods mentioned above.

It is well known that the standard JPEG algorithm causes visually disturbing blocking effects when high quantization parameter is used to obtain a high compression ratio. The standard JPEG has a default quantization table and a table of default run-length codes. In order to obtain a better compression performance, an image can be processed to yield an image-specific quantization table. All of these quantization coefficients are then entropy encoded to generate image-specific variable-length coding tables. The compression methods that use these optimal tables are called optimized JPEG algorithms [9], [11]. The opti-

Manuscript received April 7, 1998; revised April 18, 2000. This work was supported by the National Science Council (NSC), R.O.C., under Grant NSC 89-2213-E-110-024. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Brian L. Evans.

T. K. Truong is with the College of Electrical and Information Engineering, I-Shou University, Kaohsiung, Taiwan 84008, R.O.C.

L. J. Wang and W. S. Hsieh are with the Department of Computer Science and Engineering, National Sun Yat-Sen University, Kaohsiung, Taiwan 804, R.O.C.

I. S. Reed is with the Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089-2565 USA.

Publisher Item Identifier S 1057-7149(00)07590-4.

¹This algorithm in this correspondence generalizes the algorithm, in Program ART, used by America Online (AOL).