Optimal Edge Detection in Two-Dimensional Images

Richard J. Qian and Thomas S. Huang

***Abstract*— This paper presents a new edge detection scheme that detects two-dimensional (2-D) edges by a curve-segment-based detection functional guided by the zero-crossing contours of the Laplacian-of-Gaussian (LOG) to approach the true edge locations. The detection functional is shown to be optimal in terms of signal-to-noise ratio (SNR) and edge localization accuracy; it also preserves the nice scaling property held uniquely by the LOG in scale space.**

I. Introduction

The difficulties in designing an accurate and robust edge detection algorithm for two-dimensional (2-D) images mainly come from two sources. First, there are tradeoffs in choosing an operator to pursue the best overall edge detection performance. Based on Yuille and Poggio's result [11], the 2-D Laplacian-of-Gaussian (LOG) operator [6] should be used because it is the only operator that has a constrained zero-crossing behavior in 2-D scale space which, in turn, lays a necessary foundation for scale space manipulations. However, an isotropic operator like the LOG is not optimal in terms of signal-to-noise ratio (SNR) and edge localization accuracy (ELA).The Canny edge detector [2] has better SNR and ELA than the LOG. However, the local extrema of its output may have unconstrained behaviors in 2-D scale space [11]. Moreover, the 2-D version of the Canny edge detector is obtained by simply extending its one-dimensional (1-D) version based on a linear constant cross-section edge model. As a result, the 2-D Canny edge detector is not optimal even in terms of SNR and ELA, except in the case when the detected edge is a straight line having a constant intensity. More recent efforts on finding an optimal edge detector can be found, e.g., in [1] and [10]. Unfortunately, the issue of establishing a more accurate 2-D edge model has still been overlooked.

Second, it is very difficult to find a reliable combining approach in 2-D scale space, since the zero-crossings of the second derivatives of the filtered 2-D signals generally behave in a much more complex way in scale space than the 1-D zero-crossings. Recent work on detecting edges in scale space can be found, e.g., in [7] and [9], and general discussion on scale space can be found in [5]. For a more thorough discussion on the general issues related to edge detection, see our technical report [8].

In this paper, we introduce a new 2-D edge detection functional that not only achieves the optimality on SNR and ELA for detecting edges in 2-D images but also preserves the nice scaling property of the LOG in scale space. We also briefly discuss related issues including edge regularization, adaptive thresholding, and scale space combination. A more detailed version of this paper can be found in [8].

II. Optimal 2-D Edge Detection and Edge Regularization

Because, in general, the linear constant cross-section edge model can represent only a short piece of a 2-D edge in reasonable precision, edge detection algorithms assuming such a model usually detect edges by extracting edge pixels in a point-by-point manner. The drawbacks of this type of method are twofold. First, it limits the best possible edge detection performance in terms of SNR and ELA, as we will see in Section II-C. Second, it directly causes the well-known streaking

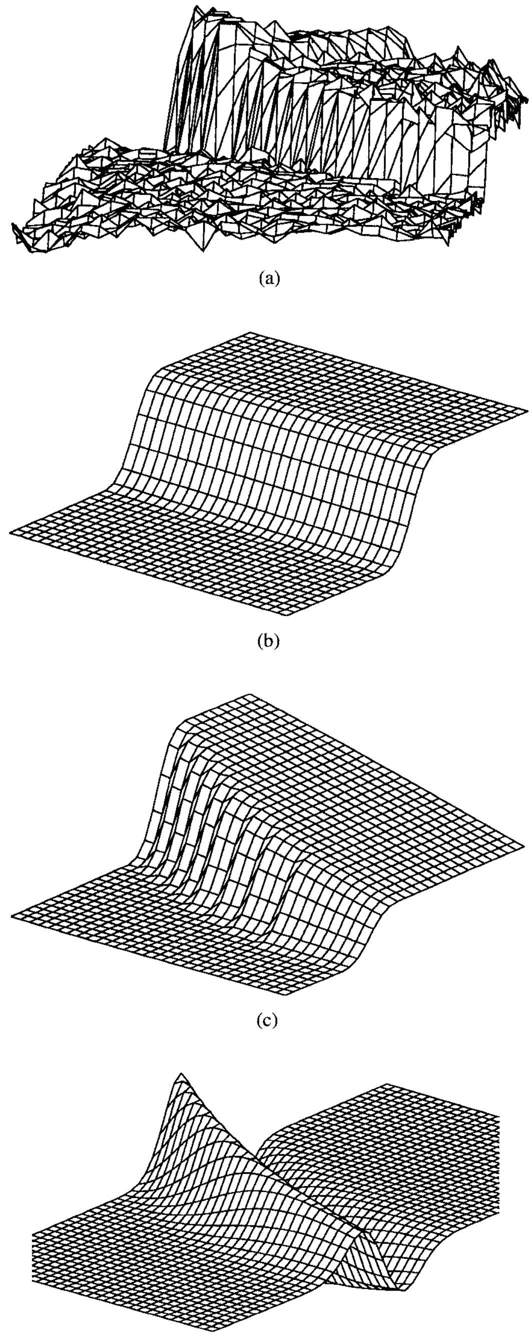


Fig. 1. Example 2-D edge and its surface models. (a) Intensity surface around a short edge segment in the Lena picture. (b) Constant cross-section edge model. (c) Parametric 2-D edge model as described in Section II-A for the edge in (b). (d) Constructed optimal 2-D edge detection functional as described in Section II-B, based on the parametric 2-D edge model in (c).

Manuscript received November 8, 1994; revised September 28, 1995. This work was supported by ARPA/ONR under Grant N00014-93-1-1167.

The authors are with the Beckman Institute and Coordinate Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: qian@ifp.uiuc.edu).

Publisher Item Identifier S 1057-7149(96)04541-1.

1057–7149/96$05.00 © 1996 IEEE

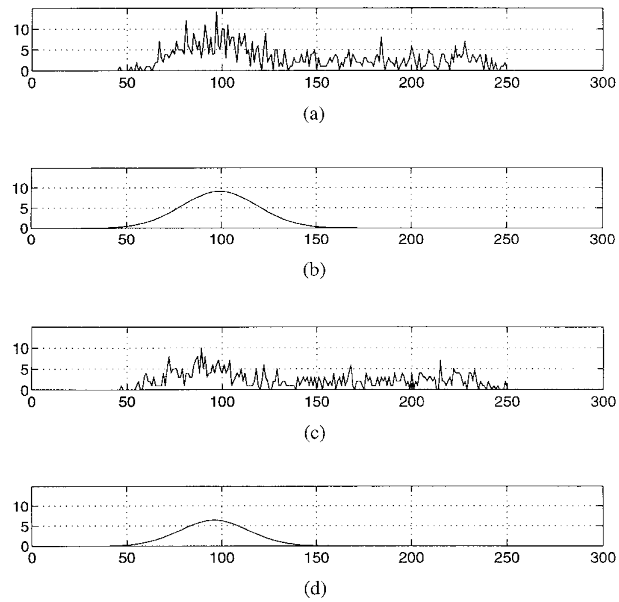


Fig. 2. Histograms of the normalized strengths of the candidate edge segments of the Lena picture in two different channels. (a) Histogram in channel 1 with $\sigma = 2.5$. (b) Fitted Gaussian distribution for noise estimation corresponding to (a). (c) Histogram in channel 2 with $\sigma = 3.2$. (d) Fitted Gaussian distribution for noise estimation corresponding to (c).

problem in edge detection results. In the following, we introduce a more accurate 2-D edge model and develop an optimal 2-D edge detection functional based on such a model. We compare the SNR and ELA of the new edge detector with that of the Canny edge detector.

A. Parametric 2-D Edge Model

Assume that a 2-D edge has a trajectory $\vec{\alpha}(u) = (x(u), y(u))^T$ in the image plane, where u is the parameter of the trajectory. Then the intensity surface around the 2-D edge can be modeled as a parameterization of the local coordinates (u, v) [3] as follows:

$$E(u,v) = A(u)P(u,v), \quad -L \le u \le L,\ -D \le v \le D \qquad (2\text{-}1)$$

where v is the parametric coordinate in the gradient direction of intensity, $A(u)$ is the amplitude function, $P(u,v)$ is the profile function of the edge, and $[-L, L] \times [-D, D]$ defines the region of support of the edge surface. Note that if we let $\vec{\alpha}(u)$ be a straight line in the image plane and $A(u)$ be a constant, the above model degenerates to the constant cross-section model. For a step edge in real image, $P(u,v)$ can be modeled as the convolution of a Gaussian function and an ideal step function, i.e.,

$$P(u,v) = \frac{1}{\sqrt{2\pi}\sigma(u)} \times \int_0^{\infty} \left[\exp\left(-\frac{(v-\xi)^2}{2\sigma^2(u)}\right) - \exp\left(-\frac{(v+\xi)^2}{2\sigma^2(u)}\right)\right] d\xi, \quad -L \le u \le L,\ -D \le v \le D \qquad (2\text{-}2)$$

where the variance $\sigma^2(u)$ of the Gaussian determines the scale of the edge profile. Note that we have subtracted the mean value of the step function before the convolution. The profile functions for other types of edges can be found in a similar way as for the step edge. See Fig. 1 for an example 2-D edge and its corresponding constant cross-section edge model and the parametric edge model.

B. An Optimal 2-D Edge Detection Functional

Let $E(u,v)$ be an edge with a trajectory $\vec{\alpha}(u)$ to be detected in an image. We assume that $\vec{\alpha}(u)$ is a regular curve, i.e., the tangent of the curve is defined at every point on the curve. We construct a 2-D edge detection functional Q_{edge} with two functions $f_t(u)$ and $f_g(u,v)$ to be optimized as follows:

$$Q_{\text{edge}}(u_0, v_0(u)) = \int_{-L}^{L} f_t(u)\left[\int_{-D}^{D} f_g(u_0-u, v) \times E(u_0-u, v_0(u_0-u)-v)dv\right]du \qquad (2\text{-}3)$$

where $v_0(u)$ is a trajectory of parameter v with respect to parameter u.

Assuming an additive white Gaussian noise with variance n_0^2 in the image and using the 2-D edge model defined by (2-1), we show in [8] that the output SNR of the edge detection functional Q_{edge} is

$$\text{SNR} = \frac{\left|\int_{-L}^{L} f_t(u)A(-u)\int_{-D}^{D} f_g(-u,v)P(-u,-v)dvdu\right|}{n_0\sqrt{\int_{-L}^{L} f_t^2(u)\int_{-D}^{D} f_g^2(-u,v)dvdu}} \qquad (2\text{-}4)$$

and the ELA is

$$\text{ELA} = \frac{n_0\sqrt{\int_{-L}^{L} f_t^2(u)\int_{-D}^{D}\left(\frac{\partial}{\partial v}f_g(-u,v)\right)^2 dvdu}}{\left|\int_{-L}^{L} f_t(u)A(-u)\int_{-D}^{D}\frac{\partial}{\partial v}f_g(-u,v)\frac{\partial}{\partial v}P(-u,-v)dvdu\right|}. \qquad (2\text{-}5)$$

If we assume that the profile functions of the 2-D edge have the same scale factor $\sigma(u)$, by recalling (2-2) the above output SNR and ELA can then be simplified as follows:

$$\text{SNR} = \frac{\left|\int_{-L}^{L} f_t(u)A(-u)du\int_{-D}^{D} f_g(0,v)P(0,-v)dv\right|}{n_0\sqrt{\int_{-L}^{L} f_t^2(u)du\int_{-D}^{D} f_g^2(0,v)dv}} \qquad (2\text{-}6)$$

and

$$\text{ELA} = \frac{n_0\sqrt{\int_{-L}^{L} f_t^2(u)du\int_{-D}^{D}\left(\frac{\partial}{\partial v}f_g(0,v)\right)^2 dv}}{\left|\int_{-L}^{L} f_t(u)A(-u)du\int_{-D}^{D}\frac{\partial}{\partial v}f_g(0,v)\frac{\partial}{\partial v}P(0,-v)dv\right|}. \qquad (2\text{-}7)$$

The edge detection functional Q_{edge} can be optimized by maximizing the SNR and minimizing the ELA. If no additional constraint is selected, the optimal solution can be obtained by using the Cauchy–Schwartz inequality. Namely, the optimal detection function along the edge trajectory $f_t^*(u)$ is

$$f_t^*(u) = A(-u), \quad -L \le u \le L \qquad (2\text{-}8)$$

and the optimal detection function in the gradient direction $f_g^*(u,v)$ for a step edge is

$$f_g^*(u,v) = P(u,-v) = \frac{1}{\sqrt{2\pi}\sigma(u)}\int_0^{\infty}\left[\exp\left(-\frac{(v+\xi)^2}{2\sigma^2(u)}\right) - \exp\left(-\frac{(v-\xi)^2}{2\sigma^2(u)}\right)\right]d\xi, \quad -L \le u \le L,\ -D \le v \le D. \qquad (2\text{-}9)$$

The above results are also recognized as the matched filters in information theory.

In practice, the exact locations of edge trajectories in an image cannot be known beforehand. That is, indeed, the purpose of edge detection. However, the edges may be approximately located by first

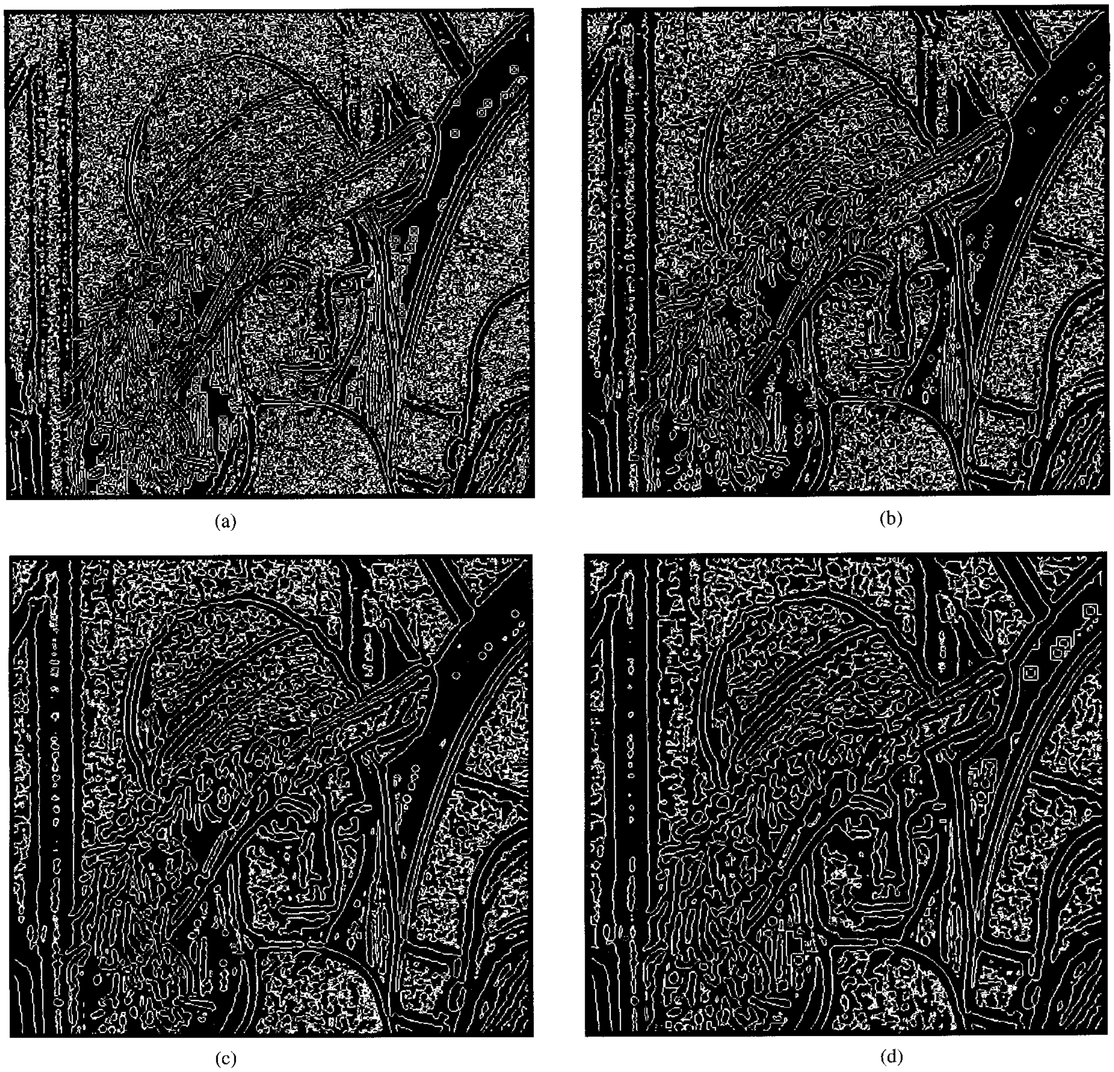


Fig. 3. LOG zero-crossing contours of the Lena picture at different scales. (a) LOG zero crossing contours at $\sigma = 2.5$. (b) LOG zero-crossing contours at $\sigma = 3.2$. (c) LOG zero-crossing contours at $\sigma = 3.9$. (d) LOG zero-crossing contours at $\sigma = 4.6$. See Fig. 5(a) for the input intensity image. From the above images, it can be seen that a zero-crossing contour that has a cleaner neighborhood usually corresponds to a more salient edge in the picture.

applying some nonoptimal edge detector, in our case the LOG, to the image. In order to determine the true location of an edge, we then search for a trajectory $v_0^*(u)$ that maximizes the response of the detection functional Q_{edge} in some neighborhood around the corresponding LOG zero-crossing contour. We choose the LOG as our preliminary edge detector because it posses the best scaling property in scale space. Using the LOG zero-crossing contours as its initial conditions, the optimal edge detection functional Q_{edge} therefore inherits the nice scaling property from the LOG. Other details on implementing the edge detection functional Q_{edge} are referred to in [8].

C. SNR and ELA Improvement over the Canny Edge Detector

Let $E(u,v)$ be an edge with a trajectory $\vec{\alpha}(u)$ to be detected in an image. Assuming $\vec{\alpha}(u)$ is curved or $E(u,v)$ has nonuniform amplitude, then we have

$$\frac{\text{SNR}^{Q_{\text{edge}}}}{\text{SNR}^{\text{Canny}}} = \frac{\text{ELA}^{\text{Canny}}}{\text{ELA}^{Q_{\text{edge}}}} = \frac{\sqrt{\int_{-L}^{L} A^2(u)du}}{A(u)}, \quad -L \leq u \leq L. \tag{2-10}$$

Namely, the edge detection functional Q_{edge} has improved the SNR and ELA simultaneously over the Canny edge detector by a factor of the number defined in (2-10). In the case where $E(u,v)$ has a curved trajectory $\vec{\alpha}(u)$ but uniform amplitude $A(u)$, the above factor is simply the square root of the total length of the edge, i.e.,

$$\frac{\text{SNR}^{Q_{\text{edge}}}}{\text{SNR}^{\text{Canny}}} = \frac{\text{ELA}^{\text{Canny}}}{\text{ELA}^{Q_{\text{edge}}}} = \sqrt{2L}, \quad -L \leq u \leq L. \tag{2-11}$$

The reason that the edge detection functional Q_{edge} has better SNR and ELA than the Canny edge detector is mainly due to the fact that Q_{edge} always detects edges piecewisely while the Canny detector is pointwise in the case where edges are curved and/or the amplitude along edges is nonuniform. The increased detecting range along edge direction improves the output SNR and lowers the probability of declaring spurious response as edges; therefore, it improves the ELA.

D. Edge Regularization

Due to noise and discretization error, a numerical optimal solution of the trajectory $v_0^*(u)$ that maximizes the edge detection functional Q_{edge} may not be continuous or smooth. To improve the continuity

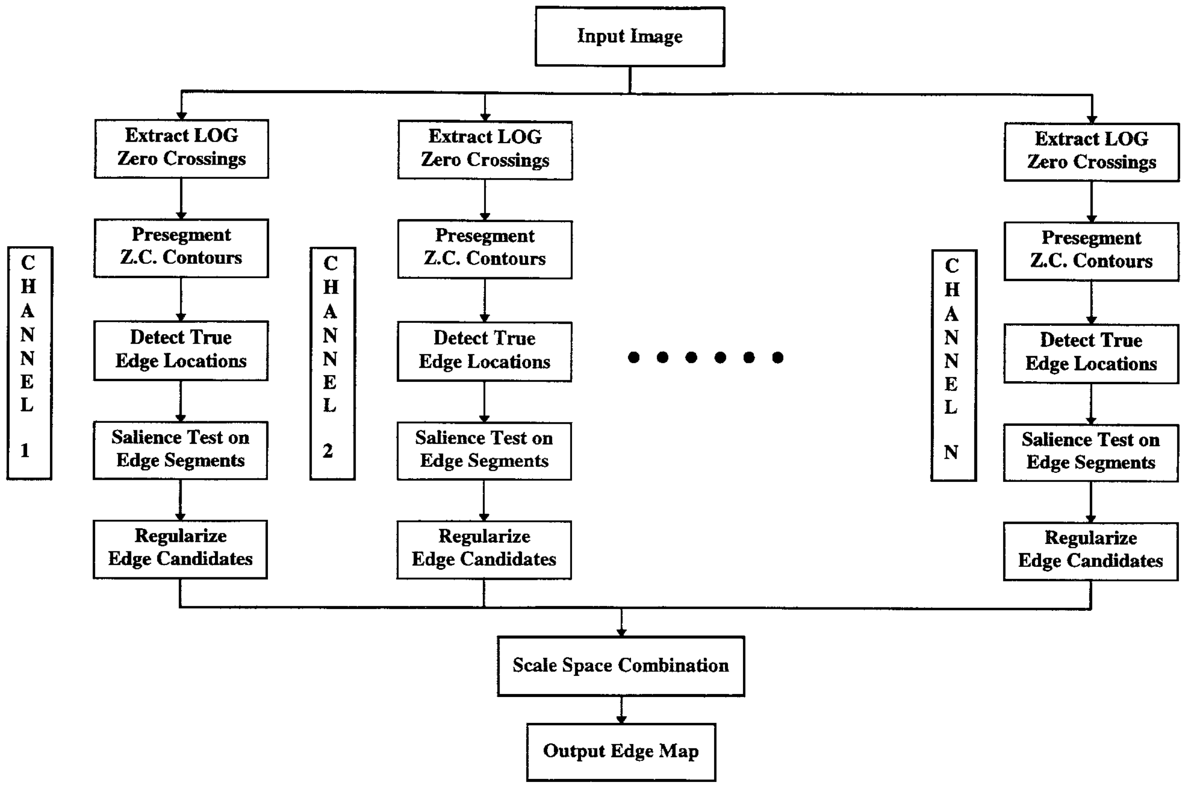


Fig. 4. Diagram of the implementation of the 2-D edge detection scheme.

and/or smoothness, we define energy functionals associated with an edge and then minimize the total energy functional of the edge to obtain an optimal trajectory that has some desired continuity and/or smoothness characteristics for that edge. The mathematical forms of the edge energy functionals and the numerical procedure of minimizing the total energy functional are similar to that reported in [4], and are referred to in [8].

III. Detecting Salient Edges at Different Scales

Edges can occur over a wide range of strengths as well as scales in real images. In order to generate a complete yet clean edge map, some thresholds are needed, and they have to be selected adaptively. Detection results from different scales are also needed to be combined. We discuss all these issues in this section.

A. Evaluating the Salience of an Edge

In order to derive an adaptive thresholding mechanism, we first need a procedure to estimate the noise level in an image. Here we introduce a global noise estimation approach. In step 1, it computes the magnitude of gradient in the Gaussian blurred image for every point on the candidate edge segments and assigns the magnitude to that point as its strength. Then all the strengths are equalized into the full scale of intensity. In step 2, the approach computes the normalized strength for each candidate edge segment. We compute the normalized strengths for all the candidate edge segments in all the channels. A histogram of the computed edge segment strengths can then be constructed for each channel. Due to random noise in an image, a Gaussian distribution can always be found at the low intensity end in the histogram constructed as above. Finally, in step 3, the approach fits the low-intensity part of the histogram into a Gaussian distribution. See Fig. 2 for an example. The mean and the variance of the Gaussian reflect the average global noise level and the spreading range of the noise. Based on these values, a thresholding mechanism is derived [8].

To generate edge detection results similar to those perceived by human visual systems, we also employ two additional thresholding criteria that were selected based on the physiological evidence from biological visual systems. The first criterion is based on the physiological evidence that unbalanced difference-of-Gaussian (DOG) operators are employed in the biological visual system. We implement this criterion into our edge detection scheme by applying a weighting function to the evaluation of the normalized strengths of candidate edge segments such that the strength of a candidate edge segment with low absolute intensity is suppressed. The second criterion is based on the visual behavior of lateral inhibition found in many biological visual systems. We implement this criterion by inferring that a zero-crossing contour that has a cleaner neighborhood may imply a higher probability to be a salient edge in the image. See Fig. 3.

B. Combining Edges from Different Scales

Our edge detection scheme generates a final edge map by combining the detected salient edge candidates from different scale channels. Since smaller scale of the detection functional always gives better ELA, and only salient edges with high SNR survive through the salience test, the final edge map always incorporates all the detected salient edge candidates from the smallest scale. The combination procedure can then continue to check if there are new salient edges in the detection results from larger scales. Our edge detection scheme preserves the nice scaling property of the LOG zero-crossing contours, and our edge candidates in large-scale channels do

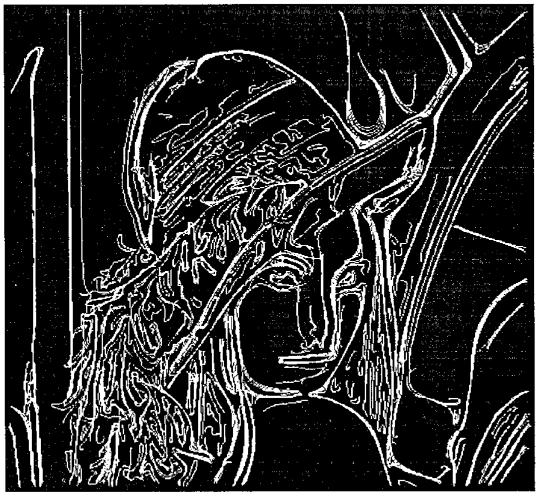
Fig. 5. Comparison of the edge detection results on the Lena picture. (a) Input intensity image. (b) Detected edges by the proposed 2-D edge detection functional $Q_{\rm edge}$ with $\sigma = 2.5$. (c) Detected edges by the 2-D Canny edge detector with $\sigma = 2.5$. (d) Detected edges by the Marr–Hildreth edge detector with $\sigma = 2.5$.

not move far away from the true edge position due to the optimal ELA performance of $Q_{\rm edge}$; this ensures that our scale space approach performs reliably on real images. Moreover, our edge candidates in each scale channel are curve segments, not points; therefore, our method is less sensitive to noise than other existing approaches.

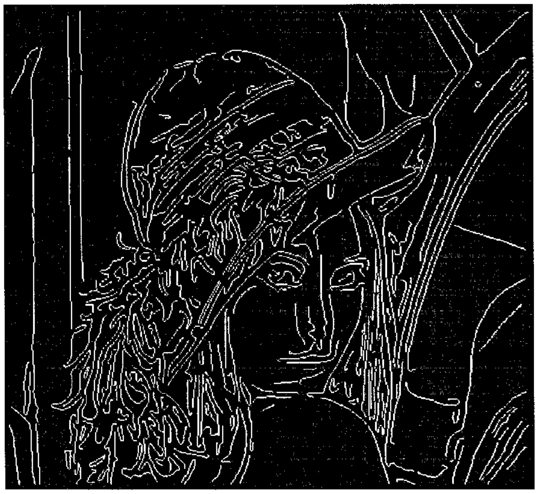
IV. Experimental Results

The proposed 2-D edge detection scheme has been implemented and tested on a large number of real images. Fig. 4 gives a diagram of our implementation. The details are referred to in [8]. The computations of the scheme consist of mainly obtaining the LOG zero-crossing contours and optimizing the contours. The cost of obtaining the LOG zero-crossing contours is mainly the cost of 2-D convolutions; the cost of optimizing one contour is mainly the cost of inverting a sparse matrix, which is of the order $O(N)$, where N is the length of the contour, due to the use of a sparse matrix technique in our implementation. The details are referred to in [8].

In Section II-C, we gave the closed-form performance comparison in terms of SNR and ELA between the proposed 2-D edge detection functional $Q_{\rm edge}$ and the Canny edge detector. Fig. 5 shows the actual edge detection results on the Lena picture by $Q_{\rm edge}$, the Canny edge detector, and the Marr–Hildreth edge detector, all with $\sigma = 2.5$. From Fig. 5, it can be seen that the edge detection result by $Q_{\rm edge}$ is clean while the edges remain continuous and smooth. On the other hand, the result by the Marr–Hildreth edge detector appears to be noisy at the selected threshold while the edges have begun to lose their continuity. The result by the Canny edge detector appears to be better than that by the Marr–Hildreth edge detector. However, while its edge map still contains more spurious edges than that by $Q_{\rm edge}$ shown in Fig. 5(b), its detected edges are not as continuous and smooth as the detected edges by $Q_{\rm edge}$ either. Fig. 6(a) shows the overlapping edge candidates on the Lena picture detected by $Q_{\rm edge}$ in all seven channels with $\sigma_{\rm min} = 2.5$ and $\sigma_{\rm max} = 6.7$. Fig. 6(b) gives the final edge map of the Lena picture generated by the scale space combination procedure implemented in our 2-D edge detection scheme.



(a)



(b)

Fig. 6. More edge detection results on the Lena picture: (a) Overlapping edge candidates detected by the proposed 2-D edge detection functional Q_{edge} in all seven channels with $\sigma_{\min} = 2.5$ and $\sigma_{\max} = 6.7$. (b) Final edge map generated by the scale space combination procedure employed in the proposed 2-D edge detection scheme.

V. Summary

We have proposed a new 2-D edge detection functional derived from an adaptive 2-D edge model. The detection functional is optimal in terms of SNR and ELA for detecting edges in 2-D images, and preserves the nice scaling behavior of the LOG operator in scale space. It detects edges based on edge segments rather than edge points. This guarantees the continuity of the detected edges and greatly reduces the impact of random noise on the detection results. The proposed edge detection scheme also employs an edge regularization procedure to enhance desired smoothness and stiffness on the detected edges. A global noise-estimation procedure and the other two physiologically based criteria have also been introduced to provide a robust estimation for the global noise level in the image, and to fine-tune the edge detection results to make them similar to what is perceived by the human visual system. Finally, a reliable scale space combining procedure has been established based on continuity in image and the scaling property in scale space of the detected edges.

References

[1] K. L. Boyer and S. Sarkar, "Comments on 'On the localization performance measure and optimal edge detection'," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 16, no. 1, pp. 106–108, 1994.

[2] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-8, no. 6, pp. 679–698, 1986.

[3] M. P. D. Carmo, *Differential Geometry of Curves and Surfaces*. Englewood Cliffs, NJ: Prentice-Hall, 1976.

[4] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vision*, pp. 321–331, 1988.

[5] T. Lindeberg, *Scale Space Theory in Computer Vision: The Kluwer International Series in Engineering and Computer Science*. Dordrecht, The Netherlands: Kluwer, 1994.

[6] D. Marr and E. Hildreth, "Theory of edge detection," in *Proc. Roy. Soc. London.*, vol. B.207, pp. 187–217, 1980.

[7] P. Perona and J. Malik, "Scale space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, no. 7, pp. 629–639, 1990.

[8] R. J. Qian and T. S. Huang, "A 2-D edge detection scheme for general visual processing," *Tech. Rep. Beckman Inst.*, UIUC, no. UIUC-BI-AI-RCV-94-04, 1994.

[9] M. Tabb and N. Ahuja, "Multiscale image segmentation using a recent transformation," in *Proc. ARPA Image Understand. Workshop*, 1994, vol. 2.

[10] H. D. Tagare and R. J. P. deFigueiredo, "On the localization performance measure and optimal edge detection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, no. 12, pp. 1186–1190, 1990.

[11] A. L. Yuille and T. A. Poggio, "Scaling theorems for zero crossings," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-8, no. 1, pp. 16–25, 1986.