EDGE DETECTION BASED ON THE MULTIRESOLUTION FOURIER TRANSFORM

Chang-Tsun Li and **Der-Chyuan Lou**
Department of Electrical Engineering,
Chung-Chen Institute of Technology,
Tashi, Taoyuan 33509, Taiwan, R.O.C.
{ctli, dclou}@cc04.ccit.edu.tw

Abstract - In this work, an edge detection technique is proposed by using the Multiresolution Fourier Transform (MFT) to analyze the local properties in the spatial frequency domain. Five major steps are adopted to implement the detection of edges. First, the Laplacian Pyramid method is used to create a high-pass filtered image. Secondly, the Multiresolution Fourier Transform (MFT) is applied to divide the high-pass filtered image into blocks and transform each of the blocks into spatial frequency domain. Thirdly, single-feature and non-single-feature blocks are differentiated. Subsequently, the blocks containing single feature are then subject to a process for estimating the orientation and the centroid of the feature in order to locate it. Finally, the accuracy of the estimated centroid of the local feature is checked. Once all the blocks are analyzed at a resolution level, the overall procedure is repeated at the next resolution level and the blocks with their father block being classified as *non-single-feature* or being rejected in the accuracy check stage at the previous level are analyzed. The algorithm stops when a specific level is reached.

1. INTRODUCTION

Edge detection has long been attempted at automated vision tasks based on the fact that luminance transitions in an image, which correspond to lines and edges, are perceptually essential to describe the visual information about the scene. For example, a football can be identified from the grass in the background by detecting the edge contour of the football (or the boundary between the football and the grass background). So, edge detection methods have been widely applied to 2-D and 3-D interpretation tasks such as the recognition of objects, stereopsis, and the inferring of 3-D structure from shape and motion.

Basically, the scene contained in an image can be seen as consisting of regions with constant or slowly varying luminance. Therefore, the local features such as lines or edges are the region boundaries, which represent the relatively rapid transitions in luminance. Physiological investigation into the functioning of the Human Visual System (HVS) also motivates the use of local features to interpret scenes. It is observed that simple cells respond to local features and their orientation [1][2][3]. These discoveries encouraged a number of theoretical proposals for vision, among the best known of these is probably the work proposed by Marr [4], in which a variety of vision problems were attacked using local features as a basis.

It is also recognized that features may appear at different scales/resolution. Therefore, edge detection methods performing at single resolution/scale is inadequate to detect features appearing at other scales. It is our intention in this

0-7803-5650-0/99/$10.00 © 1999 IEEE

work to propose a multiresolution approach to detect edges at different scales. The algorithm is based on a framework similar to the work of Calway's [5] with some new features added. The main differences between the present work and [5] are:

1) A high-pass filtering stage preceding all the stages in [5] is added to enhance the edges, which corresponds to the high frequency components of the underlying image.
2) The method employed to differentiate single and non-single blocks are different from that of [5].
3) Instead of taking the direction of the eigenvector associated with the largest eigenvalue of the "inertia tensor", the direction along which the spectral energy clusters is employed as the orientation of the edge.
4) An accuracy check stage is added to ensure that the estimates of edges are accurate to some extent. This is a stage absent form [5].

The algorithm proposed is basically a frequency domain method based on the Multiresolution Fourier Transform (MFT) [6].

The work is organized as follows: Section 2 explains the core of the proposed algorithm and steps of our edge detection method are explained. Section 3 demonstrates some experimental results. Section 4 concludes this work and suggests some future works to be done.

2. EDGE DETECTION

Five major steps are adopted to implement the detection of edges and are described as follows.

2.1 High-pass Filtering

In order to help detecting the edges, we want the luminance discontinuity to be emphasized. One way to achieve this is to high-pass filter the original image so as to reduce the influence of low frequency components of the image while preserving and enhancing the high frequency information. We utilize a method similar to Burt *et al*'s Laplacian Pyramid [7] to get a high-pass filtered version of the original image. However, in order to make the size of the weighting kernel consistent with the Multiresolution Fourier Transform (to be introduced later), we use the 4×4 unimodal Gaussian-like weighting kernel proposed by Wilson and Bhalerao [8].

2.2 Multiresolution Fourier Transform

Adopting the Multiresolution Fourier Transform (MFT) derived by Wilson *et al* [5], we create an image pyramid with K resolution levels based on the original image of size $N \times N = 2^K \times 2^K$. Each level k consists of an array of $2^k \times 2^k$ square blocks taken from the original image by sliding a sampling window of size $2^{K-k} \times 2^{K-k}$ over the image in 2^{K-k-1}-pixel wide steps in both horizontal and vertical directions. This means that the adjacent windows in either horizontal or vertical directions are half-overlapped. Since the size of the widow used in a level is a quarter of the one used at the immediate ancestor level, the pyramid thus created conforms to a quad-tree structure. The 50% overlap is aimed at reducing the artifacts. The multiresolution Fourier transform is done by transforming the windowed blocks at each level into frequency domain. The readers are referred to [6] for the theoretical details of the MFT.

2.3 Differentiation of Single-feature and Non-single-feature Blocks

Before the edges can be detected, blocks containing single feature are discriminated from the *non-single-feature* blocks so the algorithm can focus on the *single-feature* blocks in the following steps to locate the edges. It is observed that if a block contains single linear feature such as edge or line, the energy in frequency domain will concentrate along an axis oriented perpendicularly to the feature's orientation in spatial domain. Therefore, by measuring the degree of energy concentration along a specific direction, we can obtain the confidence about the existence of a linear feature in the block. Hsu and Wilson have used a pair of centroid vectors to represent a local spectrum for texture analysis and synthesis [9]. Although they did not exploit this vector to detect edges, we found that a proper modification to they method can be employed to estimate the orientation of linear features. Each of the vector pair corresponds to the centroid of the Fourier coefficients within a section of the **half-plane** spectrum as shown in Figure 1. Figure 1 illustrates the partitioning of a local spectrum, in which, θ_1 is used to divide the local spectrum into two half-planes and θ_2 is used to sub-divide the half-plane starting at θ_1 into two sections, Λ_1 and Λ_2. Because of Hermitian symmetry, to analyze the Fourier spectrum only a half plane of the local spectrum has to be involved. Fourier spectrum is analyzed in the implementation of Hsu's

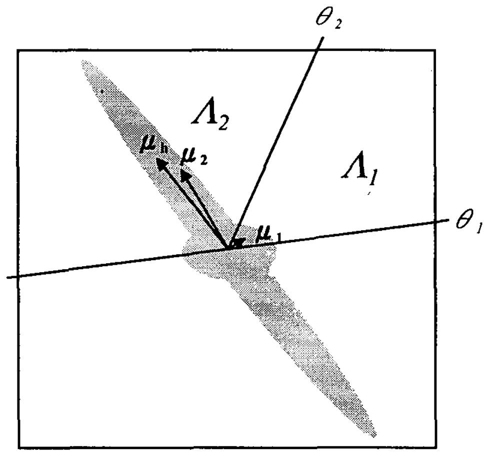


Figure 1. Division of a spectral *half-plane* into two sections by θ_1 and θ_2. Because of Hermitian symmetry, only a half plane has to be taken into account.

algorithm. Similar approach is used in this work. However, after some experiments, power spectrum, in stead of Fourier spectrum, is used to implement the algorithm in this work in order to get more accurate feature orientation.

To estimate the centroid vector pair $(\vec{\mu}_1, \vec{\mu}_2)$ of a local spectrum $\hat{f}(\vec{\omega})$, the values of two variable angles θ_1 and θ_2 (see Figure 1) are to be determined such that the *weighted sum of variance* of the two sections

$$v_s = \frac{M_1(\theta_1,\theta_2)\sigma_1^2(\theta_1,\theta_2) + M_2(\theta_1,\theta_2)\sigma_2^2(\theta_1,\theta_2)}{M_1(\theta_1,\theta_2) + M_2(\theta_1,\theta_2)} \tag{1}$$

is minimized, where

$$M_i(\theta_1,\theta_2) = \sum_{\vec{\omega}\in\Lambda_i(\theta_1,\theta_2)} \left|\hat{f}(\vec{\omega})\right|^2 \quad ,i=1,2 \tag{2}$$

$$\sigma_i(\theta_1,\theta_2) = \frac{1}{M_i(\theta_1,\theta_2)} \sum_{\vec{\omega}\in\Lambda_i(\theta_1,\theta_2)} \left|\hat{f}(\vec{\omega})\right|^2 \left\|\vec{\omega} - \vec{\mu}_i(\theta_1,\theta_2)\right\|^2 \quad ,i=1,2 \tag{3}$$

$$\vec{\mu}_i(\theta_1,\theta_2) = \frac{1}{M_i(\theta_1,\theta_2)} \sum_{\vec{\omega}\in\Lambda_i(\theta_1,\theta_2)} \left|\hat{f}(\vec{\omega})\right|^2 \vec{\omega} \quad , i=1,2 \tag{4}$$

Λ_1 and Λ_2 are the sets of coordinates in each of the two sections of the half-plane starting at angle θ_1 and divided at angle θ_2. To find the minimum v_s, several combination of angles θ_1 and θ_2 is tested. The intervals of θ_1 and θ_2 are $0 \le \theta_1 < \pi$ and $\theta_1 < \theta_2 < \theta_1 + \pi$.

After v_s, θ_1 and θ_2 in Equation (1) are found, the centroid vector $\vec{\mu}_h(\theta_1)$ of the **half-plane** starting at θ_1 is estimated in order to calculate the variance v_h of the half-plane.

$$v_h = \frac{1}{M_h(\theta_1)} \sum_{\vec{\omega}\in\Lambda_h(\theta_1)} \left|\hat{f}(\vec{\omega})\right|^2 \left\|\vec{\omega} - \vec{\mu}_h\right\|^2 \tag{5}$$

$$M_h(\theta_1) = \sum_{\vec{\omega}\in\Lambda_h(\theta_1)} \left|\hat{f}(\vec{\omega})\right|^2 \tag{6}$$

$$\vec{\mu}_h(\theta_1) = \frac{1}{M_i(\theta_1)} \sum_{\vec{\omega}\in\Lambda_i(\theta_1)} \left|\hat{f}(\vec{\omega})\right|^2 \vec{\omega} \tag{7}$$

where $\Lambda_h = \Lambda_1 \cup \Lambda_2$ is the set of coordinates of the half-plane starting at angle θ_1.

Once v_h is calculated, the variance ratio v_s/v_h is utilized to decide whether or not the local spectrum contains single feature. Given a threshold α, if

$$\frac{v_s}{v_h} \geq \alpha, \quad 0 < \alpha \leq 1 \tag{8}$$

then the block is classified as *single-feature*, otherwise, it classified as non-single-feature block and will be divided into 4 sub-blocks so as to be analyzed at the next resolution level. The physical meaning behind this idea is: if there is a linear feature in the block, the energy should concentrate along the direction orthogonal to the feature orientation. As a consequence, one of $(\vec{\mu}_1, \vec{\mu}_2)$ should point in that direction and $\vec{\mu}_h(\theta_i)$ should not deviate significantly from it (one of $(\vec{\mu}_1, \vec{\mu}_2)$).

2.4 Locating the Features

To locate a feature contained in a single-feature block, both the feature's orientation and the center of mass gravity are to be calculated. In this work, we estimate local feature orientation by calculating the argument of the centroid vector $\vec{\mu}_i$, which is closer to $\vec{\mu}_h(\theta_i)$ because it points in the direction of the axis along which the energy cluster. Since we are seeking the feature's orientation in the spatial domain, $\pi/2$ is added to the orientation of the chosen centroid to reflect the fact that feature's orientation in spatial frequency domain and in spatial domain are perpendicular to each other.

As described in [5][10], positional information in spatial domain corresponds to the phase of the Fourier transform. Therefore, the position of the center of mass gravity of a local feature in the spatial domain can be estimated by averaging phase difference over all frequencies. Given the spectrum of a local linear feature [10] as

$$\hat{f}(u,v) = |\hat{f}(u,v)|\, e^{-j(ux+vy)} \tag{9}$$

where (x, y) is the centroid of the linear feature. Note that we are using $(\vec{u}, v)$ here to represent the frequency coordinates $\vec{\omega}$ in Equations (2) to (7). The auto-correlation of the image spectrum in both u and v directions respectively are

$$\rho_u = \frac{\sum_{(u,v)\in \Lambda_h(\theta_i)} \hat{f}(u,v)\hat{f}^*(u+u',v)}{\sum_{\vec{\omega}\in \Lambda_h(\theta_i)} |\hat{f}(u,v)|^2} \tag{10}$$

$$\rho_v = \frac{\sum_{(u,v)\in \Lambda_h(\theta_i)} \hat{f}(u,v)\hat{f}^*(u,v+v')}{\sum_{\vec{\omega}\in \Lambda_h(\theta_i)} |\hat{f}(u,v)|^2} \tag{11}$$

where u' and v' are the sampling intervals in u and v directions respectively. In this work u' and v' are equal to $\frac{2\pi}{N}$. Now, by substituting Equation (9) for both Equations (10) and (11), the estimate of the centroid (x, y) can be given as

$$x_0 = \frac{Arg(\rho_u)}{u'} \tag{12}$$

$$y_0 = \frac{Arg(\rho_v)}{v'} \tag{13}$$

2.5 Accuracy Check

Once the position of the centroid of the feature has been estimated, The accuracy of the estimated position should be checked. As in [10], the synthesized local spectrum is obtained by replacing x and y of equation (9) with x_0 and y_0 respectively:

$$\tilde{f}(u,v) = |\hat{f}(u,v)|\, e^{-j(ux_0+vy_0)} \tag{14}$$

The accuracy parameter γ is then calculated by correlating the real local spectrum $\hat{f}(u,v)$ with its synthesized counterpart $\tilde{f}(u,v)$:

$$\gamma = \left| \frac{\sum_{(u,v)\in\Lambda_h(\theta_1)} \hat{f}(u,v)\tilde{f}^*(u,v)}{\sum_{\bar{\omega}\in\Lambda_h(\theta_1)} |\tilde{f}^*(u,v)|^2} \right| \tag{15}$$

From the above equation, it is easy to see that $\gamma = 1$ if the estimated centroid (x_0, y_0) is exactly equal to the real centroid (x, y), otherwise, $\gamma < 1$. The closer to γ is, the more accurate the estimate is. If γ is higher than a given threshold, all the estimated parameters of the current block are deemed accurate enough and no further estimations at next resolution level are needed for it. If γ is less than the given threshold, despite the fact that Equation (8) is satisfied, all the estimated parameters are to be discarded and further estimation is needed at next resolution level.

The block diagram of the algorithm is given in Figure 2.

3. EXPERIMENTS

We attempt to detect linear features at three different resolutions with block sizes of 32×32, 16×16, and 8×8, respectively. Figure 3 shows the original image of Lena and the edges detected with $\alpha = 0.45$ and γ of different values (0.0 and 0.7 respectively). Without checking the accuracy of estimated centroid of

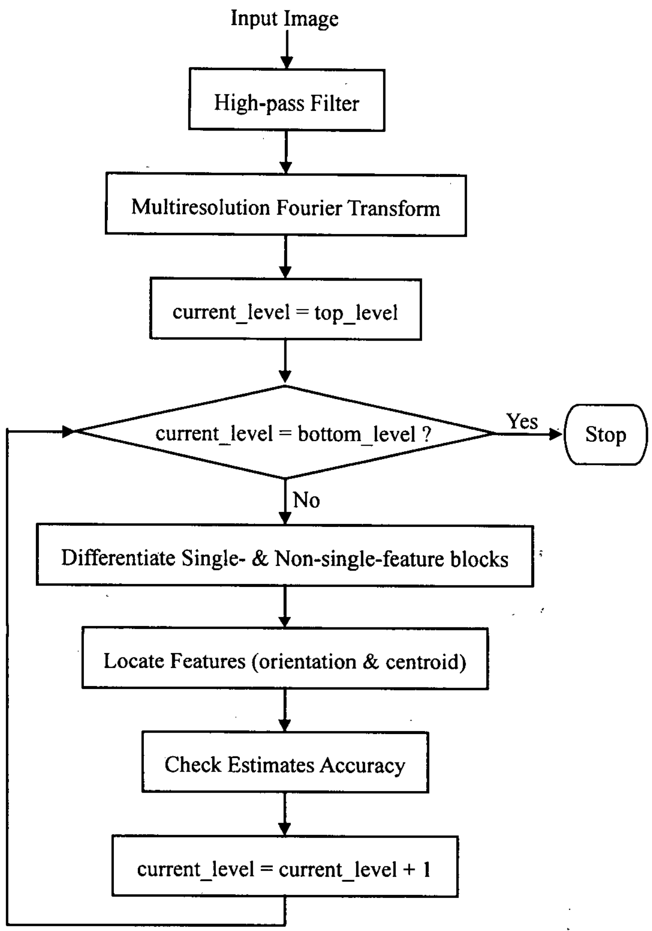
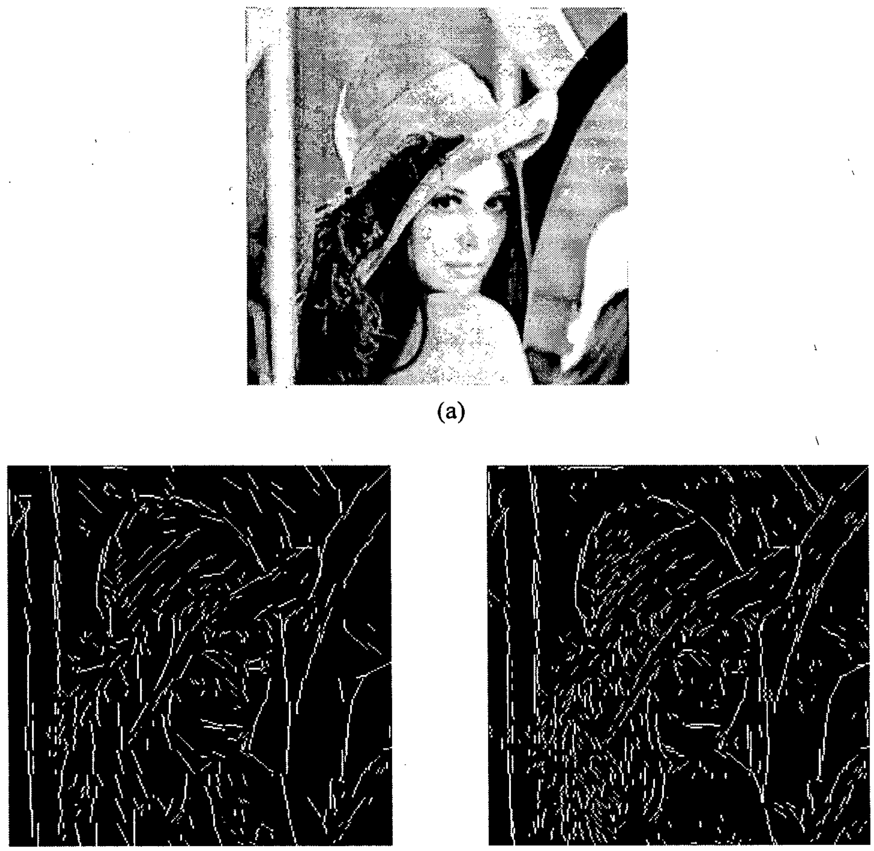


Figure 2. Block diagram of the algorithm.

the features by setting $\gamma = 0.0$, it is easy to see from Figure 3(b) that edges of large scale can be extracted with low threshold value of variance ratio $\alpha = 0.45$ without going to the next resolution level (smaller blocks). To get a smoother segmentation result, a higher value of γ is needed. The smoothness is apparent in Figure 3(c) with $\gamma = 0.7$.

Another way to get a smoother segmentation result with the same value of γ is to use higher value of α. Comparing Figure 3(b) with Figure 4(a) in which $\gamma = 0.0$, enhanced smoothness is apparent in the latter. Again the result can still be improved with higher value of γ. Figure 4(b) shows the result with $\alpha = 0.55$ and $\gamma = 0.7$. Figure 5 (a) and (b) show the segmentation result of Lena's picture with



(a)

(b) (c)

Figure 3. Edge detection results of Lena's picture with $\alpha = 0.45$. (a) orginal image of lena. (b) Edge map with $\gamma = 0.0$. (c) Edge map with $\gamma = 0.7$.

and $\gamma = 0.0$ and 0.7 respectively.

Note that because there is a 50% overlap between neighboring blocks, only the central *quarter* of each block is drawn in Figure 3 to 5. That is why the actual block sizes drawn are 16×16, 8×8, and 4×4 at the three different resolution levels, respectively.

4. CONCLUSIONS AND FUTURE WORKS

We have presented a novel edge detection algorithm based on the Multiresolution Fourier Transform (MFT) framework. The experiments show the features of different scales can be detected within the multiresolution framework