Image Edge Detection Using Hidden Markov Chain Model Based on the Non-decimated Wavelet

Renqi Zhang, Wanli Ouyang, Wai-Kuen Cham
Dept. of Electronic Engineering, The Chinese University of Hong Kong
Shatin, Hong Kong
E-mail: rqzhang@ee.cuhk.edu.hk

Abstract

Edge detection plays an important role in digital image processing. Based on the non-decimated wavelet which is shift invariant, in this paper, we develop a new edge detection technique using Hidden Markov Chain (HMC) model. With this proposed model (NWHMC), each wavelet coefficient contains a hidden state, herein, we adopt Laplacian model and Gaussian model to represent the information of the state "big" and the state "small". The model can be trained by EM algorithm, and then we employ Viterbi algorithm to reveal the hidden state of each coefficient according to MAP estimation. The detecting results of several images are provided to evaluate the algorithm. In addition, the algorithm can be applied to noisy images efficiently.

1. Introduction

Edges are one of the most important elements in image analysis and computer vision, because they play quite a significant role in many applications of image processing, in particular for machine vision. A lot of computer vision methods rely on edge detection as a pre-processing stage. However, no single edge detection algorithm can successfully discover edges for diverse images and no specific quantitative measure of the quality for edge detection is given at present.

Conventional edge detection mechanisms examine image pixels for abrupt changes by comparing pixels with their neighbors. This is often done by detecting the maximal value of gradient, such as Roberts, Prewitt, Sobel, Canny and so on, all of which are classical edge detectors. Alternatively, we can detect the zero-crossing points. Laplacian and MAR algorithms are based on the idea. Mallat established the edge detection technique in a multi-scale manner using the dyadic wavelet transform. Small-scaled transform filters are sensitive to edge signals but also prone to noise, while large-scaled filters are robust to noise but could filter out fine details.

Recently, many multi-scale transform tools based on statistics have been developed. As we know, transform coefficients exhibit statistical properties because of their intra-scale and inter-scale interdependence. In order to fuse multi-scale wavelet information, M.S Crouse developed the HMT model [1] based on the wavelet transform which pioneered a new research area on multi-scale statistical signal processing.

After that, many researchers applied and improved this HMT model, such as CHMM [2], LCHMM [3], HMT-2S [4] and so on. These HMT based models have been applied to various aspects in image analysis, including denoising, compression, fusion, segmentation, searching and so on. Edge detection algorithms based on this model have been developed [5]-[7]. In addition, new multi-scale tools, for example, Complex wavelet, Ridgelet, Contourlet, Curvelet were developed to tackle different problems of wavelet transform. In this paper, we propose a new edge detection algorithm based on the non-decimated wavelet which is shift-invariant and so able to locate edges with good precision.

Inspired by the non-decimated wavelet, we propose a novel edge detection technique NWHMC using Hidden Markov Chain (HMC) model that can fuse inter-scale dependence of wavelet coefficients. Experimental results on some natural and noisy images have demonstrated its effectiveness on edge detection.

2. Edge detection using HMC model based on non-decimated wavelet (NWHMC)

2.1. Statistical model for individual wavelet coefficients

Wavelet coefficients can be treated as random realizations, thus, they can be described by probabilistic models. Suppose they have two hidden states: the “small” one S=0 and the “big” one S=1. There exists a classical zero-mean, two-state Gaussian Mixture Model to represent every coefficient W. Therein, hidden state S=1 corresponds to a high-variance Gaussian PDF (Probability Density Function) and S=0 corresponds to a low-variance Gaussian PDF. Consequently, we model coefficients which are jointly Gaussian in this way:

$$f_W(w) = \sum_{m=0}^{1} f_{W|S}(w|S=m) f_S(S=m) \,. \tag{1}$$

However, the model conflicts with the compression property, which dictates the sparse distribution of wavelet coefficients. That is to say, only a few coefficients contain most energy and others contribute little to signal energy, as a result, PDF of wavelet coefficients should be highly concentrated around zero and heavy-tailed at both sides, all of which cannot be expressed using the previous model. The General Gaussian Distribution (GGD) model can represent peak around zero and long heavy tail approximately [8]:

$$f(y) = \frac{p}{2q\Gamma(\frac{1}{p})} \exp(-\left|\frac{y}{q}\right|^p) \tag{2}$$

GGD model has two parameters: standard deviation q and shape parameter p. Although GGD model can represent peak and heavy tail effectively, its parameter estimation is too complex. In paper [8], it proved that for “big” coefficients, the GGD is approximately a Laplacian distribution. As a result, we use individual Gaussian model for “small” coefficients and Laplacian model for “big” coefficients to simplify the GGD fit. They form a new simple zero-mean, two-state mixture model which can express wavelet coefficients effectively.

2.2. Hidden Markov Chain considering the inter-scale dependence

HMT model is a quad-tree model based on the down-sampling wavelet without shift invariance. In the paper, we propose a Hidden Markov Chain model based on the non-decimated wavelet transform which is shift invariant.

Consider an observation sequence $O = O_1, O_2 ... O_t$ and the hidden state sequence $q = q_1, q_2 ... q_t, 1 \le t \le T$, where T is the maximal wavelet decomposition scales. The parameters λ for the model are: (1) the initial state distribution π; (2) the number of states N, (3) a_{ij}, the conditional probability that q_{t+1} is in state s_j given q_t is in state s_i. (4) φ, the parameter of Laplacian model for "big" coefficients, (5) u and σ^2, the mean and variance of Gaussian model for "small" wavelet coefficients. Here we primarily focus on the case $N=2$ and $u = 0$. The observation symbol probability distribution b_i can be achieved through "big" and "small" states PDF separately as follows:

$$f(w|s=2) = \frac{1}{\sqrt{2}\varphi} \exp(-\frac{\sqrt{2}|w|}{\varphi}) \cdot \tag{3}$$

$$f(w|s=1) = \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{w^2}{2\sigma^2}) \cdot \tag{4}$$

2.3. Training parameters and Searching for hidden states

Training model parameters amounts to find the optimal parameters to maximize $P(O|\lambda)$ .We employ the EM algorithm [9] which adjust parameters λ to maximize $P(O|\lambda)$ given the observation sequence $O = O_1, O_2 ... O_t, 1 \le t \le T-1$ .We adopt the Forward-Backward algorithm to finish the E step and define the forward and backward variable as below:

$$\alpha_t(i) = P(O_1 O_2 ... O_t, q_t = s_i | \lambda) \cdot \tag{5}$$

$$\beta_t(i) = P(O_{t+1}, O_{t+2}, ..., O_T | q_t = s_i, \lambda) , \beta_T(i) = 1 \cdot \tag{6}$$

The state probability variables are:

$$\gamma_t(i) = P(q_t = s_i | O, \lambda) \cdot \tag{7}$$

$$\xi_t(i,j) = P(q_t = s_{i,} q_{t+1} = s_j | O, \lambda) \cdot \tag{8}$$

This can be expressed simply in terms of the forward-backward variables:

$$\gamma_t(i) = \frac{\alpha_t(i)\beta_t(i)}{P(O|\lambda)} = \frac{\alpha_t(i)\beta_t(i)}{\sum_{i=1}^{N} \alpha_t(i)\beta_t(i)} \cdot \tag{9}$$

$$\xi_t(i,j) = \frac{\alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)}{P(O|\lambda)}$$

$$= \frac{\alpha_t(i)a_{ij}b_j(O_{t+1})\beta_{t+1}(j)}{\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_t(i)a_{ij}b_j(O_{t+1})\beta_{t+1}(j)}. \tag{10}$$

M step is used to update parameters. Suppose there are K wavelet coefficients in total. Then $\gamma_{tk}(i)$ and $\xi_{tk}(i,j)$ represent the probability variables of the kth wavelet coefficient at the scale t.

Initial state probability:

$$\hat{\pi}_i = \frac{1}{K}\sum_{k=1}^{K}\gamma_{1k}(i). \tag{11}$$

State transition probability:

$$\hat{a}_{ij} = \frac{\sum_{k=1}^{K}\sum_{t=1}^{T-1}\xi_{tk}(i,j)}{\sum_{k=1}^{K}\sum_{t=1}^{T-1}\gamma_{tk}(i)}. \tag{12}$$

Laplacian model parameter for "big" state:

$$\hat{\varphi} = \sqrt{2}\,\frac{\sum_{k=1}^{K}\sum_{t=1}^{T}\gamma_{tk}(i)|O_{tk}|}{\sum_{k=1}^{K}\sum_{t=1}^{T}\gamma_{tk}(i)}. \tag{13}$$

Gaussian model parameter for "small" state:

$$\hat{\sigma}^2 = \frac{\sum_{k=1}^{K}\sum_{t=1}^{T}\gamma_{tk}(i)(O_{tk})^2}{\sum_{k=1}^{K}\sum_{t=1}^{T}\gamma_{tk}(i)}. \tag{14}$$

Now we attempt to adapt these model parameters to fit the observed wavelet coefficients. Viterbi algorithm [9] is employed to obtain the hidden states:

We define the quantity:

$$\delta_t(i) = \max_{q_1,q_2,\ldots q_{t-1}} p(q_1,q_2,\ldots q_t, q_t = s_i, O_1,O_2,\ldots O_t|\lambda) \tag{15}$$

as the highest probability along a single path $O_1,O_2 \ldots O_t$ at scale t. Then the complete procedure is as below:

1. Initialization:

$$\delta_1(i) = \pi_i b_i(O_1), 1 \le i \le N.$$
$$\psi_1(i) = 0, 1 \le i \le N. \tag{16}$$

2. Recursion:

$$\delta_t(j) = \max_{1\le i\le N}[\delta_{t-1}(i)a_{ij}]b_j(O_t), 2 \le t \le T. \tag{17}$$

$$\psi_t(j) = \arg\max_{1\le i\le N}[\delta_{t-1}(i)a_{ij}], 2 \le t \le T\cdot \tag{18}$$

3. Termination:

$$p^* = \max_{1\le i\le N}[\delta_T(i)]\cdot \tag{19}$$

$$q_T{}^* = \arg\max_{1\le i\le N}[\delta_T(i)]\cdot \tag{20}$$

4. State sequence backtracking:

$$q_t^* = \psi_{t+1}(q_{t+1}{}^*), t = T-1, T-2, ..., 1\cdot \tag{21}$$

The optimal hidden sequence obtained will then be used to find the edge maps.

3. Experimental results and analysis

3.1. Realization of our algorithm and parameter selection

In our experiments, natural images as shown in Fig 1 were used to test our algorithm. We use Haar wavelet to conduct the non-decimated wavelet transform of three scales. The same pixel location of every scale forms a Hidden Markov Chain, the length of which is 3. We set the initial state transition probability and state probability as below:

$$A = \begin{bmatrix} 0.95 & 0.05 \\ 0.2 & 0.8 \end{bmatrix}, \quad \pi_1(1) = \pi_1(2) = 0.5\,. \tag{22}$$

The initial parameters of Laplacian model and Gaussian model influence results a lot, because improper selection may lead to local optimal value. However, selection of these parameters is still a difficult task for the training.

For the searched hidden states using NWHMC at the two high frequency subbands: horizontal(LH) and vertical (HL), we adopt Boolean operation “and” to combine their information and get the final edge detection results.

3.2. Comparisons with Canny algorithm and WD-VHMT

Fig. 2 shows results using the proposed algorithm and they are compared with experiments using Canny detector and multi-scale WD-VHMT technique [6]. The multi-scale WD-VHMT technique is based on the down-sampled wavelet transform which is not shift invariant and uses HMT to model inter-scale dependence of wavelet coefficients. However, the proposed multi-scale algorithm adopts the non-decimated

Fig 1. **The three test images cameraman, house and peppers**

wavelet which is shift invariant and HMC model is designed to capture interdependence across different scales. It appears that edges detected using the two multi-scale methods are thicker than the one obtained from Canny, because there is an extra operation called “non-maximum suppression” to thin detected edges in Canny algorithm. However, the two multi-scale methods can capture the main objects, and remove redundant details. Main objects are more likely to be highlighted. This is because they model the statistics of both edges and non-edges and fuse information from different scales, while Canny algorithm only takes single scale information.

Comparing the detected results of WD-VHMT and our proposed NWHMC method, we can say that edge maps of our method have less isolated points that are useless. In addition, our method is more accurate at some edge points. These advantages of the proposed method are mainly because the adopted non-decimated wavelet transform. It has more useful information than the down-sampled wavelet and it is a shift-invariant transform.

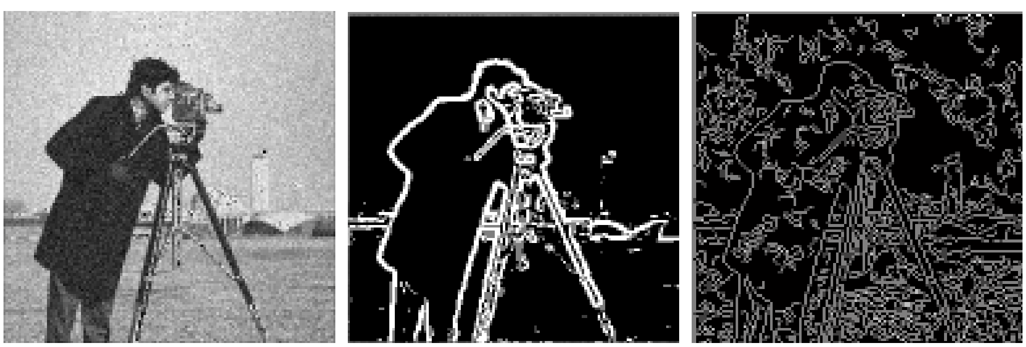
(*a*) **(b)** **(c)**

(d) (e) (f)

Fig. 2. Edge detection results (a)(d) edge maps from Canny (b)(e) edge maps from WD-VHMT (c)(f) edge maps from NWHMC

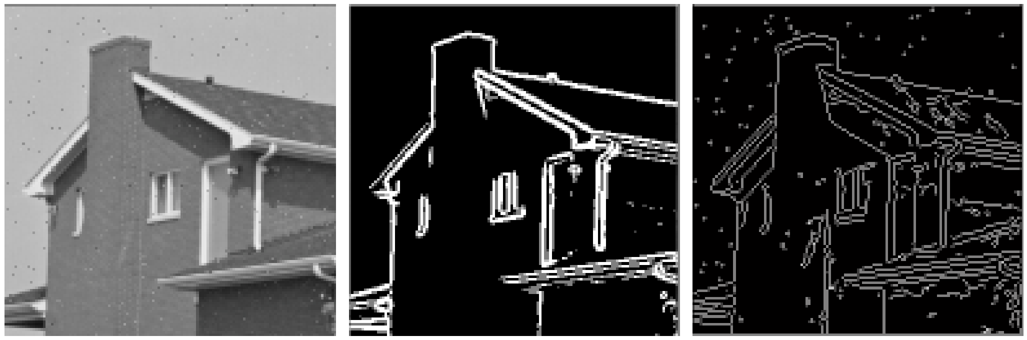
3.3. Experiments of noisy images using our algorithm

Fig. 3 and Fig. 4 show the results of the edge detectors on images corrupted by random noise and salt & pepper noise respectively. Noise disappears with the increase of wavelet scales but energy of noise overwhelms those of coefficients at the finest scale. In the proposed method, we use the information from the second finest scale to the fourth. The results reveal that the proposed method can detect edges with few errors, whilst Canny detector produces many false edges.



(a) (b) (c)

Fig. 3. Edge detection results (a) random noisy image (b)edge map with NWHMC (c) edge map with Canny algorithm



(a) (b) (c)

Fig. 4. Edge detection results (a)salt&pepper noisy image (b)edge map with NWHMC (c)edge map with Canny algorithm

4. Conclusion

The multi-resolution property of the wavelet transform has led to its efficiency in singularity detection as a multi-scale tool. In this paper, we propose a new edge detection algorithm NWHMC using HMC model based on the shift invariant non-decimated wavelet transform. Experimental results show that the proposed method is an efficient and accurate edge detecting tool for clean and noisy images.

5. References

[1] M. S. Crouse, R. D. Nowak, and R. G. Baraniuk. Wavelet- based statistical signal processing using hidden Markov models. IEEE Trans On Signal Processing. Vol 46(4). 886- 902,1998.

[2] M. S. Crouse, and R. G. Baraniuk. Contextual hidden Markov models for wavelet-domain signal processing. In:Proceedings of the 31st Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA. 95~100,1997.

[3] Fan G. L , and Xia X. G. Image denoising using local contextual hidden Markov model in the wavelet domain. IEEE Signal Processing Letters. Vol.8(5).125~128,2001.

[4] Fan G. L, and Xia X. G. Improved hidden Markov models in the wavelet-domain. IEEE Transactions On Signal Processing. Vol.49(1).115~120,2001.

[5] Sun J. X, Chen Y. Z, and Gu D. B. Multi-scale edge detection based on representation of edge characterized by hidden Markov model. Chinese Journal of Computers.Vol. 26(4). 497-501, 2003.

[6] Sun J. X, Chen Y. Z, and Zhang S. A multiscale edge detection algorithm based on wavelet domain vector hidden Markov tree model. Pattern Recognition. Vol.37(7).1315-1324, 2004.

[7] Sun Q, Hou B, and Jiao L.C. SAR Image Edge Detection Based on Contourlet- Domain Hidden Markov Models. Proc. SPIE. Vol.10(6044).218-225, 2005.

[8] P. Moulin and J. Liu. Image denoising based on scale-space mixture modeling of wavelet coefficients. In: Proc. IEEE Int. Con on Image Proc, Kobe, Japan. 1999.

[9]L. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. Proc. IEEE.Vol. 2(77).257–28, 1989.

Authors

Renqi Zhang received the B.S. degree in Electronic Engineering from Xidian University, Xi'an, China, in 2006. He is now pursuing Ph.D in the Department of Electronic Engineering, The Chinese University of Hong Kong.

Wanli Ouyang received the B.S. degree in computer science from Xiangtan University, Hunan, China, in 2003. He received the M.S. degree with the College of Computer Science and Technology, Beijing University of Technology, Beijing, China. He is now pursuing Ph.D in the Department of Electronic Engineering, The Chinese University of Hong Kong.

Wai-Kuen Cham graduated from The Chinese University of Hong Kong in 1979 in Electronics. He received his M.Sc. and Ph.D. degrees from Loughborough University of Technology, U.K., in 1980 and 1983 respectively. Since May 1985, he has been with the Department of Electronic Engineering, The Chinese University of Hong Kong. Prof. Cham is a Chartered Engineer and a senior member of IEEE.