IMAGE FUSION ALGORITHM BASED ON NEIGHBORS AND COUSINS INFORMATION IN NONSUBSAMPLED CONTOURLET TRANSFORM DOMAIN

XIAO-BO QU[1] , GUO-FU XIE[2, 3] , JING-WEN YAN[4, 1] , ZI-QIAN ZHU[5] , BEN-GANG CHEN[5]

[1]Dept. of Communication Engineering, Xiamen University, Xiamen 361005, China
[2]Key Laboratory of Computer Science (Institute of Software, The Chinese Academy of Sciences), Beijing 100080, China
[3]Dept. of Software Engineering, Xiamen University, Xiamen 361005, China
[4]Key Laboratory of Intelligent Manufacturing Technology, Ministry of Education (at Shantou University), Shantou 515063, China
[5]Research Institute of Chinese Radar Electronic Equipment, Wuxi 214063, China
E-MAIL: qxb_xmu@yahoo.com.cn, guofu85@gmail.com, yjwen@xmu.edu.cn

Abstract

Nonsubsampled contourlet transform (NSCT) provides flexible multiresolution, anisotropy and directional expansion for images. Compared with the foremost contourlet transform, it is shift-invariant and can overcome the pseudo-Gibbs phenomena around singularities. In addition, coefficients of NSCT are dependent on their neighborhood coefficients in the local window and cousin coefficients in directional subbands. In this paper, region energy and cousin correlation are defined to represent the neighbors and cousins information, respectively. Salience measure, as the combination of region energy and cousin correlation, is defined to obtain fused coefficients in the high-frequency NSCT domain. First, source images are decomposed into subimages via NSCT. Secondly, salience measure is computed. Thirdly, salience measure-maximum-based rule and average rule are employed to obtain high-frequency and low-frequency coefficients, respectively. Finally, fused image is reconstructed by inverse NSCT. Experimental results show that the proposed algorithm outperforms wavelet-based fusion algorithms and contourlet transform-based fusion algorithms.

Keywords: Image fusion; contourlet transform; wavelet transform; sparse representation

1. Introduction

Image fusion is the combination of two or more different images to form a new image by using a certain algorithm [1]. The combination of sensory data from multiple sensors can provide more reliable and accurate information [2]. It forms a rapidly developing area of research in remote sensing and computer vision [1]-[4].

Most of fusion approaches were based on combining the multiscale decompositions (MSD's) of the source images. MSD-based fusion schemes provide much better performance than the simple methods studied previously [5].Due to joint information representation at the spatial-spectral domain, the wavelet transform becomes the most popular approximation in image fusion. However, wavelet will not "see" the smoothness along the contours and separable wavelets can capture only limited directional information [6].

Contourlet transform was recently pioneered by Minh N. Do and Martin Vetterli [6] .It is a "true" two-dimensional transform that can capture the intrinsic geometrical structure, which is key in visual information. Compared with wavelet, contourlet provides different and flexible number of directions at each scale. It has been successfully employed in image enhancement, denoising and fusion [7][8][9].Unfortunately, due to downsamplers and upsamplers presented in both the laplacian pyramid and the directional filter banks(DFB), the foremost contourlet transform is not shift-invariant, which causes pseudo-Gibbs phenomena around singularities. Some shift-invariant contourlet transform have been investigated in [8][10]. Nonsubsampled contourlet transform (NSCT) [10] is a fully shift-invariant form of the contourlet transform. It leads to better frequency selectivity and regularity than contourlet transform. Further more, a directional multiscale modeling of images using the contourlet transform has been studied in [11].It demonstrates that coefficients of NSCT are strongly dependent on their neighborhood and cousin coefficients.

In this paper, region energy of neighborhood coefficients and correlation of cousin coefficients are defined to represent the dependency. Based on the two definitions, salience measure is investigated as a measurement to combine the coefficients of source images. The efficiency of salience measure in the fusion process will be presented in the following paragraphs.

2. Principle of NSCT

In the foremost contourlet transform [6] downsamplers and upsamplers are presented in both the laplacian pyramid and the DFB. Thus, it is not shift-invariant, which causes pseudo-Gibbs phenomena around singularities. NSCT is an improved form of contourlet transform. It is motivated to be employed in some applications, in which redundancy is not a major issue, i.e. image fusion.

In contrast with contourlet transform, nonsubsampled pyramid structure and nonsubsampled directional filter banks are employed in NSCT. The nonsubsampled pyramid structure is achieved by using two-channel nonsubsampled 2-D filter banks. The DFB is achieved by switching off the downsamplers/upsamplers in each two-channel filter bank in the DFB tree structure and upsampling the filters accordingly. As a result, NSCT is shift-invariant and leads to better frequency selectivity and regularity than contourlet transform.Fig.1 shows the decomposition framework of contourlet transform and NSCT.
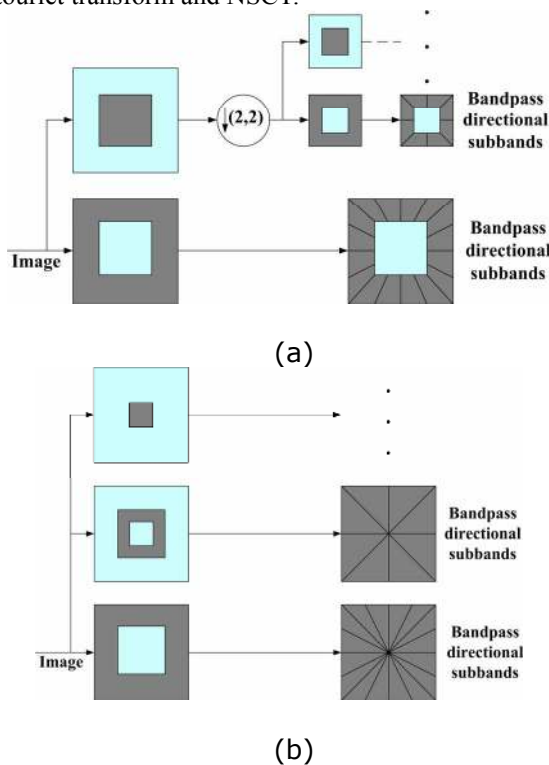

(a)

(b)

Figure 1. Decomposition framework of contourlet transform and NSCT.

3. Neighbors and cousins information in NSCT domain

As a hot topic in the sparse representation of images, coefficients characteristics of contourlet also have been studied in [11]. There are three relationships in contourlet coefficients, which are shown in Fig.5.
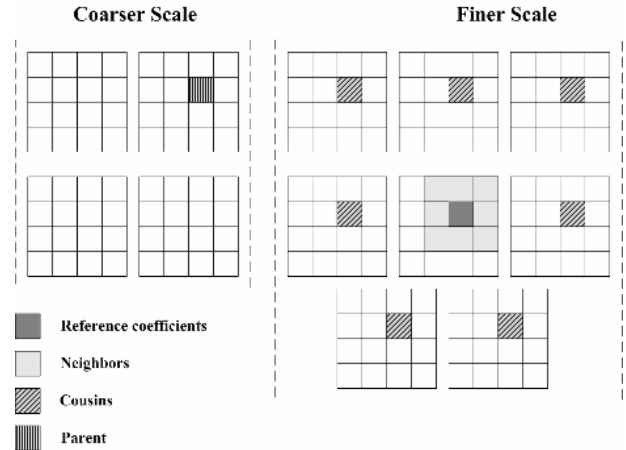

Figure 2. NSCT coefficient relationships.

The reference coefficient has eight *neighbors (NX)* in the same subband, *parent (PX)* at the same spatial location in the immediately coarser scale and *cousins (CX)* at the same scale and spatial location but in directional subbands.

In [11], mutual information is utilized as a measure of dependencies to study the joint statistics of contourlet coefficients. Suppose $I(X;Y)$ stands for the mutual information between two random variables *X* and *Y*. Estimation results in [11]show that at fine scales $I(X;NX)$ is higher than $I(X;CX)$, which is higher than $I(X;PX)$ .It indicates that the eight neighbor coefficients contain the most information about the coefficients, less information is contained in cousins and the least information is contained in the parent coefficients. Inspired by the estimation results, salience measure, based on region energy of neighborhood coefficients and correlation of cousin coefficients, is defined to combine the coefficients of source images in the fusion process in section 4.

4. NSCT-based Fusion Algorithm

In this paper, image decomposition is performed by the NSCT. We hope that predominance of NSCT, which are shift-invariant, multiresolution, localization, directionality, and anisotropy, will be more suitable for image fusion and other image processing, i.e. target recognition, object detection, etc. In the fusion process, both neighborhood coefficients and cousin coefficients information are utilized in the salience measure.

4.1. Fusion of low-frequency coefficients

Considering the images' approximate information is constructed by the low-frequency coefficients, average rule is adopted for low-frequency coefficients. Suppose $B_F(x, y)$ is the fused low-frequency coefficients, then

$$B_F(x, y) = \frac{B_1(x, y) + B_2(x, y)}{2} \tag{1}$$

where $B_1(x, y)$ and $B_2(x, y)$ denote the low-frequency coefficients of source images.

4.2. Fusion of high-frequency coefficients

High-frequency coefficients always contain edge and texture features. In order to make full use of information in the neighborhood and cousin coefficients in the NSCT domain, a salience measure, as a combination of region energy of NSCT coefficients and correlation of the cousin coefficients, is proposed for the first time.

We define region energy by computing the sum of the coefficients' square in the local window. Suppose $C_l^k(x, y)$ is the high-frequency NSCT coefficients, whose location is *(x,y)* in the subband of *k*-th direction at *l*-th decomposition scale. The region energy is defined as follows:

$$E_l^k(x, y) = \sum_{m,n \in S_{M \times N}} \left(C_l^k(x+m, y+n) \right)^2 \tag{2}$$

where $S_{M \times N}$ denotes the regional window and its size is $M \times N$ (typically 3×3). Region energy, rather than single pixel value, will be more reasonable to extract features of source images by utilizing neighbors' information. Large region energy means important image information. Note that size of region energy map is equal to size of each subband.

We define correlation of cousins, which stands for the relationship between multi-directional subbands at the same scale. Suppose $C_l^{k_p}$ and $C_l^{k_q}$ are two NSCT subbands of k_p-th and k_q-th direction at *l*-th decomposition scale. In the *l*-th decomposition scale, if number of multi-directional subbands are Q, then $k_p, k_q \in \{1, 2, \cdots, Q\}$. The correlation of cousins between k_p-th and k_q-th subbands is defined as follows:

$$\rho\left(C_l^{k_p}, C_l^{k_q}\right) = \frac{\sum_{i=0}^{I-1}\sum_{j=0}^{J-1}\left[C_l^{k_p}(i,j) - \overline{C}_l^{k_p}\right]\left[C_l^{k_q}(i,j) - \overline{C}_l^{k_q}\right]}{\sqrt{\sum_{i=0}^{I-1}\sum_{j=0}^{J-1}\left[C_l^{k_p}(i,j) - \overline{C}_l^{k_p}\right]^2 \sum_{i=0}^{I-1}\sum_{j=0}^{J-1}\left[C_l^{k_q}(i,j) - \overline{C}_l^{k_q}\right]^2}} \tag{3}$$

where $\overline{C}_l^{k_p}$ and $\overline{C}_l^{k_q}$ are the average value of $C_l^{k_p}$ and $C_l^{k_q}$, respectively. *I* and *J* denote the subband size is *I×J*.

Referring to cousins in one directional subband, a normalized correlation weight can be defined as follows:

$$W_{C_l^{k_p}}\left(C_l^{k_q}\right) = \frac{\rho\left(C_l^{k_p}, C_l^{k_q}\right)}{\sum_{k_q \in \{1,2,\cdots,Q\}} \rho\left(C_l^{k_p}, C_l^{k_q}\right)} \tag{4}$$

The normalized correlation weight indicates the dependency of cousins in multi-directional subbands. Large $W_{C_l^{k_p}}\left(C_l^{k_q}\right)$ means high dependency.

Based on the region energy of neighbors and the normalized correlation weight of cousins, salience measure can be defined as follows:

$$S_l^{k_p}(x, y) = \sum_{q \in \{1,2,\cdots,Q\}} E_l^{k_q}(x, y) W_{C_l^{k_p}}\left(C_l^{k_q}\right) \tag{5}$$

It can be seen that neighbors and cousins information are both utilized in the salience measure.

The fusion rule in high-frequency domain is salience measure-maximum-based rule. If coefficients' salience measure of one source image is larger than that of another source image, then the coefficients corresponding to maximum salience measure will be selected as the coefficients of the fused image. Suppose $C_{F, l}^{\ k}(x, y)$ is the fused high-frequency coefficients, the high-frequency fusion rule can be described as follows:

$$C_{F, l}^{\ k_p}(x, y) = \begin{cases} C_{1, l}^{\ k_p}(x, y), if: S_{1, l}^{\ k_p}(x, y) \geq S_{2, l}^{\ k_p}(x, y) \\ C_{2, l}^{\ k_p}(x, y), if: S_{1, l}^{\ k_p}(x, y) < S_{2, l}^{\ k_p}(x, y) \end{cases} \tag{6}$$

where $C_{1, l}^{\ k_p}$ and $C_{2, l}^{\ k_p}$ denote the high-frequency coefficients of source images.

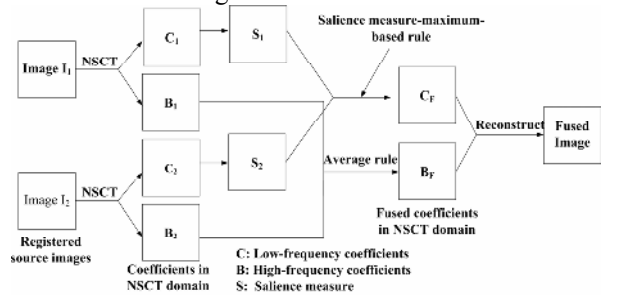The framework of the proposed NSCT-based fusion algorithm is shown in Fig.3.

Figure 3. Framework of the NSCT-based fusion algorithm

5. Experimental Results and Evaluation

The performance evaluation criteria of image fusion are still a hot topic [5]. Besides visual observation, mutual information (MI) [12] is used as objective performance evaluation criterion in this paper. Here, the new algorithm is compared with four typical ones including discrete wavelet transform with DBSS (2, 2) fusion algorithm, shift invariant discrete wavelet transform (SIDWT) with 'haar' wavelet

fusion algorithm[4][13], region-energy-based contourlet transform algorithm [9] and pixel-maximum-based NSCT algorithm.

In the experiments, two multi-focus images are used to be source images as shown in Fig.4(a)(b).Source images are in 640×480 size and 256 levels in gray value. In DBSS and SIDWT algorithms, the decomposition level is 3.Average rule and pixel-maximum-based rule are adopted in the low-frequency and high-frequency domain of wavelet transform domain, respectively. In the pixel-maximum-based NSCT algorithm, average rule and pixel-maximum rule are adopted in the low-frequency and high-frequency domain of NSCT, respectively. The fusion results of different algorithms are shown in Fig.4.The zoomed head of the student in Fig.4 (b)-(g) are shown in Fig.5.

Obviously, the fused images of three shift-invariant methods, SIDWT, NSCT and the proposed algorithm are clearer than the DBSS and SCT fused results. It proves that shift-invariant methods can overcome the pseudo-Gibbs phenomena successfully and improve the quality of the fused image. Focusing on the student head, we can find that the mistiness around the head in Fig.5 (f) is the slightest. It indicates that fused image of the proposed algorithm is clearest than other fused results and our algorithm extracts the image features the best.

(a)

(b)

(c)

(d)

(e)

(f)

(g)

Figure.4 The multi-focus source images and fused results, (a) source image 1: the clock is clear, (b) source image 2: the student is clear, (c) fused image with DBSS (2, 2) wavelet transform, (d) fused image with SIDWT, (e) fused image with contourlet transform, (f) fused image with pixel-maximum-based NSCT, (g) fused image of the proposed algorithm.
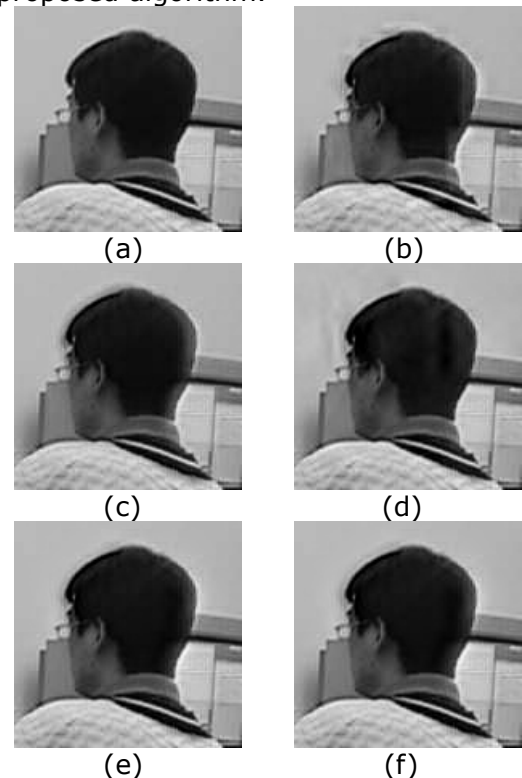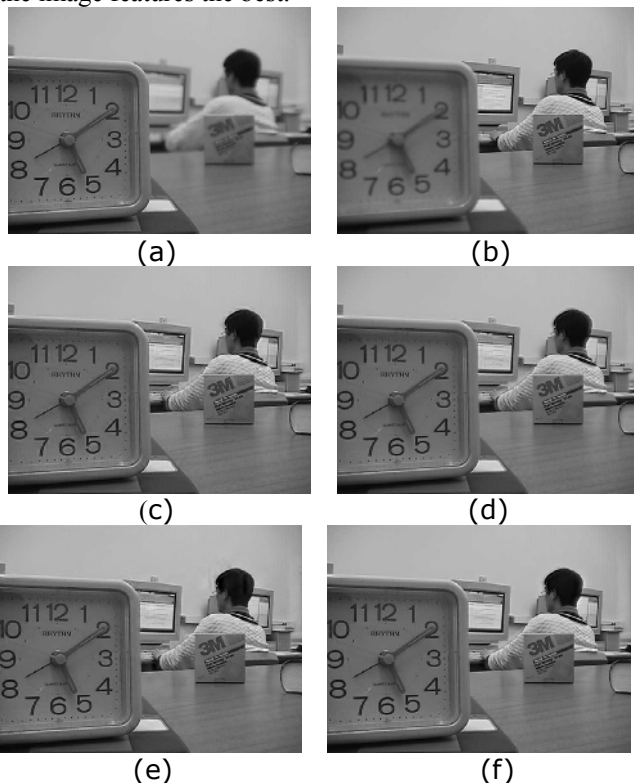
(a)

(b)

(c)

(d)

(e)

(f)

Figure.5 Zoomed head from fused images of different algorithms,(a)-(f) are gotten from Fig.4(b)-(g).(a) Clear head of source image,(b) DBSS (2,2) algorithm,(c)SIDWT algorithm,(d)Region-energy-based contourlet algorithm,(e)Pixel-maximum-based algorithm NSCT,(f) The proposed algorithm

In Fig.6, the fused results of pixel-maximum-based NSCT and the proposed algorithm are compared by subtracting the clear head of source image in Fig.4 (b).It is obvious that the remainder texture of pixel-maximum-based NSCT is more than that of the proposed algorithm, which indicates that less features are preserved in

pixel-maximum-based algorithm result. It demonstrates that the proposed algorithm can efficiently extract the features of source images. Salience measure based on coefficient characteristics in NSCT domain is reasonable and effective in fusion process.

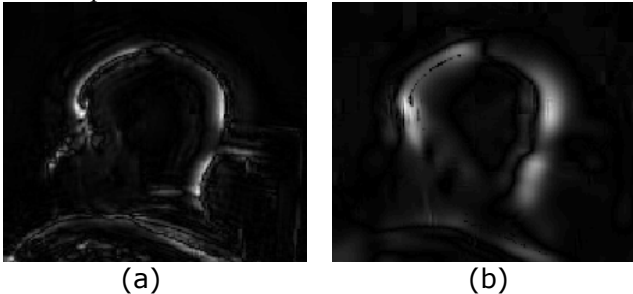(a)   (b)

Figure.6 Comparison between pixel-maximum-based NSCT algorithm and the proposed algorithm,(a)difference of head between Fig.4(f) and Fig.4(b),which is the minus result of pixel-maximum-based NSCT, (b)difference of head between Fig.4(g) and Fig.4(b),which is the minus result of proposed algorithm.

Besides visual observation, mutual information of different algorithms is compared in Tab.1.MI of the proposed algorithm is the largest. It proves that fused image of the proposed algorithm is strongly correlated with the source images and more image features are preserved in the fusion process. When all is said and done, our proposed algorithm outperforms other typical algorithms, no matter in visual observation and objective evaluation criterion.

Tab.1 Comparison on MI of different algorithms

Algorithms	DBSS	SIDWT	RECT	PMN	SMN
MI	6.6037	7.0629	6.6412	7.0695	7.1982

DBSS:discrete wavelet transform with DBSS (2, 2) fusion algorithm

SIDWT: SIDWT with 'haar' wavelet fusion algorithm

RECT:region-energy-based contourlet transform algorithm

PMN:pixel-maximum-based NSCT algorithm

SMN:salience measure-maximum-based NSCT algorithm, which is proposed in the paper.

6. Conclusion and Discussion

In this paper, we present a salience measure-maximum-based NSCT fusion algorithm. Experimental results demonstrate that the proposed fusion algorithm outperforms wavelet-based fusion algorithm, contourlet transform-based fusion algorithms and pixel-maximum-based NSCT algorithm, which indicates that the combination of neighbors and cousins information of NSCT, which is investigated by the salience measure in this paper for the first time, is reasonable and effective. As one application of NSCT, it is proven that NSCT is shift-invariant and can overcome the pseudo-Gibbs phenomena around singularities.

However, contourlet transform and NSCT is developed recently, many works are needed to explore. From our experiments, some researches could extend as follows:
(1) Besides salience measure in this paper, is there other method to measure the coefficients' dependency on their neighbors and cousins? Or is there a good way to combine the neighbors and cousins information, even parent information?
(2) Other applications of contourlet transform could be investigated, especially in those fields where wavelet transform has been employed successfully.

Acknowledgement

References

[1] David L H. An introduction to multisensor data fusion, Proc. of the IEEE, Vol.85, No.1, pp.6–23, Jan.1997.

[2] Luo R.C, Chih-Chen Yih, Kuo Lan Su. Multisensor fusion and integration: approaches, applications, and future research directions. IEEE Sensors Journal, Vol.2, No.2, pp.107–119, April. 2002.

[3] Constantinos S.P. Pattichis M.S. Micheli-Tzanakou E. Medical imaging fusion applications: An overview. 35th Asilomar Conference on Signals, Systems and Computers, pp. 1263–1267, Nov.2001.

[4] Sadjadi F. Comparative Image Fusion Analysis.2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol.3, pp.20-26, Jun.2005.

[5] Zhong Zhang, Blum R.S.A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application, Proceedings of the IEEE, Vol.87, No.8, pp.1315–1326, Aug.1999.

[6] MN Do, M Vetterli. The contourlet transform: an efficient directional multiresolution image representation, IEEE Transactions on Image Processing, Vol.14, No.12, pp.2091-2106, Dec.2005.

[7] Feng Peng, Pan Yingjun, Wei Biao,et al. Enhancing retinal image by the Contourlet transform. Pattern Recognition Letters, Vol.28, No.4. pp. 516-522, Mar.2007.

[8] Eslami Ramin, Radha Hayder. Translation-invariant contourlet transform and its application to image denoising, IEEE Transactions on Image Processing, Vol.15, No.11.pp. 3362-3374, Nov.2006.

[9] Miao Qiguang, Wang Baoshu.A Novel Image Fusion Method Using Contourlet Transform, 2006

International Conference on Communications, Circuits and Systems Proceedings,Guilin,Vol.1,pp.548-552,June.2006

[10] da Cunha Arthur, Zhou Jianping, Do Minh N. The nonsubsampled contourlet transform: Theory, design, and applications, IEEE Transactions on Image Processing, Vol.15, NO.10. pp. 3089-3101, Oct.2006.

[11] Po, D DY. Do, M N. Directional multiscale modeling of images using the contourlet transform, IEEE

Transactions on Image Processing, Vol.15, NO.6, pp. 1610-1620, Jan.2006.

[12] Guihong Qu, Dali Zhang, Pingfan Yan. Information measure for performance of image fusion. Electronics Letters,Vol.38,No.7,pp.313 − 315,Mar.2002

[13] Oliver Rockinger. Image sequence fusion using a shift-invariant wavelet transform, Proceedings of 1997 Image Processing Conference, Santa Barbara. pp. 288-291, Oct.1997.