UACEE International Journal of Advancements in Electronics and Electrical Engineering
Volume 1: Issue 1

Robust Pre-processing and Post-processing Methods for 2D Gel Electrophoresis Images using Non-Separable Quincunx Wavelet

Ratnesh Singh Sengar, Ashutosh Kumar Upadhyay, Manjit Singh
Division of Remote Handling & Robotics,
Bhabha Atomic Research Centre, Mumbai, India-400085
E-mail: ratneshss@yahoo.com
Tel/Fax: +91-22-2559-2139/+91-22-25505363

Vikram M. Gadre
Department of Electrical Engineering,
Indian Institute of Technolgy Bombay,
Mumbai, India-400076
E-mail: vmgadre@ee.iitb.ac.in

Abstract—**One of the most important tasks performed in proteomics is correct analysis and interpretation of 2D electrophoresis gel images. The nonlinearities in the gel formation and image acquisition process leads to distortions, overlapped protein spots, saturated spots, faint spots, nonlinear intensity and uneven background in the gel images. As a result, pre-processing and post processing steps are very challenging tasks for precise protein identification. The pre-processing step includes the task of removing additive or multiplicative noises and edge enhancement. The solutions available in the literature have failed to give satisfactory results. Non-separable wavelet processing methods, due to their inherent ability to be able to better represent directional information, seem to be promising for the processing of 2D gel images. In this paper, we explore approaches based on quincunx non-separable filter-banks for image pre-processing and post-processing. A novel method for edge preserved de-noising of gel images has been presented which also removes multiplicative noises. Our experiments show that non-separable wavelets are well suited for removing multiplicative noise. Our multi-scale post-processing method effectively removes the present streaks in the gel images.**

I. INTRODUCTION

Two dimensional electrophoresis is an important technique for analyzing protein expression and for enhancing data quality in the field of proteomics. Proteomics is the field that studies a multi-protein system, focusing on the interplay of multiple proteins as functional components in a biological system. By this technique a very large number of proteins can easily and simultaneously be separated, identified and characterized. This is important for understanding protein function and thus enables the development of new and more effective drugs [1] [2].

The first step in a typical proteomics analysis workflow is proteins separation, followed by quantification and differential expression analysis. Despite its limitations, 2D gel electrophoresis (2DGE) remains the most widely used protein separation method. Using 2DGE, individual proteins in a mixture are resolved in the first gel dimension according to their molecular weight and in the second dimension according to their isoelectric point [1] [4].

The objective is to extract the protein spots in gel images from the uneven background which has sharp edges, e.g. lines, artifacts and streaks in some area. Many authors report the experience typified in the following description: Due to some technical problems such as the system nonlinearities in gel formation and image acquisition, inevitably there appear overlapped protein spots, saturated spots, faint spots, nonlinear intensity and narrow lines on the gels which make the task more difficult (as shown in Fig-1) [3][4].

The noise suppression methods used in commercially available image analysis software packages are based on spatial filtering. Despite their simplicity, these filters tend to severely distort spot edges and alter the intensity values of spot pixels. This is not surprising since 2DGE images are typical examples of signals with rapidly varying local character, due to the large and unstructured variations in spot intensities and size. It is impossible to distinguish 'signal' from 'noise' in the space or frequency domain alone. It is therefore appropriate to use a joint domain. This paper deals with this issue and presents a method for de-noising gel images which also preserves and enhances the edges. Both additive and multiplicative noises have been removed.

For segmentation of gel images, several methods [5][6][7] have been referred to in the literature, they all can be seen as variants of the watershed method [8] and their

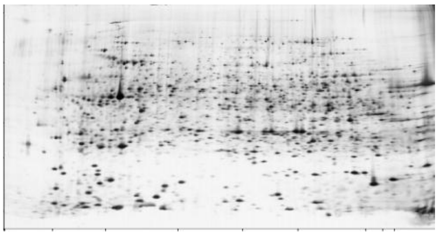


Fig. 1 Typical 2D Gel Image

results depend very much upon the sensitivity of parameter selections and post-processing thereafter. In this paper we use the segmentation method presented in [15] and then present our post processing method for streak removal.

We employ a non separable quincunx wavelet transform that outperforms the available methods of de-noising and segmentation. Even though this is a non-separable transform, the computational complexity is not significantly greater while maintaining a more general form of multiresolution character. Our investigation demonstrates that the high frequency components from this non-separable wavelet can capture more singular information than traditional wavelets. Therefore, it motivates us to apply an approach based on the non-separable wavelet to extract the features of 2D Gel electrophoresis images.

II. NON-SEPARABLE WAVELET QUINCUNX

In the context of wavelets and filter-banks, non-separable systems have non-rectangular frequency supports for the subbands in subbands, thus resulting in better frequency selectivity.

The simplest decomposition of that type, known as the Quincunx transform (QT), uses non-separable and non-oriented filters, followed by the non-separable sampling represented by the matrix Dq in (1) below. The Quincunx lowpass and highpass filters are often chosen to have diamond shaped frequency supports, as shown in Fig-2. With these diamond shaped frequency responses, the lowpass filter can preserve the high frequencies in the vertical and horizontal directions, which is a good match to the human visual system since the visual sensitivity is higher to changes in these two directions than the other directions. Due to this, Quincunx filter banks are particularly important in image processing applications. Since the det Dq=2, this transform is performed with a two-channel filter bank. At each level, the input image is decomposed with the multiresolution scale factor 2, resulting in one low-resolution subimage and one non-oriented wavelet subimage.

$$\mathbf{D_q} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \tag{1}$$

The lifting factorization is often more convenient to design and implement the Quincunx filter bank [9]. The lifting structure guarantees Perfect Reconstruction (PR), and the so-called predict and update lifting steps can be used to increase the order of the polyphase matrix (and thus of the filters) while maintaining PR. The predict and update steps ($P(z)$ and $U(z)$ respectively) involve factorizing the analysis polyphase matrix in factors of the following form:

$$P(z) = \begin{bmatrix} 1 & U(z) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -P(z) & 1 \end{bmatrix} \tag{2}$$

The corresponding synthesis matrix is given by

$$\tilde{P}(z) = \begin{bmatrix} 1 & 0 \\ P(z) & 1 \end{bmatrix} \begin{bmatrix} 1 & -U(z) \\ 0 & 1 \end{bmatrix} \tag{3}$$

It is easy to see that this filter bank is PR by construction, regardless of the specific choice of the update or predict steps $U(z)$ or $P(z)$.

By using multiple such update and predict steps, with different $P(z)$ and $U(z)$ functions, a Quincunx filter bank with higher order filters can be constructed. The lifting realization of a quincunx filter bank has the form shown in Fig. 3.

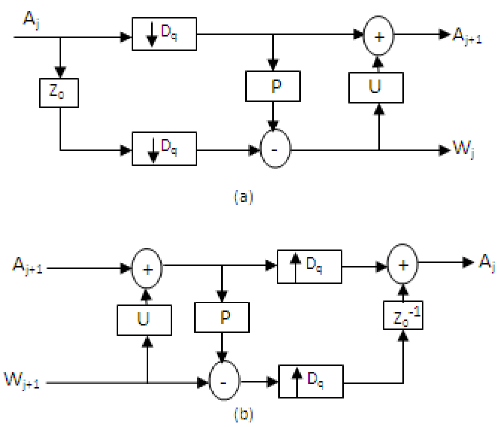


Fig. 3 Lifting Schme of Quincunx filter bank. (a) Analysis (b) Synthesis

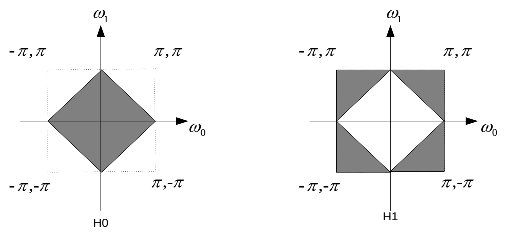


Fig. 2 Frequency supports of filters in Quincunx filter-bank

Due to the use of the lifting framework, the PR condition is automatically satisfied. The main feature of lifting is that it allows us to satisfy three properties of the filter bank separately: 1) PR: Perfect reconstruction property 2) DM: Dual vanishing moments 3) PM: Primal vanishing moments [14].

We have used quincunx interpolating filter banks designed by Kovacevic and Sweldens [10], based on the lifting scheme [9].The quincunx filter banks have truly non-separable wavelet decomposition, contrary to non-separable wavelets biased in horizontal and vertical directions.

Since the decimated transform is not shift invariant, the de-noising result will depend on the position of the discontinuities in the image. Undecimated versions of wavelet transforms are often seen to give better results [11]. In this paper, we use an undecimated version. We remove the decimation operator from the quincunx lifting scheme [10] and the N-times quincunx upsampled versions of filters are used, where N represents decomposition level. This way for higher decomposition levels the corresponding filter support widens.

III. Edge Preserving De-noising Method

Electrophoresis images can be considered as signals with many sharp features or signals containing features in a variety of frequency bands. Efficient elimination of noise, without deteriorating the significant high frequency features, can be performed in the wavelet domain (e.g. [11][12][14]). In the case of such a signal, it is impossible to distinguish between the signal and the noise component in the time domain, whereas it is possible in the time-frequency (wavelet) domain. The feature of the wavelet filtering allows an effective suppression of noise without erasing the important details in the studied signal.

We have evaluated the relative performance of separable and non-separable wavelets along with traditional approaches; and reported the results in [14]. It is concluded that non-separable wavelets have a good case for de-noising of gel images. In [14], only additive noise has been considered. The gel images are very noisy and also contain a lot of multiplicative noise along with the additive noise. In this paper, the model of an observation I is considered as:

$$I(x,y) = O(x,y).N(x,y) + W(x,y)$$

where O is noise free ideal image, N is multiplicative noise and W is additive noise.

We take the work one step ahead by introducing the method for edge enhancement along with de-noising. Bayesian threshold approach [12] has been used with slight modification for edge enhancement. In this approach, first the image I is decomposed using undecimated non-separable quincunx wavelet transform up to a desired level (say L) and detail coefficients Wj at scale (level) i are given by

$$W_i(x,y) = I * \psi_i(x,y)$$

where ψ_i is a wavelet at scale i and * denotes the convolution. The wavelet decomposition of the image I can be represented as $\{W_1, W_2, \dots \dots, W_L, A_L\}$ where A_L is approximation coefficients at last level L.

The threshold value is estimated at each scale, based on the generalized Gaussian distribution parameters:

$$threshold = \frac{\sigma^2}{\sigma_{s,i}}$$

Where σ and $\sigma_{s,i}$ are empirically calculated as:

$\sigma = 1.4826 \text{ median } (|v_1|)$

where v_1 are the wavelet coefficients from the scale 1, and

$$\sigma_{s,i} = \sqrt{\max(\sigma_i^2 - \sigma^2, 0)}$$

for $i = 1, 2, \dots, L$,

where σ_i^2 denotes the variance of the coefficients from the i-th level, and L denotes the decomposition level. A soft-thresholding policy has been applied in which all wavelet coefficients $\{v_i\}$ at each scale i having absolute value lower than the threshold are set to zero and the rest are shrunk by the threshold value.

$$v_k = \begin{cases} 0 & if\ |v_k| < threshold \\ sign(v_k)(|v_k| - threshold) & if\ |v_k| \geq threshold \end{cases}$$

For edge enhancement process, the term 'scale product' is defined as:

Definition: The scale product $P_{j,j+1}$ shows the correlation between wavelet coefficients of adjacent scales j and j+1 and defined as

$$P_{j,j+1}(x,y) = W_j(x,y).W_{j+1}(x,y)$$

In wavelet domain, it can easily be found out that the peaks due to edges propagate across scales, hence multiplying adjacent scales help to dilute the noise. In other words, by taking the scale product, edges enhance up to some level. At each scale, the scale product is calculated with its next adjacent scale, and the image is reconstructed with the following set constructed using these scale products:

$$\{P_{1,2}, P_{2,3}, P_{3,4} \dots \dots, P_{L,L+1}, A_L\}$$

Let us denote the reconstructed image as I_r. To reduce multiplicative noise, logarithmic of reconstructed image has been taken as follows:

$$IL \equiv \log(I_r(x,y)) = \log(O(x,y)) + \log(N(x,y))$$

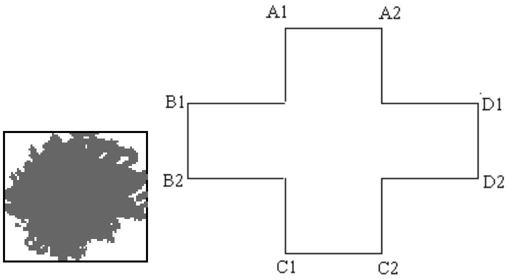
Now, multiplicative noise can be treated as well as additive noise. Therefore, logarithmic of reconstructed image is decomposed up to a level (L). Again, using same method described above, Bayes threshold is calculated at each scale and soft thresholding policy is utilized to de-noise the image. After thresholding, inverse wavelet transform is applied to reconstruct the logarithmic image. Let us denote this reconstructed image as IL_r. Now, taking exponential of the image IL_r we get final de-noised image $I_{Denoised}$.

$$I_{Denoised} = \exp(IL_r)$$

Our experiments show that this method reduces the multiplicative noise up to some satisfactory level. It also preserves more edges in comparison with the previous method which does not use scale product.

IV. Post-processing Method

In [15], a novel method for segmentation of gel images has been presented. First, the watershed transform and quincunx wavelet transform are applied on the image and then spot characteristics have been formulated through a set of rules. This approach seems to outperform the commercialized software, for which authors also have presented some results. In this paper, we have significantly improved the result by introducing our post-processing methods.



(a) (b)

Fig. 4 (a) Bounding Box (BX) around an irregular object (b) Extreme points of a plus like object. A1, A2, B1, B2, C1, C2, D1 and D2 represent top-left, top-right, left-top, left-bottom, bottom-left, bottom-right, right-top and right-bottom corner points (extreme points) respectively.

The gel image is again decomposed using the Quincunx wavelet transform and at each scale its segmented boundary information is mapped onto detail signals. For each object, detail coefficients which are in a 3x3 neighborhood of the boundary of the object are marked. Now only those marked coefficients are collected for further analysis, which are maximum in the direction of the perpendicular to the boundary at that location. These coefficients form the boundary at that scale. Thus we have multi-scale boundary information for each segmented object. Using this multi-scale boundary information, a set of bounding box *BX* and eight extreme (corner) points $\{A1,A2,B1,B2,C1,C2,D1,D2\}_j$ as shown in Fig-4 is calculated for each segmented object at each scale *j*. A1 is the top left most point within the object. First the object is searched in the top direction and when the boundary is reached, it is searched in the left direction without changing its vertical position. When the last left point at the boundary is reached at the same vertical position, that point is designated as 'top-left' and we denote it as A1. In a similar manner, all extreme points are found out. Assuming $\sum_j$ denote the sum of variable p across all scale *j's*, following variables have been derived corresponding to an object:

$$Differences\ of\ Extreme\ Points\ (DEP) = \left(\sum_j (A2 - A1) + \sum_j (B2 - B1) + \sum_j (C2 - C1) + \sum_j (D2 - D1) \right)$$

$$Sum\ of\ Perimeters\ across\ Scales\ (SPS) = \sum_j (perimeter\ of\ bounding\ box\ BX\ at\ scale\ j)$$

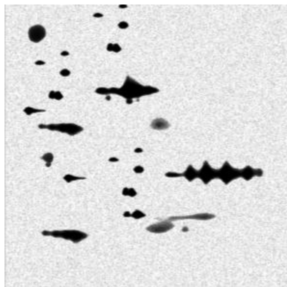
If *DEP* < *k. (SPS/12)*, then the object is considered a spot, otherwise it is marked as a streak (k is a scaling factor). By measuring through sum of parameter across scales, accuracy of detection is improved. On at a single scale, it may happen that the spot boundary seems much like the boundary of a streak. Therefore, a multi-resolution based post processing method improves the overall results.

TABLE I. COMPARISON OF PRE-PROCESSING METHOD

PSNR Values	**Noisy Image**	**Our Method**	**Method in [14]**
	PSNR	PSNR	PSNR
Image1	20.73	32.67	30.56
Image2	19.34	31.78	28.96
Image3	17.23	31.63	26.74
Image4	18.67	30.43	26.03
Image5	20.45	32.11	29.93

V. EXPERIMENTS AND RESULTS

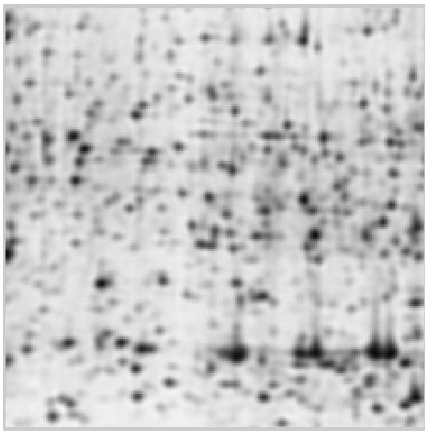
In [14], performance of the Quincunx wavelet for de-noising has been studied. It is found that for the 2d gel images, using a non-separable wavelet like the Quincunx wavelet outperforms the other methods. Here, we improved on that method and compared our results with the method presented in [14]. Results are presented in the Table-I and Fig-5. The post-processing method along with the segmentation method given in [15] is compared with a popular commercialized software package (Delta2D) for 2d gel images. Our experiments show that our method outperforms the commercialized software. Result in terms of the numbers of detected true spots and false spots has been shown in the Table-II. True spots indicates the number of spots correctly detected in the images and false spots



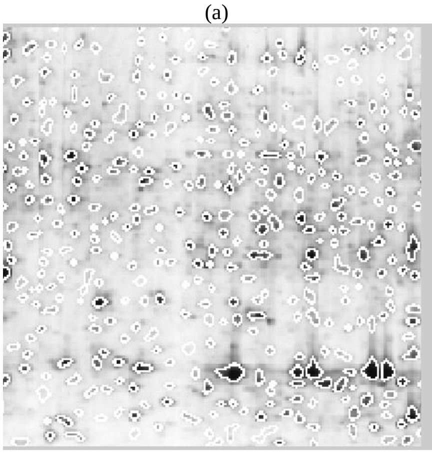
(a) Noisy Image PSNR=17.23

(b) Denoised Image PSNR=31.63

Fig. 5 A synthetic image used in experiment for measuring effectiveness of de-noising method. (a) the noisy image after introducing additive and multiplicative noise and (b) the image after de-noising using our method



(a)



(b)

Fig. 6 (a) A part of original gel image (b) segmented result of that part of gel image when pre-processed and post-processed using our method.

shows number of artifacts in background detected as spots by the specified methods. Fig-6 shows a part of original gel image before and after processing using our method.

VI. CONCLUSIONS

In this paper, we have used a shift invariant Quincunx wavelet for the pre-processing and post-processing of 2Dgel images. Lifting scheme has been employed since it enables parallel, in place and faster computation of the non-separable Quincunx wavelet transform. Our aim of research is to analyze a typical 2D gel image rapidly and accurately. This paper features new algorithms for edge preserved noise filtering, and streak detection. Our experiments show a lot of improvement in the result using our method in comparison of the other methods.

TABLE II. COMPARISON OF OUR METHOD WITH DELTA2D

	Our Method		Delta 2D	
	True Spots	False Spots	True Spots	False Spots
Image1	1250	78	1125	110
Image2	1327	34	1264	52
Image3	963	89	898	134
Image4	2134	28	2052	49
Image5	1762	79	1761	95

ACKNOWLEDGMENT

The authors would like to thank Dr M.G.R Rajan of Radiation Medicine Centre, Parel, B.A.R.C., Mumbai and Dr. H.S.Mishra, Molecular Biology Division (MBD) BARC, Mumbai who assisted the authors considerably in this work and provided important inputs.

REFERENCES

[1] P.H. O'Farrell, 'High resolution two-dimensional gel electrophoresis of proteins," *J.Biological Chem*., vol.250,pp.4007-4021,1975.

[2] L.Anderson and N.Anderson, "Some perspectives on two-dimensional protein mapping," *Clinical Chemistry*, vol.30, pp.1898-1905,Dec.1984.

[3] Pierre Marie Nugues, "Two-Dimensional Electrophoresis Image Interpretation," *IEEE Trans. On Biomedical Engineering*, Vol.40, No.8, August1993.

[4] K.Kaczmarek *et al*, "Preprocessing of two-dimensional gel electrophoresis images," *Proteomics* vol.4, pp.2377-2389, 2004.

[5] P. Cutler, G. Heald, I. R. White, and J. Ruan, "A novel approach to spot detection for two-dimensional gel electrophoresis images using pixel value collection," Proteomics, vol. 3, pp. 392–401, 2003.

[6] Yi-Sheng Liu et al, "Spot detection for a 2-DE gel image using a slice tree with confidence evaluation," Mathematical and Computer Modelling, vol.50, pp.1-14, 2009.

[7] Ramakrishnan Kazhiyur-Mannar, Dominic J. Smiraglia, Christoph Plass, Rephael Wenger, "Contour Area Filtering of two-dimensional electrophoresis images," Medical Image Analysis, vol. 10, pp. 353-365, 2006.

[8] L.Vincent and P.Soille, "Watershed in digital spaces: An efficient algorithm based on immersion simulations," *IEEE Trans. On Pattern Analysis and Machine Intelligence*, Vol.13, no.6, pp. 583-598, June 1991.

[9] W.Sweldens, "The lifting scheme: A custom-design construction of bioorthogonal wavelets," *Applied and Computational Harmonic Analysis*, vol.3,pp.186-200,1996.

[10] Jelena Kovacevic and Wim Sweldens, "Wavelet Families of Increasing Order in Arbitrary Dimensions," *IEEE Transactions on Image Processing* 9(3), pp. 480-496, March 2000.

[11] R.Coifman and D. Donoho, "Translation-invariant Denoising," *in Wavelets and Statistics*, A Antoniadis and G. Oppenheim (eds.), Springer-Verlag,pp.125-150,1995.

[12] F. F. Abramovich, T. Sapatinas, and B. Silverman,"Wavelet thresholding via a Bayesian approach," *J. Royal Statistical Society*, 1998, vol. 60, pp. 725-749.

[13] S. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*11, pp. 674–693, July 1989.

[14] Ratnesh Singh Sengar, A. K. Upadhyay, Pushkar G. Patwardhan, Manjit Singh, Vikram M. Gadre, "Approaches based on non-separable filter banks in 2D Gel electrophoresis image analysis," *Proceedings of Asia Pacific Signal and Information Processing Association (APSIPA) International Conference*, p.p. no. 387-392, December 2010.

[15] Ratnesh Singh Sengar, Ashutosh Kumar Upadhyay, Manjit Singh, Vikram M. Gadre, "Segmentation of Two Dimensional Electrophoresis Gel Image Using the Wavelet Transform and the Watershed Transform," *The eighteenth annual National Conference on Communications (NCC-2012)*, Accepted for Oral Presentation and presented, February 3-5, 2012.