

Collaborative Priors with SVD for Denoising of cDNA Microarray Images

S. Elavaar Kuzhali^{1*} and D. S. Suresh²

¹Department of Electronics and Instrumentation Engineering, Ramaiah Institute of Technology, Bangalore – 560054, Karnataka, India; kuzhalisubbnaiyah@msrit.edu

²Channabasaveshwara Institute of Technology, Gubbi, Tumkur – 572216, Karnataka, India; sureshstumkur@yahoo.co.in

Abstract

Background/Objectives: This study presents an external and internal prior guided patch based filter for minimizing Gaussian noise in complementary deoxyribonucleic acid (cDNA) microarray images. **Methods/Statistical analysis:** The proposed denoising filter is capable of taking into consideration both external and internal prior. It employs combined prior guided patch based denoising considering various distance based patch-matching methods. **Findings:** Experimental results demonstrate that the combined prior guided patch based filter outperforms the existing well-known filters in minimizing noise of cDNA microarray images. The outcome of the proposed scheme found to offer better peak signal-to-noise ratio and structural similarity index in contrast to existing filtering techniques. Effectiveness of the proposed denoising method is also assessed by estimating the spot intensities of cDNA microarray image that reflects the effect of noise reduction in the image. **Application/Improvements:** Minimization of noise is a crucial step of cDNA microarray image processing and it aids in microarray analysis by extracting valid and good quality gene expression measurements.

Keywords: Denoising, External Prior, Internal Prior, K- nearest Neighbor, Microarray Image, Singular Value Decomposition

1. Introduction

Microarray imaging technology has shown major progress in genomic research by allowing molecular biologists to monitor thousands of expression levels of genes at a time. Nowadays, the technology is a powerful tool, which is employed in the research of toxicological problems, discovering drugs, new types of diseases, diagnosis of disease and succeeds in providing promising results. In addition, the cDNA microarray technology has found application in various fields of bioinformatics. Microarrays are thousands of discrete DNA sequences printed onto glass slides by a robotic arrayer, forming spots or probes on the glass slide. Discrete DNA sequences form a two dimensional array of spots on the glass slide. Each spot on the slide represents hybridization level of cDNA from fluorescent-labelled target sample and cDNA from fluorescent-labelled reference sample. The slides are scanned to obtain two images of red channel and green channel that reflects differentially expressed genes. With the aid

of various computational techniques, the gene spots are analysed for understanding the fundamental biological phenomena. Log-intensity ratio is widely used statistic¹ for downstream processing of microarray data. Microarray experimentation and image digitization process has error prone steps and introduces pollutants to the images. The extraction of log-intensity ratios from microarray images with accuracy is a major challenge because of presence of noise in the image. Thus, need for robust algorithms to minimize noise is essential and subsequently, the downstream processing of microarray image analysis can be effectively performed.

Noise minimization is the crucial step in biomedical image processing and the step aids in downstream processing of biomedical images. Noise minimization technique on biomedical images by means of an adaptive methodology of multi-resolution² was introduced and substantiates the need for biomedical image denoising among researchers. The method of minimizing noise for cDNA microarray images at present becomes the signifi-

*Author for correspondence

cant research issue. This segment lengthens the discussion of comprehensive study³ by briefing several new techniques for minimizing noise in cDNA microarray images. The method of Chang⁴ proposes decision based filter for the elimination of impulse noises in cDNA microarray images. It works on the thought of fuzzy logic-based noise detector using particle swarm optimization algorithm to assess a corrupted pixel vector, and then utilizes the center weighted vector median filter to enhance the image quality of microarray images. An interesting attempt for removing outliers in illumine beadchip microarray image⁵ was investigated using least weighted squares estimation, and claims that estimated statistical measure is sensitive against outliers in the data.

In latest years, some researchers have utilized diverse noise minimization techniques aimed at cDNA microarray images in their methodology. The research papers⁶⁻¹⁰ witnessed the importance of denoising procedure as pre-processing step. The attempts for noise removal techniques such as histogram equalization, logarithmic transformation, arctangent hyperbolic transformation, wavelet denoising, median integrated with anisotropic diffusion filters was performed as initial step of microarray image processing.

In recent research^{11,12} histogram equalization for contrast enhancement and noise removal procedure based on Wiener filter was implemented as a pre-processing step. The study¹³ introduces application of Wiener filter for denoising of natural images. Patch based filtering in two stages frequency domain and in spatial domain was performed. Patch level-based filtering offers good methodology to minimize the presence of noise. It shows an increase in performance when noise level is less and results in a smoother denoised image for higher noise level. Popular patch based denoising methods exclusively count on input noisy image to mine features. Non Local Means - NLM¹⁴ reconstructs the noisy patch by taking the weighted average of similar reference patches from the same input image. Block-Matching 3D filtering - BM3D¹⁵ is considered as the state of art technique. It constructs upon concept of transform domain technique by extending NLM. BM3D arranges similar reference patches into a 3D structure and then processes it through the steps of 3D linear transform, thresholding, inverse 3D transform to achieve clean estimate of noisy patch. Block-Matching 3D filtering with Principal Component Analysis - BM3DPCA¹⁶ is modification of original BM3D and is intended for increase in efficiency of BM3D.

PCA on similar reference patches followed by Haar transform is employed to produce denoised image. Local Pixel Grouping LPG-PCA¹⁷ builds similar patches based on similarity in spatial structures and employs PCA on the matrix of similar patches. In Targeted Image Denoising - TID¹⁸ the similar reference patches are mined from external databases. The method exploits PCA to learn the prior and denoising quality was improved by estimating optimal spectral matrix. Following up, mining features from input noisy image and from external database has evolved in patch based approaches. Various procedures developed most recently in this category are affine regression¹⁹, internal and external correlations¹³, anchored regressors²⁰, integrating local and non-local priors²¹, external prior guided internal prior learning²² considering images corrupted with additive Gaussian noise. Although greater research interest in patch level filtering on natural images is observed, these frameworks has never been investigated with microarray images considering the complexities in the microarray image. Previous Studies indicate that several types of noise may corrupt the microarray images. Poisson²³, Gaussian^{24,25} and exponential distributions²³ have been used as additive or multiplicative noise model to characterize the noise features of cDNA microarray images. Non-additive and non-Gaussian noise can be mathematically expressed as additive Gaussian noise²⁶, so minimizing Gaussian noise is important considering the cDNA microarray images.

Matti Nykter²⁷ proposed a simulation model for microarray image with biological and statistical characteristics. The model simulates a realistic microarray image by modelling measurement and biological errors. Several error models based on Gaussian distribution, Laplacian distribution, lognormal distribution were introduced for implementation of simulation of microarray image. By adding various types and amounts of noise that characterize biological and measurement errors to the simulated images, performance of the proposed algorithm can be effectively tested. This can give a significant understanding of the efficiency of the algorithms.

2. Proposed Methodology

cDNA microarray image denoising is defined as problem of restoring a clean image from its noisy observation. Given a noise free image I^G , its noisy observation is given as $I^N = I^G + N$, where $N \in R^2$ is the additive Gaussian noise

with σ^2 variance and zero mean. Aim is to obtain a noise free estimate $I^G \in \mathbb{R}^2$ from $I^N \in \mathbb{R}^2$. An image denoising framework that integrates internal prior learned from noisy image as well as prior learned from the dataset is proposed in this article. Microarray dataset with clean images have to be created to learn the prior. However, microarray images with high quality is not achievable practically. Therefore, high quality Microarray Image Dataset (MAID) are formed by selecting visually good quality images from Stanford Microarray database. Another dataset stacked with simulated images are also constructed and the priors of these images are used to validate the proposed framework. These synthetic images are formed considering the characteristics of realistic microarray images using mamodel software²⁷. Simulated images are used to corroborate the proposed framework. Clinical investigation of two-colour microarray data is validated using the true signals, Red(R) and Green (G) or their blend of various levels of their intensity.

The article proposes an integrated method that achieves best result by adaptively selecting internal prior and external prior for denoising of noisy image. In first stage, the noisy image I^N is split into overlapping patches of size 8×8 . For every noisy patch, aim is to restore its estimate patch by the statistics obtained from internal or external reference patches. For each noisy patch, the signal to noise ratio²⁸ is calculated. Patch with low signal to noise ratio is denoised with internal reference patches information. Patch with high signal to noise ratio is denoised with external reference patches information. The reference patches for a given noisy patch is obtained by using a search algorithm, K-Nearest Neighbor (KNN) with various distance measures. The technique considers M nearest neighbors as prior data learned from noisy image or from microarray image dataset. An intuitive method of Singular Value Decomposition (SVD) is exercised to obtain the denoised patch.

The initial step in the algorithm is to perform partitioning of the noisy image to generate overlapping patches. For each noisy patch P^N , aim is to recover its particulars with the support of similar external and internal patches. Searching reference patches for a heavily polluted noisy patch from the noisy image and from external image dataset are both hard-hitting problems. Therefore, KNN based patch matching is suggested with Chebyshev, Cityblock, Euclidean, and Minkowski distance approaches to improve patch matching accuracy.

In this approach, the results prove that, by incorporating two categories of image priors, better restoration results are obtained taking the benefits from both external and internal priors. The idea behind the algorithm is to group together similar patches and a model is framed for those patches. This model is used to reduce significant noises of microarray images. The basic steps of proposed denoising algorithm is as follows.

Algo1 for Denoising Microarray image:

1. For the size of the noisy Image I^N
2. Extract overlapped noisy patch P^N of size $N1 \times N1$
3. Find signal to noise ratio (SNR) of noisy Patch
4. If SNR of noisy Patch $> C$
5. Extract similar reference patches from MAID
6. Res_pat \rightarrow Denoise (sim_patchExt, P^N)
7. Else if SNR of noisy Patch $< C$
8. Extract similar reference patches from noisy image I^N
9. Res_pat \rightarrow Denoise (sim_patchInt, P^N)
10. END

The algorithm considers the clean microarray image as ground truth image. Gaussian noise of certain noise level is added to the ground truth image to obtain the noisy image I^N . Noisy image is split into overlapping patches. Noisy patch P^N of size $N1 \times N1$ is extracted from noisy image. The equation (1) is used to estimate the signal to noise ratio of noisy patch P^N .

$$SNR \text{ of noisy patch} = \sqrt{\frac{\text{var}(P^N)}{\text{var}(n)}} \quad (1)$$

where, $\text{var}(P^N)$ is the empirical variance of the noisy patch, $\text{var}(n)$ is unknown and is approximated using one of the known denoisers. If \tilde{P} is the estimate of P^N then $\text{var}(n) \cong \text{var}(P^N - \tilde{P})$.

The Line 4 and 7 of algorithm then makes a choice of denoising the noisy patch with internal prior data or external prior data based on threshold C . The patch with low content and high noise has lower signal to noise ratio and the patch with more signal content and low noise has high signal to noise proportion. The threshold C is calculated as average of signal to noise ratio of smooth patches of noisy image corrupted with various noise levels. This step (Line 4 and 7) ensures that patches with smooth or low content prefer minimizing noise using internal prior information, whereas patches with details prefer denois-

ing using external prior information. For each noisy patch P^N the algorithm searches for M similar reference patches $\{rp_1, rp_2, \dots, rp_M\}$ from microarray image dataset (MAID) or from the noisy image I^N . In Line 5, Searching window SW of size $N_s \times N_s$ for each image in MAID is considered for obtaining the reference patches if external prior data is required. Line 5 follows Algo2 steps for learning external reference patches. Internal prior data is generated by searching for M similar reference patches $\{rp_1, rp_2, \dots, rp_M\}$ from entire noisy image I^N . Line 8 follows Algo3 steps for learning internal reference patches. The algorithm specified in Line 6 & 9 performs denoising using singular value decomposition. The restored patches are averaged together to generate the estimate image for I^G as denoising result.

Algo2: Steps for Learning External Reference Patches:
Input: Input Noisy Image I^N

1. For size of input noisy image
2. Extract overlapped noisy patch P^N of size $N_1 \times N_1$ from I^N
3. For each image in dataset MAID
4. Initialize the dimension of search window SW for an image in MAID of size $N_s \times N_s$
5. Find similar patches for noisy patch P^N , $arrayD_1 \rightarrow KNNSearch(P^N, SW, dist_{opt})$
6. Accumulate similar patches in $arrayD$ for all images of MAID
7. END
8. Select M most similar reference patches from $arrayD$ $sim_patchExt \rightarrow KNNSearch(arrayD, P^N, dist_{opt})$
9. END

Output: External reference patches $sim_patchExt$

Algo 2 starts with considering a searching window SW of size $N_s \times N_s$ for each image in MAID for obtaining the reference patches. The dimension N_s is chosen as 51 making searching window size to be 51×51 . For each noisy patch P^N in I^N , the technique tries to search its most similar patches $\{rp_1, rp_2, \dots, rp_k\}$ from correlated microarray image dataset as the denoised version of P^N . K-nearest neighbour (KNN) with $dist_{opt}$ as Euclidean distance patch matching method is chosen as the baseline-matching criterion. The data matrix $arrayD_1$ contains the reference patches $\{rp_1, rp_2, \dots, rp_k\}$ that has lesser ℓ_2 distance from input noisy patch P^N . The sum of squared differences between the database $\{rp_j\}_{j=1}^k$

and noisy patch P^N , is carried out as measurement of similarity as

$$d(P^N, rp_j) = \|P^N - rp_j\|^2 \text{ for } j=0, 1, \dots, k \quad (2)$$

Therefore, final lattice of reference patches $arrayDis$ generated considering the information lattice $arrayD_1$ of every image from the database. The algorithm considers only the M -closest reference patches i.e. $sim_patchExt (\{rp_j\}_{j=1}^M)$ leading to reduction of dimension of prior information. K-nearest neighbor algorithm with distance measures such as Chebychev, Cityblock, Euclidean, and Minkowski is exploited as a patch matching function to search for most similar M -reference patches for the noisy patch.

Algo3: Steps for Learning Internal Reference Patches:
Input: Input Noisy Image I^N

1. For size of input noisy image
2. $Dz = median(I^N)$
3. Extract overlapped noisy patch P^N of size $N_1 \times N_1$ from I^N
4. Select search window SW from the median filtered image Dz .
5. Find M most similar reference patches from Dz $sim_patchInt \rightarrow KNNSearch(P^N, SW, dist_{opt})$
6. END

Output: Internal reference patches $sim_patchInt$

Algo3 computes median of the noisy image to increase the effectiveness of search of reference patches from noisy image. Entire median filtered image Dz is considered as search window to select M most similar reference patches. The sum of squared differences, equation (2), between the patches of noisy image $\{rp_j\}_{j=1}^k$ and noisy patch P^N is carried out to arrive at M -nearest reference patches. The K-nearest neighbor algorithm with distance measures $dist_{opt}$ such as Chebychev, Cityblock, Euclidean, and Minkowski is used to build M -closest reference patches. The methodology recognizes only the nearest M reference patches i.e. $sim_patchInt (\{rp_j\}_{j=1}^M)$ resulting in internal prior information.

Algo4: Algorithm for denoising using Singular Value Decomposition
Input: $sim_patchExt/sim_patchInt$
Start

1. Assign Rp with $sim_patchExt/sim_patchInt \{rp_j\}_{j=1}^M$

2. Formulate weight vector W
3. Compute basis vectors and diagonal matrix $[A \ S] = \text{SVD}(RpWRp^T)$
4. Compute $\mathbb{S} = ((S + \sigma^2 I))^{-1}(S)$
5. The patch is denoised by evaluating Res_Pat $\rightarrow \hat{P}^G = ASA^T P^N$

End

Output: Res_Pat

Given noisy patch P^N , aim is to recover the latent clean patch P^G . The observation model for the noisy patch is $P^N = P^G + n$ where n is additive Gaussian noise with variance σ^2 .

The objective is to recover an estimate of P^G from P^N through a linear filter that minimizes $E[\|RP^N - P^G\|_2^2]$.

R is linear operator to be estimated such that there is minimum mean squared error between ground truth and estimated denoised image. R is assumed symmetric, has linearly independent basis vectors and can be factorized. These independent basis vectors will be helpful in reducing the noise if the basis vectors are sparse having high-energy concentration. Sparsity in basis matrix is achievable by finding similar reference patches for the given noisy patch using K -nearest neighbor with various distance measures. Singular value decomposition is utilized for matrix decomposition to achieve sparse basis vectors. This matrix decomposition helps in eliminating trivial noise present in the image and preserves the principal image features. Left singular matrix of R are the Eigen matrix of RR^T and represents the principal image features. The architecture of denoising using SVD primarily computes the projection matrix A and diagonal matrix S as follows:

Given $\{rp_j\}_{j=1}^M$, the SVD computes the projection matrix A using

$$[A \ S] = \text{SVD}(Rp.Rp^T) \tag{3}$$

where, $\{rp_j\}_{j=1}^M$ is represented with notation R_p , A is the basis matrix and S is the spectral coefficients diagonal matrix. The reference patches obtained have different distances to the noisy patch in consideration. The information matrix $\text{sim_patchExt}(\{rp_j\}_{j=1}^M)$ or $\text{sim_patchInt}(\{rp_j\}_{j=1}^M)$ is elevated with higher weight for the most similar patches. Weight matrix is formulated using Gaussian kernel function. The more similar, the higher

the weight it contributes to calculating matrix A and S . The distance between rp_i and the noisy patch P^N is considered to measure the similarity. The weight matrix is defined using a common Gaussian Kernel Function as follows:

$$W = \frac{1}{Z} \text{diag} \left[e^{-\frac{\|P^N - rp_1\|_2^2}{h^2}}, \dots, e^{-\frac{\|P^N - rp_M\|_2^2}{h^2}} \right] \tag{4}$$

Z is a normalization constant and h is the tuning parameter.

Then the best estimate of orthogonal basis A and spectral value matrix S is obtained by SVD using equation (3) with weight matrix. Therefore, equation (3) becomes

$$[A \ S] = \text{SVD}(RpWRp^T) \tag{5}$$

An estimation \hat{P}^G of original patch P^G is provided by

$$\text{Res_Pat} \rightarrow \hat{P}^G = ASA^T P^N \tag{6}$$

such that the estimate has minimum mean squared error with respect to the ground truth P^G and is represented with the notation Res_Pat. The best spectral matrix \mathbb{S} that minimizes the error between ground truth patch and the estimate patch is obtained¹⁸ from the equation given:

$$\mathbb{S} = ((S + \sigma^2 I))^{-1}(S) \tag{7}$$

Thus solution to the optimization problem equation (8) leads to the denoising filter

$$(A \ \mathbb{S}) = \arg \min_{A \ \mathbb{S}} E[\|ASA^T P^N - P^G\|_2^2] \tag{8}$$

Therefore, focus is to combine internal and external prior data and target in minimizing noise of microarray image patches using SVD, which is stated as collaborative prior with SVD for denoising. In patch based denoising methodologies, noisy images are decomposed into overlapping patches. A pixel in a patch is encountered for denoising procedure for several instances, so averaging of pixel estimates is considered for the final denoised pixel. The entire process of algorithmic operation ensures minimization of noise with simpler block-based approach. The proposed methodology consistently results in remarkable image quality in terms of quantitative as well as subjective assessment to prove the effectiveness of the methodology. The next section discusses results in both subjective and objective measurements for the proposed scheme.

3. Results and Discussion

The proposed method is implemented by using matlab code. Algorithm implementation depends on parameters signal to noise ratio of noisy patch, median of the noisy image to increase the effectiveness in search of reference patches from noisy image. Noise estimation is done with the high frequency component of discrete wavelet transform. Three experiments are carried out to validate the performance of the proposed methodology. In the first experimentation, performance of filters is examined on simulated microarray images degraded with different noise levels of Gaussian noise. Validation of the proposed denoising filter is testified by assessing the quality of image by adopting the performance criterions Structural Similarity Index (SSIM) and Peak Signal to Noise Ratio (PSNR). The performance of proposed technique is evaluated by comparing it with paralleled patch based filtering techniques and common filtering methods. This collaborative approach has been implemented for the patch based denoising methods. Targeted Image Denoising - TID, Non Local Means - NLM, Block-Matching 3D Filtering with

Principal Component Analysis - BM3DPCA, Local Pixel Grouping - LPGPCA, Block-Matching 3D Filtering - BM3D, median filter, Discrete Wavelet Transform (DWT) are the methods considered for comparison. The proposed approach with Chebychev, Cityblock, Euclidean, Minkowski distance achieves the best denoising results amongst the paralleled methods for various noise levels. The result for a noise level of $\sigma = 30$ is represented in Figure 1. In Figure 1(e, j, k, l, n, and o) represents the results for paralleled patch based filtering techniques. The red circles clearly show that more noise exist in the image. Figure 1(f, g, h, and i) represents results for the suggested method with various distance metrics. The red circles represent noise to a lesser extent. The integrated approach of utilizing priors from both internal and external has ensured minimization of noises. Table 1 and 2 presents the objective image qualities of the eleven denoising schemes in terms of PSNR and SSIM values for various noise level. It should be noted that in tables and figures 'Prop+' notation denotes proposed approach. Note that collaborating the priors delivered better results than the well-known common methods in cDNA microarray image denoising.

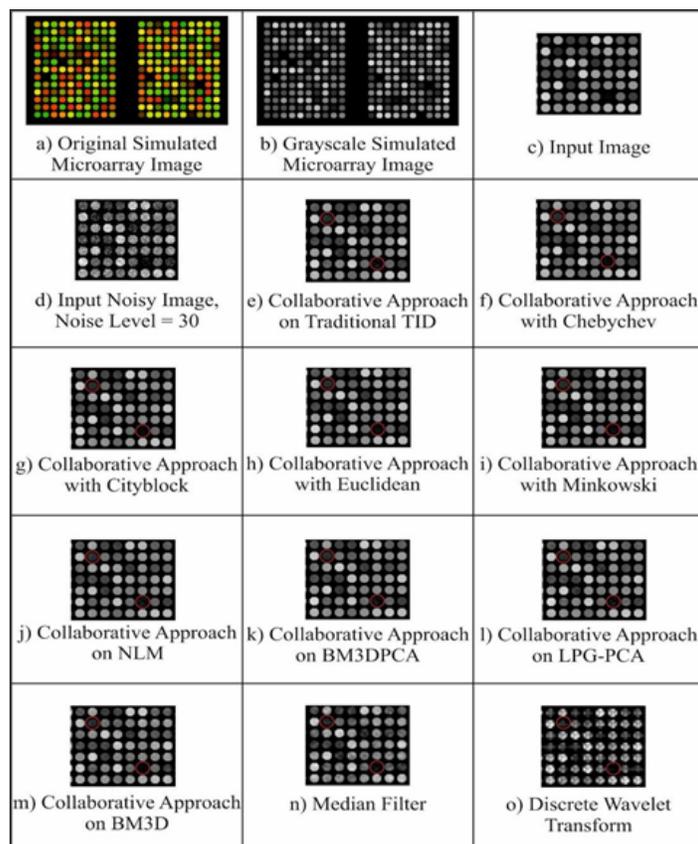


Figure 1. Denoising results - visual outcomes of simulated microarray images. Red circles highlight the noise patterns existing in the image.

Table 1. PSNR(dB) results for various noise levels on simulated microarray images

Noise Levels	10	20	30	40	50	60	70	80	90	100
Methods										
Collaborative TID	46.73	40.63	37.31	35.05	33.21	31.58	30.23	27.82	26.32	24.99
Prop+chebychev	46.97	41.20	37.92	35.47	33.00	30.71	27.26	25.28	23.85	22.91
Prop+cityblock	46.83	40.85	37.46	35.02	32.98	31.33	28.92	27.41	25.53	24.14
Prop+euclidean	46.89	41.06	37.86	35.62	33.63	32.06	30.37	27.97	26.12	24.44
Prop+minkowski	46.94	41.21	37.96	35.77	33.70	31.68	28.63	26.47	25.14	23.76
Collaborative NLM	44.72	37.62	34.35	32.56	31.21	30.06	29.02	27.10	25.70	24.27
Collaborative BM3DPCA	45.80	39.29	35.84	33.82	32.17	30.88	29.41	27.36	25.96	24.21
Collaborative LPG-PCA	42.12	36.20	32.83	30.40	28.62	27.31	25.66	24.48	23.14	21.97
Collaborative BM3D	38.10	32.10	28.29	25.81	24.11	22.91	21.81	20.72	19.79	19.07
Median	32.36	26.42	23.11	20.87	19.22	17.93	16.91	16.03	15.26	14.56
DWT	22.85	19.23	17.67	16.88	16.31	15.78	15.41	15.13	14.89	14.64

Table 2. SSIM results for various noise levels on simulated microarray images

Noise Levels	10	20	30	40	50	60	70	80	90	100
Methods										
Collaborative TID	0.999	0.997	0.994	0.989	0.982	0.975	0.967	0.954	0.936	0.916
Prop+chebychev	0.999	0.997	0.995	0.990	0.983	0.970	0.949	0.925	0.899	0.878
Prop+cityblock	0.999	0.997	0.995	0.989	0.982	0.975	0.964	0.952	0.930	0.912
Prop+euclidean	0.999	0.997	0.995	0.989	0.983	0.977	0.969	0.953	0.933	0.909
Prop+minkowski	0.999	0.998	0.995	0.991	0.985	0.977	0.956	0.930	0.910	0.889
Collaborative NLM	0.999	0.993	0.988	0.983	0.976	0.968	0.962	0.950	0.932	0.909
Collaborative BM3DPCA	0.999	0.997	0.993	0.987	0.982	0.974	0.966	0.952	0.934	0.906
Collaborative LPG-PCA	0.997	0.990	0.979	0.965	0.949	0.937	0.918	0.897	0.868	0.835
Collaborative BM3D	0.987	0.969	0.946	0.918	0.890	0.861	0.826	0.787	0.749	0.712
Median	0.961	0.889	0.814	0.742	0.673	0.613	0.560	0.513	0.471	0.432
DWT	0.849	0.720	0.636	0.575	0.523	0.477	0.451	0.428	0.408	0.391

It can be perceived from Figure 2(a) and 2(b) that, the suggested scheme evidently outperforms all other techniques for the simulated microarray image. The major reason for this trend is that patches with smooth region prefers to use internal prior and patches with edges and texture prefer to use external prior. Hence the combination of these priors provides better results compared to available existing mechanisms of denoising.

Microarray simulation model²⁷ is used for generating simulated microarray image that includes biological and hybridization errors. Errors include background noise, spot bleeding, scratches, and air bubbles etc. Hybridization and biological errors are modelled with various error models such as simple additive Gaussian noise, Dror, Hartemink, Hierarchical, Hein, Rocke etc., in the mamodel software. Simulated Microarray Image test Dataset is created with various degrees of biological

Table 3. PSNR results for simulated microarray test images with hybridization noise

Methods	Collaborative TID	Prop+ chebychev	Prop+ cityblock	Prop+ euclid	Prop+ minkowski	Collaborative NLM	Collaborative BM3DPCA	Collaborative LPG-PCA	Collaborative BM3D	Median	DWT
Image 1	23.56	23.61	23.54	23.57	23.59	23.59	23.41	23.26	22.29	19.90	19.12
Image 2	24.82	24.88	24.83	24.87	24.88	24.93	24.48	24.30	22.43	20.06	19.30
Image 3	24.44	24.49	24.44	24.44	24.46	24.39	24.26	24.08	22.79	20.31	19.19
Image 4	24.21	24.34	24.18	24.24	24.30	24.25	24.04	23.87	22.46	20.21	19.43
Image 5	25.01	25.09	24.98	25.03	25.05	25.01	24.89	24.68	23.07	20.48	19.67
Image 6	24.70	24.86	24.71	24.75	24.80	24.93	24.52	24.22	21.93	19.55	18.60
Image 7	24.34	24.40	24.32	24.38	24.40	24.30	24.14	23.98	22.46	20.24	19.33
Image 8	24.70	24.83	24.57	24.67	24.76	24.79	24.40	24.24	22.63	20.47	19.65
Average	24.47	24.56	24.45	24.49	24.53	24.53	24.26	24.08	22.51	20.15	19.29

Table 4. SSIM results for simulated microarray test images with hybridization noise

Methods	Collaborative TID	Prop+ chebychev	Prop+ cityblock	Prop+ euclid	Prop+ minko	Collaborative NLM	Collaborative BM3DPCA	Collaborative LPG-PCA	Collaborative BM3D	Median	DWT
Image 1	0.869	0.871	0.869	0.870	0.871	0.871	0.864	0.858	0.790	0.733	0.674
Image 2	0.873	0.874	0.874	0.875	0.874	0.873	0.864	0.857	0.782	0.731	0.673
Image 3	0.878	0.878	0.878	0.878	0.879	0.878	0.874	0.868	0.801	0.748	0.684
Image 4	0.858	0.862	0.856	0.860	0.861	0.861	0.850	0.844	0.775	0.729	0.676
Image 5	0.875	0.876	0.874	0.877	0.876	0.876	0.868	0.862	0.794	0.748	0.688
Image 6	0.874	0.877	0.876	0.876	0.877	0.876	0.868	0.859	0.780	0.727	0.666
Image 7	0.873	0.874	0.871	0.875	0.875	0.874	0.865	0.858	0.784	0.735	0.672
Image 8	0.837	0.841	0.836	0.838	0.840	0.840	0.827	0.821	0.752	0.711	0.657
Average	0.867	0.869	0.867	0.869	0.869	0.869	0.860	0.853	0.782	0.733	0.674

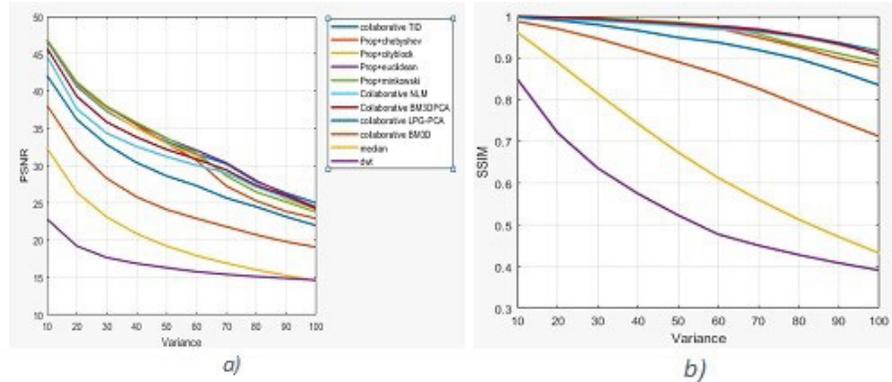


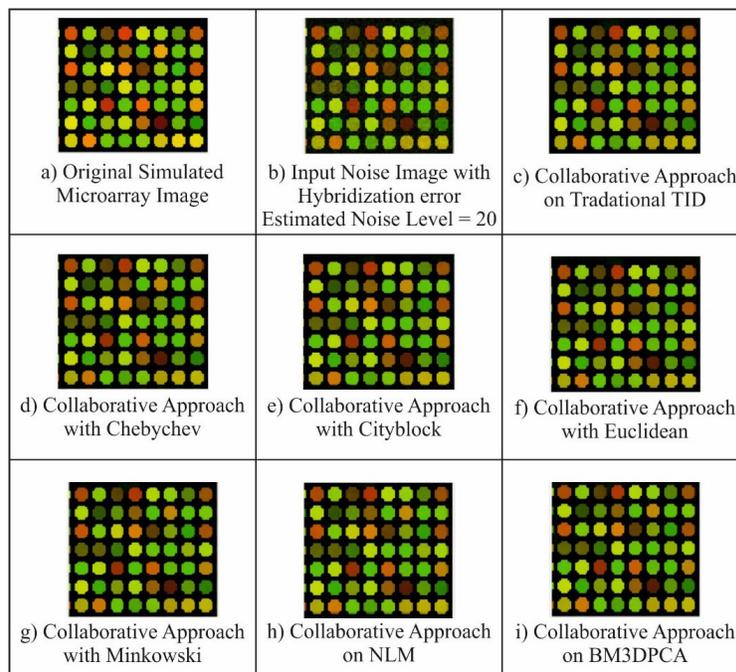
Figure 2. a) Comparative plot of PSNR values at various noise levels for simulated microarray image b) Comparative plot of SSIM values at various noise levels for simulated microarray image.

and hybridization errors included in the images using the software created by Nykter et al. To validate the efficiency of the methodology, further testing of suggested methods on simulated images included with hybridization noises is conducted and the results are tabulated in Tables 3 and 4.

The noise level in this simulated images are unknown and the simulated images that are affected with relatively very high noise level are utilized for the experimentation. The proposed algorithm depends on the noise variance (sigma) value for finding the optimal spectral coefficients (S). Noise level is estimated using the high frequency component of discrete wavelet transform.

The visual denoising results of simulated microarray test image inflicted with biological and hybridization

noise are presented in Figure 3. It can be seen clearly from visual outcome that the suggested approach with various distance metrics constantly outperforms collaborative approach for well-known patch based denoising methods and other two common schemes. The PSNR and SSIM values for the specified algorithms and for eight images in the test set are reported in Tables 3 and 4. Specifically, Figure 4(a-f) represents the visual outcomes illustrating the mean spot intensity values in red channel and green channel images. One can see that the proposed method with Chebyshev, Cityblock, Euclidean, Minkowski results in a denoised image with the spot intensity values very close to original simulated microarray image.



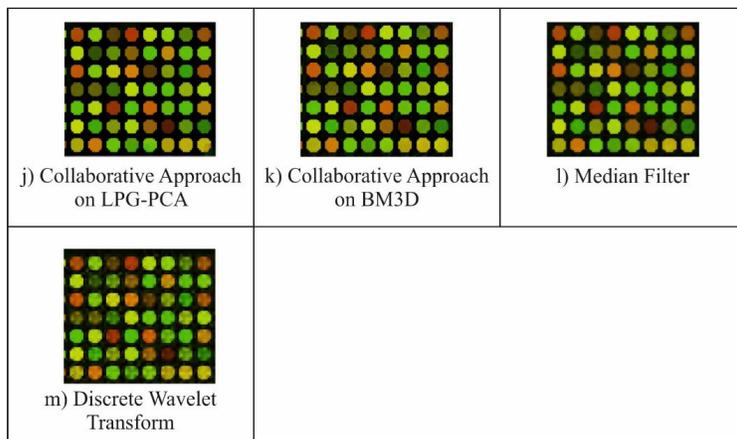
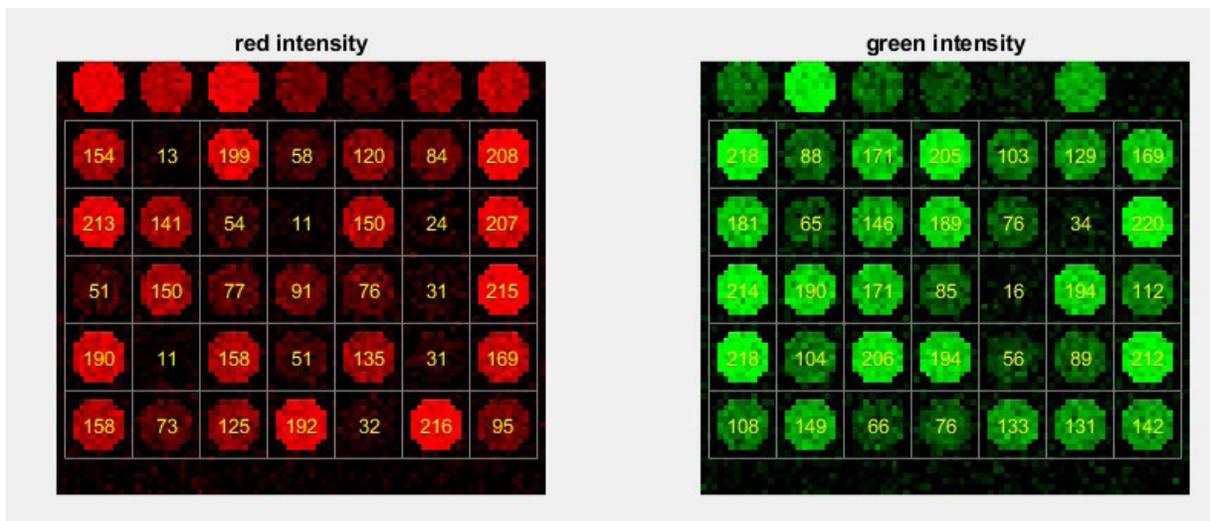


Figure 3. Denoising results of various methods on simulated microarray images with hybridization noise.



(a)



(b)



(c)



(d)



(e)

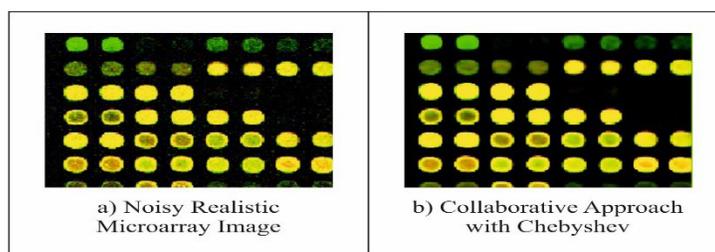


(f)

Figure 4. (a) Visual representation of simulated original microarray image and its spot intensity values for red channel and green channel images. (b) Visual representation of simulated noisy microarray image and its spot intensity values for red channel and green channel images (c) Visual representation of denoised simulated microarray image and its spot intensity values for red channel and green channel images using proposed collaborative approach with Chebyshev distance, (d) Visual representation of denoised simulated microarray image and its spot intensity values for red channel and green channel images using proposed collaborative approach with Cityblock distance, (e) Visual representation of denoised simulated microarray image and its spot intensity values for red channel and green channel images using proposed collaborative approach with Minkowski distance, (f) Visual Representation of Denoised Simulated Microarray Image and its spot intensity values for red channel and green channel images using proposed collaborative approach with Euclidean distance.

Finally, after validating our proposed method on simulated microarray images, the experiment is performed on realistic microarray image. The realistic microarray image is taken from Stanford microarray database. The visual outcomes for proposed method with various distance metric on a realistic microarray image is shown in Figure 5 (a – e). Here PSNR and SSIM evaluation on realistic microarray image cannot be carried out as there is no availability of noise free image. So the denoising results of realistic image as spot intensity values is shown in

Figure 6 (a-b) for various algorithms. From graph, it is seen that the trend lines for the suggested method is well above the noisy trend line giving us the conclusion that the suggested approach has minimized noise. Experimentation outcomes on realistic and simulated microarray images shows that the proposed method with the distance metric of Chebyshev, Cityblock, Euclidean, Minkowski retains signal quality by reducing the noise and limiting the inclusion of false patterns within the image.



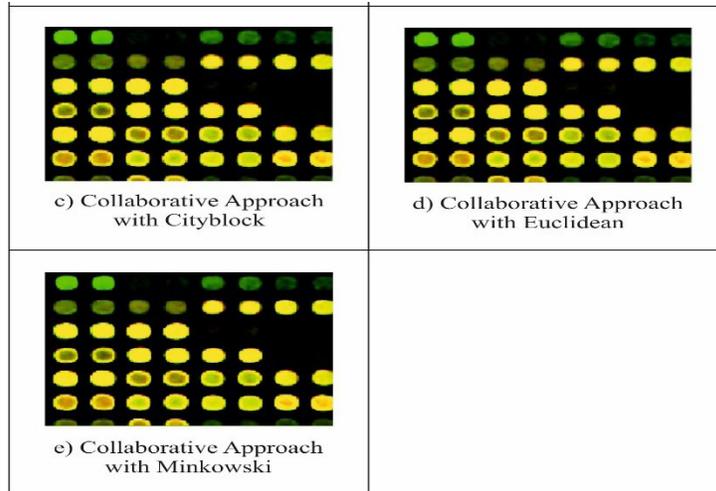
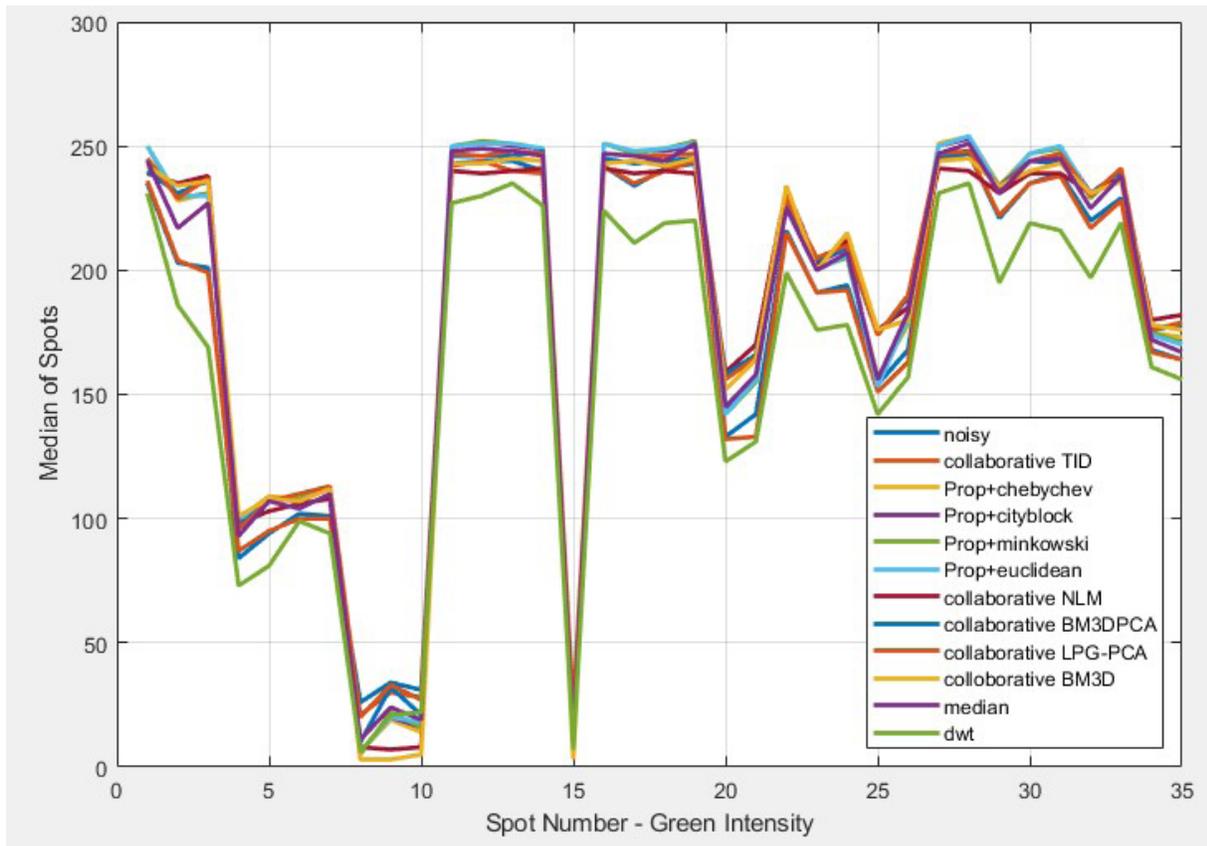
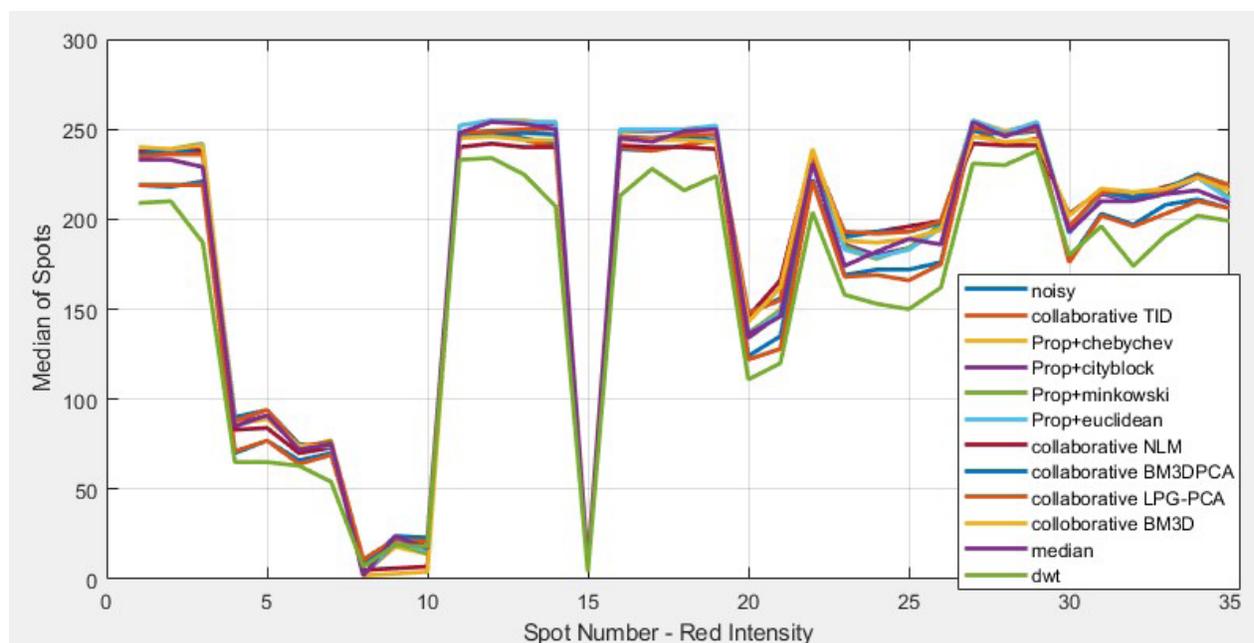


Figure 5. Denoising output of various methods on realistic microarray image.



(a)



(b)

Figure 6. (a) Comparison result of spot intensity values of various algorithms for realistic microarray green channel image, (b) Comparison result of spot intensity values of various algorithms for realistic microarray red.

4. Conclusion

In this article, an approach that combines the prior learned from noisy image and prior learned from dataset for image denoising is suggested. Firstly, algorithm considers 8×8 patches of input noisy image and based on the signal to noise ratio parameter it starts checking for its match in the inquiry window of every search image in the database or from the noisy image. The match is done considering various distance metrics. Then, the algorithm further implements singular value decomposition to enable the process of denoising. The proposed model is validated on simulated microarray image and realistic microarray image. The experimental results shows that, suggested method offers improved PSNR and SSIM in comparison to existing methods. The suggested scheme validates the effectiveness of collaborating the prior learned from noisy image and prior learned from dataset. Further, the next challenge will be sought after towards the path to improve the performance of denoising framework and to reduce the computational complexity.

5. References

1. Stekel D. Microarray Bioinformatics. *Annals of Botany*. 2004; 93(5):615–16. <https://doi.org/10.1093/aob/mch083>. PMID:PMC4242324.
2. Lahmiri S. Denoising techniques in adaptive multi-resolution domains with applications to biomedical images. *Healthcare Technology Letters*. 2017; 4(1):25–9. <https://doi.org/10.1049/htl.2016.0021>. PMID:28529760. PMID:PMC5435958.
3. Kuzhali SE, Suresh D. A Comprehensive study of Enhancement and Segmentation Techniques on Microarray Images. *i-manager's J Pattern Recognit*. 2015;2(2):46–55.
4. Chang BM, Tsai HH, Shih JS. Using fuzzy logic and particle swarm optimization to design a decision-based filter for cDNA microarray image restoration. *Engineering Applications of Artificial Intelligence*. 2014; 36:12–26. <https://doi.org/10.1016/j.engappai.2014.07.003>.
5. Kalina J. A robust pre-processing of BeadChip microarray images. *Biocybernetics and Biomedical Engineering*. 2018; 38(3):556–63. <https://doi.org/10.1016/j.bbe.2018.04.005>

6. Bariamis D, Maroulis D, Iakovidis DK. Unsupervised SVM-based gridding for DNA microarray images. *Computerized Medical Imaging and Graphics*. 2009; 34(6):418–25. <https://doi.org/10.1016/j.compmedimag.2009.09.005>. PMID:19879109
7. Belean B, Borda M, Le B, Terebes R. Computerized medical imaging and graphics FPGA based system for automatic cDNA microarray image processing. *Computerized Medical Imaging and Graphics*. 2012; 36(5):419–29. <https://doi.org/10.1016/j.compmedimag.2012.01.002>. PMID:22424667.
8. Charalambous CC, Matsopoulos GK. A new method for gridding DNA microarrays. *Computers in Biology and Medicine*. 2013; 43(10):1303–12. <https://doi.org/10.1016/j.compbiomed.2013.06.005>. PMID:24034720.
9. Helmy AK, El-taweel GS. Regular gridding and segmentation for microarray images. *Computers and Electrical Engineering*. 2013; 39(7):2173–82. <https://doi.org/10.1016/j.compeleceng.2013.07.022>
10. Wang Z, Zineddin B, Liang J, Zeng N, Li Y, Du M. cDNA microarray adaptive segmentation. *Neurocomputing*. 2014; 142:408–18. <https://doi.org/10.1016/j.neucom.2014.03.052>.
11. Khalilabad ND, Hassanpour H. Employing image processing techniques for cancer detection using microarray images. *Computers in Biology and Medicine*. 2017; 81:139–47. <https://doi.org/10.1016/j.compbiomed.2016.12.012>. PMID:28061369.
12. Khalilabad ND, Hassanpour H, Abbaszadegan MR. Fully automatic classification of breast cancer microarray images. *Journal of Electrical Systems and Information Technology*. 2016; 3(2):348–59. <https://doi.org/10.1016/j.jesit.2016.06.001>
13. Yue H, Sun X, Yang J, Wu F. Image Denoising by Exploring External and Internal Correlations. *IEEE Transactions on Image Processing*. 2015; 24(6):1967–82. <https://doi.org/10.1109/TIP.2015.2412373>. PMID:25781875.
14. Buades A, Bartomeu C, Jean MM. A non-local algorithm for image denoising. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*. 2005; 2:60–5.
15. Dabov K, Foi A. Image denoising with block-matching filtering {3D} filtering. *Proceedings of SPIE - The International Society for Optical Engineering*. 2006; 6064:354–65. <https://doi.org/10.1117/12.643267>.
16. Dabov K, Foi R, Katkovnik V, Egiazarian K. BM3D image denoising with shape-adaptive principal component analysis. *Proc Work Signal Process with Adapt Sparse Struct Represent*. 2009;6.
17. Zhang L, Dong W, Zhang D, Shi G. Two-stage image denoising by principal component analysis with local pixel grouping. *Pattern Recognition*. 2010; 43:1531–49. <https://doi.org/10.1016/j.patcog.2009.09.023>
18. Luo E, Chan SH, Nguyen TQ. Adaptive image denoising by targeted databases. *IEEE Transactions on Image Processing*. 2015; 24(7):2167–81. <https://doi.org/10.1109/TIP.2015.2414873>. PMID:25807566.
19. Pierazzo N, Facciolo G. Data adaptive dual domain denoising: A method to boost state of the art denoising algorithms. *Image Processing On Line*. 2017; 7(1):93–114. <https://doi.org/10.5201/ipol.2017.203>.
20. Timofte R. Anchored fusion for image restoration. *Proc - Int Conf Pattern Recognit*. 2017;1412–7.
21. Gu S, Zurich E, Leuven K. Integrating Local and Non-local Denoiser Priors for Image Restoration. *24th International Conference on Pattern Recognition (ICPR)*. 2018; 1:2923–38. <https://doi.org/10.1109/ICPR.2018.8545043>. PMID:29787265.
22. Xu J, Zhang L, Zhang D. External prior guided internal prior learning for real-world noisy image denoising. *IEEE Transactions on Image Processing*. 2018; 27(6):2996–3010. <https://doi.org/10.1109/TIP.2018.2811546>. PMID:29994149.
23. Chen Y, Bittner ML, Dougherty ER, Balagurunathan Y, Trent JM. Simulation of cDNA microarrays via a parameterized random signal model. *Journal of Biomedical Optics*. 2002; 7(3):507–23. <https://doi.org/10.1117/1.1486246>. PMID:12175304.
24. Adjeroh DA, Zhang Y, Parthe R. On denoising and compression of DNA microarray images. *Pattern Recognition*. 2006; 39(12):2478–93. <https://doi.org/10.1016/j.patcog.2006.02.019>.
25. Lukac R, Plataniotis KN, Smolka B, Venetsanopoulos AN. A multichannel order-statistic technique for cDNA microarray image processing. *IEEE Transactions on NanoBioscience*. 2004; 3(4):272–85. <https://doi.org/10.1109/TNB.2004.837907>. PMID:15631139.
26. Rangayyan RM, Ciuc M, Faghieh F. Adaptive-neighborhood filtering of images corrupted by signal-dependent noise. *Applied Optics*. 1998; 37(20):4477–87. <https://doi.org/10.1364/AO.37.004477>. PMID:18285899.
27. Nykter M, Aho T, Ahdesmäki M, Ruusuvaari P, Lehmissola A, Yli-Harja O. Simulation of microarray data with realistic characteristics. *BMC Bioinformatics*. 2006; 7(1):349. <https://doi.org/10.1186/1471-2105-7-349>. PMID:16848902 PMID:PMC1574357
28. Mosseri I, Zontak M, Irani M. Combining the power of Internal and External denoising. *IEEE International Conference on Computational Photography (ICCP)*. 2013; 1:1–9. <https://doi.org/10.1109/ICCP.2013.6528298>