Denoising Multi-view Images Using Non-local Means with Different Similarity Measures

Monagi H. Alkinani and Mahmoud R. El-Sakka(✉)

Computer Science Department, University of Western Ontario,
London, ON N6A 5B7, Canada
{malkinan,melsakka}@uwo.ca

Abstract. We present a stereo image denoising algorithm. Our algorithm takes as an input a pair of noisy images of an object captured from two different directions (stereo images). We use either Maximum Difference or Singular Value Decomposition similarity metrics for identifying locations of similar searching windows in the input images. We adapt the Non-local Means algorithm for denoising collected patches from the searching windows. Experimental results show that our algorithm outperforms the original Non-local Means and our previous method Stereo images denoising using Non-local Means with Structural SIMilarity (S-SSIM), and it helps to estimate more accurate disparity maps at various noise levels.

Keywords: Non-local means · Patch-based image filtering · Stereo imaging · Structural similarity index · SVD · Additive noise reduction · Disparity map

1 Introduction

Digital images are often contaminated with undesired random additive noise during data acquisition, transmission or compression phases. Additive noise is generally modelled as:

$$v(x) = u(x) + n(x), x \in \Omega \tag{1}$$

where $v(x)$ is the noisy image, $u(x)$ is the noise-free image, $n(x)$ is the additive noise, and Ω denotes the set of all pixels in the image. If $n(x)$ is a Gaussian random process, then the noise is recognized as an additive Gaussian noise. This noise varies from being almost imperceptible to being very noticeable. Image denoising schemes attempt to estimate a new image that is closer to the noise-free image.

Patch-based image filtering is mainly a proximity operation dividing the noisy image into patches, or blocks, which are then manipulated separately in order to provide an estimate of the true pixel values based on similar patches located within a searching window. Such patch-based methods include Non-local Means

A. Campilho and F. Karray (Eds.): ICIAR 2016, LNCS 9730, pp. 101–109, 2016.
DOI: 10.1007/978-3-319-41501-7_12

(NL-Means) [1], K-means and Singular Value Decomposition (K-SVD) [2], and Block Matching 3D (BM3D) [3].

The NL-Means filter is a modified version of the pixel-based bilateral filter [4]. It preserves edges and blurs homogeneous areas by exploiting similarities among the various parts of the input image. Adapting the NL-Means filter for denoising stereo images would improve the extracted depth information from noisy stereo images.

A stereo imaging system uses two cameras located within the optical axes parallel and separated by distance. This system produces two or more images called multi-view images. The depth information can be extracted from these images by analyzing the differences between the images. Utilizing multi-view images is also widespread in video denoising [5,6]. Noisy stereo images often give disappointing depth information [7–9].

In this work, NL-Means is utilized for denoising stereo images. Our proposed method extends the NL-Means searching window to search the two images when seeking similar patches. In addition, we utilize the Maximum Difference (MD) and Singular Value Decomposition (SVD) similarity metrics. To improve the speed performance of our scheme, we bounded the utilized search windows to a fixed maximum size.

2 Previous Work

Zhang *et al.* showed that using multi-view images for image denoising has an outstanding benefit over using only one-view image [10], where a noisy pixel in multi-view images is estimated based on the corresponding pixels from all other images. In addition, they extended the idea of using patch-based PCA denoising from a single image to multi-view images denoising, where similar patches are collected locally and globally from the multiple images before applying the PCA algorithm. Maximum A Posteriori-Markov Random Field (MAP-MRF) is utilized by Heo *et al.* as a model for energy minimization in order to compute the disparity maps from multi-view image [11]. They proposed an algorithm that initially restores intensity difference by adapting NL-Means algorithm. Then, the dissimilarity of support pixel distributions are calculated, where mean square error (MSE) is utilized to group similar patches. In our previous work (S-SSIM) [12], we have adapted the NL-Means method for filtering stereo images in order to improve extracting the depth information disparity maps. Yet, our previous algorithm failed to achieve encouraging results when denoising stereo images with high noise level ($\sigma > 20$), because the structural similarity index is utilized. This is due to the fact that when the noise level increases, the structural similarity index would favor those patches with a similar noise pattern not the structure of the patches.

The rest of the paper is organized as follows. Section 3 describes the utilized methodology and the proposed method. In Sect. 4, we compare the performance of our proposed method with other denoising filters. Section 5 offers concluding comments, and future work.

3 Methodology

3.1 Patch Similarity Metric

Patch similarity measures assist similarity between patches within a searching window, based on the apparent differences. Patch-based denoising methods rely on accurate patch similarity measures. We proposed using two different similarity metrics: Maximum Difference or SVD-based similarity metrics. The maximum difference intends to find similarity between patches by computing the absolute maximum difference between reference patch and other patches. Meanwhile, SVD-based measure utilizes Singular Value Decomposition (SVD) to assist the similarity between patches [13]. SVD-based similarity metric consists of two stages: in the first stage, visual quality features are extracted from patches and their singular values are calculated, whereas in the second stage a machine learning scheme is utilized to identify the most discriminant features.

3.2 Non-local Means

The original NL-Means filter divides the input images into patches before filtering each patch separately. The similarity between patches in NL-Means is assessed based on the Euclidean and the luminance distances between patches. Patches with similar grey levels are assigned larger weights when averaging. Equation 2 is used to estimate a pixel i using NL-Means filter,

$$NLMeans\,[v]_i = \sum_{j \in I} \omega(i,j)\,[v]_j \tag{2}$$

where $[v]_i$ and $[v]_j$ are pixels intensities at location i and j, respectively, and $\omega(i,j)$ is a similarity measure between pixels i and j. The similarity weight, $\omega(i,j)$, satisfies the condition $0 \leq \omega(i,j) \leq 1$ and $\sum_j \omega(i,j) = 1$. It depends on the grey level similarity and the Euclidean distance between vectors $N\,[v]_i$ and $N\,[v]_j$, where $N\,[v]_k$ denotes a square neighbourhood of fixed size and centred at a pixel k. The weights are described as,

$$\omega(i,j) = \frac{1}{Z(i)} e^{-\frac{\|(N[v]_i) - \left(N[v]_j\right)\|^2}{h^2}} \tag{3}$$

where $Z(i)$ is a normalization factor and h is a filtering parameter set depending on the noise level. NL-Means filter preserves edges, regardless of their directions. The searching windows and patches size are determined based on the noise level. The size of the patches increases when the noise level is high.

3.3 The Proposed Method

We describe a new method for improving the process of denoising stereo images. The novelty of this method is the use of the NL-Means algorithm to denoise multi-view images by using either the maximum difference or SVD similarity measurements. Similar patches from left and right stereo images are used in order to increase the number of similar patches. In order to find similar patches, the reference patch is compared with other patches by using either MD or SVD-based similarity metrics. When there are similarities between the reference patch and the other patches, patches are grouped for weighted averaging process. Figure 1 shows an example of a collection of similar patches in a stereo image pair using our proposed method.

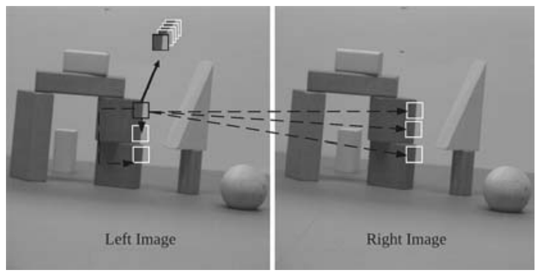


Fig. 1. Collecting similar patches from a stereo image: the patch with a black border is the reference patch, and the patches with white borders are similar patches

3.4 Algorithm Outline

Our algorithm is illustrated in Fig. 2. At each pixel k, the following procedure is performed:

1. Choose a fixed-size square patch "reference patch" $N\,[vl]_k$ centered at location k from the left image.
2. Use either MD or SVD-based similarity metrics to find the best match $N\,[vr]_q$ centered at location q for the reference patch within a bounded searching area in the right image and identify its window location.
3. Collect patches from the two windows and assign weights ω to each patch. Similar patches to the reference patch are assigned high weights. The weights are assigned as described in Eq. 3.
4. Calculate the weighted average of patches, in order to estimate the true pixel of the left image. The estimated value $NLMeans\,[vl]_i$, for a pixel i located in the left image is computed as described in Eq. 2.

4 Results

This section provides an experimental study for the performance of the proposed methods. We use a fixed 5 × 5 patch size and a fixed 9 × 9 searching window size. The fixed bounded searching area is used with the size (−20, +20). Four stereo images are used in this experiment. The four images are grey-scale images, and they are shown in Fig. 3. MatLab is used for this experiment. The

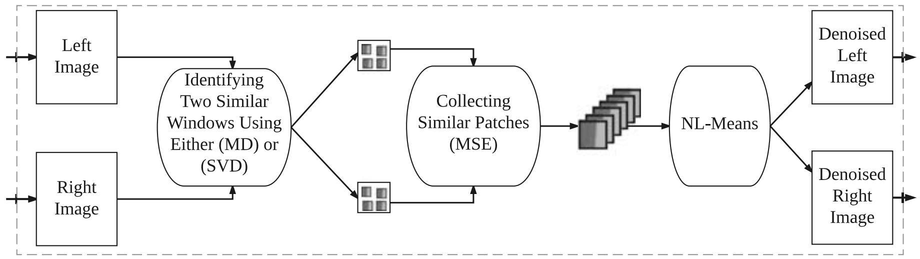


Fig. 2. A block diagram of the proposed denoising method for stereo image denoising

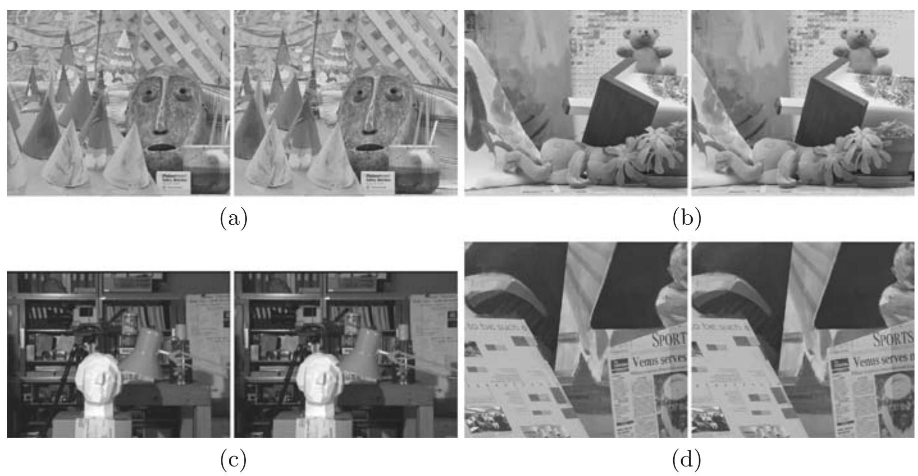


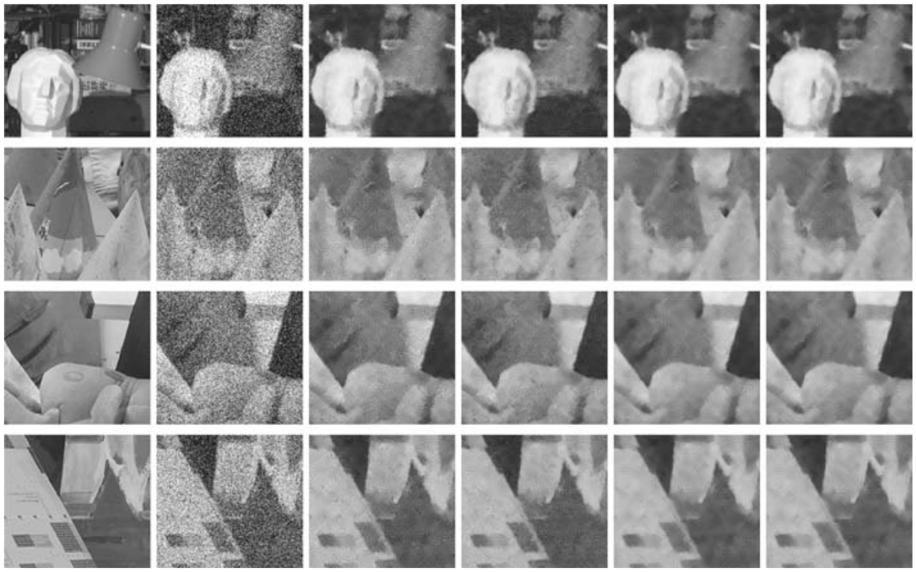
Fig. 3. The four used stereo images in the experiment: (a) *cones* images 450×375, (b) *teddy* images 450×375, (c) *tsukuba* images 384×288, and (d) *venus* images 434×383.

computers processor is Intel Core i7 (2.5 GHz). The methods are evaluated both qualitatively and quantitatively, respectively.

4.1 Qualitative Evaluation

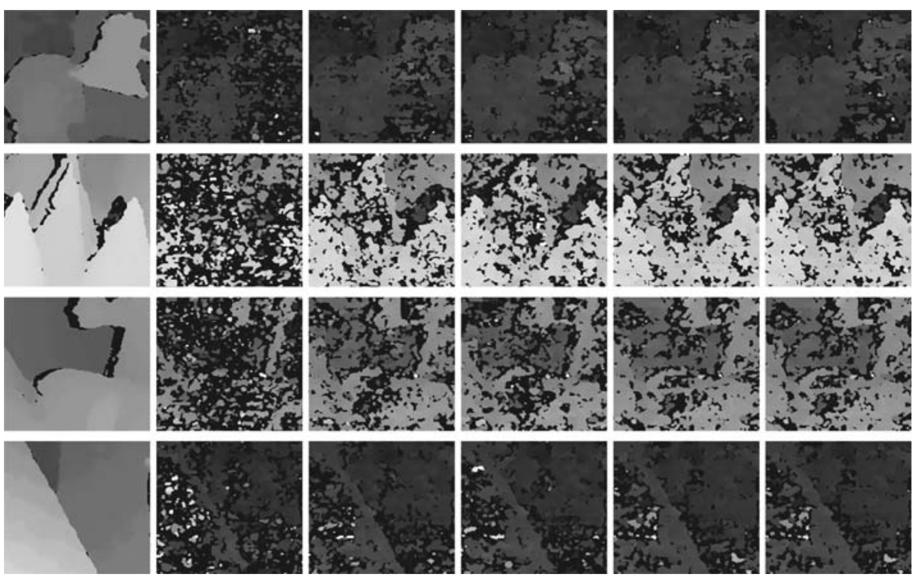
The methods are perceptively evaluated in this section. Additive White Gaussian Noisy (AWGN) stereo images with ($\sigma = 40$) are chosen to perform this evaluation. Fragments of the four noisy grayscale stereo images and the corresponding estimates are shown in Fig. 4. Each column shows the results of the same denoising method when applied to different images. Figure 5 shows the disparity maps of these denoised images.

The fragments images in Fig. 4 show that our methods outperform other methods. Our methods preserve sharp edges; e.g., the newspaper edges in `venus` image. Homogeneous regions are smoothed properly by our methods; e.g., head



(a) (b) (c) (d) (e) (f)

Fig. 4. Fragments of the noisy grayscale stereo images: (a) Original images, (b) AWGN images (σ= 40), (c) NLM, (d) S-SSIM, (e) S-MD* and (f) S-SVD*



(a) (b) (c) (d) (e) (f)

Fig. 5. Fragments of the disparity map images: (a) Original images, (b) AWGN images (σ= 40), (c) NLM, (d) S-SSIM, (e) S-MD* and (f) S-SVD*

and lamp in `tsukuba` image. The fragments of the disparity maps in Fig. 5 show that our methods outperform other methods. Our methods produce disparity maps with less errors; e.g., head and lamp in the disparity map of `tsukuba` image.

4.2 Quantitative Evaluation

Two image similarity metrics are used for the objective comparison between the results: (1) Mean SSIM (MSSIM) [14], and (2) peak signal-to-noise ratio (PSNR). The best result for SSIM is 1, while the PSNR has good results when its value is high. The experimental results of our proposed method are shown in Table 1, which compares the performance of our proposed methods (S-MD* and S-SVD*) with the original NL-Means and our previous method (S-SSIM). A bold font with a wavy under-bar highlights the highest values of SSIM, while the highest values of PSNR are highlighted with a bold font.

Figure 6 shows the performance of the methods on `tsukuba` image at various noise levels. The chart depicts that our methods are the preferable ones (from SSIM point of view). Our methods are preferable when they are used for denoising any of the four stereo images at any noise levels.

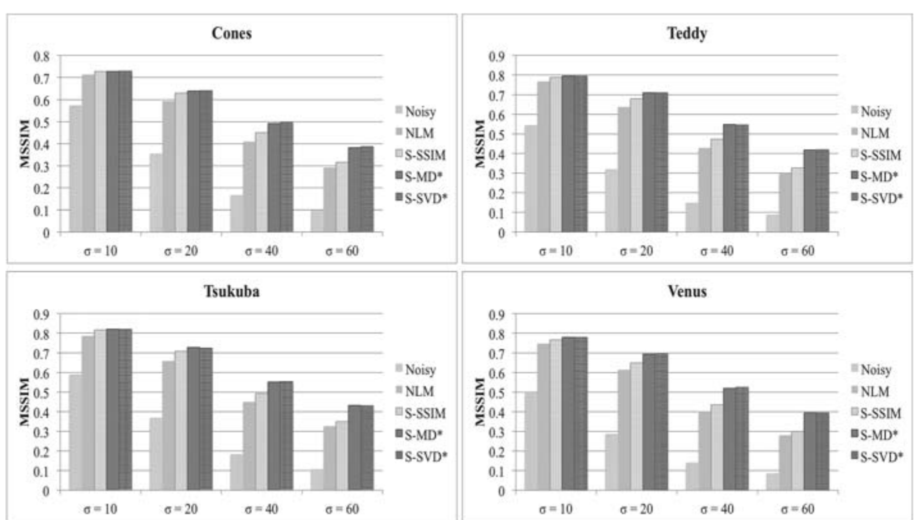


Fig. 6. The MSSIM performance of the denoising methods: NLM, S-SSIM, S-MD*, and S-SVD*, for the four stereo images at varies noise levels (σ).

Table 1. The performance of the denoising algorithms at various noise levels (σ).

	σ	$\sigma = 10$		$\sigma = 20$		$\sigma = 40$		$\sigma = 60$	
	Method	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR
Cones	Noisy	0.572	25.06	0.353	21.13	0.165	15.91	0.097	12.91
	NLM	0.714	26.58	0.593	25.05	0.409	22.59	0.292	20.63
	S-SSIM	0.727	26.76	0.630	25.47	0.450	23.12	0.316	21.1
	S-MD*	0.728	26.79	0.639	25.57	0.491	23.67	0.382	22.24
	S-SVD*	0.729	**26.83**	0.640	**25.59**	0.496	**23.73**	0.387	**22.31**
Teddy	Noisy	0.543	25.68	0.317	21.42	0.147	16.16	0.087	13.19
	NLM	0.767	27.83	0.636	26.22	0.429	23.4	0.298	21.18
	S-SSIM	0.788	28.06	0.679	26.69	0.473	24.07	0.326	21.82
	S-MD*	0.796	**28.12**	0.710	**26.92**	0.547	**24.78**	0.417	23.1
	S-SVD*	0.795	28.11	0.709	26.89	0.546	**24.78**	0.419	**23.14**
Tsukuba	Noisy	0.588	25.63	0.367	21.67	0.183	16.79	0.106	13.82
	NLM	0.787	27.51	0.659	25.83	0.449	23.07	0.327	21.06
	S-SSIM	0.817	27.99	0.708	26.47	0.492	23.69	0.349	21.54
	S-MD*	0.821	28.02	0.727	**26.58**	0.551	**24.26**	0.431	22.65
	S-SVD*	0.821	**28.04**	0.724	26.53	0.552	24.25	0.431	**22.68**
Venus	Noisy	0.494	24.90	0.285	21.13	0.140	16.23	0.084	13.27
	NLM	0.748	26.74	0.614	25.44	0.399	22.88	0.280	20.94
	S-SSIM	0.767	26.93	0.650	25.81	0.436	23.56	0.300	21.54
	S-MD*	0.780	26.98	0.694	25.99	0.519	24.09	0.394	22.58
	S-SVD*	0.780	**26.99**	0.695	**26.00**	0.525	**24.14**	0.394	**22.59**

5 Conclusions

In this paper, we looked at stereo image denoising as a multi-view image denoising process. We restored the noisy images by using either maximum difference or SVD-based similarity metrics. Empirical results show that our method achieved better denoising than the original Non-local Means and our previous work (S-SSIM). We used bounded searching areas instead of searching full areas, as in S-SSIM, for optimizing its speed. We believe that our work opens a door for future work, such as investigating the traditional similarity metrics in Non-local Means when assigning the weights between similar patches. We believe that using more accurate measurements as a similarity metric for assigning weights would help improving our algorithm.

Acknowledgment. This research is partially funded by the Natural Sciences and Engineering Research Council of Canada (NSERC). This support is greatly appreciated. This research is also partially funded by the Cultural Bureau of Saudi Arabia in Canada. This support is greatly appreciated.

References

1. Buades, A., Coll, B., Morel, J.: A non-local algorithm for image denoising. In: IEEE Conference Computer Vision and Pattern Recognition, CVPR 2005, vol. 2, pp. 60–65 (2005)
2. Aharon, M., Elad, M., Bruckstein, A.: K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. IEEE Trans. Sig. Process. **54**(11), 4311–4322 (2006)
3. Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K.: Image denoising by sparse 3-D transform-domain collaborative filtering. IEEE Trans. Image Process. **16**(8), 2080–2095 (2007)
4. Tomasi, C., Manduchi, R.: Bilateral filtering for gray and color images. In: 1998 IEEE International Conference on Computer Vision, Bombay, India, pp. 839–846 (1998)
5. Bennett, E.P., McMillan, L.: Video enhancement using per-pixel virtual exposures. In: ACM SIGGRAPH 2005 Papers, New York, NY, USA, pp. 845–852 (2005)
6. Danielyan, A., Foi, A., Katkovnik, V., Egiazarian, K.: Image and video super-resolution via spatially adaptive block matching filtering. In: Proceedings of International Workshop on Local and Non-Local Approximation in Image Processing (LNLA) (2008)
7. Leclercq, P., Morris, J.: Robustness to noise of stereo matching. In: Proceedings 12th International Conference on Image Analysis and Processing, pp. 606–611 (2003)
8. Malik, A.S., Choi, T.S., Nisar, H.: Depth Map and 3D Imaging Applications: Algorithms and Technologies, 1st edn. IGI Global, Hershey (2011)
9. Samani, A., Winkler, J., Niranjan, M.: Automatic face recognition using stereo images. In: 2006 IEEE International Conference on Speech and Signal Processing, vol. 5, p. V (2006)
10. Zhang, L., Vaddadi, S., Jin, H., Nayar, S.K.: Multiple view image denoising. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1542–1549 (2009)
11. Heo, Y.S., Lee, K.M., Lee, S.U.: Simultaneous depth reconstruction and restoration of noisy stereo images using non-local pixel distribution. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–8 (2007)
12. Alkinani, M.H., El-Sakka, M.R.: Non-local means for stereo image denoising using structural similarity. In: Kamel, M., Campilho, A. (eds.) Image Analysis and Recognition. LNCS, vol. 9164, pp. 51–59. Springer, Switzerland (2015)
13. Narwaria, M., Lin, W.: SVD-based quality metric for image and video using machine learning. IEEE Trans. Syst. Man Cybern. Part B Cybern. **42**(2), 347–364 (2012)
14. Wang, Z., Bovik, A., Sheikh, H., Simoncelli, E.: Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. **13**(4), 600–612 (2004)