A MULTIRESOLUTION APPROACH FOR IMPROVING QUALITY OF IMAGE DENOISING ALGORITHMS

Alexey Lukin, *Member, IEEE*

Moscow State University, Computer Science Faculty, Graphics & Media Lab

ABSTRACT

In this paper, we propose a multiresolution framework for improving the quality of several image and audio processing algorithms. The results of algorithms operating at different time-frequency (or space-frequency) resolutions are adaptively combined in order to achieve a variable resolution of a filter bank. Applications of the proposed model to image noise reduction algorithms are demonstrated with examples of non-local means and adaptive PCA algorithms.

1. INTRODUCTION

It is well known that signal processing algorithms dealing with multimedia information should account for properties of human perception in order to achieve better processing quality. There exist multiple studies of human auditory perception and many models of the human visual system which are extensively employed in image and audio *compression* algorithms [1]. However a multiresolution approach can also be successfully used to adjust properties of *processing* algorithms to our perception. For noise reduction, a multiresolution approach is able to deal with non-white noises by adapting sub-band noise thresholds to actual detected noise level at each scale.

In this paper, we consider a time-frequency (or space-frequency) resolution of filter banks commonly used for image and audio analysis and processing and propose a multiresolution approach that improves several existing algorithms.

We will focus on two recently proposed high-quality image denoising algorithms: non-local means [2] and adaptive principal component analysis (PCA) [3]. In section 2, we briefly describe these two algorithms and point out their strengths and deficiencies. In section 3, we suggest a general multiresolution framework that we will apply to these algorithms to improve their quality. We describe the modified algorithms in section 4, and present simulation results in section 5. Section 6 highlights some other areas where the proposed multiresolution method can be applied. We conclude by describing the useful properties of the proposed framework.

2. IMAGE DENOISING ALGORITHMS

Most existing image denoising algorithms are based on one of two approaches. The first type of algorithm operates in a single resolution and performs averaging of neighboring pixels to achieve noise smoothing. The second type of algorithm performs decomposition of a signal into sub-bands in order to apply some kind of coefficient shrinkage and then inverts the transform.

2.1. Non-local means

The state of the art in the first type of algorithms is a non-local (NL) means algorithm [2]. This algorithm is an extension to the widely used neighborhood filtering algorithms that form a denoised pixel $y_{i,j}$ as a weighted sum of the surrounding pixels $x_{i,j}$:

$$y_{i,j} = \sum_{(k,m)\in\Omega} x_{i+k,j+m} \cdot W(i,j,k,l)$$

Here $W(i,j,k,l)$ are the weights that usually depend on *geometric* distance of 2 pixel locations and *photometric* distance of 2 pixel values. The typical choice for W is

$$W(i,j,k,l) \approx \exp\left(-\frac{(x_{i,j}-x_{i+k,j+m})^2}{h^2}\right)\cdot\exp\left(-\frac{k^2+m^2}{\rho^2}\right)$$

In [2] it is pointed that such algorithms are not robust enough in presence of noisy data and tend to create impulsive noise "shocks" from a white noise source. So, [2] presents a method that averages pixels using photometric similarity of their *neighborhoods* instead of similarity of single pixel values:

$$W(i,j,k,l) \approx \exp\left(-\frac{\left\|v(x_{i,j})-v(x_{i+k,j+m})\right\|^2}{h^2}\right)$$

Here $v(x)$ is a vector of pixel values from a geometric neighborhood of pixel x, which is usually defined as a square window centered at the pixel x. The range Ω for *(k, m)* in the NL means algorithm can be as large as a whole image, hence the name "non-local".

Our experiments confirmed the conclusion of [2] that the NL means algorithm significantly outperforms other neighborhood filtering algorithms in terms of PSNR and visual quality. However the quality of the algorithm is significantly dependent on a selected size of pixel neighborhoods (“block size”) and on area of Ω (“search range”). For small blocks and small search ranges, the algorithm is limited to suppression of only high-frequency noise and cannot remove low-frequency (large-scale) noise. For large blocks and search ranges, the algorithm removes low-frequency noise effectively, but becomes less sensitive to small details (occupying a small fraction of a block) and tends to over-smooth them. In a section 3, we will propose a method of using different block sizes for denoising of different image sub-bands which reduces the described artifacts.

2.2. Adaptive principal components

Another type of image denoising methods uses transforms to split the image into sub-bands and applies coefficient shrinkage to sub-band signals. Such methods try to employ the energy compaction property of different transforms (e.g. discrete wavelet transform (DWT) or Karhunen-Loeve transform) to better separate image data from noise data.

The most widely used family of methods of this type is wavelet thresholding. Wavelets have a property of shape invariance of basis functions which allows one to control the Gibbs phenomenon via careful selection of a wavelet basis. However wavelets are not perfect in image energy compaction. The standard separable DWT does not allow good compaction of abrupt edges of different orientations: edges, being wide-band signals, are always spread into several wavelet sub-bands. There exist many modifications of the DWT with better rotational invariance, i.e. improving compaction of edges of different orientations [4].

A better approach for construction of a transform basis has been suggested in [3], where it is proposed to use a principal component analysis (PCA) to build a locally adaptive image basis which has the best possible energy compaction. After the locally optimal basis is found, the image block is transformed using this basis and transform coefficients are soft-thresholded. Then the transform is inverted and denoised blocks are overlap-added to produce the resulting image. The optimal energy compaction properties of the adaptive PCA transform provide very high quality in the resulting denoised images. However the algorithm performance is again very dependent on the transform block size. With small blocks the algorithm is unable to suppress low-frequency noise (see illustrations in [6]), and with large blocks the Gibbs phenomenon becomes stronger. The cause of the Gibbs phenomenon is the fixed-size support of all PCA basis functions: they are all defined on the same block. It would be advantageous to reduce the support of high-frequency basis functions to reduce the number of oscillations leading to the Gibbs phenomenon. In section 3, we propose a general framework for this.

3. MULTIRESOLUTION FRAMEWORK

Problems with the space-frequency or time-frequency resolution of filter banks do not arise only in image denoising algorithms. As we point in [5], the tradeoff between time resolution and frequency resolution is also critical to the performance of many audio processing algorithms based on the short-time Fourier transform (STFT). Better frequency resolution of the transform allowing finer separation of signal harmonics inevitably reduces time resolution and brings the Gibbs phenomenon leading to temporal smearing of transient events in audio.

There have been many attempts to build filter banks with variable time-frequency resolution for audio compression purposes [1]. However such attempts are limited by the fact that compression requires the critical sampling property of filter banks. This significantly restricts the freedom to vary time-frequency resolution. On the other hand, image and audio *processing* methods allow redundancy in oversampled filter banks which leads to the following proposed multiresolution framework for signal processing algorithms in Fig. 1.

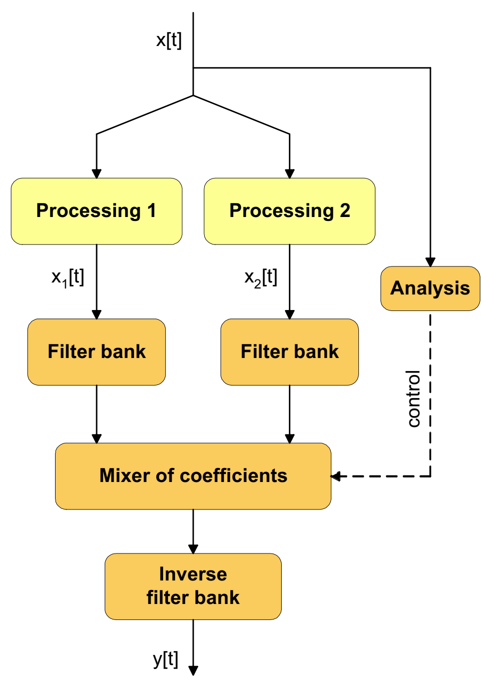

Fig. 1. Scheme of the proposed multiresolution framework.

The same processing algorithm is running in several instances with different fixed time-frequency resolutions that work in parallel on the same input data stream. Their resulting signals are combined by another filter bank with a fixed time-frequency resolution. The block where transform domain (filter bank) coefficients of different signals are mixed together will be called a “mixer of coefficients”. The process of mixing can be controlled by some prior strategy (e.g. reflecting properties of human perception) or depending on local signal features (e.g. on its stationarity).

Fig. 1 shows the example of parallel processing with just two different time-frequency resolutions, but in practice more parallel blocks can be used (depending on available computational power) to control time-frequency resolution smoother.

Since mixing of processed signals $x_1[t]$ and $x_2[t]$ is performed in the transform domain, the suggested method allows achieving of arbitrary time-frequency resolution in arbitrary areas of time-frequency (or space-frequency) plane.

4. MULTI-RESOLUTION VARIANTS OF IMAGE DENOISING ALGORITHMS

It is well known that our eye is sensitive to noise across different frequency ranges, with a maximal sensitivity at the mid-frequency range. When observing an image we first make an overall glance, and at this stage we are sensitive to large-scale details and low-frequency noise. After that we can focus our attention on particular smaller image features and at this stage we become more sensitive to smaller-scale details and higher-frequency noise. It is important that noise reduction algorithms suppress noise in a wide frequency range, especially taking into account the ever increasing pixel resolution of photo images.

4.1. Adaptive principal components

Let's apply the suggested multiresolution framework to the adaptive PCA (APCA) method from [3] in order to reduce the support area (and the number of oscillations) of the high-frequency basis vectors of the APCA. In this algorithm, the space-frequency resolution of the method is controlled by PCA block size. Let's process the input image with APCA algorithm with 2 different block sizes (we suggest 6x6 and 16x16 pixel blocks, but it can be modified depending on the image size and scale of details). As a result, we get two denoised images: one will have the Gibbs phenomenon effectively suppressed, and another will have effective low-frequency noise suppression.

To obtain the final result we simply need to combine these two images using a filter bank. We suggest using a non-decimated DWT (or a laplacian pyramid) as a mixer filter bank. Both images are transformed with the DWT (we used Haar and D4 wavelets with similar results), and two upper (high-pass) levels of wavelet coefficients are taken from the first image, while the rest of coefficients (low-pass) are taken from the second image. In this way, we take a high-quality low frequency band from the second image, and Gibbs-free high frequency band from the first image. After performing the inverse DWT we get the resulting image with significantly reduced artifacts (see [5] and [6] for demonstration).

Note that in this method we didn't use all the capabilities of the mixer of coefficients in our model. We didn't perform any image analysis and didn't adapt mixing rules to local image features. This is justified by the fact that the APCA method itself adapts the shape of its basis vectors to local image features sufficiently well.

4.2. Non-local means

The non-local means algorithm can benefit from the same multiresolution approach. In a manner similar to that described in section 4.1, we process the image with 2 versions of the NL means algorithm. One has a block size of 8x8 pixels and a search range of 5x5 pixels, and another one has a block size of 16x16 pixels and search range of 11x11 pixels. The results were combined using the same DWT approach as in section 4.1.

4.3. Optimization

A significant reduction of computational complexity is possible for both suggested denoising methods. The larger part of the complexity (typically about 70%) of these multiresolution algorithms is contained in the single-resolution processor that works with a larger block size. However, since only low frequencies are taken from the resulting image of this processor, we can perform the processing on a decimated version of the image. The original image is reduced in size by 2x2 pixel averaging. Then the processing algorithm with a large block size is applied to the downsampled image. Finally, the image is upscaled back to the original size by means of bilinear interpolation. It is important that parameters of a single-resolution algorithm are scaled to account for image downsampling. For example, the block size and search range should be reduced by 2 times to cover the same area in the image. The downsampling operation also reduces the noise level of the image (by 6 dB for white noise and 2x2 downsampling).

As a result of this modification, the overall computational cost of the multiresolution algorithms becomes only fractionally higher than the computational cost of corresponding single-resolution algorithms. The added cost is processing of the additional image with only one quarter area of the original image plus multiresolution framework overhead.

Another specific optimization is applicable to the non-local means algorithm. In the original paper [2] it was proposed to calculate the resulting image on a pixel-by-pixel basis. This is very computationally expensive, especially for large block sizes and search ranges. We suggest a significant reduction of computational complexity by averaging whole blocks instead of pixels. The size of averaged block can vary from 1x1 (the original algorithm) to the size of a similarity block $v(x)$. The dramatic reduction of computational cost is achieved at the expense of only slight decrease of visual quality and PSNR. The optimized version of the algorithm becomes suitable for realtime video processing.

5. SIMULATION RESULTS

The suggested modifications to the image denoising methods significantly improve visual quality of the resulting images by reducing the Gibbs phenomenon (ringing and noise residuals around edges) and suppressing a low-frequency noise. Fig. 2 shows the results of our modification of NL means algorithm. The results of our multiresolution adaptive PCA algorithm can be found in [5]. More detailed results can be found on our demo web-page [6]. We have tested the modified NL means method on video sequences and registered the improvements of average PSNR as shown in table 1.

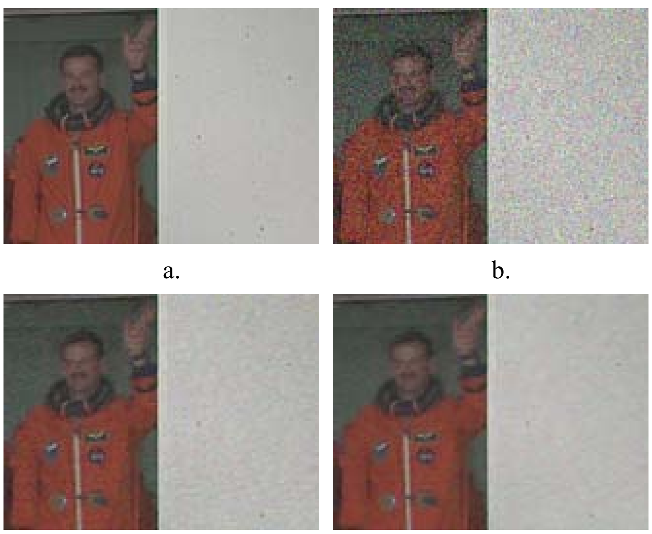

Fig. 2. Results of image denoising. **a.** original image, **b.** noisy image, **c.** non-local means method, **d.** multi-resolution variant of non-local means method.

Method	**PSNR, dB**
Noisy video	23.87
Non-local means	33.91
Multi-resolution NL means	34.28

Table 1. PSNR results for video denoising.

6. OTHER APPLICATIONS OF THE MULTIRESOLUTION APPROACH

Other applications of the multiresolution approach have been found in audio processing algorithms. In [5] we describe STFT filter banks with variable resolution for audio denoising. The adaptation of a filter bank time-frequency resolution is performed not only to the static frequency resolution of human auditory system, but also to local signal features which reduces Gibbs phenomenon near transients.

Another application has been found in analysis of audio using adaptive spectrograms. A time-frequency resolution of spectrograms can be adapted using the criterion of best compaction of local energy on a time-frequency plane. This prevents time smearing of transients and frequency smearing of stationary harmonics at the same time. The meaningfulness of a spectrogram for audio analysis is significantly increased as the new spectrogram reveals additional subtle details not present in traditional single-resolution STFT spectrograms.

7. CONCLUSION

We have proposed the multiresolution variants of two state-of-the-art image denoising algorithms and achieved the improvement in visual quality (reduction of Gibbs phenomenon and low-frequency noise) and PSNR measure. We suggest using this general framework for modification of other image and audio processing algorithms.

8. REFERENCES

[1] T. Painter, A. Spanias “A Review of Algorithms for Perceptual Coding of Digital Audio Signals”, *Proceedings of 13th International Conference on Digital Signal Processing*, 1997, vol. 1, 2-4 July 1997, pp. 179-208.

[2] A. Buades, J.M. Morel “A Non-Local Algorithm for Image Denoising”, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005, vol. 2, 20-26 June 2005, pp. 60-65.

[3] D.D. Muresan, T.W. Parks "Adaptive Principal Components and Image Denoising", *IEEE International Conference on Image Processing*, vol. 1, 14-17 Sept. 2003, pp. 101-104.

[4] F.C.A. Fernandes, R.L.C. van Spaendonck, C.S. Burrus “A Directional, Shift-Insensitive, Low-Redundancy, Wavelet Transform”, *IEEE International Conference on Image Processing*, Thessaloniki, Greece, October 7-10, 2001, vol. 1, pp. 618-621.

[5] A. Lukin, D. Kalinkina, D. Kubasov “Adaptive Multiresolution Filter Banks for Image and Audio Processing”, *Proceedings of GraphiCon International Conference on Computer Graphics & Vision*, Russia, Novosibirsk, June 2005, pp. 312-315.

[6] A. Lukin et al. “Noise Reduction for Digital Images”, *demo web page* http://audio.rightmark.org/lukin/graphics/denoising.htm

Author would like to thank our lab’s Dr. Y.M. Bayakovski, Dr. A.S. Krylov, D.K. Kubasov, D.A. Kalinkina, S.Y. Putilin, and Dr. D.S. Vatolin for helpful advice and discussions.