Neurocomputing 175 (2016) 948–965

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Efficiency of texture image enhancement by DCT-based filtering

Aleksey Rubel [a], Vladimir Lukin [a,*], Mikhail Uss [b], Benoit Vozel [c], Oleksiy Pogrebnyak [d], Karen Egiazarian [e]

[a] National Aerospace University, Department of Receivers, Transmitters and Signal Processing, Kharkov, Ukraine
[b] National Aerospace University, Department of Radioelectronic Systems Design, Kharkov, Ukraine
[c] University of Rennes 1 – Enssat, Rennes, France
[d] Instituto Politecnico Nacional, Centro de Investigacion en Computacion, Ave. Juan de Dios Batiz S/N, Mexico, D.F, Mexico
[e] Tampere University of Technology, Department of Signal Processing, Tampere, Finland

ARTICLE INFO

Article history:
Received 29 October 2014
Received in revised form
23 March 2015
Accepted 20 April 2015
Available online 6 November 2015

Keywords:
Texture analysis
Image denoising
DCT-based filtering
Visual quality

ABSTRACT

Textures or high-detailed structures as well as image object shapes contain information that is widely exploited in pattern recognition and image classification. Noise can deteriorate these features and has to be removed. In this paper, we consider the influence of textural properties on efficiency of image enhancement by noise suppression for the posterior treatment. Among possible variants of denoising, filters based on discrete cosine transform known to be effective in removing additive white Gaussian noise are considered. It is shown that noise removal in texture images using the considered techniques can distort fine texture details. To detect such situations and to avoid texture degradation due to filtering, filtering efficiency predictors, including neural network based predictor, applicable to a wide class of images are proposed. These predictors use simple statistical parameters to estimate performance of the considered filters. Image enhancement is analysed in terms of both standard criteria and metrics of image visual quality for various scenarios of texture roughness and noise characteristics. The discrete cosine transform based filters are compared to several counterparts. Problems of noise removal in texture images are demonstrated for all of them. A special case of spatially correlated noise is considered as well. Potential efficiency of filtering is analysed for both studied noise models. It is shown that studied filters are close to the potential limits.

1. Introduction

Image texture features are widely exploited in numerous applications of pattern recognition [1], remote sensing [2], similarity search in large databases [3]. In such applications, acquired (original) images are often degraded by noise that, in fact, might be the main destructive factor that prevents solving all related practical problems. Indeed, image fine texture details can be essentially disguised by the noise, and denoising is often a desired stage in the image processing chain. However, alongside with a positive effect of noise removal, the filtering can distort texture images in a larger or smaller extent. Hence, denoising should be performed more carefully in the case of texture images or texture regions of real-life images. For this reason, the noisy data should be analysed to make a decision in the followed filtering stage [1–5].

One can argue that there are many efficient image denoising techniques proposed recently [4–11]. Indeed, in the case of additive white Gaussian noise (AWGN), several filters have demonstrated good performance on different test images [11,12]. However, practically all of them run into difficulties in preserving texture. This regards partial differential equation based and total variation based denoising techniques [9,10]. Similar problems also arise for sliding window filters and other modern approaches [4,13–15]. Methods based on orthogonal transforms, in particular, discrete cosine transform (DCT) and wavelets [15–20] usually perform quite well [4]. The main reasons to apply DCT are obvious. Firstly, DCT has a good compactness of signal energy or “sparseness”. Secondly, DCT can be performed in blocks to be adapted to a local structure of processed images and to noise characteristics [4,18]. By excluding small amplitude spectrum components of transformed image data, noise removal is attained on one hand. On the other hand, the DCT-based denoising methods are, anyway, not perfect in the sense of texture preservation. While for some textures and noise intensities considerable improvement in output

* Corresponding author.
E-mail addresses: edu.rubel@gmail.com (A. Rubel), vladimlukin@yahoo.com (V. Lukin), uss_ind@mail.ru (M. Uss), benoit.vozel@univ-rennes1.fr (B. Vozel), olek@pollux.cic.ipn.mx (O. Pogrebnyak), karen.egiazarian@tut.fi (K. Egiazarian).

http://dx.doi.org/10.1016/j.neucom.2015.04.119
0925-2312/

mean square error (MSE) or peak signal to noise ratio (PSNR) is gained, there are also situations when practically no improvement is observed according to conventional criteria as output MSE or PSNR.

Another class of image denoising filters belongs to the group of so-called nonlocal filters [5,6,8,11,21–23]. These filters exploit self-similarity of image fragments (patches) and can be also equipped by other denoising mechanisms such as 3D DCT-based filtering for collected patches [21]. However, these filters also may result in distortion of image texture details, introduction of artefacts and they may run into difficulties for high-frequency data [12,18,19]. Recall that the worst efficiency of denoising has been observed for the most textural images Mandrill (also called Baboon) and Grass among the considered test images in [12,24,25]. Moreover, just for these images the potential limit of filtering efficiency for the case of additive white Gaussian noise (AWGN) is the highest (the worst). Thus, texture preservation is problematic even for the most advanced filters.

In this sense, several aspects are worth mentioning. Firstly, practically in all papers dealing with texture preservation, noise is supposed to be AWGN although it can be also spatially correlated [15,25,26]. For this reason, it is worth paying more attention to the case of spatially correlated noise. Secondly, there are different approaches to determine potential limits of noise filtering applicable for different groups of filters [12,18,26] and noise models, and analysis of these lower bounds could be interesting to understand texture images as well as the problems and limits in texture denoising efficiency. Thirdly, filter performance is mostly considered in terms of conventional criteria (metrics) as output mean square error (MSE) or peak signal to noise ratio (PSNR) [9,11–23]. Meanwhile, visual quality metrics [7,25] as well as other statistical parameters as moments [4] are worth using to characterize texture feature preservation by filters. Finally, in practice, it could be fine to predict how denoising can affect an image at hand. In particular, a natural question often arises: "Is denoising really needed for a given image or image fragment?" This question most frequently arises for texture containing images which is the main subject of our study.

First steps forward in attempts to answer this question have been made in the papers [18,27–29]. In [18,27] the authors have shown that there is a connection between simple statistics of DCT coefficients in 8x8 blocks and filtering efficiency. Due to this, these statistics can serve for predicting a parameter characterizing filtering efficiency as, e.g., the ratio of the output mean square error (MSE) and AWGN variance. Furthermore, a set of test images used to obtain such dependence did not contain enough texture images to get precise approximation and to reach the final goal of decision-making while denoising.

It has been demonstrated [25,28] that other parameters describing denoising efficiency such as improvement of PSNR (*IPSNR*) and improvement of PSNR-HVS-M (IPSNR-HVS-M) (where PSNR-HVS-M is visual quality metric [30]) can be predicted as well. Moreover, very accurate prediction can be achieved if a trained neural network (NN) is applied for approximating the dependence between statistical parameters of a noisy image and a parameter characterizing denoising efficiency [29]. Although textural images were present in the training set, thorough analysis of prediction accuracy just for texture images has not yet been performed.

Thus, there are several prime goals of this paper. First, we would like to present a wide set of simulation data characterizing denoising efficiency of two DCT-based filters [16,21] for textures with different properties and various intensities of AWGN and spatially correlated noise. Filter performance is characterized not only by conventional criterion (*PSNR*) but also by a visual quality metric. The second goal is to analyse denoising efficiency and compare it to lower bounds that can be determined for a given noise model. The third goal is to consider how well efficiency of denoising can be predicted and how to exploit prediction data.

The structure of this paper is the following. Section 2 describes image/noise models as well as DCT-based filters used in our study. Section 3 discusses how lower bounds of filtering can be determined and what they are for different images. It also describes the considered metrics of denoising efficiency. Section 4 deals with approaches to prediction of filtering efficiency, in particular, of the method that uses a trained NN for this purpose. Some details concerning numerical simulations are presented in Section 5. Analysis of the obtained results is carried out in Section 6. Then, Conclusions follow.

2. Image/noise models and considered filters

In our study, we consider a simple yet conventional image/noise model. It is assumed that an observed noisy one-component (greyscale) image is

$$I^n_{ij} = I^{tr}_{ij} + n_{ij} \quad (1)$$

where ij are pixel indices, I^{tr}_{ij} and n_{ij} are true image value and noise, respectively, $i = 1, \ldots, I_{\rm Im}$ and $j = 1, \ldots, J_{\rm Im}$, $I_{\rm Im}$ and $J_{\rm Im}$ define image size.

Concerning the true image, it is assumed to be textural since texture preservation is of our main interest in this paper. These can be real-life or artificial textures. Noise is supposed to be zero mean and Gaussian having variance σ_0^2. Note that we do not restrict ourselves by considering only white noise. Instead, we assume that noise can be both white (independent identically distributed – i.i.d.) and spatially correlated. In the latter case, it can be characterised by 2D autocorrelation function or 2D power spectrum in Fourier or other orthogonal transform basis, e.g., DCT. For spatially correlated noise, we assume the following. First, far correlation does not exist and only neighbour pixels have essential correlation of noise values. Without loss of generality, we assume in simulations that noise has the same main (vertical and horizontal) cross-sections of 2D autocorrelation function. A more important assumption is that noise spatial correlation properties are supposed to be a priori known [7] or pre-estimated with an appropriate accuracy [15]. This allows taking them into account at the filtering stage.

In this paper, two DCT-based filters are considered. Before giving their brief description, let us explain why they have been chosen for our analysis. The BM3D filter [8,21] is considered to be a state-of-the-art for suppressing AWGN. This filter is not especially suited for processing texture images but it provides practically the best results for many test images including such textural test images as Baboon and Grass [10,12] according to conventional quality metrics (output MSE or PSNR) and visual quality metrics, e.g., SSIM [31] – see data in [10]. Besides, BM3D has various modifications including those for spatially correlated noise [32,33].

The sliding DCT-based filter [7,16,43], which is a particular case of BM3D, is much simpler and faster than BM3D. Meanwhile, this filter is able to preserve texture well enough [4] and performs close to the BM3D and other advanced filters [24], especially for texture images [25]. Besides, the standard DCT-based filter can be also easily modified to take into account available information on properties of spatially correlated noise, namely, normalised DCT spectrum in 8x8 pixel blocks $W_{norm}(k,l)$ where $k,l = 0, \ldots, 7$ are indices and $k = l = 0$ relate to a direct current (DC) component.

BM3D filter exploits two denoising mechanisms [21]. First, patches (blocks) of size 8x8 pixels most similar to each given (reference) block are found. Then, these patches are collected together and the obtained 3D array (usually of size $8 \times 8 \times 2^n$) is