Computer Vision and Image Understanding xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

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

Blind video denoising via texture-aware noise estimation☆

Jinsheng Xiao[*,a,b], Hong Tian[a], Yongqin Zhang[c], Yongqiang Zhou[a], Junfeng Lei[*,a]

[a] School of Electronic Information, Wuhan University, Wuhan 430072, China
[b] Collaborative Innovation Center of Geospatial Technology, Wuhan 430079, China
[c] School of Information Science and Technology, Northwest University, Xi'an 710127, China

A R T I C L E I N F O

Keywords:
Blind video denoising
Noise estimation
Principal component analysis
Weak texture block
Gaussian noise

A B S T R A C T

Noise level is an important parameter for the design of video denoising algorithms in video processing applications. However, in practice, the noise level is unknown in most cases, but most existing denoising algorithms simply assume that the noise level is known beforehand, which severely limits their practical use. In this paper, we propose a novel blind video denoising algorithm via block-based optimal noise estimation that adaptively measures noise level by the principal component analysis. The adjacent frame images are searched to construct similar blocks by the block-matching method. The inter-frame differences of these similar blocks are used to estimate the video noise for the impact suppression of video motion, where the video noise is verified to comply with the normal distribution. The weak texture regions are selected by the thresholding function that is deduced based on the normal distribution. In addition, the proposed noise estimation approach is separately combined with several current popular non-blind video denoising methods to verify its superiority. Experimental results demonstrate that the proposed algorithm with low computational complexity not only has better estimation results, but also outperforms the state-of-the-art methods in most cases.

1. Introduction

Video signal is often contaminated by noise during the process of capturing, recording and transmitting. To meet the need of noise-free video in many practical applications, video denoising is generally used to remove the noise interference in the degraded videos as much as possible. After an overview of the recent literature, it is found that the performance of the existing video denoising methods depends heavily on the accuracy of estimated noise parameters from noisy videos. In general, video denoising can be roughly divided into two categories: blind video denoising and non-blind video denoising. The blind video denoising methods need to estimate noise level for noise removal, whereas the non-blind methods operate on the assumption that the noise level is available as a priori (Portilla et al., 2003; Tang and Jiao, 2009; Wen et al., 2008; Zhang et al., 2014). But in practice, the noise level is often unknown in many applications. The non-blind video denoising methods usually have poor results due to the inaccurate values of noise parameters in their algorithm models. Therefore, the noise estimation is a crucial issue for video denoising.

As one of the state-of-the-art image denoising methods, the dual-domain image denoising (DDID) method (Knaus and Zwicker, 2013) based on spatial and frequency information is an outstanding image denoising algorithm, which employs the bilateral filtering and short time Fourier transform (STFT) (Allen, 1977) with wavelet shrinkage for noise reduction. As for video denoising, by exploring the temporal and spatial correlation of video sequences, Varghese and Wang (2008) build a spatiotemporal Gaussian scale mixture (STGSM) model from the noisy multi-frame images to recover the clean signal from the noisy wavelet coefficients. The VBM3D method (Dabov et al., 2007) constructs the groups of similar image blocks by the block-matching approach and then filters noise of each group in the transform domain by the 3D transform-domain shrinkage and Wiener filtering. Ji et al. (2010) formulate noise removal as a low-rank matrix completion problem by grouping similar patches in both spatial and temporal domains. But these denoising methods need the known noise parameters beforehand. To overcome this drawback, Claude (Knaus and Zwicker, 2014) extended DDID (Knaus and Zwicker, 2013) to the progressive image denoising (PID) method by introducing the robust noise estimation, which

☆ The name of the Editor in chief P Kornprobst.

* Corresponding authors.
E-mail addresses: xiaojs@whu.edu.cn (J. Xiao), zhangyongqin@nwu.edu.cn (Y. Zhang), jflei@whu.edu.cn (J. Lei).

https://doi.org/10.1016/j.cviu.2017.11.012
Received 24 January 2017; Received in revised form 26 September 2017; Accepted 24 November 2017
1077-3142/

improves the results both subjectively and objectively.

Noise estimation methods can be temporal, spatial, or spatio-temporal approaches (Ghazal et al., 2007). In the spatial domain, there are three types of noise estimation: Block based methods (Fu et al., 2014), filtering based methods, and transform based methods (Liu et al., 2013). The classical noise estimation methods are based on image blocks (Mastin, 1985; Rank et al., 1999). In those methods, an input image is divided into small blocks and the standard deviation of each block is calculated. The minimum standard deviation of these image blocks is selected as the final standard deviation of the additive white Gaussian noise (AWGN) estimation. Amer and Dubois (2005) estimate the noise level by adaptively averaging the variance of image blocks with a minimum number of flat blocks. But these classical methods tend to overestimate the noise in the low-noise cases, and underestimate the noise in the condition of the high noise levels. Besides this drawback, the noise estimation based on image blocks is greatly affected by the image content and noise intensity. Pyatykh et al. (2013) proposed a noise estimation method based on the principal component analysis (PCA) with accurate results. Based on the PCA, some authors (Chen et al., 2015; Colom and Buades, 2013a; Huang et al., 2015) made further improvements for image noise estimation. Liu (2012) proposed a precise image noise level estimation method based on the singular value decomposition (SVD). The small singular values are used to estimate the noise and achieve small estimation error for various types of images, but it may be unreliable for images with rich details. Sutour et al. (2015) proposed a robust noise estimation method with promising results for different types of noisy regions except in low noise level region due to the difficulty of finding homogeneous areas.

The inter-frame correlation is considered in temporal approaches that involves with global motion. The solution is usually formulated as a motion detection or compensation method. The spatial and temporal information is exploited for video noise estimation in the literature (Ghazal et al., 2007; Yang and Tai, 2011; Zlokolica et al., 2006). Zlokolica and Philips (2004); Zlokolica et al. (2006) analyzed the coefficients of wavelet transform in the spatial and temporal domains and used the temporal and spatial gradients for the noise level estimation. But the computational complexity of this method is quite high. Ghazal et al. (2007) proposed a multi-domain method which considered domain-wise (temporal, spatial and spatial-temporal) estimations independently for the improvement of estimation reliability, where the local similarity of each domain is calculated by the Gauss–Laplace operators. Yang and Tai (2011) employed the Sobel gradients to collect homogeneous blocks and estimate the noise level from spatial, temporal-horizontal and temporal-vertical domains. Xiao et al. (2015) proposed a new video denoising method based on improved dual-domain filtering and 3D block matching, which performs good both in subjective and objective. However, these methods mentioned above have an unsatisfying accuracy of noise estimation and the high computational complexity.

To achieve the high accuracy of noise estimation, in this work, we propose a novel blind video denoising algorithm via texture-aware noise estimation. First, similar blocks are collected from the inter-frame images of video by the block-matching approach. Then the differential operation of the inter-frame similar blocks is performed for video noise estimation. Next, the well-known Normal distribution function is employed as threshold function to reduce the computational complexity. After that, an iterative index is set to make the noise estimation more accurate. In addition, the saturation effect of noise is used to avoid underestimating noise level. Experimental results show that the proposed algorithm has higher accuracy of noise estimation without the effect of the video motion and can be combined with current popular non-blind video denoising for wide application.

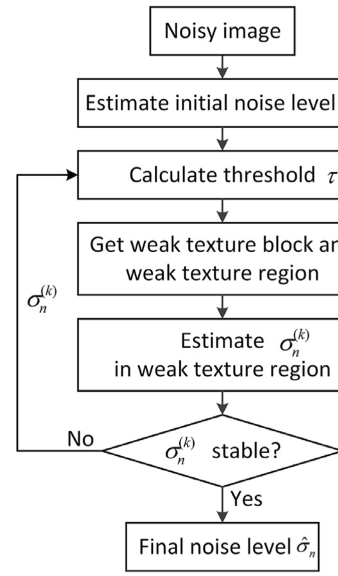

Fig. 1. Flowchart of image noise level estimation.

2. Noise estimation model via PCA

Here, Gaussian noise is assumed to be evenly distributed throughout the image. The noise estimation for constant noisy parts (weak texture block) can be represented as the noise level of the whole image. Noise estimation is an important procedure of blind video denoising. Liu et al. (2013) proposed an image noise estimation method based on PCA. Although this method is simple and effective, it tends to overestimate the noise level for images with complex structures. The noise level can be estimated accurately if the weak textured patches can be selected in noisy image. The flowchart is shown in Fig. 1.

It can be divided into selection of weak textured regions and noise estimation.

2.1. Selection of weak textured regions

From Liu et al. (2013), the image noise level can be estimated by principal component analysis of image block. An input noisy image can be decomposed into overlapping patches, whose observation model of the patches can be given as follows:

$$y_i = z_i + n_i,\ i = 1, 2, 3, \ldots, M, \tag{1}$$

where M is the number of patches, z_i is the i-th noise-free vectorized image patch of size $N^2 \times 1$, and y_i is the observed patch corrupted by the additive white Gaussian noise (AWGN) n_i with zero-mean and variance σ_n^2. The noise vectors are assumed to be uncorrelated with each other. Following the same manner of the maximum variance formulation in (Bisho, 2006), the minimum variance direction is calculable by PCA. The minimum variance direction is the eigenvector associated to the minimum eigenvalue of the covariance matrix, which is defined as:

$$\Sigma_y = \frac{1}{M}\sum_{i=1}^{M} (y_i - \mu)(y_i - \mu)^T, \tag{2}$$