A pipeline for fast video denoising

T. Ehret[1] P. Arias[1] J.-M. Morel[1]

[1] CMLA, ENS Cachan

61, avenue du Président Wilson, 94235 Cachan, FRANCE
thibaud.ehret@cmla.ens-cachan.fr

Résumé

Nous proposons un schéma de débruitage vidéo qui combine à la fois un filtrage temporel selon la trajectoire du mouvement et un algorithme de débruitage spatial ou spatiotemporel. Cette combinaison est guidée par une mesure de la qualité des trajectoires estimées.

Mots Clef

Débruitage, recalage, filtre temporel, procédé hybride.

Abstract

We propose a video denoising pipeline that combines temporal filtering along motion trajectories with a spatial or spatiotemporal denoising algorithm. The combination is driven by a weight measuring the accuracy of the estimated motion trajectories.

Keywords

Denoising, registration, temporal filtering, hybrid scheme.

1 Introduction

Video denoising is a well studied field. Most work is done to achieve the best denoising possible which usually yields extremely competitive algorithm but usually quite slow.

Currently, most state-of-the-art methods follow a patch-based approach. These patches can be only spatial [1], [3] or spatiotemporal [4]. These methods start by searching similar patches inside the video. This search is usually done in a spatiotemporal rectangular region centered on the query patch which can be motion compensated or not. Similar patches are then processed to compute clean estimate of the corresponding clean patches. This process involves for example a sparse decomposition of the patches or a Bayesian estimation.

One of the limiting factors in the performance of patch-based denoising methods is the difficulty to find the nearest patches for textures with a low-contrast in comparison to the noise [5]. For some of these cases, a global correspondence between consecutive frames (such as an optical flow) could provide more reliable matches.

Regarding the computation time, patch-based approaches are far from real-time performance. To start, searching a spatiotemporal neighborhood for similar patches is costly. To that, one needs to add the time required to filter the similar patches. As a result, these methods may take seconds to process a 1MP frame.

Some real-time denoising approaches combine temporal recursive filtering with a simple spatial denoising method. In [6] the denoised value at each pixel results from a weighted average of a temporal Kalman filter with a bilateral filter in the spatial domain. The weight depend on the magnitude of the estimated motion: if the scene is static, more weight is given to the temporal filter. A GPU real-time implementation has been presented in [7].

A similar approach has been used for burst image denoising in [8]. In this setting, several images are acquired with low-exposure times to avoid motion blur. The resulting images are registered by an homography (it is assumed that the object is planar or the scene is far from the camera) and denoised. The denoising results from a weighted average of a temporal filter and a spatiotemporal version of the NL-means algorithm [2]. The weights now depend on the variance at a pixel in the registered volume. A variance greater than the noise indicates registration errors, in which case the temporal filtering is given less weight.

In this short paper we propose a flexible video denoising pipeline by generalizing both methods in [6] and [8]. The objective of the proposed pipeline is to test several methods for the temporal filtering and spatial (or spatio-temporal) denoising within the same framework. In addition, more general image registration techniques (such as optical flow) can be used for the temporal filtering.

2 A pipeline for video denoising

In this section we present a flexible pipeline for video denoising, generalizing the denoising methods presented in [8] and [6].

First, the past video frames are registered to the current target frame. In [8] an homography is computed by matching SIFT keypoints. This approach is valid only when the scene is planar or far from the camera. Furthermore, the

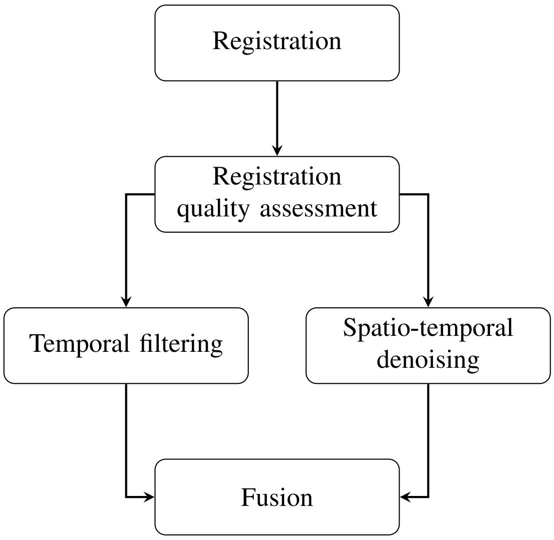

Figure 1: We propose a video denoising pipeline combining temporal filtering with a spatio-temporal denoiser. The temporal filtering is inhibited when the registration is inaccurate. Different algorithms can be used for each module, allowing to prioritize denoising quality or computation time.

SIFT registration process is quite slow. A less restrictive registration can be obtained by estimating the optical flow between consecutive frames. Even a simple stabilization algorithm might be sufficient when working with video surveillance data for example.

After the registration, the proposed scheme consists on combining (using a simple weighted average between the two methods) the results of two denoising algorithms. One of them filters the registered video in the temporal direction (a Kalman recursive filter is used in [6], and a temporal average in [8]). When the registration is accurate, a temporal averaging reduces the noise without introducing artifacts.

However motion estimation is a challenging problem, especially in the presence of noise. To handle potential registration errors a second denoising method is used. Both [6] and [8] use a spatial denoising method. Spatio-temporal denoising methods could also be used, as long as they do not depend on accurate motion estimation (e.g. [1, 4]).

The final estimate results from a weighted average of the temporal and spatial denoising methods:

$$u_{\text{hybrid}}(x) = \alpha(x)\, u_{\text{spatial}}(x) + (1 - \alpha(x))\, u_{\text{temp}}(x). \quad (1)$$

The weights $\alpha(x) \in [0, 1]$ measure the quality of the registration. The authors in [8] suggest the following weights:

$$\alpha(x) := \frac{1}{1 + \exp(c - \hat{\sigma}(x)/\sigma)}, \quad (2)$$

where $\hat{\sigma}(x)$ is the empirical variance of the registered frames along the temporal direction at location x. If the registration is accurate, the variance estimated is then the one of the noise which yields an α close to 1. In the other case, $\hat{\sigma}$ is much larger that σ which yields an α close to 0. This also explain the choice of the different algorithms and motivate the following reasoning.

This hybrid denoising scheme could also be simplified to speed up a bit the process in our case. Instead of having the dual computation for every pixels, we suggest using each algorithm separately. Indeed thanks to the large number of frame in the video, the temporal filtering scheme should suffice to compute a good estimate in the regions with no motion. The spatial denoiser, which can be more computationally expensive, is then only used when registration errors are detected. We would only need to replace equation 1 by the new split version 3 which only computes one of the estimate for each pixel instead of the two ones for the original hybrid scheme, τ being the threshold between the methods.

$$u_{\text{hybrid}} = \mathbf{1}_{\alpha > \tau} . u_{\text{spatial}} + \mathbf{1}_{\alpha < \tau} . u_{\text{temp}}. \quad (3)$$

The proposed pipeline will allow for testing different temporal and spatio-temporal denoising methods, ranging from simple filters amenable to real-time implementations to more complex algorithms if time is not a constraint.

References

[1] Dabov, K., Foi, A., and Egiazarian, K. "Video de- noising by sparse 3D transform-domain collaborative filtering". *In Proc. 15th European Signal Processing Conference*, 2007.

[2] Buades, A., Coll, B., and Morel, J-M. "A non-local algorithm for image denoising". *Computer Vision and Pattern Recognition*, 2005.

[3] Buades, A., Lisani, J. L., and Miladinovic, M. " Patch-based video denoising with optical flow estimation". *IEEE Trans. on IP, 25(6)*, 2016.

[4] P. Arias, and J.-M. Morel. "Towards a Bayesian Video Denoising Method." *Int. Conf. on AVICS*, 2015.

[5] T. Ehret, P. Arias, J.-M. Morel. "Global patch search boosts video denoising". *VISSAP* 2017.

[6] C. Zuo, Y. Liu, X. Tan, W. Wang, and M. Zhang, "Video Denoising Based on a Spatiotemporal Kalman-Bilateral Mixture Model", *The Scientific World Journal*, 2013.

[7] Pfleger, S.G., Plentz, P.D.M., Rocha, R.C.O., Pereira, A.D. and Castro, M., "Real-time video denoising on multicores and GPUs with Kalman-based and Bilateral filters fusion", *Journal of Real-Time IP*, 2017.

[8] A. Buades, et al. "A note on multi-image denoising." *Local and Non-Local Approximation in Image Processing, 2009. International Workshop on*, 2009