Optics & Laser Technology 81 (2016) 90–99

Contents lists available at ScienceDirect

Optics & Laser Technology

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

Full length article

Adaptive video compressed sampling in the wavelet domain

Hui-dong Dai [a], Guo-hua Gu [a,*], Wei-ji He [a,b,c], Qian Chen [a], Tian-yi Mao [a]

[a] Jiangsu Key Lab of Spectral Imaging & Intelligence Sense (SIIS), Nanjing University of Science and Technology, Nanjing 210094, China
[b] Science & Technology on Low-light-level Night Vision Laboratory, Xi'an 710005, China
[c] Key Laboratory of Intelligent Perception and Systems for High-Dimensional Information of Ministry of Education, Nanjing University of Science and Technology, Nanjing, 210094, China

ARTICLE INFO

Article history:
Received 1 August 2015
Received in revised form
21 December 2015
Accepted 11 January 2016
Available online 5 February 2016

Keywords:
Computational imaging
Image reconstruction techniques
Time imaging
Wavelets

ABSTRACT

In this work, we propose a multiscale video acquisition framework called adaptive video compressed sampling (AVCS) that involves sparse sampling and motion estimation in the wavelet domain. Implementing a combination of a binary DMD and a single-pixel detector, AVCS acquires successively finer resolution sparse wavelet representations in moving regions directly based on extended wavelet trees, and alternately uses these representations to estimate the motion in the wavelet domain. Then, we can remove the spatial and temporal redundancies and provide a method to reconstruct video sequences from compressed measurements in real time. In addition, the proposed method allows adaptive control over the reconstructed video quality. The numerical simulation and experimental results indicate that AVCS performs better than the conventional CS-based methods at the same sampling rate even under the influence of noise, and the reconstruction time and measurements required can be significantly reduced.

1. Introduction

Video is an important way to obtain and store information. There are a lot of redundancies in video, which provide the basis for modern video compression algorithms. However, the traditional video acquisition system samples a complete set of data and discards most of them during the compression stage for transmission or storage, which leads to a waste of resources. In addition, it can be expensive to acquire large amounts of raw video data, particularly at wavelengths where CMOS and CCD sensing technology are limited.

As a recent advancement in signal sampling, compressed sensing (CS) [1–3] offers a way to solve the above problems. It directly collects data in a compressed form, allowing acquiring and compressing simultaneously. The data to be sampled and the corresponding sampling time can be effectively reduced, as well as the pressure of the sampling system. Moreover, as a replacement for the detector array, single-pixel-based imaging device [4–6] uses a group of patterns generated by a spatial light modulator to obtain the spatial information of the scene, which provides a solution to the contradiction between growing demand for data acquisition and limited sensor resources.

A straightforward way to implement a compressed sensing video acquisition scheme is to consider each frame in the sequence independently, such as 2-D wavelet [7] and 2-D total variance (TV) [8] methods. However, they fail to address the temporal redundancy in videos. In order to take the correlations between neighboring frames into consideration, Michael B. Wakin et al. [7] consider the whole sequences as a 3-D volume and represent them by 3-D wavelet transform. Other published compressed sensing based video acquisition methods include the method based on coded apertures [9], the method based on motion estimation and compensation [10], the method based on frame difference [11] and the block-based method [12,13]. Although these CS-based methods can reduce the number of measurements required for reconstruction, the reconstruction is usually reliant on iterative algorithm to solve the optimization problems, suffering from long computation time, especially for video with a large amount of data. Recent developments get a quick preview with a few measurements in order to reduce the time for reconstruction [14,15], but they still cannot get results immediately. In this paper, we concentrate on single-pixel-based compressive video acquisition techniques that do not require computationally expensive reconstruction and are therefore able to provide near video frame rates.

Averbuch et al. proposed a novel single-pixel-based image sampling strategy based on wavelet trees and direct sampling in the wavelet domain, called adaptive compressed sampling (ACS) [16]. ACS determines the locations of significant coefficients using the parent–children relationship of wavelet trees and samples each of them with two DMD measurements. After all the

* Corresponding author.
E-mail addresses: gghnjust@mail.njust.edu.cn (G.-h. Gu), hewj@mail.njust.edu.cn (W.-j. He).

http://dx.doi.org/10.1016/j.optlastec.2016.01.021
0030-3992/

significant coefficients are sampled, images can be recovered by performing an inverse wavelet transform with low computation cost. The idea of adaptively determining the regions of significant coefficients based on wavelet trees was applied to computational ghost imaging by Aßmann and Bayer, and they proposed a method called compressed computational ghost imaging (CCGI) [17]. Inspried by CCGI, Yu et al. put forward a method named adaptive compressed ghost imaging (ACGI) [18], which also determines the significant regions based on wavelet trees, but samples them using sparse speckles. In our previous work, we extended the concept of wavelet trees by adding the sibling relationship, and proposed adaptive compressed sampling based on extended wavelet trees (EWT-ACS) [19]. Exploiting both parent–children and sibling relationships in extended wavelet trees, EWT-ACS improves the prediction accuracy of the location of significant children coefficients and therefore improves the image quality. Moreover, the number of measurements can be further reduced by digging out the redundancies in the measurements of sibling coefficients sampled previously, which makes it more suitable for single-pixel-based imaging.

In this work, we apply the idea of adaptively determining the regions of significant coefficients based on extended wavelet trees to video acquisition and propose a multiscale video acquisition framework named adaptive video compressed sampling (AVCS). The proposed method exploits sparse sampling and motion estimation in the wavelet domain to remove the spatial and temporal redundancies in video. Maintaining the advantages of EWT-ACS, our method avoids the computational overhead in CS-based methods and a video sequence can be reconstructed from compressed measurements in real time. In addition, it allows adaptive control over the reconstructed

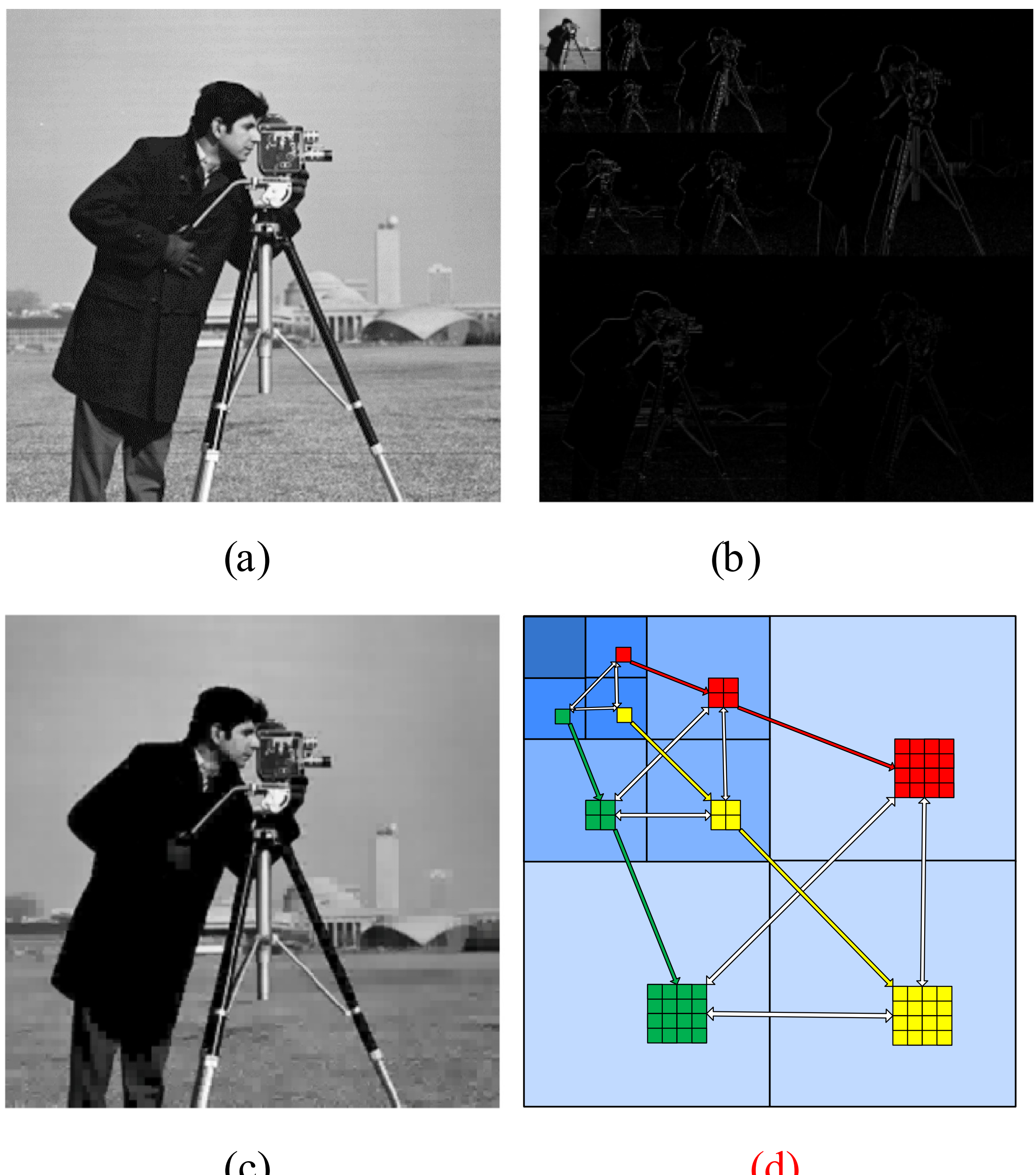

Fig. 1. Wavelet transform and the structure of extended wavelet trees. (a) A 512 × 512 pixels Cameraman image, together with (b) its sparse representation in Haar basis. (c) Compressed version with the largest 5% coefficients of (b), with a PSNR of 29.14 dB. (d) Structure of extended wavelet trees.