Neurocomputing 117 (2013) 128–137

Contents lists available at SciVerse ScienceDirect

Neurocomputing

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

Multi-frame image super-resolution reconstruction using sparse co-occurrence prior and sub-pixel registration

Beijia Ning, Xinbo Gao *

School of Electronic Engineering, Xidian University, No.2 South Taibai Road, Xi'an, 710071, China

ARTICLE INFO

Article history:
Received 15 August 2012
Received in revised form
18 January 2013
Accepted 21 January 2013
Communicated by Pingkun Yan
Available online 26 February 2013

Keywords:
Super-resolution reconstruction
Sparse representation
Sub-pixel registration
Sparse co-occurrence prior

ABSTRACT

In this paper, a hybrid super-resolution (SR) method is proposed by combining the concepts of both multi-frame and single-frame SR to generate a high-resolution (HR) image. The main contributions are in two aspects: the first one is hierarchical iterative sub-pixel registration, which provides accurate registration information of input low-resolution (LR) images or frames, to generate an initial HR image; the second one is to enhance the initial HR image with sparse co-occurrence prior, resulted from specially-designed dictionaries containing patches from both generic training images and interpolated input LR images. As a whole, the proposed hybrid SR method makes use of information from both sub-pixel registration and sparse co-occurrence prior to get reconstructed SR image with large zoom-in factor. The simulation results from synthetic images and real video frames illustrate its effectiveness and the superiority in image quality over conventional multi-frame and single-frame SR methods.

1. Introduction

Multi-frame image super-resolution (SR) is a process of generating a high resolution (HR) image from multiple low-resolution (LR) images, or video frames, from the same scene. Conventional multi-frame image SR methods normally consist of two major modules: sub-pixel motion estimation, and image fusion or interpolation. Sub-pixel registration is performed at first and then pixels from LR frames are aligned to generate an HR image [1,2]. As sub-pixel registration is the preparatory step to the SR reconstruction, some methods were proposed from different viewpoint to obtain acceptable sub-pixel shifts estimation, which include block matching [3], phase correlation [4], gradient-based method [5], PCA-based method [6] and mutual-information-based method [7]. Moreover, some methods were also proposed including simultaneous sub-pixel registration and SR reconstruction, such as maximum a posteriori probability (MAP) estimator and maximum likelihood (ML) estimator under a Bayesian framework from statistic viewpoint [1,8,9], and iterative back-projection (IBP) method [10]. Furthermore, video frame can also be utilized to reconstruct SR image [11]. Due to the ill-posedness of the SR problem, those factors such as the number of available LR images, the sizes of LR and HR images, the characteristics of images and the accuracy of sub-pixel shifts estimation etc., will affect the results of SR reconstruction. Thus regularization term was adopted into the reconstruction procedure to constrain the solution space and make the inverse problem well-posed. Such methods include projection onto convex sets (POCS) method [12] and many other regularized SR methods [13]. Nevertheless, the results of multi-frame SR reconstruction will be far less than satisfactory if only a limited number of LR images are available, or if a large image zoom-in factor is desired.

In recent years, much attention has been drawn to single-frame image SR, which can be regarded as an extreme case of multi-frame super-resolution with only one single LR image involved. Methods of single-frame SR approach are based upon machine learning, which attempt to capture the co-occurrence prior between LR and HR image patches. An example-based learning strategy was proposed [14], which is applied to generic images where the LR to HR relationship is learned via a Markov random field (MRF) solved by belief propagation. Normally, example-based or learning-based SR methods require training data which contain millions of LR and HR patch pairs. This makes the methods computationally expensive. Motivated by the theory of sparse representation from the concepts of compressed sensing [15], which suggest that the linear relationships among high-dimensional signals can be accurately recovered from their low-dimensional projections, a learning-based algorithm based on sparse coding was proposed [16], which sets up sparse representation between LR and HR image patches with dictionaries containing randomly-chosen raw patches from training images of similar statistical nature to the input image. Moreover, some researchers theoretically interpreted sparse coding as linear

* Corresponding author. Tel.: +86 29 88202845; fax: +86 29 88201620.
E-mail addresses: xbgao.xidian@gmail.com, xbgao@ieee.org, xbgao@mail.xidian.edu.cn (X. Gao).

0925-2312/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.neucom.2013.01.019

regression and proposed L_2-Boosting framework [17] and local-learning-based algorithm [18]. Also, some researcher worked on elaborate selection of training images which leads to multi-task dictionary buildup [19] and others on continuous scales of resolutions enhancement [20].

To increase the accuracy of sub-pixel shifts estimation and reduce the limitation on the performance of multi-frame SR reconstruction from the number of available LR images, the concepts of multi-frame and single-frame SR are unified and a hybrid multi-frame SR reconstruction method is proposed in this paper. First, from the viewpoint of multi-frame SR, hierarchical iterative sub-pixel registration algorithm is used to obtain shifts estimation with sufficiently good accuracy to generate an initial HR image. Then, from the viewpoint of single-frame SR, the obtained initial HR image is further enhanced on the co-occurrence prior obtained from the corresponding LR and HR patch pairs selected from dictionaries containing both generic and interpolated input LR images. So the proposed hybrid SR method incorporates advantages from not only highly accurate sub-pixel registration of multi-frame SR, but also co-occurrence prior of patch pairs from specifically designed dictionaries of single-frame SR.

The remainder of this paper is organized as follow. In Section 2, the system framework of the proposed two-stage hybrid multi-frame SR reconstruction method is presented; Section 3 explains the multi-frame SR part of the proposed method, i.e., the proposed hierarchical iterative sub-pixel registration algorithm. And Section 4 discusses the single-frame SR part of the proposed method, i.e., sparse-representation-based SR reconstruction method, especially the elaborate setup of dictionaries. Section 5 gives the simulation results from both still test images and real video, including the results from the proposed hierarchical iterative sub-pixel registration algorithm and their comparison with the results from conventional sub-pixel algorithms, the results of the proposed sparse-representation-based SR reconstruction method and their comparison with the results from other benchmark algorithms. Also, the overall system performance of the proposed hybrid multi-frame SR method is given. Finally, conclusions are drawn in Section 6.

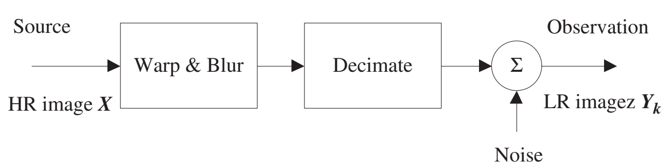

Fig. 1. Image degradation model.

2. The proposed hybrid multi-frame super-resolution method

The concept of multi-frame SR reconstruction is to fuse several observed LR images, or frames into one single HR image. The general observation model of imaging system is shown in Fig. 1.

The multiple observed LR images denoted by $\boldsymbol{Y_k}$ are used to generate a single HR image, denoted by $\boldsymbol{X}'$, which is an estimation of the original HR image $\boldsymbol{X}$. The relationship among HR image $\boldsymbol{X}$ and LR images $\boldsymbol{Y_k}$ can be modeled as

$$\boldsymbol{Y_k} = \boldsymbol{D_k H_k F_k X} + \boldsymbol{V_k} \quad k = 1,2,\ldots,N \tag{1}$$

where $\boldsymbol{D_k}$ stands for the decimation process. $\boldsymbol{H_k}$ is the blurring matrix. $\boldsymbol{F_k}$ denotes the geometric warp process and $\boldsymbol{V_k}$ represents Gaussian additive noise.

The process of traditional multi-frame SR reconstruction is reversed to that of imaging system shown in Fig. 1. And the system framework of the proposed hybrid multi-frame SR method is illustrated in Fig. 2. The whole framework contains two stages: multi-frame SR stage and single-frame SR one. In the multi-frame SR stage, the same as in the traditional process of multi-frame SR, sub-pixel shifts between input LR frames and reference frame are obtained at first, by means of the proposed hierarchical iterative sub-pixel registration method. Then, with the consideration of the limited available number of input LR frames and the random shifts between them, the fusion of LR frames in this paper is accomplished by the method of TDLI [2]. Next, to further improve the quality of that obtained HR image, in the single-frame SR stage, the concept of sparse representation is introduced, leading to the proposed sparse-representation-based SR method, including specially-designed HR and LR dictionaries. In this stage, the improvement comes from the fact that there exists co-occurrence prior between HR and LR image patches in the dictionaries which contain both generic training images and interpolated input LR ones.

3. Hierarchical iterative sub-pixel registration algorithm

The general criterion of an image registration method can be expressed as

$$s^* = \underset{s}{\mathrm{Argmax}}\, S\{\boldsymbol{R}[\boldsymbol{T_s}(x)], \boldsymbol{F}(\boldsymbol{x})\} \tag{2}$$

where $\boldsymbol{R}(\boldsymbol{x})$, $\boldsymbol{F}(\boldsymbol{x})$ are reference frame and current frame, respectively with pixel coordinate of $\boldsymbol{x}$. $\boldsymbol{T_s}$ stands for the coordinate transformation, characterized by the parameter $\boldsymbol{s}$. S is an similarity measure calculated over the overlapping region of the two frames or images. It shows that $\boldsymbol{R}(\boldsymbol{x})$ and $\boldsymbol{F}(\boldsymbol{x})$ will be registered when the similarity measure is maximized and the parameter $\boldsymbol{s}$

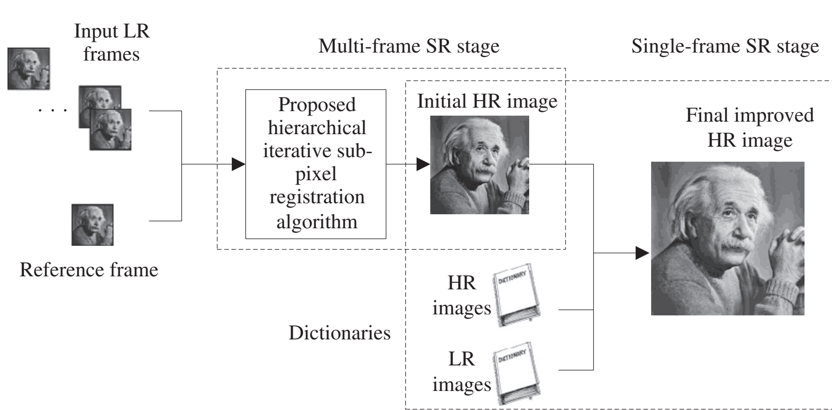

Fig. 2. Framework of the proposed hybrid multi-frame SR method.