Digital Signal Processing 50 (2016) 180–190

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

Compressive sensing super resolution from multiple observations with application to passive millimeter wave images

Wael AlSaafin [a,*], Salvador Villena [b], Miguel Vega [b], Rafael Molina [a], Aggelos K. Katsaggelos [c]

[a] Dept. of Computer Science and Artificial Intelligence, University of Granada, Spain
[b] Dept. of Languages and Information Systems, University of Granada, Spain
[c] Dept. of Electrical Engineering and Computer Science, Northwestern University, USA

ARTICLE INFO

Article history:
Available online 5 January 2016

Keywords:
Super resolution
Compressed sensing
Image reconstruction
Passive millimeter wave images

ABSTRACT

In this work we propose a novel framework to obtain high resolution images from compressed sensing imaging systems capturing multiple low resolution images of the same scene. The proposed approach of Compressed Sensing Super Resolution (CSSR), combines existing compressed sensing reconstruction algorithms with a low-resolution to high-resolution approach based on the use of a super Gaussian regularization term. The reconstruction alternates between compressed sensing reconstruction and super resolution reconstruction, including registration parameter estimation. The image estimation subproblem is solved using majorization-minimization while the compressed sensing reconstruction becomes an l_1-minimization subject to a quadratic constraint. The performed experiments on grayscale and synthetically compressed real millimeter wave images, demonstrate the capability of the proposed framework to provide very good quality super resolved images from multiple low resolution compressed acquisitions.

1. Introduction

Compressed Sensing (CS) theory offers a framework to simultaneously sense and compress signals. It establishes that a sparsely representable image/signal can be recovered from a highly incomplete set of measurements [1–3].

The design of CS image/video cameras (see [4–8]) has fostered the application of typical image processing tasks to CS observed images. CS has been applied to areas like radar analysis, face recognition, biomedical imaging, and microscopy imaging techniques [2,9,10], among others. CS measurements have also been used to recover images observed through unknown blur [11,12].

Super Resolution (SR) from a single image has also benefited from the introduction of CS theory. In [13,14] learning based SR is used to estimate a High Resolution (HR) image from the CS observation of a downsampled remote sensing image. In [15] the downsampling is incorporated in the measurement matrix, the CS image is reconstructed in the wavelet domain and the signal is deconvolved in the Fourier domain.

The recovery of an HR image from a set of unregistered LR CS observations has been scarcely addressed in the literature. To the best of our knowledge, the only reported works treating this general SR problem are [16,17]. In these papers CS and LR to HR techniques are coupled, using a fast and simple registration method, which uses the reconstructed HR images instead of the LR ones [17]. A non-robust prior model on the original image to be reconstructed was used in both papers.

This paper also deals with the reconstruction of an HR image from a group of LR CS observed images. The proposed method assumes that the HR image to be estimated is compressible and, consequently, its warped, blurred, and downsampled versions are also compressible (see [11,12]). They can then be reconstructed from their CS observations. However, instead of first recovering the LR observations and then using LR to HR techniques we propose a combined framework where LR reconstructions and HR estimation are carried out simultaneously. This proposed method is based on a sound and well founded method to estimate registration parameters in LR to HR problems and the use of a new robust sparse promoting prior for the original image.

The proposed framework has been tested, in the experimental section, on CS grayscale and Passive Millimeter Wave (PMMW) images. Without using CS measurements, the improvement of Passive Millimeter Wave (PMMW) images to perform detection tasks has

* Corresponding author.
E-mail addresses: waelsaafin@hotmail.com (W. AlSaafin), svillena@ugr.es (S. Villena), mvega@ugr.es (M. Vega), rms@decsai.ugr.es (R. Molina), aggk@eecs.northwestern.edu (A.K. Katsaggelos).

http://dx.doi.org/10.1016/j.dsp.2015.12.005

been addressed in [18–22], and the use of CS techniques to reduce the time needed to capture such images has been addressed in [5,6,23,24]. In [17], high resolution images were super resolved for the first time, from multiple CS observations of unregistered LR PMMW images. We believe that PMMW images represent an important application area where CS and LR to HR techniques can be combined to enhance the detection capabilities of current PMMWI systems.

Before going into details, the more frequently used notation in the paper is listed next

$\mathbf{y}_q$	$M \times 1$ compressed observation vector $q \in \{1, \cdots, Q\}$
$\mathbf{\Phi}$	$M \times D$ CS measurement matrix
$\mathbf{z}_q$	$D \times 1$ the q-th LR image vector
$\mathbf{r}_q$	$D \times 1$ CS observation noise vector
$\mathbf{A}$	$D \times N$ down-sampling matrix
P	zooming factor
$\mathbf{H}_q$	$N \times N$ blurring matrix
$\mathbf{C}(\mathbf{s}_q)$	$N \times N$ warping matrix formed by motion vector $\mathbf{s}_q$
$\mathbf{s}_q$	3×1 motion vector (rotation θ_q, horizontal c_q, and vertical d_q displacements)
$\mathbf{x}$	$N \times 1$ HR image vector
$\mathbf{w}_q$	$D \times 1$ HR to LR acquisition noise vector
$\mathbf{B}_q(\mathbf{s}_q)$	$D \times N$ LR acquisition model matrix
$a_q(\mathbf{s}_q), b_q(\mathbf{s}_q)$	$N \times 1$ pixel difference vectors
$\mathbf{D}_{a_q(\mathbf{s}_q)}$	$N \times N$ diagonal matrix with $a_q(\mathbf{s}_q)$ in the diagonal $\mathbf{D}_{b_q(\mathbf{s}_q)} : b_q(\mathbf{s}_q)$ in the diagonal
$\mathbf{I}$	the identity matrix
$\mathbf{L}_{\mathbf{bl}(\mathbf{s}_q)}$	bottom-left-pixel matrix ($\mathbf{br}, \mathbf{tl}, \mathbf{tr}$): (bottom-right, top-left, top-right)
$\mathbf{n}_q$	$M \times 1$ combined CS and LR acquisition noise vector
$\mathbf{W}$	$D \times D$ transformation matrix
$\mathbf{a}_q$	$D \times 1$ transformation coefficient vector
$\alpha, \beta, \eta, \tau$	non-negative parameters
$\mathbf{Q}(\mathbf{x})$	regularization term
$\omega_d^{\mathbf{x}}$	filtered image applying $\mathbf{F}_d$ on $\mathbf{x}$ in d-direction
λ_q	$D \times 1$ Lagrangian multiplier vector

The rest of this paper is organized as follows. The problem modeling and its formulation as an optimization task are presented in Sections 2 and 3, respectively. The estimation process is described in Section 4. We demonstrate the effectiveness of the proposed method in the experimental section, Section 5. Finally, conclusions are drawn in Section 6.

2. Modeling

In this work we assume that we have access to a set of Q CS LR observations of the form

$$\mathbf{y}_q = \mathbf{\Phi} \mathbf{z}_q + \mathbf{r}_q \quad q = 1, \ldots, Q, \tag{1}$$

where $\mathbf{y}_q$ is an $M \times 1$ vector representing compressed observations from the LR image $\mathbf{z}_q$, $\mathbf{\Phi}$ is a CS $M \times D$ measurement matrix, $\mathbf{z}_q$ is a column vector of size $D \times 1$ representing the q-th LR image and $\mathbf{r}_q$ represents the observation noise. We denote by R the compression ratio of the measurement system, that is $\mathrm{R} = M/D$, $R \leq 1$. The sensing matrix $\mathbf{\Phi}$ consists of either real or binary entries. The matrix used in our work is binary, since it is easier to be synthesized physically [6,23,24]. In both cases the rows/columns of $\mathbf{\Phi}$ are normalized to 1. We assume that the LR observations $\mathbf{z}_q$ are related to an HR image of size N, denoted by the column vector $\mathbf{x}$ by

$$\mathbf{z}_q = \mathbf{A}\mathbf{H}_q \mathbf{C}(\mathbf{s}_q)\mathbf{x} + \mathbf{w}_q = \mathbf{B}_q(\mathbf{s}_q)\mathbf{x} + \mathbf{w}_q, \tag{2}$$

where $\mathbf{A}$ is a $D \times N$ down-sampling matrix, $D \leq N$, which models the limited resolution of the acquisition system, when capturing the high resolution image, where $N = P^2 D$ and $\mathrm{P} \geq 1$ is the zooming factor, in each dimension of the image. $\mathbf{H}_q$ is an $N \times N$ blurring matrix, modeling the action accompanying the imaging process. In this work, $\mathbf{H}_q$ is assumed to be known. $\mathbf{C}(\mathbf{s}_q)$ is the $N \times N$ warping matrix formed by motion vector $\mathbf{s}_q = [\theta_q, c_q, d_q]^t$, where θ_q is the rotation angle, and c_q and d_q are respectively the horizontal and vertical translations of the q-th LR image with respect to the reference frame. Finally, $\mathbf{w}_q$ models the noise associated to the HR to LR acquisition process. We write $\mathbf{B}_q(\mathbf{s}_q) = \mathbf{A}\mathbf{H}_q\mathbf{C}(\mathbf{s}_q)$ for simplicity.

As explained in [25], matrices $\mathbf{C}(\mathbf{s}_q)$ can be explicitly stated as follows. Let us denote the coordinates of the reference HR grid by (u, v) and the coordinates of the q^{th} warped HR grid, after applying $\mathbf{C}(\mathbf{s}_q)$ to $\mathbf{x}$, by (u_q, v_q). Then it holds that

$$u_q = u\cos(\theta_q) - v\sin(\theta_q) + c_q \tag{3}$$

$$v_q = u\sin(\theta_q) + v\cos(\theta_q) + d_q\,. \tag{4}$$

Let us denote the displacements between the grids by $\Delta(u_q, v_q)^T = (u, v)^T - (u_q, v_q)^T$. The vector difference between the pixel at (u_q, v_q) and the pixel at its top-left position in the reference HR grid is denoted by $(a_q(\mathbf{s}_q), b_q(\mathbf{s}_q))^T$ (see Fig. 1), that is,

$$a_q(\mathbf{s}_q) = \Delta u_q - \mathrm{floor}(\Delta u_q)\,, \tag{5}$$

$$b_q(\mathbf{s}_q) = \Delta v_q - \mathrm{floor}(\Delta v_q)\,. \tag{6}$$

Using bilinear interpolation, the warped image $\mathbf{C}(\mathbf{s}_q)\mathbf{x}$ can be approximated as

$$\begin{aligned}\mathbf{C}(\mathbf{s}_q)\mathbf{x} \approx\ & \mathbf{D}_{b_q(\mathbf{s}_q)}(\mathbf{I} - \mathbf{D}_{a_q(\mathbf{s}_q)})\mathbf{L}_{\mathbf{bl}(\mathbf{s}_q)}\mathbf{x} + (\mathbf{I} - \mathbf{D}_{b_q(\mathbf{s}_q)})\mathbf{D}_{a_q(\mathbf{s}_q)}\mathbf{L}_{\mathbf{tr}(\mathbf{s}_q)}\mathbf{x} \\ & + (\mathbf{I} - \mathbf{D}_{b_q(\mathbf{s}_q)})(\mathbf{I} - \mathbf{D}_{a_q(\mathbf{s}_q)})\mathbf{L}_{\mathbf{tl}(\mathbf{s}_q)}\mathbf{x} \\ & + \mathbf{D}_{b_q(\mathbf{s}_q)}\mathbf{D}_{a_q(\mathbf{s}_q)}\mathbf{L}_{\mathbf{br}(\mathbf{s}_q)}\mathbf{x}\,,\end{aligned} \tag{7}$$

where $\mathbf{D}_{a_q(\mathbf{s}_q)}$ and $\mathbf{D}_{b_q(\mathbf{s}_q)}$ denote diagonal matrices with the vectors $a_q(\mathbf{s}_q)$ and $b_q(\mathbf{s}_q)$ in their diagonals, respectively. $\mathbf{I}$ is the identity matrix. Matrices $L_{\mathbf{z}}$ with $\mathbf{z} \in \{\mathbf{bl}(\mathbf{s}_q), \mathbf{br}(\mathbf{s}_q), \mathbf{tl}(\mathbf{s}_q), \mathbf{tr}(\mathbf{s}_q)\}$ are constructed in such a way that the product $L_{\mathbf{z}}\mathbf{x}$ produces pixels at the bottom-left, bottom-right, top-left, and top-right, locations of (u_q, v_q), respectively.

Using (1) and (2) we can write

$$\mathbf{y}_q = \mathbf{\Phi}\mathbf{B}_q(\mathbf{s}_q)\mathbf{x} + \mathbf{n}_q\,, \text{ for } q = 1, \ldots, Q\,, \tag{8}$$

where $\mathbf{n}_q$ represents the combined CS and LR acquisition noise and $\mathbf{x}$ is the HR image we want to estimate.

3. Problem formulation

Since $\mathbf{z}_q$ in (2) represents translated and rotated LR versions of the original image $\mathbf{x}$ (which are assumed to be compressible in a transformed domain), we can estimate the original HR image by first recovering the LR images using CS techniques and then recover the HR image using standard super resolution techniques on the recovered low resolution images. To be precise, if we assume that the LR images are sparse in a transformed domain with $\mathbf{z}_q = \mathbf{W}\mathbf{a}_q$, where $\mathbf{W}$ is a sparse promoting transformation of size $D \times D$, we can recover them from the model in (1) by solving

$$\hat{\mathbf{a}}_q = \arg\min_{\mathbf{a}_q} \frac{\eta}{2} \parallel \mathbf{\Phi}\mathbf{W}\mathbf{a}_q - \mathbf{y}_q \parallel^2 + \tau \parallel \mathbf{a}_q \parallel_1, \tag{9}$$

where η, τ are regularization parameters, $\parallel . \parallel$ is the Euclidean norm, and $\parallel . \parallel_1$ the $\ell 1$ norm. Then defining $\hat{\mathbf{z}}_q = \mathbf{W}\hat{\mathbf{a}}_q$ and $\mathbf{s} = (\mathbf{s}_1, \ldots, \mathbf{s}_Q)$ and using the degradation model (2), we can estimate the original image by solving

$$\hat{\mathbf{x}}, \hat{\mathbf{s}} = \arg\min_{\mathbf{x},\mathbf{s}} \frac{\beta}{2} \sum_q \parallel \mathbf{B}_q(\mathbf{s}_q)\mathbf{x} - \hat{\mathbf{z}}_q \parallel^2 + \alpha\, \mathbf{Q}(\mathbf{x}), \tag{10}$$