Computationally Efficient Image Super Resolution from Totally Aliased Low Resolution Images

A Anil Kumar, N Narendra, P Balamuralidhar and M Girish Chandra
TCS Research and Innovation, Bangalore, India.
Email:{achannaanil.kumar,n.narendra, balamurali.p, m.gchandra}@tcs.com.

***Abstract*—This paper considers the problem of super-resolution (SR) image reconstruction from a set of *totally aliased* low resolution (LR) images with different unknown sub-pixel offsets. By assuming the translational motion model, a linear compact representation between the LR image spectrums and SR image spectrum, based on multi-coset sampling is provided. Based on this model, we formulate the joint estimation of the unknown shifts and SR image spectrum as a dictionary learning problem and alternating minimization approach is employed to solve this joint estimation. Two different approaches for obtaining the SR image; one based on estimated shifts and another based on estimate SR spectrum are described. The significant advantage of the proposed approach is the smaller matrix sizes to be handled during the computation; typically on the order of number of images and enhancement factors, and is completely independent on the actual dimensions of LR and SR images, hence requiring significantly lesser resources than the current state of the art approaches. Brief simulation results are also provided to demonstrate the efficacy of this approach.**

I. Introduction

Aliasing, is a well known phenomenon caused due to sampling of signals below the Nyquist sampling rate and is usually considered as nuisance. However, aliased signals contains information about high frequency components and reconstructing the signal by resolving these high frequency components from multiple, slightly different aliased signals (also referred as low resolution (LR) signal) has many applications. One such key application being the image super-resolution (SR) which is the topic considered in this paper. Thus, a SR image is not just merely an upsampled and interpolated image, but it also contains additional details due to the incorporation of high frequency information.

Image SR finds applications in many areas such as in satellite and aerial imaging, medical imaging etc., due to which it is a well researched area and plethora of algorithms exist. Interested readers may refer to [1] for a comprehensive overview of the existing SR techniques and their applications. Essentially, the SR process comprises of two main steps: i) Registration and ii) Restoration. Registration, which is a process of transforming the images onto a common co-ordinate system at sub-pixel precision, forms the crucial step. This is because the performance of the subsequent restoration step, which is an ill-posed problem, heavily depends upon the registration accuracy. Hence, wide range of strategies exist for image registration in both the spatial and frequency domains. A summary of these registration techniques can be found in the survey articles [2]. Alternatively, iterative methods for joint registration and restoration have also been proposed such as in [3], and is shown to perform better than the disjoint two-stage SR methods; however at the cost of increased computation.

Of-late there has been a increasing trend of removing the front end anti-aliasing filter in cameras to increase the image sharpness (refer [4] for more details). Removal of this filter causes total aliasing of the LR image spectrum. Although, as outlined earlier, wide range of techniques exists, most of them considers *aliasing* as a background noise and hence the SR reconstruction performance is quiet low for this case of total aliasing. Hence, to improve the performance under aliasing, [5], [6] proposed SR techniques by including aliasing in the signal model. Whereas, [5] addressed a scenario when an aliasing-free part of the spectrum is available, [6] considered a more generic scenario when the entire signal is totally aliased. While [6] can handle total aliasing, as shall be shown later in Section II-B, this technique is computationally highly intensive as the signal is processed by vectorizing the LR images which leads to larger matrix dimensions to be handled.

In this paper, we propose a computationally efficient method for obtaining a SR image from multiple LR images that have slightly different sub-pixel offsets and which is capable of SR image reconstruction under total aliasing. The proposed approach is essentially based on multi-coset sampling framework [7], which facilitates band-wise aliasing representation in contrast to bin-wise aliasing such as in [6], thereby enabling a compact linear relationship between multiple LR images and the SR image. The relationship matrix has a Vandermonde structure and is dependent on the unknown shifts. Now, since both the offsets and the SR image are unknown, by making use of sparsity in the frequency spectrum of most of the natural images, we frame the registration step as dictionary learning problem [8] and propose to solve it using alternative minimization framework [9]. While in the dictionary updation step, we only update the parameters (here shifts) using the steepest descent iterations, in the signal estimation step we employ MUlitple Signal Classification (MUSIC) algorithm [10] to estimate the sparse spectrum. We describe two methods for obtaining the desired SR image in the subsequent restoration step; one based on the estimated sub-pixel offsets and another based on the estimated sparse frequency spectrum. Heuristic approaches for obtaining faster convergence and to improve the sub-pixel estimation accuracy are also provided. The main advantage of the proposed approach is the huge reduction in the dimension of the matrices; unlike [5], [6], the matrix dimensions are independent on the size of the LR and SR images and depends only on the number of LR images and enhancement factor. Thus, even with nominal computational resources one can obtain larger dimensional SR image. Brief simulation results are also presented to demonstrate the efficacy of the SR image reconstruction from totally aliased set of LR images and to study the sub-pixel offset estimation performance.

II. Preliminaries

In this section, we first describe the LR image model considered in this paper. This is then followed by a brief outline of the existing techniques for SR and also discuss their limitations, thus providing a motivation for the alternative approach presented in this paper.

ISBN 978-90-827970-1-5

A. LR Image Model and Problem Formulation

We restrict the motion model to translational shift model, such as in [6], [11], [12]. This kind of motion model although is applicable in scenarios where burst of images are taken in a short interval of time with small motion between the images, but is more suitable when a camera array is employed to capture LR images which is becoming more popular of-late [13], [14].

Now, the k^{th} $(1 \leq k \leq K)$ LR image $g_k(\mathbf{n})$ of size $N_x \times N_y$ can be modeled as

$$g_k(\mathbf{n}) = h(\mathbf{L}\mathbf{n} + \mathbf{c}_k) + \eta(\mathbf{n}) \tag{1}$$

where $h(\mathbf{t})$, $\mathbf{t} \in \mathbb{R}^2$ denotes the 2-D image scene , $\{ \mathbf{n} = [n_x, n_y]^T \mid \mathbf{n} \in \mathbb{Z}^2,\ 0 \leq n_x \leq N_x - 1,\ 0 \leq n_y \leq N_y - 1\}$, $\mathbf{c}_k = [c_{xk}, c_{yk}]^T$. $\mathbf{L} = \begin{pmatrix} L_x & 0 \\ 0 & L_y \end{pmatrix}$, L_x, L_y denotes the enhancement factors along the x-axis and y-axis respectively (the factors L_x, L_y are also referred to as decimation factors in the sampling literature). c_{xk} and c_{yk} which resides in the range $0 < c_{xk} < L_x$, $0 < c_{yk} < L_y$ denotes the relative translational shifts of the k^{th} LR image along x-axis and y-axis respectively and $\eta(\mathbf{n})$ denotes the additive white noise. Here, we assume the first image as the reference image and hence $c_{xk} = c_{yk} = 0$ for $k = 1$. For the sake of simplicity, we further make an assumption that the shifts are distinct and it is important to note that it resides in the sub-pixel range with respect to the LR image i.e., $c_{\theta k}/L_\theta < 1$, for $\theta \in \{x, y\}$. We refer to the ratio $c_{\theta k}/L_\theta$ as *sub-pixel offset*. Additionally, we make an assumption that the enhancement factor can exceed the number of LR images i.e., $K < L_x L_y$.

With the above mentioned LR image model, the aim here is to estimate the SR image $h(\mathbf{m})$, $\{\mathbf{m} = [m_x, m_y]^T \mid 0 < m_x < L_x n_x, 0 < m_y < L_y n_y\}$ from the K known LR images $g_k(\mathbf{n})$ having unknown shifts $\mathbf{c}_k$. Observe that in order to properly reconstruct the SR image, it is important and essential to estimate shifts with relatively good accuracy. It is to be noted from (1) that not only LR images are aliased but also due to the absence of anti-aliasing filter, the entire LR image spectrum is *totally* aliased. It is thus necessary to estimate the shifts which are in the sub-pixel range (with respect to the LR image scale) from such totally aliased spectrum. In the following section we briefly describe the existing state of the art techniques for SR image reconstruction for the aforementioned LR image model and also briefly outline their limitations.

B. Existing techniques

It is well known that the translation in time/space corresponds to simple linear phase shift in the frequency domain. While in the absence of aliasing, this relative phase shift can easily and accurately be measured, in the presence of aliasing the estimation will be erroneous. Hence, many registration methods such as Normalized Cross Power Spectrum (NCPS) methods (see [12] and the references there-in) results in erroneous estimation when employed for totally aliased LR images. [5] suggested a simple scheme of considering only the alias-free band for estimation and neglecting the other regions that are affected by aliasing. However, for the case when the entire band is affected, as in the present scenario, the same authors in [6] suggested a scheme by taking aliasing into the image model. A subspace based algorithm was also described by concatenating all the vectorized LR images. Although this method works under total aliasing, the computational complexity is enormously high which is also observed by the authors (see [6, Section VII])). For example, the dimensions of the matrix to be handled will be in the order of $N_x N_y \times L_x N_x L_y N_y$, which one can notice, that it is very large even for a nominal LR and SR image sizes. Furthermore, due to the iterative nature of the algorithm, these large dimensional matrices must be handled at each iteration.

Owing to these limitations, in the next section we present an alternative computationally efficient approach. Unlike [6], the dimensions of the matrices to be handled depends only on K and $L_x L_y$ and is independent of the sizes of the LR and SR images.

III. Proposed approach

In this section, we first provide a brief representation of LR image model of (1) in the frequency domain by making certain assumptions on the offsets. Following this, methods for the two key steps of SR, namely registration and restoration are described. At the end of this section, we outline our complete SR image algorithm process using the above described individual blocks. While in this section, the algorithms are described, in the next section we outline heuristic approaches for further reducing the computations and also briefly discuss methods for choosing parameters for guaranteed reconstruction and for performance improvement.

A. Frequency domain representation

The shifts which reside in the range $0 < c_{\theta k} < L_\theta$, can be expressed as $c_{\theta k} = \text{round}(c_{\theta k}) + \Delta c_{\theta k}$. Suppose if we neglect $\Delta c_{\theta k}$ and make an approximation that $c_{\theta k} \approx \text{round}(c_{\theta k})$ i.e., shifts rounded off to closest integer, then the LR images can be seen as obtained via the multi-coset sampling [7] of the SR image and can easily be represented in the compact framework as [7],

$$\mathbf{G}(\mathbf{f}) = \frac{1}{L_x L_y} \underbrace{\Delta(\mathbf{f})\mathbf{A}}_{\mathbf{A}_\Delta(\mathbf{f})} \mathbf{s}(\mathbf{f}) \tag{2}$$

where $\mathbf{G}(\mathbf{f}) \triangleq [G_1(\mathbf{f}), G_2(\mathbf{f}), ..., G_K(\mathbf{f})]$, for any $1 \leq k \leq K$, $G_k(\mathbf{f})$ denotes the 2-D Fourier transform of the k^{th} LR image and $\mathbf{f} \in [0, 1]^2$. The k^{th} row, q^{th} column element of matrix $\mathbf{A}$ which is of size $K \times L_x L_y$, can be expressed as

$$[\mathbf{A}]_{k,q} = e^{j2\pi \mathbf{c}_k^T \mathbf{L}^{-1} \boldsymbol{\lambda}_q} \tag{3}$$

where $\boldsymbol{\lambda}_q$ denotes the q^{th} vector of the set $\Lambda \triangleq \{0, 1, ..., L_x - 1\} \times \{0, 1, ..., L_y - 1\}$ (here $\times$ denotes the Cartesian product). $\Delta(\mathbf{f})$ is a diagonal matrix with diagonal elements $\{e^{-j2\pi \mathbf{c}_1^T \mathbf{f}}, e^{-j2\pi \mathbf{c}_2^T \mathbf{f}}, ..., e^{-j2\pi \mathbf{c}_K^T \mathbf{f}}\}$ and $\mathbf{s}(\mathbf{f}) \triangleq [H(\mathbf{f} + \mathbf{L}^{-1}\boldsymbol{\lambda}_1), H(\mathbf{f} + \mathbf{L}^{-1}\boldsymbol{\lambda}_2), ..., H(\mathbf{f} + \mathbf{L}^{-1}\boldsymbol{\lambda}_Q)]^T$, where $Q = |\Lambda| = L_x L_y$. $H(\mathbf{f} + \mathbf{L}^{-1}\boldsymbol{\lambda}_k)$ for $\mathbf{f} \in [0, 1]^2$, represents a subband of $H(\mathbf{f})$ of dimension $N_x \times N_y$ beginning from the position $\begin{pmatrix} N_x & 0 \\ 0 & N_y \end{pmatrix} \boldsymbol{\lambda}_k$, where $H(\mathbf{f})$ denotes the 2-D Fourier transform of the the SR image $h(\mathbf{m})$ and is of size $N_x L_x \times N_y L_y$.[1] Thus, LR image spectrum $G_k(\mathbf{f})$ can be seen as weighted superposition of the subbands given by rows of $\mathbf{s}(\mathbf{f})$, with weights being a function of shifts. We shall make use of this property of different weights to obtain the SR image by overcoming aliasing.

[1] Interested reader may refer to [7, Fig. 3] for a pictorial representation of the $\mathbf{s}(\mathbf{f})$ for a 1-D signal.

ISBN 978-90-827970-1-5 © EURASIP 2018

It is important to notice that the dimension of $\mathbf{A}$ is $K \times L_x L_y$ (i.e., the number of LR images $\times$ number of sub-bands) and unlike [6], the size of the matrix is *independent* of the dimensions of the LR and HR images. Now, contrary to the standard multi-coset sampling and reconstruction of [7] where the shifts are known, here both $\mathbf{s}(\mathbf{f})$ and shifts $\{\mathbf{c}_k\}_{k=1}^K$ are unknown. Hence, in the next section, we propose an iterative algorithm based on alternating minimization technique to estimate both the matrix $\mathbf{A}_\Delta(\mathbf{f})$ as well as $\mathbf{s}(\mathbf{f})$.

B. Registration and restoration algorithm

The task of the registration algorithm is to essentially estimate the shifts, while the restoration algorithm is to reconstruct the final SR image of a predefined dimension by making use of these estimated shifts. The following sections describes the details of these methods.

1) Registration: The registration step essentially estimates $\hat{\mathbf{A}}_\Delta(\mathbf{f})$ and $\hat{\mathbf{s}}(\mathbf{f})$ such that the following function is minimized:

$$\min_{\{\mathbf{c}_k\}_{k=1}^K, \hat{\mathbf{s}}(\mathbf{f})} \|\mathbf{G}(\mathbf{f}) - \frac{1}{L_x L_y}\hat{\mathbf{A}}_\Delta(\mathbf{f})\hat{\mathbf{s}}(\mathbf{f})\|_F^2. \quad (4)$$

Now, recall from Section II that we have made an assumption that $K < L_x L_y$, and hence under this scenario, (2) becomes an under-determined system of equations. To handle this scenario, we make use of the inherent frequency domain sparsity of most of the natural images. In other words, for most of the natural images, several subbands i.e., rows of $\mathbf{s}(\mathbf{f})$ will have insignificant energy and thus can be assumed to be approximately sparse. We enforce this sparsity in our estimation algorithm and then visualize estimating $\mathbf{A}_\Delta(\mathbf{f})$ as a dictionary learning problem [8], [15]. Alternating minimization technique [9] is then employed to jointly estimate the dictionary or the matrix $\hat{\mathbf{A}}_\Delta(\mathbf{f})$ and sparse $\hat{\mathbf{s}}(\mathbf{f})$ as outlined below.

a) Updating dictionary: Unlike the conventional dictionary learning problem described in [8], [15], since the structure of the dictionary or matrix $\mathbf{A}_\Delta(\mathbf{f})$ is clearly defined, similar to [16] we enforce this structure. In other words, we only update the parameters $\{\mathbf{c}_k\}_{k=1}^K$ iteratively. At the beginning of the iteration, we initialize the shifts obtained using some simpler method such as NCPS method and by considering only the lower frequencies, where the effect of aliasing is less compared to other regions. We then use the steepest descent method for updating the parameters. Now, the mean squared error (MSE) at the end of the i^{th} steepest descent iteration can be expressed as

$$F(\mathbf{A}_\Delta^{(i)}) = \frac{1}{N_x N_y}\sum_{\mathbf{f}} \|\mathbf{G}(\mathbf{f}) - \frac{1}{L_x L_y}\hat{\mathbf{A}}_\Delta^{(i)}(\mathbf{f})\hat{\mathbf{s}}^{(t-1)}(\mathbf{f})\|_2^2 \quad (5)$$

where $\hat{\mathbf{A}}_\Delta^{(i)}(\mathbf{f})$ denotes the estimated dictionary at the end of the i^{th} iteration and $\hat{\mathbf{s}}^{(t-1)}(\mathbf{f})$ denotes the estimated signal at the $(t-1)^{th}$ alternating minimization iteration.[2] Using $F(\mathbf{A}_\Delta^{(i)})$, the shifts at the $(i+1)^{th}$ iteration are updated as

$$c_{\theta k}^{(i+1)} \leftarrow c_{\theta k}^{(i)} - \mu \frac{\partial F(\mathbf{A}_\Delta^{(i)})}{\partial c_{\theta k}} \quad (6)$$

where $1 \leq k \leq K$ and μ denotes the step size. The details of the computation of the partial derivatives are not given here due to lack of space. However, computation is relatively easy since only the k^{th} row of $\mathbf{A}_\Delta(\mathbf{f})$ depends upon $c_{\theta k}$. It is important to notice that unlike in the case of [16], due to diagonal matrix $\Delta(\mathbf{f})$, the dictionary is different for each frequency bin and this must be taken into consideration while computing the gradient. With fixed $\hat{\mathbf{s}}^{(t-1)}(\mathbf{f})$, the iterations are continued till convergence i.e., till $\|F(\mathbf{A}_\Delta^{(i+1)}) - F(\mathbf{A}_\Delta^{(i)})\|^2 < \epsilon_s$. At the end of this step, the updated dictionary $\hat{\mathbf{A}}_\Delta^{(t)}(\mathbf{f})$ at the t^{th} alternating minimization iteration is formed by knowing the parameters $\{\hat{\mathbf{c}}_k^{(t)}\}_{k=1}^K$.

b) Estimation of $\hat{\mathbf{s}}^{(t)}(\mathbf{f})$: It is easy to observe that with the assumption of sparsity on $\mathbf{s}(\mathbf{f})$ and an estimate of the dictionary $\mathbf{A}_\Delta^{(t)}(\mathbf{f})$ available, (2) now reduces to sparse signal estimation from multi-coset sampling which is considered in [7]. While we employ steps similar to [7], during the implementation we adopt MUSIC algorithm [10], which is an efficient and most commonly used algorithm in array processing. The covariance matrix at the t^{th} iteration is formed as

$$\mathbf{R}^{(t)} = \sum_{\mathbf{f}} (\hat{\Delta}^{(t)}(\mathbf{f}))^H \mathbf{G}(\mathbf{f})(\mathbf{G}(\mathbf{f}))^H \hat{\Delta}^{(t)}(\mathbf{f}). \quad (7)$$

The entire vector space of dimension K is divided into signal space Λ_s of dimension α_s that is spanned by the singular vectors of $\mathbf{R}^{(t)}$ corresponding to α_s significant singular values and the other orthogonal noise subspace Λ_n spanned by the remaining $K - \alpha_s$ singular vectors. The supports i.e., the columns of $\hat{\mathbf{A}}^{(t)}(\mathbf{f})$ corresponding to the most significant α_s rows of $\hat{\mathbf{s}}(\mathbf{f})$ is estimated by projecting the atoms (i.e., columns) of $\hat{\mathbf{A}}^{(t)}(\mathbf{f})$ onto the noise subspace Λ_n and choosing α_s columns corresponding to lower α_s projected values. Subsequently, upon knowing the supports, a least square solution is taken and $\hat{\mathbf{s}}^{(t)}(\mathbf{f})$ at the t^{th} iteration is estimated. Notice here that α_s can also be interpreted as the number of significant aliasing subbands.

At the end of this t^{th} alternating minimization iteration, MSE_g, where $\text{MSE}_g \triangleq 1/(N_x N_y)\sum_{\mathbf{f}} \|\hat{\mathbf{G}}^{(t)}(\mathbf{f}) - \mathbf{G}(\mathbf{f})\|_2^2$, $\hat{\mathbf{G}}^{(t)}(\mathbf{f}) = 1/(L_x L_y)\hat{\mathbf{A}}_\Delta^{(t)}(\mathbf{f})\hat{\mathbf{s}}^{(t)}(\mathbf{f})$ is computed and the process is repeated till for some small threshold say ϵ_g, $\text{MSE}_g < \epsilon_g$. Now, at the end of this registration step, due to integer approximation (see beginning of Section III-A) we will obtain an estimate of the shifts $\{\hat{c}_{\theta k}\}_{k=1}^K$ which are integers in the range $0 < \hat{c}_{\theta k} < L_\theta$. Later in Section IV-B we briefly discuss a method for obtaining more accurate estimates of the shifts by reducing the approximation error.

2) Restoration: By using the outcomes of the aforementioned registration step, one can think of two methods using which the final desired SR image can be obtained.

a) From $\hat{\mathbf{s}}(\mathbf{f})$: Recall from the above Section III-A that rows of $\hat{\mathbf{s}}(\mathbf{f})$ provides sub-bands of the SR image. By properly arranging $\hat{\mathbf{s}}(\mathbf{f})$, the SR image spectrum can be obtained and by taking inverse Fourier transform one can get the desired SR image. However, an additional step of resizing to the predefined desired size might be necessary, since the factors L_x and L_y would be adjusted to satisfy the conditions that are stated in Section IV-B.

b) Interpolation: The other alternative method of restoration is by using non-uniform interpolation. From the previous step, an estimate of the sub-pixel offsets $\{\hat{c}_{\theta k}/L_\theta\}_{k=1}^K$ will be available. A non-uniform bi-cubic interpolation (as employed in [12], [5], [6]) can be used by placing the LR images at

[2] While the superscipt i indicates the i^{th} iteration of the steepest descent algorithm, t denotes the t^{th} iteration of the alternative minimization algorithm.

ISBN 978-90-827970-1-5 © EURASIP 2018

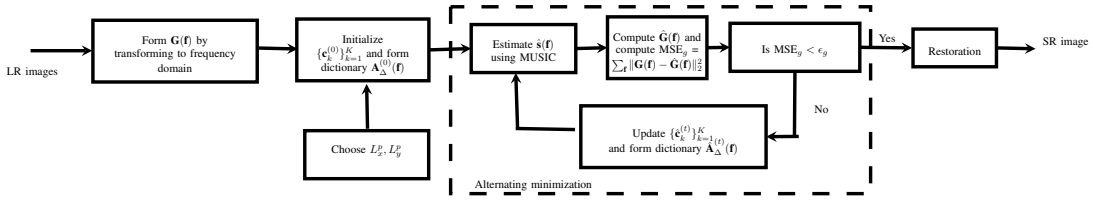


Fig. 1. Flow diagram of the proposed approach. The enhancement factors L_x and L_y are chosen based on the required precision of the offsets. Alternate minimization technique is then employed to estimate the offsets and $\hat{\mathbf{s}}(\mathbf{f})$, finally using these estimated quantities an SR image is reconstructed in the restoration step.

appropriate positions using the sub-pixel offsets on the HR grid of a desired resolution and can get the desired SR image.

C. SR image reconstruction: Summary

Fig. 1 depicts the proposed SR image reconstruction process from a set of LR images. The proposed scheme can be divided into following three main steps.
1) *Initialization:* In this step, $\mathbf{G}(\mathbf{f})$ is formed by taking the 2-D FFT of the LR images. The factors L_x and L_y is decided based on the conditions that are stated in Section IV-B. By employing an algorithm such as NCPS method, obtain an initial approximate estimate of the shifts $\{\mathbf{c}_k^{(0)}\}_{k=1}^K$.
2) *Registration:* Alternating minimization technique is employed to jointly estimate the unknowns $\hat{\mathbf{A}}_\Delta(\mathbf{f})$ as well as $\hat{\mathbf{s}}(\mathbf{f})$.
3) *Restoration:* Upon knowing $\hat{\mathbf{s}}(\mathbf{f})$ and sub-pixel offsets $\{\hat{c}_{\theta k}/L_\theta\}_{k=1}^K$ from the previous registration step, either of the aforementioned methods can be used for obtaining the desired SR image.

IV. Heuristics and choice of L_x, L_y

A. Heuristic approach

Similar to any realization of steepest descent method, the performance usually depends upon the factors such as step size i.e., μ, initialization i.e., $\{\mathbf{c}_k^{(0)}\}_{k=1}^K$ etc. Recall from Section III-A that we have made an approximation of shifts to the closest integer for (2) to hold. Hence instead of updating the shifts in smaller increments using (6), we can update to the neighboring integer depending upon the direction of the gradient. Formally, we update the shifts as shown below

$$c_{\theta k}^{(i+1)} \leftarrow \begin{cases} c_{\theta k}^{(i)} - \text{sign}(\frac{\partial F(\mathbf{A}_\Delta)}{\partial c_{\theta k}}) & 0 < c_{\theta k}^{(i)} < L_\theta, \\ & \text{and } |(\frac{\partial F(\mathbf{A}_\Delta)}{\partial c_{\theta k}})| > \epsilon_h \\ c_{\theta k}^{(i)} & \text{Otherwise} \end{cases} \quad (8)$$

where $\text{sign}(\beta) = +1$ if $\beta \geq 0$, else it is -1. The condition $|(\frac{\partial F(\mathbf{A}_\Delta)}{\partial c_{\theta k}})| > \epsilon_h$ ensures that the blind updation does not happen even when the gradient value is insignificant. In addition to aiding in faster convergence, this heuristic technique will also help in tolerating more initialization error and facilitates reaching towards global minimum.

Now, observe from (5) that $F(\mathbf{A}_\Delta)$ comprises of $N_x N_y$ summations and gradient must be computed separately for each factor since the dictionary is different for each factor. However, due to sparse nature of the frequency spectrum of images, only few frequency bins which have significant energy can be considered, while neglecting the remaining ones. These significant bins can be chosen by putting a threshold on the spectrum of LR images $G_k(\mathbf{f})$. Thus, by using the above updation rule (8) and choosing only significant bins, one can not only achieve faster convergence, but also can obtain a huge savings in computation.

B. Choice of L_x, L_y

The following proposition provides the conditions on L_x,L_y and α_s to obtain a unique SR image.

Proposition 4.1: A unique SR image reconstruction with the proposed approach can be obtained if and only if
i) L_x and L_y are coprime, i.e., $\text{GCD}(L_x, L_y) = 1$
ii) $\alpha_s < K$.

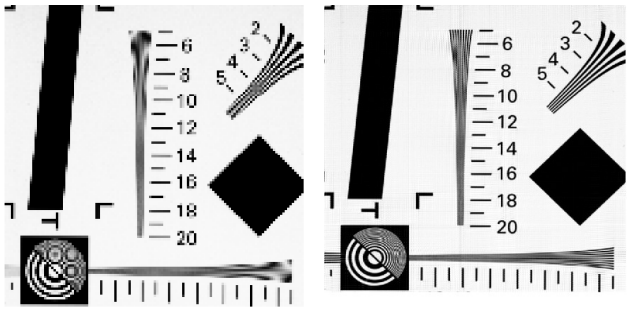
Due to lack of space, the proof of the above assertion is skipped. Recall from Section III-B1 that α_s denotes the dimension of the signal subspace, which can also be interpreted as the number of significant energy aliased subbands. The second condition implies that for the given K LR images, there is hope of reconstruction only when the number of aliased subbands (i.e., each row of $\mathbf{s}(\mathbf{f})$) having significant energy does not exceed K. It can easily be noticed that L_x and L_y has a direct influence on α_s. The higher these factors, higher is the α_s due to more aliasing bands. On the other hand, also recall that $0 < c_{\theta k} < L_\theta$, and (2) holds only when $\{c_{\theta k}\}_{k=1}^K$ are integers i.e., at the end of the registration step, if proper convergence is attained, then we will obtain the sub-pixel shift, $\{\text{round}(c_{\theta k})/L_\theta\}_{k=1}^K$. By expressing $c_{\theta k} = \text{round}(c_{\theta k}) + \Delta c_{\theta k}$, we see that the sub-pixel offset error will be $\Delta c_{\theta k}/L_\theta$. Now suppose if we increase the factor L_θ by some factor say β to βL_θ, then at the end of the convergence, we will obtain an appropriately scaled subpixel shift i.e., $\{\text{round}(\beta c_{\theta k})/(\beta L_\theta)\}_{k=1}^K$. It is easy to observe that $|\text{round}(\beta c_{\theta k})/(\beta L_\theta)$ - $(\text{round}(c_{\theta k})/L_\theta| \leq |\Delta c_{\theta k}/L_\theta|$. Thus, the factors L_x and L_y must be chosen suitably to not only satisfy the conditions stated above but they must also be as large as possible in order to have better estimate of the offsets for the given K LR images. A direct approach to optimally choose them is to include these factors as variables and have an additional step in the above alternating minimization iteration. However, this approach increases the computational complexity. Hence, a simple approach is to first use an approximate method such as NCPS to obtain the sub-pixel offsets i.e., ratios $\{c_{\theta k}^{(0)}/L_\theta\}_{k=1}^K$ for some L_θ. Subsequently, for different choices of L_x and L_y, estimate their corresponding $\{\text{round}(c_{\theta k})\}_{k=1}^K$ and form the covariance matrix using (7). By observing the eigenvalues for each of the choice of L_x and L_y, choose the factors L_x^p and L_y^p, which satisfies the above stated conditions and is larger among the choices. Although this method may not yield an optimal value, in practice we have observed that the choice obtained using this heuristic approach will be closer to the optimal value. It is important to observe that larger K admits larger values of L_x and L_y thereby leading to more accurate estimates of offsets, which is along the expected lines. The registration process is then continued as described earlier using these larger choices of L_x^p and L_y^p and subsequently in the restoration stage, SR image of desired dimension can be obtained as described in Section III-B2.

ISBN 978-90-827970-1-5 © EURASIP 2018

V. Simulation Results

Simulations were performed by fixing the number of LR image to $K = 16$. These LR images were generated by sampling $h(\mathbf{t})$ image which is subjected to translation motion with different shifts and without using any anti-aliasing filter. NCPS method was first employed to obtain an initial offset estimation, and subsequently registration using alternating minimization and restoration using non-uniform bi-cubic interpolation was performed as described in Section III. Fig. 2(a) and 2(b) shows an example of the of one of the LR image of size 64×64 and the SR image of size 320×320 (enhancement factor of 5×5) respectively with the proposed approach. A higher quality SR image (in particular observe the disc at the bottom left corner) can clearly be noticed from the figure despite total aliasing.

In the next simulation, we varied the factors L_x and L_y to study its effect on registration accuracy. Table I captures the results for different choices of enhancement factors. For the sake of comparison the result obtained using the NCPS method is also provided. The NCPS method estimates the translation by using the spectrum of LR images. From the table one can observe that as these factors increases, the registration accuracy improves which clearly corroborates the description provided in Section IV-B. It is important to note that for the choice of the parameters chosen here, enormous resources are required to apply the method of [6]; typically the matrix will be in the order of $4096 \times L_x L_y 4096$. In contrast, the size of matrices to be handled with the present approach is only in the order of $16 \times L_x L_y$ and furthermore, we found the convergence to be achieved with less than 20 alternating minimization iterations in most cases, thus demonstrating the efficacy of this approach.



(a) One of the 64×64 LR image (b) 320×320, SR image

Fig. 2. Example of an LR image and SR image resolved by a factor of 5×5.

NCPS	$L_x = 5, L_y = 6$	$L_x = 10, L_y = 11$	$L_x = 15, L_y = 16$
0.563	0.0620	0.0086	0.0024

TABLE I. Average RMSE (in pixels) comparison with NCPS method and for different choices of L_x and L_y.

VI. Conclusion

A computationally efficient approach for obtaining an SR image from a set of LR images, when the entire spectrum is affected by aliasing, is addressed in this paper. Multi-coset sampling based technique is employed to establish the relationship between the LR image spectrums and the SR image spectrum. This technique facilitates band-wise aliasing in contrast to bin-wise aliasing of the existing approaches, thereby providing a huge reduction in the the size of the relationship matrix on the order of number of LR images and number of sub-bands, and further making it independent of the dimensions of the LR and SR images. Alternating minimization framework is employed to jointly estimate sub-pixel offsets and the sparse sub-bands of SR image spectrum using the steepest-descent technique and MUSIC algorithm, respectively. Additionally, heuristic techniques are described towards reaching global minimum with fewer iterations and to improve the sub-pixel offset estimation accuracy; These are verified through simulation results. The proposed approach can easily and efficiently be deployed on mobile platforms which have limited resources and still can obtain an SR image in almost real-time.

References

[1] K. Nasrollahi and T. B. Moeslund, "Super-resolution: a comprehensive survey," *Machine vision and applications*, vol. 25, no. 6, pp. 1423–1468, 2014.

[2] B. Zitova and J. Flusser, "Image registration methods: a survey," *Image and vision computing*, vol. 21, no. 11, pp. 977–1000, 2003.

[3] Y. Tian and K.-H. Yap, "Joint image registration and super-resolution from low-resolution images with zooming motion," *ieee transactions on circuits and systems for video technology*, vol. 23, no. 7, pp. 1224–1234, 2013.

[4] V. Druchinin. Anti aliasing / low pass filter removal for sharper more detailed images. [Online]. Available: https://www.lifepixel.com/photography-gear/anti-aliasing-low-pass-filter-removal

[5] P. Vandewalle, S. Süsstrunk, and M. Vetterli, "A frequency domain approach to registration of aliased images with application to super-resolution," *EURASIP Journal on applied signal processing*, vol. 2006, pp. 233–233, 2006.

[6] P. Vandewalle, L. Sbaiz, J. Vandewalle, and M. Vetterli, "Super-resolution from unregistered and totally aliased signals using subspace methods," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3687–3703, 2007.

[7] M. Mishali and Y. C. Eldar, "Blind multiband signal reconstruction: Compressed sensing for analog signals," *IEEE Transactions on Signal Processing*, vol. 57, no. 3, pp. 993–1009, 2009.

[8] R. Rubinstein, A. M. Bruckstein, and M. Elad, "Dictionaries for sparse representation modeling," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1045–1057, 2010.

[9] C. L. Byrne, "Alternating minimization and alternating projection algorithms: A tutorial," *Sciences New York*, pp. 1–41, 2011.

[10] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *Antennas and Propagation, IEEE Transactions on*, vol. 34, no. 3, pp. 276–280, 1986.

[11] H. S. Stone, M. T. Orchard, E.-C. Chang, and S. A. Martucci, "A fast direct fourier-based algorithm for subpixel registration of images," *IEEE Transactions on geoscience and remote sensing*, vol. 39, no. 10, pp. 2235–2243, 2001.

[12] A. Scholefield and P. L. Dragotti, "Accurate image registration using approximate strang-fix and an application in super-resolution," in *Signal Processing Conference (EUSIPCO), 2014 Proceedings of the 22nd European*. IEEE, 2014, pp. 1063–1067.

[13] K. Venkataraman, D. Lelescu, J. Duparré, A. McMahon, G. Molina, P. Chatterjee, R. Mullis, and S. Nayar, "Picam: An ultra-thin high performance monolithic camera array," *ACM Transactions on Graphics (TOG)*, vol. 32, no. 6, p. 166, 2013.

[14] G. Carles, J. Downing, and A. R. Harvey, "Super-resolution imaging using a camera array," *Optics letters*, vol. 39, no. 7, pp. 1889–1892, 2014.

[15] K. Kreutz-Delgado, J. F. Murray, B. D. Rao, K. Engan, T.-W. Lee, and T. J. Sejnowski, "Dictionary learning algorithms for sparse representation," *Neural computation*, vol. 15, no. 2, pp. 349–396, 2003.

[16] M. Ataee, H. Zayyani, M. Babaie-Zadeh, and C. Jutten, "Parametric dictionary learning using steepest descent," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*. IEEE, 2010, pp. 1978–1981.

ISBN 978-90-827970-1-5