

Super-Resolution of Remotely Sensed Images With Variable-Pixel Linear Reconstruction

Maria Teresa Merino and Jorge Núñez

Abstract—This paper describes the development and applications of a super-resolution method, known as Super-Resolution Variable-Pixel Linear Reconstruction. The algorithm works combining different lower resolution images in order to obtain, as a result, a higher resolution image. We show that it can make significant spatial resolution improvements to satellite images of the Earth's surface allowing recognition of objects with size approaching the limiting spatial resolution of the lower resolution images. The algorithm is based on the Variable-Pixel Linear Reconstruction algorithm developed by Fruchter and Hook, a well-known method in astronomy but never used for Earth remote sensing purposes. The algorithm preserves photometry, can weight input images according to the statistical significance of each pixel, and removes the effect of geometric distortion on both image shape and photometry. In this paper, we describe its development for remote sensing purposes, show the usefulness of the algorithm working with images as different to the astronomical images as the remote sensing ones, and show applications to: 1) a set of simulated multispectral images obtained from a real Quickbird image; and 2) a set of multispectral real Landsat Enhanced Thematic Mapper Plus (ETM+) images. These examples show that the algorithm provides a substantial improvement in limiting spatial resolution for both simulated and real data sets without significantly altering the multispectral content of the input low-resolution images, without amplifying the noise, and with very few artifacts.

Index Terms—Image enhancement, image resolution, remote sensing, super-resolution.

I. INTRODUCTION

SPATIAL resolution enhancement is usually required in the remote sensing field, especially in satellite images taken with the aim of recognizing objects with size approaching the limiting spatial resolution scale. One way to improve resolution is to use longer focal ratios, but this requires larger and more expensive orbital platforms. Another approach is to use sensors with smaller pixel size and increased pixel density in the charge-coupled device (CCD) chips or similar—nowadays, the most common sensors in this type of project. However, there are additional technical difficulties related to high-quality manufacture and a significant increase in cost. Furthermore, the reduced

pixel size also decreases the number of photons arriving in a fixed exposure time, this being another typical constraint in the observational strategies of some projects. Moreover, the increased number of pixels per chip also increases the capacitance and produces higher noise levels, worse efficiency, and slower charge transference speed. As a result of these effects, CCD technology is approaching the spatial resolution levels currently considered optimal and, in most cases, the costs of extending the limiting resolution further are too great to be worthwhile. In this context, many recent studies have sought to develop new approaches to spatial resolution enhancement at affordable costs. Currently, the best cost/benefit ratio seems to be achieved when using digital image treatment techniques known as spatial resolution enhancement, resolution image reconstruction or simply super-resolution (see [1] and [2] for detailed reviews).

The basic premise of most super-resolution techniques is to combine several images from the same scene considered with low spatial resolution (LR) in order to produce one or several images with a higher resolution (HR). Of course, it can only be assumed that a HR can be obtained from LR images if they are undersampled and suffer from aliasing. Every LR image samples the scene as a different projection of the same scene on different sampling lattices, so they have different profiles in the aliased frequency range. Thus, none of the LR images can be obtained from the other LR ones because each one contains a certain amount of differential information from the same scene, even though it may be in the aliased frequency range. Super-Resolution techniques combine the LR images and attempt to recover as much as possible of this differential information to construct the HR result.

In the aforementioned reviews [1], [2], the most flexible definitions of the term “super-resolution” include all those methods capable of achieving a significant improvement in spatial resolution. However, there are also other more restrictive definitions that consider super-resolution as only those algorithms that also perform at some point, in some way, and in any order or combination of the following three steps.

- Step 1) Automatic coregister of the LR images.
- Step 2) Interpolation and/or recombination of the LR images onto a HR grid or image.
- Step 3) Restoration of the HR or LR images to reduce the noise, blur, and point spread function (PSF) effects.

Important advances have been made as regard Steps 1) and 3), and innovative and promising studies have been conducted. Some are based on recovering additional high-frequency information by working in the frequency or different

Manuscript received June 19, 2006; revised October 14, 2006. This work was supported in part by the Instituto Nacional De Tecnica Aeroespacial (INTA) under Contract FBG302479 and by the Spanish Ministry of Education and Science (MEC) under Grant AyA2005-08604. The work of M. T. Merino was supported by a research fellowship (FPU) awarded by the Spanish Ministry of Education and Science.

M. T. Merino is with the Departament d'Astronomia i Meteorologia, Universitat de Barcelona, 08028 Barcelona, Spain (e-mail: mmerino@am.ub.es).

J. Núñez is with the Departament d'Astronomia i Meteorologia, Universitat de Barcelona, 08028 Barcelona, Spain, and also with Observatori Fabra, 08035 Barcelona, Spain (e-mail: j.nunez@ub.edu).

Digital Object Identifier 10.1109/TGRS.2007.893271

spaces [3]–[5], using probabilistic theories [6], or performing the above three steps in mixed and/or iterative fashion [7], [8]. Among these methods are the Maximum-Likelihood (ML) super-resolution method [9], the Projection Onto Convex Sets method (POCS) [10], [11], and the Iterative Back-Projection Super-Resolution method [12]. For a deep study of the cited and other super-resolution methods, including their advantages and limitations, we refer the reader to [1] and [2].

However, despite their good results, the most powerful algorithms (which include an iterative restoration step) are difficult to be used for the typical applications of object recognition in satellite remotely sensed images, which usually have large dimensions as well as high distortions and observational effects. There are important drawbacks that prevent the most sophisticated algorithms from being used with these images (except for some very specific cases), for example, the high computational load, the critical dependence of *a priori* parameters that are not easily fixed or which vary widely throughout one image or the set of images (the stopping point of the iterations, the PSF, etc.), and the difficult automatic registration with subpixel accuracy in such complex (for example urban or mixed areas) and highly distorted images. In particular, the distortion (even if small or geometrically corrected) affects the space invariability of the PSF. If the PSF is space-variant, it is not possible to use the Fast Fourier Transform for the computations, making the use of the most modern restoration algorithms as the Richardson–Lucy–EM [13]–[15] simply impossible. In most cases with this type of image, it is preferable to use a simpler super-resolution algorithm performing only Step 2) with a low computational load, even though that means ignoring noise and blur and performing the coregister by other manual, traditional techniques. The program we present in this paper focuses on this simple but common case.

Moreover, among the remaining algorithms suitable for common object recognition of objects with size approaching the limiting spatial resolution scale from satellite remote sensing projects, most methods still perform Step 2) with traditional uniform nearest neighbor, linear, or bicubic interpolations. These interpolations mix up the values of neighboring pixels and thus they mix up the differential information contained in the aliased region of those pixels in a uniform and suboptimal way across the whole image, thus losing an important part of the possible final resolution improvement provided by the whole super-resolution method.

Given the above, it is necessary to develop a way of implementing a super-resolution algorithm (or spatial resolution enhancement in some restrictive definitions) that can be used in projects aimed at object recognition on the limiting spatial resolution scale from satellite remotely sensed images, where spatial resolution enhancement can be critical.

In this paper, we present a method known as Super-Resolution Variable-Pixel Linear Reconstruction (SRVPLR), which recombines a set of LR images in a HR image by performing Step 2) in a nonuniform optimal way and also fulfills the additional requirements of being computationally fast and versatile with the desired type of images. The algorithm core is based on the Variable-Pixel Linear Reconstruction (VPLR) algorithm developed by Fruchter and Hook [16] (known as

drizzle), a well-known method in astronomy but that needed to be implemented with some differences in order to check its suitability and then to demonstrate its good results for images as different as the ones for Earth remote sensing purposes. As stated above, that this algorithm was useful for remote sensing community was far from clear and needed a deep study and development. We have implemented the new developed standalone program SRVPLR and in this paper we show that development and some of the good results it provides to demonstrate its usefulness. Two different SRVPLR applications onto sets of simulated and real LR images show how this optimization of Step 2) produces significantly better results than the widely used traditional uniform interpolations.

The rest of this paper is organized as follows. Section II describes the SRVPLR algorithm, while the SRVPLR applications onto a set of real Landsat Enhanced Thematic Mapper Plus (ETM+) images and a set of simulated LR images created from an original Quickbird image are described in Section III. Experimental results from these applications are described and discussed in Section IV. Finally, Section V concludes this paper, and the Appendix reports the parameters used in the quantitative evaluation.

II. SUPER-RESOLUTION VARIABLE-PIXEL LINEAR RECONSTRUCTION (SRVPLR)

SRVPLR core is based on the algorithm VPLR. Fruchter and Hook [16] developed an implementation of this later algorithm (known as drizzle) for the combination of astronomical dithered undersampled images. Drizzle, designed to work with the package Space Telescope Science Data Analysis System (STSDAS) within Image Reduction and Analysis Facility (IRAF) [17], was developed for use with the Hubble Space Telescope (HST) and other astronomical instruments with dithering observation mode [18]. It should be noted that previous known drizzle implementations are designed to work only with astronomical images and we were unable to find any implementation of the VPLR algorithm suitable for use with satellite remotely sensed images like the sets used in Section III [19]. Typical astronomical images have rather different properties and needs from satellite remotely sensed images of the Earth's surface. Astronomical images consist mainly in point sources (stars, asteroids, etc.) and diffuse extended areas (nebulae, galaxies, etc.). Some of their major differences to consider are: the absence of sharp edges, shadows, etc.; the available number of images from the same scene and the multitemporal differences between them; the decisive lower signal-to-noise ratio (SNR); and types and significance of predominant noise sources and distortions. Therefore, some variations and additional functions had to be added to the core VPLR algorithm to develop an SRVPLR implementation useful not only for typical astronomical images but also for satellite remotely sensed images. In particular, some of these new added functions are as follows:

- 1) improvements in numerical processing for working with high geometrical corrections within the precision required when dealing with images with high dimensions and/or different original precisions;

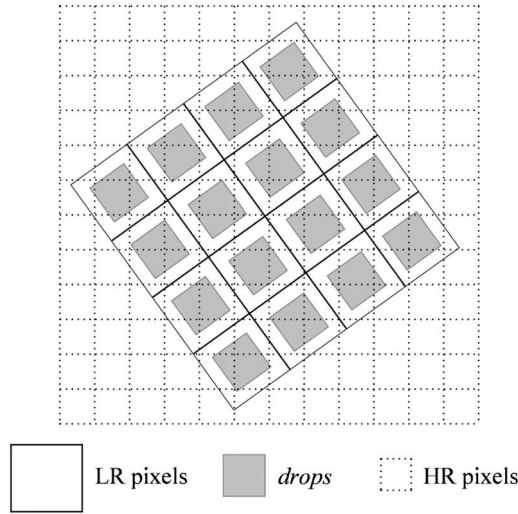


Fig. 1. Example of the elements involved in the projection of one low-resolution image (LR) onto the high-resolution grid (HR) during the variable-pixel linear reconstruction algorithm.

- 2) prior histogram matching on LR images;
- 3) proper consideration of special pixel values used in the original images as marks for special status that may be important to transmit directly or to keep in the result;
- 4) improvements to facilitate the automation of the processes with a high number of LR images with different properties;
- 5) masking and intermediate result utilities to select or evaluate more easily the area of interest in the results.

A more detailed outline of the new added functions can be found near the end of this section.

VPLR is a linear reconstruction algorithm designed to recover much of the information lost in undersampling the scene. Other commonly used linear reconstruction techniques are shift-and-add and interlacing. However, shift-and-add convolves the image again with the original pixel, thus adding to the blurring of the image and the correlation of the noise. Moreover, true interlacing is often not feasible due to poor placement of the sampling grid or the effects of geometric distortion. However, VPLR overlaps the pixels of every LR image over a resulting HR image enabling each LR pixel to be condensed prior to overlapping in order to avoid convolving the HR image with the PSF of the original LR pixel. The relative side size of the condensed pixel with respect to the LR pixel side size is given by the parameter p (pixfrac). The algorithm is closer to shift-and-add when p approaches 1 and to interlacing when p approaches 0. Moreover, a weight mask statistically optimizes this overlapping. Every condensed pixel (hereafter known as drop) is corrected by rotation, translation and bicubic geometric transformation, and finally added to the output initial HR image with a weighting proportional to the overlapping area between the drop and every HR pixel, considering the weight mask values. This algorithm not only preserves photometry but also removes the effects of geometric distortions in both image shape and photometry and minimizes shot noise effects [20]. Fig. 1 shows a graphical description of the objects involved in these projections.

Analytically, if (x_i, y_i) is a LR pixel with intensity $d_{x_i y_i}$ and weight $w_{x_i y_i}$, (x_o, y_o) is a HR pixel with intensity $I_{x_o y_o}$ and weight $W_{x_o y_o}$, and the overlapping area between the drop (x_i, y_i) and the pixel (x_o, y_o) is $a_{x_i y_i x_o y_o}$ (≤ 1), then the resulting values for the pixel $I'_{x_o y_o}$ and its weight $W'_{x_o y_o}$ are computed with the expressions

$$W'_{x_o y_o} = a_{x_i y_i x_o y_o} \cdot w_{x_i y_i} + W_{x_o y_o}$$

$$I'_{x_o y_o} = \frac{d_{x_i y_i} \cdot a_{x_i y_i x_o y_o} \cdot w_{x_i y_i} \cdot s^2 + I_{x_o y_o} W_{x_o y_o}}{W'_{x_o y_o}} \quad (1)$$

where the scale factor between HR and LR images s is introduced to conserve the superficial intensity. Once the whole LR image has been processed, each HR pixel value and weight is given by the expressions

$$W_{x_o y_o} = \sum_{x_i y_i} a_{x_i y_i x_o y_o} \cdot w_{x_i y_i}$$

$$I_{x_o y_o} = \frac{\sum_{x_i y_i} d_{x_i y_i} \cdot a_{x_i y_i x_o y_o} \cdot w_{x_i y_i} \cdot s^2}{W_{x_o y_o}}. \quad (2)$$

Note that in these summands most $a_{x_i y_i x_o y_o}$ are zero because an input LR pixel usually only affects few output HR pixels. The recombination algorithm is applied pixel by pixel throughout all the LR images and HR pixels to obtain the HR image that contains information from all those LR images already recombined and its corresponding weight map W . Note that the pixel side size of the HR image is as small as required depending on the elected s value, although this does not mean that the HR resolution is as precise as that corresponding to this size (rather, it is an intermediate value). If the drop size (given by p) is sufficiently small, not all output HR pixels will have data added to them from each input image. Thus, s and p values should be chosen in order to use a drop size small enough to avoid degrading the image, but large enough to ensure that after all the LR images are overlapped the coverage is uniform enough.

In order to work with real data, any useful implementation of the VPLR algorithm must be able to work with bicubic geometric transformation with given parameters that obtain the coregister of the LR images at a subpixel level of precision. Moreover, the weighting system (through the weight masks) has to consider other effects such as the following:

- 1) different exposure times of every LR input image;
- 2) given initial masks where the pixels are weighted depending on their reliability;
- 3) different system reference origins of LR images pixels.

SRVPLR includes all those utilities already included in previous drizzle implementations. In addition, it also contains others that are specifically necessary for working with satellite remotely sensed images of the Earth's surface:

- 1) improvements in numerical processing for working with high geometrical corrections at the precision required;
- 2) six different ways of setting all the program parameters and the properties of each image in order to facilitate the automation of the processes with a high number of LR images with different properties;

- 3) several utilities to obtain intermediate results, logs, and evaluation data if desired: histogram-matched LR images, with histograms prior and after the histogram matching, the original LR images corrected for translation, rotation, and geometrical distortions, the intermediate results of SRVPLR after every new LR recombination, etc.;
- 4) prior histogram matching on LR images when required;
- 5) proper consideration of special pixel values used in the original LR images as marks for special status that may be important to transmit or to keep throughout the process;
- 6) utilities for masking in the result the scene zones not affected by any recombination;
- 7) utilities for masking in the result the scene zones not shared by all the LR images, thus obtaining only the totally recombined area when considering every LR image;
- 8) optimization to be able to work with images with high dimensions and/or different precisions;
- 9) optional configuration of the limiting parameters by the user in order to adapt the execution to the available images and computer properties.

Finally, it should be stated that previous drizzle implementations, which are designed to work with astronomical images, are integrated into some interfaces, software platforms and libraries (IRAF, IDL, etc.) that are well known and widely used by the astrophysical community. However, in many remote sensing projects it is especially necessary to use software that is as simple, independent and easily exportable as possible. In our case, we decided to develop SRVPLR as an autonomous program written in ANSI C using only Numerical Recipes libraries [21] to meet these requirements.

III. EXAMPLES OF APPLICATIONS

SRVPLR was applied to different sets of real and simulated LR images to evaluate its performance and results [22]. We present here two of these applications, which are especially appropriate because of their representation and special difficulty.

- 1) Simulated LR images created from red, green, blue (RGB) bands of an original Quickbird image from Madrid (Spain).
- 2) Real LR RGB Landsat ETM+ images from the city of Barcelona and its surroundings (Spain) selected among the available observations over a two-year period.

We used both Quickbird simulated data and real Landsat images in order to show on one hand the gain that can be obtained in resolution with SRVPLR using almost perfect data (simulated Quickbird data) and, on the other hand, to show that the algorithm is able to obtain also very good results with real images acquired by a different remote sensing system with various operating modes. Of course, the real data experiment could be performed using a set of high-resolution Quickbird images. Nevertheless, since to obtain a set of independent images of the same region from Quickbird or IKONOS satellites is difficult and very expensive, we preferred to use the more widely available Landsat data.

The two sets of LR images have very different properties, require a different coregister process, and the SRVPLR options

and utilities used were different. Both of them are especially difficult to treat with other more sophisticated super-resolution algorithms due to the following characteristics:

- 1) huge size and a rather large number of images per set (9) to be processed;
- 2) important difficulties in determining a precise enough PSF across the image;
- 3) very different landscapes across the images (urban areas, parks, crops, etc.) with very different properties;
- 4) high geometric corrections needed to achieve subpixel precision in their coregister;
- 5) special need to work properly with special pixel values used in the original LR images as marks for special status;
- 6) temporal variability of the scene between the real LR images (including changes in the objects included, different illumination and atmospheric conditions), it being imperative to use some additional techniques such as the histogram matching.

Moreover, both Quickbird and Landsat ETM+ images are currently among the most sought-after and widely used type of satellite remotely sensed images in object recognition projects, and this makes the choice even more interesting.

In the following subsections, we describe in detail the SRVPLR applications made on each of the two sets of LR images. Fig. 2 includes outlines of the two processes described in Sections III and IV.

A. Quickbird Application

For this application, we had an original multispectral Quickbird image of an area of Madrid. We generated nine simulations from a 2000 pixel \times 2000 pixel subarray of the RGB bands with 2.4-m ground sampling distance (GSD, approximately the pixel side size over terrain). The LR images were nine 1000 pixel \times 1000 pixel RGB 4.8-m GSD simulations rotated $20^\circ \cdot n$ ($n = 0, \dots, 8$). They were created using the intermediate and additional results tools incorporated in SRVPLR. During the process, the original image was rotated and projected onto a 0.6-m GSD grid, and the resulting image was degraded with an integration of counts to simulate the 4.8-m GSD image.

The SRVPLR application presented here was performed with the 1000 pixel \times 1000 pixel simulated LR with a scale factor $s = 0.5$. Thus, it generated a 2000 pixel \times 2000 pixel HR image with a pixel side size half the pixel side size of the LR images. Therefore, the obtained HR image is the same size as the original Quickbird image from which we obtained the simulated LR. Fig. 2 (left) shows an outline of this application.

We tried several executions with different values of the pixfrac p . In this particular case with such important relative rotations, when the pixfrac value is less than the HR pixel diagonal there may be output pixels in HR that are not affected by some of the LR among other pixels affected by all of them. This situation is not desirable because the VPLR algorithm is based on the multiple overlapping of several LR images to extract part of the differential information contained in the aliasing region. Thus, in those cases where big rotations or

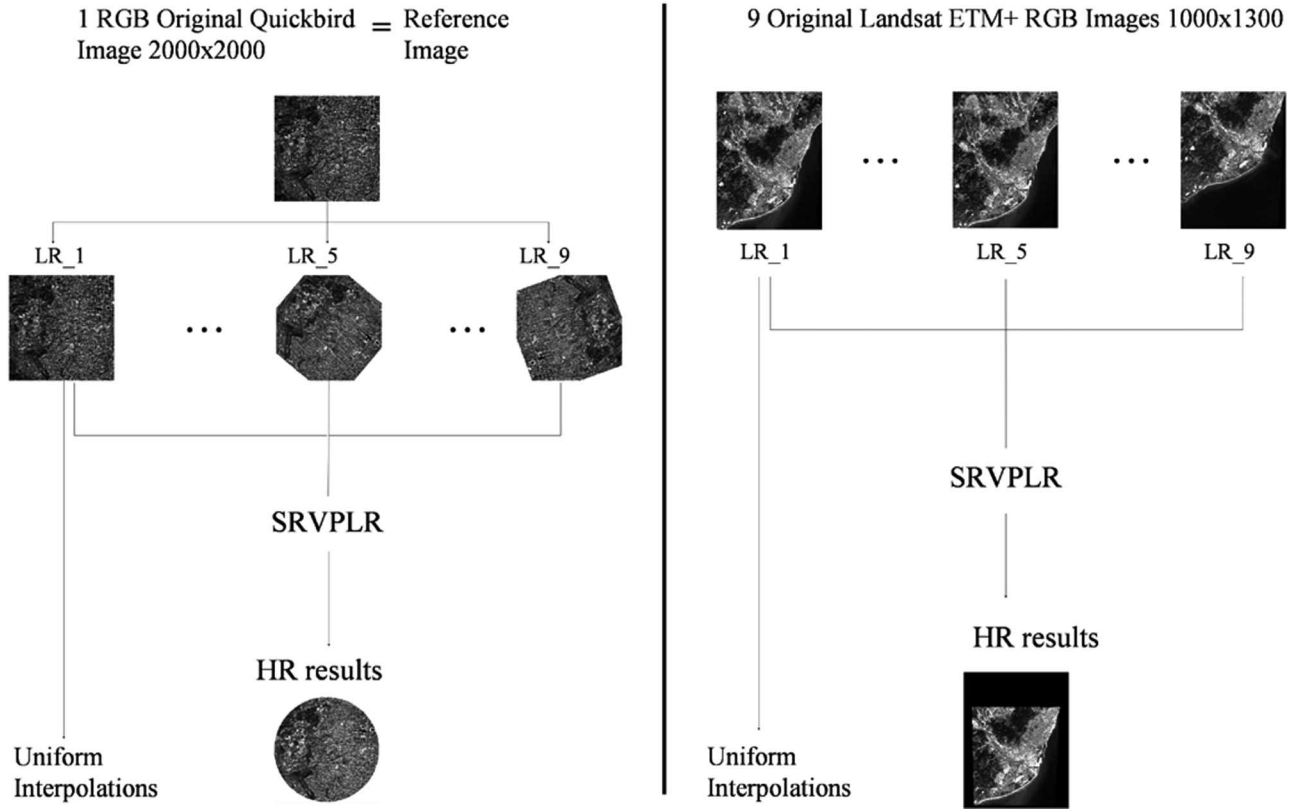


Fig. 2. Outline of the LR images sets used in each SRVPLR application. (Left) Simulated RGB Quickbird images (Section III-A). (Right) Real RGB Landsat ETM+ images (Section III-B). Further explanations in the text.

geometric distortions exist, we recommend the use of pixfrac values greater than the minimum, as follows:

$$p_{\min} = \sqrt{2 \cdot s^2} \quad (3)$$

The phenomena described above can be easily detected through the sensible use of some additional tools incorporated in SRVPLR: for extracting the intermediate results, for working with special pixel values used as marks for special status regions, and for detecting the pixels in the output image not affected by any LR projection. In theory, the minimum pixfrac value ensures that the PSF of the pixel condensation (drop) produces as little image degradation as possible. However, we recommend in practice allowing some degree of overlap with a pixfrac value slightly higher than p_{\min} in order to increase the overlapping and thus optimize the recombination. This is especially relevant when the number of LR images available is low or when the geometric correction imprecision can lead to a nonnegligible degree of deformation in the overlapping area of each output pixel with each LR image drops. As p approaches p_{\min} the overlapping decreases and produces less recombination; however, if the geometric corrections are precise there is a better localization of the recombination of the differential information from each LR. Otherwise, as p approaches 1 the overlapping increases but the convolution with the PSF drop has a greater effect on the output HR and the localization of the recombination is worse. In the particular Quickbird case,

images are simulations extracted from an original one and geometric corrections are well known. Thus, there are no big differences in a wide range of p . In Section IV, we report the results corresponding to an application of SRVPLR with $p = 0.71$, very near to p_{\min} .

In order to compute properly certain quantitative results shown in Section IV, we used an additional tool of SRVPLR to select in the output HR image only those pixels that have been overlapped by all the corrected LR image projections, the rest being masked out. In addition, we used the SRVPLR tools to deal with special pixel values used as marks for special status, to mark the pixels not affected by any LR image, to mark the pixels not affected by at least one of the LR images and to mark in the resulting HR weight image the areas not affected by all the LR images. Nevertheless, it is possible to perform different executions without some or any of these tools in order to see in the weight image the final overlapping and effects of the whole output in every region of the final HR.

B. Landsat ETM+ Application

For this application, we had a set of nine LR Landsat ETM+ multispectral real images with 30-m GSD from Barcelona and the surrounding area taken over a two-year period. We processed 1000 pixel \times 1300 pixel subarrays of the RGB bands with SRVPLR. We also had the nine corresponding 15-m GSD panchromatic (PAN) images. Their wavelength range is described in Table I.

TABLE I
LANDSAT ETM+ BANDS USED

Band	$\lambda_{\min}(\text{nm})$	$\lambda_{\max}(\text{nm})$
PAN	520	900
B	450	520
G	520	600
R	630	690

The choice of $N \times M$ dimensions with N different than M is noteworthy because some available super-resolution programs only accept images with $N = M$. Moreover, most of the programs that accept N different from M actually base the computations on $L \times L$ LR images with $L = \max(N, M)$; these contain the original LR surrounded by added pixels with a constant value. When these programs are unable to work properly with masks or with selected pixel values for marking special status, their results are contaminated by the effect of these pixel values without physical meaning. In contrast, all these problems are already solved with SRVPLR because it can work with N different from M and is designed to work properly with those selected pixel values for marking special status in the LR images and with an initial weight mask for each LR.

Coregister with subpixel precision is one of the main problems with real data. As SRVPLR only performs the recombination, it needs the parameters for the geometric correction of each LR and their reference system translations and rotations with subpixel precision for the right recombination to produce the HR. We used control points to fit by a traditional least square fitting method the following polynomials:

$$\begin{aligned} x' &= a_0 + a_1 \cdot x + a_2 \cdot y + a_3 \cdot x \cdot y \\ y' &= b_0 + b_1 \cdot x + b_2 \cdot y + b_3 \cdot x \cdot y \end{aligned} \quad (4)$$

where (x, y) and (x', y') are the reference systems to coregister and a_i and b_i are the parameters to be fitted.

The best fits obtained for each LR had root mean square error (RMSE) between 0.5 and 0.6 in pixel units. These are typical values when working with this type of satellite remotely sensed images of the Earth. However, they are larger than desirable for a super-resolution application. Nevertheless, we accepted them as valid enough for the following reasons.

- 1) In this type of image and scene, the use of polynomials with height coordinate Z parameters could obtain better fits. However, SRVPLR is not able to perform 3-D geometric corrections because this extra feature has yet to be achieved at a reasonable additional computational cost.
- 2) These RMSE values are larger than advisable but small enough to obtain a significant spatial resolution enhancement with SRVPLR. This gives them the additional interest of showing the effects of a nonperfect real case in contrast to the previous application on simulated Quickbird images.

- 3) Super-Resolution algorithms applied to real orbital images like Landsat need to work with a large amount of images, so they usually have to perform the coregister with automatic coregistration algorithms. These programs with this type of image barely achieve better RMSE values than the ones we obtained with control points and the polynomial fit explained above.

In this application, the LR images are independent real images of 30-m GSD from different days and conditions. Therefore, there is no ideal reference image as in the Quickbird application with simulated LR images. Nevertheless, Landsat ETM+ also has a 15-m GSD panchromatic band (PAN). Thus, we chose the scale factor $s = 0.5$ to produce a 15-m GSD HR result, the same GSD as the PAN Landsat ETM+ images. This enabled us to use one of them as the reference image to estimate the quantitative evaluation of this application in Section IV.

In terms of choosing an optimal pixfrac value, it should be remembered that in this case the images are real, with important geometric corrections and a relatively large RMSE. Therefore, when we used p values near to p_{\min} in (3), we could see the geometric grid-shaped pattern effect. This effect appears because the different overlapping geometry produces some pixels with more overlap than others, following a striped pattern. This pattern is more evident in homogeneous regions of the image (the sea, parks, etc.) and is more easily detected in the weight output image. It increases with geometric correction errors and decreases when the number of LR images increases. When p approaches 1, the drops are less concentrated and the overlapping increases producing less stripping effect. However, with these high p values the resolution enhancement is also decreased. After several trials, we decided to choose $p = 0.85$ as the best compromise between minimizing the stripping pattern effect and a reasonable spatial resolution enhancement.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

Fig. 3 presents several corresponding subarrays of LR, HR images that show the improvements achieved by SRVPLR. Each of the subarrays shows regions with some special interest because of their difficulties, impressive results, special effects or artifacts. Those selected subarrays are highly zoomed to allow the reader to appreciate more easily how the improvement in resolution in SRVPLR is crucial to recognize the resolved objects with sizes near the limiting spatial resolution scale of the LR images.

Moreover, we were also interested in an objective and quantitative evaluation of the results. In both applications selected for this paper, there is some kind of reference image to calculate some of the usual parameters (formulas in the Appendix) and evaluate super-resolution improvements. The resulting values (computed for each whole HR image RGB band) produced by SRVPLR are shown in Tables II and III.

Finally, we also computed (see tables) the same parameters with the three most traditional uniform interpolations (nearest neighbor, bilinear, and bicubic), these being done directly from

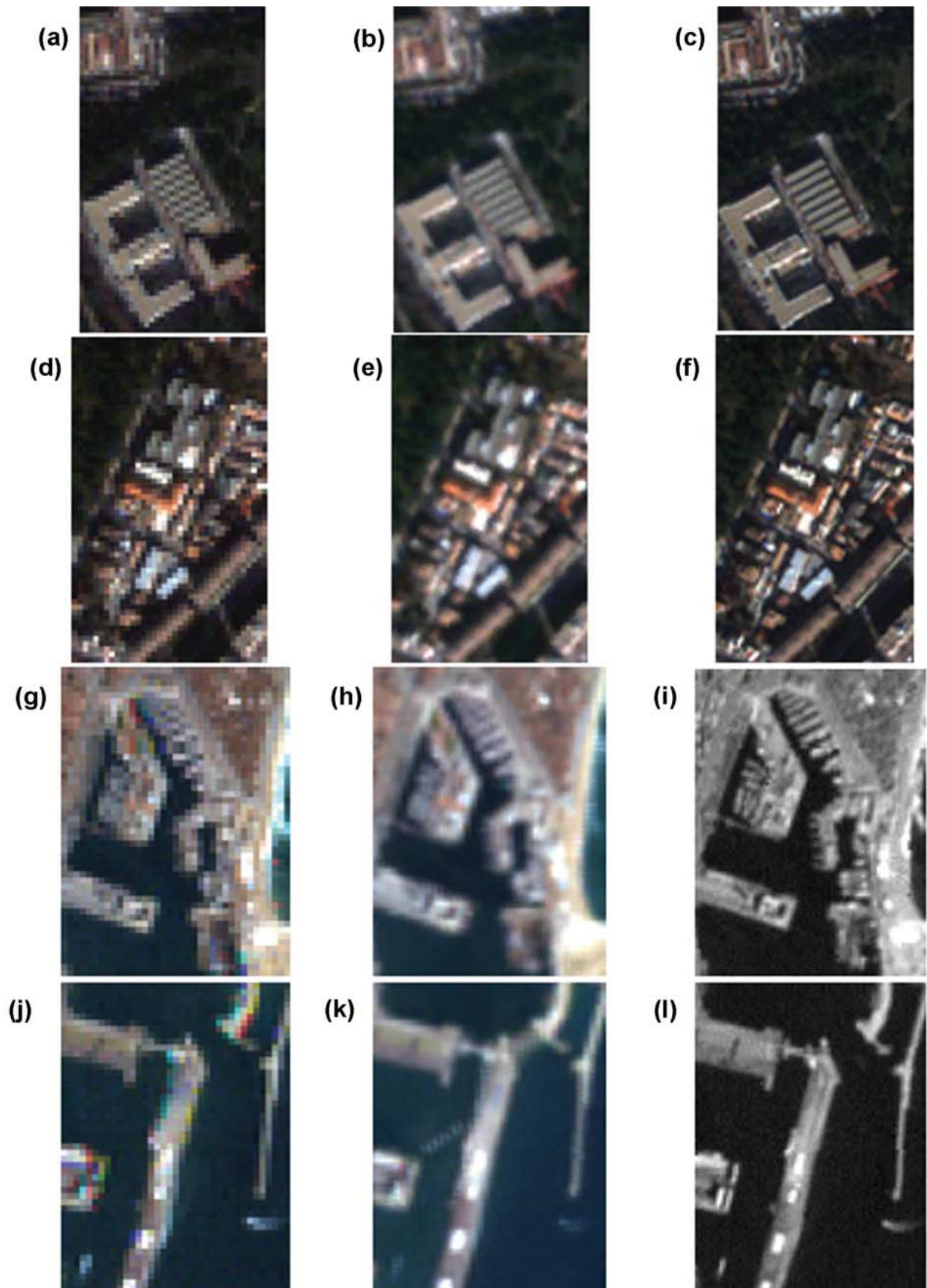


Fig. 3. Examples of graphical results for the Quickbird [(a)–(f)] and Landsat applications [(g)–(l)]. They are selected subarrays from one of the LR images used [(a), (d), (g), (j)], from the results from using SRVPLR [(b), (e), (h), (k)] and from the reference images [(c), (f), (i), (l)]. Further explanations in the text.

TABLE II

Quantitative evaluation results over the QuickBird simulated LR set					
Process	Band	RMSE	NRMSE	ρ	SNR_{dB}
SRVPLR	B	13.624563	0.070193	0.995073	23.074167
	G	23.772135	0.088112	0.992236	21.099344
	R	18.636009	0.117004	0.986310	18.636009
Nearest Neighbour Interpolation	B	13.337359	0.068713	0.995279	23.259222
	G	23.359102	0.086581	0.992504	21.251585
	R	18.819071	0.114564	0.986875	18.819071
Bilinear and Bicubic Interpolations	B	14.752393	0.076003	0.994224	22.383369
	G	26.505672	0.098243	0.990348	20.153931
	R	24.935646	0.129676	0.983184	17.742836

TABLE III

Quantitative evaluation results over the Landsat ETM+ real LR set					
Process	Band	RMSE	NRMSE	ρ	SNR_{dB}
SRVPLR	B	13.234279	0.349752	0.877673	9.124790
	G	8.566670	0.226393	0.948746	12.902721
	R	6.529128	0.172546	0.970228	15.261905
Nearest Neighbour Interpolation	B	12.425860	0.328388	0.892162	9.672267
	G	9.822371	0.259578	0.932619	11.714635
	R	9.366528	0.247530	0.938729	12.127436
Bilinear and Bicubic Interpolations	B	12.341790	0.326166	0.893616	9.731233
	G	9.842099	0.260100	0.932348	11.697208
	R	9.249114	0.244427	0.940255	12.237005

one of the LR images. Those parameters are also computed for each interpolated image band as a whole. The results for bilinear and bicubic interpolations are shown together because they produce the same values within computing precision. SRVPLR produces a HR image that directly simulates the values that a real HR sensor could detect. In contrast, the traditional interpolations obtain every HR pixel value from an approximation of the nearest LR pixels of one image without considering the change of the pixel, and they then have to be multiplied by a proportional constant factor s^2 to simulate

the actual intensity values that a real HR sensor could have collected.

It should also be pointed out that there is a positive bias in the calculation of these parameters in favor of the interpolation. This is due to the fact that these parameters are calculated using the same LR image from which we made the interpolation as the reference image, while the HR image resulting from SRVPLR is obtained by the recombination of nine independent LR images. We did not try fairer (but more complexes) methods of comparative evaluation because even with this bias favoring

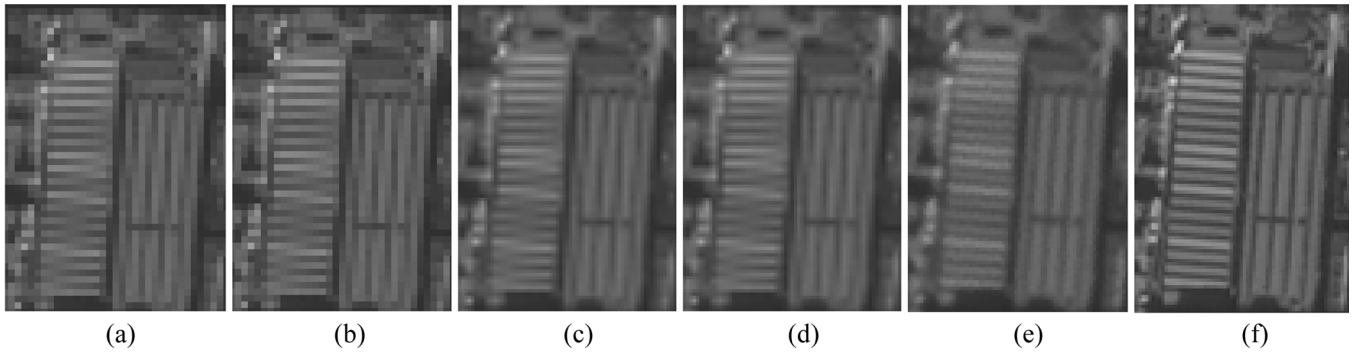


Fig. 4. Subarrays of the R bands from the following images. (a) The simulated LR image with 0° rotation. (b) Nearest neighbor interpolation of (a). (c) Bilinear interpolation of (a). (d) Bicubic interpolation of (a). (e) SRVPLR result from all nine LR images. (f) Reference original Quickbird image.

interpolation, our results show that SRVPLR obtains better results.

A. Simulated LR Set From One Quickbird Image

Fig. 3 shows some results of SRVPLR application to the simulated LR images created from a Quickbird RGB image. Each of the two selected regions shows the respective subarrays of the following images:

- 1) LR simulated image rotated 0° (a), (d);
- 2) SRVPLR result using the process described in Section III-A (b), (e);
- 3) reference image (original Quickbird image) (c), (f).

Of course, SRVPLR is unable to obtain resolution as good as in the reference image. However, the SRVPLR result must be compared with one of the LR images and the reference image as this enables us to evaluate the improvements and the artifacts.

The first region (a)–(c) shows a building with characteristic geometric edges, woods, a road with cars, and some smaller houses. First, the edges of the building show a high contrast and follow straight lines in an inclined direction, thus indicating that the improvement achieved with SRVPLR (b) is clearly significant and that there are no important artifacts resulting from the reference system axis directions. Second, the woods show how SRVPLR is also capable of making improvements in more homogeneous areas without high contrast. Finally, the houses and the road with cars have small elements with sizes near the limit of the resolution scale, which are the most difficult to recover from the LR image. Even so, the results are promising.

The second region (d)–(f) shows a dense industrial complex with different structures and buildings in several directions. There are no artifacts due to the typical multitemporal effects in the shadows, moving objects, etc., because in this case all the LR simulated images were created from the same original image.

In this application, we have chosen an s value to obtain the HR image with the same GSD as the original Quickbird RGB bands from which we obtained the simulated LR images. Thus, we can also use the original Quickbird bands as the reference images to evaluate quantitatively the improvements achieved by SRVPLR in obtaining the HR from the LR images. The

results of the parameters described in the Appendix are shown in Table II.

In this particular case, the interpolation was performed from the simulated LR image most similar to the original image (the LR rotated 0°). Fig. 4 shows a selected region with the R band subarrays from the following images:

- (a) LR simulated image rotated 0° ;
- (b) nearest neighbor interpolation of (a);
- (c) bilinear interpolation of (a);
- (d) bicubic interpolation of (a);
- (e) SRVPLR result from processing the nine LR images;
- (f) reference: original Quickbird image.

Comparison of Fig. 4 and Table II shows that the interpolation methods which best fit the histogram of the reference image and alter less the image surroundings are those with the worse evaluation parameters values (and vice versa). In this R band example, the quantitative results in Table II for the nearest neighbor interpolation image are better (lower RMSE and normalized RMSE (NRMSE), and bigger ρ and SNR) than for the bilinear and bicubic interpolation images. In contrast, in Fig. 4 subarrays the reader can easily appreciate how the nearest neighbor interpolation subarray (b) provides worse visual results than the bilinear subarray (c) or the bicubic subarray (d). The nearest neighbor interpolation method cannot recover properly the correct shape of the building on the left and clearly deforms the other smaller objects (cars, etc.) on its left side. However, the bilinear and bicubic interpolation methods recover better those shapes and objects, even though they are clearly visually worse (have more artifacts and poorer shapes) than the SRVPLR result shown in (e).

Those differences between the visual and the quantitative best interpolation method results are not contradictory. They happen because bilinear and bicubic interpolation methods achieve better qualitative (visual) fit to the interpolated image (the reference image) by mixing in a nonuniform way the information of neighbor pixels, whereas the nearest neighbor interpolation achieves a better quantitative fit to the interpolated image (the same reference image) because it just chooses the nearest pixel value.

In contrast to the interpolations of the reference image methods, the HR image obtained with SRVPLR was obtained by

the recombination of nine different LR images. These different conditions in the comparison with the reference image produce some evaluation parameter values that are slightly better for some nearest neighbor interpolations in certain bands, but at the same time the visual results are clearly better for SRVPLR, as shown in Fig. 4. In any event, and despite the disadvantages aforementioned, SRVPLR obtains better evaluation parameter results compared with the bilinear and bicubic interpolations: RMSE and NRMSE values are lower; SNR (SNR_{dB}) values are higher; and comparative correlation ρ values are nearer to 1. All these parameters suggest that the SRVPLR result is nearer to the reference image than is the bilinear and bicubic interpolations.

To sum up, SRVPLR recombines the differential information from the nine different LR images to obtain a HR image with more realistic visual results and better quantitative evaluation parameters than with traditional interpolations (excluding the nearest neighbor interpolation because of the effects previously mentioned).

B. Real LR Set of Landsat ETM+ Images

Fig. 3 shows some results of SRVPLR application to the nine real Landsat RGB images. Each of the two selected regions shows the respective subarrays of the following images:

- 1) one LR real Landsat ETM+ RGB image (g), (j);
- 2) SRVPLR result using the process described in Section III-B (h), (k);
- 3) reference image (PAN image corresponding to the LR selected above) (i), (l).

In this case, it should be noted that SRVPLR recombines nine different real images coregistered with RMSE values between 0.5 and 0.6 pixels and with very important multitemporal effects. Even so, it can be seen that SRVPLR achieves important improvements.

The first region (g)–(i) shows part of the sports harbor with the quays and other structures at the limit of spatial resolution in the LR image. SRVPLR manages to resolve many of these structures without introducing any false ones due to artifacts. Moreover, the sea is also properly maintained, despite being an almost homogeneous cover. The unavoidable and most common effect in the SRVPLR result is that the edges of some structures are a bit diffuse and slightly too curved due to the poor geometric correction resulting from the RMSE of the coregister of the LR images.

The second region (j)–(l) shows some structures at the entrance to the harbor. Here, the LR image is only one among the nine (taken on different days over a two-year period) that SRVPLR recombines. In (j) there is a boat coming into the harbor, another one inside, and all but one of the sea corridors are open. However, the boats are not present in any of the other eight LR images and the corridors are closed in some of them. This is an excellent example of the multitemporal effect, which is why in SRVPLR these elements (boat, corridors) show ghosting aspect. Their intensity is proportional to the percentage of images where they appear, so the boats are less intense than the new closed corridors. Moreover, these elements show the shaded

pattern described in Section III because the overlapping is not homogeneous enough. This artifact is more evident in these places because of the important contrast between the sea and the boat and corridor colors, and also because the elements appear in a few images. However, it is actually present throughout the image to a greater or lesser extent. Both artifacts (multitemporal effects and nonhomogeneous overlapping) could be reduced by applying SRVPLR to a high number of images, and if the geometric corrections were less intense and more accurate.

As already explained in Section III-B, in order to be able to compute the parameters established in the Appendix, our reference image in this application is one of the panchromatic images from Landsat ETM+ (hereafter called PAN). We processed our 30-m GSD LR RGB real images with $s = 0.5$ in order to produce an HR image with the same GSD as the PAN. This approach using PAN as the reference image is only approximately valid because of the following.

- 1) The images from Landsat ETM+ are in the LGPS distribution images system. This system works with the pixel centers translated 0.5 pixels in x and y from the pixel centers of the RGB HR bands resulting from SRVPLR and also from the interpolated bands from one of the LR.
- 2) The PAN image has a different frequency range and calibration from the RGB bands. Thus, the radiometry of the same scene can be very different and, moreover, this difference changes with the different kind of objects in the scene.
- 3) The reference is only one PAN image from one of the nine days and conditions for which the nine RGB LR images are available. Thus, there are multitemporal effects between the different LR images (changes in the geometry, the illumination and the objects in the scene). The PAN image was only a completely valid reference image (ignoring the other problems above) for the one RGB LR image from the same day.

Nevertheless, we decided to use PAN as the reference image, and sought to minimize its drawbacks while preserving as much as possible the object localization and using simple transforming methods that are as harmless as possible. We have no simple solution to drawback 3), but to overcome drawbacks 1) and 2) we used additional tools included in SRVPLR to perform the following transformations.

- 1) The original selected PAN image was reprojected onto an empty grid with the same side size but translated 0.5 pixels in x and y to overcome drawback 1).
- 2) Each R, G, and B band was histogram matched to the projected PAN reference image to overcome drawback 2).

Table III shows the values computed for the parameters mentioned in the Appendix.

It should be noted that the interpolations were made from the RGB LR image corresponding to the selected PAN reference image. Thus, the interpolated images do not suffer from multitemporal effects [drawback 1)] when compared with the reference image. Furthermore, they have no additional source of errors corresponding to the geometric correction, which was

described in Section III-B as being one of the main sources of errors in the HR image resulting from SRVPLR.

The results in Table III are very successful despite the important advantages that the interpolated images have compared with the HR image produced by SRVPLR. On the whole, SRVPLR produces clearly better results than interpolations in bands G and R. The worse results in band B are of not concern because we cannot expect a good fit of them since the wavelength range of the reference PAN image excludes the wavelength range of the HR B band (see Table I). Furthermore, the drawbacks present in the HR images produced by SRVPLR but not presented in the interpolated images (mentioned in the previous paragraph) are decisive in a large portion of the image where the predominant types of cover change significantly their radiance in the band B wavelength range.

To sum up, SRVPLR improves the spatial resolution limit and is a better solution than traditional interpolations, its few artifacts being mainly due to the multitemporality of the real images (effects diminishing when increasing the number of LR images processed), despite the relatively high RMSE of the coregister.

V. CONCLUSION

In this paper we have presented the super-resolution algorithm SRVPLR and two examples of applications aimed at recognition of objects with sizes approaching the limiting spatial resolution scale. Its core is based on the algorithm VPLR. Fruchter and Hook [20] developed an implementation of this later algorithm (known as drizzle) for the combination of astronomical dithered undersampled images, but it was not suitable for properly working with remote sensing data sets. Besides, astronomical and remote sensing images present so different properties that its usefulness with satellite remotely sensed data, even with additional features, needed to be demonstrated. We have developed SRVPLR as an evolution of the astronomical drizzle with new features to make it useful for both remote sensing and astronomical data sets. Moreover, we have tested and demonstrated with the applications shown in this paper that the algorithm provides very good super-resolution results with low computational cost for satellite remotely sensed images.

The first application was performed with a set of nine simulated LR RGB images extracted from one original Quickbird image, and the other application with a set of nine real RGB Landsat ETM+ images. Each case is very different from the other in terms of the application, properties of the images and artifacts. In both cases, it has been shown how SRVPLR obtains better results than traditional uniform interpolations over any of the low-resolution images (even using the low-resolution image best suited to perform the interpolation), and how significant these improvements can be for object recognition purposes with sizes near the limiting spatial resolution scale.

SRVPLR uses a nonuniform interpolation algorithm with low computational load, thus enabling real-time applications. We acknowledge that, in this approach, degradation models are limited and only theoretically applicable when the blur and noise are the same in all the LR images. In any event, and

as argued in the Introduction, this drawback is acceptable in some projects in the particular field of satellite remotely sensed images. Therefore, SRVPLR may be a very useful tool for many current projects.

Our aim now is to develop new versions of SRVPLR that are capable of supporting additional advanced tools such as the following:

- 1) automatic coregister of images with subpixel accuracy capable of working also with 3-D corrections;
- 2) correction of different illumination and shadows;
- 3) work with multispectral images in different spaces (HSI, etc.) and color band combinations;
- 4) new tools to evaluate resolution enhancement.

Finally, we are interested in developing a new full super-resolution algorithm that is also capable of dealing with noise and blur deconvolution, performing the reprojection to the HR grid with SRVPLR.

APPENDIX

The measure of success in restoration algorithms is usually error measures between the processed image and the reference image. The parameters computed in this paper as quantitative evaluators are the RMSE, the NRMSE, the comparative correlation (ρ), and the SNR in decibel units (SNR_{dB}). Each one is computed for each band separately.

When I_{ij}^n is the intensity value of the pixel (i, j) of the band n with $N \times M$ pixels and $n = \text{REF}$ is the identifier for the chosen reference image, they are computed as

$$\begin{aligned} \text{RMSE}_n &= \sqrt{\frac{\sum_{i=1}^N \sum_{j=1}^M (I_{i,j}^n - I_{i,j}^{\text{REF}})^2}{N \cdot M}} \\ \text{NRMSE}_n &= \sqrt{\frac{\sum_{i=1}^N \sum_{j=1}^M (I_{i,j}^n - I_{i,j}^{\text{REF}})^2}{\sum_{i=1}^N \sum_{j=1}^M (I_{i,j}^{\text{REF}})^2}} \\ \rho_n &= 1 - \frac{\sum_{i=1}^N \sum_{j=1}^M (I_{i,j}^n - I_{i,j}^{\text{REF}})^2}{\sum_{i=1}^N \sum_{j=1}^M (I_{i,j}^{\text{REF}})^2} \\ \text{SNR}_{\text{dB}} &= 10 \cdot \log_{10} \left(\frac{\sum_{i=1}^N \sum_{j=1}^M (I_{i,j}^{\text{REF}})^2}{\sum_{i=1}^N \sum_{j=1}^M (I_{i,j}^n - I_{i,j}^{\text{REF}})^2} \right). \end{aligned}$$

ACKNOWLEDGMENT

The authors would like to thank the Instituto Nacional de Técnica Aeroespacial (INTA) for providing the Quickbird image used in this paper. The authors would also like to thank the Departament de Geografia de la Universitat Autònoma de Barcelona and the Agència Catalana de l'Aigua (ACA) for providing the Landsat ETM+ images used also in this paper.

REFERENCES

- [1] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Process. Mag.*, vol. 20, no. 3, pp. 21–36, May 2003.
- [2] M. M. Hadhoud, F. Abd El-Samie, and S. E. El-Khamy, "New trends in high resolution image processing," in *Proc. IEEE Conf. 4th Workshop Photon. and Appl.*, May 2004, pp. 2–23.
- [3] P. Vandewalle, L. Sbaiz, M. Vetterli, and S. Sustrunk, "Super-resolution from highly undersampled images," in *Proc. IEEE ICIP*, Sep. 2005, vol. 1, pp. 889–892.
- [4] J. T. Hsu, C. C. Yen, C. C. Li, M. Sun, B. Tian, and M. Kaygusuz, "Application of wavelet-based POCS super-resolution for cardiovascular MRI image enhancement," in *Proc. IEEE 3rd Int. Conf. Image and Graphics*, Dec. 2004, pp. 572–575.
- [5] M. B. Chappalli and N. K. Bose, "Simultaneous noise filtering and super-resolution with second-generation wavelets," *IEEE Signal Process. Lett.*, vol. 12, no. 11, pp. 772–775, Nov. 2005.
- [6] H. Shen, P. Li, L. Zhang, and Y. Zhao, "A MAP algorithm to super-resolution image reconstruction," in *Proc. IEEE 3rd Int. Conf. Image and Graphics*, Dec. 2004, pp. 544–547.
- [7] D. Robinson and P. Milanfar, "Statistical performance analysis of super-resolution," *IEEE Trans. Image Process.*, vol. 15, no. 6, pp. 1413–1428, Jun. 2006.
- [8] M. D. Gupta, S. Rajaram, N. Petrovic, and T. S. Huang, "Non-parametric image super-resolution using multiple images," in *Proc. IEEE ICI*, Sep. 2005, vol. 2, pp. 89–92.
- [9] B. C. Tom and A. K. Katsaggelos, "Reconstruction of a high-resolution image by simultaneous registration, restoration, and interpolation of low-resolution images," in *Proc. IEEE Int. Conf. Image Process.*, Washington, DC, Oct. 1995, vol. 2, pp. 539–542.
- [10] J. Tian and K. Ma, "A MCMC approach for Bayesian super-resolution image reconstruction," in *Proc. IEEE ICIP*, Sep. 2005, vol. 1, pp. 45–48.
- [11] H. Stark and P. Oskoui, "High resolution image recovery from image-plane arrays, using convex projections," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 6, no. 11, pp. 1715–1726, Nov. 1989.
- [12] M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP, Graph. Models Image Process.*, vol. 53, no. 3, pp. 231–239, May 1991.
- [13] W. H. Richardson, "Bayesian-based iterative method of image restoration," *J. Opt. Soc. Amer.*, vol. 62, no. 1, pp. 55–59, Jan. 1972.
- [14] L. B. Lucy, "An iterative technique for the rectification of observed distributions," *Astron. J.*, vol. 79, no. 6, pp. 745–754, Jun. 1974.
- [15] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Stat. Soc., Ser. B*, vol. 39, no. 1, pp. 1–38, 1977.
- [16] A. Fruchter and R. N. Hook, "Dithering in the rain," *ST-ECF Newsl.*, vol. 14, pp. 9–10, Apr. 1997. [Online]. Available: http://sco.stsci.edu/newsletter/PDF/1997/newsletter.4_97.pdf
- [17] D. Tody, "The IRAF data reduction and analysis system," in *Proc. SPIE Instrum. Astronomy VI*, 1986, vol. 627, p. 733. [Online]. Available: <http://iraf.noao.edu/>
- [18] A. M. Koekemoer *et al.*, "HST Dither Handbook," in *Space Telescope Science Institute (STScI)*, Version 2, Baltimore, MD, 2002. [Online]. Available: http://www.stsci.edu/instruments/wfpc2/Wfpc2_driz/dither_handbook.html
- [19] R. N. Hook, Dec. 2003, Space Telescope Science Institute. private communication.
- [20] A. Fruchter and R. N. Hook, "Drizzle: A method for linear reconstruction of undersampled images," *Publ. Astron. Soc. Pacific*, vol. 114, no. 792, pp. 144–152, Feb. 2001.
- [21] W. H. Press *et al.*, in *Numerical recipes in C & C++* 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, Jan. 2002. [Online]. Available: <http://www.nr.com>
- [22] J. Núñez and M. Merino, "Super-resolution of remotely sensed images using drizzle and wavelets," in *Proc. 25th Asian Conf. Remote Sens.*, Chiang Mai, Thailand, 2004, pp. 262–269.



Maria Teresa Merino received the degree in physics from the University of Barcelona (UB), Barcelona, Spain, in 2002 and the M.Sc. degree in remote sensing from Institut d'Estudis Espacials de Catalunya, Barcelona, Spain, in 2004. She is currently working toward the Ph.D. degree at the Department of Astronomy and Meteorology, UB, Barcelona.

Since 2004, she has been working on image processing techniques, with respect to both astronomy and remote sensing fields, and has a special interest in super-resolution and image deconvolution.



Jorge Núñez received the degree and the Ph.D. degree in physics, both from University of Barcelona (UB), Barcelona, Spain, in 1975 and 1981, respectively.

Since 1984, he has been a Permanent Lecturer with the Department of Astronomy and Meteorology, UB, Barcelona. In 1998, he was elected member of the Academy of Sciences and Arts of Barcelona. Currently, he is conducting research on image processing including image reconstruction and restoration using Bayesian techniques and data

fusion with applications in astronomy, medicine, and remote sensing.