Superresolution Enhancement of Hyperspectral CHRIS/Proba Images With a Thin-Plate Spline Nonrigid Transform Model

Jonathan Cheung-Wai Chan, Jianglin Ma, Pieter Kempeneers, and Frank Canters

Abstract—Given the hyperspectral-oriented waveband configuration of multiangular CHRIS/Proba imagery, the scope of its application could widen if the present 18-m resolution would be improved. The multiangular images of CHRIS could be used as input for superresolution (SR) image reconstruction. A critical procedure in SR is an accurate registration of the low-resolution images. Conventional methods based on affine transformation may not be effective given the local geometric distortion in high off-nadir angular images. This paper examines the use of a nonrigid transform to improve the result of a nonuniform interpolation and deconvolution SR method. A scale-invariant feature transform is used to collect control points (CPs). To ensure the quality of CPs, a rigorous screening procedure is designed: 1) an ambiguity test; 2) the m-estimator sample consensus method; and 3) an iterative method using statistical characteristics of the distribution of random errors. A thin-plate spline (TPS) nonrigid transform is then used for the registration. The proposed registration method is examined with a Delaunay triangulation-based nonuniform interpolation and reconstruction SR method. Our results show that the TPS nonrigid transform allows accurate registration of angular images. SR results obtained from simulated LR images are evaluated using three quantitative measures, namely, relative mean-square error, structural similarity, and edge stability. Compared to the SR methods that use an affine transform, our proposed method performs better with all three evaluation measures. With a higher level of spatial detail, SR-enhanced CHRIS images might be more effective than the original data in various applications.

Index Terms—CHRIS/Proba, hyperspectral images, multiangle images, nonrigid transform, superresolution (SR) image reconstruction, thin-plate spline (TPS).

I. INTRODUCTION

CHRIS/Proba, which stands for Compact High Resolution Imaging Spectrometer onboard the Project for On-board Autonomy, acquires multiple observations of the same scene at five different angles $(+55°, +36°, 0°, -36°, -55°)$ with respect

TABLE I
SPECTRAL BANDS OF CHRIS/PROBA (IN NANOMETERS)

Band no.	Bandwidth (Bandcenter)
1	438 – 447 (442)
2	486 – 495 (490)
3	526 – 534 (530)
4	546 – 556 (551)
5	566 – 573 (570)
6	627 – 636 (631)
7	656 – 666 (661)
8	666 – 677 (672)
9	694 – 700 (697)
10	700 – 706 (703)
11	706 – 712 (709)
12	738 – 745 (742)
13	745 – 752 (748)
14	773 – 788 (781)
15	863 – 881 (872)
16	891 – 900 (895)
17	900 – 910 (905)
18	1002 – 1035 (1019)

to the fly-by positions. In mode 3 (land channel), CHRIS acquires 18 bands between 0.4 and 1 μm. Table I lists the spectral configuration of CHRIS acquisition in mode 3. The spatial resolution of CHRIS at nadir reaches 18 m, which is the highest resolution as far as spaceborne hyperspectral data are concerned. The multiangular property is useful for capturing landscape features with spectral characteristics that vary with direction. Angular spectral information has provided new insights in various applications and has been demonstrated to be effective in physical parameter extraction and more generic land-cover mapping tasks [1], [2]. Given the multiple observations and hyperspectral-oriented configuration, CHRIS's application opportunities would increase even further if the present spatial resolution could be improved. Superresolution (SR) image reconstruction is a technique which generates a high-resolution (HR) image from several low-resolution (LR) images. Various SR techniques have been tested and applied to remote sensing images such as Landsat [3], [4], SPOT [5], Quickbird [4], multiangular thermal data [6], and hyperspectral data [7]. In the case of CHRIS, its multiviewing images can be treated as LR input data for producing a HR image using SR methods.

Recent studies related to SR image reconstruction and multiviewing angular images seem to support this idea [6], [8]–[10]. In [6], projection onto convex sets (POCS), a widely used SR technique, was applied to simulated multiangle images. Different aspects were investigated such as the influences of angle, misregistration, noise, and the number of frames that can be acquired in forward looking or back scanning positions.

Their findings suggest that SR image enhancement is possible with multiple off-nadir images, but it is reckoned that the viewing angle cannot be too high (less than 48°). It was also concluded that SR is more effective with symmetric viewing (e.g., $\theta = -15°, 0°, +15°$) than with asymmetric viewing (e.g., $\theta = 0°, +15°, +30°$). A registration accuracy of less than 0.2 pixels is recommended. In [8], mainstream SR techniques were applied to multiangular CHRIS imagery. Initial assessments suggest that spatial enhancement can be achieved. In [9] and [10], SR-enhanced CHRIS images were further investigated on their usefulness for classification and unmixing applications. Their results show that substantially more classification detail can be obtained with SR-enhanced imagery and that fraction images are more accurate. Nevertheless, it was also pointed out that motion estimation with high-angle off-nadir CHRIS images is difficult and that SR implementation is not straightforward with many parameters to tune. In addition, computing time required by most SR algorithms is high.

Accurate registration of LR images is a prerequisite for the success of any SR implementation. A registration procedure can be divided into four steps: 1) feature detection; 2) feature matching; 3) transform model estimation; and 4) image resampling. Steps 1 and 2 are related to detection of correlated feature pairs [control points (CPs)] in the source and the target images. Once feature pairs have been identified, the target image can be registered with either a linear or nonlinear transform model. Global linear transform models (translation, rotation, and affine scaling) may not be suitable in situations where local geometric distortion is prevalent. In order to improve conventional SR approaches for angular CHRIS imagery, we propose the use of a nonrigid (plastic) transform function which is more effective in dealing with local geometric deformation. For a nonrigid transform model to work well, adequate ground CPs are needed. An automatic feature detection and feature matching procedure based on scale-invariant feature transform (SIFT) and multicriteria outlier detection is developed. The proposed feature detection and nonrigid transform approach is tested with a Delaunay triangulation (DT)-based nonuniform interpolation and deconvolution SR method. The results are compared with several SR methods based on an affine transform. The rest of this paper is organized as follows. Section II provides the conceptual background for SR image reconstruction and an introduction to three popular approaches, namely, nonuniform interpolation and deconvolution, iterative back projection and projection on convex sets. Section III outlines our proposed registration method which makes use of SIFT for feature detection and three methods for outlier detection: an ambiguity test, the m-estimator sample consensus (MSAC), and an iterative method based on the statistics of random errors derived from a third-order polynomial fit. Section IV describes the method of evaluation and data preparation. Results are discussed in Section V, and conclusions are drawn in Section VI.

II. CONVENTIONAL REGISTRATION AND SR METHODS

A. Background

Closely related to the problem of image restoration, SR can be considered as a second generation of image restoration techniques that also change image dimension. Upscaling and interpolation techniques can be used to increase the size of an image. However, the quality of a single LR image is limited and interpolation based on only one undersampled image does not allow recovering of the lost high-frequency information. Hence, multiple observations of the same scene are needed. The idea of SR is to fuse a sequence of LR noisy blurred images to produce a higher resolution image. The objective is to achieve the best image quality possible from several LR images. However, the application of SR algorithms is possible only if aliasing exists, which means that there are subpixel shifts between the LR images. SR techniques can be applied either in the frequency domain or in the spatial domain. The idea of SR for remote sensing images was first suggested in the frequency domain [3] with its theory built upon Fourier transforms of LR images. However, SR approaches based on the spatial domain provide better flexibility for the modeling of noise and degradation and are more suitable for our purposes [11]. Numerous SR approaches have been proposed. Conventional methods include nonuniform interpolation and deconvolution, regularization reconstruction, POCS, hybrid approaches, and others [11]. SR can also be applied on a single-frame LR image [12], which is more commonly referred to as image scaling, interpolation, zooming, and enlargement.

An important step in SR is to model the relationship between the expected HR image and a set of LR images of the same scene. If X is the HR image, Y is the set of LR images, N is the noise vector, and there are p LR images, then the relationship between the HR image and the LR image sets can be formulated as

$$Y = HX + N \tag{1}$$

where

$$Y = \begin{bmatrix} y_1 \\ y_p \end{bmatrix} \quad H = \begin{bmatrix} D_1 B_1 M_1 \\ D_p B_p M_p \end{bmatrix} \quad N = \begin{bmatrix} n_1 \\ n_p \end{bmatrix}.$$

For $1 \leq k \leq p$, y_k stands for the kth LR image written in lexicographical notion as the vector $y_k = [y_{k,1}, y_{k,2}, \ldots, y_{k,N}]^{\mathrm{T}}$, where N is the number of pixels in the image. X stands for the HR image, D_k is the subsampling matrix, B_k represents the blur matrix, M_k is a warp matrix, and n_k is the image noise. Fig. 1 shows this relationship between the LR images and the desired HR image in a flowchart as described by (1). This mathematical formulation of the SR problem is well suited for the description of image degradation, i.e., blurring, noise and subsampling.

B. Affine Transform Registration

The main procedure of SR consists of three steps: registration, interpolation, and restoration. Registration is an important step and a prerequisite for successful SR implementation. An intensity-based approach which considers intensity variations of the same location in LR image pairs is a popular choice for conventional SR methods. The observed image data with a pair of images $f(x, y)$ and $g(x, y)$ can be modeled as

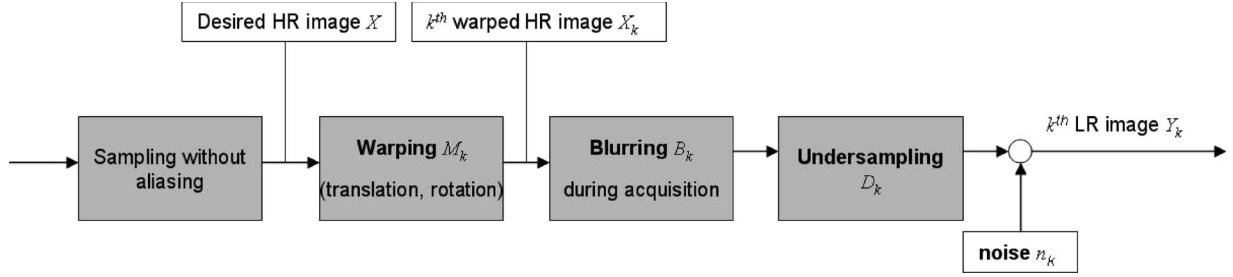$$f(x, y) = g(m_1 x + m_2 y + t_x, m_3 x + m_4 y + t_y). \tag{2}$$

Fig. 1. Flowchart showing the relationship between the desired HR image and the LR images.

The difference between $f(.)$ and $g(.)$ can be approximated using a first-order truncated Taylor series expansion. Then, the coefficients (m_i) can be calculated directly with the minimum root mean square error method. The error function being minimized, which is only an approximation, is usually further improved by a Newton–Raphson style iterative scheme. However, the derivatives of the first-order Taylor series have finite support thus limiting the motion flexibility of the estimation. To alleviate this problem, a coarse-to-fine scheme can be adopted to contend with large motions. For more details of the method, the reader is referred to [13]. Motion estimation is the first step of SR and the affine transform can be incorporated into different SR approaches.

C. Conventional SR Methods

1) Nonuniform Interpolation and Deconvolution: Interpolation and deconvolution methods treat SR as a resampling problem where a HR image is generated by interpolating the values of the LR pixels after their corresponding positions on the HR image have been determined. As the relative shifts between the LR images are arbitrary, it is natural that the interpolation is nonuniform. The final step is deconvolution, which is a deblurring process, to remove the effect of the point spread function (PSF) [14].

The nonuniform interpolation and deconvolution SR approach is the most intuitive method. Compared to other techniques, this method is cheap in computational costs and is easy to implement. However, since the errors of the interpolation process are not accounted for in the deconvolution phase, it does not guarantee an optimal solution. The approach is also restricted to cases where blur and noise effects are constant over the lower resolution images. Hence, the use of degradation models is limited.

2) Iterative Back Projection (IBP): First proposed in [15], IBP is based on a similar idea as computer-aided tomography, where a 2-D object is reconstructed from 1-D projections. The method involves a registration procedure, an iterative refinement for displacement estimation, and a simulation of the imaging process (the blurring effect) using a PSF.

The approach begins by guessing an initial HR image. The initial HR image can be generated from one of the LR images by decimating the pixels. The HR image is then down-sampled to simulate the observed LR images. Subtractions are then made between the simulated and the observed LR images. If the initial HR image is the real observed HR image, then the simulated LR images and the observed LR images will be identical and their differences zero. In the case of nonzero, the computed

difference can be "back-projected" to improve the initial guess. The back-projecting process is repeated iteratively to minimize the difference between the simulated and the observed LR images, and subsequently produces a better HR image. The iterative procedure is described by

$$X^{n+1} = X^n - G^{BP}(HX^n - Y). \tag{3}$$

In (3), G^{BP} represents the back-projection filter, X^{n+1} is the improved HR image at the $(n+1)$th iteration, and X^n is the HR image at the nth iteration.

IBP is intuitive hence easy to understand. However, its ill-posed nature means that there is no unique solution. The choice of back-projection filter is arbitrary, and it is more difficult to incorporate prior information compared with other SR approaches like regularization methods.

3) POCS: The POCS method for SR was first suggested in [16] and has been examined for multilooking imagery in [6]. It is another iterative method that uses *a priori* knowledge to impose constraints on a HR image. Every constraint must be defined as a closed convex set C. The HR image is known *a priori* to belong to the intersection C_s of m closed convex sets $C_1, C_2, \ldots, C_m$, and

$$C_s = \bigcap_{\alpha=1}^{m} C_\alpha \tag{4}$$

where C_s is found by iteratively computing projections onto the convex sets

$$f_{n+1} = P_m P_{m-1}, \ldots, P_1 f_n. \tag{5}$$

The projection operator P_i in (5) maps the current estimate f to the closest point in the set C_α using f_0 as an arbitrary starting point. Two constraints are commonly used for the formulation of POCS. The first constraint is defined as $C_D = \{F : |r| < \delta\}$ where the residual r is used to update the targeted HR image [18]. This residual r is the difference between a LR pixel value $g(i, j)$ and a corresponding blurred region in the HR estimate. It can be expressed as

$$r(i,j) = g(i,j) - \sum_k \sum_l f_n(k,l) h(i,j;k,l) \tag{6}$$

where h is the PSF operator applied on the corresponding location of the estimated HR image, and k and l are the

locations of the PSF matrix in the HR image. The projection of f onto the constraint set C_D is then

$$f_{n+1} = P_D f_n = \begin{cases} f_n + h(r - \delta)/\sum_k \sum_l h^2, & \text{if } r > \delta \\ f_n, & \text{if } -\delta \leq r \leq \delta \\ f_n + h(r + \delta)/\sum_k \sum_l h^2, & \text{if } r < \delta. \end{cases} \tag{7}$$

The parameter δ is the threshold used for the residual r which is considered large enough to cause an update (or change) to the estimated HR image.

The second constraint is called the energy consistent constraint [17] which can be defined as $C_E = \{F : 0 < f < 1\}$. The projection of f onto the constraint set C_E is then

$$f_{n+1} = P_E f_n = \begin{cases} 1, & \text{if } f > 1 \\ f, & \text{if } 1 \geq f \geq 0 \\ 0, & \text{if } f < 0. \end{cases} \tag{8}$$

III. PROPOSED REGISTRATION AND SR METHOD

An accurate registration of all the LR images is critical for the success of SR enhancement. All SR methods described in Section II require a step of motion estimation—the shift between the reference image (nadir image) and the image to be registered (off-nadir images). Typically, either a feature-based or an area-based method is used to detect possible common feature points from the image pairs [18]. Then, an affine transform or a projective transform is used for registration. However, as most linear transform models are global in nature, they are not ideal for modeling local deformations expected in angular CHRIS images. An interesting alternative would be a nonrigid transform. However, since nonrigid transforms are not consistent, they are not readily applicable for all SR methods. A nonrigid transform is more suitable for interpolation-based SR methods. For popular SR methods such as IBP and POCS which are iterative in nature, a transform model with a reversibility property, like an affine transform, is more suitable.

A successful nonrigid transform relies on an adequate number of high-quality CPs. Both area-based and feature-based methods can be used to search for CPs. While previous experiences have shown that area-based methods produce higher accuracies than feature-based methods [19], the latter are more effective for multiangular imagery because image features are more robust to variation in viewing angle [20]. Therefore, we have chosen the SIFT, a promising feature-based method, for feature detection. A rigorous selection procedure consisting of different stages of CP selection and refinement has been designed to guarantee the quality of CPs (Fig. 2). Once the CPs have been defined, a thin-plate spline (TPS) nonrigid model is used for registration.

A. SIFT CPs Selection

SIFT is a feature-based method that maintains detection stability in situations of noisy input [21]. SIFT-based descriptors are considered as robust in blurring environments [22]. The detected CPs are also associated with a qualitative descriptor
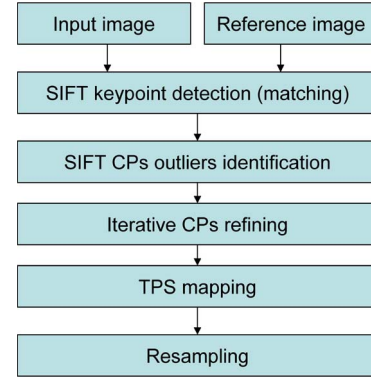

Fig. 2. Flowchart showing the processing chain for image registration.

which can be further utilized for point matching. The SIFT algorithm consists of four steps: scale-space extrema detection, keypoint localization, orientation assignment, and keypoint descriptor assignment. A detailed description of SIFT can be found in [21].

After keypoints have been identified and each keypoint descriptor has been calculated, keypoints are matched by the minimum distance method. As some of the matched keypoints are of poor quality or could be outliers, two criteria are used to filter them out. The first criterion is an ambiguity indicator

$$T_{\text{ambiguity}} = \frac{d_1}{d_2}. \tag{9}$$

In (9), d_1 is the Euclidean distance, calculated from the feature space spanned by keypoint descriptors, between a keypoint identified in the reference image and the nearest matching keypoint in the target image; d_2 represents the Euclidean distance between the same keypoint and the second nearest matching keypoint. If the values of d_1 and d_2 are similar, $T_{\text{ambiguity}}$ will be close to one indicating a high ambiguity. It means that for a certain keypoint in the input image, SIFT has detected two possible matching keypoints in the reference image. The matched pair will be deleted if $T_{\text{ambiguity}} > 0.75$. The CPs that pass the first screening are tested based on a second criterion.

The second criterion is MSAC, an improved version of the random sample consensus algorithm which has been widely used for rejecting outliers on matching points [23]. It is a spatial constraint that first estimates a projective model by using four randomly chosen points. A projective model is used because it is capable of modeling major geometric differences between multiviewing images. Then, the transform model is evaluated with regard to a fitting cost function

$$K = \sum_i \rho(e_i) \tag{10}$$

where i is the number of matched keypoint pairs and ρ is the error term defined as

$$\rho(e) = \begin{cases} e & if & e < T_m \\ T_m & if & e \geq T_m \end{cases}. \tag{11}$$

In (11), T_m is the threshold measured in pixels beyond which the keypoint pairs are considered outliers for the transform model. For a matched keypoint pair $(x_i, y_i) \leftrightarrow (x'_i, y'_i)$ with the calculated point position $(\underline{x_i}, \underline{y_i})$ and $(\underline{x'_i}, \underline{y'_i})$, the observed error function based on the transform model is defined as $e = (x_i - \underline{x_i})^2 + (y_i - \underline{y_i})^2 + (x'_i - \underline{x'_i})^2 + (y'_i - \underline{y'_i})^2$. The error term $\rho(e)$ is set to e when the error is lower than that of the threshold T_m. Every inlier has a different error value depending on how well it fits the estimated transform model. The threshold value T_m is defined to avoid extreme errors that could make some outliers to be regarded as inliers. The default setting for T_m is 64 pixels. The aforementioned procedure is repeated a certain number of times (default 500 times), and the best transform model, i.e., the one with the lowest fitting cost function value C, is identified. The CPs with values above T_m are removed. Finally, a new projective transform is estimated using all the keypoint pairs whose observed error function e values are lower than T_m.

B. CPs Refining

The CPs collected so far are potential CPs for TPS model estimation. A last step of refinement is to prune the points with large random errors. This is done by utilizing the statistical characteristics of the CPs. We first estimate a third-order polynomial transform model using all the CPs. A polynomial transform of third order is often used to deal with serious geometrical distortion. The residual stochastic characteristics of the third-order polynomial transform are well studied [24] and make the function suitable for prefitting. Using the residual distribution of CPs with respect to the polynomial transform, we implemented an iterative pruning with two constraints: a) CP pairs with residual errors larger than three standard deviations are discarded and b) pruning is halted when the three standard deviation threshold falls below 1 pixel on each side of the sample mean value.

This last refining routine can be summarized as follows.

1) Set up a third-order polynomial transform model using the least-squares method with all the CPs.
2) Calculate the noise-free point position, and obtain the model residual dx and dy in the horizontal and vertical direction, respectively. Compute the means and standard deviations of dx and dy. Eliminate the points whose dx and/or dy values are above three standard deviations.
3) Repeat the aforementioned procedure until either of the following conditions is fulfilled: The residuals in both directions are within three standard deviations for all remaining points, or the residuals fall below the -1- to 1-pixel threshold.

C. Image Warping

TPS is an interpolation function with a one-to-one mapping relationship between corresponding CPs. It is also the only spline model that can be cleanly decomposed into a global affine and a local nonaffine warping component. Thus, it can account for the local deformation caused by optical effects,
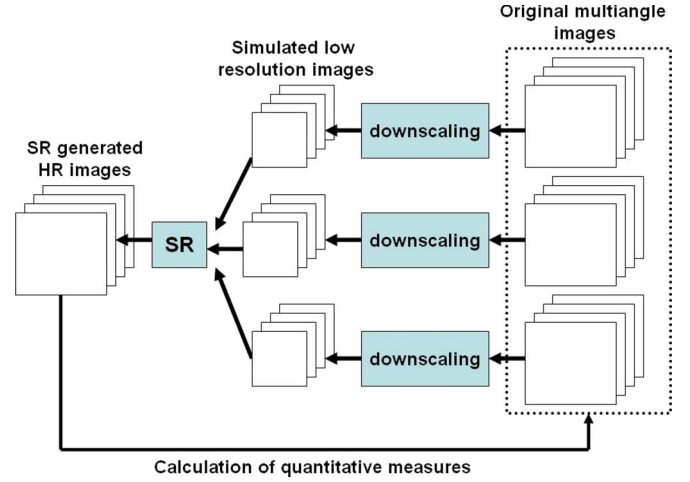

Fig. 3. Flowchart showing the validation procedure.

TABLE II
TYPE OF SR APPROACHES USED AND THE
RELATED REGISTRATION METHODS

Name	Registration Approach	Superresolution Approach
Bicubic	-	-
Non-uniform	Affine transform [13]	Delaunay-based non-uniform interpolation and deconvolution [2]
IBP	Affine transform [13]	IBP [15]
POCS	Affine transform [13]	POCS [6]
Improved non-uniform	Proposed TPS	Delaunay-based non-uniform interpolation and deconvolution [2]

relief change, and so on [25]. The TPS interpolation function can be expressed as

$$
\begin{pmatrix} f(x,y) \\ g(x,y) \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} h_{13} \\ h_{23} \end{pmatrix}
$$
$$
+ \begin{pmatrix} \sum_{i=1}^{N} F_i r_i^2 In r_i^2 \\ \sum_{i=1}^{N} G_i r_i^2 In r_i^2 \end{pmatrix} = \begin{pmatrix} x' \\ y' \end{pmatrix} \quad (12)
$$

where (x, y) represents the coordinate of the input image and $(f(x, y), g(x, y))$ that of the reference image. The coordinate (x_i, y_i) represents the detected CP position in the input image. The matrix $h_{11}, \ldots, h_{23}$ defines the affine transform. $r_i^2 = (x - x_i)^2 + (y - y_i)^2$ is the distance between (x, y) and (x_i, y_i), and F_i and G_i are the weights of the nonlinear radial interpolation function.

To solve (12) with N pairs of CPs, the following equilibrium constraints are imposed:

$$
\begin{cases} \sum_{i=1}^{N} F_i = \sum_{i=1}^{N} F_i x_i = \sum_{i=1}^{N} F_i y_i = 0 \\ \sum_{i=1}^{N} G_i = \sum_{i=1}^{N} G_i x_i = \sum_{i=1}^{N} G_i y_i = 0. \end{cases} \quad (13)
$$

With N pairs of CPs and the six equations in (13), we can solve the $2N + 6$ unknown parameters in the TPS model. A more
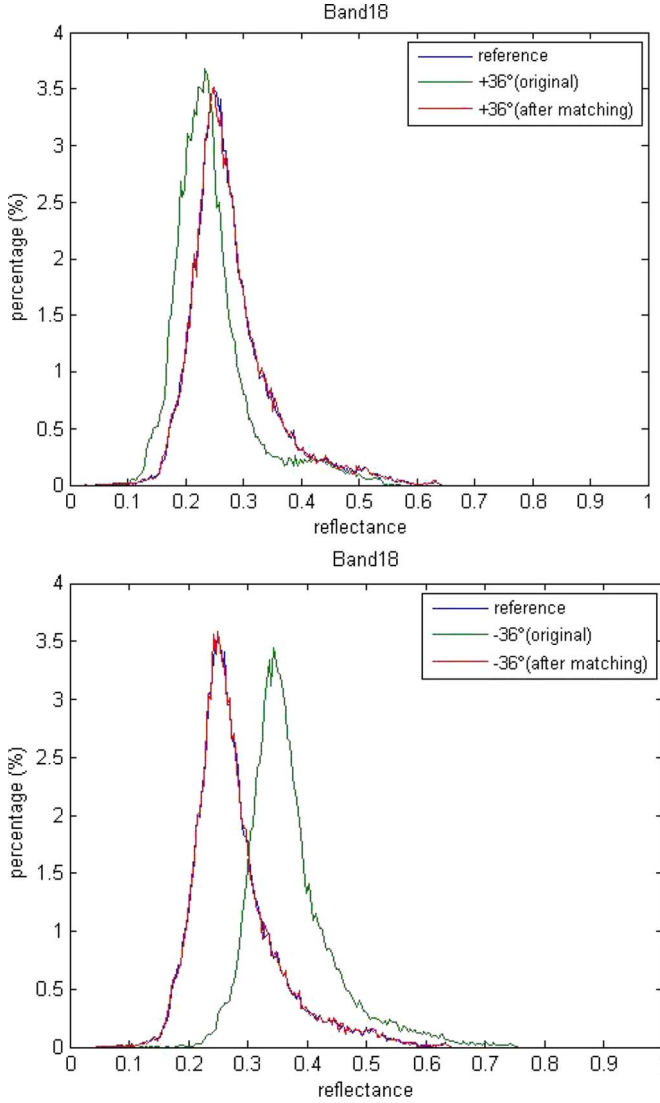
Fig. 4. Histogram matching of the off-nadir images to the nadir images. An example for band 18.

compact formulation for determining the unknown parameters can be expressed as

$$
\begin{bmatrix}
0 & 0 & 0 & 1 & 1 & \cdots & 1 \\
0 & 0 & 0 & u_1 & u_2 & \cdots & u_n \\
0 & 0 & 0 & v_1 & v_2 & \cdots & v_n \\
1 & u_1 & v_1 & 0 & r_{12}^2 Inr_{12} & \cdots & r_{1n}^2 Inr_{1n} \\
1 & u_2 & v_2 & r_{21}^2 Inr_{21} & 0 & \cdots & r_{2n}^2 Inr_{2n} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
1 & u_n & v_n & r_{n1}^2 Inr_{n1} & r_{n2}^2 Inr_{n2} & \cdots & 0
\end{bmatrix}
$$

$$
\times
\begin{bmatrix}
h_{13} & h_{23} \\
h_{11} & h_{21} \\
h_{12} & h_{22} \\
F_1 & G_1 \\
F_2 & G_2 \\
\cdots & \cdots \\
F_n & G_n
\end{bmatrix}
=
\begin{bmatrix}
0 & 0 \\
0 & 0 \\
0 & 0 \\
x_1 & y_1 \\
x_2 & y_2 \\
\cdots & \cdots \\
x_n & y_n
\end{bmatrix} . \quad (14)
$$

After the parameters have been estimated, the TPS-based warping is performed with a bilinear sampling function.

TABLE III
NUMBER OF CPs OBTAINED AFTER EACH STEP OF REFINEMENT

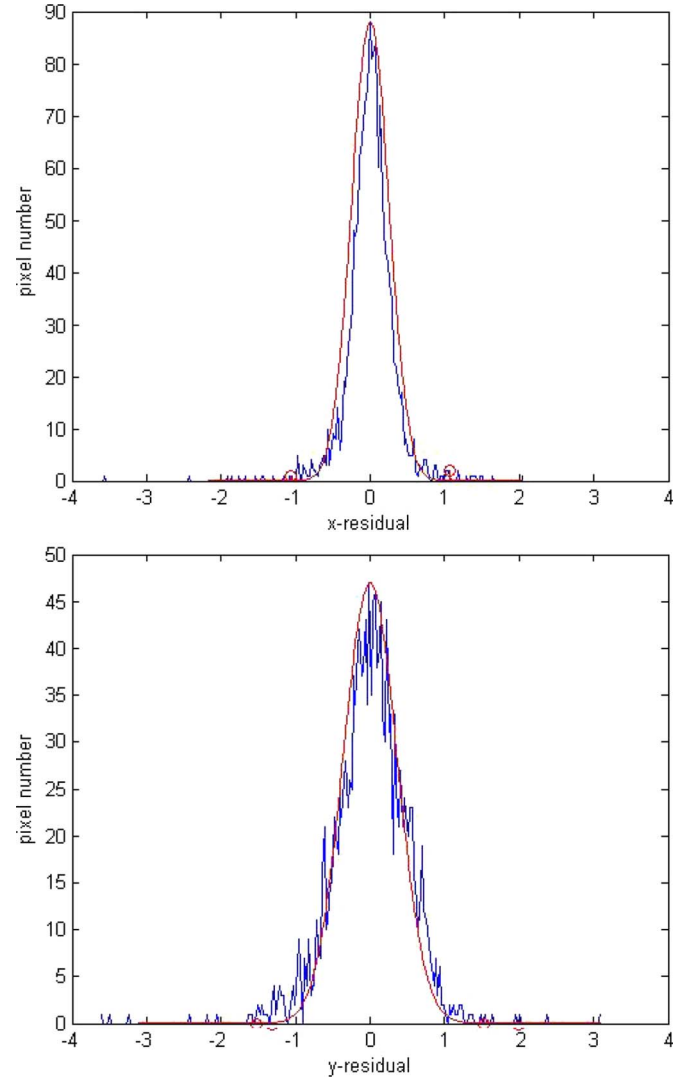| Angle	SIFT CP (ref/off-nadir)	Matched CPs after ambiguity test	After MSAC outlier detection	After iterative refining
+36°	6805/4499	1921	1639	1577
-36°	6805/6611	2270	1809	1768

Fig. 5. Example of the random error distribution for band 18 estimated by a third-order polynomial fit. The distribution resembles a (in red) normal distribution; the red circles indicate the threshold of three standard deviations. CPs with errors above the threshold value are not used.

D. DT-Based Nonuniform Interpolation and Deconvolution

DT-based interpolation has proven to be an effective interpolation technique when the spatial density and distribution of the projected data points is irregular [30]. The basic idea is to use triangular patches to locally fit the projected data. DT-based interpolation has the advantage of low complexity and stability due to its simplicity and the convexity of triangles [30]. We have chosen DT-based interpolation as part of our proposed SR method. To remove the effect of the PSF after interpolation, we adopted the Wiener deconvolution, one of the most popular deconvolution techniques. It makes minimum
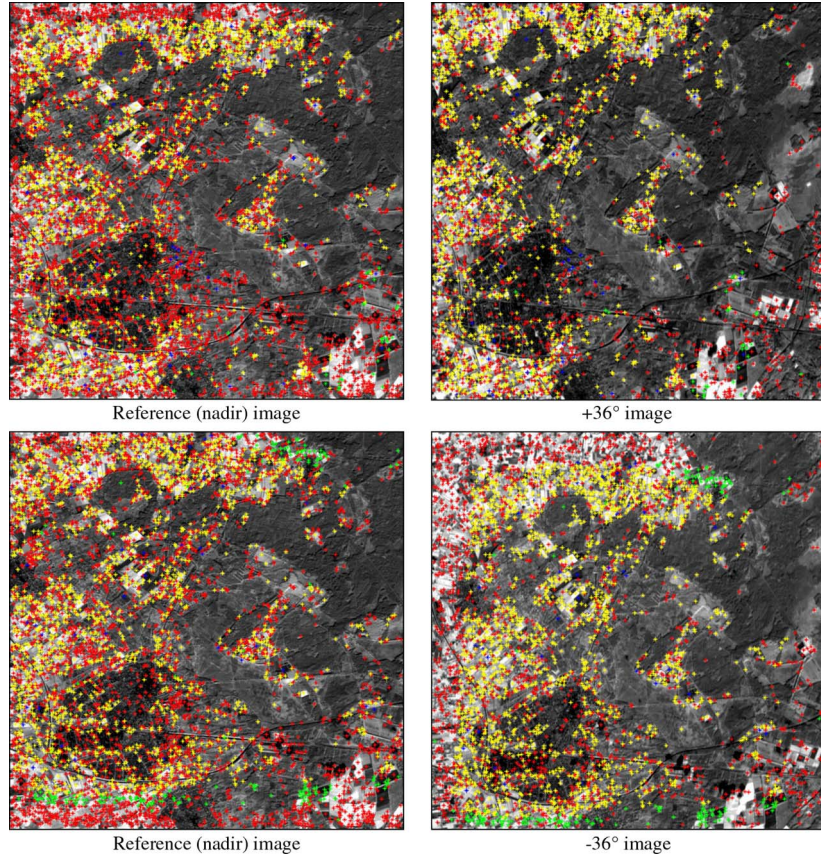
Fig. 6. Yellow markers represent the final CPs used for the nonrigid transform. Red markers represent the CPs eliminated after applying the first matching criterion—the ambiguity indicator ($T_{\text{ambiguity}}$). Green markers indicate the CPs eliminated after the MSAC test. Blue markers represent the CPs eliminated after the last iterative refining.

assumptions about image content and is robust to the impact of noise at frequencies with poor signal-to-noise ratios.

Fundamental limitations do exist for SR image enhancement. It has been shown that the number of linear systems and the volume of solutions grows fast with increments of the magnification factor; therefore, high magnification factors will result in overly smoothed solutions [31]. In the case of presence of noise and inadequate registration accuracy, the practical limit for an effective magnification factor is 1.6 according to [31]. To achieve an intended magnification factor (m), one study estimates the sufficient number of LR images as $4m^2$ [31]. Given the small number of input images available from angular CHRIS and the technical difficulty of implementing fractions of a scaling factor [D in (1)], we have chosen the decimation factor to be two in all our experiments, which is a popular choice for many SR studies using remote sensing data.

Usually, a PSF is used as an estimate for deblurring and convolution processes [B in (1)]. However, unless the blurring factor is known *a priori*, it is difficult to have an exact approximate as the real blurring tends to be spectrally and spatially variant. In [6], a 2-D circular Gaussian function was initially chosen to approximate the system PSF at the nadir position, assuming no distortion in the optics. Then, the PSFs of all off-nadir images were estimated with an affine transform function. It is common to assume a space-invariant blurring function to obtain an efficient and stable estimation [26].

In order to keep the comparison of several SR methods manageable, in this paper, the same PSF is used invariantly

for all the bands. We make the same assumption of no optical distortion so that the blurring is invariant within an image. The PSF chosen is a 2-D Gaussian filter (7 by 7 pixels) with a variance of 1.6 pixels.

IV. EVALUATION METRICS AND DATA

To evaluate the SR methods, we first simulate a set of LR images from the original CHRIS images. This is done with a convolution procedure followed by a downsampling with a factor of two (Fig. 3). Initial experiments show that incorporation of the high off-nadir images at $\pm 55°$ does not improve the SR results. Consequently, only the nadir and the $\pm 36°$ images are used. Hence, three LR images are generated. SR methods are then applied to the simulated LR images. Since the same convolution filter is used in the downscaling step of the simulation, the SR methods are assumed to have the "correct" blurring estimates in the reversing process.

SR can be applied on a single-frame LR image [12]. In order to give a critical assessment of the use of multiframe SR, we provide also the results obtained from single-frame SR using bicubic resampling. The single-frame SR is performed on the nadir images. The resulting SR images will be evaluated against the nadir images as well as the projected $\pm 36°$ images using the following quantitative measures suggested in [6] and [12].

1) Relative Mean Square Error (rmse): The rmse is the ratio of the mse of the single-frame bicubic image to that of the SR image. If x_i is the observed HR ground truth image of band
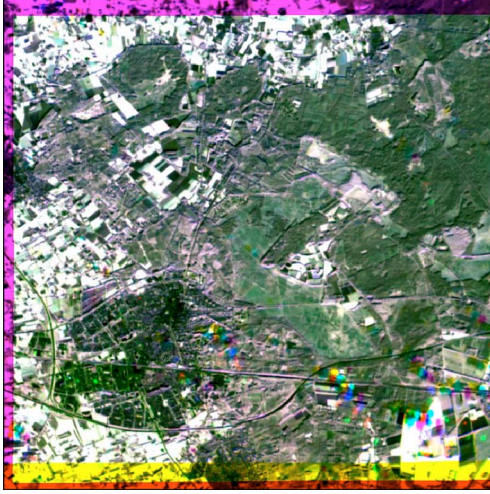
Fig. 7. Overlay of band 18 images taken from three viewing angles after registration with the TPS nonrigid transform. Results show the high quality of registration.

$i = 1, \ldots, n$, $\hat{x}_i$ is the SR image, and s is the bicubic interpolated image, then the rmse can be expressed as

$$10 \log \left(\frac{\sum\limits_{i=1}^{n} [x_i - s]^2}{\sum\limits_{i=1}^{n} [x_i - \hat{x}_i]^2} \right). \tag{15}$$

High rmse values represent improvement in image quality using multiframe SR methods as compared to the single-frame bicubic resampling SR.

2) Structural Similarity (SSIM): SSIM calculates the similarity in a linked local window by combining differences in mean, variance, and correlation [24]. The use of SSIM is recommended because when compared to mse, it has a higher correlation with visual degradation. SSIM is formulated as follows:

$$SSIM = \frac{(2m_z m_{\hat{z}} + c_1)(2\delta_{\hat{z}z} + c_2)}{(m_{\hat{z}}^2 + m_z^2 + c_1)(\delta_{\hat{z}}^2 + \delta_z^2 + c_2)}. \tag{16}$$

δ_z and $\delta_{\hat{z}}$ are the variances within the image windows z and $\hat{z}$, respectively. $\delta_{\hat{z}z}$ is the covariance between z and $\hat{z}$, and m_z and $m_{\hat{z}}$ are the respective means. The two constants c_1 and c_2 are included to avoid zeros in the denominator. The window size is set at 11 by 11 pixels, and the parameters c_1 and c_2 are 0.01 and 0.03, respectively. The local similarity measures are then averaged over all possible window offsets and all channels to obtain the mean structural similarity (MSSIM) for the whole image. The values of MSSIM range between 0 and 1. A high value means a high SSIM and hence better image quality.

3) Edge Stability (ES): Since blurring distortion is the most common distortion in SR, ES, which is most sensitive to blurring, is very suitable as an evaluation criterion [27]. ES is formulated as follows:

$$ES = \frac{1}{n} \sum_{i=1}^{n} (e\hat{z}_i - ez_i)^2. \tag{17}$$

In (17), ez is the original (HR) consecutive edge map, $e\hat{z}$ (SR) is the compared consecutive edge map, and n is the
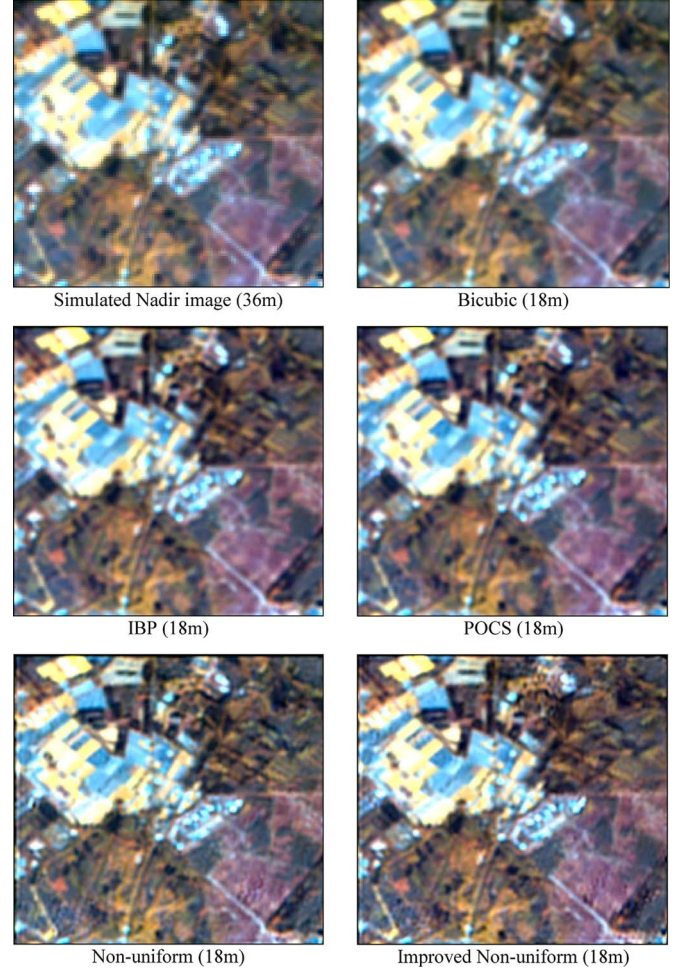

Fig. 8. Simulated SR results for four multiframe methods as compared to the single-frame bicubic resampling method.

number of edges that are detected in at least one of these two edge maps. ES measurement uses five canny edge detectors with different blur deviations to obtain an ordered set of five edge maps. The blur deviations used in this paper are 1.19, 1.44, 1.68, 2.0, and 2.38. Low ES values represent high image quality.

A CHRIS image set of Ginkelse Heide (The Netherlands) was acquired on October 20, 2007. In mode 3, the images have 18 bands between 0.4 and 1 μm. The image set is cloud free and has excellent quality. The spatial resolution of the nadir images is 18 m. The CHRIS images are first preprocessed to filter out noise and then calibrated to reflectance using algorithms described in [28] and [29]. LR images are then simulated from the atmospherically corrected original images using the procedure described in Section IV. A total of five SR methods are implemented and tested: the single-frame bicubic SR, three conventional SR approaches with affine transform, and the proposed method with nonuniform interpolation using the TPS transform (Table II).

V. RESULTS AND DISCUSSION

While CHRIS/Proba provides five acquisitions at angles $\pm 55°$, $\pm 36°$, and nadir, our initial experiments show that a better performance is obtained by using only three images: images acquired at $\pm 36°$ and nadir. This is echoed by the

TABLE IV
RESULTS OF THE QUANTITATIVE EVALUATION. FOR RMSE AND MSSIM, THE HIGHER THE VALUES THE BETTER; FOR ES, THE LOWER THE BETTER

		Bicubic	Non-uniform	IBP	POCS	Non-rigid
RMSE	nadir		0.7758	1.1657	0.5487	1.6150
	+36°		1.4106	1.9937	1.3089	2.2180
	-36°		1.9088	2.5509	1.6472	2.3799
	average		**1.3651**	**1.9034**	**1.1683**	**2.0710**
MSSIM	nadir	0.8951	0.9195	0.9176	0.9116	0.9280
	+36°	0.9134	0.9501	0.9513	0.9519	0.9575
	-36°	0.9170	0.9569	0.9585	0.9522	0.9613
	average	**0.9085**	**0.9422**	**0.9425**	**0.9386**	**0.9489**
ES	nadir	13.99	14.30	10.93	13.18	10.28
	+36°	17.87	14.70	12.57	14.29	11.28
	-36°	17.87	13.57	12.27	15.10	11.58
	average	**16.58**	**14.29**	**11.92**	**14.19**	**11.05**

simulated results in [6]. To eliminate the bias caused by different viewing angles, the off-nadir images are "corrected" by histogram matching (Fig. 4).

The SIFT method initially found 6805, 4499, and 6611 points on the nadir (reference) image, the +36° image and the −36° images, respectively. Table III lists the evolution of the number of CPs in each step. After the first screening based on the ambiguity test ($T_{\mathrm{ambiguity}}$), 1921 matched points remain for the +36° image, and 2270 for the −36° image. The second screening with MSAC further eliminates 14% (+36° image) and 16% (−36° image) of the matched points. The last iterative method eliminates CPs based on the distribution of random errors when applying a third-order polynomial fit. Fig. 5 shows the error distribution of the CPs for band 18. In the end, 1577 and 1768 CPs remain for the +36° image and the −36° image, respectively. Fig. 6 shows the geographical distribution of the CPs and their evolution through each step of the refinement. The final registration based on the TPS nonrigid transform is shown in Fig. 7. Visual inspection confirms that the registration is accurate.

Implementation of conventional SR methods described in Section II-C requires specific parameter settings. For all approaches, we use the bicubic image generated from the nadir image as the initial high-resolution image. To model the image formation process, image wrapping is used for motion estimation, a PSF is used to model the system blurring, and subsampling is fixed at 2. While image wrapping and PSF settings are the same for all SR methods, there are parameters specific to each method which require separate tuning and are defined empirically. For the IBP algorithm, the back-projection filter is the blurring kernel itself. The iteration process is continued until the change between two consecutive SR outputs is less than a threshold of 0.001 or until the maximum number of iterations (50 times) has been reached. For POCS, the parameter δ (a threshold for residual r) is set at 0.01.

SR methods are applied to simulated LR multiangle images to produce a HR image. The SR results of the simulation are shown in Fig. 8. The SR images show a clear increase in image detail. In general, image objects are better delineated. In terms of clarity and contrast, the multiframe SR methods are better than the single-frame bicubic resampling method. This confirms the results illustrated in [6], where a bilateral single-frame SR method was compared with the POCS method.

Three quantitative measures, namely, rmse, MSSIM, and ES, are calculated to assess image quality of the SR images (Table IV). The measures are calculated based on a compar-
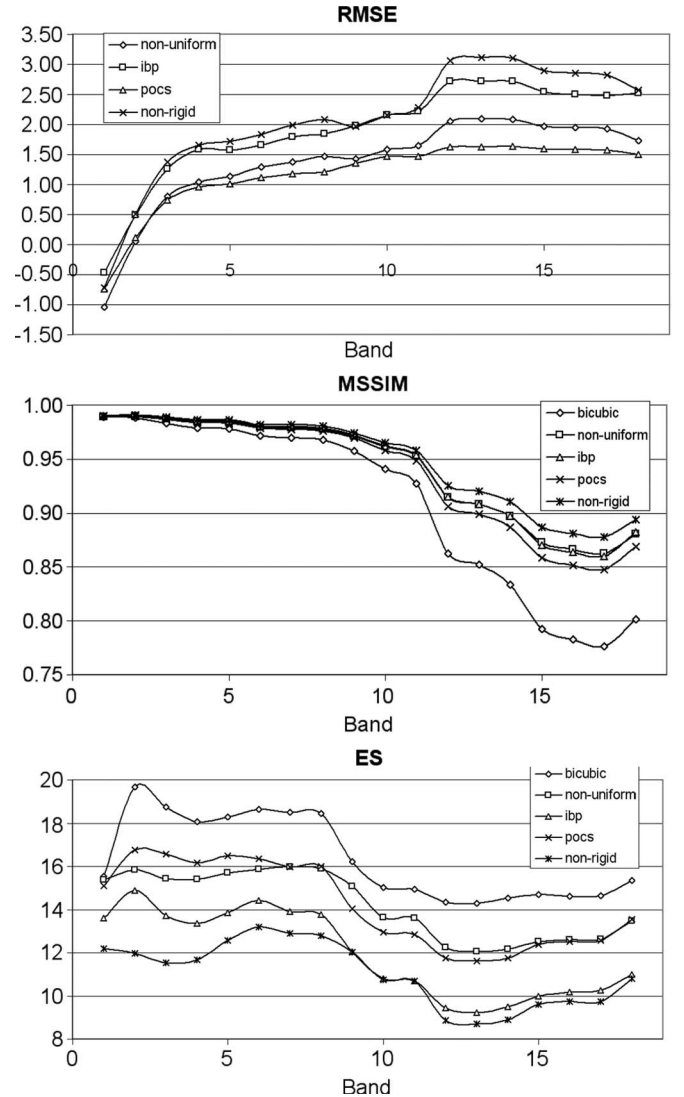

Fig. 9. Quantitative measures of image quality for different SR approaches and spectral bands.

ison between the SR image and each of the three original angular images. Each cell in Table IV represents averaged values over all the bands for each SR method. Higher rmse and MSSIM values represent higher quality, while for ES, lower values represent higher quality. Our results show that the proposed method consistently outperforms other SR methods which are based on an affine transform. The fact that the

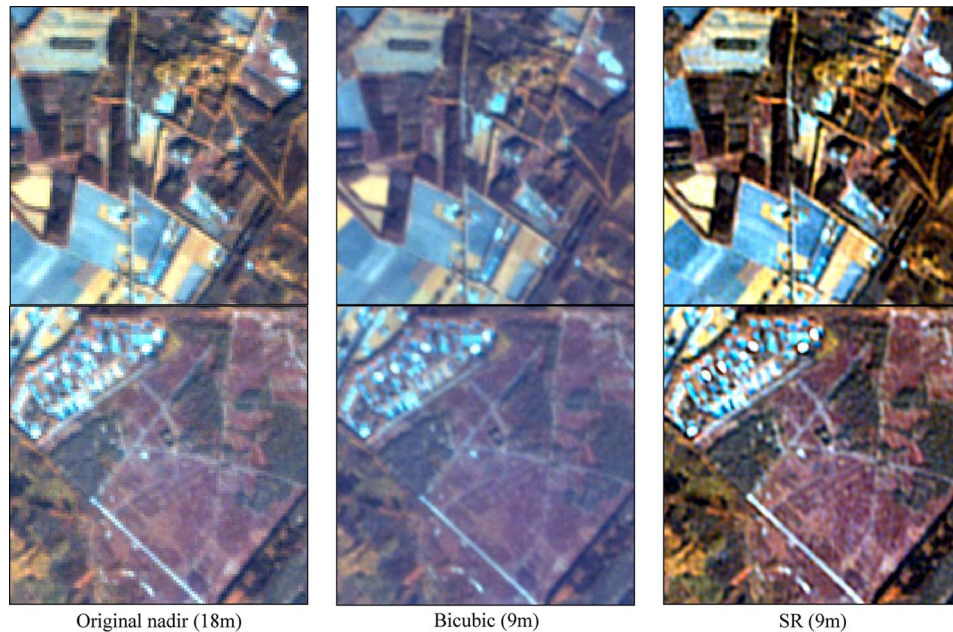Original nadir (18m)   Bicubic (9m)   SR (9m)

Fig. 10.   SR results obtained with real data using three angular CHRIS image sets. True color composites using bands 11, 5, and 2. The left column represents fragments of the original nadir image at 18-m resolution. The middle column shows the results obtained with the single-frame bicubic method. The right column corresponds to the SR results (9 m) obtained with the proposed method.

method scores best on all three quantitative measures gives a good indication that the SR images obtained with the proposed method are superior in quality. Multiframe SR methods also yield higher scores than single-frame SR on all three measures. Both qualitative visual inspection and quantitative assessment suggest that the use of multiframe SR is preferable to single-frame SR.

Fig. 9 shows the performance of different SR methods in terms of quality measures for each band. For each measure, the average value obtained for the three angular images is shown. For each of the SR methods that have been tested, the results indicate that the proposed method is superior for all bands. Multiframe SR methods also generate better results than the single-frame SR method for every band. It should be noted that the rmse results for band 1 are not better with the multiframe methods. This could be attributed to the systematic noise within the first few bands [30], [31].

In a last experiment, the proposed approach is applied on the original "real" CHRIS images (Fig. 10). For this experiment, the variance of the PSF is varied between 1.2 to 1.6 pixels to achieve the best results. Since we use a scaling factor of two, the dimensions of the image are doubled, with the number of pixels quadrupled. Apparent improvement in image quality is found with substantial increase in the amount of detail present in the HR image. Compared with the single-frame bicubic result, the HR image obtained from multiframe SR shows better contrast in terms of object boundary delineation and texture.

VI. CONCLUSION

Accurate registration is a prerequisite for successful SR image reconstruction. Multiviewing angular images pose particular difficulties for image registration due to substantial geometric distortion. To improve the performance of SR methods on angular CHRIS imagery, we proposed the use of a nonrigid transform function and a rigorous CP selection procedure to en-

sure accurate registration of multiangular imagery. Our results show that the proposed method improves the performance of a DT-based nonuniform interpolation applied in combination with a standard affine transform. The method also outperforms three other mainstream SR methods that make use of affine image registration. The SIFT-based CP selection procedure produces enough CP pairs to apply a TPS nonrigid transform. The outlier detection procedure guarantees a proper selection of CPs and a high-quality registration of off-nadir images.

Given the multiviewing and hyperspectral-oriented configuration of CHRIS/Proba, a higher spatial resolution will definitely broaden the scope of application of CHRIS data. Our results have shown that it is possible to enhance CHRIS imagery using SR methods. In this paper, we have used a spectral and spatial invariant PSF. In future efforts, it might be interesting to investigate the actual benefits of using a variant PSF. In addition, semiautomating the parameter tuning process would certainly increase the ease of use of some SR methods. For specific applications such as land-cover classification or spectral unmixing, more experiments are needed to understand the benefits of using SR-enhanced CHRIS imagery.

REFERENCES

[1] J. Verrelst, M. E. Schaepman, B. Koetz, and M. Kneubühler, "Angular sensitivity analysis of vegetation indices derived from CHRIS/PROBA data," *Remote Sens. Environ.*, vol. 112, no. 5, pp. 2341–2353, May 2008.
[2] R. Duca and F. Del Frate, "Hyperspectral and multiangle CHRIS–PROBA images for the generation of land cover maps," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 10, pp. 2857–2866, Oct. 2008.

[3] R. Y. Tsai and T. S. Huang, "Multipleframe image restoration and registration," in *Advances in Computer Vision and Image Processing*. Greenwich, CT: JAI Press, 1984, pp. 317–339.

[4] M. T. Merino and J. Núñez, "Super-resolution of remotely sensed images with variable-pixel linear reconstruction," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1446–1457, May 2007.

[5] C. Latry and B. Rouge, "Super resolution: Quincunx sampling and fusion processing," in *Proc. IGARSS*, Toulouse, France, 2003, pp. 315–317.

[6] A. Galbraith, J. Theiler, K. Thome, and R. Ziolkowski, "Resolution enhancement of multilook imagery for the multispectral thermal imager," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 9, pp. 1964–1977, Sep. 2005.

[7] T. Akgun, Y. Altunbasak, and R. M. Mersereau, "Super-resolution reconstruction of hyperspectral images," *IEEE Trans. Image Process.*, vol. 14, no. 11, pp. 1860–1875, Nov. 2005.

[8] J. C.-W. Chan, J. Ma, and F. Canters, "A comparison of superresolution reconstruction methods for multi-angle Chris/Proba images," in *Proc. SPIE Europe Remote Sens.*, Sep. 15–18, 2008, pp. 710 904-1–710 904-11.

[9] J. C.-W. Chan, J. Ma, P. Kempeneers, F. Canters, J. Vandenborre, and D. Paelinckx, "An evaluation of ecotope classification using superresolution images derived from Chris/Proba data," in *Proc. IGARSS*, Boston, MA, Jul. 6–11, 2008, vol. III, pp. 322–325.

[10] J. C.-W. Chan, J. Ma, L. Demarchi, T. Van de Voorde, and F. Canters, "Use of superresolution enhanced CHRIS/Proba images for land-cover classification and spectral unmixing," in *Proc. 6th EARSeL SIG Imaging Spectroscopy Workshop*, Tel Aviv, Israel, Mar. 16–19, 2009.

[11] 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.

[12] J. D. van Ouwerkerk, "Image super-resolution survey," *Image Vis. Comput.*, vol. 24, no. 10, pp. 1039–1052, Oct. 2006.

[13] S. Periaswamy and H. Farid, "Medical image registration with partial data," *Med. Image Anal.*, vol. 10, no. 3, pp. 452–464, Jun. 2006.

[14] H. Ur and D. Gross, "Improved resolution from subpixel shifted pictures," *CVGIP, Graph. Models Image Process.*, vol. 54, no. 2, pp. 181–186, Mar. 1992.

[15] M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP, Graph. Models Image Process.*, vol. 53, no. 3, pp. 231–239, May 1991.

[16] 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.

[17] A. M. Tekalp, M. K. Ozkan, and M. I. Sezan, "High-resolution image reconstruction for lower-resolution image sequences and space-varying image restoration," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, San Francisco, CA, Mar. 23–26, 1992, pp. III-169–III-172.

[18] B. Zitova and J. Flusser, "Image registration methods: A survey," *Image Vis. Comput.*, vol. 21, no. 11, pp. 977–1000, Oct. 2003.

[19] R. D. Eastman, J. L. Moigne, and N. S. Netanyahu, "Research issues in image registration for remote sensing," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2007, pp. 1–8.

[20] D. Capel and A. Zisserman, "A computer vision applied to super resolution," *IEEE Signal Process. Mag.*, vol. 20, no. 3, pp. 75–86, May 2003.

[21] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.

[22] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 10, pp. 1615–1630, Oct. 2005.

[23] T. Kim and Y. J. Im, "Automatic satellite image registration by combination of matching and random sample consensus," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 5, pp. 1111–1117, May 2003.

[24] H. J. Buiten and B. van Putten, "Quality assessment of remote sensing image registration—Analysis and testing of control point residuals," *ISPRS J. Photogramm. Remote Sens.*, vol. 52, no. 2, pp. 57–73, Apr. 1997.

[25] H. Chui and A. Rangarajan, "A new algorithm for non-rigid point matching," in *Proc. IEEE Conf. CVPR*, 2000, vol. 2, pp. 44–51.

[26] S. Farsiu, D. Robinson, and M. Elad, "Fast and robust multiframe super resolution," *IEEE Trans. Image Process.*, vol. 13, no. 10, pp. 1327–1344, Oct. 2004.

[27] J. F. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, no. 6, pp. 679–698, Nov. 1986.

[28] L. Guanter, L. Alonso, and J. Moreno, "A method for the surface reflectance retrieval from PROBA/CHRIS data over land: Application to ESA SPARC campaigns," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 12, pp. 2908–2917, Dec. 2005.

[29] L. Guanter, L. Alonso, and J. Moreno, "First results from the PROBA/CHRIS hyperspectral/multiangular satellite system over land and water targets," *IEEE Geosci. Remote Sens. Lett.*, vol. 2, no. 3, pp. 250–254, Jul. 2005.

[30] S. Lertrattanapanich and N. K. Bose, "High-resolution image formation from low resolution frames using Delaunay triangulation," *IEEE Trans. Image Process.*, vol. 11, no. 12, pp. 1427–1441, Dec. 2002.

[31] Z. Lin and H. Y. Shum, "On the fundamental limits of reconstruction-based superresolution algorithms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 1, pp. 83–97, Jan. 2004.
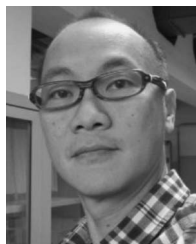
Jonathan Cheung-Wai Chan received the B.S. and M.S. degrees in geography from the Chinese University of Hong Kong, Shatin, Hong Kong, and the Ph.D. degree from the Center of Urban Planning and Environmental Management, Hong Kong University, Hong Kong, in 1999.

Between 1998 and 2001, he was a Research Scientist with the Geography Department, University of Maryland, College Park. From 2001 to 2005, he was with the Department of Electronics and Informatics, Vrije Universiteit Brussel (VUB), Brussels, Belgium, and Interuniversity Microelectronics Research Center, Leuven, Belgium, researching on topics related to humanitarian demining with remote sensing methods. Since 2005, he has been with the Department of Geography, VUB. He is part of the Cartography and GIS Research Group. He also teaches introductory and advanced courses in remote sensing. His main research interests are machine learning algorithms, land cover classification using hyperspectral data, superresolution enhancement of hyperspectral images, textural classification, change detection, and environmental monitoring using temporal remote sensing images. He conducts and supervises research on these topics within the framework of several nationally funded projects.
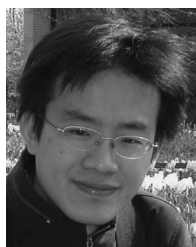
Jianglin Ma received the B.S. degree from Xidian University, Xi'an, China, in 2003, and the M.S. degree from the Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing, China, in 2006.

Currently, he is an Associate Researcher with the Cartography and GIS Research Group, Vrije Universiteit Brussel, Brussels, Belgium. His research interests are image processing and its application in remote sensing.

Pieter Kempeneers received the M.S. degree in electronic engineering from the University of Gent, Gent, Belgium, in 1994 and the Ph.D. degree in physics from the University of Antwerp, Antwerp, Belgium, in 2007.

He worked as a Researcher with the Department of Telecommunications and Information Processing (University of Gent) and in private industry (Mobile Communication Systems, Siemens). In 1999, he joined the Centre for Remote Sensing and Earth Observation Processes, Flemish Institute for Technological Research, working as a Scientist. Since 2008, he has been a Scientist with the Joint Research Centre, European Commission, Ispra, Italy. His research focus is on image processing, pattern recognition, and multi- and hyperspectral image analysis.

Frank Canters received the Diploma degrees in geography and applied computer science and the Ph.D. degree in sciences from the Vrije Universiteit Brussel, Brussels, Belgium, in 1999.

Currently, he is an Associate Professor and Head of the Geography Department, Vrije Universiteit Brussel, where he is also heading the Cartography and GIS Research Group. His main research interests are urban remote sensing, multiresolution image analysis, and modeling of uncertainty in spatial data.