High-Accuracy Urban DEM Using Permanent Scatterers

Daniele Perissin and Fabio Rocca

Abstract—**The permanent scatterers (PS) technique is a powerful operational tool that exploits a long series of synthetic aperture radar data for monitoring ground deformations with millimeter accuracy on a high spatial density grid of pointwise targets. The technique has been applied successfully to a number of applications, from subsidence and volcano monitoring to slow-landslide detection. This paper aims to analyze and demonstrate the positioning capability of the PS technique applied to the generation of urban elevation maps. The problem of the univocal identification of the PS position (discarding pixel-dependent sidelobes, both far and near) is addressed, and an easy and efficient solution is proposed. The results obtained in the Milan site allow the appreciation of the very high quality of an urban digital elevation model retrieved with the PS technique. The ground level of the city is identified with submeter accuracy, and elevated targets, where present, reveal building profiles. The estimated city street level (ranging ± 3m in 16×18 km^2) is then compared to those obtained with the same technique using a descending parallel track and an ascending one. Furthermore, the estimated PS elevation with respect to the ground has been connected to temperature-dependent elongations of high structures.**

Index Terms—**Interferometry, synthetic aperture radar (SAR), terrain mapping, urban areas.**

I. INTRODUCTION

DIFFERENTIAL synthetic aperture radar interferometry (DInSAR) is a remote sensing technology capable of measuring possible displacements of radar targets along the line of sight (LOS) by computing the difference of the phase values of two synthetic aperture radar (SAR) scenes gathered at different times over the same area of interest [1]–[5]. As well known, interferometric data can be used to recover high-resolution topographic profiles (acquiring data pairs from slightly different looking angles) [6] or to highlight possible surface-deformation phenomena (compensating the phase data for the local topography and the two acquisition geometries) [2]. Because SAR systems operate in the microwave domain (typically, the operating frequency is within the 1–10-GHz band), even subcentimeter range variations generate phase shifts that can be detected by the sensor, thus providing—at least theoretically—a powerful tool for precise geodetic surveys over large areas.

The main limitations of DInSAR technology are related to temporal and geometric decorrelation (i.e., reflectivity changes as a function of time and incidence angle of the acquisition [7]) as well as atmospheric artifacts (due to the different tropospheric and ionospheric conditions at the time of the SAR acquisitions [1]). DInSAR limitations can be overcome by adopting a multi-interferogram framework. The permanent scatterers (PSs) technique [8]–[10], which was developed in the late nineties at Politecnico di Milano, takes advantage of a long temporal series of SAR data acquired over the area of interest along the same (nominal) satellite orbit, in order to filter out atmospheric artifacts and to identify a subset of image pixels where high-precision measurements can be carried out. These pixels, which are almost unaffected by temporal and geometric decorrelation (usually but not necessarily corresponding to man-made objects), are called PSs [8], [9]. The technique has been applied successfully to a number of applications, from subsidence [9] and volcano monitoring [11] to slow-landslide detection [8], and it is currently used for both research and commercial activities.

Even if the theoretical accuracy of target positioning achievable by means of the PS technique has been already derived to be in the order of 1 m in the three dimensions [8], [10], a proof of it on a large scale has not been produced yet. This is mainly due to difficulties in the association of each coherent pixel to an actual target. In this paper, we show that an incorrect pixel-to-target association is a source of location errors. Moreover, more pixels relative to a single target give rise to the appearance of false PSs in different locations, which makes it impossible to validate the PS positioning capability. In this paper, a very simple and efficient method for associating pixels to PSs and discarding pixel-dependent sidelobes, both far and near, is proposed.

Exploiting 100 SAR images acquired by the European Space Agency (ESA) sensors European remote sensing (ERS) and Envisat over Milan (Italy), PS precise positions estimated through the interferometric phase processing are connected to the subpixel positions of the amplitude peak of the radar return of a high number of coherent targets, thus validating the theory. As a consequence, a very precise urban digital elevation model (DEM) can be obtained. The ground level of the city of Milan is identified with an expected accuracy of 1 m, and elevated targets, where present, reveal building profiles.

The suppression of location ambiguities allows the precise identification of urban radar targets. Thus, we discovered that many PSs lie on the ground. This fact can be fruitfully exploited for obtaining an estimate of the city ground level [urban digital terrain model (DTM)] with better accuracy than that achievable for a single target. By averaging the height of nearby PSs, the uncertainty can be reduced by a factor proportional to the square root of the number of PSs, thus reaching tens of centimeters.

Furthermore, a cross check of the PS elevation accuracy has been carried out by comparing three urban DTMs obtained from three different independent tracks over Milan (two descending and one ascending). Finally, the
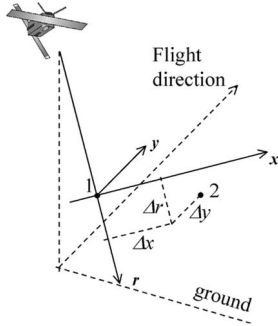
Fig. 1. Radar acquisition geometry of Master image. r axis: slant range direction; y axis: flight direction (azimuth); x axis: normal to the ry plane (cross-slant range direction). Target 1: first point target, which is located in the axes origin; Target 2: second point target with coordinates Δr, Δx, Δy.

temperature-dependent elongations of high structures have been studied as a function of the estimated elevation, showing good agreement with the theoretical behavior of composite steel–concrete buildings.

II. THREE-DIMENSIONAL (3-D) POSITIONING ACCURACY

The PS technique exploits a long series of SAR data with slightly different acquisition geometries in order to locate a target in 3-D space. The positioning accuracy achievable by means of the PS technique has been assessed in the order of 1 m in the three dimensions [8], [10]. Here, we wish to recall briefly the PS theoretical framework. We exploit the expressions derived for the coherent combination of ERS and Envisat data in [12].

Let us consider a pair of coregistered SAR images (that will be identified as Master and Slave) and the corresponding interferogram. The data are acquired with a common nominal geometry by two radar systems operating at two slightly different frequencies. Let B_n be the normal baseline and Δf_{DC} be the Doppler centroid (DC) frequency difference between the Master image, which is acquired at frequency $f_0 - \Delta f$, and the Slave, which is acquired at frequency f_0. In Fig. 1, two nearby point targets (1,2) are shown. The following orthogonal 3-D reference system, which is centered at the first point target, is then defined:

1) r axis: slant range direction of the Master image;
2) y axis: flight direction (azimuth) of the Master image;
3) x axis: normal to the ry plane (cross-slant range direction).

In this reference system, the target 2 has coordinates Δr, Δx, Δy, as shown in Fig. 1.

Under the hypothesis of no ground motion and almost no atmospheric phase contribution (close targets [1]), the interferometric phase difference between the two point scatterers is [12]

$$\Delta\phi_{12} = -\frac{4\pi}{c}\Delta f \Delta r + \frac{4\pi f_0 B_n}{cR_0}\Delta x + 2\pi\frac{\Delta f_{\mathrm{DC}}}{\mathrm{PRF}}\frac{\delta y}{\delta_{\mathrm{az}}} \quad (1)$$

where δy is the subpixel azimuth position of target 2 ($\Delta y = l_0\delta_{\mathrm{az}} + \delta y$, with δ_{az} as the azimuth sampling interval and l_0 as the integer number of pixels), assuming that target 1 is located at the center of the sampling cell. In (1), PRF is the pulse repetition frequency, R_0 is the sensor target distance, and c is the
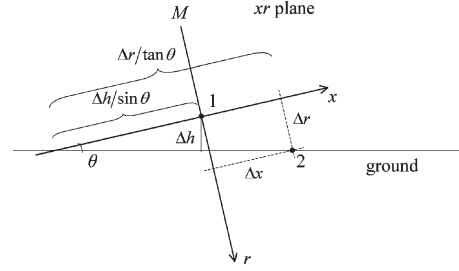

Fig. 2. Radar acquisition geometry of Master image (xr plane section). The cross-slant range distance Δx between the two targets 1 and 2 is decomposed into slant range Δr and height Δh, thus separating the contributions of flat terrain and topography. Angle θ is the incidence angle of the Master image.

light speed. Equation (1) highlights the possibility of locating a target in the 3-D space by means of SAR interferometry.

The cross-slant range distance between the two targets in the second term of (1) is usually decomposed into slant range and height, thus separating the contributions of flat terrain and topography. Defining the height difference between the two targets $\Delta h = h_2 - h_1$ as shown in Fig. 2, we have $\Delta x = (\Delta r / \tan\theta) + (\Delta h / \sin\theta)$ with θ incidence angle of the Master image. The slant range axis is sampled, thus $\Delta r = k_0\delta_{\mathrm{rg}} + \delta r$, where δ_{rg} is the slant range sampling interval, k_0 is the integer number of pixels, and δr is the subpixel residual term. Assuming that the pixel position is known, the terms of (1) can be partially compensated depending on the slant range distance. Therefore, (1) can be rewritten as

$$\Delta\phi_{12} = -\frac{4\pi}{c}\Delta f\delta r + \frac{4\pi f_0 B_n}{cR_0}\frac{\delta r}{\tan\theta}$$
$$+ \frac{4\pi f_0 B_n}{cR_0}\frac{\Delta h}{\sin\theta} + 2\pi\frac{\Delta f_{\mathrm{DC}}}{\mathrm{PRF}}\frac{\delta y}{\delta_{\mathrm{az}}}. \quad (2)$$

In the framework of a multi-interferogram analysis like the PS technique, the terms in (2) can be separated by exploiting their different behaviors in the parameter space Δf, B_n, Δf_{DC} and the three coordinates δr, Δh, and δy can be estimated. In the following text, we consider a data set formed by N ERS and M Envisat images. We assume Gaussian distribution for normal baseline and DC frequency values with variance $\sigma_{B_n}^2$ and $\sigma_{\Delta f_{\mathrm{DC}}}^2$, respectively.

Due to the 31-MHz frequency shift between ERS and Envisat, the first term in (2) changes by almost two cycles across a single-slant-range resolution cell; thus, it is extremely sensitive to the target slant range position within the resolution cell (ambiguity at about 5 m). Neglecting atmospheric effects, the variance of the associated position estimate can be calculated as [12]

$$\sigma_{\delta r}^2 = \left(\frac{c}{4\pi\Delta f}\right)^2\left(\frac{\sigma_{\Delta\phi_{\mathrm{ERS}}}^2}{N} + \frac{\sigma_{\Delta\phi_{\mathrm{Envisat}}}^2}{M}\right). \quad (3)$$

According to (3) with $N = 60$ and $M = 10$, a PS with multi-interferogram coherence $\gamma = 0.8$ ($\gamma = e^{-\sigma_{\Delta\phi}^2/2}$ [10]) can be positioned with a theoretical accuracy of about 20 cm.

The second term of (2) is usually neglected, and it is considered as a height error. Thus, the height variance is the sum of two contributions: The first one depends on the dispersion

of the normal baseline values [third term in (2)] and in first approximation can be quantified as [8], [10], [13]

$$\sigma_{\Delta h}^2 \simeq \left(\frac{\lambda R_0 \sin\theta}{4\pi}\right)^2 \frac{\sigma_{\Delta\phi}^2}{\sum_{i=1}^{N}(B_{n,i}-\bar{B}_n)^2}$$

$$\simeq \left(\frac{\lambda R_0 \sin\theta}{4\pi}\right)^2 \frac{\sigma_{\Delta\phi}^2}{N\sigma_{B_n}^2}. \tag{4}$$

For a multitemporal data set of $N = 60$ images with incidence angle $\theta = 23°$ and baseline dispersion $\sigma_{B_n} = 480$ m [10], from (4), a PS with coherence $\gamma = 0.8$ is localized with about 30 cm of elevation dispersion.

The second contribution to the height variance takes into account the bias derived from the residual slant range position δr [second term of (2)]. Given the height bias $\Delta h_{+} = \delta r \cos\theta$, its variance can be calculated assuming that the residual positions are uniformly distributed within the sampling cell δ_{rg} (~ 8 m for ERS satellites). Defining F_{int} as the possible oversampling factor used in the processing, the following expression holds:

$$\sigma_{\Delta h+}^2 = \frac{\delta_{\mathrm{rg}}^2}{12 F_{\mathrm{int}}^2} \cos^2\theta. \tag{5}$$

For an interpolation factor of 4, the standard deviation is 50 cm. Thus, most of the dispersion of the elevation values of the PS is actually due to the uncertainty in the range positioning of the scatterers, if *ad hoc* algorithms for precise range estimation of the PS are not applied. Summing (4) and (5), the resulting dispersion is 60 cm.

The variance of the azimuth position estimate [last term in (2)] is derived similarly to the height case [13] as

$$\sigma_{\delta y}^2 \simeq \left(\frac{\mathrm{PRF}\delta_{\mathrm{az}}}{2\pi}\right)^2 \frac{\sigma_{\Delta\phi}^2}{\sum_{i=1}^{N}(\Delta f_{\mathrm{DC},i}-\Delta\bar{f}_{\mathrm{DC}})^2}$$

$$\simeq \left(\frac{\mathrm{PRF}\delta_{\mathrm{az}}}{2\pi}\right)^2 \frac{\sigma_{\Delta\phi}^2}{N\sigma_{\Delta f_{\mathrm{DC}}}^2}. \tag{6}$$

Assuming a DC standard deviation $\sigma_{\Delta f_{\mathrm{DC}}} = 300$ Hz (ERS mono-gyro mode), pulse repetition frequency PRF $= 1680$ Hz, $N = 60$ images, and azimuth sampling interval $\delta_{\mathrm{az}} = 4$ m, a PS with coherence $\gamma = 0.8$ can be positioned in azimuth with a 30-cm dispersion.

III. PIXEL SELECTION

The end-to-end impulse response of a SAR system can be approximated with a cardinal sine in slant range and azimuth directions [14]. In range, the signal is in baseband, whereas in azimuth, it is modulated by the DC frequency. In both dimensions, the signal is sampled with a step slightly shorter than the resolution cell. Thus, within the main lobe of the impulse response fall at least two samples (neglecting a possible oversampling procedure).

Considering the example described at the beginning of the previous section, the expression of an acquired SAR image as a function of slant range pixel k and azimuth pixel l around

the location of target 2 (assumed pointwise) in the absence of clutter is

$$s(k,l) = Ae^{j\psi_0}\mathrm{sinc}\left(\frac{k\delta_{\mathrm{rg}}-\Delta r}{\rho_{\mathrm{rg}}}\right)$$

$$\times \mathrm{sinc}\left(\frac{l\delta_{\mathrm{az}}-\Delta y}{\rho_{\mathrm{az}}}\right)e^{j2\pi\frac{f_{\mathrm{DC}}}{\mathrm{PRF}}\left(l-\frac{\Delta y}{\delta_{\mathrm{az}}}\right)}. \tag{7}$$

In (7), A and ψ_0 are the amplitude and phase received by the sensor relative to target 2, respectively, and ρ_{az} and ρ_{rg} are the azimuth and slant range resolutions, respectively. Target 1 has been chosen as the reference as in the previous section. The last term of (7) takes into account the DC modulation in azimuth. From (7), it is immediate to recognize that the phase ψ_0 relative to the target is added to the phase of each sample of the impulse response. We are interested now in analyzing how this impacts on the localization process.

If the analyzed pixel is the closest (k_0, l_0) to the actual position of the target, then the interferometric phase is compensated by a factor proportional to $k_0\delta_{\mathrm{rg}}$ (flat terrain), (2) holds, and by means of a multi-interferogram analysis, the target can be localized as previously described. Considering the next pixel in range $(k_0 + 1, l_0)$, the compensation for flat terrain is increased by a factor proportional to δ_{rg}, which introduces a shift from the actual slant range target position. The first term in (2) is affected by an error equal to about 10 rad, aliasing the slant range position estimate. Moreover, the height is obtained with an additional error of $\Delta h_{++} = \delta_{\mathrm{rg}}\cos\theta$. For ERS typical values, the height error is more than 7 m/pixel.

Considering the next pixel in azimuth $(k_0, l_0 + 1)$, the argument of the last exponential in (7) becomes $2\pi(f_{\mathrm{DC}}/\mathrm{PRF})(l_0 + 1 - \Delta y/\delta_{\mathrm{az}}) = 2\pi(f_{\mathrm{DC}}/\mathrm{PRF})(1 - \delta y/\delta_{\mathrm{az}})$. Consequently, the azimuth position to be estimated in order to correctly compensate the associated phase term [last in (2)] becomes $\delta y - \delta_{\mathrm{az}}$. Because of the computational cost, the search range of the azimuth position is normally confined to $\pm\delta_{\mathrm{az}}$. Whenever the term to be estimated is out of range, aside from a positioning error, an additional phase noise in images with high values of DC arises.

All the mentioned problems can be solved by selecting the pixel closest to the target position. To this aim, a robust solution is the selection of the local maxima of the incoherent time average of the image amplitude (the so-called reflectivity map). However, even if the local maxima extraction is a good improvement in discarding dependent pixels, the presence of sidelobes (in particular in urbanized areas) still prevents from reaching the theoretical accuracy of the PS technique. Note that the first sidelobe in slant range is about 14 m far from the main peak and the corresponding height error is around 13 m. In azimuth, the distance between the main lobe and the first sidelobe is about 8 m.

A possible solution for the presence of sidelobes is to reduce them by means, e.g., of spectral windowing. The drawbacks of spectral windowing are loss of resolution, decrease of signal-to-clutter ratio (SCR), and the fact that sidelobes can be reduced but not totally removed (at least for strong scatterers). Here, we present a very simple and efficient solution that identifies sidelobe-dependent pixels rather than suppresses sidelobes. The identification of dependent pixels then allows the selection of a unique pixel for each PS, thus avoiding positioning artifacts.

The solution we found for sidelobe-pixel identification in a multi-interferogram analysis derives from the observation that the phase received by the satellite impacts on the whole impulse system response. Thus, considering the slant range coordinate (impulse system response in baseband), the phase difference between the main lobe and a secondary lobe in a single image is either 0 or π rad whatever the received phase. If the target keeps on being seen by the satellite and the images are well coregistered, the entire main lobe phase history in all acquisitions differs from the phase history of the secondary lobe either 0 or π rad.

Let $\phi_{k_1,i}$ be the phase of the pixel of slant range index k_1 in the ith image and $\phi_{k_2,i}$ be the phase of a nearby iso-azimuth pixel of slant range index k_2. We choose ξ_{k_1,k_2} as the correlation index between the two pixels k_1 and k_2, which is defined as

$$\xi_{k_1,k_2} = \frac{1}{N}\left|\sum_{i=1}^{N} e^{j(\phi_{k_2,i}-\phi_{k_1,i})}\right| \quad (8)$$

where N is the number of images. The index ranges between 0 (total decorrelation) and 1 (perfect correlation). Pixels with a correlation index over a certain threshold are considered as part of the same target response, and only the one with greater amplitude is retained. The threshold value must take into account the number of available images. The greater the number of images, the lower can be the threshold (0.8 for $N = 60$ images is a reasonable conservative number).

Considering the azimuth coordinate, the only difference from the previous case consists in the DC modulation. The correlation between two iso-range pixels of azimuth indexes l_1 and l_2 is thus computed as

$$\xi_{l_1,l_2} = \frac{1}{N}\left|\sum_{i=1}^{N} e^{j\left[\phi_{l_2,i}-\phi_{l_1,i}-2\pi\frac{\Delta f_{DC,i}}{PRF}(l_2-l_1)\right]}\right|. \quad (9)$$

In (9), the difference between the phase histories of the two pixels is compensated for a phase term proportional to the DC variation and to the azimuth distance between them. Again, dependent pixels can be suppressed by eliminating the ones with lower amplitude.

We derive now the probability of discarding erroneously a pixel. We consider for the sake of simplicity the case of two iso-azimuth pixels (the iso-range case is almost identical). In order to consider two iso-azimuth pixels belonging to two different targets as dependent, the correlation index ξ (8) between them must be over the threshold ξ_{th}. This means that the difference between the phase histories of the two pixels [see (8)] has variance $\sigma^2_{\Delta\phi} = -2\ln(\gamma) < -2\ln(\xi_{\text{th}})$ ($\gamma = e^{-\sigma^2_{\Delta\phi}/2}$, [10]). The phase variance is the sum of the variances of the terms, depending on phase noise (σ^2_{n1} and σ^2_{n2}), atmospheric contribution ($\sigma^2_{\alpha1}$ and $\sigma^2_{\alpha2}$), different displacement σ^2_{disp}, and different 3-D subpixel location σ^2_{3D} of the two targets (independent for hypothesis), and is given as

$$\sigma^2_{\Delta\phi} = \sigma^2_{n1} + \sigma^2_{n2} + \sigma^2_{\alpha1} + \sigma^2_{\alpha2} + \sigma^2_{\text{disp}} + \sigma^2_{3-D}. \quad (10)$$

Assuming the two targets have no relative motion and neglecting atmospheric effects, in order to produce a correlation index $\xi = 0.8$ in $N = 60$ images, both targets must have coherence $\gamma \simeq 0.92$ and their 3-D subpixel locations must be the same
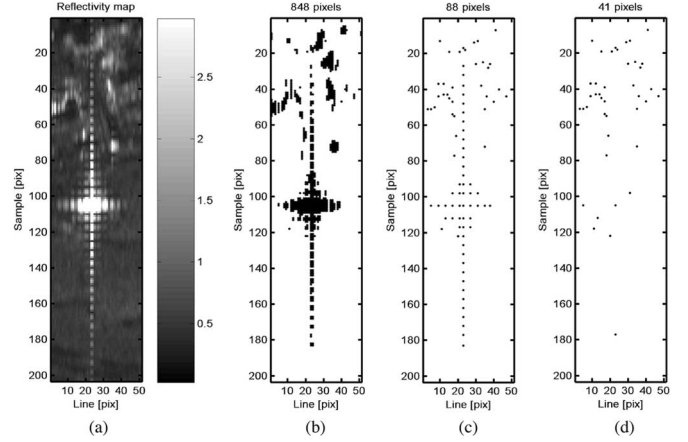

Fig. 3. Example of the PSC selection discarding dependent pixels and sidelobes. (a) Reflectivity map of the area around a very strong scatterer. (b) Eight hundred forty-eight pixels with sufficient amplitude stability have been selected. (c) Eighty-eight pixels survive after local maxima extraction. (d) Discarding sidelobes, the actual targets are only 41. The last visible lobe is more than 200 m far from the main one, and it leads to a height error of 180 m.

unless about 0.1 m (the calculation can be derived from the reversed (3), (4), and (6) with baseline dispersion $\sigma_{B_n} = 480$ m and DC standard deviation $\sigma_{\Delta f_{DC}} = 300$ Hz). Considering that azimuth and slant range subpixel positions are uniformly distributed within the respective sampling cells and height is equally probable between 0 and 30 m, the probability to have 3-D subpixel location coincident unless 0.1 m is 10^{-5}. Observing then the *a posteriori* PS coherence distribution, we can assess that the probability of discarding erroneously a pixel with a correlation threshold of 0.8 in 60 images is in the order of 10^{-6}.

In Fig. 3, a demonstrative example is reported. Fig. 3(a) shows the reflectivity map of the area around a very strong scatterer. In Fig. 3(b), 848 pixels with sufficient amplitude stability have been selected in the area of interest. In Fig. 3(c), 88 pixels survive after local maxima extraction. Fig. 3(d) shows that the actual targets are only 41 out of the initial 848. In this example, the last visible lobe is more than 200 m far from the main one, and it leads to a height error of 180 m.

Besides the improvement in discarding dependent pixels and sidelobes and the consequent univocal pixel-to-target association, the processing steps we proposed allow a conspicuous reduction of the computational costs of a PS analysis. In fact, as just shown, the number of PS candidates (PSC) can be drastically reduced at the beginning of the processing chain without any particular operation (images are required only to be coregistered).

IV. VALIDATION

The proposed technique permits the suppression of sidelobe-dependent pixels and thus makes it possible to associate each PS to an actual physical target. Under this condition, the theory previously described can be validated by exploiting scatterers whose slant range and azimuth subpixel positions can be derived from the main lobe peak of their radar response. The subpixel positions are related to the results of the interferometric analysis, showing the expected agreement.

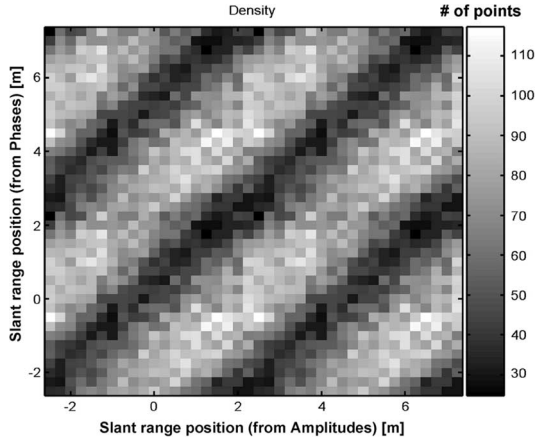The expression of the accuracy of the position estimate obtained by analyzing the amplitude of the radar response for

Fig. 4. Density of points of the scatter plot between the slant range PS location (y axis, $[m]$) derived from the phase processing [first term in (2)] and the position obtained from the amplitude peak of the radar response (x axis, $[m]$) of the analyzed PSs. For visualization purposes, both quantities are replicated at $+5$ m. The scatter plot dispersion is about 1 m.
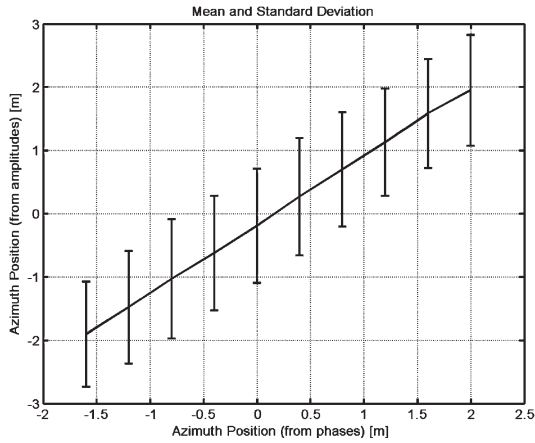

Fig. 6. PS ellipsoidal elevations estimated in Milan (Gauss–Boaga coordinates) with respect to a reference point in the middle of the scene.

Fig. 5. Average and standard deviation (sign at $\pm\sigma$) values of the PS azimuth position (obtained from the amplitude peak of the radar response) as a function of the position obtained from the phase processing [last term in (2)], both expressed in meters. The estimated dispersion is less than 1 m.

Fig. 7. PS estimated elevations as a function of the north coordinate. The average slope is in the order of 3 m/km.

strong and isolated scatterers can be expressed as a function of the system resolution ρ (see [12] or [15])

$$\sigma_{\text{pos}}^2 \approx \frac{\rho^2}{3N} \frac{\sigma_0 A_c}{\text{RCS}}. \tag{11}$$

For the slant range dimension ($\rho_{\text{rg}} \simeq 9$ m), exploiting $N = 60$ images, a PS with radar cross section RCS $= 500$ m^2 (e.g., a corner reflector with a 0.8-m side) and SCR $= 4$ (e.g., surrounding clutter with backscattering coefficient $\sigma_0 = 0$ dB distributed within the area of the ERS ground resolution cell $A_c = 125$ m^2) can be positioned with about 30 cm of uncertainty. Considering the azimuth dimension ($\rho_{\text{az}} = 5$ m), the standard deviation of the position estimate of the same PS is about 20 cm. Therefore, both measures of azimuth and slant range subpixel positions are comparable with the results of the interferometric analysis derived in Section II.

The PS analysis exploited for the validation has been carried out on a SAR data set including 100 SAR images of the town of Milan (track 208, frame 2691). The 100-SAR-image data set has been acquired by ERS-1, ERS-2, and Envisat
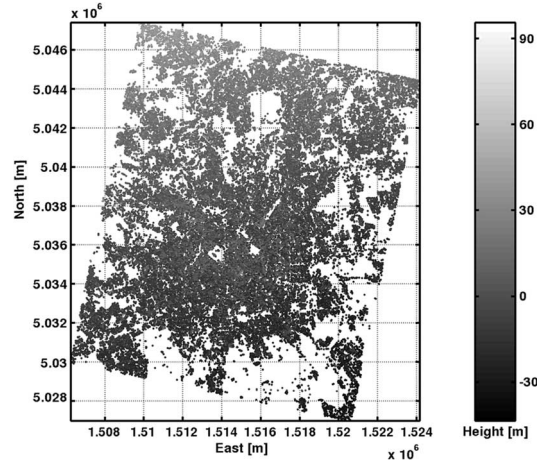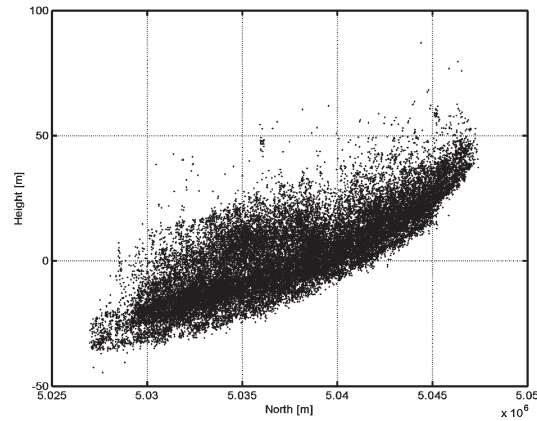
satellites from 1992 to 2005. The test area covers about 300 km^2. The ERS image acquired on January 14, 1997 is selected as the Master scene, and all the Slave acquisitions (ERS and Envisat) are resampled on the common Master grid, taking into account the different PRFs and sampling frequencies in the range direction. About 60 000 PSs have been detected with multi-interferogram coherence $\gamma > 0.7$.

In Fig. 4 [12], we show the correlation between the slant range PS location (y axis, $[m]$) derived from the phase processing [first term in (2)] and the slant range position obtained by extracting the peak of the radar response (x axis, $[m]$) of the analyzed PSs. The image reports the density of points of the scatter plot between the two measurements. As previously seen, the slant range ambiguity of the phase measure corresponds to 5 m. For visualization purposes, we replicated both quantities at $+5$ m. The good correlation found (dispersion about 1 m) is the first proof of the location capability of the PS technique.

Likewise, we proceed in the azimuth dimension. Fig. 5 shows the average and standard deviation (sign at $\pm\sigma$) values of the PS azimuth position (obtained from the amplitudes) as a function of the azimuth position obtained from the phase processing [last term in (2)], both expressed in meters. As predicted from the theory, the estimated dispersion is less than 1 m.

As a consequence, we expect to appreciate very accurate measures of the height of urban targets. In Fig. 6, ellipsoidal
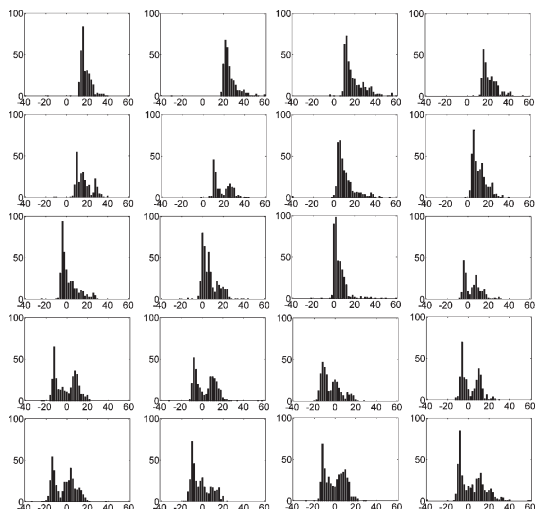
Fig. 8. PS elevation $[m]$ histograms, which are calculated on 1-km^2 areas. The top left image corresponds to northwest, and the bottom right one corresponds to southeast. From north to south, histograms are progressively shifted toward negative values, due to the slope of the plain on which Milan lies. The height distributions show different modes, and the first mode (lower elevation) has an amplitude that is always higher than the others.
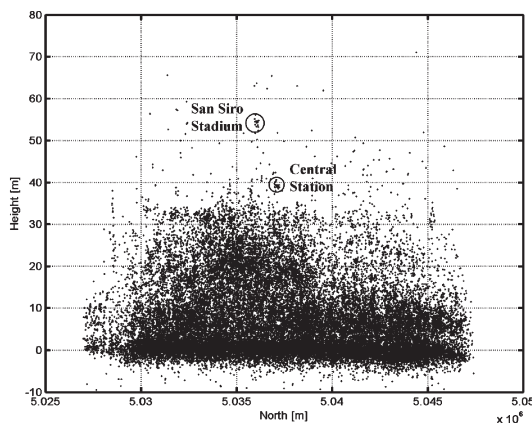

Fig. 9. North–south profile of Milan city after removing the low-pass topography. It is worth noting the thickening of high points around the center of Milan, where the buildings are higher and closer among themselves. Focusing the attention on details, some sets of adjacent points with high elevations describing high regular structures can be appreciated, e.g., the Central Station and San Siro Stadium.
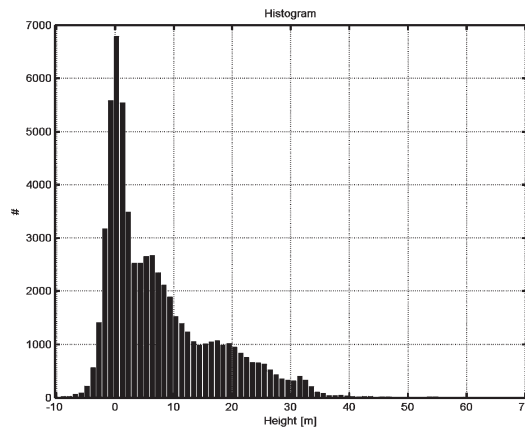

Fig. 10. Histogram of PS residual elevations (after removing the low-pass topography) in Milan. Different peaks are visible, highlighting recurrent heights (ground level, small houses, sheds, and high buildings). The peak width of targets on the ground ($\sim$1 m at -3 dB) is the actual demonstration of the PS's height accuracy.
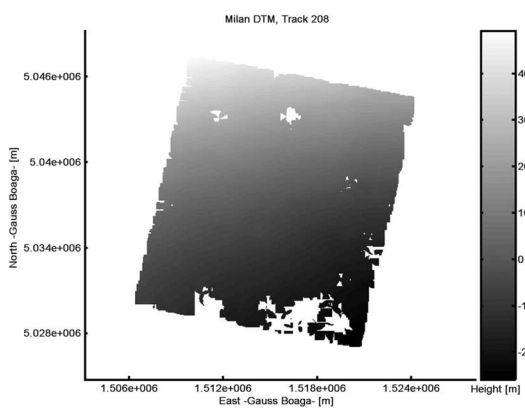

Fig. 11. PS DTM of Milan city.

elevations estimated in Milan are shown (in Gauss–Boaga coordinates). Milan is located on the plain formed by the Po river, which flows to the south of Milan approximately along the west–east direction. Thus, the plain slopes down to south, as visible in Fig. 6. The slope is more evident looking at the section of the city from north to south in Fig. 7. Its average can be quantified in the order of 3 m/km.

Looking at Figs. 6 and 7, two sets of points can be highlighted: targets at ground level following the general trend of the local topography and targets corresponding to roofs or other objects located on buildings at different elevations. A deeper analysis can be carried out by dividing the region in smaller areas and evaluating the distribution of the estimated elevations for each of them. In Fig. 8, some histograms are reported and calculated on 1 km^2 areas. The histograms have been ordered with respect to the location of the areas in the whole region: The top left image in Fig. 8 corresponds to northwest, and

the bottom right one corresponds to southeast. From north to south, the histograms are progressively shifted toward negative values due to the slope of the plain on which Milan lies. It is worth appreciating in Fig. 8 that the height distributions show different modes and that the first mode (lower elevation) has an amplitude that is always higher than the others. This means that the highest density of PSs is on the ground. The peak widths of the height distributions (histogram bins less than 2 m) are another proof of the positioning accuracy achievable with the PS technique.

V. URBAN DTM

Exploiting the high density of points at "street level," the low-varying topography of Milan can be recovered by estimating the lower envelope of the PS data. This step is very useful for two reasons. First, it is possible to obtain a very precise estimate of the ground-level slope. Second, we can get the PS elevation referred to by the local topography, namely, the actual height of the buildings.

The methodology we adopted to tackle the problem follows the observations reported at the end of the previous section. The region of interest is first divided into 1-km^2 areas. Then, the histograms of the estimated elevations are calculated in each area. The three modes with higher density are extracted for each
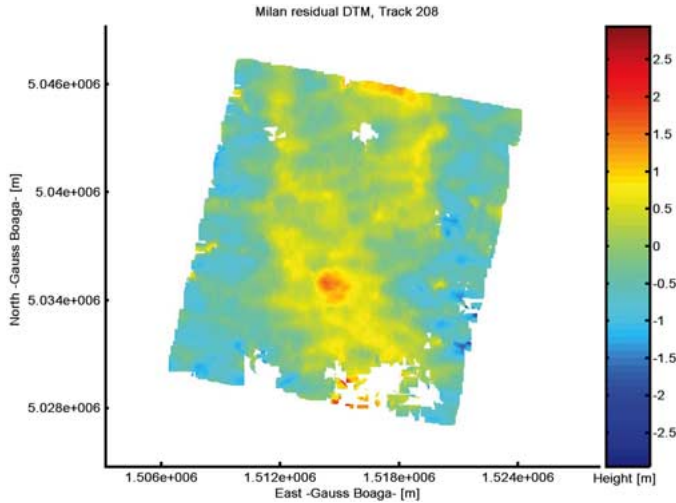
Fig. 12.   Residual PS DTM of Milan after removal of the low-pass synthetic topography (track 208).
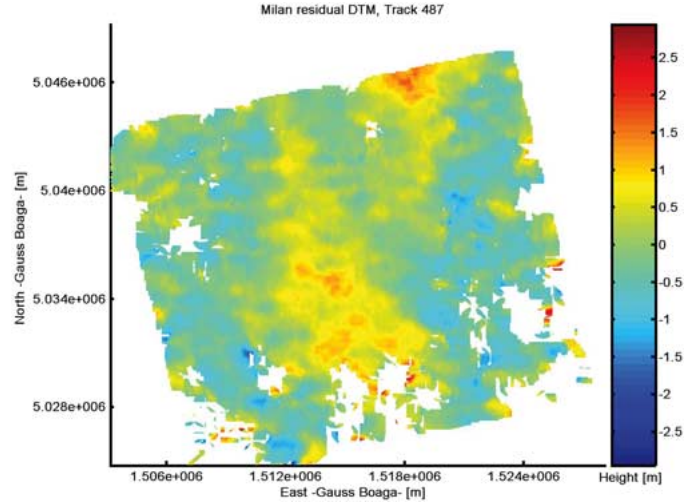

Fig. 14.   Residual PS DTM of Milan after removal of the low-pass synthetic topography (track 487).
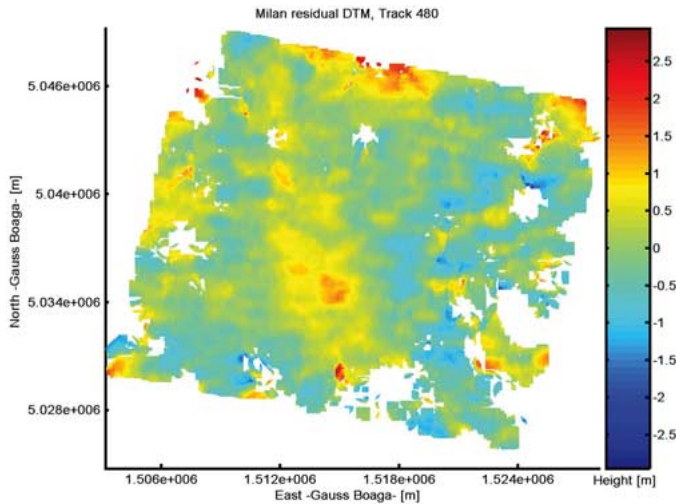

Fig. 13.   Residual PS DTM of Milan after removal of the low-pass synthetic topography (track 480).
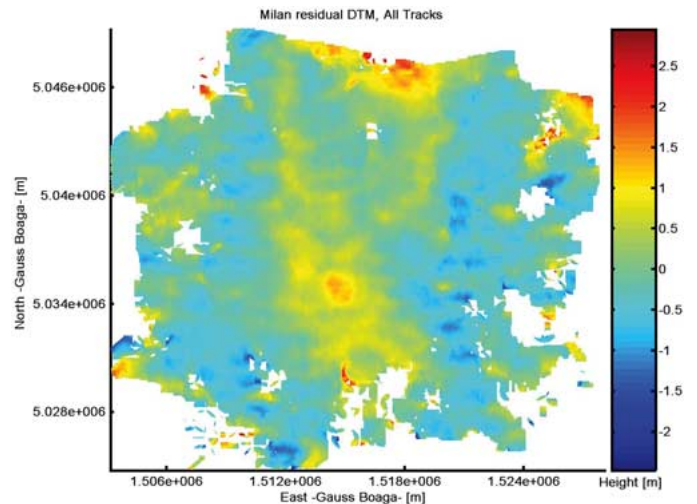

Fig. 15.   Residual PS DTM of Milan after removal of the low-pass synthetic topography, obtained by exploiting the whole cloud of PSs coming from the three tracks.

area, and that with lower elevation is retained (avoiding outliers and areas with PS density higher on roofs than on the ground). In this way, a first rough sampling (1 sample/km^2) of the terrain trend is obtained. A second-order surface is then fitted through the samples, thus obtaining a low-pass synthetic topography.

In Fig. 9, the north–south profile of Milan city after removing the low-pass topography is reported. It is worth noting the thickening of high points around the center of Milan, where the buildings are higher and closer among themselves. Focusing the attention on details, one can appreciate some sets of adjacent points with high elevations, which describe high regular structures, e.g., the Central Station and San Siro Stadium.

In Fig. 10, the histogram of residual elevations (after removing the low-pass topography) in Milan is shown. Different peaks are visible, highlighting recurrent heights (ground level, small houses, sheds, and high buildings). At high elevations, a wide variety of structures is present, whereas on the ground, a very accentuated and thin peak is visible. The width of this peak ($\sim$1 m at -3 dB) is the actual demonstration of the PS's height accuracy.

Exploiting the low-pass synthetic topography, PSs around the street level (±3 m) can be easily selected ($\sim$20 000 of 60 000 PSs). A sparse set of measurements of the ground level is thus available, and it can be used for obtaining a much more accurate estimate of its trend. The problem of resampling sparse data into a regular grid is not trivial, and in this case, it is complicated by the nonuniform density of points and by the presence of noise. The solution we adopted consists of two steps. First, data are divided into a thick regular grid, and the average elevation is calculated for each grid element (the distance between two elements is 50 m, both in the north and east coordinates). The resulting point set is then interpolated by exploiting a standard kriging process with a decorrelation distance of 300 m. Fig. 11 shows the obtained DTM of Milan city. As visible in Fig. 11, in order to respect the real data, we chose interpolation parameters that do not fill the areas where no scatterer is present. For a better appreciation of the details, in Fig. 12, the residual DTM after the removal of low-pass synthetic topography is shown. In Fig. 12, the filtering effect
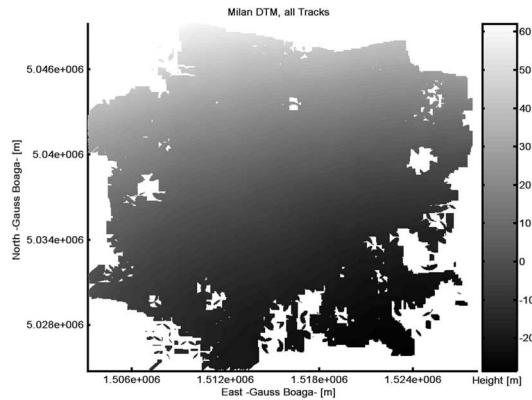
Fig. 16.   PS DTM of Milan obtained from the three superimposed tracks.

of the interpolation process is more evident. The color scale of Fig. 12 ranges between −3 and 3 m.

The accuracy of the proposed DTM can be derived by taking into account the spatial average of the PS's height that is induced by the processing steps previously described. By averaging N measures with standard deviation σ_s, the resulting accuracy is $\sigma_r = \sigma_s / \sqrt{N}$. Considering only the areas where at least one target is present, the average density of PSs on the ground in the urban site of Milan for a single track is about $70/\text{km}^2$. This means that in a $300 \times 300 \text{ m}^2$ area, the theoretical accuracy of an urban PS DTM is in the order of 10 cm.

VI. Multitrack DTM

Very interesting results can be obtained when comparing three different DTMs from two descending parallel tracks and an ascending one over Milan calculated with the technique described in the previous section. Three independent PS analyses have been carried out by exploiting the three data sets track 208, frame 2691; track 480, frame 2691; and track 487, frame 909. The number of images included in the analyses acquired by ERS-1, ERS-2, and Envisat during the time span 1992–2005 is 90, 90, and 40 in the three tracks, respectively. In order to precisely geocode all data, a unique ground control point (GCP) visible from the three different geometries has been selected. The geocoding process has been verified *a posteriori* by exploiting other multisensor targets whose 3-D ground locations estimated in the three tracks came out coincident unless 1 m. Approximately $60\,000$ PSs per track have been detected with multi-interferogram coherence $\gamma > 0.7$ ($\gamma > 0.8$ for track 487 due to the lower number of available images). By combining the data coming from the three different tracks, the average density of PSs on the ground reaches almost $250/\text{km}^2$, thus leading to a theoretical DTM accuracy of 10 cm in $200 \times 200 \text{ m}^2$.

A single DTM for each track and a global DTM obtained by exploiting the whole cloud of PSs coming from the three tracks have been calculated with the technique described in the previous section. A common low-pass synthetic topography estimated from the whole data has been used in order to make the results comparable. Figs. 12–14 are the single residual DTMs for each track after the removal of common low-pass synthetic topography. Fig. 15 reports the global residual DTM. In Figs. 12–15, the color scale ranges from −3–3 m. Besides some little differences caused by lack of PSs, the high correlation of these last results demonstrates that the actual accuracy
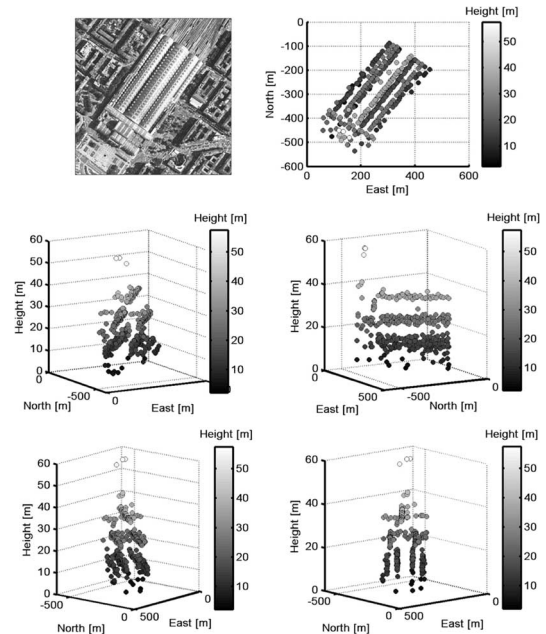

Fig. 17.   Railway central station in Milan. About 350 PSs lying on the structure have been detected from tracks 208, 480, and 487. (Upper left) Aerial photo. (Upper right) North–east coordinates of the detected PSs. The other four images are 3-D visualizations of the detected targets under different view angles. The gray scale of the plots indicates the estimated height.

achievable with the PS technique in the estimate of the terrain topography in an urban site is in the order of tens of centimeters. Finally, Fig. 16 shows the global DTM obtained from the three superimposed tracks.

It is worth noting here how radar interferometric data slowly approach the type of information currently only available from airborne laser altimetry. Laser altimetric data can be obtained using dedicated surveys, which yield positioning accuracies about one order of magnitude better than the results reported here, and spatial point densities about four orders of magnitude higher. However, laser altimetric data are costly, and in many cases, these precision levels are not necessary. The research described here can lead to a serious alternative to laser altimetry for urban topography and terrain modeling. The shortcoming of a PS DTM is clearly the coarse resolution (ERS ground resolution of about $25 \times 5 \text{ m}^2$), but as shown, the combination of different data sets can improve the resolution without losing accuracy significantly.

VII. Building Height

The obtained DTM can also be exploited for recovering the height of the buildings with respect to the ground. Moreover, the combination of PSs detected under different orbits allows a significant increase of the number of height measurements. Thus, by means of multitrack SAR PSs, after generating a DTM, the first raw urban digital surface model can also be produced. As a demonstrative example, we show in Fig. 17 the railway central station of Milan. About 350 PSs coming from the three tracks lie on the building structure. In Fig. 17, an aerial photo (on the upper left part) and the north–east coordinates (on the upper right part) of the PSs are reported. The other four images in Fig. 17 are 3-D visualizations of the detected targets under different view angles. The gray scale of the plots
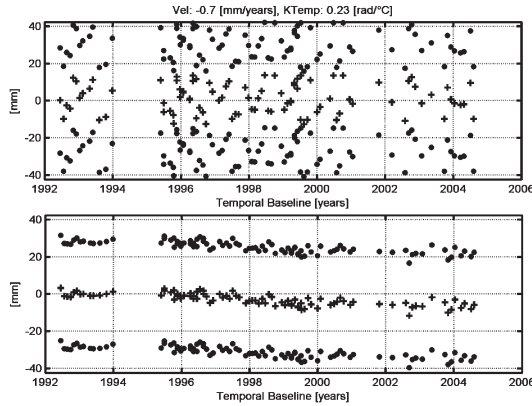
Fig. 18.   Example of PS displacement time series before and after the subtraction of the temperature-dependent term.
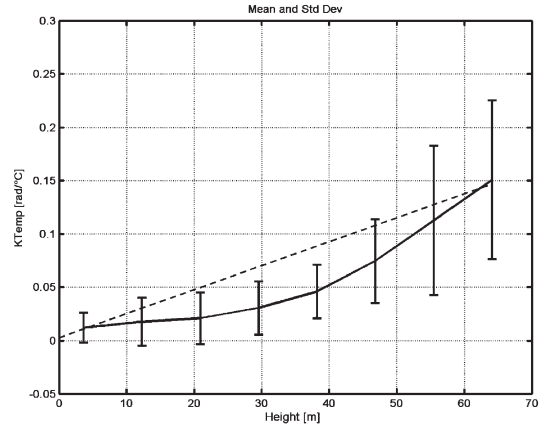

Fig. 19.   Mean value and dispersion (sign at $\pm\sigma$) of PS's thermal dilation as a function of the estimated elevation. (Dashed line) Theoretic behavior of composite steel–concrete buildings.

indicates the estimated height. The building details that can be recognized by looking at the location of the PSs are very impressive compared to radar resolution.

Knowing whether a PS is on the ground or at higher elevation is also very useful information on its physical nature. Moreover, the interpretation of PS deformation measurements depends strongly on the location of the target with respect to the ground. As a proof, we discovered a high correlation between seasonal deformations and elevation. This fact can be explained considering the thermal dilation of high composite steel–concrete buildings. The formula of linear expansion for a solid is

$$\Delta L = \alpha L_0 \Delta T. \tag{12}$$

A change in temperature ΔT causes a dilation ΔL for an object of length L_0. The proportionality constant α is called the coefficient of linear expansion and depends on the material. Assuming a thermal expansion coefficient of $1 \cdot 10^{-5}$ (composite steel–concrete buildings) and a temperature range of 30 °C, a 24-mm oscillation corresponds to an 80-m-high building. This is the case depicted in Fig. 18, where an example of a displacement time series before and after subtraction of the temperature-dependent term is shown. As the index of the oscillation range, a proportionality constant K_{temp} [rad/°C] between phase and temperature has been estimated for each PS. In Fig. 19, the constant K_{temp} is compared to the PS height with respect to the ground. The mean value of K_{temp} is plotted together with its dispersion (sign at $\pm\sigma$) as a function of the estimated height. The dashed line in Fig. 19 reports the theoretical trend of composite steel–concrete buildings. Even if real data show a more complex behavior, it is evident, recognizing from Fig. 19, that high buildings are affected by seasonal deformations. Thus, seasonal trends have to be systematically estimated and removed from interferometric phase series [9] in order to detect PSs on the buildings in an urban site.

VIII. Conclusion

The PS positioning accuracy has been demonstrated to be in the order of 1 m in the three dimensions, and experimental data on a high number of radar targets validate the result. As a consequence, very precise measures of the height of urban scatterers can be provided. Moreover, exploiting the high density of targets on the ground, the city street level can be recovered with

an accuracy of tens of centimeters. As proof, three urban DTMs of Milan obtained from three different and independent satellite tracks have been compared, showing high correlation.

References

[1] R. F. Hanssen, "Radar interferometry," in *Data Interpretation and Error Analysis.*   Dordrecht, The Netherlands: Kluwer, 2001.
[2] A. K. Gabriel, R. M. Goldstein, and H. A. Zebker, "Mapping small elevation changes over large areas: Differential radar interferometry," *J. Geophys. Res.*, vol. 94, no. B7, pp. 9183–9191, 1989.
[3] D. Massonnet and K. L. Feigl, "Radar interferometry and its application to changes in the Earth's surface," *Rev. Geophys.*, vol. 36, no. 4, pp. 441–500, Nov. 1998.
[4] P. A. Rosen, S. Hensley, I. R. Joughin, F. K. Li *et al.*, "Synthetic aperture radar interferometry," *Proc. IEEE*, vol. 88, no. 3, pp. 333–382, Mar. 2000.
[5] R. Burgmann, P. A. Rosen, and E. J. Fielding, "Synthetic aperture radar interferometry to measure Earth's surface topography and its deformation," *Annu. Rev. Earth Planet. Sci.*, vol. 28, pp. 169–209, May 2000.
[6] S. Madsen, H. A. Zebker, and J. Martin, "Topographic mapping using radar interferometry: Processing techniques," *IEEE Trans. Geosci. Remote Sens.*, vol. 31, no. 1, pp. 246–256, Jan. 1993.
[7] H. A. Zebker and J. Villasenor, "Decorrelation in interferometric radar echoes," *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 5, pp. 950–959, Sep. 1992.
[8] A. Ferretti, C. Prati, and F. Rocca, "Permanent scatterers in SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 1, pp. 8–20, Jan. 2001.
[9] ——, "Non-linear subsidence rate estimation using permanent scatterers in differential SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 5, pp. 2202–2212, Sep. 2000.
[10] C. Colesanti, A. Ferretti, F. Novali, C. Prati, and F. Rocca, "SAR monitoring of progressive and seasonal ground deformation using the permanent scatterers technique," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 7, pp. 1685–1701, Jul. 2003.
[11] S. Salvi, S. Atzori, C. Tolomei, J. Allievi, A. Ferretti, F. Rocca, C. Prati, S. Stramondo, and N. Feuillet, "Inflation rate of the Colli Albani volcanic complex retrieved by the permanent scatterers SAR interferometry technique," *Geophys. Res. Lett.*, vol. 31, no. 12, L12606, 2004.
[12] D. Perissin, C. Prati, M. Engdahl, and Y. L. Desnos, "Validating the SAR wave-number shift principle with ERS-Envisat PS coherent combination," *IEEE Trans. Geosci. Remote Sens.*, Feb. 2006, to be published.

[13] D. C. Rife and R. R. Boorstyn, "Single-tone parameter estimation from discrete-time observations," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 5, pp. 591–598, Sep. 1974.

[14] R. Bamler and P. Hartl, "Synthetic aperture radar interferometry," *Inverse Probl.*, vol. 14, no. 4, pp. R1–R54, Aug. 1998.

[15] R. Bamler and M. Eineder, "Accuracy of differential shift estimation by correlation and split-bandwidth interferometry for wideband and delta-k SAR systems," *IEEE Geosci. Remote Sens. Lett.*, vol. 2, no. 2, pp. 151–155, Apr. 2005.

Fabio Rocca received the Dottore in Ingegneria Elettronica in 1962.

From 1978 to 1988, he was with Stanford University, Stanford, CA, as a Visiting Professor. From 1975 to 1978, he was a Department Chair at Politecnico di Milano, Milan, Italy and, from 1980 to 1993, was with the Commissione d'Ateneo. He was also a cofounder of two small technological companies: Telerilevamento Europa and Aresys. He is currently a Professor of digital signal processing with Politecnico di Milano. His research interests include digital signal processing for television bandwidth compression, emission tomography, seismic data processing, and synthetic aperture radar.

Dr. Rocca was a recipient of awards from HUSPI (1979), Schlumberger (1990), Italgas Telecommunications (1995), and Eduard Rhein Foundation Technology (1999). He was also a recipient of the Special SEG Commendation (1998), Doctor Honoris Causa in Geophysics from the Institut Polytechnique de Lorraine (2001), and Best Paper awards from IGARSS (1989 and 1999) and EUSAR (2004). He was a President of the EAEG, an Honorary Member of SEG (1989) and EAGE (1998), and a President of OGS Trieste (1982–1983). He is the Coordinator of the first EEC research program in Geosciences, and a member of the Scientific Council of IN-OGS, the SAR Advisory Group of ESA, and the editorial committees of the *Journal of Seismic Exploration* and *Oil & Gas Science and Technology*.

Daniele Perissin was born in Milan, Italy, in 1977. He received the M.S. degree in telecommunications engineering and the Ph.D. degree in information technology from Politecnico di Milano, Milan, in 2002 and 2006, respectively.

He joined the research group of Prof. F. Rocca and Prof. C. Prati at Politecnico di Milano in 2002, and since then, he has been working on the permanent scatterers technique. He holds a patent on the use of urban dihedral reflectors for combining multisensor synthetic aperture radar data, which is his main interest.