#### **ORIGINAL PAPER**



# A multi-image super-resolution algorithm applied to thermal imagery

Emanuele Mandanici<sup>1</sup> · Luca Tavasci<sup>1</sup> · Francesco Corsini<sup>1</sup> · Stefano Gandolfi<sup>1</sup>

Received: 26 July 2018 / Accepted: 8 January 2019 / Published online: 1 February 2019 © The Author(s) 2019

#### **Abstract**

Although thermal imaging is a widely used technique in many applications and is under continuous development, one of its limits is the relatively coarse spatial resolution. Nevertheless, in the last years, a number of super-resolution algorithms have been developed which allow to enhance the resolution of the images. They can be divided in two main different categories: single-image or multiple image-based algorithms. In this work, a multiple image-based algorithm for the super-resolution was implemented, tested and applied to terrestrial thermal imaging with the aim to overcome the limitation of the low resolution. In particular, the method relies on the use of many images acquired from slightly different positions to obtain, thanks to the redundancy of observations, a super-resolution frame having an upsampling factor of four. Several tests were performed on synthetic datasets, and the accuracy of the obtained super-resolution images was investigated. Moreover, an original algorithm capable to identify gross errors during the image registration phase, which is one of the crucial phases, is presented and its reliability assessed. Results showed the effectiveness of the proposed method on both common visible images and thermal infrared ones, since discrepancies between reconstructed and reference values are reduced by 18 and 25% respectively, when compared with a conventional bicubic algorithm. Finally, the proposed method was tested on a case study concerning the thermal survey of the façade of a historical building in Bologna (Palazzo D'Accursio). A dataset of real thermal frames was acquired and a super-resolution image of the subject was generated through the developed algorithm. Strengths and weaknesses of the method were analysed and discussed in the paper.

Keywords Super-resolution · Thermography · Multi-image · Image registration · Resolution enhancement

### Introduction

Applications of thermal imaging have been growing in the last decades thanks to the exploitation of ground-based, airborne or spaceborne platforms. Ground-based observations can be used, for instance, for energy leak detection in buildings (Fox et al. 2016) or performance monitoring in industrial processes, whilst aerial surveys are exploited to support city planning (Bitelli et al. 2015) and precision agriculture (Ribeiro Gomes et al. 2017), among others. All these applications share the need for accurate and detailed information about surface temperature and its spatial pattern.

Despite the impressive technological development of thermal sensors fostered by the increasing demand, thermal imaging still suffers from the limitation in spatial resolution, which is far coarser if compared to common images acquired in the visible band or in the near-infrared. Excluding satellite sensors, indeed, currently available cameras operating in the long-wave thermal infrared window (8-14 microns) mount detectors that do not reach one megapixel (Luhmann et al. 2013). From the hardware point of view, further increasing the number of pixels per unit area is rather difficult, because—on the one hand, reducing the pixel size would decrease also the signal to noise ratio and—on the other hand—expanding the sensor size would be too expensive for general purpose devices (Hardie 2007; Nasrollahi and Moeslund 2014; Yue et al. 2016).

In this context, super-resolution (SR), which is the process of generating a higher resolution (HR) image starting from one or more lower resolution (LR) frames, appears a promising approach to overcome the current technological limitations and provide the capability to



 <sup>□</sup> Luca Tavasci
 luca.tavasci2@unibo.it

Department of Civil, Chemical, Environmental and Materials Engineering (DICAM), University of Bologna, Viale del Risorgimento 2, 40136 Bologna BO, Italy

analyse finer details of surface temperature patterns. Therefore, the present paper aims at developing a SR approach suitable for ground-based thermography and entirely based on LR thermal images themselves.

Many SR methods were proposed in literature based on different strategies (Borman and Stevenson 1998; Park et al. 2003; Farsiu et al. 2004; Jiji et al. 2007). The approaches to SR can be classified according to the number of input images (one or more), the domain (spatial or spectral) and the algorithm (hallucination or reconstruction). Single-image methods operate almost exclusively in the spatial domain and often implement hallucination algorithms (Baker and Kanade 2002) based on machine learning techniques, which require sets of LR and HR pairs as training (Glasner et al. 2009). Multiple-image methods, instead, assume that the differences between LR observations are produced by relative geometric or photometric displacements and, therefore, those differences can be used to reconstruct the targeted HR image. In fact, the generally ill-posed nature of the SR problem can be overdetermined by adding more LR images, provided that they carry nonredundant information (Liu et al. 2018). The imaging model, which is inverted by the reconstruction algorithm, usually consists of a warping function (describing the geometric transformations), a blurring function (accounting for the effects of atmosphere and optical system) and a downgrading factor that relates the dimensions of LR and HR images.

The major manufacturers of thermal cameras have begun to implement SR algorithms in the firmware of their most recent models. For instance, FLIR developed the UltraMax technology (FLIR 2018), which is based on the rapid acquisition of 16 thermal frames in less than 1 s. Due to the natural movements of the operator, each image is slightly offset from others. The result is a new image that has twice the original resolution, provided that there is not too much movement from the user or the target and that there is a minimum contrast in the scene. A hardware solution is instead adopted by Infratech, consisting in a fast-rotating MicroScan wheel (InfraTec GmbH 2018), which enables the acquisition of four exposures per wheel resolution, each one offset laterally by half a pixel. Other software solutions are based on a data fusion with images in the visible band. FLIR MSX for example enhances the clarity by embossing visible light details onto thermal images (TEquipmentNET 2018). This process, however, is likely to alter the thermal measurement because the finer details comes from the visible band and do not necessarily correspond to temperature variations of the observed surfaces.

Several experiments can be found also in the literature. Some authors worked on the integration of HR visible sensors and infrared ones, for satellite (Tonooka 2005;

Zhang and Huang 2013; Feng et al. 2015) and terrestrial (Chen et al. 2016; Poblete et al. 2018; Turner et al. 2014) applications. Bazzo et al. (2015) proposed a method that exploits fibre Bragg gratings and distributed optical sensors. Chikamatsu et al. (2010) adopted a system based on dual-camera acquisitions and a registration based on visible images. Other authors focussed on single frame methods, based on maximum a-posteriori algorithms (Sun et al. 2007), or regularised stochastic approaches (Panagiotopoulou and Anastassopoulos 2008), or on nonlocal means and steering kernel regression (Yu et al. 2013). Meza et al. (2015) addressed the SR and non-uniformity correction (NUC) problems with a joined solution strategy.

Generally speaking, the choice of the best SR strategy is highly dependent on the application, due to the constraints which are imposed by the specific problem to be addressed. Many common applications of ground-based thermography offer the opportunity to acquire many shots of the investigated surfaces; therefore, reconstruction methods based on multiple LR images appear promising, provided that acquisition time is compatible with the observed temperature dynamics.

Consequently, the present work proposes the application of a SR algorithm based on a number of LR images taken from slightly different points of view and evaluates its suitability for infrared thermography. The use of a multiple frame algorithm does not require hypothesis on a regularisation function nor on training databases, and it preserves a reliable geometry, a qualifying aspect for geomatics applications. In order to demonstrate the effectiveness of this method, two kinds of experiments were performed. The first one exploited a set of simulated LR images, which were generated once from a real visible HR image and once from a single thermal frame by solving directly the imaging model. The HR images were then reconstructed with the proposed SR algorithm and the discrepancies between original and reconstructed HR images were evaluated. In the second kind of experiment, the SR algorithm was applied to a real set of thermal images acquired on purpose over the façade of an historical building in Bologna city.

# Implementation of a multi-frame-based super-resolution algorithm

The purpose is the creation of a single HR frame using a sufficiently large set of LR observations. Therefore, the proposed method belongs to the so called multiple image-based algorithms (Nasrollahi and Moeslund 2014). The implementation of the algorithm follows the logic of those methods known as 'direct methods'. Their



operative scheme is usually divided into three different steps:

- Geometric registration of every LR image (targets) to the one chosen as reference. This step consists in the determination, for each LR image, of the pixel and subpixel shifts and rotations necessary to overlap the target and the reference image, through a rigid-like transformation matrix.
- 2. Projection of the pixels of the LR images on a single HR-dimensions grid.
- 3. Interpolation of the known digital numbers to recreate unknown ones.

The developed algorithm proposes for the first step the identification of the optimal reference image and the addition of a geometric registration check, carried out using a statistical approach. Then, the fusion of phases 2 and 3 in a unique step is proposed. The algorithm is described in the following sections and it is developed in the Matlab environment using its function libraries (Mathworks 2018).

# Geometric registration of low-resolution images

There is a large number of registration methods that use different approaches and can work on frequency or spatial domain. The estimation of the transformation parameters between couples of LR frames was carried out through the mutual information technique (Gao et al. 2008; Maes et al. 1997), a very efficient and general method used to compare similar images. Given a target and a reference frame having the same dimensions, the output of the registration process is a  $3 \times 3$  rigid transformation matrix that permits an accurate overlapping between the two images. The parameters contained in that matrix are the rotation angle  $\alpha_k$ , expressed in terms of direction cosines, and the translations values  $tx_k$  and  $ty_k$  along m- and n-axis respectively, where k refers to each LR image.

Since the registration proved to be more effective in images with slow intensity changes (Scarmana 2016), both the reference and the targets are filtered using a blurring filter. This allows to smooth sharp edges and small details, and therefore avoid abrupt changes of pixel digital numbers and distortions created by lossy compression protocols, e.g. jpeg. The filtering is performed through a low-pass filter, in this case a Gaussian blurring, using a sigma input value chosen after some ad hoc tuning and set to 0.5.

Once a LR frame has been chosen as the reference one, the registration algorithm can be applied to every LR observation estimating its transformation matrix. Finally, the transformation parameters  $tx_k$ ,  $ty_k$  and  $\alpha_k$  can be applied to each LR image performing a rigid-like transformation, thus generating a set of registered images.

#### Choice of the reference image

The choice of the frame to be used as the reference for the registration process may influence the final result and it is not possible to identify a priori which one of the K images is the best candidate. For this reason, the registration process was repeated K times, once for each image serving as reference. This way, K sets of transformation parameters were obtained, each one consisting of K transformation matrices. Then, to decide which set is the best for the image reconstruction, all the transformation matrices were applied and target images registered on the current reference. The deviation matrices between the reference frame and each target were determined and then aggregated in a single value Z obtained as the sum of the absolute values of all the elements of each matrix. All these aggregated values were then put in the columns of a  $K \times K$  matrix. This way, the column sums are representative of the suitability of the images to be used as the reference frame. Based on this consideration, the image providing the minimum column sum is chosen.

# **Geometric registration check**

Even after the choice of the best reference frame, some LR targets might not be properly registered due to wrong estimation of the transformation matrices, depending on the particular set of data (e.g. containing large flat areas or repeated patterns in the image). In this section, an algorithm to assess possible errors in the registration phase is described.

The starting point is the set of K-registered LR images. Each kth image has M rows and N columns, and each of its pixel has a digital number (DN)  $lr_k(m, n)$ , where m and n run from 1 to M and N respectively. Then, for each pixel and each image is possible to compute the absolute value of the normalized difference of its DNs:

$$\widehat{lr}_k(m,n) = \left| \frac{lr_k(m,n) - \overline{lr}(m,n)}{\sigma(m,n)} \right| \tag{1}$$

where,

$$\bar{lr}(m,n) = \frac{1}{K} \sum_{k} lr_k(m,n) \tag{2}$$

$$\sigma(m,n) = \sqrt{\frac{1}{K-1} \sum_{k} [lr_k(m,n) - \bar{lr}(m,n)]^2}$$
 (3)

For each LR image is now possible to compute the sum of the  $\widehat{lr}_k(m, n)$ ,  $SND_k$  (sum of normalized differences):

$$SND_k = \sum_{m} \sum_{n} \widehat{lr}_k(m, n) \tag{4}$$

The K different values of  $SND_k$ , one for each LR frame, can now be compared between themselves. Nevertheless,



these are still depending on the matrix dimensions and therefore quite inconvenient. In order to manage more suitable values, a further normalization is performed on the K values obtaining the  $\widehat{SND}_k$  ones. With the aim to exclude the images characterized by a less accurate registration, a threshold value for the  $\widehat{SND}_k$  must be chosen. Unfortunately,  $\widehat{SND}_k$  values do not belong to a Gaussian distribution, nor their real statistical distribution is known. Therefore, considering a confidence level of 95%, the threshold should be found between two and four standard deviations (Gaussian model and Chebyshev's inequality respectively). Within this range, it is to be chosen arbitrarily, depending on the experience and on the obtained results.

# **High-resolution image reconstruction**

Following the logic of SR direct methods, the next steps are the projection on a common grid and the interpolation of known data to compute the unknown ones. In this particular algorithm, a fusion of these steps is performed: transformation matrices are used to project every image on a new frame that have augmented dimensions and, at the same time, determine the most reliable digital number for every single pixel.

This step involves all the unregistered LR images that passed the registration check. Given the upsampling factor p, it is possible to determine the dimensions X and Y of the HR matrix as follows:

$$X = pM \tag{5}$$

$$Y = pN \tag{6}$$

We can define the position (x,y) of each pixel in the HR image, with x and y running from 1 to X and from 1 to Y

respectively. The parameters  $tx_k$ ,  $ty_k$  and  $\alpha_k$ , are then used to estimate the pixel positions  $(m_k^*, n_k^*)$  in the unregistered LR image starting from each pixel (x,y):

$$m_k^* = (x\cos\alpha_k - y\sin\alpha_k - ty_k * p)/p \tag{7}$$

$$n_k^* = (x \sin \alpha_k + y \cos \alpha_k - tx_k * p)/p \tag{8}$$

Note that  $m_k^*$  and  $n_k^*$  are real numbers and the integer parts of them can be written as  $\widehat{m}_k$  and  $\widehat{n}_k$ . The DNs hr(x, y) of the high-resolution image can be assigned by estimating the weighted average as follows:

$$hr(x, y) = \frac{\sum_{k} \lambda_{k, \widehat{m}, \widehat{n}} * lr_{k}(\widehat{m}_{k}, \widehat{n}_{k})}{\sum_{k} \lambda_{k, \widehat{m}, \widehat{n}}}$$
(9)

Weights are determined following a geometric criterion, i.e. by calculating the distances between the centre of the pixel  $(\widehat{n}_k, \widehat{n}_k)$  in the LR frame and the positions  $(m^*, n^*)$ :

$$d_{k,\widehat{m}} = \left| (\widehat{m}_k + 1) - m_k^* - 0.5 \right| \tag{10}$$

$$d_{k,\hat{n}} = \left| (\hat{n}_k + 1) - n_k^* - 0.5 \right| \tag{11}$$

$$d_{k,\widehat{m},\widehat{n}} = \sqrt{(d_{k,\widehat{m}})^2 + (d_{k,\widehat{n}})^2} \tag{12}$$

The distances  $d_{k,\widehat{m},\widehat{n}}$  are then used to determine the weights:

$$\lambda_{k,\widehat{m},\widehat{n}} = \left(\frac{1}{d_{k}\,\widehat{m}\,\widehat{n}}\right)^{3} \tag{13}$$

Clearly, the higher the power the more rapidly the influence of neighbour pixels decreases with distance. The exponent of 3 was chosen because a further increase did not produce significant improvements in the reconstruction. A summary of the reconstruction procedure is illustrated in Fig. 1, with some details about the subphases of every step.

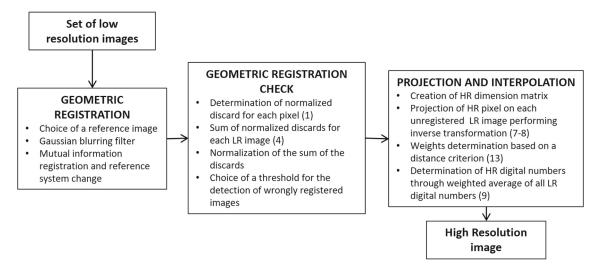


Fig. 1 Steps of the proposed SR procedure, following a modified 'direct methods' logic



#### Quality assessment of the reconstruction

After the reconstruction process is completed, the new SR image is unavoidably affected by some errors in the radiometry of the single pixels, if compared with what would have been acquired in a real HR frame. In order to estimate the a-posteriori uncertainty in the DN value of each SR pixel, a statistical parameter, which is a pseudo root mean square error, can be defined as follows:

$$\widetilde{RMSE}(x, y) = \sqrt{\frac{\sum_{k} [lr_{k}(\widehat{m}_{k}, \widehat{n}_{k}) - hr(x, y)]^{2} * \lambda_{k}, \widehat{m}, \widehat{n}}{\sum_{k} \lambda_{k}, \widehat{m}, \widehat{n}}}$$
(14)

In such way, a high value of  $\widetilde{RMSE}(x, y)$  in a certain pixel means that the values taken into account during the process are heterogeneous, thus the reconstruction is likely to be less accurate in that pixel.

# Synthetic dataset generation

In order to evaluate the accuracy of the reconstruction process, a synthetic dataset is produced by degrading an existing image. This way, a 'true HR image' exists and can be compared to the reconstructed one. To obtain a suitable set of *K* LR frames, it is possible to invert the SR logic.

A generator of LR images was then developed. A HR image is downgraded by a factor p in K different ways through a set of K different random real values for rotation angles and translations along x-axis and y-axis. Each DN in the new LR matrix is calculated using a weighted average of all the DNs falling in a window having (p+1)(p+1) dimensions and centred on the interested pixel. Weights are assigned using the following criterion: those who are completely included in the window have a unit weight, while the ones partially included have a lower weight, proportional to the overlapping area.

#### **Materials**

With the aim to evaluate the performances and the effectiveness of the algorithm, two different kind of tests were performed. The first one consists of a reconstruction that exploits two sets of simulated LR images, one from a common grayscale image and one from a thermal frame, whereas for the second test a set of real thermal images is taken into account.

The tests based on simulated LR images, generated from a HR one, have the strength to allow a comparison with a reliable 'truth'. In fact, simulated sets of images were obtained as explained in 'Synthetic dataset from a grayscale image', through the downgrading of existing HR

images. Various positions of the camera, one for each LR observation, were simulated, through the introduction of geometric displacements and little radiometric differences in every synthetic LR image.

# Synthetic dataset from a grayscale image

Here, the HR subject is a common 8-bit greyscale image portraying a car license plate<sup>1</sup> with a starting resolution of  $260 \times 160$  pixels, which is shown in Fig. 3a. It was downgraded to different images with resolution of  $65 \times 40$  pixels, following the procedure described in 'Synthetic dataset generation'.

Under ideal conditions, where shifts are well distributed along all the directions, the minimum number of LR images necessary to reconstruct the SR image would be equal to  $p^2$  (Scarmana 2016), where p is the upsampling factor. In this case, p=4 was considered a trade-off between the resolution improvement and the number of required LR images. Nevertheless, since the introduced shifts are randomly distributed, a set of 32 LR images (instead of 16) was created.

#### Synthetic dataset from a thermal image

The second image reconstruction was performed on a synthetic set of 32 thermal images obtained from a real one having a resolution of  $640 \times 480$  pixels and containing real temperature values. This image belongs to the dataset acquired for the real application test hereafter described. Each of the 32 frames produced for this test has dimensions of  $160 \times 120$  pixels and simulates what could be acquired using a thermal camera having a coarser resolution compared to the one actually used.

# Real dataset of thermal images

After the evaluation of the results on simulated thermal datasets, the last application of the SR algorithm concerned the whole real thermal set. The thermographic camera used is a FLIR P620, which is characterized by a thermal sensitivity of 60 mK at 30 °C, a spectral accuracy of  $\pm$  2 °C and a temperature range between - 40 °C and 500 °C. Images are natively stored in FLIR's radiometric jpg files that preserve the thermal sensitivity.

The subject of the thermal set of images is the façade of an historical building in Bologna, Palazzo D'Accursio. It was chosen to simulate a real scenario in which capturing

<sup>&</sup>lt;sup>1</sup>Downloaded from https://www.favcars.com/wallpapers-ac-schnitzer-99d-concept-e89-2011-184580 (last accessed on 19 June 2018).



220 Appl Geomat (2019) 11:215–228

the whole subject in a single image implies the setting of the camera to a considerable distance or the choice of an optic with a wide field of view. As a consequence, the increased ground sampling distance produces the loss of some details, which may be partially recovered through the application of SR. Since this dataset is not synthetic and little perspective or scale variations may be introduced in the survey, an overabundant number of images were acquired on purpose (50 in total). In order to capture the frames from slightly different points of view, the camera was mounted on a rack with a lever for vertical movements. The camera was then moved on ten height levels and five horizontal displacements, to minimise rotations and scale variations. The images were taken at 9 pm, 1 h after the sunset and the whole survey took about 25 min.

#### **Results and discussion**

In this section, the results obtained for the three tests introduced above are reported and discussed. The first two are related to the synthetic datasets and show the effectiveness of the method by evaluating the discrepancies with the original HR images. The latter test is an application on a real thermal dataset that actually constitute a possible application of the methodology.

#### Reconstruction of synthetic grayscale images

The first reconstruction test was performed on the simulated set described in "Synthetic dataset from a grayscale image". The 32 grayscale images were used to reconstruct a SR frame using an upsampling factor p of 4. Some preliminary tests were performed by changing different sigma values for the Gaussian blurring filter and varying the threshold values for the registration check. The most effective ones are  $\sigma =$ 0.5 for the Gaussian blurring filter and a threshold value of 2.3 for the registration control. Using these parameters, the registration algorithm presented in 'Choice of the reference image' (and 'Geometric registration check') was applied. The choice of the best reference image was done considering data shown in Fig. 2, which summarizes the content of the 32×32 matrix that represents the discards for all the mutual registrations between the considered frames. In Fig. 2, for each reference image, the mean of the cumulated discards discussed in 'Choice of the reference image' and the related standard deviations are reported on the ordinates. Several frames have similar mean values and the standard deviations are comparable in magnitude, with the lowest value for the frame 31. Also, the registration check phase was applied and no images were discarded, because of the absence of perspective or scale variations in the synthetic dataset. The reconstruction process generated the 260×160 SR image shown in Fig. 3d, whereas Fig. 3c

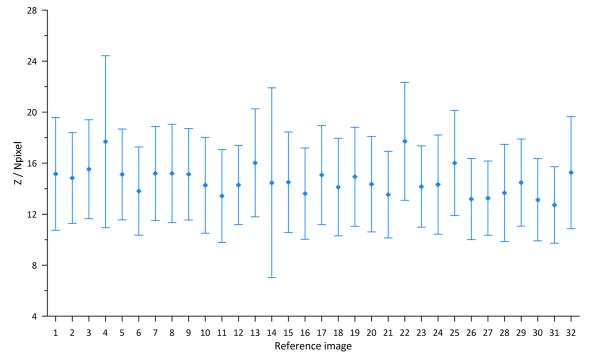
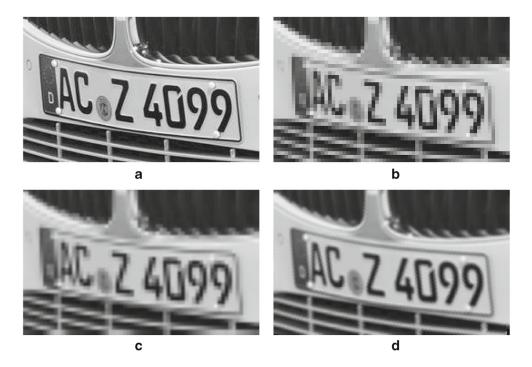


Fig. 2 Statistics on the different solutions obtained by changing the reference image. On the x-axes, the frame currently used as reference for the registration process. On the y-axes, the related aggregated value Z divided by the number of pixel contributing to its value



Fig. 3 Reconstruction of a common greyscale image: original HR image (a), one of the LR images used for the reconstruction (b), application of a conventional bicubic interpolation (c) and the SR image obtained with the proposed method (d)



is the result of the application of a bicubic interpolation (Keys 1981) (one of the most common single image-based algorithms) on the single LR frame displayed in Fig. 3b.

In order to evaluate the quality of the SR process, the differences between the original and the reconstructed images were computed pixel by pixel and the root mean square error (RMSE) in units of digital numbers was chosen as indicator of the algorithm performance. The histogram of the absolute values of the differences was analysed too. For comparison, the same differences and RMSE were computed also between the original image and one of the LR images, after a simple upsampling process by a nearest-neighbour algorithm. The same was done between the original image and the result of bicubic interpolation.

The RMSE of the differences between the original HR image and the simply upsampled LR one is 31 digital numbers. On the other hand, the RMSE value associated to the differences between original HR and SR images is 22 digital numbers, while the one between the original image and the result of bicubic interpolation is 27 digital numbers. This comparison evidences the higher effectiveness of the multi-frame-based super-resolution algorithm compared to the single-image-based one.

The histograms of the differences computed pixel by pixel with respect to the original frame are shown in Fig. 4, where solid lines mark the 95th percentile of the distributions. This percentile is located at a difference of 90 for the nearest-neighbour upsampling and at 80 for the bicubic interpolation, while it is reduced down to 65 DNs if the proposed SR algorithm is applied. It can be observed how the histogram of the multiple frame-based SR method

contains a higher number of elements in the zero class and, in general, a sharper peak than the ones related to the LR image and the bicubic interpolation. The more significant quality improvement obtained through the proposed SR is also appreciable by looking at Fig. 3b, c, d.

The *RMSE* values for the reconstructed SR image were determined for each pixel through Eq. 14 and are mapped in Fig. 5. All the edges are highlighted because the reconstruction process produces a blurring effect on them: original sharp edges like those of Fig. 3a are smoothed in Fig. 3d. Depending on the image features, i.e. sharp edges versus homogeneous areas, the algorithm can generate more or less accurate results. Indeed, such method tends to smooth peaks and troughs of DN values.

Finally, the same dataset was used to perform a test aiming to assess the effectiveness of the registration check algorithm described in 'Geometric registration check'. Slight errors in terms of translations and/or rotation were added on three of the registered images. In particular, the first frame was translated of 5 and -2 pixels along the x-axes and y-axes respectively, and a  $0.5^{\circ}$  rotation was also applied. On the second image translations of 1 and 2 pixels were imposed, whereas the third frame was rotated by a  $-2^{\circ}$  angle.

The registration check algorithm was performed on such dataset determining the  $\widehat{SND}_k$  parameter for each frame. Then, the SR-reconstructed image was created without applying any threshold to automatically remove wrongly registered frames, and the result was compared to the original one. This process was iterated removing each time the frame that had the highest  $\widehat{SND}_k$  value. The accuracy



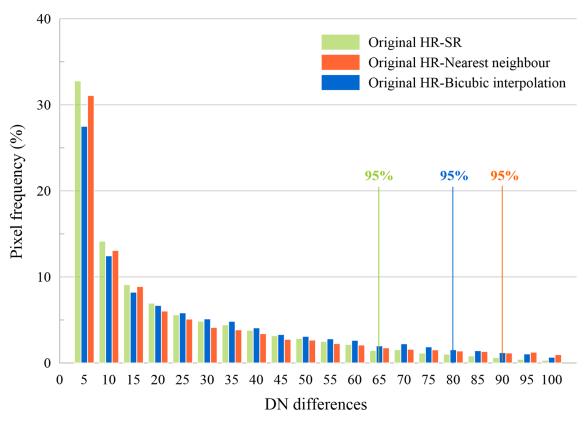


Fig. 4 Histogram of the differences in DNs between reconstructed SR and original HR images

variation of the SR images as function of the number of LR images used in the reconstruction phase is shown in Fig. 6.

It is quite evident that, after removing the first three images, the accuracy increases remarkably. These three

images are exactly the ones affected by errors in registration and their  $\widehat{SND}_k$  values are 4.04, 5.02 and 4.02 respectively, whereas the highest  $\widehat{SND}_k$  for the dataset of the remaining frames is 2.24. This test confirms the effectiveness of using

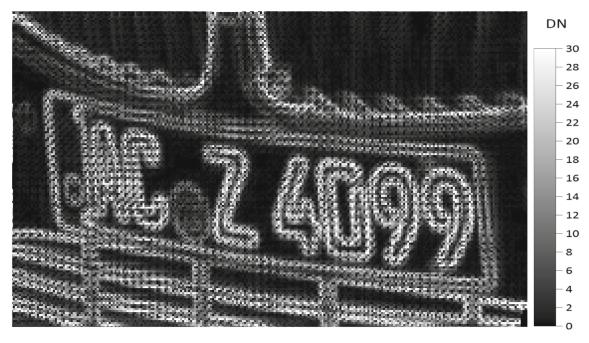
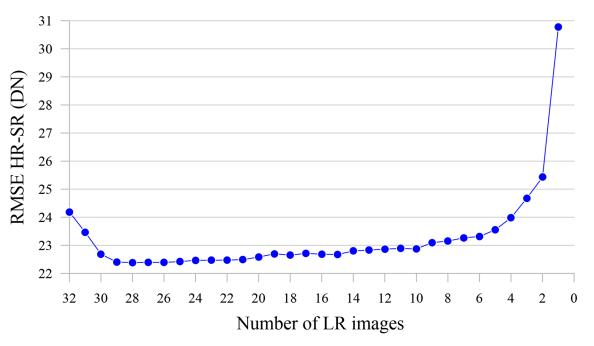


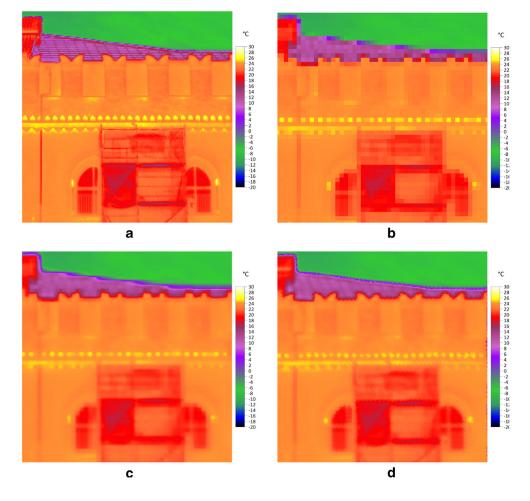
Fig. 5 Pseudo RMSE associated to the reconstruction of the image in Fig. 3d. Brighter pixels denote less reliable reconstructed values





 $\textbf{Fig. 6} \quad \text{Variation of the reconstruction accuracy depending on the number of LR images used. RMSE is computed on the DN differences between the original HR and the reconstructed SR images$ 

Fig. 7 Thermal synthetic dataset: original HR image (a), one of the LR images used for the reconstruction (b), application of bicubic interpolation (c) and SR image generated with the proposed method (d)





the registration check algorithm with a threshold set to 2.3. Changing the dataset, this value should be probably slightly changed performing an ad hoc tuning; nevertheless, the method proved its reliability. Moreover, Fig. 6 confirms that with an upsampling factor of 4 the use of at least 16 LR images is necessary to achieve a sensible enhancement in the reconstructed SR image, as postulated in 'Synthetic dataset from a grayscale image'.

#### Reconstruction of simulated thermal images

The first test on thermal images was performed using a set of 32 synthetic frames having resolution of  $160 \times 120$  pixels. Also in this case, an upsampling factor p=4, a 0.5-Gaussian blurring and a 2.3-threshold value for the registration check were used. The obtained results are displayed in Fig. 7, where every image refers to the same particular of the whole frame: (a) refers to the original HR image, (b) refers to one of the LR frames used during the process, (c) refers to the result of bicubic interpolation and (d) is related to the reconstructed SR image.

Also in this case, an improvement in terms of reduction of the RMSEs was found. In fact, while the RMSE of the differences between the LR (Fig. 7b) and the original HR (Fig. 7a) image is 1.0 °C, the RMSE between the SR and original HR images (Fig. 7a–d) is reduced to 0.6 °C.

Again, the RMSE associated to the image obtained through bicubic interpolation is about halfway, being it 0.8 °C. It is worthwhile to mention that, when choosing a non-optimal reference frame for the coregistration, the RMSE can rise up to 0.9 °C. In Fig. 8, it is possible to notice how this image is characterized by the presence of more homogeneous areas because of the higher number of elements in the 0 bin for the LR image than for the previous simulated dataset. Still, the SR algorithm generates a reconstructed image with more elements in the 0 bin than the LR one and the bicubic interpolation technique. Again, the green, blue and orange solid lines mark the 95th percentile of the distributions of the SR, bicubic interpolation and LR images respectively, confirming the consistent improvement obtained through the multi-image SR technique.

The reliability matrix determined with Eq. 14 for the reconstructed image is displayed in Fig. 9. The same edge effect observed on the car plate is identifiable also in this case.

# Reconstruction of real thermal images

In the last test, the SR algorithm was applied to the real dataset described in 'Real dataset of thermal images'. The input parameters that were used are the same defined for the synthetic datasets. In this case, the displacements

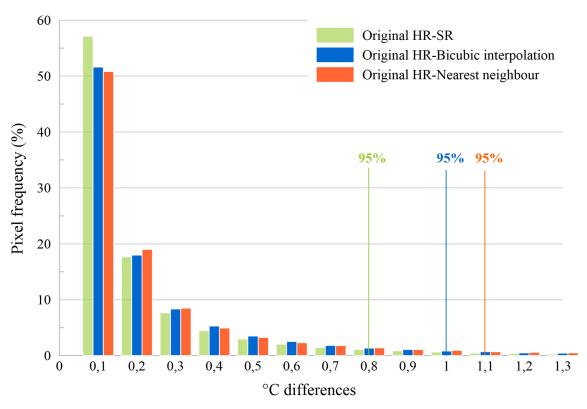


Fig. 8 Histogram of temperature differences between LR and HR images for the simulated thermal set





Fig. 9 Reliability matrix (pseudo RMSE) of the simulated thermal set. Brighter pixels denote less reliable reconstructed values

between the frames are not strictly controlled; therefore, 12 registered images were discarded during the registration check phase, probably because of excessive perspective changes or movements introduced during the survey. One of the LR frames and the reconstructed SR image are compared in Fig. 10.

Fig. 10 LR image (a) and reconstructed SR image (b), and the corresponding detail of the same area (c-d)

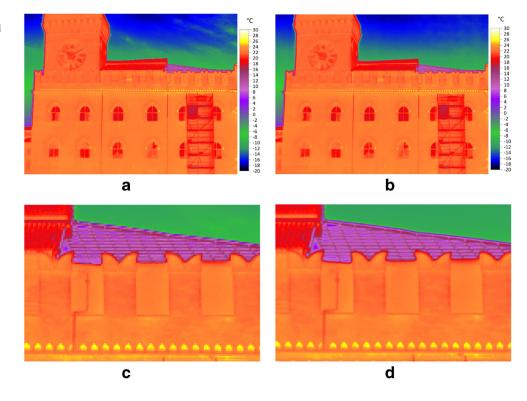




Figure 10c, d represents the same particular of the whole scene and relate to the LR and the reconstructed frames of Fig. 10a, b. The effectiveness of the algorithm is confirmed, in fact, the details in Fig. 10d appear more defined than in Fig. 10c, increasing the detail level. As for the simulated set, the same slight local blurring effect can be observed in the representation of the  $\widehat{RMSE}$  matrix in Fig. 11.

The presence in the dataset of images taken from different points of view is evident, while central areas of the image present very low  $\widehat{RMSE}$  values even on the edges, the most external ones appear brighter in Fig. 11. This effect increases with distance from the centre. The edge blurring effect is easy to recognise also in the roof in the upper right corner. Here, many edges far from the centre were reconstructed producing less reliable values.

A slight perspective change effect is identifiable in the tower corners, probably due to some portions of the building visible from certain points of view and hidden from others. As a consequence, the reliability of the reconstruction in these areas is lower than in the rest of the image. Another particular feature is the presence of pixel patterns, in particular in the cloudy area and on the façade of the building. The first is caused by clouds movements, because it is impossible to reconstruct moving objects that are changing their position in every frame. The second pattern is probably due to the cooling of the façade during the survey; temperature values of the first frames are slightly different

from the ones in last frames. This effect is not visible everywhere because of the different thermal emissivity of every material, for example, on the brickwork, where the pattern is visible, there is an apparent temperature difference of about 1 °C between the first and the last images. Conversely, the same difference does not appear on windows nor on the metal platform on the right, where the pattern is not evident. This is due to the low emissivity and consequently high reflectivity of these materials. In addition, thermal images can be affected by a residual non-uniformity effect, even after the factory calibration process that is performed by the firmware every few minutes (Mudau et al. 2011).

In any case, the pattern has no relevant effects on the geometrical aspects of the reconstruction problem, but it obviously introduces a fictitious variability on temperature values. The compensation for this effect is beyond the scope of the present work; however, it is worthwhile to point out that a faster acquisition of the LR images would greatly minimise the problem and this may be a challenge for further researches.

The whole SR process was handled by a 12-core workstation having a clock of 2.90 GHz and a 16 GB RAM. Above all, computational times strictly depend on the size of the target image and on the number of LR images used for the reconstruction, since the more expensive part of the algorithm, in terms of time, is the projection and interpola-

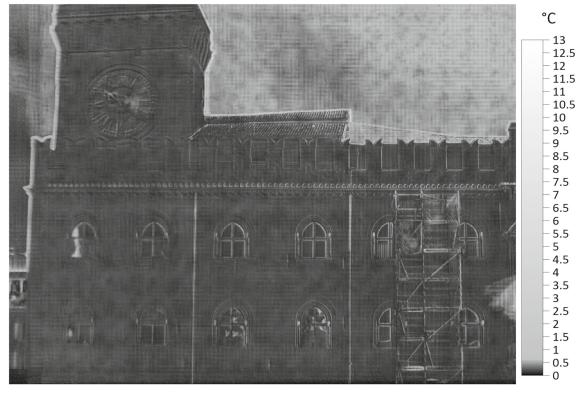


Fig. 11 Reliability matrix (pseudo RMSE) associated to the reconstruction of the thermal set of real LR images of Palazzo D'Accursio (Bologna). Brighter pixels denote less reliable reconstructed values



tion phase. The multiregistration of the whole dataset was performed using the Matlab parallel pooling feature and, with the described workstation, it took about 1 h, while the reconstruction of the 2560×1920 pixel thermal image using 50 LR frames took about 6 h. Clearly, some code optimisation may improve significantly the performances.

# **Conclusion**

This paper addresses the issue of the relatively poor spatial resolution usually affecting thermal images. For this purpose, a SR algorithm based on the multiple frame approach was implemented. Several tests were carried out on synthetic datasets in order to assess the effectiveness and reliability of the methods. Furthermore, the technique was tested on a real application by acquiring a thermal image set of the façade of a historical building in Bologna.

The first synthetic test (consisting in the reconstruction of an 8-bit grayscale image starting from 32 LR images generated by downsampling and shifting the original one) demonstrates that the proposed method produces a decrease of 29% in the RMSE on DN values, when compared to the simple nearest-neighbour upsampling, and of 18%, when compared to the bicubic algorithm. Moreover, the 95% of the pixels show errors in DNs lower than 65 considering the HR reconstructed images, whereas for the nearestneighbour upsampling there are errors up to 90 within the same percentage. Following the same procedure, a second test was performed on a thermal image. In this case, a reduction of the RMSE from 1.0 °C (nearest-neighbour) to 0.6 °C was obtained, while the 95th percentile of the errors decreases from 1.1 to 0.8 °C. The proposed method applies an upsampling factor of 4 which is higher than the solutions currently implemented by manufacturers. Moreover, it can be used on cameras without SR algorithms implemented in their firmware.

Finally, the SR algorithm was applied to the dataset acquired on the façade of the historical building obtaining a HR thermal image of dimensions 1920×2560 pixels. The comparison between the acquired 480×640 thermal frames and the reconstructed HR image evidences an effective improvement in the geometric description of the subject, especially at the edges. Even though a proper accuracy analysis is not possible in the real case, the proposed methodology includes a coherence analysis, which is able to identify the areas in which the reconstructed apparent temperature value is to be considered less reliable. This opportunity is one of the advantages of multiimage reconstruction methods over single-image ones. Furthermore, the rejection criteria implemented to identify outlier LR observations appear useful for the overall quality of the final product.

Further improvements will concern the dataset acquisition, in particular, speeding up the survey procedure and limiting as much as possible the prospective variations between frames. These precautions should limit temperature changes on the observed surfaces (which can be responsible for fictitious patterns appearing in the reconstructed image) and also improve coherence during the registration phase.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

#### References

- Baker S, Kanade T (2002) Limits on super-resolution and how to break them. IEEE Trans Pattern Anal Mach Intell 24(9):1167–1183. https://doi.org/10.1109/tpami.2002.1033210
- Bazzo JP, Pipa DR, Mezzadri F, Vagner da Silva E, Martelli C, Cardozo da Silva JC (2015) Super-resolution algorithm applied in thermal imaging of hydroelectric generators stator using hybrid sensing with DTS and FBG. In: 2015 SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference (IMOC), IEEE. https://doi.org/10.1109/imoc.2015.7369058
- Bitelli G, Conte P, Csoknyai T, Franci F, Girelli VA, Mandanici E (2015) Aerial thermography for energetic modelling of cities. Remote Sens 7(2):2152–2170. https://doi.org/10.3390/rs70202 152
- Borman S, Stevenson RL (1998) Super-resolution from image sequences-a review. In: 1998 Midwest Symposium on Circuits and Systems (Cat. No. 98CB36268), IEEE Comput. Soc. https://doi.org/10.1109/mwscas.1998.759509
- Chen X, Zhai G, Wang J, Hu C, Chen Y (2016) Color guided thermal image super resolution. In 2016 Visual Communications and Image Processing (VCIP), IEEE. https://doi.org/10.1109/vcip. 2016.7805509
- Chikamatsu S, Nakaya T, Kouda M, Kuroki N, Hirose T, Numa M (2010) Super-resolution technique for thermography with dual-camera system. In: Proceedings of 2010 IEEE international symposium on circuits and systems, IEEE. https://doi.org/10.1109/iscas.2010.5537929
- Farsiu S, Robinson D, Elad M, Milanfar P (2004) Advances and challenges in super-resolution. Int J Imaging Syst Technol 14(2):47–57. https://doi.org/10.1002/ima.20007
- Feng X, Foody G, Aplin P, Gosling SN (2015) Enhancing the spatial resolution of satellite-derived land surface temperature mapping for urban areas. Sustainable Cities and Society 19:341–348. https://doi.org/10.1016/j.scs.2015.04.007
- FLIR (2018) UltraMax the ultimate resolution. https://www.flirmedia. com/flir-instruments/industrial/technical-notes/ultramax-technical-note.html, accessed 16 November 2018
- Fox M, Goodhew S, De Wilde P (2016) Building defect detection: external versus internal thermography. Build Environ 105:317–331. https://doi.org/10.1016/j.buildenv.2016.06.011
- Gao Z, Gu B, Lin J (2008) Monomodal image registration using mutual information based methods. Image Vis Comput 26(2):164– 173. https://doi.org/10.1016/j.imavis.2006.08.002



228 Appl Geomat (2019) 11:215–228

Glasner D, Bagon S, Irani M (2009) Super-resolution from a single image. In: 2009 IEEE 12th international conference on computer vision. IEEE, Kyoto. https://doi.org/10.1109/iccv.2009.5459271

- Hardie R (2007) A fast image super-resolution algorithm using an adaptive Wiener filter. IEEE Trans Image Process 16(12):2953– 2964. https://doi.org/10.1109/tip.2007.909416
- InfraTec GmbH (2018) MicroScan significantly increases the geometrical resolution capability of the ImageIR camera series. https://www.ndt.net/search/docs.php3?id=21250&content=1, accessed 16 November 2018
- Jiji CV, Chaudhuri S, Chatterjee P (2007) Single frame image superresolution: should we process locally or globally? Multidim Syst Sign Process 18(2-3):123–152. https://doi.org/10.1007/s11045-007-0024-1
- Keys R (1981) Cubic convolution interpolation for digital image processing. IEEE Trans Acoust Speech Signal Process 29(6):1153–1160. https://doi.org/10.1109/tassp.1981.1163711
- Liu X, Chen L, Wang W, Zhao J (2018) Robust multi-frame superresolution based on spatially weighted half-quadratic estimation and adaptive BTV regularization. IEEE Transactions on Image Processing. https://doi.org/10.1109/tip.2018.2848113
- Luhmann T, Piechel J, Roelfs T (2013) Geometric calibration of thermographic cameras. In: Kuenzer C, Dech S (eds) Thermal infrared remote sensing, remote sensing and digital image processing, vol 17. Springer, Netherlands, pp 27–42. https://doi.org/10.1007/978-94-007-6639-6-2
- Maes F, Collignon A, Vandermeulen D, Marchal G, Suetens P (1997) Multimodality image registration by maximization of mutual information. IEEE Trans Med Imaging 16(2):187–198. https://doi.org/10.1109/42.563664
- Mathworks (2018) MATLAB website. https://mathworks.com/, accessed 12 December 2018
- Meza P, Machuca G, Torres S, San Martin C, Vera E (2015) Simultaneous digital super-resolution and nonuniformity correction for infrared imaging systems. Appl Opt 54(21):6508. https://doi.org/10.1364/ao.54.006508
- Mudau AE, Willers CJ, Griffith D, le Roux FPJ (2011) Nonuniformity correction and bad pixel replacement on LWIR and MWIR images. In: 2011 Saudi international electronics, communications and photonics conference (SIECPC), IEEE. https://doi.org/10.1109/siecpc.2011.5876937
- Nasrollahi K, Moeslund TB (2014) Super-resolution: a comprehensive survey. Mach Vis Appl 25(6):1423–1468. https://doi.org/10.1007/s00138-014-0623-4
- Panagiotopoulou A, Anastassopoulos V (2008) Super-resolution reconstruction of thermal infrared images. In: Proceedings of the 4th WSEAS international conference on remote sensing

- Park SC, Park MK, Kang MG (2003) Super-resolution image reconstruction: a technical overview. IEEE Signal Process Mag 20(3):21–36. https://doi.org/10.1109/msp.2003.1203207
- Poblete T, Ortega Farías S, Ryu D (2018) Automatic coregistration algorithm to remove canopy shaded pixels in UAV-borne thermal images to improve the estimation of crop water stress index of a drip-irrigated Cabernet Sauvignon vineyard. Sensors 18(2):397. https://doi.org/10.3390/s18020397
- Ribeiro Gomes K, Hernández López D, Ortega J, Ballesteros R, Poblete T, Moreno M (2017) Uncooled thermal camera calibration and optimization of the photogrammetry process for UAV applications in agriculture. Sensors 17(10):2173. https://doi.org/10.3390/s17102173
- Scarmana G (2016) An application of the least squares plane fitting interpolation process to image reconstruction and enhancement. In: 78th FIG Working Week 2016: Recovering from disaster, vol 8433. Christchurch, New Zealand
- Sun G, Li Q, Lu L (2007) MAP algorithm to super-resolution of infrared images. In: Maître H, Sun H, Liu J, Song E (eds) MIPPR 2007: Multispectral Image Processing, SPIE, vol 6787, pp 1–9. https://doi.org/10.1117/12.748453
- TEquipmentNET (2018) FLIR MSX multi-spectral dynamic imaging. https://www.tequipment.net/FlirMSXMulti-SpectralDynamicImaging.html, accessed 16 November 2018
- Tonooka H (2005) Resolution enhancement of ASTER shortwave and thermal infrared bands based on spectral similarity. In: Yasuoka Y, Ungar SG (eds) Image processing and pattern recognition in remote sensing II, SPIE, vol 5657, pp 9–19. https://doi.org/10.1117/12.580166
- Turner D, Lucieer A, Malenovský Z, King D, Robinson S (2014) Spatial co-registration of ultra-high resolution visible, multispectral and thermal images acquired with a micro-UAV over Antarctic Moss beds. Remote Sens 6(5):4003–4024. https://doi.org/10.3390/rs6054003
- Yu H, Chen Fs, Zhang Zj, Wang Cs (2013) Single infrared image super-resolution combining non-local means with kernel regression. Infrared Phys Technol 61:50–59. https://doi.org/10.1016/j. infrared.2013.07.006
- Yue L, Shen H, Li J, Yuan Q, Zhang H, Zhang L (2016) Image super-resolution: the techniques, applications, and future. Signal Process 128:389–408. https://doi.org/10.1016/j.sigpro.2016.05. 002
- Zhang H, Huang B (2013) Support vector regression-based downscaling for intercalibration of multiresolution satellite images. IEEE Trans Geosci Remote Sens 51(3):1114–1123. https://doi.org/10.1109/tgrs.2013.2243736

