

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Automatic Stabilization of Infrared Images Using Frequency Domain Methods

J. R. Martínez de Dios and A. Ollero

*Robotics, Vision and Control Research Group, University of Seville
Spain*

1. Introduction

In the last decade the decrease in the cost of infrared camera technology has boosted the use of infrared images in a growing number of applications. Traditional uses of infrared images, such as thermal analyses (Carvajal et al., 2011), nondestructive testing and predictive maintenance (Maldague, 2001), have been extended to new fields in which infrared cameras are mounted on vehicles and mobile robots and are used in applications such as border surveillance, building inspection, infrastructure maintenance, wildlife monitoring, search and rescue and surveillance, among many others. Image vibrations are harmful disturbances that perturb the performance of image-processing algorithms. Many of the aforementioned applications require having stabilized sequences of images before applying automatic image-processing techniques. Also, in cases where humans visualize the images, image vibrations induce significant stress and decrease the attention capacity of the operator.

Two main approaches for image stabilization have been developed. The first one aims at stabilizing the camera vibrations using different devices ranging from simple mechanical systems for handheld camcorders to high-performance inertial gyro-stabilized platforms. The first ones usually have low accuracy and perform “vibrations reduction” instead of “vibrations cancellation”. The high size, weight and cost of gyro-stabilized platforms constrain their use in a good number of applications.

The second main approach aims at correcting the images by applying image-processing techniques. Classical image-processing methods for image stabilization are based on detecting a set of local features –e.g. corners, lines and high-gradient points– in one image and tracking them along the images of the sequence. The relative motion of features from one image to another is used to model the motion between both images. Once the motion between images has been estimated, the second image is compensated such that no vibration can be perceived between both images. These methods have demonstrated very good performance with images with high contrast and low noise level, where a number of local features can be robustly detected. However, they do not perform well in images where these features cannot be robustly identified. This is the case of infrared images, which often have low contrast and resolution and high noise levels.

This chapter describes an automatic image-processing technique for the stabilization of sequences of images using frequency domain image representations obtained by means of

the Fourier transform. In particular, the method described in this chapter uses Fourier-Mellin transforms (FMT) and the Symmetric Phase Only Matched Filtering (SPOMF). This frequency domain representation provides advantages when stabilizing images with low contrast and high noise levels: the described method does not rely on local features but on the global structure of the image. The proposed method can correct translations, rotations and scalings between images, which is sufficient for a high number of stabilization problems. This chapter presents the main principles of the method, gives implementation details and describes its adaptation to different image stabilization applications.

The method was implemented efficiently for real-time execution with low computing resources. Several versions were tested and validated in different applications. The proposed method has also been validated with visual images but it is with infrared images where the advantages with respect to feature-based methods are more evident.

The main strengths of the proposed technique are the following:

1. it relies on the global structure of the image and thus, it is suitable for images with low contrast and resolution and high noise levels;
2. the method can be tuned balancing the stabilization accuracy and the computational burden, allowing adaptation to specific image stabilization needs.

This chapter is structured as follows:

- Section 2 introduces the problem of stabilization of sequences of infrared images.
- Section 3 presents the principles of image matching using Fourier-Mellin transform and Symmetric Phase Only Matched Filtering.
- Section 4 shows examples that illustrate the operation of the proposed method.
- Section 5 proposes some practical aspects to increase the robustness, accuracy and efficiency of the method.
- Section 6 presents various implementations in different applications.

Finally, Section 7 is devoted to the final discussions and conclusions.

2. Stabilization of infrared images

2.1 Infrared images

Infrared cameras generate images of the scene that contain the radiation intensity field within the infrared band. Infrared cameras can "see" in pitch black conditions and through smoke. They can transform radiation measurements in temperature estimations and can also generate thermograms containing temperature of the objects in the scene (Hudson, 1969). In the last decade infrared camera technology has evolved significantly. While in the decade of the 90's most infrared cameras weighted several kilograms, consumed hundreds of watts and their cryogenic cooling systems required frequent maintenance, now it is possible to find radiometric infrared cameras that weight less than 150 gr, consume less than 1,5 W and require no maintenance since they have no cooling system. These advances together with a remarkable cost decrease have motivated their use in a growing number of applications. Two infrared cameras are shown in Fig. 1, one *Mitsubishi IR-M300* (left) and one *Indigo Omega* micro-camera (right).

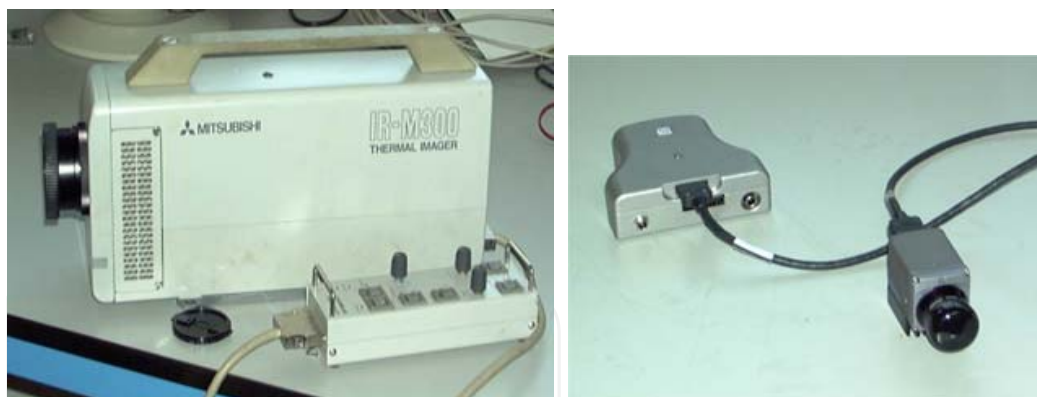


Fig. 1. Left) *Mitsubishi IR-M300* infrared camera. Right) *Indigo Omega* infrared micro-camera

However, despite these advances the quality of the images from infrared cameras is significantly lower than that of visual images due to several physical and technological reasons. Infrared camera detectors are highly affected by different types of noise including thermal noise, shot noise and flicker noise (Hudson, 1969). In fact, many cameras, particularly those operating in the mid-infrared spectral window $[3-5] \mu\text{m}$, use cooling systems -often cryogenic- to ensure that the detector operates at low and constant temperatures. Noise is crucial in infrared technology. In fact, infrared cameras usually use several characteristics to measure the influence of noise in the images, such as the *Noise Equivalent Temperature* (NET) or the *Noise Equivalent Temperature Difference* (NETD), among others (Maldague, 2001).

Also, infrared detectors usually have significantly lower sensitivity than CMOS and CCD detectors commonly used in visual cameras. This lack of sensitivity originates images with low contrast levels. Many infrared cameras compensate this lack of sensitivity by increasing camera exposure times, which originate image blurring if the camera is under vibrations. Also, infrared detectors usually have lower resolution than visual cameras. Thus, infrared images usually have lower resolutions -and details- than visual images.

Two images of the same scene are shown in Fig. 2: one taken with a visual camera (left) and one taken with a *FLIR ThermoCam P20* infrared camera (right). The differences in contrast, noise levels and detail levels are noticeable.

2.2 Brief description of image stabilization methods

Assume that we have a sequence of images taken by one camera under vibrations. Each image is represented by $Im_t(x,y)$, where t is the time when the image was captured. The camera vibrations cause relative motions between the images in the sequence. Assume image $Im_0(x,y)$ is considered the reference image. Stabilization of image $Im_t(x,y)$ consists of detecting and cancelling the motion between $Im_t(x,y)$ and $Im_0(x,y)$, such that there is no apparent motion between $Im_0(x,y)$ and the stabilized version of $Im_t(x,y)$. The process of detecting the relative motion between two images is called image matching.

There are two main groups of image-matching methods: spatial domain and frequency domain methods. Other techniques such as invariant moments (Abu-Mostafa & Psaltis, 1984) have poorer matching performance. A detailed survey can be found in (Zitová & Flusser, 2003).



Fig. 2. Two images of the same scene but taken with different cameras one visual camera (left) and one FLIR ThermoCam P20 infrared camera (right)

Spatial domain matching methods are based on local features that can be detected in both images $Im_0(x,y)$ and $Im_t(x,y)$. Assume that both images contain enough and sufficiently distributed features perceptible in both images. These feature-based methods usually have four steps. The first step is to detect local features in both images. These features are typically selected as corners (Tomasi, 1991), high-contrast points or local patterns with invariant properties, (Bay et al., 2008). The feature detector is applied to both images. The second step is to associate features detected in $Im_0(x,y)$ with features detected in $Im_t(x,y)$. Maximum likelihood criteria are often used for feature association. This step provides associations between features in both images, which can be used to estimate the relative motion between them.

The third step is to determine the motion between $Im_0(x,y)$ and $Im_t(x,y)$ using the aforementioned associations. The homography matrix is often used to model the motion between two images since it allows describing the transformations originated by changes on the location and orientation of the camera when the scene can be approximated by a plane (Hartley & Zisserman, 2004). Methods such as Least Median of Squares and RANSAC (Fischler & Bolles, 1981) are used to increase the robustness of the motion estimation. Once m_{0i} , the motion from $Im_0(x,y)$ to $Im_t(x,y)$, has been estimated, the fourth step is to apply image-processing methods to induce in $Im_t(x,y)$ a motion inverse to m_{0i} .

Spatial domain techniques have good performance if both images contain clear and robust local features but are not suitable for infrared images since a low number of local features can be robustly detected in infrared images.

On the other hand, frequency domain image-matching methods exploit the properties of images in the frequency domain. Normalized cross-correlation, see e.g. (Barnea & Silverman, 1972) and (Segeman, 1992), has been a common approach to match images with relative translations. In this case, the location of the maximum peak of the cross-correlation function between $Im_0(x,y)$ and $Im_t(x,y)$ corresponds to the relative translation between $Im_0(x,y)$ and $Im_t(x,y)$. Cross-correlation is commonly applied using the Fourier transform due to significant computational savings. This method does not have good performance if $Im_0(x,y)$ and $Im_t(x,y)$ have high noise levels: it produces broad peaks, which originate inaccuracies when

determining the location of the maximum peak. Some alternatives, such as the Symmetric Phase-Only Matched Filtering (SPOMF) (Ersoy & Zeng, 1989) produce significantly sharper peaks and are more resilient to noise, partial occlusions in the images and other defects.

Another disadvantage of Fourier transform methods is that they can only match translated images. The Fourier-Mellin transform (FMT) improves also this disadvantage since FMT methods can match translated, rotated and scaled images, (Chen et al., 1994).

2.3 Proposed frequency domain image stabilization method

The method proposed in this chapter uses image-matching methods based on frequency domain transformations. The method relies on the global structure of the image and not on local features. This fact provides significant robustness to noise and to the lack of contrast and resolution. Moreover, the proposed method is capable of stabilizing sequences which images are related through translations, rotations and scalings.

The proposed stabilization method adopts an incremental approach: the first image of the sequence, $Im_0(x,y)$, is considered reference and each of the remaining images in the sequence, $Im_t(x,y)$, is matched and corrected with respect to the stabilized version of the previous image, $Im^{S_{t-1}}(x,y)$. Section 3 briefly describes the principles of the proposed method. Section 4 illustrates the method with some examples.

Image matching can originate small -sub-pixel- errors. The accumulation of these errors in an incremental stabilization approach can originate drifts. To avoid drifts we developed another absolute stabilization approach in which all the images of the sequence are matched and corrected with respect to the same reference image $Im_0(x,y)$. This and other practical issues are discussed in Section 5.

The proposed method was customized to several problems with different stabilization requirements. In all cases the method was implemented and validated. This is described in Section 6.

3. Image matching using the Fourier-Mellin transform and Symmetric Phase-Only Matched Filtering

Assume that $Im^{S_{t-1}}(x,y)$ is the stabilized version of the image gathered at time $t-1$. In the incremental stabilization approach adopted $Im^{S_{t-1}}(x,y)$ is considered the reference image for stabilization of $Im_t(x,y)$ at time t . Assume that $Im_t(x,y)$ is a rotated, scaled and translated replica of $Im^{S_{t-1}}(x,y)$. The proposed stabilization method consists of two steps. The first one detects and corrects the rotation and scaling between $Im^{S_{t-1}}(x,y)$ and $Im_t(x,y)$. $Im^{R_t}(x,y)$ is the rotation and scaling corrected version of $Im_t(x,y)$. The second step detects and corrects the translation in axes x and y between $Im^{S_{t-1}}(x,y)$ and $Im^{R_t}(x,y)$. The stabilized version of $Im_t(x,y)$, i.e. the translation-corrected version of $Im^{R_t}(x,y)$, is denoted $Im^{S_t}(x,y)$.

Assume that $s(x,y)$ and $r(x,y)$ are the central rectangular regions of $Im_t(x,y)$ and $Im^{S_{t-1}}(x,y)$, respectively. Thus, $s(x,y)$ is a replica of $r(x,y)$ rotated with an angle α , scaled with a factor σ and translated with translational offsets (x_0, y_0) :

$$s(x,y) = r(\sigma(x \cos \alpha + y \sin \alpha) - x_0, \sigma(-x \sin \alpha + y \cos \alpha) - y_0) \quad (1)$$

The proposed method uses Symmetric Phase-Only Matched Filtering to match translated images. Below the Symmetric Phase-Only Matched Filtering and the two aforementioned steps in the algorithm are summarized.

3.1 Symmetric Phase-Only Matched Filtering

Assume that $b(x,y)$ is replica of $c(x,y)$ translated (x_0,y_0) in a noisy scene, $b(x,y)=c(x-x_0,y-y_0)+n(x,y)$, where $n(x,y)$ represents a white zero mean random noise. A traditional matched filter for $b(x,y)$ has the following transfer function:

$$H_{MF}(u,v) = \frac{1}{|n_w|^2} C^*(u,v), \quad (2)$$

where $C^*(u,v)$ is the complex conjugate of the Fourier transform of $c(x,y)$ and $|n_w|$ is the noise intensity. The output of the matched filter in (2) has a maximum peak at (x_0,y_0) . The location of this peak determines the translational offset (x_0,y_0) .

One of the main limitations of traditional matched filters is that the output is affected by the energy of the image while translation between images only influences the spectral phase. To minimize this dependence the alternative adopted in our method is to use Symmetric Phase-Only Matched Filtering (SPOMF) to determine translations between images. The SPOMF between $b(x,y)$ and $c(x,y)$ is as follows:

$$Q_{SPOMF}(u,v) = \frac{B(u,v)}{|B(u,v)|} \frac{C^*(u,v)}{|C(u,v)|} = \exp[j\varphi_B(u,v) - j\varphi_C(u,v)], \quad (3)$$

where $\varphi_B(u,v)$ and $\varphi_C(u,v)$ are the spectral phases of $b(x,y)$ and $c(x,y)$, respectively. In the absence of noise $Q_{SPOMF}(u,v) = \exp[-j2\pi(ux_0 + vy_0)]$ and the inverse Fourier transform of $Q_{SPOMF}(u,v)$ is a Dirac delta located at (x_0,y_0) . SPOMF yields to a sharp peak, significantly sharper than in case the matched filters in (2). The location of the maximum is easier to estimate and more tolerant to noise, which, as previously mentioned, is a significant advantage when dealing with infrared images.

3.2 Rotation and scaling correction step

This step aims to detect and correct the rotation and scaling between $s(x,y)$ and $r(x,y)$. The corrected version will be called $s^R(x,y)$. From (1) the Fourier transform of $s(x,y)$, $S(u,v) = \text{Fourier}\{s(x,y)\}$, can be expressed by the following expression:

$$S(u,v) = \sigma^{-2} \left| R\left(\sigma^{-1}(u \cos \alpha + v \sin \alpha), \sigma^{-1}(-u \sin \alpha + v \cos \alpha)\right) \right| \exp[-j\varphi_S(u,v)], \quad (4)$$

where $R(u,v) = \text{Fourier}\{r(x,y)\}$ and $\varphi_S(u,v)$ is the phase of $S(u,v)$. From (4) it is easy to notice that $|S(u,v)|$ is affected by rotations and scalings but is invariant to translations:

$$|S(u,v)| = \sigma^{-2} \left| R\left(\sigma^{-1}(u \cos \alpha + v \sin \alpha), \sigma^{-1}(-u \sin \alpha + v \cos \alpha)\right) \right| \quad (5)$$

The rotation and the scaling can be decoupled by converting $|R(u,v)|$ and $|S(u,v)|$ to polar coordinates. Let $r_p(\theta,\rho)$ and $s_p(\theta,\rho)$ be $|R(u,v)|$ and $|S(u,v)|$ in polar co-ordinates (θ,ρ) :

$$r_p(\theta, \rho) = |R(\rho \cos \theta, \rho \sin \theta)| \quad (6)$$

$$s_p(\theta, \rho) = |S(\rho \cos \theta, \rho \sin \theta)| \quad (7)$$

Thus, it is easy to check that:

$$\sigma^{-1}(u \cos \alpha + v \sin \alpha) = \frac{\rho}{\sigma} \cos(\theta - \alpha) \quad (8)$$

$$\sigma^{-1}(-u \sin \alpha + v \cos \alpha) = \frac{\rho}{\sigma} \sin(\theta - \alpha) \quad (9)$$

Hence, from (5), $s_p(\theta, \rho)$ can be expressed as:

$$s_p(\theta, \rho) = \sigma^{-2} r_p(\theta - \alpha, \rho / \sigma) \quad (10)$$

Thus, the rotation of an angle α has been transformed into a translation in the θ axis. Scaling can be also transformed into a translation by applying a logarithmic scale on ρ axis, i.e. $\lambda = \log(\rho)$. In log-polar coordinates (5) can be expressed as:

$$s_{pl}(\theta, \lambda) = s_{pl}(\theta, \log(\rho)) = \sigma^{-2} r_{pl}(\theta - \alpha, \lambda - \log(\sigma)), \quad (11)$$

where $s_{pl}(\theta, \lambda)$ and $r_{pl}(\theta, \lambda)$ are the log-polar versions of $|S(u, v)|$ and $|R(u, v)|$. Thus, in (11) the rotation and scaling have been transformed into translations. Let $S_{pl}(v, \omega)$ and $R_{pl}(v, \omega)$ be the Fourier transforms of $s_{pl}(\theta, \lambda)$ and $r_{pl}(\theta, \lambda)$, respectively. Using shift properties of the Fourier transform, (11) can be rewritten as:

$$|S_{pl}(v, \omega)| \exp\{-j\varphi_{S_{pl}}(v, \omega)\} = |R_{pl}(v, \omega)| \exp[-j\varphi_{S_{pl}}(v, \omega) - 2\pi(v \log(\sigma) + \omega \alpha)] \quad (12)$$

Thus, $|S_{pl}(v, \omega)| = |R_{pl}(v, \omega)|$. In (12) the rotation and the scaling appear as shifts between $\varphi_{S_{pl}}(v, \omega)$ and $\varphi_{R_{pl}}(v, \omega)$, the phases of $S_{pl}(v, \omega)$ and $R_{pl}(v, \omega)$, respectively. The phase shift is:

$$\delta = \varphi_{S_{pl}}(v, \omega) - \varphi_{R_{pl}}(v, \omega) = 2\pi(v \log(\sigma) + \omega \alpha) \quad (13)$$

Then, if we apply a SPOMF, the location of the peak of $q_r(\theta, \lambda)$, $(\theta_{\max}, \lambda_{\max})$, represents the rotation angle $\alpha = \theta_{\max}$ and the scaling factor $\lambda_{\max} = \log(\sigma)$ between $s(x, y)$ and $r(x, y)$:

$$q_r(\theta, \lambda) = \text{Fourier}^{-1}\{\exp[-j\delta]\}, \quad (14)$$

where Fourier^{-1} stands for the inverse Fourier transform. The corrected image, $Im^R_t(x, y)$, is obtained by rotating $Im_t(x, y)$ an angle $-\theta_{\max}$ and scaling by factor $\exp(\lambda_{\max})$. The central rectangular part of $Im^R_t(x, y)$, $s^R(x, y)$, is used for the translation correction step of the algorithm.

3.3 Translation correction step

Once the rotation and scaling have been corrected, $s^R(x, y)$ and $r(x, y)$ are related only by translations in axes x and y . This step detects and corrects the translational offsets between $s^R(x, y)$ and $r(x, y)$. The SPOMF can be used to determine the translations between them:

$$Q_t(u,v) = \exp[-j(\varphi_{SR}(u,v) - \varphi_R(u,v))] \quad , \quad (15)$$

where $\varphi_{SR}(u,v)$ and $\varphi_R(u,v)$ are respectively the phases of $S^R(u,v)$ and $R(u,v)$, the Fourier transforms of $s^R(x,y)$ and $r(x,y)$. The translational offset between $s^R(x,y)$ and $r(x,y)$ can be obtained by computing:

$$q_t(x,y) = \text{Fourier}^{-1} \{ Q_t(u,v) \} \quad (16)$$

The peak of $q_t(x,y)$ is located at $x=x_{max}$ and $y=y_{max}$. Thus, the translations between $s^R(x,y)$ and $r(x,y)$ in axes x and y are $x_0=x_{max}$ and $y_0=y_{max}$. The stabilized image, $Im_t^s(x,y)$, is computed by shifting $Im_t^R(x,y)$ by $-x_0$ in axis x and by $-y_0$ in axis y , respectively.

4. Experiments

This section presents some experiments that illustrate the proposed method. Two infrared images gathered from an infrared camera under vibrations are shown in Fig.3. The image at the left will be considered stabilized, $Im_t^{s_{t-1}}(x,y)$, and will be used as reference to stabilize the image at the right, $Im_t(x,y)$. The method determined and corrected the relative rotation angle, scaling factor and translations of $Im_t(x,y)$ with respect to $Im_t^{s_{t-1}}(x,y)$.



Fig. 3. Two infrared images from a sequence of images under vibrations. The image at the left is taken as reference

First, the central parts of both images, $r(x,y)$ and $s(x,y)$, are selected, see Fig. 4. Their size is chosen to be power of two in order to optimize the computational burden of the Fourier Transform FFT algorithm. The original images are 640x480 and the central parts are 256x256.

The first step of the algorithm described in Section 3 starts. First, $R(u,v)$ and $S(u,v)$, the Fourier transforms of $r(x,y)$ and $s(x,y)$ are computed. $|R(u,v)|$ and $|S(u,v)|$ are shown in Fig. 5. It can be noticed that rotation between $r(x,y)$ and $s(x,y)$ results in a rotation between $|R(u,v)|$ and $|S(u,v)|$ as predicted by (5).



Fig. 4. $r(x,y)$ and $s(x,y)$ for images in Fig. 3

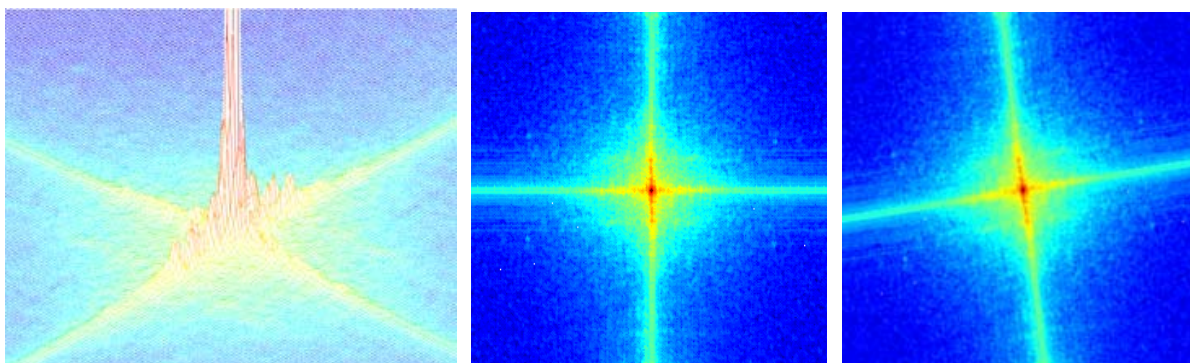


Fig. 5. Left) 3D view of $|R(u,v)|$; Center) and Right) 2D views $|R(u,v)|$ and $|S(u,v)|$, respectively

Then, $r_p(\theta, \lambda)$ and $s_p(\theta, \lambda)$, the versions of $|R(u,v)|$ and $|S(u,v)|$ in log-polar co-ordinates are computed. $r_p(\theta, \lambda)$ and $s_p(\theta, \lambda)$ are shown in Fig. 6. The colour represents the magnitude. As predicted by (11) the rotation between $|R(u,v)|$ and $|S(u,v)|$ has been transformed into a shift in the θ axis between $r_p(\theta, \lambda)$ and $s_p(\theta, \lambda)$. The value of the shift corresponds to the rotation angle between $Im^{S_{t-1}}(x,y)$ and $Im_t(x,y)$.

Next, we apply the SPOMF between $r_p(\theta, \lambda)$ and $s_p(\theta, \lambda)$ using (13) and (14). The resulting $q_r(\theta, \lambda)$ is shown in Fig. 7left. The maximum peak in $q_r(\theta, \lambda)$ is sharp. It is located at $\theta_{max} = -9.14^\circ$, $\lambda_{max} = 0$. The second peak is 50% lower in magnitude and is located at high distance from the maximum peak, which facilitates the identification of the maximum peak. Notice that $\lambda_{max} = \log(\sigma_{max}) = 0$ involves $\sigma_{max} = 1$, i.e. there is no scaling between both images. The rotation between both images is $\theta_{max} = -9.14^\circ$. Next, $Im_t(x,y)$ is corrected: $Im_t(x,y)$ is rotated an angle $-\alpha = 9.14^\circ$. $Im^R_t(x,y)$, the rotated-corrected version of $Im_t(x,y)$, is shown in Fig. 7right. The central rectangular part of $Im^R_t(x,y)$, $s^R(x,y)$, is selected for the second part of the method.

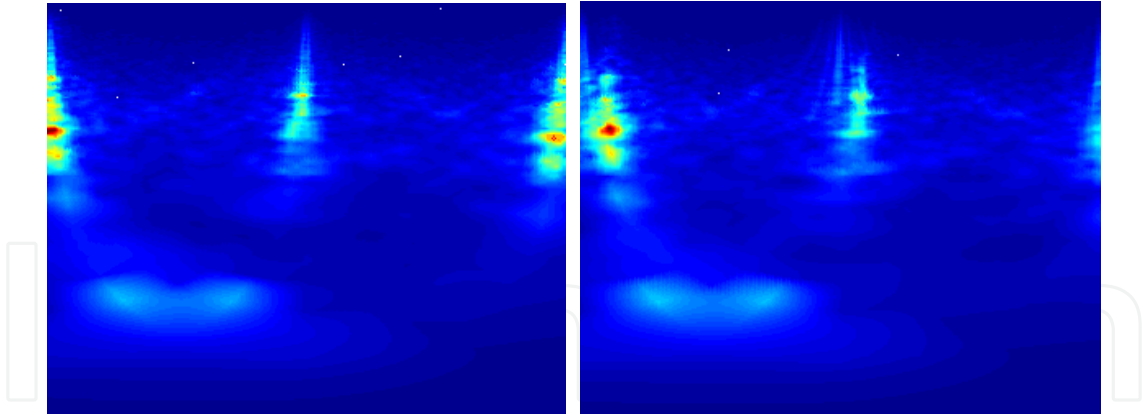


Fig. 6. Resulting $r_p(\theta, \lambda)$ (left) and $s_p(\theta, \lambda)$ (right)

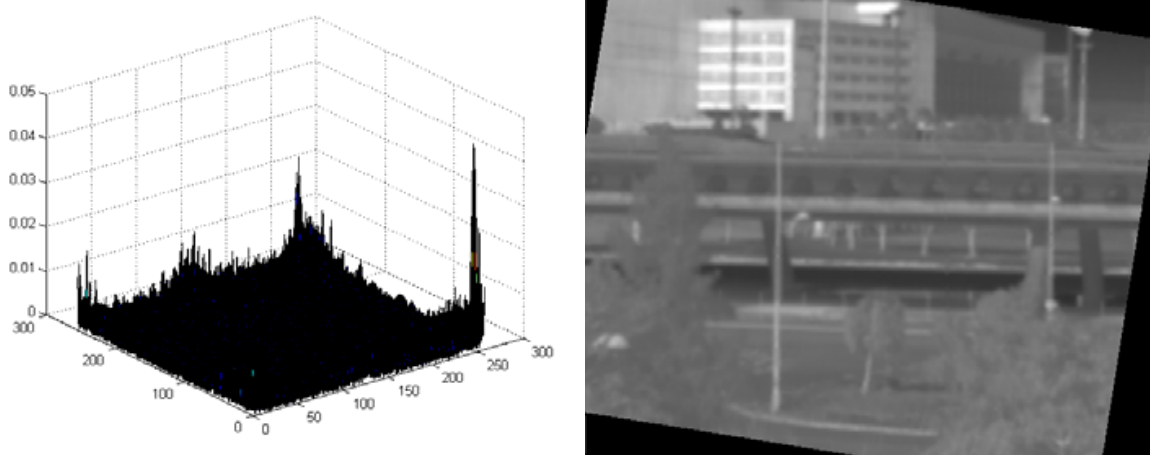


Fig. 7. Left) Resulting $q_r(\theta, \lambda)$. Right) $Im^R_t(x, y)$, i.e. $Im_t(x, y)$ after rotation correction

Once the rotation and scaling between $Im^{S_{t-1}}(x, y)$ and $Im_t(x, y)$ have been cancelled, the following step is to compute the translations between $s^R(x, y)$ and $r(x, y)$. The SPOMF is computed using (15) and (16). The resulting $q_t(x, y)$ is in Fig. 8left. The peak was very sharp. The maximum peak was located at $x_{max}=-4$ and $y_{max}=7$. Thus, $s^R(x, y)$ and $r(x, y)$ are translated $x_0=-4$ in axis x and $y_0=7$ in axis y . Next, $Im^R_t(x, y)$ is shifted by $-x_0$ and $-y_0$ in axes x and y respectively in order to compensate the vibration. Fig. 8right shows $Im^{S_t}(x, y)$, the stabilized version of $Im_t(x, y)$, i.e. the translation-corrected version of $Im^R_t(x, y)$.

5. Practical issues

This section presents practical aspects that have been developed to increase the robustness, accuracy and efficiency of the proposed method.

5.1 Image enhance

Preliminary experiments revealed that applying preprocessing image enhance methods improved significantly the performance of the proposed method. An efficient histogram stretching method was used. The objective is to transform the image levels such that the enhanced image has desired mean intensity $-MI_{ref}$ and contrast $-C_{ref}$ values. The histogram stretching implements the following linear transformation function:

$$Im_E(x,y) = \frac{C_{ref}}{C} (Im(x,y) - MI) + MI_{ref}, \quad (17)$$

where C and MI are the contrast and mean values before the histogram stretching. $Im(x,y)$ and $Im_E(x,y)$ are respectively the images before and after applying image enhance methods. Image enhance increases the level of noise in the images. Although higher noise levels have harmful effects in feature-based matching methods, the high robustness to noise of SPOMF avoids the degradation in our case. In images with very low contrast levels it is interesting to apply edge detectors. Fig. 9 shows the resulting images after applying Sobel edge detector to $r(x,y)$ and $s(x,y)$ in Fig. 4.

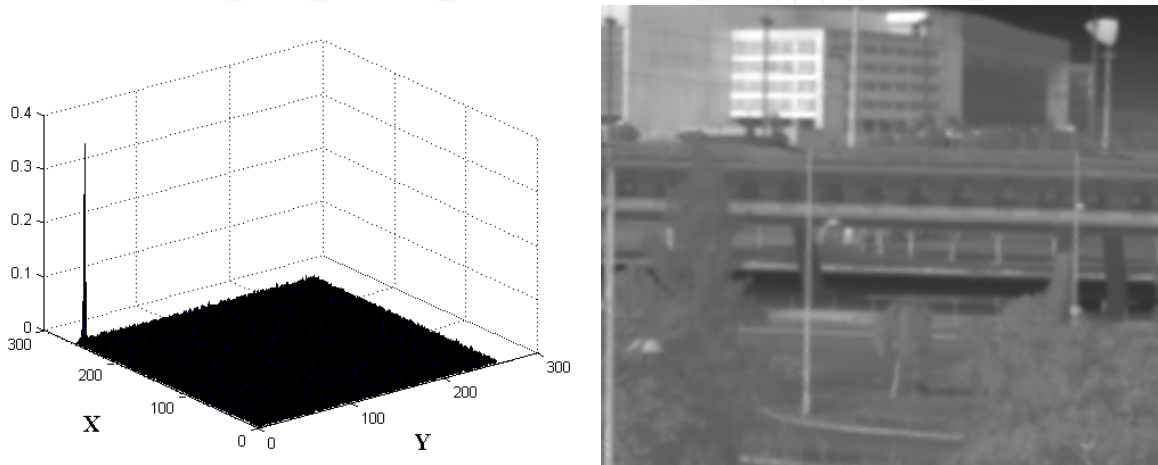


Fig. 8. Left) Resulting $q_t(x,y)$. Right) $Im^{S_t}(x,y)$, stabilized version of $Im_t(x,y)$

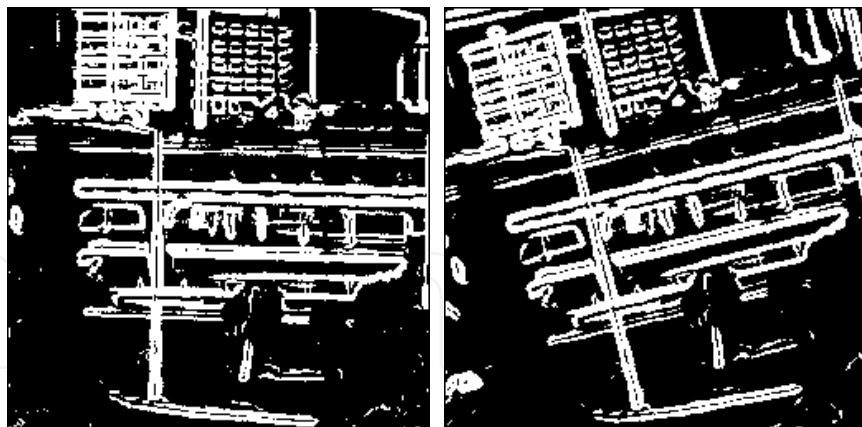


Fig. 9. Resulting images after applying Sobel edge detector to $r(x,y)$ and $s(x,y)$ in Fig. 4

5.2 Drift correction

Constrained resolutions of digital images generate small sub-pixel errors in image matching. These errors have no perceptible consequences unless they accumulate originating drifts in incremental stabilization schemes. The simplest way to avoid drifts is to adopt an absolute stabilization approach, in which images are stabilized with respect to the same reference image $Im_0(x,y)$. Different schemes were analysed to update the reference images depending

on the application. The reference image is considered valid if the overlapping between $Im_i(x,y)$ and $Im_0(x,y)$ is above a certain threshold. Otherwise, a new reference image is needed.

5.3 Operating modes

The resolution in the computation of σ , α , x_0 and y_0 is highly dependent on the size of the images. Although it is not easy to establish mathematically this dependence in real noisy images, it is possible to analytically obtain some qualitative conclusions.

Assume that the size of $r(x,y)$ and $s(x,y)$ is $M \times M$ and that the size of $s_p(\theta,\lambda)$ and $r_p(\theta,\lambda)$ is $W \times W$. The resolution of the rotation angle depends on the size of $S_p(v,\omega)$ and $R_p(v,\omega)$, i.e. $W \times W$. The minimum detectable rotation angle is $a_{min} \approx 2/W$. The value of W also has influence on the errors in the computation of the rotation angle. The higher W , the higher accuracy in the computation of the rotation angle. SPOMF produces broader peaks in $q_r(\theta,\lambda)$ in case of using smaller W . The value of W depends on the number of different radius values considered in the polar conversion, which is constrained by the size of $S(u,v)$ and $R(v,\omega)$, which size is $M \times M$.

The size of $r(x,y)$ and $s(x,y)$, i.e. M , has influence on the resolution and errors in the computation of the translations. Low values of M involve broader peaks in $q_i(x,y)$, which involves poorer peak detection.

To deal with different stabilization accuracies and computer burden, two operation modes have been selected: *Mode1* and *Mode2*. *Mode1* uses low values of M and W . *Mode1* has moderate stabilization capability. *Mode2* uses high values of M and W . *Mode2* stabilizes more accurately at the expense of higher computer burden. *Mode2* is applied when high accuracy is required or in case of vibrations of high magnitude.

5.4 Increase accuracy through sub-pixel resolution

As described in the above sub-section the resolutions in the computation of σ , α , x_0 and y_0 are limited by sizes of $q_r(\theta,\lambda)$ and $q_i(x,y)$ and thus, by the sizes of the images. Noise broadens the peaks of the SPOMF. A sub-pixel estimation method is used to determine with accuracy and robustness the location of the peak in $q_r(\theta,\lambda)$ or $q_i(x,y)$. This sub-pixel method detects the maximum of $q_r(\theta,\lambda)$ or $q_i(x,y)$ and defines a neighborhood around it. The centroid in the neighborhood around the maximum peak is considered as the location of the peak. This simple method allows incrementing the accuracy in peak localization. It also increases the robustness of the maximum peak detection in case $q_r(\theta,\lambda)$ and $q_i(x,y)$ have secondary peaks near the maximum peak.

6. Implementation

The proposed method has been implemented in various problems illustrating its flexibility.

6.1 Stabilization of images from a hovering UAV

The proposed technique has been tested with the HERO3 helicopter, see Fig. 10 developed by the Group of Robotics, Vision and Control (GRVC) at the University of Seville (Spain).

The application is to perform building inspection and particularly the detection of thermal heat losses on the building envelope. The main perception sensors on board HERO3 are an *Indigo Omega* infrared micro-camera and a visual camera. The *Indigo Omega* use infrared un-cooled detectors, with low sensitivity and high noise levels. The helicopter is equipped with an onboard PC-104 computer for stabilizing the infrared images and applying heat losses detection.



Fig. 10. HERO3 during building inspection experiments carried out at the School of Engineering of Seville (Spain)

The vibrations in the images have diverse origins. When the UAV is hovering, the images can be affected by low-frequency oscillations originated by the compensations of the UAV control systems. Also, images are affected by high-frequency vibrations induced by the UAV engine. When the helicopter is hovering at certain position and orientation, the low-magnitude distortions between consecutive images can be approximated by translations and rotations.

The method was customized for the application. It was assumed that there is no change in the scaling factor between images: it only corrected rotations and translations. The onboard infrared micro-camera weights 120 gr. and has a resolution of 160x120 pixels. $r(x,y)$ and $s(x,y)$ were selected of size 64x64. During the inspection the objective is to stabilize images when regions of interest are present in the images. An automatic function was developed to detect when an object of interest -windows in this problem- is present. At that moment the image stabilization method is triggered and this image is used to initialize the stabilization method.

Fig. 11 shows consecutive infrared images taken from an infrared camera onboard HERO3 in an experiment carried out in December 2005. The image noise level is rather high. The translations between the images can be observed. The stabilized images are in Fig. 12. The image stabilization time was 12,6 ms in a PC-104 with computing power similar to a Pentium III 800 MHz.

6.2 Stabilization of infrared images from a ship

The proposed method has been used for stabilization of images from an infrared camera on a ship. The objective was to detect other ships and obstacles using automatic computer vision methods. The stabilization method should keep the images stable despite the sea condition and ship motion. This application required a high-performance infrared camera

and optical system with high resolution and low noise level. Infrared images have some difficulties in sea scenes. Water absorbs infrared radiation: water appears in images with low radiation levels. Sky also appears in infrared images with low radiation levels. The horizon line can be well perceived in infrared images.

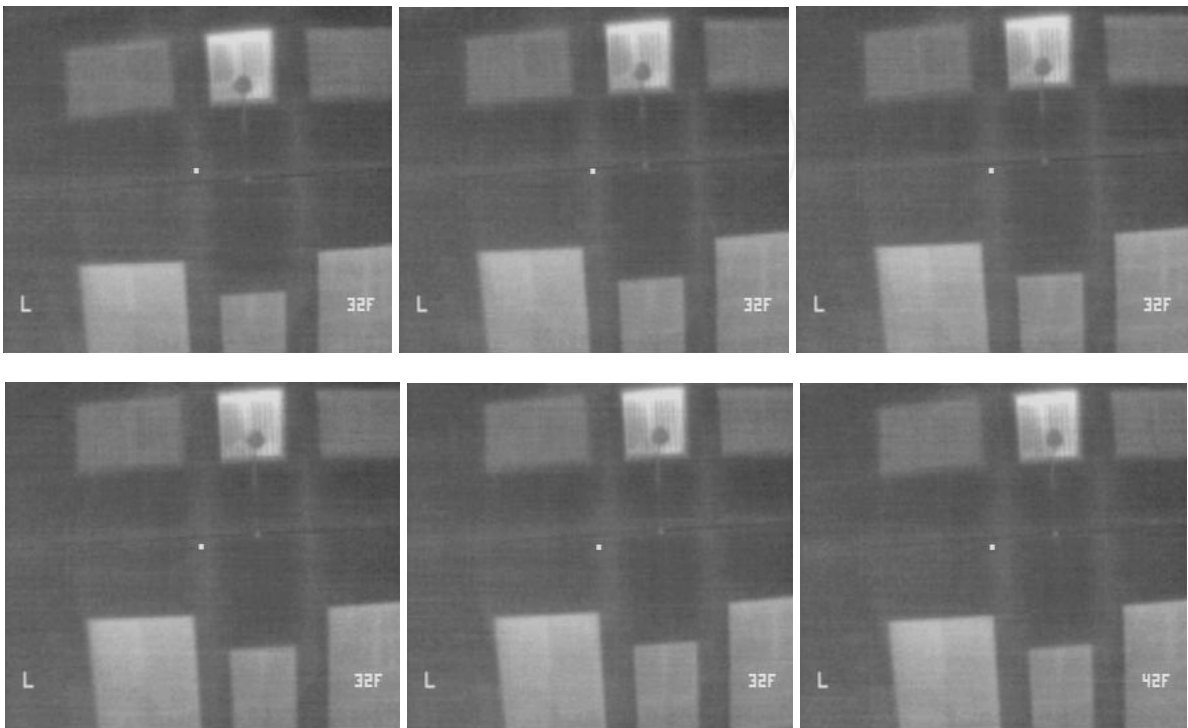


Fig. 11. Sequence of consecutive images from an *Indigo Omega* onboard HERO3

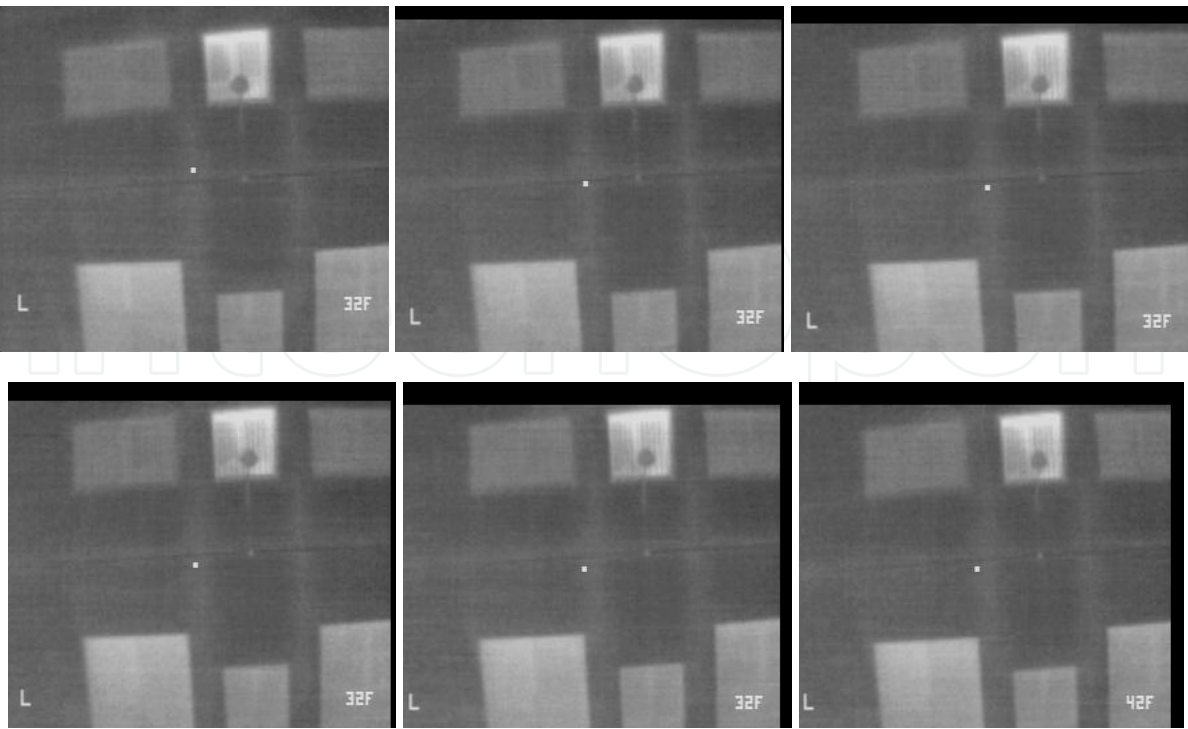


Fig. 12. Stabilized sequence of images

In images with no ships or objects, the stabilization method should greatly rely on the horizon line, what is useful to compensate rotations and translations in the vertical axis but is not enough to correct translations in the horizontal axis. If the image contains objects the method can correct also translations in the horizontal axis. Fortunately, in this application the method had to operate only if some threats for the ship navigation were found. An automatic tool was used to detect the presence of objects and to trigger the stabilization method.

The implemented version included the image enhance methods to deal with the lack of contrast. The resolution of the cameras used was 512x512. Hence, $r(x,y)$ and $s(x,y)$, input of the algorithm, were selected of size 256x256 in *Mode1* and 512x512 in *Mode2*. A short video showing the performance of the method can be found in (web1). In this video the ship was in the port and the camera was pointing at nearby forested area. Only very low magnitude vibrations can be observed. The method corrects the vibrations except the sub-pixel errors described in Section 5. The translational offsets x_0 (left) and y_0 (right) computed along one sequence of images are shown in Fig. 13. The time to stabilize one image was 102,1 ms in *Mode1* and 346,6 ms in *Mode2* in a computer with processing power similar to a Pentium III t 800 MHz.

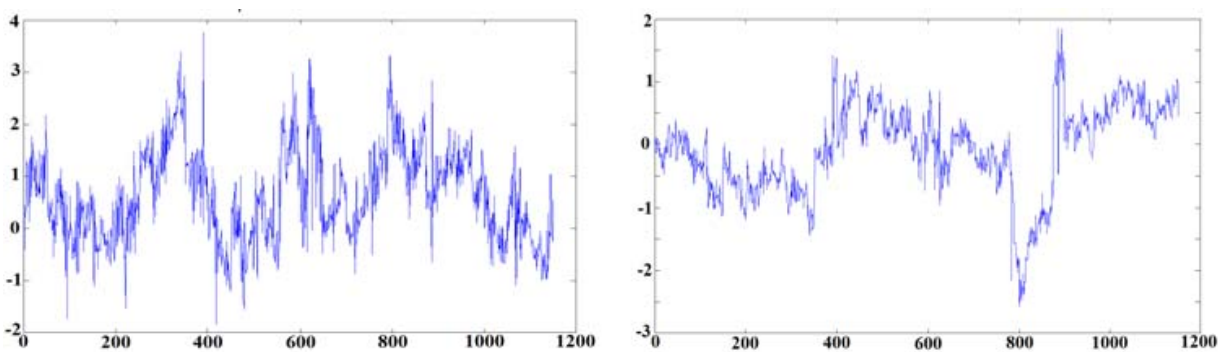


Fig. 13. Values of x_0 (left) and y_0 (right) obtained in the stabilization of a sequence of 1000 images from an infrared camera mounted on a ship

6.3 Stabilization of images from a ground vehicle in motion

The infrared camera is mounted on a ground vehicle. The camera is located on a pole at certain height and pointing such that it has a general view of an area. The vehicle can be static or moving on the ground. If the vehicle is static the method employed is similar to that used in Section 6.2. If the vehicle is moving, the stabilization method compensates rotations and vertical translations but do not cancel horizontal translations. When the vehicle is moving on the horizontal plane, even in the absence of vibrations, the images have horizontal translations. Without using any other sensors but the camera itself, it is not possible to differentiate if a horizontal translation is originated by the vehicle motion or by vibrations.

Fig. 14 shows 6 consecutive images taken from one camera under vibrations. The camera moves on the horizontal plane along the image sequence. The images resulting from the complete stabilization version of the method -taking the first image as reference- are in Fig. 15. The objects appear static along the sequence. The reduction in the overlapping between the images generates black patches of increasing size on the corrected images. The resulting

images after partial stabilization, i.e. without correcting the horizontal translations, are shown in Fig. 16. The scene appears static except for the horizontal axis.

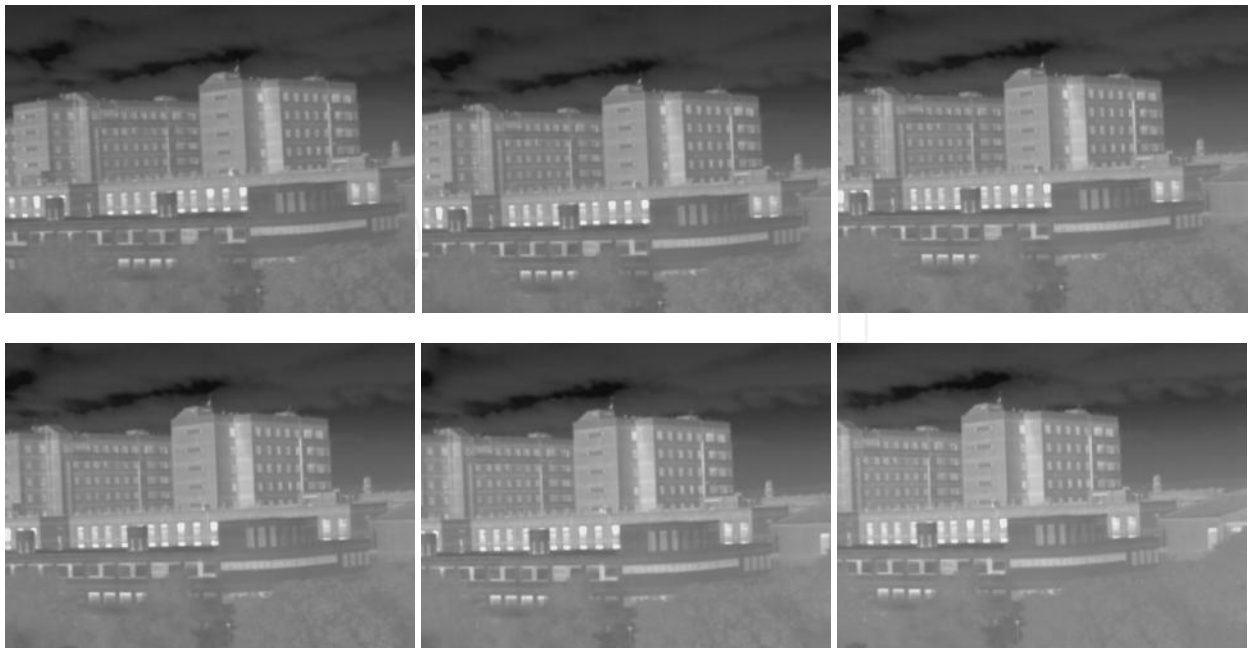


Fig. 14. Consecutive images taken from a camera under vibrations and moving on the horizontal plane

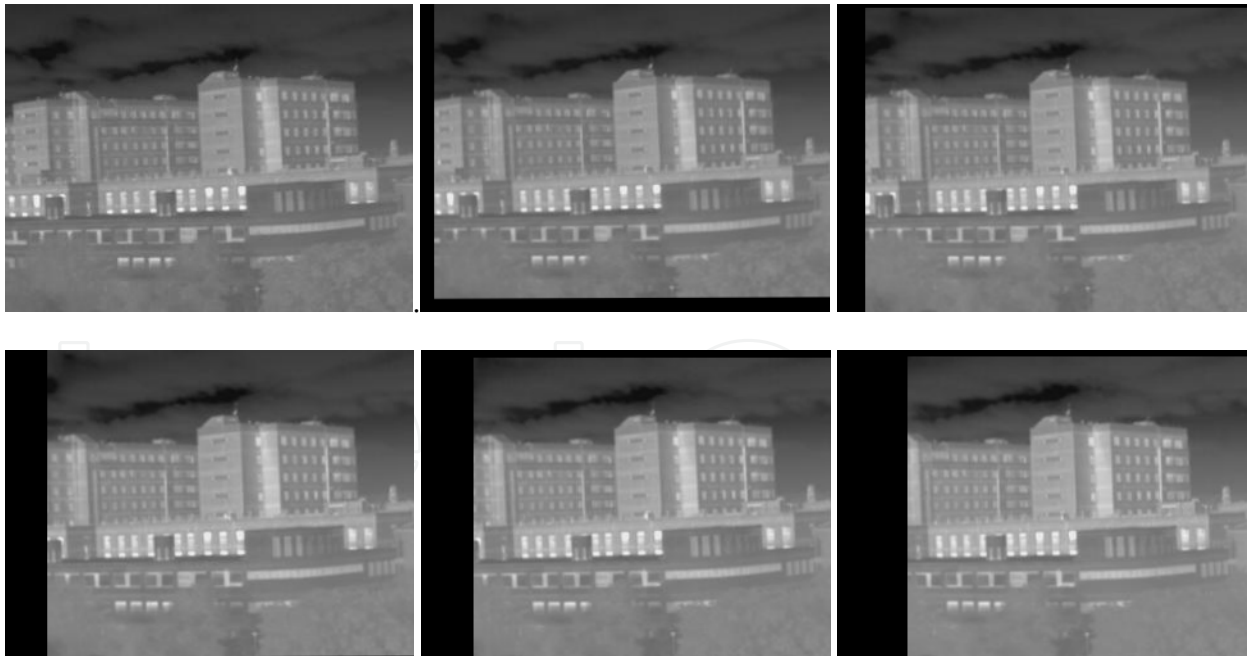


Fig. 15. Images resulting from the complete stabilization version of the method

Both versions, complete and partial stabilization, were developed, each of them with two modes: *Mode1* -image size is 256x256- and *Mode2* -image size is 512x512. An operator can change the option (complete or partial stabilization) and mode during operation. In both cases the stabilization times were similar to those in Section 6.2.

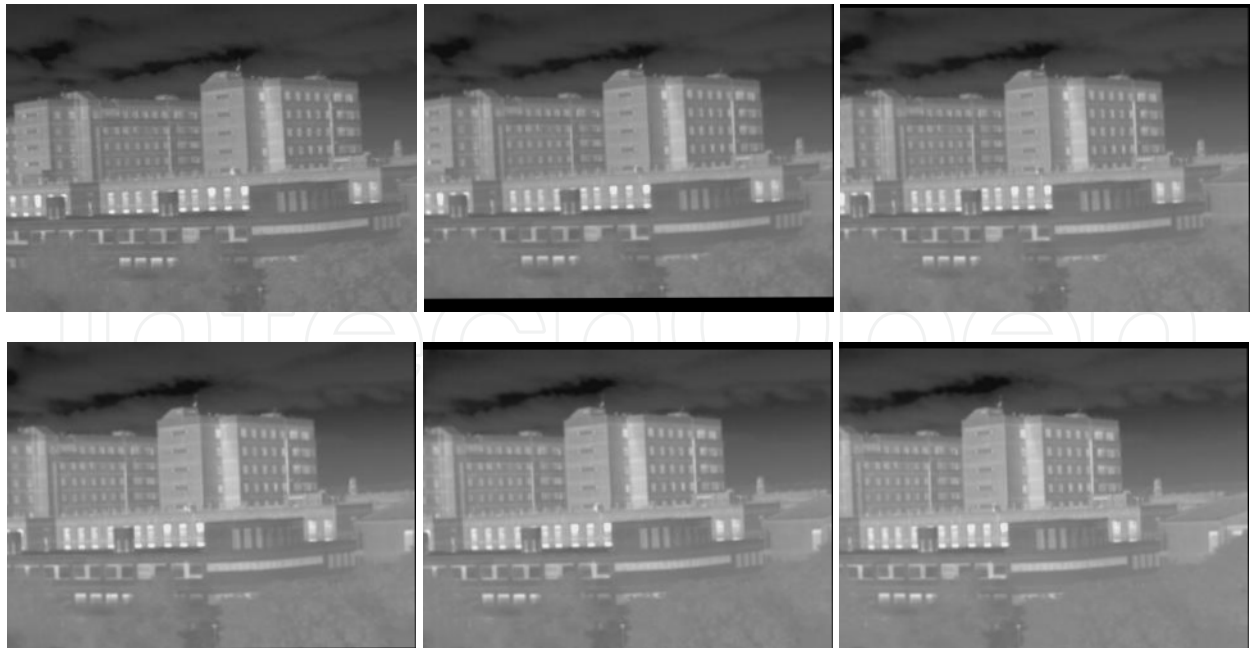


Fig. 16. Images resulting from partial stabilization. The scene appears static except for the horizontal axis

7. Conclusions

This chapter presents a technique for stabilization of sequences of infrared images using frequency domain methods. The work is motivated by the lack of local features and high noise levels commonly present in infrared images, which hampers stabilization methods based on feature matching. Instead, the proposed method relies on the global structure of the image. The described method makes extensive use of Fourier-Mellin transforms and Symmetric Phase Only Matched Filtering. The Fourier-Mellin transform allows determining the rotation, scaling and rotation between two images by converting the images to log-polar coordinates. Symmetric Phase Only Matched Filtering enhances the performance of image matching in the presence of noise.

The main strength of the proposed method is the capability to deal with images in which it is difficult to find clear and repeatable features. Its flexibility and capability for balancing between stabilization accuracy and computational cost is also remarkable and allows its customization to applications with different requirements.

The chapter also concerns several practical aspects that have been considered to increase its robustness and accuracy including image contrast enhance and the definition of modes with different performance and computer burden. The method was implemented for real-time execution with low computing resources. Different versions were implemented and validated in several applications. Three of them are briefly summarized in the chapter.

Software implementation and real-time execution have been two main requirements in the design of the methods. They originated the development of stabilization modes to balance between accuracy and computer cost. The implementation of the stabilization method in FPGA is interesting to reduce the computer burden of the main processor. In this case, *Mode2*, with better stabilization performance, could be used in all conditions.

The combination of visual and infrared images provides interesting synergies in a growing number of problems. The differences between images from cameras in different spectral bands would also hamper the application of feature-based methods. The use of the proposed method for matching images in different spectral bands is object of current work.

8. References

- Abu-Mostafa, Y.S. & Psaltis, D. (1984). Recognition aspects of moment invariants. *IEEE Trans. Pattern Anal. Mach. Intel.*, Vol. 16, No. 12, pp. 1156-1168, ISSN 0162-8828
- Barnea, D. I. & Silverman, H. F. (1972). A class of algorithms for fast image registration. *IEEE Trans. Computers* C-21, Vol. C-21, No. 6, pp. 179-186, ISSN 0018-9340
- Bay, H; Ess, A.; Tuytelaars, T; Van Gool, L. (2008). SURF: Speeded up Robust Features. *Computer Vision and Image Understanding*, Vol. 110, No. 3, pp. 346-359, ISSN 1077-3142
- Carvajal, E; Jimenez-Espadafor, F; Becerra, J.A.; Torres, M. (2011). Methodology for the Estimation of Cylinder Inner Surface Temperature in an Air-Cooled Engine. *Applied Thermal Engineering*, Vol. 31, No. 8-9, pp. 1474-1481, ISSN 1359-4311
- Chen, Q.-S.; Defrise, M.; Deconinck, F. (1994). Symmetric phase-only matched filtering of Fourier-Mellin transforms for image registration and recognition. *IEEE Trans. on Pattern Analysis and Machine Intel.*, Vol. 16, No. 12, pp. 1156-1168, ISSN 0162-8828
- Ersoy, O.K. & Zeng, M. (1989). Nonlinear matched filtering. *J. Opt. Soc. Am. A*, Vol. 6, No. 5, pp. 636-648, ISSN 1084-7529
- Fischler, M.A. & Bolles, R.C. (1981). Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM*, Vol. 24, No. 6, pp. 381-395, ISSN 0001-0782
- Hartley, R. & Zisserman, A. (2004). Multiple View Geometry in Computer Vision. *Cambridge University Press*, ISBN 0521540518, Cambridge (United Kingdom)
- Hudson, R.D. (1969). Infrared System Engineering. *John Wiley & Sons*, ISBN 0471418501, New York (USA)
- Maldague, X.P.V. (2001). Theory and Practice of Infrared Technology for Nondestructive Testing. *John Wiley & Sons*, ISBN 0471181900, New York (USA)
- Segeman, J. (1992). Fourier cross correlation and invariant transformations for an optimal recognition of functions deformed by affine groups. *J. Opt. Soc. Am. A*, Vol. 9, No. 6, pp. 895-902, ISSN 1084-7529
- Tomasi, C. (1991). Shape and motion from image streams: A factorization method, PhD Thesis Carnegie Mellon University.
- Zitová, B. & Flusser, J. (2003). Image registration methods: a survey. *Image and Vision Computing*, Vol. 21, No. 11, pp. 977-1000, ISSN 0262-8856
- <http://www.youtube.com/watch?v=inPUcDIHC4s>, Retrieved on 14/03/2012

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](https://creativecommons.org/licenses/by/3.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

IntechOpen

IntechOpen