On Quadratic Interpolation of Image Cross-Correlation for Subpixel Motion Extraction

Bian Xiong, Qinghua Zhang, Vincent Baltazart

Article

On Quadratic Interpolation of Image Cross-Correlation for Subpixel Motion Extraction

Bian Xiong [1,2], Qinghua Zhang [1,2,*] and Vincent Baltazart [1]

[1] Université Gustave Eiffel, 44344 Bouguenais, France
[2] Inria, 35042, Rennes, France
* Correspondence: Qinghua.Zhang@inria.fr

Abstract: Digital image correlation techniques are well known for motion extraction from video images. Following a two-stage approach, the pixel-level displacement is first estimated by maximizing the cross-correlation between two images, then the estimation is refined in the vicinity of the cross-correlation peak. Among existing subpixel refinement methods, quadratic surface fitting (QSF) provides good performances in terms of accuracy and computational burden. It estimates subpixel displacement by interpolating cross-correlation values with a quadratic surface. The purpose of this paper is to analytically investigate the QSF method. By means of counterexamples, it is first shown in this paper that, contrary to a widespread intuition, the quadratic surface fitted to the pixel-level cross-correlation values in the neighborhood of the cross-correlation peak does not always have a maximum. The main contribution of this paper then consists in establishing the mathematical conditions ensuring the existence of a maximum of this fitted quadratic surface, based on a rigorous analysis. Algorithm modifications for handling the failure cases of the QSF method are also proposed in this paper, in order to consolidate it for subpixel motion extraction. Experimental results based on two typical types of images are also reported.

Keywords: Subpixel motion extraction; digital image correlation; quadratic surface fitting; subpixel refinement.

This paper is an extended version of the conference paper [1].

1. Introduction

Computer vision techniques for motion extraction are widely developed in a huge variety of applications, including motion tracking, motion compensation, image registration [2], remote sensing [3], biomedicine [4], satellite imagery [5] and vibration analysis [6]. Within this scope, techniques of digital image correlation (DIC) are known to provide accurate results with a high computational efficiency, along with good robustness against noises. Various variants of DIC exist in the litterature, e.g., phase-only correlation (POC) [3], upsampling cross-correlation (UCC) [7], Fourier-based correlation [2], virtual image correlation [8].

In this paper, a well-known image correlation technique for subpixel motion extraction is analytically investigated. Subpixel accuracy is particularly important for video-based structural health monitoring (SHM) [6,9] and for aerial or satellite imagery [5]. In such applications, the region of interest may be represented by a small number of pixels in the captured images. It is then important to extract both multipixel and subpixel motions information from video images.

A large variety of correlation-based techniques for motion extraction with subpixel accuracy have been compared in [2,3]. The most widespread techniques are performed in a two-step process. At the pixel level, displacement is estimated by maximizing the cross-correlation between two images. To achieve subpixel accuracy, displacement estimation

Publisher's Note: Sensors 2022, 22, 1274.

is then refined in the vicinity of the cross-correlation peak. Among such refinement methods, quadratic surface fitting (QSF) provides a good trade-off between accuracy and computational burden, particularly suitable for video-based SHM, as reported in [10]. This method and its variant forms have also been investigated in [2,3,11–15]. Its good performance has been confirmed by our own experiments.

The purpose of this paper is to mathematically analyze the QSF method. By means of counterexamples, it will be shown that, contrary to a widespread intuition, the quadratic surface fitted to the cross-correlation values in the 3×3 pixels neighborhood of the correlation peak does not always have a maximum, despite the fact that the maximum pixel-level cross-correlation value is located at the center of this 3×3 pixels neighborhood. This absence of maximum leads to a failure of the QSF method, which should determine the subpixel displacement by maximizing the fitted quadratic surface interpolating the cross-correlation values in the vicinity of the cross-correlation peak. However, experiences reported by different authors and conducted by ourselves show that usually the QSF method produces satisfactory results. Then it is important to understand the conditions under which this method works correctly. In this paper, the conditions ensuring the existence of a maximum of the fitted quadratic surface will be formally analyzed. Then these conditions will be completed to make sure that the maximum is within the one-pixel vicinity of the pixel-level cross-correlation peak. Solutions will also be proposed to handle the failures cases of the QSF method by constrained optimization ensuring that the estimated subpixel displacement is within one pixel. These modifications apply only when a failure occurs, hence the extra numerical computation cost is insignificant.

This paper is an extended version of the conference paper [1]. It is organized as follows. The considered problem is formulated in Section 2. The QSF method is recalled in Section 3. Examples showing failures of the QSF method are presented in Section 4. The QSF method is then analyzed in Section 5. Handling of the failure cases is proposed in Section 6. Experimental results based on two typical types of images are reported in Section 7. Finally, conclusions are drawn in Section 8.

2. Problem statement

A locally rigid moving object is observed with a camera. It is assumed that the displacement of the observed object is small between successive images, with negligible rotation and negligible motion along the optical axis direction. Correlation processing will focus on a rectangular template, also known as region of interest (ROI), which includes either the whole moving object or some part of the object. The intensity of each pixel at instant t is denoted by $I(n, m, t)$ where the integer pair (n, m) indicates the position of the pixel in an image.

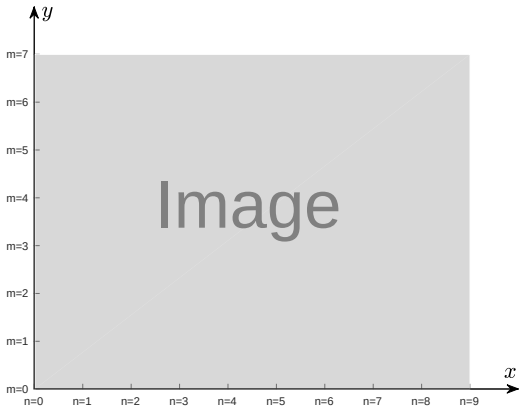

Figure 1. Illustration of the Cartesian coordinate system and integer coordinates (n, m).

Instead of the usual row and column indexes, in this paper the pair (n, m) denotes pixel integer coordinates in a Cartesian system, as illustrated in Figure 1. It will serve both to describe pixel positions in an image and to fit quadratic surfaces in the QSF method. In the second usage, the origin $(x, y) = (0, 0)$ corresponds to zero subpixel displacement. This notation choice is more suitable for the surface fitting problem formulated in the Cartesian coordinate system, in agreement with usual mathematical notations.

Motion extraction will be carried out by determining the horizontal and vertical shifts of the selected template between two images captured at instants t and $t + \Delta t$. The horizontal image shift (in number of pixels) is decomposed into an integer part $\tilde{n}$ and a fractional (subpixel) part x with $|x| < 1$, and similarly the vertical image shift is decomposed into an integer part $\tilde{m}$ and a fractional part y with $|y| < 1$, so that

$$\text{Total horizontal shift} = \tilde{n} + x \tag{1a}$$

$$\text{Total vertical shift} = \tilde{m} + y. \tag{1b}$$

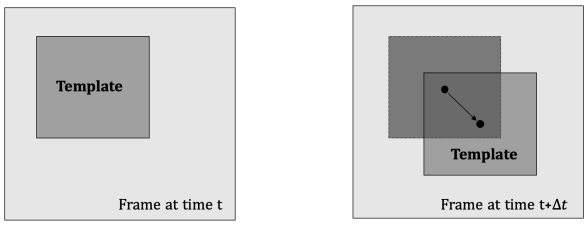

Figure 2. Template shifts between two successive images

Given two images (frames) captured at time instants t and $t + \Delta t$, as illustrated in Figure 2, template shifts are usually estimated through a two-step process [10]. The pixel level (integer) shifts $(\tilde{n}, \tilde{m})$ are first estimated by maximizing the cross-correlation between the two image templates. At the second step, the cross-correlation is somehow interpolated in the vicinity of the cross-correlation peak to estimate the subpixel shifts (x, y).

At the pixel level, let the cross-correlation be denoted by

$$r(\tilde{n}, \tilde{m}) \triangleq \sum_{(n,m)\in\mathbb{T}} I(n, m, t) I(n - \tilde{n}, m - \tilde{m}, t + \Delta t), \tag{2}$$

where $\mathbb{T}$ denotes the set of integer pairs (n, m) corresponding to the pixels belonging to the considered template. The dependences on $I(n, m, t)$ and on $I(n, m, t + \Delta t)$ are omitted in the notation $r(\tilde{n}, \tilde{m})$ for a lighter presentation. The search for the cross-correlation peak is formulated as:

$$(\tilde{n}^*, \tilde{m}^*) = \arg \max_{\substack{-\tilde{N}\le\tilde{n}\le\tilde{N} \\ -\tilde{M}\le\tilde{m}\le\tilde{M}}} r(\tilde{n}, \tilde{m}), \tag{3}$$

where $\tilde{N}$ and $\tilde{M}$ are two positive integers specifying the search ranges respectively for horizontal and vertical shifts.

In order to gain subpixel accuracy, at the second step, the cross-correlation $r(\tilde{n}, \tilde{m})$ is somehow interpolated for non integer shifts so that the correlation maximization (3) can be generalized to subpixel shifts.

In the QSF method, this interpolation is made by fitting, in the least squares sense, a second degree polynomial (or, geometrically, a quadratic surface) to the value of $r(\tilde{n}^*, \tilde{m}^*)$ and to the 8 neighboring cross-correlation values $r(\tilde{n}, \tilde{m})$, namely $r(\tilde{n}^* + n, \tilde{m}^* + m)$ with $n, m \in \{-1, 0, +1\}$. Then the maximum of the fitted polynomial yields the estimated subpixel shifts between the two templates [10].

More formally, the integer shifts $(\tilde{n}^*, \tilde{m}^*)$ being already estimated, let $p(x, y)$ denote the second degree polynomial fitted, in the least squares sense, to $r(\tilde{n}^* + x, \tilde{m}^* + y)$ for $x, y \in \{-1, 0, +1\}$. Then the subpixel shifts are estimated as

$$(x^*, y^*) = \arg \max_{(x,y)\in\mathbb{R}^2} p(x, y), \tag{4}$$

and the estimated total shifts amount to

$$(\tilde{n}^* + x^*, \tilde{m}^* + y^*). \tag{5}$$

Satisfactory experimental results of this method have been reported by different authors, e.g., [2,3,10]. Our own experiments confirm also its good performance compared to other existing methods for subpixel shift estimation. The main purpose of this paper is to consolidate the theoretical basis of this method.

More specifically, the QSF method, as recalled above, assumes implicitly that the second degree polynomial $p(x, y)$ fitted to the 9 correlation values $r(\tilde{n}^* + n, \tilde{m}^* + m)$, with $n, m \in \{-1, 0, +1\}$, always has a unique global maximum, corresponding to (x^*, y^*) located in the one-pixel vicinity of $(x, y) = (0, 0)$, so that the total shifts as expressed in (5) do not fall too far from the pixel level optimal shifts $(\tilde{n}^*, \tilde{m}^*)$. This paper will investigate the following issues.

1. Does the quadratic surface fitted in the QSF method always have a maximum in the one-pixel vicinity of the pixel-level cross-correlation peak?
2. If the answer to the first question is no, what are the conditions ensuring that the fitted quadratic surface has a maximum, and moreover, the maximum is located in the one-pixel vicinity of the pixel-level cross-correlation peak?
3. What should the algorithm do if ever the fitted quadratic surface has no maximum, or if its maximum is outside the one-pixel vicinity of the pixel-level cross-correlation peak?

3. Quadratic surface fitting for subpixel refinement

The QSF method is recalled in this section before its analysis in the next sections.

Let $(\tilde{n}^*, \tilde{m}^*)$ be resulting from the pixel level maximization (3).

The 9 integer pairs (n, m), with $n, m \in \{-1, 0, +1\}$, form a 3×3 grid[1]:

$$\mathbb{G} = \begin{bmatrix} (-1,1) & (0,1) & (1,1) \\ (-1,0) & (0,0) & (1,0) \\ (-1,-1) & (0,-1) & (1,-1) \end{bmatrix}. \tag{6}$$

Accordingly, the 9 cross-correlation values $r(\tilde{n}^* + n, \tilde{m}^* + m)$ normalized by the maximum cross-correlation value $r(\tilde{n}^*, \tilde{m}^*)$ and denoted by

$$\gamma(n, m) \triangleq \frac{r(\tilde{n}^* + n, \tilde{m}^* + m)}{r(\tilde{n}^*, \tilde{m}^*)} \tag{7}$$

form a matrix

$$\Gamma \triangleq \begin{bmatrix} \gamma(-1,1) & \gamma(0,1) & \gamma(1,1) \\ \gamma(-1,0) & \gamma(0,0) & \gamma(1,0) \\ \gamma(-1,-1) & \gamma(0,-1) & \gamma(1,-1) \end{bmatrix}. \tag{8}$$

[1] The grid $\mathbb{G}$ is formed by 9 integer pairs organized in 3 rows and 3 columns. It is also seen as a set with the integer pairs as elements, so that notations like $(-1, 1) \in \mathbb{G}$ can be used.

The central entry of Γ

$$\gamma(0,0) = \frac{r(\tilde{n}^* + 0, \tilde{m}^* + 0)}{r(\tilde{n}^*, \tilde{m}^*)} = 1. \tag{9}$$

is the maximum cross-correlation value $r(\tilde{n}^*, \tilde{m}^*)$ normalized by itself, hence this central entry $\gamma(0,0) = 1$ is the maximum value among all the 9 entries of Γ.

The second degree polynomal

$$p_\theta(x,y) = \theta_1 + [\theta_2\ \theta_3]\begin{bmatrix} x \\ y \end{bmatrix} + [x\ y]\begin{bmatrix} \theta_4 & \theta_5/2 \\ \theta_5/2 & \theta_6 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}, \tag{10}$$

with the vector $\theta \in \mathbb{R}^6$ collecting the scalar coefficients $\theta_1, \ldots, \theta_6$, is then fitted to the entry values of Γ for $(x,y) \in \mathbb{G}$, by solving the least squares problem

$$\min_{\theta \in \mathbb{R}^6} \sum_{(n,m)\in\mathbb{G}} [\, p_\theta(n,m) - \gamma(n,m)]^2, \tag{11}$$

where $\mathbb{G}$ is the grid defined in (6).

Does this fitted second degree polynomial $p_\theta(x,y)$ always have a unique global maximum?

As explained above, the central entry of the matrix Γ, namely $\gamma(0,0) = 1$, is the maximum value among all the 9 entries of Γ. It then seems reasonable to expect that the second degree polynomial (or quadratic surface) $p_\theta(x,y)$ fitted to the 9 entries of Γ has a maximum somehow close to the (maximum) central entry of the matrix Γ, which corresponds to the origin $(x,y) = (0,0)$ of the coordinate system characterizing the fitted quadratic surface.

Unfortunately, the fact that the central entry $\gamma(0,0)$ is the maximum value among the 9 entries of Γ does not really ensure that the fitted second degree polynomial $p_\theta(x,y)$ always has a global maximum, as demonstrated by the following counterexamples.

4. Counterexamples

Three examples with either synthetic or real-world images are presented below to show that the quadratic surface fitted in the QSF method does not always have a maximum in the close vicinity of the pixel-level cross-correlation peak.

Example 1

The first counterexample with synthetic images has been chosen for its simplicity so that it can be easily reproduced. The Matlab code for generating the presented result is available for download [16]. To show the robustness of this counterexample, the Matlab code can be optionally run with some random noises added to the generated synthetic images, though the result presented below is noise-free. Counterexamples based on true images (as presented in Examples 2 and 3) are also included in [16].

Consider two binary images of 18×18 pixels as shown in Figure 3, with the template (ROI) chosen as the red square window of 8×8 pixels in each image. The template in the first image contains a diagonal pattern. In the second image, this diagonal pattern is shifted by 1 pixel toward the right and also by 1 pixel toward the bottom. In these binary images, the intensity is 1 at the darker pixels and 10 at the brighter pixels.

The normalized 3×3 cross-correlation matrix around the peak as defined in (8) is

$$\Gamma = \begin{bmatrix} 0.2236 & 0.2236 & 0.8059 \\ 0.2236 & 1 & 0.2236 \\ 0.8059 & 0.2236 & 0.2236 \end{bmatrix}. \tag{12}$$

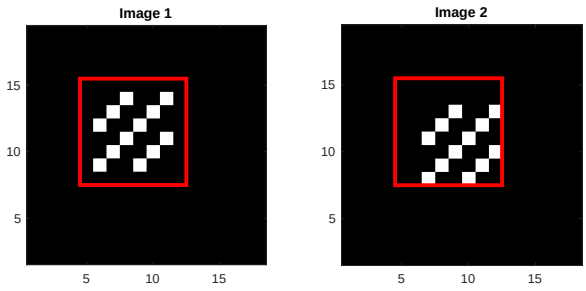

Figure 3. Two synthetic images of 18×18 pixels. The intensity is 1 at the darker pixels and 10 at the brighter pixels, The template (ROI) is selected as the red square window of 8×8 pixels.

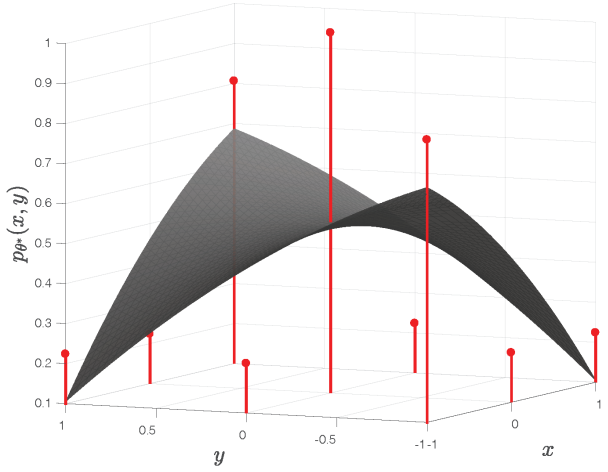

Figure 4. For the images in Figure 3, the fitted quadratic surface exhibits a saddle point. The vertical red line segments represent the entries of the normalized cross-correlation matrix Γ.

Fitting the second degree polynomial $p_\theta(x,y)$ to Γ for $(x,y) \in \mathbb{G}$ by solving the least squares problem (11) yields the solution

$$\theta^* = \begin{bmatrix} 0.5256 \\ 0.0000 \\ 0.0000 \\ -0.0647 \\ 0.2911 \\ -0.0647 \end{bmatrix}. \tag{13}$$

The corresponding quadratic surface $p_{\theta^*}(x,y)$ exhibits a saddle point, as illustrated in Figure 4. It has no global maximum, despite the fact that the central entry of the matrix Γ is its largest entry.

This counterexample clearly invalidates the widespread intuition that the fitted polynomial $p_\theta(x,y)$ in the QSF method always has a global maximum.

Example 2

This example is based on true images of the moon, included in Matlab as part of image examples. The Matlab code generating the presented results are available for download [16]. Two images of the moon with 537×358 pixels are shown in Figure 5. Like in Example 1, with some chosen template (ROI), cross-correlations between the two images are computed, and a quadratic surface is fitted to the normalized 3×3 cross-correlation matrix around the peak. In most situations, the fitted surface has a maximum

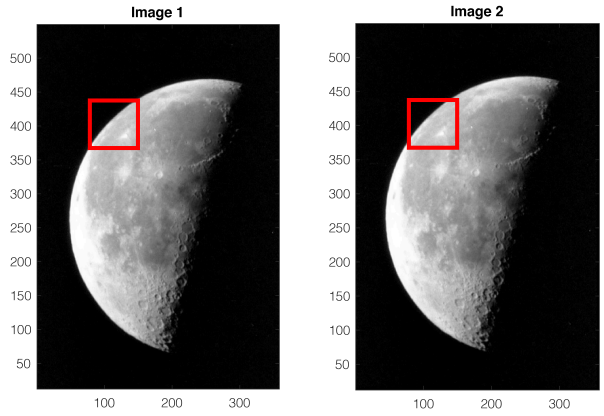

Figure 5. Two images of the moon of 537×358 pixels. The template is chosen as the red windows.

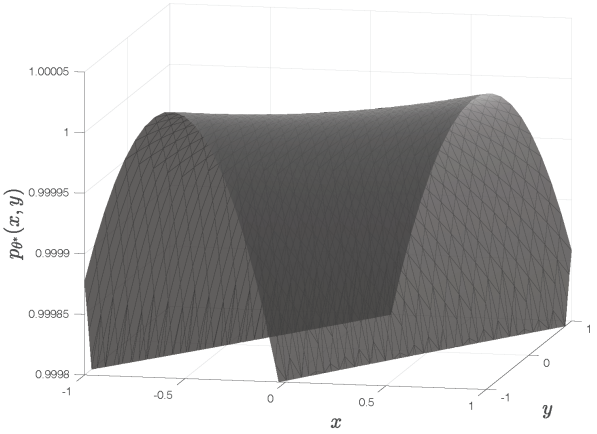

Figure 6. For the images and template in Figure 5, the fitted quadratic surface exhibits a saddle point.

close to the pixel-level cross-correlation peak, but unexpected cases do happen with some particular template choices.

For the chosen template illustrated by the red windows in Figure 5, the fitted quadratic surface is shown in Figure 6. The surface exhibiting a saddle point has no maximum. This result can be reproduced with the Matlab code downloadable from [16].

Example 3

This example is based the same images of the moon as in Example 2, but with another choice of the template for cross-correlation computation, as illustrated by the red windows in Figure 7. The fitted quadratic surface shown in Figure 8 has a maximum located at

$$(x^*, y^*) = (13.76, 4.28),$$

as indicated by the cyan dot on the surface. This maximum is far outside the square area of the one-pixel vicinity of the pixel-level cross-correlation peak delimited by the red dotted lines on the bottom plane in Figure 8. As a subpixel refinement, the maximum of the fitted surface should satisfy $|x^*| < 1$ and $|y^*| < 1$, which is not the case in this example. The details of this result are available from [16].

The examples presented in this section confirm that, in some rare situations, the quadratic surface fitted in the QSF method does not have a maximum, or has a maximum far away from the square area of the one-pixel vicinity of the pixel-level cross-correlation peak.

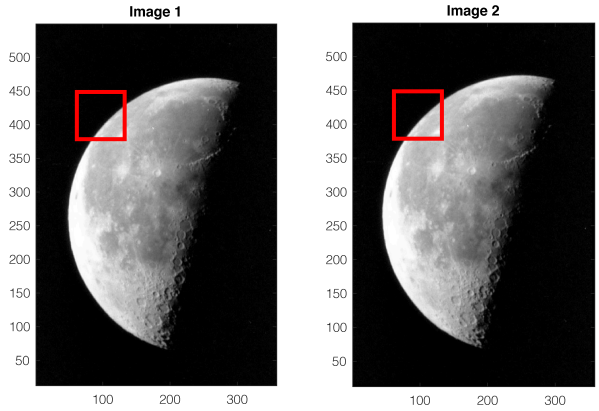

Figure 7. Two images of the moon, the same as in Figure 5, but with a different template choice shown by the red windows.

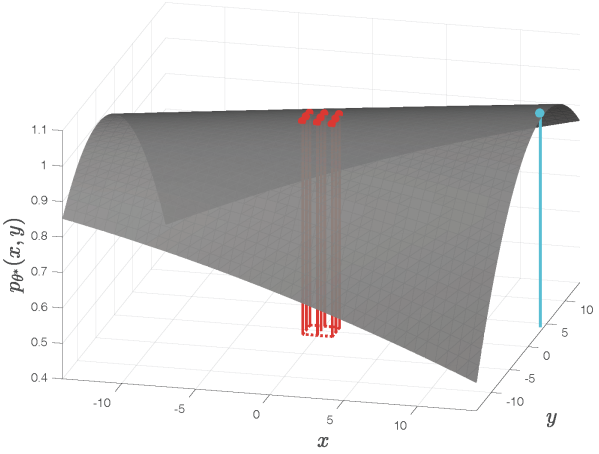

Figure 8. For the images and template in Figure 7, the fitted quadratic surface has a maximum located at $(x^*, y^*) = (13.76, 4.28)$, as shown by the cyan dot on the surface. The red dotted lines on the bottom plane delimit the square area of the one-pixel vicinity around the pixel-level cross-correlation peak. As a subpixel refinement, the maximum of the fitted surface should satisfy $|x^*| < 1$ and $|y^*| < 1$, which is not the case in this example. See [16] for the details of this example.

After the answer to the first question raised at the end of Section 2, the following sections will then answer the other two remaining questions.

5. Conditions for the existence of a maximum

In elementary algebra, it is well known that the second degree polynomial $p_\theta(x, y)$ as expressed in (10) has a unique global maximum if its Hessian matrix

$$H \triangleq \begin{bmatrix} \theta_4 & \theta_5/2 \\ \theta_5/2 & \theta_6 \end{bmatrix} \tag{14}$$

is negative definite [17]. However, this simply stated fact does not directly help to understand how the normalized cross-correlation values $\gamma(n, m)$ (those filling up the Γ matrix in (8)) should be, so that the fitted $p_\theta(x, y)$ has a global maximum. Because the polynomial coefficients $\theta_1, \ldots, \theta_6$ are determined from the values of $\gamma(n, m)$ by solving the least squares problem (11), it is straightforward to express the negative definiteness condition of H in terms of $\gamma(n, m)$. Then, in principle, the condition for the existence of a global maximum of $p_\theta(x, y)$ will be formulated in terms of the normalized cross-correlation values $\gamma(n, m)$. Nevertheless, this approach results in sophisticated conditions,

notably an inequality involving the determinant of H expressed in terms of $\gamma(n,m)$. For a better understanding, the result presented below will be formulated with simple and easily interpretable inequalities about the values $\gamma(n,m)$ filling up Γ. For instance, one of these simple inequalities states that the central entry $\gamma(0,0)$ of Γ is its largest entry. As shown by the previously presented counterexamples, this condition alone is not sufficient. It is then completed by similar simple inequalities.

Theorem 1. *If the normalized cross-correlation values $\gamma(n,m)$ filling up the matrix Γ satisfy*

$$\gamma(0,0) \geq \gamma(n,m) \text{ for all } n,m \in \{-1,0,+1\} \tag{15}$$
$$\gamma(0,m) > \gamma(n,m) \text{ for all } n,m \in \{-1,+1\} \tag{16}$$
$$\gamma(n,0) > \gamma(n,m) \text{ for all } n,m \in \{-1,+1\}, \tag{17}$$

then the second degree polynomial $p_\theta(x,y)$ fitted to the entries of Γ by solving the least squares problem (11) has a unique global maximum.

Interpretation of the conditions of Theorem 1.

- Inequalities (15): the central entry $\gamma(0,0)$ has the largest value among all the 9 entries of Γ.
- Inequalities (16): the middle entry $\gamma(0,\pm1)$ is the largest entry of the top or the bottom row of Γ.
- Inequalities (17): the middle entry $\gamma(\pm1,0)$ is the largest entry of the right or the left column of Γ.

Proof of Theorem 1.

In order to shorten lengthy equations and inequalities, let us introduce more compact notations for the normalized cross-correlation values $\gamma(m,m)$ filling up the matrix Γ defined in (8), so that Γ is rewritten as

$$\Gamma = \begin{bmatrix} a & e & b \\ f & i & g \\ c & h & d \end{bmatrix}. \tag{18}$$

Remark that the letters $a, b, d, \ldots, i$ fill Γ first at the 4 corners, then the middles of side rows and columns, before finishing at the central entry.

With these compact notations, the least squares solution (11) leads to

$$9\theta_1 = 2(e+f+g+h) + 5i - (a+b+c+d) \tag{19a}$$
$$6\theta_2 = (b-a) + (g-f) + (d-c) \tag{19b}$$
$$6\theta_3 = (a-c) + (e-h) + (b-d) \tag{19c}$$
$$-6\theta_4 = (e-a) + (e-b) + (h-c) + (h-d) + (i-f) + (i-g) \tag{19d}$$
$$4\theta_5 = (b-a+c-d) \tag{19e}$$
$$-6\theta_6 = (f-a) + (f-c) + (g-b) + (g-d) + (i-e) + (i-h). \tag{19f}$$

As already mentioned in this paper, the negative definiteness of the Hessian matrix H defined in (14) ensures that the polynomial $p_\theta(x,y)$ has a global maximum. Based on Sylvester's criterion[2], this negative definiteness will be checked through:

$$\theta_4 < 0 \tag{20}$$
$$\det(H) > 0. \tag{21}$$

According to the inequalities assumed in (16), e (or h, resp.) is the largest entry of the top (or bottom, resp.) row of Γ, then

$$e - a > 0, \quad e - b > 0 \tag{22}$$
$$h - c > 0, \quad h - d > 0 \tag{23}$$

and according to (15), i is the largest entry of Γ, then

$$i - f \geq 0, \quad i - g \geq 0. \tag{24}$$

These inequalities together with (19d) imply immediately (20).

It is more involved to check (21). The inequalities in (22) imply

$$\left(1 + \frac{3}{4}\right)(e - a) > 0 > \left(-1 + \frac{3}{4}\right)(e - b), \tag{25}$$

then

$$(e - a) + (e - b) > -\frac{3}{4}(e - a) + \frac{3}{4}(e - b) \tag{26}$$
$$= -\frac{3}{4}(b - a). \tag{27}$$

Repeat the reasoning from (25) to (27) while interchanging the positions of $(e - a)$ and $(e - b)$:

$$\left(1 + \frac{3}{4}\right)(e - b) > 0 > \left(-1 + \frac{3}{4}\right)(e - a), \tag{28}$$

leading to

$$(e - a) + (e - b) > \frac{3}{4}(b - a). \tag{29}$$

Combining (27) and (29) yields

$$(e - a) + (e - b) > \frac{3}{4}|b - a|. \tag{30}$$

This result expresses a relationship between the entries in the top row of the matrix Γ. A similar reasoning then leads to the following relationship between the entries in the bottom row of Γ:

$$(h - c) + (h - d) > \frac{3}{4}|c - d|. \tag{31}$$

According to the inequalities assumed in (15), i is the largest entry of Γ, then

$$i - f \geq 0, \quad i - g \geq 0. \tag{32}$$

[2] Usually Sylvester's criterion [18] is about the positive definiteness of a real symmetric (or complex Hermitian) matrix. It is trivial to translate this criterion to the case of negative definiteness.

Take the sums of the respective sides of the 4 inequalities in (30), (31) and (32), then

$$\begin{aligned}(e-a)+(e-b)+(h-c)+(h-d)+(i-f)+(i-g)&\\ > \frac{3}{4}|b-a|+\frac{3}{4}|c-d| &\quad (33)\\ \geq \frac{3}{4}|b-a+c-d|. &\quad (34)\end{aligned}$$

This result then implies that $-6\theta_4$ and $4\theta_5$, as expressed respectively in (19d) and (19e), satisfy

$$-6\theta_4 > \frac{3}{4}|4\theta_5|, \quad (35)$$

hence

$$-\theta_4 > \frac{1}{2}|\theta_5| \geq 0. \quad (36)$$

Following the same approach, it is then similarly shown that

$$-\theta_6 > \frac{1}{2}|\theta_5| \geq 0. \quad (37)$$

This last result can also be deduced from a certain "symmetry" between the formulae expressing θ_4 and θ_6 in (19d) and (19f).

It then follows from (14), (36) and (37) that

$$\det(H) = \theta_4\theta_6 - \frac{1}{4}\theta_5^2 > 0. \quad (38)$$

Therefore, the two inequalities (20) and (21) ensuring the negative definiteness of H are successfully checked. It is then established that the second degree polynomial $p_\theta(x,y)$ has a unique global maximum. □

When the fitted polynomial $p_\theta(x,y)$ has a unique global maximum, it may happen that this maximum is far away from the origin $(x,y) = (0,0)$ corresponding to the optimized integer shifts $(\tilde{n}^*, \tilde{m}^*)$, outside the square area of the one-pixel vicinity of the cross-correlation peak. Such situations are not desirable, since the subpixel refinement should not modify the estimated shifts by more than one pixel. The following result ensures that the maximum of $p_\theta(x,y)$ stays inside the one-pixel vicinity, under easily interpretable conditions.

Theorem 2. *If, in addition to the conditions of Theorem 1, the normalized cross-correlation values $\gamma(n,m)$ filling up the matrix Γ satisfy, for all $n,m \in \{-1,+1\}$,*

$$\gamma(0,m) - \gamma(n,m) > \frac{1}{5}[\gamma(0,m) - \gamma(-n,m)] \quad (39)$$

$$\gamma(n,0) - \gamma(n,m) > \frac{1}{5}[\gamma(n,0) - \gamma(n,-m)], \quad (40)$$

then the maximum of the fitted polynomial $p_\theta(x,y)$ is located at (x^*,y^*) such that $|x^*| < 1$ and $|y^*| < 1$.

Interpretation of the conditions of Theorem 2.

The conditions inherited from Theorem 1 ensure that the middle entry in each row or column of Γ is the largest entry of the row or column, without imposing any "degree of symmetry". For example, among the top row of Γ as expressed in (8),

inequalities formulated in (16) ensure that $\gamma(0,1)$ is the largest entry, but the ratio $[\gamma(0,1)-\gamma(-1,1)/[\gamma(0,1)-\gamma(1,1)]$ can be any positive number. Two of the inequalities in the extra condition (39) of Theorem 2 constrain this ratio between $1/5$ and 5, thus limiting the dissymmetry between $\gamma(-1,1)$ and $\gamma(1,1)$.

Proof of Theorem 2.

Based on Theorem 1 (its conditions are inherited here), the fitted polynomial $p_\theta(x,y)$ has a unique global maximum in $\mathbb{R}^2$, which is located at

$$x^* = \frac{2\theta_2\theta_6 - \theta_3\theta_5}{\theta_5^2 - 4\theta_4\theta_6} \tag{41}$$

$$y^* = \frac{2\theta_3\theta_4 - \theta_2\theta_5}{\theta_5^2 - 4\theta_4\theta_6}. \tag{42}$$

In order to prove $|x^*| < 1$, it will be shown that the numerator of $|x^*|$ is smaller than its denominator. The proof for proving $|y^*| < 1$ will be made similarly.

With the compact notations filling up the matrix Γ introduced in (18) for the normalized correlation values $\gamma(n,m)$, one of the inequalities contained in (39), namely

$$\gamma(0,1) - \gamma(-1,1) > \frac{1}{5}[\gamma(0,1) - \gamma(1,1)], \tag{43}$$

is translated into

$$e - a > \frac{1}{5}(e-b), \tag{44}$$

which is rewritten as

$$\frac{5}{2}(e-a) > \frac{1}{2}(e-b), \tag{45}$$

or in a slightly different form

$$\left(1+\frac{3}{2}\right)(e-a) > \left(-1+\frac{3}{2}\right)(e-b). \tag{46}$$

Then

$$(e-a) + (e-b) > \frac{3}{2}(e-b) - \frac{3}{2}(e-a) \tag{47}$$

$$= \frac{3}{2}(a-b). \tag{48}$$

A similar reasoning (by interchanging the positions of $(e-a)$ and $(e-b)$) leads to

$$(e-b) + (e-a) > \frac{3}{2}(e-a) - \frac{3}{2}(e-b) \tag{49}$$

$$= -\frac{3}{2}(a-b). \tag{50}$$

Combining (48) and (50) then amounts to

$$(e-a) + (e-b) > \frac{3}{2}|b-a|. \tag{51}$$

This last inequality concerns the entries in the top row of Γ as in (18). Similar reasonings about the bottom row, the left and right columns of Γ then lead to

$$(h-c)+(h-d) > \frac{3}{2}|c-d|. \tag{52}$$

$$(f-a)+(f-c) > \frac{3}{2}|a-c|. \tag{53}$$

$$(g-b)+(g-d) > \frac{3}{2}|b-d|.. \tag{54}$$

Because i is the largest entry of Γ (condition inherited from Theorem 1),

$$i-f \geq 0, \quad i-d \geq 0. \tag{55}$$

Summarizing (51), (52) and (55) then implies that θ_4 as expressed in (19d) satisfies

$$\begin{aligned} -6\theta_4 &= (e-a)+(e-b)+(h-c)+(h-d) \\ &\quad +(i-f)+(i-g) && (56)\\ &> \frac{3}{2}|b-a|+\frac{3}{2}|c-d| && (57)\\ &\geq \frac{3}{2}|b-a+c-d|. && (58) \end{aligned}$$

This result, together with (19e), then yields to

$$-\theta_4 > |\theta_5| \geq 0. \tag{59}$$

Some similar reasonings (due to a certain "symmetry" between θ_4 and θ_6 in (19)) then lead to

$$-\theta_6 > |\theta_5| \geq 0. \tag{60}$$

Add together the respective sides of (19b) and (19d), then

$$6(\theta_2-\theta_4) = 2[(e-a)+(h-c)+(i-f)]. \tag{61}$$

Every parenthesis at the right hand side of (61) is positive, due to conditions inherited from Theorem 1. Then

$$\theta_2-\theta_4 > 0. \tag{62}$$

Similarly,

$$6(\theta_2+\theta_4) = -2[(e-b)+(h-d)+(i-g)] < 0. \tag{63}$$

Based on the last 2 inequalities, $-\theta_4$ satisfies

$$-\theta_4 > \pm\theta_2, \tag{64}$$

hence

$$-\theta_4 > |\theta_2| \geq 0. \tag{65}$$

In the same manner, combining (19c) and (19f) yields

$$6(\theta_3-\theta_6) = 2[(f-c)+(g-d)+(i-h)] > 0 \tag{66}$$

$$6(\theta_3+\theta_6) = -2[(f-a)+(g-b)+(i-e)] < 0, \tag{67}$$

then

$$-\theta_6 > |\theta_3| \geq 0. \tag{68}$$

The following steps of the proof will be essentially based on

$$|\theta_4| > |\theta_5| \tag{69}$$
$$|\theta_6| > |\theta_5| \tag{70}$$
$$|\theta_4| > |\theta_2| \tag{71}$$
$$|\theta_6| > |\theta_3|, \tag{72}$$

respectively due to (59), (60), (65) and (68).

On the one hand, (71), (72) and (69) imply (the step leading to (74) below)

$$|2\theta_2\theta_6 - \theta_3\theta_5| \leq 2|\theta_2||\theta_6| + |\theta_3||\theta_5| \tag{73}$$
$$< 2|\theta_4||\theta_6| + |\theta_6||\theta_4| \tag{74}$$
$$= 3|\theta_4||\theta_6|; \tag{75}$$

and on the other hand, (69) and (70) lead to (the step leading to (78) below)

$$|\theta_5^2 - 4\theta_4\theta_6| \geq |4\theta_4\theta_6| - |\theta_5^2| \tag{76}$$
$$= 4|\theta_4||\theta_6| - |\theta_5||\theta_5| \tag{77}$$
$$> 4|\theta_4||\theta_6| - |\theta_4||\theta_6| \tag{78}$$
$$= 3|\theta_4||\theta_6|. \tag{79}$$

Therefore

$$|2\theta_2\theta_6 - \theta_3\theta_5| < |\theta_5^2 - 4\theta_4\theta_6|. \tag{80}$$

It is then concluded that x^* as expressed in (41) satisfies

$$|x^*| < 1. \tag{81}$$

In the same way, it is also proved that

$$|y^*| < 1. \tag{82}$$

The proof of Theorem 2 is thus completed. □

6. Handling failures of the QSF method

There are two possible failure cases:

- Case 1, the fitted polynomial $p_\theta(x,y)$ has no global maximum;
- Case 2, $p_\theta(x,y)$ has a global maximum, reached at (x^*,y^*), but $|x^*| \geq 1$ and/or $|y^*| \geq 1$.

In the first case, the only reasonable proposition is to retain the optimized integer shifts $(\tilde{n}^*,\tilde{m}^*)$ as the estimated total shifts.

In the second case, solutions on the boundary satisfying $|x^*| \leq 1$ and $|y^*| \leq 1$ are accepted. If $|x^*| > 1$ and/or $|y^*| > 1$, then the subpixel shifts are estimated by solving the constrained optimization problem

$$(x^\star,y^\star) = \arg \max_{|x|\leq 1,|y|\leq 1} p_\theta(x,y), \tag{83}$$

and the estimated total shifts amount to

$$(\tilde{n}^* + x^\star, \tilde{m}^* + y^\star). \tag{84}$$

As $p_\theta(x, y)$ is a second degree polynomial, the constrained optimization problem (83) can be solved by quadratic programing algorithms. In this considered Case 2, the unique unconstrained maximum of $p_\theta(x, y)$ is outside the square area constrained by $|x| \leq 1$ and $|y| \leq 1$, hence the constrained solution is certainly on the boundary of the constrained square area, on one of its sides or on one of its corners. Given the simplicity of this quadratic problem, instead of applying a general quadratic programing tool, the constrained optimization problem (83) can be solved as follows.

- Compute the values of $p_\theta(x, y)$ at the four corners of the square, namely $p_\theta(\pm 1, \pm 1)$.
- Find the maximums of $p_\theta(x, \pm 1)$ in x and $p_\theta(\pm 1, y)$ in y:

$$x^\star_\pm = \arg\max_{x \in \mathbb{R}} p_\theta(x, \pm 1) = -\frac{\theta_2 \pm \theta_5}{2\theta_4} \tag{85}$$

$$y^\star_\pm = \arg\max_{y \in \mathbb{R}} p_\theta(\pm 1, y) = -\frac{\theta_3 \pm \theta_5}{2\theta_6}. \tag{86}$$

Eliminate any result(s) not satisfying $|x^\star_\pm| \leq 1$ or $|y^\star_\pm| \leq 1$.

- Find the maximum value among the 4 corner values $p_\theta(\pm 1, \pm 1)$ and the non-eliminated values $p_\theta(x^\star_\pm, \pm 1)$ and/or $p_\theta(\pm 1, y^\star_\pm)$, if any. The solution of the constrained optimization problem (83) is then given by the location of this maximum.

7. Assessment based on two typical types of images

The original QSF method and the modified method are tested with two typical types of images for failure rate evaluation. As already stated in the introduction, the original QSF method works correctly in most situations. The occurence frequency of its failures depends on the type of processed images. Two typical examples are presented below in this section.

For the moon example already considered in Section 4, when the template window is close to a corner of the left side of the images, as shown in Figure 5, the selected region of interest has a diagonally dominant pattern, sharing a common characteristic with the synthetic example shown in Figure 3. For this reason, the QSF method is more likely to encounter difficulties with these images.

A second example comes from the publicly available DIC Challenge database [19,20] provided by the Society for Experimental Mechanics. The tested images contain uniform patterns, as shown in Figure 9.

For each trial with the original QSF method, the template window is placed at a different position in the processed images, as illustrated in Figures 5 and 9. Among all the trials, the number of cases where the fitted quadratic surface exhibits a saddle point (absence of maximum) is counted. The second type of failures, with maximum outside the one-pixel vicinity, is also counted. The results are summarized in Table 1.

As expected, the moon example leads to more failures (0.21% and 4.79% of the two types of failures over the total number of trials) due to the existence of diagonally dominant patterns. For the second type of failures, the maximum of the fitted surface can be far away from the the one-pixel vicinity, up to 52 pixels. On the other hand, the DIC Challenge example with uniform patterns has very few failures (8.74 ppm and 52.4 ppm).

Though failures of the original QSF method rarely happen, it is important to prevent them for reliable applications.

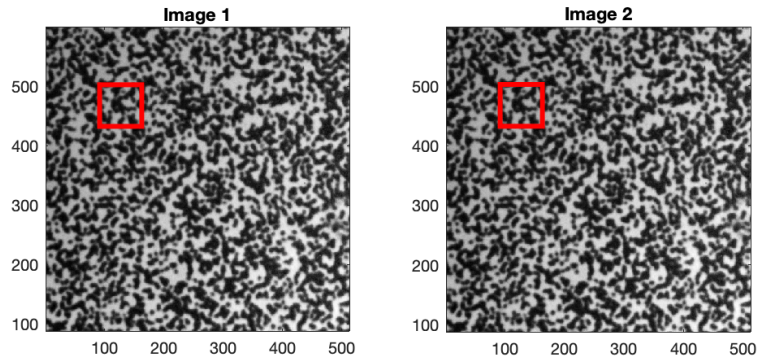

Figure 9. Two images of the DIC Challenge database [19], each with resolution 512×512 pixels. The the red windows correspond to an example of template for the QSF method.

After modifications of the QSF method as proposed in Section 6, for each the trials with the above two examples, the modified method successfully finds a subpixel displacement (x^*, y^*) such that $|x^*| < 1$ and $|y^*| < 1$.

8. Conclusion

In this paper, within the scope of digital image correlation techniques, some theoretical aspects of the QSF method have been investigated. It has been shown that, contrary to a widespread intuition, the quadratic surface fitted in the QSF method does not always have a maximum. Then for a better understanding of this method, it is analyzed by providing mathematical conditions ensuring expected results. Algorithm modifications have also been proposed to handle unexpected cases. Finally, experimental results based on two typical types of images have been reported. These results will contribute to consolidating both the theory and the practice of the QSF method.

Acknowledgments

This work has been supported by the Bretagne region, the UGE-Cosys department, and the DESDEMONA project of the EU Research Fund for Coal and Steel 2017 programme.

1. Xiong, B.; Zhang, Q.; Baltazart, V. Analysis of Quadratic Surface Fitting for Subpixel Motion Extraction from Video Images. EUSIPCO 2020 - 28th European Signal Processing Conference; , 2020; pp. 695–699. https://hal.inria.fr/hal-03129172/file/main.pdf.
2. Alba, A.; Vigueras-Gomez, J.F.; Arce-Santana, E.R.; Aguilar-Ponce, R.M. Phase correlation with sub-pixel accuracy: A comparative study in 1D and 2D. *Computer Vision and Image Understanding* **2015**, *137*, 76 – 87. doi:https://doi.org/10.1016/j.cviu.2015.03.011.
3. Tong, X.; Ye, Z.; Xu, Y.; Gao, S.; Xie, H.; Du, Q.; Liu, S.; Xu, X.; Liu, S.; Luan, K.; Stilla, U. Image Registration With Fourier-Based Image Correlation: A Comprehensive Review of

	Total trials	Absence of maximum	Maximum outside one-pixel vicinity	$\max \lvert x^* \rvert$	$\max \lvert y^* \rvert$
Moon	10938	23 (0.21%)	524 (4.79%)	35.9474	52.0165
DIC Challenge	114444	1 (8.74 ppm)	6 (52.4 ppm)	4.6726	4.7301

Table 1: Original QSF method failure occurrences in the Moon and DIC Challenge examples. Total trials: number of total trials of the QSF method. Absence of maximum: number of trials yielding a saddle point (and in percentage of total trials). Maximum outside the one-pixel vicinity: number of trials yielding a maximum of the fitted surface located at (x^*, y^*) with $|x^*| > 1$ or $|y^*| > 1$ (and ppm, part per million). $\max |x^*|$: the maximum value of $|x^*|$ among all the trials. $\max |y^*|$: the maximum value of $|y^*|$ among all the trials.

Developments and Applications. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **2019**, *PP*. doi:10.1109/JSTARS.2019.2937690.
4. Yoshikawa, H.; Yamamoto, T.; Tanaka, T.; ichi Kawabata, K.; Yoshizawa, S.; ichiro Umemura, S. Ultrasound Sub-pixel Motion-tracking Method with Out-of-plane Motion Detection for Precise Vascular Imaging. *Ultrasound in Medicine & Biology* **2020**, *46*, 782 – 795. doi: https://doi.org/10.1016/j.ultrasmedbio.2019.11.005.
5. Helmi, A.; Fkirin, M.A.; Badwai, S. Subpixel Accuracy Analysis of Phase Correlation Shift Measurement Methods Applied to Satellite Imagery. *International Journal of Advanced Computer Science and Applications* **2012**, *3*. doi:10.14569/IJACSA.2012.031233.
6. Feng, D.; Feng, M.Q. Computer vision for SHM of civil infrastructure: From dynamic response measurement to damage detection – A review. *Engineering Structures* **2018**, *156*, 105 – 117. doi:https://doi.org/10.1016/j.engstruct.2017.11.018.
7. Guizar-Sicairos, M.; Thurman, S.T.; Fienup, J.R. Efficient subpixel image registration algorithms. *Opt. Lett.* **2008**, *33*, 156–158. doi:10.1364/OL.33.000156.
8. Semin, B.; Auradou, H.; François, M. Accurate measurement of curvilinear shapes by Virtual Image Correlation. *The European Physical Journal - Applied Physics* **2011**, *56*, 10701. doi: 10.1051/epjap/2011110275.
9. Spencer, B.F.; Hoskere, V.; Narazaki, Y. Advances in Computer Vision-Based Civil Infrastructure Inspection and Monitoring. *Engineering* **2019**, *5*, 199–222.
10. Zhang, D.; Guo, J.; Lei, X.; Zhu, C. A High-Speed Vision-Based Sensor for Dynamic Vibration Analysis Using Fast Motion Extraction Algorithms. *Sensors* **2016**, *16*, 572. doi: 10.3390/s16040572.
11. Hung, P.C.; Voloshin, A.S. In-plane strain measurement by digital image correlation. *Journal of the Brazilian Society of Mechanical Sciences and Engineering* **2003**, *25*, 215–221.
12. Pan, B.; Xie, H.m.; Xu, B.q.; Dai, F.l. Performance of sub-pixel registration algorithms in digital image correlation. *Measurement Science and Technology* **2006**, *17*, 1615–1621. doi: 10.1088/0957-0233/17/6/045.
13. Pallotta, L.; Giunta, G.; Clemente, C. Subpixel SAR Image Registration Through Parabolic Interpolation of the 2-D Cross Correlation. *IEEE Transactions on Geoscience and Remote Sensing* **2020**, *58*, 4132–4144.
14. Tomás, M.B.; Ferrer, B.; Mas, D. Influence of Neighborhood Size and Cross-Correlation Peak-Fitting Method on Location Accuracy. *Sensors* **2020**, *20*.
15. Lei, X.; Jin, Y.; Guo, J.; Zhu, C. Vibration extraction based on fast NCC algorithm and high-speed camera. *Appl. Opt.* **2015**, *54*, 8198–8206. doi:10.1364/AO.54.008198.
16. Xiong, B.; Zhang, Q.; Baltazart, V. Matlab code of QSF counterexamples. Download from http://people.rennes.inria.fr/Qinghua.Zhang/QSFexample.zip or from https://sii.univ-gustave-eiffel.fr/fileadmin/redaction/SII/QSFexample.zip.
17. O'Meara, O.T. *Introduction to Quadratic Forms*; Springer-Verlag: Berlin Heidelberg, 1973.
18. Meyer, C.D. *Matrix analysis and applied linear algebra*; SIAM: Philadelphia, 2010.
19. Society for Experimental Mechanics. DIC Challenge. Previous DIC Challenge 1.0 Data, Sample3.zip. https://idics.org/challenge/.
20. Reu, P.; Toussaint, E.; Jones, E.; Bruck, H.; Iadicola, M.; Balcaen, R.; Turner, D.; Siebert, T.; Lava, P.; Simonsen, M. DIC Challenge: Developing Images and Guidelines for Evaluating Accuracy and Resolution of 2D Analyses. *Experimental Mechanics* **2017**, *58*. doi: 10.1007/s11340-017-0349-0.