Phase Amplified Correlation for Improved Sub-pixel Motion Estimation

Dimitrios Konstantinidis, Tania Stathaki, *Member, IEEE,* and Vasileios Argyriou, *Member, IEEE,*

***Abstract*—Phase correlation (PC) is widely employed by several sub-pixel motion estimation techniques in an attempt to accurately and robustly detect the displacement between two images. To achieve sub-pixel accuracy, these techniques employ interpolation methods and function-fitting approaches on the cross-correlation function derived from the PC core. However, such motion estimation techniques still present a lower bound of accuracy that cannot be overcome. To allow room for further improvements, we propose in this paper the enhancement of the sub-pixel accuracy of motion estimation techniques by employing a completely different approach: the concept of motion magnification. To this end, we propose the novel phase amplified correlation (PAC) that integrates motion magnification between two compared images inside the phase correlation part of frequency-based motion estimation algorithms and thus directly substitutes the PC core. The experimentation on magnetic resonance (MR) images and real video sequences demonstrates the ability of the proposed PAC core to make subtle motions highly distinguishable and improve the sub-pixel accuracy of frequency-based motion estimation techniques.**

***Index Terms*—Phase correlation, motion magnification, image registration, sub-pixel motion estimation**

I. Introduction

IMAGE registration concerns the task of estimating the motion between two images that are related by a geometrical transformation and is fundamental for several computer vision and video processing applications. Motion compensated prediction is widely employed for noise reduction, video compression, image super-resolution and medical image registration. In video processing, especially, motion estimation is performed on block-based partitions of input frames. One of the most successful techniques for dealing with motion estimation is phase correlation that operates in the frequency domain [1], [2]. Phase correlation is based on the shift property of the Fourier transform to accurately estimate the displacement between two compared images. Apart from the ability to accurately localize peaks that correspond to spatial displacements, PC enjoys further significant properties, such as computational speed, robustness to uniform illumination variations and insensitivity to spectral energy changes.

A significant drawback of the basic PC core, when implemented in the discrete time domain, is the fact that the computed displacements have integer accuracy as the coordinates of the maximum of the discrete cross-correlation function (i.e. inverse Fourier transform of the cross-power spectrum) will be a rounded version of the components of the true displacement vector. However, the ability to provide sub-pixel accuracies is critical for the performance of image registration techniques that are based on the PC core [3]. Sub-pixel accuracy is mainly achieved through the use of interpolation techniques that fit an analytical function (e.g. a polynomial) to the neighborhood of the maximum of the discrete cross-correlation function [4], [5].

However, since the smallest unit of counting image dimensions is the pixel, achieving the desirable sub-pixel accuracy is sometimes not feasible. To circumvent this problem, we propose a general mathematical framework to improve the sub-pixel accuracy of frequency-based motion estimation algorithms, inspired by the work in [6]. In that work, the authors proposed the concept of motion magnification, so that invisible by the human eye motions become visible in videos. Similarly, we propose in this work the magnification of motion between two compared images. By magnifying motion, we expect that subtle information about the displacement between two images becomes well observed and computed, leading to highly accurate cross-correlation peak extraction and thus improved sub-pixel motion estimation. We integrate motion magnification in the computation of phase correlation, thus bypassing the need for initially applying motion magnification between two compared images and then employing phase correlation for the computation of sub-pixel displacement. Furthermore, we propose a phase blurring technique so that the noise present in the phase of the frequency signals is diminished.

Experiments with MR images and real video sequences show that our proposed methodology consistently improves the performance of all tested motion estimation techniques that are based on the PC core. The main contributions of this work are listed below:

- We integrate the notion of motion magnification in the phase correlation procedure so that subtle image displacements become accurately and robustly estimated.
- The proposed PAC core can be integrated in any frequency-based motion estimation technique, thus taking advantage of interpolation methods and properties of the cross-power spectrum.
- We propose an amplitude-weighted blurring of the phase difference to circumvent problems with noise embedded in the phase signal.

This paper is organized as follows. In Section 2, we conduct a review of state-of-the-art sub-pixel motion estimation meth-

D. Konstantinidis and T. Stathaki are with the Department of Electrical Engineering, Imperial College London, London, SW7 2AZ, United Kingdom e-mail: d.konstantinidis12@imperial.ac.uk, t.stathaki@imperial.ac.uk.

V. Argyriou is with the Department of Computer Sciences and Mathematics, Kingston University, Surrey, KT1 2EE, United Kingdom e-mail: Vasileios.Argyriou@kingston.ac.uk.

ods that employ phase correlation. In Section 3, we discuss and analyze the mathematical framework behind motion magnification and introduce our proposed PAC core. In Section 4, we present experimental results to justify the contributions of our work, while in Section 5 we draw conclusions arising from this work.

II. Related Work

In this section, a brief review of state-of-the-art frequency-based sub-pixel image registration algorithms is presented. Furthermore, motion magnification techniques, along with the way they can be exploited for accurate sub-pixel image registration, are described. Girod was the first to theoretically and experimentally analyze the significant performance improvement of sub-pixel motion estimation for image registration [3]. Since then, a lot of studies have been made in the direction of achieving robust and accurate sub-pixel motion estimation.

Foroosh et al. in [7] came up with the observation that images with sub-pixel shifts originate from up-sampled replicas displaced by integer shifts. The authors noticed that the cross-power spectrum of downsampled images does not contain only a single peak, but rather several coherent peaks that are adjacent to each other. These peaks correspond to the polyphase transform of a filtered unit impulse centered at the point of registration. The derived filter has a rectangular frequency response and it is a 2D Dirichlet kernel that can be closely approximated by a 2D sinc function, leading to the development of a closed-form solution for the sub-pixel translation estimation. Zhou and Yu employed the same observation that the signal power of downsampled images concentrate in a few coherent peaks to develop a real-time PC method that can be applied to high-resolution images [8]. Balci and Foroosh in [9] established the exact relationship between the continuous and the discrete phase difference of two shifted images and showed that the discrete phase difference is a 2D sawtooth signal. Sub-pixel motion estimation can then be computed by counting the number of cycles of the phase difference matrix along each frequency axis and robustly estimating the non-integer fraction of the last cycle.

On the other hand, Ren et al. in [10] proposed the decomposition of the 2D translation to two 1D motions. As a result, only 1D Fourier transform is required to estimate the corresponding motion. The first two highest peaks from 1D correlation are linearly interpolated for sub-pixel accuracy. Takita et al. in [11] computed an analytical model that describes the peak of the cross-power spectrum of two slightly shifted images and how to fit this model to the data. However, their method is limited to detect only small shifts. In the case that the shift is large with respect to the image size, the pixel-level shift should be initially determined before sub-images are extracted from the original images so that the translation of these sub-images is small enough for the proposed model fitting technique.

Stone et al. in [12] was one of the first to investigate the effect of noise and aliasing in the accuracy of sub-pixel motion estimation methods. They proposed the detection of the Fourier components that are unreliable estimators of shift due to aliasing (i.e. components with small spectral magnitude) and they removed them from the shift-estimate computation. The masking out of the contributions from these unreliable spectral components, regardless of whether they occur at low or high frequencies, leads to improvements in the accuracy of sub-pixel motion estimation. The authors in [13] proposed an extension of the previous method for the estimation of planar motion. More specifically, they applied low-pass filtering on images prior to the computation of their Fourier transform. Then, the rotation was estimated from the correlation of the amplitudes of the Fourier transform of images, while the shift was calculated based on the slope of the phase difference of the images.

Hoge in [14] observed that a "noise-free" cross-power spectrum is rank-one. Thus, he proposed recasting the problem of sub-pixel motion estimation as the problem of finding the rank-one approximation of the computed cross-power spectrum. To achieve this, he employed singular value decomposition and identified the left and right dominant singular vectors that corresponded to the vertical and horizontal motion shifts respectively. The authors in [15] proposed an extension to [14] by introducing a masking operator that projects the cross-power spectrum into the space of correlation functions that result from a certain range of translations, thus attenuating the noise associated with the estimation of the phase-shifts by linear regression, especially when that noise is additive white Gaussian noise. Tong et al. in [16] proposed a more robust extension to the Hoge algorithm by replacing the least-square line fitting with the Random Sample Consensus (RANSAC) algorithm that better handles the problem of outliers. Dong et al. in [17] proposed a method for computing the rank-one approximation of the cross-power spectrum by assuming its noise can be represented by a mixture of Gaussian distributions and achieved accurate image registration results.

The effect of noise on image registration was also investigated in a series of works dealing with blur-invariant phase correlation. More specifically, Ojanvisu and Heikkila in [18] proved that by taking any even power (i.e., $2n$) of the normalized Fourier transform of an image, they can achieve invariance to centrally symmetric blur, such as motion or out-of-focus blur, as this type of noise has constant phase. The even power of the normalized Fourier transform is translated to the multiplication of the phase shift between the original and the blurred images by the factor of $2n$. Later, Pedone et al. generalized the blur-invariant phase correlation assuming the blurring function exhibits rotational symmetry [19] and both rotational and axial symmetries [20]. Interestingly, the work in [18] is closely related to ours as the authors in that paper employed phase shift amplification by even powers to improve robustness to centrally symmetric blur, while we study in this work a more general phase shift amplification in order to improve the sub-pixel accuracy of frequency-based motion estimation algorithms.

Several methodologies that incorporate robust features extracted from the compared images were also proposed. More specifically, Zhonke et al. in [21] proposed the computation of rotation between the compared images by employing Hough transform in the computed image edges. After the images are de-rotated, the translation is identified using phase correlation. Maik et al. in [22] proposed the detection of Harris corner features in order to calculate the global motion between images. After the geometric transformation of images, local motion is estimated using phase correlation in image blocks. Cheng and Menq, on the other hand, proposed a real-time image registration algorithm by employing two continuous spatial variables to measure the shift of the image and build the continuously shifting image model [23]. As the variables of the model are continuous in spatial domain, pixel-level image registration is unnecessary, thus achieving real-time tracking of the target. Park et al. in [24] proposed the identification and matching of a number of affine-invariant points in the spectrum of the source and target images and the computation of the parameters of affine transform using spectral alignment.

Argyriou and Vlachos in [25] proposed a different solution to the sub-pixel motion estimation problem by employing gradient information. The sub-pixel shifts were computed by maximizing the cross-correlation between the spatial gradient information derived from the pair of images. The proposed gradient correlation (GC) algorithm is proved to be very accurate, robust to the effect of noise and outperforms frequency-domain motion estimation methods. Later, the same authors extended the GC algorithm by computing scaling and rotation using log-polar Fourier representations of the complex gray-level edge maps of images prior to the sub-pixel motion estimation using normalized gradient correlation [26]. Moreover, the inference of the sub-pixel shifts by employing the left and right dominant singular vectors of the 2D GC matrix prior to their modeling using a generic kernel that can adapt its shape to fit the available correlation samples was proposed in [27]. A quad-tree GC algorithm that replaces block-based motion estimation with an iterative decomposition of a block to four quadrants based on the resulting motion compensated prediction error was also proposed in [28].

Recently, Argyriou and Tzimiropoulos in [29] proposed HOG-PC, a frequency-domain image registration technique based on histograms of oriented gradients (HOG). The proposed HOG representation is very dense since a descriptor is computed for each pixel and it can thus be considered as a multi-channel block representation. The authors claim that HOG-PC retains the orientation information and it is able to cope with non-overlapping regions, small deformations and noise, while achieving state-of-the-art sub-pixel motion estimation results, especially in small-sized blocks. Ye et al. in [30] took advantage of the properties of the phase of the Fourier transform and proposed a new descriptor for image matching and registration, called Histogram of Oriented Phase Congruency. On the other hand, Li, motivated by the observation that the classical PC method and the Lucas-Kanade algorithm exhibit strong complementary property between convergence range and sub-pixel accuracy, proposed a two-stage coarse-to-fine sub-pixel image registration framework that accurately computes rotation, scale and translation [31].

Most of the previously presented sub-pixel motion estimation methods concentrate on either proposing more powerful interpolation techniques [11], [27], [32] or searching and identifying interesting properties of the cross-power spectrum that can assist in the improvement of the image registration results [7], [9], [14], [33]. Unfortunately, these methods have a lower bound of sub-pixel accuracy that cannot be overcome [34]. This work is the first that proposes a different approach to the problem of sub-pixel motion estimation by employing motion magnification. Motion magnification was initially employed in [35], when the authors proposed the magnification of a band-pass filtered video in order to reveal subtle periodical changes that were not observable by the human eye. The video magnification approach was later improved by Wadhwa et al. in [6], when the authors proposed the transformation of video frames in the Fourier domain, their decomposition into image sub-bands by employing complex steerable pyramids and the phase amplification of the image sub-bands in order to magnify local motions. The authors claimed that their approach is more robust to noise and can lead to larger amplification factors than the method of [35].

Inspired by the concept of motion magnification, we propose a novel approach that improves the accuracy of sub-pixel motion estimation techniques. Our motivation lies in the belief that amplifying sub-pixel shifts can lead to a better estimation of their true values. Since all interpolation techniques are applied on a specific cross-correlation function, there is a bound on the accuracy of the detected peak based on how accurately the actual cross-correlation function can model the true peak. In other words, the cross-correlation function dictates the lower bound of accuracy that interpolation techniques can achieve since if the cross-correlation function cannot adequately describe the shift, no interpolation technique can improve the shift estimation. Our approach manages to overcome this problem by computing a new cross-correlation function based on magnified motions. We show in this work that the same interpolation techniques can achieve better estimations of the true motions on the new cross-correlation function rather than on the initial one, which is an indication of the improved shift modeling abilities of the new cross-correlation function. In this way, we bypass the problem of defining powerful interpolation techniques or searching for cross-power spectrum properties that we can take advantage of and concentrate on how robustly the new cross-correlation function that models the shift between two compared images can be computed.

III. Methodology

In this section, we initially present the well-known phase correlation core for the estimation of displacement between images and then we discuss and analyze the notion of motion magnification. Finally, we introduce our proposed methodology for merging phase correlation and motion magnification into a novel framework.

A. Phase correlation

Given an image I_A and its temporally subsequent version I_B, whose relationship with I_A is described by a relative translation or shift (i.e., $I_B = I_A(x + \Delta x, y + \Delta y)$), the PC core attempts to accurately estimate the shift between the two images. To achieve this, the PC core initially computes the Fourier transforms F_A and F_B of images I_A and I_B respectively, where it holds that $F_B = F_A e^{i(\Delta xu+\Delta yv)}$. Afterwards, the cross-power spectrum of the compared images is computed as shown in Eq. (1) and is equal to the element-wise product of the Fourier transform of the first image and the complex conjugate of the Fourier transform of the second image F_B^*, normalized element-wise to ensure that all values are in the range $[0, 1]$.

$$R = \frac{F_A \circ F_B^*}{|F_A \circ F_B^*|} = e^{-i(\Delta xu+\Delta yv)} \tag{1}$$

The inverse Fourier transform of the cross-power spectrum gives the normalized cross-correlation $r = \mathcal{F}^{-1}\{R\} = \delta(x - \Delta x, y - \Delta y)$. Ideally, the normalized cross-correlation is a Dirac function with the location of its peak defining the shift between the compared images. As a result, the shift between the compared images can be estimated by

$$(\Delta x, \Delta y) = \underset{(x,y)}{\arg\max}\{r\} \tag{2}$$

However, noise and illumination variations can affect both the magnitude and the location of the peak, thus posing difficulties in its accurate and robust extraction. Furthermore, since the normalized cross-correlation is a discrete function, achieving highly accurate sub-pixel peak measurement can be very difficult and commonly requires interpolation methods applied in the neighborhood of the detected maximum peak.

B. Motion magnification

Given an image $I_A(x, y)$, the same image translated spatially by a vector $(\Delta x, \Delta y)$ can be described as $I_B = I_A(x + \Delta x, y + \Delta y)$. The goal of motion magnification is the synthesis of a new image,

$$\begin{aligned} I_B^{'} &= I_B(x + m\Delta x, y + m\Delta y) \\ &= I_A(x + (1 + m)\Delta x, y + (1 + m)\Delta y) \end{aligned} \tag{3}$$

which has its spatial translation magnified by a factor m. Applying the Fourier transform to the previously defined images leads the initial or reference image I_A to be equal to $F_A(u, v)$ and the image I_B to be equal to $F_B = F_A(u, v)e^{i(\Delta xu+\Delta yv)}$. This means that the Fourier transform of the motion magnified image $I_B^{'}$ is equal to

$$\begin{aligned} \mathcal{F}\{I_B^{'}\} &= F_B(u, v)e^{im(\Delta xu+\Delta yv)} \\ &= F_A(u, v)e^{i(1+m)(\Delta xu+\Delta yv)} \end{aligned} \tag{4}$$

This is in par with the basic Fourier transform property that states that a shift in the time domain corresponds to a phase shift in the frequency domain. Thus, motion magnification in the time domain can be achieved by phase amplification in the frequency domain. The phase amplification factor m controls the magnitude of the motion magnification we want to achieve.

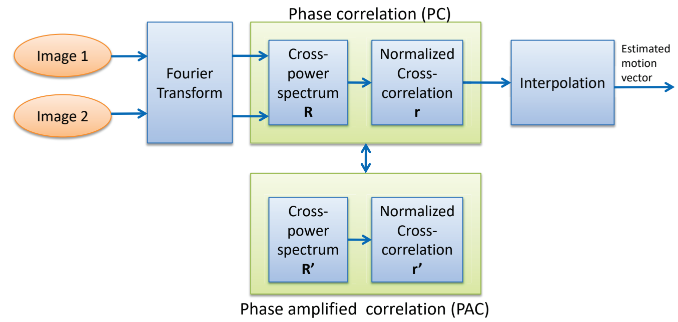


Fig. 1. A diagram of the processing stages of a general frequency-based motion estimation algorithm. The proposed PAC core directly replaces PC in such an algorithm by computing a new cross-power spectrum and cross-correlation function, from which the peak is extracted.

C. Phase amplified correlation

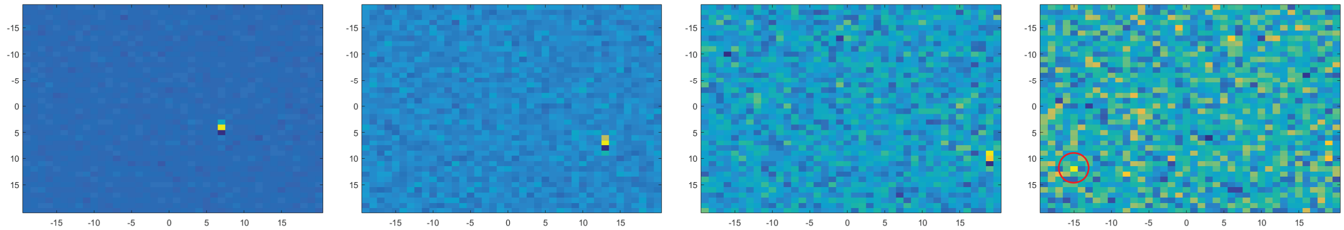
In this paper, we propose the novel phase amplified correlation core that allows the magnification of motion between two compared images and can substitute the PC core in the task of sub-pixel image registration. An outline of a general frequency-based motion estimation algorithm and the processing stage in which, our proposed PAC core is introduced, is presented in Fig. 1.

From Fig. 1, it can be observed that the PAC core replaces PC by computing a new cross-power spectrum and normalized cross-correlation function. The reason why both spectrum and cross-correlation are computed lies in the fact that there are motion estimation algorithms that process the cross-power spectrum directly, such as the Hoge algorithm [14] and we want to enable such algorithms to take advantage of the proposed PAC core as well. Furthermore, we influence interpolation techniques indirectly by feeding them with a new cross-correlation function that better models the true motion between the compared images. As a result, our method works complementary with interpolation techniques, assisting them towards the extraction of a highly accurate and reliable cross-correlation peak and therefore the computation of a motion estimation vector with improved sub-pixel accuracy.

Given the definition of motion magnification and phase correlation, the proposed PAC core computes a new cross-power spectrum between the original image I_A and the motion magnified translated one $I_B^{'}$ by introducing Eq. (4) to Eq. (1) and taking into consideration the complex exponential identity $|e^{ia}| = 1$. As a result, we obtain a new cross-power spectrum

$$\begin{aligned} R^{'} &= \frac{F_A \circ F_B^{*'}}{|F_A \circ F_B^{*'}|} = \frac{F_A \circ F_B^* e^{-im(\Delta xu+\Delta yv)}}{|F_A \circ F_B^*||e^{-im(\Delta xu+\Delta yv)}|} \\ &= e^{-i(\Delta xu+\Delta yv)}e^{-im(\Delta xu+\Delta yv)} = e^{-i(1+m)(\Delta xu+\Delta yv)} \end{aligned} \tag{5}$$

The shift between the original and motion magnified translated images is given by the location of the peak in the new normalized cross-correlation $r^{'}$, which is the result of the inverse Fourier transform of the new cross-power spectrum



(a) (b) (c) (d)

Fig. 2. Effect of the phase amplification factor m on the cross-correlation peak. The mean squared error (MSE) between the ground truth motion vector, which is equal to [-2.69,-6.04] and the estimated motion vector (MV), along with the magnitude of the peak are presented. (a) m=0, MV=[-2.88,-6.01], MSE=0.036, peak=0.729, (b) m=1, MV=[-2.86,-6.03], MSE=0.03, peak=0.389, (c) m=2, MV=[-2.77,-6.01], MSE=0.008, peak=0.168 and (d) m=3, MV=[-2.76,4.03], MSE=101.46, peak=0.102. A red circle is used to denote the location of peak in sub-figure (d).

$$\begin{aligned}(\Delta x^{'}, \Delta y^{'}) &= \arg\max_{(x,y)}\{r^{'}\} = \arg\max_{(x,y)}\{\mathcal{F}^{-1}\{R^{'}\}\} \\ &= \arg\max_{(x,y)}\{\delta(x-(1+m)\Delta x, y-(1+m)\Delta y)\}\end{aligned} \tag{6}$$

However, the computed translation $(\Delta x^{'}, \Delta y^{'})$ does not correspond to the true translation between the original and translated images since motion magnification is applied. The true translation is computed as follows in order to eliminate the effect of the phase amplification factor m

$$(\Delta x, \Delta y) = \left(\frac{\Delta x^{'}}{1+m}, \frac{\Delta y^{'}}{1+m}\right) \tag{7}$$

Eq. (7) identifies the cross-correlation peak when sub-pixel accuracies are considered. In the case that we refer to pixel-level accuracies, the rounded version of Eq. (7) holds. This means that in pixel-level accuracies the true translation Δx is computed as $\Delta x = \|\Delta x^{'}/(1+m)\|$ and similarly for Δy. One may notice that when m is equal to 0, the proposed PAC core boils down to the classical PC core. As a result, the PAC core can be considered as a general PC core for $m \neq 0$. A significant advantage of the proposed PAC core is that it can directly substitute the PC core inside any frequency-based sub-pixel motion estimation technique without the need of computing the motion magnified image. The one and only parameter that affects the performance of the PAC core is the phase amplification factor m. Larger phase amplification factors lead to larger motion magnifications and thus larger shifts of the cross-correlation peak, as it is shown in Fig. 2.

The PAC core allows subtle motions to become more easily recognizable and more accurately extracted when magnified. This effect can be observed in Fig 2 since for larger phase amplification factors m, the peak is slightly more spread to its neighborhood rather than concentrated on a single pixel and a drop in the mean squared error (MSE) between the computed and true motion vectors is observed. The drop in MSE means that the increased spread of the cross-correlation peak assists the tested sub-pixel motion estimation algorithms to more accurately identify the true displacement between the images and thus the new cross-correlation function, proposed by the PAC core, can more reliably model the true motion between the compared images. As a result, we deduce that a motion vector can be more robustly detected and evaluated on a discrete cross-correlation function on its magnified rather than on its original form. Finally, the improved sub-pixel accuracy means that we manage to indirectly overcome the lower bound of accuracy that current motion estimation algorithms face not by proposing more powerful interpolation techniques but by introducing a new and more reliable cross-correlation function to existing interpolation techniques.

D. Analysis of phase amplification factor

Unfortunately, there are limitations on the values the phase amplification factor m of the proposed PAC core can get. Firstly, an immoderate shift of the location of the cross-correlation peak can transpose the peak outside of the borders of the image or image block that is used for the motion estimation. This can lead to the peak being circularly wrapped around the image or image block and thus Eq. (7) can no longer be employed for the estimation of the true motion vector as shown in Fig. 2(d), where the motion vector cannot be reliably computed and the MSE is tremendously increased. Secondly, by magnifying the motion between two images, there is the risk of unintentionally magnifying the noise embedded in the phase of the images and noise magnification can severely affect the accuracy of sub-pixel motion estimation techniques. Finally, motion magnification increases the dissimilarity between the two compared images and thus the uncertainty of phase correlation about the fact that each image is a shifted version of the other image. This uncertainty is translated to a drop in the strength of the cross-correlation peak, which in turn means that the difference in value between the peak and the other pixels of the discrete cross-correlation function is getting smaller. A lower cross-correlation peak faces the risk that it may not be detected and recovered among the local maxima of the discrete cross-correlation function (see Fig. 2(d)).

An excessive shift of the cross-correlation peak can be avoided given an indication of the expected magnitude of the true motion vector. Given an expected motion vector $(\Delta x, \Delta y)$, the value of the phase amplification factor m is restricted by the fact that the magnified motion vector cannot overcome the boundaries of the image block, where the PAC core is applied. Given an image block of size $M \times N$, the phase amplification factor is bounded by the following set of inequalities

$$
\begin{aligned}
(1+m)\Delta x &\leq \frac{M}{2} \\
(1+m)\Delta y &\leq \frac{N}{2}
\end{aligned}
\tag{8}
$$

The inequalities of Eq. 8 can also be written in a more compact form as shown below, allowing the direct computation of an upper boundary for the phase amplification factor m.

$$
m \leq min\left(\frac{M}{2\Delta x} - 1, \frac{N}{2\Delta y} - 1\right) \tag{9}
$$

The expected motion vector is of course not known beforehand, however employing the proposed PAC core with $m = 0$ (i.e., PC core) can give an accurate estimate of the expected motion vector before applying larger amplification factors. Another alternative can be a pre-registration of the two compared images in order to restrict the motion within sub-pixel accuracies (i.e., Δx and Δy smaller than 1). An additional limitation of the phase amplification factor resides in the drop in the strength of the cross-correlation peak as the phase amplification factor increases. This drop is attributed to the fact that the shifted image resembles less the original image as their overlap decreases. A large decrease of the overlap between two compared images makes the PAC core less confident about the similarity between the images and therefore their displacement. Given the fact that we want to maintain a significant overlap between two compared images, thus enabling a reliable estimation of their displacement, a new restriction on the value of the phase amplification factor can be set. For an image block of size $M \times N$ and a shift of $(\Delta x, \Delta y)$, the phase amplification factor m is bounded by the following inequality

$$
\begin{gathered}
A_{overlap} \geq a_{th} A_B \Rightarrow \\
(M - (1+m)|\Delta x|)(N - (1+m)|\Delta y|) \geq a_{th} MN
\end{gathered}
\tag{10}
$$

where $A_{overlap}$ and A_B denote the overlapping area between the original and shifted image blocks and the area of the image block respectively. The overlap ratio a_{th} determines how large the overlapping area between the shifted image blocks should be. In order to have a reliable and accurate estimate of the true displacement, the overlap ratio should be at least 0.5. Higher values of the overlap ratio allow even more distinct cross-correlation peaks at the expense of limiting the value of the phase amplification factor, and therefore the sub-pixel accuracy of motion estimation techniques. Eqs. (9) and (10) can be considered as hard and soft constraints respectively on the value of the phase amplification factor of the proposed PAC core. Finally, the noise in the phase of the images can be suppressed by employing phase blurring techniques. In the following section, we propose such a phase blurring technique and incorporate it in the proposed PAC core.

E. Noise handling

An aspect that should be taken into account when magnifying the motion between images is noise. Noise can significantly affect the performance of sub-pixel motion estimation algorithms if magnified. The advantage of applying phase amplification in the frequency domain is that the additive noise present in the images is not magnified, but translated. More specifically, given an image I, a Gaussian distributed additive noise $N(0,\sigma)$ and the linear properties of the Fourier transform, the Fourier transform of the motion magnified image $I^{'} = I + N(0,\sigma)$ is given by

$$
\mathcal{F}\{I^{'}\} = e^{im(\Delta xu + \Delta yv)}\left(\mathcal{F}\{I\} + \mathcal{F}\{N(0,\sigma)\}\right) \tag{11}
$$

Eq. (11) states that a Gaussian distributed additive noise is translated when motion magnification is applied. Although the additive noise embedded in an image is not magnified when phase amplification is applied, this is not the case for the noise that is embedded in the phase of an image. During phase amplification, any noise present in the phase signals is amplified and it can lead to inaccuracies in the sub-pixel motion estimation. Additionally, the phase amplification can cause abrupt changes in the phase difference due to phase wrapping. To alleviate the effect of noise, we apply low-pass filtering as a simple and efficient way to increase the signal-to-noise ratio of a signal. To this end, we propose the blurring of the phase difference by employing a Gaussian kernel. The phase blurring is also based on the amplitude of the cross-power spectrum as the phase signal in regions of low amplitude is not reliable for shift estimation [12]. More specifically, given the phase difference between the compared images $\Delta\phi = \Delta xu + \Delta yv$, the blurred phase difference is given by

$$
\Delta\phi_G = \frac{(\Delta\phi\,|R|) * K_G}{|R| * K_G} \tag{12}
$$

where $|R|$ represents the amplitude of the complex cross-power spectrum and K_G is a Gaussian kernel. The Gaussian blurred phase difference $\Delta\phi_G$ replaces the original phase difference $\Delta\phi = \Delta xu + \Delta yv$ in Eq. (5) before the computation of the shift between the compared images. As we show in the experimental section, we consistently get more accurate and robust to noise results, when the proposed noise handling term is employed.

IV. Experiments

To evaluate and illustrate the efficiency of the proposed PAC core, a comparative study is performed with state-of-the-art sub-pixel motion estimation techniques that employ the PC core. More specifically, the PC core [1] with quadratic (PC+Quad) and Gaussian (PC+Gaussian) interpolation methods [36], Foroosh [7], Ren [10], Xiaohua [16] and HOG-PC [29] motion estimation algorithms are considered. Both images with ground truth motion vectors and video sequences are used for the performance evaluation of the tested sub-pixel motion estimation algorithms.

A. Experimentation with MR images

A set of 5 MR images is available by the authors in [14] for the evaluation of the sub-pixel motion estimation algorithms (see Fig. 3). The MR images depict a grapefruit that was

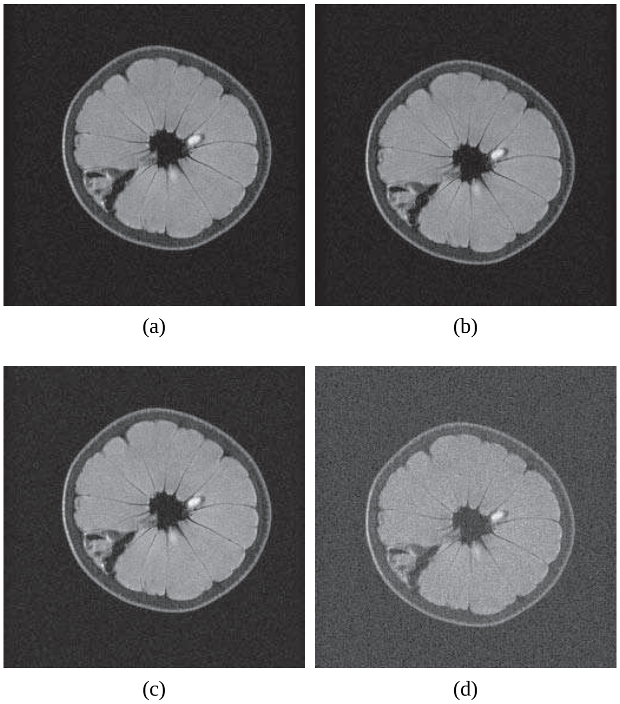


Fig. 3. Two MR images (a),(b) without noise and (c) image (a) with noise level 0.01 and (d) image (b) with noise level 0.05.

acquired using a production quality Fast Spin Echo sequence on a GE (Faireld, CT, USA) Signa Lx 1.5 Tesla MRI scanner. The images have size 256×256 and they cover a 16 cm^2 Field of View (FOV) corresponding to a 0.0625 mm square per pixel. The 5 images depict the fruit at different positions in the FOV and they were acquired by manually moving the scanner table.

Furthermore, we analyze the sub-pixel accuracy of the motion estimation algorithms under the effect of noise. This is achieved by embedding the MR images with additive white Gaussian noise of zero mean and variance in the range $[0.005 - 0.05]$ with a step of 0.005. An example of the effect of noise in the MR images is presented in Fig. 3. The 5 MR images yield 10 possible pairwise registrations and the ground truth translations are available.

The evaluation of the motion estimation algorithms is performed using the metric of mean squared error. Since ground truth is available, Eq. (13) is employed for the computation of MSE of the motion vectors (MSE_{MV}). As it is expected, the most accurate sub-pixel motion estimation algorithm should produce the lowest MSE_{MV} in the MR images. Other than the metric of MSE, we also employ the peak signal-to-noise ratio (PSNR) of the motion compensated prediction error in order to assess the performance of the motion estimation algorithms. According to [14], ground truth measurements can be significantly biased based on the image acquisition methods and thus PSNR can provide more unbiased performance evaluation.

$$MSE_{MV} = \frac{1}{B} \sum_{i=1}^{B} \|\mathbf{u}_i - \mathbf{v}_i\|^2 \tag{13}$$

In Eq. (13), $\mathbf{u}$ and $\mathbf{v}$ represent the ground truth and estimated motion vectors respectively, while B corresponds to the number of blocks that an image is divided before the computation of the local motion vector for each block. For the MR images, the whole image of size 256×256 consists of a single block, but for the video sequences presented in the next section, the video frames are divided in smaller blocks. Finally, for an image I of dimensions $M \times N$ and its motion compensated prediction I_{MC}, PSNR is computed as shown in Eq. (14).

$$\begin{aligned} &\text{PSNR} = 10log(\frac{255^2}{MSE_I}) \\ &,\text{where } MSE_I = \frac{1}{MN} \sum_{x=1}^{M} \sum_{y=1}^{N} (I(x,y) - I_{MC}(x,y))^2 \end{aligned} \tag{14}$$

The MSE_I evaluates the similarity between the original and the motion compensated images and it can be considered as a measure of the visual quality of the motion compensation. The initial experiments examine the effect of the phase amplification and the noise handling terms on the 10 pairwise registrations that the 5 MR images form. The parameter that affects the noise handling term of our proposed PAC core is the Gaussian kernel and we fix it to be of size 5×5 pixels with a standard deviation of 0.4. The reason we choose such values is in order to have a small blurring effect and therefore we do not allow a pixel to be significantly affected by its neighboring pixels (i.e. the sum of the kernel weights that correspond to the neighboring pixels contributes by only 15$\%$ to the final blurred value). Furthermore, Table I summarizes the performance of the PAC core with and without the noise handling term versus the PC core when the quadratic interpolation method is employed. The reported MSE_{MV} and PSNR correspond to the average values over all pairwise image registrations.

To further support the use of the Gaussian noise handling term, we experiment with alternative low-pass filters, such as average and median filters. Fig. 4 presents the average PSNR performance of the HOG-PC algorithm, which is considered the optimal motion estimation algorithm out of those tested in this work on all pairwise registrations of the MR images. The experiments reveal that the Gaussian filter slightly outperforms the other tested filters especially for larger values of the amplification factor, showing its ability to suppress the phase noise more effectively than the other tested filters as the noise increases. As far as the sudden changes in the phase spectrum due to phase wrapping is concerned, we observe that they do not play a significant role to the performance of the PAC core, especially for small and mediocre values of the phase amplification factor since the motion estimation accuracy of the HOG-PC algorithm improves. Moreover, we observe that the use of a median filter that is able to remove the “salt-and-pepper” effect that phase wrapping may cause is not beneficial to the performance of the PAC core.

It can be observed that the PAC core achieves better performance than the PC core for all tested values of the phase amplification factor with respect to the measures of both MSE_{MV} and PSNR. This verifies our initial statement that

TABLE I
PERFORMANCE OF PC AND PAC CORES WITH QUADRATIC INTERPOLATION AND WITH (W/ NH) AND WITHOUT (W/O NH) NOISE HANDLING FOR VARIOUS VALUES OF THE PHASE AMPLIFICATION FACTOR m ON A FEW MR IMAGE PAIRS

Method \ Image Pairs	[1,2]	[1,4]	[2,3]	[2,5]	[3,4]	[3,5]	[4,5]	MSE_{MV}	PSNR (dB)
Ground Truth	(-2.40,4.00)	(-7.20,4.32)	(-2.40,4.00)	(-4.80,8.00)	(-2.40,-3.68)	(-2.40,4.00)	(0.00,7.68)	0.0000	–
PC	(-2.03,4.01)	(-6.73,4.25)	(-2.13,3.98)	(-4.65,8.00)	(-2.25,-3.78)	(-2.27,4.00)	(-0.01,7.78)	0.3935	31.2043
PAC (m=1), w/o NH	(-2.04,4.01)	(-6.60,4.35)	(-2.17,3.98)	(-4.55,8.00)	(-2.38,-3.63)	(-2.38,4.01)	(0.00,7.66)	0.3810	31.4250
PAC (m=1), w/ NH	(-2.04,4.01)	(-6.73,4.35)	(-2.18,3.98)	(-4.55,8.00)	(-2.38,-3.63)	(-2.38,4.02)	(0.00,7.67)	0.3805	31.4287
PAC (m=2), w/o NH	(-2.04,4.02)	(-6.63,4.34)	(-2.23,3.98)	(-4.58,8.00)	(-2.34,-3.66)	(-2.34,4.01)	(0.00,7.69)	0.3606	31.4406
PAC (m=2), w/ NH	(-2.04,4.02)	(-6.63,4.33)	(-2.26,3.99)	(-4.58,8.01)	(-2.33,-3.66)	(-2.35,4.01)	(0.00,7.69)	0.3587	31.4354
PAC (m=3), w/o NH	(-2.04,4.02)	(-6.62,4.34)	(-2.21,3.98)	(-4.56,8.01)	(-2.31,-3.68)	(-2.34,4.01)	(-0.01,7.71)	0.3795	31.4321
PAC (m=3), w/ NH	(-2.05,4.03)	(-6.62,4.35)	(-2.22,3.98)	(-4.56,8.01)	(-2.31,-3.67)	(-2.33,4.01)	(-0.01,7.70)	0.3710	31.4366
PAC (m=4), w/o NH	(-2.04,4.02)	(-6.62,4.36)	(-2.21,3.98)	(-4.58,8.00)	(-2.37,-3.67)	(-2.35,4.03)	(-0.01,7.69)	0.3754	31.4313
PAC (m=4), w/ NH	(-2.05,4.03)	(-6.61,4.34)	(-2.20,3.98)	(-4.57,8.00)	(-2.36,-3.66)	(-2.35,4.03)	(-0.02,7.69)	0.3779	31.4371
PAC (m=5), w/o NH	(-2.04,4.01)	(-6.64,4.33)	(-2.20,3.99)	(-4.56,8.01)	(-2.35,-3.70)	(-2.34,4.03)	(-0.01,7.70)	0.3769	31.4233
PAC (m=5), w/ NH	(-2.03,4.02)	(-6.61,4.37)	(-2.21,3.99)	(-4.59,8.03)	(-2.34,-3.69)	(-2.35,4.02)	(-0.01,7.68)	0.3897	31.4395

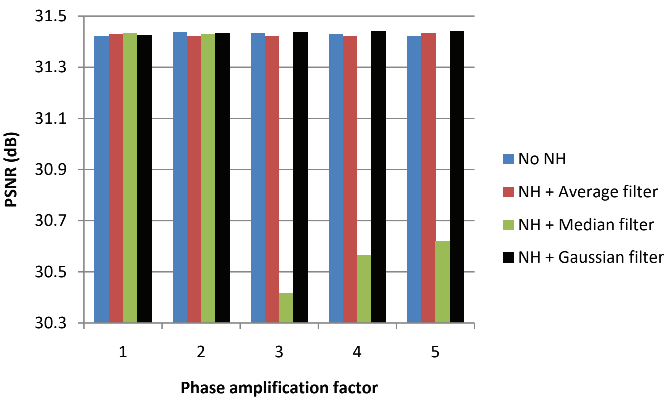


Fig. 4. Average performance of the HOG-PC algorithm on all pairwise registrations of the MR images for various values of the phase amplification factor and for different low-pass filters. NH: Noise handling.

increasing the displacement between two images enables more accurate and robust identification of the actual displacement between the images. However, although the performance of the proposed PAC core improves as the phase amplification factor m increases, a point is reached after which further increase in the value of the phase amplification factor leads to deterioration of the results. The performance degradation of the PAC core for large values of the phase amplification factor m is attributed to not only the magnified phase noise that corrupts the discrete cross-correlation function but also the drop of the magnitude of the cross-correlation peak that inhibits its detection and accurate estimation.

As far as the noise handling term is concerned, we observe that the phase difference blurring has a beneficial effect on the sub-pixel accuracy of the estimated motion vectors. More specifically, the metric of PSNR increases especially for larger values of the phase amplification factor, since in these cases phase noise is present in larger quantities. Furthermore, the metric of mean squared error (MSE_{MV}) drops for small values of the phase amplification factor, although this is not the case for larger values of the phase amplification factor m. The PAC core with the noise handling term achieves the best performance, regarding the metric of MSE_{MV}, for m=2 with a drop by 0.5% with respect to the same method without the noise handling term and an even larger drop of 8.8% with respect to the PC core. As a result, the proposed noise handling term manages to successfully suppress the noise that corrupts the phase difference between two compared images.

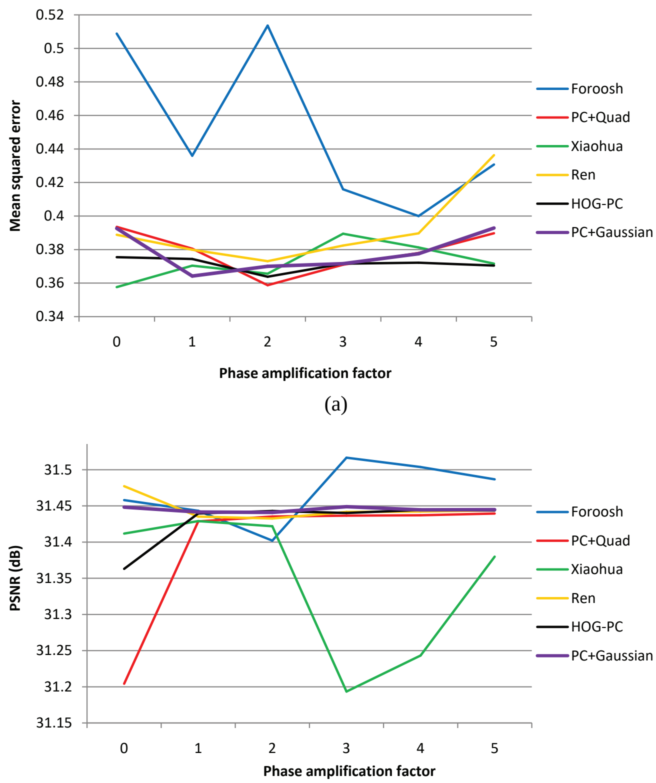
Fig. 5 presents the average values of MSE_{MV} and PSNR that the tested sub-pixel image registration algorithms achieve over all pairwise registrations among the MR images and for various values of the phase amplification factor m. It can be noticed that for $m = 2$ all tested algorithms achieve the smallest value of MSE_{MV} and the highest PSNR. An exception is the algorithm of Foroosh that achieves slightly worse results than the case of $m = 0$. However, the algorithm of Foroosh achieves the optimal performance for $m = 3$, showing that the use of the PAC core, instead of the standard PC core is beneficial for this algorithm too.

Fig 6 presents the average performance of the tested methodologies for various values of the phase amplification factor m in the case that the MR images are embedded with noise. As it is expected, the average performance of the tested sub-pixel image registration algorithms deteriorates as the noise embedded on the images increases. This happens because noise corrupts the images and poses challenges to their correct matching by the PC or PAC core. However, the performance degradation is significantly suppressed with the use of the proposed PAC core with the noise handling term and large phase amplification factors. More specifically, for the highest level of noise, the error MSE_{MV} drops by over 168% (from 10.727 to 4.001) and the PSNR is increased by 14.2% (from 19.525 to 22.299 dB) in the cases that the PC core ($m = 0$) and the PAC core with $m = 5$ are employed respectively. These results reveal one of the most significant traits of the PAC core, which is its ability to magnify the relative motion between two compared images, while it only translates the additive noise of the images. Therefore, the motion magnification enables the PAC core to more reliably and robustly estimate the shift between two images even in the presence of noise, which leads to improved robustness of all tested motion estimation algorithms when the proposed PAC core with increased values of the amplification factor are employed.

Finally, Table II compares the performance of the three best frequency-based motion estimation algorithms among those

TABLE II
PERFORMANCE OF BEST MOTION ESTIMATION ALGORITHMS WITH PC AND PAC CORES FOR ALL PAIRWISE COMBINATIONS OF MR IMAGES

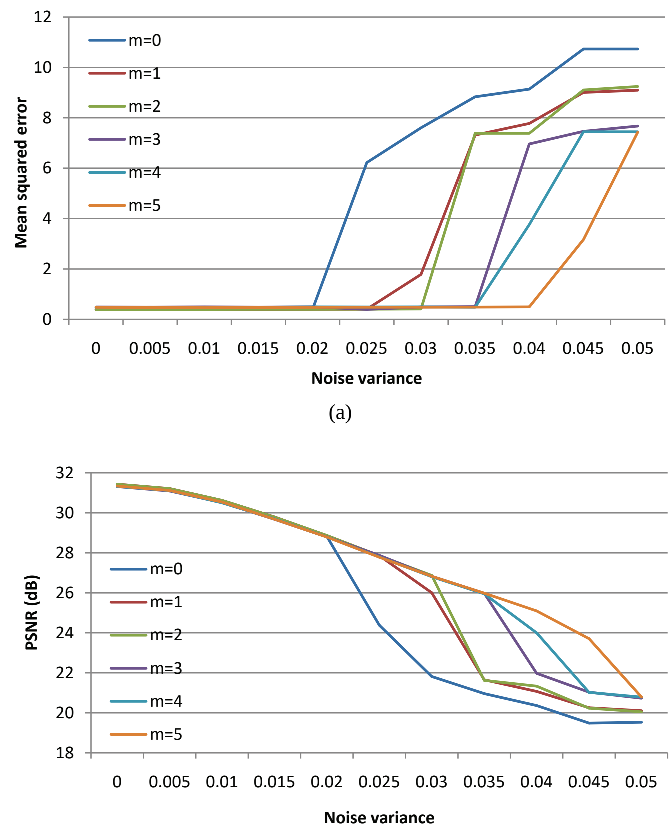
Image Pairs \ Method	Ground truth	Xiaohua [16]		Ren [10]		HOG-PC [29]	
		m=0	m=2	m=0	m=2	m=0	m=2
[1,2]	(-2.40,4,00)	(-2.03,4.00)	(-2.06,4.01)	(-2.09,4.02)	(-2.06,4.03)	(-2.04,4.02)	(-2.05,4.03)
[1,3]	(-4.80,8.00)	(-4.25,7.99)	(-4.27,8.00)	(-4.34,8.01)	(-4.26,8.03)	(-4.23,8.01)	(-4.28,8.01)
[1,4]	(-7.20,4.32)	(-6.64,4.33)	(-6.62,4.32)	(-6.58,4.38)	(-6.61,4.34)	(-6.66,4.30)	(-6.62,4.34)
[1,5]	(-7.20,12.00)	(-6.67,12.02)	(-6.64,12.00)	(-6.59,12.08)	(-6.63,12.03)	(-6.68,12.03)	(-6.63,12.03)
[2,3]	(-2.40,4.00)	(-2.20,3.99)	(-2.21,3.99)	(-2.27,3.96)	(-2.17,3.99)	(-2.17,3.98)	(-2.22,3.98)
[2,4]	(-4.80,0.32)	(-4.57,0.30)	(-4.53,0.28)	(-4.54,0.36)	(-4.56,0.31)	(-4.59,0.27)	(-4.56,0.31)
[2,5]	(-4.80,8.00)	(-4.59,8.00)	(-4.55,7.99)	(-4.54,8.01)	(-4.57,8.01)	(-4.60,8.00)	(-4.58,8.00)
[3,4]	(-2.40,-3.68)	(-2.35,-3.69)	(-2.35,-3.68)	(-2.40,-3.65)	(-2.33,-3.66)	(-2.29,-3.73)	(-2.34,-3.67)
[3,5]	(-2.40,4.00)	(-2.39,3.99)	(-2.34,3.99)	(-2.41,4.00)	(-2.36,4.01)	(-2.31,4.01)	(2.35,4.02)
[4,5]	(0.00,7.68)	(-0.04,7.71)	(-0.01,7.70)	(-0.02,7.64)	(0.00,7.70)	(-0.01,7.74)	(-0.01,7.69)
MSE_{MV}	0	0.358	0.366	0.389	0.373	0.375	0.364
PSNR (dB)	–	31.412	31.422	31.477	31.433	31.363	31.444



(a)

(b)

Fig. 5. Average performance of the tested sub-pixel image registration algorithms on all pairwise registrations of the MR images for various values of the phase amplification factor.



(a)

(b)

Fig. 6. Average performance of the tested sub-pixel image registration algorithms on all pairwise registrations of the noisy MR images for various values of the phase amplification factor.

tested as far as the metrics of MSE_{MV} and PSNR are concerned and for all pairwise MR image registrations. It can be concluded that the proposed PAC core improves either the metric of MSE_{MV} or the metric of PSNR for all three best motion estimation algorithms, thus providing a beneficial effect on their performance. Moreover, the introduction of the novel PAC core to the HOG-PC algorithm, which is considered one of the most accurate state-of-the-art motion estimation techniques according to [29], improves both the metrics of MSE_{MV} by 2.9% and PSNR by almost 0.26% with respect to the same algorithm employing the standard PC core.

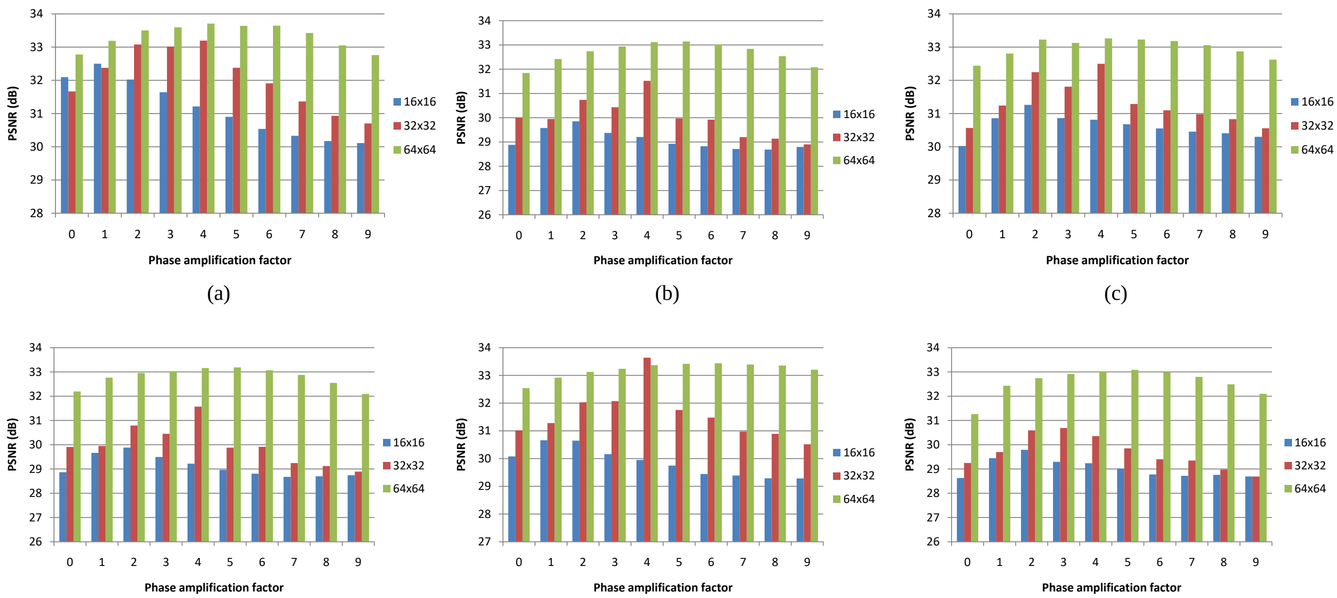
B. Experimentation with video sequences

As far as the video sequences are concerned, the well-known akiyo, container, flower, mobile_calendar and waterfall are used [37]. The first frames of the aforementioned video sequences are depicted in Fig. 7. The tested motion estimation algorithms are evaluated on these video sequences by applying block-based motion estimation. This means that each frame is divided in non-overlapping blocks of sizes 16×16, 32×32 and 64×64 pixels and the local motion compensated estimation error is evaluated for each block over all sequences.

We tested all motion estimation algorithms for various

(a) (b) (c) (d) (e)

Fig. 7. The first frames of the video sequences akiyo, container, flower, mobile_calendar and waterfall used in our evaluation process presented from left to right.



(a) (b) (c)

(d) (e) (f)

Fig. 8. Performance of (a) Foroosh, (b) PC+Quad, (c) Xiaohua, (d) Ren, (e) HOG-PC and (f) PC+Gaussian motion estimation algorithms averaged over all video sequences. The relationship between the phase amplification factor m and the block size is also depicted.

values of the phase amplification factor m over all video sequences for the first 50 frames and for various block sizes. The results expressed as PSNR and measured in dB are presented in Fig. 9. From Fig 9 a few conclusions can be drawn. Firstly, one may notice a consistent pattern as far as the effect of the phase amplification factor on the metric of PSNR is concerned. More specifically, PSNR improves as the phase amplification factor increases but a drop in PSNR is noticed when the amplification factor exceeds a certain threshold. The deterioration of the performance of the motion estimation algorithms is attributed to both the amplified phase noise and the drop in the magnitude of the cross-correlation peak, similarly to the experimentation with the MR images.

However, in the case of block-based motion estimation, another factor affects the sub-pixel accuracy of the motion estimation algorithms. As we mentioned in Section III-D, the cross-correlation peak is shifted because of the employment of the proposed PAC core. This shift did not play a significant role during the experimentation with the MR images as the block had the size of the entire image and the true motion was too small compared to the size of the block. In the experimentation with the video sequences, the block size is quite small and thus comparable to the true motion between the video frames. As a result, an excessive shift of the cross-correlation peak may move it outside of the borders of the block where it should be found. If that happens, the correct cross-correlation peak is circularly wrapped around the image block and the estimated motion vector does not correspond to the true motion between the compared image blocks. Additionally, the magnified motions cause a greater reduction of the overlapping area between smaller rather than larger image blocks. These conclusions are backed up by the fact that the threshold over which the performance of the motion estimation algorithms deteriorate is getting larger as the block size increases. Therefore, given that larger blocks are considered, there is space for the employment of larger phase amplification factors.

Furthermore, irrespectively of the video sequence and the block size, almost all sub-pixel motion estimation algorithms are improved when the proposed PAC core with small or mediocre phase amplification factors is employed and this performance improvement is enhanced when larger block sizes are considered for local motion estimation. Figure 8 depicts the relationship between the phase amplification factor and the image block size for each of the tested motion estimation algorithms. The performance of the algorithms is averaged

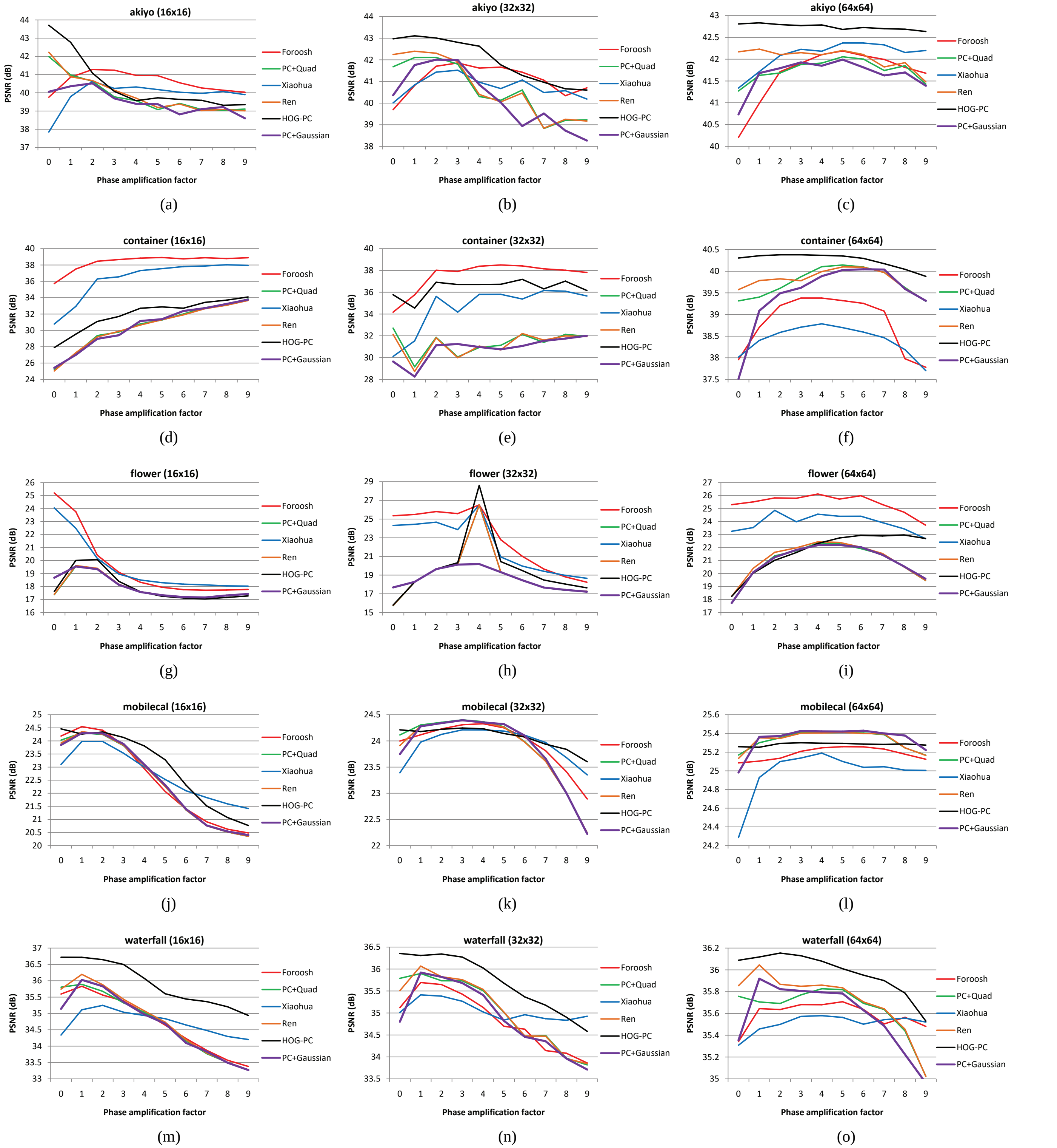


Fig. 9. The PSNR values, measured in dB, for all tested algorithms, various values of the phase amplification factor and blocks of size 16×16, 32×32 and 64×64 for the video sequences akiyo, container, flower, mobile_calendar and waterfall.

over all video sequences. Moreover, Table III presents the PSNR values of all tested algorithms averaged over all video sequences when the PC and the PAC cores are employed. The value of the phase amplification factor m that is mentioned next to the corresponding block size corresponds to the value that achieves the optimal average performance. The results of Table III verify that the motion estimation algorithms can achieve optimal performance for larger phase amplification factors when larger block sizes are considered. More specifically, a comparison between the PC and PAC cores shows that the novel PAC core provides an average improvement in the metric of PSNR by 2.48%, 6.04% and 2.98% when blocks

TABLE III
PSNR, MEASURED IN DB, OF MOTION ESTIMATION ALGORITHMS THAT EMPLOY EITHER PC OR PAC CORE WITH (W/) AND WITHOUT (W/O) THE NOISE HANDLING TERM (NH) AVERAGED OVER ALL VIDEO SEQUENCES

Block size \ Method	Foroosh	PC+Quad	Xiaohua	Ren	HOG-PC	PC+Gaussian
16×16 (m=0)	32.096	28.882	30.023	28.866	30.078	28.628
16×16 (m=2) w/o NH	31.811(-0.89%)	28.745(-0.47%)	30.390(+1.22%)	28.773(-0.32%)	30.136(+0.19%)	28.745(+0.41%)
16×16 (m=2) w/ NH	32.024(-0.22%)	29.853(+3.36%)	31.263(+4.13%)	29.881(+3.52%)	30.646(+1.89%)	29.789(+4.06%)
32×32 (m=0)	31.667	30.003	30.567	29.909	31.021	29.248
32×32 (m=4) w/o NH	32.281(+1.94%)	29.174(-2.76%)	31.011(+1.45%)	29.207(-2.35%)	31.711(+2.22%)	29.188(-0.21%)
32×32 (m=4) w/ NH	33.196(+4.83%)	31.520(+5.06%)	32.499(+6.32%)	31.571(+5.56%)	33.639(+8.44%)	30.356(+3.79%)
64×64 (m=0)	32.779	31.843	32.440	32.195	32.541	31.262
64×64 (m=5) w/o NH	33.390(+1.86%)	32.709(+2.72%)	32.979(+1.66%)	32.770(+1.79%)	33.114(+1.76%)	32.728(+4.69%)
64×64 (m=5) w/ NH	33.641(+2.63%)	33.143(+4.08%)	33.228(+2.43%)	33.189(+3.09%)	33.416(+2.69%)	33.084(+5.83%)

TABLE IV
COMPUTATIONAL COMPLEXITY IN MSEC OF MOTION ESTIMATION ALGORITHMS EMPLOYING EITHER PC OR PAC CORE WITH (W/) AND WITHOUT (W/O) THE NOISE HANDLING TERM (NH) FOR IMAGE BLOCKS OF SIZE 256×256.

Algorithm \ Method	PC	PAC w/o NH	PAC w/ NH
Foroosh	6.4	7.9(+23.44%)	11.8(+84.38%)
PC+Quad	7.2	8.7(+20.83%)	12.6(+75%)
Xiaohua	74.3	75.8(+2.02%)	79.7(+7.27%)
Ren	6.2	7.7(+24.19%)	11.6(+87.1%)
HOG-PC	8.8	10.3(+17.05%)	14.2(+61.36%)
PC+Gaussian	7.8	9.3(+19.23%)	13.2(+69.23%)

of size 16×16, 32×32 and 64×64 are employed respectively. Moreover, Table III verifies the consistent improvement in the values of PSNR when the noise handling term is applied in the PAC core.

C. Computational complexity

In this section, we examine and analyze the computational complexity of the proposed PAC core. From the perspective of computational complexity, the PAC core has to calculate a new cross-power spectrum, as shown in Eq. 5. This procedure initially involves the extraction of the phase difference between the compared images, then the amplification of the phase difference and finally the computation of the exponential term of Eq. 5. As the PAC core directly replaces the PC core employed by any sub-pixel motion estimation algorithm, the computational complexity of the PAC core is irrespective of the algorithm used for sub-pixel motion estimation. However, in order to evaluate the computational burden of the PAC core, we have to compare it with respect to the computational complexity of the tested sub-pixel motion estimation algorithms. All methodologies are implemented in Matlab and run on a quad-core processor (i7-7700HQ @ 2.80 GHz) with 8 GB of RAM. Table IV summarizes the computational complexity of the tested sub-pixel motion estimation algorithms with and without the PAC core for image blocks of size 256×256.

From Table IV, it can be concluded that in absolute numbers, the proposed PAC core does not introduce too much computational complexity (i.e. 1.5 and 5.4 msec without and with the noise handling term respectively). However, for fast sub-pixel motion estimation algorithms, such as Foroosh and Ren, the computational burden that the PAC core introduces can be significant. Furthermore, the most computationally expensive operation of the proposed PAC core is the noise handling term as it includes a convolution with a Gaussian kernel. Finally, for smaller block sizes, as the ones employed for the video sequences, the computational complexity of the proposed PAC core in absolute numbers can be negligible.

V. CONCLUSIONS

This paper presents a novel methodology that relies on the magnification of motion between two compared images to improve the accuracy and robustness of frequency-based sub-pixel motion estimation algorithms. The proposed PAC core is based on the notion that subtle motions can become more reliably estimated, once magnified. The advantages of the PAC core lie in the fact that it can be inserted in the core of any frequency-based motion estimation algorithm and directly substitute PC by calculating a new cross-power spectrum and cross-correlation function, while also taking advantage of interpolation techniques and properties of the cross-power spectrum. In this way, the PAC core overcomes the lower bound of accuracy of current motion estimation algorithms by computing a new cross-correlation function that better models the shift between the compared images prior to the extraction of the cross-correlation peak using well-established interpolation methods. Moreover, the proposed noise handling term enables the suppression of the effect of noise and the improvement of the sub-pixel motion estimation accuracy. An analysis of the motion amplification factor is also provided, along with the definition of a set of constraints that if violated, the PAC core can no longer perform optimally. Experiments with MR images and real video sequences reveal significant improvements in the sub-pixel accuracy of the tested motion estimation algorithms when the proposed PAC core is employed.

REFERENCES

[1] C. Kuglin and D. Hines, "The phase correlation image alignment method," *IEEE Conference on Cybernetics and Society*, pp. 163–165, 1975.

[2] J. Pearson, D. Hines, S. Golosman, and C. Kuglin, "Video-Rate Image Correlation Processor," in *Proc. SPIE, Applications of Digital Image Processing*, vol. 119, 1977, pp. 197–205.

[3] B. Girod, "Motion-compensating prediction with fractional-pel accuracy," *IEEE Transactions on Communications*, vol. 41, no. 4, pp. 604–612, Apr 1993.

[4] I. E. Abdou, "Practical approach to the registration of multiple frames of video images," in *Proc. SPIE, Visual Communications and Image Processing*, vol. 3653, 1998, pp. 371–382.

[5] V. Argyriou, "Sub-Hexagonal Phase Correlation for Motion Estimation," *IEEE Transactions on Image Processing*, vol. 20, no. 1, pp. 110–120, Jan 2011.

[6] N. Wadhwa, M. Rubinstein, F. Durand, and W. Freeman, "Phase-Based Video Motion Processing," *ACM Trans. Graph. (Proceedings SIGGRAPH 2013)*, vol. 32, no. 4, pp. 80:1–80:10, Jul 2013.

[7] H. Foroosh, J. B. Zerubia, and M. Berthod, "Extension of phase correlation to subpixel registration," *IEEE Transactions on Image Processing*, vol. 11, pp. 188–200, Mar 2002.

[8] H. Zhou and Y. Yu, "Planetary Image Live Stacking via Phase Correlation," in *9th International Symposium on Computational Intelligence and Design (ISCID)*, vol. 2, Dec 2016, pp. 17–21.

[9] M. Balci and H. Foroosh, "Subpixel estimation of shifts directly in the Fourier domain," *IEEE Transactions on Image Processing*, vol. 15, no. 7, pp. 1965–1972, July 2006.

[10] J. Ren, J. Jiang, and T. Vlachos, "High-Accuracy Sub-Pixel Motion Estimation From Noisy Images in Fourier Domain," *IEEE Transactions on Image Processing*, vol. 19, no. 5, pp. 1379–1384, May 2010.

[11] K. Takita, T. Aoki, Y. Sasaki, T. Higuchi, and K. Kobayashi, "High-accuracy subpixel image registration based on phase-only correlation," *IEICE Trans. Fundamentals*, vol. E86-A(8), pp. 1925–1934, 2003.

[12] H. Stone, M. Orchard, E. Chang, and S. Martucci, "A fast direct Fourier-based algorithm for subpixel registration of images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 10, pp. 2235–2243, Oct 2001.

[13] P. Vandewalle, S. Süsstrunk, and M. Vetterli, "A Frequency Domain Approach to Registration of Aliased Images with Application to Super-resolution," *EURASIP Journal on Advances in Signal Processing*, pp. 1–14, 2006.

[14] W. Hoge, "A subspace identification extension to the phase correlation method [MRI application]," *IEEE Transactions on Medical Imaging*, vol. 22, no. 2, pp. 277–280, Feb 2003.

[15] Y. Keller and A. Averbuch, "A projection-based extension to phase correlation image alignment," *Signal Processing*, vol. 87, no. 1, pp. 124–133, 2007.

[16] X. Tong, Z. Ye, Y. Xu, S. Liu, L. Li, H. Xie, and T. Li, "A Novel Subpixel Phase Correlation Method Using Singular Value Decomposition and Unified Random Sample Consensus," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 8, pp. 4143–4156, Aug 2015.

[17] Y. Dong, T. Long, W. Jiao, G. He, and Z. Zhang, "A Novel Image Registration Method Based on Phase Correlation Using Low-Rank Matrix Factorization With Mixture of Gaussian," *IEEE Transactions on Geoscience and Remote Sensing*, vol. PP, no. 99, pp. 1–15, 2017.

[18] V. Ojansivu and J. Heikkila, "Image Registration Using Blur-Invariant Phase Correlation," *IEEE Signal Processing Letters*, vol. 14, no. 7, pp. 449–452, July 2007.

[19] M. Pedone, J. Flusser, and J. Heikkilä, "Blur Invariant Translational Image Registration for N-fold Symmetric Blurs," *IEEE Transactions on Image Processing*, vol. 22, no. 9, pp. 3676–3689, Sept 2013.

[20] ——, "Registration of Images With N-Fold Dihedral Blur," *IEEE Transactions on Image Processing*, vol. 24, no. 3, pp. 1036–1045, March 2015.

[21] L. Zhongke, Y. Xiaohui, and W. Lenan, "Image registration based on Hough transform and phase correlation," in *Proceedings of the International Conference on Neural Networks and Signal Processing*, vol. 2, Dec 2003, pp. 956–959.

[22] V. Maik, E. Chae, E. Lee, G. Jo, S. Park, J. Han, J. Paik, and C. Park, "Robust sub-pixel image registration based on combination of local phase correlation and feature analysis," in *18th IEEE International Symposium on Consumer Electronics (ISCE 2014)*, June 2014, pp. 1–2.

[23] P. Cheng and C. Menq, "Real-Time Continuous Image Registration Enabling Ultraprecise 2-D Motion Tracking," *IEEE Transactions on Image Processing*, vol. 22, no. 5, pp. 2081–2090, May 2013.

[24] H. Park, G. R. Martin, and A. Bhalerao, "Local Affine Image Matching and Synthesis Based on Structural Patterns," *IEEE Transactions on Image Processing*, vol. 19, no. 8, pp. 1968–1977, Aug 2010.

[25] V. Argyriou and T. Vlachos, "Estimation of sub-pixel motion using gradient cross-correlation," *Electronics Letters*, vol. 39, no. 13, pp. 980–982, June 2003.

[26] G. Tzimiropoulos, V. Argyriou, S. Zafeiriou, and T. Stathaki, "Robust FFT-Based Scale-Invariant Image Registration with Image Gradients," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 10, pp. 1899–1906, Oct 2010.

[27] G. Tzimiropoulos, V. Argyriou, and T. Stathaki, "Subpixel Registration With Gradient Correlation," *IEEE Transactions on Image Processing*, vol. 20, no. 6, pp. 1761–1767, June 2011.

[28] V. Argyriou and T. Vlachos, "Quad-Tree Motion Estimation in the Frequency Domain Using Gradient Correlation," *IEEE Transactions on Multimedia*, vol. 9, no. 6, pp. 1147–1154, Oct 2007.

[29] V. Argyriou and G. Tzimiropoulos, "Frequency domain subpixel registration using HOG phase correlation," *Computer Vision and Image Understanding*, vol. 155, pp. 70–82, 2017.

[30] Y. Ye, J. Shan, L. Bruzzone, and L. Shen, "Robust Registration of Multimodal Remote Sensing Images Based on Structural Similarity," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 5, pp. 2941–2958, May 2017.

[31] X. Li, "High-Accuracy Subpixel Image Registration With Large Displacements," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 11, pp. 6265–6276, Nov 2017.

[32] A. Alba, J. Vigueras-Gomez, E. Arce-Santana, and R. Aguilar-Ponce, "Phase correlation with sub-pixel accuracy: A comparative study in 1D and 2D," *Computer Vision and Image Understanding*, vol. 137, pp. 76–87, 2015.

[33] A. Yousef, J. Li, and M. Karim, "High-Speed Image Registration Algorithm with Subpixel Accuracy," *IEEE Signal Processing Letters*, vol. 22, no. 10, pp. 1796–1800, Oct 2015.

[34] M. Uss, B. Vozel, V. Dushepa, V. Komjak, and K. Chehdi, "A Precise Lower Bound on Image Subpixel Registration Accuracy," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 6, pp. 3333–3345, June 2014.

[35] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. Freeman, "Eulerian Video Magnification for Revealing Subtle Changes in the World," *ACM Trans. Graph. (Proceedings SIGGRAPH 2012)*, vol. 31, no. 4, pp. 65:1–65:8, Jul 2012.

[36] V. Argyriou and T. Vlachos, "A study of sub-pixel motion estimation using phase correlation," in *British Machine Vision Conference*, Sept 2006, pp. 387–396.

[37] "Xiph.org Video Test Media [derf's collection]," https://media.xiph.org/video/derf/, [Online; accessed 11-Nov-2018].